FPSAC PROCEEDINGS 2005
ACTES SFCA 2005
Taormina
ITALIA

17th Annual International Conference
FORMAL POWER SERIES AND ALGEBRAIC COMBINATORICS
Proceedings edited by Luisa Carini, Hélène Barcelo and Jean-Yves Thibon
June 20 - June 25, 2005
Taormina – ITALIAS
ITALIA
(Universita degli studi Messina)

17ième Conférence Internationale
SERIES FORMELLES ET COMBINATOIRE ALGEBRIQUE
Actes édités par Luisa Carini, Hélène Barcelo and Jean-Yves Thibon
20 juin - 25 juin 2005
Taormina – ITALIE
(Universita degli studi Messina)

Programme Committee - Comité de programme

Christos Athanasiadis (Greece)
Hélène Barcelo (USA)
Francois Bergeron (Canada)
Sara Billey (USA)
Francesco Brenti (Italy)
Luisa Carini (Italy), chair - présidente
Bill Chen (China)
Alain Goupil (Canada)
Takayuki Hibi (Japan)
Anatol Kirillov (Russia)
Christian Krattenthaler (France)
Alain Lascoux (France)
Svante Linusson (Sweden)
Jennifer Morse (USA)
Marc Noy (Spain)
Jorn Olsson (Denmark)
Arun Ram (USA)
Domenico Senato (Italy)
Volker Strehl (Germany)
Jean-Yves Thibon, co-chair – co-président (France)
Helge Tverberg (Norway)
Michelle Wachs (USA)
Volkmar Welker (Germany)
Mike Zabrocki (Canada)

Organizing Committee - Comité d'organisation

Hélène Barcelo (USA)
Luisa Carini (Italy), chair
Vincenzo De Filippis (Italy)
Maylis Delest (France)
Jean-Marc Fédou (France)
Alain Goupil (Canada)
Enzo Maria Li Marzi (Italy)
Filippo Mignosi (Italy)
Jennifer Morse (USA)
Gaetana Restuccia (Italy)
Jean-Yves Thibon (France

FOREWARD
PREFACE

Welcome to all participants in the 17th International Conference on Algebraic Combinatorics and Formal Power Series FPSAC'05, Séries Formelles et Combinatoire Algébrique in Taormina, Sicily, Italy.

This year's conference is held in celebration of Adriano Garsia's 75th birthday. We wish to celebrate his remarkable contributions to combinatorics, as well as his legendary enthusiasm in the hope of thanking him for the constant source of inspiration he has been to many of us. The last two days of the conference are reserved for talks given by his former students, co-authors and friends culminating with a lecture by Adriano himself.

This volume contains the extended abstracts of the papers to be presented at this 17th FPSAC conference. There are five plenary lecturers given by Christine Bessenrodt, Corrado De Concini, Ronald King, Renzo Pinzani and Jeff Remmel.

We wish to thank the members of the program committee for their invaluable work in helping to select talks and posters from a pool of 115 very high-quality submissions.

We are extremely grateful to Giacomo Scarfì and Francesco Sottile for their hard and long hours of work for the creation and maintenance of the conference's website. We wish to thank the members of the organizing committee for their helpful insights. In particular, Maylis Delest who put in far more than her share of work and was responsible for the electronic submissions and (local) proceedings. We are also thankful to Jennifer Morse who, with her exemplary efficiency obtained funds from both the National Security Agency and the National Science Foundation; Christian Krattenthaler, Francesco Brenti and Andrea Brini who are generously contributing funds from the Algebraic Combinatorics in Europe network and the University of Bologna.

We are grateful to the following institutions for their substantial financial support: The University of Messina, the University of Palermo, the University of Bologna, Gruppo Nazionale per le Strutture Algebriche e Geometriche e loro Applicazioni (GNSAGA) dell'Istituto Nazionale di Alta Matematica, the National Security Agency (USA), National Science Foundation (USA) and HP-Makers.

Finally, we wish to thank the invited speakers, all those who submitted papers for consideration, and the participants.

Luisa Carini, Hélène Barcelo, Jean-Yves Thibon Messina, June 2005

TABLE OF CONTENTS
TABLE DES MATIERES

Invited Speakers - Conférenciers Invités

Papers - Articles

Posters- Affiches

Software - Logiciels

Algebra invariants for finite directed graphs with relations

Christine Bessenrodt [1]

Institut für Mathematik, Universität Hannover,
Welfengarten 1, D-30167 Hannover, Germany

April 25, 2005

1 Introduction

Finite directed graphs play an important rôle in the representation theory of finite-dimensional algebras, where they are called *quivers*. For a given quiver Q and a field K one obtains an algebra KQ by taking all paths in Q, including the trivial paths of length 0 at each vertex, as a K-basis, and defines multiplication as induced by concatenation of paths; this is called the *path algebra* to the quiver Q. The importance of the path algebras lies in a famous result by Gabriel that any finite-dimensional K-algebra over an algebraically closed field K is Morita equivalent to a factor algebra KQ/I, where I is an admissible ideal of KQ (i.e., contained in the radical square of KQ). Thus such factor algebras KQ/I are central objects of study in the representation theory of algebras; this situation is referred to as a *quiver with relations*. There are many interesting representation theoretic properties of such algebras for which one tries to find "combinatorial" methods for computing them. It is of particular importance to consider representation theoretic parameters of the algebra which are invariants for appropriate equivalence classes of algebras. In recent years a focus in the representation theory of algebras has been the investigation of derived equivalences of algebras; this is a homological notion: two algebras are derived equivalent if their derived module categories are equivalent. A lot of progress has recently been made in this very active area. It is a difficult problem to find invariants of algebras preserved by derived equivalences; only a few important representation theoretic parameters are known to be indeed invariants under derived equivalence, such as the number of simple modules, the dimension of the center of the algebra or the dimension of its Hochschild cohomology groups.

[1] In this article we present some joint work with Th. Holm; full proofs of the results in sections 2 and 3 are contained in [5].

Here, we will discuss invariants of the *Cartan matrix* of a finite-dimensional algebra $A = KQ/I$. Cartan matrices contain crucial structural information on the algebra, as their entries are the multiplicities of simple A-modules as composition factors of projective indecomposable A-modules. It is in general difficult to compute the entries of the Cartan matrix and some famous conjectures are related to their properties. The main point to note here is that the unimodular equivalence class of the Cartan matrix of a finite dimensional algebra is invariant under derived equivalence.
From a combinatorial point of view it is important that for a finite-dimensional algebra $A = KQ/I$ given by a quiver with relations, the entries of the Cartan matrix can be computed by counting paths in the quiver Q which are non-zero in the algebra A.

The algebras we study here are the (skewed-) gentle algebras which are defined combinatorially by conditions on the quiver and relations. Gentle algebras occur naturally in many places in the representation theory of finite dimensional algebras, in particular in connection with derived categories. They made their first appearance in 1981 [1] when it was shown that the algebras which are derived equivalent to hereditary algebras of type $\mathbb{A}$ are precisely the gentle algebras whose underlying undirected graph is a tree. The algebras which are derived equivalent to hereditary algebras of type $\tilde{\mathbb{A}}$ are certain gentle algebras whose underlying graph has exactly one cycle [2]. Only recently it was proved that the class of gentle algebras has the remarkable property of being closed under derived equivalence [9]. For more background on the algebraic context the reader is referred to [5].

The starting point of our investigation was a recent result by Th. Holm [7] giving an explicit combinatorial formula for the Cartan determinants of gentle algebras.
Here, we refine these results to a determination of the invariant factors of the Cartan matrix C_A of a gentle algebra $A = KQ/I$, and we also extend the formulae to skewed-gentle algebras. Indeed, the key is to refine the combinatorial analysis of the quiver and put a weight on the paths according to their lengths instead of just counting them; in our context, this makes good sense as the relations on the quiver are homogeneous (in fact, they are even more special). Taking an indeterminate q corresponding to the weight of an arrow, this gives us a q-Cartan matrix $C_A(q)$ for the algebra; this may also be considered as a so-called filtered Cartan matrix, counting the multiplicities of the simple modules in the radical layers of the algebra. Setting $q = 1$

gives the ordinary Cartan matrix C_A. We have already pointed out that the unimodular equivalence class of the Cartan matrix of a finite-dimensional algebra is invariant under derived equivalence. But unfortunately, not even the determinant of the q-Cartan matrix is in general an invariant under derived equivalence.

There are further finite-dimensional algebras given by quivers and relations for which it is possible to determine the q-Cartan matrices and obtain nice formulae for their invariants or at least their determinant. As an illustration, one such further family of quivers is considered in the final section.

2 Gentle algebras

In this section, we want to describe an extension and refinement of the result on the determinant of the Cartan matrix of a gentle algebra from [7].

First we have to give the definition of gentle algebras. They form an important subclass of the class of special biserial algebras which we now define.

Let K be an algebraically closed field. Let Q be a quiver, i.e., a finite directed graph, with set of vertices Q_0. Let I be an admissible ideal of the path algebra KQ, i.e., $I \subseteq \mathrm{rad}^2(KQ)$. Note that the radical of KQ is just the ideal generated by the arrows.

For a path p in Q we denote by $s(p)$ its start vertex and $t(p)$ its end vertex. The pair (Q, I) is called *special biserial* if the following holds:

(i) For any vertex $v \in Q_0$ the set of lengths of the paths starting in v and not being in I is finite.

(ii) Each vertex $v \in Q$ is the end point of at most two arrows and the starting point of at most two arrows.

(iii) For every arrow α there is at most one arrow β with $t(\alpha) = s(\beta)$ and $\alpha\beta \notin I$, and there is at most one arrow γ with $t(\gamma) = s(\alpha)$ and $\gamma\alpha \notin I$.

A special biserial pair (Q, I) is called *gentle*, if furthermore:

(iv) There is a generating set of I (as ideal) consisting of paths of lengths 2.

(v) For any arrow α there is at most one arrow β with $t(\alpha) = s(\beta)$ and $\alpha\beta \in I$, and there is at most one arrow γ with $t(\gamma) = s(\alpha)$ and $\gamma\alpha \in I$.

A K-algebra A is called *gentle* (resp. *special biserial*) if it is Morita equivalent to an algebra KQ/I, for (Q, I) gentle (resp. special biserial).

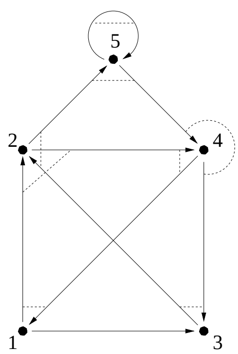
As we assume that the set of vertices Q_0 of Q is finite, condition (i) implies that the algebra KQ/I is finite-dimensional.

By condition (iv), the relations for a gentle quiver are homogeneous, so the length of a non-zero path in KQ/I is well-defined. Then we define the q-Cartan matrix of $A = KQ/I$ as $C_A(q) = (c_{ij}(q))_{i,j\in Q_0}$, where for $i, j \in Q_0$

$$c_{ij}(q) = \sum_{n\geq 0} a_n(i,j)q^n$$

with $a_n(i,j)$ the number of paths from i to j which are non-zero in A (and different in A).

Example. In the following picture, the dotted lines (or arcs) between two arrows or in a loop correspond to the generating relations for the quiver, i.e., they indicate that composing the corresponding arrows is zero in the algebra $A = KQ/I$.



Here is the q-Cartan matrix for this gentle quiver with relations:

$$C_A(q) = \begin{pmatrix} 1+q^5 & q+q^2+q^7 & q+q^4+q^6+q^9 & q^3+q^4+q^8 & q^2+q^3 \\ q^4 & 1+q^6 & q^2+q^5+q^8 & q+q^3+q^7 & q+q^2 \\ 0 & q & 1+q^3 & q^2 & 0 \\ q & q^3 & q+q^2+q^5 & 1+q^4 & 0 \\ q^2+q^3 & q^4+q^5 & q^3+q^4+q^6+q^7 & q+q^2+q^5+q^6 & 1+q \end{pmatrix}$$

The following property does not only hold for gentle algebras but for those where we have dropped the final condition (v) in the definition of gentle pairs; it gives a useful reduction tool in the proof of the main result.

Lemma 2.1 *Let $A = KQ/I$ be a special biserial algebra, where I is generated by paths of length 2. Let α be an arrow in Q, not a loop, such that there is no arrow β with $s(\alpha) = t(\beta)$ and $\beta\alpha \in I$, or there is no arrow γ with $t(\alpha) = s(\gamma)$ and $\alpha\gamma \in I$. Let Q' be the quiver obtained from Q by removing the arrow α, I' the corresponding relation ideal and $A' = KQ'/I'$. Then the q-Cartan matrices $C_A(q)$ and $C_{A'}(q)$ are unimodularly equivalent (over $\mathbb{Z}[q]$).*

For the combinatorially defined gentle algebras we can provide an explicit combinatorial description for a very nice normal form of its q-Cartan matrix:

Theorem 2.2 *Let $A = KQ/I$ be a gentle algebra, defined by a gentle pair (Q, I). Denote by c_k the number of oriented k-cycles in Q with full zero relations.*
Then the q-Cartan matrix $C_A(q)$ is unimodularly equivalent (over $\mathbb{Z}[q]$) to a diagonal matrix with entries $(1 - (-q)^k)$, with multiplicity c_k, $k \geq 1$, and all further diagonal entries being 1.

This result has some immediate nice consequences.

Corollary 2.3 *Let $A = KQ/I$ be a gentle algebra, and denote by c_k the number of oriented k-cycles in Q with full zero relations. Then the q-Cartan matrix $C_A(q)$ has determinant*

$$\det C_A(q) = \prod_{k \geq 1} (1 - (-q)^k)^{c_k} .$$

Example. In the example given before, the q-Cartan matrix has determinant

$$\det C_A(q) = 1 + q + q^3 - q^5 - q^7 - q^8 = (1 + q)(1 + q^3)(1 - q^4) .$$

Indeed, the quiver has a loop at vertex 5 (which is a 1-cycle with a zero relation), a 3-cycle with full zero relations from vertex 1 to 2 to 4 and back to 1, and a 4-cycle with full zero relations running over 2,5,4,3 and back to 2.

For an algebra $A = KQ/I$ as above, let $ec(A)$ and $oc(A)$ be the number of oriented cycles in Q with full zero relations of even and odd length, respectively. Setting $q = 1$, the corollary above immediately implies the main result from [7] which was in fact the starting point of our investigations:

Corollary 2.4 *Let $A = KQ/I$ be a gentle algebra. Then for the determinant of its ordinary Cartan matrix C_A the following holds.*

$$\det C_A = \begin{cases} 0 & \text{if } ec(A) > 0 \\ 2^{oc(A)} & \text{else} \end{cases}$$

The Theorem also gives:

Corollary 2.5 *Let $A = KQ/I$ be a gentle algebra. Then there are at most $|Q_0|$ oriented cycles with full zero relations in the quiver Q.*

Instead of deriving this from the main result, this may be proved directly. As an illustration, we give this alternative proof here.
Clearly the result holds when $|Q_0| = 1$. So we assume now that $|Q_0| > 1$, and we prove the result by induction.
We may remove all arrows going in or out of a vertex and not having a zero relation at this vertex, without losing the property of the quiver with relations being gentle and without changing the number of vertices and the number of oriented cycles with full relations. After this removal, all vertices are of degree 0, 2 or 4, and there is a zero relation at all vertices of degree 2. By induction, we may assume that there is no vertex of degree 0 and that the quiver is connected. If all vertices are of degree 4, then there are paths of arbitrary lengths, contradicting the property of being gentle. Hence there is a vertex of degree 2, which then belongs to a unique oriented cycle with full zero relations. Removing this vertex and the arrows incident to it reduces the number of vertices as well as the number of oriented cycles with full zero relations by 1 (note that any arrow in a gentle quiver belongs to at most one oriented cycle with full zero relations). Hence the result follows by induction.

Remark 2.6 Note that in our context $|Q_0|$ is the number $l(A)$ of simple A-modules, which is also invariant under derived equivalence. Hence this implies that the Cartan determinant of a gentle algebra A is at most $2^{l(A)}$.

Recall that the property of an algebra being gentle is invariant under derived equivalence [9]. Also, we have pointed out earlier that the invariant factors of the ordinary Cartan matrix $C_A = C_A(1)$ are invariants of the derived equivalence class of the algebra $A = KQ/I$. Thus we now have some easily computable invariants for gentle algebras to distinguish the derived equivalence classes.

Corollary 2.7 *Let $A = KQ/I$ and $A' = KQ'/I'$ be gentle algebras as above. If A is derived equivalent to A', then $ec(A) = ec(A')$ and $oc(A) = oc(A')$.*

Our new invariants are a quite powerful tool for distinguishing gentle algebras up to derived equivalence which cannot be separated by the more classical invariants. An illustration on how to use the invariants to tell non-equivalent gentle algebras apart is given in [5], where the 9 gentle algebras with two simples and the 18 gentle algebras with three simples and vanishing Cartan determinant are discussed in detail.

3 Skewed-gentle algebras

Also skewed-gentle algebras are defined combinatorially. They were introduced in [6]; for the notation and definition we follow here mostly [4], but we try to explain how the construction works rather than stating the technical definitions.

We start with a gentle pair (Q, I). A set Sp of vertices of the quiver Q is an admissible set of *special* vertices if the quiver with relations obtained from Q by adding loops with square zero at these vertices is again gentle; we denote this gentle pair by $(Q^{\rm sp}, I^{\rm sp})$. The triple (Q, Sp, I) is then called *skewed-gentle.*

We want to point out that the admissibility of the set Sp of special vertices is both a local as well as a global condition. Let v be a vertex in the gentle quiver (Q, I); then we can only add a loop at v if v is of degree 1 or 0 or if it is of degree 2 with a non-loop zero relation. Hence only vertices of this type are potential special vertices. But for the choice of an admissible set of special vertices we also have to take care of the global condition that after adding all loops, the pair $(Q^{\rm sp}, I^{\rm sp})$ still does not have paths of arbitrary lengths.

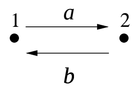
Given a skewed-gentle triple (Q, Sp, I), we now construct a new quiver with relations $(\hat{Q}, \hat{I})$ by doubling the special vertices, introducing arrows to and from these vertices corresponding to the previous such arrows and replacing a previous zero relation at the vertices by a mesh relation.

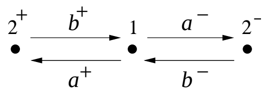
More precisely, we proceed as follows. The non-special vertices in Q are also vertices in the new quiver; any arrow between non-special vertices as well as corresponding relations are also kept. Any special vertex $v \in Sp$ is replaced by two vertices v^+ and v^- in the new quiver. An arrow a in Q from a non-special vertex w to v (or from v to w) will be doubled to arrows $a^\pm : w \to v^\pm$

(or $a^{\pm} : v^{\pm} \to w$, resp.) in the new quiver; an arrow between two special vertices v, w will correspondingly give four arrows between the pairs $v^{\pm}$ and $w^{\pm}$. We say that these new arrows lie over the arrow a. Any relation $ab = 0$ where $t(a) = s(b)$ is non-special gives a corresponding zero relation for paths of length 2 with the same start and end points lying over ab. If v is a special vertex of degree 2 in Q, then the corresponding zero relation at v, say $ab = 0$ with $t(a) = v = s(b)$, is replaced by mesh commutation relations saying that any two paths of length 2 lying over ab, having the same start and end points but running over v^+ and v^-, respectively, coincide in the factor algebra to the new quiver with relations $(\hat{Q}, \hat{I})$.
We will speak of $(\hat{Q}, \hat{I})$ as a skewed-gentle quiver *covering* the gentle pair (Q, I). Note that also here the generating relations are homogeneous.

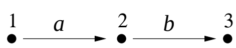
A K-algebra is then called *skewed-gentle* if it is Morita equivalent to a factor algebra $K\hat{Q}/\hat{I}$, where $(\hat{Q}, \hat{I})$ comes from a skewed-gentle triple (Q, Sp, I) as described above.

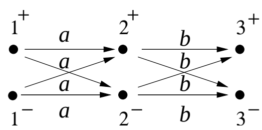
Examples. (1) We take the gentle quiver Q as shown below, with relation ideal I generated by ab and ba. Then we can take $Sp = \{2\}$, i.e., only the vertex 2 is specified as a special vertex. This gives the skewed-gentle quiver $\hat{Q}$ shown below, with relation ideal $\hat{I}$ generated by $a^+b^+ - a^-b^-$, $b^{\pm}a^{\pm}$, $b^{\pm}a^{\mp}$.

Q 

$\hat{Q}$ 

(2) We take the gentle quiver Q as shown below, with relation ideal generated by ab. This time we take $Sp = Q_0$, i.e., all vertices are chosen to be special. This gives the skewed-gentle quiver $\hat{Q}$ shown below, where for simplicity all arrows lying over a or b, respectively, are also marked a or b, respectively, but for writing down the relations generating the relation ideal $\hat{I}$ we will put signs on, so that e.g. a_+^- denotes the arrow going from 1^+ to 2^-. In this notation the generating relations are given by $a_+^+b_+^+ - a_+^-b_-^+$, $a_+^+b_+^- - a_+^-b_-^-$, $a_-^+b_+^+ - a_-^-b_-^+$, $a_-^+b_+^- - a_-^-b_-^-$.

Q 

$\hat{Q}$ 

Our result on gentle algebras generalizes nicely to skewed-gentle algebras:

Theorem 3.1 *Let (Q, I) be a gentle quiver, $(\hat{Q}, \hat{I})$ a covering skewed-gentle quiver. Let $\hat{A} = K\hat{Q}/\hat{I}$ be the corresponding skewed-gentle algebra. Denote by c_k the number of oriented k-cycles in (Q, I) with full zero relations.*
Then the q-Cartan matrix $C_{\hat{A}}(q)$ is unimodularly equivalent (over $\mathbb{Z}[q]$) to a diagonal matrix with entries $1 - (-q)^k$, with multiplicity c_k, $k \geq 1$, and all further diagonal entries being 1.

Remark 3.2 Thus, the q-Cartan matrix $C_A(q)$ for the gentle algebra A to (Q, I), and the q-Cartan matrix $C_{\hat{A}}(q)$ for a skewed-gentle cover $\hat{A}$ are unimodularly equivalent to diagonal matrices which only differ by adding as many further 1's on the diagonal as there are special vertices chosen in Q; in particular, with notation as above,

$$\det C_{\hat{A}}(q) = \det C_A(q) = \prod_{k \geq 1} (1 - (-q)^k)^{c_k}$$

and thus also for the ordinary Cartan matrices

$$\det C_A = \det C_{\hat{A}} \, .$$

4 Cycles and circulants

There are some more types of algebras for which one can determine the unimodular equivalence class of their q-Cartan matrices. As an example, we consider q-Cartan matrices to cyclic quivers with relations which occur in other interesting contexts.
In this situation, circulants make an appearance, and we first define the relevant notation.
For $x = (x_1, \dots, x_n)$ the circulant matrix to x is

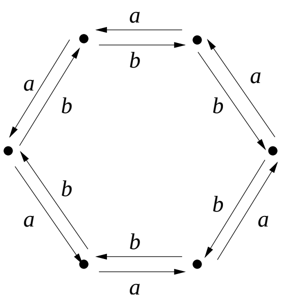
$$C = \begin{pmatrix} x_1 & x_2 & \dots & x_{n-1} & x_n \\ x_n & x_1 & \dots & x_{n-2} & x_{n-1} \\ \vdots & & & & \\ x_2 & x_3 & \dots & x_n & x_1 \end{pmatrix} .$$

Let ω be a primitive n-th root of unity in $\mathbb{C}$. Then C has the eigenvalues

$$\sum_{k=1}^{n} x_k (\omega^j)^{k-1} \, , \; j = 1, \dots, n.$$

Special circulants appear as q-Cartan matrices for cyclic quivers with monomial relations and commutation relations which are homogeneous of degree 2.

Example. Here is the corresponding cyclic quiver for the case of 6 vertices:



where we take as generating relations: a^2, b^2, $ab - ba$ (at each vertex of the quiver).

The q-Cartan matrix for such a cyclic quiver is just the circulant to $v = (1+q^2, q, 0\ldots, 0, q)$. We will describe a more precise result on these q-Cartan matrices below, but as it is easy to do we compute here the determinant. Note that as in the case of gentle quivers we have a contribution $1 - (-q)^n$ for each of the two oriented cycles with full zero relations in the quiver.

Theorem 4.1 *Let q be an indeterminate, $v = (1+q^2, q, 0\ldots, 0, q)$ (of length n), and $C(q)$ the circulant to v. Then we have*

$$\det C(q) = (1 - (-q)^n)^2\,.$$

Proof. By the above, we have

$$\begin{aligned}\det C(q) &= \prod_{j=1}^{n}(1 + q^2 + q\omega^j + q(\omega^j)^{n-1}) = \prod_{j=1}^{n}(1 + q^2 + q(\omega^j + \overline{\omega^j}))\\ &= \prod_{j=1}^{n}(q + \omega^j)(q + \overline{\omega^j}) = (\prod_{j=1}^{n}(q + \omega^j))^2\end{aligned}\,.$$

Now $-\omega^j$ is a zero of $q^n - 1$, if n is even and a zero of $q^n + 1$, if n is odd, for all j, hence we have the assertion. $\diamond$

In fact, it is not hard to transform the circulant $C(q)$ to $v = (1+q^2, q, 0\ldots, 0, q)$ into a better form:

Theorem 4.2 *Assume $n \geq 3$, and let $C(q)$ be as above.*
(i) Over $\mathbb{Z}[q]$, we can transform $C(q)$ unimodularly into the form

$$\begin{pmatrix} E_{n-2} & 0 & 0 \\ 0 & 1-(-q)^n & \big(\sum_{j=1}^{n-1} q^{2j-1}\big) - (-q)^{n-1} \\ 0 & 0 & 1-(-q)^n \end{pmatrix}$$

where E_{n-2} is the identity matrix of type $n-2$.
(ii) Over $\mathbb{Q}[q]$, we have the following unimodular equivalences:

For n odd, $$C(q) \sim \mathrm{diag}(1^{n-2}, \sum_{j=0}^{n-1}(-q)^j, q^{n+1}+q^n+q+1)\,.$$

For n even, $$C(q) \sim \mathrm{diag}(1^{n-2}, \sum_{j=0}^{(n-2)/2} q^{2j}, q^{n+2}-q^n-q^2+1)\,.$$

References

[1] I. Assem, D. Happel, Generalized tilted algebras of type $\mathbb{A}_n$, Comm. Algebra 9 (1981), no.20, 2101-2125

[2] I. Assem, A. Skowroński, Iterated tilted algebras of type $\tilde{\mathbb{A}}_n$, Math. Z. 195 (1987), 269-290

[3] V. Bekkert, H. Merklen, Indecomposables in derived categories of gentle algebras, Algebras and Representation Theory (4) 6 (2003)

[4] V. Bekkert, E.N. Marcos, H.A. Merklen, Indecomposables in derived categories of skewed-gentle algebras, Communications in Algebra (6) 31 (2003), 2615-2654

[5] C. Bessenrodt, Th. Holm, q-Cartan matrices and combinatorial invariants of derived categories for skewed-gentle algebras, preprint 2005

[6] C. Geiss, J. A. de la Peña, Auslander-Reiten components for clans, Boll. Soc. Mat. Mexicana 5 (1999), 307-326

[7] Th. Holm, Cartan determinants for gentle algebras, Archiv Math., to appear

[8] J. Rickard, Morita theory for derived categories, J. London Math. Soc. (2) 39 (1989) 436-456

[9] J. Schröer, A. Zimmermann, Stable endomorphism algebras of modules over special biserial algebras, Math. Z. 244 (2003) 515-530

Toric arrangements and lattice points in convex polytopes

Corrado De Concini

May 17, 2005

Abstract In the talk we shall introduce the notion of toric arrangement and explain how to compute the cohomology of their complements. Following ideas of Szenes and Vergne, application will be given to the computation of the number of lattice points in a rational convex polytope

THE HIVE MODEL AND THE POLYNOMIAL NATURE OF STRETCHED LITTLEWOOD-RICHARDSON COEFFICIENTS

R.C. KING

This work has been carried out in collaboration with Dr Christophe Tollu and Dr Frédéric Toumazet, Laboratoire d'Informatique de Paris-Nord, Université Paris 13

ABSTRACT. The hive model is used to explore the properties of both ordinary and stretched Littlewood-Richardson coefficients. The latter are polynomials in the stretching parameter t. It is shown that these may factorise, and that they can then be expressed as products of certain primitive polynomials. It is further shown how to determine a sequence of linear factors $(t+m)$ of the primitive polynomials, as well as bounds on their degree which are conjectured to be exact.

RÉSUMÉ. Nous utilisons le modèle des ruches pour étudier les propriétés des cœfficients dilatés de Litlewood-Richardson et des polynômes associés aux cœfficients de Littlewood-Richardson dilatés. Nous montrons que les uns et les autres peuvent se factoriser : ils s'écrivent comme des produits de coefficients (resp. polynômes) primitifs. En outre, nous montrons comment établir une suite de facteurs linéaires $(t+m)$ des polynômes primitifs , et proposons les bornes supérieure de leur degrés.

1. INTRODUCTION

Littlewood-Richardson coefficients, $c^{\nu}_{\lambda\mu}$, are interesting combinatorial objects [LR]. They are indexed by partitions λ, μ and ν, and they count the number of Littlewood-Richardson tableaux of skew shape ν/λ and weight μ. They are therefore non-negative integers. Although it is a non-trivial matter to determine whether or not $c^{\nu}_{\lambda\mu}$ is non-zero, it turns out that this is the case if and only if $|\lambda|+|\mu|=|\nu|$ and certain partial sums of the parts of λ, μ and ν satisfy what are known as Horn inequalities.

Multiplying all the parts of the partitions λ, μ and ν by a stretching parameter t, with t a positive integer, gives new partitions $t\lambda$, $t\mu$ and $t\nu$. The corresponding stretched Littlewood-Richardson coefficient are known to be polynomial in the stretching parameter t [DW2, R]. Such an LR-polynomial is defined by

$$P^{\nu}_{\lambda\mu}(t) = c^{t\nu}_{t\lambda,t\mu}, \tag{1.1}$$

and has a generating function of the form

$$F^{\nu}_{\lambda\mu}(z) = \frac{G^{\nu}_{\lambda\mu}(z)}{(1-z)^{d+1}} = \sum_{t=0}^{\infty} c^{t\nu}_{t\lambda,t\mu}\ z^t \tag{1.2}$$

where d is the degree of $P^{\nu}_{\lambda\mu}(t)$, and $G^{\nu}_{\lambda\mu}$ is a polynomial in z of degree $g \le d$.

For example, in the case $\lambda=(4,3,3,2,1)$, $\mu=(4,3,2,2,1)$ and $\nu=(7,4,4,3,2,1)$ one finds $c^{\nu}_{\lambda\mu}=13$ and

$$P^{\nu}_{\lambda,\mu}(t) = \frac{1}{10080}(t+1)(t+2)(t+3)(t+4)(t+5)(5t+21)(t^2+2t+4), \tag{1.3}$$

with

$$F^{\nu}_{\lambda\mu}(z) = \frac{1+4z+12z^2+3z^3}{(1-z)^9}. \tag{1.4}$$

It is the intention here to try to shed some light on the nature of the LR-polynomials and their generating functions. In particular we concentrate on the possible factorisation of any particular LR-polynomial as a product of simpler LR-polynomials, the degree d of an LR-polynomial, and the number of its linear factors $(t+m)$. We do not explore two particular conjectures [KTT1] to the effect that the non-zero coefficients of the polynomial $P^{\nu}_{\lambda\mu}(t)$ are always positive rational numbers, while those of $G^{\nu}_{\lambda\mu}(z)$ are all positive integers.

Our approach is based largely on the use of a hive model [BZ2, KT, B] which allows Littlewood-Richardson coefficients to be evaluated through the enumeration of integer points of certain rational polytopes. Before defining hives, puzzles and plans, that are the combinatorial constructs to be used in this context, it is worth recalling some definitions and properties of Littlewood-Richardson coefficients and LR-polynomials.

2. Definitions and properties

Let n be a fixed positive integer, let $\mathbf{x}=(x_1,x_2,\ldots,x_n)$ be a vector of indeterminates, and let $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)$ be a partition of weight $|\lambda|$ and of length $\ell(\lambda)\le n$. Thus $\lambda_k\in\mathbb{Z}^+$ for $k=1,2,\ldots,n$, with while $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_{\ell(\lambda)}>0$. and $\lambda_k=0$ for $k>\ell(\lambda)$.

Definition 2.1. *For each partition λ with $\ell(\lambda)\le n$ there corresponds a Schur function $s_\lambda(\mathbf{x})$ defined by*

$$s_\lambda(\mathbf{x}) = \frac{\left|x_i^{n+\lambda_j-j}\right|_{1\le i,j\le n}}{\left|x_i^{n-j}\right|_{1\le i,j\le n}} \tag{2.1}$$

Choosing n sufficiently large, the Littlewood-Richardson coefficients may be defined by

Definition 2.2.

$$s_\lambda(\mathbf{x})\ s_\mu(\mathbf{x}) = \sum_\nu c^{\nu}_{\lambda\mu}\ s_\nu(\mathbf{x}). \tag{2.2}$$

where the summation is over all partitions ν.

The expansion (2.2) may be effected by means of the Littlewood-Richardson rule [LR] which states that $c^{\nu}_{\lambda\mu}$ is the number of Littlewood-Richardson skew tableaux of shape ν/μ and weight λ obtained by numbering the boxes of the skew Young diagram $F^{\nu/\mu}$ with λ_i entries i for $i=1,2,\ldots,n$ that are weakly increasing across rows, strictly increasing down columns and satisfy the lattice permutation rule.

To specify the necessary and sufficient conditions on λ, μ and ν for $c^{\nu}_{\lambda\mu}$ to be non-zero it is convenient to introduce the notion of partial sums of the parts of a partition and some other notational devices.

Let n be a fixed positive integer and $N=\{1,2,\ldots,n\}$. Then for any positive integer $r\le n$ and any subset $I=\{i_1,i_2,\ldots,i_r\}$ of N of cardinality $\#I=r$, the partial sum indexed by I of any partition $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)$ of length $\ell(\lambda)\le n$ is defined to be

$$ps\,(\nu)_I = \nu_{i_1}+\nu_{i_2}+\cdots+\nu_{i_r} \tag{2.3}$$

If $i_1<i_2<\cdots<i_r$, let $\tilde I=(i_r,\ldots,i_2,i_1)$. It follows that if $\delta_r=(r,r-1,\ldots,1)$ then $\alpha(I)=\tilde I-\delta_r$ is a partition of length $\ell(\alpha(I))\le r$.

With this notation, building on a connection with the Horn conjecture [H] regarding eigenvalues of Hermitian matrices, the following theorem has been established by Klyachko [K], Knutson and Tao [KT], Knutsen, Tao and Woodward [KTW] and others. A comprehensive review of these devlopments has been provided by Fulton [F].

Theorem 2.3 (Horn inequalities)**.** *Let λ, μ and ν be partitions of lengths $\ell(\lambda),\ell(\mu),\ell(\nu)\le n$. Then $c^{\nu}_{\lambda\mu}>0$ if and only if $|\nu|=|\lambda|+|\mu|$ and for all $r=1,2,\ldots,n$*

$$ps\,(\nu)_K \le ps\,(\lambda)_I + ps\,(\mu)_J \tag{2.4}$$

for all triples $(I,J,K)\in R^n_r$, where R^n_r is the set of triples (I,J,K) with $\#I=\#J=\#K=r$ such that if $\alpha(I)=\tilde I-\delta_r$, $\beta(J)=\tilde J-\delta_r$ and $\gamma(K)=\tilde K-\delta_r$ then $c^{\gamma(K)}_{\alpha(I)\beta(J)}=1$.

Unfortunately, even for comparatively small values of r it is not a trivial matter to identify all partitions $\alpha(I), \beta(J), \gamma(K)$, and hence (I, J, K), such that the Littlewood-Richardson coefficient $c^{\gamma(K)}_{\alpha(I)\beta(J)} = 1$. We will return to this problem only after the hive model has been introduced.

Turning instead to stretched Littlewood-Richardson coefficients, the fact that all the above partial sum conditions are linear and homogeneous in the various parts of λ, μ and ν ensures the validity of the following:

Theorem 2.4 (Saturation Condition). [KT, B, DW1] *For all positive integers t*

$$c^{t\nu}_{t\lambda,t\mu} > 0 \iff c^{\nu}_{\lambda\mu} > 0. \tag{2.5}$$

Furthermore, it has been established by a variety of means that

Theorem 2.5 (Polynomial Condition). [DW2, R] *For all partitions λ, μ and ν such that $c^{\nu}_{\lambda\mu} > 0$ there exists a polynomial $P^{\nu}_{\lambda\mu}(t)$ in t such that $P^{\nu}_{\lambda\mu}(t) = c^{t\nu}_{t\lambda,t\mu}$ for all $t \in \mathbb{N}$.*

As a special case of these conditions we have the following conjecture now established as a theorem:

Theorem 2.6 (Fulton's Conjecture). [KTW] *For all positive integers t*

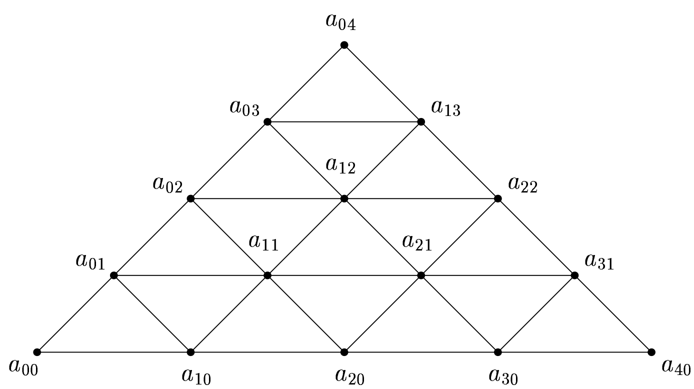
$$c^{t\nu}_{t\lambda,t\mu} = 1 \iff c^{\nu}_{\lambda\mu} = 1. \tag{2.6}$$

Thus, if $c^{\nu}_{\lambda\mu} = 1$, the corresponding polynomial $P^{\nu}_{\lambda\mu}(t) = 1$.

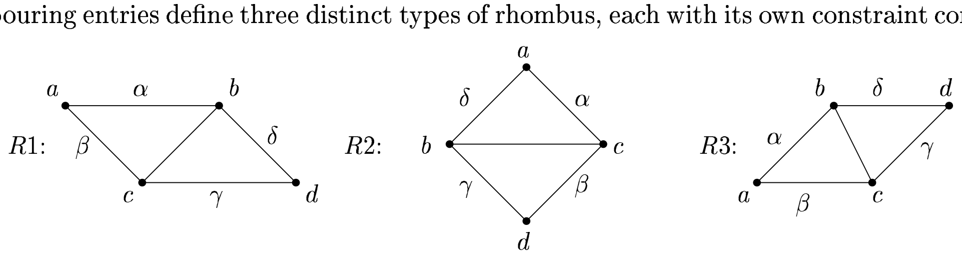
3. The hive model

The hive model arose out of the triangular arrays of Berenstein and Zeleviinsky used to specify individual contributions to Littlewood-Richardson coefficients [BZ2]. The model was then taken up by Knutsen and Tao in a manner described in an exposition by Buch [B].

An n-hive is an array of numbers a_{ij} with $0 \le i, j, i+j \le n$ placed at the vertices of an equilateral triangular graph. Typically, for $n = 4$ their arrangement is as shown below:



Such an n-hive is said to be an integer hive if all of its entries are non-negative integers. Neighbouring entries define three distinct types of rhombus, each with its own constraint condition.



In each case, with the labelling as shown, the hive condition takes the form:

$$b + c \geq a + d \tag{3.1}$$

In what follows we make use of edge labels more often than vertex labels. Each edge in the hive is labelled by means of the difference, $\epsilon = q - p$, between the labels, p and q, on the two vertices connected by this edge, with q always to the right of p. In all the above cases, with this convention, we have $\alpha + \delta = \beta + \gamma$, and the hive conditions take the form:

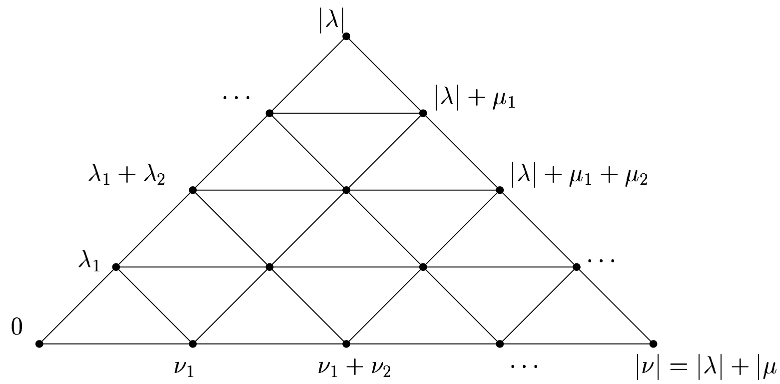
$$\alpha \geq \gamma \quad \text{and} \quad \beta \geq \delta, \tag{3.2}$$

where, of course, either one of the conditions $\alpha \geq \gamma$ or $\beta \geq \delta$ is sufficient to imply the other.

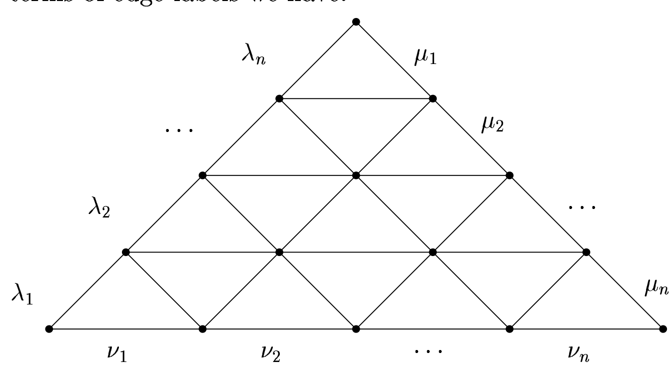
In order to enumerate contributions to Littlewood-Richardson coefficients, we require the following:

Definition 3.1. *An LR-hive is an integer n-hive, for some positive integer n, satisfying the hive conditions (3.1), or equivalently (3.2) for all its constituent rhombi of type R1, R2 and R3, with border labels determined by partitions λ, μ and ν, for which $\ell(\lambda), \ell(\mu), \ell(\nu) \leq n$ and $|\lambda| + |\mu| = |\nu|$, in such a way that $a_{00} = 0$, $a_{0,i} = ps\,(\lambda)_i$, $a_{j,n-j} = |\lambda| + ps\,(\mu)_j$ and $a_{k,0} = ps\,(\nu)_k$, for $i, j, k = 1, 2, \ldots, n$,*

Schematically, we have



Alternatively, in terms of edge labels we have:

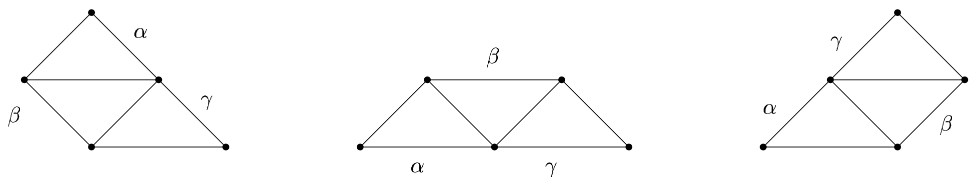


Proposition 3.2. [B] *The Littlewood-Richardson coefficient $c^{\nu}_{\lambda\mu}$ is the number of LR-hives with border labels determined as above by λ, μ and ν.*

Proof There exists a bijection between Littlewood-Richardson diagrams, D, of shape determined by ν/λ and of weight μ and LR-hives, H with border labels specified by λ, μ and ν. An illustration of this bijection is given below for a typical Littlewood-Richardson diagram, D, in the case $n = 3$, $\lambda = (3, 2)$, $\mu = (2, 1)$ and $\nu = (4, 3, 1)$. In D, which has overall shape ν, the portion of shape λ has been signified by entries 0, while the other entries correspond to the parts of the weight μ arranged in accordance with the Littlewood-Richardson rules. The first step is to form a sort of generalised Gelfand-Zetlin pattern G by writing down a list of partitions describing the shapes of subdiagrams of D formed by restricting the entries to be no more than k for $k = 3, 2, 1, 0$. Then one adds a diagonal of zeros and forms cumulative sums to arrive at an array Z. The lower right triangular portion of Z is then reoriented to give an LR-hive H, where for display purposes the hive edges have been omitted.

$$
\begin{aligned}
&D = \begin{array}{|c|c|c|c|}\hline 0&0&0&1\\\hline 0&0&2&\multicolumn{1}{c}{}\\\cline{1-3} 1&\multicolumn{1}{c}{}&\multicolumn{1}{c}{}&\multicolumn{1}{c}{}\\\cline{1-1}\end{array}
\quad\Longleftrightarrow\quad G = \begin{array}{cccccccc}4&&3&&1&&&\\&4&&3&&1&&\\&&4&&2&&1&\\&&&3&&2&&0\end{array}\\
&\Longleftrightarrow\quad Z = \begin{array}{cccccccccc}0&&4&&7&&8&&&\\&0&&4&&7&&8&&\\&&0&&4&&6&&7&\\&&&0&&3&&5&&5\end{array}
\quad\Longleftrightarrow\quad H = \begin{array}{ccccccc}&&&5&&&\\&&5&&7&&\\&3&&6&&8&\\0&&4&&7&&8\end{array}
\end{aligned}
\tag{3.3}
$$

When expressed in terms of edge labels, the hive conditions (3.2) for all constituent rhombi of types R1, R2 and R3 imply that in every LR-hive the edge labels along any line parallel to the north-west, north-east and southern boundaries of the hive are weakly decreasing in the north-east, south-west and easterly directions, respectively. This can be seen from the following 5-vertex sub-diagrams.



The edge conditions on the overlapping pairs of rhombi (R1,R2), (R1,R3) and (R2,R3) in the above diagrams give in each case $\alpha \geq \beta$ and $\beta \geq \gamma$, so that $\alpha \geq \gamma$ as claimed. This is of course consistent with the fact that edges of the three north-west, north-east and southern boundaries of each LR-hive are specified by partitions λ, μ and ν, respectively.

4. Factorisation

It was noted in the work of Berenstein and Zelevinsky [BZ1] that some Kostka coefficients may factorise. Although rather easy to prove using semistandard tableaux, this factorisation property may be established through the use of K-hives [KTT1]. A full account of this is presented here in the poster session [KTT2]. The same methods may then be used to show that some Littlewood-Richardson coefficients may also factorise.

In order to state a conjecture for the precise conditions under which $c^{\nu}_{\lambda\mu}$ factorises it is convenient to introduce some further notation. As usual let n be a fixed positive integer and let $N = \{1, 2, \ldots, n\}$. Then for any $I = \{i_1, i_2, \ldots, i_r\} \subseteq N$, with $i_1 < i_2 < \cdots < i_r$ and $1 \leq r \leq n$, let $\tilde{I} = N \backslash I$ be the complement of I in N. In addition, for any partition or weight $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_n)$ let $\kappa_I = (\kappa_{i_r}, \ldots, \kappa_{i_2}, \kappa_{i_1})$. With this notation we make the following:

Conjecture 4.1. *Let λ, μ and ν be partitions of lengths $\ell(\lambda), \ell(mu), \ell(\nu) \leq n$. If $c^{\nu}_{\lambda\beta} > 0$ and there exists proper subsets $I, J, K \subset N$ with $(I, J, K) \in R^n_r$ such that $ps\,(\mu)_K = ps\,(\lambda)_I + ps\,(\mu)_J$ then*

$$
c^{\nu}_{\lambda\mu} = c^{\nu_K}_{\lambda_I \mu_J}\; c^{\nu_{\tilde{K}}}_{\lambda_{\tilde{I}} \mu_{\tilde{J}}}. \tag{4.1}
$$

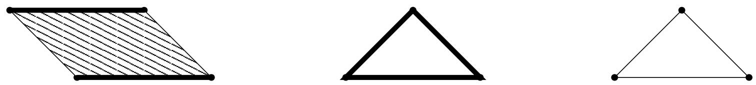
This means that if any one of Horn's inequalities (2.4) is an equality for $1 \leq r < n$ then $c^{\nu}_{\lambda\mu}$ factorises. We say that a Littlewood-Richardson coefficient $c^{\nu}_{\lambda\mu}$ is *primitive* if it cannot be factorised, that is to say all Horn's inequalities (2.4) are strict inequalities. Repeated use of the above conjecture would allow any non-vanishing Littlewood-Richardson coefficient to be written as a product of primitive Littlewood-Richardson coefficients. Furthermore, since the partial sum conditions are preserved under scaling by any positive number t, we have the following:

Conjecture 4.2. *Let λ, μ and ν be partitions of lengths $\ell(\lambda), \ell(mu), \ell(\nu) \leq n$. If $c^{\nu}_{\lambda\beta} > 0$ and there exists proper subsets I, J and K of N with $(I, J, K) \in R^n_r$ such that $ps\,(\mu)_K = ps\,(\lambda)_I + ps\,(\mu)_J$*

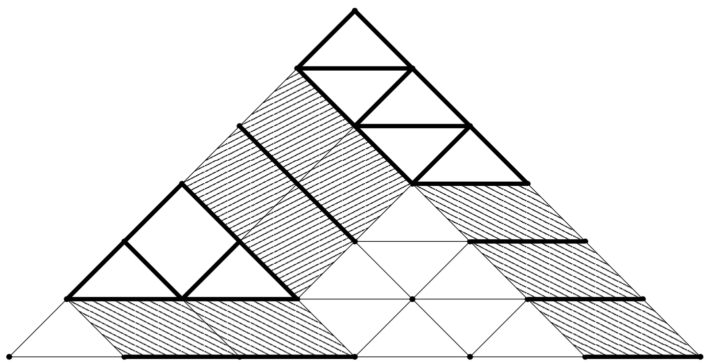
then

$$P^{\nu}_{\lambda\mu}(t) = P^{\nu_K}_{\lambda_I\mu_J}(t)\ P^{\nu_{\bar K}}_{\lambda_{\bar I}\mu_{\bar J}}(t). \tag{4.2}$$

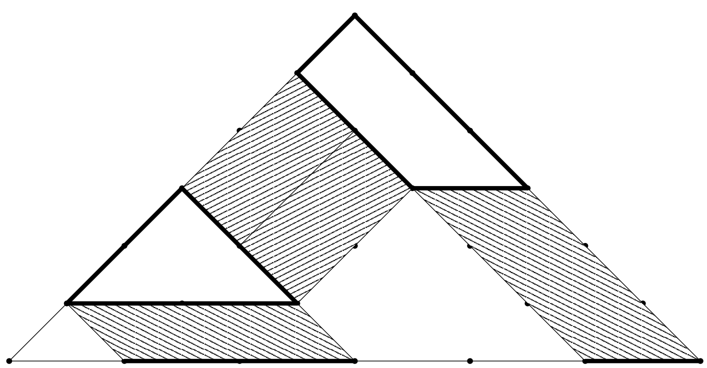
The origin of these conjectures is a study of the properties of certain puzzles introduced by Knutson *at al* [KTW]. These are triangular diagrams on a hive lattice consisting of three elementary pieces: a dark triangle, a light triangle and a shaded rhombus with its edges either dark or light according as they are to the right or left, respectively of an acute angle of the rhombus, when viewed from its interior:



The puzzle is to put these together, oriented in any manner, so as to form a hive shape with all the edges matching. For example, one such puzzle takes the form shown below:

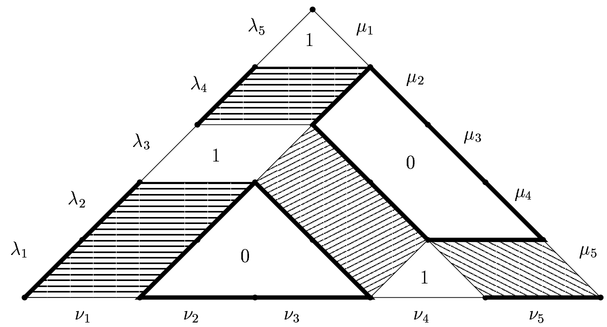


As pointed out by Danilov and Koshevoy [DK], this can be simplified, without loss of information, to give a labyrinth or hive plan by deleting all interior edges of the three types of region: shaded corridors in the form of parallelograms consisting of rhombi of just one type, either R1, or R2 or R3, and dark rooms and light rooms that are convex polygons consisting solely of just dark triangles and just light triangles, respectively.



It is a remarkable fact [KTW] that for each positive integer $r \leq n$ and triple $(I, J, K) \in R^n_r$, there exists a unique puzzle, and correspondingly a unique hive plan of the above type. In this hive plan the dark edges on the boundary are those specified by I, J and K. In connection with the above Conjecture 4.1, the thick edges on the boundary of each LR-hive are then labelled by the parts of λ_I, μ_J andν_K, and the thin edges by the parts of $\lambda_{\bar I}$, $\mu_{\bar J}$ and $\nu_{\bar k}$.

To take a different example with $n = 5$, $r = 3$, $I = \{1, 2, 4\}$, $J = \{2, 3, 4\}$, $K = \{2, 3, 5\}$, we have $\alpha(I) = (1, 0, 0)$, $\beta(J) = (1, 1, 1)$ and $\gamma(K) = (2, 1, 1)$. The fact that $(I, J, K) \in R^5_3$ then follows from the observation that $c^{211}_{1,111} = 1$. Superposing the corresponding hive plan on the LR-hives with boundaries specified by λ, μ and ν then gives



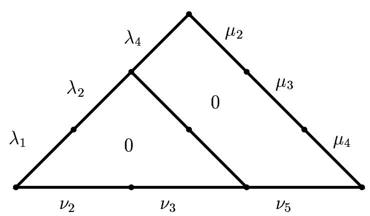
The validity of Conjecture 4.1 would imply that if

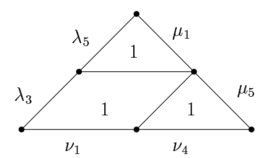
$$\nu_2 + \nu_3 + \nu_5 = \lambda_1 + \lambda_2 + \lambda_4 + \mu_2 + \mu_3 + \mu_4. \tag{4.3}$$

then $c^{\nu}_{\lambda\mu}$ must factorise as follows

$$c^{\nu}_{\lambda\mu} = c^{(\nu_2,\nu_3,\nu_5)}_{(\lambda_1,\lambda_2,\lambda_4),(\mu_2,\mu_3,\mu_4)}\; c^{(\nu_1,\nu_4)}_{(\lambda_3,\lambda_5),(\mu_1,\mu_5)} \tag{4.4}$$

To see how this comes about one just deletes the corridors from the initial LR-hive and glues together all the dark rooms, labelled 0, and all the light rooms, labelled 1, to create two smaller LR-hives, as shown below:





To prove the validity of such a factorisation, one has to show that the equality (4.3) leads to a bijection between all the large LR-hives and all pairs of small LR-hives obtained by the deletion and glueing processes. The proof is a matter of confirming that in such a case all the relevant hive conditions are satisfied not only after carrying out the deletion and glueing, but also after carrying out their inverses, cutting a pair of small LR-hives and inserting corridors to create a large LR-hive. A procedure for confirming this will be discussed in the accompanying talk.

An explicit illustration of the result is provided by the case $\lambda = (7,5,3,0,0)$, $\mu = (7,4,2,0,0)$ and $\nu = (8,8,8,2,2)$ for which $\lambda_I = (7,5,0)$, $\mu_J = (4,2,0)$ and $\nu_K = (8,8,2)$ with $\lambda_{\tilde{I}} = (3,0)$, $\mu_{\tilde{J}} = (7,0)$ and $\nu_{\tilde{K}} = (8,2)$. In this case

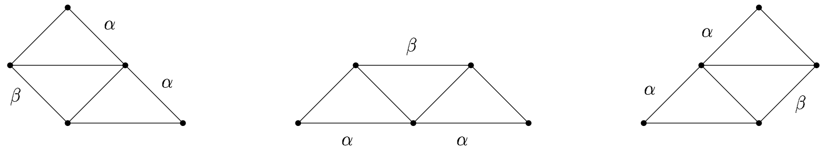
$$c^{88822}_{753,742} = c^{882}_{75,42}\; c^{82}_{3,7} = 1 \cdot 1 = 1. \tag{4.5}$$

Although this trivial result may not appear very exciting, we shall see that it has some important features. First of all the Littlewood-Richardson coeficient has turned out to be 1. It follows from Fulton's Conjecture that the corresponding stretched LR-polynomial is also just 1, that is to say its degree is 0. The question is, could this have been predicted, and what more generally can one say about the degrees of such polynomials? It is in this setting that the above factorisation is important.

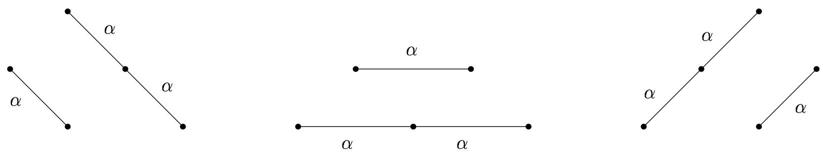
5. Degrees of stretched polynomials

Even given a certain amount of factorisation that reduces the evaluation of stretched Littlewood-Richardson polynomials to that of calculating these LR-polynomials in the primitive case, their evaluation may be combinatorially formidable. In any given case a knowledge of the degree of the polynomial would be extremely advantageous. Here we establish an upper bound on this degree by means of the following rather innocuous looking obervation.

Taking $\alpha = \gamma$ in each of the 5-vertex diagrams encountered earlier, gives



In each case the rhombus constraints give $\alpha \geq \beta \geq \alpha$ so that we must have $\beta = \alpha$. This result can be displayed more simply by suppressing all the labels on the vertices of the hives and inserting an edge between pairs of vertices whose labels differ by the same integer α. This gives the diagrams:

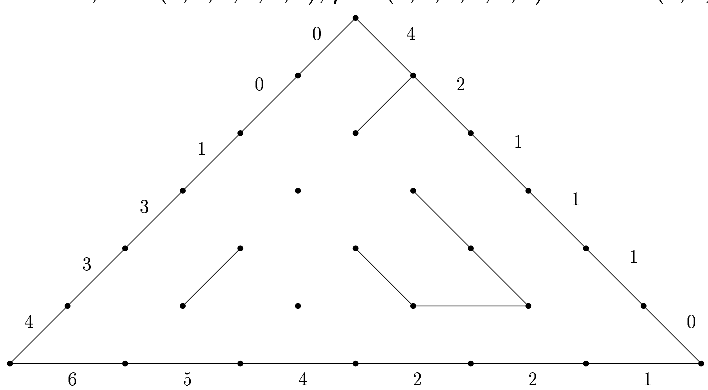


where in each case the equality of neighbouring differences α in a linear sequence of three vertices forces an identical difference α between two vertices in the neighbouring line.

Applying these notions to our LR-hives with boundaries of length n and with border labels determined by λ, μ and ν, it follows from the above that any equalities of successive parts of these partitions propagate as equalities of differences in hive entries within each LR-hive. To be more precise let all the λ-boundary edges be labelled by the parts of λ. If any sequence of parts of λ share the same value, say α, then we can identify an equilateral sub-hive having the sequence of equally labelled edges as one boundary, with its other boundaries parallel to the μ and ν-boundaries of the original hive. Within this sub-hive all the vertices along lines parallel to the λ-boundary are to be connected by edges indicating that in any LR-hive the differences in values between neighbouring entries along these lines are all α.

This process is to be repeated first for all sequences of equal edge labels along the λ-boundary, and then for all sequences of equal edge labels along the μ and ν boundaries. Finally, all neighbouring vertices on all three boundaries are to be connected by edges. In this way we arrive at a skeletal graph $G_{n;\lambda\mu\nu}$ of the hive.

For example, for $n = 6$, $\lambda = (4, 3, 3, 1, 0, 0)$, $\mu = (4, 2, 1, 1, 1, 0)$ and $\nu = (6, 5, 4, 2, 2, 1)$ we have



The important of such skeletal graphs is that they indicates constraints on LR-hive entries that are implied by the specification of the boundary labels. These constraints on the interior vertex labels reduce the total number of degrees of freedom of such labels. This leads to the following:

Proposition 5.1. *Let λ, μ and ν be partitions such that $c^{\nu}_{\lambda\mu} > 0$. Let $\deg(P(t))$ be the degree of the corresponding stretched LR-polynomial $P(t) = c^{t\nu}_{t\lambda,t\mu}$. Let $d(G_{n;\lambda\mu\nu})$ be the number of connected interior components of the graph $G_{n;\lambda\mu\nu}$ that are not connected to the boundary. Then*

$$\deg(P(t)) \leq d(G_{n;\lambda\mu\nu}). \tag{5.1}$$

Proof The application of the stretching parameter t leaves $G_{n;\lambda\mu\nu}$ unaltered, so that the number of degrees of freedom in assigning entries to the stretched LR-hives is still $d(G_{n;\lambda\mu\nu})$. For each interior connected component that is not connected to the boundary we can select any one convenient vertex. The value a_{ij} of each such selected interior vertex label may or may not be fixed by the hive constraints. However, it will be subject to linear inequalities of the form $p \le a_{ij} \le q$ arising from the hive conditions. As the boundary vertex and edge labels are scaled by t, then all the parameters specifying these linear inequalities are also scaled by t to give $tp \le a_{ij} \le tq$. Hence, in enumerating all possible LR-hives in the stretched case, the freedom in assigning a_{ij} gives rise to a contribution to $P_{\lambda\mu}(t)$ that is at most linear in t. It follows that the degree of this polynomial is at most $d(G_{n;\lambda\mu\nu})$. □

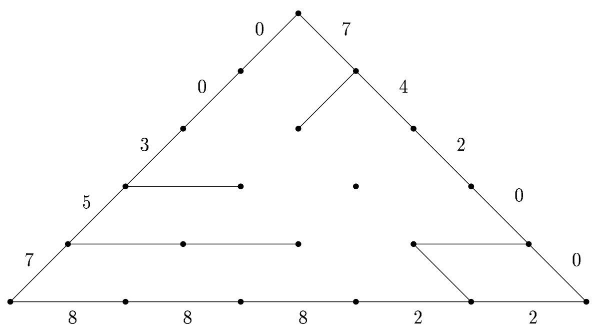
Unfortunately, the interior edges arising from sequences of equal parts in λ, μ and ν may intersect at common vertices. This makes it difficult to arrive at a formula for $d(G_{n;\lambda\mu\nu})$.

In the above example for $n = 6$ with $\lambda = (4,3,3,1,0,0)$, $\mu = (4,2,1,1,1,0)$ and $\nu = (6,5,4,2,2,1)$ we have $d(G_{n;\lambda\mu\nu}) = 4$ and the corresponding polynomial is given by

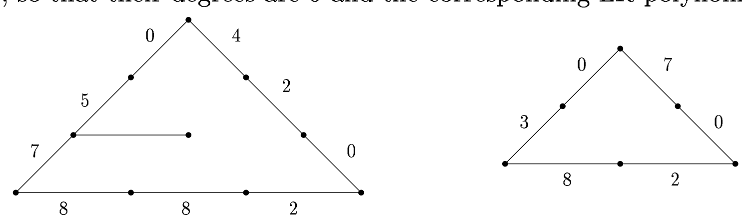
$$P^{\nu}_{\lambda\mu}(t) = \frac{1}{24}(t+1)(t+2)(t+3)(t+4), \tag{5.2}$$

so that in this case the above bound on the polynomial degree is staurated.

On the otherhand, for our previous example in the case $n = 5$ and $\lambda = (7,5,3,0,0)$, $\mu = (7,4,2,0,0)$ and $\nu = (8,8,8,2,2)$, the graph $G_{n;\lambda\mu\nu}$ takes the form:



As can be seen there is just one interior vertex that is not connected by means of edges to the boundary. It follows that $d(G_{n;\lambda\mu\nu}) = 1$. However, in this case we know that $c^{\nu}_{\lambda\mu} = 1$ so that by Fulton's Conjecture we have $P^{\nu}_{\lambda\mu}(t) = 1$and the degree of the LR-polynomial is 0. Thus the above degree bound is not saturated. The explanation for this can be seen in the skeletal diagrams of the factors arising in this non-primitive case. In these skeletal diagrams all vertices are connected to the boundary, so that their degrees are 0 and the corresponding LR-polynomials are both 1.



Encourgaged by these results and many other examples we conjecture that in the primitive case the bound in Proposition 5.1 is saturated, that is we have:

Conjecture 5.2. *If $c^{\nu}_{\lambda\mu}$ is primitive and $P(t) = c^{t\nu}_{t\lambda,t\mu}$ then*

$$\deg(P(t)) = d(G_{n;\lambda\mu\nu}), \tag{5.3}$$

and conversely, if $\deg(P(t)) < d(G_{n;\lambda\mu\nu})$ then $c^{\nu}_{\lambda\mu}$ is not-primitive.

6. Linear factors

It will have been noted that in our illustrative examples (1.3) and (5.2) the stretched Littlwood-Richardson polynomials $P^{\nu}_{\lambda\mu}(t)$ contain factors $(t+m)$ for some sequence of values $m = 1, 2, \ldots, M$ for some positive integer M. This is no accident since $P^{\nu}_{\lambda\mu}(t)$ is nothing other than an Ehrhart quasi-polynomial $i(\mathcal{P}, t)$ of a rational complex polytope $\mathcal{P}$ defined by the set of linear inequalities corresponding to the LR-hive conditions. This quasi-polynomial is actually a polynomial, but whether this is the case or not, the reciprocity theorem for Ehrhart quasi-polynomials [S] states that $i(\mathcal{P},t)$ is defined for all integers t and that for $t=-m$ with m a positive integer $i(\mathcal{P},-m) = (-1)^d\,\bar{i}(\mathcal{P},m)$, where d is the dimension of the polytope $\mathcal{P}$, and $\bar{i}(\mathcal{P},m)$ is the number of integer points inside $m\mathcal{P}$. This number of integer points may be zero, thereby giving rise to a zero of $P(t)=i(\mathcal{P},t)$ at $t=-m$. Moreover, if this number is zero for $m=M$ and non-zero for $m=M+1$ it follows from its geometric interpretation that it is zero for $m=1,2,\ldots,M$ and non-zero for all $m>M$. In such a case $P^{\nu}_{\lambda\mu}(t)$ necessarily contains $(t+1)(t+2)\cdots(t+M)$ as a factor.

In this section we describe one particular approach to the determination of M, based on a conjecture regarding the continuation of $P^{\nu}_{\lambda\mu}(t)$ to negative integer values $t=-m$. For t a positive integer, we certainly have

$$s_{t\lambda}(x_1,x_2,\ldots,x_n) = \frac{\left|x_i^{t\lambda_j+n-j}\right|}{\left|x_i^{n-j}\right|}. \tag{6.1}$$

This may be readily extended to the case $t=-m$ with m a positive integer, to give

$$s_{-m\lambda}(x_1,x_2,\ldots,x_n) = \frac{\left|x_i^{-m\lambda_j+n-j}\right|}{\left|x_i^{n-j}\right|} = \frac{\left|x_i^{-m\lambda_{n-k+1}-n+k}\right|}{\left|x_i^{-n+k}\right|}, \tag{6.2}$$

where first x_i^{n-1} has been extracted as a common factor from the ith row of each determinant for $i=1,2,\ldots,n$ and cancelled from numerator and denominator, and then j replaced by $k=n-j+1$ with an appropriate reversal of order of the columns in both determinants. If we now set $\overline{x}_i = x_i^{-1}$ for $i=1,2,\ldots,n$. this gives

$$s_{-m\lambda}(x_1,x_2,\ldots,x_n) = \frac{\left|\overline{x}_i^{m\lambda_{n-k+1}+n-k}\right|}{\left|\overline{x}_i^{n-k}\right|} = s_{m\lambda_n,\ldots,m\lambda_2,m\lambda_1}(\overline{x}_1,\overline{x}_2,\ldots,\overline{x}_n). \tag{6.3}$$

To simplify the notation, for any partition $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)$ let $\tilde{\lambda}=(\lambda_n,\ldots,\lambda_2,\lambda_1)$ be the vector obtained by reversing the order of its parts. Then reverting to the indeterminates $\mathbf{x} = (x_1,x_2,\ldots,x_n)$ we have:

$$s_{m\tilde{\lambda}}(\mathbf{x}) = s_{m\lambda_n,\ldots,m\lambda_2,m\lambda_1}(x_1,x_2,\ldots,x_n) = \frac{\left|x_i^{m\lambda_{n-k+1}+n-k}\right|}{\left|x_i^{n-k}\right|}. \tag{6.4}$$

This allows us to make the following definition, which amounts to an extension of stretched Littlewood-Richardson coefficients to the domain of a negative stretching parameter $t=-m$:

Definition 6.1. *For any λ, μ and ν such that $c^{\nu}_{\lambda\mu}>0$, let $c^{-m\nu}_{-m\lambda,-m\mu} = c^{m\tilde{\nu}}_{m\tilde{\lambda},m\tilde{\mu}}$, for any positive integer m, where*

$$s_{m\tilde{\lambda}}(\mathbf{x})\ s_{m\tilde{\mu}}(\mathbf{x}) = \sum_{\nu} c^{m\tilde{\nu}}_{m\tilde{\lambda},m\tilde{\mu}}\ s_{m\tilde{\nu}}(\mathbf{x}). \tag{6.5}$$

With this definition, the consideration of numerous examples, suggests the validity of the following:

Conjecture 6.2. *Let $c^{\nu}_{\lambda\mu}>0$ be primitive and let the corresponding LR-polynomial be $P^{\nu}_{\lambda\mu}(t)$. Then the value of this LR-polynomial at negative integer values $t=-m$ coincides with the corresponding negatively stretched Littlewood-Richardson coefficients, that is to say*

$$P^{\nu}_{\lambda\mu}(-m) = c^{-m\nu}_{-m\lambda,-m\mu}. \tag{6.6}$$

To exploit this it is necessary that Schur functions such as $s_{m\tilde{\lambda}}(\mathbf{x})$, as defined by (6.4), be standardised. This may be carried out by reordering the columns of the numerator determinant. However there are two quite different possible outcomes: either $s_{m\tilde{\lambda}}(\mathbf{x})=0$ or $s_{m\tilde{\lambda}}(\mathbf{x}) = \eta_\rho\, s_\rho(\mathbf{x})$ for some partition ρ with $\eta_\rho=\pm1$. Similar results apply to $s_{m\tilde{\mu}}(\mathbf{x})$ and $s_{m\tilde{\nu}}(\mathbf{x})$. The validity of the above conjecture would then imply that

$$P^{\nu}_{\lambda\mu}(-m) = \eta^{\rho}_{\lambda\mu} c^{\tau}_{\rho\sigma} \tag{6.7}$$

where $\eta^\nu_{\lambda\mu} = 0$ if any one of $s_{m\tilde\lambda}(\mathbf{x})$, $s_{m\tilde\mu}(\mathbf{x})$ or $s_{m\tilde\nu}(\mathbf{x})$ is identically zero, and is ± 1 in all other cases, while ρ, σ and τ are defined by the identities $s_{m\tilde\lambda}(\mathbf{x}) = \eta_\rho\, s_\rho(\mathbf{x})$, $s_{m\tilde\mu}(\mathbf{x}) = \eta_\sigma\, s_\sigma(\mathbf{x})$ and $s_{m\tilde\nu}(\mathbf{x}) = \eta_\tau\, s_\tau(\mathbf{x})$.

It follows that we can expect two types of zero of $P^\nu_{\lambda\mu}(t)$ for $t = -m$: type (i) associated with $\eta^\nu_{\lambda\mu} = 0$ and and type (ii) associated with the vanishing of $c^\tau_{\rho\sigma}$.

To see this in an example consider the case $n = 7$, $\lambda = (4, 3, 3, 2, 1)$, $\mu = (4, 3, 2, 2, 1)$ and $\nu = (7, 4, 4, 4, 3, 2, 1)$ for which $P(t) = P^\nu_{\lambda\mu}(t)$ has already been given in (1.3). Here we find that $s_{m\tilde\lambda} = 0$ for $m = 1, 2$, $s_{m\tilde\mu}(\mathbf{x}) = 0$ for $m = 1, 2$ and $s_{m\tilde\nu}(\mathbf{x}) = 0$ for $m = 1, 2, 3$. This accounts for the three zeros associated with the factors $(t + 1)$, $(t + 2)$ and $(t + 3)$. For all $m \geq 4$ we have $s_{m\tilde\lambda}(\mathbf{x}) = s_\rho(\mathbf{x})$, $s_{m\tilde\mu}(\mathbf{x}) = s_\sigma(\mathbf{x})$ and $s_{m\tilde\nu}(\mathbf{x}) = s_\tau(\mathbf{x})$ for some partitions ρ, σ and τ. It then remains to be seen whether or not $c^\tau_{\rho\sigma} = 0$.

Starting with $m = 4$ it is found that $\rho = (10, 9, 9, 8, 6, 5, 5)$, $\sigma = (10, 8, 7, 7, 6, 5, 5)$ and $\tau = (22, 14, 14, 14, 14, 12, 10)$. Since $\rho_1 + \sigma_1 = 20 < 22 = \tau_1$ it follows that $c^\tau_{\rho\sigma} = 0$. This accounts for a factor of $(t + 4)$ in $P(t)$. Similarly with $m = 5$ it is found that $\rho = (14, 12, 12, 10, 7, 5, 5)$, $\sigma = (14, 12, 9, 9, 7, 5, 5)$ and $\tau = (19, 18, 18, 18, 17, 14, 11)$. This time since $\rho_1 + \sigma_1 = 28 < 29 = \tau_1$ it again follows that $c^\tau_{\rho\sigma} = 0$, thereby accounting for a factor of $(t + 5)$ in $P(t)$. On the other hand for $m = 6$ it is found that $\rho = (18, 15, 15, 12, 8, 5, 5)$, $\sigma = (18, 14, 11, 11, 8, 5, 5)$ and $\tau = (36, 22, 22, 22, 20, 16, 12)$. This time it is found that $c^\tau_{\rho\sigma} = 3$. This implies that there is no factor $(t + 6)$ in $P(t)$. Indeed it is easy to check from (1.3) that $P(-6) = 3$. In the same way we can derive the fact that $P(-7) = 39$ and $P(-8) = 247$, in perfect agreement with (1.3).

Thus, as a corollary of Conjecture 6.2, we are lead by virtue of (6.7) to:

Conjecture 6.3. *If $c_{\lambda\mu} > 0$ is primitive, then the LR-polynomial $P^\nu_{\lambda\mu}(t) = c^{t\nu}_{t\lambda,t\mu}$ contains a factor $(t + m)$ if and only if either $\eta^\nu_{\lambda\mu} = 0$ or $c^\tau_{\rho\sigma} = 0$.*

It can be shown that for sufficiently large m both $\eta^\nu_{\lambda\mu}$ and $c^\tau_{\rho\sigma}$ in (6.7) are positive, provided that the original $c^\nu_{\lambda\mu}$ is primitive. Moreover, it can be shown in such a case that if $\eta^\nu_{\lambda\mu} c^\tau_{\rho\sigma} = 0$ for some psoitive integer m, then the same is true for all smaller positive integers. This means that it is possible to identify M such that the right hand side of (6.7) is zero for all $m \leq M$ and non-zero for all $m > M$. This is entirely, consistent with the remarks made at the beginning of this section regarding the zeros of Ehrhart quasi-polynomials.

As a final conjecture we offer

Conjecture 6.4. *Let λ, μ and ν be partitions such that $c^\nu_{\lambda\mu} > 0$. Then $c^\nu_{\lambda\mu}$ is primitive if and only if $c^{-m\nu}_{-m\lambda,-m\mu}$ is non-zero for some positive integer m.*

References

[BZ1] A.D. Berenstein and A.V. Zelevinsky, *When is the multiplicity of a weight equal to 1?* Funct. Anal. Appl. **24** (1990) 259–269.

[BZ2] A.D. Berenstein and A.V. Zelevinsky, *Triple multiplicities for $sl(r+1)$ and the spectrum of the exterior algebra of the adjoint representation* J. Alg. Comb. **1** (1992) 7-22.

[B] A.S. Buch, *The saturation conjecture (after A. Knutson and T. Tao). With an appendix by W. Fulton* Enseign. Math. **46** (2000) 43–60.

[DK] V.I. Danilov and G.A. Koshevoy, *Diskretnaya vognutost' i ermitovy matritsy* Trudy Math. Inst. Steklova **241** (2003) 68–90, in Russian. Title translation: *Discrete convexity and Hermitian matrices.*

[DW1] H. Derksen and J. Weyman, *Semi-invariants of quivers and saturation for Littlewood-Richardson coefficients* J. Amer. Math. Soc. **13** (2000) 467–479.

[DW2] H. Derksen and J. Weyman, *On the Littlewood-Richardson plynomials* J. Algebra. **255** (2002) 247–257.

[H] A. Horn, *Eigenvalues of sums of Hermitian matrices* Pacific J. Math. **12** (1962) 225–241.

[F] W. Fulton, *Eigenvalues, invariant factors, highest weights, and Schubert calculus*, Bull. Amer. Math. Soc. **37** (2000) 209–249.

[KTT1] R.C. King, C. Tollu and F. Toumazet, *Stretched Littlewood-Richardson and Kostka coefficients* CRM Proceedings and Lecture Notes, Vol. 34, Amer. Math. Soc, 2004, pp99–112.

[KTT2] R.C. King, C. Tollu and F. Toumazet, *The hive model and the factorisation of Kostka coefficients* To be presented at SSPCM'05, 2005.

[K] A.A.Klyachko, *Stable bundles, representation theory and Hermitian operators* Selecta Math. New Ser. **4** (1998) 419–445.

[KT] A. Knutson and T. Tao, *The honeycomb model of $GL_n(\mathbb{C})$ tensor products. I: Proof of the saturation conjecture*, J. Amer. Math. Soc. **12** (1999) 1055–1090.

[KTW] A. Knutson, T. Tao and C. Woodward, *The honeycomb model of $GL_n(\mathbb{C})$ tensor products. II: Puzzles determine facets of the Littlewood-Richardson cone*, J. Amer. Math. Soc. **17** (2004) 19–48.

[LR] D.E. Littlewood and A. Richardson, *Group characters and algebra* Phil. Trans. Roy. Soc. London A **233** (1934) 99–141

[R] E. Rassart, *A polynomial prperty of Littlewood-Richardson coefficienst* J Comb. Theo. Ser. A **107** (2004) 161–179

[S] R.P. Stanley, *Enumerative Combinatorics*, Vol. I, Cambridge University Press, Cambridge, 1999.

School of Mathematics, University of Southampton, Southampton SO17 1BJ, U.K.
E-mail address: `R.C.King@maths.soton.ac.uk`

ECO's carol

Renzo Pinzani

April 28, 2005

Among the general methods usually employed in combinatorics (such as generating functions, theory of species, linear operator methods, order theory, incidence algebras, Hopf algebras, umbral calculus, and so on), I will deal with a particular one, which was born in Florence in the last decade of the past century, thanks to the work of a group of researchers in discrete mathematics and theoretical computer science. This method is usually referred to as the *ECO method* (ECO is an acronym, standing for Enumeration of Combinatorial Objects). The roots of the ECO method can be traced back to the paper [CGHK], where the authors study Baxter permutations. For the first time, a combinatorial construction is presented which can be described by a *generating tree*, as it usually happens for many ECO-construction. However, we have to wait for the master degree thesis of Alberto Del Lungo [DL], in 1992, to see the birth of the ECO method, whose first application was of a theoretical computer science nature, concerning the exhaustive generation of directed animals. However it became quickly evident that the range of applicability of such a method was much wider than suggested by a mere, specific problem of generation. In 1995 the first enumerative application of the ECO method was presented [BDLPP1], and various problems connected with the enumeration of k-coloured Motzkin paths were solved. Soon after this first result, a general methodology for plane tree enumeration was settled down [BDLPP2], and finally a general survey was published [BDLPP] where many enumerative examples are treated in great detail.

The ECO method consists basically of a way of constructing the objects of a class of combinatorial structures. Consider a class of objects $\mathcal{O}$ in which a concept of *size* is introduced. This means that a partition $\{\mathcal{O}_n \mid n \in \mathbf{N}\}$ of $\mathcal{O}$ is defined, such that $\mathcal{O}_n$ is the subset of the objects of size n. Our main problem is to determine the numerical sequence $f_n = |\mathcal{O}_n|$. The ECO method gives an answer to this question in several cases, providing a recursive construction of the objects of the class under consideration.

Let $\vartheta : \mathcal{O} \longrightarrow 2^{\mathcal{O}}$ be a function[1] such that, for any $n \in \mathbf{N}$, if $O \in \mathcal{O}_n$, then $\vartheta(O) \subseteq \mathcal{O}_{n+1}$. We say that ϑ is an *ECO operator* when the family of

[1]for a given set A, we denote by 2^A the power set of A.

sets $\{\vartheta(O) \mid O \in \mathcal{O}_n\}$ is a partition of $\mathcal{O}_{n+1}$. The ECO method consists precisely of the effective construction of an ECO operator $\vartheta = \vartheta_{\mathcal{O}}$ for the class of objects $\mathcal{O}$; it is clear that the enumeration of $\mathcal{O}$ is possible only if such a construction is sufficiently regular. Typically, the construction of an ECO operator allows to find a functional equation satisfied by the (ordinary) generating function of $\mathcal{O}$ (or a recursive relation satisfied by the sequence $f_n = |\mathcal{O}_n|$), whose solution is often provided by the application of suitable analytical tools.

The purpose of my talk is threefold. First of all, I would like to run through the main stages of ECO by showing its birth and rapid development from a historical and chronological point of view. Next I will try to give a detailed survey of what have been done with the ECO method. More precisely, I will show how ECO can be fruitfully used in enumerative and algebraic combinatorics, in bijective combinatorics and in the random and exhaustive generation of combinatorial objects. This will be done by first presenting the algebraic foundations of ECO, mainly interpreting it in a linear algebra context, and then by describing several enumerative and bijective results concerning disparate combinatorial objects, such as plane trees, lattice paths, pattern avoiding permutations, polyominoes, restricted set partitions. In particular, the generating tree associated with a given ECO construction provides a fundamental tool for finding bijections which also preserve several statistics. As far as generation is concerned, I will show how the ECO method naturally suggests a general method to randomly and exhaustively generate several classes of combinatorial objects, depending on the nature of their generating function (rational, algebraic, transcendental). Finally, I intend to provide some ideas for further work, by proposing open problems which naturally arise in this context. This suggestions will be scattered throughout the talk, each of them appearing where it seemed to me more suitable.

References

[BDLPP1] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani *A construction for enumerating k-coloured Motzkin paths*, Lecture Notes in Comput. Sci. 959 (1995) 254-263.

[BDLPP2] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani *Towards a methodology for tree enumeration* proceedings of FPSAC 1995, Marne-la Vallée, later appeared on Discrete Math. 180 (1998) 45-64.

[BDLPP] E. Barcucci, A. Del Lungo E. Pergola, R. Pinzani, *ECO: a Methodology for the Enumeration of Combinatorial Objects*, J. Differ. Equations Appl. 5 (1999) 435-490.

[CGHK] F. R. K. Chung, R. L. Graham, V. E. Hoggatt, M. Kleiman *The number of Baxter permutations*, J. Combin. Theory Ser. A 24 (1978) 382-394.

[DL] A. Del Lungo *Algoritmi per la generazione di animali direzionati*, master degree thesis, 1992.

Title: Homomorphisms, Symmetric Functions, and Permutation Enumeration.

Jeffrey. B. Remmel
Department of Mathematics
University of California, San Diego

Abstract. A large number of generating functions for permutation statistics for the symmetric group S_n, the hyperoctahedral group B_n, and wreath products of the form $G \wr S_n$ can be derived by applying a ring homomorphism on the ring of symmetric functions $\Gamma(x_1, x_2, \ldots)$ to simple symmetric function identities. This idea goes back to 1993 paper of Brenti and has been exploited in a number of papers by Beck, Langley, Wagner, Mendes and the speaker. In this talk, we will introduce some new families of bases of symmetric functions and show how one can derive some old and new generating functions for permutation statistics by applying ring homomorphisms to some simple indentities involving these new families of bases for $\Gamma(x_1, x_2, \ldots)$.

BERGMAN COMPLEXES, COXETER ARRANGEMENTS, AND GRAPH ASSOCIAHEDRA

FEDERICO ARDILA, VICTOR REINER, AND LAUREN WILLIAMS

Abstract

Tropical varieties play an important role in algebraic geometry. The Bergman complex $\mathcal{B}(M)$ and the positive Bergman complex $\mathcal{B}^+(M)$ of an oriented matroid M generalize to matroids the notions of the tropical variety and positive tropical variety associated to a linear ideal. Our main result is that if $\mathcal{A}$ is a Coxeter arrangement of type Φ with corresponding oriented matroid M_Φ, then $\mathcal{B}^+(M_\Phi)$ is dual to the graph associahedron of type Φ, and $\mathcal{B}(M_\Phi)$ equals the nested set complex of $\mathcal{A}$.

Résumé. Les variétés tropicales jouent un rôle important en géométrie algébrique. Le complexe de Bergman $\mathcal{B}(M)$ et le complexe de Bergman positif $\mathcal{B}^+(M)$ d'un matroïde orienté M étendent aux matroïdes les notions de variété tropicale et de variété tropicale positive associées à un idéal linéaire. Notre résultat principal est que si $\mathcal{A}$ est un arrangement de Coxeter de type Φ, et si M_Φ est le matroïde orienté correspondant, alors $\mathcal{B}^+(M_\Phi)$ est le dual de l'associaèdre du graphe de type Φ, et $\mathcal{B}(M_\Phi)$ est le complexe des ensembles imbriqués de $\mathcal{A}$.

1. Introduction

In this paper we study the Bergman complex and the positive Bergman complex of a Coxeter arrangement, and we relate them to the nested set complexes that arise in De Concini and Procesi's wonderful arrangement models [8, 9], and to the graph associahedra introduced by Carr and Devadoss [6], by Davis, Januszkiewicz, and Scott [7], and by Postnikov [14].

The *Bergman complex* of a matroid is a pure polyhedral complex which can be associated to any matroid. It was first defined by Sturmfels [18] in order to generalize to matroids the notion of a tropical variety associated to a linear ideal. The Bergman complex can be described in terms of the lattice of flats of the matroid, and is homotopy equivalent to a wedge of spheres, as shown by Ardila and Klivans [1].

The *positive Bergman complex* $\mathcal{B}^+(M)$ of an oriented matroid M is a subcomplex of the Bergman complex of the underlying unoriented matroid $\underline{M}$. It generalizes to oriented matroids the notion of the positive tropical variety associated to a linear ideal. $\mathcal{B}^+(M)$ depends on a choice of acyclic orientation of M, and as one varies this acyclic orientation, one gets a covering of the Bergman complex of $\underline{M}$. The positive Bergman complex can be described in terms of the Las Vergnas face lattice of M and it is homeomorphic to a sphere, as shown by Ardila, Klivans, and Williams [2].

Graph associahedra are polytopes which generalize the associahedron, which were discovered independently by Carr and Devadoss [6], by Davis, Januszkiewicz, and Scott [7], and by Postnikov [14]. There is an intrinsic tiling by associahedra of the Deligne-Knudsen-Mumford compactification of the real moduli space of curves $\overline{M_0^n(\mathbb{R})}$, a space which is related to the Coxeter complex of type A. The motivation for Carr and Devadoss' work was the desire to generalize this phenomenon to all simplicial Coxeter systems.

Let $\mathcal{A}_\Phi$ be the Coxeter arrangement corresponding to the (possibly infinite, possibly non-crystallographic) root system Φ associated to a Coxeter system (W, S) with diagram Γ; see Section 4 below. Choose a region R of the arrangement, and let M_Φ be the oriented matroid associated to $\mathcal{A}_\Phi$ and R. In this paper we prove:

Theorem 1.1. *The positive Bergman complex $\mathcal{B}^+(M_\Phi)$ of the arrangement $\mathcal{A}_\Phi$ is dual to the graph associahedron $P(\Gamma)$.*

In particular, the cellular sphere $\mathcal{B}^+(M_\Phi)$ is actually a simplicial sphere, and a *flag* (or *clique*) complex.

This result is also related to the *wonderful model of a hyperplane arrangement* and to *nested set complexes*. The wonderful model of a hyperplane arrangement is obtained by blowing up the non-normal crossings of the arrangement, leaving its complement unchanged. De Concini and Procesi [8] introduced this model in order to study the topology of this complement. They showed that the *nested sets* of the arrangement encode the underlying combinatorics. Feichtner and Kozlov [9] gave an abstract notion of the *nested set complex* for any meet-semilattice, and Feichtner and Müller [10] studied its topology. Recently, Feichtner and Sturmfels [11] studied the relation between the Bergman fan and the nested set complexes (see Section 5 below).

In this paper we also prove:

Theorem 1.2. *The Bergman complex $\mathcal{B}(M_\Phi)$ of $\mathcal{A}_\Phi$ equals its nested set complex.*

In particular, the cell complex $\mathcal{B}(M_\Phi)$ is actually a simplicial complex.

2. The Bergman complex and the positive Bergman complex

Our goal in this section is to explain the notions of the Bergman complex of a matroid and the positive Bergman complex of an oriented matroid which were studied in [1] and [2]. In order to do so we must review a certain operation on matroids and oriented matroids.

Definition. Let M be a matroid or oriented matroid of rank r on the ground set $[n]$, and let $\omega \in \mathbb{R}^n$. Regard ω as a weight function on M, so that the weight of a basis $B = \{b_1, \dots, b_r\}$ of M is given by $\omega_B = \omega_{b_1} + \omega_{b_2} + \cdots + \omega_{b_r}$. Let B_ω be the collection of bases of M having minimum ω-weight. (If M is oriented, then bases in B_ω inherit orientations from bases of M.) This collection is itself the set of bases of a matroid (or oriented matroid) which we call M_ω.

It is not obvious that M_ω is well-defined. However, when M is an unoriented matroid, we can see this by considering the matroid polytope of M: the face that minimizes the linear functional ω is precisely the matroid polytope of M_ω. For a proof that M_ω is well-defined when M is oriented, see [2].

Notice that M_ω will not change if we translate ω or scale it by a positive constant. We can therefore restrict our attention to the sphere $S^{n-2} := \{\omega \in \mathbb{R}^n \ : \ \omega_1 + \cdots + \omega_n = 0\,, \omega_1^2 + \cdots + \omega_n^2 = 1\}$. The Bergman complex of M will be a certain subset of this sphere.

The matroid M_ω depends only on a certain flag associated to ω.

Definition. Given $\omega \in \mathbb{R}^n$, let $\mathcal{F}(\omega)$ denote the unique flag of subsets

$$\emptyset = F_0 \subset F_1 \subset \cdots \subset F_k \subset F_{k+1} = [n] \tag{1}$$

such that ω is constant on each set $F_i \setminus F_{i-1}$ and satisfies $\omega|_{F_i \setminus F_{i-1}} < \omega|_{F_{i+1} \setminus F_i}$. We call $\mathcal{F}(\omega)$ the *flag* of ω, and we say that the *weight class* of ω or of the flag $\mathcal{F}$ is the set of vectors ν such that $\mathcal{F}(\nu) = \mathcal{F}$.

It is shown in [1] that M_ω depends only on the flag $\mathcal{F} := \mathcal{F}(\omega)$; specifically

$$M_\omega = \bigoplus_{i=1}^{k+1} F_i / F_{i-1} \tag{2}$$

where F_i/F_{i-1} is obtained from the matroid restriction of M to F_i by quotienting out the flat F_{i-1}. Hence we we also refer to this oriented matroid M_ω as $M_{\mathcal{F}}$.

Definition/ Theorem 2.1. [1] *The Bergman complex of a matroid M on the ground set $[n]$ is the set*

$$\begin{aligned} \mathcal{B}(M) &= \{\omega \in S^{n-2} \ : \ M_{\mathcal{F}(\omega)} \text{ has no loops}\} \\ &= \{\omega \in S^{n-2} \ : \ \mathcal{F}(\omega) \text{ is a flag of flats of } M\} \end{aligned}$$

Since the matroid M_ω depends only on the weight class that ω is in, the Bergman complex of M is the disjoint union of the weight classes of flags $\mathcal{F}$ such that $M_{\mathcal{F}}$ has no loops. We say that the weight class of a flag $\mathcal{F}$ is *valid* for M if $M_{\mathcal{F}}$ has no loops.

There are two polyhedral subdivisions of $\mathcal{B}(M)$, one of which is clearly finer than the other.

Definition. The *fine subdivision* of $\mathcal{B}(M)$ is the subdivision of $\mathcal{B}(M)$ into valid weight classes: two vectors u and v of $\mathcal{B}(M)$ are in the same class if and only if $\mathcal{F}(u) = \mathcal{F}(v)$. The *coarse subdivision* of $\mathcal{B}(M)$ is the subdivision of $\mathcal{B}(M)$ into M_ω-equivalence classes: two vectors u and v of $\mathcal{B}(M)$ are in the same class if and only if $M_u = M_v$.

The fine subdivision gives the following corollary of Theorem 2.1.

Corollary 2.2. [1] *Let M be a matroid. The fine subdivision of the Bergman complex $\mathcal{B}(M)$ is a geometric realization of $\Delta(L_M - \{\hat{0}\,, \hat{1}\,\})$, the order complex of the proper part of the lattice of flats of M. It follows that $\mathcal{B}(M)$ is homotopy equivalent to a wedge of spheres.*

There are positive analogues of all of the above definitions and theorems. First we must give the definition of *positive covectors* and *positive flats.*

Definition. Let M be an acyclic oriented matroid on the ground set $[n]$. We say that a covector $v \in \{+,-,0\}^n$ of M is *positive* if each of its entries is $+$ or 0. We say that a flat of M is *positive* if it is the 0-set of a positive covector.

Observation 2.3. *If M is the acyclic oriented matroid corresponding to a hyperplane arrangement $\mathcal{A}$ and a specified region R, then the positive flats are in correspondence with the faces of R.*

For example, consider the braid arrangement A_3, consisting of the six hyperplanes $x_i = x_j, 1 \leq i < j \leq 4$ in $\mathbb{R}^4$. Figure 1 illustrates this arrangement, when intersected with the hyperplane $x_4 = 0$ and the sphere $x_1^2 + x_2^2 + x_3^2 = 1$. Let R be the region specified by the inequalities $x_1 \geq x_2 \geq x_3 \geq x_4$, and let M_{A_3} be the oriented matroid corresponding to the arrangement A_3 and the region R. Then the positive flats are $\emptyset, 1, 4, 6, 124, 16, 456$ and 123456.

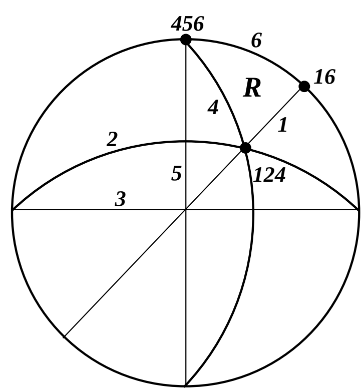


FIGURE 1. The braid arrangement A_3.

Definition/ Theorem 2.4. [2] *The positive Bergman complex of M is*

$$\begin{aligned} \mathcal{B}^+(M) &= \{\omega \in S^{n-2} \ : \ M_{\mathcal{F}(\omega)} \text{ is acyclic}\} \\ &= \{\omega \in S^{n-2} \ : \ \mathcal{F}(\omega) \text{ is a flag of positive flats of } M\} \end{aligned}$$

Within each equivalence class of the coarse subdivision of $\mathcal{B}(M)$, the vectors ω give rise to the same unoriented M_ω. Since the orientation of M_ω is inherited from that of M, they also give rise to the same oriented matroid M_ω. Therefore each coarse cell of $\mathcal{B}(M)$ is either completely contained in or disjoint from $\mathcal{B}^+(M)$. Thus $\mathcal{B}^+(M)$ inherits the coarse and the fine subdivisions from $\mathcal{B}(M)$, and each subdivision of $\mathcal{B}^+(M)$ is a subcomplex of the corresponding subdivision of $\mathcal{B}(M)$.

Recall that the *Las Vergnas face lattice* $\mathcal{F}_{\ell v}(M)$ is the lattice of positive flats of M, ordered by containment. Note that the lattice of positive flats of the oriented matroid M sits inside L_M, the lattice of flats of M. By Observation 2.3, if M is the oriented matroid of the arrangement $\mathcal{A}$ and the region R, then $\mathcal{F}_{\ell v}(M)$ is the face poset of R.

Corollary 2.5. [2] *Let M be an oriented matroid. Then the fine subdivision of $\mathcal{B}^+(M)$ is a geometric realization of $\Delta(\mathcal{F}_{\ell v}(M) - \{\hat{0}, \hat{1}\})$, the order complex of the proper part of the Las Vergnas face lattice of M. It follows that the positive Bergman complex of an oriented matroid is homeomorphic to a sphere.*

Therefore $\mathcal{B}^+(M)$ is one of the spheres in $\mathcal{B}(M)$.

Example 2.6. *Let M be the oriented matroid from Figure 1. The positive flats of M are $\{\emptyset, 1, 4, 6, 16, 124, 456, 123456\}$. The lattice of positive flats of M is shown in bold in Figure 2, within the lattice of flats of M.*

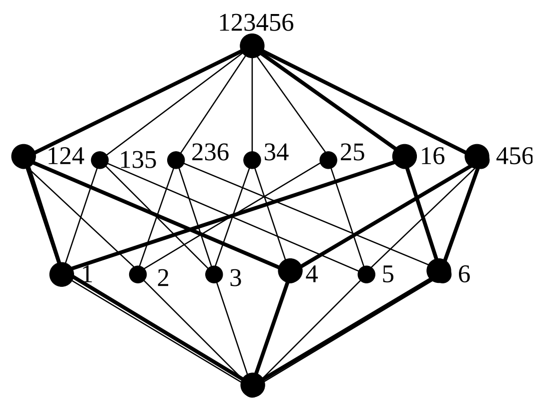


FIGURE 2. The lattice of positive flats within the lattice of flats.

We close this section with some observations about when two flags of flats in M correspond to the same cell of the coarse subdivision of $\mathcal{B}(M)$. Recall that the *connected components* of matroid M are the equivalence classes for the following equivalence relation on the ground set E of M: say $e \sim e'$ for two elements e, e' in E whenever they lie in a common circuit of M, and then take the transitive closure of $\sim$. Recall also that every connected component is a flat of M, and M decomposes (uniquely) as the direct sum of its connected components.

Definition. To each flag $\mathcal{F}$ of flats of a matroid M indexed as in (1), associate a forest $T_{\mathcal{F}}$ of rooted trees, in which each vertex v is labelled by a flat $F(v)$, as follows:

- For each connected component F of the matroid M, create a rooted tree (as specified below) and label its root vertex with F.
- For each vertex v already created, and already labelled by some flat $F(v)$ which is a connected component of some flat F_j in the flag $\mathcal{F}$, create children of v labelled by each of the connected components of F_{j-1} which are contained *properly* in $F(v)$.

Alternatively, one can construct the forest $T_{\mathcal{F}}$ by listing all the connected components of all the flats in $\mathcal{F}$, and partially ordering them by inclusion.

Proposition 2.7. *For any flag $\mathcal{F}$ of flats in a matroid M, the labelled forest $T_{\mathcal{F}}$ determines the matroid $M_{\mathcal{F}}$.*

In general, the converse of this proposition does *not* hold; one can have $M_{\mathcal{F}} = M_{\mathcal{F}'}$ without $T_{\mathcal{F}} = T_{\mathcal{F}'}$. For example (cf. [11, Example 1.2]), in the matroid M on ground set $E = \{1,2,3,4,5\}$ having rank 3 and circuits $\{123, 145, 2345\}$, the two flags

$$\mathcal{F} := (\emptyset \subset 1 \subset 123 \subset 12345)$$
$$\mathcal{F}' := (\emptyset \subset 1 \subset 145 \subset 12345).$$

exhibit this possibility.

However, we can give at least one nice hypothesis that allows one to reconstruct $T_{\mathcal{F}}$ from $M_{\mathcal{F}}$. Given a base B of a matroid M on ground set E, and any element $e \in E\backslash B$, there is a unique circuit of M contained in $B \cup \{e\}$, called the *basic circuit* $\mathrm{circ}(B, e)$. Note that the flat spanned by $\mathrm{circ}(B, e)$ will always be a connected flat.

Definition. Say that a base B of a matroid M is *circuitous* if every connected flat spanned by a subset of B is spanned by the basic circuit $\mathrm{circ}(B, e)$ for some $e \in E\backslash B$.

Note that the basic circuit $\mathrm{circ}(B, e)$ spanning the connected flat F must be $(F \cap B) \cup e$. Before we state our proposition, we give two useful lemmas.

Lemma 2.8. *Let F be a flat in a matroid, spanned by some independent set I. Then every connected component of F is spanned by some subset of I, namely, by the intersection of that component with I.*

Lemma 2.9. *Let $F \subset G$ be flats of a matroid that are spanned by subsets of a circuitous base B. If G is connected, then G/F is also connected.*

Proposition 2.10. *Let B be a circuitous base of a matroid M. Then for any two flags $\mathcal{F}, \mathcal{F}'$ of flats spanned by subsets of B, one has $M_{\mathcal{F}} = M_{\mathcal{F}'}$ if and only if $T_{\mathcal{F}} = T_{\mathcal{F}'}$.*

It will turn out that the simple roots Δ of a root system Φ always form a circuitous base for the associated matroid M_{Φ}; see Proposition 4.3(iii) below.

Remark 2.11. When the matroid M is connected, the forest $T_{\mathcal{F}}$ constructed above is a rooted tree. It coincides with the tree constructed by Feichtner and Sturmfels in [11, Proposition 3.1] when they choose the minimal building set for their lattice. In this way, Proposition 2.7 follows from [11, Theorem 4.4].

3. Graph associahedra

Graph associahedra are polytopes which generalize the associahedron, which were discovered independently by Carr and Devadoss [6], Davis, Januszkiewicz, and Scott [7], and Postnikov [14]. There is an intrinsic tiling by associahedra of the Deligne-Knudsen-Mumford compactification of the real moduli space of curves $\overline{M_0^n}(\mathbb{R})$, a space which is related to the Coxeter complex of type A. The motivation for Carr and Devadoss' work was the desire to generalize this phenomenon to all Coxeter systems.

In order to define graph associahedra, we must introduce the notions of tubes and tubings. We follow the presentation of [6].

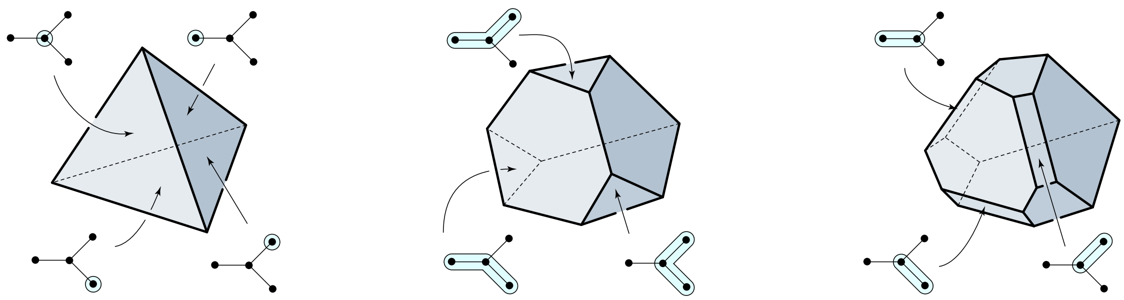


FIGURE 3. $P(D_4)$ ©Satyan Devadoss

Definition. Let Γ be a graph. A *tube* is a proper nonempty set of nodes of Γ whose induced graph is a proper, connected subgraph of Γ. There are three ways that two tubes can interact on the graph:

- Tubes are *nested* if $t_1 \subset t_2$.
- Tubes *intersect* if $t_1 \cap t_2 \neq \emptyset$ and $t_1 \not\subset t_2$ and $t_2 \not\subset t_1$.
- Tubes are *adjacent* if $t_1 \cap t_2 = \emptyset$ and $t_1 \cup t_2$ is a tube in Γ.

Tubes are *compatible* if they do not intersect and they are not adjacent. A *tubing* T of Γ is a set of tubes of Γ such that every pair of tubes in T is compatible. A *k-tubing* is a tubing with k tubes.

Graph-associahedra are defined via a construction which we will now describe.

Definition. Let Γ be a graph on n nodes. (Note that Γ need not be the graph of a Coxeter system.) Let Δ_Γ be the $n-1$ simplex in which each facet corresponds to a particular node. Note that each proper subset of nodes of Γ corresponds to a unique face of Δ_Γ, defined by the intersection of the faces associated to those nodes. The empty set corresponds to the face which is the entire polytope Δ_Γ. For a given graph Γ, truncate faces of Δ_Γ which correspond to 1-tubings in increasing order of dimension (i.e. first truncate vertices, then edges, then 2-faces, ...). The resulting polytope $P(\Gamma)$ is the *graph associahedron* of Carr and Devadoss.

Figure 3 illustrates the construction of the graph associahedron of a Coxeter diagram of type D_4. We start with a simplex, whose four facets correspond to the vertices of the diagram. In the first step, we truncate three of the vertices, to obtain the second polytope shown. We then truncate three of the edges, to obtain the third polytope shown. In the final step, we truncate the four facets which all correspond to tubes. This step is not shown in Figure 3, since it does not affect the combinatorial type of the polytope.

When the graph Γ is the n-element chain, the polytope $P(\Gamma)$ is the associahedron A_{n-1}. One can see this by considering an easy bijection between valid tubings and parenthesizations of a word of length $n-1$, as illustrated in Figure 4.

We thank Satyan Devadoss for allowing us to reproduce in our Figures 3 and 4, two of his figures from [6].

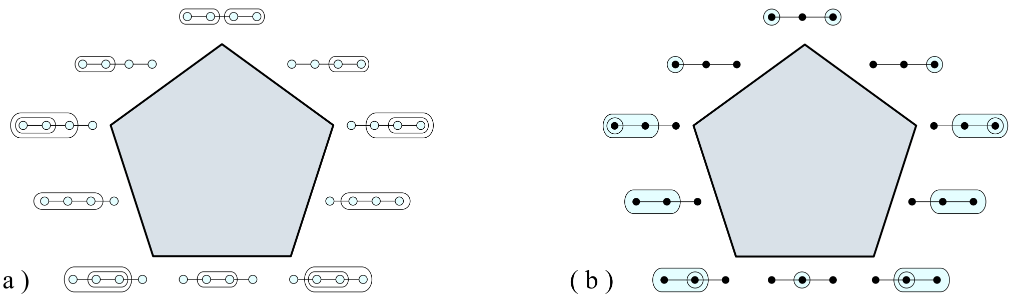


FIGURE 4. The associahedron A_2 is the graph associahedron of a 3-element chain.

Carr and Devadoss proved that the face poset of $P(\Gamma)$ can be described in terms of valid tubings.

Theorem 3.1. [6] *The face poset of* $P(\Gamma)$ *is isomorphic to the set of valid tubings of* Γ*, ordered by reverse containment:* $T < T'$ *if* T *is obtained from* T' *by adding tubes.*

Corollary 3.2. [6] *When* Γ *is a path with* $n-1$ *nodes,* $P(\Gamma)$ *is the associahedron* A_n *of dimension* n*. When* Γ *is a cycle with* $n-1$ *nodes,* $P(\Gamma)$ *is the cyclohedron* W_n*.*

4. The positive Bergman complex of a Coxeter arrangement

In this section we prove that the positive Bergman complex of a Coxeter arrangement of type Φ is dual to the graph associahedron of type Φ. More precisely, both of these objects are homeomorphic to spheres of the same dimension, and their face posets are dual. We begin by reviewing our conventions about Coxeter systems and the related arrangements and matroids.

A *Coxeter system* is a pair (W, S) consisting of a group W and a set of generators $S \subset W$, subject only to relations of the form

$$(ss')^{m(s,s')} = 1,$$

where $m(s,s) = 1$, $m(s,s') = m(s',s) \geq 2$ for $s \neq s'$ in S. In case no relation occurs for a pair s, s', we make the convention that $m(s,s') = \infty$. We will always assume that S is finite.

Note that to specify a Coxeter system (W,S), it is enough to draw the corresponding *Coxeter diagram* Γ: this is a graph on vertices indexed by elements of S, with vertices s and s' joined by an edge labelled $m(s,s')$ whenever this number (∞ allowed) is at least 3.

Remark 4.1. In what follows, the reader should note that nothing will turn out to depend on the edge labels $m(s,s')$ of Γ; the positive Bergman complex, the Bergman complex, or the graph associahedron associated with Γ will depend only upon the undirected graph underlying Γ.

Although an arbitrary Coxeter system (W,S) need not have a faithful representation of W as a group generated by orthogonal reflections with respect to a positive definite inner product, there exists a reasonable substitute, called its *geometric representation* [12, Sec.

5.3, 5.13], which we recall here. Let $V := \mathbb{R}^{|S|}$ with a basis of *simple roots* $\Delta := \{\alpha_s : s \in S\}$. Define an $\mathbb{R}$-valued bilinear form $(\cdot, \cdot)$ on V by

$$(\alpha_s, \alpha_{s'}) := -\cos\left(\frac{\pi}{m(s,s')}\right)$$

and let s act on V by the "reflection" that fixes $\alpha_s^{\perp}$ and negates α_s:

$$s(v) := v - 2(v, \alpha_s)\alpha_s.$$

This turns out to extend to a faithful representation of W on V, and one defines the *root system* Φ and *positive roots* Φ^+ by

$$\Phi := \{w(\alpha_s) : w \in W, s \in S\}$$
$$\Phi^+ := \{\alpha \in \Phi : \alpha = \sum_{s \in S} c_s \alpha_s \text{ with } c_s \geq 0\}$$

It turns out that $\Phi = \Phi^+ \sqcup \Phi^-$ where $\Phi^- := -\Phi^+$. We use M_Φ to denote the matroid represented by Φ^+ in V, which is of finite rank $r = |S|$, but has ground set E of possibly (countably) infinite cardinality. Its lattice of flats L_{M_Φ} may be infinite, although of finite rank r, and is well-known (see, e.g. [3]) to be isomorphic to the poset of *parabolic subgroups*

$$\{wW_J w^{-1} : w \in W, J \subseteq S\}$$

ordered by inclusion. In other words, every flat F is spanned by $w(\Phi_J^+)$ for some standard parabolic subroot system Φ_J^+ and $w \in W$.

Definition. Given a root $\alpha \in \Phi$, expressed uniquely in terms of the simple roots Δ as $\alpha = \sum_{s \in S} c_s \alpha_s$, define the *support* of α (written $\operatorname{supp} \alpha$) to be the vertex-induced subgraph of the Coxeter diagram Γ on the set of vertices $s \in S$ for which $c_s \neq 0$.

We will need the following well-known lemma about supports of roots. A proof of its first assertion for the Coxeter systems associated to Kac-Moody Lie algebras can be found in [13, Lemma 1.6]; we will need the assertion in general.

Lemma 4.2. *Let (W,S) be an arbitrary Coxeter system with Coxeter graph Γ. Then for any root $\alpha \in \Phi$ the graph $\operatorname{supp} \alpha$ is connected, and conversely, every connected subgraph Γ' of Γ occurs as $\operatorname{supp} \alpha$ for some positive root α.*

If one wants to think of the oriented matroid M_Φ as the oriented matroid of a hyperplane arrangement (as opposed to the oriented matroid of the configuration of vectors Φ^+), one must work with the contragredient representation V^* [12, 5.13]. Let $\{\delta_s : s \in S\}$ denote the basis for V^* dual to the basis of simple roots Δ for V. Then the *(closed) fundamental chamber* R is the nonnegative cone spanned by $\{\delta_s : s \in S\}$ inside V^*. The *Tits cone* is the union $\bigcup_{w \in W} w(R)$, a (possibly proper, not necessarily closed nor polyhedral) convex cone inside V^*. Every positive root $\alpha \in \Phi^+$ gives an oriented hyperplane H_α in V^* with nonnegative half-space $\{f \in V^* : f(\alpha) \geq 0\}$. These hyperplanes and half-spaces decompose the Tits cone[1] into (closed) regions that turn out to be simplicial cones which are exactly

[1]We should point out that when W is infinite, only part of the hyperplane or its nonnegative half-space lies inside the Tits cone, so we only consider their intersection with the Tits cone.

the images $w(R)$ as w runs through W; the tope (maximal covector) in the oriented matroid M_Φ associated to $w(R)$ will have the sign $+$ on the roots $\Phi^+ \cap w^{-1}(\Phi^+)$ and the sign $-$ on the roots $\Phi^+ \cap w^{-1}(\Phi^-)$.

Proposition 4.3. *Let (W, S) be an arbitrary Coxeter system, with root system Φ and Coxeter diagram Γ.*

(i) *Positive flats in the oriented matroid M_Φ correspond to subsets $J \subset S$.*
(ii) *Connected positive flats in the oriented matroid M_Φ correspond to subsets $J \subset S$ such that the vertex-induced subgraph Γ_J is connected, that is, to tubes in Γ.*
(iii) *The simple roots Δ form a circuitous base for the matroid M_Φ.*
(iv) *If $F \subset G$ are flats in M_Φ with G connected, then the matroid quotient G/F is connected.*

Proof. (i): The hyperplanes bounding the base region/tope R are $\{H_{\alpha_s} : s \in S\}$, so positive flats are those spanned by sets of the form $\{\alpha_s : s \in J\}$ for subsets $J \subset S$. We denote such a positive flat by $\mathrm{cl}(J)$.
(ii): Let $J \subset S$ with subgraph Γ_J, and consider its associated positive flat $\mathrm{cl}(J)$. The first assertion of Lemma 4.2 shows that $\mathrm{cl}(J)$ will not be connected if Γ_J is disconnected. To see this, represent the flat $\mathrm{cl}(J)$ by a matrix in which the rows correspond to simple roots of $\mathrm{cl}(J)$, i.e. vertices of Γ_J, and the columns express each positive root in $\mathrm{cl}(J)$ as a combination of simple roots. By permuting columns, one can obtain a matrix which is a block-direct sum of two smaller matrices, and hence $\mathrm{cl}(J)$ will not be connected.

On the other hand, if Γ_J is connected, then the second assertion of Lemma 4.2 shows that there is a positive root α with $\mathrm{supp}\,\alpha = \Gamma_J$, and consequently $\{\alpha_s : s \in J\} \cup \{\alpha\}$ gives a circuit in M_Φ spanning this flat, so it is connected.
(iii): This follows from the argument in (ii); given $J \subset S$ with Γ_J connected, the basic circuit $\mathrm{circ}(\Delta, \alpha)$ where $\mathrm{supp}\,\alpha = \Gamma_J$ spans the connected flat corresponding to J.
(iv): Let F, G correspond to the parabolic subgroups uW_Ju^{-1}, vW_Kv^{-1}, or equivalently, assume they are spanned by $u\Phi_J^+, v\Phi_K^+$. One can make the following reductions:

- Translating by v^{-1}, one can assume that v is the identity.
- Since (W_K, K) itself forms a Coxeter system with root system Φ_K, one can assume $M_\Phi = G$ and $K = S$. In particular, M_Φ is connected.
- Replacing the Coxeter system (W, S) by the system (W, uSu^{-1}), one can assume that u is the identity.

In other words, F is the positive flat corresponding to some subgraph Γ_J of Γ, and we must show M_Φ/F is a connected matroid. This is a consequence of (iii) and Lemma 2.9. □

We now give our main result.

Theorem 1.1. *Let (W, S) be an arbitrary Coxeter system, with root system Φ, Coxeter diagram Γ, and associated oriented matroid M_Φ. Then the face poset of the coarse subdivision of $\mathcal{B}^+(M_\Phi)$ is dual to the face poset of the graph associahedron $P(\Gamma)$.*

Proof. By Theorem 3.1, we need to show that the face poset of (the coarse subdivision of) $\mathcal{B}^+(M_\Phi)$ is equal to the poset of tubings of Γ, ordered by containment. We begin by describing a map Ψ from flags of positive flats to tubings of Γ.

By Proposition 4.3, positive flats of M_Φ correspond to subsets $J \subset S$ or subgraphs Γ_J of the Coxeter graph Γ. Furthermore, a positive flat is connected if and only if Γ_J is a tube, and hence an arbitrary positive flat corresponds to a disjoint union of compatible tubes, no two of which are nested. Since an inclusion of flats corresponds to an inclusion of the subsets J, a flag $\mathcal{F}$ of positive flats corresponds to a nested chain of such unions of non-nested compatible tubes, that is, to a tubing $\Psi(\mathcal{F})$. Furthermore, in this correspondence, inclusion of flags corresponds to containment of tubings.

We claim that the map from flags to tubings is surjective. Given some tubing of Γ, linearly order its tubes $J_1, \dots, J_k$ by any linear extension of the inclusion partial ordering, and then the flag $\mathcal{F}$ of positive flats having F_i spanned by $\{\alpha_s : s \in J_1 \cup J_2 \cup \dots \cup J_i\}$ will map to this tubing.

Lastly, we show that Ψ is actually a well-defined injective map when regarded as a map on cells of the coarse subdivision of $\mathcal{B}^+(M_\Phi)$. To do so, it is enough to show that two flags $\mathcal{F}, \mathcal{F}'$ of positive flats give the same tubing if and only if $M_{\mathcal{F}}$ and $M_{\mathcal{F}'}$ coincide. By Lemma 4.3(iv) and Proposition 2.10, we need to show that $\Psi(\mathcal{F})$ and $\Psi(\mathcal{F}')$ coincide if and only if $T_{\mathcal{F}}$ and $T_{\mathcal{F}'}$ coincide. But this is clear, because by construction, the rooted forest $T_{\mathcal{F}}$ ignores the ordering within the flag, and only records the data of the tubes which appear, that is, the tubing.

$\square$

Corollary 4.4. *The Bergman complex and the positive Bergman complex of a Coxeter arrangement $\mathcal{A}$ are both simplicial. The latter is furthermore a flag simplicial sphere.*

Another corollary of our proof is a new realization for the positive Bergman complex of a Coxeter arrangement: we can obtain it from a simplex by a sequence of stellar subdivisions.

5. The Bergman complex of a Coxeter arrangement

Nested set complexes are simplicial complexes which are the combinatorial core of De Concini and Procesi's subspace arrangement models [8], and of the resolution of singularities in toric varieties [9]. We now recall the definition of the minimal nested set complex of a meet-semilattice L, which we will simply refer to as **the** nested set complex of L, and denote $\mathcal{N}(L)$.

Say an element y of L is *irreducible* if the lower interval $[\hat{0}, y]$ cannot be decomposed as the product of smaller intervals of the form $[\hat{0}, x]$. The nested set complex $\mathcal{N}(L)$ of L is a simplicial complex whose vertices are the irreducible elements of L. A set X of irreducibles is *nested* if for any antichain $\{x_1, \dots, x_k\}$ in X, $x_1 \vee \dots \vee x_k$ is not irreducible. These nested sets are the faces of $\mathcal{N}(L)$.

If M is a matroid and L_M is its lattice of flats, we will also call $\mathcal{N}(L_M)$ the nested set complex of M, and denote it $\mathcal{N}(M)$. (Recall that the irreducible elements of L_M are the connected flats of M.) It turns out that when we are considering the oriented matroid M_Φ of a Coxeter arrangement of type Φ, the Bergman complex $\mathcal{B}(M_\Phi)$ and the nested set complex $\mathcal{N}(M_\Phi)$ are equal.

To prove this theorem, we use a result of Feichtner and Sturmfels [11]. They showed that, for any matroid M, the order complex of $\mathcal{N}(M)$ refines the coarse subdivision of the

Bergman complex $\mathcal{B}(M)$ and is refined by its fine subdivision. Moreover, they proved the following theorem.

Theorem 5.1. [11] *The nested set complex $\mathcal{N}(M)$ equals the Bergman complex $\mathcal{B}(M)$ if and only if the matroid G/F is connected for every pair of flats $F \subset G$ in which G is connected.*

Combining their Theorem 5.1 with Proposition 4.3(iv) immediately yields the following result.

Theorem 1.2. *For any Coxeter system (W, S) and associated root system Φ, the coarse subdivision of the Bergman complex $\mathcal{B}(M_\Phi)$ of the Coxeter arrangement of type Φ is equal to the nested set complex $\mathcal{N}(M_\Phi)$.*

ACKNOWLEDGMENTS We are very grateful to Satyan Devadoss for allowing us to reproduce two of his figures from [6]. These are our Figures 3 and 4. Additionally, we thank Bernd Sturmfels and Eva Feichtner for useful discussions.

REFERENCES

[1] F. Ardila and C. Klivans. *The Bergman complex of a matroid and phylogenetic trees.* Preprint, arXiv:math.CO/0311370, 2003.
[2] F. Ardila, C. Klivans, and L. Williams. *The positive Bergman complex of an oriented matroid*, to appear in the European Journal of Combinatorics.
[3] H. Barcelo and E. Ihrig, *Lattices of parabolic subgroups in connection with hyperplane arrangements*, J. Algebraic Combin. **9** (1999), 5–24.
[4] A. Björner and F. Brenti, *The combinatorics of Coxeter groups*, in preparation.
[5] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, G. Ziegler. *Oriented matroids.* Encyclopedia of Mathematics and Its Applications, vol. 46. Cambridge University Press, Cambridge, 1993.
[6] M. Carr, S. Devadoss. *Coxeter complexes and graph associahedra*, preprint.
[7] M. Davis, T. Januszkiewicz, and R. Scott, *Fundamental groups of minimal blow-ups*, Adv. Math. **177** (2003) 115-179.
[8] C. De Concini, C. Procesi. *Wonderful models of subspace arrangements*, Selecta Math. (N.S.) 1, no. 3 (1995), 459-494.
[9] E. Feichtner, D. Kozlov. *Incidence combinatorics of resolutions*, to appear in Selecta Math. (N.S.).
[10] E. Feichtner, I. Müller, *On the topology of nested set complexes*, to appear in Proc. Amer. Math. Soc.
[11] E. Feichtner, B. Sturmfels. *Matroid polytopes, nested sets and Bergman fans*, preprint.
[12] J. Humphreys. *Reflection groups and Coxeter groups.* Cambridge University Press, 1990.
[13] V. Kac. *Infinite dimensional Lie algebras.* Cambridge University Press, 1990.
[14] A. Postnikov. *Permutahedra, associahedra, and beyond*, in preparation.
[15] D. Speyer and B. Sturmfels. *The tropical Grassmannian.* Advances in Geometry, **4** (2004), no.3, p. 389-411.
[16] D. Speyer and L. Williams. *The tropical totally positive Grassmannian,* to appear in the Journal of Algebraic Combinatorics.
[17] R. P. Stanley. *Enumerative combinatorics, vol.1.* Cambridge University Press, New York, 1986.
[18] B. Sturmfels. *Solving systems of polynomial equations.* CBMS Regional Conference Series in Mathematics, vol. 97. American Mathematical Society, Providence, RI, 2002.
[19] G. Ziegler. *Lectures on polytopes.* Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995.

MICROSOFT, MINNESOTA, MIT

A Structure Theory of the Sandpile Monoid for Digraphs (EXTENDED ABSTRACT)

L. Babai* E. Toumpakari†

May, 2005

Abstract

The Abelian Sandpile Model is a diffusion process on graphs, studied, under various names, in statistical physics, theoretical computer science, and algebraic graph theory. The model takes a rooted directed multigraph $\mathcal{X}^*$, the *ambient space*, in which the root is accessible from every vertex, and associates with it a commutative monoid $\mathcal{M}$, a commutative semigroup $\mathcal{S}$, and an abelian group $\mathcal{G}$ as follows. For vertices i, j, let a_{ij} denote the number of $i \to j$ edges and let $\deg(i)$ denote the out-degree of i in $\mathcal{X}^*$. Let V be the set of ordinary (non-root) vertices. With each $i \in V$ associate a symbol x_i and consider the relations $\deg(i)x_i = \sum_{j \in V} a_{ij}x_j$. Let $\mathcal{M}$, $\mathcal{S}$, and $\mathcal{G}$ be the commutative monoid, semigroup and group, respectively, generated by $\{x_i : i \in V\}$ subject to these defining relations. $\mathcal{M}$ is the *sandpile monoid*, $\mathcal{S}$ is the *sandpile semigroup*, and $\mathcal{G}$ is the *sandpile group* associated with $\mathcal{X}^*$. We observe that $\mathcal{G}$ is the unique minimal ideal of $\mathcal{M}$.

We establish connections between the algebraic structure of $\mathcal{M}$, $\mathcal{S}$, $\mathcal{G}$, and the combinatorial structure of the underlying ambient space $\mathcal{X}^*$. $\mathcal{M}$ is a distributive lattice of semigroups each of which has a unique idempotent. The distributive lattice in question is the lattice $\mathcal{L}$ of idempotents of $\mathcal{M}$; $\mathcal{L}$ turns out to be isomorphic to the dual of the lattice of ideals of the poset of normal strong components of $\mathcal{X}^*$ (strong components which contain a cycle). The $\mathcal{M} \to \mathcal{L}$ epimorphism defines the *finest* semilattice congruence of $\mathcal{M}$; therefore $\mathcal{L}$ is the *universal semilattice* of $\mathcal{M}$.

We characterize the directed graphs $\mathcal{X}^*$ for which $\mathcal{S}$ has a unique idempotent; this includes the important case when the digraph induced on the ordinary vertices is strongly connected. If the idempotent in $\mathcal{S}$ is unique then the Rees quotient $\mathcal{S}/\mathcal{G}$ (obtained by contracting $\mathcal{G}$ to a zero element) is nilpotent. Let, in this case, k denote the nilpotence class of $\mathcal{S}/\mathcal{G}$. Our main result establishes the existence of functions ψ_1 and ψ_2 such that $|\mathcal{S}/\mathcal{G}| \leq \psi_1(k)$ and $\mathcal{G}$ contains a cyclic subgroup of index $\leq \psi_2(k)$. This result is a corollary to our asymptotic characterization of the ambient spaces with bounded k: every sufficiently large directed multigraph with this property can be described as a "circular tollway system of bounded effective volume."

1 The Abelian Sandpile Model

The Abelian Sandpile Model is a diffusion process on graphs, studied, under various names, in statistical physics, theoretical computer science, and algebraic graph theory[1]. The model takes a finite directed multigraph ("digraph") $\mathcal{X}^*$ with a special vertex called the *sink* as its *ambient space*, and associates with it a finite commutative monoid $\mathcal{M}$, a finite commutative semigroup $\mathcal{S}$, and a finite abelian group $\mathcal{G}$, called the *sandpile monoid*, the *sandpile semigroup* and the *sandpile group* of $\mathcal{X}^*$, respectively. We assume that the sink is accessible from every vertex and has out-degree zero. Vertices other than the sink will be called *ordinary*. A *state* of the game is an assignment of an integer $h_i \geq 0$ to each ordinary vertex i. The integer h_i may

*Department of Computer Science, University of Chicago. Email: laci@cs.uchicago.edu

†Department of Mathematics, University of Chicago, Email: evelint@math.uchicago.edu

[1]The Abelian Sandpile Model is identical with the "dollar-game" in Biggs' terminology [3], and is a variant of the "chip-firing game" studied in computer science [5].

be thought of as the number of sandgrains (the *height* of the sandpile) at *site* i. A state is *stable* if for all ordinary vertices i, $0 \le h_i < \deg(i)$, where deg denotes the out-degree in $\mathcal{X}^*$. If $h_i \ge \deg(i)$ for an ordinary vertex i, the "pile" at i may be "toppled," sending one grain through each edge leaving i. So h_i is reduced by $\deg(i)$, and for each ordinary vertex j, the height h_j increases by a_{ij}, the number of edges from i to j. The sink "collects" the grains "falling off" the ordinary vertices and never topples. Starting with any state and toppling unstable ordinary vertices in succession, we arrive at a stable state in a finite number of steps, since the sink is accessible from every vertex.

By a Jordan-Hölder argument (the "Diamond Lemma," cf. [14]), the order in which the topplings occur does not matter [5, 7]; given an initial state $\mathbf{h}$, every stabilizing sequence ("avalanche") leads to the same stable state $\sigma(\mathbf{h})$; hence the term "abelian."

2 The Sandpile Monoid

Our standard reference to semigroup theory is Grillet [9].

The *sandpile monoid* is defined as the set of stable states under the operation of pointwise addition and stabilization. We denote this operation by $\oplus$. So, for stable states $\mathbf{h}_1$ and $\mathbf{h}_2$ we set $\mathbf{h}_1 \oplus \mathbf{h}_2 := \sigma(\mathbf{h}_1 + \mathbf{h}_2)$. The all-zero state 0 is the identity in $\mathcal{M}$. The subsemigroup of $\mathcal{M}$ generated by the non-zero states is the *sandpile semigroup* $\mathcal{S}$. Clearly, $\mathcal{M} = \mathcal{S} \cup \{0\}$.

3 The Sandpile Group

Fact 3.1 *Every finite monoid has a unique minimal ideal.*

Fact 3.2 *The minimal ideal of a finite commutative monoid is a group.*

Definition 3.3 The unique minimal ideal of the sandpile monoid is called the *sandpile group*.

Note that this definition is more concise but equivalent to the definitions occurring in the literature (Dhar [7], Creutz [6]).

The common approach to defining the sandpile group is the following: We say that the stable state $\mathbf{h}_1$ is *accessible* from the stable state $\mathbf{h}_2$ if $(\exists \mathbf{h} \in \mathcal{M})(\mathbf{h}_1 = \mathbf{h}_2 \oplus \mathbf{h})$. We say that a state $\mathbf{h}_1$ is accessible from a state $\mathbf{h}_2$ if $\sigma(\mathbf{h}_1)$ is accessible from $\sigma(\mathbf{h}_2)$. A stable state is called *recurrent* (or "critical") if it is accessible from every state. The set of recurrent states is then defined to be the sandpile group. (It is identical with the unique minimal ideal of $\mathcal{M}$.)

4 Generators and relations for the Sandpile Monoid

Let us fix some notation. The digraph $\mathcal{X}^* = (V^*, E^*)$ denotes our ambient space; a vertex is designated as the sink; recall that a_{ij} is the number of edges from vertex i to vertex j. Finally, $\mathcal{X} = (V, E)$ denotes the subgraph of $\mathcal{X}^*$ induced on the set V of ordinary (non-sink) vertices.

Proposition 4.1

(i) The sandpile monoid is the commutative monoid generated by the symbols $\{x_i : i \in V\}$ subject to the set of defining relations $\mathcal{R} = \{\deg(i)x_i = \sum_{j \in V} a_{ij}x_j : i \in V\}$.

(ii) The sandpile semigroup is the commutative semigroup generated by the symbols $\{x_i : i \in V\}$ subject to the set of defining relations $\mathcal{R} = \{\deg(i)x_i = \sum_{j \in V} a_{ij}x_j : i \in V\}$.

5 Generators and relations for the Sandpile Group

Definition 5.1 Let $\mathcal{M}$ be a monoid, $\mathcal{G}$ a group and $\phi : \mathcal{M} \to \mathcal{G}$ a homomorphism. We say that $(\phi, \mathcal{G})$ is the *universal group* of $\mathcal{M}$ if every homomorphism from $\mathcal{M}$ to a group factors through ϕ.

Observation 5.2 *If $\langle W|R\rangle$ is a presentation of a monoid $\mathcal{M}$, then $\langle W|R\rangle$ is also a presentation of the universal group $\mathcal{G}$ of $\mathcal{M}$ as a group.*

Fact 5.3 *Let $\mathcal{M}$ be a finite commutative monoid and let $\mathcal{G}$ be the minimal ideal of $\mathcal{M}$. Let $e \in \mathcal{G}$ be the identity in $\mathcal{G}$ and let $\phi : \mathcal{M} \to \mathcal{G}$ be defined by $\phi(x) := e + x$. Then $(\phi, \mathcal{G})$ is the universal group of $\mathcal{M}$.*

Corollary 5.4 *The sandpile group is the universal group of the sandpile monoid (under the homomorphism described in Fact 5.3).*

Corollary 5.5 (Dhar[7]) *The sandpile group is isomorphic to the quotient $\mathbb{Z}^V/\Lambda$, where Λ is the lattice spanned by the rows of the reduced Laplacian (see Definition 5.7).*

Definition 5.6 The *Laplacian* $L = (L_{ij})_{i,j\in V^*}$ of $\mathcal{X}^*$ is the $|V^*| \times |V^*|$ matrix defined by

$$L_{ij} := \begin{cases} \deg(i) - a_{ii} & \text{if } i = j, \\ -a_{ij} & \text{otherwise.} \end{cases} \tag{1}$$

Definition 5.7 The *reduced Laplacian* $\Delta = (\Delta_{ij})_{i,j\in V}$ of $\mathcal{X}^*$ is defined as the matrix obtained from the Laplacian L by deleting the row and the column corresponding to the sink.

The digraph version of Kirchhoff's [12] classical Matrix-Tree Theorem (Tutte [17], cf. [13]) now implies:

Corollary 5.8 (Dhar[7]) *The order of the sandpile group is the number of directed spanning trees of the ambient space $\mathcal{X}^*$ directed towards the sink.*

6 The universal lattice of the Sandpile Monoid

Definition 6.1 Let $\mathcal{S}$ be a semigroup, $\mathcal{L}$ a semilattice, and $\phi : \mathcal{S} \to \mathcal{L}$ a homomorphism. We say that $(\phi, \mathcal{L})$ is the *universal semilattice* of $\mathcal{S}$ if every homomorphism from $\mathcal{S}$ to a semilattice factors through ϕ.

Fact 6.2 *Every semigroup has a universal semilattice.*

Fact 6.3 *For a finite monoid, the universal semilattice is a lattice.*

Fact 6.4 *Every finite lattice is the universal lattice of a finite commutative monoid (namely, of itself).*

Fact 6.5 *For a finite commutative monoid $\mathcal{M}$, the universal lattice is isomorphic to the lattice of idempotents of $\mathcal{M}$ and $\mathcal{M}$ is a lattice of semigroups with unique idempotent.*

Theorem 6.6 *For a finite lattice $\mathcal{L}$, the following are equivalent:*

(i) $\mathcal{L}$ is the universal lattice of a sandpile monoid.

(ii) $\mathcal{L}$ is distributive.

The proof of Theorem 6.6 goes through a description of the semilattice of idempotents of the sandpile monoid in terms of the strong components of $\mathcal{X}$:

Definition 6.7

(i) A vertex is *normal* if it belongs to a cycle; and *abnormal* otherwise.

(ii) A strong component of a digraph is *normal* if it contains a cycle.

So, an abnormal strong component consists of a single abnormal vertex. All vertices of a normal strong component are normal.

Theorem 6.8 *The following lattices are isomorphic:*

(i) The lattice of idempotents of the sandpile monoid corresponding to the ambient space $\mathcal{X}^*$.

(ii) The dual lattice of ideals of the accessibility partial order on the set of normal strong components of $\mathcal{X}$.

Corollary 6.9 *The sandpile semigroup has a unique idempotent if and only if*

(i) $\mathcal{X}$ is a directed acyclic graph (DAG) with at least one vertex of $\deg \geq 2$ (*degree, as always, relative to $\mathcal{X}^*$), or*

(ii) $\mathcal{X}$ has a unique normal strong component.

7 Bounded nilpotence class

Fact 7.1 *If $\mathcal{S}$ is a finite semigroup with a unique idempotent and $\mathcal{G}$ is the minimal ideal of $\mathcal{S}$ then the Rees quotient (obtained by contracting $\mathcal{G}$ to a zero element) is nilpotent.*

In this section we consider the case where the sandpile semigroup $\mathcal{S}$ has a unique idempotent. This is the case, in particular, when $\mathcal{X}$ is strongly connected.

Let k denote the nilpotence class of $\mathcal{S}/\mathcal{G}$. We observe that $k-1$ is the maximum weight of any (not necessarily stable) transient (non-recurrent) state $\mathbf{h} \in \mathbb{N}^V$.

We need to treat DAGs separately.

Proposition 7.2 $\mathcal{M} = \mathcal{G}$ *if and only if $\mathcal{X}$ is a DAG.*

Our main result (Theorem 7.10) will asymptotically characterize the ambient spaces corresponding to bounded nilpotence class. The main corollary to the result asserts, somewhat surprisingly, that the boundedness of the nilpotence class of $\mathcal{S}/\mathcal{G}$ implies the boundedness of $\mathcal{S}/\mathcal{G}$ itself, and has strong structural implications on the sandpile group $\mathcal{G}$.

Corollary 7.3 *There exist functions ψ_1 and ψ_2 such that if the sandpile quotient $\mathcal{S}/\mathcal{G}$ has nilpotence class k then its order is $|\mathcal{S}/\mathcal{G}| \leq \psi_1(k)$; and if $\mathcal{X}$ is not a DAG then the sandpile group $\mathcal{G}$ contains a cyclic subgroup of index $\leq \psi_2(k)$.*

Definition 7.4

(i) We define the *strong degree* of the vertex v (denoted by $\deg_s(v)$) to be the number of edges from v to the vertices in the strong component of v. So, an ordinary vertex v is abnormal if and only if $\deg_s(v) = 0$.

(ii) Let $\mathcal{X}^*$ be an ambient space and let A be the set of vertices that belong to all cycles in $\mathcal{X}$. We define the *effective volume* of $\mathcal{X}^*$ to be

$$\operatorname{vol}(\mathcal{X}^*) := \prod_{v \in A} \deg_s(v) \prod_{v \in V \setminus A} \deg(v). \tag{2}$$

Theorem 7.5 *For a class of ambient spaces $\mathcal{C}$, the following are equivalent:*

(i) The nilpotence class of the Rees quotients $\mathcal{S}(\mathcal{X}^*)/\mathcal{G}(\mathcal{X}^*)$ is bounded ($\mathcal{X}^* \in \mathcal{C}$).

(ii) The number of transient states, $|\mathcal{M}(\mathcal{X}^*) \setminus \mathcal{G}(\mathcal{X}^*)|$, is bounded ($\mathcal{X}^* \in \mathcal{C}$).

(iii) The effective volume $\mathrm{vol}(\mathcal{X}^*)$ *is bounded ($\mathcal{X}^* \in \mathcal{C}$).*

Now we describe the asymptotic structure of $\mathcal{X}^*$ for bounded k.

Definition 7.6 Let $\mathcal{Y}$ be a directed acyclic graph (DAG) and let v be a vertex.

(i) The vertex v is an *entrance* if every vertex of $\mathcal{Y}$ is accessible from v.

(ii) The vertex v is an *exit* if v is accessible from every vertex of $\mathcal{Y}$.

(iii) $\mathcal{Y}$ is a *rest area* if it has an entrance and an exit. Note that the entrance and the exit are unique.

(iv) The *interior* vertices of a rest area are the vertices other than the entrance and the exit. The set of interior vertices of the rest area $\mathcal{Y}$ is denoted by $\mathrm{int}(\mathcal{Y})$.

Definition 7.7 We say that an ordinary vertex $v \in V$ is *thin* if $\deg_s(v) = 1$.

Definition 7.8

(i) A *circular highway* is a strong component of $\mathcal{X}$ which is a thin directed cycle (every vertex of the cycle is thin).

(ii) We now construct a *circular tollway.* We start with a circular highway on which we designate edges $(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)$ in this cyclic order ($n \geq 0$). For $i = 1, \ldots, n$, we delete the edge (u_i, v_i) and we glue a rest area R_i between u_i and v_i so that $\mathrm{entrance}(R_i) = u_i$ and $\mathrm{exit}(R_i) = v_i$. (The R_i are disjoint.)

(iii) We say that an ambient space $\mathcal{X}^*$ is a *circular tollway system* if $\mathcal{X}$ has a unique normal strong component and this strong component is a circular tollway.

Definition 7.9 Let $\mathcal{X}^*$ be an ambient space.

(i) We call an ordinary vertex v *relevant* if $\deg(v) \geq 1$.

(ii) An edge (v, w) is *relevant* if its tail v is a relevant vertex. (In this definition, $v \in V$ and $w \in V^*$.)

Theorem 7.10 *For a class of ambient spaces $\mathcal{C}$, the following are equivalent:*

(a) Any of the equivalent conditions (i), (ii) and (iii) of Theorem 7.5.

(b) $(\exists n_0 \geq 0)$(if $\mathcal{X}^* \in \mathcal{C}$ has more than n_0 relevant edges then $\mathcal{X}^*$ is a circular tollway system of bounded effective volume).

We defer the proofs to an expanded version of this paper.

References

[1] P. Bak, C. Tang, K. Wiesenfeld: Self-organized Criticality. *Phys. Rev. A* **38** (1988) pp. 364–374

[2] N. Biggs: Algebraic Potential Theory on Graphs. *Bull. London Math. Soc.* **29** (1997) pp. 641–682

[3] N. Biggs: Chip-Firing and the Critical Group of a Graph. *Journal of Algebraic Combinatorics* **9** (1999) pp. 25–45

[4] A. Bjørner, L. Lovász: Chip-firing Games on Directed Graphs. *Journal of Algebraic Combinatorics* **1** (1992) pp. 305–328

[5] A. Bjørner, L. Lovász, P. Shor: Chip-firing Games on Graphs. *Europ. J. Combinatorics* **12** (1991) pp. 283–291

[6] M. Creutz: Abelian Sandpile. *Computers in Physics* **5** (1991) pp. 198–203

[7] D. Dhar: Self-organized Critical State of Sandpile Automaton Models. *Phys. Rev. Lett.* **64** (1990) pp. 1613–1616

[8] D. Dhar, P. Ruelle, S. Sen, D. Verma: Algebraic Aspects of Abelian Sandpile Models. *J. Phys. A* **28** (1995) no.4 pp. 805–831

[9] P.A. Grillet: *Semigroups.* Pure and Applied Mathematics **193**, M. Dekker (1995)

[10] P.A. Grillet: Nilsemigroups on trees. *Semigroup Forum* **43** (1991) pp. 187-201.

[11] H. Jensen: *Self-Organized Criticality.* Cambridge Lecture Notes in Physics **10** (1998)

[12] G. Kirchhoff: Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanisher Ströme geführt wird. *Ann. Phys. Chem.* **72** (1847) pp. 497-508

[13] L. Lovász: *Combinatorial Problems and Exercises.* 2nd Edition (1993) Akadémiai Kiadó, Budapest, and Elsevier

[14] M. H. A. Newmann: On theories with a combinatorial notion of "equivalence." *Ann. Math.* **43** (1942) pp. 211–264

[15] E. Speer: Asymmetric Abelian Sandpile Models. *Journal of Statistical Physics* **71** nos.1/2 (1993)

[16] E. Toumpakari: On the sandpile group of regular trees. *Europ. J. Combinatorics*, to appear

[17] W. T. Tutte: The dissection of equilateral triangles into equilateral triangles. *Proc. Cambridge Phil. Soc.* **44** (1948) pp. 463–482

DECREASING SUBSEQUENCES IN PERMUTATIONS AND WILF EQUIVALENCE FOR INVOLUTIONS

MIREILLE BOUSQUET-MÉLOU AND EINAR STEINGRÍMSSON

Abstract. In a recent paper, Backelin, West and Xin describe a map ϕ^* that recursively replaces all occurrences of the pattern $k \cdots 21$ in a permutation σ by occurrences of the pattern $(k-1)\cdots 21k$. The resulting permutation $\phi^*(\sigma)$ contains no decreasing subsequence of length k. We prove that, rather unexpectedly, the map ϕ^* commutes with taking the inverse of a permutation.

In the BWX paper, the definition of ϕ^* is actually extended to full rook placements on a Ferrers board (the permutations correspond to square boards), and the construction of the map ϕ^* is the key step in proving the following result. Let T be a set of patterns starting with the prefix $12\cdots k$. Let T' be the set of patterns obtained by replacing this prefix by $k\cdots 21$ in every pattern of T. Then for all n, the number of permutations of the symmetric group $\mathcal{S}_n$ that avoid T equals the number of permutations of $\mathcal{S}_n$ that avoid T'.

Our commutation result, generalized to Ferrers boards, implies that the number of *involutions* of $\mathcal{S}_n$ that avoid T is equal to the number of involutions of $\mathcal{S}_n$ avoiding T', as recently conjectured by Jaggard.

Version française. Dans un article récent, Backelin, West et Xin ont défini une transformation ϕ^* qui détruit récursivement toutes les sous-suites décroissantes de longueur k d'une permutation (k est fixé). Cette transformation s'obtient en itérant une transformation élémentaire ϕ qui détruit *une* sous-suite décroissante de longueur k.

Ces deux transformations peuvent être étendues à des objets plus généraux que les permutations : des placements de tours sur des diagrammes de Ferrers. Le trio BWX s'est servi de ϕ^* pour démontrer le résultat suivant. Soit T un ensemble de motifs commençant tous par le préfixe $12\cdots k$. Soit T' l'ensemble de motifs obtenu en remplaçant ce préfixe par $k\cdots 21$ dans chacun des motifs de T. Alors pour tout n, le nombre de permutations de $\mathcal{S}_n$ qui évitent T est égal au nombre de permutations de $\mathcal{S}_n$ qui évitent T'.

Le résultat principal de notre travail est que, très curieusement, la transformation itérée ϕ^* commute avec l'inversion des permutations (alors que ça n'est pas le cas de la transformation élémentaire ϕ). Plus généralement, elle commute avec la symétrie diagonale des placements de tours.

Un corollaire est l'analogue du résultat de BWX pour les involutions : avec des notations évidentes, $I_n(T) = I_n(T')$, où T et T' sont définis comme ci-dessus. Ce résultat avait été conjecturé par Jaggard.

1. Introduction

Let $\pi = \pi_1\pi_2\cdots\pi_n$ be a permutation of length n. Let $\tau = \tau_1\cdots\tau_k$ be another permutation. An *occurrence* of τ in π is a subsequence $\pi_{i_1}\cdots\pi_{i_k}$ of π that is order-isomorphic to τ. For instance, 246 is an occurrence of $\tau = 123$ in $\pi = 251436$. We say that π *avoids* τ if π contains no occurrence of τ. For instance, the above

Date: 12th November 2004.

Key words and phrases. Pattern avoiding permutations, Wilf equivalence, involutions, decreasing subsequences, prefix exchange.

Both authors were partially supported by the European Commission's IHRP Programme, grant HPRN-CT-2001-00272, "Algebraic Combinatorics in Europe".

permutation π avoids 1234. The set of permutations of length n is denoted by $\mathcal{S}_n$, and $\mathcal{S}_n(\tau)$ denotes the set of τ-avoiding permutations of length n.

The idea of systematically studying pattern avoidance in permutations appeared in the mid-eighties [20]. The main problem in this field is to determine $S_n(\tau)$, the cardinality of $\mathcal{S}_n(\tau)$, for any given pattern τ. This question has subsequently been generalized and refined in various ways (see for instance [1, 4, 8, 17], and [16] for a recent survey). However, relatively little is known about the original question. The case of patterns of length 4 is not yet completed, since the pattern 1324 still remains unsolved. See [5, 9, 22, 21, 25] for other patterns of length 4.

For length 5 and beyond, all the solved cases follow from three important generic results. The first one, due to Gessel [9, 10], gives the generating function of the numbers $S_n(12\cdots k)$. The second one, due to Stankova and West [23], states that $S_n(231\tau) = S_n(312\tau)$ for any pattern τ on $\{4, 5, \ldots, k\}$. The third one, due to Backelin, West and Xin [3], shows that $S_n(12\cdots k\tau) = S_n(k\cdots 21\tau)$ for any pattern τ on the set $\{k+1, k+2, \ldots, \ell\}$. In the present paper an analogous result is established for pattern-avoiding *involutions*. We denote by $\mathcal{I}_n(\tau)$ the set of involutions avoiding τ, and by $I_n(\tau)$ its cardinality.

The systematic study of pattern avoiding involutions was also initiated in [20], continued in [9, 11] for increasing patterns, and then by Guibert in his thesis [12]. Guibert discovered experimentally that, for a surprisingly large number of patterns τ of length 4, $I_n(\tau)$ is the nth Motzkin number:

$$M_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!}{k!(k+1)!(n-2k)!}.$$

This was already known for $\tau = 1234$ (see [18]), and consequently for $\tau = 4321$, thanks to the properties of the Schensted correspondence [19]. Guibert explained all the other instances of the Motzkin numbers, except for two of them: 2143 and 3214. However, he was able to describe a two-label generating tree for the class $\mathcal{I}_n(2143)$. Several years later, the Motzkin result for the pattern 2143 was at last derived from this tree: first in a bijective way [13], then using generating functions [6]. No simple generating tree could be described for involutions avoiding 3214, and it was only in 2003 that Jaggard [15] gave a proof of this final conjecture, inspired by [2]. More generally, he proved that for $k = 2$ or 3, $I_n(12\cdots k\tau) = I_n(k\cdots 21\tau)$ for all τ. He conjectured that this holds for all k, which we prove here.

We derive this from another result, which may be more interesting than its implication in terms of forbidden patterns. This result deals with a transformation ϕ^* that was defined in [3] to prove that $S_n(12\cdots k\tau) = S_n(k\cdots 21\tau)$. This transformation acts not only on permutations, but on more general objects called *full rook placements on a Ferrers shape* (see Section 2 for precise definitions). The map ϕ^* may, at first sight, appear as an *ad hoc* construction, but we prove that it has a remarkable, and far from obvious, property: it commutes with the inversion of a permutation, and more generally with the corresponding diagonal reflection of a full rook placement. (By the inversion of a permutation π we mean the map that sends π, seen as a bijection, to its inverse.)

The map ϕ^* is defined by iterating a transformation ϕ, which chooses a certain occurrence of the pattern $k\cdots 21$ and replaces it by an occurrence of $(k-1)\cdots 21k$. The map ϕ itself does *not* commute with the inversion of permutations, and our proof of the commutation theorem is actually quite complicated.

This strongly suggests that we need a better description of the map ϕ^*, on which the commutation theorem would become obvious. By analogy, let us recall what

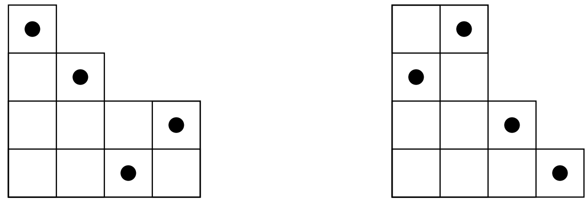


FIGURE 1. A full rook placement on a Ferrers board, and its inverse.

happened for the Schensted correspondence: the fact that the inversion of permutations exchanges the two tableaux only became completely clear with Viennot's description of the correspondence [24].

Actually, since the Schensted correspondence has nice properties regarding the monotone subsequences of permutations, and provides one of the best proofs of the identity $I_n(12\cdots k) = I_n(k\cdots 21)$, we suspect that the map ϕ^* might be related to this correspondence, or to an extension of it to rook placements.

2. Wilf equivalence for involutions

One of the main implications of this paper is the following.

Theorem 1. *Let $k \geq 1$. Let T be a set of patterns, each starting with the prefix $12\cdots k$. Let T' be the set of patterns obtained by replacing this prefix by $k\cdots 21$ in every pattern of T. Then, for all $n \geq 0$, the number of involutions of $\mathcal{S}_n$ that avoid T equals the number of involutions of $\mathcal{S}_n$ that avoid T'.*

In particular, the involutions avoiding $12\cdots k\tau$ and the involutions avoiding $k\cdots 21\tau$ are equinumerous, for any permutation τ of $\{k+1, k+2, \ldots, \ell\}$.

This theorem was proved by Jaggard for $k = 2$ and $k = 3$ [15]. It is the analogue, for involutions, of a result recently proved by Backelin, West and Xin for permutations [3]. Thus it is not very suprising that we follow their approach. This approach requires looking at pattern avoidance for slightly more general objects than permutations, namely, *full rook placements on a Ferrers board.*

Let λ be an integer partition, which we represent as a Ferrers board (Figure 1). A full rook placement on λ, or a *placement* for short, is a distribution of dots on this board, such that every row and column contains exactly one dot. This implies that the board has as many rows as columns.

Each cell of the board will be denoted by its coordinates: in the first placement of Figure 1, there is a dot in the cell $(1, 4)$. If the placement has n dots, we associate with it a permutation π of $\mathcal{S}_n$, defined by $\pi(i) = j$ if there is a dot in the cell (i, j). The permutation corresponding to the first placement of Figure 1 is $\pi = 4312$. This induces a bijection between placements on the $n \times n$ square and permutations of $\mathcal{S}_n$.

The *inverse* of a placement p on the board λ is the placement p' obtained by reflecting p and λ with respect to the main diagonal; it is thus a placement on the *conjugate* of λ, usually denoted by λ'. This terminology is of course an extension to placements of the classical terminology for permutations.

Definition 2. Let p be a placement on the board λ, and let π be the corresponding permutation. Let τ be a permutation of $\mathcal{S}_k$. We say that p *contains* τ if there exists in π an occurrence $\pi_{i_1}\pi_{i_2}\cdots\pi_{i_k}$ of τ such that the corresponding dots are contained in a *rectangular* sub-board of λ. In other words, the cell with coordinates $(i_k, \max_j \pi_{i_j})$ must belong to λ.

The placement of Figure 1 contains the pattern 12, but avoids the pattern 21, even though the associated permutation $\pi = 4312$ contains several occurrences of 21. We denote by $\mathcal{S}_\lambda(\tau)$ the set of placements on λ that avoid τ. If λ is self-conjugate, we denote by $\mathcal{I}_\lambda(\tau)$ the set of *symmetric* (that is, self-inverse) placements on λ that avoid τ. We denote by $S_\lambda(\tau)$ and $I_\lambda(\tau)$ the cardinalities of these sets.

In [2, 3, 23], it was shown that the notion of pattern avoidance in placements is well suited to deal with prefix exchanges in patterns. This was adapted by Jaggard [15] to involutions:

Proposition 3. *Let α and β be two involutions of $\mathcal{S}_k$. Let T_α be a set of patterns, each beginning with α. Let T_β be obtained by replacing, in each pattern of T_α, the prefix α by β. If, for every self-conjugate shape λ, $I_\lambda(\alpha) = I_\lambda(\beta)$, then $I_\lambda(T_\alpha) = I_\lambda(T_\beta)$ for every self-conjugate shape.*

Hence Theorem 1 will be proved if we can prove that $I_\lambda(12\cdots k) = I_\lambda(k\cdots 21)$ for any self-conjugate shape λ. A simple induction on k, combined with Proposition 3, shows that it is actually enough to prove the following:

Theorem 4. *Let λ be a self-conjugate shape. Then $I_\lambda(k\cdots 21) = I_\lambda((k-1)\cdots 21k)$.*

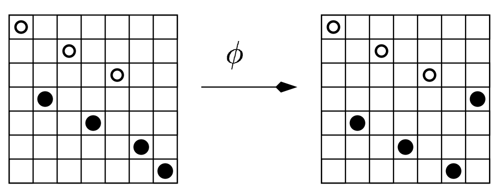
A similar result was proved in [3] for general (asymmetric) placements: for every shape λ, one has $S_\lambda(k\cdots 21) = S_\lambda((k-1)\cdots 21k)$. The proof relies on the description of a recursive bijection between the sets $\mathcal{S}_\lambda(k\cdots 21)$ and $\mathcal{S}_\lambda((k-1)\cdots 21k)$. What we prove here is that this complicated bijection actually *commutes with the inversion of a placement*, and this implies Theorem 4.

But let us first describe (and slightly generalize) the transformation defined by Backelin, West and Xin [3]. This transformation depends on k, which from now on is supposed to be fixed. Since Theorem 4 is trivial for $k = 1$, we assume $k \geq 2$.

Definition 5 (The transformation ϕ). Let p be a placement containing $k\cdots 21$, and let π be the associated permutation. To each occurrence of $k\cdots 21$ in p, there corresponds a decreasing subsequence of length k in π. The $\mathcal{A}$*-sequence* of p, denoted by $\mathcal{A}(p)$, is the smallest of these subsequences for the lexicographic order.

The corresponding dots in p form an occurrence of $k\cdots 21$. Rearrange these dots cyclically so as to form an occurrence of $(k-1)\cdots 21k$. The resulting placement is defined to be $\phi(p)$.

If p avoids $k\cdots 21$, we simply define $\phi(p) := p$. The transformation ϕ is also called the $\mathcal{A}$*-shift*.



$7\ \underline{4}\ 6\ \underline{3}\ 5\ \underline{2}\ \underline{1}$ $\qquad$ $7\ 3\ 6\ 2\ 5\ 1\ 4$

FIGURE 2. The $\mathcal{A}$-shift on the permutation 7 4 6 3 5 2 1, when $k = 4$.

An example is provided by Figure 2 (the letters of the $\mathcal{A}$-sequence are underlined, and the corresponding dots are black). It is easy to see that the $\mathcal{A}$-shift decreases the inversion number of the permutation associated with the placement (details will be given in the proof of Corollary 11). This implies that after finitely many iterations of

ϕ, there will be no more decreasing subsequences of length k in the placement. We denote by ϕ^* the iterated transformation, that recursively transforms *every* pattern $k \cdots 21$ into $(k-1) \cdots 21k$. For instance, with the permutation $\pi = 7\ 4\ 6\ 3\ 5\ 2\ 1$ of Figure 2 and $k = 4$, we find

$$\pi = 7\ \underline{4}\ 6\ \underline{3}\ 5\ \underline{2}\ \underline{1} \longrightarrow \underline{7}\ \underline{3}\ 6\ \underline{2}\ 5\ \underline{1}\ 4 \longrightarrow 3\ 2\ 6\ 1\ 5\ 7\ 4 = \phi^*(\pi).$$

The main property of ϕ^* that was proved and used in [3] is the following:

Theorem 6 (The BWX bijection). *For every shape λ, the transformation ϕ^* induces a bijection from $\mathcal{S}_\lambda((k-1) \cdots 21k)$ to $\mathcal{S}_\lambda(k \cdots 21)$.*

The key to our paper is the following rather unexpected theorem.

Theorem 7 (Global commutation). *The transformation ϕ^* commutes with the inversion of a placement.*

For instance, with π as above, we have

$$\pi^{-1} = 7\ \underline{6}\ \underline{4}\ \underline{2}\ 5\ 3\ \underline{1} \longrightarrow \underline{7}\ \underline{4}\ \underline{2}\ \underline{1}\ 5\ 3\ 6 \longrightarrow 4\ 2\ 1\ 7\ 5\ 3\ 6 = \phi^*(\pi^{-1})$$

and we observe that

$$\phi^*(\pi^{-1}) = (\phi^*(\pi))^{-1}.$$

Note, however, that $\phi(\pi^{-1}) \neq (\phi(\pi))^{-1}$. Indeed, $\phi(\pi^{-1}) = 7\ 4\ 2\ 1\ 5\ 3\ 6$ while $(\phi(\pi))^{-1} = 6\ 4\ 2\ 7\ 5\ 3\ 1$, so that the elementary transformation ϕ, that is, the $\mathcal{A}$-shift, does not commute with the inversion.

Theorems 6 and 7 together imply that ϕ^* induces a bijection from $\mathcal{I}_\lambda((k-1) \cdots 21k)$ to $\mathcal{I}_\lambda(k \cdots 21)$, for every self-conjugate shape λ. This proves Theorem 4, and hence Theorem 1. The rest of the paper is devoted to proving Theorem 7, which we call the theorem of *global commutation.* By this, we mean that the inversion commutes with the global tranformation ϕ^* (but not with the elementary transformation ϕ).

Remarks

1. At first sight, our definition of the $\mathcal{A}$-sequence (Definition 5), does not seem to coincide with the definition given in [3]. Let $a_k \cdots a_2 a_1$ denote the $\mathcal{A}$-sequence of the placement p, with $a_k > \cdots > a_1$. We identify this sequence with the corresponding set of dots in p. The dot a_k is the lowest dot that is the leftmost point in an occurrence of $k \cdots 21$ in p. Then a_{k-1} is the lowest dot such that $a_k a_{k-1}$ is the beginning of an occurrence of $k \cdots 21$ in p, and so on.

However, in [3], the dot a_k is chosen as above, but then each of the next dots $a'_{k-1}, \ldots, a'_1$ is chosen to be as far *left* as possible, and not as *low* as possible. Let us prove that the two procedures give the same sequence of dots. Assume not, and let $a_j \neq a'_j$ be the first (leftmost) point where the two sequences differ. By definition, a_j is lower than a'_j, and to the right of it. But then the sequence $a_{k-1} \cdots a_{j+1} a'_j a_j \cdots a_2 a_1$ is an occurrence of the pattern $k \cdots 21$ in p, which is smaller than $a_k \cdots a_2 a_1$ for the lexicographic order, a contradiction.

The fact that the $\mathcal{A}$-sequence can be defined in two different ways will be used very often in the paper.

2. At this stage, we have reduced the proof of Theorem 1 to the proof of the global commutation theorem, Theorem 7.

3. From local commutation to global commutation

In order to prove that ϕ^* commutes with the inversion of placements, it would naturally be tempting to prove that ϕ itself commutes with the inversion. However, this is not the case, as shown above. Given a placement p and its inverse p', we thus want to know how the placements $\phi(p)$ and $\phi(p')'$ differ.

Definition 8. For any shape λ and any placement p on λ, we define $\psi(p)$ by

$$\psi(p) := \phi(p')'.$$

Thus $\psi(p)$ is also a placement on λ.

Note that $\psi^m(p) = (\phi^m(p'))'$, so that the theorem of global commutation, Theorem 7, can be restated as $\psi^* = \phi^*$.

Combining the above definition of ψ with Definition 5 gives an alternative description of ψ.

Lemma 9 (The transformation ψ). *Let p be a placement containing $k \cdots 21$. Let $b_1, b_2, \ldots, b_k$ be defined recursively as follows: For all j, b_j is the leftmost dot such that $b_j \cdots b_2 b_1$ ends an occurrence of $k \cdots 21$ in p. We call $b_k \cdots b_2 b_1$ the $\mathcal{B}$*-sequence *of p, and denote it by $\mathcal{B}(p)$.*

Rearrange the k dots of the $\mathcal{B}$-sequence cyclically so as to form an occurrence of $(k-1) \cdots 21k$: the resulting placement is $\psi(p)$.

If p avoids $k \cdots 21$, then $\psi(p) = p$. The transformation ψ is also called the $\mathcal{B}$-shift.

According to the first remark that concludes Section 2, we can alternatively define b_j, for $j \geq 2$, as the *lowest* dot such that $b_j \cdots b_2 b_1$ ends an occurrence of $k \cdots 21$ in p.

We have seen that, in general, ϕ does not commute with the inversion. That is, $\phi(p) \neq \psi(p)$ in general. The above lemma tells us that $\phi(p) = \psi(p)$ if and only if the $\mathcal{A}$-sequence and the $\mathcal{B}$-sequence of p coincide. If they do *not* coincide, then we still have the following remarkable property, whose proof is deferred to the very end of the paper.

Theorem 10 (Local commutation). *Let p be a placement for which the $\mathcal{A}$- and $\mathcal{B}$-sequences do not coincide. Then $\phi(p)$ and $\psi(p)$ still contain the pattern $k \cdots 21$, and*

$$\phi(\psi(p)) = \psi(\phi(p)).$$

For instance, for the permutation of Figure 2 and $k = 4$, we have the following commutative diagram, in which the underlined (resp. overlined) letters correspond to the $\mathcal{A}$-sequence (resp. $\mathcal{B}$-sequence):

$$\begin{tikzcd}
 & \overline{7}\,\overline{\underline{4}}\,6\,\overline{\underline{3}}\,5\,\overline{\underline{2}}\,\underline{1} \arrow[dl, "\phi"'] \arrow[dr, "\psi"] & \\
\overline{7}\,\overline{3}\,6\,\overline{2}\,5\,\overline{1}\,4 \arrow[dr, "\psi"'] & & \underline{4}\,\underline{3}\,6\,\underline{2}\,5\,7\,\underline{1} \arrow[dl, "\phi"] \\
 & 3\,2\,6\,1\,5\,7\,4 &
\end{tikzcd}$$

A classical argument, which is sometimes stated in terms of *locally confluent* and *globally confluent rewriting systems* (see [14] and references therein), will show that Theorem 10 implies $\psi^* = \phi^*$, and actually the more general following corollary.

Corollary 11. *Let p be a placement. Any iterated application of the transformations ϕ and ψ yields ultimately the same placement, namely $\phi^*(p)$. Moreover, all the minimal sequences of transformations that yield $\phi^*(p)$ have the same length.*

Before we prove this corollary, let us illustrate it. We think of the set of permutations of length n as the set of vertices of an oriented graph, the edges of which are given by the maps ϕ and ψ. Figure 3 shows a connected component of this graph. The dotted edges represent ϕ while the plain edges represent ψ. The dashed edges correspond to the cases where ϕ and ψ coincide. We see that all the paths that start at a given point converge to the same point.

Proof. For any placement p, define the inversion number of p as the inversion number of the associated permutation π (that is, the number of pairs (i,j) such that $i<j$ and $\pi_i>\pi_j$). Assume p contains at least one occurrence of $k\cdots 21$, and let $i_1<\cdots<i_k$ be the positions (abscissae) of the elements of the $\mathcal{A}$-sequence of p. A careful examination of the inversions of p and $\phi(p)$ shows that

$$\operatorname{inv}(p)-\operatorname{inv}(\phi(p))=k-1+2\sum_{m=1}^{k-1}\text{Card}\,\{i: i_m<i<i_{m+1}\text{ and }\ \pi_{i_1}>\pi_i>\pi_{i_{m+1}}\}.$$

In particular, $\operatorname{inv}(\phi(p))<\operatorname{inv}(p)$. By symmetry, together with the fact that $\operatorname{inv}(\pi^{-1})=\operatorname{inv}(\pi)$, it follows that $\operatorname{inv}(\psi(p))<\operatorname{inv}(p)$ too.

We encode the compositions of the maps ϕ and ψ by words on the alphabet $\{\phi,\psi\}$. For instance, if u is the word $\phi\psi^2$, then $u(p)=\phi\psi^2(p)$. Let us prove, by induction on $\operatorname{inv}(p)$, the following two statements:
1. If u and v are two words such that $u(p)$ and $v(p)$ avoid $k\cdots 21$, then $u(p)=v(p)$.
2. Moreover, if u and v are minimal for this property (that is, for any non-trivial factorization $u=u_0u_1$, the placement $u_1(p)$ still contains an occurrence of $k\cdots 21$ — and similarly for v), then u and v have the same length.

If the first property holds for p, then $u(p)=v(p)=\phi^*(p)$. If the second property holds, we denote by $L(p)$ the length of any minimal word u such that $u(p)$ avoids $k\cdots 21$.

If π is the identity, then the two results are obvious. They remain obvious, with $L(p)=0$, if p does not contain any occurrence of $k\cdots 21$.

Now assume p contains such an occurrence, and $u(p)$ and $v(p)$ avoid $k\cdots 21$. By assumption, neither u nor v is the empty word. Let f (resp. g) be the rightmost letter of u (resp. v), that is, the *first* transformation that is applied to p in the evaluation of $u(p)$ (resp. $v(p)$). Write $u=u'f$ and $v=v'g$.

If $f(p)=g(p)$, let q be the placement $f(p)$. Given that $\operatorname{inv}(q)<\operatorname{inv}(p)$, and that the placements $u(p)=u'(q)$ and $v(p)=v'(q)$ avoid $k\cdots 21$, both statements follow by induction.

If $f(p)\neq g(p)$, we may assume, without loss of generality, that $f=\phi$ and $g=\psi$. Let $q_1=\phi(p)$, $q_2=\psi(p)$ and $q=\phi(\psi(p)=\psi(\phi(p))$ (Theorem 10). The induction

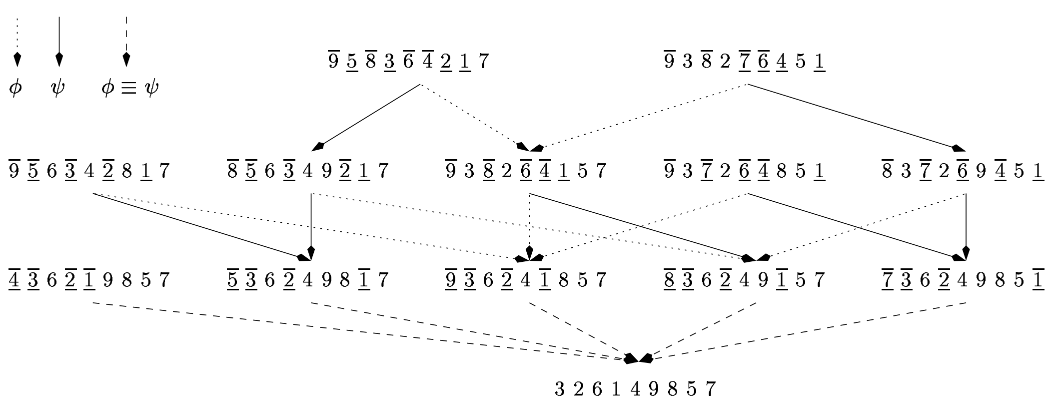
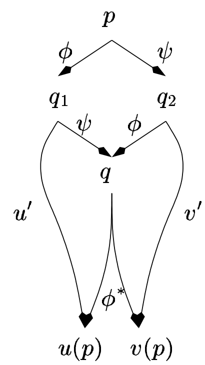


FIGURE 3. The action of ϕ and ψ on a part of $\mathcal{S}_9$, for $k=4$.

hypothesis, applied to q_1, gives $u'(q_1) = \phi^*(\psi(q_1)) = \phi^*(q)$, that is, $u(p) = \phi^*(q)$ (see the figure below). Similarly, $v'(q_2) = \phi^*(q_2) = \phi^*(q)$, that is, $v(p) = \phi^*(q)$. This proves the first statement. If u and v are minimal for p, then so are u' and v' for q_1 and q_2 respectively. By the first statement of Theorem 10, q_1 and q_2 still contain the pattern $k \cdots 21$, so $L(q) = L(q_1) - 1 = L(q_2) - 1$, and the words u' and v' have the same length. Consequently, u and v have the same length too. □



4. The local commutation theorem

We have reduced the proof of Theorem 1 to the proof of the local commutation theorem, Theorem 10. The rest of the paper is devoted to this proof, which turns out to be unexpectedly complicated. There is no question that one needs to find a more illuminating description of ϕ^*, or of $\phi \circ \psi$, which makes Theorems 7 and 10 clear.

In the full version of the paper, available on the arXiv [7], this theorem is first proved for permutations, and then extended to placements. In this extended abstract, we simply study one big example, and use it to describe the main steps of the proof (for permutations). This example is illustrated in Figure 4.

Example. Let π be the following permutation of length 21:

$$\pi = 17\ 21\ 20\ 16\ 19\ 18\ 13\ 15\ 11\ 14\ 12\ 8\ 10\ 9\ 7\ 4\ 2\ 6\ 5\ 3\ 1.$$

1. Let $k = 12$. The $\mathcal{A}$-sequence of π is

$$\mathcal{A}(\pi) = 17\ 16/15\ 14\ 12\ 10\ 9\ 7/6\ 5\ 3\ 1,$$

while its $\mathcal{B}$-sequence is

$$\mathcal{B}(\pi) = 21\ 20\ 19\ 18/15\ 14\ 12\ 10\ 9\ 7/4\ 2.$$

Observe that the intersection of $\mathcal{A}(\pi)$ and $\mathcal{B}(\pi)$ (delimited by '/') consists of the letters 15 14 12 10 9 7, and that they are consecutive both in $\mathcal{A}(\pi)$ and $\mathcal{B}(\pi)$. Also, $\mathcal{B}$ contains more letters than $\mathcal{A}$ before this intersection, while $\mathcal{A}$ contains more letters than $\mathcal{B}$ after the intersection. In the full version of the paper, we prove that this is always true.

2. Let us now apply the $\mathcal{B}$-shift to π. One finds:

$$\psi(\pi) = 17\ 20\ 19\ 16\ 18\ 15\ 13\ 14\ 11\ 12\ 10\ 8\ 9\ 7\ 4\ 2\ 21\ 6\ 5\ 3\ 1.$$

The new $\mathcal{A}$-sequence is now $\mathcal{A}(\psi(\pi)) = 17\ 16/15\ 13\ 11\ 10\ 8\ 7/6\ 5\ 3\ 1$. Observe that all the letters of $\mathcal{A}(\pi)$ that were before or after the intersection with $\mathcal{B}(\pi)$ are still in the new $\mathcal{A}$-sequence, as well as the first letter of the intersection. We prove that this is always true. In this example, the last letter of the intersection is still in the new $\mathcal{A}$-sequence, but this is not true in general.

By symmetry with respect to the main diagonal, after the $\mathcal{A}$-shift, the letters of $\mathcal{B}$ that were before or after the intersection are in the new $\mathcal{B}$-sequence, as well as the first letter of $\mathcal{A}$ following the intersection. This can be checked on our example:

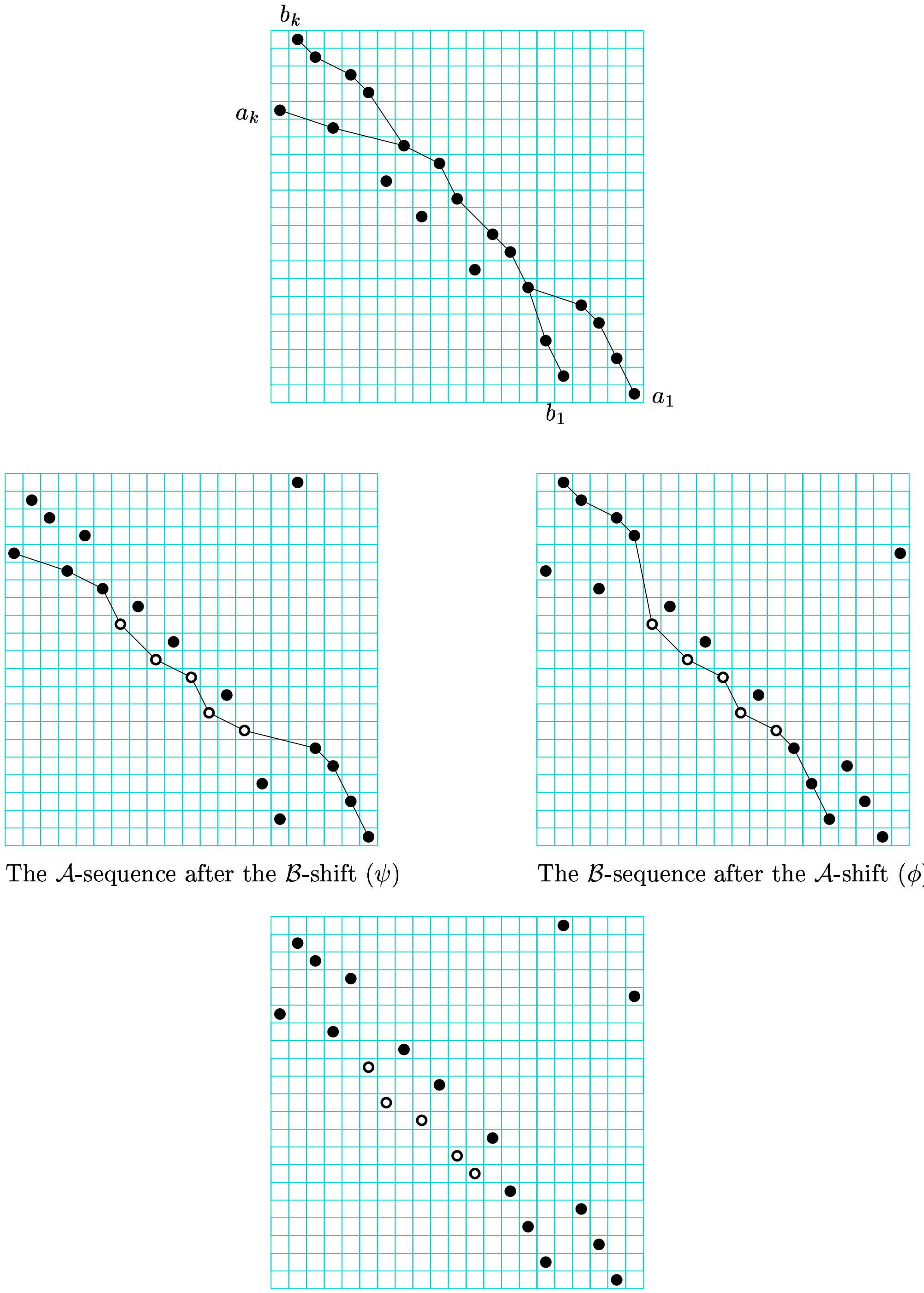
$$\phi(\pi) = 16\ 21\ 20\ 15\ 19\ 18\ 13\ 14\ 11\ 12\ 10\ 8\ 9\ 7\ 6\ 4\ 2\ 5\ 3\ 1\ 17,$$

and the new $\mathcal{B}$-sequence is $\mathcal{B}(\phi(\pi)) = 21\ 20\ 19\ 18/13\ 11\ 10\ 8\ 7\ 6\ /4\ 2$.

3. Let $a_i = b_j$ denote the first (leftmost) point in $\mathcal{A}(\pi) \cap \mathcal{B}(\pi)$, and let $a_d = b_e$ be the last point in this intersection. We have seen that after the $\mathcal{B}$-shift, the new $\mathcal{A}$-sequence begins with $a_k \cdots a_i = 17\ 16\ 15$, and ends with $a_{d-1} \cdots a_1 = 6\ 5\ 3\ 1$. The letters in the center of the new $\mathcal{A}$-sequence, that is, the letters replacing $a_{i-1} \cdots a_d$, are $x_{i-1} \cdots x_d = 13\ 11\ 10\ 8\ 7$. Similarly, after the $\mathcal{A}$-shift, the new $\mathcal{B}$-sequence begins with $b_k \cdots b_{j+1} = 21\ 20\ 19\ 18$, and ends with $a_{d-1} b_{e-1} \cdots b_1 = 6\ 4\ 2$. The central letters are again $x_{i-1} \cdots x_d = 13\ 11\ 10\ 8\ 7$! (See Figure 4). This is not a coincidence; we prove in [7] that this always holds.

4. We finally combine all these properties to describe explicitly how the maps $\phi \circ \psi$ and $\psi \circ \phi$ act on a permutation π, and conclude that they yield the same permutation if the $\mathcal{A}$- and $\mathcal{B}$-sequences of π do not coincide.

Acknowledgements. We thank Yves Métivier and Gérard Huet for references on the confluence of rewriting systems, and Olivier Guibert for interesting discussions on pattern avoiding involutions.



The $\mathcal{A}$-sequence after the $\mathcal{B}$-shift (ψ)

The $\mathcal{B}$-sequence after the $\mathcal{A}$-shift (ϕ)

FIGURE 4. Top: A permutation π, with its $\mathcal{A}$- and $\mathcal{B}$-sequences shown. Left: After the $\mathcal{B}$-shift. Right: After the $\mathcal{A}$-shift. Bottom: After the composition of ϕ and ψ.

References

[1] E. Babson and E. Steingrímsson. Generalized permutation patterns and a classification of the Mahonian statistics. *Sém. Lothar. Combin.*, 44:Art. B44b, 18 pp. (electronic), 2000.

[2] E. Babson and J. West. The permutations $123p_4 \cdots p_m$ and $321p_4 \cdots p_m$ are Wilf-equivalent. *Graphs Combin.*, 16(4):373–380, 2000.

[3] J. Backelin, J. West, and G. Xin. Wilf-equivalence for singleton classes. In H. Barcelo and V. Welker, editors, *Proceeedings of the 13th Conference on Formal Power Series and Algebraic Combinatorics*, pages 29–38, Arizona State University, May 2001. To appear in *Adv. in Appl. Math.*

[4] E. Barcucci, A. Del Lungo, E. Pergola, and R. Pinzani. Some permutations with forbidden subsequences and their inversion number. *Discrete Math.*, 234(1-3):1–15, 2001.

[5] M. Bóna. Exact enumeration of 1342-avoiding permutations: a close link with labeled trees and planar maps. *J. Combin. Theory Ser. A*, 80(2):257–272, 1997.

[6] M. Bousquet-Mélou. Four classes of pattern-avoiding permutations under one roof: Generating trees with two labels. *Electron. J. Combin.*, 9(2):Research Paper 19, 2003.

[7] M. Bousquet-Mélou and E. Steingrímsson. Decreasing subsequences in permutations and Wilf equivalence for involutions. Technical report, ArXiv math.CO/0405334, 2004.

[8] A. Claesson and T. Mansour. Enumerating permutations avoiding a pair of Babson-Steingrímsson patterns. *Ars Combinatoria*, to appear.

[9] I. Gessel. Symmetric functions and P-recursiveness. *J. Combin. Theory Ser. A*, 53(2):257–285, 1990.

[10] I. Gessel, J. Weinstein, and H. S. Wilf. Lattice walks in $\mathbf{Z}^d$ and permutations with no long ascending subsequences. *Electron. J. Combin.*, 5(1):Research Paper 2, 11 pp., 1998.

[11] D. Gouyou-Beauchamps. Standard Young tableaux of height 4 and 5. *European J. Combin.*, 10(1):69–82, 1989.

[12] O. Guibert. *Combinatoire des permutations à motifs exclus, en liaison avec mots, cartes planaires et tableaux de Young.* PhD thesis, LaBRI, Université Bordeaux 1, 1995.

[13] O. Guibert, E. Pergola, and R. Pinzani. Vexillary involutions are enumerated by Motzkin numbers. *Ann. Comb.*, 5(2):153–174, 2001.

[14] G. Huet. Confluent reductions: abstract properties and applications to term rewriting systems. *J. Assoc. Comput. Mach.*, 27(4):797–821, 1980.

[15] A. D. Jaggard. Prefix exchanging and pattern avoidance for involutions. *Electron. J. Combin.*, 9(2):Research Paper 16, 2003.

[16] S. Kitaev and T. Mansour. A survey of certain pattern problems. Submitted, 2004.

[17] C. Krattenthaler. Permutations with restricted patterns and Dyck paths. *Adv. in Appl. Math.*, 27(2-3):510–530, 2001.

[18] A. Regev. Asymptotic values for degrees associated with strips of Young diagrams. *Adv. in Math.*, 41(2):115–136, 1981.

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

[20] R. Simion and F. W. Schmidt. Restricted permutations. *European J. Combin.*, 6(4):383–406, 1985.

[21] Z. E. Stankova. Forbidden subsequences. *Discrete Math.*, 132(1-3):291–316, 1994.

[22] Z. E. Stankova. Classification of forbidden subsequences of length 4. *European J. Combin.*, 17(5):501–517, 1996.

[23] Z. E. Stankova and J. West. A new class of Wilf-equivalent permutations. *J. Algebraic Combin.*, 15(3):271–290, 2002.

[24] G. Viennot. Une forme géométrique de la correspondance de Robinson-Schensted. In *Combinatoire et représentation du groupe symétrique (Actes Table Ronde CNRS, Univ. Louis-Pasteur, Strasbourg, 1976)*, pages 29–58. Lecture Notes in Math., Vol. 579. Springer, Berlin, 1977.

[25] J. West. *Permutations with forbidden subsequences, and stack-sortable permutations.* PhD thesis, MIT, 1990.

CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France and Matematik, Chalmers tekniska högskola och Göteborgs universitet S-412 96 Göteborg, Sweden

E-mail address: `mireille.bousquet@labri.fr, einar@math.chalmers.se`

STABLE GROTHENDIECK POLYNOMIALS AND K-THEORETIC FACTOR SEQUENCES

ANDERS SKOVSTED BUCH, ANDREW KRESCH, MARK SHIMOZONO, HARRY TAMVAKIS, AND ALEXANDER YONG

Abstract. We give a nonrecursive combinatorial formula for the expansion of a stable Grothendieck polynomial in the basis of stable Grothendieck polynomials for partitions. The proof is based on a generalization of the Edelman-Greene insertion algorithm. This result is applied to prove a number of formulas and properties for K-theoretic quiver polynomials and Grothendieck polynomials. In particular we formulate and prove a K-theoretic analogue of Buch and Fulton's factor sequence formula for the cohomological quiver polynomials.

1. Introduction

1.1. **Stable Grothendieck polynomials.** For each permutation w there is a symmetric power series $G_w = G_w(x_1, x_2, \dots)$ called the stable Grothendieck polynomial for w. These power series were defined by Fomin and Kirillov [13, 12] as a limit of the ordinary Grothendieck polynomials of Lascoux and Schützenberger [18]. We recall this definition in Section 2. The term of lowest degree in G_w is the Stanley function (or stable Schubert polynomial) F_w. The *Stanley coefficients* in the Schur expansion of a Stanley function are interesting combinatorial invariants which generalize the Littlewood-Richardson coefficients. It was proved by Edelman and Greene [10] and Lascoux and Schützenberger [19] that Stanley coefficients are nonnegative. Various combinatorial rules have been given for these coefficients [11, 15, 23].

Given a partition $\lambda = (\lambda_1 \geq \cdots \geq \lambda_k \geq 0)$, the *Grassmannian permutation* w_λ for λ is uniquely defined by the requirement that $w_\lambda(i) = i + \lambda_{k+1-i}$ for $1 \leq i \leq k$ and $w_\lambda(i) < w_\lambda(i+1)$ for $i \neq k$. The power series $G_\lambda := G_{w_\lambda}$ play a role in combinatorial K-theory similar to the role of Schur functions in cohomology. Buch has shown [3] that any stable Grothendieck polynomial G_w can be written as a finite linear combination

$$G_w = \sum_\lambda c_{w,\lambda} G_\lambda \tag{1}$$

of stable Grothendieck polynomials indexed by partitions, using integer coefficients $c_{w,\lambda}$ that generalize the Stanley coefficients [2]. Lascoux gave a recursive formula for stable Grothendieck polynomials which confirms a conjecture that these coefficients have signs that alternate with degree, i.e. $(-1)^{|\lambda|-\ell(w)} c_{w,\lambda} \geq 0$ [17]. The central result of this paper is a new formula for the coefficients $c_{w,\lambda}$ which generalizes Fomin and Greene's rule [11] for Stanley coefficients.

Date: November 28, 2004.

To state our formula, we need the 0-*Hecke monoid*, which is the quotient of the free monoid of all finite words in the alphabet $\{1, 2, \dots\}$ by the relations

$$p\,p \equiv p \qquad \text{for all } p \tag{2}$$

$$p\,q\,p \equiv q\,p\,q \qquad \text{for all } p, q \tag{3}$$

$$p\,q \equiv q\,p \qquad \text{for } |p-q| \geq 2. \tag{4}$$

There is a bijection between the 0-Hecke monoid and the infinite symmetric group $S_\infty = \bigcup_{n \geq 1} S_n$. Given any word a there is a unique permutation $w \in S_\infty$ such that $a \equiv b$ for some (or equivalently every) reduced word b of w. In this case we write $w(a) = w$ and say that a is a *Hecke word* for w. Notice that the reduced words for w are precisely the Hecke words for w that are of minimum length. The *Hecke product* $u \cdot v$ of two permutations $u, v \in S_\infty$ is the element $w(ab) \in S_\infty$ where a and b are words such that $w(a) = u$ and $w(b) = v$.

We use the English notation for partitions and tableaux. A *decreasing tableau*[1] is a Young tableau whose rows decrease strictly from left to right, and whose columns decrease strictly from top to bottom. The (row reading) *word* of a tableau is obtained by reading the rows of the tableau from left to right, starting with the bottom row, followed by the next-to-bottom row, etc. We shall identify a tableau with its word.

Theorem 1. *For any permutation w we have*

$$G_w = \sum_\lambda c_{w,\lambda} G_\lambda$$

where $c_{w,\lambda}$ is equal to $(-1)^{|\lambda|-\ell(w)}$ times the number of decreasing tableaux T of shape λ such that $w(T) = w$.

Example 2. Let $w = 31524$. The decreasing tableaux that are Hecke words for w are:

$$\begin{array}{|c|c|}\hline 4 & 3 \\ \hline 2 & 1 \\ \hline \end{array} \qquad \begin{array}{|c|c|c|}\hline 4 & 3 & 1 \\ \hline 2 \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|c|c|}\hline 4 & 3 & 1 \\ \hline 2 & 1 \\ \cline{1-2} \end{array}$$

So $G_w = G_{22} + G_{31} - G_{32}$.

When the permutation w is 321-avoiding, Theorem 1 also generalizes Buch's rule for the coefficients $c_{w,\lambda}$ in terms of *set-valued* tableaux [3], in the sense that there is an explicit bijection between the relevant decreasing and set-valued tableaux. As a consequence, we obtain a new proof of the set-valued Littlewood-Richardson rule for K-theoretic Schubert structure constants on Grassmannians, as well as an alternative rule based on decreasing tableaux.

1.2. **Hecke insertion.** Fomin and Kirillov proved that the monomial coefficients of (stable) Grothendieck polynomials are counted by combinatorial objects called resolved wiring diagrams (also known as FK-graphs, pipe dreams, or nonreduced RC-graphs) [13, 12]. This formula was used in [3] to express the monomial coefficients of stable Grothendieck polynomials for partitions in terms of set-valued tableaux. We prove Theorem 1 by exhibiting an explicit bijection between the set of FK-graphs for a permutation w and the set of pairs (T, U) where T is a decreasing tableau with $w(T) = w$ and U is a set-valued tableau of the same shape as T. This

[1] The use of decreasing tableaux rather than increasing, is merely for convenience in the definition of a K-theoretic factor sequence.

bijection, called *Hecke insertion*, is the technical core of our paper. It is a subtle extension of the Edelman-Greene insertion algorithm from the set of reduced words to the set of all (Hecke) words.

Hecke insertion allows us to define a product of decreasing tableaux $(T_1, T_2) \mapsto T_1 \cdot T_2$ (see section 3.5). This product is used in the definition of K-theoretic factor sequences in the next section.

1.3. **Quiver coefficients.** Our main application of Theorem 1 concerns quiver coefficients. A sequence of vector bundle morphisms $E_0 \to E_1 \to \cdots \to E_n$ over a variety X together with a set of *rank conditions* $r = \{r_{ij}\}$ for $0 \leq i \leq j \leq n$ define a *quiver variety* $\Omega_r \subset X$ of points where each composition of bundle maps $E_i \to E_j$ has rank at most r_{ij}. We demand that the rank conditions can *occur*, which is equivalent to the requirement that $r_{ii} = e_i := \operatorname{rank}(E_i)$ for all i, $0 \leq r_{ij} \leq \min(r_{i,j-1}, r_{i+1,j})$ for $0 \leq i < j \leq n$, and $r_{ij} + r_{i-1,j+1} \geq r_{i-1,j} + r_{i,j+1}$ for $0 < i \leq j < n$. We also demand that the bundle maps are generic, so that the quiver variety Ω_r obtains its expected codimension $d(r) = \sum_{i<j}(r_{i,j-1} - r_{ij})(r_{i+1,j} - r_{ij})$. Buch and Fulton proved a formula for the cohomology class of Ω_r [5] which was later generalized to K-theory by Buch [2]. The K-theory version states that the Grothendieck class of Ω_r is given by

$$(5) \qquad [\mathcal{O}_{\Omega_r}] = \sum_{\mu} c_\mu(r)\, G_{\mu_1}(E_1 - E_0) G_{\mu_2}(E_2 - E_1) \cdots G_{\mu_n}(E_n - E_{n-1})\,,$$

where the sum is over sequences $\mu = (\mu_1, \ldots, \mu_n)$ of partitions μ_i such that $\sum |\mu_i| \geq d(r)$ and each partition μ_i can be contained in the rectangle $e_i \times e_{i-1}$ with e_i rows and e_{i-1} columns. The coefficients $c_\mu(r)$ in this formula are integers called *quiver coefficients*. When $\sum |\mu_i| = d(r)$, the coefficient $c_\mu(r)$ also appears in the cohomology formula from [5] and is called a *cohomological quiver coefficient*. It was conjectured that cohomological quiver coefficients are nonnegative, while K-theoretic quiver coefficients have signs that alternate with degree, i.e. $(-1)^{\sum |\mu_i| - d(r)} c_\mu(r) \geq 0$. The conjecture for cohomological quiver coefficients was proved by Knutson, Miller, and Shimozono [16], after which the general case was proved by Buch [4] and Miller [21]. Both of the latter proofs are based on the *ratio formula* from [16], which expresses the Grothendieck class of a quiver variety as a ratio of two double Grothendieck polynomials.

A more precise conjecture for cohomological quiver coefficients was posed in [5], which asserts that any such coefficient $c_\mu(r)$ counts the number of *factor sequences* of tableaux with shapes given by the sequence of partitions μ. A factor sequence is a sequence of semistandard Young tableaux that can be obtained by performing a series of plactic factorizations and multiplications of chosen tableaux arranged in a *tableau diagram*. For a specific choice of tableau diagram, this more precise conjecture was also proved by Knutson, Miller and Shimozono [16]. It appears, however, that the original definition of factor sequences from [5] has no natural generalization to K-theory.

In this paper, we prove that K-theoretic quiver coefficients are counted by a new type of factor sequence. These sequences are constructed from a tableau diagram of decreasing tableaux using the same algorithm that defines the original factor sequences, except that the plactic product is replaced with a product $(U, T) \mapsto U \cdot T$ of decreasing tableaux which respects Hecke words (see section 3.5).

For each $0 \leq i < j \leq n$ let R_{ij} be a rectangle with $r_{i+1,j} - r_{ij}$ rows and $r_{i,j-1} - r_{ij}$ columns. Let U_{ij} be the unique decreasing tableau of shape R_{ij} such that the lower left box contains the number $r_{i,j-1}$, and each box is one larger than the box below it and one smaller than the box to the left of it. For example, if $r_{i,j-1} = 6$, $r_{i+1,j} = 5$, and $r_{ij} = 2$ then

$$U_{ij} \;=\; \begin{array}{|c|c|c|c|}\hline 8&7&6&5\\\hline 7&6&5&4\\\hline 6&5&4&3\\\hline\end{array}\;.$$

These tableaux U_{ij} can be arranged in a triangular tableau diagram as in [5, §4]. We define a K*-theoretic factor sequence* for the rank conditions r by induction on n. If $n = 1$ then the only factor sequence is the sequence (U_{01}) consisting of the only tableau in the tableau diagram. If $n \geq 2$ then the numbers $\overline{r} = \{\overline{r}_{ij} : 0 \leq i \leq j \leq n-1\}$ defined by $\overline{r}_{ij} = r_{i,j+1}$ form a valid set of rank conditions corresponding to a sequence of $n-1$ bundle maps. In this case, a factor sequence for r is any sequence of the form $(U_{01} \cdot A_1, \dots, B_{i-1} \cdot U_{i-1,i} \cdot A_i, \cdots, B_{n-1} \cdot U_{n-1,n})$, for a choice of decreasing tableaux A_i and B_i such that $(A_1 \cdot B_1, \dots, A_{n-1} \cdot B_{n-1})$ is a factor sequence for $\overline{r}$.

Theorem 3. *The K-theoretic quiver coefficient $c_\mu(r)$ is equal to $(-1)^{\sum |\mu_i| - d(r)}$ times the number of K-theoretic factor sequences $(T_1, \dots, T_n)$ for the rank conditions r, such that T_i has shape μ_i for each i.*

Using results about Demazure characters it was proved in [16] that cohomological quiver coefficients are special cases of the Stanley coefficients associated to the *Zelevinsky permutation* $z(r)$ [24, 16]. We recall the definition of this permutation in Section 4. In this paper we prove more generally that the K-theoretic quiver coefficients are special cases of the coefficients $c_{z(r),\lambda}$ in the expansion (1) of the stable Grothendieck polynomial for $z(r)$. This result also sharpens the fact from [4, 8] that quiver coefficients are special cases of the decomposition coefficients of Grothendieck polynomials studied in [6] (see section 1.4.1). Given a sequence of partitions $\mu = (\mu_1, \dots, \mu_n)$ such that μ_i is contained in the rectangle $e_i \times e_{i-1}$, let $\lambda(\mu)$ be the partition obtained by concatenating the partitions $(e_0 + \cdots + e_{n-2-i})^{e_i} + \mu_{n-i}$ for $i = 0, \dots, n-1$. Our proof of the following identity is based on a bijection between the K-theoretic factor sequences for r and the decreasing tableaux representing $z(r)$.

Theorem 4. *For any set of rank conditions r and sequence of partitions μ we have $c_\mu(r) = c_{z(r),\lambda(\mu)}$.*

Central to the proof of the nonnegativity of cohomological quiver coefficients given in [16] is the *stable component formula*, which writes the cohomology class of a quiver variety as a sum of products of Stanley functions. This sum is over all *lace diagrams* representing the rank conditions r, which have the smallest possible number of crossings. The K-theoretic version of the component formula from [4, 21] states that

$$(6) \qquad [\mathcal{O}_{\Omega_r}] = \sum_{(w_1,\dots,w_n)} (-1)^{\sum \ell(w_i) - d(r)} G_{w_1}(E_1 - E_0) \cdots G_{w_n}(E_n - E_{n-1})$$

where the sum is over a generalization of minimal lace diagrams, which was named *KMS-factorizations* in [4]. We recall this definition in Section 4. The K-theoretic factor sequences also have the following characterization.

Theorem 5. *A sequence of decreasing tableaux $(T_1, \dots, T_n)$ is a K-theoretic factor sequence for the rank conditions r if and only if $(w(T_1), \dots, w(T_n))$ is a KMS-factorization for r.*

We will use the statement of Theorem 5 as our definition of K-theoretic factor sequences. When this definition is used, Theorem 3 is an immediate consequence of Theorem 1 combined with the K-theoretic stable component formula (6). The above inductive construction of factor sequences is then derived from a similar construction of KMS-factorizations proved in [4].

1.4. **Other applications.** We list other applications for the methods presented in this extended abstract that are not developed here but which will appear in the full version of this paper.

1.4.1. *Decomposition coefficients of Grothendieck polynomials.* Fulton's *universal Schubert polynomials* from [14] describe certain quiver varieties associated to a sequence of vector bundles $E_1 \to \cdots \to E_{n-1} \to E_n \to F_n \to F_{n-1} \to \cdots \to F_1$ over X, such that $\operatorname{rank}(E_i) = \operatorname{rank}(F_i) = i$ for each i. Given a permutation $w \in S_{n+1}$, we let $\Omega_w \subset X$ be the degeneracy locus of points where the rank of each composed map $E_q \to F_p$ is at most equal to the number of integers $1 \le i \le p$ such that $w(i) \le q$. The quiver formula (5) can be applied to give a formula

$$[\mathcal{O}_{\Omega_w}] = \sum_{\mu} c^{(n)}_{w,\mu} G_{\mu_1}(E_2 - E_1) \cdots G_{\mu_n}(F_n - E_n) \cdots G_{\mu_{2n-1}}(F_1 - F_2) \tag{7}$$

for the Grothendieck class of Ω_w, where the coefficients $c^{(n)}_{w,\mu}$ are special cases of quiver coefficients. It was shown in [2] that the coefficients $c_{w,\lambda}$ of the expansion (1) of the stable Grothendieck polynomial for w can be obtained as the specializations $c^{(n)}_{w,(\emptyset^{n-1},\lambda,\emptyset^{n-1})}$, where $\emptyset^{n-1}$ denotes a sequence of $n-1$ empty partitions. More generally, it was proved in [6, Thm. 4] that the coefficients $c^{(n)}_{w,\lambda}$ can be used to expand a double Grothendieck polynomial as a linear combination of products of stable Grothendieck polynomials applied to disjoint intervals of variables. In [6], the formula (7) was also used to prove that

$$[\mathcal{O}_{\Omega_w}] = \sum (-1)^{\ell(u_1 \cdots u_{2n-1} w)} G_{u_1}(E_2 - E_1) \cdots G_{u_n}(F_n - E_n) \cdots G_{u_{2n-1}}(F_1 - F_2)$$

where this sum is over all sequences of permutations $(u_1, \dots, u_{2n-1})$ such that $u_i \in S_{\min(i,2n-i)+1}$ and w is equal to the Hecke product $u_1 \cdot u_2 \cdots u_{2n-1}$. Combining this with Theorem 1, we obtain the following generalization of [7, Thm. 1].

Theorem 6. *The coefficient $c^{(n)}_{w,\mu}$ of (7) is equal to $(-1)^{\sum |\mu_i| - \ell(w)}$ times the number of sequences $(T_1, \dots, T_{2n-1})$ of decreasing tableaux of shapes $(\mu_1, \dots, \mu_{2n-1})$, such that the entries of T_i are at most $\min(i, 2n-i)$ and $w(T_1 T_2 \cdots T_n) = w$.*

1.4.2. *Expansion of Grothendieck polynomials.* Theorem 1 may be refined to give an expansion of Grothendieck polynomials. The cohomological analogue is the combinatorial rule [20, 22, 23] for the expansion of a Schubert polynomial as a positive sum of Demazure characters [9]. Taking a suitable limit, the Schubert polynomial becomes a Stanley function and each Demazure character becomes a Schur function. Using divided difference operators, one may introduce a new basis of $\mathbb{Z}[x_1, x_2, \dots]$ called Grothendieck-Demazure polynomials. We have a conjectural expansion of a Grothendieck polynomial as an alternating sum of Grothendieck-Demazure polynomials. In the limit this expansion becomes that in Theorem 1.

2. Grothendieck polynomials

2.1. **Definition.** Lascoux and Schützenberger's original definition of Grothendieck polynomials was based on divided difference operators [19]. In this paper we will use Fomin and Kirillov's construction of these polynomials [12], in notation that generalizes Billey, Jockusch, and Stanley's formula for Schubert polynomials [1].

Define a *compatible pair* to be a pair (a, i) of words $a = a_1 a_2 \cdots a_p$ and $i = i_1 i_2 \cdots i_p$, such that $i_1 \leq i_2 \leq \cdots \leq i_p$, and so that $i_j < i_{j+1}$ whenever $a_j \leq a_{j+1}$. For $w \in S_\infty$, the *stable Grothendieck polynomial* for w is given by [12]

$$G_w = \sum_{(a,i)} (-1)^{\ell(i)-\ell(w)} x^i \tag{8}$$

where the sum is over all compatible pairs (a, i) such that $w(a) = w$. Here $\ell(i)$ is the common length of a and i, and $x^i = x_{i_1} x_{i_2} \cdots x_{i_{\ell(i)}}$. The ordinary *Grothendieck polynomial* $\mathfrak{G}_w$ is given by the same sum (8), but only including the compatible pairs (a, i) for which $a_j \geq i_j$ for each j. The *Schubert polynomial* for w is equal to the lowest term of $\mathfrak{G}_w$, while the *Stanley function* F_w is the lowest term of G_w.

We also require Buch's formula [3] for the stable Grothendieck polynomial G_λ. A *set-valued tableau* of shape λ is a filling of the boxes of λ with finite nonempty sets of positive integers, such that these sets are weakly increasing along rows and strictly increasing down columns. In other words, all integers in a box must be smaller than or equal to the integers in the box to the right of it, and strictly smaller than the integers in the box below it. For a set-valued tableau S, let x^S denote the monomial where the exponent of x_i is equal to the number of boxes containing the integer i, and let $|S|$ be the degree of this monomial. We have [3]

$$G_\lambda = \sum_S (-1)^{|S|-|\lambda|} x^S \tag{9}$$

where S runs over all set-valued tableaux of shape λ.

2.2. **The required bijection.** For any permutation $w \in S_n$, it follows from (8) and the symmetry of stable Grothendieck polynomials that $G_w = G_{w_0 w^{-1} w_0}$, where $w_0 = w_0^{(n)}$ is the longest permutation in S_n. Theorem 1 is therefore equivalent to the following statement. Define an *increasing tableau* to be a Young tableau with strictly increasing rows and columns.

Theorem 7. *The coefficient $c_{w,\lambda}$ is equal to $(-1)^{|\lambda|-\ell(w)}$ times the number of increasing tableaux T of shape λ such that $w(T) = w^{-1}$.*

In light of (8) and (9), to prove this theorem, it suffices to establish a bijection $(a, i) \mapsto (T, U)$ between all compatible pairs (a, i) such that $w(a) = w$, and all pairs of tableaux (T, U) of the same shape, such that T is increasing with $w(T) = w^{-1}$ and U is set-valued. In addition, this bijection must satisfy that $x^U = x^i$.

3. Hecke Insertion

Let M_1 be the set of pairs (Y, x) where Y is an increasing tableau and x is a letter. Let M_2 be the set of triples (Z, r, α) where Z is an increasing tableau, Z has a corner cell (r, c) in the r-th row, and $\alpha \in \{0, 1\}$. *Hecke insertion* defines a bijection

$$\begin{aligned} \Phi : M_1 &\to M_2 \\ (Y, x) &\mapsto (Z, r, \alpha) \end{aligned} \tag{10}$$

such that either

(1) $\alpha = 0$ and shape$(Z) =$ shape(Y).
(2) $\alpha = 1$ and shape$(Z) =$ shape$(Y) \bigsqcup \{(r, c)\}$.

This bijection defines the Hecke insertion of x into the increasing tableau Y, resulting in the increasing tableau Z, ending at the corner cell (r, c) of Z. Unlike ordinary Schensted insertion, it is possible for a Hecke insertion not to add a cell to the tableau: a new cell is created if and only if $\alpha = 1$.

3.1. **Hecke (Row) Insertion.** Let $(Y, x') \in M_1$. Inductively for $i \geq 1$, suppose that the first $i - 1$ rows of Y have been suitably modified, and that the number x is being inserted into the i-th row. Let y be the smallest entry in the i-th row such that $y > x$. If no such element exists, set $y = \infty$ and let z be the last entry in the i-th row. If the i-th row is empty then set $z = 0$.

(H1) If $y = \infty$ and $z = x$: The insertion terminates. Let (r, c) be the corner cell in the same column as z and $\alpha = 0$.
(H2) If $y = \infty$ and $z < x$: The insertion terminates. (a) If adjoining x to the end of the i-th row, results in an increasing tableau, then do so, with $r = i$ and $\alpha = 1$. (b) If not (and this can only happen if $i > 1$), let (r, c) be the corner cell in the same column as z and $\alpha = 0$.
(H3) If $y < \infty$, replace y by x if the result is an increasing tableau and otherwise leave the row unchanged. Continue by inserting y into the $(i + 1)$-th row.

Let Z be the resulting increasing tableau. This algorithm produces a triple $\Phi(Y, x') = (Z, r, \alpha) \in M_2$. Write $Z = (Y \xleftarrow{H} x')$.

Remark 8. A increasing tableau is produced in cases (H2) and (H3) unless x would be placed just to the right of, or just below, another x.

Example 9.

$$\begin{ytableau} 1 & 2 & 3 & 5 \\ 2 & 5 \\ 3 & 6 \\ 4 \end{ytableau} \xleftarrow{H} 3 = \begin{ytableau} 1 & 2 & 3 & 5 \\ 2 & 5 \\ 3 & 6 \\ 4 \end{ytableau}$$

3 is inserted into the first row, which contains 3. So 5 is inserted into the second row, whose largest value is $z = 5$. This is case (H1). Then $\alpha = 0$ and $r = 3$, since $(3, 2)$ is the cell at the bottom of the column of z.

Example 10.

$$\begin{ytableau} 1 & 2 & 4 \\ 2 & 5 \end{ytableau} \xleftarrow{H} 2 = \begin{ytableau} 1 & 2 & 4 \\ 2 & 4 \\ 5 \end{ytableau}$$

2 is inserted into the first row, which contains 2. 4 is inserted into the second row, displacing the 5. The 5 is inserted into the third row, where it comes to rest. This is case (H2a). Then $\alpha = 1$ and $r = 3$.

Example 11.

$$\begin{array}{|c|c|c|}\hline 1&2&4\\\hline 2&3&\\\cline{1-2} 3&4&\\\cline{1-2} 5&&\\\cline{1-1}\end{array} \xleftarrow{H} 2 = \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 2&3&\\\cline{1-2} 3&4&\\\cline{1-2} 5&&\\\cline{1-1}\end{array}$$

2 is inserted into the first row, which contains a 2. 4 is inserted into the second row, which has largest entry $z = 3$. 4 can't come to rest at the cell $(2, 3)$ since that is just below the 4 in cell $(1, 3)$. Case (H2b) holds. Then $\alpha = 0$ and $r = 3$ because $(3, 2)$ is the cell at the bottom of the column of z.

Example 12.

$$\begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline 3&5\\\hline\end{array} \xleftarrow{H} 1 = \begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline 3&5\\\hline 5&\\\cline{1-1}\end{array}$$

1 is inserted into the first row, which already contains a 1. So 3 is inserted into the second row. It would have replaced 4, but this replacement would place a 3 directly below another 3, violating the increasing tableau condition. So the second row is unchanged and 4 is inserted into the third row. Similarly 4 cannot replace 5. So 5 is inserted into the fourth row, where it comes to rest in the cell $(4, 1)$ with $\alpha = 1$.

3.2. **Reverse Hecke insertion.** The inverse map $\Psi : M_2 \to M_1$ is defined as follows. Let $(Z, r, \alpha) \in M_2$, (r, c) the corner cell in the r-th row of Z, and $y = Z_{r,c}$. If $\alpha = 1$ then remove y. In any case, reverse insert y up into the previous row.

Whenever a value y is reverse inserted into a row, let x be the largest entry in the row such that $x < y$. If replacing x by y yields an increasing tableau then do so; otherwise leave the row unchanged. In any case, reverse insert x into the previous row.

Eventually a value x' reverse inserted out of the first row, leaving behind an increasing tableau Y. Call x' the output value. Define $\Psi(Z, r, \alpha) = (Y, x')$.

Remark 13. Note that the only obstructions for replacing x by y, are when the entry below or to the right of x already contains y.

Example 14. Let us apply reverse Hecke insertion to the tableau computed in Example 12 at the cell $(4, 1)$ with $\alpha = 1$. Since $\alpha = 1$ the entry 5 in cell $(4, 1)$ is removed. Then 5 is reverse inserted into the third row. Since 5 is already in the third row, the third row remains unchanged and 3 is reverse inserted into the second row. 3 cannot replace 2 because this would place a 3 directly atop a 3, creating a vertical violation of the increasing tableau condition. The second row is unchanged and 2 is reverse inserted into the first row. 2 cannot replace 1 for the same reason. The first row is unchanged and 1 is the output value. This recovers the initial tableau of Example 12.

Proposition 15. *The maps Φ and Ψ are mutually inverse bijections.*

3.3. **Properties of Hecke insertion.** Hecke insertion respects Hecke words.

Lemma 16. *Suppose reverse Hecke insertion of the tableau T at some corner cell results in the tableau T' and the output value x. Then $w(T) = w(T'x)$.*

Hecke insertion also satisfies the following Pieri property.

Lemma 17. *Suppose we first reverse Hecke insert starting from one corner C_1 of T, and then reverse Hecke insert from a corner C_2 of the modification of T. Then the first output value is strictly smaller than the second output value if and only if C_1 is strictly lower than C_2.*

3.4. **Proof of Theorem 7 via column Hecke Robinson-Schensted.** In this section we give the bijection that was sought in section 2.2. We may define Hecke column insertion by switching the roles of rows and columns in Hecke row insertion. Write $\Phi' : M_1 \to M_2$ for this bijection.

Let (a, i) be as in section 2.2 with $a = a_1 a_2 \cdots a_p$ and $i = i_1 i_2 \cdots i_p$. We start with the empty tableau pair $(T_0, U_0) = (\varnothing, \varnothing)$. If (T_{j-1}, U_{j-1}) has been defined for some $j \geq 1$, let $(T_j, s_j, \alpha_j) = \Phi'(T_{j-1}, a_j)$. Let U_j be obtained from U_{j-1} by adjoining a new cell to the end of the s_j-th row containing the singleton set $\{i_j\}$ if $\alpha_j = 1$. Otherwise U_j is obtained from U_{j-1} by putting i_j into the existing set in the corner cell in row s_j. Define $(T, U) = (T_p, U_p)$. The map $(a, i) \mapsto (T, U)$ has the desired properties. U is a set-valued tableau by Lemma 17 and $x^i = x^U$ by definition. The fact that $w(T) = w^{-1}$ follows from Lemma 16 combined with the fact that the reversal of a word gives a bijection between the Hecke words for w and those for w^{-1}. This proves Theorems 7 and 1.

3.5. **Product of decreasing tableaux.** For use with factor sequences, we define the product of the decreasing tableaux T_1 and T_2. Consider the variant of Hecke insertion in which larger numbers bump smaller numbers. In other words, we reverse the order of the positive integers in the algorithm of Section 3.1. Let $T_1 \cdot T_2$ be the decreasing tableau obtained by inserting the word of T_2 into T_1 using this variant of Hecke insertion. More precisely, if $a_1 a_2 \cdots a_p$ is the word of T_2 then $T_1 \cdot T_2 = (((T_1 \xleftarrow{H} a_1) \xleftarrow{H} a_2) \cdots) \xleftarrow{H} a_p$. This product has the following properties.

Lemma 18. *(1) For decreasing tableaux T_1, T_2 we have $w(T_1 \cdot T_2) = w(T_1) \cdot w(T_2)$.*

(2) Suppose a decreasing tableau T is cut along a vertical line into T_{left} and T_{right}. Then $T = T_{left} \cdot T_{right}$.

(3) Suppose T is cut along a horizontal line into tableaux T_{bottom} and T_{top}. Then $T = T_{bottom} \cdot T_{top}$.

Our applications to factor sequences require that the product of decreasing tableaux satisfies the properties of this lemma. When the concatenation of the words of T_1 and T_2 is a reduced word of a permutation, then these conditions imply that $T_1 \cdot T_2$ agrees with the Coxeter-Knuth product, but no such uniqueness statement holds in general. The product $T_1 \cdot T_2$ also fails to be associative.

4. Quiver varieties

Let $r = \{r_{ij}\}$ be a set of rank conditions for $0 \leq i, j \leq n$, and set $N = e_0 + \cdots + e_n$ where $e_i = r_{ii}$. A result of Zelevinsky shows that when the base variety X is a product of matrix spaces, the quiver variety $\Omega_r \subset X$ is identical to a dense open subset of a Schubert variety [24]. The *Zelevinsky permutation* corresponding to this Schubert variety was used in [16] to prove the ratio formula for quiver varieties.

With the notation from [4], the Zelevinsky permutation can be constructed as a product of permutations as follows (see [16, Prop. 1.6] for a different construction). Extend the rank conditions $r = \{r_{ij}\}$ by setting $r_{ij} = e_j + \cdots + e_i$ for $0 \leq j < i \leq n$.

Then define decreasing tableaux U_{ij} as in the introduction, but for all $0 \le i < n$ and $0 < j \le n$. The corresponding permutation $W_{ij} = w(U_{ij})$ is given by

$$W_{ij}(p) = \begin{cases} p + r_{i,j-1} - r_{ij} & \text{if } r_{ij} < p \le r_{i+1,j} \\ p - r_{i+1,j} + r_{ij} & \text{if } r_{i+1,j} < p \le r_{i+1,j} + r_{i,j-1} - r_{ij} \\ p & \text{otherwise.} \end{cases}$$

The Zelevinsky permutation can now be defined by $z(r) = \prod_{j=1}^{n} \prod_{i=0}^{n-1} W_{ij}$.

For each $1 \le j \le n-1$ we set $\delta_j = W_{jj} W_{j+1,j} \cdots W_{n-1,j} \in S_N$. A *KMS-factorization* for the rank conditions r is any sequence $(w_1, \dots, w_n)$ of permutations with $w_i \in S_{e_{i-1}+e_i}$, such that the Zelevinsky permutation $z(r)$ is equal to the Hecke product

$$w_1 \cdot \delta_1 \cdot w_2 \cdot \delta_2 \cdots \delta_{n-1} \cdot w_n \,.$$

These sequences of permutations generalize the notion of a minimal lace diagram from [16] and give the index set in the K-theoretic stable component formula (6) from [4, 21].

We define a *K-theoretic factor sequence* for the rank conditions r to be any sequence $(T_1, \dots, T_n)$ of decreasing tableaux, such that $(w(T_1), \dots, w(T_n))$ is a KMS-factorization for r. As noted in the introduction, this definition means that Theorem 3 is a consequence of Theorem 1 combined with the stable component formula (6). To obtain the inductive definition of factor sequences we need the following result proved in [4, Thm. 7], which shows that KMS-factorizations can themselves be defined as 'factor sequences'.

Theorem 19. *(a) If $(w_1, \dots, w_n)$ is a KMS-factorization for r, then each permutation w_i has a reduced factorization $w_i = v_{i-1} \cdot W_{i-1,i} \cdot u_i$ with $v_{i-1} \in S_{e_{i-1}}$ and $u_i \in S_{e_i}$, such that $v_0 = u_n = 1$.*

(b) Let $u_1, v_1, \dots, u_{n-1}, v_{n-1}$ be permutations with $u_i, v_i \in S_{e_i}$. Then the sequence $(W_{01} \cdot u_1, v_1 \cdot W_{12} \cdot u_2, \dots, v_{n-1} \cdot W_{n-1,n})$ is a KMS-factorization for r if and only if $(u_1 \cdot v_1, u_2 \cdot v_2, \dots, u_{n-1} \cdot v_{n-1})$ is a KMS-factorization for $\overline{r}$.

We also need the following statement.

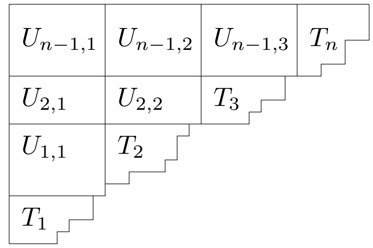
Lemma 20. *Let T be any decreasing tableau such that $w(T) \in S_m$, and for some integers $a, b < m$ we have $w(T)(p) \le b$ for all $a < p \le m$. Then T contains the rectangle $R = (m-a) \times (m-b)$ in its upper left corner. The upper-left box of R equals $m-1$, and the boxes of R decrease by one for each step down or to the right.*

Let $(U, T) \mapsto U \cdot T$ be the product of decreasing tableaux defined in section 3.5.

Corollary 21. *A sequence of decreasing tableaux $(T_1, \dots, T_n)$ is a K-theoretic factor sequence for the rank conditions r if and only if there exist decreasing tableaux A_i, B_i for $1 \le i \le n-1$, such that $T_i = B_{i-1} \cdot U_{i-1,i} \cdot A_i$ for each i (with $B_0 = A_n = \emptyset$) and $(A_1 \cdot B_1, \dots, A_{n-1} \cdot B_{n-1})$ is a K-theoretic factor sequence for $\overline{r}$.*

Given a sequence $(T_1, \dots, T_n)$ of decreasing tableaux, such that each tableau T_i can be contained in the rectangle $e_i \times e_{i-1}$ and all entries of T_i are smaller than $e_{i-1} + e_i$, we let $\Phi(T_1, \dots, T_n)$ denote the decreasing tableau constructed from this sequence as well as the tableaux U_{ij} for $i \geq j$ as follows.

$\Phi(T_1, \dots, T_n) \quad =$



Notice that the upper-left box of $U_{n-1,1}$ is equal to $N-1$, and the boxes in the union of tableaux U_{ij} decrease by one for each step down or to the right. Theorem 4 follows from the following proposition combined with Theorems 1 and 3.

Proposition 22. *The map $(T_1, \dots, T_n) \mapsto \Phi(T_1, \dots, T_n)$ gives a bijection of the set of all K-theoretic factor sequences for r with the set of all decreasing tableaux representing $z(r)$.*

Proof. Since the permutation of a decreasing tableau can be defined as the south-west to north-east Hecke product of the simple reflections given by the boxes of the tableau, it follows from the definition of KMS-factorizations that $(T_1, \dots, T_n)$ is a factor sequence if and only if $\Phi(T_1, \dots, T_n)$ represents the Zelevinsky permutation $z(r)$. It remains to show that any decreasing tableau T representing $z(r)$ contains the arrangement of rectangular tableaux U_{ij} in its upper-left corner, and has no boxes strictly south-east of the tableaux U_{ii} for $1 \leq i \leq n-1$. The inclusion of the tableaux U_{ij} in T follows from Lemma 20 because $z(r) \in S_N$ and for each $0 < i \leq n$ and $p > r_{ni}$ we have $z(r)(p) \leq r_{i0}$, see [16, Prop 1.6] or [4, Lemma 3.1].

To see that T contains no boxes strictly south-east of U_{ii}, we use that the Grothendieck polynomial $\mathfrak{G}_{\widehat{z}(r)}(x_1, \dots, x_N)$ is separately symmetric in each group of variables $\{x_p \mid r_{n,i} < p \leq r_{n,i-1}\}$, where $\widehat{z}(r) = w_0^{(N)} z(r)^{-1} w_0^{(N)}$ and $w_0^{(N)}$ is the longest permutation in S_N. This is true because the descent positions of $\widehat{z}(r)$ are contained in the set $\{r_{nj} \mid 0 < j \leq n\}$. It follows that the exponent of $x_{r_{ni}+1}$ in any monomial of $\mathfrak{G}_{\widehat{z}(r)}(x_1, \dots, x_N)$ is less than or equal to $N - r_{n,i-1} = r_{i-2,0}$. Now T can be used to construct a unique compatible pair (a, k) for $\widehat{z}(r)$, such that T contains the integer p in some box of row q if and only if $(a_l, k_l) = (N-p, q)$ for some l. Since this pair contributes the monomial x^k to $\mathfrak{G}_{\widehat{z}(r)}(x_1, \dots, x_N)$, it follows that row $r_{ni}+1$ of T has at most $r_{i-2,0}$ boxes. This means exactly that T contains no boxes south-east of $U_{i-1,i-1}$, as required. □

References

1. S. Billey, W. Jockusch, and R. P. Stanley, *Some combinatorial properties of Schubert polynomials*, J. Algebraic Combin. **2** (1993), no. 4, 345–374.
2. A. S. Buch, *Grothendieck classes of quiver varieties*, Duke Math. J. **115** (2002), no. 1, 75–103. MR 2003m:14018
3. ———, *A Littlewood-Richardson rule for the K-theory of Grassmannians*, Acta Math. **189** (2002), no. 1, 37–78. MR 2003j:14062
4. ———, *Alternating signs of quiver coefficients*, to appear in J. Amer. Math. Soc., 2003.
5. A. S. Buch and W. Fulton, *Chern class formulas for quiver varieties*, Invent. Math. **135** (1999), no. 3, 665–687. MR 2000f:14087
6. A. S. Buch, A. Kresch, H. Tamvakis, and A. Yong, *Grothendieck polynomials and quiver formulas*, to appear in Amer. J. Math., 2003.
7. ———, *Schubert polynomials and quiver formulas*, Duke Math. J. **122** (2004), 125–143.

8. A. S. Buch, F. Sottile, and A. Yong, *Quiver coefficients are Schubert structure constants*, preprint, 2004.
9. M. Demazure, *Une nouvelle formule des caractères*, Bull. Sci. Math. (2) **98** (1974), no. 3, 163–172.
10. P. Edelman and C. Greene, *Balanced tableaux*, Adv. in Math. **63** (1987), no. 1, 42–99. MR 88b:05012
11. S. Fomin and C. Greene, *Noncommutative Schur functions and their applications*, Discrete Math. **193** (1998), no. 1-3, 179–200, Selected papers in honor of Adriano Garsia (Taormina, 1994). MR 2000c:05149
12. S. Fomin and A. N. Kirillov, *Grothendieck polynomials and the Yang-Baxter equation*, Proc. Formal Power Series and Alg. Comb. (1994), 183–190.
13. ———, *The Yang-Baxter equation, symmetric functions, and Schubert polynomials*, Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993), vol. 153, 1996, pp. 123–143. MR 98b:05101
14. W. Fulton, *Universal Schubert polynomials*, Duke Math. J. **96** (1999), no. 3, 575–594. MR 99m:14093
15. M. D. Haiman, *Dual equivalence with applications, including a conjecture of Proctor*, Discrete Math. **99** (1992), no. 1-3, 79–113.
16. A. Knutson, E. Miller, and M. Shimozono, *Four positive formulae for type A quiver polynomials*, preprint, 2003.
17. A. Lascoux, *Transition on Grothendieck polynomials*, Physics and combinatorics, 2000 (Nagoya), World Sci. Publishing, River Edge, NJ, 2001, pp. 164–179. MR 2002k:14082
18. A. Lascoux and M.-P. Schützenberger, *Polynômes de Schubert*, C. R. Acad. Sci. Paris Sér. I Math. **294** (1982), no. 13, 447–450. MR 83e:14039
19. ———, *Structure de Hopf de l'anneau de cohomologie et de l'anneau de Grothendieck d'une variété de drapeaux*, C. R. Acad. Sci. Paris Sér. I Math. **295** (1982), no. 11, 629–633. MR 84b:14030
20. ———, *Noncommutative Schubert polynomials*, Funktsional. Anal. i Prilozhen. **23** (1989), no. 3, 63–64.
21. E. Miller, *Alternating formulae for K-theoretic quiver polynomials*, to appear in Duke Math. J., 2003.
22. V. Reiner and M. Shimozono, *Key polynomials and a flagged Littlewood-Richardson rule*, J. Combin. Theory Ser. A **70** (1995), no. 1, 107–143.
23. ———, *Plactification*, J. Algebraic Combin. **4** (1995), no. 4, 331–351.
24. A. V. Zelevinskiĭ, *Two remarks on graded nilpotent classes*, Uspekhi Mat. Nauk **40** (1985), no. 1(241), 199–200. MR 86e:14027

Matematisk Institut, Aarhus Universitet, Ny Munkegade, 8000 Århus C, Denmark
E-mail address: abuch@imf.au.dk

Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail address: kresch@maths.warwick.ac.uk

Department of Mathematics, Virginia Tech, 460 McBryde Hall, Blacksburg, VA 24061-0123
E-mail address: mshimo@vt.edu

Department of Mathematics, Brandeis University - MS 050, P. O. Box 9110, Waltham, MA 02454-9110, USA
E-mail address: harryt@brandeis.edu

Department of Mathematics, University of California at Berkeley, 970 Evans Hall, Berkeley, CA 94720-3840, USA
E-mail address: ayong@math.berkeley.edu

The octahedron recurrence and combinatorics of arrays

V. I. Danilov and G.A. Koshevoy *
CEMI RAS, Nahimovskii pr. 47, 117418 Moscow, and
Poncelet Lab. (UMI 2615 of CNRS and Independent University of Moscow)
e-mail: vdanilov43@mail.ru, koshevoy@cemi.rssi.ru

1 Introduction

In [1] it was proposed a functional approach to combinatorics of Young tableaux. This approach proved itself extremely useful for combinatorics of Littlewood-Richardson coefficients and solutions to the Horn problem, see [2, 3, 8, 11]. Due to this line of approach, LR-coefficients count integer-valued discrete concave functions on a triangle grid in $\mathbb{Z}^2$ with prescribed boundary values, and the Hermitian matrices A, B and $A+B$ with spectra α, β and γ, respectively, exist iff there exists a discrete concave function on a triangular grid with the increments α, β and γ along the sides of the triangle ([3, 11]).

An interesting bijection between two sets constituted from some pairs of discrete concave functions was constructed in [12]. This bijection is based upon the octahedron recurrence, or the tropicalization of the discrete Hirota equation. Specifically the octahedron recurrence gradually propagates a pair of discrete concave functions at two adjoint faces (grids) of a tetrahedron to a pair of functions on other two faces. In [6] we have shown that the resulting functions are indeed discrete concave, and this is because of a relation between the octahedron recurrence and discrete concavity in dimension 3.

Here we show that the modified RSK for arrays[1] can be obtained by the octahedral recurrence. Specifically, to an array we associate a supermodular function on $\mathbb{Z}^2_+$. We locate this function at a rectangular modular face of a prism, and we set the null function at the non-modular face, and again the null function at the "bottom" (with respect to the propagation vector) triangle face. These initial data we propagate using the octahedron recurrence to

*A partial support from the grant NSch-1939.2003.6 is acknowledged. G.A. Koshevoy also thanks for support the Foundation of Support of Russian Science.

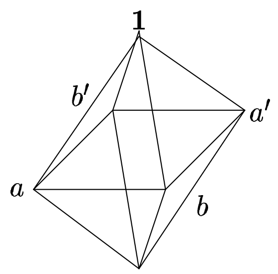
[1]For integer-valued arrays, the modified (crystal) RSK differs from the the RSK by changing Q-symbol by the Schutzenberger involution of it. This slight modification is forced by tensor product on crystals (see [4, 5]).

the whole prism, and get a polarized function on the prism (discrete concave by modulo adding two one-dimensional concave functions). Restrictions of this function to two other modular faces are so-called vertically strip concave (VS-concave) and horizontally strip concave functions (HS-concave), respectively. This construction is invertible (the octahedron recurrence with the opposite propagation vector). Thus we get a bijection, the *functional RSK*, between the set of arrays and pairs of VS-concave and HS-concave functions of equal "shape". The set of integer-valued VS-concave (HS-concave) functions is in bijection with the set of semi-standard Young tableaux [4]. Thus, for the integer-valued setup, this construction provides a bijection between arrays and pairs of semi-standard Young tableaux of the same shape. This, of course, reminds the RSK correspondence, but indeed this bijection is the modified RSK. Note, that the Schutzenberger involution also might be obtained by the octahedron recurrence. Let us remark an advantage of the functional form of RSK: the direct and inverse bijections are done symmetrically.

Finally, the complexity of the functional RSK (direct and reverse) is $O(\max(n,m)^3)$, and the usual RSK algorithm has complexity $O((\sum_{ij}(a(i,j))^{3/2})$.

2 Octahedron recurrence

The main idea of the octahedron recurrence is rather transparent. Specifically, consider the octahedron



Picture 1.

with the vertexes $\mathbf{0}, a, a', b, b'$ and $\mathbf{1}$. Let f be a real-valued function given at the points $\mathbf{0}, a, a', b, b'$. Then we can *propagate* f to the point $\mathbf{1}$ by the following rule

$$f(\mathbf{1}) = \max(f(a) + f(a'), f(b) + f(b')) - f(\mathbf{0}).$$

We refer to [13] for justifications of this rule and interesting examples of appearances of this rule in combinatorics. Rather unexpectedly this related to flips in [7]. We want to point out a relation of this rule to concavity.

Specifically, suppose $f(b) + f(b') = \max(f(a) + f(a'), f(b) + f(b'))$. Then, we have $f(\mathbf{0}) + f(\mathbf{1}) = f(b) + f(b')$. This means that the restriction of the function to the rhombus $\mathbf{0}, b, \mathbf{1}, b'$ coincides with the restriction of an affine function h. Moreover, $h(a) + h(a') = 2h((a + a')/2) = 2h((b + b')/2) = 2f((b + b')/2) = f(b) + f(b') \geq f(a) + f(a')$. And we can choose h such that $f(a) \leq h(a)$ and $f(a') \leq h(a')$ hold true. This means that the function f is sub-affine on the octahedron, i.e. f looks alike a *concave* function. Moreover, the rhombus $\mathbf{0}, b, \mathbf{1}, b'$ is an affinity set of f.

In other words, we propagate the function f to the point $\mathbf{1}$ in order to get a concave (discretely) function on the octahedron, such that an affinity area (the convex hull of the affinity set) contains the vector $\mathbf{01}$, the *propagation vector*.

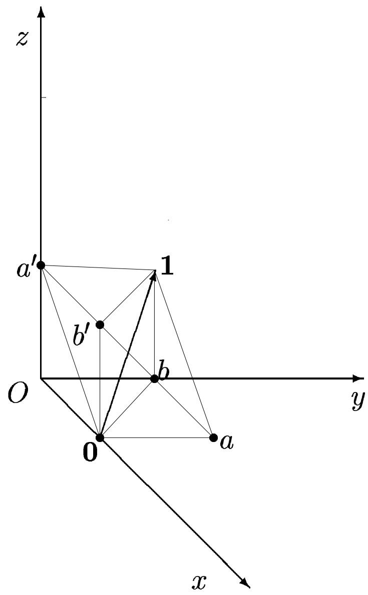
Now, using this rule, which is called the *octahedron recurrence (OR)*, we can propagate a function given at some domain to a large domain. Here is one of possible initial domains (see [13]). Let us consider the set L of points (n, i, j) with integers i, j, n, $n \geq 0$ and $n = i + j \pmod 2$. Suppose a function f is given at a subset of L constituted from points of the form $(\cdot, \cdot, 0)$ and $(\cdot, \cdot, 1)$. Then using the octahedron recurrence with the propagation vector $(0, 0, 2)$ we can propagate the function to points of L of the form $(\cdot, \cdot, 2)$, and so on to the whole L.

Of course, the initial data can be given at more sophisticated subsets, see [13] and [9].

We will display the octahedron recurrence in a slightly different manner[2]. Namely, for us it will be convenient to consider the octahedron recurrence with the propagation vector $(-1, 1, 1)$ and locate the initial data at the qudrants Oxy (*ground*) and Oxz (*front wall*). The modular flats take the form $x = a$, $y = b$, $z = c$ and $x + y + z = d$, where $a, b, c, d \in \mathbb{Z}$. If we cut $\mathbb{R}^3$ by these planes, we get a decomposition of $\mathbb{R}^3$ into primitive tetrahedrons and octahedrons. All octahedrons are parallel, that is one can be obtained by an integer translation of another.

In this set-up, the unit octahedron is of the form depicted at Picture 2.

[2]Sometimes it is convenient to draw pictures for the OR with different propagation vectors. In order to set the octahedron recurrence, we have to choose a unimodular set in the lattice $\mathbb{Z}^3$, say, $\{e_1, e_2, e_3, e_1 - e_2, e_1 - e_3, e_2 - e_3\}$ and the propagation vector $e_3 - e_1 + e_2$, where e_1, e_2, e_3 is a basis in the lattice $\mathbb{Z}^3$. Then the primitive octahedron becomes the convex hull of the points 0, $e_3 - e_1$, e_2, e_3, $e_2 - e_1$, $e_3 - e_1 + e_2$. An integer translation of a plane, spanned by a triple of vectors in the unimodular set, is a *modular flat*. A *non-modular flats* are parallel to planes spanned either by the pair $(e_3 - e_1, e_2)$ or $(e_2 - e_1, e_3)$.



Picture 2.

Thus, a primitive propagation takes the following form: given values at the points $\mathbf{0}$, a and b at the ground flour and two values at the points a' and b' at the first flour, due to the OR we get a value at the third point $\mathbf{1}$ at the first flour.

We claim that functions, which we get as an output of the octahedron recurrence, inherit some concavity properties of input functions. The next two sections are devoted to this issue.

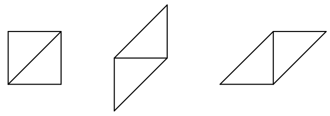
3 Discrete concave functions on 2D-grids

We consider functions on $\mathbb{Z}^2$ defined on finite sets of special form. We call such sets grids and they are specified as follows. A finite subset $T \subset \mathbb{Z}^2$ is a *grid* if i) T has no holes, i.e. $T = co(T) \cap \mathbb{Z}^2$, and ii) any edge of the convex hull $co(T)$ is parallel to one of the vectors $(1,0)$, $(0,1)$ $(1,1)$. (Obviously, a grid has a hexagonal shape, which might degenerated to a pentagon, a trapezoid, a parallelogram or a triangle.)

Let $f : T \to \mathbb{R}$ be a function on a grid T. A primitive triangle in T is either a triple x, $x + (0,1)$ and $x + (1,1)$ of points of T, or a triple x, $x + (1,0)$, $x + (1,1)$. Convex hulls of these primitive triangles constitute a simplicial decomposition of $\mathrm{co}(T)$ (if T is not one-dimensional). We uniquely interpolate the function f by affinity to the triangles on this decomposition of $\mathrm{co}T$, and get a function $\tilde{f} : co(T) \to \mathbb{R}$.

Definition. A function f on a grid T is said to be *discrete concave*, if the interpolation $\tilde{f}$ is a concave function on $\mathrm{co}(T)$.

We can reformulate discrete concavity of a function f without using the interpolation $\tilde{f}$. Namely we have to require validity of three types of "rhombus" inequalities. Consider "primitive" rhombus in T of the form



Then discrete concavity is equivalent to validity of three types of "rhombus" inequalities. The inequalities require that sum at two points of drawn diagonal is greater or equal to the sum at two points of non-drawn diagonal.

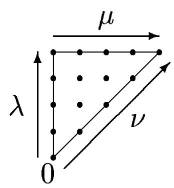
(i) $f(i,j) + f(i+1,j+1) \geq f(i+1,j) + f(i,j+1)$;

(ii) $f(i,j+1) + f(i+1,j+1) \geq f(i+1,j+2) + f(i,j)$;

(iii) $f(i+1,j) + f(i+1,j+1) \geq f(i,j) + f(i+2,j+1)$.

Note, that if only the requirement (i) is valid, then a function is called *supermodular.* If a function is supermodular and the requirement (ii) is valid, then the function is discrete concave on every vertical strip of the unit length, and we call such functions *vertically-strip concave* (VS-concave). Analogously, if (i) and (iii) are valid, a function is called *horizontally strip-concave* (HS-concave).

Mostly, we will be interested in functions on the triangle grid with the vertexes $(0,0)$, $(0,n)$, (n,n); denoted by Δ_n. On the next picture we depicted the grid Δ_4.



Consider a discrete concave function f on the grid Δ_n[3] and consider its restriction to each side of the triangle: the left-hand side, the top of the triangle and the hypotenuse. Specifically, we orient these sides as depicted on the previous picture and consider increments of the function on each unit segment. Then, increments along the left-hand side constitute an n-tuple

$$\lambda(1) = f(0,1) - f(0,0), \lambda(2) = f(0,2) - f(0,1), ..., \lambda(n) = f(0,n) - f(0,n-1).$$

It is easy follows from the rhombus inequalities of the type (i) and (iii) that

$$\lambda(1) \geq \lambda(2) \geq ... \geq \lambda(n).$$

Analogously, we define n-tuple μ ($\mu(i) = f(i,n) - f(i-1,n)$, $i = 1,...,n$) and ν ($\nu(k) = f(k,k) - f(k-1,k-1)$, $k = 1,...,n$), which are also decreasing tuples. We call these n-tuples *increments* of the function f on the corresponding sides of the triangle grid. Obviously, the increments are invariant

[3]Such a discrete concave function was called a *hive* in [11, 12].

under adding a constant to f. Therefore, we have to consider functions modulo adding a constant or to require $f(0,0) = 0$.

Let us briefly say about main roles of discrete concave functions in combinatorics and representation theory. We let to denote $DC_n(\lambda, \mu, \nu)$ the set of discrete concave functions on the grid Δ_n with increments λ, μ, ν. This set is a polytope (probably empty) in the space of all functions on Δ_n. If this polytope is non-empty, when the n-tuples λ, μ, ν are decreasing and there holds $|\lambda| + |\mu| = |\nu|$. For $n > 2$, we need more relations in order to get a non-empty $DC_n(\lambda, \mu, \nu)$. The necessary and sufficient conditions for non-emptyness of $DC_n(\lambda, \mu, \nu)$ (so-called Horn inequalities) are in [11], see also [8, 10, 3]. Moreover, $DC_n(\lambda, \mu, \nu)$ is non-empty if and only if there exist Hermitian matrices A and B, such that A, B, $A + B$ have spectra λ, μ, ν, respectively (a solution to the Horn problem).

We let to denote $DC_n^{\mathbb{Z}}(\lambda, \mu, \nu)$ the set of integer-valued discrete concave functions on the grid Δ_n, of course the tuples λ, μ, ν have to be integer-valued as well. The cardinality of this set coincides with the Littlewood-Richardson coefficient, the multiplicity of the irreducible representation V_ν (of $GL(n)$) in the tensor product irreps $V_\lambda \otimes V_\mu$.

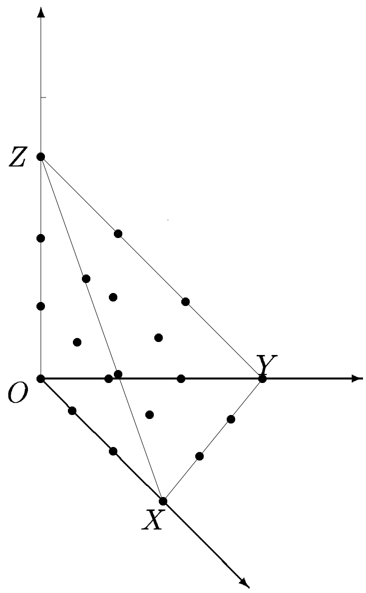
4 Functions on 3D-grids

Recall that we consider the octahedron recurrence with the propagation vector $(-1, 1, 1)$ and locate the initial data at the qudrants OXZ (*ground*) and OXY (*front wall*). The modular flats take the form $x = a$, $y = b$, $z = c$ and $x+y+z = d$, where $a, b, c, d \in \mathbb{Z}$. All primitive octahedrons are parallel, and each octahedron has three diagonals parallel to vectors $(1, 1, -1)$, $(1, -1, 1)$ and $(-1, 1, 1)$, respectively, and corresponding three pairs of antipodal vertexes. Any two of this diagonal vectors span a non-modular flat.

The diagonal being parallel to the propagation vector $(-1, 1, 1)$, we call the *mail* diagonal.

Definition. A function $F : \mathbb{Z}^3 \to \mathbb{R} \cup \{-\infty\}$ is said to be *polarized*, if, for any primitive octahedron, sum of values of F at the vertexes of the main diagonal is equal to the maximum of the sum of values of F at the antipodal vertexes of two others diagonals.

We denote $\Delta_n(OXYZ)$ the three-dimensional grid, constituted from the non-negative integer points (x, y, z), such that $x + y + z \le n$ (see Picture 3 with $\Delta_3(XYZ)$). It is easy to see that, for any initial data given at the *ground* $\Delta_n(OXY)$ and the *front wall* $\Delta_n(OXZ)$, there exists a unique polarized function with domain $\Delta_n(OXYZ)$ and these initial data. This is due to the OR. However, we can set initial data at the *shadow wall* $\Delta_n(OYZ)$ and the *slope wall* $\Delta_n(XYZ)$ and get a polarized function. In that case, we have to apply the OR with the reverse propagation vector $(1, -1, -1)$.



Picture 3.

The fundamental property of the octahedron recurrence is that if the initial data (at the ground and the front wall) are discrete concave function, then the corresponding polarized function on the grid $\Delta_n(OXYZ)$ is a kind of three-dimensional discrete concave function. Without going in details of discrete concave functions in $\mathbb{Z}^n$, we give relevant notions for purposes this paper.

Discrete concavity on **2D**-grids is equivalent to fulfilling of three kinds of rhombus inequalities. In dimension 3, we have four kinds of modular flats. In each such a 2-dimensional flat we have rhombuses, which corresponds to triangular decomposition of the flat by cutting it by three others kinds of modular flats. We require validity of the rhombus inequality for each such a rhombus in each flat: the sum of values at the "short" diagonal is greater or equal to the sum at the "long" diagonal.

Definition. A function $F : \mathbb{Z}^3 \to \mathbb{R} \cup \{-\infty\}$ is a *polarized discrete concave* function if F is polarized and all kinds of rhombus inequalities in each modular flat are fulfilled.

Let us denote by PDC_n the set of polarized discrete concave functions on the three-dimensional grid $\Delta_n(OXYZ)$.

Theorem 1. *Let F be a polarized function on the three-dimensional grid $\Delta_n(OXYZ)$. Suppose the restriction of F to the ground face $\Delta_n(OXY)$ and to the front wall face $\Delta_n(OXZ)$ are 2-dimensional discrete concave functions. Then $F \in PDC_n$.*

For a proof see [6]. Note, that this theorem is equivalent to the following corollary (a sketch of proof of which is in [9]).

Corollary 1. *If the restrictions of a polarized function to the ground and the front wall faces are discrete concave, then the restriction to the shadow wall and the slope wall are also discrete concave.*

Proof. In fact, any rhombus located on the slope or shadow wall is also a rhombus for three-dimensional grid, and, therefore, the corresponding rhombus inequality is valid. □

Corollary 2 *Let the restriction of a polarized function to the ground be discrete concave and the restriction to the front wall be HS-concave. Then the restrictions to two other faces are HS-concave.*

Proof. In fact, we can add to F an appropriate function $\varphi(z)$ of the vertical variable z, in order to get a discrete concave function on the front wall. $F+\varphi(z)$ is not changed on the ground, therefore by Corollary 1, $F+\varphi(z)$ is discrete concave at the other two wall, therefore F is HS-concave on these walls. □

Corollary 3. *Suppose the restriction of a polarized function to the ground is HS-concave and the restriction to the front wall is VS-concave (here we consider horizontal being parallel to the segment XY). Then F is VS-concave on the shadow wall.*

Proof. As above, having add to F an appropriate separable function on variables x and y, we get a polarized function $G = F + \varphi(x) + \psi(y)$, which will be discrete concave on the ground and the front wall. By Corollary 1, G is discrete concave on the shadow wall. Therefore, F is VS-concave on this wall. □

Corollary 4. *Suppose a polarized function F is discrete concave on the ground and VS-concave on the front wall. Then F is discrete concave on the shadow wall.*

Proof. In fact, having add an appropriate function on x to F, we get a discrete concave function on the front wall. On the ground this function will be also discrete concave. But this function remains the same on the shadow wall, and by Corollary 1 the function on this wall is discrete concave. □

Now let us consider the polarized functions (or the octahedron recurrence) on the prism $\Delta_n(OXY)\times\{0,1,...,m\}$.

At points outside the prism, we set functions equal $-\infty$. Therefore, on the non-modular face $\Delta_n(XY)\times\{0,1,...,m\}$, a polarized function F turns out to be a separable function (on variables $x+y$ and z).

Corollary 5 *Let F be a polarized function on the prism $\Delta_n(OXY)\times\{0,1,...,m\}$. Suppose the restriction of F to the ground face $\Delta_n(OXY)\times\{0\}$ and the restriction to the front wall $\Delta_n(OX)\times\{0,1,...,m\}$ are discrete concave functions. Then F is polarized discrete concave function on the prism* *(and, in particular, F is discrete concave on the shadow wall $\Delta_n(OY) \times \{0,1,...,m\}$ and on the ceiling $\Delta_n(OXY) \times \{m\}$).*

Proof. It is easy to see that it suffices to prove the corollary in the case $m = 2$.

In the beginning we consider the case $m = 1$. Let us extend the ground to the size of $n+1$, that is we add to $\Delta_n(OXY)$ new points $(n+1, 0, 0), ..., (0, n+1, 0)$. Let us extend F to these points such that we get a discrete concave function on the extended ground $\Delta_{n+1}(OXY) \times \{0\}$ and a discrete concave function on the "extended" front wall. We can always to do that by setting small values ($<< 0$) to these new points. Let us denote $\tilde{F}$ this extension. By Theorem 1, the function $\tilde{F}$ is a polarized discrete concave function. We claim, that the restriction of this function to the prism is a polarized discrete concave function. In fact, it suffices to check that $\tilde{F}$ coincides with F on the non-modular face $\Delta_n(XY) \times \{0,1\}$. But this holds since we assigned small values to the new points. Thus F and $\tilde{F}$ coincide on the prism. Since $\tilde{F}$ is discrete concave function, F is discrete concave too.

Let us move to the case $m = 2$. We have to check all rhombus inequalities for all rhombuses in the prism $\Delta_n(OXY) \times \{0, 2\}$. Let us first consider the rhombuses of the vertical size 2. It is easy to see that these rhombuses belong to the tetrahedron of size $n + 1$. Then the corresponding rhombus inequality is valid, since they are valid for $\tilde{F}$. Other rhombuses are located either in the prism $\Delta_n(OXY) \times \{0,1\}$, or in the prism $\Delta_n(OXY) \times \{1,2\}$. For the first prism, the corresponding inequality follows due to the above case with $m = 1$. Moreover, we get that F is discrete concave on the triangle $\Delta_n(OXY) \times \{1\}$. Now, again applying the case $m = 1$ to the prism $\Delta_n(OXY) \times \{1,2\}$, we get validity of rhombus inequalities in this prism. □

Using similar reasonings one can get the following

Corollary 6 *Suppose a polarized function F on the prism $\Delta_n(OXY) \times \{0,1,...,m\}$ has discrete concave restrictions to the ceiling $\Delta_n(OXY) \times \{m\}$ and the shadow wall $\Delta_n(OY) \times \{0,1,...,m\}$. Then F is a polarized discrete concave function and its restrictions to the front wall $\Delta_n(OX) \times \{0,1,...,m\}$ and the ground $\Delta_n(OXY) \times \{0\}$ are discrete concave functions.*

5 Functional form of RSK

Consider the rectangle $[0, n] \times [0, m]$ on the plane with natural n and m, constituted from unit squares with the centers at the points $(i-1/2, j-1/2)$, $i = 1, ..., n$, $j = 1, ..., m$, we call such squares boxes. An array is a filling of each box (i, j) with a non-negative "mass" $a(i, j)$.

To each array a we associate a function $f = f_a$ on the rectangular grid

$\{0,1,...,n\} \times \{0,1,...,m\}$ by setting to the point (i,j) the value

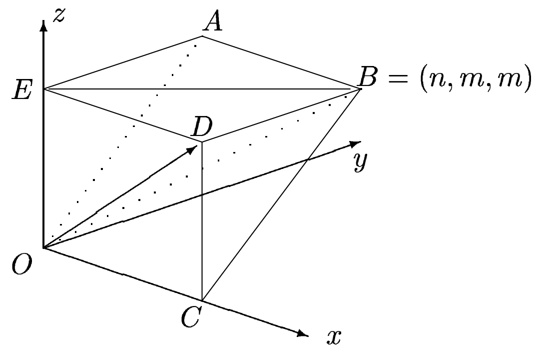
$$f_a(i,j) = \sum_{i' \le i, j' \le j} a(i',j').$$

In other words, this value is equal to the mass of all boxes to the south-west to the point (i,j). This is a reason to denote by $\iint a$ the function f_a. On the bottom and the left boundary of the rectangle the function equals 0. For other (i,j), we obviously have

$$f(i,j) - f(i-1,j) - f(i,j-1) + f(i-1,j-1) = a(i,j).$$

From this $a(i,j)$ might be understand as the mixed derivative of f $(a = \partial\partial f)$, or as a break of f along the common edge $[(i-1,j-1),(i,j)]$ of two affinity areas. Since $a(i,j) \ge 0$, the function $\iint a$ is supermodular.

Here it will be convenient to consider the octahedron recurrence with the propagation vector $(1,0,1)$. On Picture 4 with $n = m$ we depicted modular and non-modular flats: the modular flats are parallel to the faces of the tetrahedron $OEAB$, and the non-modular flats parallel to the face $OEDC$ and the plane passing through ODB (the propagation vector is parallel to the vector OD).



Picture 3.

Let us locate the null function at the face OEA; at the slope rectangle $OABC$ we locate the function $f_a = \iint a$. Specifically, we assign the value $\iint a(i,j))$ to the point (i,j,j). Now, we propagate these data by the octahedron recurrence to the prism. From Corollary 5 (Section 4), we get a VS-concave function at the top face rectangle $EABD$ and HS-concave function at the right face triangle CDB (the vertical is the y-axe in the first case, and the z-axe in the second case).

Thus, we have

Theorem 2. *Let a be an array, and let F be a function on the prism obtained by the octahedron recurrence from the following initial data: the zero values at the faces $OEDC$ and OEA, and $\iint a$ at $OABC$. Then the* *restriction of F to $EABD$ is a VS-concave function and the restriction to CDB is an HS-concave function.*

Let us note, that the latter two functions coincide at the edge DB.

Remark. We can also propagate data in the reverse direction. Specifically, assume we are given a function f on the top face $EABD$ and a function g on the triangle CDB. Suppose there hold

a) f is VS-concave and equals 0 at the edges EA and ED;

b) g is HS-concave and equals 0 at the edge CD;

c) the functions f and g coincide at the edge DB.

Then having applied the OR (with the propagation vector $(-1, 0, -1)$) to these data, we get a pair of functions on the triangle OEA and the slope rectangle $OABC$. Due to Corollary 6, we get a discrete concave function on OEA and a supermodular function on $OABC$. Moreover, we get the null function on the triangle OEA. This is because this function equals 0 at the edge EA (due to the item a)) and at the edge OE (this follows from b) and separability of the OR on the non-modular face). But these boundary values force nullity of a discrete concave function.

Now, we get an array a as the mixed derivatives of the resulting supermodular function on $OABC$. Thus, this octahedron recurrence provides us with a functional form of the RSK (indeed modified RSK [4]).

Namely, due to a bijection between VS-concave functions (HS-concave) functions and semi-standard Young tableaux ([4]), the above Theorem and Remark establish a bijection between the set of arrays and and the set of pairs of SSYT of equal shape (the increments of the functions along the edge DB is exactly the shape of the tableaux).

Let us briefly explain the mapping from VS-concave functions to SSYT. Namely, let f be a VS-concave function. Let us consider the array $\partial\partial f$. To get the corresponding semi-standard Young tableaux we have to read this array from left to right and from bottom to top. Reading a row gives us filling of the corresponding row in the Young tableau, the mass $a(i, j)$ exhibits the multiplicity of repetitions of the letters i in the j-th row of the Young tableaux (we consider the French style of drawing Young diagrams and tableaux). Consider an example. Let f be given by $\begin{matrix} 5 & 10 & 12 & 23 \\ 5 & 10 & 12 & 20 \\ 5 & 6 & 8 & 12 \end{matrix}$.

Then $\partial\partial f = \begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 4 & 0 & 4 \\ 5 & 1 & 2 & 4 \end{pmatrix}$, and the corresponding semi-standard Young tableau is

$$\begin{matrix} 4 & 4 & 4 \\ 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 & 1 & 2 & 3 & 3 & 4 & 4 & 4 & 4 \end{matrix}$$

References

[1] *Berenstein A. and A.Zelevinsky*, Involutions on the Gel'fand-Tsetlin patterns and muliplicities in the skew GL_n-modules, *Doklady AN SSSR* (in Russian), 300(1988), 1291–1294, English translation in Soviet Math. Dokl. 37 (1988), 799–802

[2] *Buch A.S.* The saturation conjecture (after A.Knutson and T.Tao). With an appendix by W. Fulton// Enseig. Math. 2000 (2), 46, 43–60

[3] *Danilov V.I. and G.A. Koshevoy*, Discrete convexity and Hermitian matrices. Proc. V.A.Steklov Inst. Math.; Vol 241 (2003), 68–90 (English transl. in Proc. V.A.Steklov Inst. Math.; Vol 241, 2003)

[4] *Danilov V.I. and G.A. Koshevoy*, Arrays and Combinatorics of Young Tableaux, *Uspehi Mathematicheskih Nauk*, 60:2(2005) (in Russian), (English transl. in *Russian Mathematical Survey*)

[5] *Danilov V., Koshevoy G.* Bicrystals and crystal $GL(V), GL(W)$-duality. Preprint RIMS-1485 (2004) Kyoto Univ., Kyoto

[6] *Danilov V., Koshevoy G.* A simple proof of the associativity and commutativity of LR-coefficients (or the hive ring). arXiv: math.CO/0409447.

[7] *Fock V.V., Goncharov A.B.* Moduli spaces of local systems and higher Teichmüller theoty. arXiv:math.AG/0311149

[8] *Fulton W.* Eigenvalues, Invariant Factors, Highest Weights, and Schubert Calculus//Bull. Amer. Math. Soc. 2000, vol. 37, pp. 209–249.

[9] *Henriques A., Kamnitzer J.* The octahedron recurrence and gl_n cristals, arXiv:math.CO/0408114.

[10] *Karzanov A.V.* Concave cocirculations in a triangular grid, to appear in *Linear Algebra and Appl.*, arXiv:math.CO/0304289.

[11] *Knutson A., Tao T.* The honeycomb model of $GL_n(\mathbb{C})$ tensor products I: proof of the saturation conjecture//J. of Amer. Math. Soc. 1999, 12, no. 4, 1055–1090.

[12] *Knutson A., Tao T., Woodward C.* A positive proof of the Littlewood-Richardson rule using the octahedron recurrence//Electron.J.Combin. 2004. v. 11, Research Paper 61; arXive:math.CO/0306274.

[13] *Speyer D.E.* Perfect matchings and the octahedron recurrence. arXiv:math.CO/0402452.

NEW EXPLICIT EXPRESSION FOR $A_n^{(1)}$ SUPERNOMIALS

LIPIKA DEKA AND ANNE SCHILLING

Abstract. A new fermionic formula for type $A_{n-1}^{(1)}$ supernomials is presented. This formula is different from the one given by Hatayama et al. [6]. A new set of unrestricted rigged configurations is introduced which is in bijection with the unrestricted crystal paths.

Résumé. On présente une nouvelle formule fermionique pour les supernomiales de type $A_{n-1}^{(1)}$. Cette formule est différente de celle donnée par Hatayama et al. [6]. On présente un nouvel ensemble de configurations 'gréées' sans restriction qui est en bijection avec l'ensemble des chemins cristallins sans restriction.

1. INTRODUCTION

Supernomial coefficients were first introduced in [15] as q-analogs of the coefficient of x^a in the expansion of $\prod_{j=1}^{N}(1+x+x^2+\cdots+x^j)^{L_j}$. They are generalizations of q-multinomial coefficients and were used to prove Bose-Fermi or Rogers–Ramanujan-type identities. The supernomials of [15] can be naturally associated with the algebra $A_1^{(1)}$. Motivated by crystal base theory, supernomials can be defined for any affine Kac–Moody Lie algebra as generating functions of unrestricted paths with energy statistics [16, 6, 7, 5]. An explicit formula for the $A_n^{(1)}$ supernomials for completely symmetric and completely antisymmetric crystals was proved in [6]. This formula is called fermionic as it is a manifestly positive expression. The purpose of this note is to give a new explicit expression for supernomials of type $A_n^{(1)}$.

Our motivation to seek an explicit expression for supernomials is their appearance in generalizations of the Bailey lemma [2]. Bailey's lemma is a very powerful method to prove Rogers–Ramanujan-type identities. In [16] a type A_n generalization of Bailey's lemma was conjectured which was subsequently proven in [18]. A type A_2 Bailey chain, which yields an infinite family of identities, was given in [1].

Recently, fermionic expressions for generating functions of unrestricted paths for type $A_1^{(1)}$ have also surfaced in connection with box-ball systems. Takagi [17] establishes a bijection between box-ball systems and a new set of rigged configurations to prove a fermionic formula for the q-binomial coefficient. This bijection extends a bijection of Kirillov and Reshetikhin [9] between semi-standard Young tableaux and rigged configurations to unrestricted paths.

In this note we define a new set of unrestricted rigged configurations for type $A_n^{(1)}$. A bijection between this new set and the set of unrestricted crystal paths is given which preserves the statistics. In particular this yields a new fermionic expression for the supernomial coefficients of type $A_n^{(1)}$. Subsequently, a crystal structure on the new set of rigged configurations has been defined [12] which can be used to establish the bijection and the correct properties of the statistics. These results generalize to other affine simply-laced root systems. In this note we give an algorithmic definition of the bijection by extending the definition in [10]. Details will be available in [4].

This paper is structured as follows. In section 2 we review crystals of type $A_n^{(1)}$, the definition of unrestricted paths and the definition of supernomials as generating functions of unrestricted paths with energy statistics. In section 3 we give our new definition of unrestricted rigged configurations and derive from this a fermionic expression for the generating function of unrestricted rigged configurations graded by cocharge. Our main results are stated in section 4. The fermionic formula

Date: November 2004; revised April 2005.

Partially supported by NSF grant DMS-0200774.

of section 3 yield an explicit expression for the supernomials. This result is based on a bijection between unrestricted paths and unrestricted rigged configurations.

2. UNRESTRICTED PATHS AND SUPERNOMIALS

$A^{(1)}_{n-1}$-supernomials were first introduced in [16] as generating functions of unrestricted paths graded by an energy function. An unrestricted path is an element in the tensor product of crystals $B = B^{r_k,s_k} \otimes B^{r_{k-1},s_{k-1}} \otimes \cdots \otimes B^{r_1,s_1}$. As a set the crystal $B^{r,s}$ of type $A^{(1)}_{n-1}$ is the set of all column-strict Young tableaux of shape (s^r) over the alphabet $\{1, 2, \ldots, n\}$. Kashiwara [8] introduced the notion of crystals and crystal graphs as a combinatorial means to study representations of quantum algebras. In particular, there are Kashiwara operators e_i, f_i defined on the elements in $B^{r,s}$ for $0 \le i \le n$. However, for the purpose of this note, we do not require the explicit action of e_i and f_i.

Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ be an n-tuple of nonnegative integers. The set of **unrestricted paths** is defined as

$$\mathcal{P}(B, \lambda) = \{b \in B \mid \mathrm{wt}(b) = \lambda\}.$$

Here $\mathrm{wt}(b) = (w_1, \ldots, w_n)$ is the weight of b where w_i counts the number of letters i in b.

Example 2.1. For $B = B^{1,1} \otimes B^{2,2} \otimes B^{3,1}$ of type A_3 and $\lambda = (2, 3, 1, 2)$ the path

$$b \ = \ \begin{array}{|c|}\hline 2\\\hline\end{array} \otimes \begin{array}{|c|c|}\hline 1&2\\\hline 2&4\\\hline\end{array} \otimes \begin{array}{|c|}\hline 1\\\hline 3\\\hline 4\\\hline\end{array}$$

is in $\mathcal{P}(B, \lambda)$.

There exists a crystal isomorphism $R : B^{r,s} \otimes B^{r',s'} \to B^{r',s'} \otimes B^{r,s}$, called the **combinatorial** R**-matrix**. Combinatorially it is given as follows. Let $b \in B^{r,s}$ and $b' \in B^{r',s'}$. The product $b \cdot b'$ of two tableaux is defined as the Schensted insertion of b' into b. Then $R(b \otimes b') = \tilde{b}' \otimes \tilde{b}$ is the unique pair of tableaux such that $b \cdot b' = \tilde{b}' \cdot \tilde{b}$.

The **local energy function** $H : B^{r,s} \otimes B^{r',s'} \to \mathbb{Z}$ is defined as follows. For $b \otimes b' \in B^{r,s} \otimes B^{r',s'}$, $H(b \otimes b')$ is the number of boxes of the shape of $b \cdot b'$ outside the shape obtained by concatenating (s^r) and $({s'}^{r'})$.

Example 2.2. For

$$b \otimes b' = \begin{array}{|c|c|}\hline 1&2\\\hline 2&4\\\hline\end{array} \otimes \begin{array}{|c|}\hline 1\\\hline 3\\\hline 4\\\hline\end{array}$$

we have

$$b \cdot b' = \begin{array}{|c|c|c|}\hline 1&1&3\\\hline 2&2&4\\\hline 4\\\cline{1-1}\end{array} = \begin{array}{|c|}\hline 1\\\hline 2\\\hline 4\\\hline\end{array} \cdot \begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline\end{array} = \tilde{b}' \cdot \tilde{b}.$$

so that

$$R(b \otimes b') = \tilde{b}' \otimes \tilde{b} = \begin{array}{|c|}\hline 1\\\hline 2\\\hline 4\\\hline\end{array} \otimes \begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline\end{array}.$$

Since the concatentation of $\begin{array}{|c|c|}\hline \ & \ \\\hline \ & \ \\\hline\end{array}$ and $\begin{array}{|c|}\hline \ \\\hline \ \\\hline \ \\\hline\end{array}$ is $\begin{array}{|c|c|c|}\hline \ & \ & \ \\\hline \ & \ & \ \\\hline \ \\\cline{1-1}\end{array}$, the local energy function $H(b \otimes b') = 0$.

Now let $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$ be a k-fold tensor product of crystals. The **tail energy function** $\overleftarrow{D} : B \to \mathbb{Z}$ is given by

$$\overleftarrow{D} = \sum_{1 \le i < j \le k} H_{j-1} R_{j-2} \cdots R_{i+1} R_i,$$

where H_i (resp. R_i) is the local energy function (resp. combinatorial R-matrix) acting on the i-th and $(i+1)$-th tensor factors.

Definition 2.3. The q-supernomial coefficient is the generating function of unrestricted paths graded by the tail energy function

$$S_{B,\lambda}(q) = \sum_{b \in \mathcal{P}(B,\lambda)} q^{\overleftarrow{D}(b)}.$$

3. Unrestricted rigged configurations and fermionic formula

Rigged configurations are combinatorial objects invented to label the solutions of the Bethe equations, which give the eigenvalues of the Hamiltonian of the underlying physical model [3]. Motivated by the fact that representation theoretically the eigenvectors and eigenvalues can also be labelled by Young tableaux, Kirillov and Reshetikhin [9] gave a bijection between tableaux and rigged configurations. This result and generalizations thereof were proven in [10].

In terms of crystal base theory, the bijection is between highest weight paths and rigged configurations. The new result of this note is an extension of this bijection to a bijection between unrestricted paths and a new set of rigged configurations, which we define in this section. In [12] we also define a crystal structure on this new set of rigged configurations.

Let $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$ and denote by $L = (L_i^{(a)} \mid (a,i) \in \mathcal{H})$ the multiplicity array of B, where $L_i^{(a)}$ is the multiplicity of $B^{a,i}$ in B. Here $\mathcal{H} = \overline{I} \times \mathbb{Z}_{>0}$ and $\overline{I} = \{1, 2, \ldots, n-1\}$ is the index set of the Dynkin diagram A_{n-1}. The sequence of partitions $\nu = \{\nu^{(a)} \mid a \in \overline{I}\}$ is a (L,λ)-**configuration** if

$$\sum_{(a,i)\in\mathcal{H}} i m_i^{(a)} \alpha_a = \sum_{(a,i)\in\mathcal{H}} i L_i^{(a)} \Lambda_a - \lambda, \tag{3.1}$$

where $m_i^{(a)}$ is the number of parts of length i in partition $\nu^{(a)}$. Note that we do not require λ to be a dominant weight here. The **(quasi-)vacancy number** of a configuration is defined as

$$p_i^{(a)} = \sum_{j\geq 1} \min(i,j) L_j^{(a)} - \sum_{(b,j)\in\mathcal{H}} (\alpha_a|\alpha_b) \min(i,j) m_j^{(b)}.$$

Here $(\cdot|\cdot)$ is the normalized invariant form on the weight lattice P such that $(\alpha_i|\alpha_j)$ is the Cartan matrix. Let $\mathrm{C}(L,\lambda)$ be the set of all (L,λ)-configurations. We call $p_i^{(a)}$ quasi-vacancy number to indicate that they can actually be negative in our setting. For the rest of the paper we will simply call them vacancy numbers.

In the usual setting a rigged configuration (ν, J) consists of a configuration $\nu \in \mathrm{C}(L,\lambda)$ together with a double sequence of partitions $J = \{J^{(a,i)} \mid (a,i) \in \mathcal{H}\}$ such that the partition $J^{(a,i)}$ is contained in a $m_i^{(a)} \times p_i^{(a)}$ rectangle. In particular this requires that $p_i^{(a)} \geq 0$. For unrestricted paths we need a bigger set, where the lower bound on the parts in $J^{(a,i)}$ 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 $c_k = \lambda_{k+1} + \lambda_{k+2} + \cdots + \lambda_n$. We also set $c_0 = c_1$. Let $\mathcal{A}(\lambda')$ be the set of tableaux of shape λ' such that the entries in column k are from the set $\{1, 2, \ldots, c_{k-1}\}$ and are strictly decreasing along each column.

Example 3.1. For $n = 4$ and $\lambda = (0,1,1,1)$, the set $\mathcal{A}(\lambda')$ consists of the following tableaux

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

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 $t_{j,a}$ denotes the entry in row j and column a of t, and $\chi(S) = 1$ 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.

Definition 3.2. An **unrestricted rigged configuration** (ν, J) is a configuration $\nu \in \mathrm{C}(L,\lambda)$ together with a sequence $J = \{J^{(a,i)} \mid (a,i) \in \mathcal{H}\}$ where $J^{(a,i)}$ is a $(M_i^{(a)}(t), p_i^{(a)}, m_i^{(a)})$-quasipartition for some $t \in \mathcal{A}(\lambda')$. Denote the set of all unrestricted rigged configurations corresponding to (L,λ) by $\mathrm{RC}(L,\lambda)$.

Remark 3.3.

(1) Note that this definition is similar to the definition of level-restricted rigged configurations [13, Definition 5.5]. Whereas for level-restricted rigged configurations the vacancy number had to be modified according to tableaux in a certain set, here the lower bounds are modified.

(2) For type A_1 we have $\lambda = (\lambda_1, \lambda_2)$ so that $\mathcal{A} = \{t\}$ contains just the single tableau

$$t = \begin{array}{|c|}\hline \lambda_2 \\ \hline \lambda_2 - 1 \\ \hline \vdots \\ \hline 1 \\ \hline \end{array}.$$

In this case $M_i(t) = -\sum_{j=1}^{\lambda_2} \chi(i \geq t_{j,1}) = -i$. This agrees with the findings of [17].

The quasipartition $J^{(a,i)}$ is called **singular** if it has a part of size $p_i^{(a)}$. It is often useful to view an (unrestricted) rigged configuration (ν, J) as a sequence of partitions ν where the parts of size i in $\nu^{(a)}$ are labeled by the parts of $J^{(a,i)}$. The pair (i,x) where i is a part of $\nu^{(a)}$ and x is a part of $J^{(a,i)}$ is called a **string** of the a-th rigged partition $(\nu,J)^{(a)}$. The label x is called a **rigging**.

Example 3.4. Let $n = 4$, $\lambda = (2,2,1,1)$, $L_1^{(1)} = 6$ and all other $L_i^{(a)} = 0$. Then

$$(\nu, J) \;=\; \begin{array}{|c|c|c|l}\cline{1-3} & & & -2 \\ \cline{1-3} & \multicolumn{3}{l}{0} \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|l}\cline{1-2} & & 0 \\ \cline{1-2} \end{array} \quad \begin{array}{|c|l}\cline{1-1} & -1 \\ \cline{1-1} \end{array}$$

is an unrestricted rigged configuration in $\mathrm{RC}(L,\lambda)$, where we have written the parts of $J^{(a,i)}$ next to the parts of length i in partition $\nu^{(a)}$. To see that the riggings form quasipartitions, let us write the vacancy numbers $p_i^{(a)}$ next to the parts of length i in partition $\nu^{(a)}$:

$$\begin{array}{|c|c|c|l}\cline{1-3} & & & 0 \\ \cline{1-3} & \multicolumn{3}{l}{3} \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|l}\cline{1-2} & & 0 \\ \cline{1-2} \end{array} \quad \begin{array}{|c|l}\cline{1-1} & -1 \\ \cline{1-1} \end{array}.$$

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

$$\begin{array}{|c|c|c|l}\cline{1-3} & & & -2 \\ \cline{1-3} & \multicolumn{3}{l}{-1} \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|l}\cline{1-2} & & 0 \\ \cline{1-2} \end{array} \quad \begin{array}{|c|l}\cline{1-1} & -1 \\ \cline{1-1} \end{array},$$

which are less or equal to the riggings in (ν, J).

The following statistics can be defined on the set of unrestricted rigged configurations. For $(\nu, J) \in \mathrm{RC}(L,\lambda)$ let

$$cc(\nu, J) = cc(\nu) + \sum_{(a,i) \in \mathcal{H}} |J^{(a,i)}|,$$

where $|J^{(a,i)}|$ is the sum of all parts of the quasipartition $J^{(a,i)}$ and

$$cc(\nu) = \frac{1}{2} \sum_{a,b \in \overline{I}} \sum_{j,k \geq 1} (\alpha_a \mid \alpha_b) \min(j,k) m_j^{(a)} m_k^{(b)}.$$

Definition 3.5. The RC polynomial is defined as

$$\mathrm{RC}_{L,\lambda}(q) = \sum_{(\nu,J)\in \mathrm{RC}(L,\lambda)} q^{cc(\nu,J)}.$$

The RC polynomial is in fact S_n-symmetric in the weight λ. This is not obvious from its definition as both (3.1) and the lower bounds are not symmetric with respect to λ.

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

$$\bigcup_{S\in\mathcal{SA}(\lambda')} (-1)^{|S|+1}\{J \mid J^{(a,i)} \text{ is a } (M_i^{(a)}(S), p_i^{(a)}, m_i^{(a)})\text{-quasipartition}\}.$$

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 $\mathrm{RC}_{L,\lambda}(q)$ may be rewritten as

$$\mathrm{RC}_{L,\lambda}(q) = \sum_{S\in\mathcal{SA}(\lambda')} (-1)^{|S|+1} \sum_{\nu\in \mathrm{C}(L,\lambda)} q^{cc(\nu)+\sum_{(a,i)\in\mathcal{H}} m_i^{(a)} M_i^{(a)}(S)} \times \prod_{(a,i)\in\mathcal{H}} \begin{bmatrix} m_i^{(a)} + p_i^{(a)} - M_i^{(a)}(S) \\ m_i^{(a)} \end{bmatrix}$$

called **fermionic formula**.

4. Main results

In this section we relate the fermionic formula for the RC polynomial of section 3 and the q-supernomial coefficients of section 2.

Theorem 4.1. *If L is the multiplicity array for the crystal B, then $S_{B,\lambda}(q) = \mathrm{RC}_{L,\lambda}(q)$.*

This theorem follows immediately from the following result.

Theorem 4.2. *There exists a bijection $\Phi : \mathcal{P}(B,\lambda) \to \mathrm{RC}(L,\lambda)$ which preserves the statistics, that is, $\overleftarrow{D}(b) = cc(\Phi(b))$ for all $b \in \mathcal{P}(B,\lambda)$.*

A proof of Theorem 4.2 is given in [4, 12].

In [12] a crystal structure is defined on the set of unrestricted rigged configurations which is the same as the crystal structure on paths. The highest weight elements are given by the usual rigged configurations and highest weight paths, respectively, for which Theorem 4.2 is known to hold by [10]. Since the statistics is constant on all classical crystal components, the proof of Theorem 4.2 follows in general. It should be noted that the results in [12] hold for all for all simply-laced types, not just type $A_{n-1}^{(1)}$. Hence Theorem 4.2 holds whenever there is a corresponding bijection for the highest weight elements (for example for type $D_n^{(1)}$ for symmetric powers [14] and antisymmetric powers [11]). It is expected that using virtual crystals and the method of folding Dynkin diagrams, these results can be extended to other affine root systems.

In this note, which is a summary of [4], we take a different approach and define the map Φ algorithmically which generalizes the bijection of [10]. To define Φ we first need to define certain maps on paths and rigged configurations. These maps correspond to the following operations on crystals:

(1) If $B = B^{1,1} \otimes B'$, let $\mathrm{lh}(B) = B'$. This operation is called **left-hat**.

(2) If $B = B^{r,s} \otimes B'$ with $s \geq 2$, let $\mathrm{ls}(B) = B^{r,1} \otimes B^{r,s-1} \otimes B'$. This operation is called **left-split**.
(3) If $B = B^{r,1} \otimes B'$ with $r \geq 2$, let $\mathrm{lb}(B) = B^{1,1} \otimes B^{r-1,1} \otimes B'$. This operation is called **box-split**.

In analogy we define $\mathrm{lh}(L)$ (resp. $\mathrm{ls}(L)$, $\mathrm{lb}(L)$) to be the multiplicity array of $\mathrm{lh}(B)$ (resp. $\mathrm{ls}(B)$, $\mathrm{lb}(B)$), if L is the multiplicity array of B. The corresponding maps on crystal elements are given by:

(1) Let $b = c \otimes b' \in B^{1,1} \otimes B'$. Then $\mathrm{lh}(b) = b'$.
(2) Let $b = c \otimes b' \in B^{r,s} \otimes B'$, where $c = c_1 c_2 \cdots c_s$ and c_i denotes the i-th column of c. Then $\mathrm{ls}(b) = c_1 \otimes c_2 \cdots c_s \otimes b'$.
(3) Let $b = \begin{array}{|c|}\hline b_1 \\ \hline b_2 \\ \hline \vdots \\ \hline b_r \\ \hline \end{array} \otimes b' \in B^{r,1} \otimes B'$, where $b_1 < \cdots < b_r$. Then $\mathrm{lb}(b) = \begin{array}{|c|}\hline b_r \\ \hline \end{array} \otimes \begin{array}{|c|}\hline b_1 \\ \hline \vdots \\ \hline b_{r-1} \\ \hline \end{array} \otimes b'$.

In the next subsection we define the corresponding maps on rigged configurations, and give the bijection in subsection 4.2.

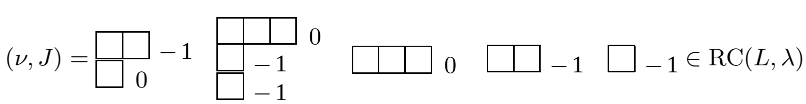
4.1. **Operations on rigged configurations.** Suppose $L_1^{(1)} > 0$. The main algorithm on rigged configurations as defined in [9, 10] for admissible rigged configurations can be extended to our setting. For a tuple of nonnegative integers $\lambda = (\lambda_1, \ldots, \lambda_n)$, let λ^- be the set of all nonnegative tuples $\mu = (\mu_1, \ldots, \mu_n)$ such that $\lambda - \mu = e_r$ for some $1 \leq r \leq n$ where e_r is the canonical r-th unit vector in $\mathbb{Z}^n$. Define $\delta : \mathrm{RC}(L, \lambda) \to \bigcup_{\mu \in \lambda^-} \mathrm{RC}(\mathrm{lh}(L), \mu)$ by the following algorithm. Let $(\nu, J) \in \mathrm{RC}(L, \lambda)$. Set $\ell^{(0)} = 1$ and repeat the following process for $a = 1, 2, \ldots, n-1$ or until stopped. Find the smallest index $i \geq \ell^{(a-1)}$ such that $J^{(a,i)}$ is singular. If no such i exists, set $\mathrm{rk}(\nu, J) = a$ and stop. Otherwise set $\ell^{(a)} = i$ and continue with $a+1$. Set all undefined $\ell^{(a)}$ to ∞.

The new rigged configuration $(\tilde{\nu}, \tilde{J}) = \delta(\nu, J)$ is obtained by removing a box from the selected strings and making the new strings singular again. Explicitly

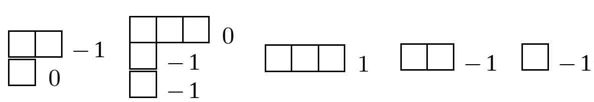
$$m_i^{(a)}(\tilde{\nu}) = m_i^{(a)}(\nu) + \begin{cases} 1 & \text{if } i = \ell^{(a)} - 1 \\ -1 & \text{if } i = \ell^{(a)} \\ 0 & \text{otherwise.} \end{cases}$$

The partition $\tilde{J}^{(a,i)}$ is obtained from $J^{(a,i)}$ by removing a part of size $p_i^{(a)}(\nu)$ for $i = \ell^{(a)}$, adding a part of size $p_i^{(a)}(\tilde{\nu})$ for $i = \ell^{(a)} - 1$, and leaving it unchanged otherwise. Then $\delta(\nu, J) \in \mathrm{RC}(\mathrm{lh}(L), \mu)$ where $\mu = \lambda - e_{\mathrm{rk}(\nu, J)}$.

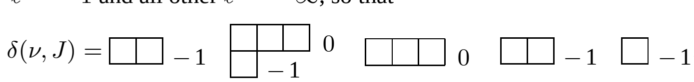
Example 4.3. Let L be the multiplicity array of $B = B^{1,1} \otimes B^{2,1} \otimes B^{2,3}$ and $\lambda = (2, 2, 2, 1, 1, 1)$. Then



Writing the vacancy numbers next to each part instead of the riggings we get



Hence $\ell^{(1)} = \ell^{(2)} = 1$ and all other $\ell^{(a)} = \infty$, so that



Also $cc(\nu, J) = 2$.

Let $s \geq 2$. Suppose $B = B^{r,s} \otimes B'$ and L the corresponding multiplicity array. Note that $\mathrm{C}(L,\lambda) \subset \mathrm{C}(\mathrm{ls}(L),\lambda)$. Under this inclusion map, the vacancy number $p_i^{(a)}$ for ν increases by $\delta_{a,r}\chi(i<s)$. Hence there is a well-defined injective map $i : \mathrm{RC}(L,\lambda) \to \mathrm{RC}(\mathrm{ls}(L),\lambda)$ given by $i(\nu,J) = (\nu,J)$.

Suppose $r \geq 2$ and $B = B^{r,1} \otimes B'$ with multiplicity array L. Then there is an injection $j : \mathrm{RC}(L,\lambda) \to \mathrm{RC}(\mathrm{lb}(L),\lambda)$ defined by adding singular strings of length 1 to $(\nu,J)^{(a)}$ for $1 \leq a < r$. Moreover the vacancy numbers stay the same.

4.2. **Bijection.** The map $\Phi : \mathcal{P}(B,\lambda) \to \mathrm{RC}(L,\lambda)$ is defined by various commutative diagrams. Note that it is possible to go from $B = B^{r_k,s_k} \otimes B^{r_{k-1},s_{k-1}} \otimes \cdots \otimes B^{r_1,s_1}$ to the empty crystal via successive application of lh, ls and lb.

Definition 4.4. Define that map $\Phi : \mathcal{P}(B,\lambda) \to \mathrm{RC}(L,\lambda)$ such that the empty path maps to the empty rigged configuration, and:

(1) Suppose $B = B^{1,1} \otimes B'$. Then the diagram

$$\begin{array}{ccc} \mathcal{P}(B,\lambda) & \xrightarrow{\Phi} & \mathrm{RC}(L,\lambda) \\ {\scriptstyle \mathrm{lh}}\big\downarrow & & \big\downarrow{\scriptstyle \delta} \\ \bigcup\limits_{\mu \in \lambda^-} \mathcal{P}(\mathrm{lh}(B),\mu) & \xrightarrow[\Phi]{} & \bigcup\limits_{\mu \in \lambda^-} \mathrm{RC}(\mathrm{lh}(L),\mu) \end{array}$$

commutes.

(2) Suppose $B = B^{r,s} \otimes B'$ with $s \geq 2$. Then the following diagram commutes:

$$\begin{array}{ccc} \mathcal{P}(B,\lambda) & \xrightarrow{\Phi} & \mathrm{RC}(L,\lambda) \\ {\scriptstyle \mathrm{ls}}\big\downarrow & & \big\downarrow{\scriptstyle i} \\ \mathcal{P}(\mathrm{ls}(B),\lambda) & \xrightarrow[\Phi]{} & \mathrm{RC}(\mathrm{ls}(L),\lambda) \end{array}$$

(3) Suppose $B = B^{r,1} \otimes B'$ with $r \geq 2$. Then the following diagram commutes:

$$\begin{array}{ccc} \mathcal{P}(B,\lambda) & \xrightarrow{\Phi} & \mathrm{RC}(L,\lambda) \\ {\scriptstyle \mathrm{lb}}\big\downarrow & & \big\downarrow{\scriptstyle j} \\ \mathcal{P}(\mathrm{lb}(B),\lambda) & \xrightarrow[\Phi]{} & \mathrm{RC}(\mathrm{lb}(L),\lambda) \end{array}$$

It is shown in [4] that the map Φ of Definition 4.4 is indeed a well-defined bijection.

Example 4.5. Let $B = B^{1,1} \otimes B^{2,1} \otimes B^{2,3}$ and $\lambda = (2,2,2,1,1,1)$. Then

$$b = \boxed{3} \otimes \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array} \otimes \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ \hline \end{array} \in \mathcal{P}(B,\lambda)$$

and $\Phi(b)$ is the rigged configuration (ν,J) of Example 4.3. We have $\overleftarrow{D}(b) = cc(\nu,J) = 2$.

Example 4.6. Let $n = 4$, $B = B^{2,2} \otimes B^{2,1}$ and $\lambda = (2,2,1,1)$. Then the multiplicity array is $L_1^{(2)} = 1, L_2^{(2)} = 1$ and $L_i^{(a)} = 0$ for all other (a,i). There are 7 possible unrestricted paths in $\mathcal{P}(B,\lambda)$. For each path $b \in \mathcal{P}(B,\lambda)$ the corresponding rigged configuration $(\nu,J) = \Phi(b)$ together with the tail energy and cocharge is summarized below.

$$b = \begin{array}{|c|c|}\hline 1&1\\\hline 2&2\\\hline\end{array}\otimes\begin{array}{|c|}\hline 3\\\hline 4\\\hline\end{array} \qquad (\nu,J) = \square\ 0 \quad \begin{array}{l}\square\ -1\\ \square\ -1\end{array} \quad \square\ 0 \qquad \overleftarrow{D}(b)=0=cc(\nu,J)$$

$$b = \begin{array}{|c|c|}\hline 1&1\\\hline 2&4\\\hline\end{array}\otimes\begin{array}{|c|}\hline 2\\\hline 3\\\hline\end{array} \qquad (\nu,J) = \square\ -1 \quad \begin{array}{l}\square\ 0\\ \square\ 0\end{array} \quad \square\ 0 \qquad \overleftarrow{D}(b)=1=cc(\nu,J)$$

$$b = \begin{array}{|c|c|}\hline 1&2\\\hline 2&3\\\hline\end{array}\otimes\begin{array}{|c|}\hline 1\\\hline 4\\\hline\end{array} \qquad (\nu,J) = \square\ 0 \quad \begin{array}{l}\square\ 0\\ \square\ 0\end{array} \quad \square\ -1 \qquad \overleftarrow{D}(b)=1=cc(\nu,J)$$

$$b = \begin{array}{|c|c|}\hline 1&2\\\hline 2&4\\\hline\end{array}\otimes\begin{array}{|c|}\hline 1\\\hline 3\\\hline\end{array} \qquad (\nu,J) = \square\ 0 \quad \begin{array}{l}\square\ 0\\ \square\ -1\end{array} \quad \square\ 0 \qquad \overleftarrow{D}(b)=1=cc(\nu,J)$$

$$b = \begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline\end{array}\otimes\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array} \qquad (\nu,J) = \square\ 0 \quad \begin{array}{l}\square\ 0\\ \square\ 0\end{array} \quad \square\ 0 \qquad \overleftarrow{D}(b)=2=cc(\nu,J)$$

$$b = \begin{array}{|c|c|}\hline 1&1\\\hline 2&3\\\hline\end{array}\otimes\begin{array}{|c|}\hline 2\\\hline 4\\\hline\end{array} \qquad (\nu,J) = \square\ -1 \quad \square\square\ 0 \quad \square\ -1 \qquad \overleftarrow{D}(b)=0=cc(\nu,J)$$

$$b = \begin{array}{|c|c|}\hline 1&2\\\hline 3&4\\\hline\end{array}\otimes\begin{array}{|c|}\hline 1\\\hline 2\\\hline\end{array} \qquad (\nu,J) = \square\ -1 \quad \square\square\ 1 \quad \square\ -1 \qquad \overleftarrow{D}(b)=1=cc(\nu,J)$$

The supernomial in this case is $\mathrm{RC}_{L,\lambda}(q) = 2+4q+q^2 = S_{B,\lambda}(q)$.

REFERENCES

[1] G.E. Andrews, A. Schilling, S.O. Warnaar, *An A_2 Bailey lemma and Rogers-Ramanujan-type identities*, J. Amer. Math. Soc. **12** (1999) 677–702.

[2] W. N. Bailey, *Identities of the Rogers–Ramanujan type*, Proc. London Math. Soc. (2) **50** (1949) 1–10.

[3] H. A. Bethe, *Zur Theorie der Metalle, I. Eigenwerte und Eigenfunktionen der linearen Atomkette*, Z. Physik **71** (1931) 205–231.

[4] L. Deka, A. Schilling, *New fermionic formula for unrestricted Kostka polynomials*, in preparation.

[5] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, Z. Tsuboi, *Paths, crystals and fermionic formulae*, Prog. Math. Phys. **23** (2002) 205–272, Birkhäuser Boston, Boston, MA.

[6] G. Hatayama, N. Kirillov, A. Kuniba, M. Okado, T. Takagi, Y. Yamada, *Character formulae of $\hat{sl}_n$-modules and inhomogeneous paths*, Contemp. Math. **248** (1999) 243–291.

[7] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, Y. Yamada, *Remarks on fermionic formula*, Contemp. Math. **248** (1999) 243–291.

[8] M. Kashiwara, *Crystalizing the q-analogue of universal enveloping algebras*, Commun. Math. Phys. **133** (1990) 249–260.

[9] A. N. Kirillov, N. Y. Reshetikhin, *The Bethe ansatz and the combinatorics of Young tableaux*, (Russian) Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **155** (1986), Differentsialnaya Geometriya, Gruppy Li i Mekh. VIII, 65–115, 194; translation in J. Soviet Math. **41** (1988), no. 2, 925–955.

[10] A. N. Kirillov, A. Schilling, M. Shimozono, *A bijection between Littlewood-Richardson tableaux and rigged configurations*, Selecta Mathematica (N.S.) **8** (2002) 67–135.

[11] A. Schilling, *A bijection between type $D_n^{(1)}$ crystals and rigged configurations*, J. Algebra **285** (2005) 292–334.

[12] A. Schilling, *Crystal structure on rigged configurations*, in preparation.

[13] A. Schilling, M. Shimozono, *Fermionic formulas for level-restricted generalized Kostka polynomials and coset branching functions*, Commun. Math. Phys. **220** (2001) 105–164.

[14] A. Schilling, M. Shimozono, *$X = M$ for symmetric powers*, J. Algebra, to appear (math.QA/0412376).

[15] A. Schilling, S.O. Warnaar, *Supernomial coefficients, polynomial identities and q-series*, The Ramanujan Journal **2** (1998) 459–494.

[16] A. Schilling, S.O. Warnaar, *Inhomogeneous lattice paths, generalized Kostka polynomials and A_{n-1} supernomials*, Commun. Math. Phys. **202** (1999) 359–401.

[17] T. Takagi, *Inverse scattering method for a soliton cellular automaton*, Nuclear Phys. B **707** (2005) 577–601.

[18] S.O. Warnaar, *The Bailey lemma and Kostka polynomials*, J. Algebraic Combin. **20** (2004) 131–171.

Department of Mathematics, University of California, One Shields Avenue, Davis, CA 95616-8633, U.S.A.
E-mail address: deka@math.ucdavis.edu
URL: http://www.math.ucdavis.edu/~deka

E-mail address: anne@math.ucdavis.edu
URL: http://www.math.ucdavis.edu/~anne

(−1)–ENUMERATION OF SELF–COMPLEMENTARY PLANE PARTITIONS

THERESIA EISENKÖLBL

Fakultät für Mathematik, Universität Wien,
Nordbergstraße 15, A-1090 Wien, Austria.
E-mail: Theresia.Eisenkoelbl@univie.ac.at

Abstract. We prove a product formula for the remaining cases of the weighted enumeration of self–complementary plane partitions contained in a given box where adding one half of an orbit of cubes and removing the other half of the orbit changes the weight by -1. We use nonintersecting lattice path families to express this enumeration as a Pfaffian which can be expressed in terms of the known ordinary enumeration of self–complementary plane partitions.

1. Introduction

A plane partition P can be defined as a finite set of integer points (i,j,k) with $i,j,k>0$ and if $(i,j,k)\in P$ and $1\le i'\le i$, $1\le j'\le j$, $1\le k'\le k$ then $(i',j',k')\in P$. We interpret these points as midpoints of cubes and represent a plane partition by stacks of cubes (see Figure 1). If we have $i\le a$, $j\le b$ and $k\le c$ for all cubes of the plane partition, we say that the plane partition is contained in a box with sidelengths a,b,c.

Plane partitions were first introduced by MacMahon. One of his main results is the following [10, Art. 429, $x\to 1$, proof in Art. 494]:

The number of all plane partitions contained in a box with sidelengths a,b,c equals

$$B(a,b,c)=\prod_{i=1}^{a}\prod_{j=1}^{b}\prod_{k=1}^{c}\frac{i+j+k-1}{i+j+k-2}=\prod_{i=1}^{a}\frac{(c+i)_b}{(i)_b}, \tag{1}$$

where $(a)_n:=a(a+1)(a+2)\dots(a+n-1)$ is the rising factorial.

MacMahon also started the investigation of the number of plane partitions with certain symmetries in a given box. These numbers can also be expressed as product formulas similar to the one given above. In [14], Stanley introduced additional complementation symmetries giving six new combinations of symmetries which led to more conjectures all of which were settled in the 1980's and 90's (see [14, 8, 3, 17]).

Many of these theorems come with q–analogs, that is, weighted versions that record the number of cubes or orbits of cubes by a power of q and give expressions containing q–rising factorials instead of rising factorials (see [1, 2, 11]). For plane partitions with complementation symmetry, it seems to be difficult to find natural q–analogs. However,

2000 *Mathematics Subject Classification.* Primary 05A15; Secondary 05B45 52C20.

Key words and phrases. lozenge tilings, rhombus tilings, plane partitions, determinants, pfaffians, nonintersecting lattice paths.

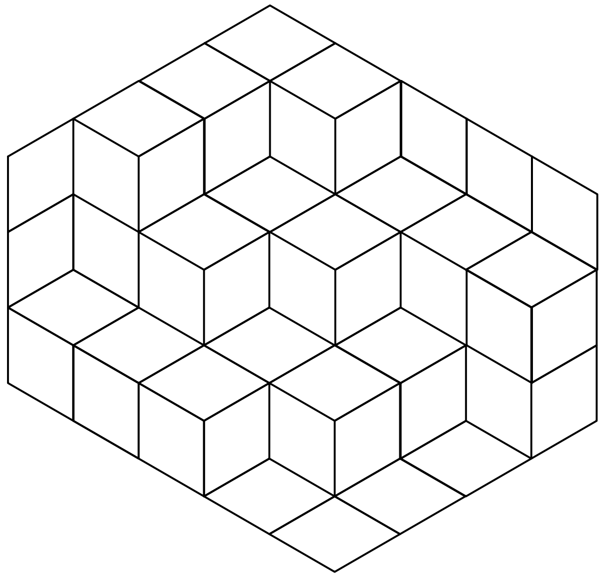


Figure 1. A self–complementary plane partition

in Stanley's paper a q–analog for self–complementary plane partitions is given (the weight is not symmetric in the three sidelengths, but the result is). Interestingly, upon setting $q = -1$ in the various q–analogs, one consistently obtains enumerations of other objects, usually with additional symmetry restraints. This observation, dubbed the "'(-1)–phenomenon" has been explained for many but not all cases by Stembridge (see [15] and [16]).

In [7], Kuperberg defines a (-1)–enumeration for all plane partitions with complementation symmetry which admits a nice closed product formula in almost all cases. These conjectures were solved in Kuperberg's own paper and in the paper [4] except for one case without a nice product formula and the case of self-complementary plane partitions in a box with some odd sidelengths which will be the main theorem of this paper. We start with the precise definitions.

A plane partition P contained in the box $a \times b \times c$ is called *self–complementary* if $(i, j, k) \in P \Leftrightarrow (a + 1 - i, b + 1 - j, c + 1 - k) \notin P$ for $1 \leq i \leq a$, $1 \leq j \leq b$, $1 \leq k \leq c$. This means that the box consists exactly of the plane partition and the image obtained from it by symmetry with respect to the central point of the box.

A convenient way to look at a self–complementary plane partition is the projection to the plane along the $(1, 1, 1)$–direction (see Figure 1). A plane partition contained in an $a \times b \times c$–box becomes a rhombus tiling of a hexagon with sidelengths a, b, c, a, b, c. It is easy to see that self-complementary plane partitions correspond exactly to those rhombus tilings with a $180°$ rotational symmetry.

The (-1)–weight is defined as follows: A self–complementary plane partition contains exactly one half of each orbit under the operation $(i, j, k) \mapsto (a+1-i, b+1-j, c+1-k)$. Let a move consist of removing one half of an orbit and adding the other half. Two plane partitions are connected either by an odd or by an even number of moves, so it is possible to define a relative sign. The sign becomes absolute if we assign weight 1 to the half-full plane partition (see Figure 2).

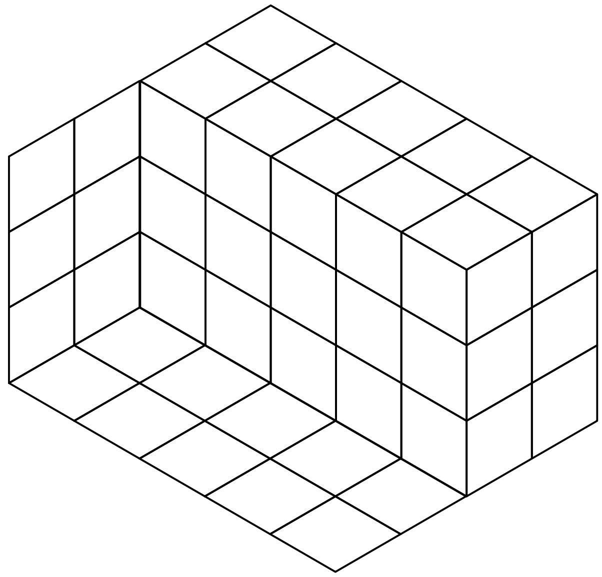


FIGURE 2. A plane partition of weight 1.

Therefore, this weight is $(-1)^{n(P)}$ where $n(P)$ is the number of cubes in the "left" half of the box and we want to evaluate $\sum_P (-1)^{n(P)}$. For example, the plane partition in Figure 1 has weight $(-1)^{10} = 1$.

In order to be able to state the result for the (-1)–enumeration more concisely, Stanley's result on the ordinary enumeration of self–complementary plane partitions is needed. It will also be used as a step in the proof of the (-1)–enumeration.

Theorem 1 (Stanley [14]). *The number $SC(a,b,c)$ of self–complementary plane partitions contained in a box with sidelengths a,b,c can be expressed in terms of $B(a,b,c)$ in the following way:*

$$\begin{array}{rl} B\left(\frac{a}{2},\frac{b}{2},\frac{c}{2}\right)^2 & \textit{for } a,b,c \textit{ even,} \\ B\left(\frac{a}{2},\frac{b+1}{2},\frac{c-1}{2}\right) B\left(\frac{a}{2},\frac{b-1}{2},\frac{c+1}{2}\right) & \textit{for } a \textit{ even and } b,\ c \textit{ odd,} \\ B\left(\frac{a+1}{2},\frac{b}{2},\frac{c}{2}\right) B\left(\frac{a-1}{2},\frac{b}{2},\frac{c}{2}\right) & \textit{for } a \textit{ odd and } b,\ c \textit{ even,} \end{array}$$

where $B(a,b,c) = \prod_{i=1}^{a} \frac{(c+i)_b}{(i)_b}$ is the number of all plane partitions in an $a \times b \times c$–box.

Note that a self-complementary plane partition contains exactly half of all cubes in the box. Therefore, there are no self-complementary plane partitions in a box with three odd sidelengths.

Now we can express the (-1)–enumeration of self–complementary plane partitions in terms of $SC(a,b,c)$, the ordinary enumeration of self–complementary plane partitions.

Theorem 2. *The enumeration of self–complementary plane partitions in a box with sidelengths a, b, c counted with weight $(-1)^{n(P)}$ equals*

$$\begin{aligned} B\left(\frac{a}{2}, \frac{b}{2}, \frac{c}{2}\right) &\quad \text{for } a, b, c \text{ even,} \\ SC\left(\tfrac{a}{2}, \tfrac{b+1}{2}, \tfrac{c-1}{2}\right) SC\left(\tfrac{a}{2}, \tfrac{b-1}{2}, \tfrac{c+1}{2}\right) &\quad \text{for } a \text{ even and } b,\ c \text{ odd} \\ SC\left(\tfrac{a+1}{2}, \tfrac{b}{2}, \tfrac{c}{2}\right) SC\left(\tfrac{a-1}{2}, \tfrac{b}{2}, \tfrac{c}{2}\right) &\quad \text{for } a \text{ odd and } b,\ c \text{ even} \end{aligned}$$

where $SC(a, b, c)$ is given in Theorem 1 in terms of the numbers of plane partitions contained in a box and $n(P)$ is the number of cubes in the plane partition P that are not in the half-full plane partition (see Figure 2).

Remark. *Since the sides of the box play symmetric roles this covers all cases. (For three odd sidelengths there are no self-complementary plane partitions.) The case of three even sidelengths has already been proved in* [4].

Note that Theorem 1 and Theorem 2 are very analogous. In the even case, the (-1)–enumeration is the square root of the ordinary enumeration. In the other cases, it is still true that there are half as many linear factors in the (-1)–enumeration (viewed as a polynomial in c, say).

In Stanley's paper [14], the theorem actually gives a q–enumeration of plane partitions. The case $q = -1$ gives the same expression as Theorem 2 above if at least one side has odd length, but this does not give a proof of Theorem 2 because the weights of individual plane partitions are different. In the case of only even sidelengths, Stanley's theorem gives $SC\,(a/2, b/2, c/2)^2$ (analogously to Theorem 1) which does not equal the result $B\,(a/2, b/2, c/2)$ in Theorem 2.

Stanley's proof uses a special case of the Littlewood-Richardson rule, the expansion of the product of Schur functions $s_\lambda(x_1, \ldots, x_n) \cdot s_\mu(x_1, \ldots, x_n)$ where λ and μ are rectangular partitions whose sidelengths differ by at most one. It is possible to give an alternative proof of Theorem 2 using a modification of Stanley's weight leading to the expansion of $s_\lambda(x_1, \ldots, x_{n+1}) \cdot s_\mu(x_1, \ldots, x_n)$ where λ and μ are rectangular partitions. (Note the different numbers of variables.) See also the remark at the end of the paper.

2. Outline of the proof

For brevity, we just give the proof in the case a even, b, c odd and $(c - b)/2$ even.

Step 1: From plane partitions to families of nonintersecting lattice paths.

We use the projection to the plane along the $(1, 1, 1)$–direction and get immediately that self–complementary plane partitions contained in an $a \times b \times c$–box are equivalent to rhombus tilings of a hexagon with sides a, b, c, a, b, c invariant under 180°–rotation. A tiling of this kind is clearly determined by one half of the hexagon.

Since the sidelengths a, b, c play a completely symmetric role and two of them must have the same parity we assume without loss of generality that $c - b$ is even and $b \le c$. The result turns out to be symmetric in b and c, so we can drop the last condition in the statement of Theorem 2. Write x for the positive integer $(c - b)/2$ and divide the hexagon in half with a line parallel to the side of length a (see Figure 3). As shown in the same figure, we find a bijection between these tiled halves and families of nonintersecting lattice paths.

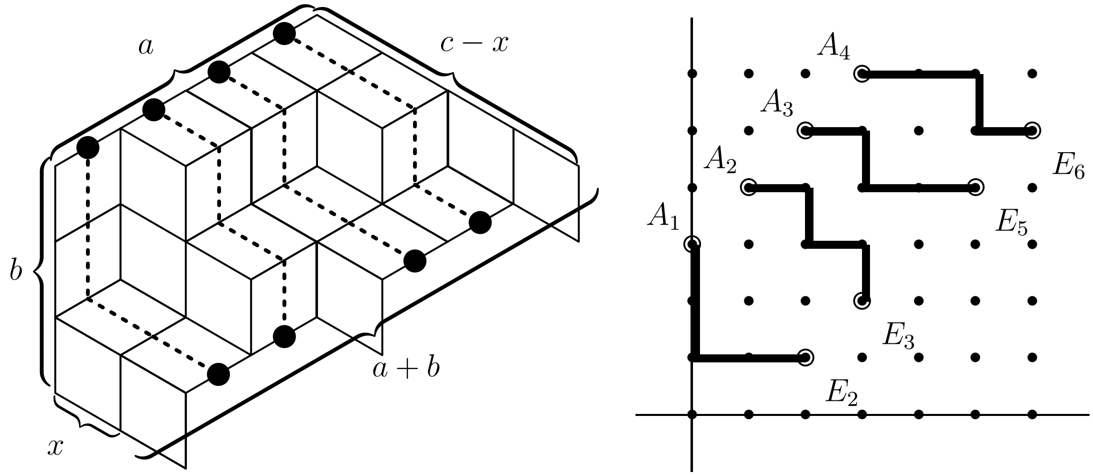


FIGURE 3. The paths for the self–complementary plane partition in Figure 1 and the orthogonal version. ($x = \frac{c-b}{2}$)

The starting points of the lattice paths are the midpoints of the edges on the side of length a. The end points are the midpoints of the edges parallel to a on the opposite boundary. This is a symmetric subset of the midpoints on the cutting line of length $a + b$.

The paths always follow the rhombi of the given tiling by connecting midpoints of parallel rhombus edges. It is easily seen that the resulting paths have no common points (i.e. they are nonintersecting) and the tiling can be recovered from a nonintersecting lattice path family with unit diagonal and down steps and appropriate starting and end points. Of course, the path families will have to be counted with the appropriate (-1)–weight.

After changing to an orthogonal coordinate system (see Figure 3), the paths are composed of unit South and East steps and the coordinates of the starting points are

$$A_i = (i-1, b+i-1) \quad \text{for } i = 1, \dots, a. \tag{2}$$

The end points are a points chosen symmetrically among

$$E_j = (x+j-1, j-1) \quad \text{for } j = 1, \dots, a+b. \tag{3}$$

Here, symmetrically means that if E_j is chosen, then $E_{a+b+1-j}$ must be chosen as well.

Note that the number $a+b$ of potential end points on the cutting line is always odd. Therefore, there is a middle one which is in no path family for even a (see Figure 3).

Now the (-1)–weight has to be defined for the paths. For a path from A_i to E_j we can use the weight $(-1)^{\text{area}(P)}$ where $\text{area}(P)$ is the area between the path and the x–axis and then multiply the weights of all the paths in the family. We have to check that the weight changes sign if we replace a half orbit with the complementary half orbit. If one of the affected cubes is completely inside the half shown in Figure 3, $\sum_P \text{area}(P)$ changes by one. If the two affected cubes are on the border of the figure, two symmetric endpoints, say E_j and $E_{a+b+1-j}$, are changed to E_{j+1} and E_{a+b-j} or vice versa. It is easily checked that in this case $\sum_P \text{area}(P)$ changes by $j + (a+b-j)$ which is odd.

It is straightforward to check that the weight for the "half-full" plane partition (see Figure 2) equals $(-1)^{a(a-2)/8}$ for a even, b, c odd. Therefore, we have to multiply the path enumeration by this global sign.

Step 2: From lattice paths to a sum of determinants

This weight can be expressed as a product of weights on individual steps (the exponent of (-1) is just the height of the step), so the following lemma is applicable. By the main theorem on nonintersecting lattice paths (see [9, Lemma 1] or [5, Theorem 1]) the weighted count of such families of paths can be expressed as a determinant.

Lemma 3. *Let $A_1, A_2, \ldots, A_n, E_1, E_2, \ldots, E_n$ be integer points meeting the following condition: Any path from A_i to E_l has a common vertex with any path from A_j to E_k for any i, j, k, l with $i < j$ and $k < l$.*

Then we have

$$\mathcal{P}(\mathbf{A} \to \mathbf{E}, \text{nonint.}) = \det_{1 \le i,j \le n} (\mathcal{P}(A_i \to E_j)), \tag{4}$$

where $\mathcal{P}(A_i \to E_j)$ denotes the weighted enumeration of all paths running from A_i to E_j and $\mathcal{P}(\mathbf{A} \to \mathbf{E}, \text{nonint.})$ denotes the weighted enumeration of all families of nonintersecting lattice paths running from A_i to E_i for $i = 1, \ldots, n$.

The condition on the starting and end points is fulfilled in our case because the points lie on diagonals, so we have to find an expression for $T_{ij} = \mathcal{P}(A_i \to E_j)$, the weighted enumeration of all single paths from A_i to E_j in our problem.

It is well-known that the enumeration of paths of this kind from (x, y) to (x', y') is given by the q–binomial coefficient $\begin{bmatrix} x'-x+y-y' \\ x'-x \end{bmatrix}_q$ if the weight of a path is q^e where e is the area between the path and a horizontal line through its endpoint.

The q–binomial coefficient (see [13, p. 26] for further information) can be defined as

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{\prod_{j=n-k+1}^{n} (1 - q^j)}{\prod_{j=1}^{k} (1 - q^j)}.$$

Although it is not obvious from this definition, the q–binomial coefficient is a polynomial in q. So it makes sense to put $q = -1$.

It is easy to verify that

$$\begin{bmatrix} n \\ k \end{bmatrix}_{-1} = \begin{cases} 0 & n \text{ even, } k \text{ odd,} \\ \binom{\lfloor n/2 \rfloor}{\lfloor k/2 \rfloor} & \text{else.} \end{cases} \tag{5}$$

Taking also into account the area between the horizontal line through the endpoint and the x–axis, we obtain

$$T_{ij} = \mathcal{P}(A_i \to E_j) = (-1)^{(x+j-i)(j-1)} \begin{bmatrix} b + x \\ b + i - j \end{bmatrix}_{-1}.$$

Now we apply Lemma 3 to all possible sets of end points. Thus, the (-1)–enumeration can be expressed as a sum of determinants which are minors of the $a \times (a + b)$–matrix T:

Lemma 4. *The (-1)–enumeration can be written as*

$$(-1)^{a(a-2)/8} \sum_{1 \le k_1 < \cdots < k_{a/2} \le (a+b-1)/2} \det(T_{k_1}, \ldots, T_{k_{a/2}}, T_{a+b+1-k_{a/2}}, \ldots, T_{a+b+1-k_1})$$

for a even and b, c odd,

where T_{ij} is $(-1)^{(x+j-i)(j-1)} \begin{bmatrix} b+x \\ b+i-j \end{bmatrix}_{-1}$ and T_j denotes the jth column of T which has length a.

Remark. *The same argument works for the ordinary enumeration, we just have to replace T_{ij} by the ordinary path enumeration $\binom{b+x}{b+i-j}$.*

Step 3: The sum of determinants is a single Pfaffian

Recall that the Pfaffian of a skew–symmetric $2n \times 2n$–matrix M is defined as

$$\operatorname{Pf} M = \sum_m \operatorname{sgn} m \prod_{\substack{\{i,j\} \in m \\ i<j}} M_{ij},$$

where the sum runs over all $m = \{\{m_1, m_2\}, \{m_3, m_4\}, \ldots, \{m_{2n-1}, m_{2n}\}\}$ with the conditions $\{m_1, \ldots, m_{2n}\} = \{1, \ldots, 2n\}$, $m_{2k-1} < m_{2k}$ and $m_1 < m_3 < \cdots < m_{2n-1}$. The term $\operatorname{sgn} m$ is the sign of the permutation $m_1 m_2 m_3 \ldots m_{2n}$.

Specifically, $(\operatorname{Pf} M)^2 = \det M$.

Our sums of determinants can be simplified by a theorem of Ishikawa and Wakayama [6, Theorem 1(1)] which we use to express the sum as a Pfaffian. We use the specialization $A = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ on the version given in [12, Corollary 3.2] and obtain the following lemma.

Lemma 5. *Let S be a $2m \times 2n$–matrix with $m \le n$ and S^* be the matrix*

$$(S_1, \ldots, S_n, S_{2n}, \ldots, S_{n+1})$$

where S_j denotes the jth column of S. Let A be the matrix $\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$. Then the following identity holds:

$$\sum_{1 \le k_1 < \cdots < k_m \le n} \det(S_{k_1}, \ldots, S_{k_m}, S_{2n+1-k_m}, \ldots, S_{2n+1-k_1}) = \operatorname{Pf}(S^* A ({}^t S^*))$$

$$= \operatorname{Pf}_{1 \le i,j \le 2m} \left(\sum_{k=1}^{n} (S_{ik} S_{j,2n+1-k} - S_{jk} S_{i,2n+1-k}) \right).$$

Now we apply this lemma to our sums.

Lemma 6. *The Pfaffian for the (-1)–enumeration for $b \le c$ is*

$$(-1)^{a(a-2)/8} \operatorname{Pf}_{1 \le i,j \le a} \left(\sum_{k=1}^{\frac{a+b-1}{2}} (T_{ik} T_{j,a+b+1-k} - T_{jk} T_{i,a+b+1-k}) \right),$$

for a even and b, c odd,

where $T_{ij} = (-1)^{(x-i)(j-1)} \begin{bmatrix} b+x \\ b+i-j \end{bmatrix}_{-1}$ *(and* $x = (c-b)/2$*).*

Proof. Apply the lemma with $2m = a$, $2n = a + b - 1$ and

$$S = (T_1, \ldots, T_{\frac{a+b-1}{2}}, T_{\frac{a+b+3}{2}}, \ldots, T_{a+b})$$

to obtain

$$\operatorname{Pf}_{1\le i,j\le a}\left(\sum_{k=1}^{\frac{a+b-1}{2}} (T_{ik}T_{j,a+b+1-k} - T_{jk}T_{i,a+b+1-k})\right).$$

□

Lemma 7. *The Pfaffian for the ordinary enumeration* $SC(a,b,c)$ *for* $b \le c$ *is*

$$\operatorname{Pf}_{1\le i,j\le a}\left(\sum_{k=1}^{\lfloor\frac{a+b}{2}\rfloor}\left(\binom{b+x}{b+i-k}\binom{b+x}{j+k-a-1} - \binom{b+x}{b+j-k}\binom{b+x}{i+k-a-1}\right)\right)$$

for a *and* $c-b$ *even.*

Proof. Replace T_{ij} by the ordinary enumeration of the respective paths. This replaces (-1)–binomial coefficients by ordinary ones. (Doing the same thing for the analogous expressions in Section 9 of [4] gives the result for the case of even sidelengths.) □

Remark. *Of course, the closed form of this Pfaffian is known by Stanley's theorem (see Theorem 1). Therefore, we can use them to evaluate the Pfaffian for the* (-1)*–enumeration.*

Step 4: Evaluation of the Pfaffian

Now, the Pfaffian of Lemma 6 can be reduced to products of the known Pfaffians in Lemma 7 corresponding to the ordinary enumeration. The calculations have to be done separately for different parities of the parameters and we present only the case a, x even, b, c odd.

For M_{ij} in Pf M we can write

$$\sum_{k=1}^{(a+b-1)/2} (-1)^{(k+1)(i+j)}\left(\binom{(b+x-1)/2}{\lfloor(b+i-k)/2\rfloor}\binom{(b+x-1)/2}{\lfloor(j+k-a-1)/2\rfloor} - \binom{(b+x-1)/2}{\lfloor(b+j-k)/2\rfloor}\binom{(b+x-1)/2}{\lfloor(i+k-a-1)/2\rfloor}\right)$$

with $1 \le i, j \le a$.

Splitting the sum into terms $k = 2l$ and $k = 2l - 1$ gives

$$\sum_{l=1}^{\lfloor (a+b-1)/4\rfloor} (-1)^{i+j} \Bigg(\binom{(b+x-1)/2}{(b-1)/2+\lfloor (i-1)/2\rfloor - l+1}\binom{(b+x-1)/2}{\lfloor (j-1)/2\rfloor + l - a/2} \\ - \binom{(b+x-1)/2}{(b-1)/2+\lfloor (j-1)/2\rfloor - l+1}\binom{(b+x-1)/2}{\lfloor (i-1)/2\rfloor + l - a/2}\Bigg) \\ + \sum_{l=1}^{\lceil (a+b-1)/4\rceil} \Bigg(\binom{(b+x-1)/2}{(b-1)/2+\lfloor i/2\rfloor - l+1}\binom{(b+x-1)/2}{\lfloor j/2\rfloor + l - a/2-1} \\ - \binom{(b+x-1)/2}{(b-1)/2+\lfloor j/2\rfloor - l+1}\binom{(b+x-1)/2}{\lfloor i/2\rfloor + l - a/2-1}\Bigg) \quad (6)$$

Now we apply some row and column operations to our matrix M. Start with $row(1)$, then write the differences $row(2i+1) - row(2i)$ for $i = 1, \ldots, a/2-1$, and finally $row(2i-1)+row(2i)$ for $i = 1, \ldots, a/2$. Now apply the same operations to the columns, so that the resulting matrix is still skew–symmetric. The new matrix has the same Pfaffian only up to sign $(-1)^{(a/2)(a/2-1)/2}$ which cancels with the global sign in Lemma 6.

Computation gives:

$$M_{2i+1,j} - M_{2i,j} = -\sum_{l=1}^{\lfloor (a+b-1)/4\rfloor} (-1)^{j} \Bigg(\binom{(b+x-1)/2+1}{(b-1)/2+i-l+1}\binom{(b+x-1)/2}{\lfloor (j-1)/2\rfloor + l - a/2} \\ - \binom{(b+x-1)/2}{(b-1)/2+\lfloor (j-1)/2\rfloor - l+1}\binom{(b+x-1)/2+1}{i + l - a/2}\Bigg)$$

Thus, apart from the first row and column, the left upper corner looks like

$$M_{2i+1,2j+1} - M_{2i,2j+1} - M_{2i+1,2j} + M_{2i,2j} \\ = \sum_{l=1}^{\lfloor (a+b-1)/4\rfloor} \Bigg(\binom{(b+x-1)/2+1}{(b-1)/2+i-l+1}\binom{(b+x-1)/2+1}{j + l - a/2} \\ - \binom{(b+x-1)/2+1}{(b-1)/2+j-l+1}\binom{(b+x-1)/2+1}{i + l - a/2}\Bigg), \quad (7)$$

where $i, j = 1, \ldots a/2-1$. Note how similar this is to the original matrix, only the (-1)–binomial coefficients are now replaced with ordinary binomial coefficients. The goal is to identify two blocks in the matrix which correspond to ordinary enumeration of self–complementary plane partitions.

The right upper corner is zero (of size $(a/2-1) \times a/2$).

Furthermore,

$$M_{2i-1,j} + M_{2i,j} = \sum_{l=1}^{\lceil (a+b-1)/4\rceil} \Bigg(\binom{(b+x-1)/2+1}{(b-1)/2+i-l+1}\binom{(b+x-1)/2}{\lfloor j/2\rfloor + l - a/2-1} \\ - \binom{(b+x-1)/2}{(b-1)/2+\lfloor j/2\rfloor - l+1}\binom{(b+x-1)/2+1}{i + l - a/2-1}\Bigg)$$

Therefore, we get for the right lower corner of the matrix

$$\begin{aligned} M_{2i-1,2j-1} &+ M_{2i,2j-1} + M_{2i-1,2j} + M_{2i,2j} \\ &= \sum_{l=1}^{\lceil (a+b-1)/4 \rceil} \left(\binom{(b+x-1)/2+1}{(b-1)/2+i-l+1} \binom{(b+x-1)/2+1}{j+l-a/2-1} \right. \\ &\qquad \left. - \binom{(b+x-1)/2+1}{(b-1)/2+j-l+1} \binom{(b+x-1)/2+1}{i+l-a/2-1} \right), \end{aligned} \tag{8}$$

where $i, j = 1, \ldots, a/2$.

This is almost a block diagonal matrix, only the first row and column spoil the picture.

Example ($a = 8, b = 3, c = 7$):

$$\left(\begin{array}{cccc|cccc} 0 & 0 & 1 & 5 & 0 & 0 & -1 & -5 \\ 0 & 0 & 3 & 12 & 0 & 0 & 0 & 0 \\ -1 & -3 & 0 & 9 & 0 & 0 & 0 & 0 \\ -5 & -12 & -9 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 & -6 & -15 \\ 0 & 0 & 0 & 0 & 1 & 0 & -9 & -18 \\ 1 & 0 & 0 & 0 & 6 & 9 & 0 & -9 \\ 5 & 0 & 0 & 0 & 15 & 18 & 9 & 0 \end{array} \right)$$

If **$(a/2)$ is even**, the right lower corner is an $(a/2) \times (a/2)$–matrix with non-zero determinant, as we will see later, thus, we can use the last $a/2$ rows to annihilate the second half of the first row. This potentially changes the entry 0 in position $(1, 1)$, but leaves everything else unchanged. We can use the same linear combination on the last $a/2$ columns to annihilate the second half of the first column. The resulting matrix is again skew–symmetric which means that the entry $(1, 1)$ has returned to the value 0. Since simultaneous row and column manipulations of this kind leave the Pfaffian unchanged, it remains to find out the Pfaffian of the right lower corner $(a/2 \times a/2)$ and the Pfaffian of the left upper corner $(a/2 \times a/2)$.

The right lower block is given by Equation (8). This corresponds exactly to the ordinary enumeration of self–complementary plane partitions in Lemma 7. Therefore, the Pfaffian of this block is $SC(a/2, (b+1)/2, (c+1)/2)$ (which is non-zero as claimed).

The left upper $a/2 \times a/2$ block (including the first row and column) is

$$\left(\begin{array}{c|c} 0 & M_{1,2j+1} - M_{1,2j} \\ \hline M_{2i+1,1} - M_{2i,1} & \displaystyle\sum_{l=1}^{\lfloor \frac{a+b-1}{4} \rfloor} \left(\binom{(b+x+1)/2}{(b-1)/2+i-l+1} \binom{(b+x+1)/2}{j+l-a/2} - \binom{(b+x+1)/2}{(b-1)/2+j-l+1} \binom{(b+x+1)/2}{i+l-a/2} \right) \end{array} \right),$$

where i, j run from 0 to $a/2 - 1$ and

$$M_{2i+1,1} - M_{2i,1} = \sum_{l=1}^{\lfloor (a+b-1)/4 \rfloor} \left(\binom{(b+x+1)/2}{(b-1)/2+i-l+1} \binom{(b+x-1)/2}{l-a/2} - \binom{(b+x-1)/2}{(b-1)/2-l+1} \binom{(b+x+1)/2}{i+l-a/2} \right)$$
$$M_{1,2j+1} - M_{1,2j} = \sum_{l=1}^{\lfloor (a+b-1)/4 \rfloor} \left(\binom{(b+x-1)/2}{(b-1)/2-l+1} \binom{(b+x+1)/2}{j+l-a/2} - \binom{(b+x+1)/2}{(b-1)/2+j-l+1} \binom{(b+x-1)/2}{l-a/2} \right).$$

Note that the exceptional row and column almost fit the general pattern. We just have sometimes $(b+x-1)/2$ instead of $(b+x+1)/2$. Replace $row(i)$ with $row(i)-row(i-1)$ for $i = 1, 2, \ldots, a/2-1$ in that order. Then do the same thing for the columns. In the resulting matrix all occurrences of $(b+x+1)/2$ have been replaced with $(b+x-1)/2$.

After shifting the indices by one, we get

$$\sum_{l=1}^{\lfloor \frac{a+b-1}{4} \rfloor} \left(\binom{(b+x-1)/2}{(b-1)/2+i-l} \binom{(b+x-1)/2}{j+l-a/2-1} - \binom{(b+x-1)/2}{(b-1)/2+j-l} \binom{(b+x-1)/2}{i+l-a/2-1} \right),$$

for $i, j = 1, \ldots, a/2$.

The Pfaffian of this matrix can easily be identified as $SC(\frac{a}{2}, \frac{b-1}{2}, \frac{c-1}{2})$ by Lemma 7. Using Theorem 1, we obtain for the (-1)–enumeration

$$SC(\tfrac{a}{2}, \tfrac{b+1}{2}, \tfrac{c+1}{2}) SC(\tfrac{a}{2}, \tfrac{b-1}{2}, \tfrac{c-1}{2}) = SC\left(\tfrac{a}{2}, \tfrac{b+1}{2}, \tfrac{c-1}{2}\right) SC\left(\tfrac{a}{2}, \tfrac{b-1}{2}, \tfrac{c+1}{2}\right),$$

which proves the main theorem in this case.

If $\boldsymbol{(a/2)}$ **is odd**, we move the first row and column to the $(a/2)$th place (which does not change the sign). Now we have an $(a-2)/2 \times (a-2)/2$–block matrix in the left upper corner which has non-zero determinant and thus can be used to annihilate the first half of the exceptional row and column similar to the previous case. By Equation (7) and Lemma 7 this is clearly $SC((a-2)/2, (b+1)/2, (c+1)/2)$.

For the right lower $(a+2)/2 \times (a+2)/2$–block, note that the relevant half of the exceptional column is

$$M_{2i-1,1} + M_{2i,1} = \sum_{l=1}^{\lceil (a+b-1)/4 \rceil} \left(\binom{(b+x+1)/2}{(b+1)/2+i-l} \binom{(b+x-1)/2}{l-a/2-1} \right.$$
$$\left. - \binom{(b+x-1)/2}{(b+1)/2-l} \binom{(b+x+1)/2}{i+l-a/2-1} \right).$$

We use again row and column operations of the type $row(i)-row(i-1)$. This changes all occurrences of $(b+x+1)/2$ to $(b+x-1)/2$ and the extra row and column now fit the pattern in Equation (8) with $i, j = 0$. After shifting i, j to $i-1, j-1$, we identify this Pfaffian as $SC((a+2)/2, (b-1)/2, (c-1)/2)$. Again, by Theorem 1, the product of the two terms is exactly $SC(a/2, (b-1)/2, (c+1)/2) SC(a/2, (b+1)/2, (c-1)/2)$ as claimed in the theorem. □

Remark. *In the Pfaffian expression for the various enumerations, it is possible to replace the occurrences of x with different variables and still get a nice factorisation. For example, the Pfaffian in Lemma 7 is changed to*

$$\mathrm{Pf}_{1\le i,j\le a}\left(\sum_{k=1}^{\left\lfloor\frac{a+b}{2}\right\rfloor}\left(\binom{b+y_1}{b+i-k}\binom{b+y_2}{j+k-a-1}-\binom{b+y_1}{b+j-k}\binom{b+y_2}{i+k-a-1}\right)\right).$$

The Pfaffian of this matrix is still a product of linear factors each involving only one of the two variables y_1 and y_2. Each of these groups of factors corresponds to one of the $B(\cdot,\cdot,\cdot)$–factors in Theorem 1.

This new matrix corresponds to self-complementary plane partitions with certain restrictions which leads to a modification of Stanley's proof giving an alternative proof of Theorem 2 using the expansion of $s_\lambda(x_1,\ldots,x_{n+1})\cdot s_\mu(x_1,\ldots,x_n)$ where λ and μ are rectangular partitions. (Note the different numbers of variables.)

References

[1] G. E. Andrews, *Plane partitions (II): The equivalence of the Bender–Knuth and the MacMahon conjectures*, Pacific J. Math., **72** (1977), 283–291.

[2] G. E. Andrews, *Plane partitions (I): The MacMahon conjecture*, Adv. in Math. Suppl. Studies, **1** (1978), 131–150.

[3] G. E. Andrews, *Plane partitions V: The t.s.s.c.p.p. conjecture*, J. Combin. Theory Ser. A, **66** (1994), 28–39.

[4] T. Eisenkölbl, (-1)-*enumeration of plane partitions with complementation symmetry*, Adv. in Appl. Math. **30**, (2003), no. 1-2, 53–95, arXiv:math.CO/0011175.

[5] I.M. Gessel, X. Viennot, *Determinant, paths and plane partitions*, Preprint, (1989).

[6] M. Ishikawa and M. Wakayama, *Minor summation formula of Pfaffians*, Linear and Multilinear Algebra **39** (1995), 285–305.

[7] G. Kuperberg, *An exploration of the permanent-determinant method*, Electron. J. Combin. **5**, (1998), #R46, arXiv:math.CO/9810091.

[8] G. Kuperberg, *Symmetries of plane partitions and the permanent determinant method*, J. Combin. Theory Ser. A **68** (1994), 115–151.

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

[10] P.A. MacMahon, *Combinatory Analysis*, vol. 2, Cambridge University Press, (1916); reprinted by Chelsea, New York, (1960).

[11] W. H. Mills, D. H. Robbins and H. Rumsey, *Proof of the Macdonald conjecture*, Inventiones Math. **66**, (1982), 73–87.

[12] S. Okada, *Applications of minor summation formulas to rectangular-shaped representations of classical groups*, J. Algebra, **205**, (1998), 337–367.

[13] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, Wadsworth & Brooks/Cole, Pacific Grove, California, (1986).

[14] R.P. Stanley, *Symmetries of plane partitions*, J. Combin. Theory Ser A **43** (1986), 103–113; Erratum **44** (1987), 310.

[15] J.R. Stembridge, *Some hidden relations involving the ten symmetry classes of plane partitions*, J. Combin. Theory Ser. A **68** (1994), 372–409.

[16] J.R. Stembridge, *On minuscule representations, plane partitions and involutions in complex Lie groups*, Duke Math. J. **73** (1994), 469–490.

[17] J. R. Stembridge, *The enumeration of totally symmetric plane partitions*, Adv. in Math. **111** (1995), 227–245.

COUNTING UNROOTED MAPS USING TREE-DECOMPOSITION

ÉRIC FUSY

Abstract. We present a new method to count unrooted maps on the sphere up to orientation-preserving homeomorphism. It is based on tree-decomposition and turns out to be very efficient to enumerate unrooted 2-connected and unrooted 3-connected maps. In particular, our method improves significantly on the best-known complexity to enumerate unrooted 3-connected maps, also called oriented convex polyedra.

Résumé. Nous présentons une nouvelle méthode pour compter les cartes non-enracinées sur la sphère orientée. La méthode est basée sur la notion de décomposition en arbre et s'avère très efficace pour énumérer les cartes 2-connexes et 3-connexes non-enracinées. En particulier, notre méthode améliore significativement les meilleurs résultats de complexité pour énumérer les cartes 3-connexes non enracinées, aussi appelées polyèdres convexes orientés.

Introduction

The enumeration of unrooted maps has been a well-studied problem for more than 20 years. Liskovets [4] was the first one to develop a general method for the enumeration of unrooted maps on the sphere up to orientation-preserving homeomorphism. It is based on two main tools: Burnside formula and study of the quotient maps.

With an adaptation of Burnside (orbit counting) lemma, the enumeration of unrooted maps comes down to enumerating rooted maps with a symmetry (rotation) of order k: for a family of maps enumerated according to the number n of edges, we write respectively c_n, c'_n and $c_n^{(k)}$ for the number of unrooted maps, rooted maps and rooted maps with a symmetry of order k; then c_n can be computed with the formula:

$$c_n = \frac{1}{2n}\left(c'_n + \sum_{k=2}^{n} \phi(k) c_n^{(k)}\right) \tag{1}$$

and a similar formula exists for the enumeration according to the number of vertices and faces, see Section 1. We represent rooted maps with a symmetry of order k as k-rooted maps, which are maps with k undistinguishable roots. Then, the *quotient map* of such a symmetric map is essentially a rooted map with two marked cells (a vertex, or the middle of a face or of an edge). The enumeration of such maps is easy to handle for the family of unconstrained maps [4], and we use these results in our article . Their approach can also be used for families of constrained maps, such as loopless maps [7], eulerian and unicursal maps [6] and 2-connected maps [5], but their treatment is less easy for these cases.

In this article, we introduce a new method for the enumeration of unrooted maps of a constrained family, based on the concept of tree-decomposition. Using this method, we carry out the enumeration of unrooted 2-connected and, above all, of unrooted 3-connected maps (also done by Walsh [13]). A first tree-decomposition "by multiple edges", allows (basically) to repercute a symmetry of order k of a k-rooted map on a symmetry of order k of a k-rooted 2-connected map. Hence it allows to find equations linking generating functions of k-rooted 2-connected maps and generating functions of k-rooted maps, which are easy to obtain from the method of quotient map. Then a second tree-decomposition "by separating 4-cycles" allows to find equations linking generating functions of k-rooted 3-connected maps, and generating functions of k-rooted 2-connected maps, which have already been obtained thanks to the first tree-decomposition. Finally, using Equation 1, we can enumerate unrooted 2-connected and unrooted 3-connected maps.

Main results Two results are obtained: a theorem about the *algebraic structure* of k-rooted maps and a theorem giving the complexity of enumeration of unrooted 2-connected and unrooted 3-connected maps. First we need a few notations:

Given a series $\alpha(t)$, a series $f(t)$ is said α-rational if there exists a rational function $\mathrm{R}(\mathrm{T})$ such that $f(t) = \mathrm{R}(\alpha(t))$. Given two series in two variables $\alpha_1(t_\bullet, t_\circ)$ and $\alpha_2(t_\bullet, t_\circ)$, a series in two variables $f(t_\bullet, t_\circ)$ is said (α_1, α_2)-rational if there exists a rational expression $\mathrm{R}(\mathrm{T}_1, \mathrm{T}_2)$ in two variables such that $f(t_\bullet, t_\circ) = \mathrm{R}(\alpha_1(t_\bullet, t_\circ), \alpha_2(t_\bullet, t_\circ))$.

Now we introduce the three "easily" algebraic series in one variable (they correspond to families of trees) $\beta(x)$, $\eta(y)$ and $\gamma(z)$ [1] given by

$$\beta(x) = x + 3\beta(x)^2, \qquad \eta(y) = \frac{y}{(1-\eta(y))^2}, \qquad \gamma(z) = z(1+\gamma(z))^2$$

and their versions in two variables $\beta_{1,2}(x_\bullet, x_\circ)$, $\eta_{1,2}(y_\bullet, y_\circ)$, and $\gamma_{1,2}(z_\bullet, z_\circ)$ (corresponding to bicolored trees of the respective families) given by

$$\left\{\begin{array}{lcl} \beta_1 &=& x_\bullet + \beta_1^2 + 2\beta_1\beta_2 \\ \beta_2 &=& x_\circ + \beta_2^2 + 2\beta_1\beta_2 \end{array}\right., \qquad \left\{\begin{array}{lcl} \eta_1 &=& \frac{y_\bullet}{(1-\eta_2)^2} \\ \eta_2 &=& \frac{y_\circ}{(1-\eta_1)^2} \end{array}\right., \qquad \left\{\begin{array}{lcl} \gamma_1 &=& z_\bullet(1+\gamma_2)^2 \\ \gamma_2 &=& z_\circ(1+\gamma_1)^2 \end{array}\right..$$

Theorem 1. *All series of k-rooted maps, k-rooted 2-connected maps and k-rooted 3-connected maps counted according to the number of edges of their quotient map are respectively β-rational, η-rational, and γ-rational.*

All series of k-rooted maps, k-rooted 2-connected maps and k-rooted 3-connected maps counted according to the number of vertices and faces (two parameters) of their quotient map are respectively (β_1, β_2)-rational, (η_1, η_2)-rational and (γ_1, γ_2)-rational.

In particular, all these series are algebraic.

Using algebraicity of the series of k-rooted maps, methods of computer algebra can be used to quickly extract their initial coefficients. Using Equation 1 (and its version in two variables if counting is done according to the number of vertices and faces), enumeration of unrooted maps can be performed very efficiently (using Maple, several hundreds of initial coefficients are easily computed):

Theorem 2. *For the enumeration of unrooted 2-connected and unrooted 3-connected maps according to the number of edges, to obtain the* N *first coefficients, we need $\mathcal{O}(\mathrm{N}\log(\mathrm{N}))$ operations.*

For the enumeration of unrooted 2-connected and unrooted 3-connected maps according to the number of vertices and faces, to obtain the table of the first coefficients with indices (i,j) with $i+j \leqslant \mathrm{N}$, we need $\mathcal{O}(\mathrm{N}^2)$ operations.

The arithmetical operations involved in the calculations are, as in [14]*, the multiplication of a '"large" integer with $\mathcal{O}(\mathrm{N})$ digits and of a "small" integer with $\mathcal{O}(\log(\mathrm{N}))$ digits.*

In particular, for the case of unrooted 3-connected maps, which is interesting as these objects correspond to polyedral maps, our complexity, in $\mathcal{O}(\mathrm{N}\log(\mathrm{N}))$ for one parameter and $\mathcal{O}(\mathrm{N}^2)$ for two parameters, improves significantly on the best known complexity obtained by Walsh [14]. Indeed, he had a complexity of $\mathcal{O}(\mathrm{N}^3)$ for one parameter and a complexity of $\mathcal{O}(\mathrm{N}^5)$ for two parameters.

Acknowledgments. The author would like to thank Gilles Schaeffer for his invaluable help in developping this new method. In particular he pointed the idea of tree-decomposition and helped to do some calculations and to correct the article.

1. Definitions and enumeration scheme

1.1. **Maps.** A map is a proper embedding of a connected graph (with possibly loops and multiple edges) on a closed oriented surface, where *proper* means that edges are smooth arcs that do not cross. All maps treated in this article are on the sphere. For enumeration, maps are considered up to all orientation-preserving homeomorphisms of the sphere, which also correspond to a continuous deformation of the sphere.

[1] We use three different variable names x, y, z, because they will later be linked by relations of change of variable.

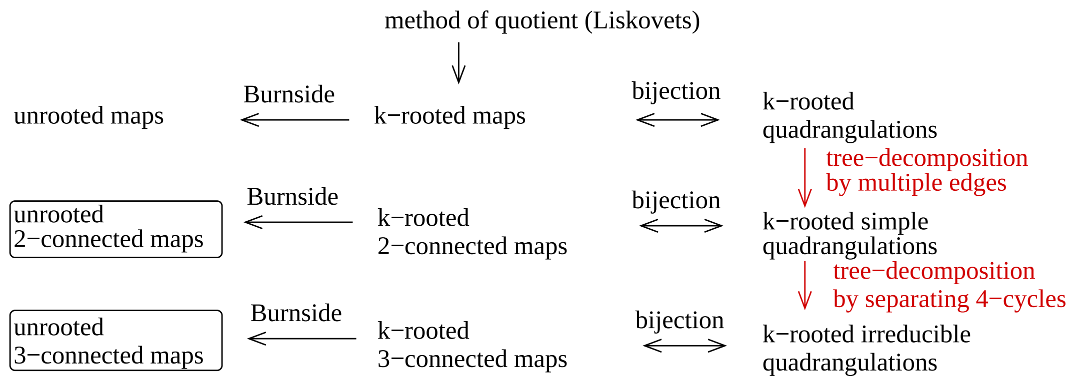


FIGURE 1. The scheme of the method to enumerate unrooted 2-connected and unrooted 3-connected maps

A map is said *2-connected* (or non-separable) if it has no loops and at least 2 of its vertices have to be removed to disconnect the map. A map is said *3-connected* if it has no loops nor multiple edges and at least 3 of its vertices have to be removed to disconnect the map.

A map is *rooted* by marking and orienting one of its edges. This operation suffices to eliminate all non trivial homeomorphism of the map. Hence, enumeration of rooted maps is more easy as we can use the root to start a recursive decomposition.

A *k-rooted map* (with $k \geqslant 2$) is a map with k **undistinguishable** roots. This means that the k objects obtained by marking differently (say, in blue) one of the k roots are equal. Rooted maps endowed with an automorphism of order k are in bijection with k-rooted maps (see [4] for more details). As k-rooted maps are easier to handle for our purpose, we will manipulate them rather than rooted maps with an automorphism of order k.

1.2. **Quadrangulations.** A *quadrangulation* is a map whose all faces have degree 4. A quadrangulation is said *simple* if it has no multiple edge. A quadrangulation is said *irreducible* if each 4-cycle of edges of the quadrangulation is the contour of one of its faces.

For each quadrangulation, its vertices can be colored in black and white so that each edge connects a black and a white vertex. Such a bicoloration is unique up to the choice of the colors. A quadrangulation endowed with such a bicoloration is said *bicolored*.

1.3. **Structure of k-rooted maps and method of quotient maps.** It was observed by Liskovets [4] that a k-rooted map can be realized by an embedding on the geometrical sphere, so that the embedding is invariant by a certain rotation of angle $2\pi/k$ of the sphere [2]. In addition, the points of the sphere crossed by the rotation-axis are either a vertex or the centre of a face, and can also be the middle of an edge if $k = 2$. These points are called the *poles* of the k-rooted map. The *type* of a k-rooted map is the type of its two poles. For example, if the two poles are a vertex and a face, then the k-rooted map is said to have type face-vertex.

Then, if we cut the sphere of the symmetrical embedding along two meridians forming a dihedral angle of $2\pi/k$, we can extract a sector of the map borded by these two meridians. By pasting together the two meridians, the sector becomes a map on the sphere. The symmetry of order k of the initial geometrical embedding ensures that this map is independant of the choice of the two meridians. We call this map the *quotient-map* of the k-rooted map. Observe that this quotient map has one root and two marked cells (the poles of the k-rooted map). The method of quotient maps developed by Liskovets consists in counting

[2]This point of view is not topologically relevant but it helps to have a geometrical intuition and it allows to define nicely the quotient of a k-rooted map.

k-rooted maps of a family by studying the structure of their quotient map. In the case of unconstrained maps, it works very well, as quotient maps are essentially rooted maps with two marked cells.

1.4. **Burnside formula adapted to unrooted maps.** Consider a family of maps on the sphere (for example the family of 2-connected maps). Let c_n, c'_n and $c_n^{(k)}$ denote respectively the number of unrooted, rooted and k-rooted maps of the family with n edges. Let c_{ij}, c'_{ij} and $c_{ij}^{(k)}$ denote respectively the number of unrooted, rooted and k-rooted maps of the family with $i+1$ vertices and $j+1$ faces. Then, Burnside (orbit counting) formula was adapted by Liskovets [4] to give the two following enumerative formulas for unrooted maps, where $\phi()$ is Euler totient function.

$$(2) \qquad 2nc_n = c'_n + \sum_k \phi(k) c_n^{(k)} \qquad\qquad 2(i+j)c_{ij} = c'_{ij} + \sum_k \phi(k) c_{ij}^{(k)}$$

As a consequence, enumeration of unrooted maps in one parameter (resp. two parameters) comes down to the enumeration of rooted maps (already done for 2-connected and 3-connected maps, see [8]) and of k-rooted maps of the family with one parameter (resp. two parameters).

1.5. **Bijection between maps and quadrangulations.** A classical result in map theory is a bijection between maps and bicolored quadrangulations, that we shall refer to as Tutte's bijection. We just detail its properties here. Tutte's bijection is a bijection between maps with n edges (resp. with i vertices and j faces) and bicolored quadrangulations with n faces (resp. with i black and j white vertices). Indeed, by this bijection, vertices, faces and edges of a map correspond respectively to black vertices, white vertices and faces of the bicolored quadrangulation.

In addition, under Tutte's bijection, rooted maps are in bijection with rooted quadrangulations and k-rooted maps are in bijection with so called *k-rooted bicolored quadrangulations*, which are defined as k-rooted quadrangulations such that the origins of the k roots have the same color when the quadrangulation is bicolored. We will only deal with such k-rooted quadrangulations and will shortly call them k-rooted quadrangulations. Observe that the type of a k-rooted map and the type of its associated k-rooted quadrangulation are linked by the above mentioned correspondance (for example 2-rooted maps with type edge-face are in bijection with 2-rooted quadrangulations with type face-white vertex), so that a k-rooted quadrangulation can only have type vertex-vertex if $k > 2$, and can also have type face-face and type face-vertex if $k = 2$.

Moreover, Tutte's bijection has the nice property that 2-connected maps are in bijection with bicolored simple quadrangulations and 3-connected maps are in bijection with bicolored irreducible quadrangulations. As a consequence, thanks to Tutte's bijection, the enumeration of k-rooted 2-connected maps by number of edges (resp. by numbers of vertices and faces) comes down to the enumeration of k-rooted simple quadrangulations by number of faces (resp. by numbers of black vertices and white vertices). The situation is the same for 3-connected maps, but with irreducible quadrangulations instead of simple quadrangulations, see Figure 1.

1.6. **Notations.** We will use the letters F, g and q to denote respectively generating functions of k-rooted, k-rooted simple and k-rooted irreducible quadrangulations. We will use the subscripts f, v, b and w to denote respectively a pole which is a face, a vertex, a black vertex and a white vertex. The subscripts b and w are only used for generating functions with two parameters, where we have to take the bicoloration into account. For example, $g_{vv}^{(k)}(y)$ is the series counting k-rooted simple quadrangulations of type vertex-vertex by the number of faces in their quotient map, and $q_{bw}^{(k)}(z_\bullet, z_\circ)$ is the series counting k-rooted irreducible quadrangulations, whose poles are a black and a white vertex, by the number of black and white vertices in their quotient map (and without counting the two axial vertices).

Lemma 3. *All generating functions of k-rooted quadrangulations in one (resp. two) variable are β-rational (resp. (β_1, β_2)-rational).*

Proof. From the method of quotient-map of Liskovets, the quotient-map of a k-rooted quadrangulation is essentially a quadrangulation with two marked cells (these cells can be a vertex or also a face if $k = 2$).

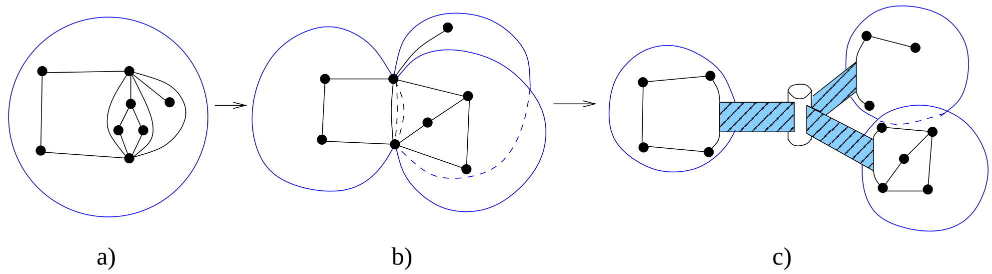


Figure 2. The tree-decomposition by multiple edges of a quadrangulation.

Hence the series counting these objects involve the first and second derivatives (or partial derivatives for two variables) of the series F counting rooted quadrangulations. This series is well-known to be β-rational in one variable [2] and (β_1, β_2)-rational in two variables [1] (see [10] for a combinatorial explanation). In addition, the fact of being β-rational (resp. (β_1, β_2)-rational) can easily be proved to be stable under derivation. Indeed, $d\mathrm{F}/dx = (d\mathrm{F}/d\beta)/(dx/d\beta)$ is the quotient of two β-rational expressions, and we can proceed similarly for two variables. The result follows. □

2. Tree-decompositions

2.1. **Tree-decomposition by multiple edges.** We explain here how to transform an unrooted quadrangulation Q (that may have multiple edges) into a tree with two kinds of nodes: nodes representing multiple edges and nodes representing simple quadrangulations.

One way to see this decomposition is as follows. Take a multiple edge of Q of multiplicity d. Cut the sphere along each of the d edges forming the multiple edge. In this way we obtain d sectors, each sector being delimited by two consecutive edges of the multiple edge. Now, for each sector, identify the two meridians corresponding to the two edges delimiting the sector by pasting them together. Thus we make out of each sector a map on the sphere and we can link these d maps, at their edge corresponding to the initial multiple edge, around a new node: this will be the node of the tree corresponding to the multiple edge. Now we can carry on recursively the tree-decomposition for each of the d maps, until all multiple edges have been split into nodes of the tree.

Another way to see this decomposition is to imagine that we do not cut along the edges of the multiple edge, but that we "blow" equally, from the interior of the sphere, each of the d sectors delimited by the multiple edge. We obtain thus d components drawn each on a sphere, where the d spheres are connected (glued) at the multiple edge, see Figure 2b. We can then represent this multiple edge as a rigid link (see Figure 2c) around which the d components are linked via their unique edge belonging to the multiple edge. We can also here carry on the decomposition for each of the d components.

2.2. **Tree-decomposition by separating 4-cycles.** In this section we transform a simple quadrangulation with at least 3 faces into a tree with two kinds of nodes: so-called axis-nodes and nodes corresponding to irreducible quadrangulations. The description of this tree-decomposition can also be found in [3]. We describe first the tree-decomposition for rooted objects and we will see then that we can also see the tree-decomposition on unrooted objects.

Let us first define the axis-map with k faces ($k \geqslant 3$) as the simple quadrangulation consisting of two pole-vertices linked by k parallel chains of 2 edges, each couple of two consecutive paths forming one of the k faces of the axis-map, see Figure 3a.

Now we state the following lemma of decomposition of a rooted simple quadrangulation Q with at least 3 faces:

Lemma 4. *There exists a unique rooted quadrangulation* Q_0*, with maximal possible number* $k+1$ *of faces such that:*

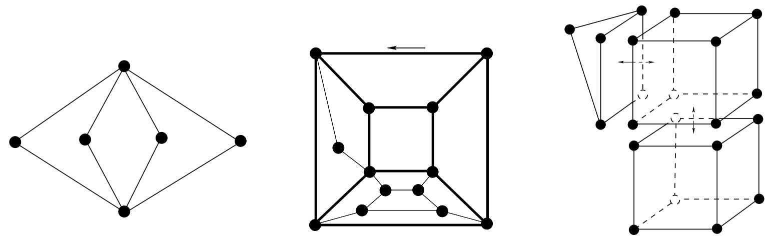


FIGURE 3. An axis-map with 4 faces (a). The tree-decomposition of a quadrangulation by separating 4-cycles, performed with a root (b) or without a root (c).

- Q_0 *is an axis-map or an irreducible quadrangulation.*
- *There are k rooted simple quadrangulations* $\mathrm{Q}_1, \ldots, \mathrm{Q}_k$ *with at least 2 faces such that* Q *can be seen as the quadrangulation* Q_0 *where each of the k non root faces* f_i *of* Q_0 *is substituted in a canonical way by one of the* Q_i, $1 \leqslant i \leqslant k$, *the contour of* f_i *being replaced by the contour of the root face of* Q_i.

Proof. If there exists an internal chain of length 2 between two opposite vertices of the outer face of Q, take the sequence of all chains of length 2 (including the 2 outer ones) between these two vertices. Forgetting all other edges, we get an axis-map. Hence Q can be seen as this axis-map where each non root face is substituted by a quadrangulation.

Otherwise, define a proper 4-cycle of Q as a 4-cycle of edges different from the contour of the root face of Q. Here we have to see Q as drawn on the plane with its root face as infinite face, so that we can distinguish interior and exterior. A proper 4-cycle is said maximal if it is not strictly included in the interior of any other proper 4-cycle. It can easily be shown (see [8]) that the interiors of maximal proper 4-cycles partition the interior of Q. Let Q_0 be the rooted quadrangulation obtained from Q by keeping the contour of the root face and of the maximal proper 4-cycles of Q. The quadrangulation Q_0 is trivially irreducible by maximallity of the 4-cycles of which we have kept the contour. Hence we are in the case where Q can be seen as a rooted irreducible quadrangulation where each inner face is substituted by a rooted quadrangulation. □

The first (resp. second) case of Lemma 4 correspond to the case where the root node of the (rooted) decomposition-tree is an axis-node (resp. a node which is an irreducible quadrangulation). For example, on Figure 3b, the rooted quadrangulation can be seen as a (rooted) cube where two faces are substituted by another cube and an axis-map with 3 faces.

Remark We make the following distinction for the case of an axis-node: if the parallels chains of length 2 are incident to the origin of the root, the root node of the tree is said a *vertical* axis-node, otherwise, it is said an *horizontal* axis-node.

Now we can carry on the tree-decomposition for each rooted quadrangulation Q_i with $1 \leqslant i \leqslant k$. Thus, we get finally a (rooted) decomposition-tree with axis nodes and nodes which are irreducible quadrangulations. Observe that, if Q_0 and the root node of one of the Q_i are simultaneously axis-nodes, then they are stretched in perpendicular directions by maximallity of the number of faces of Q_0.

Observe that the preceding decomposition on rooted objects ensures that, as in Section 2.1, we can “blow” from the interior of the sphere to “sculpt” the quadrangulation Q in a tree with nodes which are irreducible quadrangulations and axis-nodes, these nodes being connected (glued) at so-called *interconnection-faces*, see Figure 3c. Hence we can say that an unrooted simple quadrangulation “is” its tree-decomposition (after a judicious deformation of the sphere). We see thus that the geometrical shape of the tree in the space does not depend on the face of the quadrangulation where we choose to place the root to start the tree-decomposition.

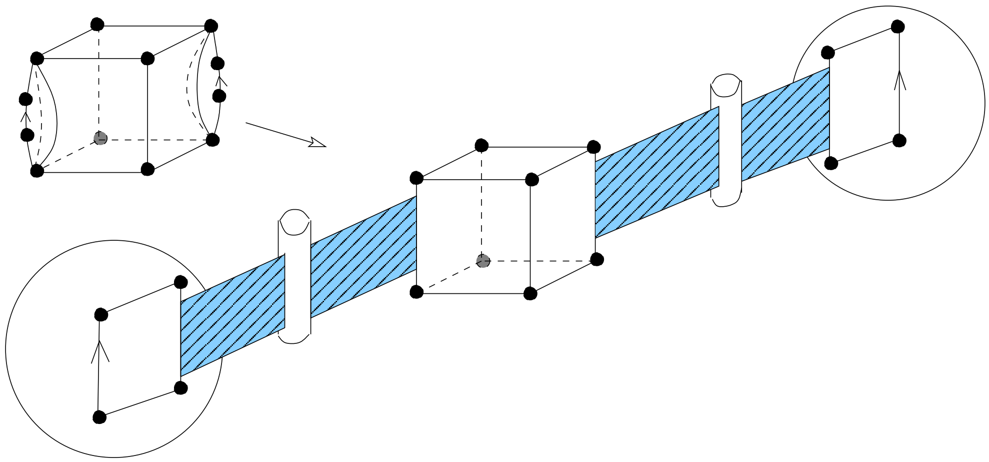


FIGURE 4. Repercussion of the symmetry of a k-rooted quadrangulation on its decomposition-tree.

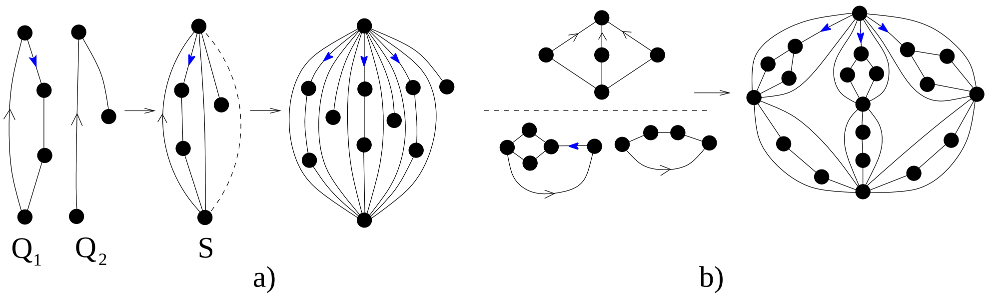


FIGURE 5. Construction of a k-rooted quadrangulation of type a (Figure a), and of type b (Figure b).

2.3. **Centre of a tree.** The *centre* of a tree T is defined in the following recursive way. If T is reduced to an edge or a node, then the centre of T is this edge (resp. this node). Otherwise, remove all leaves of T to obtain a (shrinked) tree $\widetilde{\mathrm{T}}$. Then the centre of T is defined to be the centre of $\widetilde{\mathrm{T}}$.

The important point is that the definition does not require that T is rooted. Hence the centre is invariant under any symmetry of T.

3. Using the tree-decomposition by multiple edges to enumerate unrooted 2-connected maps

3.1. **Repercussion of the symmetry of a k-rooted quadrangulation on its decomposition-tree.** As we have seen, the tree-decomposition by multiple edges of a quadrangulation Q can be seen as a deformation of the sphere on which Q is drawn and by splitting multiple edges into links so as to form a decomposition-tree "living" in the 3D-space. In addition, if Q is k-rooted, then its decomposition-tree is invariant under the symmetry (rotation) of order k induced by its k-root. Hence, the centre of the tree is fixed by the symmetry, see Figure 4. This centre can be a node or an edge of the tree. However, the case of an edge is excluded because an edge of the tree always links a node of type "multiple edge" and a node of type "simple quadrangulation", hence an edge of the tree can not be invariant under a non-trivial symmetry of the tree. As a consequence, the centre is a node and there are two cases: either it is a node of type multiple edge -we say that Q has type a- or it is a node of type simple quadrangulation -we say that Q has type b-.

3.2. **Case where the centre is a multiple edge (type a).** First we need to define a *simply rooted* quadrangulation as a quadrangulation whose root edge does not belong to a multiple edge. We also define a bi-rooted quadrangulation as a quadrangulation having a secondary root which is differently marked (say in blue).

Now we explain how to construct a k-rooted quadrangulation whose centre of the decomposition-tree is a multiple edge with multiplicity $k \cdot d$ ($d \geqslant 1$), see Figure 5a. Take a bi-rooted, simply-rooted (i.e. whose primary root is a simple edge) quadrangulation Q_1. Cut it along its primary root-edge, thus transforming Q_1 into a sector with two bording meridians. Among these two meridians, we call root-meridian the one corresponding to the right part of the cutted edge (we imagine that the edge we have cut along has a "width").

Now take $d-1$ simply rooted quadrangulations $\mathrm{Q}_2, \dots, \mathrm{Q}_d$ and perform the same cutting operation as for Q_1. Then paste the root meridian of Q_2 with the non-root meridian of Q_1, the pasting operation being such that the orientations of the roots of the two sectors coincide. Then, iteratively, for each $i \leqslant d$, paste the root meridian of Q_i with the non-root meridian of Q_{i-1}.

We obtain finally a big sector S whose root meridian is the root meridian of Q_1. Now make k copies $\mathrm{S}_1, \dots, \mathrm{S}_k$ of S and, for each $1 \leqslant i \leqslant k$, paste the root meridian of S_i with the non-root meridian of S_{i-1}. In this way we obtain finally a quadrangulation consisting of k identical sectors, each carrying a blue root (the secondary root of Q_1). By erasing the mark of the primary root of Q_1 and of the roots of $\mathrm{Q}_2 \dots \mathrm{Q}_d$ in each sector, we obtain a k-rooted quadrangulation of type a. Observe that each k-rooted quadrangulation of type a is obtained exactly twice by this construction. Indeed, the inverse operation consists in choosing an extremity v (two possibilities) of the central multiple edge and then orienting all edges of the multiple edge toward v.

Writing $f(x)$ for the series counting simply rooted quadrangulations by their number of faces, this construction gives the series counting k-rooted quadrangulations of type a:

$$\frac{1}{2}(4xf'(x)).\frac{1}{1-f(x)}$$

In addition, all objects constructed in this way have clearly type vertex-vertex.

3.3. **Case where the centre is a simple quadrangulation (type b).** Here we give a construction of k-rooted quadrangulations of type b as composed objects, see Figure 5b. Take a k-rooted simple quadrangulation Q_s. For the k-orbite of root edges, either leave its k edges untouched (Case 1) or perform the following operation (Case 2): take a bi-rooted bicolor-consistent quadrangulation $\widetilde{\mathrm{Q}}$. Then cut Q_s along each of its k root edges and cut $\widetilde{\mathrm{Q}}$ along its primary root edge, transforming $\widetilde{\mathrm{Q}}$ into a sector S bordered by two meridians. Take k copies of S and for each (cutted) root-edge e of Q_s, place a copy of S in the empty sector of Q_s leaved by the cutting of e. This placement is made by pasting the two meridians of S with the two border-edges of Q_s created by the cutting of e, and by making the orientation of e and of the primary root edge of S coincide.

Proceed similarly for each k-orbite of non-root edges of Q_s, with the only difference that the quadrangulation used for the substitution is not bi-rooted but just rooted. Finally, keep only the k marks of the roots of Q_s if we are in Case 1 (i.e. no substitution at the root edges of Q_s), and keep only the marks of the secondary roots of the k copies of $\widetilde{\mathrm{Q}}$ if we are in Case 2. Thus, we obtain a k-rooted quadrangulation Q of type b.

Observe that k-rooted quadrangulations of type b obtained by this construction always have the following property: their k root edges are simple if their incident face (the face on their right) belongs to the central simple quadrangulation (because this case corresponds to Case 1 where there is no substitution at the root edges of Q_s). The missing k-rooted quadrangulations of type b are obtained by the same construction, with the difference that we always cut the k root edges of Q_s. Then the other difference is that the first substituted quadrangulation $\widetilde{\mathrm{Q}}$ is not bi-rooted but just rooted. At the end of this construction, we only keep the mark of the k roots of Q_s

Similarly as in Section 3.2, these two complementary constructions allow to obtain all k-rooted quadrangulations of type b exactly twice. Writing $\mathrm{F}(x)$ for the series counting rooted quadrangulations by their number of faces and $\mathrm{E}(x) = 2x\mathrm{F}'(x)+\mathrm{F}(x)+1$, this construction gives the following three series, depending on the type of Q_s

$$\frac{\mathrm{E}(x)}{1+\mathrm{F}(x)} g_{vv}^{(k)}\left((1+\mathrm{F}(x))^2\right), \quad \mathrm{E}(x) g_{fv}^{(2)}\left((1+\mathrm{F}(x))^2\right), \quad \mathrm{E}(x)(1+\mathrm{F}(x)) g_{ff}^{(2)}\left((1+\mathrm{F}(x))^2\right).$$

3.4. **Obtaining the equations.** As k-rooted quadrangulations are partitioned in two sets whether the centre of their decomposition-tree is a multiple edge or a simple quadrangulation, we obtain the following equations by taking the sum of the series obtained in Section 3.2 and Section 3.3:

$$\mathrm{F}_{vv}^{(k)}(x) = 2\frac{xf'(x)}{1-f(x)} + \frac{\mathrm{E}(x)}{1+\mathrm{F}(x)} g_{vv}^{(k)}\left((1+\mathrm{F}(x))^2\right) \tag{3}$$

$$\mathrm{F}_{fv}^{(2)}(x) = \mathrm{E}(x) g_{fv}^{(2)}\left((1+\mathrm{F}(x))^2\right) \tag{4}$$

$$\mathrm{F}_{ff}^{(2)}(x) = \mathrm{E}(x)(1+\mathrm{F}(x)) g_{ff}^{(2)}\left((1+\mathrm{F}(x))^2\right) \tag{5}$$

where the only unknown series are $g_{vv}^{(k)}$, $g_{fv}^{(2)}$ and $g_{ff}^{(2)}$.

Similar equations can be easily obtained in two variables by taking the bicoloration of vertices into account. Writing $df(x_\bullet, x_\circ) = \frac{d}{dt} f(tx_\bullet, tx_\circ)_{t=1}$ and adapting E in two variables as $\mathrm{E}(x_\bullet, x_\circ) = 2\frac{d}{dt}\mathrm{F}(tx_\bullet, tx_\circ)_{t=1} + \mathrm{F}(x_\bullet, x_\circ) + 1$, Equation 3 becomes for example:

$$\left\{ \begin{array}{rcl} \mathrm{F}_{bw}^{(k)}(x_\bullet, x_\circ) & = & 2\frac{df}{1-f} + \frac{\mathrm{E}}{1+\mathrm{F}} g_{bw}^{(k)}\left(x_\bullet(1+\mathrm{F})^2, x_\circ(1+\mathrm{F})^2\right) \\ \mathrm{F}_{bb}^{(k)}(x_\bullet, x_\circ) & = & \frac{\mathrm{E}}{1+\mathrm{F}} g_{bb}^{(k)}\left(x_\bullet(1+\mathrm{F})^2, x_\circ(1+\mathrm{F})^2\right) \\ \mathrm{F}_{ww}^{(k)}(x_\bullet, x_\circ) & = & \frac{\mathrm{E}}{1+\mathrm{F}} g_{ww}^{(k)}\left(x_\bullet(1+\mathrm{F})^2, x_\circ(1+\mathrm{F})^2\right) \end{array} \right. \tag{6}$$

where all series (including f and F) have two variables, one for the number of black vertices, the other one for the number of white vertices.

Observe that, as the series $\mathrm{F}_{vv}^{(k)}$ (in one or two variables) does not depend on k as was observed in Lemma 3, it follows from the form of Equation 3 and 6 that the series $g_{vv}^{(k)}$ does not depend on k, hence the exponent (k) can be ommited.

Lemma 5. *The series g counting rooted simple quadrangulations and all series of k-rooted simple quadrangulations in one variable (resp. two variables) are η-rational (resp. (η_1, η_2)-rational).*

Proof. Using Lemma 3, we know that $\mathrm{F}(x)$, $\mathrm{F}_{vv}(x)$, $\mathrm{F}_{fv}(x)$ and $\mathrm{F}_{ff}(x)$ are β-rational, and so are x (because $x = \beta - 3\beta^2$), $f(x)$ (because $\mathrm{F} = f/(1-f)$), and $\mathrm{E}(x)$. Hence it follows from Equations 3, 4 and 5 that $g_{vv}\left(x(1+\mathrm{F})^2\right)$, $g_{fv}\left(x(1+\mathrm{F})^2\right)$ and $g_{ff}\left(x(1+\mathrm{F})^2\right)$ are β-rational. Now we have to make the change of variable $y = x(1+\mathrm{F})^2$. It can easily be proved (or found in [2]) that $\beta(x) = \eta(y)/(1+3\eta(y))$ when y and x are linked by the change of variable $y = x(1+\mathrm{F})^2$. Hence, replacing $\beta(x)$ by $\eta(y)/(1+3\eta(y))$ in the β-rational expression of $g_{vv}\left(x(1+\mathrm{F})^2\right)$, $g_{fv}\left(x(1+\mathrm{F})^2\right)$ and $g_{ff}\left(x(1+\mathrm{F})^2\right)$, we obtain η-rational expressions for $g_{vv}(y)$, $g_{fv}(y)$ and $g_{ff}(y)$. Finally, $g(y)$ is η-rational from [2].

We can proceed similarly in two variables, using the fact that $\beta_1(x_\bullet, x_\circ)$ and $\beta_2(x_\bullet, x_\circ)$ have a rational expression in terms of $\eta_1(y_\bullet, y_\circ)$ and $\eta_2(y_\bullet, y_\circ)$ when $(y_\bullet, y_\circ)$ and $(x_\bullet, x_\circ)$ are linked by the change of variable $(y_\bullet, y_\circ) = (x_\bullet(1+\mathrm{F})^2, x_\circ(1+\mathrm{F})^2)$. □

Lemma 6. *The* N *initial coefficients counting unrooted 2-connected maps according to their number of edges can be computed with $\mathcal{O}(\mathrm{N}\log(\mathrm{N}))$ operations.*

The table of initial coefficients with indices (i, j) and $i + j \leqslant \mathrm{N}$ counting unrooted 2-connected maps according to their number of vertices and faces can be computed with $\mathcal{O}(\mathrm{N}^2)$ operations.

Proof. First we use the following notation. For a series f in one variable (resp. two variables), we denote by $\mathcal{C}_\mathrm{N}(f)$ the number of operations necessary to extract its N initial coefficients (resp. its coefficients with indices (i, j) and $i + j \leqslant \mathrm{N}$). Writing c_n (resp. c_{ij}) for the number of unrooted 2-connected maps with n

edges (resp. $i+1$ vertices and $j+1$ faces), Equation 2 (Burnside formula) can be easily transposed in the following equations on series:

$$\begin{aligned}\sum_n 2nc_n y^n &= g(y) + y g_{fv}(y^2) + y^2 g_{ff}(y^2) + \sum_{k\geqslant 2}\phi(k) g_{vv}(y^k)\\ \sum_{i,j} 2(i+j)c_{ij} y_\bullet^i y_\circ^j &= g(y_\bullet, y_\circ) + y_\bullet g_{fb}(y_\bullet^2, y_\circ^2) + y_\circ g_{fw}(y_\bullet^2, y_\circ^2) + y_\bullet y_\circ g_{ff}(y_\bullet^2, y_\circ^2)\\ &\quad + \sum_{k\geqslant 2}\phi(k)\left(\frac{y_\bullet}{y_\circ} g_{bb}(y_\bullet^k, y_\circ^k) + g_{bw}(y_\bullet^k, y_\circ^k) + \frac{y_\circ}{y_\bullet} g_{ww}(y_\bullet^k, y_\circ^k)\right)\end{aligned}$$

According to Lemma 5, $g(y)$, $g_{fv}(y)$, $g_{ff}(y)$ and $g_{vv}(y)$ are η-rational, hence they are algebraic (because they live in the algebraic extension of the algebraic series $\eta(y)$). As a consequence, they are differentiably finite (see [11]), i.e. solution of a linear differential equation with polynomial coefficients. Taking coefficient $[y^n]$ in this differential equation yields that the coefficients of these series verify a linear recurrence with polynomial coefficients. As a consequence, the N initial coefficients of these series can be computed with $\mathcal{O}(\mathrm{N}\log(\mathrm{N}))$ "arithmetical" operations, which are the multiplication of a "small" integer with $\mathcal{O}(\log(\mathrm{N}))$ bits and of a "large" integer with $\mathcal{O}(\mathrm{N})$ bits (same operations as in [14]). Hence, $\mathcal{C}_\mathrm{N}\left(\sum 2nc_n\right) = \mathcal{C}_\mathrm{N}(g) + \mathcal{C}_{\mathrm{N}/2}(g_{fv}+g_{ff}) + \sum_{k=2}^{\mathrm{N}}\mathcal{C}_{\mathrm{N}/k}(g_{vv}) = \mathcal{O}(\mathrm{N}) + \mathcal{O}(\mathrm{N}/2) + \sum_{k=2}^{\mathrm{N}}\mathcal{O}(\mathrm{N}/k) = \mathcal{O}(\mathrm{N}\log(\mathrm{N}))$.

Similarly, the coefficients of an algebraic series in two variables "essentially" verify a linear recurrence, this time with two indices. As a consequence, if $f(y_\bullet, y_\circ)$ is algebraic, then $\mathcal{C}_\mathrm{N}(f) = \mathcal{O}(\mathrm{N}^2)$. As series of k-rooted simple quadrangulations in two variables are (η_1,η_2)-rational, they are algebraic. Hence, $\mathcal{C}_\mathrm{N}\left(\sum_{i,j}2(i+j)c_{ij}\right) = \mathcal{C}_\mathrm{N}(g) + \mathcal{C}_{\mathrm{N}/2}(g_{ff}+g_{fb}+g_{fw}) + \sum_{k=2}^{\mathrm{N}}\mathcal{C}_{\mathrm{N}/k}(g_{bb}+g_{bw}+g_{ww}) = \mathcal{O}(\mathrm{N}) + \mathcal{O}((\mathrm{N}/2)^2) + \sum_{k=2}^{\mathrm{N}}\mathcal{O}((\mathrm{N}/k)^2) = \mathcal{O}(\mathrm{N}^2)$ where we use the fact that $\sum_k 1/k^2$ converges. $\square$

4. Using the tree-decomposition by separating 4-cycles to enumerate unrooted 3-connected maps

4.1. Repercussion of the symmetry of a k-rooted simple quadrangulation on its decomposition-tree. First we introduce the families $\mathcal{W}$ of rooted simple quadrangulations with at least two faces and the family $\mathcal{G}$ consisting of the objects of $\mathcal{W}$ whose root node of the decomposition tree is not an horizontal axis-node. We write $\mathrm{W}(y)$ and $\mathrm{G}(y)$ for the series counting these two families by their number of faces (notations of [3]). Observe that $\mathrm{W}(y) = g(y) - 2y$ and $\mathrm{W}(y)/y = \frac{\mathrm{G}(y)/y}{1-\mathrm{G}(y)/y}$. We define also the families $\mathcal{W}'$ and $\mathcal{G}'$ of objects of $\mathcal{W}$ and $\mathcal{G}$ having a secondary root incident to a face different from the root face. The series counting objects of $\mathcal{W}'$ and $\mathcal{G}'$ by their number of faces are respectively $4\mathrm{C}(y)$ and $4\mathrm{B}(y)$ where $\mathrm{C}(y) = y^2\frac{d}{dy}(\mathrm{W}(y)/y)$ and $\mathrm{B}(y) = y^2\frac{d}{dy}(\mathrm{G}(y)/y)$.

Let Q be a simple k-rooted quadrangulation with at least 3 faces. Here we work with $k \geqslant 3$. The case $k=2$ is more difficult (for example a symmetry of order 2 of an axis-map can exchange its poles), but can also be thoroughly treated, see the full version. As in Section 3.1, the decomposition tree of Q is invariant under the symmetry of order k induced by the k-root of Q. Hence, the centre of the tree (which is a node because $k>2$) is invariant by the symmetry. Also here two cases arise: either the centre is an axis-node -we say that Q has type $\mathfrak{a}$- or it is an irreducible quadrangulation -we say that Q has type $\mathfrak{b}$-.

4.2. Construction of k-rooted simple quadrangulations of type $\mathfrak{a}$. Similarly as in Section 3.2, we construct a k-rooted simple quadrangulation, whose centre of the decomposition tree is an axis-map with $k\cdot d$ faces, as a composed object. Take a k-rooted axis-map with $k\cdot d$ faces and whose all roots point toward a pole of the axis-map, that we call the *north pole*. Then take k copies of an object Q_1 of $\mathcal{G}'$ and substitute each root face of the axis-map by one of these copies, making the primary root of the copies of Q_1 be oriented toward the north pole of the axis-map. Proceed similarly for each k-orbite of non-root faces of the axis-map, with the only difference that the substituted objects are k copies of an object of $\mathcal{G}$ instead of $\mathcal{G}'$. Finally keep only the marks of the secondary root of the k copies of Q_1.

As in Section 3.2, each k-rooted simple quadrangulation of type $\mathfrak{a}$ is obtained exactly twice by this construction. The series counting k-rooted simple quadrangulations of type $\mathfrak{a}$ is:

$$2\frac{\mathrm{B}(y)}{y}\frac{1}{1-\mathrm{G}(y)/y}$$

and all these objects have type vertex-vertex.

4.3. **Construction of k-rooted simple quadrangulations of type $\mathfrak{b}$.** As precedently, we give a construction of k-rooted simple quadrangulations of type $\mathfrak{b}$ as composed objects. Take a k-rooted irreducible quadrangulation Q_{irr} (remark that Q_{irr} has type vertex-vertex because $k>2$). Take k copies of an object Q_1 of $\mathcal{W}'$ and substitute each root face of Q_{irr} by one of the copies of Q_1 in a "canonical" way, e.g. by superposing the primary root edge of Q_1 with the root edge of the face where the substitution takes place. Then proceed similarly for each k-orbite of non-root faces of Q_{irr}, with the difference that the substituted objects are k copies of an object of $\mathcal{W}$ instead of $\mathcal{W}'$. Finally keep only the marks of the secondary root of the k copies of Q_1.

By this construction, all k-rooted simple quadrangulations of type $\mathfrak{b}$ are obtained exactly 4 times. Indeed, as a quadrangular face has 4 sides, there are 4 possibilities to guess the primary root edge of the k copies of Q_1. We obtain the following series counting k-rooted simple quadrangulations of type $\mathfrak{b}$:

$$\frac{\mathrm{C}(y)}{\mathrm{W}(y)}q_{vv}^{(k)}(\mathrm{W}(y)/y)$$

4.4. **Obtaining the equations.** As k-rooted simple quadrangulations are partitioned in two sets whether the center of their decomposition tree is an axis-node or an irreducible quadrangulation, summing the series obtained in Section 4.2 and Section 4.3, we obtain the following equation linking series of k-rooted simple quadrangulations with series of k-rooted irreducible quadrangulations, for $k>2$:

$$g_{vv}^{(k)}(y) = 2\frac{\mathrm{B}(y)}{y}\frac{1}{1-\mathrm{G}(y)/y} + \frac{\mathrm{C}(y)}{\mathrm{W}(y)}q_{vv}^{(k)}(\mathrm{W}(y)/y) \tag{7}$$

Similar equations can be easily obtained in two variables by taking the bicoloration of Q into account. Writing $\mathrm{C}(y_\bullet, y_\circ) = y_\bullet\frac{\partial \mathrm{W}}{\partial y_\bullet} + y_\circ\frac{\partial \mathrm{W}}{\partial y_\circ} - \mathrm{W}$ and $\mathrm{B}(y_\bullet, y_\circ) = y_\bullet\frac{\partial \mathrm{G}}{\partial y_\bullet} + y_\circ\frac{\partial \mathrm{G}}{\partial y_\circ} - \mathrm{G}$ for the versions in two variables of $\mathrm{C}(y)$ and $\mathrm{B}(y)$, the version in two variables of Equation 7 becomes:

$$\begin{aligned}
g_{bb}^{(k)}(y_\bullet, y_\circ) &= \frac{\mathrm{B}}{y_\bullet}\frac{1}{1-\mathrm{G}/y_\bullet} + \frac{\mathrm{C}}{\mathrm{W}}q_{bb}^{(k)}(\mathrm{W}/y_\circ, \mathrm{W}/y_\bullet)\\
g_{ww}^{(k)}(y_\bullet, y_\circ) &= \frac{\mathrm{B}}{y_\circ}\frac{1}{1-\mathrm{G}/y_\circ} + \frac{\mathrm{C}}{\mathrm{W}}q_{ww}^{(k)}(\mathrm{W}/y_\circ, \mathrm{W}/y_\bullet)\\
g_{bw}^{(k)}(y_\bullet, y_\circ) &= \frac{\mathrm{C}}{\mathrm{W}}q_{bw}^{(k)}(\mathrm{W}/y_\circ, \mathrm{W}/y_\bullet)
\end{aligned}$$

Observe that these equations are the same for all values of k. As we have already seen that $g_{vv}^{(k)}(y)$ does not depend on k, then $q_{vv}^{(k)}(z)$ does not depend on k so that exponent (k) can be ommited.

Lemma 7. *All series of k-rooted irreducible quadrangulation in one variable (resp. two variables) are γ-rational (resp. (γ_1, γ_2)-rational).*

Proof. Similar to the proof of Lemma 5. In one variable, we use the form of Equation 7 to see that $q_{vv}^{(k)}(\mathrm{W}(y)/y)$ is η-rational. Then we use the fact [8] that $\eta(y)$ has a rational expression in terms of $\gamma(z)$ when z and y are linked by the change of variable $z = \mathrm{W}(y)/y$. Substituting η by this expression in the η-rational expression of $q_{vv}^{(k)}(\mathrm{W}(y)/y)$, we obtain a γ-rational expression for $q_{vv}^{(k)}(z)$.

The proof for two variables is similar, using in particular the fact that $\eta_1(y_\bullet, y_\circ)$ and $\eta_2(y_\bullet, y_\circ)$ have a rational expression in terms of $\gamma_1(z_\bullet, z_\circ)$ and $\gamma_2(z_\bullet, z_\circ)$ when $(z_\bullet, z_\circ)$ and $(y_\bullet, y_\circ)$ are linked by the change of variable $(z_\bullet, z_\circ) = (\mathrm{W}/y_\circ, \mathrm{W}/y_\bullet)$. □

Lemma 8. *The* N *initial coefficients counting unrooted 3-connected maps according to their number of edges can be computed with* $\mathcal{O}(\mathrm{N}\log(\mathrm{N}))$ *operations.*

The table of initial coefficients with indices (i, j) *and* $i + j \leqslant \mathrm{N}$ *counting unrooted 3-connected maps according to their number of vertices and faces can be computed with* $\mathcal{O}(\mathrm{N}^2)$ *operations.*

Proof. Using the algebraicity of the generating function of k-rooted irreducible quadrangulations, we can perform the same treatment as in the proof of Lemma 6. □

Finally, Lemma 6 and 8 yield Theorem 2. Using Tutte's bijection between k-rooted objects (see also Figure 1), Lemma 3, 5 and 7 yield Theorem 1.

5. Conclusion

We have proposed an original and efficient method to enumerate unrooted maps. In particular, we have improved significantly on the complexity of counting oriented convex polyedra (unrooted 3-connected maps).

Our method is flexible and can be adapted to enumerate other families of unrooted maps. For example, a similar scheme can be used to count unrooted loopless and then unrooted maps without loops and multiple edges. This time, a first tree decomposition, said "by loops" allows to obtain enumeration of k-rooted loopless maps from k-rooted maps. Then the tree decomposition by multiple edges (this time on k-rooted maps instead of k-rooted quadrangulations as in this article) allows to enumerate k-rooted maps without loop and multiple edge from loopless k-rooted maps.

Another very interesting problem is the enumeration of unrooted 3-connected maps on the sphere up to all homeomorphisms (including orientation-reversing). Indeed according to Whitney's Theorem, 3-connected planar graphs have a unique toplogical embedding on the sphere, so that these unrooted 3-connected maps exactly correspond to unlabelled 3-connected planar graphs. In this case, a Burnside formula is also available, letting the problem come down to the enumeration of oriented k-rooted 3-connected maps, but also orientation-reversing ones such as 2-rooted 3-connected maps representing a reflexion. The tree-decomposition by separating 4-cycles can be used to obtain an equation linking 2-rooted 2-connected maps and 2-rooted 3-connected maps of type reflexion. Hence, the method of tree decomposition is also here promising.

References

[1] D. Arques. Relations fonctionnelles et dénombrement des cartes pointées sur le tore. *J. Comb. Theory Ser. B* 43:253-274, 1987

[2] I.P. Goulden, D.M. Jackson. Combinatorial enumeration. John Wiley and Sons Inc., New York. 1983

[3] S. Kunz-Jacques, G. Schaeffer. The asymptotic number of prime alternating links. *Proceedings of the International Conference FPSAC'01, Phoenix.*

[4] V.A. Liskovets. A census of non-isomorphic planar maps, *Coll. Math. Soc. J. Bolyai, Proc. Conf. Algebr. Meth. in Graph Th.* 25:p. 2 479-494, 1981

[5] V.A. Liskovets, T.R.S. Walsh. The enumeration of non-isomorphic 2-connected planar maps. *Canad. J. Math.* 3:417-435, 1983

[6] V.A. Liskovets, T.R.S. Walsh. Enumeration of eulerian and unicursal planar maps. *Discr. Math.*, to appear.

[7] V.A. Liskovets, T.R.S. Walsh. Counting unrooted loopless planar maps. *Proceedings of the International Conference FPSAC'04.*

[8] R.C. Mullin, P.J. Schellenberg. The enumeration of c-nets via quadrangulations. *J. Combinatorial Theory*, 4:259-276, 1968

[9] B. Salvy, P. Zimmermann. Gfun: A Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Transactions on Mathematical Software*, 20(2):163–177, June 1994.

[10] G. Schaeffer. Bijective census and random generation of Eulerian planar maps with prescribed vertex degrees. *Electron. J. Combin.* **4** 1:Research Paper 20. 1997

[11] R. P. Stanley. Differentiably Finite Power Series. *Europ. J. Combinatorics* **1**:175-188, 1980

[12] W.T. Tutte. A census of planar maps, *Canad. J. Math.*, 15:249-271, 1963

[13] T.R.S. Walsh. Counting Non-isomorphic Three-Connected Planar Maps. *J. Combinatorial Theory*, Series B 32:33-44, 1982

[14] T.R.S. Walsh. Efficient enumeration of sensed maps, manuscript.

Éric Fusy, INRIA Rocquencourt, Projet ALGO BP 105, 78153 Le Chesnay, and LIX (Ecole Polytechnique)
E-mail address: Eric.Fusy@inria.fr

A q-analogue of the Kostant partition function and twisted representations

Victor GUILLEMIN
Department of Mathematics, Massachusetts Institute of Technology

and

Etienne RASSART
School of Mathematics, Institute for Advanced Study

November 21, 2004

Abstract

We discuss a family of representations of Lie groups related to quantization with respect to the Dirac signature operator. The combinatorics of these twisted representations is similar to that of the usual irreducible representations, but involve a specialization of a q-analogue of the Kostant partition function. In particular, we prove signature analogues of the Kostant formula for weight multiplicities and the Steinberg formula for tensor product multiplicities. Using symmetric functions, we also find, for type A, analogues of the Weyl branching rule and the Gelfand-Tsetlin theorem.

Résumé

Nous étudions une famille de représentations de groupes de Lie liée à la quantisation par rapport à l'opérateur de signature de Dirac. Ces représentations obéissent à des règles combinatoires semblables à celles qui régissent le cas classique des représentations irréductibles, mais font appel à une spécialisation d'un q-analogue de la fonction de partition de Kostant. Nous donnons des analogues des formules de Kostant, pour les multiplicités de poids, et de Steinberg, pour les multiplicités de facteurs dans les produits tensoriels. À l'aide de fonctions symétriques, nous trouvons aussi en type A des analogues de la règle de bifurcation de Weyl et de la théorie de Gelfand-Tsetlin.

Introduction

The results described in this note are closely related to an article of Guillemin, Sternberg and Weitsman [1] on signature quantization.

A symplectic manifold (M, ω) is *pre-quantizable* if the cohomology class of ω is an integral class, i.e. is in the image of the map $H^2(M, \mathbb{Z}) \to H^2(M, \mathbb{R})$. This assumption implies the existence of a *pre-quantum structure* on M: a line bundle, $\mathbb{L}$, and a connection, ∇, such that $\operatorname{curv}(\nabla) = \omega$. If g is a Riemannian metric compatible with ω, then, from g and ω, one gets an elliptic operation $\not{\partial}_{\mathbb{C}} : S^+ \to S^-$, the *spin-$\mathbb{C}$ Dirac operator*, and, by twisting this operator with $\mathbb{L}$, an operator $\not{\partial}_{\mathbb{C}}^{\mathbb{L}} : S^+ \otimes \mathbb{L} \to S^- \otimes \mathbb{L}$. If M is compact one can "quantize" it by associating with it the virtual vector space

$$Q(M) = \operatorname{Index} \not{\partial}_{\mathbb{C}}^{\mathbb{L}} . \tag{1}$$

Moreover if G is a compact Lie group and τ a Hamiltonian action of G on M one gets from τ a representation of G on $Q(M)$ which is well-defined up to isomorphism (independent of the choice of g).

The results described in this note are closely related to two theorems in the article [1]. In this article the authors study the signature analogue of spin-$\mathbb{C}$ quantization: i.e. they define the virtual vector space (1) by replacing $\not{\partial}_{\mathbb{C}}$ by the signature operator $\not{\partial}_{\mathrm{sig}}$, and prove signature versions of a number of standard theorems about quantized symplectic manifolds. The two theorems we'll be concerned with in this paper are the following.

1. Let $G = \left(S^1\right)^n$ and let M be a $2n$-dimensional toric variety with moment polytope $\Delta \subseteq \mathbb{R}^n$. Then, for spin-$\mathbb{C}$ quantization, the weights of the representation of G on $Q(M)$ are the lattice points, $\beta \in \Delta \cap \mathbb{Z}^n$, and each weight occurs with multiplicity 1. For signature quantization the weights are the same; however, the weight β occurs with multiplicity 2^n if β lies in $\operatorname{Int}(\Delta)$, with multiplicity 2^{n-1} if it lies on a facet, and, in general, with multiplicity 2^{n-i} if it lies on i facets. Further details can be found in the work of Agapito [2].

2. Let G be a compact simply connected Lie group, λ a dominant weight and $O_\lambda = M$ the coadjoint orbit of G through λ. In the spin-$\mathbb{C}$ theory, the representation of G on $Q(M)$ is the unique irreducible representation V_λ of G with highest weight λ; however, in the signature theory, it is the representation

$$\widetilde{V}_\lambda = V_{\lambda-\rho} \otimes V_\rho\,, \tag{2}$$

where ρ is half the sum of the positive roots. (This is modulo the proviso that $\lambda - \rho$ be dominant.)

The article [1] also contains a signature version of the Kostant multiplicity formula. We recall that the Kostant multiplicity formula computes the multiplicity with which a weight, μ, of T occurs in V_λ by the formula

$$\sum_{\sigma\in\mathcal{W}} (-1)^{|\sigma|}\, K(\sigma(\lambda+\rho) - (\mu+\rho)) \tag{3}$$

where $\mathcal{W}$ is the Weyl group, $|\sigma|$ is the length of σ in $\mathcal{W}$, and K, the *Kostant partition function* (described below in Definition 1). The signature version of the Kostant multiplicity formula computes the multiplicity $\widetilde{m}_\lambda(\mu)$ with which the weight μ appears in $\widetilde{V}_\lambda$ by a similar formula:

$$\widetilde{m}_\lambda(\mu) = \sum_{\sigma\in\mathcal{W}} (-1)^{|\sigma|}\, K_2(\sigma(\lambda) - \mu) \tag{4}$$

where K_2 is the $q = 2$ specialization of a new q-analogue of the Kostant partition function, described below.

Our initial goal in writing this paper was to give a purely algebraic derivation of this result; however we noticed that there are $\widetilde{V}_\lambda$ analogues of a number of other basic formulas in the representation theory of compact semisimple Lie groups, in particular, an analogue of the Steinberg formula and, for $\mathrm{GL}_k\mathbb{C}$, analogues of the Weyl branching rule and the Gelfand-Tsetlin theorem. Some of the proofs are sketched but details can be found in [3].

The Kostant partition function and its q-analogues

We start by introducing the Kostant partition function.

Definition 1 *The* Kostant partition function *for a root system Φ, given a choice of positive roots Φ_+, is the function*

$$K(\mu) = \left|\left\{(k_\alpha)_{\alpha\in\Phi_+} \in \mathbb{N}^{|\Phi_+|} \;:\; \sum_{\alpha\in\Phi_+} k_\alpha \alpha = \mu\right\}\right|, \tag{5}$$

i.e. $K(\mu)$ is the number of ways that μ can be written as a sum of positive roots (see [4]).

Note that $K(\mu)$ can also be computed as the number of integer points inside the polytope

$$Q_\mu = \left\{(k_\alpha)_{\alpha\in\Phi_+} \in \mathbb{R}_{\geq 0}^{|\Phi_+|} \;:\; \sum_{\alpha\in\Phi_+} k_\alpha \alpha = \mu\right\}. \tag{6}$$

We can write down a generating function for the $K(\mu)$ that is very similar to Euler's generating function for the number of partitions (see [4, Section 25.2]):

$$\sum_\mu K(\mu) e^\mu = \prod_{\alpha\in\Phi_+} \frac{1}{1-e^\alpha}\,. \tag{7}$$

The classical q-analogue $\widehat{K}_q(\mu)$ of $K(\mu)$, due to Lusztig [5], keeps track of how many times the roots appear:

$$\widehat{K}_q(\mu) = \sum_{(k_\alpha)_\alpha \in Q_\mu} q^{\sum k_\alpha}\,, \tag{8}$$

corresponding to the generating function

$$\sum_\mu \widehat{K}_q(\mu) e^\mu = \prod_{\alpha\in\Phi_+} \left(\sum_{m\geq 0} q^m e^{m\alpha}\right) = \prod_{\alpha\in\Phi_+} \frac{1}{1-qe^\alpha}\,. \tag{9}$$

The q-analogue $K_q(\mu)$ that interests us here is the one that counts the integer points of Q_μ according to how many of the k_α's are nonzero:

$$K_q(\mu) = \sum_{(k_\alpha)_\alpha \in Q_\mu} q^{|\{k_\alpha > 0\}|} . \tag{10}$$

In terms of generating functions, this translates to

$$\sum_\mu K_q(\mu) e^\mu = \prod_{\alpha\in\Phi_+} \left(1 + q \sum_{m\geq 1} e^{\alpha}\right) = \prod_{\alpha\in\Phi_+} \frac{1+(q-1)e^\alpha}{1-e^\alpha} . \tag{11}$$

The representations $\widetilde{V}_\lambda = V_{\lambda-\rho}\otimes V_\rho$

We are working in the context of a complex semisimple Lie algebra $\mathfrak{g}$ with root system Φ, choice of positive roots Φ_+ , and Weyl group $\mathcal{W}$; ρ is half the sum of the positive roots (or the sum of the fundamental weights). For a dominant weight λ, we denote by V_λ the irreducible representation of $\mathfrak{g}$ with highest weight λ. We will call a weight λ *strictly dominant* if $\lambda-\rho$ is dominant. We will use the notation Λ^+ for the set of dominant weights, and Λ^+_S for the set of strictly dominant weights. For a strictly dominant weight, we define the representation

$$\widetilde{V}_\lambda = V_{\lambda-\rho}\otimes V_\rho \tag{12}$$

and its character

$$\widetilde{\chi}_\lambda = \chi_{V_{\lambda-\rho}\otimes V_\rho} = \chi_{\lambda-\rho}\cdot\chi_\rho . \tag{13}$$

The following theorem of Guillemin, Sternberg, and Weitsman [1] provides a formula for the multiplicities of the weights in the weight space decomposition of $\widetilde{V}_\lambda$. This formula is very similar to the Kostant multiplicity formula (3), but uses the $q=2$ specialization of the q-analogue of the Kostant partition function $K_q(\mu)$ introduced above, instead of the usual Kostant partition function. The formula for the $\widetilde{V}_\lambda$ multiplicities further distinguishes itself from the Kostant formula by being free of the ρ factors.

An analogue of the Kostant multiplicity formula for the $\widetilde{V}_\lambda$

Theorem 2 (Guillemin-Sternberg-Weitsman [1]) *Let λ be a strictly dominant weight. Then the multiplicity of the weight ν in the tensor product $\widetilde{V}_\lambda = V_{\lambda-\rho}\otimes V_\rho$ is given by*

$$\widetilde{m}_\lambda(\nu) = \dim\left(\widetilde{V}_\lambda\right)_\nu = \sum_{\omega\in\mathcal{W}} (-1)^{|\omega|} K_2(\omega(\lambda)-\nu) , \tag{14}$$

where $|\omega|$ is the length of ω in the Weyl group.

Proof. We give a simple proof here using the Weyl character formula. This formula expresses the character χ_λ of V_λ as the quotient

$$\chi_\lambda = \frac{A_{\lambda+\rho}}{A_\rho} , \tag{15}$$

where $A_\mu = \sum_{\omega\in\mathcal{W}} (-1)^{|\omega|} e^{\omega(\mu)}$. For ρ, we get the nice expression [4, Lemma 24.3]

$$A_\rho = \prod_{\alpha\in\Phi_+} \left(e^{\alpha/2} - e^{-\alpha/2}\right) = e^\rho \prod_{\alpha\in\Phi_+} \left(1-e^{-\alpha}\right) , \tag{16}$$

which means, in particular, that we get

$$\chi_\rho = \frac{A_{2\rho}}{A_\rho} = \frac{e^{2\rho}\prod_{\alpha\in\Phi_+}\left(1-e^{-2\alpha}\right)}{e^{\rho}\prod_{\alpha\in\Phi_+}\left(1-e^{-\alpha}\right)} = e^\rho \prod_{\alpha\in\Phi_+}\left(1+e^{-\alpha}\right) . \tag{17}$$

Thus, for λ strictly dominant,

$$\widetilde{\chi}_\lambda = \chi_{\lambda-\rho}\cdot\chi_\rho = \sum_{\omega\in\mathcal{W}}(-1)^{|\omega|}\,e^{\omega(\lambda)}\prod_{\alpha\in\Phi_+}\frac{1+e^{-\alpha}}{1-e^{-\alpha}} \tag{18}$$

$$= \sum_{\omega\in\mathcal{W}}(-1)^{|\omega|}\,e^{\omega(\lambda)}\sum_{\mu}K_2(\mu)\,e^{-\mu}$$

$$= \sum_{\mu}\sum_{\omega\in\mathcal{W}}(-1)^{|\omega|}\,K_2(\mu)\,e^{\omega(\lambda)-\mu}\,. \tag{19}$$

Extracting the coefficient of e^ν on both sides gives (14). ■

The next step will be to use a formula due to Atiyah and Bott for the characters of the V_λ and $\widetilde{V}_\lambda$ to break down $\widetilde{V}_\lambda$ into its irreducible components and find their multiplicities. The Atiyah-Bott formula [6, 7] gives the character of V_μ as

$$\chi_\mu = \sum_{\omega\in\mathcal{W}}e^{\omega(\mu)}\prod_{\alpha\in\Phi_+}\frac{1}{1-e^{-\omega(\alpha)}}\,. \tag{20}$$

Remark 3 *We can deduce this formula from the Weyl character formula (equation* (15)*) by first observing that*

$$\prod_{\alpha\in\Phi_+}\left(1-e^{-\omega(\alpha)}\right) = (-1)^{|\omega|}e^{\sum\{\alpha\in\Phi_+\,:\,\omega(\alpha)\in\Phi_-\}}\prod_{\alpha\in\Phi_+}\left(1-e^{-\alpha}\right) \tag{21}$$

Also,

$$\rho-\omega(\rho) = \sum\{\alpha\in\Phi_+\;:\;\omega(\alpha)\in\Phi_-\}\,. \tag{22}$$

Combining (21) *with* (22) *gives*

$$\prod_{\alpha\in\Phi_+}\left(1-e^{-\omega(\alpha)}\right) = (-1)^{|\omega|}e^{\rho-\omega(\rho)}\prod_{\alpha\in\Phi_+}\left(1-e^{-\alpha}\right)\,, \tag{23}$$

and we can translate Weyl's character formula into the Atiyah-Bott formula using this equation.

For any $\omega\in\mathcal{W}$,

$$\chi_\rho = e^\rho\prod_{\alpha\in\Phi_+}\left(1+e^{-\alpha}\right)$$

$$= e^{\omega(\rho)}\prod_{\alpha\in\Phi_+}\left(1+e^{-\omega(\alpha)}\right)\,, \tag{24}$$

since characters are invariant under the Weyl group action. Using this and the Atiyah-Bott formula, we can write[1]

$$\widetilde{\chi}_\lambda = \chi_{\lambda-\rho}\cdot\chi_\rho = \sum_{\omega\in\mathcal{W}}e^{\omega(\lambda)}\prod_{\alpha\in\Phi_+}\frac{1+e^{-\omega(\alpha)}}{1-e^{-\omega(\alpha)}} \tag{25}$$

$$= \sum_{\omega\in\mathcal{W}}e^{\omega(\lambda)}\prod_{\alpha\in\Phi_+}\frac{1}{1-e^{-\omega(\alpha)}}\sum_{I\subseteq\Phi_+}e^{-\omega(\alpha_I)}$$

[1] Alternatively, we can obtain equation (25) from equation (18) by observing that for $\omega\in\mathcal{W}$,

$$\omega\cdot\left(\prod_{\alpha\in\Phi_+}\frac{1+e^{-\alpha}}{1-e^{-\alpha}}\right) = \prod_{\alpha\in\Phi_+}\frac{1+e^{-\omega(\alpha)}}{1-e^{-\omega(\alpha)}} = (-1)^{|\omega|}\prod_{\alpha\in\Phi_+}\frac{1+e^{-\alpha}}{1-e^{-\alpha}}\,.$$

where as before, $\alpha_I = \sum_{\alpha \in I} \alpha$. This gives

$$\widetilde{\chi}_\lambda = \sum_{I \subseteq \Phi_+} \left(\sum_{\omega \in \mathcal{W}} e^{\omega(\lambda - \alpha_I)} \prod_{\alpha \in \Phi_+} \frac{1}{1 - e^{-\omega(\alpha)}} \right) . \tag{26}$$

Letting, $\lambda_I = \lambda - \alpha_I$, we observe that if λ_I is dominant, the Atiyah-Bott formula tells us that

$$\sum_{\omega \in \mathcal{W}} e^{\omega(\lambda - \alpha_I)} \prod_{\alpha \in \Phi_+} \frac{1}{1 - e^{-\omega(\alpha)}} \tag{27}$$

is the character χ_{λ_I} of the irreducible representation V_{λ_I}, so that

$$\widetilde{\chi}_\lambda = \sum_{I \subseteq \Phi_+} \chi_{\lambda_I} \qquad \text{and} \qquad \widetilde{V}_\lambda = V_{\lambda - \rho} \otimes V_\rho = \bigoplus_{I \subseteq \Phi_+} V_{\lambda_I} \tag{28}$$

if all the λ_I are dominant.

Finally, since α_I and $\alpha_{I'}$ can be equal for different subsets I and I', certain highest weights appear multiple times in the above sums. For the weight $\mu = \lambda_I = \lambda - \alpha_I$, we will get V_μ as many times as we can write $\alpha_I = \lambda - \mu$ as a sum of positive roots, where each positive root appears at most once. Hence

$$\widetilde{V}_\lambda = \sum_\mu P(\lambda - \mu)\, V_\mu \,, \tag{29}$$

where the sum is over all μ such that $\mu = \lambda_I$ for some I, and $P(\nu)$ is given by

$$\sum_\nu P(\nu) e^\nu = \prod_{\alpha \in \Phi_+} (1 + e^\alpha) \ . \tag{30}$$

Remark 4 *David Vogan pointed out to us that this decomposition is well-known and can be deduced from the Steinberg formula. For type A_n, the number of distinct μ's in the above sum is the number of forests of labelled unrooted tree on $n+1$ vertices [8, 9].*

A tensor product formula for the $\widetilde{V}_\lambda$

We will derive here an analogue of the Steinberg formula for the $\widetilde{V}_\lambda$. Given two representations $\widetilde{V}_\lambda$ and $\widetilde{V}_\mu$, the problem is to determine whether their tensor product $\widetilde{V}_\lambda \otimes \widetilde{V}_\mu$ can be decomposed in terms of $\widetilde{V}_\nu$'s. This is readily seen to be the case, as

$$\widetilde{V}_\lambda \otimes \widetilde{V}_\mu \quad = \quad (V_{\lambda-\rho} \otimes V_\rho) \otimes (V_{\mu-\rho} \otimes V_\rho) \quad = \quad (V_{\lambda-\rho} \otimes V_\rho \otimes V_{\mu-\rho}) \otimes V_\rho \,. \tag{31}$$

Breaking up $V_{\lambda-\rho} \otimes V_\rho \otimes V_{\mu-\rho}$ into irreducibles V_γ and tensoring each factor with V_ρ yields factors $V_\gamma \otimes V_\rho = \widetilde{V}_{\gamma+\rho}$. Thus for strictly dominant weights λ and μ, we can write

$$\widetilde{V}_\lambda \otimes \widetilde{V}_\mu = \sum_{\nu \in \Lambda_S^+} \widetilde{N}_{\lambda\mu}^\nu \widetilde{V}_\nu \tag{32}$$

for some nonnegative integers $\widetilde{N}_{\lambda\mu}^\nu$.

Theorem 5 *For λ, μ and ν strictly dominant weights, the tensor product multiplicity $\widetilde{N}_{\lambda\mu}^\nu$ of $\widetilde{V}_\nu$ in $\widetilde{V}_\lambda \otimes \widetilde{V}_\mu$ is given by*

$$\widetilde{N}_{\lambda\mu}^\nu = \sum_{\omega \in \mathcal{W}} \sum_{\sigma \in \mathcal{W}} (-1)^{|\omega\sigma|}\, K_2(\omega(\lambda) + \sigma(\mu) - \nu) \,. \tag{33}$$

Proof. Starting from the equation $\widetilde{V}_\lambda \otimes \widetilde{V}_\mu = \sum_{\nu\in\Lambda_S^+} \widetilde{N}_{\lambda\mu}^\nu \widetilde{V}_\nu$, we can use equation (18) to write

$$\sum_{\omega\in\mathcal{W}} (-1)^{|\omega|} e^{\omega(\lambda)} \prod_{\alpha\in\Phi_+} \frac{1+e^{-\alpha}}{1-e^{-\alpha}} \cdot \widetilde{\chi}_\mu = \sum_{\nu\in\Lambda_S^+} \widetilde{N}_{\lambda\mu}^\nu \sum_{\tau\in\mathcal{W}} (-1)^{|\tau|} e^{\tau(\nu)} \prod_{\alpha\in\Phi_+} \frac{1+e^{-\alpha}}{1-e^{-\alpha}} \,.$$

Cancelling terms and using Theorem 2 to write down the character $\widetilde{\chi}_\mu$ yields

$$\begin{aligned} \sum_{\omega\in\mathcal{W}} (-1)^{|\omega|} e^{\omega(\lambda)} \cdot \sum_{\beta} \sum_{\sigma\in\mathcal{W}} (-1)^{|\sigma|} K_2(\sigma(\mu)-\beta)\, e^{\beta} &= \sum_{\nu\in\Lambda_S^+} \widetilde{N}_{\lambda\mu}^\nu \sum_{\tau\in\mathcal{W}} (-1)^{|\tau|} e^{\tau(\nu)} \\ \sum_{\beta} \sum_{\omega\in\mathcal{W}} \sum_{\sigma\in\mathcal{W}} (-1)^{|\omega|+|\sigma|}\, K_2(\sigma(\mu)-\beta)\, e^{\omega(\lambda)+\beta} &= \sum_{\nu\in\Lambda_S^+} \sum_{\tau\in\mathcal{W}} (-1)^{|\tau|}\, \widetilde{N}_{\lambda\mu}^\nu\, e^{\tau(\nu)} \end{aligned}$$

Substituting $\gamma = \omega(\lambda)+\beta$ on the left hand side, and $\gamma = \tau(\nu)$ on the right hand side gives

$$\sum_{\gamma} \sum_{\omega\in\mathcal{W}} \sum_{\sigma\in\mathcal{W}} (-1)^{|\omega\sigma|}\, K_2(\sigma(\mu)+\omega(\lambda)-\gamma)\, e^{\gamma} = \sum_{\substack{\gamma \text{ conjugate}\\ \text{to a strictly}\\ \text{dominant weight}}} \sum_{\tau\in\mathcal{W}} (-1)^{|\tau|}\, \widetilde{N}_{\lambda\mu}^{\tau^{-1}(\gamma)}\, e^{\gamma}\,,$$

and extracting the coefficient of e^γ on both sides yields

$$\sum_{\omega\in\mathcal{W}} \sum_{\sigma\in\mathcal{W}} (-1)^{|\omega\sigma|}\, K_2(\sigma(\mu)+\omega(\lambda)-\gamma) = \sum_{\tau\in\mathcal{W}} (-1)^{|\tau|}\, \widetilde{N}_{\lambda\mu}^{\tau^{-1}(\gamma)}\,. \tag{34}$$

Now, since $\widetilde{N}_{\lambda\mu}^{\tau^{-1}(\gamma)}$ vanishes unless $\tau^{-1}(\gamma)$ is strictly dominant, all the terms in the sum on the right hand side vanish except for the one where τ is the identity (i.e. the term where $\gamma=\nu$), and we get the result. ■

If we denote by $N_{\lambda\mu}^\nu$ the multiplicities of the irreducible representations V_ν in the tensor product $V_\lambda \otimes V_\mu$, defined by

$$V_\lambda \otimes V_\mu = \sum_{\nu\in\Lambda^+} N_{\lambda\mu}^\nu\, V_\nu\,, \tag{35}$$

then we can write down the tensor product multiplicities $\widetilde{N}_{\lambda\mu}^\nu$ for the decomposition of $\widetilde{V}_\lambda \otimes \widetilde{V}_\mu$ into $\widetilde{V}_\nu$'s in terms of the $N_{\lambda\mu}^\nu$ as follows:

$$\begin{aligned} \widetilde{V}_\lambda \otimes \widetilde{V}_\mu &= V_{\lambda-\rho} \otimes V_\rho \otimes V_{\mu-\rho} \otimes V_\rho \\ &= \left(\left(\sum_{\beta\in\Lambda^+} N_{\lambda-\rho,\rho}^\beta\, V_\beta \right) \otimes V_{\mu-\rho} \right) \otimes V_\rho \\ &= \left(\sum_{\beta\in\Lambda^+} \sum_{\gamma\in\Lambda^+} N_{\lambda-\rho,\rho}^\beta\, N_{\beta,\mu-\rho}^\gamma\, V_\gamma \right) \otimes V_\rho \\ &= \sum_{\beta\in\Lambda^+} \sum_{\gamma\in\Lambda^+} N_{\lambda-\rho,\rho}^\beta\, N_{\beta,\mu-\rho}^\gamma\, \widetilde{V}_{\gamma+\rho} \\ &= \sum_{\nu\in\Lambda_S^+} \sum_{\beta\in\Lambda^+} N_{\lambda-\rho,\rho}^\beta\, N_{\beta,\mu-\rho}^{\nu-\rho}\, \widetilde{V}_\nu\,, \end{aligned}$$

so that for strictly dominant ν,

$$\widetilde{N}_{\lambda\mu}^\nu = \sum_{\beta\in\Lambda^+} N_{\lambda-\rho,\rho}^\beta\, N_{\beta,\mu-\rho}^{\nu-\rho}\,. \tag{36}$$

Remark 6 *In type A, there is a combinatorial interpretation for the coefficients $N_{\lambda\mu}^\nu$ in terms of shifted Young tableaux: they are given by a shifted analogue of the Littlewood-Richardson rule (see [10]*

Links with symmetric functions in type A

As for the weight multiplicities and Clebsch-Gordan coefficients, there is a link between the character products $\widetilde{\chi}_\lambda = \chi_{\lambda-\rho} \cdot \chi_\rho$ and symmetric functions in type A, again in terms of Schur functions.

The character of the irreducible polynomial representation V_λ of $\mathrm{GL}_k\mathbb{C}$, where we now think of λ as a partition with k parts (allowing the empty part) is the Schur function $s_\lambda(x_1, \ldots, x_k)$. We will call a partition *strict* if all its parts are distinct (corresponding to a strictly dominant weight). Thus we have that, for $\mathrm{GL}_k\mathbb{C}$,

$$\widetilde{\chi}_\lambda \quad = \quad \chi_{\lambda-\rho} \cdot \chi_\rho \quad = \quad s_{\lambda-\rho}(x_1, \ldots, x_k)\, s_\rho(x_1, \ldots, x_k)\,, \tag{37}$$

for any strict partition λ. It is readily checked that the weight ρ corresponds to the partition $(k-1, k-2, \ldots, 1, 0)$.

Remark 7 *We can also write the characters of $\widetilde{V}_\lambda$ in terms of Hall-Littlewood polynomials (see [11, III, 1. and 2.]). The results of the following sections can be deduced from this link with Hall-Littlewood polynomials, but we will rather use the Schur function expression* (37) *for the characters. This makes the proofs a bit more technical but avoids the heavier machinery of Hall-Littlewood polynomials.*

A branching rule for the $\widetilde{V}_\lambda$ in type A

We have seen that the representations $\widetilde{V}_\lambda$ behave somewhat like irreducible representations, in that tensor products of them can be broken down into direct sums of $\widetilde{V}_\nu$'s again, and that the multiplicities in those decompositions as well as in the weight space decomposition are given by formulas very similar to those of Kostant and Steinberg in the irreducible case. The Weyl branching rule (see [4] for example) describes how to restrict a representation V_λ from $\mathrm{GL}_k\mathbb{C}$ to $\mathrm{GL}_{k-1}\mathbb{C}$. This rule can be applied iteratively and provides a way to index one-dimensional subspaces of V_λ by diagrams (Gelfand-Tsetlin diagrams [12]) that is compatible with the weight space decomposition. It is natural to ask whether the representations $\widetilde{V}_\lambda$ of $\mathrm{GL}_k\mathbb{C}$ are also well-behaved under restriction, or in another words, if there is an analogue of the Weyl branching rule for the $\widetilde{V}_\lambda$ in type A.

For two partitions $\mu = (\mu_1, \ldots, \mu_m)$ and $\gamma = (\gamma_1, \ldots, \gamma_{m-1})$, we say that γ interlaces μ, and write $\gamma \lhd \mu$, if

$$\mu_1 \geq \gamma_1 \geq \mu_2 \geq \gamma_2 \geq \mu_3 \geq \cdots \geq \mu_{m-1} \geq \gamma_{m-1} \geq \mu_m\,.$$

For two such partitions μ and γ such that $\gamma \lhd \mu$, we define

$$\nabla(\mu, \gamma) = \Big|\big\{ i \in \{1, 2, \ldots, m-1\} \ :\ \mu_i > \gamma_i > \mu_{i+1} \big\}\Big|\,. \tag{38}$$

In other words, $\nabla(\mu, \gamma)$ is the number of γ_i that are wedged strictly between μ_i and μ_{i+1}.

Theorem 8 *The decomposition of the restriction of the representation $\widetilde{V}_\lambda$ of $\mathrm{GL}_k\mathbb{C}$ to $\mathrm{GL}_{k-1}\mathbb{C}$ into irreducible representations of $\mathrm{GL}_{k-1}\mathbb{C}$ is given by*

$$\mathrm{Res}^{\mathrm{GL}_k\mathbb{C}}_{\mathrm{GL}_{k-1}\mathbb{C}}\, \widetilde{V}_\lambda = \bigoplus_{\nu \in \Lambda^+_S \,:\, \nu \lhd \lambda} 2^{\nabla(\lambda,\nu)}\, \widetilde{V}_\nu\,. \tag{39}$$

Proof. We give here a sketch of the proof. We argue using characters and the fact that those can be written in terms of Schur functions. We saw above (equation (37)) that the character of the representation $\widetilde{V}_\lambda$ of $\mathrm{GL}_k\mathbb{C}$ is the product of Schur functions $s_{\lambda-\rho}(x_1, \ldots, x_k)\, s_\rho(x_1, \ldots, x_k)$. We obtain the character of the restriction of $\widetilde{V}_\lambda$ to $\mathrm{GL}_{k-1}\mathbb{C}$ by setting the last variable x_k equal to 1. Using well-known identities on Schur functions (see [13, Section 7.15] for example), we have that

$$s_\lambda(x_1, \ldots, x_{k-1}, 1) = \sum_{\mu \lhd \lambda} s_\mu(x_1, \ldots, x_{k-1})\,. \tag{40}$$

and

$$s_\rho(x_1, \ldots, x_k) = \prod_{1 \leq i < j \leq k} (x_i + x_j)\,. \tag{41}$$

Thus,

$$s_{\lambda-\rho}(x_1, \ldots, x_{k-1}, 1) s_\rho(x_1, \ldots, x_{k-1}, 1) = \sum_{\mu \lhd \lambda - \rho} s_\mu(x_1, \ldots, x_{k-1}) \prod_{1 \leq i < j \leq k-1} (x_i + x_j) \prod_{i=1}^{k-1} (x_i + 1)\,. \tag{42}$$

We recognize the product $\prod_{1\leq i<j\leq k-1}(x_i+x_j)$ as the Schur function $s_\rho(x_1,\ldots,x_{k-1})$ (where ρ now corresponds to the partition $(k-2,k-3,\ldots,1,0)$ with $k-1$ parts), and the product $\prod_{i=1}^{k-1}(x_i+1)$ as the sum $(e_0+e_1+\cdots+e_{k-1})$ of elementary symmetric functions in the variables $x_1,\ldots,x_{k-1}$. A dual version of the Pieri rule [13, Section 7.15] describes how to break down the product of a Schur function with an elementary symmetric function into Schur functions:

$$s_\mu\, e_m = \sum_\nu s_\nu\,, \tag{43}$$

where the sum is over all ν obtained from μ by adding a vertical strip of size m, i.e. over the ν such that $\mu\subseteq\nu$ and the skew-shape ν/μ consists of m boxes, no two of which are in the same row. As we are working in $k-1$ variables, the s_ν with more than $k-1$ parts vanish, so we can add the further constraint that the vertical strip be confined to the first $k-1$ rows (we will say such a vertical strip has height at most $k-1$). This gives

$$\begin{aligned} s_{\lambda-\rho}(x_1,\ldots,x_{k-1},1)s_\rho(x_1,\ldots,x_{k-1},1) &= \sum_{\mu\,\lhd\,\lambda-\rho}\sum_\nu s_\nu(x_1,\ldots,x_{k-1})\,s_\rho(x_1,\ldots,x_{k-1}) \\ \widetilde{\chi}_\lambda(x_1,\ldots,x_{k-1},1) &= \sum_{\mu\,\lhd\,\lambda-\rho}\sum_\nu \widetilde{\chi}_{\nu+\rho}(x_1,\ldots,x_{k-1}) \end{aligned} \tag{44}$$

where the sum is over all the ν that can be obtained from μ by adding a vertical strip of size and height at most $k-1$. We can rewrite this as

$$\widetilde{\chi}_\lambda(x_1,\ldots,x_{k-1},1) = \sum_{\mu\,\lhd\,\lambda-\rho}\sum_\nu \widetilde{\chi}_\nu(x_1,\ldots,x_{k-1}) \tag{45}$$

where the sum is over all strict partitions ν such that $\nu-\rho$ can be obtained from μ by adding a vertical strip of size and height at most $k-1$. Since the $s_\nu s_\rho$ are linearly independent, we can lift this to the level of representations to get

$$\mathrm{Res}^{\mathrm{GL}_k\mathbb{C}}_{\mathrm{GL}_{k-1}\mathbb{C}}\,\widetilde{V}_\lambda = \bigoplus_{\mu\,\lhd\,\lambda-\rho}\bigoplus_\nu \widetilde{V}_\nu\,, \tag{46}$$

with the sum over the same set of ν as before.

In order to compute the multiplicity of a given $\widetilde{V}_\nu$ in $\mathrm{Res}^{\mathrm{GL}_k\mathbb{C}}_{\mathrm{GL}_{k-1}\mathbb{C}}\,\widetilde{V}_\lambda$, we define, for strict partitions λ and ν, $n(\lambda,\nu)$ to be the number of ways that $\nu-\rho$ can be obtained by adding a vertical strip of size and height at most $k-1$ to some partition μ such that $\mu\ \lhd\ \lambda-\rho$, so that

$$\widetilde{V}_\lambda = \bigoplus_{\nu\in\Lambda_S^+} n(\lambda,\nu)\,\widetilde{V}_\nu\,. \tag{47}$$

It can be checked that

$$n(\lambda,\nu) = \begin{cases} 2^{\nabla(\lambda,\nu)} & \text{if } \nu\ \lhd\ \lambda \text{ and } \nu\in\Lambda_S^+\,,\\ 0 & \text{otherwise.}\end{cases} \tag{48}$$

■

Gelfand-Tsetlin theory for the $\widetilde{V}_\lambda$

After restricting to $\mathrm{GL}_{k-1}\mathbb{C}$, we can further restrict to $\mathrm{GL}_{k-2}\mathbb{C}$. From now on, we will assume that all partitions are strict. We can write

$$\begin{aligned} \mathrm{Res}^{\mathrm{GL}_k\mathbb{C}}_{\mathrm{GL}_{k-2}\mathbb{C}}\,\widetilde{V}_\lambda &= \mathrm{Res}^{\mathrm{GL}_{k-1}\mathbb{C}}_{\mathrm{GL}_{k-2}\mathbb{C}}\left(\mathrm{Res}^{\mathrm{GL}_k\mathbb{C}}_{\mathrm{GL}_{k-1}\mathbb{C}}\,\widetilde{V}_\lambda\right) \\ &= \mathrm{Res}^{\mathrm{GL}_{k-1}\mathbb{C}}_{\mathrm{GL}_{k-2}\mathbb{C}}\left(\bigoplus_{\nu\,\lhd\,\lambda} 2^{\nabla(\lambda,\nu)}\,\widetilde{V}_\nu\right) \\ &= \bigoplus_{\nu\,\lhd\,\lambda} 2^{\nabla(\lambda,\nu)}\,\mathrm{Res}^{\mathrm{GL}_{k-1}\mathbb{C}}_{\mathrm{GL}_{k-2}\mathbb{C}}\,\widetilde{V}_\nu && (49)\\ &= \bigoplus_{\nu\,\lhd\,\lambda} 2^{\nabla(\lambda,\nu)}\left(\bigoplus_{\mu\,\lhd\,\nu} 2^{\nabla(\nu,\mu)}\,\widetilde{V}_\mu\right) && (50)\\ &= \bigoplus_{\mu\,\lhd\,\nu\,\lhd\,\lambda} 2^{\nabla(\lambda,\nu)+\nabla(\nu,\mu)}\,\widetilde{V}_\mu\,. && (51)\end{aligned}$$

Denoting by $\lambda^{(m)} = \lambda_1^{(m)} \geq \cdots \geq \lambda_m^{(m)} \geq 0$ the strict partitions indexing the representations $\widetilde{V}$ of $\mathrm{GL}_m\mathbb{C}$, we can iterate the branching rule until we get to $\mathrm{GL}_1\mathbb{C}$:

$$\mathrm{Res}_{\mathrm{GL}_1\mathbb{C}}^{\mathrm{GL}_k\mathbb{C}} \widetilde{V}_\lambda = \bigoplus_{\lambda^{(1)} \lhd \cdots \lhd \lambda^{(k)} = \lambda} 2^{\nabla(\lambda^{(k)},\lambda^{(k-1)})+\nabla(\lambda^{(k-1)},\lambda^{(k-2)})+\cdots+\nabla(\lambda^{(2)},\lambda^{(1)})} V_{\lambda^{(1)}} \,. \tag{52}$$

We will call a sequence of strict partitions of the form $\lambda^{(1)} \lhd \cdots \lhd \lambda^{(k)} = \lambda$ a *twisted Gelfand-Tsetlin diagram* for λ, which can be viewed schematically as

$$\begin{array}{ccccccccc}
\lambda_1^{(k)} & & \lambda_2^{(k)} & & \cdots & & \lambda_{k-1}^{(k)} & & \lambda_k^{(k)} \\
& \lambda_1^{(k-1)} & & \lambda_2^{(k-1)} & & \cdots & & \lambda_{k-1}^{(k-1)} & \\
& & \ddots & & \vdots & & \iddots & & \\
& & & \lambda_1^{(2)} & & \lambda_2^{(2)} & & & \\
& & & & \lambda_1^{(1)} & & & &
\end{array} \tag{53}$$

with $\lambda_j^{(k)} = \lambda_j$ and each $\lambda_j^{(i)}$ is a nonnegative integer satisfying

$$\lambda_j^{(i)} > \lambda_{j+1}^{(i)} \tag{54}$$

and

$$\lambda_j^{(i+1)} \;\geq\; \lambda_j^{(i)} \;\geq\; \lambda_{j+1}^{(i+1)} \tag{55}$$

for all $1 \leq j \leq i$, $1 \leq i \leq k-1$.

Let $\widetilde{V}_{\mathcal{D}}$ be the subspace of $\widetilde{V}_\lambda$ corresponding to a twisted Gelfand-Tsetlin diagram $\mathcal{D}$. This subspace has dimension $2^{\nabla(\mathcal{D})}$, where

$$\nabla(\mathcal{D}) = \nabla(\lambda^{(k)},\lambda^{(k-1)}) + \nabla(\lambda^{(k-1)},\lambda^{(k-2)}) + \cdots + \nabla(\lambda^{(2)},\lambda^{(1)}) \,. \tag{56}$$

We can also think of $\nabla(\mathcal{D})$ as the number of triangles

$$\begin{array}{ccc}
\lambda_j^{(i)} & & \lambda_{j+1}^{(i)} \\
& \lambda_j^{(i+1)} &
\end{array}$$

with strict inequalities $\lambda_j^{(i+1)} > \lambda_j^{(i)} > \lambda_{j+1}^{(i+1)}$ in the diagram $\mathcal{D}$.

We show here that $\widetilde{V}_{\mathcal{D}}$ lies completely within the same weight space of the weight space decomposition of $\widetilde{V}_\lambda$.

We will think of the groups $\mathrm{GL}_k\mathbb{C}$ as included into one another by identifying $\mathrm{GL}_m\mathbb{C}$ with

$$\left(\begin{array}{c|c}
\mathrm{GL}_m\mathbb{C} & \mathbf{0} \\
\hline
\mathbf{0} & id_{k-m}
\end{array}\right)$$

Consider the element $I \in \mathfrak{gl}_m\mathbb{C}$ and a representation $\widetilde{V}_\mu$ of $\mathrm{GL}_m\mathbb{C}$. We have the representation $\mathrm{GL}_k\mathbb{C} \to \mathfrak{gl}(V_\mu \otimes V_\rho)$. For $v \in V_{\mu-\rho}$ and $w \in V_\rho$, we have

$$\begin{aligned}
I \cdot (v \otimes w) &= (I \cdot v) \otimes w + v \otimes (I \cdot w) \\
&= \left(\left(\sum_{j=1}^m (\mu-\rho)_j\right) v\right) \otimes w + v \otimes \left(\left(\sum_{j=1}^m \rho_j\right) w\right) \\
&= \left(\sum_{j=1}^m \left((\mu-\rho)_j + \rho_j\right)\right) v \otimes w \\
&= \left(\sum_{j=1}^m \mu_j\right) v \otimes w \,,
\end{aligned}$$

since $V_{\mu-\rho}$ has highest weight $\mu-\rho$ and V_ρ has highest weight ρ. So $I \in \mathfrak{gl}_m\mathbb{C}$ gets represented as $(\sum_{j=1}^m \mu_j)\, I$ in $\widetilde{V}_\mu$.

In general, for

$$\mathrm{Res}^{\mathrm{GL}_k\mathbb{C}}_{\mathrm{GL}_m\mathbb{C}}\, \widetilde{V}_\lambda = \bigoplus_{\lambda^{(m)} \lhd \cdots \lhd \lambda^{(k)}=\lambda} 2^{\nabla(\lambda^{(k)},\lambda^{(k-1)})+\nabla(\lambda^{(k-1)},\lambda^{(k-2)})+\cdots+\nabla(\lambda^{(m+1)},\lambda^{(m)})}\, \widetilde{V}_{\lambda^{(m)}}\,,$$

we will find that $I \in \mathfrak{gl}_m\mathbb{C}$ gets represented as $(\sum_{i=1}^m \lambda_i^{(m)})\, I$ in $\widetilde{V}_{\lambda^{(m)}}$.

Therefore, in the basis $I_1,\,\ldots\,, I_k$, the subspace $\widetilde{V}_{\mathcal{D}}$ corresponding to a twisted Gelfand-Tsetlin diagram $\mathcal{D}$ has weight

$$\Big(\sum_{i=1}^{1}\lambda_i^{(1)}\,,\,\sum_{i=1}^{2}\lambda_i^{(2)}\,,\,\ldots\,,\,\sum_{i=1}^{k}\lambda_i^{(k)}\Big)$$

or

$$\Big(\sum_{i=1}^{1}\lambda_i^{(1)}\,,\,\sum_{i=1}^{2}\lambda_i^{(2)}-\sum_{i=1}^{1}\lambda_i^{(1)}\,,\,\ldots\,,\,\sum_{i=1}^{k}\lambda_i^{(k)}-\sum_{i=1}^{k-1}\lambda_i^{(k-1)}\Big)$$

in the usual basis $J_1,\ldots,J_k$.

In other words, $\widetilde{V}_{\mathcal{D}} \subseteq \big(\widetilde{V}_\lambda\big)_\beta$ if

$$\beta_m = \sum_{i=1}^{m}\lambda_i^{(m)} - \sum_{i=1}^{m-1}\lambda_i^{(m-1)}\,, \tag{57}$$

or, equivalently,

$$\beta_1+\cdots+\beta_m = \sum_{i=1}^{m}\lambda_i^{(m)}\,. \tag{58}$$

Hence twisted Gelfand-Tsetlin diagrams for λ correspond to the same weight if all their row sums are the same. So we have proved the following analogue of the Gelfand-Tsetlin theorem [12].

Theorem 9 *Let $\lambda=(\lambda_1,\ldots,\lambda_k)$ be a strictly dominant weight. The dimension of the representation $\widetilde{V}_\lambda$ of $\mathrm{GL}_k\mathbb{C}$ is given by*

$$\dim\,\widetilde{V}_\lambda = \sum_{\mathcal{D}} 2^{\nabla(\mathcal{D})} \tag{59}$$

where the sum is over all twisted Gelfand-Tsetlin diagrams with top row λ.

Furthermore, the multiplicity $\widetilde{m}_\lambda(\beta)$ of the weight β in $\widetilde{V}_\lambda$ is given by

$$\widetilde{m}_\lambda(\beta) = \dim\,\big(\widetilde{V}_\lambda\big)_\beta = \sum_{\mathcal{D}} 2^{\nabla(\mathcal{D})} \tag{60}$$

where the sum is over all twisted Gelfand-Tsetlin diagrams with top row λ and row sums satisfying equation (57) (or (58)).

Remark 10 *We can also prove that $\widetilde{V}_{\mathcal{D}}$ lies completely within a weight space of $\widetilde{V}_\lambda$ using characters and Schur function identities.*

Acknowledgements

We would like to thank Richard Stanley for suggesting that the Schur function approach might work, rather than our more complicated approach in terms of Hall-Littlewood polynomials, and also for the observation that the tensor product of two twisted representations can be written as a positive sum (rather than as a virtual sum) of twisted representations. We would also like to thank Shlomo Sternberg and David Vogan for useful discussions and comments.

References

[1] V. Guillemin, S. Sternberg, and J. Weitsman. Signature quantization. *Proc. Natl. Acad. Sci. USA*, 100(22):12559–12560, 2003.

[2] Jose Agapito. A weighted version of quantization commutes with reduction principle for a toric manifold. arXiv:math.SG/0307318, 2003.

[3] Etienne Rassart. *Geometric approaches to computing Kostka numbers and Littlewood-Richardson coefficients*. PhD thesis, Massachusetts Institute of Technology, 2004.

[4] William Fulton and Joe Harris. *Representation Theory*, volume 129 of *GTM*. Springer, 1991.

[5] George Lusztig. Singularities, character formulas, and a q-analog of weight multiplicities. In *Analysis and topology on singular spaces, II, III (Luminy, 1981)*, Astérisque 101-102, pages 208–229. Soc. Math. France, Paris, 1983.

[6] M. F. Atiyah and R. Bott. A Lefschetz fixed point formula for elliptic complexes. I. *Ann. of Math. (2)*, 86:374–407, 1967.

[7] M. F. Atiyah and R. Bott. A Lefschetz fixed point formula for elliptic complexes. II. *Ann. of Math. (2)*, 88:451–491, 1968.

[8] Daniel J. Kleitman and Kenneth J. Winston. Forests and score vectors. *Combinatorica*, 1:49–54, 1981.

[9] Richard P. Stanley. Decompositions of rational convex polytopes. *Annals of Discrete Mathematics*, 6:333–342, 1980.

[10] John R. Stembridge. Shifted tableaux and the projective representations of the symmetric groups. *Advances in Mathematics*, 74:87–134, 1989.

[11] I. G. Macdonald. *Symmetric functions and Hall polynomials, second edition*. Oxford Mathematical Monographs. Clarendon Press, Oxford Science Publications, 1995.

[12] I. M. Gel'fand and M. L. Cetlin. Finite-dimensional representations of the group of unimodular matrices. *Dokl. Akad. Nauk. SSSR*, 71:825–828, 1950. (In Russian; an English translation appeared in Gelfand's *Collected papers*, Springer-Verlag, 1987-1989).

[13] Richard P. Stanley. *Enumerative Combinatorics, volume 2*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

SELF-AVOIDING WALKS CROSSING A SQUARE

A. J. GUTTMANN, I. JENSEN, AND M. BOUSQUET-MÉLOU

Abstract. We study a restricted class of self-avoiding walks (SAW) which start at the origin $(0,0)$, end at (L,L), and are entirely contained in the square $[0,L]\times[0,L]$ on the square lattice $\mathbb{Z}^2$. The number of distinct walks is known to grow as $\lambda^{L^2+o(L^2)}$. We give a precise estimate for λ as well as obtaining upper and lower bounds. We give exact results for the number of SAW of length $2L+2K$ for $K=0,1,2$ and asymptotic results for $K=o(L^{1/3})$.

We also consider the model in which a weight or *fugacity* x is associated with each step of the walk. This gives rise to a canonical model of a *phase transition*. For $x<1/\mu$ the average length of a SAW is proportional to L, while for $x>1/\mu$ it is proportional to L^2. Here μ is the growth constant of unconstrained SAW in $\mathbb{Z}^2$. For $x=1/\mu$ we provide numerical evidence, but no proof, that the average walk length is $O(L^{4/3})$.

We also consider Hamiltonian walks under the same restrictions. These grow as $\tau^{L^2+o(L^2)}$ on the same $L\times L$ lattice. We give precise estimates for τ, as well as upper and lower bounds, and prove $\tau<\lambda$.

Nous étudions les chemins auto-évitants (CAE) du réseau carré qui partent de l'origine $(0,0)$, finissent en (L,L), et sont entièrement contenus dans le carré $[0,L]\times[0,L]$. On sait que le nombre de tels chemins croît comme $\lambda^{L^2+o(L^2)}$. Nous donnons une estimation précise, ainsi que des bornes supérieures et inférieures pour λ. Nous donnons le nombre exact de CAE de longueur $2L+2K$ traversant le carré de côté L, pour $K=0,1,2$, et le comportement asymptotique de ce nombre pour $K=o(L^{1/3})$.

On associe ensuite un poids x à chaque pas d'un chemin, ce qui mène à une modèle présentant une transition de phase. Si μ désigne la constante de croissante des CAE non contraints, alors pour $x<1/\mu$, la longueur moyenne d'un CAE traversant le carré de côté L est proportionnelle à L, tandis qu'elle est proportionnelle à L^2 lorsque $x>\mu$. Pour $x=\mu$, nos données numériques suggèrent que la longueur moyenne est en $O(n^{3/4})$.

Nous considérons aussi des chemins hamiltoniens traversant un carré. Le nombre de tels chemins croît comme $\tau^{L^2+o(L^2)}$. Nous donnons une estimation précise et des bornes supérieures et inférieures pour τ, et nous prouvons que $\tau<\lambda$.

1. Introduction

We are considering the problem of self-avoiding walks on the square lattice $\mathbb{Z}^2$. For walks on an infinite lattice, it is generally accepted [9] that the number c_n of such walks of length n, considered up to a translation, grows as $c_n \sim const.\mu^n n^{\gamma-1}$, with metric properties, such as mean-square radius of gyration or mean-square end-to-end distance growing as $\langle R^2\rangle_n \sim const.n^{2\nu}$, where $\gamma=43/32$ and $\nu=3/4$. The growth constant μ is lattice dependent, and for the square lattice is not known exactly, but is indistinguishable numerically from the unique positive root of the equation $13x^4-7x^2-581=0$. We denote the generating function by $C(x):=\sum_n c_n x^n$, and it will be useful to define a second generating function for those SAW which start at the origin $(0,0)$ and end at a given point (u,v), as $G_{(0,0;u,v)}(x)$. In terms of this generating function, the mass $m(x)$ is defined [9] to be the rate of decay of G along a

Date: May 11, 2005.

MBM was partially supported by the European Commission's IHRP Programme, grant HPRN-CT-2001-00272, "Algebraic Combinatorics in Europe".

coordinate axis,

$$m(x) := \lim_{n\to\infty} \frac{-\log G_{(0,0;n,0)}(x)}{n}. \tag{1}$$

Here, we are interested in a restricted class of square lattice SAW which start at the origin $(0,0)$, end at (L,L), and are entirely contained in the square $[0,L]\times[0,L]$. A fugacity x is associated with each step of the walk. Historically, this problem seems to have led two largely independent lives. One as a problem in combinatorics (in which case the fugacity has been implicitly set to $x=1$), and one in the statistical mechanics literature where the behaviour as a function of fugacity x has been of considerable interest, as there is a fugacity dependent phase transition.

The problem seems to have first been mentioned by Knuth [7], within the framework of a discussion on how to estimate large numbers. The first full discussion as a mathematical problem seems to be by Abbott and Hanson [1] in 1978, many of whose results and methods are still useful today. In [10] there is mention of a version of the problem being due to earlier work of Hammersley. A key question considered in [1] and in this paper, is the number of distinct SAW on the constrained lattice, and their growth as a function of the size of the lattice. Let $c_n(L)$ denote the number of n-step SAW which start at the origin $(0,0)$, end at (L,L) and are entirely contained in the square $[0,L]\times[0,L]$. Further, let $C_L(x) := \sum_n c_n(L)x^n$. Then $C_L(1)$ is the number of distinct walks from the origin to the diagonally opposite corner of an $L\times L$ lattice. In [1], and independently in [13] it was proved that $C_L(1) = \lambda^{L^2+o(L^2)}$. The value of λ is not known, though bounds and estimates have been given in [1, 13]. One of our purposes in this paper is to improve on both the bounds and the estimate.

In the statistical mechanics literature, the problem appears to have been introduced by Whittington and Guttmann [13] in 1990, who were particularly interested in the phase transition that takes place as one varies the fugacity associated with the walk length. At a critical value, x_c the average walk length of a path on an $L\times L$ lattice changes from being proportional to L to being proportional to L^2. In [13] the critical fugacity proved to be $\geq 1/\mu$, and conjectured to be $x_c = 1/\mu$. In [8] the conjecture was proved.

In [1] the more general problem of SAW constrained to an $L\times M$ lattice was considered, where the analogous question was asked: how many self-avoiding paths are there from $(0,0)$ to (L,M)?

If one denotes the number of such paths by $C_{L,M}$, it is clear that, for M fixed, the paths can be generated by a finite dimensional transfer matrix, and hence that the generating function is rational. Indeed, in [1] it was proved that

$$G_2(z) = \sum_{L\geq 0} C_{L,2}z^L = \frac{1-z^2}{1-4z+3z^2-2z^3-z^4}, \tag{2}$$

(where here we have corrected a typographical error). It follows that $C_{L,2} \sim \text{const.}\lambda_2^{2L}$, where $\lambda_2 = \sqrt{\frac{2}{\sqrt{13}-3}} = 1.81735\ldots$.

In this paper we also consider two further problems which can be seen as generalisations of the stated problem. Firstly, we consider the problem where SAWs are allowed to start anywhere on the left edge of the square and terminate anywhere on the right edge; so these are walks *spanning* the rectangle from left to right. We denote by T_L the number of such SAWs on an $L\times L$ lattice. Secondly, we consider the problem in which there may be several independent self- and mutually-avoiding walks, each such walk starting and ending on the perimeter of the square. The SAW are not allowed to take steps along the edges of the perimeter. Such walks partition the rectangle into distinct regions and by colouring the regions alternately black and white we get a *cow-patch* pattern. We denote by P_L the number of such configurations of SAWs on an $L\times L$ lattice. Each problem is illustrated in figure 1. These generalisations are introduced as they allow us to establish rigorous bounds on λ, which we do below.

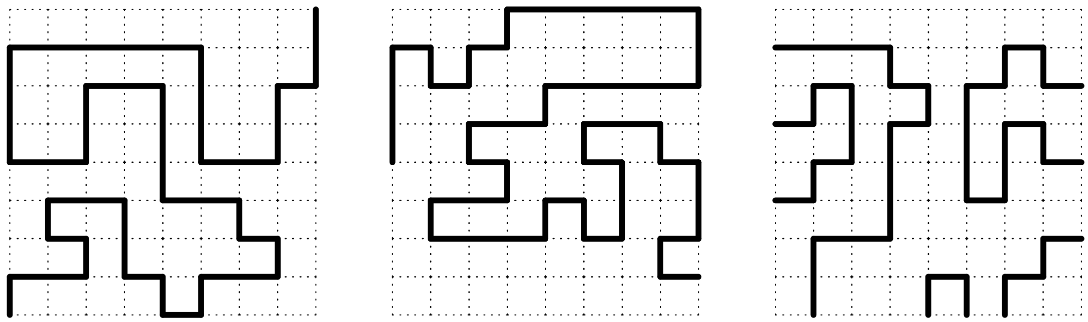


FIGURE 1. An example of a SAW configuration crossing a square (left panel), spanning a square from left to right (middle panel) and a cow-patch (right panel).

Following the work in [13], Madras in [8] proved a number of theorems. In fact, most of Madras's results were proved for the more general d-dimensional hypercubic lattice, but here we will quote them in the more restricted two-dimensional setting.

Theorem 1. *The following limits,*

$$\mu_1(x) := \lim_{L\to\infty} C_L(x)^{1/L} \quad \textit{and} \quad \mu_2(x) := \lim_{L\to\infty} C_L(x)^{1/L^2},$$

are well-defined in $\mathbb{R} \cup \{+\infty\}$.

More precisely,

(*i*) $\mu_1(x)$ *is finite for* $0 < x \le 1/\mu$, *and is infinite for* $x > 1/\mu$. *Moreover,* $0 < \mu_1(x) < 1$ *for* $0 < x < 1/\mu$ *and* $\mu_1(1/\mu) = 1$.

(*ii*) $\mu_2(x)$ *is finite for all* $x > 0$. *Moreover,* $\mu_2(x) = 1$ *for* $0 < x \le 1/\mu$ *and* $\mu_2(x) > 1$ *for* $x > 1/\mu$.

The average length of (weighted) walks crossing the $L \times L$ square is defined to be

$$\langle n(x)\rangle_L := \sum_n n c_n(L) x^n \Big/ \sum_n c_n(L) x^n. \tag{3}$$

Let $a(x)$ and $b(x)$ be two functions of some variable x. We write that $a(x) = \Theta(b(x))$ as $x \to x_0$ if there exist two positive constants κ_1 and κ_2 such that, for x sufficiently close to x_0,

$$\kappa_1\, b(x) \le a(x) \le \kappa_2\, b(x).$$

Theorem 2. *For* $0 < x < 1/\mu$, *we have that* $\langle n(x)\rangle_L = \Theta(L)$ *as* $L \to \infty$, *while for* $x > 1/\mu$, *we have* $\langle n(x)\rangle_L = \Theta(L^2)$.

The situation at $x = 1/\mu$ is unknown. We provide compelling numerical evidence that in fact $\langle n(1/\mu)\rangle_L = \Theta(L^{1/\nu})$, where $\nu = 3/4$, in accordance with an intuitive suggestion in [8].

Theorem 3. *For* $x > 0$, *define* $f_1(x) = \log \mu_1(x)$ *and* $f_2(x) = \log \mu_2(x)$.

(*i*) *The function* f_1 *is a strictly increasing, negative-valued convex function of* $\log x$ *for* $0 < x < 1/\mu$, *and* $f_1(x) = \Theta(-m(x))$ *as* $x \to 1/\mu^-$, *where* $m(x)$ *is the* mass, *defined by* (1).

(*ii*) *The function* f_2 *is a strictly increasing, convex function of* $\log x$ *for* $x > 1/\mu$, *and satisfies* $0 < f_2(x) \le \log \mu + \log x$.

Some, but not all of the above results were previously proved in [13], but these three theorems elegantly capture all that is rigorously known.

2. Bounds on the growth constant λ

For the more general problem of SAW going from $(0,0)$ to (L, M) on an $L \times M$ lattice, it was proved in [1] that

Theorem 4. *For each fixed M, $\lim_{L\to\infty} C_{L,M}^{\frac{1}{LM}} = \lambda_M$ exists.*

Further, Abbott and Hanson state that a similar proof can be used to establish that $\lim_{L\to\infty} C_{L,L}^{\frac{1}{L^2}} := \lambda$ exists. This was proved rather differently in [13].

2.1. Upper bounds on λ

In [1] an upper bound on the growth constant λ was obtained by recasting the problem in a matrix setting. We give below an alternative method for establishing upper bounds, based on defining a superset of paths. We then show that these two methods are in fact essentially identical.

Following [1], consider any non-intersecting path crossing the $L \times L$ square. Label each unit square in the $L \times L$ lattice by 1 if it lies to the right of the path, and by 0 if it lies to the left. This provides a one-to-one correspondence between paths and a subset of $L \times L$ matrices with elements 0 or 1. Matrices corresponding to allowed paths are called *admissible,* otherwise they are *inadmissible.* Since the total number of $L \times L$ $0-1$ matrices is 2^{L^2}, we immediately have the weak bound $C_{L,L} \leq 2^{L^2}$. Of the 16 possible 2×2 matrices, only 14 can correspond to portions of non-intersecting lattice paths. Lote that there are only 12 actual paths from $(0,0)$ to $(2,2)$, but a further two matrices may correspond to paths that are embedded in a larger lattice. Thus we find the bound $C_{L,L} \leq 14^{(L/2)^2}$, so $\lambda \leq 1.9343...$ Similarly, for 3×3 lattices we find 320 admissible matrices (out of a possible 512), so $\lambda \leq 320^{1/9} = 1.8982..$ For 4×4 lattices, [1] claims that there are 22662 admissible matrices, but we believe the correct number to be 22816, giving the bound $\lambda \leq 1.8723...$ We have made dramatic extensions of this work, using a combination of finite-lattice methods and transfer matrices, as described below, and have determined the number of admissible matrices up to 19×19. There are 3.5465202×10^{90} such matrices, giving the bound

$$\lambda \leq 1.781684.$$

This bound is fully equivalent to the bound $\lambda \leq (2P_L)^{1/L^2}$, where P_L denotes the number of cow-patch configurations on the $L \times L$ lattice. This equivalence follows if one colours cow-patches by two colours, such that adjacent regions have different colours. Labelling the two colours 0 and 1 produces a $0-1$ matrix representation.

2.2. Lower bounds on λ

In [1] the useful bound

$$\lambda > \lambda_M^{\frac{M}{M+1}}$$

is proved.

The above evaluation of λ_2, see (2), immediately yields $\lambda > 1.4892\ldots$.

Based on exact enumeration, we have found the exact generating functions $G_M(z) = \sum_L C_{L,M} z^L$ for $M \leq 6$. For $M = 3$ we find:

$$G_3(z) = \frac{[1, -4, -4, 36, -39, -26, 50, 6, -15, 1]}{[1, -12, 54, -124, 133, 16, -175, 94, 69, -40, -12, 4, 1]},$$

where we denote by $[a_0, a_1, \ldots, a_n]$ the polynomial $a_0 + a_1 z + \cdots + a_n z^n$. As explained above, all the generating functions $G_M(z)$ are rational. For $M = 4, 5, 6$, their numerator and denominators are found to have degree $(26,27),(71,75)$ and $(186,186)$ respectively, in an obvious notation.

From these, we find the following values: $\lambda_3 = 1.76331\ldots$, $\lambda_4 = 1.75146\ldots$, $\lambda_5 = 1.74875\ldots$ and $\lambda_6 = 1.74728\ldots$ from which we obtain the bound $\lambda > 1.61339\ldots$.

However, an alternative lower bound can be obtained from spanning SAWs, defined in Section 1. If T_L denotes the number of spanning SAW on the $L \times L$ lattice, then we prove in the full version of this paper that

$$\lambda \geq T(L)^{1/((L+1)(L+2))}. \tag{4}$$

This gives the improved bound $\lambda > 1.6284$.

Combining our results for lower and upper bounds finally gives

$$1.6284 < \lambda < 1.781684.$$

3. Computer enumeration

In the following we give a fairly detailed description of the algorithm we use to enumerate the number of walks crossing a square and briefly outline how this basic algorithm is modified in order to include a step fugacity, study SAWs spanning a square and the cow-patch configurations.

3.1. The algorithm

The basic algorithm used to enumerate self-avoiding walks crossing a square is based on the method of Conway et al. [2] for enumerating ordinary self-avoiding walks. The number of walks crossing an $L \times M$ rectangle is counted using a transfer matrix algorithm. The transfer matrix technique involves drawing a boundary line through the rectangle intersecting

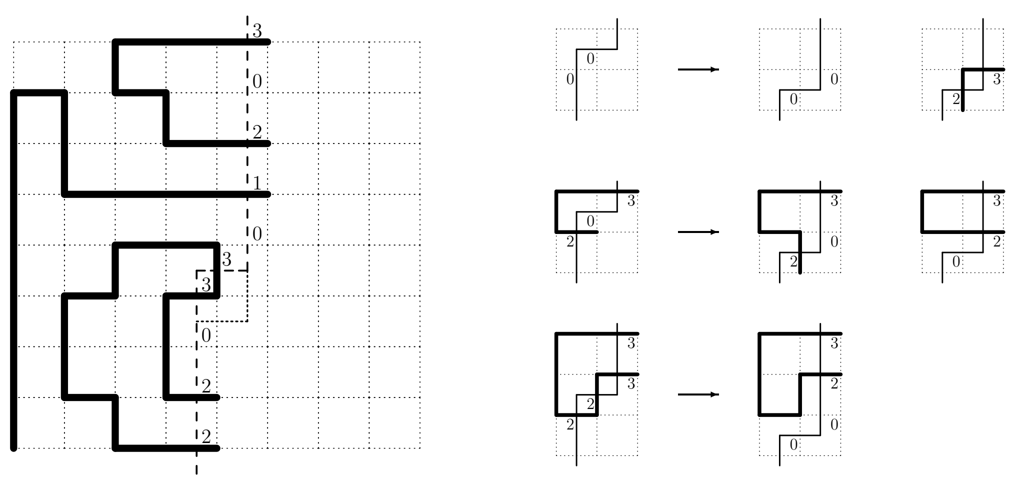


FIGURE 2. The left panel shows a snapshot of the intersection (dashed line) during the transfer matrix calculation. Walks within a rectangle are enumerated by successive moves of the kink in the boundary, as exemplified by the position given by the dotted line, so that the $L \times M$ rectangle is built up one vertex at a time. To the left of the boundary we have drawn an example of a partially completed walk. Numbers along the boundary indicate the encoding of this particular configuration. The right panel shows some of the local configurations which occur as the kink in the intersection is moved one step.

up to $M+2$ edges. For each configuration of occupied or empty edges we maintain a count of partially completed walks intersecting the boundary in that pattern. Walks in rectangles are counted by moving the boundary, adding one vertex at a time (see figure 2). Rectangles are built up column by column with each column constructed one vertex at a time. Configurations are represented by lists of states $\{\sigma_i\}$, where the value of the state σ_i must indicate if the ith edge of the boundary is occupied or empty. An empty edge is indicated by $\sigma_i = 0$. An occupied edge is either *free* (that is, not connected to other edges of the boundary by a path located to the left of the boundary) or connected to exactly one such edge. We indicate this by $\sigma_i = 1$ for a free end, $\sigma_i = 2$ for the lower end of a loop and $\sigma_i = 3$ for the upper end of loop connecting two edges. Since we are studying self-avoiding walks on a two-dimensional lattice the compact encoding given above uniquely specifies which ends are paired. Read from the bottom the configuration along the intersection in figure 2 is $\{2203301203\}$ (prior to the move) and $\{2300001203\}$ (after the move).

There are major restrictions on the possible configurations and their updating rules. Firstly, since the walk has to cross the rectangle there is exactly one free end in any configuration. Secondly, all remaining occupied edges are connected by a path to the left of the intersection and we cannot close a loop. It is therefore clear that the total number of 2's equals the total number of 3's. Furthermore, as we look through the configuration from the bottom the number of 2's is never smaller than the number of 3's (so that configurations can be seen as well-balanced parentheses systems). We also have to ensure that the graphs we construct have only one connected component. In the following we shall briefly show how this is achieved.

TABLE 1. The various 'input' states and the 'output' states which arise as the boundary line is moved in order to include one more vertex. Each panel contains up to three possible 'output' states or other allowed actions.

Bottom \ Top	0	1	2	3
0	00 23	01 10 Res	02 20	03 30
1	01 10 Res		$\overset{\frown}{00}$	$\overset{\frown}{00}$
2	02 20	$\overset{\frown}{00}$	$\overline{00}$	
3	03 30	$\overset{\frown}{00}$	00	$\overline{00}$

We call the configuration before and after the move the 'source' and 'target', respectively. Initially we have just one configuration with a single '1' at position 0 (all other entries '0') thus ensuring that we start in the bottom-left corner. As the boundary line is moved one step, we run through all the existing sources. Each source gives rise to one or two targets and the count of the source is added to the count of the target (the initial count of a target being zero). After a source has been processed it can be discarded since it will make no further contribution. Table 1 lists the possible local 'input' states and the 'output' states which arise as the kink in the boundary is propagated one step, and the various symbols are explained below. Firstly, the values of the 'Bottom' and 'Top' table entries refer to the edge-states of the kink prior to the move. The Top (Bottom) entry is the state of the edge intersected by (below) the horizontal part of the boundary. Some of the update rules are illustrated further in figure 2. The topmost panels represent the input state '00' having the allowed output states '00' and '23' corresponding to leaving the edges empty or inserting a new loop, respectively. The middle panels represents the input state '20' with output states '20' and '02' from the two ways of continuing the loop-end (note that the loop has to be continued since we would otherwise generate an additional free end not located at the allowed positions in the corners). The bottommost panels represents the input state '22' as part of the configuration $\{02233\}$. In this case we connect two loop-ends and we thus join two separate loops into a single larger loop. The matching upper end of the innermost loop becomes the new lower end of the joined loop. The relabeling of the matching loop-end when connecting two '2's (or two '3's) is denoted by over-lining in Table 1. When we join loop-ends to a free end (inputs '12', '21', '13', and '31') we have to relabel the matching loop-end as a free end. This type of relabeling is indicated by the symbol $\widehat{00}$. The input state '11' never occurs since there is only one free end. The input state '23' is not allowed since connecting the two ends results in a closed loop. Finally, we have marked two outputs, from the inputs '01' and '10' with 'Res', indicating situations where we terminate free ends. This results in completed partial walks and is only allowed if there are no other occupied edges in the source (otherwise we would produce graphs with separate pieces) and if we are at the top-most vertex (otherwise we would not cross the rectangle). The count for this configuration is the number of walks crossing a rectangle of height M and length L equal to the number of completed columns.

3.2. Complexity

The time required to obtain the number of walks on $L \times M$ rectangles grows exponentially with M and linearly with L. Time and memory requirements are basically proportional to the maximal number of distinct configurations along the boundary line. When there is no kink in the intersection (a column has just been completed) we can calculate this number, $N_{\text{conf}}(M)$, exactly. Obviously the free end cuts the boundary line configuration into two separate pieces. Each of these pieces consists of '0's and an equal number of '2's and '3's with the latter forming a well-balanced parenthesis system.

Each piece thus corresponds to a *Motzkin path* [12, Ch. 6] (just map 0 to a horizontal step, 2 to a north-east step, and 3 to a south-east step). The number of Motzkin paths M_n with n steps is easily derived from the generating function $\mathcal{M}(x) = \sum_n M_n x^n$, which satisfies $\mathcal{M} = 1 + x\mathcal{M} + x^2\mathcal{M}^2$, so that

$$\mathcal{M}(x) = [1 - x - ((1+x)(1-3x))^{1/2}]/2x^2. \tag{5}$$

The number of configurations $N_{\text{conf}}(M)$ for a rectangle of height M is simply obtained by inserting a free end between two Motzkin paths, so that the generating function $\sum_M N_{\text{conf}}(M)x^M$ is simply $x\mathcal{M}(x)^2$. The Lagrange inversion formula gives

$$N_{\text{conf}}(M) = 2\sum_{i\geq 0} \frac{(M+1)!}{i!(i+2)!(M-2i)!}.$$

When the boundary line has a kink the number of configurations exceeds $N_{\text{conf}}(M)$ but clearly is less than $N_{\text{conf}}(M+1)$. From (5) we see that asymptotically $N_{\text{conf}}(M)$ grows like 3^M (up to a power of M). So the same is true for the maximal number of boundary line configurations and hence for the computational complexity of the algorithm. Note that the total number a walks grows like λ^{LM} and our algorithm thus leads to a better than exponential improvement over direct enumeration.

The integers occurring in the expansion become very large so the calculation was performed using modular arithmetic [6]. This involves performing the calculation modulo various prime numbers p_i and then reconstructing the full integer coefficients at the end. We used primes of the form $p_i = 2^{30} - r_i$, where r_i are distinct integers, less than 1000, such that p_i is a (different) prime for each value of i. The Chinese remainder theorem ensures that any integer has a unique representation in terms of residues. If the largest integer occurring in the final expansion is m, then we have to use a number of primes k such that $p_1 p_2 \cdots p_k > m$.

3.3. Extensions of the algorithm

The algorithm is easily generalised to include a step fugacity x. The count associated with the boundary line configuration has to be replaced by a generating function for partial walks. Since we only use this generalisation to study walks crossing a square, the generating function is a polynomial of degree (at most) M^2 in x. The coefficient of x^n in this polynomial is the number of walks of length n intersecting the boundary line in the pattern specified by the configuration. When the boundary is updated, if m additional steps are inserted, the generating function of the source is multiplied by x^m and added to the generating function of the target. Not all M^2 terms in the polynomials need be retained. Firstly, only terms with n even are non-zero and only these are retained. Secondly, in order to construct a given boundary line configuration, a certain minimal number of steps $n_{\min}$ are required and terms with $n < n_{\min}$ can be discarded.

The generalisation to spanning walks is also quite simple. Firstly, we have $M+1$ initial configurations which are empty except for a free end at position $0 \leq j \leq M$. This corresponds to the $M+1$ possible starting positions for the walk on the left boundary. Secondly, we have to change how we produce the final counts. The easiest way to ensure that a walk spans the rectangle and that only single component graphs are counted is as follows: When column $L+1$ has been completed we look at the $M+1$ configurations with a single free end and add the counts from all of them. This is the number of walks spanning an $L \times M$ rectangle.

The generalisation to cow-patch patterns is more complicated. Graphs can now have many separate components and there can be many free ends in a boundary line configuration. Note also that each free end has to start and terminate with a step perpendicular to the border of the rectangle and there are no steps along the edges of the borders of the rectangle. There are 2^{M-1} initial configurations since any of the edges in the first column from position 1 to $M-1$ can be occupied by a free end or be empty. There is an extra updating rule in the bulk in that we can have the local input '11' (joining of two free ends) with the only possible output being '00'. Also the updating rules at the upper and lower borders of the rectangle are different in this case. At the upper border we only have the input '00' with the outputs '00' and '10' corresponding to the insertion of a free end on a vertical edge at the upper border. There is no '23' or '01' outputs since these would produce an occupied edge along the upper border. At the lower border we have inputs '00', '01', and '02' and in each case the only possible output is '00' (with the appropriate relabeling in the '02' case). Finally, the count of the number of cow-patch patterns is obtained by summing over all boundary line configurations after the completion of a column.

3.4. Results

As discussed above, in order to obtain the exact value of the number of SAW crossing a square, some of which are integers with nearly 100 digits, we performed the enumerations several times, each time *modulo* a different prime. The enumerations were then reconstructed using the Chinese Remainder Theorem. Each run for a 19×19 lattice took about 72 hours using 8 processors of a multiprocessor 1 GHz Compaq Alpha computer. Ten such runs were needed to uniquely specify the resultant numbers.

Proceeding as above, we have calculated $c_n(L)$ for all n for $L \leq 17$. In other words, we have obtained the polynomials $C_L(x)$ for $L \leq 17$. In addition, we have computed $C_{18}(1)$ and $C_{19}(1)$, the total number of SAW crossing an 18×18 and 19×19 square respectively. We have also computed the corresponding quantities for cow-patch and spanning SAWs, denoted $P_L(1)$ and $T_L(1)$ respectively, for $L \leq 19$.

Finally, in [1] the question was asked whether $C_{L,M}^{\frac{1}{LM}}$ is decreasing in both L and M. We can answer this in the negative, based on our enumerations.

4. Numerical analysis

It has been proved [1, 13] that $\lim_{L\to\infty} C_{L,L}^{\frac{1}{L^2}} = \lambda$ exists. From this it is reasonable to expect (but not a logical consequence) that $R_L = C_{L+1,L+1}/C_{L,L} \sim \lambda^{2L}$ so the generating function $\mathcal{R}(x) = \sum_L R_L x^L$ has a radius of convergence $x_c = 1/\lambda^2$, which we can estimate accurately using differential approximants [4]. We estimate in this way that for the crossing problem $x_c = 0.32858(5)$, for the spanning problem $x_c = 0.3282(6)$ and for the cow-patch problem $x_c = 0.328574(2)$. So we see that λ is the same for the three problems, and we estimate that $\lambda = 1.744550(5)$. It is not difficult to prove that λ defined for SAW crossing a square, and for spanning walks takes the same value. For cow-patch walks, this is somewhat more difficult, but we have done so (see the full version of this paper).

We now speculate on the sub-dominant terms. For SAW on an infinite lattice, it is widely accepted (but not proved) that $c_n \sim const.\mu^n n^g$, where c_n is the number of n step SAW equivalent up to a translation.

It seems reasonable to speculate that, the number of SAWs crossing an $L \times L$ lattice is equivalent to $A\lambda^{L^2+bL}L^\alpha$. We have investigated this possibility numerically, and found it to be well supported by the data.

For cow-patches we find $b \approx 0.8558$ and $\alpha \approx -0.500$. For transverse walks and walks crossing a square b is quite small, possibly zero. For transverse walks we find $\alpha \approx 1.75$ while for walks crossing the square $\alpha \approx 0$. This suggests asymptotic behaviours $A_P\lambda^{L^2+0.8558L}/\sqrt{L}$, $A_T\lambda^{L^2}L^{7/4}$ and $A_W\lambda^{L^2}\log L$ respectively, where A_P, A_T, and A_W can be estimated, and the $\log L$ term (or some power of a logartihm) would follow if α were exactly zero.

As remarked in the introduction, we have also studied (numerically) the behaviour of $\langle n(1/\mu)\rangle_L$, by a log-log plot as well as other numerical methods. The results are totally consistent with the conjecture [8], that $\langle n(1/\mu)\rangle_L$ is proportional to $L^{1/\nu}$, where $\nu = 3/4$.

5. Asymptotics for walks of "small" length crossing a square

We now consider walks of length $2L + 2K$ crossing an $L \times L$ square. Note that walks of length $2L$ are the minimal possible length. With $K = 0$ the number of possible walks is $\binom{2L}{L}$. This result is obvious, as there are $2L$ steps in the path, of which L must be in the positive x (and of course positive y) direction. Note that this has the asymptotic expansion

$$\binom{2L}{L} = \frac{4^L}{\sqrt{L\pi}}\left(1 - \frac{1}{4L} + \frac{1}{128L^2} + \frac{5}{1024L^3} + \mathrm{O}\left(L^{-4}\right)\right).$$

With $K = 1$ we have proved that the number of possible paths is given by $2L\binom{2L}{L+2}$. This result has the asymptotic expansion

$$2L\binom{2L}{L+2} = \frac{L4^L}{\sqrt{L\pi}}\left(2 - \frac{33}{4L} + \frac{1345}{64L^2} - \frac{23835}{512L^3} + \mathrm{O}\left(L^{-4}\right)\right).$$

For $K = 2$ we have proved that the number of possible paths is given by

$$\frac{2(2L)!}{L!(L+4)!}(48 + 90L + 8L^2 - 28L^3 - 3L^4 + 4L^5 + L^6) - 4.$$

This has asymptotic expansion

$$\frac{L^2 4^L}{\sqrt{L\pi}}\left(2 - \frac{49}{4L} + \frac{2913}{64L^2} - \frac{92971}{512L^2} + \mathrm{O}\left(L^{-3}\right)\right).$$

Our technique can be, in theory, extended to count walks of length $2L + 2K$, for any given value of K. It proves that the sequence of numbers thus obtained is always *P-recursive*. That is to say, it satisfies a linear recurrence relation with polynomial coefficients [11]. But, even for $K = 3$, the number of special cases that must be treated becomes very large. We have resorted to a numerical study for higher values of K, and for $K = 3$ we found

$$\frac{L^3 4^L}{\sqrt{L\pi}} \left(\frac{4}{3} - \frac{49}{6L} + \frac{1931 \pm 1}{64L^2} + \mathrm{O}(L^{-3}) \right),$$

while the corresponding result for $K = 4$ is

$$\frac{L^4 4^L}{\sqrt{L\pi}} \left(\frac{2}{3} + \frac{11}{4L} + \mathrm{O}(L^{-2}) \right).$$

We can give an heuristic argument for the general form of the leading term in the asymptotic expansion of the case $K = k$, which leads to the leading order term $\frac{4^L}{\sqrt{L\pi}} \frac{(2L)^k}{k!}$. Here the first term is given by the number of ways of choosing the "backbone", $\binom{2L}{L} \sim \frac{4^L}{\sqrt{L\pi}}$ and the second is given by the number of ways of placing k defects (or backward steps) on a path of length $2L$, which is just $(2L)^k$. The defects are indistinguishable, introducing the factor $k!$.

This argument can be refined into a proof, for $K = \mathrm{o}(L^{1/3})$ by following the steps, *mutatis mutandis* in the proof of a similar result given in [3].

6. Hamiltonian walks crossing a square

Hamiltonian walks can only exist on $2L \times 2L$ lattices. For lattices with an odd number of edges, one site must be missed. A Hamiltonian walk is of length $4L(L+1)$ on a $2L \times 2L$ lattice. The number of such walks grows as τ^{4L^2}, where we find $\tau \approx 1.472$ based on exact enumeration up to 17×17 lattices. This is about 20% less than λ, the growth constant for all SAWs. In [5] the estimate of the growth constant for Hamiltonian SAW on the unconstrained square lattice 1.472801 ± 0.00001 was given. This should be precisely the same as the corresponding result for Hamiltonian walks on an $L \times L$ lattice, in the large L limit. In [1] it is proved that $2^{1/3} \leq \tau \leq 12^{1/4}$, that is to say, $1.260 \leq \tau \leq 1.861$. We can improve on these bounds as follows: we define generalized cow-patch walks to be Hamiltonian if every vertex of the square not belonging to the border of the square belongs to one of the SAWs of the cow-patch. Then the upper bounds given above translate verbatim into upper bounds for τ, while lower bounds are given by Hamiltonian spanning walks and (4). In this way we find $1.429 \leq \tau \leq 1.52999$. As we have shown above that $1.6284 < \lambda$, this proves that $\tau < \lambda$.

References

[1] H. L. Abbott and D. Hanson, A lattice path problem, *Ars Combinatoria* **6**, 163-178 (1978).
[2] A. R. Conway, I. G. Enting and A. J. Guttmann, Algebraic techniques for enumerating self-avoiding walks on the square lattice, *J. Phys. A: Math.Gen.* **26**, 1519–1534 (1993).
[3] I.G Enting, A.J. Guttmann, L. B. Richmond and N. C. Wormald, Enumeration of almost convex polygons on the square lattice, *Random Structures and Algorithms* **3** (4), 445-461, (1992).
[4] A. J. Guttmann, Asymptotic analysis of power-series expansions in *Phase Transitions and Critical Phenomena vol. 13* (eds. C Domb and J L Lebowitz) (New York: Academic), (1989).
[5] J. L. Jacobsen, Frustration and Disorder in Discrete Lattice Models, *Ph. D, Univ. of Arhus*, July (1998).
[6] D. E. Knuth, *Seminumerical Algorithms. The Art of Computer Programming, Vol 2.* (Reading, Mass: Addison Wesley) (1969).
[7] D. E. Knuth, Mathematics and Computer Science: Coping with Finiteness. *Science* **194**, 1235-42 (1976).
[8] N. Madras, Critical behaviour of self-avoiding walks that cross a square, *J. Phys. A: Math. Gen,* **28**, 1535-1547, (1995).
[9] N. Madras and G. Slade, *The Self-Avoiding Walk* (Boston: Birkhäuser) (1993).
[10] J. G. Oxley and D. J. A. Welsh, On some percolation problems of J. M. Hammersley, *J. Appl. Prob.*, **16**, 526-40, (1979).
[11] R. P. Stanley, Differentiably finite power series, *European J. Combin.,* **1**,175-188, (1980).
[12] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, (1999).

[13] S. G. Whittington and A. J. Guttmann, Self-avoiding walks which cross a square, *J. Phys A:Math. Gen*, **23**, 5601-9, (1990).

ARC Centre of Excellence for Mathematics and Statistics of Complex Systems, Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia and CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France

E-mail address: tonyg,I.Jensen@ms.unimelb.edu.au and mireille.bousquet@labri.fr

The Genesis of the Macdonald Polynomial Statistics

J. Haglund

Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104-6395
jhaglund@math.upenn.edu

May 9, 2005

Abstract

Recently M. Haiman, N. Loehr and the author [HHL05b], [HHL05a] proved that for μ a partition of n, the modified Macdonald polynomial $\tilde{H}_\mu[z_1,\ldots,z_n;q,t]$ can be expressed as a sum of monomials in the z_i times certain nonnegative integral powers of q,t with direct combinatorial descriptions (i.e. statistics). These powers are generalizations of the classical permutation statistics maj and inv. The result was first conjectured by the author [Hag04a], and was partially motivated by a conjectured formula for the character of the space of diagonal harmonics [HHL$^+$05c]. Beyond giving a long-sought after combinatorial formula for Macdonald polynomials, this result has many nice corollaries, including a simple proof of Lascoux and Schützenberger's formula, involving the statistic cocharge, for Hall-Littlewood polynomials. In this paper we describe the sequence of experimental steps and Maple calculations which led to the discovery of the Macdonald polynomial statistics.

In un recente lavoro in collaborazione con M. Haiman, N. Loehr [HHL05b], [HHL05a] dimostriamo che per una partizione di n, μ, il polinomio di Macdonald modificato $\tilde{H}_\mu[z_1,\ldots,z_n;q,t]$ puo' essere espresso come un polinomio nelle variabili z_i i cui cofficienti sono potenze non negative di q,t. Questo risultato ha una diretta interpretazione combinatoria (o meglio statistica). Gli esponenti dei coefficienti sono una generalizzazione delle classiche statistiche di permutazione maj e inv. L'ipotesi di un tale risultato fu avanzata in [Hag04a] e in parte fu motivata dalla formula, in forma di congettura, per i caratteri dello spazio delle armoniche diagonali [HHL$^+$05c]. Questo risultato non solo da' un'interpretazione combinatoria, cercata da lungo tempo, dei polinomi di Macdonald, ma permette di derivare anche altri risultati rilevanti. Per esempio e' possibile dare una dimostrazione semplificata della formula di Lascoux and Schützenberger riguardante la statistica cocharge per i polinomi di Hall-Littlewood. In questo articolo descriviamo la sequenza di passaggi sperimentali e di calcoli fatti utilizzando Maple che portano alla scoperta della statistica dei polinomi di Macdonald.

1 Introduction

Given a sequence $\mu = (\mu_1, \mu_2, \ldots)$ of nonincreasing, nonnegative integers with $\sum_i \mu_i = n$, we say μ is a partition of n, denoted by either $|\mu| = n$ or $\mu \vdash n$. By adding or subtracting parts of size 0 if necessary, we will always assume partitions of n have exactly n parts. We let $\eta(\mu) = \sum_i (i-1)\mu_i$, and if λ is another partition, set $\tilde{K}_{\lambda,\mu}(q,t) = t^{\eta(\mu)} K_{\lambda,\mu}(q,1/t)$, where $K_{\lambda,\mu}(q,t)$ is Macdonald's q,t-Kostka polynomial [Mac95, p.354]. We call $\tilde{H}_\mu[Z;q,t] = \sum_{\lambda \vdash |\mu|} s_\lambda \tilde{K}_{\lambda,\mu}(q,t)$ the modified Macdonald polynomial, where $s_\lambda = s_\lambda[Z]$ is the Schur function, the sum is over all $\lambda \vdash |\mu|$, and $Z = z_1, \ldots, z_n$. The $\tilde{H}_\mu[Z;q,t]$ can be easily transformed by a plethystic substitution into Macdonald's original symmetric functions $P_\mu[Z;q,t]$. Macdonald defined the P_μ in terms of orthogonality with respect to a scalar product, and conjectured $K_{\lambda,\mu}(q,t) \in \mathbb{N}[q,t]$ [Mac95, p. 355]. (From their definition, all one can infer is that the $K_{\lambda,\mu}(q,t)$ are rational functions in q,t). This conjecture in turn led Garsia and Haiman to the "n! conjecture" [GH93], which was proved by Haiman in 2000 [Hai01]. This result implies that $\tilde{K}_{\lambda,\mu}(q,t) \in \mathbb{N}[q,t]$, and moreover that $\tilde{H}_\mu[X;q,t]$ is the character of a certain S_n-module $V(\mu)$, and thus gives a representation-theoretical interpretation for the coefficients of $\tilde{K}_{\lambda,\mu}(q,t)$. Macdonald also posed the problem of finding a combinatorial rule to describe the $K_{\lambda,\mu}(q,t)$, which is still open.

Recently the author introduced a conjectured combinatorial formula for the coefficient of a monomial in $\tilde{H}_\mu[Z;q,t]$ [Hag04a]. The formula, described in Section 2, is manifestly in $\mathbb{N}[q,t,z_1,\ldots,z_n]$. Shortly after its introduction, the conjecture was proved by M. Haiman, N. Loehr and the author. Some comments on the proof, which is surprisingly simple, are made in Section 3. As a corollary they obtain a new, short proof of a famous result of Lascoux and Schützenberger, namely that the coefficients in the Schur function expansion of the Hall-Littlewood polynomial $\tilde{H}_\mu[Z;0,t]$ can be expressed as a sum, over semi-standard young tableaux W, of q to a statistic known as cocharge(W). This new proof is discussed in more detail in Section 4. Also, since the matrix expressing the Schur basis in terms of the monomial basis is upper unitriangular, another consequence is a new proof that the $\tilde{K}_{\lambda,\mu}(q,t) \in Z[q,t]$. Another consequence is a new prrof of Knop and Sahi's formula for Jack symmetric functions [KS97], as well as an extension of their formula to Macdonald's integral form $J_\mu[Z;q,t]$. In addition a simple formula for the $\tilde{K}_{\lambda,\mu}(q,t)$ is obtained when μ has two columns. It is hoped that further refinements will result in combinatorial formulas for the $\tilde{K}_{\lambda,\mu}(q,t)$ for general μ.

The idea for this article was suggested by A. Garsia, who thought it would be beneficial to have a detailed description of the sequence of steps which led the author to the discovery of the Macdonald polynomial statistics. In Section 5 we overview the pioneering work of Garsia, Haiman and others on the n! conjecture, the space of diagonal harmonics DH_n, and the q,t-Catalan numbers. These numbers were defined by Garsia and Haiman as a complicated sum of rational functions in q,t, which they conjectured simplified to a polynomial in $\mathbb{N}[q,t]$ [GH96]. We then discuss a series of discoveries by the author, M. Haiman and others involving combinatorial formulas for the q,t-Catalan numbers and other sequences connected to DH_n, which culminated in the introduction of the "shuffle conjecture" by Haiman, Haglund, Loehr, Remmel and Ulyanov [HHL$^+$05c], which gives a combinatorial formula for the character of DH_n.

In Section 6 we discuss how various special cases of the shuffle conjecture led the author to suspect an analogous combinatorial formula for the monomial expansion of $\tilde{H}_\mu[X;q,t]$ could be found. We then include an explanation of the various stages and Maple calculations in the experimental process which resulted in the discovery of the statistics.

2 The Formula

We assign (row,column)-coordinates to squares in the first quadrant, obtained by permuting the (x,y) coordinates of the upper right-hand corner of the square, so the lower left-hand square has coordinates $(1,1)$, the square above it $(2,1)$, etc.. For a square w, we call the first coordinate of w the *row value* of w, denoted row(w), and the second coordinate of w the *column value* of w, denoted col(w). Given $\mu \vdash n$, we let μ also stand for the Ferrers diagram of μ (French convention), consisting of the set of n squares with coordinates (i,j), with $1 \le i \le n$, $1 \le j \le \mu_i$.

Let T be a finite set of squares in the first quadrant. A subset of squares of T consisting of all those $w \in T$ with a given row value is called a row of T, and a subset of squares of T consisting of all those $w \in T$ with a given column value is called a column of T. Furthermore, we let $T(i)$ denote the ith square of T encountered if we read across rows from left to right, starting with the squares of largest row value and working downwards. Given a square $w \in T$, define the leg of w, denoted $\mathrm{leg}(w)$, to be the number of squares in T which are strictly above and in the same column as w, and the arm of w, denoted $\mathrm{arm}(w)$, to be the number of squares in T strictly to the right and in the same row as w. Also, if w has coordinates (i, j), we let $\mathrm{south}(w)$ denote the square with coordinates $(i-1, j)$.

A word σ of length n in an alphabet $\mathcal{A}$ is a linear sequence $\sigma_1\sigma_2\cdots\sigma_n$, with $\sigma_i \in \mathcal{A}$. Note that repeats are allowed. If the letter i occurs α_i times in σ, for each $i \geq 1$, we say σ has content α, denoted $\mathrm{content}(\sigma) = \alpha$. We call a pair (σ, T), where σ is a word of positive integers and T is a set of squares in the first quadrant, a *filling*. We represent (σ, T) geometrically by placing σ_i in square $T(i)$, for $1 \leq i \leq n$. For $w \in T$, we let $\sigma(w)$ denote the element of σ placed in square w, and we call $\sigma_1 \cdots \sigma_n$ the *reading word* of (σ, T). A *descent* of (σ, T) is a square $w \in T$, with $\mathrm{south}(w) \in T$ and $\sigma(w) > \sigma(\mathrm{south}(w))$.

Let $\mathrm{Des}(\sigma, T)$ denote the set of all descents of (σ, T). For partitions μ, define a generalized major index statistic $\mathrm{maj}(\sigma, \mu)$ via

$$\mathrm{maj}(\sigma, \mu) = \sum_{w \in \mathrm{Des}(\sigma,\mu)} 1 + \mathrm{leg}(w). \tag{1}$$

An inversion of (σ, T) is a pair of squares (a, b) with $a, b \in T$, $a(\sigma) > b(\sigma)$, and either

$$\begin{cases} \mathrm{row}(a) = \mathrm{row}(b) \text{ and } \mathrm{col}(a) < \mathrm{col}(b), & \text{or} \\ \mathrm{row}(a) = \mathrm{row}(b) + 1 \text{ and } \mathrm{col}(a) > \mathrm{col}(b) \ . \end{cases} \tag{2}$$

Let $\mathrm{Inv}(\sigma, T)$ denote the set of all inversions of (σ, T), and define the inversion statistic $\mathrm{inv}(\sigma, T)$ via

$$\mathrm{inv}(\sigma, T) = |\mathrm{Inv}(\sigma, T)| - \sum_{w \in \mathrm{Des}(\sigma,T)} \mathrm{arm}(w), \tag{3}$$

where $|T|$ denotes the cardinality of a set T. For example, if (σ, μ) is the filling in Figure 1, then representing squares by their coordinates,

$$\mathrm{Des}(\sigma, \mu) = \{(2, 1), (4, 2), (2, 2)\}, \tag{4}$$
$$\mathrm{Inv}(\sigma, \mu) = \{((3, 1), (3, 2)), ((2, 2), (2, 3)), ((2, 2), (1, 1)), ((2, 3), (1, 1)), ((2, 3), (1, 2))\}, \tag{5}$$

so $\mathrm{maj}(\sigma, \mu) = 3 + 1 + 3 = 7$, $\mathrm{inv}(\sigma, \mu) = 5 - (2 + 0 + 1) = 2$.

2	2	
2	1	
3	5	3
1	1	4

Figure 1: A filling of the partition $(3, 3, 2, 2)$ by the word 2221353114.

Note that, if 1^n denotes a column of n cells, then

$$\mathrm{maj}(\sigma, 1^n) = \sum_{i \in \mathrm{Des}(\sigma,1^n)} i, \tag{6}$$

the usual major index statistic on the word σ, while

$$\mathrm{inv}(\sigma,(n)) = \sum_{\substack{1\le i<j\le n\\ \sigma_i>\sigma_j}} 1, \tag{7}$$

the usual inversion statistic.

For $\mu \vdash n$, define

$$\tilde{C}_\mu[Z;q,t] = \sum_\sigma t^{\mathrm{maj}(\sigma,\mu)} q^{\mathrm{inv}(\sigma,\mu)} z^\sigma, \tag{8}$$

where $z^\sigma = \prod_{i=1}^n z_{\sigma_i}$ is the "weight" of σ and the sum is over all words σ of n positive integers satisfying $1 \le \sigma_i \le n$ for $1 \le i \le n$. The following result was conjectured by the author [Hag04a] and proven by Haglund, Haiman, Loehr [HHL05b], [HHL05a]).

Theorem 1 *For all partitions μ,*

$$\tilde{C}_\mu[Z;q,t] = \tilde{H}_\mu[Z;q,t]. \tag{9}$$

Given a set T of squares and a subset $S \subseteq T$, define

$$F_T[Z;q,S] = \sum_{\substack{\sigma\\ \mathrm{Des}(\sigma,T)=S}} q^{\mathrm{inv}(\sigma,T)} z^\sigma. \tag{10}$$

In [Hag04a] the following result is obtained.

Theorem 2 *For all S,T, $F_T[Z;q,S]$ is a symmetric function in the z_i.*

Given $S \subseteq \mu$, let

$$P(S) = \sum_{w\in S} 1 + \mathrm{leg}(w). \tag{11}$$

Note that by the definition of $\mathrm{maj}(\sigma,\mu)$,

$$\tilde{C}_\mu[Z;q,t] = \sum_{S\subseteq\mu} t^{P(S)} F_\mu[Z;q,S]. \tag{12}$$

In [HHL05a] it is shown that the $F_T[Z;q,t]$ are special cases of polynomials introduced by Lascoux, Leclerc and Thibon [LLT97], commonly known as LLT polynomials. These are symmetric polynomials whose description involves a tuple of arbitrary skew shapes and whose coefficients depend on q. It has been an open conjecture of theirs that the coefficients of these polynomials when expanded in the Schur basis are in $\mathbb{N}[q]$. Thus we now have the expansion of $\tilde{H}_\mu[Z;q,t]$ into LLT polynomials, and furthermore understanding the positivity of the coefficients of the $\tilde{K}_{\lambda,\mu}(q,t)$ is reduced to understanding the special case of the positivity of LLT polynomials (when each skew-shape in the tuple is a ribbon).

Definition 1 *Given a word σ of content $(\gamma_1,\gamma_2,\ldots)$, construct a permutation σ', the standardization of σ, by replacing the γ_1 1's in σ by the numbers $1,\ldots,\gamma_1$, the γ_2 2's in σ by the numbers $\gamma_1+1,\ldots,\gamma_1+\gamma_2$, etc., in such a way that, for $i<j$, $\sigma_i \le \sigma_j$ if and only if $\sigma'_i < \sigma'_j$. For example, if $\sigma = 224123114$ then $\sigma' = 458167239$.*

Remark 1 *At first glance it may seem that $inv(\sigma,T)$ may not always be nonnegative, but given a square $u \in Des(\sigma,T)$, for each square v in the same row as u and to the right of u, either $\sigma(u) > \sigma(v)$, or $\sigma(v) > \sigma(south(u))$, or both. Assume for the moment that σ has distinct entries. If we adopt the convention that for a square $w \notin T$, $\sigma(w) = \infty$, it follows that $inv(\sigma,T)$ equals the number of triples of squares u,v,w, where $u \in T$, $v \in T$, $row(u) = row(v)$, $col(u) < col(v)$, and $south(u) = w$, and if we draw a circle through u,v,w, and read in the σ values of u,v,w in counterclockwise order around the circle, starting at the smallest value, then the three values form a strictly increasing sequence. If σ has repeated entries, first standardize then count triples in (σ',T) as above.*

3 The Proof

Let $p_k(Z) = \sum_i z_i^k$ be the kth power sum. Given a real parameter w, let $p_k[Z(1-w)] = \sum_i z_i^K(1-w^k)$ and $p_k[Z/(1-w)] = \sum_i z_i^k/(1-w^k)$. These are both special cases of plethystic notation, indicated by the square brackets around $Z(1-w)$ and $Z/(1-w)$. For an arbitrary symmetric function $F(Z)$, let $F[Z(1-w)]$ (respectively $F[Z/(1-w)]$) be the result of first expressing $F(Z)$ as a polynomial in the $p_k(Z)$, then replacing each $p_k(Z)$ by $p_k[Z(1-w)]$ (respectively $p_k[Z/(1-w)]$).

The polynomial $\tilde{H}_\mu[Z;q,t]$ can be defined [Hai03] as the unique polynomial satisfying the following axioms, where $\lambda \leq \mu$ refers to the dominance order $\lambda_1 + \ldots + \lambda_i \leq \mu_1 + \ldots + \mu_i - 1$ for $1 \leq i \leq n$:

(M1) $\tilde{H}_\mu[Z(q-1)] = \sum_{\lambda \leq \mu'} a_{\lambda,\mu} m_\lambda$ for some $a_{\lambda,\mu} \in \mathbb{Q}(q,t)$

(M2) $\tilde{H}_\mu[Z(t-1)] = \sum_{\lambda \leq \mu} a_{\lambda,\mu} m_\lambda$ for some $a_{\lambda,\mu} \in \mathbb{Q}(q,t)$

(M3) The coefficient of z_1^n in the expansion of $\tilde{H}_\mu[Z;q,t]$ into monomials equals 1.

We call a filling of μ by arbitrary nonzero integers a "super filling". We usually represent negative letters, such as $-2, -4, -7$ by "barred letters" $\overline{2}, \overline{4}, \overline{7}$. Assume the letters satisfy the ordering

$$1 < \overline{1} < 2 < \overline{2} < \cdots < n < \overline{n}. \tag{13}$$

Given a super filling σ, we define the standardization σ' of σ to be the unique permutation satisfying $\sigma'_i < \sigma'_j$ whenever $\sigma_i < \sigma_j$, or whenever $i < j$ and $\sigma_i = \sigma_j$ with σ_i a positive letter, or whenever $i > j$ and $\sigma_i = \sigma_j$ with σ_i a negative letter. We then extend the definition of maj and inv to super fillings by letting $\mathrm{maj}(\sigma,\mu) = \mathrm{maj}(\sigma',\mu)$ and $\mathrm{inv}(\sigma,\mu) = \mathrm{inv}(\sigma',\mu)$. Furthermore, we let $|\sigma|$ be the filling obtained by replacing each σ_i by its absolute value.

Using some general results about quasi-symmetric functions and the superization of a symmetric function derived in [HHL$^+$05c], one easily obtains the fact that

$$\tilde{C}_\mu[Z(q-1);q,t] = \sum_\sigma (-1)^{m(\sigma)} q^{p(\sigma)+\mathrm{inv}(\sigma,\mu)} t^{\mathrm{maj}(\sigma,\mu)} z^{|\sigma|} \tag{14}$$

$$\tilde{C}_\mu[Z(t-1);q,t] = \sum_\sigma (-1)^{m(\sigma)} q^{\mathrm{inv}(\sigma,\mu)} t^{p(\sigma)+\mathrm{maj}(\sigma,\mu)} z^{|\sigma|} \tag{15}$$

$$\tag{16}$$

where the sum is over all super fillings σ of μ, $m(\sigma)$ is the number of negative letters in σ, and $p(\sigma)$ is the number of positive letters in σ.

We say two squares u, v of a Ferrers shape attack each other if they are either in the same row, or if u is one row above v and in a column strictly to the right of v. Otherwise u, v are said to be nonattacking. We call a super filling (σ,μ) nonattacking if for all $u, v \in \mu$, $|\sigma(u)| = |\sigma(v)|$ implies u, v are nonattacking.

The first step in the proof of Theorem 1 involves the construction of a sign-reversing involution on super fillings of μ which cancels most of the terms in (14). The involution looks for an attacking pair of squares containing 1's or $\overline{1}$'s. If more than one such pair exists, it chooses the last such pair in the reading word, and switches the sign on the first element of the pair in the reading word. One checks that the q,t-weights are preserved. If no attacking pairs containing $1,\overline{1}$'s exist, then you search for attacking pairs containing $2,\overline{2}$'s, etc. The fixed points are super fillings with no attacking pairs, which are easily seen to satisfy the triangularity condition (M1).

So far we have assumed (13) holds, but in fact (14) and (15) hold for *any* fixed total ordering of the alphabet of positive and negative letters. We now construct a second involution assuming the ordering

$$1 < 2 < \cdots < n < \overline{n} < \cdots < \overline{2} < \overline{1}. \tag{17}$$

Search for the first occurrence of a 1 or $\overline{1}$ in the reading word, ignoring any such letters in the bottom row. If such a 1, $\overline{1}$ exists, then switch its sign. If there is no such 1 or $\overline{1}$, look for the first occurrence of a 2 or $\overline{2}$, ignoring any letters in the bottom two rows, etc. As in the first involution, the q, t weights are preserved. The fixed points are now super fillings with no 1's or $\overline{1}$'s above the bottom row, no 2's or $\overline{2}$'s above the bottom two rows, etc.. Hence the weight $z^{|\sigma|}$ must satisfy $\lambda_1 \leq \mu_1$, $\lambda_1 + \lambda_2 \leq \mu_1 + \mu_2$, etc., and (M2) is satisfied. The proof is completed by noting that if σ is the filling of all 1's, $\mathrm{inv}(\sigma, \mu) = \mathrm{maj}(\sigma, \mu) = 0$, which implies (M3).

4 The Cocharge Formula

In this section we briefly highlight one of the consequences of Theorem 1, which is described in more detail in [HHL05a]. By considering fillings with no inversions, we obtain the following famous result of Lascoux and Schützenberger, where a semi-standard Young tableau of shape λ and weight μ (or $SSYT(\lambda, \mu)$) is a filling of λ by μ_1-1's, μ_2-2's, etc., with entries that are weakly increasing across rows and strictly increasing up columns.

Theorem 3

$$\tilde{H}_\mu[Z; 0, t] = \sum_\lambda s_\lambda \sum_{T \in SSYT(\lambda, \mu)} t^{cocharge(read(T))}, \tag{18}$$

where the statistic cocharge is described (in the proof) below.

Proof. We begin by describing the set of fillings with no inversions. Let M_i be an arbitrary multiset of μ_i positive integers, for $1 \leq i \leq \mu_1'$, and consider those fillings of μ where the elements of M_i are placed in the ith row of μ, in any arbitrary order, for each i in the range $1 \leq i \leq \mu_1'$. It turns out there is exactly one of these fillings with no inversions. This filling can be constructed by first filling the bottom row of μ with the elements of M_1, in nondecreasing order left to right. Then in square $(2, 1)$, place the smallest integer in M_2 which is strictly bigger than $\sigma_{\mathbf{M}}(1, 1)$, if it exists. If all elements of M_2 are less than or equal to $\sigma_{\mathbf{M}}(1, 1)$, then place the smallest element of M_2 in $(2, 1)$. It is easy to see that this will force any triple of squares of the form $\{(2, 1), (2, j), (1, 1)\}$, $j \geq 2$, to form a clockwise circle in the sense of Remark 1. Next remove $\sigma_{\mathbf{M}}(2, 1)$ from M_2 to form M_2', and place the smallest element of M_2' larger than $\sigma_{\mathbf{M}}(1, 2)$ in square $(2, 2)$, if it exists. If not, place the smallest element of M_2' in square $(2, 2)$. Now iterate the process, moving left to right, in each square of row two placing the smallest remaining element larger than the element in the square just below, if it exists, and otherwise placing the smallest remaining element in the square. The same process is then applied to row 3, comparing elements of M_3 to elements in the second row of $\sigma_{\mathbf{M}}$, then moving on to row 4, etc.. If $\mu = 5531$, $M_1 = \{1, 1, 3, 6, 7\}$, $M_2 = \{1, 2, 4, 4, 5\}$, $M_3 = \{1, 2, 3\}$, and $M_4 = \{2\}$, then $\sigma_{\mathbf{M}}$ is the filling in Figure 2.

2				
3	1	2		
2	4	4	1	5
1	1	3	6	7

Figure 2: A filling with no inversions.

We now construct a word associated to a filling σ we call the cocharge word $\mathrm{cword}(\sigma)$. Initialize $\mathrm{cword}(\sigma)$ to the empty string, then scan through the reading word of σ, starting at the beginning. Whenever a 1 is encountered, adjoin the number of the row containing this 1 to the left of $\mathrm{cword}(\sigma)$. After reaching the end of the reading word, scan through the reading word again, from the beginning, this time looking for 2's. Whenever a 2 is encountered, adjoin the row number to the left of $\mathrm{cword}(\sigma)$ as before. After finishing the 2's, loop back and scan through the reading word for 3's, etc. For example, if σ is the filling in Figure 2, $\mathrm{cword}(\sigma) = 11222132341123$.

Next we translate the statistic $\mathrm{maj}(\sigma, \mu)$ into a statistic on $\mathrm{cword}(\sigma)$. Note that $\sigma(1,1)$ corresponds to the rightmost 1 in $\mathrm{cword}(\sigma)$ - denote this 1 by w_{11}. If $\sigma(2,1) > \sigma(1,1)$, $\sigma(2,1)$ corresponds to the rightmost 2 which is left of w_{11}, otherwise it corresponds to the rightmost 2 (in $\mathrm{cword}(\sigma)$). In any case denote this 2 by w_{12}. More generally, the element in $\mathrm{cword}(\sigma)$ corresponding to $\sigma(i,1)$ is the first i encountered when travelling left from $w_{1,i-1}$, looping around and starting at the beggining of $\mathrm{cword}(\sigma)$ if neccessary. To find the subword $w_{21}w_{22}\cdots w_{2\mu_2'}$ corresponding to the second column of σ, we do the same algorithm on the word $\mathrm{cword}(\sigma)'$ obtained by removing the elements $w_{11}w_{12}\cdots w_{1\mu_1'}$ from $\mathrm{cword}(\sigma)$, then remove $w21w_{22}\cdots w_{2\mu_2'}$ and apply the same process to find $w_{31}w_{32}\cdots w_{3\mu_3'}$ etc..

Clearly $\sigma(i,j) \in \mathrm{Des}(\sigma, \mu)$ if and only if w_{ij} occurrs to the left of $w_{i,j-1}$ in $\mathrm{cword}(\sigma)$. Thus $\mathrm{maj}(\sigma, \mu)$ is transparently equal to the statistic cocharge(cword(σ)) described in [Man01, pp.48-49]. Well-known properties of the RSK algorithm now imply Theorem 3. □

5 Diagonal Harmonics and the n! Conjecture

Assume $\mu \vdash n$, $w_1, \ldots, w_n$ is an arbitrary ordering of the squares of μ, and set

$$\Delta_\mu(x_1, \ldots, x_n, y_1, \ldots, y_n) = \left| x_i^{\mathrm{row}(w_j)} y_i^{\mathrm{col}(w_j)} \right|_{1 \le i,j \le n}. \tag{19}$$

For example,

$$\Delta_{221} = \begin{vmatrix} 1 & y_1 & x_1 & x_1y_1 & x_1^2 \\ 1 & y_2 & x_2 & x_2y_2 & x_2^2 \\ 1 & y_3 & x_3 & x_3y_3 & x_3^2 \\ 1 & y_4 & x_4 & x_4y_4 & x_4^2 \\ 1 & y_5 & x_5 & x_5y_5 & x_5^2. \end{vmatrix} \tag{20}$$

Note that Δ_{1^n} is, up to sign, the Vandermonde determinant in $x_1, \ldots, x_n$.

Next define the Garsia-Haiman module $V(\mu)$ as the linear span over $\mathbb{C}$ of Δ_μ and its partial derivatives of all orders. An element $\sigma = \sigma_1 \cdots \sigma_n$ in the symmetric group S_n acts on a polynomial $f \in V(\mu)$ via the "diagonal action"

$$\sigma f(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(x_{\sigma_1}, \ldots, x_{\sigma_n}, y_{\sigma_1}, \ldots, y_{\sigma_n}). \tag{21}$$

Note that $V(\mu) = \bigoplus_{i,j \ge 0} V(\mu)^{(i,j)}$, where $V(\mu)^{(i,j)}$ is the portion of $V(\mu)$ of bihomogeneous (x,y)-degree (i,j), and that the diagonal action respects this bigrading.

In 2000 Haiman [Hai01] proved the "$n!$ conjecture", first posed in [GH93], which says that the dimension of $V(\mu)$ equals $|\mu|!$. It had previously been shown [Hai99] that the $n!$ conjecture implies the coefficient of $q^i t^j$ in $\tilde{K}_{\lambda,\mu}(q,t)$ equals the multiplicity of the irreducible S_n character χ^λ in the character of $V(\mu)^{(i,j)}$ under the diagonal action, or equivalently that $\tilde{H}_\mu[Z; q, t]$ equals the image of the character of $V(\mu)$ under the Frobenius map which sends χ^λ to s_λ. Macdonald's conjecture that $\tilde{K}_{\lambda,\mu}(q,t) \in \mathbb{N}[q,t]$ follows.

The definition of the statistic $|\mathrm{Inv}(\sigma, \mu)|$ was motivated by the "dinv" statistic occurring in recent work on the combinatorics of the space DH_n of diagonal harmonics [HHL$^+$05c],[HL]. This space is defined [Hai94] as the linear span over $\mathbb{Q}$ of the set of all polynomials $f(x_1, \ldots, x_n, y_1, \ldots, y_n)$ satisfying

$$\sum_{i=1}^{n} \partial^h x_i \, \partial^k y_i \, f = 0, \forall h + k > 0. \tag{22}$$

We can write $DH_n = \bigoplus_{i,j \ge 0} DH_n^{(i,j)}$, and the diagonal action (21) respects this bigrading. The modules $V(\mu)$ are S_n-submodules of DH_n.

Let ∇ be the linear operator on symmetric functions defined on the $\tilde{H}_\mu[Z;q,t]$ basis via

$$\nabla \tilde{H}_\mu[Z;q,t] = t^{n(\mu)} q^{n(\mu')} \tilde{H}_\mu[Z;q,t]. \tag{23}$$

Another famous result of Haiman [Hai02] is that the bigraded character of DH_n under the diagonal action is given by ∇e_n, where e_n is the nth elementary symmetric function. This was first conjectured by Garsia and Haiman in the early 1990's, who made a special study of $< \nabla e_n, s_{1^n} >$, the coefficient of s_{1^n} in the expansion of ∇e_n into Schur functions. They called $< \nabla e_n, s_{1^n} >$ the "q,t-Catalan sequence", denoted $C_n(q,t)$, since they were able to prove that $C_n(1,1) = \binom{2n}{n}/(n+1)$, the nth Catalan number. As in the case of $\tilde{K}_{\lambda,\mu}(q,t)$, all one can infer from its definition is that $C_n(q,t)$ is a complicated sum of rational functions in q,t.

A Dyck path is a lattice path, consisting of north $(0,1)$ and east $(1,0)$ steps, starting at $(0,0)$ and ending at (n,n), which never goes below the diagonal $x = y$. The first major step in the path to the Macdonald polynomial statistics was made in 2000, when Haglund discovered empirically that $C_n(q,t)$ appeared to be expressible as the sum of $q^{\text{area}} t^{\text{bounce}}$ over Dyck paths ifor combinatorial statistics area, bounce [Hag03]. Shortly after, this conjecture was independently discovered by Haiman in the following form, which is more convenient for generalization.

$$C_n(q,t) = \sum_D q^{\text{dinv}(D)} t^{\text{area}(D)}. \tag{24}$$

Here area(D) is the number of complete squares below D but strictly above $y = x$. We let $a_i = a_i(D)$ denote the number of these squares in the ith row, from the bottom of the square $n \times n$ grid, and define dinv(D), the number of "diagonal" inversions of D, as the number of pairs (i,j) with

$$1 \le i < j \le n \text{ and } a_i = a_j \text{ or } a_i = a_j + 1. \tag{25}$$

For example, for the path in Figure 3 (ignore the numbers in the grid for the moment), we have area $=$ $9, \text{dinv} = 13$.

The Haglund-Haiman conjecture for $C_n(q,t)$ was proven shortly after by Garsia and Haglund [GH01], [GH02], by a complicated application of plethystic symmetric function identities previously developed by Garsia, Bergeron and others through a series of papers [GT96], [BGHT99],[GHT99]. More recently Haglund [Hag04b] used the same techniques to prove a more general result conjectured by Egge, Haglund, Kremer and Killpatrick [EHKK03] giving the coefficient of an arbitrary hook shape in ∇e_n in terms of statistics on lattice paths with diagonal $(1,1)$ steps also allowed (known as Schröder paths).

A consequence of Haiman's formula for the character of DH_n is the fact that the dimension of DH_n, as a $\mathbb{Q}$-vector space, is $(n+1)^{n-1}$. This number is well-known to equal the number of parking functions on n cars, which can be represented geometrically by starting with a Dyck path D, then placing the numbers 1 through n immediately to the right of the North steps of D, with strict decrease down columns. Shortly after the proof of the combinatorial formula for $C_n(q,t)$, Haglund and Loehr [HL] conjectured that the Hilbert series of DH_n equals the sum of $q^{\text{dinv}(P)} t^{\text{area}(P)}$ over all parking functions P on n cars. Here the statistic area(P) is simply area(D) for the underlying Dyck path D. If we refer to the number in the j-row as car$_j$, then dinv(P) is the number of pairs (i,j) which satisfy the conditions (25), and in addition, if $a_i = a_j$, then $\text{car}_j > \text{car}_i$, while if $a_i = a_j + 1$, then $\text{car}_i > \text{car}_j$, The Haglund-Loehr conjecture is still open.

The research into the combinatorics of DH_n culminated with the discovery of the "shuffle conjecture" about two years ago by Haglund, et. al [HHL⁺05c]. This gives a formula for ∇e_n in terms of statistics on "word parking functions", which are placements of positive integers in the columns as before, but we allow repeats (but still require strict decrease down columns, and require cars to be no larger than n). A given object P of this type is weighted by $z^P q^{\text{dinv}(P)} t^{\text{area}(P)}$, where area$(P)$ is again the area of the underlying Dyck path. The description of dinv(P) involves the *diagonal reading word* $\sigma = \sigma(P)$ of P, which is obtained by reading in the cars along diagonals (lines parallel to $y = x$), outside to in and top to bottom. Next standardize σ as in Remark 1, regard σ' as the diagonal reading word of some P', then set dinv$(P) = $ dinv(P'). For example, if P is the word parking function in Figure 3, we have $\sigma = 64641532$,

$\sigma' = 74851632$, area $= 9$, and dinv $= 6$, with inversion "pairs" (i, j) of rows equal to $(3, 8)$, $(4, 8)$, $(1, 7)$, $(2, 7)$, $(5, 6)$ and $(3, 4)$ each contributing 1 to dinv. Also, z^P is defined as z^σ.

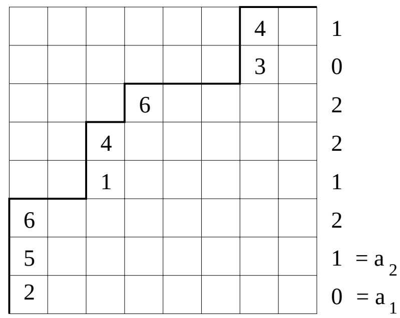


Figure 3: A word parking function with the a_i on the right

Conjecture 1 *(The shuffle conjecture - [HHL+05c])*

$$\nabla e_n = \sum_P z^P q^{dinv(P)} t^{area(P)}, \tag{26}$$

where the sum is over all word parking functions P with n cars.

An equivalent formulation of the conjecture is that the coefficient of the monomial $z_1^{\lambda_1} z_2^{\lambda_2} \cdots$ in ∇e_n equals the sum of $q^{\mathrm{dinv}} t^{\mathrm{area}}$ over all parking functions whose diagonal reading word is a shuffle of increasing sequences of lengths $\lambda_1, \lambda_2, \ldots$, hence the phrase "shuffle conjecture". As is well known, the Hilbert series is the coefficient of $z_1 \cdots z_n$ in the expansion of the character into monomials, hence the shuffle conjecture contains Haglund and Loehr's conjectured formula for the Hilbert series of DH_n as a special case. In [HL] it is shown that it also implies the formula for $C_n(q, t)$, as well as the formula for hook shapes in terms of Schröder paths.

6 Water on Mars

Recall that $\tilde{H}_\mu[Z; q, t]$ is the character of the S_n module $V(\mu)$, which is a submodule of DH_n. Thus we could hope that some version of the shuffle conjecture would hold for $\tilde{H}_\mu[Z; q, t]$. Furthermore, since the statistics in the shuffle conjecture are defined solely in terms of the statistics for the Hilbert series of DH_n, together with the simple operation of standardization, we could hope that discovering statistics for the Hilbert series of $V(\mu)$ would result in a combinatorial formula for the monomial expansion of $\tilde{H}_\mu[Z; q, t]$ through a similar standardization process.

Let $\mathrm{Hilb}_\mu(q, t)$ denote the bigraded Hilbert series of $V(\mu)$. The problem of finding combinatorial statistics for $\mathrm{Hilb}_\mu(q, t)$ was studied by Garsia and Haiman, who obtained statistics when μ has two rows or is a hook shape [GH95]. The author had also made some previous unsuccessful attempts at this problem, but the reasoning in the above paragraph increased the stakes dramatically, and so the author decided to give a more determined attack on the problem.

Since the dimension of $V(\mu)$ is $n!$, it is natural to view the problem in terms of searching for a pair of permutation statistics, which depend on μ, to generate $\mathrm{Hilb}_\mu(q, t)$. Since μ has n squares, permutations are in bijection with fillings of μ by distinct integers, and furthermore from a result of Macdonald [Mac95, p.365,Ex.6] one can easily deduce that

$$\mathrm{Hilb}_\mu(1, t) = \sum_{\sigma \in S_n} t^{\mathrm{maj}(\sigma, \mu)}. \tag{27}$$

We now assume that there exists some q-statistic to match with this maj t-statistic, and try to find it.

For a given Dyck path D, let $t^{\text{area}(D)}F_D(Z;q)$ denote the restriction of the sum on the right-hand-side of (26) to those word parking functions for D. It is proven in [HHL+05c] that $F_D(Z;q)$ is a symmetric function (and morevover is a constant power of q times an LLT polynomial, where the elements of the tuple are vertical columns associated with the lengths of the vertical segments of D). Let μ be a partition, and abbreviate μ_1 by ℓ. While working with the F_D, the author proved that if $D = D(\mu)$ is the special type of path consisting of μ'_ℓ vertical steps, followed by μ'_ℓ horizontal steps, followed by $\mu'_{\ell-1}$ vertical steps, followed by $\mu'_{\ell-1}$ horizontal steps, etc., then

$$F_{D(\mu)}(Z;q) = q^{\text{mindinv}(D(\mu))} \sum_{\lambda \vdash n} s_\lambda K_{\lambda',\mu'}(q), \tag{28}$$

where $K_{\lambda,\mu}(q) = q^{\eta(\mu)}\tilde{K}_{\lambda,\mu}(q^{-1})$ is the charge version of the polynomial from Section 4, and mindinv(D) equals the minimum value of dinv over all word parking functions for D. This can be proven from recurrences for the $F_{D(\mu)}$, obtained by placing the largest car n at the top of columns, then showing these recurrences are equivalent to those obtained by Garsia and Procesi [GP92] for the Hall-Littlewood polynomials. (An equivalent result is contained in [SSW03], where it is deduced from recurrences of this type for general LLT polynomials).

Since

$$\tilde{H}_\mu[Z;0,q] = \sum_\lambda s_\lambda \tilde{K}_{\lambda,\mu}(q), \tag{29}$$

from the first it seemed that any statistics for the $\tilde{H}_\mu[Z;q,t]$ had to be connected to cocharge in some way. Now by (28) the monomial expansion of

$$\sum_{\lambda \vdash n} s_\lambda K_{\lambda',\mu'}(q) \tag{30}$$

can be obtained by summing $q^{-\text{mindinv}(D(\mu))+\text{dinv}(W)}z^W$ over all word parking functions W for $D(\mu)$, so one might suspect that a power of q related to dinv is the right thing to insert into (27). A parking function P for $D(\mu)$ can be transformed into a filling (σ,μ) by pushing all squares in the ith column from the left down $i-1$ squares, removing all empty columns, and finally rotating 180 degrees, as in Figure 4. Furthermore this sends dinv(σ) to $|Inv(\sigma,\mu)|$. Note that area gets sent to maj, since we have descents everywhere. We are thus led to

$$\sum_{\sigma \in S_n} t^{\text{maj}(\sigma,\mu)} q^{-\text{mindinv}(D(\mu))+|Inv(\sigma,\mu)|} \tag{31}$$

as a candidate for Hilb$_\mu(q,t)$.

Maple computations verified that (31) correctly predicts the coefficient of $t^{\eta(\mu)}$ (the highest power of t) in Hilb$_\mu(q,t)$, and that the coefficient of t^0 would be correct if we didn't have the $-\text{mindinv}(D((\mu))$ in the q-power. This led to the hypothesis that $|Inv(\sigma,\mu)|$ formed an upper bound for the correct q-statistic, and that the key was to find the right thing to subtract from it, something possibly depending on the descent set. The final stage involved a series of Maple calculations, each testing a different idea for something to subtract, combined with various maj-like choices for the t-statistic. Many things worked up through $n=4$ and for various special cases like $\mu = 2,1,1,1,...$, but continually failed for $n=5$. Here is a sample of the author's daily journal summarizing the experiments:

Journal entry:

On 3/29/04, ran pistolbad.map, which computes inversion in columns for tstat and dinv minus mindinv for qstat. Bombs for $\mu = [2,2,1]$. See bomb.out Tried a modification, where comaj of the columns replaced the tstat, same qstat. This also bombed in general, but surprisingly works for $\mu = 221$, $\mu = 222$, $\mu = 2221$, and $\mu = 2222$. Bombs for $\mu = 2211$ though.

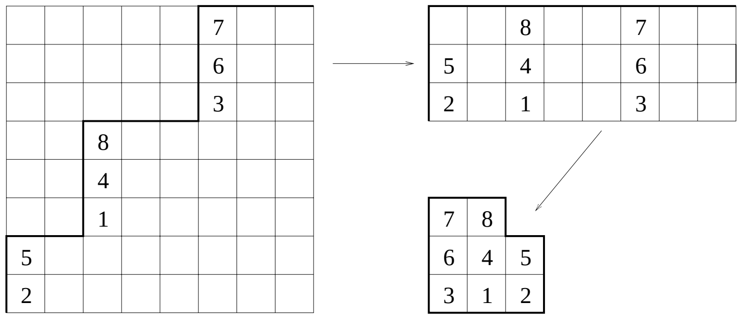


Figure 4: Transforming a parking function into a filling

At this point the author planned to drop the problem, and return to less speculative projects. One last idea occurred to him though, namely that mindinv($D(\mu)$) equals the sum of the arm lengths of the squares not in the bottom row of μ, which led to the idea of subtracting the arm of each descent from $|Inv(\sigma, \mu)|$. A computer run incorporating this idea then successfully generated $\mathrm{Hilb}_\mu(q,t)$ for all $|\mu| \leq 8$. A subsequent calculations a few days later extended this to $|\mu| \leq 9$, and morevover verified that applying the Hilbert series statistics to the standardization of words generated the entire monomial expansion of $\tilde{H}_\mu[Z;q,t]$. In a phone conversation a few months later, after the conjecture was made public, A. Garsia told the author “You found water on Mars”.

Journal Entry:

On 4/2/04 ran pistolbbad.map, which computes inversions in columns for tstat and dinv minus mindinv for qstat, where mindinv is defined as the sum over all descents, of the number of entries to the right of the top element of the decent pair. Seems to work for μ having two rows, but bombs for $\mu = 2, 2, 1$.

Journal Entry:

Also on 4/2/04, ran pistolb.map, which computes maj on columns for tstat, and uses mindinv as described in paragraph just above. Surprisingly, works for all μ with $|\mu| <= 8$! The run for $n = 9$ took over a week, but finally finished successfully.

Journal Entry:

On 4/6/04 ran program pistolc.map, which computes the Macdonald poly from shuffles using the hilbert series inv and maj stats. Correctly generates the Macdonald poly $\tilde{H}_\mu$ for all μ with $|\mu| \leq 8$! From 4/7 thru 4/16 ran tempc.map, which also verifies shuffle conjecture for n=9.

References

[BGHT99] F. Bergeron, A. M. Garsia, M. Haiman, and G. Tesler. Identities and positivity conjectures for some remarkable operators in the theory of symmetric functions. *Methods Appl. Anal.*, 6:363–420, 1999.

[EHKK03] E. Egge, J. Haglund, K. Killpatrick, and D. Kremer. A Schroder generalization of Haglund’s statistic on Catalan paths. *Electron. J. Combin.*, 10:Research Paper 16, 21 pp. (electronic), 2003.

[GH93] A. M. Garsia and M. Haiman. A graded representation model for Macdonald polynomials. *Proc. Nat. Acad. Sci. U.S.A.*, 90(8):3607–3610, 1993.

[GH95] A. M. Garsia and M. Haiman. Factorizations of Pieri rules for Macdonald polynomials. *Discrete Math.*, 139(1-3):219–256, 1995. Formal power series and algebraic combinatorics (Montreal, PQ, 1992).

[GH96] A. M. Garsia and M. Haiman. A remarkable q,t-Catalan sequence and q-Lagrange inversion. *J. Algebraic Combin.*, 5(3):191–244, 1996.

[GH01] A. M. Garsia and J. Haglund. A positivity result in the theory of Macdonald polynomials. *Proc. Nat. Acad. Sci. U.S.A.*, 98:4313–4316, 2001.

[GH02] A. M. Garsia and J. Haglund. A proof of the q,t-Catalan positivity conjecture. *Discrete Math.*, 256:677–717, 2002.

[GHT99] A. M. Garsia, M. Haiman, and G. Tesler. Explicit plethystic formulas for Macdonald q,t-Kostka coefficients. *Sém. Lothar. Combin.*, 42:Art. B42m, 45 pp. (electronic), 1999. The Andrews Festschrift (Maratea, 1998).

[GP92] A. M. Garsia and C. Procesi. On certain graded S_n-modules and the q-Kostka polynomials. *Adv. Math.*, 94(1):82–138, 1992.

[GT96] A. M. Garsia and G. Tesler. Plethystic formulas for Macdonald q,t-Kostka coefficients. *Adv. Math.*, 123:144–222, 1996.

[Hag03] J. Haglund. Conjectured statistics for the q,t-Catalan numbers. *Adv. Math.*, 175(2):319–334, 2003.

[Hag04a] J. Haglund. A combinatorial model for the Macdonald polynomials. *Proc. Nat. Acad. Sci. U.S.A.*, 101:16127–16131, 2004.

[Hag04b] J. Haglund. A proof of the q,t-Schroder conjecture. *Internat. Math. Res. Notices*, 11:525–560, 2004.

[Hai94] M. Haiman. Conjectures on the quotient ring by diagonal invariants. *J. Algebraic Combin.*, 3:17–76, 1994.

[Hai99] M. Haiman. Macdonald polynomials and geometry. In *New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97)*, volume 38 of *Math. Sci. Res. Inst. Publ.*, pages 207–254. Cambridge Univ. Press, Cambridge, 1999.

[Hai01] M. Haiman. Hilbert schemes, polygraphs, and the Macdonald positivity conjecture. *J. Amer. Math. Soc.*, 14:941–1006, 2001.

[Hai02] M. Haiman. Vanishing theorems and character formulas for the Hilbert scheme of points in the plane. *Invent. Math.*, 149:371–407, 2002.

[Hai03] Mark Haiman. Combinatorics, symmetric functions, and Hilbert schemes. In *Current developments in mathematics, 2002*, pages 39–111. Int. Press, Somerville, MA, 2003.

[HHL05a] J. Haglund, M. Haiman, and N. Loehr. A combinatorial formula for Macdonald polynomials. *Jour. Amer. Math. Soc., to appear*, 2005.

[HHL05b] J. Haglund, M. Haiman, and N. Loehr. Combinatorial theory of Macdonald polynomials I: Proof of Haglund's formula. *Proc. Nat. Acad. Sci. U.S.A.*, 102(8):2690–2696, 2005.

[HHL$^+$05c] J. Haglund, M. Haiman, N. Loehr, J. B. Remmel, and A. Ulyanov. A combinatorial formula for the character of the diagonal coinvariants. *Duke J. Math.*, 126:195–232, 2005.

[HL] J. Haglund and N. Loehr. A conjectured combinatorial formula for the Hilbert series for Diagonal Harmonics. *Discrete Math., to appear.* Proceedings of the FPSAC 2002 Conference, preprint available at www.math.upenn.edu/∼jhaglund/.

[KS97] Friedrich Knop and Siddhartha Sahi. A recursion and a combinatorial formula for Jack polynomials. *Invent. Math.*, 128(1):9–22, 1997.

[LLT97] Alain Lascoux, Bernard Leclerc, and Jean-Yves Thibon. Ribbon tableaux, Hall-Littlewood functions, quantum affine algebras, and unipotent varieties. *J. Math. Phys.*, 38(2):1041–1068, 1997.

[Mac95] I. G. Macdonald. *Symmetric Functions and Hall Polynomials.* Oxford Mathematical Monographs, second ed., Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1995.

[Man01] Laurent Manivel. *Symmetric functions, Schubert polynomials and degeneracy loci*, volume 6 of *SMF/AMS Texts and Monographs.* American Mathematical Society, Providence, RI, 2001. Translated from the 1998 French original by John R. Swallow, Cours Spécialisés [Specialized Courses], 3.

[SSW03] A. Schilling, M. Shimozono, and D. E. White. Branching formula for q-Littlewood-Richardson coefficients. *Adv. in Appl. Math.*, 30(1-2):258–272, 2003. Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).

COMBINATORIAL DEFORMATIONS OF THE FULL TRANSFORMATION SEMIGROUP

JOHN HALL

Abstract. We define two deformations of the Full Transformation Semigroup algebra. One makes the algebra "as semisimple as possible", while another leads to an eigenvalue result involving Schur functions.

Preliminaries

The *Full Transformation Semigroup on n letters*, denoted T_n, is the semigroup of all set maps $w : [n] \to [n]$, where $[n] = \{1, 2, \ldots, n\}$ and the multiplication is the usual composition. Such maps can be depicted in several ways; we will most often use one-line notation, for example $w = 214442$ denotes the map sending 1 to 2, 2 to 1, 3 to 4, etc.

Maps in T_n are indexed by triples (π, P, ϕ), where P is the image of the map, π is the set partition of $[n]$ whose blocks are the inverse images of the elements of P, and ϕ is the permutation describing which block is mapped to which element of the image. In what follows, $\pi = \{\pi_1, \pi_2, \ldots\}$ will always denote a set partition of $[n]$ with blocks ordered by increasing smallest element. Similarly in writing $P = \{p_1, p_2, \ldots\}$ a subset of $[n]$ we shall always intend $p_1 < p_2 < \ldots$. Permutations will be written in cycle notation.

With these conventions, we shall let $w_{\pi,P,\phi} \in T_n$ denote the map taking $x \in \pi_i$ to $p_{\phi(i)}$.

Example 1. *For* $w = 214442 \in T_6$ *we have* $\pi(w) = 16|2|345, P(w) = \{1, 2, 4\}$, *and* $\phi(w) = (12)(3)$, *the transposition exchanging* 1 *and* 2 *and fixing* 3. *The permutation* ϕ *is most easily visualized in the following diagram of* w.

$$\begin{tikzcd}[row sep=small]
\pi\text{:} & 16 \arrow[dr, dash] & 2 \arrow[dl, dash] & 345 \arrow[d, dash] \\
P\text{:} & 1 & 2 & 4
\end{tikzcd}$$

The invertible elements of T_n, i.e., the bijective maps, form a subsemigroup isomorphic to the Symmetric Group S_n. Thus the elements of T_n can be thought of as generalized permutations, and we can ask which of the many combinatorial aspects of the Symmetric Group can be extended in a meaningful way to the Full Transformation Semigroup.

Let $\mathbb{C}T_n$ denote the Full Transformation Semigroup algebra, consisting of complex linear combinations of elements of T_n. $\mathbb{C}T_n$ has a chain of two-sided ideals

$$\mathbb{C}T_n = I_n \supseteq I_{n-1} \supseteq \ldots \supseteq I_1 \supseteq I_0 = 0,$$

where for $1 \le k \le n$, I_k as a vector space is the complex span of the maps of rank less than or equal to k (the *rank* of a map is the cardinality of its image). For

Date: May 3, 2005.

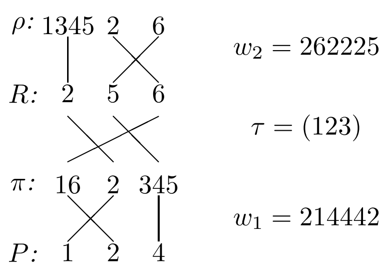
$1 \le k \le n$ define the algebras $A_{n,k} = I_k/I_{k-1}$. We can think of $A_{n,k}$ as being the algebra spanned by the maps of rank k, where two maps multiply to zero if their composition has rank less than k. The top quotient $A_{n,n}$ is isomorphic to $\mathbb{C}S_n$, the group algebra of the Symmetric Group, and is therefore semisimple. However $A_{n,k}$ is not semisimple for $k < n$, meaning that the radical $\sqrt{A_{n,k}}$ is non-trivial.

It is known that the irreducible modules for $A_{n,k}$ are indexed by partitions $\lambda \vdash k$. In fact Hewitt and Zuckerman give a calculation in [5] that generates all irreducible matrix representations for $A_{n,k}$. However, their methods are difficult to apply in practice and do not even determine the dimensions of the representations. These dimensions are known, thanks to a more recent character result of Putcha [8]. Regardless of the approach, it is clear that the non-semisimplicity of $A_{n,k}$ causes great difficulties. This has led us to define several deformations of $A_{n,k}$, with the aim of making the algebra generically semisimple.

The first deformation

Let $w_1 = w_{\pi,P,\phi}$ and $w_2 = w_{\rho,R,\psi}$ be two maps of rank k. Notice that in order for the product $w_1 w_2$ to be nonzero in $A_{n,k}$ each element of $R = P(w_2)$ must lie in a different block of $\pi = \pi(w_1)$. In this situation we can associate to the maps w_1 and w_2 the permutation $\tau \in S_k$ defined by the condition $r_i \in \pi_{\tau(i)}$.

Example 2. *Let* $w_1 = 214442$ *and* $w_2 = 262225$. *Then* $P(w_2) = \{2, 5, 6\}$ *and* $\pi(w_1) = 16|2|345$. *The smallest element* 2 *of* $P(w_2)$ *is in the second block of* $\pi(w_1)$, *the next-smallest element* 5 *is in the third block, and the largest* 6 *is in the first block. Thus* $\tau = (123)$, *as can be seen in the following diagram.*



Now define a new multiplication in $A_{n,k}$ by

$$w_1 * w_2 := x^{\mathrm{inv}(\tau)} w_1 w_2,$$

where $\mathrm{inv}(\tau)$ is the number of inversions of τ. It is not difficult to show that this multiplication is associative.

Example 3. *Taking* w_1 *and* w_2 *as above we have*

$$w_1 * w_2 = x^2 w_1 w_2 = x^2 121114.$$

Let $A_{n,k}(x)$ denote the algebra with the multiplication $*$. Setting $x = 1$ recovers the original multiplication in $A_{n,k}$. As we shall see, there is a sense in which $A_{n,k}(x)$ is "as semisimple as possible" for generic x.

Definition 1. *Let* A *be an associative algebra. The* Munn matrix algebra $\mathcal{A} = \mathcal{M}(A; m, n; \Pi)$ *as a vector space is the set of all* $m \times n$ *matrices with entries in* A. Π *is an* $n \times m$ *matrix over* A, *called the* sandwich matrix, *and multiplication is defined by* $X \cdot Y := X \Pi Y$.

Fact 1. *$A_{n,k}$ is isomorphic to the Munn matrix algebra $\mathcal{M}(\mathbb{C}S_k; \binom{n}{k}, S(n,k); \Pi_{n,k})$, where $\Pi_{n,k}$ is the $S(n,k) \times \binom{n}{k}$ sandwich matrix*

$$(\Pi_{n,k})_{\pi,P} = \begin{cases} \tau & \textit{if } p_i \in \pi_{\tau(i)}, 1 \le i \le k, \textit{ and} \\ 0 & \textit{otherwise.} \end{cases}$$

Example 4. *For $n = 4$ and $k = 3$ we have*

$$\Pi_{4,3} = \begin{pmatrix} (id) & (id) & 0 & 0 \\ 0 & (id) & (id) & 0 \\ 0 & 0 & (id) & (id) \\ (id) & 0 & 0 & (123) \\ (id) & 0 & (23) & 0 \\ 0 & (id) & 0 & (12) \end{pmatrix},$$

where the ordering on the columns is $123, 124, 134, 234$, and the ordering on the rows is $1|2|34, 1|23|4, 12|3|4, 14|2|3, 1|24|3, 13|2|4$.

Note that the non-zero entries of $\Pi_{n,k}$ are precisely the permutations τ that arise in the $*$ multiplication. Hence the first deformation preserves the Munn matrix algebra structure.

Proposition 1. *$A_{n,k}(x)$ is isomorphic to the Munn matrix algebra $\mathcal{M}(\mathbb{C}S_k; \binom{n}{k}, S(n,k); \Pi_{n,k}(x))$, where $\Pi_{n,k}(x)$ is the $S(n,k) \times \binom{n}{k}$ sandwich matrix defined by*

$$(\Pi_{n,k}(x))_{\pi,P} = \begin{cases} x^{\mathrm{inv}(\tau)}\tau & \textit{if } p_i \in \pi_{\tau(i)}, 1 \le i \le k, \textit{ and} \\ 0 & \textit{otherwise} \end{cases}$$

When $k = 1$ the parameter x does not show up at all, and the semisimple part of $A_{n,1}$ is only one-dimensional. For the remainder of this section we shall assume $k > 1$.

Since we want our Munn matrix algebra to be semisimple, it is natural to ask in what way the semisimplicity of $\mathcal{A}$ depends on the sandwich matrix Π. The following result of Clifford and Preston ([2], Theorem 5.19) provides an answer.

Theorem 1. *(Clifford and Preston) A Munn matrix algebra of the form $\mathcal{A} = \mathcal{M}(\mathbb{C}G; m, n; \Pi)$ is semisimple if and only if Π is non-singular, i.e., if and only if $m = n$ and Π is a unit in the ring of $m \times m$ matrices over $\mathbb{C}G$.*

Note in particular that for semisimplicity we need the matrices to be square. But our sandwich matrix is $S(n,k) \times \binom{n}{k}$. What can we do?

One idea is to define the *rank* of Π to be the largest non-singular minor of Π. (So, in particular, $\mathrm{rank}(\Pi) \le \min(m,n)$.) A result of McAlister [7] implies that this rank is intimately related to the size of $\sqrt{\mathcal{A}}$. To state McAlister's result we first need to define a technical condition known as *suitability*.

Definition 2. *Let P be an $n \times m$ matrix over A with rank r. Let R and S be permutation matrices over A such that*

$$RPS = \begin{pmatrix} M & P_{12} \\ P_{21} & P_{22} \end{pmatrix},$$

where M is an invertible $r \times r$ submatrix of P, and let

$$Q = S \begin{pmatrix} M^{-1} & 0 \\ 0 & 0 \end{pmatrix} R.$$

Then we say that P is suitable *if $PQP - P \in (\sqrt{A})_{n\times m}$.*

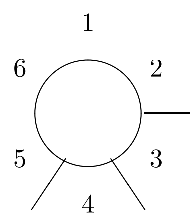
We note here in passing that the suitability condition is trivially satisfied when a matrix has full rank.

Theorem 2. *(McAlister) Let $\mathcal{A} = \mathcal{M}(A; m, n; \Pi)$ be a Munn matrix algebra, and let Π be suitable of rank r. Then $\mathcal{A}/\sqrt{\mathcal{A}} \cong \left(A/\sqrt{A}\right)_r$, the algebra of all $r \times r$ matrices with entries in $A/\sqrt{A}$.*

The suitability condition does not hold for the undeformed algebra $A_{n,k}$. But what about for $A_{n,k}(x)$? We show that $A_{n,k}(x)$ has full rank for generic x, by considering the submatrix formed by the rows corresponding to a special set of $\binom{n}{k}$ partitions of $[n]$.

Definition 3. *A set partition π of $[n]$ is* cyclically contiguous *if the blocks of π are intervals, with the possible exception of the first block, which may be of the form $\{1, 2, \ldots, i\} \cup \{j, j+1, \ldots, n\}$, i.e., the union of an initial segment and a terminal segment. If the first block is also an interval, we say that π is* contiguous. *(Note that contiguous implies cyclically contiguous, not the other way around.)*

We use the term "cyclically contiguous" for such a partition π because if we think of the elements of $[n]$ as being arranged in (clockwise) order around a circle, then in a sense *all* of the blocks of π are intervals. For example, $\pi = 1256|3|4$ is cyclically contiguous, as shown in the following diagram.



$\pi = 1256|3|4$

For $k > 1$ there is an obvious bijection between cyclically contiguous partitions of $[n]$ into k blocks and k-subsets of $[n]$. Let $\Pi^c(x)$ be the $\binom{n}{k} \times \binom{n}{k}$ submatrix of $\Pi(x)$ consisting of the rows corresponding to cyclically contiguous partitions. We show that $\Pi^c(x)$ is nonsingular for generic x, and hence $\Pi(x)$ is suitable of rank $\binom{n}{k}$. Thus we have

Theorem 3. *For $k > 1$ and generic x, $A_{n,k}(x)/\sqrt{A_{n,k}(x)} \cong (\mathbb{C}S_k)_{\binom{n}{k}}$.*

Note that the rank of Π cannot be any larger than its width $\binom{n}{k}$. So no deformation of $A_{n,k}$ that preserves the Munn matrix algebra structure can be any more semisimple than $A_{n,k}(x)$.

Corollary 1. *For $k > 1$,*

$$\dim\left(\sqrt{A_{n,k}(x)}\right) = \left(S(n,k) - \binom{n}{k}\right)\binom{n}{k}k!.$$

Since this dimension formula is relatively simple one might hope to find a nice combinatorial basis for $\sqrt{A_{n,k}(x)}$, as Garsia and Reutenauer did in [3] for Solomon's Descent Algebra. So far we have found several families of elements in the radical, but they are not in general independent, and do not form a spanning set.

Let $A^c_{n,k}(x)$ be the subalgebra of $A_{n,k}(x)$ spanned by the maps whose partitions are cyclically contiguous. $A^c_{n,k}(x)$ is also a Munn matrix algebra, with sandwich matrix $\Pi^c(x)$.

Grood [4] has generalized the classical Specht-module construction for the symmetric group (see [6]) to describe the irreducible modules of the *rook monoid* R_n, another semigroup containing the symmetric group. As it turns out, $\mathbb{C}R_n$ has a similar tower of ideals

$$\mathbb{C}R_n = J_n \supseteq J_{n-1} \supseteq \ldots \supseteq J_1 \supseteq J_0 = 0,$$

and defining $B_{n,k} = J_k/J_{k-1}$ we have $B_{n,k} \cong (\mathbb{C}S_k)_{\binom{n}{k}}$. We have extended the first deformation to an algebra $A_{n,k}(x, \mathbf{y})$ with canonical isomorphisms $A^c_{n,k}(1, \mathbf{1}) \cong A^c_{n,k}$ and $A^c_{n,k}(1, \mathbf{0}) \cong B_{n,k}$. We are currently attempting to modify Grood's approach to explicitly construct the irreducible modules for the generic algebra $A^c_{n,k}(x, \mathbf{y})$.

The second deformation

There is another associative multiplication we can define on $A_{n,k}$. The symmetric group S_k acts on the maps of rank k by

$$\sigma w_{\pi,P,\phi} := w_{\pi,P,\sigma\phi}.$$

Now define

$$w_1 \circ w_2 := \sum_{\sigma \in S_k} p_{\rho(\sigma)} \sigma w_1 w_2,$$

where $\rho(\sigma) \vdash k$ is the cycle type of σ, and $p_{\rho(\sigma)}$ is the corresponding power-sum symmetric function in the variables $x_1, \ldots, x_k$.

Example 5. *Let* $w_1 = 1442$ *and* $w_2 = 3134$. *Then* $w_1 w_2 = 4142$, *and*

$$\begin{aligned} w_1 \circ w_2 &= p_{1^3} 4142 + p_{21}(1412 + 2124 + 4241) + p_3(1214 + 2421) \\ &= (x_1 + x_2 + x_3)^3 4142 + (x_1^2 + x_2^2 + x_3^2)(x_1 + x_2 + x_3)(1412 + 2124 + 4241) \\ &\qquad + (x_1^3 + x_2^3 + x_3^3)(1214 + 2421) \end{aligned}$$

Note that if we choose values for the x_i so that $p_1 = 1$ and $p_i = 0$ for all $i \geq 2$, we recover the original multiplication in $A_{n,k}$. Such a specialization of the x_i must exist because the p_i are algebraically independent.

Let $A_{n,k}(\mathbf{x})$ denote the algebra with the multiplication $\circ$. Explicit calculations for small values of n and k suggest that $A_{n,k}(\mathbf{x})$ is no more semisimple than $A_{n,k}$, i.e., that even for generic values of the x_i we have $\dim\sqrt{A_{n,k}(\mathbf{x})} = \dim\sqrt{A_{n,k}}$. However something interesting does come out of this multiplication.

The following fact gives a useful characterization of the radical of an algebra.

Fact 2. *Let A be a finite-dimensional associative algebra and $\{v_1, \ldots, v_n\}$ a basis of A. Identify A as a vector space with $\mathbb{C}^n$, and define the $n \times n$ Gram matrix M for A by $(M)_{i,j} = tr(v_i v_j)$. Then the nullspace of M is $\sqrt{A}$.*

If we define $M_{n,k}(\mathbf{x})$ to be the Gram matrix for $A_{n,k}(\mathbf{x})$ we have

Proposition 2.

$$(M_{n,k}(\mathbf{x}))_{i,j} = \begin{cases} S(n,k)k! \sum_{\lambda \vdash k} \frac{k!}{f^\lambda} \chi^\lambda(\mu) s_\lambda^2 & w_i w_j \text{ induces a permutation of cycle type } \mu \text{ on the image of } w_i \\ 0 & \text{otherwise.} \end{cases}$$

Here for λ a partition of k, χ^λ is the corresponding irreducible character of S_k, $f^\lambda = \chi^\lambda(1)$ its dimension, and s_λ the associated Schur function.

In the sequel we will always normalize the Gram matrix $M_{n,k}(\mathbf{x})$ by dividing by the constant $S(n,k)k!$.

In the semisimple case $k = n$ one can use Frobenius' factorization of the group determinant (see [1]) to derive the following result about the normalized matrix $M_{n,n}(\mathbf{x})$.

Theorem 4. *The eigenvalues of $M_{n,n}(\mathbf{x})$ are $\pm\left(\frac{n!}{f^\lambda}s_\lambda\right)^2$, $\lambda \vdash n$, where the positive values appear with multiplicity $\binom{f^\lambda+1}{2}$ and the negative values appear with multiplicity $\binom{f^\lambda}{2}$.*

Corollary 2. *The algebra $A_{n,n}(\mathbf{x})$ is semisimple if and only if the values of the parameters x_i avoid the zeros of the Schur functions s_λ.*

For $k = 1$ there is always a unique non-zero eigenvalue ns_1^2, but the analogous result for $1 < k < n$ is so far only conjectural.

Conjecture 1. *The non-zero eigenvalues of $M_{n,k}(\mathbf{x})$, $k < n$, are of the form cs_λ^2 for $\lambda \vdash k$, where c is an algebraic scalar. The "multiplicity" of s_λ^2, i.e., the sum of the multiplicities of the cs_λ^2, is $(\binom{n}{k}f^\lambda)^2$ for $\lambda \vdash k$, $\lambda \neq 1^k$, and $\binom{n-1}{k-1}^2$ for $\lambda = 1^k$.*

This conjecture is difficult to check even by computer for $n > 4$. The following table gives some sample data.

n	k	eigenvalues
3	2	$(-12s_{1^2}^2)^1, (-8s_2^2)^2, (-4s_2^2)^1, 0^5, (4s_2^2)^3, (8s_2^2)^2, (12s_{1^2}^2)^3, (16s_2^2)^1$
4	2	$(-32s_{1^2}^2)^3, (-\sqrt{448}s_2^2)^5, (-8s_2^2)^{10}, 0^{39}, (8s_2^2)^{15}, (\sqrt{448}s_2^2)^5, (32s_{1^2}^2)^6, (56s_2^2)^1$

(The conjectured multiplicities come from Putcha's results; they are the dimensions of the irreducible characters for $A_{n,k}$.)

References

[1] C. Curtis, Pioneers of representation theory: Frobenius, Burnside, Schur, and Brauer, History of Mathematics, **15**, Amer. Math. Soc., 1999.

[2] A. H. Clifford and G. P. Preston The Algebraic Theory of Semigroups, Vol. I, Amer. Math. Soc., 1961

[3] A. M. Garsia and C. Reutenauer, *A decomposition of Solomon's descent algebra*, Adv. Math. **77** (1989), no. 2, 189–262.

[4] C. Grood, *A Specht module analog for the rook monoid*, Electron. J. Combin. **9** (2002), #R2.

[5] E. Hewitt and H. Zuckerman, *The irreducible representations of a semigroup related to the symmetric group*, Illinois J. Math. **1** (1957), 188–213.

[6] G. D. James, The Representation Theory of the Symmetric Groups, Lecture Notes in Mathematics, **682**, Springer-Verlag, 1978.

[7] D. B. McAlister, *Rings Related to Completely 0-Simple Semigroups*, J. Austral. Math. Soc. **12** (1971), 257–274.

[8] M. Putcha, *Complex representations of finite monoids*, Pro. London Math. Soc. **73** (1996), no. 3, 623–641.

School of Mathematics, University of Minnesota, Minneapolis, MN 55455

E-mail address: jhall@math.umn.edu

LOCAL ACTION OF THE SYMMETRIC GROUP AND GENERALIZATIONS OF QUASI-SYMMETRIC FUNCTIONS

EXTENDED ABSTRACT

FLORENT HIVERT

Abstract. This paper is a tentative to better understand the result of Garsia and Wallach showing that the ring of quasi-symmetric functions is free as a module over the ring of symmetric functions. We conjecture that their result is a particular case of a more general situation whose construction is decribed here.

We consider a certain class of actions of the symmetric group $\mathfrak{S}_n$ on polynomials $\mathbb{K}[X_n] = \mathbb{K}[x_1, \ldots, x_n]$ called *local actions*, which are very similar to the one of [6, 7]. After classifying these actions, we study the sub-class of these actions whose set of fixed polynomial is a sub-algebra of $\mathbb{K}[X_n]$. This gives rise to an infinite hierarchy of sub-Hopf-algebras of QSym, interpolating between QSym and Sym. We conjecture that these algebra are free modules over Sym as is suggested by an explicit formula giving the Hilbert series of the quotient.

Contents

Date: Id:GenQSym.tex,v 1.4 2004/09/26 17:21:15 averell Exp .

1. Introduction

The symmetric polynomials Sym are by definition the polynomials of $\mathbb{K}[X_n] := \mathbb{K}[x_1, \dots, x_n]$ which are invariant under the action of the symmetric group $\mathfrak{S}_n$ on polynomials. They form a sub-algebra Sym $:= \mathbb{K}[X_n]^{\mathfrak{S}_n}$ of $\mathbb{K}[X_n]$, and consequently, polynomials can be considered as a module over Sym. It is well known that this module is free of rank $n!$ and this result is now widely interpreted for example in the cohomology or the Grothendieck ring of the flag manifold.

In his study of descents of permutations, Gessel defined a generalization of the notion of symmetric functions namely the quasi-symmetric functions QSym [4]. This algebra as been extensively studied: it can also be endowed with a natural Hopf algebra structure which is dual to the Hopf algebra **NCSF** of noncommutative symmetric functions [11, 3]. Quasi-symmetric functions are defined as the set of polynomials with a certain partial symmetry property. However, in [6, 7], the author described a new action of the symmetric group and its Hecke algebra on the space of polynomials whose invariant are exactly the quasi-symmetric polynomials. At this point, it is important to realize that the action is not compatible with the product, so that there is no way to deduce from it that quasi-symmetric polynomials form an algebra.

A further step in the study of quasi-symmetric polynomials is the result of Garsia and Wallach [2] showing that QSym is free as a module on Sym. This was conjectured by F. Bergeron and C. Reutenauer and independently by J.-Y. Thibon when he gave me my thesis subject ! Unfortunately I was not able to solve it.

As in the case of polynomials the module of QSym over Sym is of rank $n!$, which can be easily deduced from the generating series. However the proof of Garsia and Wallach relies on a very subtle analyze of the generating series and do not gives no good explanation of the link between the $n!$ and the symmetric group. The present paper is actually a tentative to understand their result. Though there is apparently no link between the rank $n!$ of the module and the existence of the action of the symmetric group, both of these properties seem to generalize. This give a new light, which allows us to give a combinatorial interpretation of the Hilbert series of the module, assuming its freeness:

$$\mathrm{Hilb}_t(\mathrm{QSym}^r(X_n)) = \sum_{\sigma\in\mathfrak{S}_n} \frac{t^{\mathrm{Maj}(\sigma)+r(n-\mathrm{Fix}(\sigma))}}{(1-t)(1-t^2)\cdots(1-t^n)}\,, \tag{1}$$

where Maj is the major index ant Fix is the number of fixed points. Such a result is quite unexpected even in the case of $QSym$.

The paper is structured as follows. In a first part, we study a class of actions of the symmetric group on polynomials generalizing those of [6, 7]. In these actions, the elementary transposition σ_i acts only on the variables x_i and x_{i+1}, and therefore we call these action *local*. Then, we investigate the case where the set of fixed points under one of these action form a sub-algebra of the algebra of polynomials. This gives rise to an infinite hierarchy of sub-algebras of QSym, interpolating between QSym and Sym. We further show that, when the number of variables is infinite, these algebras are actually sub-Hopf-algebras of QSym and describe the dual Hopf algebras as quotient of **NCSF**. Finally, we show that, as in the case of QSym these algebras are free, and that, in the finite number of variable case, they seem to be Cohen-Macaulay as indicated by their generating series. It seems that Garsia and Wallach's proof could be adapted to this general case, and further work is in progress in this direction.

2. Background

2.1. **Quasi-symmetric functions.** Let $X := \{x_1 < x_2 < \cdots < x_n\}$ denote a totally ordered set of commutative indeterminates. X is called the *alphabet*. By $\mathcal{P}(X)$ (resp. $\mathcal{P}_k(X)$), we mean the set of the subsets (resp. k-elements subsets) of the alphabet X.

Let m be the monomial $x_1^{m_1} \dots x_n^{m_n}$ where the m_i's are possibly zero. For readability, we identify m with the integer vector $[m_1, m_2, \dots, m_n]$. We define the *support* of m as the set $A \in \mathcal{P}(X)$ of the x_i's whose exponent are non-zero, as well as the composition I obtained by removing the zeros in the sequence $(m_1, m_2, \dots, m_n)$. In the sequel we write A^I in place of the monomial m. For example if $X = \{x_1 < x_2 < x_3 < x_4\}$, we write $x_1^2 x_3 = [2, 0, 1, 0] = \{x_1, x_3\}^{(2,1)}$ and $x_1^3 x_2^5 x_4 = [3, 5, 0, 1] = \{x_1, x_2, x_4\}^{(3,5,1)}$.

A polynomial $f \in \mathbb{C}[X]$ is said to be *quasi-symmetric* if and only if for each composition $I = (i_1, \dots, i_r)$ the coefficient of the monomial

$$x_{j_1}^{i_1} x_{j_2}^{i_2} \dots x_{j_r}^{i_r} \tag{2}$$

is independent of the choice of $j_1 < j_2 < \cdots < j_r$. With our support and exponent notations it says that the coefficient of A^I is independent of the set of variables $A \in \mathcal{P}_r(X)$.

The quasi-symmetric polynomials form a subalgebra of $\mathbb{C}[X]$ denoted by QSym_n. It is often convenient to let n tends toward ∞ and to take the inverse limit in the category of graded ring. Then, we get an algebra called the algebra of *quasi-symmetric functions* [4]. Such functions can be seen as formal sums of monomials on an infinite alphabet $X := \{x_1 < x_2 < \cdots < x_n < \cdots\}$.

It is clear that the family of so-called *quasi-monomial functions* defined by

$$M_I := \sum_{A \in \mathcal{P}_r(X)} A^I = \sum_{j_1 < \cdots < j_r} x_{j_1}^{i_1} \cdots x_{j_r}^{i_r} = \sum_{k \to I} x^k \tag{3}$$

labeled by compositions $I := (i_1 \dots, i_r)$ form a basis of QSym, and the last sum is extended to all integer vectors $k \to I$ of length n obtained by inserting zeros in the composition I. For example,

$$M_{(2,1)} = \{x_1, x_2\}^{(2,1)} + \{x_1, x_3\}^{(2,1)} + \{x_1, x_4\}^{(2,1)} + \{x_2, x_3\}^{(2,1)} + \\ \{x_2, x_4\}^{(2,1)} + \{x_3, x_4\}^{(2,1)}$$

which can be written more concisely as:

$$M_{(2,1)} = [2,1,0,0] \; + \; [2,0,1,0] \; + \; [2,0,0,1] \; + \; [0,2,1,0] \; + \\ [0,2,0,1] \; + \; [0,0,2,1]$$

instead of $M_{(2,1)} = x_1^2x_2 + x_1^2x_3 + x_1^2x_4 + x_2^2x_3 + x_2^2x_4 + x_3^2x_4$.

2.2. **Noncommutative symmetric functions.** The algebra of *noncommutative symmetric functions* [3] is the free associative algebra $\mathbf{NCSF} = \mathbb{C}\langle S_1, S_2, \dots\rangle$ generated by an infinite sequence of noncommutative indeterminates S_k, called *complete* symmetric functions. For a composition $I = (i_1, i_2, \dots, i_r)$, one sets $S^I := S_{i_1}\, S_{i_2}\, \dots\, S_{i_r}$. The family (S^I) is a linear basis of **NCSF**. A useful realization can be obtained by taking an infinite alphabet $A = \{\, a_1,\, a_2, \dots\}$ and defining its complete homogeneous symmetric functions by the generating function

$$\sum_{n\geq 0} t^n\, S_n(A) = \; (1 - t\, a_1)^{-1}\, (1 - t\, a_2)^{-1}\, (1 - t\, a_3)^{-1}\, \dots \tag{4}$$

Then, $S_n(A)$ appears as the sum of all nondecreasing words of length n. Note that these functions are not symmetric in the usual sense. They are invariant for a more subtle action of the symmetric group of the alphabet due to Lascoux and Schützenberger [9], (see also [3]). The role of Schur functions is played by the noncommutative *ribbon Schur functions* R_I defined by

$$R_I := \sum_{J \preceq I} (-1)^{\ell(I)-\ell(J)}\, S^J\, . \tag{5}$$

The family (R_I) forms a basis of **NCSF**. In the realization of **NCSF** given by Equation (4), R_I reduces to the sum of all words of shape I [3].

The duality pairing $\langle \cdot \,|\, \cdot \rangle$ between QSym and **NCSF** is defined by $\left\langle M_I | S^J \right\rangle = \delta_{IJ}$ or equivalently $\langle F_I | R_J\rangle = \delta_{IJ}$ (*cf.* [11, 3]). This duality can be interpreted as the canonical duality between the Grothendieck groups respectively associated with finite dimensional and projective modules over 0-Hecke algebras (*cf.* [1, 8]).

3. Local actions

3.1. **Definitions.**

Definition 1. A *local action* of $\mathfrak{S}_n$ on $\mathbb{K}[X]$ is an action of $\mathfrak{S}_n$ on $\mathbb{K}[X]$ *i.e.* a morphism $\rho : \mathfrak{S}_n \longrightarrow \mathrm{End}(\mathbb{K}[X])$ which satisfies the following conditions:

(1) the elementary transposition σ_i acts only on the variables x_i and x_{i+1}, the other variables behaving as scalars;
(2) the action of an elementary transposition σ_i on a monomial $x_i^a x_{i+1}^b$ is independent of i and, depending on the values of a and b, either exchanges the variables or leaves the monomial invariant: for all i

$$\sigma_i(x_i^a x_{i+1}^b) = \begin{cases} x_i^b x_{i+1}^a & \text{for certain values of } a \text{ and } b\,, \\ x_i^a x_{i+1}^b & \text{for the other values}\,. \end{cases}$$

Consequently, the action of σ_1 on all monomials in x_1, x_2 fully determines the operation ρ.

Example 1. Let $m := [m_1, m_2, \dots m_n]$ be a monomial. Let also

$$\rho(\sigma_i)(m) := \begin{cases} [\dots, m_i, m_{i+1}, \dots] & \text{if } m_i \equiv m_{i+1} \pmod 2, \\ [\dots, m_{i+1}, m_i, \dots] & \text{otherwise.} \end{cases}$$

The hypothesis (1) and (2) are verified. We will see in the sequel that this actually defines a local action of the symmetric group.

The following proposition is an immediate consequence of the definition:

Proposition 3.1. *Let ρ a local action of $\mathfrak{S}_n$ on $\mathbb{K}[x_1, \dots, x_n]$. Then the restriction of ρ to $\mathfrak{S}_{n-1}$ is a local action on $\mathbb{K}[x_1, \dots, x_{n-1}]$.*

Let us give an other characterization of local actions:

Definition 2. Let ρ be a map of $\mathfrak{S}_n$ to $\mathrm{End}\,\mathbb{K}[X]$ which satisfies conditions (1) and (2) of the definition 1. To ρ we associate the relation $\mathcal{R}_\rho$ on the integers defined by

$$u \,\mathcal{R}_\rho\, v \quad \text{iff} \quad \sigma_1[u, v] = [u, v]. \tag{6}$$

Note that we do not suppose that ρ is a morphism, consequently it does not surely define an action. The conditions (1) and (2) ensure that, for all i

$$\rho(\sigma_i)(m) = \begin{cases} [\dots, m_i, m_{i+1}, \dots] & \text{if } m_i \,\mathcal{R}_\rho\, m_{i+1}, \\ [\dots, m_{i+1}, m_i, \dots] & \text{otherwise.} \end{cases} \tag{7}$$

Conversely, if $\mathcal{R}$ is a reflexive relation on the set of integers, Equation (7) defines a map from $\mathfrak{S}_n$ to $\mathrm{End}\,\mathbb{K}[X]$ which satisfies both conditions (1) and (2).

Proposition 3.2. *Let ρ be a map from $\mathfrak{S}_n$ to $\mathrm{End}\,\mathbb{K}[X]$ which satisfies both conditions (1) and (2) of definition 1. Let $\mathcal{R}_\rho$ be the associated relation. The following properties are equivalent:*

(1) ρ *is a morphism from* $\mathfrak{S}_n$ *to* $\mathrm{End}(\mathbb{K}[X])$,
(2) $\mathcal{R}_\rho$ *is an equivalence relation.*

Example 2. Let us go back to the former example. There are two equivalence classes: E for even and O for odd. To the monomial

$$m = [\ 0,\ 2,\ 1,\ 4,\ 3,\ 5,\ 1,\ 1, 2\ ]$$

we associate the word

$$C(m) = [\,\mathrm{E},\ \mathrm{E},\ \mathrm{O},\ \mathrm{E},\ \mathrm{O}, \mathrm{O}, \mathrm{O}, \mathrm{O}, \mathrm{E}\,]$$

and the two sequences of exponents

$$E_{\mathrm{E}} = (0, 2, 4, 2) \quad \text{and} \quad E_{\mathrm{O}} = (1, 3, 5, 1, 1).$$

Let $\sigma = 743652198$. Then,

$$\sigma(C(m)) = [\,\mathrm{O},\ \mathrm{O},\ \mathrm{O},\ \mathrm{E},\ \mathrm{O},\ \mathrm{E},\ \mathrm{E},\ \mathrm{E}, \mathrm{O}].$$

We finally get

$$\rho(\sigma)m = [\ 1,\ 3,\ 5,\ 0,\ 1,\ 2,\ 4,\ 2, 1\ ].$$

3.2. **Characteristic and generalized Temperley-Lieb algebras.** In the previous subsection, we defined a family of actions of the symmetric group $\mathfrak{S}_n$ on polynomials. In this section we describe more precisely these actions in terms of the representations theory. To describe a representation of $\mathfrak{S}_n$, rather than giving its character we prefer the equivalent but easier to handle Frobenius characteristic. We only describe briefly this tool and refer to [10] for notations and details.

Recall that the direct sum $\bigoplus_{n\geq 0} R(\mathfrak{S}_n)$ of Grothendieck rings of all symmetric groups is in natural isomorphism with the ring of symmetric functions, by the so-called Frobenius characteristic map ch. It sends the irreducible character χ^λ to the Schur function s_λ. The product of symmetric functions corresponds to induction from $\mathfrak{S}_n \times \mathfrak{S}_p$ to $\mathfrak{S}_{n+p}$. One can get the value of the character χ on the conjugacy class (cycle type) indexed by the partition $\mu = (\mu_1, \ldots, \mu_p)$ by the scalar product:

$$\chi(\mu) = \langle \mathrm{ch}(\chi) | p_\mu \rangle\,. \tag{8}$$

where p_μ is the product of power sum symmetric functions.

The r-actions are compatible with the usual grading of polynomial rings. Hence, one can define the graded characteristic ch_t of the representation on $\mathbb{C}[X]$ as the generating series of the characteristics of the representations on the homogeneous components $\mathbb{C}_i[X]$:

$$\mathrm{ch}_t(\mathbb{C}[X]) = \sum_{i=0}^{\infty} \mathrm{ch}(\mathbb{C}_i[X])t^i. \tag{9}$$

The main result of this section is the following

Theorem 3.3. *Let $(\rho_n)_n$ be the family of local actions of $\mathfrak{S}_n$ related to an equivalence relation $\mathcal{R}_\rho$. Denote by $C := \mathbb{N}/\mathcal{R}_\rho$ the set of equivalence classes. The generating series of the graded characteristic of the actions ρ_n of $\mathfrak{S}_n$ with respect to n is given by*

$$\sum_n \mathrm{ch}_t(\mathbb{C}[X_n]_\rho)u^n = \prod_{c\in C} H\left(\sum_{i\in c} t^i u\right), \tag{10}$$

where,

$$H(u) := \sum_j h_j u^j. \tag{11}$$

In particular, if the number s of equivalence classes in C is finite, then only the products of at most s terms h_i appear in this characteristic. As a consequence, only the irreducible representation V_λ for λ of length smaller than s appear. Thus the action is in fact an action of a quotient of the algebra of $\mathfrak{S}_n$. This can be described explicitly:

Theorem 3.4. *Suppose that the set of equivalence classes associated to a local action ρ of $\mathfrak{S}_n$ is finite of cardinal $s < n$.*

The image of $\mathbb{C}[\mathfrak{S}_n]$ in $\mathrm{End}(\mathbb{C}[X])$ is the quotient of $\mathbb{C}[\mathfrak{S}_n]$ by the ideal generated by

$$\nabla_i := \sum_{\sigma\in\mathfrak{S}^i_{s+1}} -1^{\ell(\sigma)}\sigma, \tag{12}$$

where $\mathfrak{S}^i_{s+1}$ is the shifted symmetric group $\mathfrak{S}_{s+1}$ which acts on the set $\{i, i+1, \dots, i+s+1\}$.

The operators $\rho(\sigma)$ where σ runs along the sets of permutations that avoid the pattern $(s+1)s\dots21$ form a basis of the image of $\mathbb{C}[\mathfrak{S}_n]$ in $\mathrm{End}(\mathbb{C}[X])$.

Recall that the permutation avoiding $s+1s\dots21$ are also the permutations such that the shape of the tableaux obtained by Robinson-Schensted map is a partition with at most s rows.

The proof is then the same as in [7], by comparison on dimension.

3.3. **Fixed polynomials and main theorem.**

Definition 3. Let ρ be a local action. A polynomial f is *fixed by ρ* or *ρ-symmetric*, if it is invariant under the action of ρ, that is,

$$\rho(\sigma)f = f \qquad \text{for all } \sigma\in\mathfrak{S}_n. \tag{13}$$

The set of ρ-symmetric polynomials is denoted by $\mathbb{K}[X_n]^{\rho(\mathfrak{S}_n)}$.

We now want to know for which local actions the set of fixed polynomials form a sub-algebra of the polynomial algebra $\mathbb{K}[x1,\dots,x_n]$. Equivalently, we want to know under which conditions on $\mathcal{R}$ the product of two fixed polynomials is always a fixed polynomial.

Remark 1. Let ρ be a local action of $\mathfrak{S}_n$ on the space of polynomials. Let $f(x_1, \ldots, x_n)$ be a ρ-symmetric polynomial. Then $f(x_1, \ldots, x_{n-1}, 0)$ is μ-symmetric under the restriction μ of ρ to $\mathfrak{S}_{n-1}$.

We now want to classify the local-actions whose symmetric polynomials form an algebra.

Definition 4. Let r be an integer or infinite. The r-action of $\mathfrak{S}_n$ on polynomials denoted by ρ_r is defined by

$$\rho_r(\sigma_i)(m) = \begin{cases} [\ldots, m_i, m_{i+1}, \ldots] & \text{if } m_i \geq r \text{ and } m_{i+1} \geq r, \\ [\ldots, m_{i+1}, m_i, \ldots] & \text{otherwise.} \end{cases} \tag{14}$$

Note that the trivial action corresponds to $r = 0$, the classical action corresponds to r infinite and the quasi-symmetrizing action of [6, 7] corresponds to $r = 1$.

The main theorem says that they are the only actions whose invariants are stable by multiplication :

Theorem 3.5. *Let ρ be a local action of $\mathfrak{S}_n$. The set $\mathbb{K}[X_n]^{\rho(\mathfrak{S}_n)}$ of fixed polynomials is a sub-algebra of $\mathbb{K}[X]$ if an only if there exists an r, integer or infinite, such that $\rho = \rho_r$.*

Let $\mathrm{QSym}^r(X_n) := \mathbb{K}[X_n]^{\rho_r(\mathfrak{S}_n)}$ denote the sub-algebra of $\mathbb{K}[X_n]$ spanned by r-quasi-symmetric polynomial. It is clear that the restriction morphisms $(n > p)$

$$\begin{aligned} \phi_p^n \ : \ \mathrm{QSym}^r(x_1, \ldots, x_n) &\longrightarrow \mathrm{QSym}^r(x_1, \ldots, x_p) \\ f(x_1, \ldots, x_n) &\longmapsto f(x_1, \ldots, x_p, 0, \ldots, 0) \end{aligned} \tag{15}$$

are compatible (*i.e.* $\phi_n^m \circ \phi_p^n = \phi_p^m$). Hence, it makes sense to take the reverse limit of this projective system in the category of *graded* algebras. We emphasize on the graded condition because we only want to deal with finite degree expressions rather than series. The limit is called the algebra of r-quasi-symmetric functions denoted by QSym^r. We will see in Section 4 than it can actually be equipped naturally with a Hopf algebra structure. For the moment we only concentrate on the algebra structure.

As in the case of classical symmetric and quasi-symmetric functions (*cf.* [10, 4]), QSym^r can be viewed as an algebra of power series of finite degree on an infinite set of variables.

The following inclusions hold:

$$\begin{aligned} \mathbb{K}[X_n] = \mathrm{QSym}^0(X_n) \supset \mathrm{QSym}^1(X_n) \supset \mathrm{QSym}^2(X_n) \supset \\ \cdots \supset \mathrm{QSym}^r(X_n) \supset \cdots \supset \mathrm{QSym}^\infty(X_n) = \mathrm{Sym}(X_n)\,, \end{aligned} \tag{16}$$

and for the infinite alphabet case:

$$\mathrm{QSym}^0 \supset \mathrm{QSym}^1 \supset \cdots \supset \mathrm{QSym}^r \supset \cdots \supset \mathrm{QSym}^\infty = \mathrm{Sym}\ . \tag{17}$$

A basis of QSym^r is given by the analog of the quasi-monomial basis *i.e.* by the orbit-sums. It is indexed by pairs of composition I in parts at least r and a partition λ whose parts are strictly smaller than r. We call such a pair an *r-composition* and write it (I, λ). A monomial whose exponent is an r-composition with possibly added trailing zeroes is called a r-dominant monomial.

For example, here is the list of the 24 3-compositions of 8

$$(8), (7,1), (6,2), (6,11), (53), (5,21), (5,111), (44), (43,1), (4,22),$$
$$(4,211), (4,1111), (35), (34,1), (33,2), (33,11), (3,221), (3,2111),$$
$$(3,11111), (2222), (22211), (221111), (2111111), (11111111)\,.$$

For later reference let us state the following obvious lemma:

Lemma 3.6. *For a fixed r, there is only one r-dominant monomial in each orbit under the r-action.*

The r-dominant monomial corresponding to the orbit of the monomial $m = [m_1, \ldots, m_i, \ldots]$ is the monomial obtained by putting the part m_i strictly smaller than r in decreasing order and the end of m.

Take $I = (I_1, \ldots, I_k)$ and $\lambda = (\lambda_1, \ldots, \lambda_l)$. The monomial function $M^r_{(I,\lambda)}$ is defined as the sum of the orbit of the r-dominant monomial

$$X^{(I,\lambda)} := x_1^{I_1} x_2^{I_2} \ldots x_k^{I_k} x_{k+1}^{\lambda_1} x_{k+2}^{\lambda_2} \ldots x_{k+l}^{\lambda_l} \tag{18}$$

under the r-action, that is

$$M^r_{(I,\lambda)}(X) := \sum_{\substack{i_1 \neq i_2 \neq \cdots \neq i_{k+l} \\ i_1 < i_2 < \cdots < i_k}} x_{i_1}^{I_1} x_{i_2}^{I_2} \ldots x_{i_k}^{I_k} x_{i_{k+1}}^{\lambda_1} x_{i_{k+2}}^{\lambda_2} \ldots x_{i_{k+l}}^{\lambda_l} \tag{19}$$

One easily sees that

$$M^r_{(I,\lambda)} = \sum_{K \in I \sqcup\!\sqcup \lambda^\sigma} M_K\,, \tag{20}$$

where λ^σ denotes the set of all distinct reordering of λ and $M_K = M^1_K$ is the classical quasi-symmetric functions. For example

$$\begin{aligned} M^3_{(4,6),(2,1)} &= M_{4,6,1,2} + M_{4,1,6,2} + M_{4,1,2,6} + M_{1,4,6,2} + M_{1,4,2,6} + M_{1,2,4,6} \\ &\quad + M_{4,6,2,1} + M_{4,2,6,1} + M_{4,2,1,6} + M_{2,4,6,1} + M_{2,4,1,6} + M_{2,1,4,6} \end{aligned}$$

The previous expression allows to deduce the product rule of QSym^r from the one of QSym. Recall that the product of Classical quasi-monomial function is given by the quasi-shuffle:

$$M_I M_J = \sum_K \langle K | I \uplus J \rangle\, M_K\,, \tag{21}$$

where $I \uplus J$ is defined recursively by

$$\epsilon \uplus I \;=\; I \uplus \epsilon = I \tag{22}$$
$$I \uplus J \;=\; i_1 \triangleright (I' \uplus J) + j_1 \triangleright (I \uplus J') + (i_1 + j_1) \triangleright (I' \uplus J')\,, \tag{23}$$

where ϵ is the empty composition, $\triangleright$ denotes the operation of adding a part at the beginning of a composition, and finally $I = i_1 \triangleright I'$ and $J = j_1 \triangleright J'$. Then

Proposition 3.7. *Let $r > 0$ and (I, λ) and J, μ two r-compositions. Then,*

$$M^r_{(I,\lambda)} M^r_{(J,\mu)} = \sum_{(K,\nu)} \sum_{\substack{A \in I \sqcup\!\sqcup \lambda^\sigma \\ B \in J \sqcup\!\sqcup \mu^\sigma}} \langle (K,\nu) | A \uplus B \rangle \, M^r_{(K,\nu)} \tag{24}$$

where λ^σ and μ^σ run through the set of the reordering of λ and μ as in Equation (20) and $\langle (K,\nu) | A \uplus B \rangle$ is the coefficient of (K,ν) in the quasi-shuffle of A with B.

3.4. Generating series.

Theorem 3.8. *The generating series of the graded characteristic of the r-action on $\mathfrak{S}_n$ with respect to n is given by*

$$\sum_n \mathrm{ch}_t(\mathbb{C}[X_n]) a^n = H\left(\frac{at^r}{1-t}\right) \prod_{i=0}^{r-1} H(at^i)\,, \tag{25}$$

where as usual

$$H(u) := \sum_i h_i u^i = \sigma_u(A) = \sum_i h_i(A) u^i\,. \tag{26}$$

A consequence of this theorem, which can also be proved directly, is the generating series of the homogeneous dimension of QSym^r.

Theorem 3.9. *The generating series of the dimensions of the homogeneous components of $\mathrm{QSym}^r(X_n)$ is given by:*

$$\sum_{d,n} \dim_d(\mathrm{QSym}^r(X_n)) a^n t^d = \frac{1-t}{1-t-at^r} \prod_{i=0}^{r-1} \frac{1}{1-at^i} \tag{27}$$

The generating series of the dimensions of the homogeneous components of QSym^r is given by:

$$\mathrm{Hilb}_t(\mathrm{QSym}^r) := \sum_{d,n} \dim_d(\mathrm{QSym}^r) t^d = \frac{1-t}{1-t-t^r} \prod_{i=1}^{r-1} \frac{1}{1-t^i} \tag{28}$$

Here are the first values:

	0	1	1	3	4	5	6	7	8	9	10
QSym	1	1	2	4	8	16	32	64	128	256	512
QSym^2	1	1	2	3	5	8	13	21	34	55	89
QSym^3	1	1	2	3	5	7	11	16	24	35	52
QSym^4	1	1	2	3	5	7	11	15	22	31	44
Sym	1	1	2	3	5	7	11	15	22	30	42

4. Hopf algebra structure

There are two equivalent ways to endow QSym^r with a Hopf algebra structure. The first way is to prove that QSym^r is a sub-Hopf algebra of QSym; that is that the coproduct of an r-Quasi-Symmetric element is an r-quasi-symmetric element. We prefer a slightly less direct but in our opinion more instructive way, namely to use the classical alphabet doubling trick to define the coproduct.

Let $X = \{x_1, x_2, \dots\}$ and $Y = \{y_1, y_2, \dots\}$ be two totally ordered infinite alphabets. The ordered sum $X \sqcup Y$ of alphabet is, as a set, the union of X and Y, where we keep the order inside X and Y and we choose that the elements of X are smaller than the elements of Y. Clearly, as soon as an alphabet Z is totally ordered and infinite $\mathrm{QSym}^r \simeq \mathrm{QSym}^r(Z)$; in particular,

$$\mathrm{QSym}^r(X) \simeq \mathrm{QSym}^r(X \sqcup Y)\,. \tag{29}$$

But now its clear from the restriction property (Remark 1) that a r-quasi-symmetric function in $X \sqcup Y$ is in particular r-quasi-symmetric in X and in Y. Thus one has a natural algebra morphism

$$\mathrm{QSym}^r(X) \mapsto \mathrm{QSym}^r(X) \otimes \mathrm{QSym}^r(Y)\,. \tag{30}$$

which can be seen as a coproduct. The co-associativity is an easy consequence of the associativity of the ordered sum of alphabets. The co-unit corresponds to the restriction to an empty alphabet, *i.e.* the mapping which extracts the coefficient of the constant monomial.

Hence the following theorem

Theorem 4.1. *The r-quasi-symmetric functions form an infinite hierarchy of graded sub-Hopf-algebras:*

$$\mathrm{QSym} = \mathrm{QSym}^1 \supset \cdots \supset \mathrm{QSym}^r \supset \cdots \supset \mathrm{QSym}^\infty = \mathrm{Sym}\,. \tag{31}$$

Let us express this explicitly.

Proposition 4.2. *Let $r > 0$ and (I, λ) be a r-composition. Then, the coproduct of $M^r_{(I,\lambda)}$ is given by*

$$\delta(M^r_{(I,\lambda)}) = \sum_{(K,K')=I;\ \lambda \in \nu \sqcup\!\sqcup \nu'} M^r_{(K,\nu)} \otimes M^r_{(K',\nu')}\,. \tag{32}$$

For example

$$\begin{aligned}
\delta(M^2_{635211}) = {} & M^2_{635211} \otimes 1 + M^2_{63521} \otimes M^2_1 + M^2_{63511} \otimes M^2_2 + M^2_{6351} \otimes M^2_{21} + \\
& M^2_{635} \otimes M^2_{211} + M^2_{63211} \otimes M^2_5 + M^2_{6321} \otimes M^2_{51} + M^2_{6311} \otimes M^2_{52} + \\
& M^2_{631} \otimes M^2_{521} + M^2_{63} \otimes M^2_{5211} + M^2_{6211} \otimes M^2_{35} + M^2_{621} \otimes M^2_{351} + \\
& M^2_{611} \otimes M^2_{352} + M^2_{61} \otimes M^2_{3521} + M^2_{6} \otimes M^2_{35211} + M^2_{211} \otimes M^2_{635} + \\
& M^2_{21} \otimes M^2_{6351} + M^2_{11} \otimes M^2_{6352} + M^2_{1} \otimes M^2_{63521} + 1 \otimes M^2_{635211}
\end{aligned}$$

4.1. **Dual Hopf algebra and noncommutative symmetric functions.** Since the Hopf algebra QSym^r is a sub-algebra the algebra of quasi-symmetric functions, its dual is naturally a quotient of **NCSF**. The goal of this section is to describe explicitly this duality.

Recall that **NCSF** is the free associative algebra over an infinite set of generators S_i with basis (S_i) where

$$S^I = S^{(I_1,I_2,\ldots,I_k)} := S_{I_1}S_{I_2}\ldots S_{I_k}\,. \tag{33}$$

The bases (M_I) and (S^I) are two dual bases of respectively QSym and **NCSF**. Since the basis $M^r_{(I,\lambda)}$ of QSym^r is formed by disjoints sums of elements of (M_I), the dual of QSym^r can be described by identifying some elements of the basis (S^I).

The dual of QSym^r can be described explicitly as follows:

Theorem 4.3. *The dual of* QSym^r *is the quotient of the Hopf algebra* **NCSF** *by the relations*

$$S_iS_j = S_jS_i \qquad \textit{for all } i < r \textit{ and } j \in \mathbb{N}. \tag{34}$$

A basis of this quotient is given by $(S^{(I,\lambda)})$, where (I,λ) goes along the set of all r-compositions.

4.2. **Primitive Elements.** The Hopf algebra QSym^r is graded connected and co-commutative. Hence, by the Milnor-Moore theorem it is isomorphic to the universal enveloping algebra of the Lie algebra of its primitive elements. The goal of this section is to explicitly describe these primitive elements.

In [3] (see also [11]), an explicit construction of several generating families, called non-commutative power sums of the Lie algebra Prim(**NCSF**) is given. Let us recall briefly one of them. We use a slightly modified base: our Φ_n corresponds to $n\Phi_n$ of [3] or equivalently to the base P_n^* of [11]. They are defined by means of generating series, the relation

$$\sum_{k\geq 1} t^k\Phi_k = \log\left(1+\sum_{k\geq 1} S_kt^k\right) \tag{35}$$

defining uniquely a sequence of primitive elements $(\Phi_k)_{k\geq 1}$ in **NCSF**. Explicitly this gives:

$$\Phi_n = \sum_{K\vDash n} \frac{(-1)^{\ell(K)-1}}{\ell(k)} S^K\,. \tag{36}$$

An important consequence is that **NCSF** is freely generated by (Φ_n); moreover as a Hopf algebra it is isomorphic to the universal enveloping algebra $\mathrm{U}(\mathcal{L}(\Phi_i; i>0))$ where $\mathcal{L}(\Phi_i; i>0)$ is the free Lie algebra generated by $(\Phi_i)_{i>0}$.

By projecting these elements into $\mathbf{NCSF}^r$ we get an explicit description of the structure of the primitive Lie algebra $\mathrm{Prim}(\mathbf{NCSF}^r)$ and of the Hopf algebra $\mathbf{NCSF}^r$.

Theorem 4.4. $\mathbf{NCSF}^r$ *is the quotient of the Hopf algebra* $\mathbf{NCSF}$ *by the relations*

$$[\Phi_i, \Phi_j] = 0 \qquad \text{for all } i < r \text{ and } j \in \mathbb{N}. \tag{37}$$

The Lie algebra $\mathrm{Prim}(\mathbf{NCSF}^r)$ *is generated by* $(\Phi_i)_{i>0}$. *Moreover it is isomorphic to the central extension of the free Lie algebra with generators* $(\Phi_i)_{i\geq r}$ *by the trivial commutative Lie algebra with basis* $(\Phi_i)_{i<r}$

Hence we get that the set of the elements

$$\Phi^{(I,\lambda)} = \Phi_{I_1} \dots \Phi_{I_k} \Phi_{\lambda_1} \dots \Phi_{\lambda_k}\,, \tag{38}$$

where (I, λ) goes along all r-compositions, is a multiplicative basis of $\mathbf{NCSF}^r$ generated by primitive elements.

5. Free algebra structure

This last property shows that the dual basis $\phi_{(I,\lambda)}$ has a very nice product rule. As a consequence, we elucidate the structure of QSym^r. First let us give explicitly the dual basis of Φ^I of $\mathbf{NCSF}^r$. In the classical case of the duality $(\mathrm{QSym}, \mathbf{NCSF})$ the dual bases of the Φ^I is the basis ϕ_I given by

$$\phi_I = \sum_{I \succeq K} \frac{1}{(\#(I,K))!} M_K\,, \tag{39}$$

where $(\#(I,K))! = \ell(J_1)!\,\ell(J_2)! \dots \ell(J_r)!$, and $J_1 \dots J_r$ are compositions such that $I = I_1 \cdot I_2 \cdots I_r$ and $K = (|I_1|, |I_2|, \dots, |I_r|)$. In particular ϕ_i is equal to the symmetric power sum $p_i = \sum x_k^i$. Since the Φ_i's are primitive the product of the ϕ_I is given by the shuffle product

$$\phi_I \phi_J = \sum_K \langle K | I \sqcup\!\sqcup J \rangle\, \phi_K\,. \tag{40}$$

Let $\phi'^r_{(I,\lambda)}$ be the dual basis of $\Phi^{(I,\lambda)}$. The formula which expresses $\phi_{(I,\lambda)}$ in terms of the ϕ_J is the same as the formula which expresses $M_{(I,\lambda)}$ in terms of the M_J's $\phi'^r_{(I,\lambda)} = \sum_{K \in I \sqcup\!\sqcup \lambda^\sigma} \phi_K$. We prefer a slightly different normalization: write λ in exponential notation $\lambda = (1^{m_1} 2^{m_1} \dots)$, and set $z_\lambda := \prod m_i!$. The basis $\phi^r_{(I,\lambda)}$ is defined as

$$\phi^r_{(I,\lambda)} := z_\lambda \phi'^r_{(I,\lambda)} = z_\lambda \sum_{K \in I \sqcup\!\sqcup \lambda^\sigma} \phi_K = \phi_I \phi_{\lambda_1} \phi_{\lambda_2} \dots \phi_{\lambda_l}\,. \tag{41}$$

Then the product of two such functions is given by

$$\phi^r_{(I,\lambda)} \phi^r_{(J,\mu)} = \sum_K \langle K | I \sqcup\!\sqcup J \rangle\, \phi^r_{(K,\nu)}\,, \tag{42}$$

where ν is the partition obtained by sorting the concatenation of the partitions λ and μ. Now it is known that the shuffle algebra is free on the Lyndon words (see [12] for a definition of Lyndon words). Hence we have the following theorem

Theorem 5.1. *The algebra* QSym^r *is the free algebra generated by the set*

- ϕ_k *for* $k < r$*;*
- ϕ_L *for all Lyndon words* $L = (l_1 \dots l_m)$ *with* $l_i \geq r$*.*

As a simple consequence:

Corollary 5.2. QSym^r *is a free* Sym*-module with generators* $(\phi_L)_L$ *where* L *goes along the set of Lyndon words* $L = (l_1 \dots l_m)$ *with* $m > 1$ *and* $l_i \geq r$*.*

It seems that the same property holds when the number of variables is finite. This is our main conjecture.

Conjecture 1. Let n be finite and $X_n := \{x_1, \dots, x_n\}$. As a $\mathrm{Sym}(X_n)$-module $\mathrm{QSym}^r(X_n)$ is free of dimension $n!$.

As proved recently by Garsia and Wallach [2], this conjecture holds for in the case of QSym that is for $r = 1$. It seems that their proof can be adapted to the generalized case, but we have not worked out this proof. However, we believe that the knowledge of the generalized case and the following proposition allows us to simplify the proof. Assuming this conjecture, we get the following Hilbert series for the quotient:

Proposition 5.3. *The quotient of the Hilbert series of* QSym^r *and* Sym *is given by*

$$\mathrm{Hilb}_t(\mathrm{QSym}^r(X_n)) / \mathrm{Hilb}_t(\mathrm{Sym}(X_n)) = \sum_{\sigma \in \mathfrak{S}_n} t^{\mathrm{Maj}(\sigma) + r(n - \mathrm{Fix}(\sigma))}, \tag{43}$$

where $\mathrm{Fix}(\sigma)$ *is the number of fixed points of the permutation* σ *and* $\mathrm{Maj}(\sigma)$ *is the* Major index *of* σ *that is the sum if the descents of* σ*.*

As far as we now, such a formula was unknown even for QSym. It follows from a result of Gessel and Reutenauer [5] (Theorem (8.4)).

This identity strongly suggests that the freeness of as well as the $n!$ result for each $Qsym^r[X_n]$ should be derivable by means of some action of S_n, in the same manner these results are established for the the polynomial ring. To this date no such a proof has been found even in the $r = 1$ case since the Garsia-Wallach proof follows a completely different path. Identifying the action that yields such a proof would be an outstanding result that would deeply increase our understanding of these remarkable modules.

References

[1] Duchamp G., Krob D., Leclerc B. ,Thibon J.-Y., *Fonctions quasi-symétriques, fonctions symétriques non commutatives, et algèbres de Hecke à $q = 0$*, C.R. Acad. Sci. Paris, **322**, 107–112, 1996.

[2] A. Garsia et N. Wallach, *Qsym over Sym is free*, J. Combin. Theory, Ser. A, **104** (2003), no. 2, 217–263.

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

[4] 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.

[5] I. Gessel, C. Reutenauer, Counting permutations with cycle and descent structure, J. Combin. Theory, Ser A, **64** (1993), 189–215.

[6] F. Hivert, *Combinatoire des fonctions quasi-symétriques*, Phd Thesis, University of Marne-La-Vallée, (1999).

[7] F. Hivert, *Hecke Algebras, Differences Operators, and Quasi-Symmetric Functions*, Adv. Math. **(155)**, (2000), 181-238.

[8] Krob D., Thibon J.-Y., *Noncommutative symmetric functions IV : Quantum linear groups and Hecke algebras at $q = 0$*, J. Alg. Comb., **6** (1997), no. 4, 339–376.

[9] A. Lascoux, and M. P. Schützenberger, Le Monode plaxique, *Quad. del. Ric. Sci.* **109** (1981), 129–156.

[10] I. G. Macdonald, Symmetric functions and Hall polynomials, *Clarendon Press*, (1995).

[11] Malvenuto C., Reutenauer C., *Duality between quasi-symmetric functions and Solomon descent algebra*, J. Algebra, **177**, 967–982, 1995.

[12] C. Reutenauer, *Free Lie Algebras*, Oxford science publications, (1993).

THE COMBINATORICS OF TWISTED INVOLUTIONS IN COXETER GROUPS

AXEL HULTMAN

Abstract. The open intervals in the Bruhat order on twisted involutions in a Coxeter group are shown to be PL spheres. This implies results conjectured by F. Incitti and sharpens the known fact that these posets are Gorenstein* over $\mathbb{Z}_2$.

We also introduce a Boolean cell complex which is an analogue for twisted involutions of the Coxeter complex. Several classical Coxeter complex properties are shared by our complex. When the group is finite, it is a shellable sphere, shelling orders being given by the linear extensions of the weak order on twisted involutions. Furthermore, the h-polynomial of the complex coincides with the polynomial counting twisted involutions by descents. In particular, this gives a type independent proof that the latter is symmetric.

1. Introduction

Let (W, S) be a finite Coxeter system with an involutive automorphism θ. In [25], Springer studied the combinatorics of the twisted involutions $\mathfrak{I}(\theta)$. Together with Richardson, he refined his results in [23, 24] and put them to use in the study of the subposet of the Bruhat order on W induced by $\mathfrak{I}(\theta)$. One of their tools was another partial order on $\mathfrak{I}(\theta)$ which they called the *weak order* for reasons that will be explained later. Their motivation was an intimate connection between the Bruhat order on $\mathfrak{I}(\theta)$ and Bruhat decompositions of certain symmetric varieties.

The purpose of the present article is to investigate the properties of the Bruhat order and the weak order on $\mathfrak{I}(\theta)$ in an arbitrary Coxeter system. The Bruhat order and the two-sided weak order on a Coxeter group appear as special cases of these posets, and many properties carry over from the special cases to the general situation.

Specifically, we prove for the Bruhat order on $\mathfrak{I}(\theta)$ that the order complex of every open interval is a PL sphere. When $\theta = \mathrm{id}$, this is the principal consequence of a conjecture of Incitti [21] predicting that the poset is EL-shellable. The result sharpens [17, Theorem 4.2] which asserts that every interval in the Bruhat order on $\mathfrak{I}(\theta)$ is Gorenstein* over $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$.

Regarding the weak order on $\mathfrak{I}(\theta)$, we construct from it a Boolean cell complex Δ_θ analogously to how the Coxeter complex Δ_W is constructed from the weak order on W. Counterparts for Δ_θ of several important properties of Δ_W are proved. When W is finite, Δ_θ is a shellable sphere; any linear extension of the weak order on $\mathfrak{I}(\theta)$ is a shelling order. Moreover, the h-polynomial of Δ_θ turns out to coincide with the

Supported by the European Commission's IHRP Programme, grant HPRN-CT-2001-00272, "Algebraic Combinatorics in Europe".

generating function counting the elements of $\mathfrak{I}(\theta)$ by their number of descents. This yields a uniform proof that these polynomials are symmetric.

The remainder of the paper is organized as follows. In Section 2, we review necessary background material from combinatorial topology and the theory of Coxeter groups. Thereafter, in Section 3, we prove some facts about the combinatorics of $\mathfrak{I}(\theta)$ that we need in the sequel. Most of these are extensions to arbitrary Coxeter groups of results from [23, 24, 25]. The Bruhat order on $\mathfrak{I}(\theta)$ is studied in Section 4. Finally, in Section 5, we focus on the weak order on $\mathfrak{I}(\theta)$ and the aforementioned analogue of the Coxeter complex.

Acknowledgement. The author would like to thank Mark Dukes for valuable comments on an earlier version of the paper.

2. Preliminaries

Here, we collect some background terminology and facts for later use.

2.1. **Posets and combinatorial topology.** A poset is *bounded* if it has unique maximal and minimal elements, denoted $\hat{1}$ and $\hat{0}$, respectively. If P is bounded, then its *proper part* is the subposet $\overline{P} = P \setminus \{\hat{0}, \hat{1}\}$.

Definition 2.1. *A poset P is* Eulerian *if it is bounded, graded and finite, and its Möbius function μ satisfies*

$$\mu(p, q) = (-1)^{r(q)-r(p)}$$

for all $p \leq q \in P$, where r is the rank function of P.

To any poset P, we associate the *order complex* $\Delta(P)$. It is the simplicial complex whose simplices are the chains in P. When we speak of topological properties of a poset P, we have the properties of $\Delta(P)$ in mind.

Definition 2.2. *A poset P is* Gorenstein* *if it is bounded, graded and finite, and every open interval in P has the homology of a sphere of top dimension.*

It follows from the correspondence between the Möbius function and the Euler characteristic (the theorem of Ph. Hall) that the Gorenstein* property implies the Eulerian property.

Definition 2.3. *A simplicial complex is a* piecewise linear*, or* PL*,* sphere *if it admits a subdivision which is a subdivision of the boundary of a simplex.*

In particular, PL spheres are of course homeomorphic to spheres. The former class is more well-behaved under certain operations, and this sometimes facilitates inductive arguments. Our interest in PL spheres, rather than spheres in general, will be confined to such situations.

A *simplicial poset* is a finite poset equipped with $\hat{0}$ in which every interval is isomorphic to a Boolean lattice. Such a poset is the *face poset* (i.e. poset of cells ordered by inclusion) of a certain kind of regular CW complex called *Boolean cell complex.* Thus, Boolean cell complexes are slightly more general than simplicial ones, since we allow simplices to share vertex sets. Such complexes were first considered by Björner [1] and by Garsia and Stanton [16].

A Boolean cell complex is *pure* if its *facets* (inclusion-maximal cells) are equidimensional. A pure complex is *thin* if every cell of codimension 1 is contained in exactly two facets.

A *shelling order* of a pure Boolean cell complex Δ is an ordering $F_0, \ldots, F_t$ of the facets of Δ such that the subcomplex $F_i \cap (\cup_{\alpha<i} F_\alpha)$ is pure of codimension 1 for all $i \in [t] = \{1, \ldots, t\}$. If Δ has a shelling order, then it is *shellable*. The next result can be found in [1].

Proposition 2.4. *If Δ is a pure, thin, finite and shellable Boolean cell complex, then Δ is homeomorphic to a sphere.*

Let f_i be the number of i-dimensional cells in Δ (including $f_{-1} = 1$). The *f-polynomial* is the generating function of the f_i:

$$f_\Delta(x) = \sum_{i\geq 0} f_{i-1} x^i.$$

An equivalent, often more convenient, way of encoding this information is the *h-polynomial*. Setting $d = \dim(\Delta)$, it is defined by

$$h_\Delta(x) = (1-x)^{d+1} f_\Delta\left(\frac{x}{1-x}\right).$$

We define coefficients h_i by

$$h_\Delta(x) = \sum_{i=0}^{d+1} h_i x^i.$$

The next result follows from [26, Proposition 4.4].

Proposition 2.5 (Dehn-Sommerville equations)**.** *Suppose Δ is a pure Boolean cell complex of dimension d which is homeomorphic to a sphere. Then, for all $i \in [d+1]$,*

$$h_{d+1-i} = h_i.$$

If Δ is shellable, the h_i have nice combinatorial interpretations that we now describe. Suppose $F_1, \ldots, F_t$ is a shelling order of Δ. It can be proved that for all $i \in [t]$, $\cup_{\alpha\leq i} F_i$ contains a unique minimal cell which is not contained in $\cup_{\alpha<i} F_\alpha$. Let r_i denote the dimension of this cell. Then, for all j,

$$h_j = |\{i \in [t] \mid r_i = j-1\}|.$$

2.2. **Properties of Coxeter groups.** We now briefly review some important features of Coxeter groups for later use. We refer the reader to [4] or [18] for material which may not be familiar.

Henceforth, let (W, S) be a Coxeter system with $|S| < \infty$. The Coxeter length function is $\ell : W \to \mathbb{N}$. If $w = s_1 \ldots s_k \in W$ and $\ell(w) = k$, the word $s_1 \ldots s_k$ is called a *reduced expression* for w. Here and in what follows, symbols of the form s_i are always assumed to be elements in S. We do not distinguish notationally between a word in the free monoid over S and the element in W that it represents; we trust the context to make the meaning clear.

Given $w \in W$, we define the *left* and *right descent sets*, respectively, by

$$D_L(w) = \{s \in S \mid \ell(sw) < \ell(w)\}$$

and

$$D_R(w) = \{s \in S \mid \ell(ws) < \ell(w)\}.$$

Observe that $s \in D_R(w)$ (resp. $D_L(w)$) iff w has a reduced expression ending (resp. beginning) with s.

The following two results are equivalent formulations of an important structural property of Coxeter groups. The first has an obvious analogous formulation for left descents.

Proposition 2.6 (Exchange property)**.** *Suppose $s_1 \dots s_k$ is a reduced expression for $w \in W$. If $s \in D_R(w)$, then $ws = s_1 \dots \widehat{s_i} \dots s_k$ for some $i \in [k]$ (where the hat denotes omission).*

Proposition 2.7 (Deletion property)**.** *If $w = s_1 \dots s_k$ and $\ell(w) < k$, then $w = s_1 \dots \widehat{s_i} \dots \widehat{s_j} \dots s_k$ for some $1 \leq i < j \leq k$.*

The most important partial order on W is probably the one we now define. Denote by T the set of *reflections* in W, i.e.

$$T = \{wsw^{-1} \mid w \in W \text{ and } s \in S\}.$$

Definition 2.8. *The* Bruhat order *on W is the partial order $\leq$ defined by $v \leq w$ iff there exist $t_1, \dots, t_k \in T$ such that $w = vt_1 \dots t_k$ and $\ell(vt_1 \dots t_{i-1}) < \ell(vt_1 \dots t_i)$ for all $i \in [k]$. We denote this poset by* $\mathrm{Br}(W)$*.*

Although not immediately obvious from Definition 2.8, the Bruhat order is graded with rank function ℓ. It is topologically well-behaved:

Theorem 2.9 ([1, 6, 15])**.** *The open intervals in* $\mathrm{Br}(W)$ *are PL spheres.*

Innocent as it seems, the next property is nevertheless the key to many results on $\mathrm{Br}(W)$. It follows from Deodhar [12, Theorem 1.1]. Again, there is an obvious formulation for left descents.

Proposition 2.10 (Lifting property)**.** *Let $v, w \in W$ with $v \leq w$ and suppose $s \in D_R(w)$. Then,*

(i) *$vs \leq w$.*
(ii) *$s \in D_R(v) \Rightarrow vs \leq ws$.*

Next, we define another fruitful way to order W. It is readily seen that the following is a weaker order than $\mathrm{Br}(W)$:

Definition 2.11. *The* right weak order *on W is the partial order $\leq_R$ defined by $v \leq_R w$ iff $w = vu$ for some $u \in W$ with $\ell(u) = \ell(w) - \ell(v)$.*

There is of course also a left weak order defined in the obvious way; we denote it by $\leq_L$. Clearly, both weak orders are graded with rank function ℓ.

3. The combinatorics of twisted involutions

Let (W, S) be a Coxeter system, and $\theta : W \to W$ a group automorphism such that $\theta^2 = \mathrm{id}$ and $\theta(S) = S$. In other words, θ is induced by an involutive automorphism of the Coxeter graph of W.

Definition 3.1. *The set of* twisted involutions *is*

$$\mathfrak{I}(\theta) = \{w \in W \mid \theta(w) = w^{-1}\}.$$

Note that $\mathfrak{I}(\mathrm{id})$ is the set of ordinary involutions in W.

In this section, we show that the combinatorics of $\mathfrak{I}(\theta)$ is strikingly similar to that of W. Our results in this section have been developed for finite W by Springer [25] and by Richardson and Springer [23, 24], but their proofs (specifically, of the crucial [23, Lemma 8.1]) do not hold in the general case, since they make use of the existence of a longest element in W.

Example 3.2 (cf. Example 10.1 in [23])**.** Let W be any Coxeter group, and consider the automorphism $\theta : W \times W \to W \times W$ given by $(v, w) \mapsto (w, v)$. It is easily seen that

$$\mathfrak{I}(\theta) = \{(w, w^{-1}) \mid w \in W\},$$

so that we have a natural bijection $\mathfrak{I}(\theta) \longleftrightarrow W$. This construction allows many properties of $\mathfrak{I}(\theta)$ to be seen as generalizations of Coxeter group properties.

Consider the set of symbols $\underline{S} = \{\underline{s} \mid s \in S\}$. The free monoid over $\underline{S}$ acts from the right on the set W by

$$w\underline{s} = \begin{cases} ws & \text{if } \theta(s)ws = w, \\ \theta(s)ws & \text{otherwise,} \end{cases}$$

and $w\underline{s}_1 \ldots \underline{s}_k = (\ldots((w\underline{s}_1)\underline{s}_2)\ldots)\underline{s}_k$. Observe that $w\underline{s}\underline{s} = w$ for all $w \in W$, $s \in S$. By abuse of notation, we write $\underline{s}_1 \ldots \underline{s}_k$ instead of $e\underline{s}_1 \ldots \underline{s}_k$, where $e \in W$ is the identity element.

Remark 3.3. In [23], a slightly different monoid action is used. It satisfies the relation $\underline{s}\underline{s} = \underline{s}$ rather than $\underline{s}\underline{s} = 1$. We use our formulation since it makes the results easier to state and the similarity to the situation in Coxeter groups more transparent.

In $\mathfrak{I}(\theta)$ it is sometimes more convenient to use the following equivalent definition of the action:

Lemma 3.4. *Suppose $w \in \mathfrak{I}(\theta)$ and $s \in S$. Then,*

$$w\underline{s} = \begin{cases} ws & \textit{if } \ell(\theta(s)ws) = \ell(w), \\ \theta(s)ws & \textit{otherwise.} \end{cases}$$

Proof. Obviously, $\theta(s)ws = w$ implies $\ell(\theta(s)ws) = \ell(w)$. Conversely, suppose that $\ell(\theta(s)ws) = \ell(w)$. If $\theta(s)ws \neq w$, w must have a reduced expression which begins with $\theta(s)$ or ends with s. Assume without loss of generality that $\theta(s)s_1 \ldots s_k$ is such an expression. Since $\theta(w) = w^{-1}$, we have $\ell(ws) < \ell(w)$. No reduced expression for

w can both begin with $\theta(s)$ and end with s; the Exchange property therefore implies $ws = s_1 \dots s_k$, so that $\theta(s)ws = w$. □

Our interest in this action stems from the fact that the orbit of the identity element is precisely $\mathfrak{I}(\theta)$, as the next proposition shows.

Proposition 3.5.

(i) *For all* $\underline{s}_1, \dots, \underline{s}_k \in \underline{S}$, *we have* $\underline{s}_1 \dots \underline{s}_k \in \mathfrak{I}(\theta)$.

(ii) *Given* $w \in \mathfrak{I}(\theta)$, *there exist symbols* $\underline{s}_1, \dots, \underline{s}_k \in \underline{S}$ *such that* $w = \underline{s}_1 \dots \underline{s}_k$.

Proof. It is readily checked that $w\underline{s} \in \mathfrak{I}(\theta)$ iff $w \in \mathfrak{I}(\theta)$. This proves (i). Noting that $\ell(ws) < \ell(w) \Leftrightarrow \ell(w\underline{s}) < \ell(w)$, (ii) follows by induction over the length. □

Motivated by this proposition, we define the *rank* $\rho(w)$ of a twisted involution $w \in \mathfrak{I}(\theta)$ to be the minimal k such that $w = \underline{s}_1 \dots \underline{s}_k$ for some $\underline{s}_1, \dots, \underline{s}_k \in \underline{S}$. The expression $\underline{s}_1 \dots \underline{s}_k$ is then called a *reduced* $\underline{S}$*-expression* for w.

Example 3.6. Let W and θ be as in Example 3.2. If $s_1 \dots s_k$ is a reduced expression for $w \in W$, then $\underline{(s_1, e)} \dots \underline{(s_k, e)}$ is a reduced $\underline{S}$-expression for the corresponding twisted involution $(w, w^{-1}) \in \mathfrak{I}(\theta)$. To see that it is reduced, note that by construction ρ is always at least half the length, and that the length of (w, w^{-1}) in $W \times W$ is $2k$.

We write $\mathrm{Br}(\mathfrak{I}(\theta))$ for the subposet of $\mathrm{Br}(W)$ induced by $\mathfrak{I}(\theta)$. The study of $\mathrm{Br}(\mathfrak{I}(\theta))$ was initiated in [23, 24] because of its connection to Bruhat decompositions of certain symmetric varieties.

It follows from [17, Theorem 4.8] that $\mathrm{Br}(\mathfrak{I}(\theta))$ is graded with rank function ρ and that $\rho(w) = (\ell(w) + \ell^\theta(w))/2$ for all $w \in \mathfrak{I}(\theta)$, where ℓ^θ is the twisted absolute length function (see [17] for the definition). When W is finite, a different way to define the rank function was provided by Richardson and Springer [23]; the equivalence between the two formulations is due to Carter [11, Lemma 2]. In the case of W being a classical Weyl group and $\theta = \mathrm{id}$, Incitti [19, 20, 21] found the rank function using combinatorial arguments.

Example 3.7. With W and θ as in Example 3.2, it is readily seen that $\mathrm{Br}(\mathfrak{I}(\theta)) \cong \mathrm{Br}(W)$. Thus, ordinary Bruhat orders are a special case of this construction.

We now proceed to prove a number of facts that are completely analogous to familiar results from the theory of Coxeter groups.

Lemma 3.8. *For all* $w \in \mathfrak{I}(\theta)$, $s \in S$, *we have* $\rho(w\underline{s}) = \rho(w) \pm 1$, *and* $\rho(w\underline{s}) = \rho(w) - 1 \Leftrightarrow s \in D_R(w)$.

Proof. Since $w\underline{s}\underline{s} = w$, $|\rho(w\underline{s}) - \rho(w)| \leq 1$. Using Lemma 3.4, it is straightforward to verify from Definition 2.8 that $w > w\underline{s}$ iff $s \in D_R(w)$, and otherwise $w < w\underline{s}$. The lemma now follows from the fact that ρ is the rank function of $\mathrm{Br}(\mathfrak{I}(\theta))$. □

Lemma 3.9 (Lifting property for $\underline{S}$)**.** *Let* $s \in S$ *and* $v, w \in W$ *with* $v \leq w$, *and suppose* $s \in D_R(w)$. *Then,*

(i) $v\underline{s} \leq w$.

(ii) $s \in D_R(v) \Rightarrow v\underline{s} \leq w\underline{s}$.

Proof. Suppose $s \in D_R(v)$. If $v\underline{s} = vs$ and $w\underline{s} = \theta(s)ws$, Proposition 2.10 shows first that $vs \leq ws$, then $v = \theta(s)vs \leq ws$ and, finally, $v\underline{s} = \theta(s)v \leq \theta(s)ws = w\underline{s}$. The other cases admit similar proofs. ☐

Proposition 3.10 (Exchange property for $\mathfrak{I}(\theta)$)**.** *Suppose $\underline{s}_1 \dots \underline{s}_k$ is a reduced $\underline{S}$-expression and that $\rho(\underline{s}_1 \dots \underline{s}_k \underline{s}) < k$. Then, $\underline{s}_1 \dots \underline{s}_k \underline{s} = \underline{s}_1 \dots \widehat{\underline{s}_i} \dots \underline{s}_k$ for some $i \in [k]$.*

Proof. Let $w = \underline{s}_1 \dots \underline{s}_k$ and $v = \underline{s}_1 \dots \underline{s}_k \underline{s}$. We have $w > v$. Let $i \in [k]$ be maximal such that $v\underline{s}_k \dots \underline{s}_i > v\underline{s}_k \dots \underline{s}_{i+1}$ (it exists; otherwise we would have $\rho(v\underline{s}_k \dots \underline{s}_1) < 0$). Repeated application of Lemma 3.9 shows that $w\underline{s}_k \dots \underline{s}_{i+1} \geq v\underline{s}_k \dots \underline{s}_i$. Since $\rho(w\underline{s}_k \dots \underline{s}_{i+1}) = \rho(v\underline{s}_k \dots \underline{s}_i)$ and ρ is the rank function of $\mathrm{Br}(\mathfrak{I}(\theta))$, this implies $w\underline{s}_k \dots \underline{s}_{i+1} = v\underline{s}_k \dots \underline{s}_i$. Thus, $v = w\underline{s}_k \dots \underline{s}_{i+1}\underline{s}_i \dots \underline{s}_k = \underline{s}_1 \dots \widehat{\underline{s}_i} \dots \underline{s}_k$. ☐

Proposition 3.11 (Deletion property for $\mathfrak{I}(\theta)$)**.** *If $\rho(\underline{s}_1 \dots \underline{s}_k) < k$, then we have $\underline{s}_1 \dots \underline{s}_k = \underline{s}_1 \dots \widehat{\underline{s}_i} \dots \widehat{\underline{s}_j} \dots \underline{s}_k$ for some $1 \leq i < j \leq k$.*

Proof. Let $j \in [k]$ be minimal such that $\underline{s}_1 \dots \underline{s}_j$ is not reduced and apply Proposition 3.10 to this expression. ☐

4. The Bruhat order on twisted involutions

We now turn our attention to the poset $\mathrm{Br}(\mathfrak{I}(\theta))$. Incitti [19, 20, 21] showed that $\mathrm{Br}(\mathfrak{I}(\mathrm{id}))$ is EL-shellable and Eulerian whenever W is a classical Weyl group. He conjectured the same to hold for any Coxeter group W (when W is infinite, it should hold for every interval in $\mathrm{Br}(\mathfrak{I}(\mathrm{id}))$). In [17], it was proved that every interval in $\mathrm{Br}(\mathfrak{I}(\theta))$ is Gorenstein* over $\mathbb{Z}_2$ (in particular Eulerian) for arbitrary W and θ. The purpose of this section is to strengthen this result by showing that every interval in $\mathrm{Br}(\mathfrak{I}(\theta))$ is a PL sphere. Thus, the main consequence of Incitti's conjecture holds, although it is still not proved that the spheres actually are shellable.

A key poset property is that of admitting a *special matching*:

Definition 4.1. *Let P be a poset and $M : P \to P$ an involution such that for all $x \in P$, either x covers $M(x)$ or $M(x)$ covers x. Then, M is called a* special matching *iff for all $x, y \in P$ such that $M(x) \neq y$, it holds that*

$$y \text{ covers } x \Rightarrow M(x) < M(y).$$

The term "special matching" is due to Brenti, see [8]. In Eulerian posets, special matchings are equivalent to *compression labellings* in the sense of du Cloux [13]. The next result follows from [13, Theorem 3.5], which, in turn, is a reformulation of a result from Dyer's thesis [15]. It was reproved in the setting of Bruhat orders by Reading [22] (and his proof is easily adapted to the general situation).

Theorem 4.2 ([13, 15, 22])**.** *Let P be an Eulerian poset with a special matching M. If $(\hat{0}, M(\hat{1}))$ is a PL sphere, then so is $\overline{P}$.*

Corollary 4.3. *Suppose P is an Eulerian poset in which every lower interval $[\hat{0}, x]$, $x \neq \hat{0}$, has a special matching. Then, every open interval in P is a PL sphere.*

Proof. Links in PL spheres are PL spheres (see [5, Theorem 4.7.21.iv]). Therefore, it is enough to prove that $\overline{P}$ is a PL sphere. This follows from Theorem 4.2 by induction over the rank. $\square$

Remark 4.4. Explicitly requiring P to be Eulerian in Corollary 4.3 is not important. In fact, if every lower interval in a bounded poset P has a special matching, then P is necessarily Eulerian [9].

Theorem 4.5. *Let $w \in \mathfrak{I}(\theta)$, and suppose $s \in D_R(w)$. Then, the map $v \mapsto v\underline{s}$ is a special matching on the interval $[e, w] \subseteq \mathrm{Br}(\mathfrak{I}(\theta))$.*

Proof. Part (i) of Lemma 3.9 shows that $v \mapsto v\underline{s}$ maps $[e, w]$ to itself, and the fact that $v\underline{s}\underline{s} = v$ for all v shows that the map is an involution. As in the proof of Lemma 3.8, v and $v\underline{s}$ are always comparable. By rank considerations, one of them must therefore cover the other.

Now, pick $x, y \in [e, w]$ such that y covers x and consider Lemma 3.9. If $x\underline{s} < x$, part (ii) shows that $x\underline{s} < y\underline{s}$. If $x\underline{s} > x$ and $y\underline{s} < y$, we must have $x\underline{s} = y$ by part (i). Finally, if $x\underline{s} > x$ and $y\underline{s} > y$, then $x\underline{s} < y\underline{s}$, again by part (i). Thus, $v \mapsto v\underline{s}$ is a special matching. $\square$

Corollary 4.6. *The open intervals in* $\mathrm{Br}(\mathfrak{I}(\theta))$ *are PL spheres.*

It follows from Theorem 4.5 that the intervals in $\mathrm{Br}(\mathfrak{I}(\theta))$ are *accessible* posets as defined in [13]. It is known that not all accessible posets are Bruhat intervals, i.e. intervals in some $\mathrm{Br}(W)$. Interestingly, the smallest counterexamples (called *Dyer obstructions* in [13]) coincide with $\mathrm{Br}(\mathfrak{I}(\theta))$ when $W = A_3$ and θ is either the identity or the unique non-trivial Coxeter graph automorphism, respectively. This makes it natural to ask whether or not all accessible posets arise as intervals in Bruhat orders on twisted involutions.

5. A Coxeter complex analogue for twisted involutions

In this section, we construct a cell complex Δ_θ whose relationship with $\mathfrak{I}(\theta)$ has many features in common with the connection between the Coxeter complex Δ_W and W. Although some results also make sense for infinite groups, our main interest here is in the finite setting. Therefore, throughout the rest of the paper, (W, S) will be a *finite* Coxeter system with an involutive automorphism θ.

We define a graph G_θ on the vertex set $\mathfrak{I}(\theta)$ with edges labelled by elements in S as follows: there is an edge with label s between v and w iff $v\underline{s} = w$.

If we direct all edges according to decreasing ρ-values and merge multiple edges, we obtain from G_θ the Hasse diagram of the following partial order which was first defined in [23]:

Definition 5.1. *The* weak order *on $\mathfrak{I}(\theta)$ is the partial order $\preceq$ defined by $v \preceq w$ iff there exist $\underline{s}_1, \ldots, \underline{s}_k \in \underline{S}$ such that $v\underline{s}_1 \ldots \underline{s}_k = w$ and $\rho(v) = \rho(w) - k$. We denote this poset by $Wk(\theta)$.*

Observe that $\mathrm{Wk}(\theta)$ is a subposet of $\mathrm{Br}(\mathfrak{I}(\theta))$, i.e. the identity map $\mathrm{Wk}(\theta) \to \mathrm{Br}(\mathfrak{I}(\theta))$ is order-preserving.

It should be noted that $\mathrm{Wk}(\theta)$ does not in general coincide with the order induced by the (left or right) weak order on W. In fact, while the former is clearly graded with rank function ρ, the latter is not a graded poset in general.

Example 5.2. Return to the situation in Example 3.2. Observe that

$$(w, w^{-1})\underline{(s, e)} = (ws, (ws)^{-1}) \preceq (w, w^{-1}) \Leftrightarrow s \in D_R(w) \Leftrightarrow ws \leq_R w$$

and

$$(w, w^{-1})\underline{(e, s)} = (sw, (sw)^{-1}) \preceq (w, w^{-1}) \Leftrightarrow s \in D_L(w) \Leftrightarrow sw \leq_L w.$$

Hence, in this setting, $\mathrm{Wk}(\theta)$ is isomorphic to the transitive closure of the union of the left and right weak orders on W. This poset is sometimes called the *two-sided weak order* on W. It was studied by Björner in [3].

Given $J \subseteq S$, consider the subgraph of G_θ obtained by removing all edges with labels not in J. For $w \in \mathfrak{I}(\theta)$, let wC_J be the connected component which contains w in this subgraph. It should be stressed that we regard wC_J as an edge-labelled graph, not merely as a set of vertices. Define

$$P_\theta = \{wC_J \mid w \in \mathfrak{I}(\theta) \text{ and } J \subseteq S\}.$$

The elements of P_θ are partially ordered by reverse inclusion, i.e. $g_1 \leq g_2$ iff g_2 is a (labelled) subgraph of g_1.

Proposition 5.3. *The poset P_θ is the face poset of a pure Boolean cell complex Δ_θ of dimension $|S| - 1$.*

Proof. The bottom element of P_θ is G_θ. The maximal elements are the twisted involutions. Let $w \in \mathfrak{I}(\theta)$. The map $wC_J \mapsto J$ is easily seen to be a poset isomorphism from the interval $[G_\theta, w] = [wC_S, wC_\emptyset] \subseteq P_\theta$ to the dual of the Boolean lattice of subsets of S. $\square$

Remark 5.4. We briefly indicate why we regard Δ_θ as a Coxeter complex analogue. Suppose we replace $\mathfrak{I}(\theta)$ with W and G_θ with the Cayley graph of W (with respect to the generating set S). The connected component wC_J would then become the subgraph induced by the parabolic coset $w\langle J\rangle$. Ordering the set of such cosets by reverse inclusion produces the face poset of the Coxeter complex Δ_W. We refer to Brown's book [10] for a thorough background on Coxeter complexes.

Remark 5.5. The complex Δ_θ is not in general a simplicial complex. For example, if s and $\theta(s) \neq s$ commute for some $s \in S$, we have $\underline{s} = \overline{\theta(s)} = s\theta(s)$. Thus, there are two edges between e and $s\theta(s)$ in G_θ, implying that the facets in Δ_θ indexed by e and $s\theta(s)$ share two codimension 1 cells. Similar examples exist when $\theta = \mathrm{id}$.

Lemma 5.6. *Let $J \subseteq S$ and $w \in W$. Then, the vertex set of wC_J contains a unique Wk(θ)-minimal element* $\min(w, J)$.

Proof. Given $s \in J$, a $\mathrm{Wk}(\theta)$-minimal element in wC_J must clearly not have a reduced $\underline{S}$-expression ending in $\underline{s}$. Conversely, if $v \in wC_J$ is not $\mathrm{Wk}(\theta)$-minimal, there exists $s \in J$ such that $\rho(v\underline{s}) < \rho(v)$. Since $v\underline{s}\underline{s} = v$, we obtain a reduced $\underline{S}$-expression for v ending in $\underline{s}$ by attaching $\underline{s}$ to any reduced $\underline{S}$-expression for $v\underline{s}$. Thus, the $\mathrm{Wk}(\theta)$-minimal elements in wC_J are precisely the elements that have no reduced $\underline{S}$-expressions ending in $\underline{s}$ for all $s \in J$. It is clear that at least one such element exists.

Now suppose u and v are two $\mathrm{Wk}(\theta)$-minimal elements in wC_J. Let $\underline{s}_1 \dots \underline{s}_k$ be a reduced $\underline{S}$-expression for v. We may write

$$u = v\underline{s}_{k+1} \dots \underline{s}_{k+l} = \underline{s}_1 \dots \underline{s}_{k+l},$$

for some $s_{k+1}, \dots, s_{k+l} \in J$. By Proposition 3.11, this expression contains a reduced subexpression for u. Since it cannot end in any $\underline{s}_t$, $k+1 \le t \le k+l$, it must be a subexpression of $\underline{s}_1 \dots \underline{s}_k$. By symmetry, this subexpression must, in turn, contain a reduced subexpression for v. Thus, $u = v$. $\square$

The completely analogous Coxeter complex version of the next result is due to Björner [2, Theorem 2.1].

Theorem 5.7. *Any linear extension of $Wk(\theta)$ is a shelling order for Δ_θ. In particular, Δ_θ is shellable.*

Proof. Suppose $w_1, \dots, w_k$ (where $k = |\mathfrak{I}(\theta)|$) is a linear extension of $\mathrm{Wk}(\theta)$. Let $j \in [k]$, and suppose g is a cell in $w_i \cap w_j$ for some $i < j$. We must show that g is contained in a codimension 1 cell in $w_j \cap (\cup_{\alpha<j} w_\alpha)$. In terms of the graphs that represent the cells, the situation is this: $g = w_iC_J = w_jC_J$ for some $J \subseteq S$. It must be shown that g contains an edge connecting w_j with some w_α, $\alpha < j$. If J contains a right descent s of w_j, we can use $w_\alpha = w_j\underline{s}$. Otherwise, $w_j = \min(w, J)$. By Lemma 5.6, this implies $w_j \preceq w_i$, contradicting our choice of linear extension. $\square$

Corollary 5.8. *The complex Δ_θ is homeomorphic to the $(|S|-1)$-dimensional sphere.*

Proof. Since codimension 1 cells in Δ_θ correspond to edges in G_θ, Δ_θ is thin. The corollary now follows from Proposition 2.4. $\square$

We now define an analogue of the W-Eulerian polynomial (i.e. the generating function counting the elements of W with respect to the number of descents) for twisted involutions:

$$\mathrm{des}_\theta(x) = \sum_{w \in \mathfrak{I}(\theta)} x^{|D_R(w)|}.$$

Example 5.9. Again, consider the setting of Example 3.2. Let ϕ denote the natural bijection $W \to \mathfrak{I}(\theta)$. From the argument in Example 5.2, it follows that $|D_R(\phi(w))| = |D_R(w)| + |D_L(w)|$ for all $w \in W$. Thus, we have in this setting

$$\mathrm{des}_\theta(x) = \sum_{w \in W} x^{|D_R(w)|+|D_L(w)|} = \sum_{w \in W} x^{|D_R(w)|+|D_R(w^{-1})|}.$$

This can be viewed as a two-sided analogue of the W-Eulerian polynomial.

For Coxeter complexes, the counterpart of the next result is [7, Theorem 2.3].

Theorem 5.10. *The h-polynomial of Δ_θ coincides with $des_\theta(x)$.*

Proof. Consider the shelling order of Δ_θ given in the proof of Theorem 5.7. The unique minimal new cell introduced in the ith shelling step is $w_i C_J$, where J is the set of right ascents (i.e. non-descents) of w_i. The dimension of this cell is $|S|-|J|-1 = |D_R(w_i)| - 1$. For the h-vector of Δ_θ, this means that

$$h_j = |\{w \in \mathfrak{I}(\theta) \mid |D_R(w)| = j\}|.$$

Thus,

$$\mathrm{des}_\theta(x) = \sum_{j=0}^{|S|} h_j x^j.$$

□

A polynomial $P \in \mathbb{Z}[x]$ is called *symmetric* if $x^d P(x^{-1}) = P(x)$, where $d = \deg(P)$.

Corollary 5.11. *The polynomial $des_\theta(x)$ is symmetric.*

Proof. This is immediate from Theorem 5.10 and the Dehn-Sommerville equations. □

Remark 5.12. Suppose W is irreducible, and let w_0 denote the longest element in W. It is known (see [4, Exercise 4.10]) that $ww_0 = w_0 w$ for all $w \in W$ unless W is of one of the types $I_2(2n+1), A_n$, D_{2n+1} and E_6. Thus, in all other cases $w \mapsto ww_0$ is an involution $\mathfrak{I}(\theta) \to \mathfrak{I}(\theta)$ which sends ascents to descents, proving Corollary 5.11 for these cases. When $W = A_n$, $\theta = \mathrm{id}$, Corollary 5.11 is due to Strehl [27]. See also [14].

References

[1] A. Björner, Posets, regular CW complexes and Bruhat order, *European J. Combin.* **5** (1984), 7–16.

[2] A. Björner, Some combinatorial and algebraic properties of Coxeter complexes and Tits buildings, *Adv. Math.* **52** (1984), 173–212.

[3] A. Björner, The Möbius function of factor order, *Theoretical Computer Sci.* **117** (1993), 91–98.

[4] A. Björner and F. Brenti, *Combinatorics of Coxeter groups*, Graduate Texts in Mathematics, Springer Verlag, to appear.

[5] A. Björner, M. Las Vergnas, B. Sturmfels, N. White and G. Ziegler, *Oriented matroids (Second edition)*, Encyclopedia of mathematics and its applications **46**, Cambridge Univ. Press, 1999.

[6] A. Björner and M. Wachs, Bruhat order of Coxeter groups and shellability, *Adv. Math.* **43** (1982), 87–100.

[7] F. Brenti, q-Eulerian polynomials arising from Coxeter groups, *European J. Combin.* **15** (1994), 417–441.

[8] F. Brenti, Kazhdan-Lusztig polynomials: History, problems, and combinatorial invariance, *Sémin. Lothar. Combin.* **49** (2002), B49b, 30pp.

[9] F. Brenti, F. Caselli and M. Marietti, Kazhdan-Lusztig polynomials for partially ordered sets, in preparation.

[10] K. S. Brown, *Buildings*, Springer Monographs in Mathematics, New York, NY, 1989.

[11] R. W. Carter, Conjugacy classes in the Weyl group, *Comp. Math.* **25** (1972), 1–59.

[12] V. V. Deodhar, Some characterizations of Bruhat ordering on a Coxeter group and determination of the relative Möbius function, *Invent. Math.* **39** (1977), 187–198.
[13] F. du Cloux, An abstract model for Bruhat intervals, *European J. Combin.* **21** (2000), 197–222.
[14] W. M. B. Dukes, Permutation statistics for involutions, preprint 2004.
[15] M. J. Dyer, *Hecke algebras and reflections in Coxeter groups*, Ph. D. thesis, University of Sydney, 1987.
[16] A. Garsia and D. Stanton, Group actions on Stanley-Reisner rings and invariants of permutations groups, *Adv. Math.* **51** (1984), 107–201.
[17] A. Hultman, Fixed points of involutive automorphisms of the Bruhat order, *Adv. Math.*, to appear.
[18] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge studies in advanced mathematics **29**, Cambridge Univ. Press, 1990.
[19] F. Incitti, The Bruhat order on the involutions of the symmetric group, *J. Algebraic Combin.*, to appear.
[20] F. Incitti, The Bruhat order on the involutions of the hyperoctahedral group, *European J. Combin.* **24** (2003), 825–848.
[21] F. Incitti, *Bruhat order on the involutions of classical Weyl groups*, Ph. D. thesis, Università di Roma "La Sapienza", 2003.
[22] N. Reading, The cd-index of Bruhat intervals, *Electron. J. Combin.* **11(1)** (2004), R74, 25pp.
[23] R. W. Richardson and T. A. Springer, The Bruhat order on symmetric varieties, *Geom. Dedicata* **35** (1990), 389–436.
[24] R. W. Richardson and T. A. Springer, Complements to: The Bruhat order on symmetric varieties, *Geom. Dedicata* **49** (1994), 231–238.
[25] T. A. Springer, Some results on algebraic groups with involutions, *Advanced Studies in Pure Math.* **6**, 525–543, Kinokuniya/North-Holland, 1985.
[26] R. P. Stanley, f-vectors and h-vectors of simplicial posets, *J. Pure Appl. Algebra* **71** (1991), 319–331.
[27] V. Strehl, Symmetric Eulerian distributions for involutions, *Sémin. Lothar. Combin.* **1**, Strasbourg 1980. Publications de l'I.R.M.A. 140/S-02 , Strasbourg, 1981.

Department of mathematics, KTH, SE-100 44 Stockholm, Sweden
E-mail address: axel@math.kth.se

COMBINATORIAL INVARIANCE OF KAZHDAN-LUSZTIG POLYNOMIALS FOR SHORT INTERVALS IN THE SYMMETRIC GROUP

FEDERICO INCITTI

Abstract. The well-known *combinatorial invariance conjecture* states that, given a Coxeter group W, ordered by Bruhat order, and given two elements $x, y \in W$, with $x < y$, the Kazhdan-Lusztig polynomial, or equivalently the R-polynomial, associated with (x, y) supposedly depends only on the poset structure of the interval $[x, y]$.

In this paper we solve the conjecture for the first open cases, showing that it is true for intervals of length 5 and 6 in the symmetric group. The main tool is a pictorial way for describing the Bruhat order in the symmetric group, namely the *diagram* of a pair of permutations. It is shown how the diagram of (x, y) allows to get information about the poset structure of $[x, y]$, and about the R-polynomial associated with (x, y). As a parallel result, we obtain expressions of the R-polynomials for some general classes of pairs of permutations.

1. Introduction

In their fundamental paper [10] Kazhdan and Lusztig defined, for every Coxeter group W, a family of polynomials with integer coefficients, indexed by pairs of elements of W. These polynomials, known as *Kazhdan-Lusztig polynomials* of W, are related to the algebraic geometry and topology of Schubert varieties. They also play a crucial role in representation theory. In order to prove the existence of these polynomials Kazhdan and Lusztig used another family of polynomials which arises from the multiplicative structure of the Hecke algebra associated with W. These polynomials are known as *R-polynomials* of W.

A famous conjecture concerning Kazhdan-Lusztig polynomials, the so-called *combinatorial invariance conjecture*, states that, given two Coxeter groups W_1 and W_2, partially ordered by Bruhat order, and given elements $x, y \in W_1$, with $x < y$, and $u, v \in W_2$, with $u < v$, if the two intervals $[x, y]$ and $[u, v]$ are isomorphic as posets, then $P_{x,y}(q) = P_{u,v}(q)$. This conjecture is known to be true if $[x, y]$ is a lattice and holds for intervals up to length 4. Recently it has been proved that it is true for $x = u = e$ (see [3]).

In this paper we solve the conjecture for the first open cases, showing that it is true for intervals of length 5 and 6 in the symmetric group. The main tool is a pictorial way for describing the Bruhat order in the symmetric group, namely the *diagram* of a pair of permutations.

In §2 we collect some basic notions and in §3 we state the main result. Then we introduce the diagram of a pair of permutations (§4). We develop this tool, already introduced in [9], showing how the diagram of (x, y) allows to get information about the poset structure of $[x, y]$ (§5), and about the R-polynomial associated with (x, y) (§6). As a parallel result, we obtain expressions of the R-polynomials for some general classes of pairs of permutations. After a classification of the intervals of the symmetric group up to length 5 (§7), we give the proof of the main result in §8.

2. Preliminaries

We let $\mathbf{N} = \{1, 2, 3, \dots\}$. For $n \in \mathbf{N}$ we let $[n] = \{1, 2, \dots, n\}$ and for $n, m \in \mathbf{N}$, with $n \leq m$, we let $[n, m] = \{n, n+1, \dots, m\}$. We refer to [11] for poset theory. Given a poset P, we denote by $\lhd$ the covering relation. The *Hasse diagram* of P is the directed graph having P as vertex set and such that there is an edge from x to y if and only if $x \lhd y$. 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 refer to [8] for basic notions about Coxeter groups. Given a Coxeter group W, with set of generator S, the set of *reflections* 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 can be written as a product of k generators. The *Bruhat graph* of W, denoted by $BG(W)$ (or simply BG) 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 $\ell(x) < \ell(y)$. The edge is often supposed labelled by the reflection t:

$$x \xrightarrow{t} y.$$

Finally, the *Bruhat order* of W is the partial order which is the transitive closure of BG: given $x, y \in W$, $x \leq y$ in the Bruhat order if and only if there is a chain $x = x_0 \to x_1 \to \cdots \to x_k = y$. It is known that W, partially ordered by the Bruhat order, is a graded poset with rank function given by the length. Given $x, y \in W$, with $x < y$, we set $\ell(x, y) = \ell(y) - \ell(x)$.

We now define R-polynomials and Kazhdan-Lusztig polynomials.

Theorem 2.1. *There exists a unique family of polynomials $\{R_{x,y}(q)\}_{x,y\in W} \subseteq \mathbb{Z}[q]$ satisfying the following conditions:*

(1) $R_{x,y}(q) = 0$, *if $x \not\leq y$;*
(2) $R_{x,y}(q) = 1$, *if $x = y$;*
(3) *if $x < y$ and $s \in S$ is such that $ys \lhd y$ then*

$$R_{x,y}(q) = \begin{cases} R_{xs,ys}(q), & \text{if} \quad xs \lhd x, \\ (q-1)R_{x,ys}(q) + qR_{xs,ys}(q), & \text{if} \quad xs \rhd x. \end{cases}$$

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 [8, §§7.4, 7.5]. The polynomials whose existence and uniqueness are guaranteed by Theorem 2.1 are called the *R-polynomials* of W.

Theorem 2.2. *There exists a unique family of polynomials $\{P_{x,y}(q)\}_{x,y\in W} \subseteq \mathbb{Z}[q]$ satisfying the following conditions:*

(1) $P_{x,y}(q) = 0$, *if $x \not\leq y$;*
(2) $P_{x,y}(q) = 1$, *if $x = y$;*
(3) *if $x < y$ then* $\deg(P_{x,y}(q)) < \ell(x,y)/2$ *and*

$$q^{\ell(x,y)} P_{x,y}\left(q^{-1}\right) - P_{x,y}(q) = \sum_{x<z\leq y} R_{x,z}(q)\, P_{z,y}(q).$$

A proof of Theorem 2.2 appears in [8, §§7.9, 7.10, 7.11]. The polynomials whose existence and uniqueness are guaranteed by Theorem 2.2 are called the *Kazhdan-Lusztig polynomials* of W. Theorem 2.2 ensures that knowing the R-polynomials is equivalent to knowing the Kazhdan-Lusztig polynomials. In fact part (3) can be recursively used to compute one family from the other, by induction on $\ell(x, y)$.

The R-polynomials satisfy the following relation (see, e.g., [1, Exercise 5.11]).

Proposition 2.3. *Let $x, y \in W$, with $x < y$. Then*

$$\sum_{x\leq z\leq y} (-1)^{\ell(x,z)} R_{x,z}(q) R_{z,y}(q) = 0.$$

In particular, if $\ell(x, y)$ is even, we have

$$R_{x,y}(q) = \frac{1}{2} \sum_{\substack{x<z<y \\ \ell(x,z) \text{ odd}}} R_{x,z}(q)R_{z,y}(q) - \frac{1}{2} \sum_{\substack{x<z<y \\ \ell(x,z) \text{ even}}} R_{x,z}(q)R_{z,y}(q) \tag{1}$$

Note that equation (1) allows to compute the R-polynomial associated with an interval $[x, y]$ of even length ℓ, once the R-polynomials associated with the subintervals of $[x, y]$ of length $< \ell$ are known. In particular, if the combinatorial invariance of the R-polynomials is true for intervals of length $< \ell$, then it is true also for intervals of length ℓ.

In order to give a combinatorial interpretation of the R-polynomials, another family of polynomials, known as the *$\widetilde{R}$-polynomials*, has been introduced.

Proposition 2.4. *Let $x, y \in W$. Then there is a unique polynomial $\widetilde{R}_{x,y}(q) \in \mathbb{N}[q]$ such that*

$$R_{x,y}(q) = q^{\frac{\ell(x,y)}{2}} \widetilde{R}_{x,y}\left(q^{\frac{1}{2}} - q^{-\frac{1}{2}}\right).$$

The advantage of the $\widetilde{R}$-polynomials over the R-polynomials is that they have non-negative integer coefficients. In fact there is a nice combinatorial interpretation for them. In order to give that, we introduce some notation.

Given $x, y \in W$, with $x < y$, we denote by $Paths(x, y)$ the set of paths in BG from x to y. The *length* of $\Delta = (x_0, x_1, \dots, x_k) \in Paths(x, y)$, denoted by $|\Delta|$, is the number k of its edges. Now let $\prec$ be a fixed *reflection ordering* on the set T of reflections (see, e.g., [1, §5.2] for the definition and for a proof of existence). A path $\Delta = (x_0, x_1, \dots, x_k) \in Paths(x, y)$, 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 the order $\prec$ if $t_1 \prec t_2 \prec \cdots \prec t_k$. Denote by $Paths^{\prec}(x, y)$ the set of all paths in $Paths(x, y)$ which are increasing with respect to $\prec$. The following result is due to Dyer [5].

Theorem 2.5. *Let W be a Coxeter and let $x, y \in W$, with $x < y$. Set $\ell = \ell(x, y)$. Then*

$$\widetilde{R}_{x,y}(q) = \sum_{k=1}^{\ell} c_k q^k, \tag{2}$$

where

$$c_k = |\{\Delta \in Paths^{\prec}(x, y) : |\Delta| = k\}|, \quad \textit{for every } k \in [\ell].$$

Equation (2) can be refined. Every path from x to y in BG necessarily has a length of the same parity of ℓ, so $c_k = 0$ if $k \not\equiv \ell\,(2)$. Furthermore, by the EL-shellability of the Bruhat order (see, e.g., [5]), in BG there is exactly one increasing path from x to y of length ℓ, thus $c_\ell = 1$.

Finally, let us introduce the notion of absolute length of a pair. We recall that the *absolute length* of $x \in W$, denoted by $a\ell(x)$, is defined as the minimum k such that x can be written as the product of k reflections. By [6, Theorem 1.2], $a\ell(x)$ is the (oriented) distance between e and x in BG. And by [6, Theorem 2.3], the coefficient of q^k in $\widetilde{R}_{x,y}(q)$ is non-zero if and only if there is a path in BG of length k. These facts suggest the following definition.

Definition 2.6. *Let $x, y \in W$, with $x < y$. The* absolute length *of (x, y), denoted by $a\ell(x, y)$, is the (oriented) distance between x and y in BG.*

By the results in [6], this notion generalizes that of absolute length of an element. In fact, given $x \in W$, we have $a\ell(x) = a\ell(e, x)$. On the other hand, by [2, Proposition 6.1], there exists $m \equiv \ell(x, y)$ (2) such that the coefficient of q^k in $\widetilde{R}_{x,y}(q)$ is non-zero if and only if $k \in [m, \ell(x, y)]$. It turns out that this m is exactly $a\ell(x, y)$. Putting all together, we have the following.

Corollary 2.7. *Let W be a Coxeter and let $x, y \in W$, with $x < y$. Set $\ell = \ell(x, y)$ and $a\ell = a\ell(x, y)$. Then*

$$\widetilde{R}_{x,y}(q) = q^{\ell} + c_{\ell-2}\, q^{\ell-2} + \cdots + c_{a\ell+2}\, q^{a\ell+2} + c_{a\ell}\, q^{a\ell}.$$

where

$$c_k = |\{\Delta \in Paths^{\prec}(x, y) : |\Delta| = k\}| \geq 1$$

for every $k \in [a\ell, \ell - 2]$, with $k \equiv \ell$ (2).

It is worth noting that the relation $x \to y$ is determined by the poset structure of $[x, y]$ (see [4]). Combining that with the results in [6], we have the following.

Proposition 2.8. *Let W be a Coxeter group and let $x, y \in W$, with $x < y$. The absolute length $a\ell(x, y)$ is a combinatorial invariant, that is, it depends only on the poset structure of $[x, y]$.*

The following is the well-known *combinatorial invariance conjecture*, which was stated by Lusztig and independently by Dyer in the 1980s, and since then it has remained unsolved.

Conjecture 2.9. *Let W_1 and W_2 be two Coxeter groups and let $x, y \in W_1$, with $x < y$, and $u, v \in W_2$, with $u < v$. Then*

$$[x, y] \cong [u, v] \quad \Rightarrow \quad P_{x,y}(q) = P_{u,v}(q).$$

In other words, the Kazhdan-Lusztig polynomial associated with (x, y) supposedly depends only on the poset structure of $[x, y]$. The combinatorial invariance conjecture is equivalent to the analogous statement for the R-polynomials, by Theorem 2.2, and for the $\widetilde{R}$-polynomials, by Proposition 2.4. It is known to be true if $[x, y]$ is a lattice (see [2, Theorem 6.3]) and holds for intervals up to length 4 (see, e.g., [1]). In this paper we prove that this is true for intervals of the symmetric group of length 5 and 6.

3. Main result

As usual, we denote by S_n the *symmetric group* over n elements, that is the set of all bijections of $[n]$ onto itself, and call its elements *permutations*. The symmetric group S_n is known to be a Coxeter group, with generators given by the simple transpositions $(i, i+1)$, for $i \in [n-1]$.

Given a poset P and $x, y \in P$, with $x < y$, we denote by $a(x, y)$ and $c(x, y)$, respectively, the number of *atoms* and *coatoms* of the interval $[x, y]$, that is

$$a(x, y) = |\{z \in [x, y] : x \lhd z\}| \quad \text{and} \quad c(x, y) = |\{z \in [x, y] : z \lhd y\}|,$$

and by $cap(x, y)$ its *capacity*, that is

$$cap(x, y) = \min\{a(x, y), c(x, y)\}.$$

We also denote by $\mathcal{B}_k$ the *boolean algebra* of rank k, that is the family $\mathcal{P}([k])$ of all subsets of $[k]$ partially ordered by inclusion. The following is the main result of this paper.

Theorem 3.1. *Let $x, y \in S_n$, for some n, with $x < y$ and $\ell(x, y) = 5$. Set $a = a(x, y)$, $c = c(c, y)$ and $cap = cap(x, y)$. Then*

$$\widetilde{R}_{x,y}(q) = \begin{cases} q^5 + 2q^3 + q, & \text{if} \quad \{a, c\} = \{3, 4\}, \\ q^5 + 2q^3, & \text{if} \quad a = c = 3, \\ q^5 + q^3, & \text{if} \quad cap \in \{4, 5\} \text{ but } [x, y] \not\cong \mathcal{B}_5, \\ q^5, & \text{if} \quad cap \in \{6, 7\} \text{ or } [x, y] \cong \mathcal{B}_5. \end{cases}$$

A sketch of the proof of Theorem 3.1 will be given in §8. As a consequence of it, we have the combinatorial invariance of Kazhdan-Lusztig polynomials for intervals of length 5 and 6 intervals in the symmetric group, as we state in the following.

Corollary 3.2. *Let $x, y \in S_n$, with $x < y$, and $u, v \in S_m$, with $u < v$, for some n and m, be such that $\ell(x, y) = \ell(u, v) \in \{5, 6\}$. Then*

$$[x, y] \cong [u, v] \quad \Rightarrow \quad P_{x,y}(q) = P_{u,v}(q).$$

Proof. Theorem 3.1 and Proposition 2.4 imply the combinatorial invariance of the R-polynomials for intervals of length 5. For intervals of length 6, it follows from equation (1). Finally, by Theorem 2.2, the assertion for the Kazhdan-Lusztig polynomials follows. □

4. Diagram of a pair of permutations

To denote a permutation $x \in S_n$ we often use the *one-line notation*: we write $x = x_1 x_2 \dots x_n$, to mean that $x(i) = x_i$ for every $i \in [n]$. The *diagram* of a permutation $x \in S_n$ is the subset of $\mathbf{N}^2$ so defined:

$$Diag(x) = \{(i, x(i)) : i \in [n]\}$$

The symmetric group is a Coxeter group and there is a nice combinatorial characterization of the Bruhat order relation in it. In order to give that, we introduce the following notation: for $x \in S_n$ and $(h, k) \in [n]^2$, we set

$$x[h, k] = |\{i \in [h] : x(i) \in [k, n]\}, \tag{3}$$

that is, $x[h,k]$ is the number of points of the diagram of x lying in the upper-left quarter plane with origin at (h,k). And given $x, y \in S_n$ and $(h,k) \in [n]^2$, we set

$$(x,y)[h,k] = y(h,k) - x(h,k). \tag{4}$$

The characterization is the following.

Theorem 4.1. *Let $x, y \in S_n$. Then*

$$x \leq y \quad \Leftrightarrow \quad (x,y)[h,k] \geq 0, \quad \text{for every } (h,k) \in [n]^2.$$

It is useful to extend the notation introduced in (3) and (4) to every $(h,k) \in \mathbf{R}^2$. We call the mapping $(h,k) \mapsto (x,y)[h,k]$, which associates with every $(h,k) \in \mathbf{R}^2$ the integer $(x,y)[h,k]$, the *multiplicity mapping* of the pair (x,y). Theorem 4.1 can be reformulated as follows:

$$x \leq y \quad \Leftrightarrow \quad (x,y)[h,k] \geq 0, \quad \text{for every } (h,k) \in \mathbf{R}^2.$$

Definition 4.2. *Let $x, y \in S_n$. The* diagram *of the pair (x,y) is the collection of:*

(1) *the diagram of x;*
(2) *the diagram of y;*
(3) *the multiplicity mapping $(h,k) \mapsto (x,y)[h,k]$.*

We pictorially represent the diagram of a pair (x,y) with the following convention: the diagram of x (respectively, y) is denoted by black dots (respectively, white dots) and, if $x < y$, then 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 with a lower integer corresponds a lighter grey. For example, the diagram of the pair $(315472986, 782496315)$ is depicted in Figure 1.

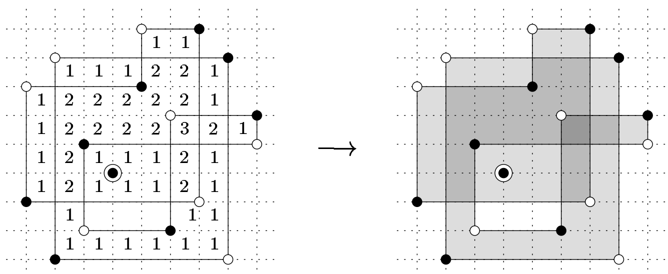


FIGURE 1. Diagram of a pair of permutations.

Finally, it is useful to give the following definition.

Definition 4.3. *Let $x, y \in S_n$, with $x < y$. The* support *of the pair (x,y) is*

$$\Omega(x,y) = \{(h,k) \in \mathbf{R}^2 : (x,y)[h,k] > 0\}.$$

5. FROM THE DIAGRAM TO THE POSET STRUCTURE

In this section we show how it is possible, starting from the diagram of a pair (x,y), with $x < y$, to get information about the poset structure of the interval $[x,y]$.

5.1. **Symmetries.** In general Coxeter groups, it is known that the mapping $x \mapsto x^{-1}$ is an isomorphism of the Bruhat order. Also, finite Coxeter groups always have a maximum, usually denoted w_0, and the mappings $x \mapsto xw_0$ and $x \mapsto w_0x$ are anti-isomorphisms. It follows that $x \mapsto w_0xw_0$ is an isomorphism.

In the symmetric group these facts can be described in a nice pictorial way. In fact the maximum of S_n is $w_0 = n\, n-1 \, \ldots \, 2\, 1$, and given $x \in S_n$, the diagrams of x^{-1}, xw_0, w_0x can be respectively obtained from the diagram of x by a reflection with respect to the lines $\{(h,k) \in \mathbf{R}^2 : h = k\}$, $\{(h,k) \in \mathbf{R}^2 : h = (n+1)/2\}$ and $\{(h,k) \in \mathbf{R}^2 : k = (n+1)/2\}$. The diagram of w_0xw_0 can be obtained from that of x by a reflection with respect to the point $((n+1)/2, (n+1)/2)$.

So, given $x, y \in S_n$, with $x < y$, there are four trivial associated intervals belonging to the same isomorphism class, and four belonging to the dual class. This means that, in order to study all possible isomorphism types that can occur in the symmetric group, it is enough to consider diagrams of pairs of permutations up to these symmetries.

5.2. **Atoms and coatoms.** A combinatorial characterization of the covering relation in the Bruhat order of the symmetric group is given in terms of free rises. We recall that, given $x \in S_n$, a *free rise* of x is a pair (i, j), with $i < j$ and $x(i) < x(j)$, such that there is no $k \in \mathbf{N}$, with $i < k < j$ and $x(i) < x(k) < x(j)$. It is known that, given $x, y \in S_n$, then $x \lhd y$ if and only if $y = x(i, j)$, where (i, j) is a free rise of x.

If (i, j) is a free rise of x, then the *rectangle associated* with (i, j) is:

$$Rect_x(i,j) = \{(h,k) \in \mathbf{R}^2 : i \leq h < j \,,\, x(i) < k \leq x(j)\}.$$

Now let $x, y \in S_n$, with $x < y$. In order to describe the atoms of the interval $[x, y]$, we say that a free rise (i, j) of x is *good* with respect to y if

$$Rect_x(i,j) \subseteq \Omega(x,y).$$

The following gives a characterization of the atoms of an interval.

Proposition 5.1. *Let $x, y \in S_n$, with $x < y$. Then z is an atom of $[x, y]$ if and only if $z = x(i, j)$, where (i, j) is a free rise of x good with respect to y.*

Proof. Let (i, j) be a free rise of x, and let $z = x(i, j)$. We know that $x \lhd z$. Thus z is an atom of $[x, y]$ if and only if $z \leq y$. As it can be easily checked, for every $(h, k) \in \mathbf{R}^2$, we have

$$(z,y)[h,k] = \begin{cases} (x,y)[h,k] - 1, & \text{if } (h,k) \in Rect_x(i,j), \\ (x,y)[h,k], & \text{otherwise.} \end{cases}$$

Then, by Theorem 4.1, $z \leq y$ if and only if $(x, y)[h, k] \geq 1$ for every $(h, k) \in Rect_x(i, j)$, that is if and only if $Rect_x(i, j) \subseteq \Omega(x, y)$. □

In a specular way, we can define a *free inversion* of y, as a pair (i, j), with $i < j$ and $y(i) > y(j)$, such that there is no $k \in \mathbf{N}$, with $i < k < j$ and $y(i) > y(k) > y(j)$. Note that $x \lhd y$ if and only if $x = y(i, j)$, where (i, j) is a free inversion of y. The *rectangle associated* with (i, j) is:

$$Rect_y(i,j) = \{(h,k) \in \mathbf{R}^2 : i \leq h < j \,,\, y(j) < k \leq y(i)\}.$$

Given $x, y \in S_n$, with $x < y$, we say that a free inversion (i, j) of y is *good* with respect to x if

$$Rect_y(i,j) \subseteq \Omega(x,y).$$

Next result gives a characterization of the coatoms of an interval, and its proof is completely specular to that of Proposition 5.1.

Proposition 5.2. *Let $x, y \in S_n$, with $x < y$. Then w is a coatom of $[x, y]$ if and only if $w = y(i, j)$, where (i, j) is a free inversion of y good with respect to x.*

Going back to the example shown in Figure 1, the free rises of x good with respect to y, and the free inversions of y good with respect to x are illustrated in Figure 2. We can conclude that

$$a(x,y) = c(x,y) = 11.$$

6. From the diagram to the $\widetilde{R}$-polynomial

By Theorem 2.5, computing an $\widetilde{R}$-polynomial means computing the number of increasing chains in BG with respect to a given reflection ordering on the reflections. In this section we show how it can be done for an interval $[x, y]$ of the symmetric group, in terms of the diagram of (x, y).

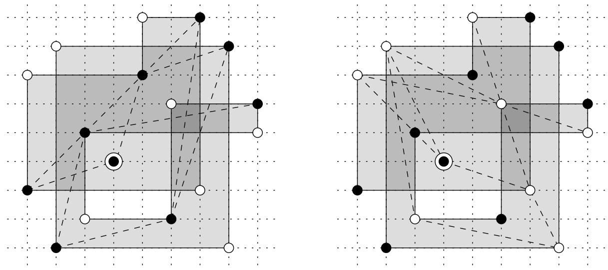


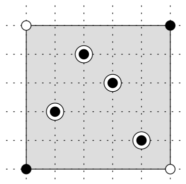
FIGURE 2. Atoms and coatoms.

6.1. **Symmetries.** Given a Coxeter group W, and given $x, y \in W$, with $x < y$, it is known (see, e.g., [1]) that $\widetilde{R}_{x,y}(q) = \widetilde{R}_{x^{-1},y^{-1}}(q)$ and, if W is finite, $\widetilde{R}_{x,y}(q) = \widetilde{R}_{yw_0,xw_0}(q) = \widetilde{R}_{w_0y,w_0x}(q) = \widetilde{R}_{w_0xw_0,w_0yw_0}(q)$. In the symmetric group this means that the eight pairs obtained from (x, y) by the symmetries previously described, have all the same $\widetilde{R}$-polynomial associated. Thus, in order to list all possible $\widetilde{R}$-polynomials that can occur in the symmetric group, it is enough to consider diagrams up to those symmetries.

6.2. **The *stair method*.** It is known that in the symmetric group S_n the reflections are the transpositions ($T = \{(i,j) : i,j \in [n]\}$) and that a possible reflection ordering on them is the lexicographic order (see, e.g., [1]). From now on we always assume this order fixed on the transpositions. For instance, in S_4:

$$(1,2) \prec (1,3) \prec (1,4) \prec (2,3) \prec (2,4) \prec (3,4).$$

Let $x, y \in S_n$, with $x < y$. We start considering the case in which (x, y) is an edge of BG, that is $y = x(i, j)$ for some transposition (i, j). In this case the support is a rectangle. A possible diagram of (x, y) is, for example, the following:



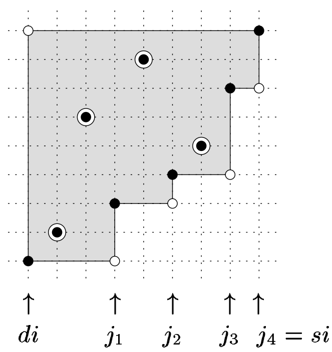
Now let (x, y) be not necessarily an edge of BG. There is a nice way to describe the increasing paths in BG from x to y. Suppose given such a path:

$$x = x_0 \xrightarrow{t_1} x_1 \xrightarrow{t_2} \cdots \xrightarrow{t_k} x_k = y. \tag{5}$$

We recall that the *difference index* of x with respect to y, denoted by di, is the minimal k such that $x(k) \neq y(k)$. We also call $x^{-1}y(di)$ the *stair index* of x with respect to y, and denote it by si. Note that, by definition, $di < si$. Also note that in (5) necessarily $t_1 = (di, j_1)$, for some j_1, otherwise the path would not be increasing. We now consider the case $t_h = (di, j_h)$, for every $h \in [k]$, with $j_k = si$:

$$x = x_0 \xrightarrow{(di,j_1)} x_1 \xrightarrow{(di,j_2)} \cdots \xrightarrow{(di,j_{k-1})} x_{k-1} \xrightarrow{(di,si)} x_k = y.$$

We call such a path a *stair path*, because of the shape of the support: it is a stair with k steps. Here is an example:



In the general case, the increasing path in (5) is a sequence of stair paths. It is useful to give the following definitions, which formalize the notion of stair.

Definition 6.1. *Let $x \in S_n$. A* stair *of x is a sequence $s = (j_0, j_1, \ldots, j_k) \in [n]^k$, which is increasing and such that $(x(j_0), x(j_1), \ldots, x(j_k))$ is also increasing. The* stair area *associated with s is the subset of $\mathbf{R}^2$ so defined:*

$$Stair_x(s) = \bigcup_{i \in [k]} \{(a, b) \in \mathbf{R}^2 : j_{i-1} \leq a < j_i \,,\, x(j_{i-1}) < b \leq x(j_i)\}.$$

We also say that the permutation $x(j_0, j_k, \ldots, j_1)$ is obtained *from x by* performing *the stair s.*

Definition 6.2. *Let $x, y \in S_n$, with $x < y$. A stair s of x is said to be* good *with respect to y if*

$$Stair_x(s) \subseteq \Omega(x, y)$$

Then we have the following, whose proof is similar to that of Proposition 5.1.

Proposition 6.3. *Let $x, y \in S_n$, with $x < y$. Let $s = (j_0, j_1, \ldots, j_k)$ be a stair of x. Then $x(j_0, j_k, \ldots, j_1) \leq y$ if and only if s is good with respect to y.*

An increasing path from x to y is obtained by performing subsequent stairs, choosing at each step one of the leftmost good stairs. Next definition gives the last tool we need.

Definition 6.4. *Let $x, y \in S_n$, with $x < y$. An* initial stair *of (x, y) is a stair s of x good with respect to y which starts with di and ends with si:*

$$s = (di, j_1, j_2, \ldots, j_{k-1}, si).$$

Examples of initial stairs will be shown in Figure 3. Note that an initial stair of (x, y) always exists. One can be obtained, for example, by choosing its indices as small as possible.

We can finally give a general algorithm, which allows to generate all possible increasing paths in BG from x to y, and thus to compute the coefficients of the $\widetilde{R}$-polynomial. We call it the *stair method*. Starting from the diagram of (x, y), all the increasing paths from x to y can be recursively constructed as follows:

(1) choose an initial stair of (x, y);
(2) call x_1 the permutation obtained from x by performing the choosen stair (note that $x_1 \leq y$, by Proposition 6.3);
(3) recursively apply the procedure on (x_1, y).

An example is shown in Figure 3. Here we take the example introduced in Figure 1, and apply the algorithm to that, making some choices. The corresponding increasing path in BG which arises has length 9 and is:

$$x \xrightarrow{(1,4)} \bullet \xrightarrow{(1,5)} x_1 \xrightarrow{(2,3)} \bullet \xrightarrow{(2,8)} x_2 \xrightarrow{(3,6)} x_3 \xrightarrow{(4,5)} x_4 \xrightarrow{(5,7)} x_5 \xrightarrow{(6,8)} \bullet \xrightarrow{(6,9)} y.$$

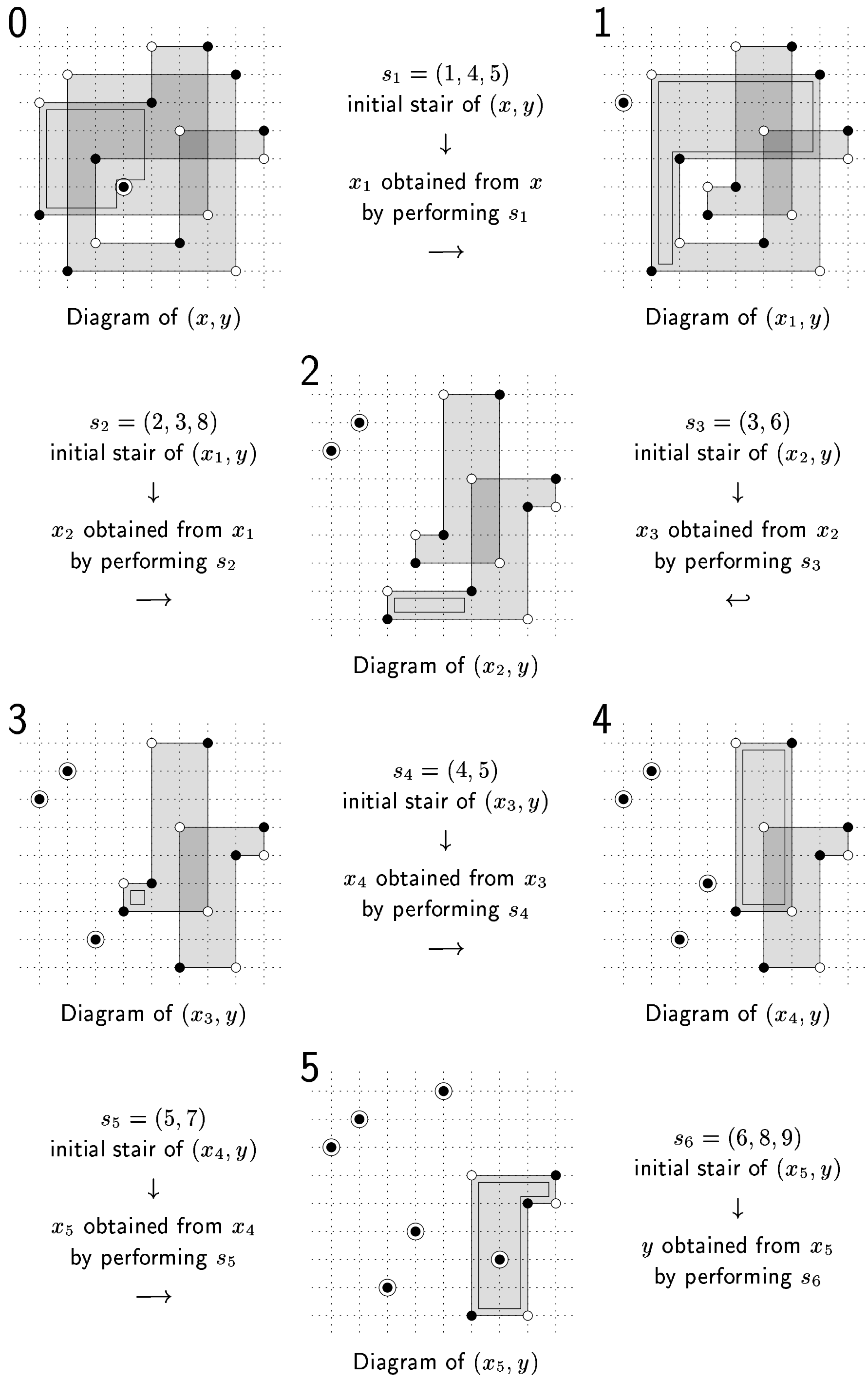


FIGURE 3. The stair method.

We could have choosen, for example, the initial stairs $(1, 5)$ or $(1, 3, 5)$ of (x, y), instead of $(1, 4, 5)$, or the initial stair $(6, 7, 8, 9)$ of (x_5, y), instead of $(6, 8, 9)$. Considering all possible choices, we can obtain all increasing paths from x to y, and it turns out that they are 10 and that:

$$\widetilde{R}_{x,y}(q) = q^{13} + 4q^{11} + 4q^9 + q^7.$$

Note that the only increasing path in BG from x to y of length $\ell(x,y)$, whose existence and uniqueness is guaranteed by the EL-shellability, is obtained by choosing in the algorithm always the lexicographically minimal stair.

6.3. **Special cases.** The stair method allows to compute the $\widetilde{R}$-polynomial for some general classes of pairs (x,y).

Definition 6.5. *Let $x,y \in S_n$, with $x < y$. We say that*

(1) (x,y) *has the* 01-multiplicity property *if*
$$(x,y)[h,k] \in \{0,1\} \quad \text{for every } (h,k) \in \mathbf{R}^2;$$
(2) (x,y) *is* simple *if it has the* 01*-multiplicity property and* $\{i \in [n] : x(i) = y(i)\} = \emptyset$*;*
(3) (x,y) *is a* permutaomino *if it is simple and* $\Omega(x,y)$ *is connected.*

Examples are shown in Figure 4.

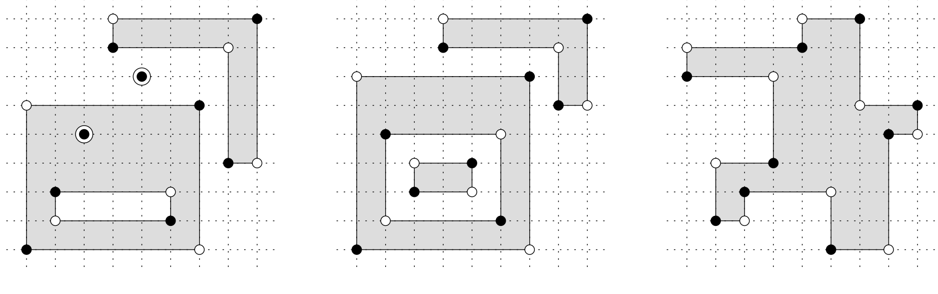


FIGURE 4. Special pairs of permutations.

Definition 6.6. *Let $x,y \in S_n$, with $x < y$, be such that (x,y) satisfies the* 01*-multiplicity property. Then the* fixed point multiplicity *of (x,y) is*
$$fpm(x,y) = |\{i \in [n] : x(i) = y(i),\ (x,y)[i,x(i)] = 1\}|.$$

Here we have the expressions of the $\widetilde{R}$-polynomial for these classes of pairs.

Proposition 6.7. *Let $x,y \in S_n$, with $x < y$.*

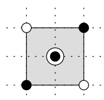
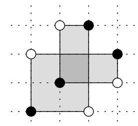
(1) *If (x,y) has the* 01*-multiplicity property, then*
$$\widetilde{R}_{x,y}(q) = (q^2+1)^{fpm(x,y)} q^{a\ell(x,y)},$$
(2) *If (x,y) is simple then*
$$\widetilde{R}_{x,y}(q) = q^{\ell(x,y)},$$
(3) *If (x,y) is a permutaomino then*
$$\widetilde{R}_{x,y}(q) = q^{n-1},$$

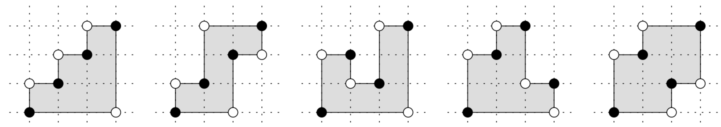
A particular case in which (x,y) has the 01-multiplicity property, is when (x,y) is an edge of BG. Then we get the following.

Proposition 6.8. *Let $x,y \in S_n$, with $x \to y$ and $\ell(x,y) = 2p+1$. Then*
$$\widetilde{R}_{x,y} = q(q^2+1)^p.$$

7. Intervals of length ≤ 5

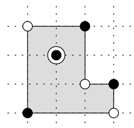
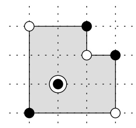
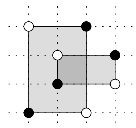
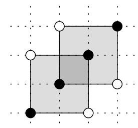
7.1. **Length 3.** The diagrams of pairs (x, y), with $\ell(x, y) = 3$, which are not trivially reducible, are, up to symmetries, exactly the following:

the *eye* diagram:  the *rabbit* diagram: 

permutaominos with 8 edges: 

It is known that only the so called k-crowns can occur as intervals of length 3 in Coxeter groups, and in particular in the symmetric group there are only 2, 3 and 4-crowns. Namely, the eye diagram corresponds to a 2-crown, and the last permutaomino to a 4-crown. The other permutaominoes and the rabbit diagram correspond to a 3-crown. Finally, $\widetilde{R}_{x,y}(q) = q^3 + q$, for the eye diagram, and q^3 in all other cases.

7.2. **Length 4.** Among the diagrams corresponding to length 4 intervals, we mention all the permutaominos with 10 edges. And the following four:

1.  2.  3.  4. 

We call them the *essential diagrams* of length 4. They are, up to symmetries, the only ones obtained by "enlarging" the eye diagram in S_4, which are not trivially reducible.

For general length 4 intervals, as poset types we have the products of the 2, 3 and 4-crown times $\{0, 1\}$, and two irreducible ones. And the $\widetilde{R}$-polynomial can be either $q^4 + q^2$ or q^2.

7.3. **Length 5.** For completeness, even if we don't use it in the proof of our main result, we mention that in [7] all the intervals of length 5 occurring in the symmetric group have been listed. They have been generated using a Maple package by J. R. Stembridge.

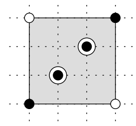
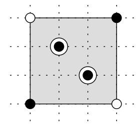
8. Sketch of proof of the main result

Sketch of proof of Theorem 3.1. Suppose known the poset structure of $[x, y]$, with $\ell(x, y) = 5$. We want show that it allows to determine the polynomial $\widetilde{R}_{x,y}(q)$.

By Proposition 2.8, the poset structure of $[x, y]$ determines $a\ell(x, y) \in \{1, 3, 5\}$.

If $a\ell(x, y) = 5$, then $\widetilde{R}_{x,y}(q) = q^5$ is determined. Note that in this case the poset $[x, y]$ is a lattice and it is known that this implies either $cap(x, y) \geq 6$, or $[x, y] \cong \mathcal{B}_5$.

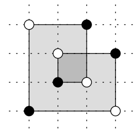
If $a\ell(x, y) = 1$, then (x, y) is an edge of BG. So the diagram of (x, y) is one of the following:

In this case, we have $\{a(x, y), c(x, y)\} = \{3, 4\}$ and, by Proposition 6.8

$$\widetilde{R}_{x,y}(q) = q^5 + 2q^3 + q.$$

Finally suppose $a\ell(x, y) = 3$. In this case $\widetilde{R} = q^5 + bq^3$, for some $b \in \mathbf{N}$. The only diagrams in S_4 of length 5 are, up to symmetries, exactly two: the one corresponding to the edges of BG, already considered, and the following:

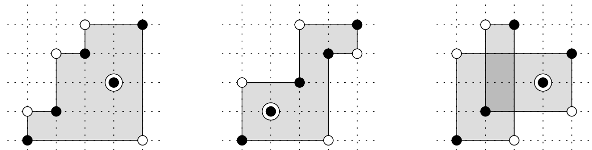
the *heart* diagram: 

In this case $a(x, y) = c(x, y) = 3$, and applying the stair method we obtain

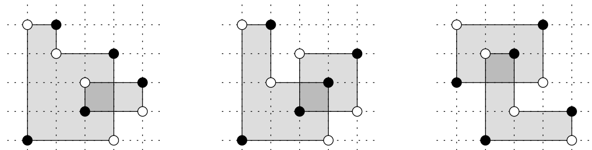
$$\widetilde{R}_{x,y}(q) = q^5 + 2q^3.$$

All other cases can be obtained by "enlarging" the essential diagrams of length 4 in S_5. Considering only non trivially reducible cases, it turns out that two general situations can occur:

(1) the diagram of (x, y) is obtained from one of the diagrams of length 3 corresponding to a 3-crown, by adding one fixed point with multiplicity 1. Here are some examples:



(2) the diagram of (x, y) is the overlapping of a permutaomino with 4 edges and one with 6 edges. A few examples:



After a patient enumeration of all possible cases, and using the interpretation of atoms and coatoms in terms of the diagram, it turns out that in all these cases $cap(x, y) \in \{4, 5\}$, but the boolean algebra $\mathcal{B}_5$ never occurs. And applying the stair method, we get

$$\widetilde{R}_{x,y}(q) = q^5 + q^3.$$

Putting all together, we get our result. □

References

[1] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, Springer-Verlag, New York, to appear.

[2] F. Brenti, A combinatorial formula for Kazhdan-Lusztig polynomials, *Invent. Math.* **118** (2004), 371–394.

[3] F. Brenti, F. Caselli and M. Marietti, Special matchings and Kazhdan-Lusztig polynomials, *Adv. Math.*, to appear.

[4] M. J. Dyer, *Hecke algebras and reflections in Coxeter groups*, Ph. D. Thesis, University of Sydney, 1987.

[5] M. J. Dyer, Hecke algebras and shellings of Bruhat intervals, *Compositio Math.*, **89** (1993), 91–115.

[6] 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).

[7] A. Hultman, *Combinatorial complexes, Bruhat intervals and reflection distances*, Ph.D. thesis, Kungl Tekniska Högskolan, Stockholm, 2003.

[8] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Univ. Press, Cambridge, 1990.

[9] C. Kassel, A. Lascoux and C. Reutenauer, The singular locus of a Schubert variety, *J. Algebra* **269** (2003), 74–108.

[10] D. Kazhdan and G. Lusztig, Representations of Coexeter groups and Hecke algebras, *Invent. Math.* **53** (1979), 165–184.

[11] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1986.

Institut Mittag-Leffler, Auravägen 17, 18260, Djursholm, Sweden

E-mail address: `incitti@math.kth.se`

A Proof of Stanley's Open Problem

Masao Ishikawa

Department of Mathematics, Faculty of Education, Tottori University

Abstract

In the open problem session of the FPSAC'03, R.P. Stanley gave an open problem about a certain sum of the Schur functions (See [19]). The purpose of this paper is to give a proof of this open problem. The proof consists of three steps. At the first step we express the sum by a Pfaffian as an application of our minor summation formula ([7]). In the second step we prove a Pfaffian analogue of Cauchy type identity which generalize [22]. Then we give a proof of Stanley's open problem in Section 4. We also present certain corollaries obtained from this identity involving the Big Schur functions and some polynomials arising from the Macdonald polynomials, which generalize Stanley's open problem.

Résumé

Dans la session de problèmes de SFCA'03, Stanley a posé un problème ouvert sur certaine somme de fonctions de Schur (voir [19]). Le but de cet article est de résoudre ce problème ouvert. La preuve consiste en trois étapes. Premièrement on exprime cette somme comme un Pfaffien en appliquant notre formule de sommation de mineurs [7]. Deuxièmement on démontre un analogue Pfaffien de l'identité de type Cauchy, qui généralise une identité de Sunquist [22]. Et puis on résoud le probème ouvert de Stanley dans la Section 4. On présente aussi quelques corollaires de cette identité impliquant les grandes fonctions de Schur et des polynômes apparaissant dans l'étude des polynômes de Macdonald, qui généralise le problème originel de Stanley.

1 Introduction

In the open problem session of the 15th Anniversary International Conference on Formal Power Series and Algebraic Combinatorics (Vadstena, Sweden, 25 June 2003), R.P. Stanley gave an open problem on a sum of Schur functions with a weight including four parameters, i.e. Theorem 1.1 (See [19]). The purpose of this paper is to give a proof of this open problem. In the process of our proof, we obtain a Pfaffian identity, i.e. Theorem 3.1, which generalize the Pfaffian identities in [22]. Note that certain determinant and Pfaffian identities of this type first appeared in [15], and applied to solve some alternating sign matrices enumerations under certain symmetries stated in [11]. Certain conjectures which extensively generalize the determinant and Pfaffian identities of this type were stated in [17], and a proof of the conjectured determinant and Pfaffian identities was given in [6]. Our proof proceeds by three steps. In the first step we utilize the minor summation formula ([7]) to express the sum of Schur functions into a Pfaffian. In the second step we express the Pfaffian by a determinant using a Cauchy type Pfaffian formula (also see [16], [17] and [6]), and try to simplify it as much as possible. In the final step we complete our proof using a key proposition, i.e. Proposition 4.1 (See [18] and [21]).

We follow the notation in [13] concerning the symmetric functions. In this paper we use a symmetric function f in n variables $(x_1, \dots, x_n)$, which is usually written as $f(x_1, \dots, x_n)$, and also a symmetric function f in countably many variables $x = (x_1, x_2, \dots)$, which is written as $f(x)$ (for detailed description of the ring of symmetric functions in countably many variables, see [13], I, sec.2). To simplify this notation we express the n-tuple $(x_1, \dots, x_n)$ by X_n, and sometimes simply write $f(X_n)$ for $f(x_1, \dots, x_n)$. When the number of variables is finite and there is no fear of confusion what this number is, we simply write X for X_n in abbreviation. Thus $f(x)$ is in countably many variables, but $f(X)$ is in finite variables and the number of variables is clear from the assumption.

Given a partition λ, define $\omega(\lambda)$ by

$$\omega(\lambda) = a^{\sum_{i\geq 1}\lceil \lambda_{2i-1}/2\rceil} b^{\sum_{i\geq 1}\lfloor \lambda_{2i-1}/2\rfloor} c^{\sum_{i\geq 1}\lceil \lambda_{2i}/2\rceil} d^{\sum_{i\geq 1}\lfloor \lambda_{2i}/2\rfloor},$$

where a, b, c and d are indeterminates, and $\lceil x\rceil$ (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.

$$\begin{array}{|c|c|c|c|c|}
\hline
a & b & a & b & a \\
\hline
c & d & c & d & \multicolumn{1}{c}{} \\
\cline{1-4}
a & b & a & b & \multicolumn{1}{c}{} \\
\cline{1-4}
c & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\
\cline{1-1}
\end{array}$$

Let $s_\lambda(x)$ denote the Schur function corresponding to a partition λ. R. P. Stanley gave the following conjecture in the open problem session of FPSAC'03.

Theorem 1.1. Let

$$z = \sum_{\lambda} \omega(\lambda) s_\lambda.$$

Here the sum runs over all partitions λ. Then we have

$$\log z - \sum_{n\geq 1} \frac{1}{2n} a^n (b^n - c^n) p_{2n} - \sum_{n\geq 1} \frac{1}{4n} a^n b^n c^n d^n p_{2n}^2 \\ \in \mathbb{Q}[[p_1, p_3, p_5, \dots]]. \tag{1.1}$$

Here $p_r = \sum_{i\geq 1} x_i^r$ denote the rth power sum symmetric function.

As direct consequence of this theorem, we obtain the following corollary. Let $S_\lambda(x;t) = \det\left(q_{\lambda_i - i + j}(x;t)\right)_{1\leq i,j\leq \ell(\lambda)}$ denote the big Schur function corresponding the partitions λ, where $q_r(x;t) = Q_{(r)}(x;t)$ denote the Hall-Littlewood functions (See [13], III, sec.2).

Corollary 1.2. Let

$$Z(x;t) = \sum_{\lambda} \omega(\lambda) S_\lambda(x;t),$$

Here the sum runs over all partitions λ. Then we have

$$\log Z(x;t) - \sum_{n\geq 1} \frac{1}{2n} a^n (b^n - c^n)(1-t^{2n}) p_{2n} - \sum_{n\geq 1} \frac{1}{4n} a^n b^n c^n d^n (1-t^{2n})^2 p_{2n}^2 \in \mathbb{Q}[[p_1, p_3, p_5, \dots]]. \tag{1.2}$$

This corollary is also generalized to the two parameter polynomials defined by I. G. Macdonald. Define

$$T_\lambda(x;q,t) = \det \left(Q_{(\lambda_i - i + j)}(x;q,t) \right)_{1\leq i,j\leq \ell(\lambda)}$$

where $Q_\lambda(x;q,t)$ stands for the Macdonald polynomial corresponding to the partition λ, and $Q_{(r)}(x;q,t)$ is the one corresponding to the one row partition (r) (See [13], IV, sec.4). Then we obtain the following corollary:

Corollary 1.3. Let

$$Z(x;q,t) = \sum_\lambda \omega(\lambda) T_\lambda(x;q,t),$$

Here the sum runs over all partitions λ. Then we have

$$\log Z(x;q,t) - \sum_{n\geq 1} \frac{1}{2n} a^n (b^n - c^n) \frac{1-t^{2n}}{1-q^{2n}} p_{2n} - \sum_{n\geq 1} \frac{1}{4n} a^n b^n c^n d^n \frac{(1-t^{2n})^2}{(1-q^{2n})^2} p_{2n}^2 \in \mathbb{Q}[[p_1, p_3, p_5, \dots]]. \tag{1.3}$$

In the rest of this section we briefly recall the definition of Pfaffians. For the detailed explanation of Pfaffians, the reader can consult [9] and [20]. Let n be a non-negative integer and assume we are given a $2n$ by $2n$ skew-symmetric matrix $A = (a_{ij})_{1\leq i,j\leq 2n}$, (i.e. $a_{ji} = -a_{ij}$), whose entries a_{ij} are in a commutative ring. The *Pfaffian* of A is, by definition,

$$\operatorname{Pf}(A) = \sum \epsilon(\sigma_1, \sigma_2, \dots, \sigma_{2n-1}, \sigma_{2n})\, a_{\sigma_1\sigma_2} \dots a_{\sigma_{2n-1}\sigma_{2n}}.$$

where the summation is over all partitions $\{\{\sigma_1,\sigma_2\}_<, \dots, \{\sigma_{2n-1},\sigma_{2n}\}_<\}$ of $[2n]$ into 2-elements blocks, and where $\epsilon(\sigma_1,\sigma_2,\dots,\sigma_{2n-1},\sigma_{2n})$ denotes the sign of the permutation

$$\begin{pmatrix} 1 & 2 & \cdots & 2n \\ \sigma_1 & \sigma_2 & \cdots & \sigma_{2n} \end{pmatrix}.$$

2 Pfaffian Expressions

Given a partition $\lambda = (\lambda_1, \dots, \lambda_m)$ satisfying $\ell(\lambda) \leq m$, we associate a decreasing sequence $\lambda + \delta_m$ which is usually denoted by $l = (l_1, \dots, l_m)$, where $\delta_m = (m-1, m-2, \dots, 0)$. The key observation to prove Theorem 1.1 is the following theorem, which shows that the weight $\omega(\lambda)$ can be expressed by a Pfaffian:

Theorem 2.1. Let n be a non-negative integer. Let $\lambda = (\lambda_1, \dots, \lambda_{2n})$ be a partition such that $\ell(\lambda) \leq 2n$, and put $l = (l_1, \dots, l_{2n}) = \lambda + \delta_{2n}$. Define a $2n$ by $2n$ skew-symmetric matrix $A = (\alpha_{ij})_{1\leq i,j\leq 2n}$ by

$$\alpha_{ij} = a^{\lceil (l_i-1)/2 \rceil} b^{\lfloor (l_i-1)/2 \rfloor} c^{\lceil l_j/2 \rceil} d^{\lfloor l_j/2 \rfloor}$$

for $i < j$, and as $\alpha_{ji} = -\alpha_{ij}$ holds for any $i, j \geq 0$. Then we have

$$\operatorname{Pf}\left[A\right]_{1\leq i,j\leq 2n} = (abcd)^{\binom{n}{2}} \omega(\lambda).$$

The essential idea to prove Theorem 2.1 is the following lemma which already appeared as Lemma 7 in Section 4 of [7].

Lemma 2.2. Let x_i and y_j be indeterminates, and let n is a non-negative integer. Then

$$\operatorname{Pf}[x_i y_j]_{1\le i<j\le 2n} = \prod_{i=1}^{n} x_{2i-1} \prod_{i=1}^{n} y_{2i}. \ \square \tag{2.1}$$

Theorem 2.1 shows that the weight $\omega(\lambda)$ can be expressed by the Pfaffians of submatrices of a certain matrix, and the row/column indices of the submatrices are determined by the partition λ. This shows that the weighted sum of the Schur functions is a sum of minors multiplied by the "sub-Pfaffians". Thus we need a minor summation formula from [7].

Let m, n and r be integers such that $r \le m, n$ and let T be an m by n matrix. For any index sets $I = \{i_1, \dots, i_r\}_< \subseteq [m]$ and $J = \{j_1, \dots, j_r\}_< \subseteq [n]$, let $\Delta^I_J(A)$ denote the sub-matrix obtained by selecting the rows indexed by I and the columns indexed by J. If $r = m$ and $I = [m]$, we simply write $\Delta_J(A)$ for $\Delta^{[m]}_J(A)$. Similarly, if $r = n$ and $J = [n]$, we write $\Delta^I(A)$ for $\Delta^I_{[n]}(A)$. For any finite set S and a non-negative integer r, let $\binom{S}{r}$ denote the set of all r-element subsets of S. We cite a theorem from [7] which we call a minor summation formula:

Theorem 2.3. Let n and N be non-negative integers such that $2n \le N$. Let $T = (t_{ij})_{1\le i\le 2n, 1\le j\le N}$ be a $2n$ by N rectangular matrix, and let $A = (a_{ij})_{1\le i,j\le N}$ be a skew-symmetric matrix of size N. Then

$$\sum_{I\in\binom{[N]}{2n}} \operatorname{Pf}\left(\Delta^I_I(A)\right) \det\left(\Delta_I(T)\right) = \operatorname{Pf}\left(TA\,{}^tT\right).$$

If we put $Q = (Q_{ij})_{1\le i,j\le 2n} = TA\,{}^tT$, then its entries are given by

$$Q_{ij} = \sum_{1\le k<l\le N} a_{kl} \det\left(\Delta^{ij}_{kl}(T)\right), \qquad (1 \le i, j \le 2n).$$

Here we write $\Delta^{ij}_{kl}(T)$ for $\Delta^{\{ij\}}_{\{kl\}}(T) = \begin{vmatrix} t_{ik} & t_{il} \\ t_{jk} & t_{jl} \end{vmatrix}$. $\square$

First we restrict our attention to the finite variables case. As an application of the minor summation formula, i.e. Theorem 2.3, we can express the sum with a Pfaffian.

Theorem 2.4. Let n be a positive integer and let $\omega(\lambda)$ be as defined in Section 1. Let

$$z_n = z_n(X_{2n}) = \sum_{\ell(\lambda)\le 2n} \omega(\lambda) s_\lambda(X_{2n}) = \sum_{\ell(\lambda)\le 2n} \omega(\lambda) s_\lambda(x_1, \dots, x_{2n}) \tag{2.2}$$

be the sum restricted to $2n$ variables. Then we have

$$z_n(X_{2n}) = \frac{(abcd)^{-\binom{n}{2}}}{\prod_{1\le i<j\le 2n}(x_i - x_j)} \operatorname{Pf}\left(p_{ij}\right)_{1\le i<j\le 2n}, \tag{2.3}$$

where p_{ij} is defined by

$$p_{ij} = \frac{\begin{vmatrix} x_i + ax_i^2 & 1 - a(b+c)x_i - abcx_i^3 \\ x_j + ax_j^2 & 1 - a(b+c)x_j - abcx_j^3 \end{vmatrix}}{(1 - abx_i^2)(1 - abx_j^2)(1 - abcdx_i^2x_j^2)}. \tag{2.4}$$

Proof. By Theorem 2.3 it is enough to compute

$$\beta_{ij} = \sum_{k \geq l \geq 0} a^{\lceil (k-1)/2 \rceil} b^{\lfloor (k-1)/2 \rfloor} c^{\lceil l/2 \rceil} d^{\lfloor l/2 \rfloor} \begin{vmatrix} x_i^k & x_i^l \\ x_j^k & x_j^l \end{vmatrix}.$$

Let $f_{kl}^{ij} = a^{\lceil (k-1)/2 \rceil} b^{\lfloor (k-1)/2 \rfloor} c^{\lceil l/2 \rceil} d^{\lfloor l/2 \rfloor} \begin{vmatrix} x_i^k & x_i^l \\ x_j^k & x_j^l \end{vmatrix}$, then, this sum can be divided into four cases, i.e.

$$\beta_{ij} = \sum_{\substack{k=2r+1,\, l=2s \\ r \geq s \geq 0}} f_{kl}^{ij} + \sum_{\substack{k=2r,\, l=2s \\ r \geq s \geq 0}} f_{kl}^{ij} + \sum_{\substack{k=2r+1,\, l=2s+1 \\ r \geq s \geq 0}} f_{kl}^{ij} + \sum_{\substack{k=2r+2,\, l=2s+1 \\ r \geq s \geq 0}} f_{kl}^{ij}.$$

We compute each case:

(i) If $k = 2r+1$ and $l = 2s$ for $r \geq s \geq 0$, then

$$\begin{aligned}
\sum_{\substack{k=2r+1,\, l=2s \\ r \geq s \geq 0}} f_{kl}^{ij} &= \sum_{r \geq s \geq 0} a^r b^r c^s d^s \begin{vmatrix} x_i^{2r+1} & x_i^{2s} \\ x_j^{2r+1} & x_j^{2s} \end{vmatrix} \\
&= \sum_{r \geq s \geq 0} c^s d^s \begin{vmatrix} \frac{a^s b^s x_i^{2s+1}}{1-abx_i^2} & x_i^{2s} \\ \frac{a^s b^s x_j^{2s+1}}{1-abx_j^2} & x_j^{2s} \end{vmatrix} \\
&= \frac{(x_i - x_j)(1 + abx_i x_j)}{(1-abx_i^2)(1-abx_j^2)(1-abcdx_i^2 x_j^2)}.
\end{aligned}$$

In the same way we obtain the followings by straight forward computations.

(ii) If $k = 2r$ and $l = 2s$ for $r \geq s \geq 0$, then

$$\sum_{\substack{k=2r,\, l=2s \\ r \geq s \geq 0}} f_{kl}^{ij} = \frac{a(x_i^2 - x_j^2)}{(1-abx_i^2)(1-abx_j^2)(1-abcdx_i^2 x_j^2)}.$$

(iii) If $k = 2r+1$ and $l = 2s+1$ for $r \geq s \geq 0$, then

$$\sum_{\substack{k=2r+1,\, l=2s+1 \\ r \geq s \geq 0}} f_{kl}^{ij} = \frac{abcx_i x_j(x_i^2 - x_j^2)}{(1-abx_i^2)(1-abx_j^2)(1-abcdx_i^2 x_j^2)}.$$

(iv) If $k = 2r+2$ and $l = 2s+1$ for $r \geq s \geq 0$, then

$$\sum_{\substack{k=2r+2,\, l=2s+1 \\ r \geq s \geq 0}} f_{kl}^{ij} = \frac{acx_i x_j(x_i - x_j)(1 + abx_i x_j)}{(1-abx_i^2)(1-abx_j^2)(1-abcdx_i^2 x_j^2)}.$$

Summing up these four identities, we obtain

$$\beta_{ij} = \frac{(x_i - x_j)\{1 + abx_i x_j + a(x_i + x_j) + abcx_i x_j(x_i + x_j) + acx_i x_j(1 + abx_i x_j)\}}{(1-abx_i^2)(1-abx_j^2)(1-abcdx_i^2 x_j^2)}.$$

It is easy to see the numerator is written by the determinant, and this completes the proof. □

3 Cauchy Type Pfaffians

The aim of this section is to prove (3.4). In the next section we will use this identity to prove Stanley's open problem. First we prove a fundamental Pfaffian identity, i.e. Theorem 3.1, and deduce all the identities in this section to this theorem. An intensive generalization was conjectured in [17] and proved in [6]. There is a certain Pfaffian-Hafnian analogue of Borchardt's identity in [5].

First we fix our notation. Let n be an non-negative integer. 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. Set $V_{ij}^n(X, Y; A, B)$ to be

$$\begin{cases} a_i x_i^{n-j} y_i^{j-1} & \text{if } 1 \le j \le n, \\ b_i x_i^{2n-j} y_i^{j-n-1} & \text{if } n+1 \le j \le 2n, \end{cases}$$

for $1 \le i \le 2n$, and define $V^n(X, Y; A, B)$ by

$$V^n(X, Y; A, B) = \det \left(V_{ij}^n(X, Y; A, B) \right)_{1 \le i,j \le 2n}.$$

For example, if $n = 1$, then we have $V^1(X, Y; A, B) = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, and if $n = 2$, then $V^2(X, Y; A, B)$ looks as follows:

$$V^2(X, Y; A, B) = \begin{vmatrix} a_1 x_1 & a_1 y_1 & b_1 x_1 & b_1 y_1 \\ a_2 x_2 & a_2 y_2 & b_2 x_2 & b_2 y_2 \\ a_3 x_3 & a_3 y_3 & b_3 x_3 & b_3 y_3 \\ a_4 x_4 & a_4 y_4 & b_4 x_4 & b_4 y_4 \end{vmatrix}.$$

The main result of this section is the following theorem.

Theorem 3.1. Let n be a positive integer. Let $X = (x_1, \dots, x_{2n})$, $Y = (y_1, \dots, y_{2n})$, $A = (a_1, \dots, a_{2n})$, $B = (b_1, \dots, b_{2n})$, $C = (c_1, \dots, c_{2n})$ and $D = (d_1, \dots, d_{2n})$ be $2n$-tuples of variables. Then

$$\operatorname{Pf} \left[\frac{\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} \cdot \begin{vmatrix} c_i & d_i \\ c_j & d_j \end{vmatrix}}{\begin{vmatrix} x_i & y_i \\ x_j & y_j \end{vmatrix}} \right]_{1 \le i < j \le 2n} = \frac{V^n(X, Y; A, B) V^n(X, Y; C, D)}{\prod_{1 \le i < j \le 2n} \begin{vmatrix} x_i & y_i \\ x_j & y_j \end{vmatrix}}. \tag{3.1}$$

The following proposition is obtained easily by elementary transformations of the matrices and we omit the proof.

Proposition 3.2. Let n be a positive integer. Let $X = (x_1, \dots, x_{2n})$ be $2n$-tuples of variables and let t be an indeterminate. Then

$$V^n(X, \mathbf{1} + tX^2; X, \mathbf{1}) = (-1)^{\binom{n}{2}} t^{\binom{n}{2}} \prod_{1 \le i < j \le 2n} (x_i - x_j), \tag{3.2}$$

where $\mathbf{1}$ denotes the $2n$-tuple $(1, \dots, 1)$, and $\mathbf{1} + tX^2$ denotes the $2n$-tuple $(1 + tx_1^2, \dots, 1 + tx_{2n}^2)$.

Let t be an arbitrary indeterminate. If we set $y_i = 1 + tx_i^2$ in (3.1), then

$$\begin{vmatrix} x_i & 1 + tx_i^2 \\ x_j & 1 + tx_j^2 \end{vmatrix} = (x_i - x_j)(1 - tx_i x_j)$$

and (3.2) immediately implies the following corollary.

Corollary 3.3. Let n be a non-negative integer. Let $X = (x_1, \dots, x_{2n})$, $A = (a_1, \dots, a_{2n})$, $B = (b_1, \dots, b_{2n})$, $C = (c_1, \dots, c_{2n})$ and $D = (d_1, \dots, d_{2n})$ be $2n$-tuples of variables. Then

$$\operatorname{Pf}\left[\frac{(a_ib_j - a_jb_i)(c_id_j - c_jd_i)}{(x_i - x_j)(1 - tx_ix_j)}\right]_{1\le i<j\le 2n} = \frac{V^n(X, \mathbf{1} + tX^2; A, B)V^n(X, \mathbf{1} + tX^2; C, D)}{\prod_{1\le i<j\le 2n}(x_i - x_j)(1 - tx_ix_j)}. \tag{3.3}$$

In particular, we have

$$\operatorname{Pf}\left[\frac{a_ib_j - a_jb_i}{1 - tx_ix_j}\right]_{1\le i<j\le 2n} = (-1)^{\binom{n}{2}}t^{\binom{n}{2}}\frac{V^n(X, \mathbf{1} + tX^2; A, B)}{\prod_{1\le i<j\le 2n}(1 - tx_ix_j)}. \square \tag{3.4}$$

Now we give a sketch of a proof of Theorem 3.1. Let n and r be integers such that $2n \ge r \ge 0$. Let $X = (x_1, \dots, x_{2n})$ be a $2n$-tuple of variables and let $1 \le k_1 < \cdots < k_r \le 2n$ be a sequence of integers. Let $X^{(k_1,\dots,k_r)}$ denote the $(2n - r)$-tuple of variables obtained by removing the variables $x_{k_1}, \dots, x_{k_r}$ from X_{2n}. The key to prove Theorem 3.1 is the following lemma:

Lemma 3.4. Let n be a positive integer. Let $X = (x_1, \dots, x_{2n})$, $A = (a_1, \dots, a_{2n})$ and $C = (c_1, \dots, c_{2n})$ be $2n$-tuples of variables. Then the following identity holds.

$$\begin{aligned}&\sum_{k=1}^{2n-1}\frac{\prod_{\substack{i=1\\i\ne k}}^{2n-1}(x_k - x_i)}{x_k - x_{2n}}(a_k - a_{2n})(c_k - c_{2n})\\
&\times V^{n-1}(X^{(k,2n)}, \mathbf{1}^{(k,2n)}; A^{(k,2n)}, \mathbf{1}^{(k,2n)})V^{n-1}(X^{(k,2n)}, \mathbf{1}^{(k,2n)}; C^{(k,2n)}, \mathbf{1}^{(k,2n)})\\
&= \frac{V^n(X, \mathbf{1}; A, \mathbf{1})V^n(X, \mathbf{1}; C, \mathbf{1})}{\prod_{i=1}^{2n-1}(x_i - x_{2n})}.\end{aligned}$$

Here $\mathbf{1}$ denotes the $2n$-tuples $(1, \dots, 1)$.

This lemma and the expansion of the Pfaffians along the last row/column implies Theorem 3.1 by a direct computation.

4 A Proof of Stanley's Open Problem

The key idea of our proof is the following proposition, which the reader can find in [18], Exercise 7.7, or [21], Section 3.

Proposition 4.1. Let $f(x_1, x_2, \dots)$ be a symmetric function with infinite variables. Then $f \in \mathbb{Q}[p_\lambda :$ all parts $\lambda_i > 0$ are odd$]$ if and only if

$$f(t, -t, x_1, x_2, \dots) = f(x_1, x_2, \dots).\square$$

Our strategy is simple. If we set $v_n(X_{2n})$ to be

$$\log z_n(X_{2n}) - \sum_{k\ge1}\frac{1}{2k}a^k(b^k - c^k)p_{2k}(X_{2n}) - \sum_{k\ge1}\frac{1}{4k}a^kb^kc^kd^kp_{2k}(X_{2n})^2 \tag{4.1}$$

then we claim it satisfies

$$v_{n+1}(t, -t, X_{2n}) = v_n(X_{2n}). \tag{4.2}$$

This will eventually prove Theorem 1.1. As an immediate consequence of (2.3), (2.4) and (3.4), we obtain the following theorem:

Theorem 4.2. Let $X = (x_1, \dots, x_{2n})$ be a $2n$-tuple of variables. Then

$$z_n(X_{2n}) = (-1)^{\binom{n}{2}} \frac{V^n(X^2, \mathbf{1} + abcdX^4; X + aX^2, \mathbf{1} - a(b+c)X^2 - abcX^3)}{\prod_{i=1}^{2n}(1 - abx_i^2) \prod_{1 \le i < j \le 2n}(x_i - x_j)(1 - abcdx_i^2 x_j^2)}, \tag{4.3}$$

where $X^2 = (x_1^2, \dots, x_{2n}^2)$, $\mathbf{1} + abcdX^4 = (1 + abcdx_1^4, \dots, 1 + abcdx_{2n}^4)$, $X + aX^2 = (x_1 + ax_1^2, \dots, x_{2n} + ax_{2n}^2)$ and $\mathbf{1} - a(b+c)X^2 - abcX^3 = (1 - a(b+c)x_1^2 - abcx_1^3, \dots, 1 - a(b+c)x_{2n}^2 - abcx_{2n}^3)$. $\square$

The (4.3) is key expression to prove that $v_n(X_{2n})$ satisfies (4.2). Once one knows (4.3), then it is straightforward computation to prove Stanley's open problem. The following proposition is the first step.

Proposition 4.3. Let $X = (x_1, \dots, x_{2n})$ be a $2n$-tuple of variables. Put

$$f_n(X_{2n}) = V^n(X^2, \mathbf{1} + abcdX^4; X + aX^2, \mathbf{1} - a(b+c)X^2 - abcX^3).$$

Then $f_n(X_{2n})$ satisfies

$$\begin{aligned} &f_{n+1}(t, -t, X_{2n}) \\ &= (-1)^n 2t(1 - abt^2)(1 - act^2) \prod_{i=1}^{2n}(t^2 - x_i^2) \prod_{i=1}^{2n}(1 - abcdt^2 x_i^2) \cdot f_n(X_{2n}). \end{aligned} \tag{4.4}$$

From Theorem 4.2 and Proposition 4.3 we obtain the following proposition.

Proposition 4.4. Let $X = (x_1, \dots, x_{2n})$ be a $2n$-tuple of variables. Then

$$z_{n+1}(t, -t, X_{2n}) = \frac{1 - act^2}{(1 - abt^2)(1 - abcdt^4) \prod_{i=1}^{2n}(1 - abcdt^2 x_i^2)} z_n(X_{2n}). \tag{4.5}$$

Now the proof of Theorem 1.1 is straightforward computation. We omit the details.

5 Open Problems

The author tried to find an analogous formula when the sum runs over all distinct partitions by computer experiments using Stembridge's SF package (cf. [2] and [3]). But the author could not find any conceivable formula when the sum runs over all distinct partitions.

He also checked Hall-Littlewood functions case, and could not find in the general case, but found some nice formulas if we substitute -1 into t. These are byproducts found by our computer experiments.

Conjecture 5.1. Let

$$w(x; t) = \sum_{\lambda} \omega(\lambda) P_\lambda(x; t),$$

where $P_\lambda(x; t)$ denote the Hall-Littlewood function corresponding to the partition λ, and the sum runs over all partitions λ. Then

$$\begin{aligned} \log w(x; -1) + \sum_{n \ge 1 \text{ odd}} \frac{1}{2n} a^n c^n p_{2n} + \sum_{n \ge 2 \text{ even}} \frac{1}{2n} a^{\frac{n}{2}} c^{\frac{n}{2}} (a^{\frac{n}{2}} c^{\frac{n}{2}} - 2b^{\frac{n}{2}} d^{\frac{n}{2}}) p_{2n} \\ \in \mathbb{Q}[[p_1, p_3, p_5, \dots]]. \end{aligned}$$

would hold. $\square$

We might replace the Hall-Littlewood functions $P_\lambda(x;t)$ by the Macdonald polynomials $P_\lambda(x;q,t)$ in this conjecture. Let $P_\lambda(x;q,t)$ denote the Macdonald polynomial corresponding to the partition λ (See [13], IV, sec.4).

Conjecture 5.2. Let

$$w(x;q,t)=\sum_{\lambda}\omega(\lambda)P_\lambda(x;q,t).$$

Here the sum runs over all partitions λ. Then

$$\log w(x;q,-1)+\sum_{n\geq 1\text{ odd}}\frac{1}{2n}a^nc^np_{2n}+\sum_{n\geq 2\text{ even}}\frac{1}{2n}a^{\frac{n}{2}}c^{\frac{n}{2}}(a^{\frac{n}{2}}c^{\frac{n}{2}}-2b^{\frac{n}{2}}d^{\frac{n}{2}})p_{2n}\\ \in\mathbb{Q}[[p_1,p_3,p_5,\dots]]$$

would hold. □

6 Four-Parameter Partition Identities

In [2] C.E. Boulet gave a bijective proof of the following partition identities, i.e. Theorem 6.1 and Theorem 6.7. The aim of this section is to give another proof of these identities. To make our arguments easier, we first consider the strict partitions case.

Theorem 6.1. (Boulet)

$$\sum_{\mu\text{ strict partitions}}\omega(\mu)=\prod_{j=1}^{\infty}\frac{(1+a^jb^{j-1}c^{j-1}d^{j-1})(1+a^jb^jc^jd^{j-1})}{1-a^jb^jc^{j-1}d^{j-1}}.\tag{6.1}$$

Here the sum runs over all strict partitions μ.

To prove this theorem, we need the following lemma, which can be derived from Lemma 2.2 by exactly the same method as we proved Lemma 2.1. Note that any strict partition μ can be written as $\mu_1>\cdots>\mu_{2n}\geq 0$ for a uniquely determined integer n. Let $\ell(\mu)$ denote the length of the strict partition μ, which is the number of nonzero parts of μ. For example, the length of $\mu=(10,8,7,5,3)$ is five.

Lemma 6.2. Let n be a nonnegative integer. Let $\mu=(\mu_1,\dots,\mu_{2n})$ be a strict partition such that $\mu_1>\cdots>\mu_{2n}\geq 0$. Define a skew-symmetric matrix $A=(\alpha_{ij})_{1\leq i,j\leq 2n}$ by

$$\alpha_{ij}=\begin{cases}a^{\lceil\mu_i/2\rceil}b^{\lfloor\mu_i/2\rfloor}c^{\lceil\mu_j/2\rceil}d^{\lfloor\mu_j/2\rfloor}z, & \text{if }\mu_j=0,\\ a^{\lceil\mu_i/2\rceil}b^{\lfloor\mu_i/2\rfloor}c^{\lceil\mu_j/2\rceil}d^{\lfloor\mu_j/2\rfloor}z^2, & \text{if }\mu_j>0,\end{cases}$$

for $i<j$, and as $\alpha_{ji}=-\alpha_{ij}$ holds for any $i,j\geq 0$. Then we have

$$\operatorname{Pf}[A]=\omega(\mu)z^{\ell(\mu)}.\ \square$$

Let J_n denote the square matrix of size n whose (i,j)-entry is $\delta_{i,n+1-j}$. We simply write J for J_n when there is no fear of confusion on the size n. The following lemma can be obtained from Theorem 3 of Section 3 in [7]. (cf. Theorem of Section 4 in [23]). The prototype of this type identity first appeared in [14].

Lemma 6.3. Let n be a positive integer. Let $A = (a_{ij})_{1\le i,j\le n}$ and $B = (b_{ij})_{1\le i,j\le n}$ be skew symmetric matrices of size n. Then

$$\sum_{t=0}^{r} z^{r} \sum_{I\in\binom{[n]}{2t}} \gamma^{|I|} \operatorname{Pf}\left(\Delta_I^I(A)\right) \operatorname{Pf}\left(\Delta_I^I(B)\right) = \operatorname{Pf}\begin{bmatrix} J\,{}^tAJ & J \\ -J & C \end{bmatrix}, \tag{6.2}$$

where $|I| = \sum_{i\in I} i$ and $C = (C_{ij})_{1\le i,j}$ is given by

$$C_{ij} = \gamma^{i+j} b_{ij} z.$$

Theorem 6.4. Let n be a positive integer. Then

$$\sum_{\substack{\mu \text{ strict partitions}\\ \mu_1\le n}} \omega(\mu) z^{\ell(\mu)} = \operatorname{Pf}\begin{bmatrix} S & J_{n+1} \\ -J_{n+1} & B \end{bmatrix}, \tag{6.3}$$

where $S = (1)_{0\le i<j\le n}$ and $B = (\beta_{ij})_{0\le i<j\le n}$ with

$$\beta_{ij} = \begin{cases} a^{\lceil j/2\rceil} b^{\lfloor j/2\rfloor} c^{\lceil i/2\rceil} d^{\lfloor i/2\rfloor} z & \text{if } 0 = i < j \le n, \\ a^{\lceil j/2\rceil} b^{\lfloor j/2\rfloor} c^{\lceil i/2\rceil} d^{\lfloor i/2\rfloor} z^2 & \text{if } 0 < i < j \le n. \end{cases}$$

For example, if $n = 3$, then the Pfaffian in the right-hand side of (6.3) is

$$\operatorname{Pf}\left[\begin{array}{cccc|cccc} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & az & abz & a^2bz \\ 0 & 0 & -1 & 0 & -az & 0 & abcz^2 & a^2bcz^2 \\ 0 & -1 & 0 & 0 & -abz & -abcz^2 & 0 & a^2bcdz^2 \\ -1 & 0 & 0 & 0 & -a^2bz & -a^2bcz^2 & -a^2bcdz^2 & 0 \end{array}\right],$$

and this is equal to $1 + a(1 + b + ab)z + abc(1 + a + ad)z^2 + a^3bcdz^3$. Meanwhile, 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 $\ell(\mu) = 4$ 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 correspond to the above Pfaffian.

Let $\psi_n = \psi_n(a, b, c, d; z) = \operatorname{Pf}\begin{bmatrix} S & J_{n+1} \\ -J_{n+1} & B \end{bmatrix}$ denote the right-hand side of (6.3) for a nonnegative integer n. For example, we have $\psi_0 = 1$, $\psi_1 = 1+az$, $\psi_2 = 1+a(1+b)z+abcz^2$ and $\psi_3 = 1+a(1+b+ab)z+abc(1+a+ad)z^2+a^3bcdz^3$. By elementary transformations and expansions along rows/columns of Pfaffians, we obtain the following recursion formula.

Proposition 6.5. Let $\psi_n = \psi_n(a,b,c,d;z)$ be as above. Then we have

$$\psi_{2n} = (1+b)\psi_{2n-1} + (a^n b^n c^n d^{n-1} z^2 - b)\psi_{2n-2},$$
$$\psi_{2n+1} = (1+a)\psi_{2n} + (a^{n+1} b^n c^n d^n z^2 - b)\psi_{2n-1},$$

for any positive integer j.

From this recurrence relation we immediately obtain the following corollary.

Corollary 6.6. Set $q = abcd$, $x_n = \psi_{2n}$ and $y_n = \psi_{2n+1}$ then

$$x_{n+1} = \left\{1 + ab + a(1+bc)z^2 q^n\right\} x_n - ab(1 - z^2 q^n)(1 - acz^2 q^{n-1})x_{n-1},$$
$$y_{n+1} = \left\{1 + ab + abc(1+ad)z^2 q^n\right\} y_n - ab(1 - z^2 q^n)(1 - acz^2 q^n)y_{n-1},$$

where $x_0 = 1$, $y_0 = 1 + az$, $x_1 = 1 + a(1+b)z + abcz^2$ and

$$y_1 = 1 + a(1+b+ab)z + abc(1+a+ad)z^2 + a^3bcdz^3.$$

There is no enough space here to describe the details of the results, but, when $z = 1$, we can identify them with the three-term relation of the Al-Salam Chihara polynomials and the solution is expressed by an appropriate basic hypergeometric series, i.e.

$$x_n(a,b,c,d;1) = (-a;q)_n\, {}_2\phi_1\left(\begin{matrix} q^{-n}, -c \\ -a^{-1}q^{-n+1} \end{matrix}\,\middle|\, q; -bq\right),$$
$$y_n(a,b,c,d;1) = (1+a)(-abc;q)_n\, {}_2\phi_1\left(\begin{matrix} q^{-n}, -acd \\ -(abc)^{-1}q^{-n+1} \end{matrix}\,\middle|\, q; -c^{-1}q\right).$$

This method works to prove Theorem 6.1.

Finally let us mention that a similar argument also works to prove the following ordinary partition identity.

Theorem 6.7. (Boulet)

$$\sum_{\lambda \text{ partitions}} \omega(\lambda) = \prod_{j=1}^{\infty} \frac{(1 + a^j b^{j-1} c^{j-1} d^{j-1})(1 + a^j b^j c^j d^{j-1})}{(1 - a^j b^j c^j d^j)(1 - a^j b^j c^{j-1} d^{j-1})(1 - a^j b^{j-1} c^j d^{j-1})}. \tag{6.4}$$

Here the sum runs over all partitions λ.

References

[1] G.E. Andrews, "On a Partition Function of Richard Stanley", *Electron. J. Combin.* **11** (2) (2004), R1.

[2] C.E. Boulet, "A four-parameter partition identity", `arXiv:math.CO/0308012`, to appear in Ramanujan Journal.

[3] C.E. Boulet, "A four-parameter partition identity", *Formal Power Series and Algebraic Combinatorics (Vancouver 2004)*, extended abstract.

[4] W. Fulton and J. Harris, *Representation theory, a first course*, Springer-Verlag (1991).

[5] M. Ishikawa, H. Kawamuko and S. Okada, "A Pfaffian-Hafnian analogue of Borchardt's identity", `arXiv:math.CO/0408364`.

[6] M. Ishikawa, H. Tagawa, S. Okada and J. Zeng, "Generalizations of Cauchy's determinant and Schur's Pfaffian", `arXiv:math.CO/0411280`.

[7] M. Ishikawa and M. Wakayama, "Minor summation formula of Pfaffians", *Linear and Multilinear Alg.* **39** (1995), 285-305.

[8] M. Ishikawa and M. Wakayama, "Applications of minor summation formula III, Plücker relations, lattice paths and Pfaffian identities", `arXiv:math.CO/0312358`.

[9] D. Knuth, "Overlapping pfaffians", *Electronic J. of Combi.* **3**, 151–163.

[10] C. Krattenthaler, "Advanced determinant calculus', *Seminaire Lotharingien Combin.* **42** *("The Andrews Festschrift")* (1999), Article B42q, 67.

[11] G. Kuperberg, "Symmetry classes of alternating-sign matrices under one roof", *Ann. of Math.* (2) **156** (2002), 835-866.

[12] A. Lascoux, *Symmetric functions & Combinatorial operators on polynomials*, CBMS/AMS Lectures Notes **99**, (2003).

[13] I.G. Macdonald, *Symmetric functions and Hall polynomials, 2nd Edition*, Oxford University Press, (1995).

[14] S. Okada, *On the generating functions for certain classes of plane partitions*, J. Combin. Th. Ser. A **51** (1989), 1–23

[15] S. Okada, *Applications of minor-summation formulas to rectangular-shaped representations of classical groups*, J. Alg. **205** (1998), 337–367.

[16] S. Okada, "Enumeration of symmetry classes of alternating sign matrices and characters of classical groups", `arXiv:math.CO/0308234`.

[17] S. Okada, Determinant and Pfaffian formulae of Cauchy type and their applications, in "Aspects of Combinatorial Representation Theory", Sūrikaisekikenkyūsho Kōkyūroku **1382** (2004), 198–215 (in Japanese).

[18] R.P. Stanley, *Enumerative combinatorics, Volume II*, Cambridge University Press, (1999).

[19] R.P. Stanley, "Open problem", *International Conference on Formal Power Series and Algebraic Combinatorics (Vadstena 2003)*, June 23 - 27, 2003, available from http://www-math.mit.edu/ rstan/trans.html.

[20] J. Stembridge, "Nonintersecting paths, Pfaffians and plane partitions", *Adv. Math.* **83** (1990), 96–131.

[21] J. Stembridge, "Enriched P-partitions", *Trans. Amer. Math. Soc.* **349** (1997), 763–788.

[22] T. Sundquist "Two variable Pfaffian identities and symmetric functions" *J. Alg. Combin.* **5** (1996), 135–148.

[23] C.A. Tracy and H. Widom, "A limit theorem for shifted Schur measures" `arXiv:math.PR/0210255`,

Masao ISHIKAWA, Department of Mathematics, Faculty of Education,
Tottori University, Tottori 680 8551, Japan
E-mail address: ishikawa@fed.tottori-u.ac.jp

Generalizations of Cauchy's Determinant and Schur's Pfaffian

Masao Ishikawa* Soichi Okada† Hiroyuki Tagawa‡ Jiang Zeng§

Abstract

We present several generalizations of Cauchy's determinant $\det(1/(x_i + y_j))$ and Schur's Pfaffian $\mathrm{Pf}((x_j - x_i)/(x_j + x_i))$ by considering matrices whose entries involve some generalized Vandermonde determinants. Special cases of our formulae include previous formulae due to S. Okada and T. Sundquist. As an application, we give a relation for the Littlewood–Richardson coefficients involving a rectangular partition.

Résumé

On présente plusieurs généralisations du déterminant de Cauchy $\det(1/(x_i + y_j))$ et du Pfaffian de Schur $\mathrm{Pf}((x_j - x_i)/(x_j + x_i))$ en considérant des matrices dont les coefficients impliquent des déterminants de Vandermonde généralisés. Des cas particuliers de nos formules contiennent celles obtenues précédemment par S. Okada et T. Sundquist. Comme une application, on donne une relation pour les coefficients de Littlewood–Richardson associés aux trois partitions dont une est de forme rectangle.

1 Introduction

Identities for determinants and Pfaffians are of great interest in many branches of mathematics. Some people need relations among minors or subPfaffians of a general matrix, others have to evaluate special determinants or Pfaffians. In combinatorics and representation theory, an important role is played by Cauchy's determinant identity [3]

$$\det\left(\frac{1}{x_i + y_j}\right)_{1\le i,j\le n} = \frac{\prod_{1\le i<j\le n}(x_j - x_i)(y_j - y_i)}{\prod_{i,j=1}^{n}(x_i + y_j)}, \tag{1.1}$$

and Schur's Pfaffian identity [20]

$$\mathrm{Pf}\left(\frac{x_j - x_i}{x_j + x_i}\right)_{1\le i,j\le 2n} = \prod_{1\le i<j\le 2n}\frac{x_j - x_i}{x_j + x_i}. \tag{1.2}$$

Also their variations and generalizations have many applications. See, for example, [5], [8], [12], [16], [17], [22], [23]. Also see [10] and [11] for a survey of determinant evaluations.

In this article, we establish several identities of Cauchy-type determinants and Schur-type Pfaffians involving generalized Vandermonde determinants. Let $\boldsymbol{x} = (x_1, \cdots, x_n)$ and $\boldsymbol{a} = (a_1, \cdots, a_n)$ be two vectors of variables of length n. For nonnegative integers p and q with $p + q = n$, we define a generalized Vandermonde matrix $V^{p,q}(\boldsymbol{x}; \boldsymbol{a})$ to be the $n \times n$ matrix with ith row

$$(1, x_i, \cdots, x_i^{p-1}, a_i, a_i x_i, \cdots, a_i x_i^{q-1}).$$

We introduce another generalized Vandermonde matrix $W^n(\boldsymbol{x}; \boldsymbol{a})$ as the $n \times n$ matrix with ith row

$$(1 + a_i x_i^{n-1}, x_i + a_i x_i^{n-2}, \cdots, x_i^{n-1} + a_i).$$

---

*Faculty of Education, Tottori University, e-mail:ishikawa@fed.tottori-u.ac.jp

†Graduate School of Mathematics, Nagoya University, e-mail:okada@math.nagoya-u.ac.jp

‡Faculty of Education, Wakayama University, e-mail:tagawa@math.edu.wakayama-u.ac.jp

§Institut Girard Desargues, Université Claude Bernard Lyon-I, e-mail:zeng@igd.univ-lyon1.fr

If $q=0$, then $V^{n,0}(\boldsymbol{x};\boldsymbol{a})=\left(x_i^{j-1}\right)_{1\le i,j\le n}$ and the determinant $\det V^{n,0}(\boldsymbol{x};\boldsymbol{a})=\prod_{1\le i<j\le n}(x_j-x_i)$ is the usual Vandermonde determinant.

The main purpose of this paper is to prove the following identities for the determinants and Pfaffians involving these generalized Vandermonde determinants. In the later sections, we also give several variants of these determinants and Pfaffians.

Theorem 1.1. (a) Let n be a positive integer and let p and q be nonnegative integers. For six vectors of variables

$$\boldsymbol{x}=(x_1,\cdots,x_n),\ \boldsymbol{y}=(y_1,\cdots,y_n),\ \boldsymbol{a}=(a_1,\cdots,a_n),\ \boldsymbol{b}=(b_1,\cdots,b_n),$$
$$\boldsymbol{z}=(z_1,\cdots,z_{p+q}),\ \boldsymbol{c}=(c_1,\cdots,c_{p+q}),$$

we have

$$\det\left(\frac{\det V^{p+1,q+1}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}{y_j-x_i}\right)_{1\le i,j\le n}$$
$$=\frac{(-1)^{n(n-1)/2}}{\prod_{i,j=1}^{n}(y_j-x_i)}\det V^{p,q}(\boldsymbol{z};\boldsymbol{c})^{n-1}\det V^{n+p,n+q}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z};\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}).\quad(1.3)$$

(b) Let n be a positive integer and let p, q, r, s be nonnegative integers. For seven vectors of variables

$$\boldsymbol{x}=(x_1,\cdots,x_{2n}),\ \boldsymbol{a}=(a_1,\cdots,a_{2n}),\ \boldsymbol{b}=(b_1,\cdots,b_{2n}),$$
$$\boldsymbol{z}=(z_1,\cdots,z_{p+q}),\ \boldsymbol{c}=(c_1,\cdots,c_{p+q}),$$
$$\boldsymbol{w}=(w_1,\cdots,w_{r+s}),\ \boldsymbol{d}=(d_1,\cdots,d_{r+s}),$$

we have

$$\mathrm{Pf}\left(\frac{\det V^{p+1,q+1}(x_i,x_j,\boldsymbol{z};a_i,a_j,\boldsymbol{c})\det V^{r+1,s+1}(x_i,x_j,\boldsymbol{w};b_i,b_j,\boldsymbol{d})}{x_j-x_i}\right)_{1\le i,j\le 2n}$$
$$=\frac{1}{\prod_{1\le i<j\le 2n}(x_j-x_i)}\det V^{p,q}(\boldsymbol{z};\boldsymbol{c})^{n-1}\det V^{r,s}(\boldsymbol{w};\boldsymbol{d})^{n-1}$$
$$\times\det V^{n+p,n+q}(\boldsymbol{x},\boldsymbol{z};\boldsymbol{a},\boldsymbol{c})\det V^{n+r,n+s}(\boldsymbol{x},\boldsymbol{w};\boldsymbol{b},\boldsymbol{d}).\quad(1.4)$$

(c) Let n be a positive integer and let p be a nonnegative integer. For six vectors of variables

$$\boldsymbol{x}=(x_1,\cdots,x_n),\ \boldsymbol{y}=(y_1,\cdots,y_n),\ \boldsymbol{a}=(a_1,\cdots,a_n),\ \boldsymbol{b}=(b_1,\cdots,b_n),$$
$$\boldsymbol{z}=(z_1,\cdots,z_p),\ \boldsymbol{c}=(c_1,\cdots,c_p),$$

we have

$$\det\left(\frac{\det W^{p+2}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}{(y_j-x_i)(1-x_iy_j)}\right)_{1\le i,j\le n}$$
$$=\frac{1}{\prod_{i,j=1}^{n}(y_j-x_i)(1-x_iy_j)}\det W^{p}(\boldsymbol{z};\boldsymbol{c})^{n-1}\det W^{2n+p}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z};\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}).\quad(1.5)$$

(d) Let n be a positive integer and let p and q be nonnegative integers. For seven vectors of variables

$$\boldsymbol{x}=(x_1,\cdots,x_{2n}),\ \boldsymbol{a}=(a_1,\cdots,a_{2n}),\ \boldsymbol{b}=(b_1,\cdots,b_{2n}),$$
$$\boldsymbol{z}=(z_1,\cdots,z_p),\ \boldsymbol{c}=(c_1,\cdots,c_p),$$
$$\boldsymbol{w}=(w_1,\cdots,w_q),\ \boldsymbol{d}=(d_1,\cdots,d_q),$$

we have

$$\mathrm{Pf}\left(\frac{\det W^{p+2}(x_i,x_j,\boldsymbol{z};a_i,a_j,\boldsymbol{c})\det W^{q+2}(x_i,x_j,\boldsymbol{w};b_i,b_j,\boldsymbol{d})}{(x_j-x_i)(1-x_ix_j)}\right)_{1\le i,j\le 2n}$$
$$=\frac{1}{\prod_{1\le i<j\le 2n}(x_j-x_i)(1-x_ix_j)}\det W^{p}(\boldsymbol{z};\boldsymbol{c})^{n-1}\det W^{q}(\boldsymbol{w};\boldsymbol{d})^{n-1}$$
$$\times\det W^{2n+p}(\boldsymbol{x},\boldsymbol{z};\boldsymbol{a},\boldsymbol{c})\det W^{2n+q}(\boldsymbol{x},\boldsymbol{w};\boldsymbol{b},\boldsymbol{d}).\quad(1.6)$$

These identities were conjectured by S. Okada in [18]. If we put $p = q = 0$ in (1.3) or $p = q = r = s = 0$ in (1.4), then the identities read

$$\det\left(\frac{b_j - a_i}{y_j - x_i}\right)_{1\le i,j\le n} = \frac{(-1)^{n(n-1)/2}}{\prod_{i,j=1}^{n}(y_j - x_i)} \det V^{n,n}(\boldsymbol{x},\boldsymbol{y};\boldsymbol{a},\boldsymbol{b}), \tag{1.7}$$

$$\mathrm{Pf}\left(\frac{(a_j - a_i)(b_j - b_i)}{x_j - x_i}\right)_{1\le i,j\le 2n} = \frac{1}{\prod_{1\le i<j\le 2n}(x_j - x_i)} \det V^{n,n}(\boldsymbol{x};\boldsymbol{a}) \det V^{n,n}(\boldsymbol{x};\boldsymbol{b}). \tag{1.8}$$

These special cases, as well as the identities (1.5) with $p = 0$ and (1.6) with $p = q = 0$, are first given by S. Okada [16, Theorems 4.2, 4.7, 4.3, 4.4] in his study of rectangular-shaped representations of classical groups. Another special case of the identity (1.5) with $p = 1$ is given in [17] and applied to the enumeration of vertically and horizontally symmetric alternating sign matrices. These special cases are the starting point of our study.

Under the specialization

$$x_i \leftarrow x_i^2, \quad y_i \leftarrow y_i^2, \quad z_i \leftarrow z_i^2, \quad w_i \leftarrow w_i^2, \quad a_i \leftarrow x_i, \quad b_i \leftarrow y_i, \quad c_i \leftarrow z_i, \quad d_i \leftarrow w_i,$$

one can deduce from (1.3) and (1.4) the following identities:

$$\det\left(\frac{s_{\delta(k)}(x_i, y_j, \boldsymbol{z})}{x_i + y_j}\right)_{1\le i,j\le n} = \frac{\prod_{1\le i<j\le n}(x_j - x_i)(y_j - y_i)}{\prod_{i,j=1}^{n}(x_i + y_j)} s_{\delta(k)}(\boldsymbol{z})^{n-1} s_{\delta(k)}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}), \tag{1.9}$$

$$\begin{aligned} &\mathrm{Pf}\left(\frac{x_j - x_i}{x_j + x_i} s_{\delta(k)}(x_i, x_j \boldsymbol{z}) s_{\delta(l)}(x_i, x_j, \boldsymbol{w})\right)_{1\le i,j\le 2n} \\ &= \prod_{1\le i<j\le 2n} \frac{x_j - x_i}{x_j + x_i} s_{\delta(k)}(\boldsymbol{z})^{n-1} s_{\delta(l)}(\boldsymbol{w})^{n-1} s_{\delta(k)}(\boldsymbol{x},\boldsymbol{z}) s_{\delta(l)}(\boldsymbol{x},\boldsymbol{w}), \end{aligned} \tag{1.10}$$

where s_λ denotes the Schur function corresponding to a partition λ and $\delta(k) = (k, k-1, \cdots, 1)$ denotes the staircase partition. If we take $k = 0$ in (1.9) and $k = l = 0$ in (1.10), we obtain Cauchy's determinant identity (1.1) and Schur's Pfaffian identity (1.2). Another special case of (1.9) with $k = l = 1$ is the rational case of Frobenius' identity [4]. Also, if we take $k = l = 1$ in (1.10), we obtain the rational case of an elliptic generalization of (1.2) given in [19].

2 A Sketch of the Proof of Our Main Theorem

In this section, we give an outline of the proof of Theorem 1.1. First S. Okada presented the identities in Theorem 1.1 at the workshop on "Aspects of Combinatorial Representation Theory" and "2nd East Asian Conference on Algebra and Combinatorics". At the point they were conjectures, and we tried a couple of methods to prove some of these identities, for example, the inductions, the Desnanot–Jacobi formula, and the complex analysis. Among such methods, the best and simplest way we found is to use the Desnanot–Jacobi formula (Lemma 2.1) and a homogeneous version $U^{p,q}$ of the generalized Vandermonde matrix $V^{p,q}$.

The proof of Theorem 1.1 consists of two parts. In the first part, we prove (1.4) by applying the Desnanot–Jacobi formula for Pfaffians to reduce the general case to the case $n = 2$, and then by using the induction on $p + q + r + s$ to show the case $n = 2$. In the second part, we translate (1.4) into the homogeneous version (2.5) and derive (1.3), (1.5), (1.6) from this 'master' identity. Only a sketch of the proof is given below, and the details can be found in our paper [6].

2.1 Proof of (1.4)

For the first part, we recall the Desnanot–Jacobi formulae for determinants and Pfaffians. Given a square matrix A and indices $i_1, \cdots, i_r, j_1, \cdots, j_r$, we denote by $A^{i_1,\cdots,i_r}_{j_1,\cdots,j_r}$ the matrix obtained by removing the rows $i_1, \cdots, i_r$ and the columns $j_1, \cdots, j_r$ of A.

Lemma 2.1. (1) If A is a square matrix, then we have

$$\det A^1_1 \cdot \det A^2_2 - \det A^1_2 \cdot \det A^2_1 = \det A \cdot \det A^{1,2}_{1,2}. \tag{2.1}$$

(2) If A is a skew-symmetric matrix, then we have

$$\operatorname{Pf} A_{1,2}^{1,2} \cdot \operatorname{Pf} A_{3,4}^{3,4} - \operatorname{Pf} A_{1,3}^{1,3} \cdot \operatorname{Pf} A_{2,4}^{2,4} + \operatorname{Pf} A_{1,4}^{1,4} \cdot \operatorname{Pf} A_{2,3}^{2,3} = \operatorname{Pf} A \cdot \operatorname{Pf} A_{1,2,3,4}^{1,2,3,4}. \tag{2.2}$$

This Pfaffian analogue of Desnanot–Jacobi formula is given in [9], [8], and is called the Plücker relation in [8].

By applying Desnanot–Jacobi formula for Pfaffians to the skew-symmetric matrix on the left hand side of (1.4) and using the induction on n, we can see that the proof of (1.4) is reduced to the case $n = 2$ with $\boldsymbol{z}$, $\boldsymbol{c}$, $\boldsymbol{w}$, $\boldsymbol{d}$ replaced by

$$\boldsymbol{z} \leftarrow (\boldsymbol{x}^{(1,2,3,4)}, \boldsymbol{z}), \quad \boldsymbol{c} \leftarrow (\boldsymbol{a}^{(1,2,3,4)}, \boldsymbol{c}), \quad \boldsymbol{w} \leftarrow (\boldsymbol{x}^{(1,2,3,4)}, \boldsymbol{w}), \quad \boldsymbol{d} \leftarrow (\boldsymbol{b}^{(1,2,3,4)}, \boldsymbol{d}),$$

respectively, where $\boldsymbol{x}^{(1,2,3,4)}$ denotes the vector obtained by removing x_1, x_2, x_3, x_4 from $\boldsymbol{x}$. Then the identity (1.4) in the case $n = 2$ can be proven by the induction on $p + q + r + s$ with the help of the following relations between $\det V^{p,q}$ and $\det V^{p-1,q}$ (or $\det V^{q,p}$).

Lemma 2.2. (1) If $p \geq q$ and $p \geq 1$, then we have

$$\det V^{p,q}(\boldsymbol{x}; \boldsymbol{a}) = \prod_{i=1}^{p+q-1} (x_{p+q} - x_i) \cdot \det V^{p-1,q}(x_1, \cdots, x_{p+q-1}; a'_1, \cdots, a'_{p+q-1}), \tag{2.3}$$

where we put
$$a'_i = \frac{a_i - a_{p+q}}{x_i - x_{p+q}} \quad (1 \leq i \leq p + q - 1).$$

(2) For nonnegative integers p and q, we have

$$\det V^{p,q}(\boldsymbol{x}; \boldsymbol{a}) = (-1)^{pq} \prod_{i=1}^{p+q} a_i \cdot \det V^{q,p}(\boldsymbol{x}; \boldsymbol{a}^{-1}), \tag{2.4}$$

where $\boldsymbol{a}^{-1} = (a_1^{-1}, \cdots, a_{p+q}^{-1})$.

Remark 2.3. We can also reduce the proof of the other identities (1.3), (1.5) and (1.6) in Theorem 1.1 to the case of $n = 2$ by using Desnanot-Jacobi formulae. It is easy to show the case of $n = 2$ of (1.3) by using the relations in Lemma 2.2 and the induction on $p + q$. We can prove (1.5) (resp. (1.6)) in the case of $n = 2$, by regarding the both sides as polynomials in z_{p+q} (resp. z_p) and showing that the values coincide at appropriate points by a brute force. Also the special cases of these identities can be obtained by regarding the both sides as meromorphic functions and computing the principal parts at their poles.

2.2 Homogeneous version and proof of (1.3), (1.5) and (1.6)

For the second part of the proof, we introduce a homogeneous version of the matrix $V^{p,q}(\boldsymbol{x}; \boldsymbol{a})$. For vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{a}$, $\boldsymbol{b}$ of length n and nonnegative integers p, q with $p+q = n$, we set $U^{p,q}\begin{pmatrix} \boldsymbol{x} & \boldsymbol{a} \\ \boldsymbol{y} & \boldsymbol{b} \end{pmatrix}$ to be the $n \times n$ matrix with ith row

$$(a_i x_i^{p-1}, a_i x_i^{p-2} y_i, \cdots, a_i y_i^{p-1}, b_i x_i^{q-1}, b_i x_i^{q-2} y_i, \cdots, b_i y_i^{q-1}).$$

Then we have the following relation among $\det U^{p,q}$, $\det V^{p,q}$ and $\det W^p$. Here use the following notation for vectors $\boldsymbol{x} = (x_1, \cdots, x_n)$ and $\boldsymbol{y} = (y_1, \cdots, y_n)$:

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \ldots, x_n + y_n), \quad \boldsymbol{x}\boldsymbol{y} = (x_1 y_1, \ldots, x_n y_n),$$

and, for integers k and l,

$$\boldsymbol{x}^k = (x_1^k, \ldots, x_n^k), \quad \boldsymbol{x}^k \boldsymbol{y}^l = (x_1^k y_1^l, \ldots, x_n^k y_n^l).$$

Lemma 2.4.

$$U^{p,q}\begin{pmatrix}\boldsymbol{x} & \boldsymbol{a} \\ \boldsymbol{y} & \boldsymbol{b}\end{pmatrix} = \prod_{k=1}^{p+q} a_k x_k^{p-1} \cdot V^{p,q}\left(\boldsymbol{x}^{-1}\boldsymbol{y}; \boldsymbol{a}^{-1}\boldsymbol{b}\boldsymbol{x}^{q-p}\right), \tag{2.5}$$

$$V^{p,q}(\boldsymbol{x};\boldsymbol{a}) = U^{p,q}\begin{pmatrix}\boldsymbol{1} & \boldsymbol{1} \\ \boldsymbol{x} & \boldsymbol{a}\end{pmatrix}, \tag{2.6}$$

$$\det U^{n,n}\begin{pmatrix}\boldsymbol{x} & \boldsymbol{1}+\boldsymbol{a}\boldsymbol{x} \\ \boldsymbol{1}+\boldsymbol{x}^2 & \boldsymbol{x}+\boldsymbol{a}\end{pmatrix} = (-1)^{n(n-1)/2}\det W^{2n}(\boldsymbol{x};\boldsymbol{a}), \tag{2.7}$$

$$\det U^{n,n+1}\begin{pmatrix}\boldsymbol{x} & \boldsymbol{1}+\boldsymbol{a}\boldsymbol{x}^2 \\ \boldsymbol{1}+\boldsymbol{x}^2 & \boldsymbol{1}+\boldsymbol{a}\end{pmatrix} = (-1)^{n(n-1)/2}\det W^{2n+1}(\boldsymbol{x};\boldsymbol{a}), \tag{2.8}$$

where $\boldsymbol{1} = (1, \cdots, 1)$.

We can "homogenize" the identity (1.4). It follows from (2.5) and (2.6) that the following theorem is equivalent to (1.4).

Theorem 2.5. Let n be a positive integer and let p, q, r and s be nonnegative integers. Suppose that the vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$ have length $2n$, the vectors $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ have length $p+q$, and the vectors $\boldsymbol{\zeta}$, $\boldsymbol{\omega}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ have length $r+s$. Then we have

$$\begin{aligned}
&\operatorname{Pf}\left(\frac{\det U^{p+1,q+1}\begin{pmatrix}x_i, x_j, \boldsymbol{\xi} & a_i, a_j, \boldsymbol{\alpha} \\ y_i, y_j, \boldsymbol{\eta} & b_i, b_j, \boldsymbol{\beta}\end{pmatrix}\det U^{r+1,s+1}\begin{pmatrix}x_i, x_j, \boldsymbol{\zeta} & c_i, c_j, \boldsymbol{\gamma} \\ y_i, y_j, \boldsymbol{\omega} & d_i, d_j, \boldsymbol{\delta}\end{pmatrix}}{\det\begin{pmatrix}x_i & x_j \\ y_i & y_j\end{pmatrix}}\right)_{1\le i<j\le 2n} \\
&= \frac{1}{\prod_{1\le i<j\le 2n}\det\begin{pmatrix}x_i & x_j \\ y_i & y_j\end{pmatrix}}\det U^{p,q}\begin{pmatrix}\boldsymbol{\xi} & \boldsymbol{\alpha} \\ \boldsymbol{\eta} & \boldsymbol{\beta}\end{pmatrix}^{n-1}\det U^{r,s}\begin{pmatrix}\boldsymbol{\zeta} & \boldsymbol{\gamma} \\ \boldsymbol{\omega} & \boldsymbol{\delta}\end{pmatrix}^{n-1} \\
&\qquad\times\det U^{n+p,n+q}\begin{pmatrix}\boldsymbol{x}, \boldsymbol{\xi} & \boldsymbol{a}, \boldsymbol{\alpha} \\ \boldsymbol{y}, \boldsymbol{\eta} & \boldsymbol{b}, \boldsymbol{\beta}\end{pmatrix}\det U^{n+r,n+s}\begin{pmatrix}\boldsymbol{x}, \boldsymbol{\zeta} & \boldsymbol{c}, \boldsymbol{\gamma} \\ \boldsymbol{y}, \boldsymbol{\omega} & \boldsymbol{d}, \boldsymbol{\delta}\end{pmatrix}.
\end{aligned} \tag{2.9}$$

The special case of $p = q = r = s = 0$ of this identity (2.9) is given by M. Ishikawa [5, Theorem 3.1], and is one of the key ingredients of his proof of Stanley's conjecture on a certain weighted summation of Schur functions. (See [21].)

In this setting, a homogeneous version of (1.3) is a direct consequence of (2.9). A key is the following relation between determinant and Pfaffian. If A is any $m\times(2n-m)$ matrix, then we have

$$\operatorname{Pf}\begin{pmatrix}O & A \\ -{}^tA & O\end{pmatrix} = \begin{cases}(-1)^{n(n-1)/2}\det A & \text{if } m = n, \\ 0 & \text{if } m \ne n.\end{cases} \tag{2.10}$$

Corollary 2.6. Let n be a positive integer and let p and q be fixed nonnegative integers. For vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$, $\boldsymbol{w}$, $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$ of length n, and vectors $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ of length $p+q$, we have

$$\begin{aligned}
&\det\left(\frac{\det U^{p+1,q+1}\begin{pmatrix}x_i, z_j, \boldsymbol{\xi} & a_i, c_j, \boldsymbol{\alpha} \\ y_i, w_j, \boldsymbol{\eta} & b_i, d_j, \boldsymbol{\beta}\end{pmatrix}}{\det\begin{pmatrix}x_i & z_j \\ y_i & w_j\end{pmatrix}}\right)_{1\le i,j\le n} \\
&= \frac{(-1)^{n(n-1)/2}}{\prod_{1\le i,j\le n}\det\begin{pmatrix}x_i & z_j \\ y_i & w_j\end{pmatrix}}\det U^{p,q}\begin{pmatrix}\boldsymbol{\xi} & \boldsymbol{\alpha} \\ \boldsymbol{\eta} & \boldsymbol{\beta}\end{pmatrix}^{n-1}\det U^{n+p,n+q}\begin{pmatrix}\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\xi} & \boldsymbol{a}, \boldsymbol{c}, \boldsymbol{\alpha} \\ \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{\eta} & \boldsymbol{b}, \boldsymbol{d}, \boldsymbol{\beta}\end{pmatrix}.
\end{aligned} \tag{2.11}$$

Proof. In (2.9), we take $r = s = 0$ and put

$$\begin{aligned}
c_1 &= \cdots = c_n = 1, & c_{n+1} &= \cdots = c_{2n} = 0, \\
d_1 &= \cdots = d_n = 0, & d_{n+1} &= \cdots = d_{2n} = 1.
\end{aligned} \tag{2.12}$$

Then we can apply (2.10) to obtain (2.11). □

Now the identity (1.3) follows from this corollary (2.11), (2.6) and an appropriate replacement of variables. Also the remaining identities (1.5) and (1.6) are immediate from (2.11) and (2.9) by using the relations (2.7) and (2.8). This completes the proof of Theorem 1.1.

3 A variation of the determinant and Pfaffian identities

In this section, we give a variation of the identities in Theorem 1.1, which can be regarded as a generalization of an identity of T. Sundquist [23].

Let n be a positive integer and let p and q be nonnegative integers with $p+q=n$. Let $\boldsymbol{x}=(x_1,\cdots,x_n)$ and $\boldsymbol{a}=(a_1,\cdots,a_n)$ be vectors of variables. For partitions λ and μ with $l(\lambda)\le p$ and $l(\mu)\le q$, we define a matrix $V^{p,q}_{\lambda,\mu}(\boldsymbol{x};\boldsymbol{a})$ to be the $n\times n$ matrix with ith row

$$(x_i^{\lambda_p}, x_i^{\lambda_{p-1}+1}, x_i^{\lambda_{p-2}+2}, \cdots, x_i^{\lambda_1+p-1}, a_i x_i^{\mu_q}, a_i x_i^{\mu_{q-1}+1}, a_i x_i^{\mu_{q-2}+2}, \cdots, a_i x_i^{\mu_1+q-1}).$$

For example, if $\lambda=\mu=\emptyset$, then we have $V^{p,q}_{\emptyset,\emptyset}(\boldsymbol{x};\boldsymbol{a})=V^{p,q}(\boldsymbol{x};\boldsymbol{a})$. Let $\mathcal{P}_n$ denote the set of partitions λ with length $\le n$ which are of the form $\lambda=(\alpha_1,\cdots,\alpha_r|\alpha_1+1,\cdots,\alpha_r+1)$ in the Frobenius notation. We define

$$F^{p,q}(\boldsymbol{x};\boldsymbol{a})=\sum_{\lambda\in\mathcal{P}_p,\,\mu\in\mathcal{P}_q}(-1)^{(|\lambda|+|\mu|)/2}\det V^{p,q}_{\lambda,\mu}(\boldsymbol{x};\boldsymbol{a}).$$

The main result of this section is the following theorem.

Theorem 3.1. (a) Let n be a positive integer and let p and q be nonnegative integers. For six vectors of variables

$$\boldsymbol{x}=(x_1,\cdots,x_n),\quad \boldsymbol{y}=(y_1,\cdots,y_n),\quad \boldsymbol{z}=(z_1,\cdots,z_{p+q}),$$
$$\boldsymbol{a}=(a_1,\cdots,a_n),\quad \boldsymbol{b}=(b_1,\cdots,b_n),\quad \boldsymbol{c}=(c_1,\cdots,c_{p+q}),$$

we have

$$\det\left(\frac{F^{p+1,q+1}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}{(y_j-x_i)(1-x_iy_j)}\right)_{1\le i,j\le n}$$
$$=\frac{(-1)^{n(n-1)/2}}{\prod_{i,j=1}^n(y_j-x_i)(1-x_iy_j)}F^{p,q}(\boldsymbol{z};\boldsymbol{c})^{n-1}F^{n+p,n+q}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z};\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}). \tag{3.1}$$

(b) Let n be a positive integer and let p, q, r, s be nonnegative integers. For seven vectors of variables

$$\boldsymbol{x}=(x_1,\cdots,x_{2n}),\quad \boldsymbol{a}=(a_1,\cdots,a_{2n}),\quad \boldsymbol{b}=(b_1,\cdots,b_{2n}),$$
$$\boldsymbol{z}=(z_1,\cdots,z_{p+q}),\quad \boldsymbol{c}=(c_1,\cdots,c_{p+q}),$$
$$\boldsymbol{w}=(w_1,\cdots,w_{r+s}),\quad \boldsymbol{d}=(d_1,\cdots,d_{r+s}),$$

we have

$$\operatorname{Pf}\left(\frac{F^{p+1,q+1}(x_i,x_j,\boldsymbol{z};a_i,a_j,\boldsymbol{c})F^{r+1,s+1}(x_i,x_j,\boldsymbol{w};b_i,b_j,\boldsymbol{d})}{(x_j-x_i)(1-x_ix_j)}\right)_{1\le i,j\le 2n}$$
$$=\frac{1}{\prod_{1\le i<j\le 2n}(x_j-x_i)(1-x_ix_j)}F^{p,q}(\boldsymbol{z};\boldsymbol{c})^{n-1}F^{r,s}(\boldsymbol{w};\boldsymbol{d})^{n-1}$$
$$\times F^{n+p,n+q}(\boldsymbol{x},\boldsymbol{z};\boldsymbol{a},\boldsymbol{c})F^{n+r,n+s}(\boldsymbol{x},\boldsymbol{w};\boldsymbol{b},\boldsymbol{d}). \tag{3.2}$$

In particular, by putting $p=q=r=s=0$ and $b_i=x_i$ for $1\le i\le 2n$ in (3.2), we obtain Sundquist's identity [23, Theorem 2.1].

Corollary 3.2.

$$\operatorname{Pf}\left(\frac{a_j-a_i}{1-x_ix_j}\right)_{1\le i<j\le 2n}=\frac{(-1)^{n(n-1)/2}}{\prod_{1\le i<j\le 2n}(1-x_ix_j)}\sum_{\lambda,\mu\in\mathcal{P}_n}(-1)^{(|\lambda|+|\mu|)/2}\det V^{n,n}_{\lambda,\mu}(\boldsymbol{x};\boldsymbol{a}). \tag{3.3}$$

The key ingredient to prove Theorem 3.1 and Corollary 3.2 is the following relation between $F^{p,q}(\boldsymbol{x};\boldsymbol{a})$ and $\det V^{p,q}(\boldsymbol{y};\boldsymbol{b})$.

Proposition 3.3. We have

$$\begin{aligned} F^{p,q}(\boldsymbol{x};\boldsymbol{a}) &= (-1)^{\binom{p}{2}+\binom{q}{2}} \prod_{i=1}^{p+q} x_i^{p-1} \cdot \det V^{p,q}(\boldsymbol{x}+\boldsymbol{x}^{-1};\boldsymbol{a}\boldsymbol{x}^{q-p}), \\ &= (-1)^{\binom{p}{2}+\binom{q}{2}} \det U^{p,q} \left(\begin{array}{c|c} \boldsymbol{x} & \boldsymbol{1} \\ \boldsymbol{1}+\boldsymbol{x}^2 & \boldsymbol{a} \end{array} \right). \end{aligned} \tag{3.4}$$

This proposition can be proven by the Cauchy-Binet formula and the computation of minors in the following lemma.

Lemma 3.4. Let D_r be the following $r\times(2r-1)$ matrix with columns indexed by $0,1,\cdots,2r-2$:

$$D_r = \begin{array}{c} \begin{matrix} 0 & \quad & r-2 & r-1 & r & \quad & 2r-2 \end{matrix} \\ \begin{pmatrix} & & & 1 & & & \\ & & 1 & & 1 & & \\ & \iddots & & & & \ddots & \\ 1 & & & & & & 1 \end{pmatrix} \end{array}.$$

Then the minor of D_r corresponding to a partition λ is given by

$$\det \Delta^{[r]}_{I(\lambda)}(D_r) = \begin{cases} (-1)^{r(r-1)/2+|\lambda|/2} & \text{if } \lambda\in\mathcal{P}_r, \\ 0 & \text{otherwise,} \end{cases}$$

where $\Delta^{[r]}_{I(\lambda)}(D_r)$ is the $r\times r$ submatrix of D_r consisting of columns $\lambda_r, \lambda_{r-1}+1,\cdots,\lambda_1+(r-1)$.

Concluding this section, we should remark that, in Theorem 3.1, we can replace $F^{p,q}(\boldsymbol{x};\boldsymbol{a})$ by the following alternatives:

$$\begin{aligned} G^{p,q}(\boldsymbol{x};\boldsymbol{a}) &= \sum_{\lambda\in\mathcal{Q}_p,\,\mu\in\mathcal{Q}_q} (-1)^{(|\lambda|+|\mu|)/2} \det V^{p,q}_{\lambda,\mu}(\boldsymbol{x};\boldsymbol{a}), \\ H^{p,q}(\boldsymbol{x};\boldsymbol{a}) &= \sum_{\lambda\in\mathcal{R}_p,\,\mu\in\mathcal{R}_q} (-1)^{(|\lambda|+p(\lambda)+|\mu|+p(\mu))/2} \det V^{p,q}_{\lambda,\mu}(\boldsymbol{x};\boldsymbol{a}), \end{aligned}$$

where $\mathcal{Q}_n$ (resp. $\mathcal{R}_n$) is the set of partitions λ with length $\le n$ which are of the form $\lambda=(\alpha+1|\alpha)$ (resp. $\lambda=(\alpha|\alpha)$) in the Frobenius notation. Then we obtain similar relations between $G^{p,q}(\boldsymbol{x};\boldsymbol{a})$, $H^{p,q}(\boldsymbol{x};\boldsymbol{a})$ and $\det V^{p,q}(\boldsymbol{y};\boldsymbol{b})$.

4 Another generalization of Cauchy's determinant identity

In this section, we give another type of generalized Cauchy's determinant identities involving $\det V^{p,q}$ and $\det W^p$. The following is the main result of this section:

Theorem 4.1.

$$\begin{aligned} &\det\left(\frac{1}{\det V^{p+1,q+1}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}\right)_{1\le i,j\le n} \\ &= \frac{(-1)^{n(n-1)/2}\prod_{1\le i<j\le n}\det V^{p+1,q+1}(x_i,x_j,\boldsymbol{z};a_i,a_j,\boldsymbol{c})\det V^{p+1,q+1}(y_i,y_j,\boldsymbol{z};b_i,b_j,\boldsymbol{c})}{\prod_{i,j=1}^n \det V^{p+1,q+1}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}, \end{aligned} \tag{4.1}$$

$$\begin{aligned} &\det\left(\frac{1}{\det W^{p+2}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}\right)_{1\le i,j\le n} \\ &= \frac{(-1)^{n(n-1)/2}\prod_{1\le i<j\le n}\det W^{p+2}(x_i,x_j,\boldsymbol{z};a_i,a_j,\boldsymbol{c})\det W^{p+2}(y_i,y_j,\boldsymbol{z};b_i,b_j,\boldsymbol{c})}{\prod_{i,j=1}^n \det W^{p+2}(x_i,y_j,\boldsymbol{z};a_i,b_j,\boldsymbol{c})}. \end{aligned} \tag{4.2}$$

If $p = q = 0$, then the identity (4.1) becomes

$$\det\left(\frac{1}{b_j - a_i}\right)_{1\le i,j\le n} = \frac{(-1)^{n(n-1)/2}\prod_{1\le i<j\le n}(a_j - a_i)(b_j - b_i)}{\prod_{i,j=1}^{n}(b_j - a_i)},$$

which is equivalent to Cauchy's determinant identity (1.1).

Proof. It follows from Desnanot–Jacobi formula (2.1) that it suffices to show the identities in the case $n = 2$. If we put

$$f(x, y; a, b) = \det V^{p+1,q+1}(x, y, \boldsymbol{z}; a, b, \boldsymbol{c}), \quad \text{or} \quad \det W^{p+2}(x, y, \boldsymbol{z}; a, b, \boldsymbol{c}),$$

then the case $n = 2$ is equivalent to the following quadratic relation:

$$f(x_1, x_2; a_1, a_2)f(y_1, y_2; b_1, b_2) - f(x_1, y_1; a_1, b_1)f(x_2, y_2; a_2, b_2) + f(x_1, y_2; a_1, b_2)f(x_2, y_1; a_2, b_1) = 0. \tag{4.3}$$

This relation can be obtained by the Plücker relation for determinants. □

5 A hyperpfaffian expression

The purpose of this section is to obtain a hyperpfaffian expression of $\det V^{p,q}(\boldsymbol{x}; \boldsymbol{a})$ when $p = q$ is even. First we recall the definition of hyperpfaffians ([1], see also [14]). Let n and r be positive integers. Define a subset $\mathcal{E}_{rn,n}$ of the symmetric groups $\mathcal{S}_{rn}$ by

$$\mathcal{E}_{rn,n} = \{\sigma \in \mathcal{S}_{rn} : \sigma(n(i-1)+1) < \sigma(n(i-1)+2) < \cdots < \sigma(ni) \text{ for } 1 \le i \le n\}.$$

For example, if $n = r = 2$, then $\mathcal{E}_{4,2}$ is composed of the following 6 elements:

$$\mathcal{E}_{4,2} = \{(1, 2, 3, 4), (1, 3, 2, 4), (1, 4, 2, 3), (3, 4, 1, 2), (2, 4, 1, 3), (2, 3, 1, 4)\}.$$

Let $a = (a_{i_1\dots i_n})_{1\le i_1<\dots<i_n\le nr}$ be an alternating tensor, i.e. $a_{i_{\sigma(1)}\cdots i_{\sigma(n)}} = \operatorname{sgn}(\sigma)a_{i_1\dots i_n}$ for any permutations $\sigma \in \mathcal{S}_{nr}$. The hyperpfaffian of a is, by definition,

$$\operatorname{Pf}^{[n]}(a) = \frac{1}{r!}\sum_{\sigma\in\mathcal{E}_{nr,n}}\operatorname{sgn}(\sigma)\prod_{i=1}^{r}a_{\sigma(n(i-1)+1),\dots,\sigma(ni)}.$$

An alternating 2-tensor a is a skew-symmetric matrix and the hyperpfaffian $\operatorname{Pf}^{[2]}(a)$ is the usual Pfaffian of the skew-symmetric matrix.

The main result of this section is the following Theorem. (Similar expressions are obtained when n is odd.)

Theorem 5.1. If n is even, then

$$\det V^{n,n}(\boldsymbol{x}; \boldsymbol{a}) = \operatorname{Pf}^{[n]}\left[\left(1 + \prod_{s=1}^{n} a_{i_s}\right)\prod_{1\le s<t\le n}(x_{i_t} - x_{i_s})\right]_{1\le i_1<\cdots<i_n\le 2n}, \tag{5.1}$$

$$\det U^{n,n}\begin{pmatrix}\boldsymbol{x} & \boldsymbol{a}\\ \boldsymbol{y} & \boldsymbol{b}\end{pmatrix} = \operatorname{Pf}^{[n]}\left[\left(\prod_{s=1}^{n} a_{i_s} + \prod_{s=1}^{n} b_{i_s}\right)\prod_{1\le s<t\le n}\det\begin{pmatrix}y_{i_s} & x_{i_s}\\ y_{i_t} & x_{i_t}\end{pmatrix}\right]_{1\le i_1<\cdots<i_n\le 2n}. \tag{5.2}$$

The essential part of the proof of the theorem is to compute the following special Pfaffian and hyperpfaffian.

Lemma 5.2. Let n and r be positive integers and assume $n = 2m$ is even. Then we have

$$\operatorname{Pf}\left(\frac{(x_j^m - x_i^m)^2}{x_j - x_i}\right)_{1\le i,j\le nr} = \begin{cases}\prod_{1\le i<j\le n}(x_j - x_i) & \text{if } r = 1,\\ 0 & \text{if } r \ge 2,\end{cases} \tag{5.3}$$

$$\operatorname{Pf}^{[n]}\left[\prod_{1\le s<t\le n}(x_{i_t} - x_{i_s})\right]_{1\le i_1<\cdots<i_n\le nr} = \begin{cases}\prod_{1\le i<j\le n}(x_j - x_i) & \text{if } r = 1,\\ 0 & \text{if } r \ge 2.\end{cases} \tag{5.4}$$

Proof. The first identity (5.3) is obtained from (1.4) by putting $p = q = r = s = 0$ and $a_i = b_i = x_i^m$. The second identity (5.4) follows from the first one (5.3) and the composition of hyperpfaffians given in [14, Eq. (82)]. $\square$

6 Application to Littlewood–Richardson coefficients

In this section, we use the Pfaffian identity (1.5) in Theorem 1.1 and the minor-summation formula [7] to derive a relation between Littlewood–Richardson coefficients.

For three partitions λ, μ and ν, we denote by $\mathrm{LR}^{\lambda}_{\mu,\nu}$ the Littlewood-Richardson coefficient. These numbers $\mathrm{LR}^{\lambda}_{\mu,\nu}$ appear in the following expansions (see [15]) :

$$s_\mu(X)s_\nu(X) = \sum_{\lambda} \mathrm{LR}^{\lambda}_{\mu,\nu}\, s_\lambda(X),$$
$$s_{\lambda/\mu}(X) = \sum_{\nu} \mathrm{LR}^{\lambda}_{\mu,\nu}\, s_\nu(X),$$
$$s_\lambda(X,Y) = \sum_{\mu,\nu} \mathrm{LR}^{\lambda}_{\mu,\nu}\, s_\mu(X)s_\nu(Y).$$

We are concerned with the Littlewood–Richardson coefficients involving rectangular partitions. Let $\square(a,b)$ denote the partition whose Young diagram is the rectangle $a \times b$, i.e.

$$\square(a,b) = (b^a) = (\underbrace{b,\dots,b}_{a}).$$

For a partition $\lambda \subset \square(a,b)$, we define $\lambda^\dagger = \lambda^\dagger(a,b)$ by

$$\lambda^\dagger_i = b - \lambda_{a+1-i} \quad (1 \le i \le a).$$

Okada [16] used the special case of the identities (1.3) and (1.4) (i.e., the case of $p = q = 0$ and $p = q = r = s = 0$) to prove the following proposition. (This proposition is also derived by the combinatorial algorithm called Littlewood–Richardson rule.)

Proposition 6.1. Let n be a positive integer and let e and f be nonnegative integers.

(1) For partitions μ, ν, we have

$$\mathrm{LR}^{\square(n,e)}_{\mu,\nu} = \begin{cases} 1 & \text{if } \nu = \mu^\dagger(n,e), \\ 0 & \text{otherwise.} \end{cases} \tag{6.1}$$

(2) For a partition λ of length $\le 2n$, we have

$$\mathrm{LR}^{\lambda}_{\square(n,e),\square(n,f)} = \begin{cases} 1 & \text{if } \lambda_{n+1} \le \min(e,f) \text{ and } \lambda_i + \lambda_{2n+1-i} = e+f \ (1 \le i \le n), \\ 0 & \text{otherwise.} \end{cases} \tag{6.2}$$

The main result of this section is the following theorem, which generalizes (6.2). It would be interesting to find a bijective proof of the equality (6.5).

Theorem 6.2. Let n be a positive integer and let e and f be nonnegative integers. Let λ and μ be partitions such that $\lambda \subset \square(2n, e+f)$ and $\mu \subset \square(n,e)$. Then we have

(1) $\mathrm{LR}^{\lambda}_{\mu,\square(n,f)} = 0$ unless

$$\lambda_n \ge f \quad \text{and} \quad \lambda_{n+1} \le \min(e,f). \tag{6.3}$$

(2) If λ satisfies the above condition (6.3) and we define two partitions α and β by

$$\alpha_i = \lambda_i - f, \quad \beta_i = e - \lambda_{2n+1-i}, \quad (1 \le i \le n), \tag{6.4}$$

then we have

$$\mathrm{LR}^{\lambda}_{\mu,\square(n,f)} = \mathrm{LR}^{\beta}_{\alpha,\mu^\dagger(n,e)}. \tag{6.5}$$

In particular, $\mathrm{LR}^{\lambda}_{\mu,\square(n,f)} = 0$ unless $\alpha \subset \beta$.

In particular, if $\mu = \square(n,e)$ is a rectangle, then this theorem reduces to (6.2). If μ is a near-rectangle, then we have the following corollary by using Pieri's rule [15, (5.16), (5.17)].

Corollary 6.3. Suppose that a partitions $\lambda \subset \square(2n, e+f)$ satisfies the condition (6.3) in Theorem 6.2. Define two partitions α and β by (6.4). Then we have

$$\mathrm{LR}^{\lambda}_{(e^{n-1},e-k),(f^n)} = \begin{cases} 1 & \text{if } \beta/\alpha \text{ is a horizontal strip of length } k, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathrm{LR}^{\lambda}_{(e^{n-k},(e-1)^k),(f^n)} = \begin{cases} 1 & \text{if } \beta/\alpha \text{ is a vertical strip of length } k, \\ 0 & \text{otherwise.} \end{cases}$$

In order to prove Theorem 6.2, we substitute

$$a_i = x_i^{e+p+n}, \quad b_i = x_i^{f+r+n}, \quad c_i = z_i^{e+p+n}, \quad d_i = w_i^{f+r+n} \tag{6.6}$$

in the Pfaffian identity (1.4). By the bi-determinant definition of Schur functions, we have

$$\det V^{p,q}(\boldsymbol{x}; \boldsymbol{x}^k) = \begin{cases} s_{\square(q,k-p)}(\boldsymbol{x})\Delta(\boldsymbol{x}) & \text{if } k \geq p, \\ 0 & \text{if } k < p, \end{cases}$$

where $\Delta(\boldsymbol{x}) = \prod_{1\leq i<j\leq n}(x_j - x_i)$. Hence, under the substitution (6.6), the identity (1.4) gives us the following Pfaffian identity.

Proposition 6.4. We have

$$\frac{1}{\Delta(\boldsymbol{x})} \operatorname{Pf}\left((x_j - x_i) s_{\square(q+1,e+n-1)}(x_i, x_j, \boldsymbol{z}) s_{\square(s+1,f+n-1)}(x_i, x_j, \boldsymbol{w})\right)_{1\leq i,j\leq 2n}$$
$$= s_{\square(q,e+n)}(\boldsymbol{z})^{n-1} s_{\square(s,f+n)}(\boldsymbol{w})^{n-1} s_{\square(n+q,e)}(\boldsymbol{x}, \boldsymbol{z}) s_{\square(n+s,f)}(\boldsymbol{x}, \boldsymbol{w}). \tag{6.7}$$

Remark 6.5. If we substitute

$$a_i = x_i^{e+p+n}, \quad b_i = y_i^{e+p+n} \quad (1 \leq i \leq n)$$

in the determinant identity (1.3), then we have

$$\frac{1}{\Delta(\boldsymbol{x})\Delta(\boldsymbol{y})} \det\left(s_{\square(q+1,e+n-1)}(x_i, y_j, \boldsymbol{z})\right)_{1\leq i,j\leq n}$$
$$= (-1)^{n(n-1)/2} s_{\square(q,e+n)}(\boldsymbol{z})^{n-1} s_{\square(q+n,e)}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}). \tag{6.8}$$

The special case ($q = e+n-1$) of this identity is given in [13, Proposition 8.4.3], and the proof there works in the general case of (6.8).

If we take $q = s = 0$ in (6.7), we have

$$\frac{1}{\Delta(\boldsymbol{x})} \operatorname{Pf}\left((x_j - x_i) h_{e+n-1}(x_i, x_j, \boldsymbol{z}) h_{f+n-1}(x_i, x_j, \boldsymbol{w})\right)_{1\leq i,j\leq 2n}$$
$$= s_{\square(n,e)}(\boldsymbol{x}, \boldsymbol{z}) s_{\square(n,f)}(\boldsymbol{x}, \boldsymbol{w}), \tag{6.9}$$

where h_r denotes the rth complete symmetric function. We use the minor-summation formula [7] to expand the left hand side in the Schur function bases $\{s_\lambda(\boldsymbol{x})\}$.

Lemma 6.6. Let $b_{k,l}$ be the coefficient of $x^k y^l$ in

$$(y - x) h_{e+n-1}(x, y, \boldsymbol{z}) h_{f+n-1}(x, y, \boldsymbol{w}).$$

Then we have $b_{kl} = -b_{lk}$, and b_{kl}, $k < l$, is given by

$$b_{kl} = \sum_{i,j} h_i(\boldsymbol{z}) h_j(\boldsymbol{w}),$$

where the sum is taken over all pairs of integers (i, j) satisfying

$$i + j = (e+n-1) + (f+n-1) + 1 - k - l, \quad 0 \leq i \leq (e+n-1) - k, \quad 0 \leq j \leq (f+n-1) - k.$$

Here we recall the minor summation formula [7].

Lemma 6.7. Let X be a $2n \times N$ matrix and A be an $N \times N$ skew-symmetric matrix. Then we have

$$\sum_I \operatorname{Pf} \Delta_I^I(A) \det \Delta_I^{[2n]}(X) = \operatorname{Pf}(XA^tX),$$

where I runs over all $2n$-element subsets of $[N]$, and $\Delta_J^I(M)$ denotes the submatrix of a matrix M obtained by picking up the rows indexed by I and the columns indexed by J.

By applying this minor-summation formula, we obtain

Proposition 6.8. Let $B = (b_{ij})_{0 \le i,j \le e+f+2n-1}$ be the skew-symmetric matrix, whose entries b_{ij} are given in Lemma 6.6. Then, for a partition $\lambda \subset \Box(2n, e+f)$, we have

$$\sum_{\substack{\mu \subset \Box(n,e) \\ \nu \subset \Box(n,f)}} \mathrm{LR}^{\lambda}_{\mu,\nu}\, s_{\mu^\dagger(n,e)}(\boldsymbol{z}) s_{\nu^\dagger(n,f)}(\boldsymbol{w}) = \operatorname{Pf} \Delta^{I(\lambda)}_{I(\lambda)}(B), \tag{6.10}$$

where $I(\lambda) = \{\lambda_{2n}, \lambda_{2n-1}+1, \dots, \lambda_2+2n-2, \lambda_1+2n-1\}$.

Now we can finish the proof of Theorem 6.2.

Proof of Theorem 6.2. In the above argument, we take $p \ge n$ and $r = 0$. In this case, the variables $\boldsymbol{w}$ disappear and we see that

$$b_{kl} = \begin{cases} h_{(e+n-1)+(f+n-1)+1-k-l}(\boldsymbol{z}) & \text{if } 0 \le k \le \min(e+n-1, f+n-1) \text{ and } l \ge f+n-1, \\ 0 & \text{otherwise} \end{cases}$$

and the equation (6.10) becomes

$$\sum_{\mu \subset \Box(n,e)} \mathrm{LR}^{\lambda}_{\mu,\Box(n,f)}\, s_{\mu^\dagger(n,e)}(\boldsymbol{z}) = \operatorname{Pf} \Delta^{I(\lambda)}_{I(\lambda)}(B).$$

The skew-symmetric matrix B has the form $B = \begin{pmatrix} O & C \\ -{}^tC & O \end{pmatrix}$ with

$$C = \left(h_{e+n-1-i-j}(\boldsymbol{z})\right)_{0 \le i \le f+n-1, 0 \le j \le e+n-1}.$$

¿From the relation (2.10), we see that the subPfaffian $\operatorname{Pf} B_{I(\lambda)}$ vanishes unless

$$\lambda_{n+1} \le \min(e,f), \quad \lambda_n \ge f.$$

If these conditions are satisfied, then we have

$$\begin{aligned} \operatorname{Pf} B_{I(\lambda)} &= (-1)^{n(n-1)/2} \det \left(h_{\beta_i - \alpha_{n+1-j} - i + (n+1-j)}(\boldsymbol{z})\right)_{1 \le i,j \le n} \\ &= s_{\beta/\alpha}(\boldsymbol{z}). \end{aligned}$$

Hence we have

$$\sum_{\mu \subset \Box(n,e)} \mathrm{LR}^{\lambda}_{\mu,\Box(n,f)}\, s_{\mu^\dagger(n,e)}(\boldsymbol{z}) = s_{\beta/\alpha}(\boldsymbol{z}).$$

Comparing the coefficients of $s_{\mu^\dagger(n,e)}(\boldsymbol{z})$ completes the proof. □

References

[1] A. I. Barvinok, New algorithms for linear k-matroid intersection and matroid k-parity problems, Math. Program. **69** (1995), 449–470.

[2] D. M. Bressoud, "Proofs and Confirmations : The Story of the Alternating Sign Matrix Conjecture", Cambridge Univ. Press, 1999.

[3] A. L. Cauchy, Mémoire sur les fonctions alternées et sur les sommes alternées, Exercices Anal. et Phys. Math. **2** (1841), 151–159.

[4] G. Frobenius, Über die elliptischen Funktionen zweiter Art, J. Reine und Angew. Math. **93** (1882), 53–68.

[5] M. Ishikawa, Minor summation formula and a proof of Stanley's open problem, `arXiv:math.CO/0408204`.

[6] M. Ishikawa, S. Okada, H. Tagawa and J. Zeng Generalizations of Cauchy's determinant and Schur's Pfaffian", `arXiv:math.CO/0411280`.

[7] M. Ishikawa and M. Wakayama, Minor summation formulas of Pfaffians, Linear and Multilinear Algebra **39** (1995), 285–305.

[8] M. Ishikawa and M. Wakayama, Applications of the minor summation formula III: Plücker relations, lattice paths and Pfaffians, `arXiv:math.CO/0312358`.

[9] D. Knuth, Overlapping Pfaffians, Electron. J. Combin. **3** (2) ("The Foata Festschrift") (1996), 151–163.

[10] C. Krattenthaler, Advanced determinant calculus, Sem. Lothar. Combin. **42** (1999), B42q.

[11] C. Krattenthaler, Advanced determinant calculus : a complement, `arXiv:math.CO/0503507`.

[12] D. Laksov, A. Lascoux and A. Thorup, On Giambelli's theorem on complete correlations, Acta Math. **162** (1989), 143–199.

[13] A. Lascoux, Symmetric Functions and Combinatorial Operators on Polynomials, CBMS Reg. Conf. Ser. Math. **99**, AMS, 2003.

[14] J.-G. Luque and J.-Y. Thibon, Pfaffian and Hafnian identities in shuffle algebras, Adv. Appl. Math. **29** (2002), 620–646.

[15] I. G. Macdonald, "Symmetric Functions and Hall Polynomials (2nd ed.),"Oxford Univ. Press, 1995.

[16] S. Okada, Applications of minor-summation formulas to rectangular-shaped representations of classical groups, J. Alg. **205** (1998), 337–367.

[17] S. Okada, Enumeration of symmetry classes of alternating sign matrices and characters of classical groups, `arXiv:math.CO/0408234`.

[18] S. Okada, Determinant and Pfaffian formulae of Cauchy type and their applications, in "Aspects of Combinatorial Representation Theory," Sūrikaisekikenkyūsho Kōkyūroku **1382** (2004), 198–215 (in Japanese).

[19] S. Okada, An elliptic generalization of Schur's Pfaffian identity, `arXiv:math.CA/0412038`.

[20] I. Schur, Über die Darstellung der symmetrischen und der alternirenden Gruppe durch gebrochene lineare Substitutionen, J. Reine Angew. Math. **139** (1911), 155–250.

[21] R. P. Stanley, Open problem, presented at International Conference on Formal Power Series and Algebraic Combinatorics (Vadstena 2003), June 23 – 27, 2003, available from `http://www-math.mit.edu/ rstan/trans.html`.

[22] J. R. Stembridge, Non-intersecting paths, Pfaffians and plane partitions, Adv. Math. **83** (1990), 96–131.

[23] T. Sundquist, Two variable Pfaffian identities and symmetric functions, J. Algebraic Combin. **5** (1996), 135–148.

AFFINE STANLEY SYMMETRIC FUNCTIONS

THOMAS LAM

Abstract. We define a new family $\tilde{F}_w(X)$ of generating functions for $w \in \tilde{S}_n$ which are affine analogues of Stanley symmetric functions. We establish basic properties of these functions such as their symmetry and conjecture certain positivity properties. As an application, we relate these functions to the k-Schur functions of Lapointe, Lascoux and Morse as well as the cylindric Schur functions of Postnikov.

In [Sta84], Stanley introduced a family $\{F_w(X)\}$ of symmetric functions now known as *Stanley symmetric functions*. He used these functions to study the number of reduced decompositions of permutations $w \in S_n$. Later, the functions $F_w(X)$ were found to be stable limits of Schubert polynomials. Another fundamental property of Stanley symmetric functions is the fact that they are Schur-positive ([EG, LS]).

This extended abstract describes work in progress on an analogue of Stanley symmetric functions for the affine symmetric group $\tilde{S}_n$ which we call *affine Stanley symmetric functions*. Our first main theorem is that these functions $\tilde{F}_w(X)$ are indeed symmetric functions. Most of the other main properties of Stanley symmetric functions established in [Sta84] also have analogues in the affine setting.

Our definition of affine Stanley symmetric functions is motivated by relations with two other classes of symmetric functions which have received attention lately. Lapointe, Lascoux and Morse [LLM] initiated the study of *k-Schur functions*, denoted $s_\lambda^{(k)}(X)$, in their study of Macdonald polynomial positivity. Lapointe and Morse have more recently connected k-Schur functions with the Verlinde algebra of $SL(n)$. Separately, *cylindric Schur functions* were defined by Postnikov [Pos] in connection with the quantum cohomology $QH^*(Gr_{m,n})$ of the Grassmannian (see also [GK]). We shall connect these two classes of symmetric functions via affine Stanley symmetric functions. More precisely, we show that when $w \in \tilde{S}_n$ is a "Grassmannian" affine permutation then $\tilde{F}_w(X)$ is "dual" to the k-Schur functions $s_\lambda^{(k)}(X)$. We call these functions $\tilde{F}_w(X)$ *affine Schur functions*. Affine Schur functions were earlier defined by Lapointe and Morse who called them *dual k-Schur functions*. In analogy with the usual Stanley symmetric function case, conjecture that all affine Stanley symmetric functions expand positively in terms of affine Schur functions. We then show that cylindric Schur functions are special cases of certain *skew affine Schur functions* and correspond to 321-avoiding affine permutations.

The non-affine case suggests that our work may be connected with the affine flag variety and objects that might be called "affine Schubert polynomials". Shimozono has conjectured a precise relationship between k-Schur functions and the homology of the affine Grassmannian. The dual conjecture ([MS]) is that affine Schur functions represent Schubert classes in the cohomology $H^*(\mathcal{G}/\mathcal{P})$ of the affine Grassmannian.

In section 1, we establish some notation for permutations and affine permutations, and for symmetric functions. In section 2 we recall the definition of Stanley symmetric functions, give their main properties and explain the relationship with Schubert polynomials. In section 3, we define affine Stanley symmetric functions and prove that they are symmetric. In section 4, we

---

Date: November, 2004; revised February, 2005.

give basic properties of affine Stanley symmetric functions, imitating the results of [Sta84]. In section 5, we define affine Schur functions and relate them to k-Schur functions. In section 6, we connect skew affine Schur functions with cylindric Schur functions. In section 7, we make a number of positivity conjectures concerning the expansion of affine Stanley symmetric functions in terms of affine Schur functions. Finally, in section 8, we discuss relations with the affine flag variety and a generalisation to affine stable Grothendieck polynomials.

We should remark that Stanley symmetric functions for the hyperoctahedral group have also been defined; see for example [LTK].

A full version of this paper can be found on the archive: `math.CO/0501335`.

Acknowledgements. I thank Jennifer Morse and Mark Shimozono for interesting discussions about k-Schur functions. I am grateful to Alex Postnikov for introducing toric Schur functions in his class. I am also indebted to my advisor, Richard Stanley for guidance over the last few years.

1. Preliminaries

1.1. **Symmetric group.** A positive integer $n \geq 2$ will be fixed throughout the paper. Let $\tilde{S}_n$ denote the affine symmetric group with simple generators $s_0, s_1, \ldots, s_{n-1}$ satisfying the relations

$$\begin{aligned} s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1} && \text{for all } i \\ s_i^2 &= 1 && \text{for all } i \\ s_i s_j &= s_j s_i && \text{for } |i-j| \geq 2. \end{aligned}$$

Here and elsewhere, the indices will be taken modulo n without further mention. There are many different explicit constructions of $\tilde{S}_n$, see for example [BB]. The symmetric group S_n embeds in $\tilde{S}_n$ as the subgroup generated by $s_1, s_2, \ldots, s_{n-1}$.

For an element $w \in \tilde{S}_n$ let $R(w)$ denote the set of *reduced words* for w. A word $\rho = (\rho_1\rho_2\cdots\rho_l) \in [0, n-1]^l$ is a reduced word for w if $w = s_{\rho_1}s_{\rho_2}\cdots s_{\rho_l}$ and l is the smallest possible integer for such a decomposition exists. The integer $l = l(w)$ is called the *length* of w. If $\rho, \pi \in R(w)$ for some w, then we write $\rho \sim \pi$. If ρ is an arbitrary word with letters from $[0, n-1]$ then we write $\rho \sim 0$ if it is not a reduced word of any affine permutation. If $w, u \in \tilde{S}_n$ then we say that w *covers* u and write $w \gtrdot u$ if $w = s_i \cdot u$ and $l(w) = l(u) + 1$. The transitive closure of $\gtrdot$ is called the *weak Bruhat order* and denoted $\succcurlyeq$.

1.2. **Symmetric functions.** We will follow mostly [Mac, Sta99] for our symmetric function notation. Let Λ denote the ring of symmetric functions. Usually, our symmetric functions will have variables $x_1, x_2, \ldots$ and will be written as $f(x_1, x_2, \ldots)$ or $f(X)$. If we need to emphasize the variable used, we write Λ_X. We use λ, μ and ν to denote partitions. We will use m_λ, p_λ, e_λ, h_λ and s_λ to denote the *monomial*, *power sum*, *elementary*, *homogeneous* and *Schur* bases of Λ. Let $\langle .,. \rangle$ denote the *Hall inner product* of Λ satisfying $\langle h_\lambda, m_\mu\rangle = \langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$. For $f \in \Lambda$, write $f^\perp : \Lambda \to \Lambda$ for the linear operator adjoint to multiplication by f with respect to $\langle .,. \rangle$. We let $\omega : \Lambda \to \Lambda$ denote the algebra involution of Λ sending h_n to e_n.

If $f(X) \in \Lambda$ then $f(X, Y) = \sum_i f_i(X) \otimes g_i(Y) \in \Lambda_X \otimes \Lambda_Y$ for some f_i and g_i. This is the *coproduct* of f, written $\Delta f = \sum_i f_i \otimes g_i \in \Lambda \otimes \Lambda$. We have the following formula for the coproduct ([Mac]):

$$\Delta f = \sum_\lambda s_\lambda^\perp f \otimes s_\lambda. \tag{1}$$

Let $\mathcal{P}ar^n$ denote the set $\{\lambda \mid \lambda_1 \leq n-1\}$ of partitions with no row longer than $n-1$. The following two subspaces of Λ will be important to us:

$$\Lambda^{(n)} = \mathbb{C}\left\langle m_\lambda \mid \lambda \in \mathcal{P}ar^n\right\rangle$$
$$\Lambda_{(n)} = \mathbb{C}\left\langle h_\lambda \mid \lambda \in \mathcal{P}ar^n\right\rangle = \mathbb{C}\left\langle e_\lambda \mid \lambda \in \mathcal{P}ar^n\right\rangle = \mathbb{C}\left\langle p_\lambda \mid \lambda \in \mathcal{P}ar^n\right\rangle.$$

If $f \in \Lambda_{(n)}$ and $g \in \Lambda^{(n)}$ then define $\langle f, g\rangle$ to be their usual Hall inner product within Λ. Thus $\{h_\lambda\}$ and $\{m_\lambda\}$ with $\lambda \in \mathcal{P}ar^n$ form dual bases of $\Lambda_{(n)}$ and $\Lambda^{(n)}$. Note that $\Lambda_{(n)}$ is a subalgebra of Λ but $\Lambda^{(n)}$ is not closed under multiplication. Instead, $\Lambda^{(n)}$ is a coalgebra; it is closed under comultiplication.

2. Stanley symmetric functions

Let $w \in S_n$ with length $l = l(w)$. Define the generating function $F_{w^{-1}}(X)$ by

$$F_{w^{-1}}(x_1, x_2, \ldots) = \sum_{a_1 a_2 \cdots a_l \in R(w)} \sum_{\substack{1 \leq b_1 \leq b_2 \leq \cdots \leq b_l \\ a_i > a_{i+1} \Rightarrow b_{i+1} > b_i}} x_{b_1} x_{b_2} \cdots x_{b_l}.$$

We have indexed the $F_{w^{-1}}(X)$ by the inverse permutation to agree with the definition we shall give later. These generating functions, known as *Stanley symmetric functions*, were shown in [Sta84] to be symmetric. Stanley also studied these functions under the action of ω, the action of $s_1^\perp$ and also proved that the Schur expansions of $F_w(X)$ possess *dominant* terms. Edelman and Greene [EG] and Lascoux and Schützenberger [LS] showed that Stanley symmetric functions are Schur positive so that if

$$F_w(X) = \sum_\lambda a_{w\lambda} s_\lambda(X)$$

then $a_{w\lambda} \geq 0$. Note that the length $l(w)$ is equal to the degree of F_w and the number $|R(w)|$ of reduced decompositions of w is given by the coefficient of $x_1 x_2 \cdots x_l$ in F_w. We now give a different formulation of the definition in a manner similar to [FS].

Let $\mathbb{C}[S_n]$ denote the group algebra of the symmetric group equipped with a inner product $\langle w, v\rangle = \delta_{wv}$. Define linear operators $u_i : \mathbb{C}[S_n] \to \mathbb{C}[S_n]$ for $i \in [1, n-1]$ by

$$u_i.w = \begin{cases} s_i.w & \text{if } l(s_i.w) > l(w), \\ 0 & \text{otherwise.} \end{cases}$$

The operators satisfy the braid relations $u_i u_{i+1} u_i = u_{i+1} u_i u_{i+1}$ together with $u_i^2 = 0$ and $u_i u_j = u_j u_i$ for $|i-j| \geq 2$. They generate an algebra known as the *nilCoxeter algebra*. Note that the action on $\mathbb{C}[S_n]$ is a faithful representation of these relations.

Let $A_k(u) = \sum_{b_1 > b_2 > \cdots > b_k} u_{b_1} u_{b_2} \cdots u_{b_k}$. Then the Stanley symmetric functions can be written as

$$F_w(X) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle A_{a_t}(u) A_{a_{t-1}}(u) \cdots A_{a_1}(u) \cdot 1, w\right\rangle x_1^{a_1} x_2^{a_2} \cdots x_t^{a_t} \tag{2}$$

where the sum is over all compositions a.

For completeness, we explain briefly the relationship between $F_w(X)$ and the Schubert polynomials of Lascoux and Schützenberger. For $w \in S_n$, we have a Schubert polynomial $\mathfrak{S}_w \in \mathbb{C}[x_1, x_2, \ldots, x_{n-1}]$. If $w \in S_n$, then $w \times 1^s \in S_{n+s}$ denotes the corresponding permutation of S_{n+s} acting on the elements $[1, n]$ of $[1, n+s]$. Similarly, $1^s \times w \in S_{n+s}$ denotes the corresponding permutation acting on the elements $[s+1, n+s]$ of $[1, n+s]$. Schubert polynomials have the important stability property $\mathfrak{S}_w = \mathfrak{S}_{w \times 1^s}$. Stanley symmetric functions $F_w(X)$ are obtained by taking the other limit: $F_w = \lim_{s\to\infty} \mathfrak{S}_{1^s\times w}$. The limit is taken by treating both sides as formal power series and taking the limit of each coefficient.

3. Affine Stanley symmetric functions

Our first definition of affine Stanley symmetric functions will imitate the definition (2) above. Let $\mathcal{U}_n$ be the *affine nilCoxeter algebra* generated over $\mathbb{C}$ by generators $u_0, u_1, \ldots, u_{n-1}$ satisfying

$$\begin{aligned} u_i^2 &= 0 && \text{for all } i \in [0,n],\\ u_iu_{i+1}u_i &= u_{i+1}u_iu_{i+1} && \text{for all } i \in [0,n],\\ u_iu_j &= u_ju_i && \text{for all } i,j \in [0,n] \text{ satisfying } |i-j| \geq 2. \end{aligned}$$

Here and henceforth the indices are to be taken modulo n. A basis of $\mathcal{U}_n$ is given by the elements $u_w = u_{\rho_1}u_{\rho_2}\cdots u_{\rho_l}$ where $\rho = (\rho_1\rho_2\cdots\rho_l)$ is some reduced word for w.

Define $h_k(\mathbf{u}) \in \mathcal{U}_n$ for $k \in [0, n-1]$ by

$$h_k(\mathbf{u}) = \sum_{A \in \binom{[0,n-1]}{k}} u_A$$

where for a k-subset $A = \{a_1, a_2, \ldots, a_k\} \subset [0, n-1]$ the element $u_A \in \mathcal{U}_n$ is defined as any expression $u_{a_1}u_{a_2}\cdots u_{a_k}$ where if i and $i+1$ (modulo n) are both in A then u_{i+1} must precede u_i. All such expressions are equal within $\mathcal{U}_n$. For example if $n = 9$ and $A = \{0, 2, 4, 5, 6, 8\}$ then $u_A = u_0u_8u_2u_6u_5u_4 = u_2u_6u_5u_4u_0u_8 = \cdots$. A similar definition of $h_k(\mathbf{u})$ was given by Postnikov [Pos], in the context of the *affine nil-Temperley-Lieb algebra.*

Define a representation of $\mathcal{U}_n$ on $\mathbb{C}[\tilde{S}_n]$ by

$$u_i.w = \begin{cases} s_i.w & \text{if } l(s_i.w) > l(w),\\ 0 & \text{otherwise.}\end{cases}$$

It is easy to see that this is indeed a representation of $\mathcal{U}_n$. Equip $\mathbb{C}[\tilde{S}_n]$ with the inner product $\langle w, v\rangle = \delta_{wv}$. The following definition was heavily influenced by [FG].

Definition 1. Let $w \in \tilde{S}_n$. Define the *affine Stanley symmetric functions* $\tilde{F}_w(X)$ by

$$\tilde{F}_w(X) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle h_{a_t}(u)h_{a_{t-1}}(u)\cdots h_{a_1}(u)\cdot 1, w\right\rangle x_1^{a_1}x_2^{a_2}\cdots x_t^{a_t},$$

where the sum is over compositions of $l(w)$ satisfying $a_i \in [0, n-1]$.

The seemingly more general "skew" affine Stanley symmetric functions

$$\tilde{F}_{w/v}(X) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle h_{a_t}(u)h_{a_{t-1}}(u)\cdots h_{a_1}(u)\cdot v, w\right\rangle x_1^{a_1}x_2^{a_2}\cdots x_t^{a_t}$$

are actually equal to the usual affine Stanley symmetric functions $\tilde{F}_{wv^{-1}}(X)$.

Our first proposition follows from the definition.

Proposition 2. *Suppose* $w \in S_n \subset \tilde{S}_n$. *Then* $\tilde{F}_w(X) = F_w(X)$.

The main theorem of this section is the following.

Theorem 3. *The generating functions* $\tilde{F}_w(X)$ *are symmetric.*

Theorem 3 follows immediately from Proposition 5. In the following, intervals $[a, b]$ are to be taken in the cyclic fashion within $[0, n-1]$. Also, max and min of a cyclic interval is meant to be taken modulo n in the obvious manner. So if $n = 6$ then $[4, 1] = \{4, 5, 0, 1\}$ and $\max([4, 1]) = 1$ and $\min([4, 1]) = 4$. We will need a technical lemma first.

Lemma 4. *We have the following identities for reduced words.*

(1) *Let $a, b \in [0, n-1]$ with $a \neq b-1$. Then $a(a-1)(a-2)\cdots ba(a-1)(a-2)\cdots b \sim 0$.*
(2) *Let $a, b, c \in [0, n-1]$ satisfying $a \neq b-1$; $c \neq b$ and $c \in [b, a]$. Then $a(a-1)(a-2)\cdots bc \sim (c-1)a(a-1)(a-2)\cdots b$.*

Proof. Both results can be calculated by induction. $\square$

So for example, the element $(s_4s_3s_2)(s_4s_3s_2)$ is not reduced and we have $(s_6s_5s_4s_3s_2s_1)s_4 = s_3(s_6s_5s_4s_3s_2s_1)$.

Proposition 5. *The elements $h_k(\mathbf{u})$ for $k \in [0, n-1]$ commute.*

Proof. For each $w \in \tilde{S}_n$ satisfying $l(w) = x + y$, we calculate the coefficient of u_w in $h_x(\mathbf{u})h_y(\mathbf{u})$ and $h_y(\mathbf{u})h_x(\mathbf{u})$. We assume that x and y are both not equal to 0 for otherwise the result is obvious. Let $u_w = u_Au_B$ where $|A| = x$ and $|B| = y$. We need to exhibit a bijection between reduced decompositions of this form and those of the form $u_w = u_Cu_D$ with $|C| = y$ and $|D| = x$. We assume for simplicity (though it is not crucial to our proof) that $A \cup B = [0, n-1]$ for otherwise we are in the non-affine case and the proposition follows from results of Stanley [Sta84] or Fomin-Greene [FG]. Let $A = \bigcup_i A_i$ and $B = \bigcup_i B_i$ be minimal decompositions of A and B into cyclic intervals. If $A_i \subset B_j$ for some pair (i, j) then we call A_i an *inner* interval and similarly for $B_k \subset A_l$. Otherwise the interval is called *outer*.

Using Lemma 4 and our assumption that $A \cup B = [0, n-1]$ we can describe the outer intervals in an explicit manner. Each outer interval A_i *touches* an outer interval $\mathrm{rn}(A_i) = B_k$ called the *right neighbour* of A_i, for a unique k, so that $\min(A_i) = \max(B_k) + 1$. Also A_i overlaps with an outer interval $\mathrm{ln}(A_i) = B_l$ for a unique l, so that $\max(A_i) \geq \min(B_l) - 1$ called the *left neighbour*. If $\mathrm{rn}(A_i) = B_k$ then we also write $A_i = \mathrm{ln}(B_k)$ and similarly for $\mathrm{rn}(B_k)$. Note that it is possible that $\mathrm{rn}(A_i) = \mathrm{ln}(A_i)$ since we are working cyclically.

Our bijection will depend only locally on each pair of an outer interval A^* and its right neighbour $B^* = \mathrm{rn}(A^*)$. We call the interval $I = [\min(B^*), \min(\mathrm{ln}(A^*)) - 1]$ a *critical interval*. Critical intervals cover $[0, n-1]$ in a disjoint manner. For example, suppose $n = 10$ and $A = \{1, 2, 3, 6, 7, 8, 9\}$ and $B = \{0, 2, 4, 5, 7, 9\}$ (Figure 1), so that we have $u_Au_B = u_9u_8u_7u_6u_3u_2u_1u_0u_9u_7u_5u_4u_2u_0$. Then $A_1 = [1, 3]$ and $A_2 = [6, 9]$ are both outer intervals. Also $B_1 = [2, 5]$, $B_2 = \{7\}$ and $B_3 = [9, 0]$. Only B_2 is an inner interval. The left neighbour of A_1 is $\mathrm{ln}(A_1) = B_1$ and the right neighbour is $\mathrm{rn}(A_1) = B_3$. The critical intervals are $[9, 1]$ and $[2, 8]$.

Let $a = \min(\mathrm{ln}(A^*)) - 1$ and $b = \min(B^*)$. Let $c = |[b, a]|$, $d = |A \cap [b, a]|$ and $e = |B \cap [b, a]|$. Renaming for convenience, we let $S_1, S_2, \ldots, S_r$ be the inner intervals (of B) contained in A^* and $T_1, \ldots, T_t$ be those contained in B^*, arranged so that $S_k > S_{k+1}$ for all k within $[b, a]$ and similarly $T_k > T_{k+1}$. We now define a subset $U \subset [b, a]$ satisfying $|U| = d$. The algorithm begins with $U = [b, a]$ and a changing index i set to $i := a$ to begin with. The index i decreases from a to b and at each step the element i may be removed from U according to the rule:

(1) If $i \in A^*$ then we remove it from U unless $i \in S_k$ for some $k \in [1, r]$.
(2) If $i \in B^*$ then we remove it from U unless $i \in T_k + 1$ for some $k \in [1, t]$.
(3) Otherwise we do not remove i from U and set $i := i - 1$. Repeat.

When $|U| = d$ we stop the algorithm. The algorithm always terminates with $|U| = d$ since there are at least $c - d = |[b, a]| - (A \cap [b, a])$ elements to remove. In fact the algorithm terminates before $i = b$ since $\cup_i S_i \neq A^* \cap I$. We will denote the result of the algorithm by $\phi(A^* \cup_i T_i, B^* \cup_i S_i) := U$. Note that $\min(U) = b$.

The bijection $u_Au_B \mapsto u_Cu_D$ is obtained by letting $D \subset [0, n-1]$ be the subset obtained from B by changing $B \cap I$ in each critical interval I to U. By the definition of U we see that

$|D| = |A|$. We claim that $u_A u_B = u_C u_D$ or alternatively $s_A s_B (s_D)^{-1} = s_C$ for some C satisfying $|C| = |B|$ (here it is slightly more convenient to calculate within the affine symmetric group, which is legal since our words are all reduced). We can calculate this locally on each critical interval since the $s_{D \cap I}$ commute as I varies over critical intervals. Note that U always has the form of a disjoint union $S_1 \cup S_2 \cup \cdots \cup S_{r'} \cup [b, a']$ for some $r' \leq r$ where $a' > \max(B^*)$ or the form $S_1 \cup \cdots \cup S_r \cup \{T_1 + 1\} \cup \{T_2 + 1\} \cup \cdots \cup \{T_{t'} + 1\} \cup [b, a']$ where $a' \leq \max(B^*)$.

Let us assume that U has the first form. Focusing on $I = [b, a] = [\min(B^*), \min(\ln(A^*)) - 1]$ we are interested in

$$\underline{s} = s_{A^* \cap I} s_{T_1} \cdots s_{T_t} s_{S_1} \cdots s_{S_r} s_{B^*} (s_{[b,a']})^{-1} (s_{S_{r'}})^{-1} \cdots (s_{S_1})^{-1}.$$

Then we get

$$\begin{aligned} \underline{s} &= s_{A^* \cap I} s_{T_1} \cdots s_{T_t} s_{S_{r'+1}} \cdots s_{S_r} (s_{[\max(B^*)+1, a']})^{-1} \\ &= s_{S_{r'+1}-1} \cdots s_{S_r - 1} s_{T_1} \cdots s_{T_t} s_{A^* \cap I} (s_{[\max(B^*)+1, a']})^{-1} \\ &= s_{S_{r'+1}-1} \cdots s_{S_r - 1} s_{T_1} \cdots s_{T_t} s_{[a'+1, a]} \qquad \text{using } \max(B^*) + 1 = \min(A^*). \end{aligned}$$

We used Lemma 4 repeatedly and also the fact that the certain intervals do not "touch" and so commute. Let U' be the disjoint union $[a'+1, a] \cup \{S_{s'+1} - 1\} \cup \cdots \cup \{S_s - 1\} \cup T_1 \cup \cdots T_t$. Note that it is always the case that $\max(U') = a$. The other form of U involves a similar calculation. One checks that we can combine this argument for each critical interval showing that $s_A s_B (s_D)^{-1}$ is indeed equal to s_C for some C.

Finally, we need to show that this map is a bijection. Again we work locally on a critical interval and assume that U has the first form. If we replace A^* (more precisely $A^* \cap I$) by U' and B^* by U, then our internal intervals are $S'_1 = S_1, \ldots S'_{r'} = S_{r'}$ and $T'_1 = S_{r'+1} - 1, \ldots, T'_{r-r'} = S_{r'} - 1, T'_{r-r'+1} = T_1, \ldots T'_{r-r'+t} = T_t$. We now show that $B^* \cup_i S_i = \phi(U', U)$ from which the bijectivity will follow. Note that since $\min(U) = b$ and $\max(U') = a$ the critical intervals of $u_C u_D$ are the same as those of $u_A u_B$. By definition $\phi(U', U)$ keeps $S'_1, S'_2, \ldots$ and keeps $T'_1 + 1, T'_2 - 1, \ldots, T'_{r-r'} + 1$, removing all other values up to this point. At this point the algorithm stops since $\phi(U', U)$ is of the correct size. We see that we obtain $\phi(U', U) = B^* \cup_i S_i$ back in this way. A similar argument works for the second form of U. $\square$

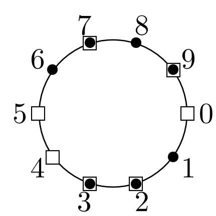


FIGURE 1. Dots represent elements of A. Squares represent elements of B.

Example 1. We illustrate the map $U = \phi(A^* \cup_i T_i, B^* \cup_i S_i)$ of the proof. Suppose $[b, a] = [2, 20]$ and $A^* = [14, 20]$, $B^* = [2, 13]$. Let $S_1 = [16, 18]$ and $T_1 = [8, 11]$ and $T_2 = \{5\}$ be the inner intervals. Then $d = 12$ and $U = \{2, 3, 4, 5, 6, 9, 10, 11, 12, 16, 17, 18\}$. We can compute that

$$s_{A^*} s_{11} s_{10} s_9 s_8 s_5 s_{B^*} s_{18} s_{17} s_{16} s_2 s_3 s_4 s_5 s_6 s_9 s_{10} s_{11} s_{12} s_{16} s_{17} s_{18} = s_{A^*} s_{[7,13]} s_5$$

so that $U' = [7, 20] \cup \{5\}$. Finally one checks that $B^* \cup_i S_i = \phi(U', U)$.

We end this section by giving an alternative description of the affine Stanley symmetric functions. Let $w \in \tilde{S}_n$. Let $a = (a_1, \ldots, a_l) \in R(w)$ be a reduced word and $b = (b_1 \geq b_2 \cdots \geq b_l)$ be an positive integer sequence. Then (a, b) is called a *compatible pair for w* if whenever

$b_i = b_{i+1} = \cdots = b_j$ and $\{k, k+1\} \subset \{a_i, a_{i+1}, \ldots, a_j\}$ then we have that $k+1$ precedes k (for any i, j, k). Two compatible pairs (a, b) and (a', b') are equivalent if $b = b'$ and for any maximal interval $[i, j] \subset [1, l]$ satisfying $b_i = b_{i+1} = \cdots = b_j$ we have that $a_i a_{i+1} \cdots a_j$ and $a'_i a'_{i+1} \cdots a'_j$ are reduced words for the same affine permutation. Then

$$F_w(X) = \sum_{\overline{(a,b)}} x_{b_1} x_{b_2} \cdots x_{b_l}$$

where the sum is over equivalence classes $\overline{(a,b)}$ of compatible pairs for w.

4. Basic properties

We give a number of basic properties for the functions $\tilde{F}_w$. The first main property follows immediately from the definition.

Proposition 6. *We have $\tilde{F}_w \in \Lambda^{(n)}$ for each $w \in \tilde{S}_n$.*

In fact we shall see later that they span the subspace $\Lambda^{(n)}$.

Theorem 7 (Coproduct formula)**.** *The following coproduct expansion holds:*

$$\tilde{F}_w(x_1, x_2, \ldots, y_1, y_2, \ldots) = \sum_{uv=w} \tilde{F}_v(x_1, x_2, \ldots)\tilde{F}_u(y_1, y_2, \ldots).$$

In particular we have

$$s_1^{\perp} \tilde{F}_w = \sum_{w \gtrdot v} \tilde{F}_v.$$

Proof. The first formula follows immediately from the definition and the fact that $\tilde{F}_{w/v}(Y) = \tilde{F}_{wv^{-1}}(Y)$. To obtain the second formula, we first write, using the first formula and (1),

$$\sum_{uv=w} \tilde{F}_v(X) \otimes \tilde{F}_u(Y) = \sum_{\lambda} s_\lambda^{\perp} \tilde{F}_w(X) \otimes s_\lambda(Y).$$

The terms of the formula are to be interpreted within Λ, even though the sum is an element of $\Lambda^{(n)}$. Now take the inner product of both sides with $s_1(Y)$ to get

$$s_1^{\perp} \tilde{F}_w(X) = \sum_{uv=w} \tilde{F}_v(X) \left\langle \tilde{F}_u(Y), s_1(Y) \right\rangle.$$

Now $\left\langle \tilde{F}_u(Y), s_1(Y) \right\rangle = 0$ unless $u = s_i$ is a simple reflection for some i, in which case $\tilde{F}_{s_i}(Y) = s_1(Y)$. This gives the second formula. $\square$

Define $\omega : \Lambda_{(n)} \to \Lambda_{(n)}$ as usual by $\omega : h_i \mapsto e_i$. Define $\omega^+ : \Lambda^{(n)} \to \Lambda^{(n)}$ by requiring that $\langle \omega(f), \omega^+(g) \rangle = \langle f, g \rangle$. Alternatively, we require that $\{e_\lambda \mid \lambda \in \mathcal{P}ar^n\}$ and $\{\omega^+(m_\lambda) \mid \lambda \in \mathcal{P}ar^n\}$ form dual bases. The map ω^+ is clearly an involution but it does not agree with ω (see for example [Sta99, Chapter 7, Ex. 9]).

Denote by $w \mapsto w^*$ the involution of $\tilde{S}_n$ given by $s_i \mapsto s_{n-1-i}$.

Theorem 8 (Conjugacy formula)**.** *Let $w \in \tilde{S}_n$. Then $\omega^+(\tilde{F}_w) = \tilde{F}_{w^*}$.*

We shall prove Theorem 8 by calculating within the subalgebra $\Lambda_{(n)}(\mathbf{u})$ of $\mathcal{U}_n$ generated by $\{h_k(\mathbf{u})\}$. By Proposition 5, this subalgebra is naturally the homomorphic image to $\Lambda_{(n)}$. In fact $\Lambda_{(n)}(\mathbf{u})$ is abstractly isomorphic to $\Lambda_{(n)}$. For an element $f \in \Lambda_{(n)}$, we let $f(\mathbf{u})$ denote the corresponding image in $\Lambda_{(n)}(\mathbf{u})$.

Theorem 9. *The elements $e_k(\mathbf{u}) \in \mathcal{U}_n$ are given by*

$$e_k(\mathbf{u}) = \sum_{A \in \binom{[0,n-1]}{k}} \tilde{u}_A,$$

where for a k-subset $A = \{a_1, a_2, \ldots, a_k\} \subset [0, n-1]$ the element $\tilde{u}_A \in \mathcal{U}_n$ is defined as any expression $u_{a_1} u_{a_2} \cdots u_{a_k}$ where if i and $i+1$ (modulo n) are both in A then u_i must precede u_{i+1} within $\tilde{u}_A$.

Sketch of proof. We verify this using the relation $e_k = h_k - h_{k-1}e_1 + \cdots \pm h_1 e_{k-1}$. Since $k \leq n$ we can restrict our attention to proper subsets of $[0, n-1]$, one at a time. After this, the theorem follows by induction. □

Thus $u_i \mapsto u_{n-1-i}$ is an involution of $\mathcal{U}_n$ sending $h_k(\mathbf{u})$ to $e_k(\mathbf{u})$. More generally, when λ is a hook so that $s_\lambda \in \Lambda_{(n)}$ then $s_\lambda(\mathbf{u})$ can be written as a sum over the reading words of certain tableaux (see [Lam]). We shall not need this generality, however see Conjecture 15.

Theorem 8 follows by writing

$$\Omega^{(n)} = \sum_{\lambda \in \mathcal{P}ar^n} h_\lambda(\mathbf{u}) m_\lambda = \sum_{\lambda \in \mathcal{P}ar^n} e_\lambda(\mathbf{u}) w^+(m_\lambda)$$

and using the fact that $\tilde{F}_w(X) = \left\langle \Omega^{(n)} \cdot 1, w \right\rangle$.

5. Affine Schur functions

We now describe another representation of $\tilde{S}_n$ and $\mathcal{U}_n$. Let $\mathcal{P}$ denote the set of doubly infinite $(0,1)$-sequences $p = (\ldots, p_{-2}, p_{-1}, p_0, p_1, p_2, \ldots)$ and let $\mathbb{C}[\mathcal{P}]$ denote the space of formal $\mathbb{C}$- linear combinations of such sequences. Let $\tilde{S}_n$ act on $\mathcal{P}$ by letting s_i act on $p = (\ldots, p_{-2}, p_{-1}, p_0, p_1, p_2, \ldots)$ by swapping p_{kn+i} and p_{kn+i+1} for each $k \in \mathbb{Z}$. One can check directly that this defines a representation of $\tilde{S}_n$.

A subrepresentation $\mathbb{C}[\mathcal{P}^*]$ of $\mathbb{C}[\mathcal{P}]$ is given by taking only those bit sequences p satisfying $p_N = 1$ for sufficiently small $N << 0$ and $p_N = 0$ for $N >> 0$. These sequences correspond to the *edge sequences* of a partition (see [Sta99, vL]). The edge sequence is obtained by drawing the partition in the English notation and reading the "edge" of the partition from bottom left to top right – writing a 1 if you go up and writing a 0 if you go to the right. Let $p(\lambda)$ denote the edge sequence associated to λ normalised so that $p(\emptyset)_i = 1$ for $i \leq 0$ and $p(\emptyset)_i = 0$ for $i \geq 1$. For example, $p(32) = (\ldots, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, \ldots))$. It is easy to see that $\mathbb{C}[\mathcal{P}^*]$ is indeed a subrepresentation, but it is by no means irreducible. Let $\mathcal{P}^n$ denote the set $\left\{\tilde{S}_n \cdot (\emptyset)\right\}$ of the orbit of the edge sequence of the empty partition. This orbit can be described very naturally when thought of as partitions: it is the set of n-cores [Las]. From now on we will identify $\mathcal{P}^n$ with the set of n-cores. Since the stabiliser of the empty partition is $S_n \subset \tilde{S}_n$, the set $\mathcal{P}^n$ of n-cores is in fact isomorphic to $\tilde{S}_n / S_n$ where here S_n is generated by $s_1, s_2, \ldots, s_{n-1}$.

Now let $\mathcal{U}_n$ act on $\mathbb{C}[\mathcal{P}^n]$ by

$$u_i \cdot \nu = \begin{cases} s_i \cdot \nu & \text{if } s_i \cdot \nu \text{ is obtained from } \nu \text{ by adding boxes.} \\ 0 & \text{otherwise.} \end{cases}$$

One checks ([Las, LM]) that $s_i \cdot \nu$ is always obtained from ν by either adding boxes or removing boxes (never both) when $\nu \in \mathcal{P}^n$. The fact that this defines an action of $\mathcal{U}_n$ is easy to verify. Equip $\mathcal{P}^n$ with the inner product $\langle \nu, \mu \rangle = \delta_{\nu\mu}$.

Definition 10. The *skew affine Schur functions* $\tilde{F}_{\nu/\mu}(X)$ are given by

$$\tilde{F}_{\nu/\mu}(X) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle h_{a_t}(u)h_{a_{t-1}}(u)\cdots h_{a_1}(u)\cdot\mu,\nu\right\rangle x_1^{a_1}x_2^{a_2}\cdots x_t^{a_t}.$$

The *affine Schur functions* are given by $\tilde{F}_\nu(X) = \tilde{F}_{\nu/\emptyset}(X)$.

By Proposition 5, these functions are actually symmetric. One can view affine Schur functions as the generating functions for certain semistandard tableaux built on n-cores. These tableaux are called "k-tableaux" by Lapointe and Morse [LM]. Affine Schur functions had earlier been defined by Lapointe and Morse, and were called *dual k-Schur functions*.

The following proposition is immediate.

Proposition 11. *If $w\cdot\emptyset = \nu \in \mathcal{P}^n$ and w is a minimum length representative in its coset of $\tilde{S}_n/S_n$ then*

$$\tilde{F}_\nu(X) = \tilde{F}_w(X)$$

so that affine Schur functions are special cases of affine Stanley symmetric functions. We write $w = w(\nu)$.

We will call affine permutations of the proposition *Grassmannian*. Note that all the weak Bruhat orders corresponding to $\tilde{S}_n$ modulo a maximal parabolic subgroup are isomorphic so that we lose no generality considering only this maximal parabolic subgroup. This is unlike the non-affine case, where Grassmannian permutations for different maximal parabolics are significantly different.

Theorem 12. *The affine Schur functions form a basis for $\Lambda^{(n)}$.*

We sketch a proof of this theorem. In fact we show that the transition matrix between affine Schur functions and monomial symmetric functions is unitriangular. A more general statement is true for affine Stanley symmetric functions: the monomial (and also affine Schur) expansion contains a unique *dominant* term (see [Sta84] for the non-affine version of this result).

Affine Grassmannian permutations are also naturally indexed by partitions $\lambda \in \mathcal{P}ar^n$. An easy bijection is given by the *code* or *affine inversion table* [BB, Las]. This is a vector $c = (c_0, c_1, \ldots, c_{n-1}) \in \mathbb{N}^n - \mathbb{P}^n$ of non-negative entries with at least one 0. It is shown in [BB] that there is a bijection between codes and affine permutations. The action of the simple generator s_i on the code $c = (c_0, c_1, \ldots, c_{n-1})$ can be described by

$$s_i\cdot(c_0,\ldots,c_{i-1},c_i,\ldots,c_{n-1}) = (c_0,\ldots,c_i+1,c_{i-1},\ldots,c_{n-1})$$

whenever $c_i > c_{i-1}$. To each affine permutation w we let $\mu(w)$ denote the partition conjugate to the decreasing permutation of its code $c(w)$. Grassmannian permutations correspond to weakly increasing codes, so that $w \mapsto \mu(w)$ is a bijection for Grassmannian permutations. From hereon, we will label an affine Schur function with partitions $\lambda \in \mathcal{P}ar^n$, so that $\tilde{F}_\nu(X) = \tilde{F}_\lambda(X)$ if $\nu \in \mathcal{P}^n$ and $\lambda = \mu(w(\nu))$.

We observe that applying a term of $h_k(\mathbf{u})$ to w will increase k different entries of $c(w)$ by 1 (assuming the result is non-zero). For $w \in \tilde{S}_n$, let $\alpha_{w\lambda}$ be given by $\tilde{F}_w(X) = \sum_{\lambda\in\mathcal{P}ar^n}\alpha_{w\lambda}m_\lambda$.

We obtain the following theorem, which implies Theorem 12.

Theorem 13. *Let $w \in \tilde{S}_n$. Then*

- *If $\alpha_{w\lambda} \neq 0$ then $\lambda \preceq \mu(w)$.*
- *We have $\alpha_{w\mu(w)} = 1$.*

We now describe the relationship between affine Schur functions and k-Schur functions (with $k = n-1$). k-Schur functions were originally used to investigate Macdonald polynomial positivity and were defined as symmetric functions with coefficients in $\mathbb{C}(t)$. There are a number of different definitions of k-Schur functions [LLM, LM] which conjecturally agree. The form of the k-Schur functions that we will use are (conjecturally) the $t = 1$ specialisations of the original definition. Suppose $\tilde{F}_\lambda(X) = \sum_\mu K^{(n)}_{\lambda\mu} m_\mu$ where $\lambda \in \mathcal{P}ar^n$ and the sum is over μ satisfying $\mu \in \mathcal{P}ar^n$. Then the k-Schur functions $s^{(k)}_\lambda(X) \in \Lambda_{(n)}$ are given by requiring that

$$h_\mu = \sum_\lambda K^{(n)}_{\lambda\mu} s^{(k)}_\lambda(X).$$

A form of this definition is called the "k-Pieri" rule in [LM]. Affine Schur functions and k-Schur functions are dual in the sense that $\left\langle s^{(k)}_\mu, \tilde{F}_\nu \right\rangle = \delta_{\mu\nu}$. This can be seen by writing the *affine Cauchy kernel*

$$\begin{aligned}\Omega^{(n)} &= \sum_{\mu:\ \mu\in\mathcal{P}ar^n} h_\mu(X) m_\mu(Y) = \sum_{\mu:\ \mu\in\mathcal{P}ar^n} \left(\sum_{\lambda:\ \lambda\in\mathcal{P}ar^n} K^{(n)}_{\lambda\mu} s^{(k)}_\lambda(X) \right) m_\mu(Y) \\ &= \sum_{\lambda:\ \lambda\in\mathcal{P}ar^n} s^{(k)}_\lambda(X) \left(\sum_{\mu:\ \mu\in\mathcal{P}ar^n} K^{(n)}_{\lambda\mu} m_\mu(Y) \right) = \sum_{\lambda:\ \lambda\in\mathcal{P}ar^n} s^{(k)}_\lambda(X) \tilde{F}_\lambda(Y).\end{aligned}$$

6. Relation with cylindric Schur functions

We have seen that affine Schur functions correspond to affine Stanley symmetric functions for Grassmannian permutations.

Cylindric Schur functions [GK] are special cases of skew affine Schur functions. One can define them in the same way as skew affine Schur functions by letting $\mathcal{U}_n$ act on periodic bit sequences $p = (\ldots, p_{-2}, p_{-1}, p_0, p_1, p_2, \ldots)$ satisfying $p_i = p_{i+n}$. It is clear that periodic bit sequences are closed under the action of $\tilde{S}_n$ and in fact form $n + 1$ finite orbits depending on the value of $p_1 + p_2 + \cdots + p_n \in [0, n]$. They can be thought of as edge sequences of partitions lying on a cylinder.

It is more convenient to work with cylindric partitions instead of the periodic edge sequences, and let the cylindric partitions be drawn on the plane (satisfying certain invariance conditions under translation) so that many cylindric partitions may have the same edge sequence but be considered distinct – we may translate a cylindric partition within the plane to obtain a different one. Formally, a cylindric partition λ is an infinite lattice path in $\mathbb{Z}^2$, consisting only of moves upwards and to the right, invariant under the translation by a vector $(k, n-k)$ for some $k \in [0, n]$. We denote the set of cylindric partitions by $\mathcal{P}^c$.

If λ is a cylindric partition then $s_i \cdot \lambda$ is the cylindric partition obtained from λ by either adding boxes at all corners along diagonals congruent to i mod n, or removing boxes, or doing nothing. Define $u_i : \mathbb{C}[\mathcal{P}^c] \to \mathbb{C}[\mathcal{P}^c]$ by

$$u_i \cdot \lambda = \begin{cases} s_i \cdot \lambda & \text{if } s_i \cdot \lambda \text{ is obtained from } \lambda \text{ by adding boxes.} \\ 0 & \text{otherwise.} \end{cases}$$

This defines a representation of $\mathcal{U}_n$ on $\mathbb{C}[\mathcal{P}^c]$, and equipping $\mathbb{C}[\mathcal{P}^c]$ with the natural inner product one can check that for $\lambda, \mu \in \mathcal{P}^c$ the functions $\tilde{F}^c_{\lambda/\mu}$ given by

$$\tilde{F}^c_{\lambda/\mu}(X) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle h_{a_t}(u) h_{a_{t-1}}(u) \cdots h_{a_1}(u) \cdot \mu, \lambda \right\rangle x_1^{a_1} x_2^{a_2} \cdots x_t^{a_t}$$

are the cylindric Schur functions of [McN, Pos]. They are chains $\lambda = \lambda^0 \subset \lambda^1 \subset \cdots \subset \lambda^r \subset \mu$ of cylindric partitions such that each λ^i/λ^{i-1} has at most one box in each column. Note that $\lambda \subset \mu$ means that μ lies to the southeast of λ when considered as lattice paths. If $w \in \tilde{S}_n$ is an affine permutation of minimal length satisfying $w \cdot \mu = \lambda$ then $\tilde{F}^c_{\lambda/\mu}(X) = \tilde{F}_w(X)$. The element w will necessarily be 321-avoiding. Conversely, any affine Stanley symmetric function labelled by a 321-avoiding permutation is equal to a cylindric Schur function.

7. Positivity

We conjecture that affine Schur functions generalise Schur functions for Stanley symmetric function positivity.

Conjecture 14. *The affine Stanley symmetric functions $\tilde{F}_w(X)$ expand positively in terms of the affine Schur functions $\tilde{F}_\lambda(X)$.*

This conjecture seems to be consistent with all the known behaviour of k-Schur functions and cylindric Schur functions. In fact, it has been conjectured that the multiplicative constants for k-Schur functions are non-negative, which by duality would imply that the skew affine Schur functions expand positively in terms of affine Schur functions. Similarly, the conjecture seems to be consistent with Postnikov's result [Pos] that "toric" Schur polynomials (in finitely many variables) expand positively into Schur polynomials. The fact that in infinitely many variables the cylindric Schur functions are nearly never Schur positive can probably be reconciled via affine Schur positivity. See also McNamara's work on cylindric Schur positivity [McN].

Since $s^{(k)}_\lambda(X) \in \Lambda_{(n)}$ we have an element $s^{(k)}_\lambda(\mathbf{u}) \in \mathcal{U}_n$ (as before $k = n-1$). The following conjecture is inspired by the paper of Fomin-Greene [FG]. J. Morse has communicated to the author that a similar conjecture was studied by L. Lapointe and herself.

Conjecture 15. *The "non-commutative" k-Schur function $s^{(k)}_\lambda(\mathbf{u})$ can be written as a non-negative sum of monomials in $u_0, \ldots, u_{n-1}$.*

Proposition 16. *Conjecture 15 implies Conjecture 14.*

Proof. We compute using the affine Cauchy kernel that

$$\tilde{F}_w(X) = \sum_{\lambda \in \mathcal{P}ar^n} \langle h_\lambda(\mathbf{u}) \cdot 1, w \rangle \, m_\lambda(X) = \left\langle \Omega^{(n)} \cdot 1, w \right\rangle = \sum_{\lambda \in \mathcal{P}ar^n} \left\langle s^{(k)}_\lambda(\mathbf{u}) \cdot 1, w \right\rangle \tilde{F}_\lambda(X).$$

Since u_i acts with non-negative coefficients, Conjecture 15 now implies that the coefficients $\left\langle s^{(k)}_\lambda(\mathbf{u}) \cdot 1, w \right\rangle$ are non-negative. □

8. Final comments

8.1. **The affine flag variety, quantum cohomology and fusion ring.** The connections with k-Schur functions and with cylindric Schur functions indicate that our definition of affine Stanley symmetric functions are indeed the correct definitions. Shimozono has conjectured that the multiplication of k-Schur functions calculate the homology multiplication of the affine Grassmannian. Multiplication of k-Schur functions is related to co-multiplication of affine Schur functions which are special cases of affine Stanley symmetric functions (for Grassmannian affine permutations). Thus it seems plausible that affine Stanley symmetric functions in general should be related to the affine flag variety. The direction towards *affine Schubert polynomials* seems to be the most fruitful one to take.

Note that our results show directly that k-Schur functions and cylindric Schur functions are related. This was already known if we combine Postnikov's work on cylindric Schur functions and Gromov-Witten invariants of the Grassmannian with Lapointe and Morse's work on k-Schur functions and the fusion ring (also known as the Verlinde algebra). Finally it is known that the fusion ring agrees with the quantum cohomology of the Grassmannian at $q = 1$. This suggests that there may be an interesting q-analogue of our theory.

8.2. **Affine stable Grothendieck polynomials.** Whereas Schubert polynomials are representatives for the cohomology of the flag variety, *Grothendieck polynomials* are representatives for the K-theory of the flag variety. In the same way that Stanley symmetric functions are stable Schubert polynomials, one can define *stable Grothendieck polynomials*. Our definition of affine Stanley symmetric functions naturally generalises to a definition of affine stable Grothendieck polynomials.

Let $\tilde{\mathcal{U}}_n$ be the algebra obtained from $\mathcal{U}_n$ by replacing the relation $u_i^2 = 0$ with $u_i^2 = u_i$. Define $\tilde{h}_k(\mathbf{u}) \in \tilde{\mathcal{U}}_n$ for $k \in [0, n-1]$ with the same formula as for $h_k(\mathbf{u})$.

Definition 17. Let $w \in \tilde{S}_n$. The affine stable Grothendieck polynomial $\tilde{G}_w(X)$ is

$$\tilde{G}_w(X) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle \tilde{h}_{a_t}(u)\tilde{h}_{a_{t-1}}(u)\cdots\tilde{h}_{a_1}(u)\cdot 1, w \right\rangle x_1^{a_1}x_2^{a_2}\cdots x_t^{a_t},$$

where the sum is over compositions of $l(w)$ satisfying $a_i \in [0, n-1]$.

References

[BB] A. Björner and F. Brenti, Affine permutations of type A, *Electronic Journal of Combinatorics*, **3**(2), 1996, Research paper 18.

[EG] P. Edelman and C. Greene, Balanced tableaux, *Adv. in Math.* **63**1 (1987), 42-99.

[FG] S. Fomin and C. Greene, Noncommutative Schur functions and their applications, *Discrete Mathematics* **193** (1998), 179-200.

[FS] S. Fomin and R. Stanley, Schubert polynomials and the nilCoxeter algebra, *Adv. Math.* **103** (1994), 196-207.

[GK] I. Gessel and C. Krattenthaler, Cylindric Partitions, *Trans. Amer. Math. Soc.* **349** (1997), 429-479.

[LTK] T.K. Lam, B_n Stanley symmetric functions, *Discrete Mathematics*, **157** (1996), 241-270.

[Lam] T. Lam, Ribbon Schur operators, preprint, 2004; math.CO/0409463.

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

[LM] L. Lapointe and J.Morse, Tableaux on k+1-cores, reduced words for affine permutations, and k-Schur expanions, preprint, 2004.

[Las] A. Lascoux, Ordering the affine symmetric group, in Algebraic Combinatorics and Applications (Gossweinstein, 1999), Springer, Berlin (2001), 218-231.

[LS] A. Lascoux and M. Schützenberger, Schubert polynomials and the Littlewood-Richardson rule, *Lett. Math. Phys.* **10**(2-3) (1985) 111-124.

[vL] M. van Leeuwen, Edge sequences, ribbon tableaux, and an action of affine permutations, *Europ. J. Combinatorics*, **20** (1999) 179-195.

[Mac] I. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, 1995.

[McN] P. McNamara, Cylindric Skew Schur Functions, preprint, 2004; math.CO/0410301.

[MS] J. Morse and M. Shimozono, talk at Park City Mathematical Institute, July 2004.

[Pos] A. Postnikov, Affine approach to quantum Schubert calculus, preprint, 2002; math.CO/0205165.

[Sta84] R. Stanley, On the number of reduced decompositions of elements of Coxeter groups, *European J. Combinatorics* **5** (1984), 359-372.

[Sta99] R. Stanley, *Enumerative Combinatorics, Vol 2*, Cambridge, 1999.

E-mail address: thomasl@math.mit.edu

Department of Mathematics, M.I.T., Cambridge, MA 02139

EHRHART POLYNOMIALS OF CYCLIC POLYTOPES

FU LIU

Abstract. The Ehrhart polynomial of an integral convex polytope counts the number of lattice points in dilates of the polytope. In [1], the authors conjectured that for any cyclic polytope with integral parameters, the Ehrhart polynomial of it is equal to its volume plus the Ehrhart polynomial of its lower envelope and proved the case when the dimension $d = 2$. In our article, we prove the conjecture for any dimension.

1. Introduction

For any *integral convex polytope* P, that is, a convex polytope whose vertices have integral coordinates, any positive integer $m \in \mathbb{N}$, we denote by $i(P, m)$ the number of lattice points in mP, where $mP = \{mx \mid x \in P\}$ is *the mth dilate polytope* of P. In our paper, we will focus on a special class of polytopes, cyclic polytopes, which are defined in terms of the moment curve:

Definition 1.1. The **moment curve** in $\mathbb{R}^d$ is defined by

$$\nu_d : \mathbb{R} \to \mathbb{R}^d, t \mapsto \nu_d(t) = \begin{pmatrix} t \\ t^2 \\ \vdots \\ t^d \end{pmatrix}.$$

Let $T = \{t_1, \ldots, t_n\}_<$ be a linearly ordered set. Then the **cyclic polytope** $C_d(T) = C_d(t_1, \ldots, t_n)$ is the convex hull $\operatorname{conv}\{v_d(t_1), v_d(t_2), \ldots, v_d(t_n)\}$ of $n > d$ distinct points $\nu_d(t_i), 1 \leq i \leq n$, on the moment curve.

The main theorem in our article is the one conjecured in [1, Conjecture 1.5]:

Theorem 1.2. *For any integral cyclic polytope* $C_d(T)$,

$$i(C_d(T), m) = \operatorname{Vol}(mC_d(T)) + i(C_{d-1}(T), m).$$

Hence,

$$i(C_d(T), m) = \sum_{k=0}^{d} \operatorname{Vol}_k(mC_k(T)) = \sum_{k=0}^{d} \operatorname{Vol}_k(C_k(T)) m^k,$$

where $\operatorname{Vol}_k(mC_k(T))$ *is the volume of* $mC_k(T)$ *in* k*-dimensional space, and we let* $\operatorname{Vol}_0(mC_0(T)) = 1$.

One direct result of Theorem 1.2 is that $i(C_d(T), m)$ is always a polynomial in m. This result was already shown by Eugène Ehrhart for any integral polytope in 1962 [2]. Thus, we call $i(P, m)$ the Ehrhart polynomial of P when P is an integral polytope. There is much work on the coefficients of Ehrhart polynomials. For instance it's well known that the leading and second coefficients of $i(P, m)$ are the normalized volume of P and one half of the normalized volume of the boundary of P. But there is no known explicit method of describing all the coefficients of Ehrhart polynomials of general integral polytopes. However, because of some special properties that cyclic polytopes have, we are able to calculate the Ehrhart polynomial of cyclic polytopes in the way described in Theorem 1.2.

In this paper, we use a standard triangulation decomposition of cyclic polytopes, and careful counting of lattice points to reduce Theorem 1.2 to the case n=d+1, (Theorem 2.9). We then prove Theorem 2.9 with the use of certain linear transformations and decompositions of polytopes containing our cyclic polytopes.

2. Statements and Proofs

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. Also, We denote by ∂P and $I(P)$ the boundary and the interior of P, respectively.

For simplicity, for any region $R \subset \mathbb{R}^d$, we denote by $\mathcal{L}(R) := R \cap \mathbb{Z}^d$ the set of lattice points in R.

Consider the projection $\pi : \mathbb{R}^d \to \mathbb{R}^{d-1}$ that forgets the last coordinate. In [1, Lemma 5.1], the authors showed that the inverse image under π of a lattice point $y \in C_{d-1}(T) \cap \mathbb{Z}^{d-1}$ is a line that intersects the boundary of $C_d(T)$ at integral points. By a similar argument, it's easy to see that this is true when we replace the cyclic polytopes by their dilated polytopes. Note that $\pi(mC_d(T)) = mC_{d-1}(T)$, so for any lattice point y in $mC_{d-1}(T)$ the inverse image under π intersects the boundary at lattice points.

Definition 2.1. For any x in a real space, let $l(x)$ denote the last coordinate of x.

For any polytope $P \subset \mathbb{R}^d$ and any point $y \in \mathbb{R}^{d-1}$, let $\rho(y, P) = \pi^{-1}(y) \cap P$ be the intersection of P with the inverse image of y under π. Let $p(y, P)$ and $n(y, P)$ be the point in $\rho(y, P)$ with the largest and smallest last coordinate, respectively. If $\rho(y, P)$ is the empty set, i.e., $y \notin \pi(P)$, then let $p(y, P)$ and $n(y, P)$ be empty sets as well. Clearly, $p(y, P)$ and $n(y, P)$ are on the boundary of P. Also, we let $\rho^+(y, P) = \rho(y, P) \setminus n(y, P)$, and for any $S \subset \mathbb{R}^{d-1}$, $\rho^+(S, P) = \cup_{y \in S} \rho^+(y, P)$.

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 *nonnegative part* of P.

For any facet F of P, if F has an interior point in the positive boundary of P, (it's easy to see that $F \subset PB(P)$) then we call F a *positive facet* of P and define the sign of F as $+1 : \mathrm{sign}(F) = +1$. Similarly, we can define the *negative facets* of P with associated sign -1.

By the argument we gave before Definition 2.1, π induces a bijection of lattice points between $NB(mC_d(T))$ and $\pi(mC_d(T)) = mC_{d-1}(T)$. Hence, Theorem 1.2 is equivalent to the following Proposition:

Proposition 2.2. $\mathrm{Vol}(mC_d(T)) = |\mathcal{L}(\Omega(mC_d(T)))|$.

From now on, we will consider any polytopes or sets as *multisets* which allow *negative multiplicities.* We can consider any element of a multiset as a pair (x, m), where m is the multiplicity of element x. (A multiplicity zero for an element x is used when x does not appear at all in the multiset.) Then for any multisets M_1, M_2 and any integers m, n and i, we define the following operators:

(i) Scalar product: $iM_1 = i \cdot M_1 = \{(x, im) \mid (x, m) \in M_1\}$.
(ii) Addition: $M_1 \oplus M_2 = \{(x, m+n) \mid (x, m) \in M_1, (x, n) \in M_2\}$.
(iii) Subtraction: $M_1 \ominus M_2 = M_1 \oplus ((-1) \cdot M_2)$.

It's clear that the following holds:

Lemma 2.3.

a) $\forall R_1, \dots, R_k \subset \mathbb{R}^d$, $\forall i_1, \dots, i_k \in \mathbb{Z} : \mathcal{L}(\bigoplus_{j=1}^k i_j R_j) = \bigoplus_{j=1}^k i_j \mathcal{L}(R_j)$.
b) *For any polytope* $P \subset \mathbb{R}^d$, $\forall R_1, \dots, R_k \subset \mathbb{R}^{d-1}$, $\forall i_1, \dots, i_k \in \mathbb{Z}$:

$$\rho^+\left(\bigoplus_{j=1}^k i_j R_j, P\right) = \bigoplus_{j=1}^k i_j \rho^+(R_j, P).$$

Let P be a convex polytope. For any y an interior point of $\pi(P)$, since π is a continous 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. We hope that $\rho(y, P)$ always has only one point, so we define the following polytopes and discuss several properties of them.

Definition 2.4. We call a convex polytope P a *nice* polytope with respect to π if for any $y \in \partial\pi(P)$, $|\rho(y, P)| = 1$ and for any lattice point $y \in \pi(P)$, $\pi^{-1}(y)$ intersects ∂P at lattice points.

Lemma 2.5. *A nice polytope P has the following properties:*

(i) *For any $y \in I(\pi(P))$, $\pi^{-1}(y) \cap \partial P = \{p(y, P), n(y, P)\}$. In particular, if y is a lattice point, then $p(y, P)$ and $n(y, P)$ are each lattice points.*
(ii) *For any $y \in \partial\pi(P)$, $\pi^{-1}(y) \cap \partial P = \rho(y, P) = p(y, P) = n(y, P)$, so $\rho^+(y, P) = \emptyset$. In particular, when y is a lattice point, $\rho(y, P)$ is a lattice point as well.*
(iii) *$\mathcal{L}$ and ρ^+ commute: for any $R \subset \mathbb{R}^{d-1}$, $\mathcal{L}(\rho^+(R, P)) = \rho^+(\mathcal{L}(R), P)$.*
(iv) *Let R be a region containing $I(\pi(P))$. Then*

$$\Omega(P) = \rho^+(R, P) = \bigoplus_{y \in R} \rho^+(y, P).$$

Moreover,

$$|\mathcal{L}(\Omega(P))| = \sum_{y \in \mathcal{L}(R)} l(p(y, P)) - l(n(y, P)).$$

(By convention, if $y \notin \pi(P)$, we let $l(p(y, P)) - l(n(y, P)) = 0$.)
(v) *If P is decomposed into nice polytopes $P_1, \dots, P_k$, i.e., $P = P_1 \cup \cdots \cup P_k$ and $I(P_i) \cap I(P_j) = \emptyset$ for any distinct i, j, then $\Omega(P) = \bigoplus_{i=1}^k \Omega(P_i)$, so $\mathcal{L}(\Omega(P)) = \bigoplus_{i=1}^k \mathcal{L}(\Omega(P_i))$.*
(vi) *The set of facets of P are partitioned into the set of positive facets and the set of negative facets, i.e., every facet is either positive or negative but not both.*

Proof. The first three and last properties are immediately true. And the fourth one follows directly from the second one. The fifth property can be checked by considering the definition of Ω. $\square$

By using these properties, we are able to give the following proposition about a nice convex polytope:

Proposition 2.6. *Let P be a nice convex polytope with respect to π such $\pi(P)$ is also nice, and all the points in P have nonnegative last coordinate. Suppose further that for any facet F of P, $\pi(F)$ is a nice polytope with respect to π. Then*

$$\Omega(P) = \bigoplus_{F:\text{ a facet of } P} \operatorname{sign}(F)\rho^+(\Omega(\pi(F)), \operatorname{conv}(F, \pi(F))),$$

where $\operatorname{conv}(F, \pi(F))$ *denotes the convex hull of the set* $F \cup \{(y', 0)' \mid y \in \pi(F)\}$*, i.e. the region between F and its projection onto the hyperplane* $\{(x_1, \ldots, x_d)' \mid x_d = 0\}$*. (Note, for any vector v, we use v' to denote its transpose. So for a vertical vector y, $(y', 0)'$ is just the vector obtained from y by attaching a zero to the bottom of y.)*

Proof. A special case of Lemma 2.5/(iv) is when $R = \Omega(\pi(P))$, so we have

$$\Omega(P) = \rho^+(\Omega(\pi(P)), P) = \bigoplus_{y \in \Omega(\pi(P))} \rho^+(y, P).$$

Now for any points a and b, we use $(a, b]$ to denote the half-open line segment between a(excluding) and b(including). Then, $\rho^+(y, P) = (n(y, P), p(y, P)] = (((y', 0)', p(y, P)] \ominus ((y', 0)', n(y, P)])$. Therefore,

$$\begin{aligned}\Omega(P) &= \bigoplus_{y \in \Omega(\pi(P))} (((y', 0)', p(y, P)] \ominus ((y', 0)', n(y, P)]) \\ &= \left(\bigoplus_{y \in \Omega(\pi(P))} ((y', 0)', p(y, P)] \right) \bigoplus \left(\bigoplus_{y \in \Omega(\pi(P))} (-1) \cdot ((y', 0)', n(y, P)] \right).\end{aligned}$$

Let $F_1, F_2, \ldots, F_\ell$ be all the positive facets of P and $F_{\ell+1}, \ldots, F_k$ be all the negative facets. Then it's clear that $\pi(F_1) \cup \pi(F_2) \cup \cdots \cup \pi(F_\ell)$ and $\pi(F_{\ell+1}) \cup \cdots \cup \pi(F_k)$ both give a decomposition of $\pi(P)$. Therefore by Lemma 2.5/(v), we have that $\Omega(\pi(P)) = \bigoplus_{i=1}^{\ell} \Omega(\pi(F_i)) = \bigoplus_{j=\ell+1}^{k} \Omega(\pi(F_j))$. Hence,

$$\begin{aligned}\bigoplus_{y \in \Omega(\pi(P))} ((y', 0)', p(y, P)] &= \bigoplus_{i=1}^{\ell} \bigoplus_{y \in \Omega(\pi(F_i))} ((y', 0)', p(y, P)] \\ &= \bigoplus_{i=1}^{\ell} \rho^+(\Omega(\pi(F_i)), \operatorname{conv}(F_i, \pi(F_i))).\end{aligned}$$

Similarly, we will have

$$\bigoplus_{y \in \Omega(\pi(P))} (-1) \cdot ((y', 0)', n(y, P)] = \bigoplus_{j=\ell+1}^{k} (-1)\rho^+(\Omega(\pi(F_j)), \operatorname{conv}(F_j, \pi(F_j))).$$

Thus, by putting them together, we get

$$\Omega(P) = \bigoplus_{F\,:\text{ a facet of } P} \operatorname{sign}(F)\rho^+(\Omega(\pi(F)), \operatorname{conv}(F, \pi(F))).$$

□

In the last proposition, we used a new notation $\operatorname{conv}(F, \pi(F))$ to denote certain polytopes. For polytopes that can be written in this way, we have the following lemma, whose proof is trivial:

Lemma 2.7. *Let H be a hyperplane in $\mathbb{R}^d$ such that $\pi(H) = \mathbb{R}^{d-1}$. Let $S_1 \subset S_2$ be two convex polytopes inside H and the last coordinates of all of their points are non-negative. Then for any $y \in \pi(S_1)$, $\rho^+(y, \operatorname{conv}(S_1, \pi(S_1))) = \rho^+(y, \operatorname{conv}(S_2, \pi(S_2)))$.*

Having discussed some properties of nice polytopes with respect to π, we come back to the dilated cyclic polytopes which are our main interest and show that they are nice:

Lemma 2.8. *$mC_d(T)$ is a nice polytope with respect to π.*

Proof. We already argued that $mC_d(T)$ satisfies the second condition to be nice. So it's left to check that $|\rho(y, C_d(T))| = 1$ for any $y \in \partial C_{d-1}(T)$.

Let $y = (y_1, y_2, \dots, y_{d-1})'$ and suppose y is on a facet F of $mC_{d-1}(T)$ and without loss of generality, let $m\nu_{d-1}(t_1), m\nu_{d-1}(t_2), \dots, m\nu_{d-1}(t_{d-1})$ be the $d-1$ vertices of F. Then there exist $\lambda_1, \dots, \lambda_{d-1} \in \mathbb{R}_{\geq 0}$ such that $y = \sum_{j=1}^{d-1} \lambda_j m\nu_{d-1}(t_j)$ and $\sum_{j=1}^{d-1} \lambda_j = 1$.

Let $x \in \pi^{-1}(y) \cap mC_d(T)$. There exist $\lambda'_1, \dots, \lambda'_n \in \mathbb{R}_{\geq 0}$ such that $x = \sum_{j=1}^{n} \lambda'_j m\nu_d(t_j)$ and $\sum_{j=1}^{n} \lambda'_j = 1$. Then $y = \pi(x) = \sum_{j=1}^{n} \lambda'_j m\nu_{d-1}(t_j)$. Since y is on the facet F, $\lambda'_j = 0$ unless $1 \leq j \leq d-1$. Thus $y = \sum_{j=1}^{d-1} \lambda'_j m\nu_{d-1}(t_j)$ and $\sum_{j=1}^{d-1} \lambda'_j = 1$. Therefore $\lambda_j = \lambda'_j, 1 \leq j \leq d-1$. Hence $x = \sum_{j=1}^{d-1} \lambda_j m\nu_d(t_j)$ is the only point in $\pi^{-1}(y) \cap mC_d(T)$. □

We know that for any cyclic polytope $C_d(T)$ with $n = |T| > d+1$, we can decompose it into $n-d$ cyclic polytopes $P_1 \cup \dots \cup P_{n-d}$, which is a triangulation of $C_d(T)$ and where P_i's are all defined by $(d+1)$-element integer sets. E.g., the pulling triangulation of [4] has this property. Therefore by the fourth property in Lemma 2.5, we have that $\mathcal{L}(\Omega(C_d(T))) = \bigcup_{i=1}^{n-d} \mathcal{L}(\Omega(P_i))$. Thus $|\mathcal{L}(\Omega(P))| = \sum_{i=1}^{k} |\mathcal{L}(\Omega(P_i))|$. Note that we also have $\operatorname{Vol}(C_d(T)) = \sum_{i=1}^{n-d} \operatorname{Vol}(P_i)$. We conclude that to prove Proposition 2.2, it is enough to prove the following:

Theorem 2.9. *For any integer sets T with $n = |T| = d+1$, $\operatorname{Vol}(mC_d(T)) = |\mathcal{L}(\Omega(mC_d(T)))|$.*

Definition 2.10. A map $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ is *structure perserving* if it preserves volume and it commutes with the following operations:

(i) $\mathcal{L}$: taking lattice points of a region $R \subset \mathbb{R}^d$;
(ii) conv : taking the convex hull of a collection of points;
(iii) Ω : taking the nonnegative part of a convex polytope;
(iv) PB : taking the positive boundary of a convex polytope;
(v) NB : taking the negative boundary of a convex polytope.

Remark 2.11. Here φ commuting with conv implies (or is equivalent to) that for any set of points $x_1, \dots, x_k \in \mathbb{R}^d$, and for any $\lambda_1, \dots, \lambda_k \in \mathbb{R}^{\geq 0}$ with $\sum_{i=1}^{k} \lambda_i = 1$, $T(\sum_{i=1}^{k} \lambda_i x_i) = \sum_{i=1}^{k} \lambda_i T(x_i)$. Therefore φ is an affine transformation, which can be defined by a $d \times d$ matrix A and a vector $u \in \mathbb{R}^d : T(x) = Ax + u$. Moreover, φ commuting with PB and NB implies that φ preserves the positive facets and negative facets of a convex polytope.

Lemma 2.12. *Let A be a $d \times d$ integral lower triangular matrix with 1's on its diagonal, and u be an integral vector in $\mathbb{R}^d$. Then $\varphi : x \mapsto Ax + u$ gives a map which is structure preserving, and so does φ^{-1}. Therefore, φ is a bijection from $\mathbb{Z}^d$ to $\mathbb{Z}^d$. Hence, for any subset $S \in \mathbb{R}^d$, $|\mathcal{L}(S)| = |\mathcal{L}(\varphi(S))|$.*

Moreover, for any $y \in \mathbb{R}^{d-1}$, if we define $\tilde{\varphi}(y) = \tilde{A}y + \tilde{u}$, where $\tilde{A}$ is the right upper $(d-1)\times(d-1)$ matrix of A and $\tilde{u} = \pi(u)$, then $\rho^+(\tilde{\varphi}(y), \varphi(P)) = \varphi(\rho^+(y, P))$, for any polytope P.

Proof. The determinant of A is 1, hence φ is volume preserving. It's easy to check that φ commutes with $\mathcal{L}$ and conv. To show that φ commutes with Ω, PB and NB, it suffices to show that for all $x_1, x_2 \in \mathbb{R}^d$ with $\pi(x_1) = \pi(x_2)$ and $l(x_1) > l(x_2)$, then $\pi(\varphi(x_1)) = \pi(\varphi(x_2))$ and $l(\varphi(x_1)) > l(\varphi(x_2))$. This is not hard to check using the fact that A is a lower triangular matrix with 1's on its diagonal. Hence, φ is structure preserving.

Note that φ^{-1} maps x to $A^{-1}x - A^{-1}u$. But we know that A^{-1} is also an integral lower triangular matrix with 1's on its diagonal and $-A^{-1}u$ is an integral vector. So φ^{-1} is structure preserving as well.

It's clear that $\tilde{\varphi} = \pi \circ \varphi \circ \pi^{-1}$, which implies that $\pi^{-1} \circ \tilde{\varphi} = \varphi \circ \pi^{-1}$. So we have

$$x \in \varphi(\rho^+(y, P)) \Leftrightarrow \varphi^{-1}(x) \in \rho^+(y, P) = \pi^{-1}(y) \cap P$$
$$\Leftrightarrow x \in \varphi(\pi^{-1}(y)) \cap \varphi(P) = \pi^{-1}(\tilde{\varphi}(y)) \cap \varphi(P) \Leftrightarrow x \in \rho^+(\tilde{\varphi}(y), \varphi(P)).$$

□

Now for any real numbers $r_1, r_2, \dots, r_d$, we consider the $d \times d$ lower triangular matrices

$$A_{r_1,\dots,r_d}(i,j) = \begin{cases} (-1)^{i-j} e_{i-j}(r_1, \dots, r_i), & i \geq j \\ 0, & i < j \end{cases}$$

and

$$B_{r_1,\dots,r_d}(i,j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \& i < d \\ (-1)^{i-j} e_{i-j}(r_1, \dots, r_i), & j \neq i = d \end{cases}$$

where $e_k(r_1, \dots, r_l) = \sum_{i_1 < i_2 < \cdots < i_k} r_{i_1} r_{i_2} \dots r_{i_k}$ is the kth elementary symmetric function in $r_1, \dots, r_l$.

For simplicity, we allow a map originally defined on $\mathbb{R}^d$ to work in higher dimension, by applying the map to the first d coordinates. Then it's not hard to see that $A_{r_1,\dots,r_d} = A_{r_1,\dots,r_{d-1}} B_{r_1,\dots,r_d} = B_{r_1,\dots,r_d} A_{r_1,\dots,r_{d-1}}$.

We also define vectors

$$u_{r_1,\dots,r_d} = \begin{pmatrix} -r_1 \\ r_1 r_2 \\ -r_1 r_2 r_3 \\ \vdots \\ (-1)^d r_1 r_2 \dots r_d \end{pmatrix} = \begin{pmatrix} -e_1(r_1) \\ e_2(r_1, r_2) \\ -e_3(r_1, r_2, r_3) \\ \vdots \\ (-1)^d e_d(r_1, r_2, \dots, r_d) \end{pmatrix},$$

and

$$v_{r_1,\dots,r_d} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ (-1)^d r_1 r_2 \dots r_d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ (-1)^d e_d(r_1, r_2, \dots, r_d) \end{pmatrix}.$$

Similarly, we allow the addition operation between two vectors of different dimensions by adding the lower dimension one to the first corresponding coordinates of the higher one. Thus, $u_{r_1,\ldots,r_d} = u_{r_1,\ldots,r_{d-1}} + v_{r_1,\ldots,r_d}$.

Now we define maps $\varphi_{r_1,\ldots,r_d} : x \mapsto A_{r_1,\ldots,r_d}x + u_{r_1,\ldots,r_d}$ and $\phi_{r_1,\ldots,r_d} : x \mapsto B_{r_1,\ldots,r_d}x + v_{r_1,\ldots,r_d}$. Unlike $\varphi_{r_1,\ldots,r_d}$, $\phi_{r_1,\ldots,r_d}$ does not depend on the order of r_i's. In other words, for any permutation $\sigma \in S_d$, $\phi_{r_1,\ldots,r_d} = \phi_{r_{\sigma(1)},\ldots,r_{\sigma(d)}}$.

Note that $\phi_{r_1,\ldots,r_d}$ only changes the dth coordinate of a vector, so we have the following lemma:

Lemma 2.13. $\varphi_{r_1,\ldots,r_d} = \varphi_{r_1,\ldots,r_{d-1}} \circ \phi_{r_1,\ldots,r_d}$.

Remark 2.14. When we consider $\varphi_{r_1,\ldots,r_d}$ and $\phi_{r_1,\ldots,r_d}$ operating on the moment curve, we have

$$\varphi_{r_1,\ldots,r_d}(\nu_d(t)) = A_{r_1,\ldots,r_d}\begin{pmatrix} t \\ t^2 \\ \vdots \\ t^d \end{pmatrix} + u_{r_1,\ldots,r_d} = \begin{pmatrix} (t-r_1) \\ (t-r_1)(t-r_2) \\ \vdots \\ (t-r_1)(t-r_2)\cdots(t-r_d) \end{pmatrix},$$

$$\phi_{r_1,\ldots,r_d}(\nu_d(t)) = B_{r_1,\ldots,r_d}\begin{pmatrix} t \\ t^2 \\ \vdots \\ t^d \end{pmatrix} + v_{r_1,\ldots,r_d} = \begin{pmatrix} t \\ t^2 \\ \vdots \\ t^{d-1} \\ (t-r_1)(t-r_2)\cdots(t-r_d) \end{pmatrix}.$$

Remark 2.15. When $r_1,\ldots,r_d$ are integers, $\varphi_{r_1,\ldots,r_d}, \phi_{r_1,\ldots,r_d}$ and their inverse maps are structure preserving by Lemma 2.12.

Now by using ϕ's (or φ's), we are able to determine the sign of the facets of dilated cyclic polytopes:

Proposition 2.16. *Let $P = mC_d(T)$, where $m \in \mathbb{N}$ and $T = \{t_1, t_2, \ldots, t_n\}_<$ an integral ordered set. Let F be a facet of P determined by vertices $\nu_d(t_{i_1}), \nu_d(t_{i_2}), \ldots, \nu_d(t_{i_d})$. Let k be the smallest element of the set $\{1,2,\ldots,n\} \setminus \{i_1,\ldots,i_d\}$, then* $\mathrm{sign}(F) = (-1)^{d-k}$. *In particular, when $|T| = n = d+1$, let F_k be the facet of P determined by all the vertices of P except $\nu_d(t_{i_k})$, then for $k \in [d]$,* $\mathrm{sign}(F_k) = \mathrm{sign}(\sigma_k)$, *where $\sigma_k = (k, k+1, \cdots, d) \in S_d$ and* $\mathrm{sign}(F_{d+1}) = -1$.

Proof. We first consider the case when $m = 1$, i.e. P is a cyclic polytope. Without loss of generality, we assume that $i_1 < i_2 < \cdots < i_d$. Consider the polytope $Q = \phi_{t_{i_1},\ldots,t_{i_d}}(P)$. For $j = 1,2,\ldots,n$, the last coordinate of the vertex of Q which mapped from $\nu_d(t_j)$ is $l(\phi_{t_{i_1},\ldots,t_{i_d}}(\nu_d(t_j))) = (t_j - t_{i_1})(t_j - t_{i_2})\cdots(t_j - t_{i_d})$. Hence the last coordinates of the vertices of $\phi_{t_{i_1},\ldots,t_{i_d}}(F)$ are all 0's. So $\phi_{t_{i_1},\ldots,t_{i_d}}(F)$ is on the hyperplane obtained by setting the last coordinate to 0. Since k is the smallest element not in $\{i_1,\ldots,i_d\}$, $i_1 = 1, i_2 = 2, \ldots, i_{k-1} = k-1, i_k > k$. So $t_k - t_{i_l} > 0$ when $l = 1,2,\ldots,k-1$; and $t_k - t_{i_l} < 0$ when $l = k, k+1, \ldots, d$. Therefore $\mathrm{sign}(l(\phi_{t_{i_1},\ldots,t_{i_d}}(\nu_d(t_k))) = (-1)^{d-k+1}$. By using Gale's evenness condition [3], it's not hard to see that $\mathrm{sign}(l(\phi_{t_{i_1},\ldots,t_{i_d}}(\nu_d(t_l))) = (-1)^{d-k+1}$, for all $l \notin \{i_1,\ldots,i_d\}$. Thus we can conclude that $l(\phi_{t_{i_1},\ldots,t_{i_d}}(P))$ is nonnegative if $d-k$ is odd, and is nonpositive if $d-k$ is even. Hence $\phi_{t_{i_1},\ldots,t_{i_d}}(F)$ and F are negative facets if $d-k$ is odd, and positive facets if $d-k$ is even. So $\mathrm{sign}(F) = (-1)^{d-k}$. For $n = d+1$, it's easy to see that $\mathrm{sign}(\sigma_k) = (-1)^{d-k} = \mathrm{sign}(F_k)$.

For $m > 1$, we just need to consider the map $x \mapsto B_{t_{i_1},\dots,t_{i_d}} x + m v_{t_{i_1},\dots,t_{i_d}}$ instead of $\phi_{t_{i_1},\dots,t_{i_d}}$, and then we will have similar results.

□

Lemma 2.17. *For all $d \in \mathbb{R}^+$, for all $s_1, \dots, s_d \in \mathbb{N}$, let $x_0 = 1$ and $P_{s_1,\dots,s_d} = \{(x_1, \dots, x_d) \in \mathbb{R}^d \mid \forall i \in [d] : 0 \le x_i \le s_i x_{i-1}\}$, $R_{s_1,\dots,s_d} = \Omega(P_{s_1,\dots,s_d})$. Then $R_{s_1,\dots,s_d} = P_{s_1,\dots,s_d} \cap \{x_d > 0\}$ and for all $d \ge 2 : R_{s_1,\dots,s_d} = \rho^+(R_{s_1,\dots,s_{d-1}}, P_{s_1,\dots,s_d})$.*

Moreover, the vertices of $P_{s_1,\dots,s_d}$ are

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} s_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} s_1 \\ s_1 s_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} s_1 \\ s_1 s_2 \\ s_1 s_2 s_3 \\ \vdots \\ s_1 s_2 \cdots s_d \end{pmatrix}$$

and the positive boundary of $P_{s_1,\dots,s_d}$ is just the convex hull of the first $d-1$ vertices and the last one. Note the first $d-1$ vertices span a $(d-2)$-dimensional space $\{(x_1, \dots, x_d)' \mid x_d = x_{d-1} = 0\}$. Hence $PB(P_{s_1,\dots,s_d})$ is in the hyperplane spanned by this $(d-2)$-dimensional space and the last vertex.

Proof. The first result is immediate by considering the definition of Ω.

We have $R_{s_1,\dots,s_{d-1}} \subset P_{s_1,\dots,s_{d-1}}$, so

$$\begin{aligned}\rho^+(R_{s_1,\dots,s_{d-1}}, P_{s_1,\dots,s_d}) \subset \rho^+(P_{s_1,\dots,s_{d-1}}, P_{s_1,\dots,s_d}) &= \rho^+(\pi(P_{s_1,\dots,s_d}), P_{s_1,\dots,s_d}) \\ &= \Omega(P_{s_1,\dots,s_d}) = R_{s_1,\dots,s_d}.\end{aligned}$$

But for $x = (x_1, \dots, x_d) \in R_{s_1,\dots,s_d}$, we have that $x_d > 0$ which implies that $s_d x_{d-1} > 0$, so $x_{d-1} > 0$. Therefore $\pi(x) \in R_{s_1,\dots,s_{d-1}}$. Thus, $x \in \rho^+(R_{s_1,\dots,s_{d-1}}, P_{s_1,\dots,s_d})$. Now we can conclude that $R_{s_1,\dots,s_d} = \rho^+(R_{s_1,\dots,s_{d-1}}, P_{s_1,\dots,s_d})$.

□

Theorem 2.18. *Let $d \in \mathbb{N}$ and $T = \{t_1, t_2, \dots, t_{d+1}\}_<$ be an integral ordered set, then*

$$\Omega(C_d(T)) = \bigoplus_{\sigma \in S_d} \operatorname{sign}(\sigma) \varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d)}} (R_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d)}}).$$

Proof. We proceed by induction on d. When $d = 1$, $C_d(T)$ is just the interval $[t_1, t_2]$. Then the only element $\sigma \in S_1$ is the identity map. $R_{t_2 - t_1} = (0, t_2 - t_1]$. And $\varphi_{t_1} : x \mapsto x - t_1$, so $\varphi^{-1}_{t_1} : x \mapsto x + t_1$. Thus $\varphi^{-1}_{t_1}((0, t_2 - t_1]) = (t_1, t_2] = \Omega([t_1, t_2])$.

Now we assume the theorem is true for dimensions less than d, and we will prove the case of dimension $d (\ge 2)$. Let $P = \phi_{t_1,\dots,t_d}(C_d(T))$, and let $v_i = \phi_{t_1,\dots,t_d}(\nu_d(t_i))$, $i \in [d+1]$, be the vertices of P. Then for $i \in [d]$, $v_i = \begin{pmatrix} \nu_{d-1}(t_i) \\ 0 \end{pmatrix}$ and for $i = d+1$, $v_{d+1} = \begin{pmatrix} \nu_{d-1}(t_{d+1}) \\ \prod_{i=1}^d (t_{d+1} - t_i)) \end{pmatrix}$. Since $\prod_{i=1}^d (t_{d+1} - t_i)) > 0$, the last coordinates of all the points in P are nonnegative. By Proposition 2.6, we have that

$$\Omega(P) = \bigoplus_{F\text{: a facet of } P} \operatorname{sign}(F) \rho^+(\Omega(\pi(F)), \operatorname{conv}(F, \pi(F))).$$

As in Proposition 2.16, we let F_k be the facet of $C_d(T)$ determined by all the vertices of $C_d(T)$ except $\nu_d(t_{i_k})$, then

$$\Omega(P) = \bigoplus_{k\in[d+1]} \operatorname{sign}(\phi_{t_1,\dots,t_d}(F_k))\rho^+(\Omega(\pi(\phi_{t_1,\dots,t_d}(F_k))), \operatorname{conv}(\phi_{t_1,\dots,t_d}(F_k), \pi(\phi_{t_1,\dots,t_d}(F_k)))).$$

For $k = d+1$, $\tilde{F} = \phi_{t_1,\dots,t_d}(F_{d+1}) = \operatorname{conv}(\{v_i\}_{i=1}^d)$ is on the hyperplane $H_0 = \{(x_1,\dots,x_d)' \in \mathbb{R}^d \mid x_d = 0\}$. So $\operatorname{conv}(\tilde{F}, \pi(\tilde{F}))$ is just $\tilde{F}$. Thus $\rho^+(\Omega(\pi(\tilde{F})), \operatorname{conv}(\tilde{F}, \pi(\tilde{F})))$ is an empty set.

And for $k \in [d]$, by Proposition 2.16, $\operatorname{sign}(F_k) = \operatorname{sign}(\sigma_k)$, where $\sigma_k = (k, k+1, \cdots, d) \in S_d$. Let $T_k = T \setminus \{t_k\}$, then $\pi(\phi_{t_1,\dots,t_d}(F_k)) = \pi(F_k) = C_{d-1}(T_k)$, because $\phi_{t_1,\dots,t_d}$ just changes the last coordinates. It's easy to see that

$$\operatorname{conv}(\phi_{t_1,\dots,t_d}(F_k), \pi(\phi_{t_1,\dots,t_d}(F_k))) = \operatorname{conv}(\{v_i\}_{i\neq k} \cup \{v'_{d+1}\}),$$

where $v'_{d+1} = \begin{pmatrix} \nu_{d-1}(t_{d+1}) \\ 0 \end{pmatrix}$ is the projection of v_{d+1} to the hyperplane H_0.

Hence,

$$\Omega(P) = \bigoplus_{k\in[d]} \operatorname{sign}(\sigma_k)\rho^+(\Omega(C_{d-1}(T_k)), \operatorname{conv}(\{v_i\}_{i\neq k} \cup \{v'_{d+1}\})).$$

For any $k \in [d]$, $T_k = \{t_{\sigma_k(1)}, t_{\sigma_k(2)}, \dots, t_{\sigma_k(d-1)}, t_{d+1}\}_<$. By the induction hypothesis, we have that

$$\Omega(C_{d-1}(T_k)) = \bigoplus_{\tau\in S_{d-1}} \operatorname{sign}(\tau)\varphi^{-1}_{t_{\sigma_k(\tau(1))},\dots,t_{\sigma_k(\tau(d-1))}}(R_{t_{d+1}-t_{\sigma_k(\tau(1))},\dots,t_{d+1}-t_{\sigma_k(\tau(d-1))}}).$$

So,

$$\begin{aligned}
&\operatorname{sign}(\sigma_k)\phi^{-1}_{t_1,\dots,t_d}(\Omega(C_{d-1}(T_k))) \\
= &\bigoplus_{\tau\in S_{d-1}} \operatorname{sign}(\sigma_k)\operatorname{sign}(\tau)\varphi^{-1}_{t_{\sigma_k(\tau(1))},\dots,t_{\sigma_k(\tau(d-1))}}(R_{t_{d+1}-t_{\sigma_k(\tau(1))},\dots,t_{d+1}-t_{\sigma_k(\tau(d-1))}}) \\
= &\bigoplus_{\sigma\in S_d:\sigma(d)=k} \operatorname{sign}(\sigma)\varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(R_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)}}). \quad (\text{let } \sigma = \sigma_k\tau)
\end{aligned}$$

Let H_k be the hyperplane determined by $\phi_{t_1,\dots,t_d}(F_k)$, and $H_k^+ = \{x \in H_k \mid l(x) \geq 0\}$. We claim that for all $\sigma \in S_d$ with $\sigma(d) = k$, we have

$$\varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(PB(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}})) \subset H_k^+.$$

Given this, we can pick a convex polytope $S_k \subset H_k$, such that

a) The last coordinates of the points in S_k are nonnegative;
b) S_k contains $\varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(PB(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}}))$, for all $\sigma \in S_d$ with $\sigma(d) = k$;
c) S_k contains $\phi_{t_1,\dots,t_d}(F_k)$.

Note that $\pi(H_k)$ contains $\pi(\phi_{t_1,\dots,t_d}(F_k)) = \pi(F_k) = C_{d-1}(T_k)$, which has dimension $d-1$. So $\pi(H_k) = \mathbb{R}^{d-1}$.

Hence, by Lemma 2.7

$$
\begin{aligned}
& \Omega(C_d(T)) \\
= \; & \phi^{-1}_{t_1,\dots,t_d}(\Omega(P)) \\
= \; & \bigoplus_{k\in[d]} \operatorname{sign}(\sigma_k)\phi^{-1}_{t_1,\dots,t_d}(\rho^+(\Omega(C_{d-1}(T_k)), \operatorname{conv}(\{v_i\}_{i\neq k}\cup\{v'_{d+1}\}))) \\
= \; & \bigoplus_{k\in[d]} \operatorname{sign}(\sigma_k)\phi^{-1}_{t_1,\dots,t_d}(\rho^+(\Omega(C_{d-1}(T_k)), \operatorname{conv}(S_k,\pi(S_k)))) \\
= \; & \bigoplus_{k\in[d]} \operatorname{sign}(\sigma_k)\phi^{-1}_{t_1,\dots,t_d}(\rho^+(\bigoplus_{\tau\in S_{d-1}} \operatorname{sign}(\tau)\varphi^{-1}_{t_{\sigma_k(\tau(1))},\dots,t_{\sigma_k(\tau(d-1))}}(R_{t_{d+1}-t_{\sigma_k(\tau(1))},\dots,t_{d+1}-t_{\sigma_k(\tau(d-1))}}), \\
& \qquad \operatorname{conv}(S_k,\pi(S_k)))) \\
= \; & \bigoplus_{k\in[d]} \bigoplus_{\sigma\in S_d,\sigma(d)=k} \operatorname{sign}(\sigma)\phi^{-1}_{t_1,\dots,t_d}(\rho^+(\varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(R_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)}}), \\
& \qquad \varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}}))) \\
= \; & \bigoplus_{\sigma\in S_d} \operatorname{sign}(\sigma)\phi^{-1}_{t_1,\dots,t_d}\varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(\rho^+(R_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)}}, \\
& \qquad P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}})) \\
= \; & \bigoplus_{\sigma\in S_d} \operatorname{sign}(\sigma)\varphi^{-1}_{t_{\sigma(1)},\dots,t_{\sigma(d)}}(R_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}}).
\end{aligned}
$$

Thus the claim implies the theorem.

Showing the claim is equivalent to showing that

$$PB(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}}) \subset \varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(H_k^+).$$

Both $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}$ and its inverse only work on the first $d-1$ coordinates of any point in $\mathbb{R}^d$. Thus $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(H_k^+)$ is just $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(H_k)\cap\{x\in\mathbb{R}^d \mid l(x)\geq 0\}$. But it's clear that $PB(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}})$ is in $\{x\in\mathbb{R}^d \mid l(x)\geq 0\}$. So it's enough to show that

$$PB(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}}) \subset \varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(H_k).$$

By Lemma 2.17, $PB(P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}})$ lies in the hyperplane H which is spanned by $\{(x_1,\dots,x_d)' \mid x_d = x_{d-1} = 0\}$ and $\begin{pmatrix} t_{d+1}-t_{\sigma(1)} \\ (t_{d+1}-t_{\sigma(1)})(t_{d+1}-t_{\sigma(2)}) \\ (t_{d+1}-t_{\sigma(1)})(t_{d+1}-t_{\sigma(2)})(t_{d+1}-t_{\sigma(3)}) \\ \vdots \\ (t_{d+1}-t_{\sigma(1)})(t_{d+1}-t_{\sigma(2)})\cdots(t_{d+1}-t_{\sigma(d)}) \end{pmatrix}$.

So we need show that $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(H_k) = H$. Since H_k is the hyperplane containing $\phi_{t_1,\dots,t_d}(F_k)$, it's enough to show that $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d-1)}}(\phi_{t_1,\dots,t_d}(F_k)) = \varphi_{t_{\sigma(1)},\dots,t_{\sigma(d)}}(F_k)$ is contained in H. However, $F_k = \operatorname{conv}(\nu_d(T_k))$. Meanwhile, by remark 2.14, we have

$$\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d)}}(\nu_d(t)) = \begin{pmatrix} (t-t_{\sigma(1)}) \\ (t-t_{\sigma(1)})(t-t_{\sigma(2)}) \\ \vdots \\ (t-t_{\sigma(1)})(t-t_{\sigma(2)})\cdots(t-t_{\sigma(d)}) \end{pmatrix}.$$

Since $\sigma(d) = k$, for any $i \in [d], i \neq k$, $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d)}}(\nu_d(t_i))$ has the last two coordinates equal to 0. And for $i = d+1$, $\varphi_{t_{\sigma(1)},\dots,t_{\sigma(d)}}(\nu_d(t_{d+1}))$ is exactly the last vertex of $P_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d-1)},t_{d+1}-t_{\sigma(d)}}$, which completes the proof the claim and hence the theorem.

$\square$

Remark 2.19. If we define $\varphi_{m,r_1,\dots,r_d} : x \mapsto A_{r_1,\dots,r_d}x + mu_{r_1,\dots,r_d}$, then similarly we can prove that

$$\Omega(mC_d(T)) = \bigoplus_{\sigma\in S_d} \operatorname{sign}(\sigma)\varphi^{-1}_{m,t_{\sigma(1)},\dots,t_{\sigma(d)}}(mR_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d)}}).$$

Corollary 2.20.

$$\mathcal{L}(\Omega(mC_d(T))) = \bigoplus_{\sigma\in S_d} \operatorname{sign}(\sigma)\mathcal{L}(\varphi^{-1}_{m,t_{\sigma(1)},\dots,t_{\sigma(d)}}(mR_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d)}})).$$

Hence,

$$|\mathcal{L}(\Omega(mC_d(T)))| = \sum_{\sigma\in S_d} \operatorname{sign}(\sigma)|\mathcal{L}(mR_{t_{d+1}-t_{\sigma(1)},\dots,t_{d+1}-t_{\sigma(d)}})|.$$

It's easy to see that $mR_{s_1,\dots,s_d} = R_{ms_1,s_2,\dots,s_d}$. Moreover,

$$|L(R_{s_1,\dots,s_d})| = \sum_{x_1=1}^{s_1}\sum_{x_2=1}^{s_2x_1}\cdots\sum_{x_n=1}^{s_nx_{n-1}} 1.$$

Therefore, it's natural to look at the following:

Lemma 2.21. *For any nonnegative integers $a_1, a_2, \dots, a_n$, let*

$$h(a_1, a_2, \dots, a_n) = \sum_{x_1=1}^{a_1}\sum_{x_2=1}^{a_2x_1}\cdots\sum_{x_n=1}^{a_nx_{n-1}} 1.$$

Then the only highest degree term of h is $\frac{1}{n!}a_1^n a_2^{n-1} a_3^{n-2} \dots a_n$. This is also true when we consider h as a polynomial just in the variable a_1.

Proof of Lemma 2.21: We will prove it by induction on n.

When $n = 1$, $h(a_1) = \sum_{x_1=1}^{a_1} 1 = a_1$. Thus the lemma holds.

Assume the lemma is true for n, and note that $h(a_1, a_2, \dots, a_{n+1}) = \sum_{x_1=1}^{a_1} h(a_2x_1, a_3, \dots, a_{n+1})$. By assumption, $\frac{1}{n!}a_2^n a_3^{n-1} \dots a_{n+1}x_1^n$ is the only highest degree term of $h(a_2x_1, a_3, \dots, a_{n+1})$ when we consider it as polynomial both in $y = a_2x_1, a_3, \dots, a_{n+1}$ and in y. This implies that $\frac{1}{n!}a_2^n a_3^{n-1} \dots a_{n+1}x_1^n$ is the only highest degree term of $h(a_2x_1, a_3, \dots, a_{n+1})$ when we consider it both in $a_2, a_3, \dots, a_{n+1}$ and in x_1. Then our lemma immediately follows from the fact that the highest degree term of $\sum_{x_1=1}^{a_1} x_1^n$ is $\frac{1}{n+1}a_1^{n+1}$.

$\square$

Proposition 2.22. *For any nonnegative integers $a_1, a_2, \dots, a_n$, let $\mathcal{H}_m(a_1, a_2, \dots, a_n) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)h(ma_{\sigma(1)}, a_{\sigma(2)} \dots, a_{\sigma(n)})$. Then*

$$\mathcal{H}_m(a_1, a_2, \dots, a_n) = \frac{m^n}{n!}\prod_{i=1}^n a_i \prod_{1\le i<j\le n}(a_i - a_j).$$

Proof of Proposition 2.22:

Clearly if any of a_i's is 0, then $\mathcal{H}_m(a_1,\ldots,a_n)=0$. Also for $1\le i<j\le n$, $\mathcal{H}_m$ changes sign when we switch a_i and a_j, i.e.,

$$\mathcal{H}_m(\ldots,a_i,\ldots,a_j,\ldots)=-\mathcal{H}_m(\ldots,a_j,\ldots,a_i,\ldots).$$

Therefore, $\mathcal{H}_m(a_1,\ldots,a_n)$ must be a multiple of

$$\prod_{i=1}^{n} a_i \prod_{1\le i<j\le n}(a_i-a_j),$$

which has degree $\frac{1}{2}n(n+1)$.

So now it's enough to show that $\mathcal{H}_m(a_1,\ldots,a_n)$ is of degree $\frac{1}{2}n(n+1)$ and the coefficient of $a_1^n a_2^{n-1} a_3^{n-2}\ldots a_n$ in $\mathcal{H}_m(a_1,\ldots,a_n)$ is $\frac{m^n}{n!}$, which follows from Lemma 2.21. □

Proof of Theorem 2.9: By Corollary 2.20,

$$\begin{aligned}
|\mathcal{L}(\Omega(mC_d(T)))| &= \sum_{\sigma\in S_d}\operatorname{sign}(\sigma)|\mathcal{L}(mR_{t_{d+1}-t_{\sigma(1)},\ldots,t_{d+1}-t_{\sigma(d)}})| \\
&= \mathcal{H}_m(t_{d+1}-t_{\sigma(1)},t_{d+1}-t_{\sigma(2)},\ldots,t_{d+1}-t_{\sigma(d)}) \\
&= \frac{m^d}{d!}\prod_{i=1}^{d}(t_{d+1}-t_i)\prod_{1\le i<j\le d}(t_i-t_j) \\
&= \frac{m^d}{d!}\prod_{1\le i<j\le d+1}(t_i-t_j)=\operatorname{Vol}(mC_d(T)).
\end{aligned}$$

□

As we argued earlier in our paper, the proof of Theorem 2.9 completes the proof of Proposition 2.2 and thus proof of our main Theorem 1.2.

Acknowledgements. I would like to thank Richard Stanley for showing me the conjecture in [1].

References

[1] M. Beck, J.A. De Loera, M. Develin, J. Pfeifle, and R. P. Stanley, *Coefficients and roots of Ehrhart polynomials*, arXiv:math.CO/0402148, preprint.

[2] E. Ehrhart, *Sur les polyèdres rationnels homothétiques à n dimensions*, C. R. Acad. Sci. Paris **254** (1962), 616–618.

[3] D. Gale, *Neighborly and cyclic polytopes*, in: "Convexity" (V. Klee, ed.), Proc. Symposia in Pure Mathematics, Vol. VII Amer. Math. Soc., Providence RI (1963), 225–232.

[4] C. W. Lee, *Subdivisions and triangulations of polytopes*, in: "Handbook of Discrete and Computational Geometry" (J.E. Goodman & J. O'Rourke, eds.), CRC Press, Boca Raton (1997), 271–290.

SQUARE q,t-LATTICE PATHS AND $\nabla(p_n)$

NICHOLAS A. LOEHR AND GREGORY S. WARRINGTON

ABSTRACT. The *combinatorial q,t-Catalan numbers* are weighted sums of Dyck paths introduced by J. Haglund and studied extensively by Haglund, Haiman, Garsia, Loehr, and others. The q,t-Catalan numbers, besides having many subtle combinatorial properties, are intimately connected to symmetric functions, algebraic geometry, and Macdonald polynomials. In particular, the n'th q,t-Catalan number is the Hilbert series for the module of diagonal harmonic alternants in $2n$ variables; it is also the coefficient of s_{1^n} in the Schur expansion of $\nabla(e_n)$. Using q,t-analogues of labelled Dyck paths, Haglund et al. have proposed combinatorial conjectures for the monomial expansion of $\nabla(e_n)$ and the Hilbert series of the diagonal harmonics modules.

This article extends the combinatorial constructions of Haglund et al. to the case of lattice paths contained in squares. We define and study several q,t-analogues of these lattice paths, proving combinatorial facts that closely parallel corresponding results for the q,t-Catalan polynomials. We also conjecture an interpretation of our combinatorial polynomials in terms of the nabla operator. In particular, we conjecture combinatorial formulas for the monomial expansion of $\nabla(p_n)$, the "Hilbert series" $\langle\nabla(p_n), h_{1^n}\rangle$, and the sign character $\langle\nabla(p_n), s_{1^n}\rangle$.

1. INTRODUCTION

In 1996, A. Garsia and M. Haiman introduced a two-variable analogue of the Catalan numbers called the *q,t-Catalan numbers* [7]. Garsia and Haiman's definition of the q,t-Catalan, which arose from their study of Macdonald polynomials and diagonal harmonics, was quite complicated. Several years later, J. Haglund [8] conjectured an elementary combinatorial definition of the q,t-Catalan numbers as weighted sums of Dyck paths relative to two statistics called area and bounce. Shortly thereafter, Haiman proposed an equivalent combinatorial interpretation involving area and a third statistic called dinv. Garsia and Haglund eventually proved that the two combinatorial definitions were equivalent to the original definition of Garsia and Haiman [5, 6]. Haiman proved many of the conjectures relating the q,t-Catalan numbers to the representation theory of diagonal harmonics modules and the algebraic geometry of the Hilbert scheme [17, 18]. Meanwhile, various authors studied the subtle combinatorial properties of the combinatorial q,t-Catalan numbers and their generalizations [4, 9, 13, 14, 19, 20, 21, 22, 23, 24]. Surveys of different aspects of this research can be found in [15, 16, 19], and especially [11].

This article discusses a generalization of the combinatorial q,t-Catalan numbers in which Dyck paths are replaced by lattice paths inside squares. We develop the combinatorial theory of these "square q,t-lattice paths," which closely parallels the corresponding theory for the q,t-Catalan numbers. We also conjecture algebraic interpretations for our combinatorial generating functions in terms of the nabla operator introduced by F. Bergeron and Garsia [1, 2, 3]. In particular, we conjecture a combinatorial formula for the monomial expansion of $\nabla(p_n)$ that is quite similar to a formula for $\nabla(e_n)$ conjectured in [13].

To motivate and organize our work on lattice paths inside squares, we begin by quickly reviewing the combinatorial and algebraic results associated with the combinatorial q,t-Catalan numbers.

2000 *Mathematics Subject Classification.* Primary 05E10; Secondary 05A30, 20C30.

Key words and phrases. square lattice paths, diagonal harmonics, Catalan numbers, Catalan paths.

Both authors' research was supported by National Science Foundation Postdoctoral Research Fellowships.

The main body of the paper discusses the corresponding results and conjectures for our square q,t-lattice paths.

1.1. **Combinatorial Aspects of the q,t-Catalan Numbers.** This section reviews the essential definitions and combinatorial results involving the q,t-Catalan numbers. More details can be found in [11, 19] and in various papers listed in the bibliography.

(1) *Lattice Paths and Dyck Paths.* A *lattice path* in a $c \times d$ rectangle is a path from $(0,0)$ to (c,d) consisting of c east steps and d north steps of length 1. Such a path can be represented as a word $w = w_1 \cdots w_{c+d}$ with d zeroes (encoding north steps) and c ones (encoding east steps). Let $\mathcal{R}_{c,d}$ be the set of lattice paths from $(0,0)$ to (c,d). A *Dyck path of order n* is a lattice path in an $n \times n$ rectangle that never visits any point (x,y) with $y < x$. Let $\mathcal{D}_n$ be the set of Dyck paths of order n.

(2) *Statistics on Paths.* In addition to a classical area statistic on Dyck paths, there are two main statistics relevant to this paper: a dinv statistic introduced by Haiman and a bounce statistic introduced by Haglund. We will refer to these three statistics throughout this section, but omit their definitions as they arise as special cases of the corresponding statistics introduced in §2.1 for square q,t-lattice paths.

(3) *Combinatorial q,t-Catalan Numbers.* Haglund [8] defined the combinatorial q,t-Catalan numbers by the formula

$$C_n(q,t) = \sum_{D \in \mathcal{D}_n} q^{\mathrm{area}(D)} t^{\mathrm{bounce}(D)}.$$

There exists a bijection $\alpha : \mathcal{D}_n \to \mathcal{D}_n$ that maps the ordered pair of statistics (area, bounce) to (dinv, area). Therefore, we have

$$C_n(q,t) = \sum_{D \in \mathcal{D}_n} q^{\mathrm{dinv}(D)} t^{\mathrm{area}(D)},$$

which is Haiman's formula for the combinatorial q,t-Catalan numbers.

(4) *Univariate and Joint Symmetry.* The existence of the bijection α implies that

$$\sum_{D \in \mathcal{D}_n} q^{\mathrm{area}(D)} = \sum_{D \in \mathcal{D}_n} q^{\mathrm{dinv}(D)} = \sum_{D \in \mathcal{D}_n} q^{\mathrm{bounce}(D)}$$

and that $C_n(q,1) = C_n(1,q)$. This fact is called *univariate symmetry* of the q,t-Catalan numbers. A stronger result called *joint symmetry* states that $C_n(q,t) = C_n(t,q)$. This result is a corollary of Garsia and Haglund's long proof linking $C_n(q,t)$ to the nabla operator [5, 6]; there is no known direct bijective proof of joint symmetry.

(5) *Recursion.* For $0 \le k \le n$, let $\mathcal{D}_{n,k}$ consist of all Dyck paths $D \in \mathcal{D}_n$ ending with exactly k east steps. It is equivalent to require that $h_0(D) = k$. Set

$$C_{n,k}(q,t) = \sum_{D \in \mathcal{D}_{n,k}} q^{\mathrm{area}(D)} t^{\mathrm{bounce}(D)}. \tag{1}$$

By considering the length $r = h_1(D)$ of $H_1(D)$ for paths $D \in \mathcal{D}_{n,k}$, Haglund [8] proved the recursion

$$C_{n,k}(q,t) = q^{k(k-1)/2} t^{n-k} \sum_{r=0}^{n-k} \begin{bmatrix} r+k-1 \\ r, k-1 \end{bmatrix}_q C_{n-k,r}(q,t) \text{ for } 1 \le k \le n \tag{2}$$

with initial conditions $C_{n,0}(q,t) = \chi(n=0)$ for all $n \ge 0$. Since $C_n(q,t) = t^{-n} C_{n+1,1}(q,t)$, this recursion uniquely determines the q,t-Catalan numbers.

(6) *Fermionic Formula.* By iterating the recursion, Haglund [8] derived an explicit "fermionic" formula for $C_n(q,t)$ as a sum over compositions of n:

$$C_n(q,t) = \sum_{w_0+\cdots+w_s=n,\ w_i>0} q^{\sum_i \binom{w_i}{2}} t^{\sum_i i w_i} \prod_{i=0}^{s-1} \begin{bmatrix} w_{i+1}+w_i-1 \\ w_{i+1}, w_i-1 \end{bmatrix}_q. \tag{3}$$

The summand indexed by $(w_0,\dots,w_s)$ counts those Dyck paths D such that $v_i(D)=w_i$ for $0\le i\le s$.

(7) *Specialization at $t=1/q$.* Using the recursion for $C_{n,k}(q,t)$, Haglund [8] showed that

$$q^{\binom{n}{2}} C_{n,k}(q,1/q) = \frac{[k]_q}{[n]_q} \begin{bmatrix} 2n-k-1 \\ n-k, n-1 \end{bmatrix}_q q^{(k-1)n}.$$

Later, Loehr [19, 24] gave algebraic and bijective proofs of the equivalent formula

$$q^{\binom{n+1}{2}-nk} C_{n,k}(q,1/q) \;=\; \begin{bmatrix} 2n-k-1 \\ n-k, n-1 \end{bmatrix}_q - q^k \begin{bmatrix} 2n-k-1 \\ n-k-1, n \end{bmatrix}_q. \tag{4}$$

Using $C_n(q,t)=t^{-n}C_{n+1,1}(q,t)$ in these formulas, we obtain

$$q^{\binom{n}{2}} C_n(q,1/q) = \frac{1}{[n+1]_q} \begin{bmatrix} 2n \\ n, n \end{bmatrix}_q = \begin{bmatrix} 2n \\ n, n \end{bmatrix}_q - q \begin{bmatrix} 2n \\ n-1, n+1 \end{bmatrix}_q. \tag{5}$$

Garsia and Haiman proved (5) for their original definition of the q,t-Catalan numbers, using completely different methods [7].

(8) *Statistics for Labelled Paths.* The area and dinv statistics for Dyck paths extend naturally to *labelled Dyck paths*, but we omit their definitions. A *labelled Dyck path* of order n is a path $D\in\mathcal{D}_n$ in which each vertical step is assigned a label between 1 and n. We require that the labels of vertical steps in the same column strictly increase from bottom to top. Let $\mathcal{P}_n$ denote the set of all such objects with distinct labels; let $\mathcal{Q}_n$ denote the set of all such objects where labels may be repeated (subject to the increasing-column condition). We can represent a labelled Dyck path $Q\in\mathcal{Q}_n$ by a pair of vectors $(g(Q),r(Q))$, where $g(Q)=(g_0(Q),\dots,g_{n-1}(Q))$ is the area vector of the path (ignoring labels), and $r(Q)=(r_0(Q),\dots,r_{n-1}(Q))$ is the sequence of labels in Q from bottom to top. We call $r(Q)$ the *label vector* of Q. Define the *content function* for Q by letting $c_Q(j)$ be the number of j's in the label vector $r(Q)$.

area and dinv are easily extended to labelled Dyck paths. Haglund and Loehr [14] studied the generating function

$$H_n(q,t) = \sum_{P\in\mathcal{P}_n} q^{\mathrm{area}(P)} t^{\mathrm{dinv}(P)},$$

obtaining a fermionic formula and other results. Later, Loehr [21] defined a third statistic pmaj on $\mathcal{P}_n$, which generalizes Haglund's bounce statistic to labelled paths. The pmaj statistic was used to derive other results about $H_n(q,t)$, such as univariate symmetry, a recursion for $H_n(q,t)$, and the specialization $q^{n(n-1)/2}H_n(q,1/q)=[n+1]_q^{n-1}$. The larger collection of objects $\mathcal{Q}_n$ was first introduced in [13]; its significance is discussed below.

(9) *Combinatorial Extensions.* Haglund's basic idea of studying area, bounce, and dinv statistics on Dyck paths has many fruitful combinatorial generalizations. Besides the labelled Dyck paths just mentioned, one can introduce statistics for Schröder paths, lattice paths inside triangles of different shapes, lattice paths inside trapezoids, labelled versions of these paths, etc. These generalizations share many important properties, like univariate symmetry, joint symmetry (often conjectural), and nice specializations when $t=1/q$. Most of the generalizations also have conjectured algebraic interpretations involving Macdonald

polynomials. We will not enter into any details here but merely refer the reader to the papers in the bibliography. Of course, the goal of the present paper is to introduce yet another extension of the basic combinatorial setup to lattice paths inside squares.

1.2. **Algebraic Aspects of the q,t-Catalan Numbers.** This section reviews the principal theorems and conjectures connecting Haglund's combinatorial q,t-Catalan numbers (and their extensions) to the theory of Macdonald polynomials, diagonal harmonics modules, and symmetric functions. We assume the reader is familiar with the basic definitions and results in symmetric function theory and representation theory; see [25, 26] for more details. We begin by recalling some necessary notation and definitions.

(1) *Partitions.* Let $\mu = (\mu_1 \geq \cdots \geq \mu_s > 0)$ be a partition. Define $|\mu| = \mu_1 + \cdots + \mu_s$, $\mu \vdash n$ iff $|\mu| = n$, $\ell(\mu) = s$, $n(\mu) = \sum_{i=1}^{s}(i-1)\mu_i$, $\mu_i = 0$ for all $i > s$, and write μ' for the transpose of μ. If $\lambda, \mu \vdash n$, write $\lambda \geq \mu$ iff λ dominates μ iff $\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i$ for all i. We draw the Ferrers diagram D of μ in the first quadrant of the xy-plane, left-justified, with the longest row appearing at the bottom. With this convention, for each cell $c \in D$ we define the *arm*, *coarm*, *leg*, and *coleg* of c (denoted $a(c)$, $a'(c)$, $l(c)$, and $l'(c)$) to be the number of cells strictly east, strictly west, strictly north, and strictly south of c in D. Also define the following elements in the polynomial ring $\mathbb{Q}[q,t]$:

$$\begin{aligned}
M &= (1-q)(1-t)\\
B_\mu &= \sum_{c\in\mu} q^{a'(c)}t^{l'(c)}\\
\Pi_\mu &= \prod_{\substack{c\in\mu\\ (a'(c),l'(c))\neq(0,0)}} (1-q^{a'(c)}t^{l'(c)})\\
T_\mu &= q^{\sum_{c\in\mu} a(c)}t^{\sum_{c\in\mu} l(c)} = q^{n(\mu')}t^{n(\mu)}\\
w_\mu &= \prod_{c\in\mu}\left[(q^{a(c)}-t^{l(c)+1})(t^{l(c)}-q^{a(c)+1}).\right]
\end{aligned}$$

For example, if $\mu = (3,2)$, then $|\mu| = 5$, $\ell(\mu) = 2$, $n(\mu) = 2$, $\mu' = (2,2,1)$, $n(\mu') = 4$, $B_\mu = 1+q+q^2+t+qt$, $\Pi_\mu = (1-q)(1-q^2)(1-t)(1-qt)$, $T_\mu = q^4t^2$, and

$$w_\mu = (q^2-t^2)(q-t^2)(1-t)(q-t)(1-t)(t-q^3)(t-q^2)(1-q)(1-q^2)(1-q).$$

(2) *Symmetric Functions.* Let F be the field $\mathbb{Q}(q,t)$. Let Λ_F^n denote the set of symmetric functions homogeneous of degree n in the variables $z_1,\ldots,z_N$ (where $N \geq n$) with coefficients in F. Λ_F^n is an F-vector space whose bases are indexed by partitions of n. We will use the five classical bases for Λ_F^n [25, 26]: the *monomial basis* $\{m_\mu : \mu \vdash n\}$, the *homogeneous basis* $\{h_\mu : \mu \vdash n\}$, the *elementary basis* $\{e_\mu : \mu \vdash n\}$, the *power-sum basis* $\{p_\mu : \mu \vdash n\}$, and the *Schur basis* $\{s_\mu : \mu \vdash n\}$. The *Hall scalar product* is defined on Λ_F^n by requiring that the Schur basis be orthonormal. Then the power-sum basis is orthogonal relative to this scalar product, and the monomial basis is dual to the homogeneous basis.

(3) *Macdonald Polynomials.* Besides the five classical bases, there are also five "Macdonald-type" bases for Λ_F^n [25, 16, 15]: the *Macdonald polynomials* $\{P_\mu : \mu \vdash n\}$, the *dual Macdonald polynomials* $\{Q_\mu : \mu \vdash n\}$, the *integral Macdonald polynomials* $\{J_\mu : \mu \vdash n\}$, the *transformed integral Macdonald polynomials* $\{H_\mu : \mu \vdash n\}$, and the *modified Macdonald polynomials* $\{\tilde{H}_\mu : \mu \vdash n\}$. We will only use the modified Macdonald polynomials, which can be defined quickly as follows. Let ϕ_q be the unique F-linear map on Λ_F^n defined on the basis $\{p_\mu\}$ by $\phi_q(p_\mu) = (\prod_{i=1}^{\ell(\mu)}[1-q^{\mu_i}])p_\mu$. Similarly, define a linear map ϕ_t by requiring that $\phi_t(p_\mu) =$

$(\prod_{i=1}^{\ell(\mu)}[1-t^{\mu_i}])p_\mu$. Then there exists a unique basis $\{\tilde{H}_\mu : \mu \vdash n\}$ of Λ_F^n characterized by the axioms:

(1) $\phi_q(\tilde{H}_\mu) = \sum_{\lambda \geq \mu} a_{\lambda,\mu} s_\lambda$ for some $a_{\lambda,\mu} \in F$

(2) $\phi_t(\tilde{H}_\mu) = \sum_{\lambda \geq \mu'} b_{\lambda,\mu} s_\lambda$ for some $b_{\lambda,\mu} \in F$

(3) $\langle \tilde{H}_\mu, s_n \rangle = 1$.

Haglund recently conjectured an explicit combinatorial formula for $\tilde{H}_\mu$ [10], and this conjecture was proved by Haglund, Haiman, and Loehr by verifying the three axioms [12].

(4) *Nabla Operator.* The nabla operator, introduced by F. Bergeron and Garsia [1, 2, 3], is the unique F-linear map on Λ_F^n defined on the basis $\{\tilde{H}_\mu\}$ by $\nabla(\tilde{H}_\mu) = T_\mu \tilde{H}_\mu$.

(5) *Diagonal Harmonics.* For $n \geq 1$, define $R_n = \mathbb{Q}[x_1, \ldots, x_n, y_1, \ldots, y_n]$. Define the diagonal action of the symmetric group S_n on R_n by $\pi \cdot x_i = x_{\pi(i)}$ and $\pi \cdot y_i = y_{\pi(i)}$ for $\pi \in S_n$. The S_n-module R_n is doubly graded by total degree in the x-variables and total degree in the y-variables. Define the module of *diagonal harmonics* to be

$$DH_n = \left\{ f \in R_n : \sum_{i=1}^{n} \partial x_i^h \partial y_i^k f = 0 \text{ for all } h,k \text{ with } h+k \geq 1 \right\}.$$

Define the module of *diagonal harmonic alternants* to be

$$DHA_n = \{f \in DH_n : \pi \cdot f = \operatorname{sgn}(\pi) f \text{ for all } \pi \in S_n\}.$$

Both DH_n and DHA_n are bihomogeneous submodules of R_n. For any bihomogeneous submodule V of R_n, let $V^{h,k}$ denote the elements of V that are homogenous of degree h in the x-variables and homogeneous of degree k in the y-variables. The *Hilbert series of* V is defined to be

$$\operatorname{Hilb}(V) = \sum_{h,k \geq 0} q^h t^k \dim(V^{h,k}).$$

Writing each $V^{h,k}$ as a direct sum of irreducible S_n-submodules and replacing each such submodule by the associated Schur function, we obtain the *Frobenius series of* V, denoted $\operatorname{Frob}(V)$, which is an element of Λ_F^n. In symbols,

$$\operatorname{Frob}(V) = \sum_{h,k \geq 0} q^h t^k \operatorname{Frob}(V^{h,k}),$$

where $\operatorname{Frob}(V^{h,k})$ is the image of the character of $V^{h,k}$ under the classical Frobenius map.

We can now state some of the theorems and conjectures giving the algebraic significance of the q,t-Catalan numbers and their extensions.

(1) *Master Theorem for the q,t-Catalan Numbers:* For all $n \geq 1$, the following five elements of $\mathbb{Q}(q,t)$ are all equal (and are, therefore, elements of $\mathbb{N}[q,t]$):

(a) $\sum_{D \in \mathcal{D}_n} q^{\operatorname{area}(D)} t^{\operatorname{bounce}(D)}$ (Haglund's combinatorial formula)

(b) $\sum_{D \in \mathcal{D}_n} q^{\operatorname{dinv}(D)} t^{\operatorname{area}(D)}$ (Haiman's combinatorial formula)

(c) $\langle \nabla(e_n), s_{1^n} \rangle$ (nabla formula)

(d) $\sum_{\mu \vdash n} T_\mu^2 M B_\mu \Pi_\mu / w_\mu$ (Garsia-Haiman's rational-function formula)

(e) $\operatorname{Hilb}(DHA_n)$ (representation-theoretical formula)

The equality of (a) and (b) follows from the bijection ϕ mentioned earlier. Formula (d) arises from the expansion of e_n in terms of the basis $\{\tilde{H}_\mu\}$, namely

$$e_n = \sum_{\mu \vdash n} (M B_\mu \Pi_\mu / w_\mu) \tilde{H}_\mu \tag{6}$$

(Theorem 2.4 in [7]). Applying the definition of ∇ and the fact that $\langle \tilde{H}_\mu, s_{1^n} \rangle = T_\mu$, it follows that (c) equals (d). The equality of (d) and (e) follows from a difficult theorem of Mark Haiman giving a complete character formula for DH_n [17, 18]. The equality of (a) and (d) is also a hard result, due to Garsia and Haglund, whose proof uses intricate symmetric function identities and plethystic machinery [5, 6].

(2) *Hilbert Series Conjecture:* For all $n \geq 1$, the following six elements of $\mathbb{Q}(q,t)$ are all equal (and are, therefore, elements of $\mathbb{N}[q,t]$):
   (a) $\sum_{P \in \mathcal{P}_n} q^{\mathrm{dinv}(P)} t^{\mathrm{area}(P)}$ (first combinatorial formula)
   (b) $\sum_{P \in \mathcal{P}_n} q^{\mathrm{area}(P)} t^{\mathrm{pmaj}(P)}$ (second combinatorial formula)
   (c) $\langle \nabla(e_n), h_{1^n} \rangle$ (nabla formula)
   (d) $\sum_{\mu \vdash n} \langle \tilde{H}_\mu, h_{1^n} \rangle T_\mu M B_\mu \Pi_\mu / w_\mu$ (first rational-function formula)
   (e) $\sum_{\mu \vdash n+1} M \Pi_\mu B_\mu^{n+1} / w_\mu$ (second rational-function formula)
   (f) $\mathrm{Hilb}(DH_n)$ (representation-theoretical formula)

   At the time of this writing, the following equalities have been proved: (a)=(b) holds by a bijection in [19, 21]; (c)=(d) follows easily from (6); (c)=(e) was proved by Haglund in [9]; (d)=(f) follows from Haiman's results on the character of DH_n [17, 18]. The open conjecture states that the combinatorial formulas (a) and (b) equal the algebraic formulas (c) through (f).

(3) *Shuffle Conjecture* [13]*:* For all $n \geq 1$, the following five elements of Λ_F^n are all equal (and are, therefore, Schur-positive):
   (a) $\sum_{Q \in \mathcal{Q}_n} q^{\mathrm{dinv}(Q)} t^{\mathrm{area}(Q)} z_1^{c_Q(1)} \cdots z_n^{c_Q(n)}$ (combinatorial formula)
   (b) $\nabla(e_n)$ (nabla formula)
   (c) $\sum_{\mu \vdash n} \tilde{H}_\mu T_\mu M B_\mu \Pi_\mu / w_\mu$ (Macdonald polynomial formula)
   (d) $\sum_{\lambda \vdash n} s_\lambda \left(\sum_{\mu \vdash n+1} M \Pi_\mu B_\mu s_{\lambda'}[B_\mu] / w_\mu \right)$ (rational-function Schur expansion)
   (e) $\mathrm{Frob}(DH_n)$ (representation-theoretical formula)

   Here, $s_{\lambda'}[B_\mu]$ denotes the value of the Schur function $s_{\lambda'}(z_1, \ldots, z_n)$ at $z_i = q^{a'(c_i)} t^{l'(c_i)}$, where $c_1, \ldots, c_n$ are the cells of μ in any order. It is known that (b)=(c) by (6), (b)=(d) by a result of Haglund [9], and (c)=(e) by results of Haiman [17, 18]. The conjecture states that these four algebraic quantities are given by the combinatorial formula (a). If the conjecture is true, then (a) gives the monomial expansion of $\nabla(e_n)$. It has been proved that (a) is a symmetric function in the variables $z_1, \ldots, z_n$ [13].

(4) *Extensions of these Conjectures.* Many of the algebraic conjectures and results extend to the more general combinatorial objects mentioned earlier. For example, the generating functions for q,t-Schröder paths are related to the polynomials $\langle \nabla(e_n), e_d h_{n-d} \rangle$ [4, 9]. Generating functions for unlabelled and labelled m-Dyck paths (paths from $(0,0)$ to (mn,n) never going below the line $x = my$) are conjectured to give information about the sign character and monomial expansion of $\nabla^m(e_n)$ [13, 19, 22, 23]. We again refer the reader to [11] and the papers in the bibliography for more information.

2. Combinatorics of Square q,t-Lattice Paths

2.1. Statistics for Square Paths. A *square lattice path of order* n is a lattice path in an $n \times n$ square. Let $\mathcal{SQ}_n$ denote the set of square lattice paths of order n. We now define three statistics on paths in $\mathcal{SQ}_n$ generalizing the area, dinv, and bounce statistics defined in §1.1.

(1) *Square* area *Statistic.* Let $S \in \mathcal{SQ}_n$. Set $\ell = \ell(S)$ to be the minimum possible value such that S stays weakly above the line $y = x - \ell$. We call ℓ the *deviation* of the path S. Since S begins at the origin and ends at (n,n), we see that $0 \leq \ell \leq n$. Define the *area vector* $g(S) = (g_0(S), \ldots, g_{n-1}(S))$ by requiring that $g_i(S) + n - i$ be the number of complete boxes in the i'th row from the bottom between S and the line $x = n$. Note that the entries of this vector can be negative, but that this area vector reduces to the area vector in §1.1 when

S is a Dyck path. We define $\text{area}(S) = \sum_{i=0}^{n-1}(\ell + g_i(S))$. This can be interpreted as the number of complete boxes to the right of S and to the left of the line $y = x - \ell$.

(2) *Square* dinv *Statistic.* Suppose $S \in \mathcal{SQ}_n$ has $(g_0(S), \ldots, g_{n-1}(S))$ as its area vector. Define

$$\text{dinv}(S) = \sum_{i<j} \chi(g_i(S) - g_j(S) \in \{0,1\}) + \sum_i \chi(g_i(S) < -1).$$

If S is a Dyck path, then the condition $g_i(S) < -1$ never holds, and this formula for $\text{dinv}(S)$ reduces to the formula given in §1.1.

(3) *Square* bounce *Statistic.* Let $S \in \mathcal{SQ}_n$ have deviation ℓ. The *break point* of S, $(\ell_x(S), \ell_y(S))$, is the leftmost point along the path S lying on the line $y = x - \ell$.

We now proceed to define a bounce path $\text{bpath}(S)$ in analogy with the bounce paths defined for Dyck paths. The bounce path for S consists of two pieces: a *positive part* located northeast of the break point, and a *negative part* located southwest of the break point. First consider the positive part. A ball starts at (n,n) and makes an initial vertical move V_{-1} of length $v_{-1} = \ell$ ending at $(n, n-\ell)$. The ball then makes alternating horizontal and vertical moves $H_0, V_0, H_1, V_1, \ldots, H_s, V_s$ until it reaches the break point. We let h_i and v_i denote the length of the moves H_i and V_i, respectively. We determine h_i and v_i for each $i \geq 0$ as follows. First, the ball moves west h_i units until it is blocked by the north step of S ending at the horizontal level occupied by the ball. Second, the ball moves south $v_i = h_i$ units to return to the line $y = x - \ell$. As before, the steps that block the ball's westward motion are called *blocking north steps.*

The negative part of the bounce path traces the motion of a second bouncing ball that starts at the origin and moves northeast towards the break point. This ball makes an initial horizontal move H_{-1} of length $h_{-1} = \ell$ from $(0,0)$ to $(\ell, 0)$. It then makes alternating vertical and horizontal moves $V_{-2}, H_{-2}, V_{-3}, H_{-3}, \ldots, V_u, H_u$ until it reaches the break point. For each $i < -1$, the ball moves north v_i units until it is blocked by the east step of S *ending* at the vertical line occupied by the ball. (Note that this is *not* just a reflected version of the bounce algorithm in the positive part.) The ball then moves east $h_i = v_i$ units to return to the line $y = x - \ell$. The east steps that block the ball's northward motion are called *blocking east steps.*

Finally, we define the *bounce statistic* for any path $S \in \mathcal{SQ}_n$. Let $V_u, \ldots, V_s$ be the nonzero vertical moves in $\text{bpath}(S)$, where $u \leq 0 \leq s$. Set $\text{bounce}(S) = \sum_{i=u}^{s}(i-u)v_i$. Also set $\text{bmin}(S) = u$ and $\text{bmax}(S) = s$.

For a Dyck path D, the deviation ℓ is 0, the break point is $(0,0)$, the positive part of the bounce path coincides with the bounce path described in §1.1, and the negative part of the bounce path is empty. In this case, we have $\text{bmin}(D) = 0$ (we ignore the zero moves V_{-1} and H_{-1}), and the bounce statistic just defined reduces to the formula used in §1.1.

For example, Figure 1.1 illustrates a path $S \in \mathcal{SQ}_{15}$ and its bounce path. For this path, $\ell(S) = 3$, the break point is $(8,5)$,

$$g(S) = (0, -1, -2, -1, -1, -3, -2, -2, -2, -3, -2, -1, -1, 0, 1),$$

$\text{area}(S) = 25$, $\text{dinv}(S) = 52$, $\text{bmin}(S) = -4$, $\text{bmax}(S) = 2$, $(v_{-4}, \ldots, v_2) = (h_{-4}, \ldots, h_2) = (1,2,2,3,3,2,2)$, and $\text{bounce}(S) = 49$.

2.2. **Comparison of the Statistics.**

Theorem 1. *There exists a bijection* $\phi : \mathcal{SQ}_n \to \mathcal{SQ}_n$ *such that* $\text{area}(\phi(S)) = \text{dinv}(S)$ *and* $\text{bounce}(\phi(S)) = \text{area}(S)$. *The deviation of* $\phi(S)$ *is the number of* -1*'s in* $g(S)$*; the break point of* $\phi(S)$ *is*

$$(\ell_x(\phi(S)), \ell_y(\phi(S))) = (|\{j : g_j(S) < 0\}|, |\{j : g_j(S) < -1\}|);$$

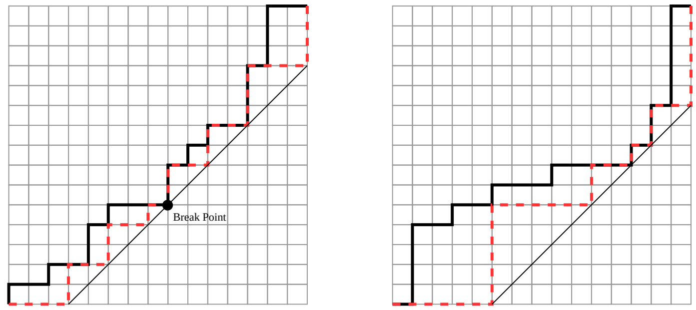


FIGURE 1. The lefthand picture is a path S in $\mathcal{SQ}_{15}$ (solid path) along with bpath(S) (dotted path). The righthand picture is the image of S under ϕ.

$\mathrm{bmin}(\phi(S)) = \min_j g_j(S)$; *and* $\mathrm{bmax}(\phi(S)) = \max_j g_j(S)$. *Moreover,* $\phi(S)$ *ends with an east step iff* S *begins with a north step, and* $\phi|_{\mathcal{D}_n} : \mathcal{D}_n \to \mathcal{D}_n$ *is the inverse of the bijection* α *from §1.1.*

For example, let S be the path from Figure 1.1. Figure 1.2 shows $\phi(S)$ along with its bounce path. Note that $g(S)$ has two -3's, five -2's, five -1's, two 0's, and one 1, so that

$$(v_{-3}(\phi(S)), \ldots, v_1(\phi(S))) = (2, 5, 5, 2, 1).$$

Furthermore, $\mathrm{area}(S) = 25 = \mathrm{bounce}(\phi(S))$ and $\mathrm{dinv}(S) = 52 = \mathrm{area}(\phi(S))$.

2.3. **Symmetry Properties.** For all $n \geq 1$, define

$$S_n(q,t) = \sum_{S \in \mathcal{SQ}_n} q^{\mathrm{area}(S)} t^{\mathrm{bounce}(S)} = S_n(q,t) = \sum_{S \in \mathcal{SQ}_n} q^{\mathrm{dinv}(S)} t^{\mathrm{area}(S)}.$$

(The second equality follows from the bijection ϕ.) Letting $q = 1$ or $t = 1$ here, we obtain the following *univariate symmetry* properties.

Corollary 2.

$$\sum_{S \in \mathcal{SQ}_n} q^{\mathrm{dinv}(S)} = \sum_{S \in \mathcal{SQ}_n} q^{\mathrm{bounce}(S)} = \sum_{S \in \mathcal{SQ}_n} q^{\mathrm{area}(S)}.$$

Conjecture 3. The *joint symmetry* property $S_n(q,t) = S_n(t,q)$ holds for all n.

This conjecture has been confirmed by computer for $1 \leq n \leq 11$.

Computing $S_n(q,t)$ for small values of n, one sees that the polynomial $S_n(q,t)$ is always divisible by 2. Our next goal is to explain this property. Define $\mathcal{SQ}_n^N$, $\mathcal{SQ}_n^E$, ${}^N\mathcal{SQ}_n$, and ${}^E\mathcal{SQ}_n$ to be the paths in $\mathcal{SQ}_n$ that end with a north step, end with an east step, begin with a north step, and begin with an east step, respectively. Set

$$S_n^\gamma(q,t) = \sum_{S \in \mathcal{SQ}_n^\gamma} q^{\mathrm{area}(S)} t^{\mathrm{bounce}(S)}; \qquad {}^\gamma S_n(q,t) = \sum_{S \in {}^\gamma\mathcal{SQ}_n} q^{\mathrm{dinv}(S)} t^{\mathrm{area}(S)}.$$

for $\gamma \in \{E, N\}$. We will show that $S_n^N(q,t) = S_n^E(q,t) = S_n(q,t)/2$. Since ϕ sends paths with initial north steps to paths with terminal east steps and vice versa, it also follows that ${}^N S_n(q,t) = {}^E S_n(q,t) = S_n(q,t)/2$. We call these identities *pair-symmetries*.

To prove the pair-symmetries, it suffices to construct a bijection $\psi : \mathcal{SQ}_n^E \to \mathcal{SQ}_n^N$ preserving area and bounce. We begin by introducing a cyclic shift map $\mathrm{cyc} : \mathcal{SQ}_n \to \mathcal{SQ}_n$. Let $S \in \mathcal{SQ}_n$ be encoded by the word $w_1 w_2 \cdots w_{2n} \in \{0,1\}^{2n}$. Define $\mathrm{cyc}(S)$ to be the path encoded by the word $w_{2n} w_1 w_2 \cdots w_{2n-1}$.

Lemma 4. *For* $S \in \mathcal{SQ}_n$, $\mathrm{area}(S) = \mathrm{area}(\mathrm{cyc}^i(S))$ *for all integers* i.

Theorem 5. *There is a bijection* $\psi : \mathcal{SQ}_n^E \to \mathcal{SQ}_n^N$ *preserving* area *and* bounce. *Consequently,*

$$S_n^N(q,t) = S_n^E(q,t) = S_n(q,t)/2 = {}^N S_n(q,t) = {}^E S_n(q,t).$$

We close this section with an alternate formula for ${}^E S_n(q,t) = S_n(q,t)/2$.

Theorem 6. *Let* $\mathrm{dinv}_0(S) = \sum_{i<j} \chi(g_i(S) - g_j(S) \in \{0,1\}) + \sum_i \chi(g_i(S) < 0)$. *Then*

$${}^E S_n(q,t) = \sum_{S \in {}^E\mathcal{SQ}_n} q^{\mathrm{dinv}(S)} t^{\mathrm{area}(S)} = \sum_{U \in \mathcal{SQ}_n^E} q^{\mathrm{dinv}_0(U)} t^{\mathrm{area}(U)}.$$

2.4. **Recursion for Square Paths.** In §1.1, we saw that the generating functions $C_{n,k}(q,t)$ for Dyck paths of order n with $h_0 = k$ satisfied the recursion (2). Now we prove a similar recursion for square q,t-lattice paths. The idea is to remove the "earliest" bounce in a square q,t-lattice path, namely the negative bounce arriving at the break point.

Formally, for $n > 0$ and $1 \leq k \leq n$, we set

$$R_{n,k}(q,t) = \sum_{S \in \mathcal{SQ}_n} q^{\mathrm{area}(S)} t^{\mathrm{bounce}(S)} \chi(h_{\mathrm{bmin}(S)} = k,\ \ell(S) > 0).$$

The condition $\ell(S) > 0$ means that S is not a Dyck path, while $h_{\mathrm{bmin}(S)} = k$ means that the last horizontal move in the negative part of the bounce path (arriving at the break point) has length k. To take care of the Dyck paths in $\mathcal{SQ}_n$, we define $R_{n,0}(q,t) = C_n(q,t) = t^{-n} C_{n+1,1}(q,t)$ for $n \geq 0$. For $k = n \geq 0$, we have $R_{n,n}(q,t) = q^{\binom{n}{2}}$ since the only path that contributes is the one that goes east n steps and then north n steps. Clearly, $S_{n,k}(q,t) = \sum_{k=0}^{n} R_{n,k}(q,t)$.

Theorem 7. *For* $0 < k < n$,

$$R_{n,k}(q,t) = q^{\binom{k}{2}} t^{n-k} \sum_{r=1}^{n-k} \begin{bmatrix} r+k \\ r,k \end{bmatrix}_q C_{n-k,r}(q,t) + q^{\binom{k}{2}} t^{n-k} \sum_{r=1}^{n-k} q^k \begin{bmatrix} r+k-1 \\ r-1,k \end{bmatrix}_q R_{n-k,r}(q,t). \tag{7}$$

Note that recursions (2) and (7), and the initial conditions, uniquely determine the quantities $R_{n,k}(q,t)$ and $S_n(q,t)$ and provide an efficient method for computing them.

2.5. **Fermionic Formula.** We now obtain a fermionic formula for $S_n(q,t)$ in analogy with (3).

Theorem 8. *For* $n \geq 1$,

$$\begin{aligned} S_n(q,t) = 2q^{\binom{n}{2}} + \sum_{\substack{w_0+\cdots+w_s=n \\ w_i>0;s\geq 1}} \Bigg(& q^{\mathrm{pow}_1} t^{\mathrm{pow}_2} \prod_{j=0}^{s-1} \begin{bmatrix} w_j + w_{j+1} - 1 \\ w_j - 1, w_{j+1} \end{bmatrix}_q + \\ & \sum_{a=0}^{s-1} q^{\mathrm{pow}_3} t^{\mathrm{pow}_2} \begin{bmatrix} w_a + w_{a+1} \\ w_a, w_{a+1} \end{bmatrix}_q \prod_{j=0}^{a-1} \begin{bmatrix} w_j + w_{j+1} - 1 \\ w_j, w_{j+1} - 1 \end{bmatrix}_q \prod_{j=a+1}^{s-1} \begin{bmatrix} w_j + w_{j+1} - 1 \\ w_j - 1, w_{j+1} \end{bmatrix}_q \Bigg), \end{aligned} \tag{8}$$

where $\mathrm{pow}_1 = \sum_{j=0}^{s} \binom{w_j}{2}$, $\mathrm{pow}_2 = \sum_{j=0}^{s} j w_j$, *and* $\mathrm{pow}_3 = \mathrm{pow}_1 + \sum_{0 \leq j < a} w_j$.

For $0 < k < n$, there is a similar fermionic formula for $R_{n,k}(q,t)$. We simply use the second line of (8), summing over all $(w_0, \ldots, w_s)$ and all a such that $w_0 + \cdots + w_s = n$, $w_i > 0$, $s_i \geq 1$, $0 \leq a \leq s-1$, *and* fixing $w_0 = k$.

2.6. **Specialization at $t = 1/q$.** Our next goal is to derive explicit formulas for $R_{n,k}(q, 1/q)$ and $S_n(q, 1/q)$ similar to the formulas for $C_{n,k}(q, 1/q)$ and $C_n(q, 1/q)$ from §1.1.

Theorem 9. *For $1 \leq k \leq n$,*

$$q^{\binom{n-k+1}{2}-\binom{k}{2}} R_{n,k}(q, 1/q) = \begin{bmatrix} 2n-k-1 \\ n-1, n-k \end{bmatrix}_q + q^k \begin{bmatrix} 2n-k-1 \\ n, n-k-1 \end{bmatrix}_q. \tag{9}$$

For $k = 0$, $q^{\binom{n}{2}} R_{n,0}(q, 1/q) = q^{\binom{n}{2}} C_n(q, 1/q)$ is given by formula (5).

Theorem 10. *For all $n \geq 1$,*

$$q^{\binom{n}{2}} S_n(q, 1/q) = 2 \begin{bmatrix} 2n-1 \\ n, n-1 \end{bmatrix}_q = \frac{2}{1+q^n} \begin{bmatrix} 2n \\ n, n \end{bmatrix}_q.$$

We remark that the methods in [24] can be used to mechanically translate the preceding algebraic manipulations into bijective proofs of the same results. However, because of all the subtractions involved, the bijections will be extremely complicated.

3. Algebraic Conjectures for Square Paths

We now give some conjectures connecting square q, t-lattice paths to Macdonald polynomials and the nabla operator. These conjectures closely resemble the corresponding results for the q, t-Catalan numbers from §1.2.

3.1. **Unlabelled Paths.** *Master Conjecture for Square q, t-Lattice Paths:* For all $n \geq 1$, the following elements of $\mathbb{Q}(q, t)$ are all equal (and are, therefore, elements of $\mathbb{N}[q, t]$):

(a) $\sum_{S \in \mathcal{SQ}_n^N} q^{\mathrm{area}(S)} t^{\mathrm{bounce}(S)}$
(b) $\sum_{S \in \mathcal{SQ}_n^E} q^{\mathrm{area}(S)} t^{\mathrm{bounce}(S)}$
(c) $\sum_{S \in {}^N\mathcal{SQ}_n} q^{\mathrm{dinv}(S)} t^{\mathrm{area}(S)}$
(d) $\sum_{S \in {}^E\mathcal{SQ}_n} q^{\mathrm{dinv}(S)} t^{\mathrm{area}(S)}$
(e) $\sum_{S \in \mathcal{SQ}_n^E} q^{\mathrm{dinv}_0(S)} t^{\mathrm{area}(S)}$
(f) $(-1)^{n-1} \langle \nabla(p_n), s_{1^n} \rangle$
(g) $\sum_{\mu \vdash n} (1-t^n)(1-q^n) \Pi_\mu T_\mu^2 / w_\mu = \sum_{\mu \vdash n} MB_{(n^n)} \Pi_\mu T_\mu^2 / w_\mu$

We have already seen that (a) through (e) are equal, using the bijections ψ, ϕ, and cyc^{-1}. To see that (f) equals (g), we use the expansion of p_n in terms of the basis $\{\tilde{H}_\mu\}$, namely

$$p_n = \sum_{\mu \vdash n} ((-1)^{n-1}(1-t^n)(1-q^n)\Pi_\mu / w_\mu) \tilde{H}_\mu. \tag{10}$$

This expansion follows immediately from Corollary 2.4 in [7] and the definition of plethysm. Applying ∇ replaces each $\tilde{H}_\mu$ by $T_\mu \tilde{H}_\mu$, and taking the scalar product with s_{1^n} turns $\tilde{H}_\mu$ into another factor T_μ. Hence, (f) equals (g). The main conjecture, asserting that (a) equals (f), has been tested for $1 \leq n \leq 8$.

3.2. **Labelled Paths.** Fix n and N with $n \leq N \leq \infty$. Let $\mathcal{SQF}_n$ denote the set of all pairs (S, r), where: S is a path in $\mathcal{SQ}_n^E$ (so that S ends with an east step); and $r = r_0 \ldots r_{n-1}$ is a *label vector* with $r_i \in \{1, 2, \ldots, N\}$ such that $g_{i+1}(S) = g_i(S) + 1$ implies $r_i < r_{i+1}$. If we attach the labels r_i to the vertical steps of S as we did for Dyck paths, then the last condition means that labels in each column must strictly increase from bottom to top. Let $\mathcal{SQH}_n$ denote the subset of $\mathcal{SQF}_n$ such that $r_0 \dots r_{n-1}$ is a permutation of $\{1,2,\dots,n\}$. Given $(S,r) \in \mathcal{SQF}_n$, define $\operatorname{area}(S,r) = \operatorname{area}(S)$ and

$$\operatorname{dinv}_0(S,r) = \sum_{i<j} \chi((g_i(S) - g_j(S) = 0 \text{ and } r_i < r_j) \text{ or } \\ (g_i(S) - g_j(S) = 1 \text{ and } r_i > r_j)) + \sum_i \chi(g_i(S) < 0).$$

(It is equivalent to use all labelled paths *beginning* with an east step, replacing $\chi(g_i(S) < 0)$ by $\chi(g_i(S) < -1)$ in the definition of dinv_0.)

Hilbert series conjecture for square q,t-lattice paths: For all $n \geq 1$,

$$(-1)^{n-1}\langle \nabla(p_n), h_{1^n} \rangle = \sum_{(S,r)\in \mathcal{SQH}_n} q^{\operatorname{area}(S,r)} t^{\operatorname{dinv}_0(S,r)}.$$

Frobenius series conjecture for square q,t-lattice paths: For all $n \geq 1$,

$$(-1)^{n-1}\nabla(p_n[z_1,\dots,z_N]) = \sum_{(S,r)\in \mathcal{SQF}_n} q^{\operatorname{area}(S,r)} t^{\operatorname{dinv}_0(S,r)} \prod_{i=0}^{n-1} z_{r_i}.$$

Applying ∇ to (10), we see that

$$(-1)^{n-1}\nabla(p_n) = \sum_{\mu \vdash n} ((1-t^n)(1-q^n)\Pi_\mu T_\mu / w_\mu) \tilde{H}_\mu.$$

We remark that the same arguments used in [13] show that the Frobenius series conjecture implies both of the preceding conjectures, along with shuffle-type formulas for any scalar product of the form $\langle \nabla(p_n), h_\mu e_\eta \rangle$. It is an open problem to find a naturally occurring doubly-graded S_n-module M_n that has $(-1)^{n-1}\nabla(p_n)$ as its Frobenius series. Since elements of $\mathcal{SQH}_n$ encode functions $f : \{1,2,\dots,n\} \to \{1,2,\dots,n\}$ in the obvious way ($f^{-1}(\{i\})$ are the labels in column i), we should have $\dim(M_n) = |\mathcal{SQH}_n| = n^n$.

3.3. **More Nabla Conjectures.** So far, we have seen combinatorial formulas that are conjectured to give the monomial expansions of $\nabla(e_n)$ and $(-1)^{n-1}\nabla(p_n)$. Since $e_n = m_{(1^n)}$ and $p_n = m_{(n)}$, these results suggest that $\nabla(m_\mu)$ may have a nice monomial expansion for any $\mu \vdash n$. In fact, an even stronger statement appears to be true.

Conjecture 11. For all $n \geq 1$ and $\mu, \nu \vdash n$,

$$(-1)^{n-\ell(\mu)}\langle \nabla(m_\mu), s_\nu \rangle \in \mathbb{N}[q,t].$$

The conjecture has been tested for $1 \leq n \leq 8$. If the conjecture is true, it readily follows that $\nabla(m_\mu)|_{m_\nu} \in \mathbb{N}[q,t]$ for all μ and ν. In [3], Bergeron, Garsia, Haiman, and Tesler made the analogous conjecture

$$\iota(\mu')\langle \nabla(s_\mu), s_\nu \rangle \in \mathbb{N}[q,t],$$

where $\iota(\mu) = \binom{\ell(\mu)}{2} + \sum_{\mu_i < (i-1)} (i-1-\mu_i)$. This second conjecture implies that $\nabla(s_\mu)|_{m_\nu} \in \mathbb{N}[q,t]$ for all μ and ν. Because of the signs, it is not clear whether either conjecture easily implies the other one.

References

[1] F. Bergeron and A. Garsia, "Science Fiction and Macdonald Polynomials," *CRM Proceedings and Lecture Notes AMS* VI **3** (1999), 363—429.

[2] F. Bergeron, N. Bergeron, A. Garsia, M. Haiman, and G. Tesler, "Lattice Diagram Polynomials and Extended Pieri Rules," *Adv. in Math.* **2** (1999), 244—334.

[3] F. Bergeron, A. Garsia, M. Haiman, and G. Tesler, "Identities and Positivity Conjectures for some remarkable Operators in the Theory of Symmetric Functions," *Methods and Applications of Analysis* VII **3** (1999), 363—420.

[4] E. Egge, J. Haglund, D. Kremer, and K. Killpatrick, "A Schröder generalization of Haglund's statistic on Catalan paths," *Electronic Journal of Combinatorics* **10** (2003), R16, 21 pages.
[5] 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.
[6] A. Garsia and J. Haglund, "A positivity result in the theory of Macdonald polynomials," *Proc. Nat. Acad. Sci.* **98** (2001), 4313—4316.
[7] A. Garsia and M. Haiman, "A remarkable q,t-Catalan sequence and q-Lagrange Inversion," *J. Algebraic Combinatorics* **5** (1996), 191—244.
[8] J. Haglund, "Conjectured Statistics for the q,t-Catalan numbers," *Advances in Mathematics* **175** (2003), 319—334.
[9] J. Haglund, "A proof of the q,t-Schröder conjecture," *Intl. Math. Res. Notices* **11** (2004), 525—560.
[10] J. Haglund, "A combinatorial model for the Macdonald polynomials," *Proc. Nat. Acad. Sci. USA*, **101**, 46 (2004), 16127–16131.
[11] J. Haglund, *The q,t-Catalan Numbers and the Space of Diagonal Harmonics.* AMS University Lecture Series, to appear.
[12] J. Haglund, M. Haiman, and N. Loehr, "A combinatorial formula for Macdonald polynomials," arXiv:math.CO/0409538, 2004.
[13] J. Haglund, M. Haiman, N. Loehr, J. Remmel, and A. Ulyanov, "A combinatorial formula for the character of the diagonal coinvariants," to appear in *Duke Math. J.*
[14] J. Haglund and N. Loehr, "A Conjectured Combinatorial Formula for the Hilbert Series for Diagonal Harmonics." *Proceedings of FPSAC 2002 Conference* (Melbourne, Australia), to appear in *Discrete Mathematics.*
[15] M. Haiman, "Notes on Macdonald Polynomials and the Geometry of Hilbert Schemes," *Symmetric Functions 2001: Surveys of Developments and Perspectives,* Proceedings of the NATO Advaced Study Institute. Sergey Fomin, ed. Kluwer, Dordrecht (2002) 1—64.
[16] M. Haiman, "Combinatorics, symmetric functions, and Hilbert schemes," *CDM 2002: Current Developments in Mathematics*, Intl. Press Books (2003), 39—112.
[17] M. Haiman, "Hilbert schemes, polygraphs, and the Macdonald positivity conjecture." *J. Amer. Math. Soc.* **14** (2001), 941—1006.
[18] M. Haiman. "Vanishing theorems and character formulas for the Hilbert scheme of points in the plane," *Invent. Math.* **149** (2002), 371—407.
[19] N. Loehr, *Multivariate Analogues of Catalan Numbers, Parking Functions, and their Extensions.* Ph.D. thesis, University of California at San Diego, June 2003.
[20] N. Loehr, "Trapezoidal Lattice Paths and Multivariate Analogues," *Adv. in Appl. Math.* **31** (2003), 597—629.
[21] N. Loehr, "Combinatorics of q,t-Parking Functions," to appear in *Adv. in Appl. Math.*.
[22] N. Loehr, "Conjectured Statistics for the Higher q,t-Catalan Sequences." Preprint.
[23] N. Loehr and J. Remmel, "Conjectured Combinatorial Models for the Hilbert Series of Generalized Diagonal Harmonics Modules," *Electronic Journal of Combinatorics* **11** (2004), R68.
[24] N. Loehr, "The Major Index Specialization of the q,t-Catalan," to appear in *Ars Combinatoria.*
[25] I. G. Macdonald, *Symmetric Functions and Hall Polynomials.* 2nd ed. Oxford University Press, 1995.
[26] B. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions.* Wadsworth and Brooks/Cole, 1991.

Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104
E-mail address: nloehr@math.upenn.edu

Department of Mathematics, Wake Forest University, Winston-Salem, NC 27109
E-mail address: warrings@wfu.edu

POSITIVITY QUESTIONS FOR CYLINDRIC SKEW SCHUR FUNCTIONS

PETER MCNAMARA

Abstract. Recent work of A. Postnikov shows that cylindric skew Schur functions, which are a generalisation of skew Schur functions, have a strong connection with a problem of considerable current interest: that of finding a combinatorial proof of the non-negativity of the 3-point Gromov-Witten invariants. After explaining this motivation, we study cylindric skew Schur functions from the point of view of Schur-positivity. Using a result of I. Gessel and C. Krattenthaler, we generalise a formula of A. Bertram, I. Ciocan-Fontanine and W. Fulton, thus giving an expansion of an arbitrary cylindric skew Schur function in terms of skew Schur functions. While we show that no non-trivial cylindric skew Schur functions is Schur-positive, we conjecture that this can be reconciled using the new concept of cylindric Schur-positivity.

Résumé. Les travaux récents de A. Postnikov montrent que les fonctions gauches cylindriques de Schur, qui sont une généralisation des fonctions gauches de Schur, ont un lien étroit avec un problème actuellement très étudié : trouver une preuve combinatoire de la non-negativité des invariants de Gromov-Witten de 3-pointes. Après avoir expliqué cette motivation, nous étudions les fonctions gauches cylindriques de Schur du point de vue de la Schur-positivité. En utilisant un résultat de I. Gessel et C. Krattenthaler, nous généralisons une formule de A. Bertram, I. Ciocan-Fontanine et W. Fulton, donnant ainsi une expansion d'une fonction gauche cylindrique de Schur arbitraire en termes de fonctions gauches de Schur. Tandis que nous prouvons qu'aucune fonction gauche cylindrique de Schur non triviale n'est Schur-positive, nous introduisons le concept de Schur-positivité cylindrique et conjecturons que tout fonction gauche cylindrique de Schur est Schur-positive cylindrique.

1. Introduction

The Schur functions $s_\lambda(x)$, where λ runs over all partitions, form an important basis for the ring of symmetric functions in infinitely many variables $x = (x_1, x_2, \ldots)$. In particular, the skew Schur function $s_{\lambda/\mu}(x)$ can be expanded in terms of Schur functions as

$$s_{\lambda/\mu}(x) = \sum_{\nu} c^{\lambda}_{\mu\nu} s_\nu(x),$$

where $c^{\lambda}_{\mu\nu}$ denotes the ubiquitous *Littlewood-Richardson* coefficients. It is well known that Littlewood-Richardson coefficients are non-negative and a skew Schur function is thus one of the most famous examples of a *Schur-positive* function, i.e. a symmetric function whose expansion as a linear combination of Schur functions has all positive coefficients. It is worth mentioning that Schur-positivity has a particular representation-theoretic significance: if a homogeneous symmetric function of degree N is Schur-positive, then it is known to arise as the Frobenius image of some representation of the symmetric group S_N. This is one of the reasons why questions of Schur-positivity have received, and continue to receive, much attention in recent times.

As their name suggests, *cylindric* skew Schur functions are a natural generalisation of skew Schur functions. In particular, cylindric skew Schur functions are symmetric functions, and skew Schur functions are themselves cylindric skew Schur functions. While this is enough to give them combinatorial appeal, cylindric skew Schur functions have recently been shown to play a central role in the study of the quantum cohomology ring of the Grassmannian. More specifically, this contemporary motivation for cylindric skew Schur functions involves the fundamental open problem of finding a combinatorial proof of the non-negativity of the 3-point Gromov-Witten invariants. While it will be our starting point in Section 3, no knowledge of quantum cohomology will be assumed and our emphasis will be combinatorial. Gromov-Witten invariants are connected to the topic of cylindric skew Schur functions

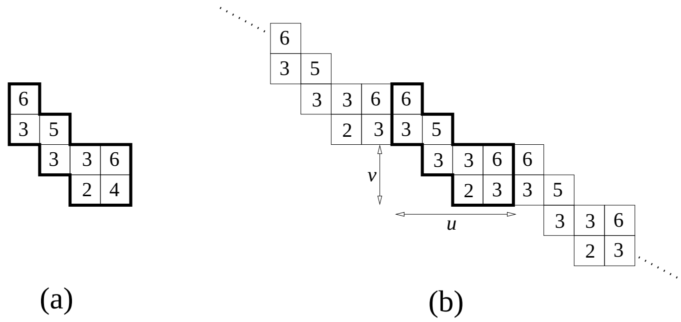


FIGURE 1.

via a theorem of A. Postnikov [12]. Since cylindric skew Schur functions are symmetric, they can be expanded in terms of Schur functions and Postnikov's theorem states that the Gromov-Witten invariants appear as particular coefficients in this expansion. The fundamental open problem mentioned above then becomes a question about the Schur-positivity of cylindric skew Schur functions. In short, cylindric skew Schur functions play a similar role for the 3-point Gromov-Witten invariants as that played by skew Schur functions for the Littlewood-Richardson coefficients. Rather than tackling the fundamental open problem directly, however, our goal is to give a general study of the Schur-positivity of cylindric skew Schur functions.

At this point, it makes sense to give an informal introduction to cylindric skew shapes, semistandard cylindric tableaux and cylindric skew Schur functions. Suppose we are given a skew shape and a semistandard Young tableau of that shape, such as the one shown in French notation in Figure 1(a). We can think of cylindric skew shapes as coming from skew shapes that have been wrapped around a cylinder so that boxes at the bottom of the rightmost column are now directly to the left of boxes at the top of the leftmost column. We will represent this cylindric skew shape C as a skew shape together with its images under repeated applications of some translation $(-u, v)$, as shown in Figure 1(b). Notice that we have used our previous semistandard Young tableau but have had to modify one of the entries to ensure that the entries continue to be weakly increasing in the rows. The result is an example of a semistandard cylindric tableau. The definition of the cylindric skew Schur function $s_C(x)$ is then completely analogous to that of a skew Schur function. This example motivates the formal introduction that is the subject of Section 2.

The geometric definition of Gromov-Witten invariants, combined with Postnikov's theorem, tells us that cylindric skew Schur functions in a restricted number of variables are Schur-positive. In Section 4, we show that, except for trivial cases, cylindric skew Schur functions are never Schur-positive in infinitely many variables. Since they play a crucial role in our proof of this result, we investigate the class of "cylindric ribbons," determining the form of the Schur expansion of their corresponding cylindric skew Schur functions. We conclude Section 4 with a discussion of the minimum number of variables in which a cylindric skew Schur function will not be Schur-positive.

In Section 5, we develop a tool for expanding cylindric skew Schur functions as a signed sum of skew Schur functions. A result of I. Gessel and C. Krattenthaler [5] serves as the foundation for our tool, while our formulation is inspired by a result of A. Bertram, I. Ciocan-Fontanine and W. Fulton [2].

Since cylindric skew Schur functions are such a natural generalisation of skew Schur functions, one might ask if there is some way to extend the Schur-positivity of skew Schur functions to the cylindric setting. In Section 6, we define cylindric Schur-positivity as an analogue of Schur-positivity and we give evidence in favour of a conjecture that every cylindric skew Schur functions is cylindric Schur-positive.

Before beginning in earnest, we introduce terminology and notation that we will use throughout. We will denote the sets of integers, non-negative integers and positive integers by $\mathbb{Z}$, $\mathbb{N}$ and $\mathbb{P}$ respectively. For $N \in \mathbb{P}$, we will write $[N]$ to denote the set $\{1, 2, \ldots, N\}$. For symmetric function notation, we will follow [8].

We will allow a partition $\lambda = (\lambda_1, \ldots, \lambda_k)$ to have parts equal to zero and we identify λ with the sequence $(\lambda_1, \ldots, \lambda_k, 0, 0, \ldots)$. We write $l(\lambda)$ for the number of non-zero parts (length) of λ and $|\lambda|$ for the sum of the parts of λ. We use a^k in the list of parts of a partition to denote a sequence of k a's. For example, $\lambda = (j, 1^k)$ has one part of size j and k parts of size 1. We let $\emptyset$ denote the unique partition with length 0. We can represent a partition λ by its Young diagram, and then the *conjugate partition* $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$ of λ is the partition obtained by reading the column lengths of λ from left to right.

If μ is another partition then we say that $\mu \subseteq \lambda$ if $\mu_i \leq \lambda_i$ for all i; this is equivalent to saying that the diagram of μ is contained in the diagram of λ. If $\mu \subseteq \lambda$, then we define the *skew shape* λ/μ to be the set of boxes in the diagram of λ that remain after we remove those boxes corresponding to the partition μ. A *ribbon* (or *rim hook* or *border strip*) is an edgewise connected skew shape that contains no 2×2 block of boxes. An *n-ribbon* is then simply a ribbon with n boxes.

Acknowledgements

This extended abstract is an abbreviated version of [11], which results from work begun while the author was a graduate student at MIT. I am grateful to my advisor, Richard Stanley, and to Alex Postnikov for several interesting discussions on the topic. François Bergeron and Christophe Reutenauer, my mentors at LaCIM, have both made valuable suggestions, and their expertise and enthusiasm have been of considerable assistance.

2. Cylindric skew Schur functions

Cylindric skew shapes are not a new idea and there are three references in particular that are of great relevance to our work. The first of these is [5], which will play a significant role in Section 5. Semistandard cylindric tableaux, which we will shortly define, appear under the name "proper tableaux" in [2]. The main result of [12] serves as the starting point for our results. For the following introduction to the notation and definition of cylindric skew shapes, we will largely follow [12].

Fix positive integers u and v. We define the *cylinder* $\mathfrak{C}_{vu}$ to be the following quotient of $\mathbb{Z}^2$:

$$\mathfrak{C}_{vu} = \mathbb{Z}^2/(-u, v)\mathbb{Z}.$$

In other words, $\mathfrak{C}_{vu}$ is the quotient of the integer lattice $\mathbb{Z}^2$ modulo a shifting action which sends (i, j) to $(i-u, j+v)$. For $(i, j) \in \mathbb{Z}^2$, we let $\langle i, j\rangle = (i, j) + (-u, v)\mathbb{Z}$ denote the corresponding element of $\mathfrak{C}_{vu}$. $\mathfrak{C}_{vu}$ inherits a natural partial order $\leq_{\mathfrak{C}}$ from $\mathbb{Z}^2$ which is generated by the relations $\langle i, j\rangle <_{\mathfrak{C}} \langle i+1, j\rangle$ and $\langle i, j\rangle <_{\mathfrak{C}} \langle i, j+1\rangle$.

Note that this partial order is antisymmetric since u and v are positive. Recall that a subposet Q of a poset P is said to be *convex* if, for all elements $x < y < z$ in P, we have $y \in Q$ whenever we have $x, z \in Q$.

Definition 2.1. A *cylindric skew shape* is a finite convex subposet of the poset $\mathfrak{C}_{vu}$.

Example 2.2. We can regard skew shapes λ/μ as a special case of cylindric skew shapes. Suppose λ/μ fits inside a box of height v and width u. We embed λ/μ in $\mathfrak{C}_{vu}$ by mapping the box in the ith row and jth column of λ/μ to $\langle i, j\rangle$. Figure 2 shows the resulting image of λ/μ in $\mathbb{Z}^2$, with one representative of λ/μ shown in bold. Notice that elements of different representatives of λ/μ are always incomparable in $\mathbb{Z}^2$. Of course, we could also embed λ/μ in $\mathfrak{C}_{v'u'}$ where $v' \geq v$ and $u' \geq u$.

Example 2.3. The class of *cylindric ribbons* will play an important role and they are defined in the analogous way to ribbons in the classical case. As we just did for skew shapes, we will identify any cylindric skew shape with its corresponding set of boxes in $\mathbb{Z}^2$. Note that the skew shapes from the previous example can be edgewise connected when viewed as subsets of $\mathfrak{C}_{vu}$. However, they are not edgewise connected when viewed as subsets of $\mathbb{Z}^2$, as in the figure.

Definition 2.4. A *cylindric ribbon* is a cylindric skew shape which, when viewed as a subset $\mathbb{Z}^2$, is edgewise connected and contains no 2×2 block of boxes.

The cylindric skew shape in Figure 1(b) is an example of a cylindric ribbon once we delete the boxes filled with 6's.

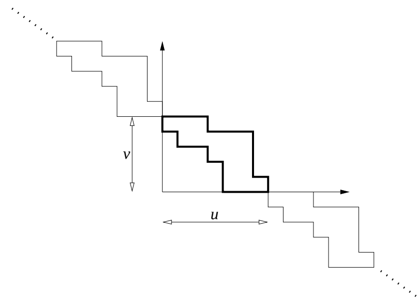


FIGURE 2. Skew shapes are cylindric skew shapes

Suppose C is a cylindric skew shape which is a subposet of the cylinder $\mathfrak{C}_{vu}$. Let us define what we mean by the rows and columns of C. The *p-th row* is the set $\{\langle i,j\rangle \in C \mid j=p\}$ and the *q-th column* is the set $\{\langle i,j\rangle \in C \mid i=q\}$. So the rows only depend on $p \bmod v$ and the columns only depend on $q \bmod u$. Thus the cylinder $\mathfrak{C}_{vu}$ has exactly v rows and u columns.

Definition 2.5. For a cylindric skew shape C, a *semistandard cylindric tableau* of shape C is a map $T : C \to \mathbb{P}$ that weakly increases in the rows of C and strictly increases in the columns.

See Figure 1(b) for an example. We are now ready to define our main object of study.

Definition 2.6. For a cylindric skew shape C, the *cylindric skew Schur function $s_C(x)$* in the variables $x = (x_1, x_2, \ldots)$ is defined by

$$s_C(x) = \sum_T \prod_{c\in C} x_{T(c)} = \sum_T x_1^{\#T^{-1}(1)} x_2^{\#T^{-1}(2)} \cdots ,$$

where the sums are over all semistandard cylindric tableaux T of shape C.

The terminology "cylindric skew Schur function" is partially justified by the following two observations:

Example 2.7. Because of Example 2.2, skew Schur functions and, in particular, Schur functions are all examples of cylindric skew Schur functions.

Theorem 2.8. *For any cylindric skew shape C, $s_C(x)$ is a symmetric function.*

We omit the proof as it is basically the same as the proof from [1], which also appears as [13, Theorem 7.10.2], that the skew Schur function $s_{\lambda/\mu}(x)$ is symmetric.

3. Cylindric skew Schur functions from Gromov-Witten invariants

As mentioned in the introduction, there is a good reason for recent interest in cylindric skew Schur functions. This motivation is centred around the main result of [12]. A nice introduction, with emphasis on the context and the importance of Postnikov's result can be found in [14]. Here, however, we merely extract from these two references the minimum amount of background necessary to show how Postnikov's work ties together cylindric skew Schur functions and an open problem of considerable interest.

Given k and n with $n > k \geq 1$, we let Gr_{kn} denote the manifold of k-dimensional subspaces of $\mathbb{C}^n$. Gr_{kn} is a complex projective variety known as the *Grassmann variety* or *Grassmannian.* For a partition λ, we will write $\lambda \subseteq k\times(n-k)$ if the Young diagram for λ has at most k rows and at most $n-k$ columns. In this case, we let $\lambda^\vee$ denote the partition $(n-k-\lambda_k, \ldots, n-k-\lambda_1)$. Given $\lambda, \mu, \nu \subseteq k \times (n-k)$, we let $C_{\mu\nu}^{\lambda,d}$ denote the *(3-point) Gromov-Witten invariant*, defined geometrically as the number of rational curves of degree d in Gr_{kn} that meet fixed generic translates of the Schubert varieties $\Omega_{\lambda^\vee}, \Omega_\mu$ and Ω_ν, provided that this number is finite. This last condition implies that $C_{\mu\nu}^{\lambda,d}$ is defined if $|\mu|+|\nu| = nd+|\lambda|$, and otherwise we set $C_{\mu\nu}^{\lambda,d} = 0$. If $d = 0$, then a degree 0 curve is just a point in Gr_{kn} and we get the geometric interpretation of the Littlewood-Richardson coefficient $c_{\mu\nu}^{\lambda} = C_{\mu\nu}^{\lambda,0}$. While we do not claim

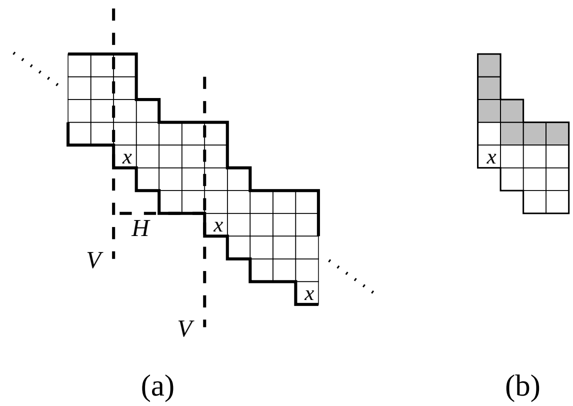


FIGURE 3. Describing C as $\lambda/d/\mu$

that this paragraph is sufficient to give a firm understanding of $C_{\mu\nu}^{\lambda,d}$, we do claim that it is clear from this geometric definition that $C_{\mu\nu}^{\lambda,d} \geq 0$. No algebraic or combinatorial proof of this inequality is known and, as stated in [14], it is a fundamental open problem to find such a proof.

Postnikov's result shows that the Gromov-Witten invariants $C_{\mu\nu}^{\lambda,d}$ appear as the coefficients when we expand certain cylindric skew Schur functions in terms of Schur functions. It follows that improving our understanding of this expansion could lead to a solution of the open problem.

Before stating his result, we need to introduce some notation that will allow us to write any cylindric skew shape in the form $\lambda/d/\mu$, where λ and μ are partitions and $d \in \mathbb{N}$. From this point on, unless otherwise stated, all of our cylindric skew shapes C will be subposets of the cylinder $\mathfrak{C}_{vu}$ with $v = k$ and $u = n - k$.

Suppose we are given any cylindric skew shape C. The process for finding λ, d and μ is best understood from a figure, and we will use the cylindric skew shape shown in Figure 3(a) as a running example. The boxes labelled x are identified, so that $k = 3$ and $n - k = 4$ in this example. First, we must choose a set of representatives for the elements of C. A convenient way to do this is to take the elements between two adjacent representatives of a vertical line V. Now draw a horizontal line segment H running below our representatives of C. We regard the intersection of V and the left end of H in Figure 3(a) as our origin. The partition μ is now the partition whose Young diagram is outlined by H, V and the lower boundary of C. In our example, $\mu = (2, 1)$.

Next, consider just our set of representatives for the elements of C as in Figure 3(b). Define a partition Λ by supposing the resulting skew shape is Λ/μ. Therefore, in our example, $\Lambda = (4, 4, 4, 4, 2, 1, 1)$. If $\Lambda \subseteq k \times (n-k)$ then set $d = 0$, $\lambda = \Lambda$ and we are done. Otherwise, let $\Lambda[-1]$ denote the unique partition ν that makes Λ/ν an n-ribbon with $n - k$ non-empty columns. In other words, $\Lambda[-1]$ is obtained by removing an n-ribbon along the top of Λ, starting in Λ's leftmost column and ending in Λ's rightmost column. It is not difficult to see that such a ribbon always has $k+1$ non-empty rows. In our example, we remove the shaded boxes in Figure 3(b) and $\Lambda[-1] = (4, 4, 4, 1)$. We can see that $\Lambda[-1]$ is well-defined by referring back to Figure 3(a). Effectively what we are doing is removing the cylindric ribbon that runs all the way along the top of C. We see that this cylindric ribbon must have n elements.

Now if $\Lambda[-1] \subseteq k \times (n-k)$, then we set $d = 1$ and $\lambda = \Lambda[-1]$. Otherwise, obtain $\Lambda[-2]$ from $\Lambda[-1]$ in the same way that $\Lambda[-1]$ was obtained from Λ: remove an n-ribbon from the top of $\Lambda[-1]$, starting in the leftmost column and ending in the rightmost column. Repeating this procedure, we can construct $\Lambda[-e]$, stopping as soon as $\Lambda[-e] \subseteq k \times (n-k)$. We then set $d = e$ and $\lambda = \Lambda[-e]$. In our example, we see that $\Lambda[-2] = (3, 3) \subseteq k \times (n-k)$ and so $d = 2$, $\lambda = (3, 3)$ and $\lambda/d/\mu = (3, 3)/2/(2, 1)$.

Remark 3.1. There are several things to note about $\lambda/d/\mu$:

(i) For a given C, $\lambda/d/\mu$ is clearly not unique and depends on our choice of origin.

(ii) μ is not necessarily contained in λ. For example, moving our origin 1 square down and 1 square to the left, the reader is encouraged to verify that $\mu = (4,3,2,1)$, $\Lambda = (4,4,4,4,4,3,2,2)$ and that $\lambda/d/\mu = (3,3)/3/(4,3,2,1)$.
(iii) We always have $\lambda \subseteq k \times (n-k)$ and it is always possible to choose our origin so that $\mu \subseteq k \times (n-k)$.
(iv) We could alternatively have defined λ by saying it is the *n-core* of Λ, where the n-core is defined in the following manner. Given a partition τ, successively remove n-ribbons from τ so that after each ribbon removal, the resulting shape is a partition. Stop when no more n-ribbons can be removed. It is a well-known fact (see, for example, [8, I.1, Example 8]) that the resulting partition λ is independent of the choice of ribbons removed, and λ is said to be the *n-core* of τ.
(v) Our notation $\lambda/d/\mu$ is equivalent to that in [12], but our explanation of it is very different. We choose this description in terms of removal of ribbons because it will be useful in later sections.

For any formal power series f in the variables $x = (x_1, x_2, \ldots)$, we will write $f(x_1, \ldots, x_k)$ to denote the specialization $f(x_1, x_2, \ldots, x_k, 0, 0, \ldots)$. We are finally ready to state [12, Theorem 6.3].

Theorem 3.2. *For any two partitions $\lambda, \mu \subseteq k \times (n-k)$ and a non-negative integer d, we have*

$$s_{\lambda/d/\mu}(x_1, \ldots, x_k) = \sum_{\nu \subseteq k \times (n-k)} C_{\mu\nu}^{\lambda,d} s_\nu(x_1, \ldots, x_k). \tag{3.1}$$

Since we are restricting to k variables, the left-hand side is a sum over semistandard cylindric tableaux T that map $\lambda/d/\mu$ to the set $[k]$. Since T must increase in the columns of $\lambda/d/\mu$, this implies that $s_{\lambda/d/\mu}(x_1, x_2, \ldots, x_k)$ is non-zero only if all the columns of $\lambda/d/\mu$ contain at most k elements. One can check that this is equivalent to all the rows of $\lambda/d/\mu$ containing at most $n-k$ elements. In this case, we follow Postnikov in saying that $\lambda/d/\mu$ is a *toric* shape. While we take this opportunity to note that toric shapes are the shapes that are most relevant to the Gromov-Witten invariants, we will continue to work with general cylindric skew shapes.

Since will be mostly interested in the case of infinitely many variables $x = (x_1, x_2, \ldots)$, we make a few quick remarks about $s_{\lambda/d/\mu}(x)$ and $s_{\lambda/d/\mu}(x_1, \ldots, x_k)$. First, since all the entries in any column of a semistandard cylindric tableau are distinct, the monomial $x_1^{a_1} x_2^{a_2} \cdots$ appears with coefficient 0 in $s_{\lambda/d/\mu}(x)$ if $a_i > n-k$ for some i. It follows that we have the useful fact that

$$s_{\lambda/d/\mu}(x) = \sum_\nu c_\nu s_\nu(x) = \sum_{\nu:\nu_1 \leq n-k} c_\nu s_\nu(x). \tag{3.2}$$

From this, we conclude

$$s_{\lambda/d/\mu}(x_1, \ldots, x_k) = \sum_{\nu:l(\nu) \leq k} c_\nu s_\nu(x_1, \ldots, x_k) = \sum_{\nu \subseteq k \times (n-k)} c_\nu s_\nu(x_1, \ldots, x_k),$$

explaining why the sum in (3.1) is only over $\nu \subseteq k \times (n-k)$. Finally, we note that $s_{\lambda/d/\mu}(x_1, \ldots, x_k)$ is essentially obtained from $s_{\lambda/d/\mu}(x)$ by removing all those terms involving s_ν with $l(\nu) > k$. In fact, in the sections that follow, we will be focusing most of our attention on these terms s_ν with $l(\nu) > k$.

Since we know from the geometric definition of Gromov-Witten invariants that $C_{\mu\nu}^{\lambda,d} \geq 0$, we conclude from (3.1) that $s_{\lambda/d/\mu}(x_1, \ldots, x_k)$ is Schur-positive. On the other hand, we observe that $s_{\lambda/d/\mu}(x)$ may not be Schur-positive. For example, when $k = n-k = 2$ and $\lambda/d/\mu = (1,0)/1/(1,0)$,

$$s_{\lambda/d/\mu} = m_{22} + 2m_{211} + 4m_{1111} = s_{22} + s_{211} - s_{1111}.$$

In the next section, we answer the following question:

Question 3.3. *For what cylindric skew shapes C is $s_C(x)$ Schur-positive?*

4. Schur-positivity

We saw in Example 2.2 that the skew shape λ/μ can be regarded as a cylindric skew shape C when λ/μ fits inside a box of height k and width $n-k$. In this case, we then know that s_C is Schur-positive. The following theorem, which is the main result of this section, states that these are the only

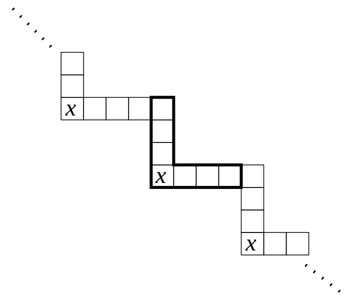


FIGURE 4. $H_{4,3}$

Schur-positive cylindric skew Schur functions. To state the theorem, we need to define what it means for cylindric skew shapes to be isomorphic. Recall that every cylindric skew shape can be viewed as a subposet of $\mathfrak{C}_{vu}$, for some positive numbers v and u. For cylindric skew shapes to be isomorphic, we certainly need the corresponding subposets to be isomorphic, but we also need to preserve certain "strictness" relations in the columns. Therefore, we will say that cylindric skew shapes C_1 and C_2 are isomorphic if there exists a poset isomorphism $f : C_1 \to C_2$ such that, for $x, y \in C_1$, x and y are in the same column of C_1 if and only if $f(x)$ and $f(y)$ are in the same column of C_2.

Theorem 4.1. *Let C be a cylindric skew shape. Then $s_C(x)$ is Schur-positive if and only if C is isomorphic to a skew shape.*

In other words, s_C is never Schur-positive except in the trivial case of C being a skew shape.

Let us say as a few words about the proof of this result. The first of two main steps is to prove it to be true for cylindric ribbons. The second step involves using the outer coproduct for the ring of quasisymmetric functions, as defined in [9, 10], to deduce the result for general cylindric skew shapes.

Cylindric ribbons are also interesting in their own right, and they will be our next topic of discussion. We begin with a special class.

Example 4.2. A cylindric ribbon is said to be a *cylindric hook* if it has a unique minimal element (when viewed as a subposet of $\mathfrak{C}_{k,n-k}$). See Figure 4 for an example. We see that, unlike hooks in the classical case, cylindric hooks have just one maximal element. Also note that $\mathfrak{C}_{k,n-k}$ has just one cylindric hook as a subposet, up to isomorphism. We denote this cylindric hook by $H_{k,n-k}$. A cylindric hooks is the simplest example of a cylindric skew shape that is not toric. It follows that $s_{H_{k,n-k}}(x_1, \ldots, x_k) = 0$. This is also evident in the following result which shows that the Schur expansion of $s_{H_{k,n-k}}(x)$ is a nice alternating sum of Schur functions of hooks.

Lemma 4.3. *With all functions in the variables $x = (x_1, x_2, \ldots)$, we have*

$$s_{H_{k,n-k}} = s_{(n-k,1^k)} - s_{(n-k-1,1^{k+1})} + \cdots + (-1)^{n-k-2} s_{(2,1^{n-2})} + (-1)^{n-k-1} s_{(1^n)}.$$

Using this result, we can now describe the Schur expansions of general cylindric ribbons.

Proposition 4.4. *Let C by a cylindric ribbon which is a subposet of the cylinder $\mathfrak{C}_{k,n-k}$. Then*

$$s_C(x) = \left(\sum_{\nu \subseteq k \times (n-k)} c_\nu s_\nu(x) \right) + s_{H_{k,n-k}}(x),$$

with c_ν a non-negative integer for all $\nu \subseteq k \times (n-k)$.

Let C be a cylindric skew shape that is not isomorphic to a skew shape. We know from Theorem 3.2 that s_C in k variables is Schur-positive. On the other hand, by Theorem 4.1, s_C in an infinite number of variables is not Schur-positive. We conclude this section with a discussion of the minimum number of variables in which s_C fails to be Schur-positive.

If C is a cylindric ribbon, we deduce from Proposition 4.4 and Lemma 4.3 that s_C remains Schur-positive in $k+1$ variables but always fails to be Schur-positive in $k+2$ variables. By looking at coproducts, we can use this fact to say something about a general cylindric skew shape C. Let $C[-1]$ denote the cylindric skew shape that results when we remove the cylindric ribbon that runs all the way along the top of C.

Proposition 4.5. *Let C be a cylindric skew shape that is not isomorphic to a skew shape and that is a subposet of $\mathfrak{C}_{k,n-k}$. If m denotes the maximum number of elements in a column of $C[-1]$, then s_C is not Schur-positive in $m+k+2$ variables.*

We do not claim, and it is not true, that $m+k+2$ is the best possible value. In other words, it can be the case that s_C is not Schur-positive in some number of variables that is less than $m+k+2$. For toric shapes, it is clear that $m \leq k-1$, and so we get the following result.

Corollary 4.6. *Let C be a toric shape that is not isomorphic to a skew shape and that is a subposet of $\mathfrak{C}_{k,n-k}$. Then s_C is not Schur-positive in $2k+1$ variables.*

5. From cylindric skew shapes to skew shapes

So far, we have not discussed any tools for dealing with cylindric skew Schur functions. The subject of this section is a rule for expressing any cylindric skew Schur function as a signed sum of skew Schur functions. Our rule is based on a result of Gessel and Krattenthaler from [5], with our reformulation modelled on a result from [2]. We begin with an exposition of these two results, starting with the latter.

By saying that a partition τ is obtained from λ by adding d n-ribbons, we mean that there is a sequence of partitions

$$\lambda = \nu_0 \subseteq \nu_1 \subseteq \cdots \subseteq \nu_d = \tau \tag{5.1}$$

such that ν_i/ν_{i-1} is an n-ribbon for $i = 1, \ldots, d$. We say that the *width* of a ribbon is its number of non-empty columns. If $\tau_1 \leq n-k$, then we define

$$\varepsilon(\tau/\lambda) = (-1)^{\sum_{i=1}^{d}(n-k-\mathrm{width}(\nu_i/\nu_{i-1}))}.$$

It can be shown that $\varepsilon(\tau/\lambda)$ is independent of the choice of the sequence in (5.1).

The result of interest from [2] is the following:

Theorem 5.1. *Suppose we have $\lambda, \mu, \nu \subseteq k \times (n-k)$ with $|\mu| + |\nu| = |\lambda| + dn$ for some $d \geq 0$. Then the Gromov-Witten invariant $C_{\mu\nu}^{\lambda,d}$ can be expressed in terms of Littlewood-Richardson coefficients as*

$$C_{\mu\nu}^{\lambda,d} = \sum_{\tau} \varepsilon(\tau/\lambda) c_{\mu\nu}^{\tau}, \tag{5.2}$$

where the sum is over all τ with $\tau_1 \leq n-k$ that can be obtained from λ by adding d n-ribbons.

Formulas for $C_{\mu\nu}^{\lambda,d}$ similar to (5.2) have appeared in different contexts in [4, 6, 7, 16]. Multiplying both sides of (5.2) by $s_\nu(x_1, \ldots, x_k)$, summing over all $\nu \subseteq k \times (n-k)$, and applying Theorem 3.2, we get:

Corollary 5.2. *For any cylindric skew shape $\lambda/d/\mu$ with $\lambda, \mu \subseteq k \times (n-k)$, we have*

$$s_{\lambda/d/\mu}(x_1, \ldots, x_k) = \sum_{\tau} \varepsilon(\tau/\lambda) s_{\tau/\mu}(x_1, \ldots, x_k), \tag{5.3}$$

where the sum is over all τ with $\tau_1 \leq n-k$ that can be obtained from λ by adding d n-ribbons.

From our point of view, the obvious disadvantage of Corollary 5.2 is that it only gives certain terms in the expansion of $s_{\lambda/d/\mu}(x)$. For example, for cylindric shapes that are not toric, both sides of (5.3) will be zero. Gessel and Krattenthaler's setting does not have this limitation. To apply their result to get an expression for $s_{\lambda/d/\mu}(x)$, we first have some work to do. Their basic result, Proposition 1, is stated in terms of lattice paths. In their Section 9, they show how to apply Proposition 1 to obtain expressions for Schur functions. Mimicking their approach, we first obtain an expression for $s_{\lambda/d/\mu}$ in terms of the elementary symmetric functions. Recall from page 5 that, for a given $\lambda/d/\mu$, Λ is the

r	$\Lambda' + rn$	resulting partition	sign
$(0,0,0,0)$	$(7,5,4,4)$	$(7,5,4,4)$	$+$
$(-1,0,0,1)$	$(0,5,4,11)$	$(8,5,4,3)$	$-$
$(-1,0,1,0)$	$(0,5,11,4)$	$(9,5,3,3)$	$+$
$(-1,1,0,0)$	$(0,12,4,4)$	$(11,3,3,3)$	$-$
$(0,-1,0,1)$	$(7,-2,4,11)$	$(8,8,4,0)$	$+$
$(0,-1,1,0)$	$(7,-2,11,4)$	$(9,8,3,0)$	$-$
$(1,-1,0,0)$	$(14,-2,4,4)$	$(14,3,3,0)$	$+$
$(-1,-1,1,1)$	$(0,-2,11,11)$	$(9,9,2,0)$	$+$
$(-1,-1,0,2)$	$(0,-2,4,18)$	$(15,3,2,0)$	$-$
$(-1,-1,2,0)$	$(0,-2,18,4)$	$(16,2,2,0)$	$+$

Table 1. Applying Theorem 5.3

unique partition satisfying $\Lambda[-d] = \lambda$. In this case, we also write $\lambda[d] = \Lambda$ and we see that Λ is obtained from λ by adding d n-ribbons, each starting in $\lambda's$ rightmost column (column $n-k$) and ending in column 1. We get that

$$s_{\lambda/d/\mu}(x) = \sum_{\substack{r_1+\cdots+r_{n-k}=0\\ r_i\in\mathbb{Z}}} \det\left(e_{r_s n+\Lambda'_s-\mu'_t-s+t}(x)\right)_{s,t=1}^{n-k}, \tag{5.4}$$

where, as usual, we set $e_0 = 1$ and $e_i = 0$ for $i < 0$. We wish to simplify this expression using the dual Jacobi-Trudi identity and, to do so, we must introduce the modification rule

$$s_{(\tau_1,\ldots,\tau_i,\tau_{i+1},\ldots,\tau_{n-k})'/\mu}(x) = -s_{(\tau_1,\ldots,\tau_{i-1},\tau_{i+1}-1,\tau_i+1,\tau_{i+2},\ldots,\tau_{n-k})'/\mu}(x), \tag{5.5}$$

which allows us to interpret $s_{\tau'/\mu}$ when $\tau = (\tau_1,\ldots,\tau_{n-k})$ is not necessarily a partition containing μ. For example,

$$s_{(7,-2,11,4)'/(2,1)} = -s_{(7,10,-1,4)'/(2,1)} = s_{(7,10,3,0)'/(2,1)} = -s_{(9,8,3,0)'/(2,1)}.$$

On the other hand,

$$s_{(-7,5,4,18)'/(2,1)} = 0,$$

since no number of applications of the rule (5.5) will result in a skew Schur function. We can further simplify (5.4) by letting $\Lambda' + rn$ denote the integer sequence $(\Lambda'_1 + r_1 n, \ldots, \Lambda'_{n-k} + r_{n-k} n)$.

Putting this all together, (5.4) becomes

Theorem 5.3. [5] *For any cylindric shape $\lambda/d/\mu$ that is a subposet of $\mathfrak{C}_{k,n-k}$, we have*

$$s_{\lambda/d/\mu}(x) = \sum_{\substack{r_1+\cdots+r_{n-k}=0\\ r_i\in\mathbb{Z}}} s_{(\Lambda'+rn)'/\mu}(x) \tag{5.6}$$

where $\Lambda = \lambda[d]$ and the right-hand side is interpreted in terms of the modification rule (5.5).

Example 5.4. Consider $\lambda/d/\mu = (3,3)/2/(2,1)$ as depicted in Figure 3. We see that $n = 7$, $n-k = 4$, $\Lambda' = (7,5,4,4)$ and $\mu = (2,1,0,0)$. The values of $r = (r_1,\ldots,r_{n-k})$ that make $s_{(\Lambda'+rn)'/\mu} \neq 0$ are listed in the first column of Table 1. We conclude that

$$\begin{aligned} s_{(3,3)/2/(2,1)}(x) &= s_{(7,5,4,4)'/(2,1)}(x) - s_{(8,5,4,3)'/(2,1)}(x) + s_{(9,5,3,3)'/(2,1)}(x)\\ &\quad -s_{(11,3,3,3)'/(2,1)}(x) + s_{(8,8,4,0)'/(2,1)}(x) - s_{(9,8,3,0)'(2,1)}(x)\\ &\quad +s_{(14,3,3,0)'/(2,1)}(x) + s_{(9,9,2,0)'/(2,1)}(x) - s_{(15,3,2,0)'/(2,1)}(x)\\ &\quad +s_{(16,2,2,0)'/(2,1)}(x). \end{aligned}$$

Using Theorem 5.3, we can actually show that Corollary 5.2 extends to the case of infinitely many variables. This is the main result of this section.

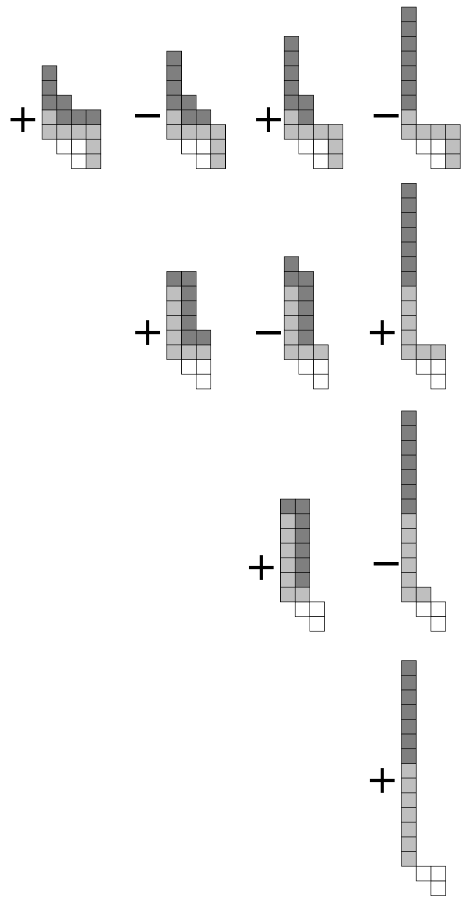


FIGURE 5.

Theorem 5.5. *For any cylindric skew shape $\lambda/d/\mu$ that is a subposet of $\mathfrak{C}_{k,n-k}$, we have*

$$s_{\lambda/d/\mu}(x) = \sum_{\tau} \varepsilon(\tau/\lambda) s_{\tau/\mu}(x), \tag{5.7}$$

where the sum is over all τ with $\tau_1 \leq n-k$ that can be obtained from λ by adding d n-ribbons.

Note 5.6. While $\lambda \subseteq k \times (n-k)$ by definition, we do not require that $l(\mu) \leq k$, unlike in Theorem 5.1 and Corollary 5.2.

It does not seem that the proof of Theorem 5.1 from [2] can be easily modified to prove Theorem 5.5, and our proof of Theorem 5.5 is somewhat technical.

Example 5.7. Again, consider $\lambda/d/\mu = (3,3)/2/(2,1)$ as depicted in Figure 3. Figure 5 shows the set of all possible $\varepsilon(\tau/\lambda)\tau/\mu$ with $\tau_1 \leq n-k$ such that τ can be obtained from $(3,3)$ by adding 2 7-ribbons. The positioning of the partitions in the figure is supposed to be helpful, as it is determined by the rightmost column of the added ribbons. There can be more than one way to add ribbons to λ and get a particular τ, but this does not affect our expression for $s_{\lambda/d/\mu}$.

We see that we get the same result as in Example 5.4. While the result obtained from Theorem 5.3 is more compact to write, we find the graphical description of $s_{\lambda/d/\mu}$ in Theorem 5.5 preferable, especially from the point of view of intuition. We make much use of Theorem 5.5 in proving the results of the next section.

Remark 5.8. Because the expression of a cylindric skew shape C in the form $\lambda/d/\mu$ is not unique, Theorem 5.5 can be used to give a host of identities among skew Schur functions. For example, consider

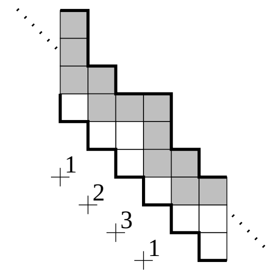


FIGURE 6.

the cylindric skew shape C shown in Figure 6 with $k = n - k = 3$. By choosing the origins labelled 1, 2 and 3 respectively, we see that C can be written as $(3,3,1)/1/(2,1)$, $(3,2,2)/1/(2,1)$ or $(1)/2/(2,1)$. It follows that

$$\begin{aligned} s_C(x) &= s_{333211/21} - s_{3322111/21} + s_{331111111/21} \\ &= s_{33331/21} - s_{32221111/21} + s_{322111111/21} \\ &= s_{33322/21} - s_{3222211/21} + s_{3211111111/21} + s_{2222221/21} - s_{22111111111/21}. \end{aligned}$$

6. CYLINDRIC SCHUR-POSITIVITY

Before presenting the conjecture which is the main subject of this section, we begin with a relevant application of Theorem 5.5.

In the same way that Schur functions are those skew Schur functions $s_{\lambda/\mu}(x)$ with $\mu = \emptyset$, we will say that *cylindric Schur functions* are those cylindric skew Schur functions $s_{\lambda/d/\mu}(x)$ with $\mu = \emptyset$. While the Schur functions are known to be a basis for the symmetric functions, we have the following result for the cylindric Schur functions.

Proposition 6.1. *For a given $k, n-k$, the cylindric Schur functions of the form $s_{\lambda/d/\emptyset}(x)$, with $\lambda/d/\emptyset$ a subposet of $\mathfrak{C}_{k,n-k}$, are linearly independent.*

We might next ask if every cylindric skew Schur function $s_{\lambda/d/\mu}(x)$ with $\lambda/d/\mu$ a subposet of $\mathfrak{C}_{k,n-k}$ can be expressed as a linear combination of cylindric Schur functions of the form $s_{\nu/e/\emptyset}(x)$, where each $\nu/e/\emptyset$ is also a subposet of $\mathfrak{C}_{k,n-k}$. As we shall see, an affirmative answer to this question would also imply Conjecture 6.3 below.

Definition 6.2. Suppose $\lambda/d/\mu$ is a cylindric skew shape that is a subposet of $\mathfrak{C}_{k,n-k}$. We say that $s_{\lambda/d/\mu}(x)$ in the variables $x = (x_1, x_2, \ldots)$ is *cylindric Schur-positive* if it can be expressed as a linear combination of cylindric Schur functions $s_{\nu/e/\emptyset}(x)$ with positive coefficients, where each such $\nu/e/\emptyset$ is also a subposet of $\mathfrak{C}_{k,n-k}$.

As an analogue of the fact that every skew Schur function is Schur-positive, we propose the following conjecture.

Conjecture 6.3. *Every cylindric skew Schur function is cylindric Schur-positive.*

Proposition 4.4 implies this conjecture is true for cylindric ribbons. The rest of this section will be devoted to other evidence in favour of the conjecture.

It follows from (3.2) that we can split $s_{\lambda/d/\mu}(x)$ into two sums as follows:

$$s_{\lambda/d/\mu}(x) = \sum_{\nu \subseteq k \times (n-k)} a_\nu s_\nu(x) + \sum_{\substack{\nu : \nu_1 \leq n-k \\ l(\nu) > k}} b_\nu s_\nu(x). \tag{6.1}$$

When $\nu \subseteq k \times (n-k)$, we know that $s_\nu(x)$ is a cylindric Schur function. Furthermore, we know from Theorem 3.2 that $a_\nu \geq 0$ for all $\nu \subseteq k \times (n-k)$. Therefore, the first sum is cylindric Schur-positive.

Now consider the second sum, which we denote by $B(\lambda/d/\mu, x)$. We know that $s_{\lambda/d/\mu}(x)$ is cylindric Schur-positive when $d = 0$. Therefore, we can assume by induction that $s_{\lambda/(d-1)/\mu}(x)$ is cylindric Schur-positive:

$$s_{\lambda/(d-1)/\mu}(x) = \sum_{\substack{\nu,e \\ \nu\subseteq k\times(n-k)}} c_{\nu,e} s_{\nu/e/\emptyset}(x), \tag{6.2}$$

where $c_{\nu,e} \geq 0$ for all ν, e, and e is a always non-negative integer. (For $s_{\nu/e/\emptyset}(x) \neq 0$, we require that $ne = |\lambda| - |\mu| + n(d-1) - |\nu|$.) We conjecture, in fact, that $B(\lambda/d/\mu, x)$ can be expressed exactly in terms of $s_{\lambda/(d-1)/\mu}(x)$ as:

$$B(\lambda/d/\mu, x) = \sum_{\substack{\nu,e \\ \nu\subseteq k\times(n-k)}} c_{\nu,e} s_{\nu/e+1/\emptyset}(x).$$

Plugging this into (6.1), we get

$$s_{\lambda/d/\mu}(x) = \sum_{\nu\subseteq k\times(n-k)} a_\nu s_\nu(x) + \sum_{\substack{\nu,e \\ \nu\subseteq k\times(n-k)}} c_{\nu,e} s_{\nu/e+1/\emptyset}(x), \tag{6.3}$$

where $a_\nu, c_{\nu,e} \geq 0$ for all ν, e. This expression is a strong refinement of Conjecture 6.3 as it gives much information about the form of the cylindric Schur-positive expansion of $s_{\lambda/d/\mu}(x)$. Using [3, 15], we have verified (6.3) for all $\lambda/d/\mu$ with $k, n-k, d \leq 5$.

As promised, we can also reformulate Conjecture 6.3 into a seemingly easier statement.

Theorem 6.4. *Conjecture 6.3 holds if and only if every cylindric skew Schur function $s_{\lambda/d/\mu}(x)$ with $\lambda/d/\mu$ a subposet of $\mathfrak{C}_{k,n-k}$ can be expressed as a linear combination of cylindric Schur functions $s_{\nu/e/\emptyset}(x)$, where each $\nu/e/\emptyset$ is also a subposet of $\mathfrak{C}_{k,n-k}$.*

In other words, to prove Conjecture 6.3, we don't have to show that the coefficients are positive.

References

[1] Edward A. Bender and Donald E. Knuth. Enumeration of plane partitions. *J. Combinatorial Theory Ser. A*, 13:40–54, 1972.
[2] Aaron Bertram, Ionuţ Ciocan-Fontanine, and William Fulton. Quantum multiplication of Schur polynomials. *J. Algebra*, 219(2):728–746, 1999.
[3] Anders S. Buch. Littlewood-Richardson calculator, 1999. Available from `http://home.imf.au.dk/abuch/lrcalc/`.
[4] C. J. Cummins. su(n) and sp($2n$) WZW fusion rules. *J. Phys. A*, 24(2):391–400, 1991.
[5] Ira M. Gessel and Christian Krattenthaler. Cylindric partitions. *Trans. Amer. Math. Soc.*, 349(2):429–479, 1997.
[6] Frederick M. Goodman and Hans Wenzl. Littlewood-Richardson coefficients for Hecke algebras at roots of unity. *Adv. Math.*, 82(2):244–265, 1990.
[7] Victor G. Kac. *Infinite-dimensional Lie algebras*. Cambridge University Press, Cambridge, third edition, 1990.
[8] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995.
[9] Claudia Malvenuto. *Produits et coproduits des fonctions quasi-symétriques et de l'algèbre des descentes*, volume 16 of *Publications du Laboratoire de Combinatoire et d'Informatique Mathématique*. Laboratoire de Combinatoire et d'Informatique Mathématique, Université du Québec à Montréal, 1993. Ph.D. thesis.
[10] Claudia Malvenuto and Christophe Reutenauer. Duality between quasi-symmetric functions and the Solomon descent algebra. *J. Algebra*, 177(3):967–982, 1995.
[11] Peter McNamara. Cylindric skew Schur functions. Submitted, 2004. `http://www.arxiv.org/abs/math.CO/0410301`.
[12] Alexander Postnikov. Affine approach to quantum Schubert calculus. *Duke Math. J.*, to appear. `http://www.arxiv.org/abs/math.CO/0205165`.
[13] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.
[14] Richard P. Stanley. Recent developments in algebraic combinatorics. *Israel J. Math.*, 143:317–339, 2004.
[15] John R. Stembridge. SF, posets and coxeter/weyl. Available from `http://www.math.lsa.umich.edu/~jrs/maple.html`.
[16] Mark A. Walton. Fusion rules in Wess-Zumino-Witten models. *Nuclear Phys. B*, 340(2-3):777–790, 1990.

Laboratoire de Combinatoire et d'Informatique Mathématique (LaCIM), Université du Québec à Montréal, Case Postale 8888, succursale Centre-ville, Montréal (Québec) H3C 3P8, Canada
E-mail address: `mcnamara@lacim.uqam.ca`

Counting non-equivalent coverings and non-isomorphic maps for Riemann surfaces *

Alexander Mednykh

Sobolev Institute of Mathematics, Novosibirsk State University,

630090, Novosibirsk, Russia

mednykh@math.nsc.ru

Abstract

The main result of the paper is a new formula for the number of conjugacy classes of subgroups of a given index in a finitely generated group. As application of this result a simple proof of the formula for the number of non-equivalent coverings over surface (orientable or not, bordered or not) is given. Another application is a formula for the number of non-isomorphic unrooted maps on an orientable closed surface with a given number of edges.

Keywords: *number of subgroups, conjugacy class of subgroups, surface coverings, unrooted maps*

1 Introduction

Let $M_\Gamma(n)$ denote the number of subgroups of index n in a group Γ, and $N_\Gamma(n)$ be the number of conjugacy classes of such groups. The last function counts the isomorphism classes of transitive permutation representations of degree n of Γ and hence, also the equivalence classes of n-fold unbranched connected coverings of a topological space with fundamental group Γ.

M. Hall [4] determined the numbers of subgroups $M_\Gamma(n)$ for a free group $\Gamma = \mathrm{F}_r$ of rank r. Later V. Liskovets [9] developed a new method for calculation of $N_\Gamma(n)$ for the same group. Both functions $M_\Gamma(n)$ and $N_\Gamma(n)$ for the fundamental group Γ of a closed surface were obtained in [15] and [16] for orientable and non-orientable surfaces, respectively. See also [17] and [3] for the case of the fundamental group of the Klein bottle and a survey [8] for related problems. In all these cases the problem of calculating of $M_\Gamma(n)$ was solved essentially by using representation theory of symmetric groups, contributed by Hurwitz and Frobenius, as the main tool ([6], [7]). The solution for the problem to finding $N_\Gamma(n)$ was based on the further development of the Liskovets

*Supported by the RFBR (grants 03-01-00104 and 02-01-22004) and by INTAS (grant 03-51-3663)

method ([9], [10]). In [11] and [12], these ideas were applied to determine $M_\Gamma(n)$ for the fundamental groups of some Seifert spaces. Asymptotic formulas for $M_\Gamma(n)$ in many important cases were obtained in series of papers by T. W. Müller and his collaborators ([19],[20],[21]). An excellent exposition of the above results and related subjects is given in the book [14].

In the present paper, a new formula for the number of conjugacy classes of subgroups of given index in an arbitrary finitely generated group is obtained.

The main counting principle suggested in Section 2 of the paper is rather universal. It can be applied to Fuchsian groups to calculate the number of non-equivalent surface coverings (Section 3) as well as the number of unrooted maps on the surface (Section 4). Remark that the results of Section 3 were obtained in [9, 15, 16] by making use of cumbersome combinatorial technique. In the present paper they are rederived as simple consequences of Theorem 1 in Section 2. Another application of Theorem 1 is given in Section 4 where a new approach to determination of the number of unrooted maps on the closed oriented surface with given number of edges is suggested. Earlier, in more complicated way this result was obtained in [18].

2 Counting conjugacy classes of subgroup

Denote by $\mathrm{Epi}(K,\,\mathbb{Z}_\ell)$ the set of epimorphisms of a group K onto the cyclic group $\mathbb{Z}_\ell$ of order ℓ and by $|E|$ the cardinality of a set E.

The main result of this paper is the following counting principle.

Theorem 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*

$$N_\Gamma(n) = \frac{1}{n} \sum_{\substack{\ell \mid n \\ \ell m = n}} \sum_{K \underset{m}{<} \Gamma} |\mathrm{Epi}(K,\,\mathbb{Z}_\ell)|,$$

where the sum $\sum\limits_{K \underset{m}{<} \Gamma}$ *is taken over all subgroups K of index m in the group Γ.*

Proof: Let P be a subgroup in Γ and $N(P,\Gamma)$ is the normalizer of P in the group Γ. We need the following two elementary lemmas.

Lemma 1 *The number of conjugacy classes of subgroups of index n in the group Γ is given by the formula*

$$N_\Gamma(n) = \frac{1}{n} \sum_{P \underset{n}{<} \Gamma} |N(P,\Gamma)/P|.$$

Proof: Let E be a conjugacy class of subgroups of index n in the group Γ. We claim that

$$\sum_{P\in E}|N(P,\Gamma)/P| = n.$$

Indeed, let $P'\in E$. Then $|E| = |\Gamma : N(P',\Gamma)|$ and for any $P\in E$ the groups $N(P,\Gamma)/P$ and $N(P',\Gamma)/P'$ are isomorphic. We have

$$\sum_{P\in E}|N(P,\Gamma)/P| = |E||N(P',\Gamma)/P'| = |\Gamma : N(P',\Gamma)||N(P',\Gamma) : P'| = |\Gamma : P'| = n.$$

Hence,

$$n\,N(n) = \sum_{E} n = \sum_{E}\sum_{P\in E}|N(P,\Gamma)/P| = \sum_{P\underset{n}{<}\Gamma}|N(P,\Gamma)/P|,$$

where the sum $\sum\limits_{E}$ is taken over all conjugacy classes E of subgroups of index n in the group Γ. □

Lemma 2 *Let P be a subgroup of index n in the group Γ. Then*

$$|N(P,\Gamma)/P| = \sum_{\substack{\ell|n\\ \ell m=n}}\ \sum_{P\underset{Z_\ell}{\lhd}K\underset{m}{<}\Gamma}\phi(\ell),$$

where $\phi(\ell)$ is the Euler function and the second sum is taken over all subgroups K of index m in Γ containing P as a normal subgroup with $K/P\cong \mathsf{Z}_\ell$. The sum vanishes if there are no such subgroups.

Proof: Set $G = N(P,\Gamma)/P$. Since $P\lhd N(P,\Gamma)<\Gamma$ and $P\underset{n}{<}\Gamma$, the order of any cyclic subgroup of G is a divisor of n.

Note that there is a one-to-one correspondence between cyclic subgroups Z_ℓ in the group G and subgroups K satisfying $P\underset{Z_\ell}{\lhd}K\underset{m}{<}\Gamma$, where $\ell m = n$.

Given a cyclic subgroup $\mathsf{Z}_\ell < G$ there are exactly $\phi(\ell)$ elements of G which generate Z_ℓ.

Hence,

$$|G| = \sum_{\ell|n}\phi(\ell)\sum_{\mathsf{Z}_\ell<G}1 = \sum_{\ell|n}\phi(\ell)\sum_{P\underset{Z_\ell}{\lhd}K\underset{m}{<}\Gamma}1 = \sum_{\ell|n}\ \sum_{P\underset{Z_\ell}{\lhd}K\underset{m}{<}\Gamma}\phi(\ell).$$

□

We finish the proof of the theorem by applying Lemma 1 and Lemma 2 for $\ell m = n$:

$$n\,N(n) = \sum_{P \underset{n}{<} \Gamma} |N(P,\Gamma)/P| =$$

$$\sum_{P \underset{n}{<} \Gamma} \sum_{\ell | n} \sum_{P \underset{Z_\ell}{\triangleleft} K \underset{m}{<} \Gamma} \phi(\ell) = \sum_{\ell | n} \sum_{P \underset{n}{<} \Gamma} \sum_{P \underset{Z_\ell}{\triangleleft} K \underset{m}{<} \Gamma} \phi(\ell) =$$

$$\sum_{\ell | n} \sum_{K \underset{m}{<} \Gamma} \sum_{P \underset{Z_\ell}{\triangleleft} K} \phi(\ell) = \sum_{\ell | n} \sum_{K \underset{m}{<} \Gamma} |\mathrm{Epi}(K, \mathsf{Z}_\ell)|.$$

The last equality is a consequence of the following observation. Given subgroup P, $P \underset{Z_\ell}{\triangleleft} K$ there are exactly $\phi(\ell)$ epimorphisms $\psi : K \to \mathsf{Z}_\ell$, with $\mathrm{Ker}\,(\psi) = \mathrm{P}$. Indeed, any two of them differ one from other by an element of the group $\mathrm{Aut}(\mathsf{Z}_\ell)$ consisting of $\phi(\ell)$ elements. □

Denote by $\mathrm{Hom}(\Gamma,\, \mathsf{Z}_\ell)$ the set of homomorphisms of a group Γ into the cyclic group Z_ℓ of order ℓ. Following G. Jones [1] we note that $|\mathrm{Hom}(\Gamma,\, \mathsf{Z}_\ell)| = \sum_{d|\,\ell} |\mathrm{Epi}(\Gamma,\, \mathsf{Z}_d)|$. Hence, by the Möbius inversion formula [2, §8.3, p. 148] we have the following result

Lemma 3

$$|\mathrm{Epi}(\Gamma,\, Z_\ell)| = \sum_{d|\,\ell} \mu(\frac{\ell}{d}) |\mathrm{Hom}(\Gamma,\, Z_d)|,$$

where $\mu(n)$ is the Möbius function.

This lemma essentially simplifies the calculation of $|\mathrm{Epi}(\Gamma,\, \mathsf{Z}_\ell)|$ for a finitely generated group Γ. Indeed, let $\mathrm{H}_1(\Gamma) = \Gamma/[\Gamma,\, \Gamma]$ be the first homology group of Γ. Suppose that $\mathrm{H}_1(\Gamma) = \mathsf{Z}_{\mathrm{m}_1} \oplus \mathsf{Z}_{\mathrm{m}_2} \oplus \ldots \oplus \mathsf{Z}_{\mathrm{m}_\mathrm{s}} \oplus \mathsf{Z}^\mathrm{r}$. Then we have

Lemma 4

$$|\mathrm{Epi}(\Gamma,\, Z_\ell)| = \sum_{d|\,\ell} \mu(\frac{\ell}{d})\,(m_1, d)\,(m_2, d) \ldots (m_s, d)\, d^r,$$

where (m, d) is the greatest common divisor of m and d.

Proof: Note that $|\mathrm{Hom}(\mathsf{Z}_m,\, \mathsf{Z}_d)| = (m, d)$ and $|\mathrm{Hom}(\mathsf{Z},\, \mathsf{Z}_d)| = d$. Since the group Z_d is Abelian, we obtain

$$|\mathrm{Hom}(\Gamma,\, \mathsf{Z}_d)| = |\mathrm{Hom}(\mathrm{H}_1(\Gamma),\, \mathsf{Z}_d)| = (m_1, d)\,(m_2, d) \ldots (m_s, d)\, d^r.$$

Then the result follows from Lemma 3. □

In particular, we have

Corollary 1

(i) *Let* $\mathrm{F_r}$ *be a free group of rank* r. *Then* $\mathrm{H_1(F_r)} = \mathbb{Z}^{\mathrm{r}}$ *and* $|\mathrm{Epi}(\mathrm{F}_r,\, \mathbb{Z}_\ell)| \;=\; \sum\limits_{d|\,\ell} \mu(\frac{\ell}{d})\, d^r$.

(ii) *Let* $\Phi_g = \langle a_1, b_1, \ldots, a_g, b_g : \prod\limits_{i=1}^{g} [a_i,\, b_i] = 1\rangle$ *be the fundamental group of a closed orientable surface of genus* g. *Then* $\mathrm{H_1(\Phi_g)} = \mathbb{Z}^{2\mathrm{g}}$ *and* $|\mathrm{Epi}(\Phi_r,\, \mathbb{Z}_\ell)| = \sum\limits_{d|\,\ell} \mu(\frac{\ell}{d})\, d^{2g}$.

(iii) *Let* $\Lambda_p = \langle a_1, a_2, \ldots, a_p : \prod\limits_{i=1}^{p} a_i^2 = 1\rangle$ *be the fundamental group of a closed non-orientable surface of genus* p. *Then* $\mathrm{H_1(\Lambda_p)} = \mathbb{Z}_2 \oplus \mathbb{Z}^{\mathrm{p}-1}$ *and* $|\mathrm{Epi}(\Lambda_p,\, \mathbb{Z}_\ell)| = \sum\limits_{d|\,\ell} \mu(\frac{\ell}{d})\, (2, d)\, d^{p-1}$.

Note that the fundamental group of any compact surface (orientable or not, possibly, with non-empty boundary) is one of the three groups F_r, Φ_g and Λ_p listed in Corollary 1. In the next two sections we identify the number of conjugacy classes of subgroups of index n in the group Γ and the number of equivalence classes of n-fold unbranched connected coverings of a manifold with fundamental group Γ.

3 Counting surface coverings

Recall that the fundamental group $\pi_1(\mathcal{B})$ of a bordered surface $\mathcal{B}$ of Euler characteristic $\chi = 1 - r$, $r \geq 0$, is a free group F_r of rank r. An example of such a surface is the disc $\mathcal{D}_r$ with r holes. As the first application of the counting principle (Theorem 1) we have the following result obtained earlier by V. Liskovets [9]

Theorem 2 *Let* $\mathcal{B}$ *be a bordered surface with the fundamental group* $\pi_1(\mathcal{B}) = \mathrm{F}_r$. *Then the number of non-equivalent* $n-$*fold coverings of* $\mathcal{B}$ *is given by the formula*

$$N(n) = \frac{1}{n} \sum_{\substack{\ell|\,n \\ \ell\, m=n}} \sum_{d|\,\ell} \mu(\frac{\ell}{d})\, d^{(r-1)m+1} M(m),$$

where $M(m)$ *is the number of subgroups of index* m *in the group* F_r.

Proof: Note that all subgroups of index m in F_r are isomorphic to $\Gamma_m = \mathrm{F}_{(r-1)m+1}$. By Theorem 1 we have

$$N(n) = \frac{1}{n} \sum_{\substack{\ell|\,n \\ \ell\, m=n}} |\mathrm{Epi}(\Gamma_m,\, \mathbb{Z}_\ell)| \cdot M(m).$$

By Corollary 1(i) we get

$$|\mathrm{Epi}(\Gamma_m,\, \mathsf{Z}_\ell)| = \sum_{d|\,\ell} \mu(\frac{\ell}{d})\, d^{(r-1)m+1}$$

and the result follows. □

By the M. Hall recursive formula [4] the number of subgroups of index m in the group F_r is equal to $M(m) = \dfrac{t_{m,r}}{(m-1)!}$, where

$$t_{m,r} = m!^r - \sum_{j=1}^{m-1} \binom{m-1}{j-1} (m-j)!^r\, t_{j,r}, \;\; t_{1,r} = 1.$$

The next result was obtained in [15] in a rather complicated way.

Theorem 3 *Let $\mathcal{S}$ be a closed orientable surface with the fundamental group $\pi_1(\mathcal{S}) = \Phi_g$. Then the number of non-equivalent $n-$fold coverings of $\mathcal{S}$ is given by the formula*

$$N(n) = \frac{1}{n} \sum_{\substack{\ell|\, n \\ \ell\, m = n}} \sum_{d|\,\ell} \mu(\frac{\ell}{d})\, d^{2(g-1)m+2} M(m),$$

where $M(m)$ is the number of subgroups of index m in the group Φ_g.

Proof: Recall that any subgroup K_m of index m in the group Φ_g is isomorphic to $\Phi_{g'}$, where g' and g are related by the Riemann-Hurwitz formula [6] $2g' - 2 = m(2g-2)$. Hence, $K_m = \Phi_{(g-1)m+1}$. By the main counting principle (Theorem 1) we have

$$N(n) = \frac{1}{n} \sum_{\substack{\ell|\, n \\ \ell\, m = n}} |\mathrm{Epi}(K_m,\, \mathsf{Z}_\ell)| \cdot M(m),$$

where

$$|\mathrm{Epi}(K_m,\, \mathsf{Z}_\ell)| = \sum_{d|\,\ell} \mu(\frac{\ell}{d})\, d^{2(g-1)m+2}$$

is given by Corollary 1(ii) . □

Let $\mathcal{N}$ be a closed non-orientable surface of genus p with the fundamental group $\pi_1(\mathcal{N}) = \Lambda_p$. Denote by $\mathcal{N}_m^+$ and $\mathcal{N}_m^-$ an orientable and non-orientable $m-$fold coverings of $\mathcal{N}$, respectively and set $\Gamma_m^+ = \pi_1(\mathcal{N}_m^+)$ and $\Gamma_m^- = \pi_1(\mathcal{N}_m^-)$. For simplicity, we will refer to Γ_m^+ and Γ_m^- as orientable and non-orientable subgroups of index m in Λ_p, respectively. By the Riemann-Hurwitz formula we get

$$2\,\mathrm{genus}\,(\mathcal{N}_m^+) - 2 = m(p-2) \;\text{ and }\; \mathrm{genus}\,(\mathcal{N}_m^-) - 2 = m(p-2),$$

where $p = \text{genus}(\mathcal{N})$. Hence $\Gamma_m^+ = \Phi_{\frac{m}{2}(p-2)+1}$ and $\Gamma_m^- = \Lambda_{m(p-2)+2}$.

By the main counting principle, the number of non-equivalent $n-$fold coverings of N is given by the formula

$$N(n) = \frac{1}{n} \sum_{\substack{\ell \mid n \\ \ell m = n}} (|\text{Epi}(\Gamma_m^+, \mathsf{Z}_\ell)| \cdot M^+(m) + |\text{Epi}(\Gamma_m^-, \mathsf{Z}_\ell)| \cdot M^-(m)),$$

where $M^+(m)$ and $M^-(m)$ are the numbers of orientable and non-orientable subgroups of index m in the group Λ_p, respectively.

By Corollary 1(ii) and Corollary 1(iii), we have

$$|\text{Epi}(\Gamma_m^+, \mathsf{Z}_\ell)| = \sum_{d \mid \ell} \mu(\frac{\ell}{d})\, d^{m(p-2)+2} \ \text{ and } \ |\text{Epi}(\Gamma_m^-, \mathsf{Z}_\ell)| = \sum_{d \mid \ell} \mu(\frac{\ell}{d})\,(2,d)\, d^{m(p-2)+1}.$$

As a result, we have proved the following theorem obtained earlier in [16] by making use of a cumbersome combinatorial technique.

Theorem 4 *Let n be a closed orientable surface with the fundamental group $\pi_1(\mathcal{N}) = \Lambda_p$. Then the number of non-equivalent $n-$fold coverings of $\mathcal{N}$ is given by the formula*

$$N(n) = \frac{1}{n} \sum_{\substack{\ell \mid n \\ \ell m = n}} \sum_{d \mid \ell} \mu(\frac{\ell}{d})\,(d^{m(p-2)+2} M^+(m) + (2,d)\, d^{m(p-2)+1} M^-(m)),$$

where $M^+(m)$ and $M^-(m)$ are the numbers of orientable and non-orientable subgroups of index m in the group Λ_p, respectively.

For completeness note that ([15], [16]) if $\Gamma = \Phi_g$ or Λ_p then the number $M(m)$ of subgroups of index m in the group Γ is equal to

$$R_\nu(m) = m \sum_{s=1}^{m} \frac{(-1)^{s+1}}{s} \sum_{\substack{i_1+i_2+\ldots+i_s=m \\ i_1, i_2, \ldots, i_s \geq 1}} \beta_{i_1}\beta_{i_2}\cdots\beta_{i_s},$$

where $\beta_k = \sum_{\chi \in D_k} \big(\frac{k!}{f^\chi}\big)^\nu$, D_k is the set of irreducible representations of a symmetric group S_k, f^χ is the degree of the representation χ, $\nu = 2g-2$ for $\Gamma = \Phi_g$ and $\nu = p-2$ for $\Gamma = \Lambda_p$. Moreover, in the latter case, $M^+(m) = 0$ if m is odd, $M^+(m) = R_{2\nu}(\frac{m}{2})$ if m is even, and $M^-(m) = M(m) - M^+(m)$. Also, the number of subgroups can be found by the following recursive formula

$$M(m) = m\,\beta_m - \sum_{j=1}^{m-1} \beta_{m-j} M(j), \quad M(1) = 1.$$

4 Non-isomorphic maps on surface

In this section we deal with the problem of enumeration of oriented unrooted maps of given genus g. From now on *a surface* is a connected, orientable surface without a border. A *map* is a 2-cell decomposition of a surface. Standardly, maps on surfaces are described as 2-cell embeddings of graphs. An embedded graph is a 4-tuple (D, V, I, L), where D and V are disjoint sets of *darts* and *vertices*, respectively, I is an incidence function $I : D \to V$ assigning to each dart an initial vertex, and L is the dart-reversing involution. Edges of a graph are orbits of L. Note that some edges may be incident just with one vertex, such edges will be called *semiedges*. In what follows we shall deal with the category of *oriented maps*, that means one of the two global orientations of the underlying surface is fixed. Recall, that a map is called *rooted* if it has one distinguished dart x called a *root*. An *isomorphism* between rooted maps takes root onto root. Recall that if (M, x) and (M, y) are two rooted maps based on the same map with a dart set D then the number of isomorphism classes for (M, x) and (M, y) is the same. There is a 1-1 correspondence between isomorphism classes of rooted maps defined in the category of oriented maps, and isomorphism classes of rooted maps in the category of maps on orientable surfaces as they are defined, for instance, in monograph [13, page 7].

We fix the set of darts D and consider different maps based on D. We will determine the number of isomorphism classes of (unrooted) maps with n darts and of given genus g. This number will be denoted by $NUM_g(n)$.

Denote by $Orb(S_g/Z_\ell)$ the set of all orbifolds arising as cyclic quotients under some action of Z_ℓ on maps on a surface of genus g and by $NRM_O(m)$ the number of rooted quotient maps for a given orbifold type O which lift onto maps on a surface of genus g, having $n = \ell m$ darts. We note that if the map contains no semi-edges then the number of darts $n = 2\,e$, where e is the number of edges of the map.

Let S_g be an orientable surface of genus g and Z_ℓ be a cyclic group of automorphisms of S_g. Denote by $(\gamma; m_1, m_2, \ldots, m_r)$, $2 \le m_1 \le m_2 \le \ldots \le m_r \le \ell$, the signature of orbifold $O = S_g/Z_\ell$. That is, the underlying space of O is an oriented surface of genus γ and the regular cyclic covering $S_g \to O = S_g/Z_\ell$ is branched over r points of O with branch indices $m_1, m_2, \ldots, m_r$, respectively. W.J. Harvey [5] obtained necessary and sufficient conditions for an existence of a cyclic orbifold S_g/Z_ℓ with signature $(\gamma; m_1, m_2, \ldots, m_r)$.

Given orbifold O of the signature $(\gamma; m_1, m_2, \ldots, m_r)$ define an orbifold fundamental group $\pi_1(O)$ to be an F-group generated by 2γ generators $a_1, b_1, a_2, b_2, \ldots, a_\gamma, b_\gamma$ and by r generators e_j, $j = 1, \ldots, r$ satisfying the relations

$$\prod_{i=1}^{\gamma}[a_i, b_i]\prod_{j=1}^{r} e_j = 1,\ e_j^{m_j} = 1 \text{ for every } \ j = 1, \ldots, r,$$

where $[a, b] = aba^{-1}b^{-1}$.

An epimorphism $\pi_1(O) \to Z_\ell$ onto a cyclic group of order ℓ is called *order preserving* if it preserves the orders of generators e_j, $j = 1, \ldots, r$. Equivalently, an order preserving epimorphism $\pi_1(O) \to Z_\ell$ has a torsion-free kernel. We denote by $Epi_0(\pi_1(O), Z_\ell)$ the number of order preserving epimorphisms $\pi_1(O) \to Z_\ell$.

An elementary consideration based on Theorem 1 enables us to prove the following theorem [18].

Theorem 5 *With the above notation the following enumeration formula holds:*

$$NUM_g(n) = \frac{1}{n} \sum_{\ell|n, n=\ell m} \sum_{O \in Orb(S_g/Z_\ell)} Epi_0(\pi_1(O), Z_\ell) NRM_O(m).$$

The number of rooted maps on the orbifold as well as the number of order preserving epimorphisms are explicitly determined in joint paper with R. Nedela [18]. Numerical tables for the number of non-isomorphic (unrooted) maps on closed surface of genus 1, 2 and 3 with up to 30 edges are also given in the paper.

Acknowledgement: I gratefully thank Prof. Jin Ho Kwak for helpful comments and Com^{2}MaC, Pohang University of Science and Technology (Korea) for hospitality during preparation of the present paper. It is also my pleasure to send warm regards to referees for careful reading of the paper and making valuable remarks and suggestions.

References

[1] G. A. Jones, Enumeration of homomorphisms and surface coverings, *Quart. J. Math. Oxford* **46**:2 (1995), 485–507.

[2] G. A. Jones and J. M. Jones, Elementary Number Theory. Berlin: Springer-Verlag, 1998.

[3] G. A. Jones and V. Liskovets, Enumerating of subgroups of finite index, In: *Mini-Workshop on Hurwitz Theory and Ramifications* (ed. by J. H. Kwak et A. D. Mednykh), Combinatorial and Computational Mathematics Center, Pohang University of Science and Technology, 2003, 185–188.

[4] M. Hall, Jr., Subgroups of finite index in free groups, *Canad. J. Math.,* **1**, No.1 (1949), 187–190.

[5] W.J. Harvey, Cyclic groups of automorphisms of a compact Riemann surface, Quart. J. Math. **17** (1966), 86-97.

[6] A. Hurwitz, Uber Riemann'sche Flächen mit gegebenen Verzweigungspunkten, *Math. Ann.* **39** (1891), 1–60.

[7] A. Hurwitz, Uber die Anzahl der Riemann'schen Flächen mit gegebenen Verzweigungspunkten, *Math. Ann.* **55** (1902), 53–66.

[8] J. H. Kwak and J. Lee, Enumeration of graph coverings, surface branched coverings and related group theory, In: *Combinatorial and Computational Mathematics: Present and Future* (ed. by S. Hong, J. H. Kwak et al.), World Scientific, Singapore, 2001, 97–161.

[9] V. Liskovets, Towards the enumeration of subgroups of a free group, *Dokl. Akad. Nauk BSSR* **15** (1971), 6–9 (in Russian).

[10] V. Liskovets, Recursive enumeration under mutually orthogonal group action, *Acta-Appl. Math.* **52** (1998), 91–120.

[11] V. Liskovets and A. Mednykh, Enumeration of subgroups in the fundamental groups of orientable circle bundles over surfaces, *Comm. in Algebra*, **28**, No.4, (2000), 1717–1738.

[12] V. Liskovets and A. Mednykh, The number of subgroups in the fundamental groups of some non-oriented 3-manifofds, In: *Formal power series and algebraic combinatorics. Proceedings of the 12th international conference, FPSAC'00, Moscow, Russia, June 26-30, 2000*, Springer, Berlin, 2000, 276–287.

[13] Y.P. Liu, "Enumerative theory of maps", Kluwer Academic Publishers, London 1999.

[14] A. Lubotzky and D. Segal, Subgroup Growth, Birkhäuser Verlag, Basel-Boston-Berlin, 2003.

[15] A.D. Mednykh, Hurwitz problem on the number of nonequivalent coverings of a compact Riemann surface, *Siber. Math. J.* **23**, No.3, (1982), 415-420.

[16] A.D. Mednykh and G.G. Pozdnyakova, Number of nonequivalent coverings over a nonorientable compact surface, *Siber. Math. J.* **27**, No.1, (1986), 99-106.

[17] A.D. Mednykh, On the number of subgroups in the fundamental group of a closed surfaces, *Comm. in Algebra*, **16**, No.10, (1988), 2137–2148.

[18] A .D. Mednykh and R. Nedela, Enumeration of unrooted maps with given genus (2004), `http://www.savbb.sk/mu/articles/4_2004_nedela.pdf`. Submitted to *J. Combinatorial Theory (B).*

[19] T. W. Müller, Enumerating representations in finite wreath products, *Adv. Math.*, **153**, No.1, (2000), 118–154.

[20] T. W. Müller and J. Shareshian, Enumerating representations in finite wreath products. II: Explicit formulae, *Adv. Math.* , **171**, No.2, (2002), 276–331.

[21] T. W. Müller and J.-C. Puchta, Character theory of symmetric groups and subgroup growth of surface groups, *J. London Math. Soc. (2)*, **66** (2002), 623–640.

PERMUTOHEDRA, ASSOCIAHEDRA, AND BEYOND

ALEXANDER POSTNIKOV

Abstract. The volume and the number of lattice points of the permutohedron P_n are given by certain multivariate polynomials that have remarkable combinatorial properties. We give 3 different formulas for these polynomials. We also study a more general class of polytopes that includes the permutohedron, the associahedron, the cyclohedron, the Stanley-Pitman polytope, and various generalized associahedra related to wonderful compactifications of De Concini-Procesi. These polytopes are constructed as Minkowski sums of simplices. We calculate their volumes and describe their combinatorial structure. The coefficients of monomials in $\operatorname{Vol} P_n$ are certain positive integer numbers, which we call the mixed Eulerian numbers. These numbers are equal to the mixed volumes of hypersimplices. Various specializations of these numbers give the usual Eulerian numbers, the Catalan numbers, the numbers $(n+1)^{n-1}$ of trees (or parking functions), the binomial coefficients, etc. We calculate the mixed Eulerian numbers using certain binary trees. Many results are extended to an arbitrary Weyl group.

1. Permutohedron

Let $x_1, \ldots, x_{n+1}$ be real numbers. The *permutohedron* $P_n(x_1, \ldots, x_{n+1})$ is the convex polytope in $\mathbb{R}^{n+1}$ defined as the convex hull of all permutations of the vector $(x_1, \ldots, x_{n+1})$:

$$P_n(x_1, \ldots, x_{n+1}) := \mathrm{ConvexHull}((x_{w(1)}, \ldots, x_{w(n+1)}) \mid w \in S_{n+1}),$$

where S_{n+1} is the symmetric group. Actually, this is an n-dimensional polytope that lies in some hyperplane $H \subset \mathbb{R}^{n+1}$. More generally, for a Weyl group W, we can define the *weight polytope* as a convex hull of a Weyl group orbit:

$$P_W(x) := \mathrm{ConvexHull}(w(x) \mid w \in W),$$

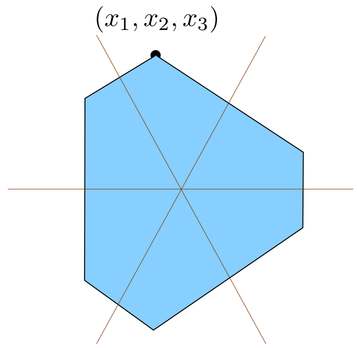
where x is a point in the weight space on which the Weyl group acts.

For example, for $n = 2$ and distinct x_1, x_2, x_3, the permutohedron $P_2(x_1, x_2, x_3)$ (type A_2 weight polytope) is the hexagon shown below. If some of the numbers x_1, x_2, x_3 are equal to each other then the permutohedron degenerates into a triangle, or even a single point.

Date: 11/21/2004; minor updates 02/28/2005; based on transparencies dated 07/26/2004.

Key words and phrases. Permutohedron, associahedron, Minkowski sum, mixed volume, Weyl's character formula, Hall's marriage theorem, wonderful compactification, nested families, generalized Catalan numbers, Stanley-Pitman polytope, parking functions, hypersimplices, mixed Eulerian numbers, binary trees, shifted tableaux, Gelfand-Zetlin patterns.

$P_2(x_1, x_2, x_3) =$



Question: What is the volume of the permutohedron $V_n := \mathrm{Vol}\, P_n$?

Since P_n does not have the full dimension in $\mathbb{R}^{n+1}$, one needs to be careful with definition of the volume. We assume that the volume form Vol is normalized so that the volume of a unit parallelepiped formed by generators of the integer lattice $\mathbb{Z}^{n+1} \cap H$ is 1. More generally, we can ask the following question.

Question: What is the number of lattice points $N_n := P_n \cap \mathbb{Z}^{n+1}$?

We will see that both V_n and N_n are polynomials of degree n in the variables $x_1, \ldots, x_{n+1}$. The polynomial V_n is the top homogeneous part of N_n. The *Ehrhart polynomial* of P_n is $E(t) = N_n(tx_1, \ldots, tx_n)$, and V_n is its top degree coefficient.

We will give 3 totally different formulas for these polynomials.

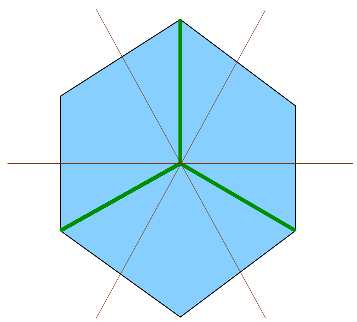
Let us first mention the special case $(x_1, \ldots, x_{n+1}) = (n+1, \ldots, 1)$. The polytope

$$P_n(n+1, n, \ldots, 1) = \mathrm{ConvexHull}((w(1), ..., w(n+1)) \mid w \in S_{n+1})$$

is *the most symmetric* permutohedron. It is invariant under the action of the symmetric group S_{n+1}. For example, for $n = 2$, it is the regular hexagon:

regular hexagon

subdivided into 3 rhombi



The polytope $P_n(n+1, \ldots, 1)$ is a *zonotope,* i.e., Minkowski sum of line intervals. It is well known that

- $V_n(n+1, \ldots, 1) = (n+1)^{n-1}$ is the number of trees on $n+1$ labelled vertices. Indeed, $P_n(n+1, \ldots, 1)$ can be subdivided into parallelepipeds of unit volume associated with trees. This follows from a general result about zonotopes.
- $N_n(n+1, \ldots, 1)$ is the number of forests on $n+1$ labelled vertices.

In general, for arbitrary $x_1, \ldots, x_{n+1}$, the permutohedron $P_n(x_1, \ldots, x_{n+1})$ is not a zonotope. We cannot easily calculate its volume by subdividing it into parallelepipeds.

2. First Formula

Theorem 2.1. *Fix any distinct numbers $\lambda_1, \ldots, \lambda_{n+1}$ such that $\lambda_1 + \cdots + \lambda_{n+1} = 0$. We have*

$$V_n(x_1, \ldots, x_{n+1}) = \frac{1}{n!} \sum_{w \in S_{n+1}} \frac{(\lambda_{w(1)}x_1 + \cdots + \lambda_{w(n+1)}x_{n+1})^n}{(\lambda_{w(1)} - \lambda_{w(2)})(\lambda_{w(2)} - \lambda_{w(3)}) \cdots (\lambda_{w(n)} - \lambda_{w(n+1)})}.$$

Notice that all λ_i's in the right-hand side are canceled after the symmetrization.

More generally, let W be the Weyl group associated with a rank n root system, and let $\alpha_1, \ldots, \alpha_n$ be a choice of simple roots.

Theorem 2.2. *Let λ be any regular weight. The volume of the weight polytope is*

$$\operatorname{Vol} P_W(x) = \frac{f}{|W|} \sum_{w \in W} \frac{(\lambda, w(x))^n}{(\lambda, w(\alpha_1)) \cdots (\lambda, w(\alpha_n))},$$

where the volume is normalized so that the volume of the parallelepiped generated by the simple roots α_i is 1, and f is the index of the root lattice in the weight lattice.

Idea of Proof. Let us use Khovansky-Pukhlikov's method [PK]. Instead of counting the number of lattice points in P, calculate the sum $\Sigma[P]$ of formal exponents e^a over lattice points $a \in P \cap \mathbb{Z}^n$. We can work with unbounded polytopes. For example, for a simplicial cone C, the sum $\Sigma[C]$ is given by a simple rational expression. Any polytope P can be explicitly presented as an alternating sum of simplicial cones $\Sigma[P] = \Sigma[C_1] \pm \Sigma[C_2] \pm \cdots$ over vertices of P.

Applying this method to the weight polytope, we obtain the following claim.

Theorem 2.3. *For a dominant weight μ, the sum of exponents over lattice points of the weight polytope $P_W(\mu)$ equals*

$$\Sigma[P_W(\mu)] := \sum_{a \in P_W(\mu) \cap (L+\mu)} e^a = \sum_{w \in W} \frac{e^{w(\mu)}}{(1 - e^{-w(\alpha_1)}) \cdots (1 - e^{-w(\alpha_n)})},$$

where L be the root lattice.

Compare this claim with Weyl's character formula! If we replace the product over simple roots α_i in the right-hand side of Theorem 2.3 by a similar product over all positive roots, we obtain exactly Weyl's character formula.

Theorem 2.2 and its special case Theorem 2.1 and be deduced from Theorem 2.3 in the same way as Weyl's dimension formula can be deduced from Weyl's character formula. $\square$

Remark 2.4. The sum of exponents $\Sigma[P_W(\mu)]$ over lattice points of the weight polytope is obtained from the character $ch\ V_\mu$ of the irreducible representation V_μ of the associated Lie group by replacing all nonzero coefficients in $ch\, V_\mu$ with 1. For example, in type A, $ch\, V_\mu$ is the Schur polynomial s_μ and $\Sigma[P_n(\mu)]$ is obtained from the Schur polynomial s_μ by erasing the coefficients of all monomials.

3. Second Formula

Let us use the coordinates $y_1, \ldots, y_{n+1}$ related to $x_1, \ldots, x_{n+1}$ by

$$\begin{cases} y_1 = -x_1 \\ y_2 = -x_2 + x_1 \\ y_3 = -x_3 + 2x_2 - x_1 \\ \cdots \\ y_{n+1} = -\binom{n}{0} x_n + \binom{n}{1} x_{n-1} - \cdots \pm \binom{n}{n} x_1 \end{cases}$$

Write $V_n = \mathrm{Vol}\ P_n(x_1, \ldots, x_{n+1})$ as a polynomial in the variables $y_1, \ldots, y_{n+1}$.

Theorem 3.1. *We have*

$$V_n = \frac{1}{n!} \sum_{(S_1,\ldots,S_n)} y_{|S_1|} \cdots y_{|S_n|},$$

where the sum is over ordered collections of subsets $S_1, \ldots, S_n \subset [n+1]$ such that, for any distinct $i_1, \ldots, i_k$, we have $|S_{i_1} \cup \cdots \cup S_{i_k}| \geq k+1$.

This theorem implies that $n!\,V_n$ is a polynomial in $y_2, \ldots, y_{n+1}$ with positive integer coefficients. For example, $V_1 = \mathrm{Vol}\,([(x_1, x_2), (x_2, x_1)]) = x_1 - x_2 = y_2$ and $2V_2 = 6\,y_2^2 + 6\,y_2\,y_3 + y_3^2$.

Remark 3.2. The condition on subsets $S_1, \ldots, S_n$ in Theorem 3.1 is very similar to the condition in Hall's marriage theorem. One just needs to replace the inequality $\geq k+1$ with $\geq k$ to obtain Hall's marriage condition.

It is not hard to prove the following analog of Hall's theorem.

Dragon marriage problem: *There are $n+1$ brides and n grooms living in a medieval village. A dragon comes to the village and takes one of the brides. We are given a collection G of pairs of brides and grooms that can marry each other. Find the condition on G that ensures that, no matter which bride the dragon takes, it will be possible to match the remaining brides and grooms.*

Proposition 3.3. (Dragon marriage theorem) *Let $S_1, \ldots, S_n \subset [n+1]$. The following three conditions are equivalent:*

1. *For any distinct $i_1, \ldots, i_k$, we have $|S_{i_1} \cup \cdots \cup S_{i_k}| \geq k+1$.*
2. *For any $j \in [n{+}1]$, there is a system of distinct representatives in $S_1, \ldots, S_n$ that avoids j. (This is a reformulation of the dragon marriage problem.)*
3. *One can remove some elements from S_i's to get the edge set of a spanning tree in K_{n+1}.*

Theorem 3.1 can be extended to a larger class of polytopes discussed below.

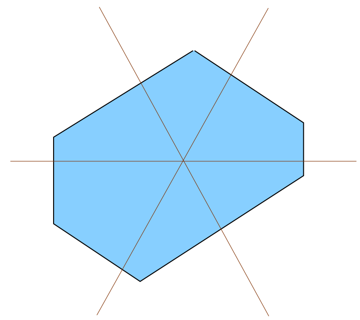
4. Generalized permutohedra

Let $\Delta_{[n+1]} = \mathrm{ConvexHull}(e_1, \ldots, e_{n+1})$ be the standard coordinate simplex in $\mathbb{R}^{n+1}$. For a subset $I \subset [n+1]$, let $\Delta_I = \mathrm{ConvexHull}(e_i \mid i \in I)$ denote the face of the coordinate simplex $\Delta_{[n+1]}$. Let $\mathbf{Y} = \{Y_I\}$ be a collection of parameters $Y_I \geq 0$ for all nonempty subsets $I \subset [n+1]$. Let us define the *generalized permutohedron* $P_n(Y)$ as the Minkowski sum of the simplices Δ_I scaled by factors Y_I:

$$P_n(\mathbf{Y}) := \sum_{I \subset [n+1]} Y_I \cdot \Delta_I \qquad \text{(Minkowski sum)}$$

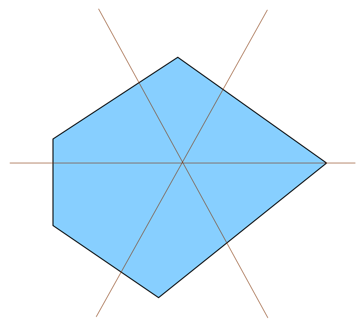
Generalized permutohedra are obtained from usual permutohedra by moving their faces while preserving all angles. So, instead of n degrees of freedom in usual permutohedra, we have $2^{n+1}-2$ degrees of freedom in generalized permutohedra.

a generalized permutohedron



this is also

a generalized permutohedron



The combinatorial type of $P_n(\mathbf{Y})$ depends only on the set of $B \subset 2^{[n+1]}$ of I's for which $Y_I \neq 0$. Here are some interesting examples of generalized permutohedra.

- If $Y_I = y_{|I|}$, i.e., the variables Y_I are equal to each other for all subsets of the same cardinality, then $P_n(\mathbf{Y})$ is the usual permutohedron P_n.
- If $B = \{\{i, i+1, \ldots, j\} \mid 1 \leq i \leq j \leq n\}$ is the set of consecutive intervals, then $P_n(\mathbf{Y})$ is the *associahedron*, also known as the *Stasheff polytope.* The polytope $P_n(\mathbf{Y})$ can be equivalently defined as the Newton polytope of $\prod_{1\leq i \leq j \leq n+1}(x_i + x_{i+1} + \cdots + x_j)$. This is exactly *Loday's realization* of the associahedron, see [L].
- If B is the set of cyclic intervals, then $P_n(\mathbf{Y})$ is a *cyclohedron.*
- If B is the set of connected subsets of nodes of a Dynkin diagram, then $P_n(\mathbf{Y})$ the polytope related to De Concini-Procesi's *wonderful compactification,* see [DP], [DJS].
- If $B = \{[i] \mid i = 1, \ldots, n+1\}$ is the complete flag of initial intervals, then $P_n(\mathbf{Y})$ is the *Stanley-Pitman polytope* from [SP].

Theorem 3.1 can be extended to generalized permutohedra, as follows.

Theorem 4.1. *The volume of the generalized permutohedron $P_n(\mathbf{Y})$ is given by*

$$\operatorname{Vol} P_n(\mathbf{Y}) = \frac{1}{n!} \sum_{(S_1, \ldots, S_n)} Y_{S_1} \cdots Y_{S_n},$$

where the sum is over ordered collections of subsets $S_1, \ldots, S_n \subset [n+1]$ such that, for any distinct $i_1, \ldots, i_k$, we have $|S_{i_1} \cup \cdots \cup S_{i_k}| \geq k+1$.

Theorem 4.2. *The number of lattice points in the generalized permutohedron $P_n(\mathbf{Y})$ is*

$$P_n(\mathbf{Y}) \cap \mathbb{Z}^{n+1} = \frac{1}{n!} \sum_{(S_1,\dots,S_n)} \{Y_{S_1} \cdots Y_{S_n}\},$$

where the summation is over the same collections $(S_1, \dots, S_n)$ as before, and

$$\left\{ \prod_I Y_I^{a_I} \right\} := (Y_{[n+1]}+1)^{\{a_{[n+1]}\}} \prod_{I \neq [n+1]} Y_I^{\{a_I\}}, \text{ where } Y^{\{a\}} = Y(Y+1)\cdots(Y+a-1).$$

In other words, the formula for the number of lattice points in $P_n(\mathbf{Y})$ is obtained from the formula for the volume by replacing usual powers in all terms by raising powers.

These formulas generalize formulas from [SP] for the volume and the number of lattice points in the Stanley-Pitman polytope. In this case, collections $(S_1, \dots, S_n)$ of *initial intervals* $S_i = [s_i]$ that satisfy the dragon marriage condition, see Proposition 3.3, are in one-to-one correspondence with *parking functions* $(s_1, \dots, s_n)$. The volume $P_n(\mathbf{Y})$ is given by the sum over parking functions.

5. Nested families and generalized Catalan numbers

In this section, we describe the combinatorial structure of the class of generalized permutohedra $P_n(\mathbf{Y})$ such that the set $B \subset 2^{[n+1]}$ of subsets I with nonzero Y_I satisfies the following connectivity condition:

(B1) If $I, J \in B$ and $I \cap J \neq \emptyset$, then $I \cup J \in B$.

All examples of generalized permutohedra mentioned in the previous section satisfy this additional condition. Without loss of generality we will also assume that

(B2) B contains all singletons $\{i\}$, for $i \in [n+1]$.

Indeed, the Minkowski sum of a polytope with $\Delta_{\{i\}}$, which is a single point, is just a parallel translation of the polytope.

Sets $B \subset 2^{[n+1]}$ that satisfy conditions (B1) and (B2) are called *building sets*. Note that condition (B1) implies that all maximal by inclusion elements in B are pairwise disjoint.

Let us say that a subset N in a building set B is a *nested family* if it satisfies the following conditions:

(N1) For any $I, J \in N$, we have either $I \subseteq J$, or $J \subseteq I$, or I and J are disjoint.
(N2) For any collection of $k \geq 2$ disjoint subsets $I_1, \dots, I_k \in N$, their union $I_1 \cup \cdots \cup I_k$ is not in B.
(N3) N contains all maximal elements of B.

Let $\mathcal{N}(B)$ be the poset of all nested families in B ordered by reverse inclusion.

Theorem 5.1. *The poset of faces of the generalized permutohedron $P_n(\mathbf{Y})$ ordered by inclusion is isomorphic to $\mathcal{N}(B)$.*

This claim was independently discovered by E. M. Feichtner and B. Sturmfels [FS].

For a graph G on the set of vertices $[n+1]$, let B_G be the collection of nonempty subsets $I \subset [n+1]$ such that the induced graph $G|_I$ is connected. Then B_G satisfies conditions (B1) and (B2) of a building set. The generalized permutohedron associated with B_G is combinatorially equivalent to the *graph associahedron* constructed by Carr and Devadoss [CD] using blow-ups. In this case, it is enough to require condition (N2) only for pairs of subsets, in the definition of a nested family.

Remark 5.2. Since our generalized permutohedra include the associahedron, one can also call them *generalized associahedra*. However this name is already reserved for a different generalization of the associahedron studied by Fomin, Chapoton, and Zelevinsky [FCZ].

Let $f_B(q)$ be the *f-polynomial* of the generalized permutohedron:

$$f_B(q) = \sum_{i=0}^{n} f_i\, q^i = \sum_{N \in \mathcal{N}(B)} q^{n+1-|N|},$$

where f_i is the number of i-dimensional faces of $P_n(\mathbf{Y})$.

Let us say that a building set B is *connected* if it has a unique maximal element. Each building set B is a union of pairwise disjoint connected building sets, called the *connected components* of B. For a subset $S \subset [n+1]$, the *induced building set* is defined as $B|_S = \{I \in B \mid I \subset S\}$. In the case of building sets B_G associated with graphs G, notions of connected components and induced building sets correspond to similar notions for graphs.

Theorem 5.3. *The f-polynomial $f_B(q)$ is determined by the following recurrence relations:*

(1) *If B consists of a single singleton, then $f_B(q) = 1$.*

(2) *If B has connected components $B_1, \dots, B_k$, then*

$$f_B(q) = f_{B_1}(q) \cdots f_{B_k}(q).$$

(3) *If $B \subset 2^{[n+1]}$ is a connected building and $n \geq 1$, then*

$$f_B(q) = \sum_{S \subsetneq [n+1]} q^{n-|S|} f_{B|_S}(q).$$

Define the *generalized Catalan number*, for a building set B, as the number $C(B) = f_B(0)$ of vertices of the corresponding generalized permutohedron $P_n(\mathbf{Y})$. These numbers are given by the recurrence relations similar to the relations in Theorem 5.3. (But in (3) we sum only over subsets $S \subset [n+1]$ of cardinality n).

For a graph G, let $C(G) = C(B_G)$. Let T_{pqr} be the graph that has a central node with 3 attached chains of with p, q and r vertices. For example, T_{111} is the Dynkin diagram of type D_4. The above recurrence relations imply the following expression for the generating function for generalized Catalan numbers:

$$\sum_{p,q,r \geq 0} C(T_{pqr})\, x^p y^q z^r = \frac{C(x)\, C(y)\, C(z)}{1 - x\, C(x) - y\, C(y) - z\, C(z)},$$

where $C(x) = \sum_{n \geq 0} C_n x^n = \frac{1-\sqrt{1-4x}}{2x}$ is the generating function for the usual Catalan numbers.

Proposition 5.4. *For the Dynkin diagram of type A_n, we have $C(A_n) = C_n = \frac{1}{n+1}\binom{2n}{n}$ is the usual Catalan number. For the extended Dynkin diagram of type $\hat{A}_n$, we have $C(\hat{A}_n) = \binom{2n}{n}$. For the Dynkin diagram of type D_n, the corresponding Catalan number is*

$$C(D_n) = 2\, C_n - 2\, C_{n-1} - C_{n-2}.$$

Remark 5.5. One can define the generalized Catalan number for any Lie type. However this number does not depend on multiplicity of edges in the Dynkin diagram. Thus Lie types B_n and C_n give the same (usual) Catalan number as type A_n.

6. Mixed Eulerian numbers

Let us return to the usual permutohedron $P_n(x_1, \ldots, x_{n+1})$. Let us use the coordinates $z_1, \ldots, z_n$ related to $x_1, \ldots, x_{n+1}$ by

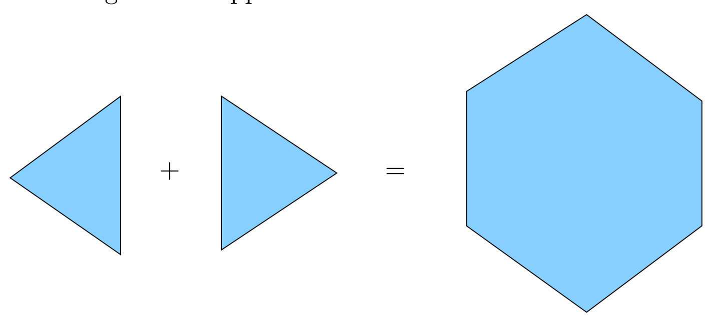
$$z_1 = x_1 - x_2,\ z_2 = x_2 - x_3,\ \cdots,\ z_n = x_n - x_{n+1}$$

This coordinate system is canonically defined for an arbitrary Weyl group as the coordinate system in the weight space given by the fundamental weights.

The permutohedron P_n can be written as the Minkowski sum

$$P_n = z_1\,\Delta_{1n} + z_2\,\Delta_{2n} + \cdots + z_n\,\Delta_{nn}$$

of the *hypersimplices* $\Delta_{kn} = P_n(1, \ldots, 1, 0, \ldots, 0)$ (with k 1's). For example, the hexagon can be express as the Minkowski sum of the hypersimplices Δ_{12} and Δ_{22}, which are two triangles with opposite orientations:



This implies that the volume of P_n can be written as

$$\mathrm{Vol}\ P_n = \sum_{c_1,\ldots,c_n} A_{c_1,\ldots,c_n} \frac{z_1^{c_1}}{c_1!} \cdots \frac{z_n^{c_n}}{c_n!},$$

where the sum is over $c_1, \ldots, c_n \geq 0$, $c_1 + \cdots + c_n = n$, and

$$A_{c_1,\ldots,c_n} = \mathrm{MixedVolume}(\Delta_{1n}^{c_1}, \ldots, \Delta_{nn}^{c_n}) \in \mathbb{Z}_{>0}$$

is the mixed volume of hypersimplices. In particular, $n!\,V_n$ is a positive integer polynomial in $z_1, \ldots, z_n$. Let us call the integers $A_{c_1,\ldots,c_n}$ the *Mixed Eulerian numbers*.

Examples: The mixed Eulerian numbers are marked in bold.

$$V_1 = \mathbf{1}\,z_1;$$
$$V_2 = \mathbf{1}\,\frac{z_1^2}{2} + \mathbf{2}\,z_1 z_2 + \mathbf{1}\,\frac{z_2^2}{2};$$
$$V_3 = \mathbf{1}\,\frac{z_1^3}{3!} + \mathbf{2}\,\frac{z_1^2}{2} z_2 + \mathbf{4}\,z_1 \frac{z_2^2}{2} + \mathbf{4}\,\frac{z_2^3}{3!} + \mathbf{3}\,\frac{z_1^2}{2} z_3 + \mathbf{6}\,z_1 z_2 z_3 +$$
$$+\mathbf{4}\,\frac{z_2^2}{2} z_3 + \mathbf{3}\,z_1 \frac{z_3^2}{2} + \mathbf{2}\,z_2 \frac{z_3^2}{2} + \mathbf{1}\,\frac{z_3^3}{3!}.$$

Theorem 6.1. *Mixed Eulerian numbers have the following properties:*

1. $A_{c_1,\ldots,c_n}$ *are positive integers defined for* $c_1, \ldots, c_n \geq 0$, $c_1 + \cdots + c_n = n$.
2. $\sum \frac{1}{c_1!\cdots c_n!} A_{c_1,\ldots,c_n} = (n+1)^{n-1}$.
3. $A_{0,\ldots,0,n,0,\ldots,0}$ *(n in the k-th position) is the usual Eulerian number* A_{kn}, *i.e., the number of permutations in* S_n *with* k *descents.*

(4) $A_{0,\dots,0,i,n-i,0,\dots,0}$ *(with* k *0's in front of* $i, n-i$*) is equal to the number of permutations* $w \in S_{n+1}$ *with* k *descents and* $w(n+1) = i+1$*.*
(5) $A_{1,\dots,1} = n!$
(6) $A_{k,0,\dots,0,n-k} = \binom{n}{k}$
(7) $A_{c_1,\dots,c_n} = 1^{c_1} 2^{c_2} \cdots n^{c_n}$ *if* $c_1 + \cdots + c_i \geq i$*, for* $i = 1, \dots, n$*.*
There are exactly $C_n = \frac{1}{n+1}\binom{2n}{n}$ *such sequences* $(c_1, \dots, c_n)$*.*
(8) $\sum A_{c_1,\dots,c_n} = n!\, C_n$*.*

Property (3) follows from the well-know fact that $A_{kn} = n! \operatorname{Vol} \Delta_{kn}$; and property (4) follows from the result of [ERS] about the mixed volume of two adjacent hypersimplices. Property (8) was numerically noticed by Richard Stanley. Moreover, he conjectured the following claim.

Theorem 6.2. *Let us write* $(c_1, \dots, c_n) \sim (c'_1, \dots, c'_n)$ *whenever* $(c_1, \dots, c_n, 0)$ *is a cyclic shift of* $(c'_1, \dots, c'_n, 0)$*. Then, for fixed* $(c_1, \dots, c_n)$*, we have*

$$\sum_{(c'_1,\dots,c'_n)\sim(c_1,\dots,c_n)} A_{c'_1,\dots,c'_n} = n!$$

In other words, the sum of mixed Eulerian numbers in each equivalence class is $n!$*.*
There are exactly the Catalan number $\frac{1}{n+1}\binom{2n}{n}$ *equivalence classes of sequences.*

For example, we have $A_{1,\dots,1} = n!$ and $A_{n,0,\dots,0} + A_{0,n,0,\dots,0} + A_{0,0,n,\dots,0} + \cdots + A_{0,\dots,0,n} = n!$, because the sum of usual Eulerian numbers $\sum_k A_{kn}$ is $n!$.

Remark 6.3. Every equivalence class contains exactly one sequence $(c_1, \dots, c_n)$ such that $c_1 + \cdots + c_i \geq i$, for $i = 1, \dots, n$. For this special sequence, the mixed Eulerian number is given by the simple product $A_{c_1,\dots,c_n} = 1^{c_1} \cdots n^{c_n}$; see Theorem 6.1.(7).

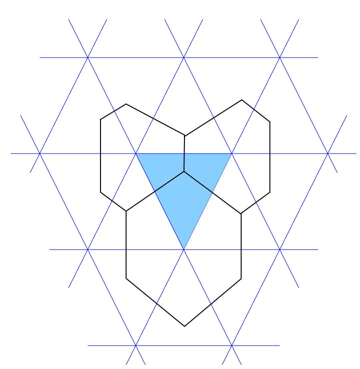
Theorem 6.2 follows from the following claim.

Theorem 6.4. *Let* $\hat{U}_n(z_1, \dots, z_{n+1}) = \operatorname{Vol} P_n$*. (It does not depend on* z_{n+1}*.)*

$$\hat{U}_n(z_1, \dots, z_{n+1}) + \hat{U}_n(z_{n+1}, z_1, \dots, z_n) + \cdots + \hat{U}_n(z_2, \dots, z_{n+1}, z_1) = \\ = (z_1 + \cdots + z_{n+1})^n$$

This claim extends to any Weyl group W. It has a simple geometric proof using alcoves of the associated affine Weyl group. Cyclic shifts come from symmetries of type A extended Dynkin diagram.

Idea of Proof. The volume of the fundamental alcove times $|W|$ equals the sum of volumes of $n+1$ adjacent permutohedra. For example, the 6 areas of the blue triangle on the following picture is the sum of the areas of three hexagons. □



Corollary 6.5. *Fix $z_1, \dots, z_{n+1}$, $\lambda_1, \dots, \lambda_{n+1}$ such that $\lambda_1 + \cdots + \lambda_{n+1} = 0$. Symmetrizing the expression*

$$\frac{1}{n!}\,\frac{(\lambda_1 z_1 + (\lambda_1+\lambda_2)z_2 + \cdots (\lambda_1+\cdots+\lambda_{n+1})z_{n+1})^n}{(\lambda_1-\lambda_2)\cdots(\lambda_n-\lambda_{n+1})}$$

with respect to $(n+1)!$ permutations of $\lambda_1, \dots, \lambda_{n+1}$ and $(n+1)$ cyclic permutations of $z_1, \dots z_{n+1}$, we obtain

$$(z_1 + \cdots + z_{n+1})^n.$$

It would be interesting to find a direct proof of this claim.

7. Third formula

Let us give a combinatorial interpretation for the mixed Eulerian numbers based on plane binary trees.

Let T be an *increasing plane binary tree* with n nodes labelled $1, \dots, n$. It is well-known that the number of such trees is $n!$. Let v_i be the node of T labelled i, for $i = 1, \dots, n$. In particular, v_1 is the top node of T. Let us define a different labelling of the nodes $v_1, \dots, v_n$ of T by numbers $d_1, \dots, d_n \in [n]$ based on the *depth-first search algorithm*. This labelling is uniquely characterized by the following condition: For any node v_i in T and any v_j in the left (respectively, right) branch of v_i, we have $d_j < d_i$ (respectively, $d_i < d_j$). In particular, for the left-most node v_l in T, we have $d_l = 1$ and, for the right-most node v_r, we have $d_r = n$. Then, for any node v_i, the numbers d_j, for all descendants v_j of v_i (including v_i), form a *consecutive interval* $[l_i, r_i]$ of integers. (In particular, $l_i \le d_i \le r_i$.)

Remark 7.1. For a plane binary tree T, the collection N of intervals $[l_i, r_i]$, $i = 1, \dots, n$, is a maximal nested family for the building set B formed by all consecutive intervals in $[n]$, i.e., the building set for the usual associahedron, see Section 5. The map $T \mapsto N$, is a bijection between plane binary trees and vertices of the associahedron in Loday's realization [L].

Let $\mathbf{i} = (i_1, \dots, i_n) \in [n]^n$ be a sequence of integers. Let us say that an increasing plane binary tree T is $\mathbf{i}$-*compatible* if, $i_k \in [l_k, r_k]$, for $k = 1, \dots, n$. For a node v_k in such a tree, define its weight as

$$wt(v_k) = \begin{cases} \frac{i_k - l_k + 1}{d_k - l_k + 1} & \text{if } i_k \le d_k \\ \frac{r_k - i_k + 1}{r_k - d_k + 1} & \text{if } i_k \ge d_k \end{cases}$$

Define the $\mathbf{i}$-*weight* of an $\mathbf{i}$-compatible tree T as

$$wt(\mathbf{i}, T) = \prod_{k=1}^{n} wt(v_k).$$

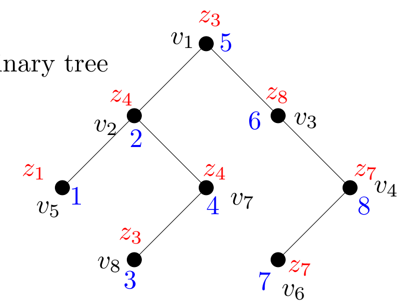
It is not hard to see that $n!\, wt(\mathbf{i}, T)$ is always a positive integer.

Here is an example of an $\mathbf{i}$-compatible tree T, for $\mathbf{i} = (3, 4, 8, 7, 1, 7, 4, 3)$. The labels d_k of the nodes v_k are $5, 2, 6, 8, 1, 7, 4, 3$. (They are shown on picture in blue color.) The intervals $[l_k, r_k]$ are $[1, 8]$, $[1, 4]$, $[6, 8]$, $[7, 8]$, $[1, 1]$, $[7, 7]$, $[3, 4]$, $[3, 3]$. We also marked each node v_k by the variable z_{i_k}. The $\mathbf{i}$-weight of this tree is $wt(\mathbf{i}, T) = \frac{3}{5} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} \cdot \frac{1}{1} \cdot \frac{2}{2} \cdot \frac{1}{1}$.

an $\mathbf{i}$-compartible increasing plane binary tree

$T =$

$\mathbf{i} = (3, 4, 8, 7, 1, 7, 4, 3)$



Theorem 7.2. *The volume of the permutohedron is equal to*

$$V_n = \sum_{\mathbf{i} \in [n]^n} z_{i_1} \cdots z_{i_n} \sum_{T \text{ is } \mathbf{i}\text{-compatible}} wt(\mathbf{i}, T).$$

where the first sum is over n^n sequences $\mathbf{i} = (i_1, \ldots, i_n)$ and second sum is over $\mathbf{i}$-compatible increasing plane binary trees with n nodes.

Let us give a combinatorial interpretation for the mixed Eulerian numbers.

Theorem 7.3. *Let $(i_1, \ldots, i_n)$ be any sequence such that $z_{i_1} \cdots z_{i_n} = z_1^{c_1} \cdots z_n^{c_n}$. Then*

$$A_{c_1,\ldots,c_n} = \sum_{T \text{ is } \mathbf{i}\text{-compatible}} n!\, wt(\mathbf{i}, T),$$

where the sum is over $\mathbf{i}$-compatible increasing plane binary trees with n nodes.

Note that all terms $n!\, wt(T)$ in this formula are positive integers. Actually, this theorem gives not just one but $\binom{n}{c_1,\ldots,c_n}$ different combinatorial interpretations of the mixed Eulerian numbers $A_{c_1,\ldots,c_n}$ for each way to write $z_1^{c_1} \cdots z_n^{c_n}$ as $z_{i_1} \cdots z_{i_n}$.

The proof of this theorem is based on the following recurrence relation for the volume of the permutohedron. Let us write $\operatorname{Vol} P_n$ as a polynomial $U_n(z_1, \ldots, z_n)$.

Proposition 7.4. *For any $i = 1, \ldots, n$, we have*

$$\frac{\partial}{\partial z_i} U_n(z_1, \ldots, z_n) = \sum_{k=1}^{n} \binom{n+1}{k} wt_{i,k,n}\, U_{k-1}(z_1, \ldots, z_{k-1})\, U_{n-k-1}(z_{k+1}, \ldots, z_n),$$

where

$$wt_{i,k,n} = \frac{1}{n} \begin{cases} i(n-k) & \text{if } 1 \le i \le k; \\ k(n-k) & \text{if } k \le i \le n. \end{cases}$$

Idea of Proof. The derivative $\partial U_n / \partial z_i$ is the rate of change of the volume $\operatorname{Vol} P_n$ of the permutohedron as we move its generating vertex in the direction of the coordinate z_k. (More generally, in the direction of the k-th fundamental weight.) It can be written as a sum of areas of facets of P_n scaled by some factors. Each facet of P_n is a product $P_{k-1} \times P_{n-k-1}$ of two smaller permutohedra. There are exactly $\binom{n+1}{k}$ facets like this. The corresponding factor $wt_{i,k,n}$ tells how fast the facet moves as we shift the generating vertex of P_n.

This formula can be extended to the weight polytope for an arbitrary Weyl group W. In general, the coefficient $wt_{i,k,n}$ equals the inner product of two fundamental weights (ω_i, ω_k). $\square$

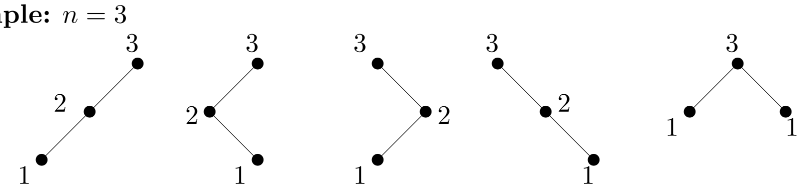
Comparing different formulas for V_n, we obtain many interesting combinatorial identities. For example, we have the following claim.

Corollary 7.5. *We have*

$$(n+1)^{n-1} = \sum_T \frac{n!}{2^n} \prod_{v \in T} \left(1 + \frac{1}{h(v)}\right),$$

where is sum is over unlabeled plane binary trees T on n nodes, and $h(v)$ denotes the "hook-length" of a node v in T, i.e., the number of descendants of the node v (including v).

Example: $n = 3$



hook-lengths of binary trees

The identity says that

$$(3+1)^2 = 3 + 3 + 3 + 3 + 4.$$

This identity was combinatorially proved by Seo [S].

References

[CD] M. Carr, S. Devadoss: Coxeter complexes and graph associahedra, `math.QA/0407229`.

[DJS] M. Davis, T. Januszkiewicz, R. Scott: Nonpositive curvature of blow-ups. *Selecta Math. (N.S.)* **4** (1998), no. 4, 491–547.

[DP] C. De Concini, C. Procesi: Wonderful models for subspace arrangements, *Selecta Math. (N.S.)* **1** (1995), 459–494.

[ERS] R. Ehrenborg, M. Readdy, E. Steingrímsson: Mixed volumes and slices of the cube, *J. Combin. Theory Ser. A* **81** (1998), no. 1, 121–126.

[FS] E. M. Feichtner, B. Sturmfels: Matroid polytopes, nested sets and Bergman fans, `math.CO/0411260`.

[FCZ] S. Fomin, F. Chapoton, A. Zelevinsky: Polytopal realizations of generalized associahedra, *Canadian Mathematical Bulletin* **45** (2002), 537-566.

[L] J.-L. Loday: Realization of the Stasheff polytope, `math.AT/0212126`.

[PK] A. V. Pukhlikov, A. G. Khovanskii: Finitely additive measures of virtual polyhedra, *St. Petersburg Math. J.* **4** (1993), no. 2, 337–356.

[S] S. Seo: A combinatorial proof of Postnikov's identity and a generalized enumeration of labeled trees, `math.CO/0409323`.

[SP] R. P. Stanley, J. Pitman: A polytope related to empirical distributions, plane trees, parking functions, and the associahedron, *Discrete Comput. Geom.* **27** (2002), no. 4, 603–634.

Department of Mathematics, M.I.T., Cambridge, MA 02139

E-mail address: `apost (at) math (dot) mit (dot) edu`

URL: `http://www.math-mit.edu/~apost/`

KAZHDAN-LUSZTIG IMMANANTS AND PRODUCTS OF MATRIX MINORS

BRENDON RHOADES AND MARK SKANDERA

Abstract. We define a family of polynomials of the form $\sum f(\sigma)x_{1,\sigma(1)}\cdots x_{n,\sigma(n)}$ in terms of the Kazhdan-Lusztig basis $\{C'_w(1)\,|\,w\in S_n\}$ for the symmetric group algebra $\mathbb{C}[S_n]$. Using this family, we obtain nonnegativity properties of polynomials of the form $\sum c_{I,I'}\Delta_{I,I'}(x)\Delta_{\overline{I},\overline{I'}}(x)$. In particular, we show that the application of certain of these polynomials to Jacobi-Trudi matrices yields symmetric functions which are nonnegative linear combinations of Schur functions.

Résumé. Nous definissons une famille de polynômes $\sum f(\sigma)x_{1,\sigma(1)}\cdots x_{n,\sigma(n)}$ en terme de la base $\{C'_w(q)\,|\,w\in S_n\}$ de Kazhdan-Lusztig pour l'algébra $\mathbb{C}[S_n]$. En utilisant cette famille, nous obtenons quelques propriétées des polynômes totalement non négatifs de la forme $\sum c_{I,I'}\Delta_{I,I'}(x)\Delta_{\overline{I},\overline{I'}}(x)$. En particulier, nous démontrons que l'application des certains de ces polynômes aux matrices de Jacobi-Trudi rapporte des fonctions symmétriques qui sont des combinaisons linéaires non négatives des fonctions de Schur.

1. Introduction

Since its introduction in [14], the Kazhdan-Lusztig basis $\{C'_w(q)\,|\,w\in S_n\}$ of the Hecke algebra $H_n(q)$ has found many applications related to algebraic geometry, combinatorics, and Lie theory. One such application, due to Haiman [12], clarifies three nonnegativity properties of certain polynomials which arise in the representation theory of $H_n(q)$. Years later, two of these nonnegativity properties were observed in a family of polynomials which arise in the study of inequalities satisfied by minors of totally nonnegative matrices [4, 18]. Building upon the arguments of Haiman [12], we will show that this family posesses the third nonnegativity property as well.

The nonnegativity properties are as follows. Let $x=(x_{ij})$ be a generic square matrix. For each pair (I,I') of subsets of $[n]=\{1,\ldots,n\}$, define $\Delta_{I,I'}(x)$ to be the (I,I') minor of x, i.e., the determinant of the submatrix of x corresponding to rows I and columns I'. A real matrix is called *totally nonnegative* (TNN) if each of its minors is nonnegative. A polynomial $p(x)=p(x_{1,1},\ldots,x_{n,n})$ in n^2 variables is called

Date: February 22, 2005.

totally nonnegative if for every TNN matrix A, the number

$$p(A) \underset{def}{=} p(a_{1,1}, \dots, a_{n,n})$$

is nonnegative. Much current work in total nonnegativity is motivated by problems in quantum Lie theory. (See e.g. [7, 15, 28].) The strong connection between total nonnegativity and Jacobi-Trudi matrices leads to more nonnegativity properties. (See [17, 20] for information on Jacobi-Trudi matrices, and [8] for connections to total nonnegativity.) We will call the polynomial $p(x)$ *Schur nonnegative* (SNN) if for every $n \times n$ Jacobi-Trudi matrix A, the symmetric function $p(A)$ is equal to a nonnegative linear combination of Schur functions. We will also call such a symmetric function Schur nonnegative. Much current work in Schur nonnegativity is motivated by problems concerning the cohomology ring of the Grassmannian variety. (See e.g. [2, 6].) In analogy to Schur nonnegativity, we will call $p(x)$ *monomial nonnegative* (MNN) if for every $n \times n$ Jacobi-Trudi matrix A, $p(A)$ is equal to a nonnegative linear combination of monomial symmetric functions. We will also call such a symmetric function monomial nonnegative. Since each Schur function is itself monomial nonnegative, any SNN polynomial must also be MNN.

Some nontrivial classes of polynomials possessing the TNN, SNN and MNN properties are contained in the complex span of the monomials $\{x_{1,w(1)} \cdots x_{n,w(n)} \,|\, w \in S_n\}$. We will call such polynomials *immanants*. In particular, for every function $f : S_n \to \mathbb{C}$ we define the *f-immanant* (as in [21, Sec. 3]) by

$$\mathrm{Imm}_f(x) \underset{\mathrm{def}}{=} \sum_{w \in S_n} f(w) x_{1,w(1)} \cdots x_{n,w(n)}.$$

Some familiar immanants are those of the form $\mathrm{Imm}_{\chi^\lambda}(x)$, where χ^λ is an irreducible character of S_n. Goulden and Jackson conjectured [10] and Greene proved [11] these immanants to be MNN. Stembridge then conjectured [26] these immanants to be TNN and SNN, and he [23] and Haiman [12] proved these two conjectures. (See [12, 13, 22, 23, 24] for related conjectures and results.) Other immanants of the form

$$\Delta_{J,J'}(x)\Delta_{\overline{J},\overline{J'}}(x) - \Delta_{I,I'}(x)\Delta_{\overline{I},\overline{I'}}(x) \tag{1.1}$$

characterize the inequalities satisfied by products of two minors of TNN matrices. (Equivalently, these characterize the inequalities satisfied by products of two entries of the exterior algebra representation of TNN elements of $GL_n(\mathbb{C})$.) Fallat, Gekhtman and Johnson characterized [4] the TNN immanants of the form (1.1), in the principal minor case ($I = I'$, etc.) A characterization of the general case followed in [16, 18], as did a proof that all such TNN immanants are MNN.

In Section 2 we define more immanants in terms of the Kazhdan-Lusztig basis of $\mathbb{C}[S_n]$. We then use the Schur nonnegativity of these Kazhdan-Lusztig immanants in Section 3 to prove the Schur nonnegativity of all TNN immanants of the form

(1.1). More properties of the Kazhdan-Lusztig immanants and open problems follow in Section 4.

2. Kazhdan-Lusztig immanants

Let q be a formal parameter and define the *Hecke algebra* $H_n(q)$ to be the $\mathbb{C}[q^{1/2}, q^{-1/2}]$-algebra generated by elements $T_{s_1}, \ldots, T_{s_{n-1}}$, 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 each permutation w we 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 $q = 1$ gives 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

$$C'_v(q) = \sum_{u \leq v} P_{u,v}(q) q^{-\ell(v)/2} T_u, \tag{2.1}$$

where the comparison of permutations is in the Bruhat order, and

$$\{P_{u,v}(q) \,|\, u, v \in S_n\}$$

are certain polynomials in q, known as the *Kazhdan-Lusztig polynomials* [14]. Solving the equations (2.1) for T_v, we have

$$T_v = \sum_{u \leq v} (-1)^{\ell(v)-\ell(u)} P_{w_0 v, w_0 u}(q) q^{\ell(u)/2} C'_u(q),$$

where w_0 is the longest permutation in S_n.

For each permutation v in S_n define the function $f_v : S_n \to \mathbb{C}$ by

$$f_v(w) = (-1)^{\ell(w)-\ell(v)} P_{w_0 w, w_0 v}(1).$$

Extending these functions linearly to $\mathbb{C}[S_n]$, we see that they are dual to the Kazhdan-Lusztig basis in the sense that

$$f_v(C'_w(1)) = \delta_{v,w}.$$

We will denote the f_v-immanant by

$$\mathrm{Imm}_v(x) \underset{\text{def}}{=} \sum_{w \geq v} f_v(w) x_{1,w(1)} \cdots x_{n,w(n)},$$

and will call these immanants the *Kazhdan-Lusztig immanants*. In the case that v is the identity permutation, we obtain the determinant.

Results in [12, 23] imply that the Kazhdan-Lusztig immanants are TNN and SNN. To give brief proofs, we shall consider the following elements of $H_n(q)$. Given indices $1 \leq i \leq j \leq n$, define $z_{[i,j]}$ to be the element of $H_n(q)$ which is the sum of elements T_w corresponding to permutations w in the subgroup of S_n generated by $s_i, \ldots, s_{j-1}$.

Proposition 2.1. *Let z be an element of $H_n(q)$ of the form*

$$z = z_{[i_1,j_1]} \cdots z_{[i_r,j_r]}. \tag{2.2}$$

Then we have

$$z = \sum_{w \in S_n} p_{z,w}(q) C'_w(q),$$

where the expressions $p_{z,w}(q)$ are Laurent polynomials in $q^{1/2}$ with nonnegative coefficients. In particular, an element of the form (2.2) in $\mathbb{C}[S_n]$ is equal to a nonnegative linear combination of the Kazhdan-Lusztig basis elements $\{C'_w(1) \,|\, w \in S_n\}$.

Proof. Let $s_{[i,j]}$ be the longest permutation in the subgroup generated by $s_i, \ldots s_{j-1}$. By [12, Prop. 3.1], we have

$$z_{[i,j]} = q^{\ell(w)/2} C'_{s_{[i,j]}}(q).$$

A result of Springer [19] implies that for every pair (u, v) of permutations in S_n, we have

$$C'_u(q) C'_v(q) = \sum_{w \in S_n} f^w_{u,v}(q) C'_w(q),$$

where the expressions $f^w_{u,v}(q)$ are Laurent polynomials in $q^{1/2}$ with nonnegative coefficients. (See also [12, Appendix].) $\square$

Proposition 2.2. *For each permutation w in S_n, the Kazhdan-Lusztig immanant $\mathrm{Imm}_w(x)$ is totally nonnegative.*

Proof. For any complex matrix A and any function $f : S_n \to \mathbb{C}$ we have

$$\mathrm{Imm}_f(A) = \sum_z c_z f(z),$$

where the sum is over elements z of $\mathbb{C}[S_n]$ of the form (2.2), and the coefficients c_z depend on A. If A is a totally nonnegative matrix, then these coefficients are real and nonnegative. (See, e.g., [16, Lem. 2.5], [23, Thm. 2.1].)

Let A be a TNN matrix. By Propositiion 2.1 we have

$$\begin{aligned}\mathrm{Imm}_w(A) &= \sum_z c_z f_w(z) \\ &= \sum_z c_z \sum_v p_{z,v}(1) f_w(C'_v(1)) \\ &= \sum_z c_z p_{z,w}(1) \\ &\geq 0.\end{aligned}$$

□

The following easy consequence of [12, Thm. 1.5] implies the Schur nonnegativity of the Kazhdan-Lusztig immanants. Following [12], we define a *generalized Jacobi-Trudi matrix* to be a finite matrix whose i, j entry is the homogeneous symmetric function $h_{\mu_i - \nu_i}$, where $\mu = (\mu_1, \ldots, \mu_n)$ and $\nu = (\nu_1, \ldots, \nu_n)$ are weakly decreasing nonnegative sequences, and by convention $h_m = 0$ if m is negative. Thus each generalized Jacobi-Trudi matrix is constructed from an ordinary Jacobi-Trudi matrix by repeating some rows and/or columns.

Proposition 2.3. *For each permutation w in S_n, and each $n \times n$ generalized Jacobi-Trudi matrix A, the symmetric function* $\mathrm{Imm}_w(A)$ *is Schur nonnegative.*

Proof. By [12, Thm. 1.5], we have

$$\sum_{v \in S_n} a_{1,v(1)} \cdots a_{n,v(n)} v = \sum_u g_{v,u}(A) C'_u(1),$$

where $g_{v,u}(A)$ is a Schur nonnegative symmetric function which depends upon A. Applying the function f_w to both sides of this equations, we have

$$\begin{aligned}\mathrm{Imm}_w(A) &= \sum_u g_{w,u}(A) f_w(C'_u(1)) \\ &= g_{w,w}(A).\end{aligned}$$

□

3. Main Results

Studying inequalities satisfied by products of principal minors of TNN matrices, Fallat, Gekhtman and Johnson [4, Thm. 4.6] characterized all TNN immanants of the form

$$\Delta_{J,J}(x)\Delta_{\overline{J},\overline{J}}(x) - \Delta_{I,I}(x)\Delta_{\overline{I},\overline{I}}(x),$$

(where $\overline{I} = [n] \smallsetminus I$, $\overline{J} = [n] \smallsetminus J$) and more generally, all TNN polynomials of the form

$$\Delta_{J,J}(x)\Delta_{L,L}(x) - \Delta_{I,I}(x)\Delta_{K,K}(x).$$

This result was generalized in [18, Thm. 3.2] as follows.

Proposition 3.1. *Let I, J, K, L be subsets of $[n]$ and let I', J', K', L' be subsets of $[n']$, and define the subsets I'', J'', K'', L'' of $[n+n']$ by*

$$\begin{aligned} I'' &= I \cup \{n+n'+1-i \,|\, i \in K'\}, \\ K'' &= K \cup \{n+n'+1-i \,|\, i \in I'\}, \\ J'' &= J \cup \{n+n'+1-i \,|\, i \in L'\}, \\ L'' &= L \cup \{n+n'+1-i \,|\, i \in J'\}. \end{aligned} \tag{3.1}$$

Then the polynomial

$$\Delta_{J,J'}(x)\Delta_{L,L'}(x) - \Delta_{I,I'}(x)\Delta_{K,K'}(x) \tag{3.2}$$

is totally nonnegative if and only if the sets $I, \ldots, L$, $I', \ldots, L'$ satisfy

$$\begin{aligned} I \cup K = J \cup L, \qquad I' \cup K' = J' \cup L', \\ I \cap K = J \cap L, \qquad I' \cap K' = J' \cap L', \end{aligned} \tag{3.3}$$

and for each subinterval B of $[n+n']$ the sets $I'', \ldots, K''$ satisfy

$$\max\{|B \cap J''|, |B \cap L''|\} \leq \max\{|B \cap I''|, |B \cap K''|\}. \tag{3.4}$$

The proof in [18] shows that these polynomials are MNN as well. (See [16, Cor. 6.1].) Two combinatorial alternatives to the system of inequalities (3.4) are given in [16, Thms. 5.2, 5.4]. The second of these proves the total nonnegativity of the polynomials (3.2) by relating them to TNN immanants defined in terms of the Temperley-Lieb algebra.

Given a formal parameter ξ, we define the *Temperley-Lieb algebra* $TL_n(\xi)$ to be the $\mathbb{C}[\xi]$-algebra generated by elements $t_1, \ldots, t_{n-1}$ subject to the relations

$$\begin{aligned} t_i^2 &= \xi t_i, && \text{for } i = 1, \ldots, n-1, \\ t_i t_j t_i &= t_i, && \text{if } |i-j| = 1, \\ t_i t_j &= t_j t_i, && \text{if } |i-j| \geq 2. \end{aligned}$$

The rank of $TL_n(\xi)$ as a $\mathbb{C}[\xi]$-module is well-known to be $\frac{1}{n+1}\binom{2n}{n}$, and a natural basis is given by the elements of the form $t_{i_1} \cdots t_{i_\ell}$, where $i_1 \cdots i_\ell$ is a reduced word for a 321-avoiding permutation in S_n. We shall call these elements the standard basis elements of $TL_n(\xi)$, or simply *the basis elements* of $TL_n(\xi)$.

The Temperley-Lieb algebra may be realized as a quotient of the Hecke algebra by

$$H_n(q)/(z_{[1,3]}) \cong TL_n(q^{1/2} + q^{-1/2}),$$

where the element $z_{[1,3]}$ of $H_n(q)$ is defined as before Proposition 2.1. We will let θ be the homomorphism

$$\begin{aligned} H_n(q) &\to TL_n(q^{1/2}+q^{-1/2}) \\ q^{-1/2}(T_{s_i}+1) &\mapsto t_i. \end{aligned}$$

(See e.g. [5], [9, Sec. 2.1, Sec. 2.11], [27, Sec. 7].)

Immanants called *Temperley-Lieb immanants* in [16] were defined in terms of the homomorphism θ, specialized at $q = 1$. For each basis element τ of $TL_n(2)$, let $f_\tau : S_n \to \mathbb{R}$ be the function defined by

$$f_\tau(v) = \text{ coefficient of } \tau \text{ in } \theta(T_v),$$

and let

$$\mathrm{Imm}_\tau(x) = \sum_{w \in S_n} f_\tau(w) x_{1,w(1)} \cdots x_{n,w(n)}$$

be the corresponding immanant. By [16, Thm. 3.1], the Temperley-Lieb immanants are TNN. Furthermore, the following result shows that the Temperley-Lieb immanants are Kazhdan-Lusztig immanants. To prove this, we define for each 321-avoiding permutation w in S_n an element $D_w(q)$ of $H_n(q)$ as follows. For any reduced word $i_1 \cdots i_\ell$ for w, define

$$D_w(q) \underset{\mathrm{def}}{=} q^{-1/\ell}(T_{s_{i_1}}+1)\cdots(T_{s_{i_\ell}}+1).$$

(This element does not depend upon the particular reduced word.) The element $D_w(q)$ satisfies

$$\theta(D_w(q)) = t_{i_1} \cdots t_{i_\ell},$$

and it follows that the set

$$\{\theta(D_w(q)) \,|\, w \text{ a 321-avoiding permutation } \}$$

is equal to the standard basis of $TL_n(q^{1/2}+q^{-1/2})$.

Proposition 3.2. *Let w be any 321-avoiding permutation in S_n, and define $\tau = \theta(D_w(1))$. Then the Temperley-Lieb immanant $\mathrm{Imm}_\tau(x)$ is equal to the Kazhdan-Lusztig immanant $\mathrm{Imm}_w(x)$.*

Proof. Let v be any permutation in S_n. Then we have

$$v = \sum_{u \le v} (-1)^{\ell(v)-\ell(u)} P_{w_0v,w_0u}(1) C'_u(1).$$

The coefficient of $x_{1,v(1)} \cdots x_{n,v(n)}$ in $\mathrm{Imm}_\tau(x)$ is equal to $f_\tau(v)$, which is the coefficient of τ in

$$\theta(v) = \sum_{u \le v} (-1)^{\ell(v)-\ell(u)} P_{w_0v,w_0u}(1) \theta(C'_u(1)). \tag{3.5}$$

A result of Fan and Green [5, Thm. 3.8.2] implies that we have

$$\theta(C'_w(q)) = \begin{cases} \theta(D_w(q)) & \text{if } w \text{ is 321-avoiding,} \\ 0 & \text{otherwise.} \end{cases}$$

(See also [3, Thm. 4].) We may therefore assume that each permutation u appearing in (3.5) is 321-avoiding, and we may rewrite the sum as

$$\theta(v) = \sum_{u \leq v} (-1)^{\ell(v)-\ell(u)} P_{w_0 v, w_0 u}(1) \theta(D_u(1)).$$

The coefficient of $\tau = \theta(D_w(1))$ in this expression is $(-1)^{\ell(v)-\ell(w)} P_{w_0 v, w_0 u}(1)$. But this is precisely the coefficient of $x_{1,v(1)} \cdots x_{n,v(n)}$ in $\mathrm{Imm}_w(x)$. □

Thus the Temperley-Lieb immanants are precisely the Kazhdan-Lusztig immanants corresponding to 321-avoiding permutations. From the Schur nonnegativity of the Kazhdan-Lusztig immanants, it then follows that all TNN polynomials of the form (3.2) are SNN.

Theorem 3.3. *Let I, J, K, L be subsets of $[n]$, let I', J', K', L' be subsets of $[n']$, and suppose that these satisfy the conditions of Proposition 3.1. Then the polynomial*

$$\Delta_{J,J'}(x)\Delta_{L,L'}(x) - \Delta_{I,I'}(x)\Delta_{K,K'}(x) \tag{3.6}$$

is Schur nonnegative.

Proof. Define $r = |I| + |K|$, and let $k_1 \leq \cdots \leq k_r$ be the nondecreasing rearrangement of the elements of I and K, including repeated elements. Define $k'_1, \ldots, k'_r$ analogously, and let y be the $r \times r$ matrix whose i, j entry is the variable x_{k_i, k'_j}. Thus y is the matrix obtained from x by duplicating rows whose indices belong to $I \cap K$ and columns whose indices belong to $I' \cap K'$.

By Proposition 3.1, the polynomial (3.6) is TNN, and by [16, Prop. 5.3, Thm. 5.4] we have

$$\Delta_{J,J'}(x)\Delta_{L,L'}(x) - \Delta_{I,I'}(x)\Delta_{K,K'}(x) = \sum_{\tau} \mathrm{Imm}_\tau(y),$$

where the sum is over a subset of basis elements of $TL_r(2)$. By Proposition 3.2 this is a sum of Kazhdan-Lustig immanants,

$$\Delta_{J,J'}(x)\Delta_{L,L'}(x) - \Delta_{I,I'}(x)\Delta_{K,K'}(x) = \sum_{w} \mathrm{Imm}_w(y), \tag{3.7}$$

for an appropriate set of 321-avoiding permutations w in S_r.

Now let A be an arbitrary $n \times n'$ Jacobi-Trudi matrix, and let B be the generalized Jacobi-Trudi matrix whose i, j entry is a_{k_i, k'_j}. Then the evaluation of the left-hand side of (3.7) at $x = A$ is equal to the evaluation of the right-hand side at $y = B$.

By Proposition 2.3, the resulting symmetric function on the right-hand side is SNN. Thus the polynomial $\Delta_{J,J'}(x)\Delta_{L,L'}(x) - \Delta_{I,I'}(x)\Delta_{K,K'}(x)$ is SNN. $\square$

Theorem 3.3 provides new machinery for proving that certain symmetric functions of the form $s_{\alpha/\kappa}s_{\beta/\lambda} - s_{\gamma/\mu}s_{\delta/\nu}$ are SNN. For example, the combinatorial test in [16, Thm. 4.2] makes it easy to see that for

$$J = \{i \in [n] \mid i \text{ odd } \}$$

and for any subsets I, I' of $[n]$, the immanant

$$\Delta_{J,J}(x)\Delta_{\overline{J},\overline{J}}(x) - \Delta_{I,I'}(x)\Delta_{\overline{I},\overline{I'}}(x)$$

is SNN. Choosing $n = 6$ and $I = \{1, 3, 4\}$, $I' = \{1, 2, 4\}$, we may apply this immanant,

$$\Delta_{135,135}(x)\Delta_{246,246}(x) - \Delta_{134,124}(x)\Delta_{256,356}(x)$$

to the Jacobi-Trudi matrix indexed by the skew shape 766655/22211,

$$\begin{bmatrix} h_5 & h_6 & h_7 & h_9 & h_{10} & h_{12} \\ h_3 & h_4 & h_5 & h_7 & h_8 & h_{10} \\ h_2 & h_3 & h_4 & h_6 & h_7 & h_9 \\ h_1 & h_2 & h_3 & h_5 & h_6 & h_8 \\ 0 & 1 & h_1 & h_3 & h_4 & h_6 \\ 0 & 0 & 1 & h_2 & h_3 & h_5 \end{bmatrix},$$

to see the Schur nonnegativity of the symmetric function

$$s_{864/32}s_{875/42} - s_{755/22}s_{855/31}.$$

4. Open Questions

The Littlewood-Richardson coefficients, defined by

$$s_\lambda s_\mu = \sum_\nu c^\nu_{\lambda,\mu} s_\nu,$$

and the inequalities satisfied by these coefficients have have interesting interpretations in algebraic geometry and representation theory. (See, e.g., [1, 6, 25].) A basic open question about these inequalities may be stated as follows.

Question 4.1. For what conditions on partitions α, β, γ, δ is the symmetric function $s_\alpha s_\beta - s_\gamma s_\delta$ Schur nonnegative? Equivalently, what conditions on these four partitions imply that $c^\nu_{\alpha,\beta} \geq c^\nu_{\gamma,\delta}$ for all ν?

Some conjectured sufficient conditions are given by Fomin, Fulton, Li and Poon [6, Conj. 2.8, Conj. 5.1]. Generalizing the second of these conjectures, Bergeron, Biagioli and Rosas [2, Conj. 2.9] have conjectured sufficient conditions for Schur nonnegativity of symmetric functions of the form

$$s_{\alpha/\kappa}s_{\beta/\lambda} - s_{\gamma/\mu}s_{\delta/\nu}. \tag{4.1}$$

It would be interesting to determine which of the conjectured sufficient conditions can be derived from Theorem 3.3. On the other hand it would be interesting to find symmetric functions of the form (4.1) for which Schur nonnegativity follows from Theorem 3.3, but not from the conjectured sufficient conditions.

The fact that Theorem 3.3 may be applied to *generalized* Jacobi-Trudi matrices highlights an important difference between the determinant and other Kazhdan-Lusztig immanants. Specifically, Kazhdan-Lusztig immanants do not vanish on a matrix having a pair of equal rows. It was shown in [16, Prop. 3.14] that Temperley-Lieb immanants vanish on matrices having *three* equal rows. This fact generalizes nicely to arbitrary Kazhdan-Lusztig immanants.

Proposition 4.1. *Let w be a permutation in S_n and suppose that the one-line notation $w(1)\cdots w(n)$ contains no decreasing subsequence of length k. Then $\mathrm{Imm}_w(x)$ vanishes on any $n\times n$ matrix having k equal rows or columns.*

Proof. Omitted. □

It would be interesting to generalize other determinantal formulas and identities to Kazhdan-Lusztig immanants.

Some work on immanants related to representations of S_n has led to the study of certain elements of $\mathbb{C}[S_n]$ associated to total nonnegativity. Following Stembridge [23], we define the *cone of total nonnegativity* to be the smallest cone in $\mathbb{C}[S_n]$ containing the set

$$\{\sum_w a_{1,w(1)}\cdots a_{n,w(n)}w \,|\, A \text{ TNN } \}.$$

We shall denote this cone by $\mathcal{C}_{TNN}$. (We omit the number n from this notation, although the cone obviously depends upon n.) Dual to $\mathcal{C}_{TNN}$ is the cone of TNN immanants, which we shall denote by $\check{\mathcal{C}}_{TNN}$,

$$\check{\mathcal{C}}_{TNN} = \{\mathrm{Imm}_f(x) \,|\, f(z) \geq 0 \text{ for all } z \in \mathcal{C}_{TNN}\}.$$

No simple description of the extremal rays of these cones is known. However, Stembridge showed [23, Thm. 2.1] that $\mathcal{C}_{TNN}$ is contained in the cone whose extremal rays are elements of $\mathbb{C}[S_n]$ of the form (2.2). Furthermore, Stembridge showed that this containment is proper for $n \geq 4$. We shall denote this third cone by $\mathcal{C}_{INT}$.

Define C_{KL} to be the cone whose extremal rays are the Kazhdan-Lusztig basis elements $\{C'_w(1) \,|\, w \in S_n\}$. It is not difficult to show that $\mathcal{C}_{INT}$ is contained in $\mathcal{C}_{KL}$

and that this containment is proper for $n \geq 4$. Thus we have the proper containment of the dual cones
$$\check{\mathcal{C}}_{KL} \subset \check{\mathcal{C}}_{INT} \subset \check{\mathcal{C}}_{TNN}.$$
For small n, many of the Kazhdan-Lusztig immanants seem to be extremal rays in $\check{\mathcal{C}}_{TNN}$. In the case $n = 4$ we have the following.

Proposition 4.2. *Let a totally nonnegative immanant* $\mathrm{Imm}_f(x)$ *in* V_4 *have coordinates* $\{d_w \,|\, w \in S_4\}$ *with respect to the basis of Kazhdan-Lusztig immanants,*
$$\mathrm{Imm}_f(x) = \sum_{w \in S_4} d_w \mathrm{Imm}_w(x).$$
Then d_w *can be negative only for* $w \in \{3412, 4231\}$.

Proof. Omitted. □

Question 4.2. To restate the previous proposition for arbitrary n, would it suffice to say that d_w can be negative only when the Schubert variety Γ_w in $\mathcal{F}_n$ is not smooth? (i.e. when w avoids the patterns $3412, 4231$?)

5. Acknowledgements

The authors are grateful to Sara Billey, Francesco Brenti, Fabrizio Caselli, Fokko Ducloux, Kenneth Fan, John Stembridge, Peter Trapa, and referees for helpful conversations, comments, and computer programs.

References

[1] H. Barcelo and A. Ram. Combinatorial representation theory. In *New Perspectives in Algebraic Combinatorics (Berkeley, CA 1996-97)* (G. C. Rota, ed.), vol. 38 of *Math. Sci. Res. Inst. Publ.*. Cambridge University Press, Cambridge, 1999 pp. 23–90.

[2] F. Bergeron, R. Biagioli, and M. Rosas. On a conjecture concerning Littlewood-Richardson coefficients, 2003. In progress.

[3] S. C. Billey and G. Warrington. Kazhdan-Lusztig polynomials for 321-hexagon avoiding permutations. *J. Algebraic Combin.*, **13** (2001) pp. 111–136.

[4] S. M. Fallat, M. I. Gekhtman, and C. R. Johnson. Multiplicative principal-minor inequalities for totally nonnegative matrices. *Adv. Appl. Math.*, **30** (2003) pp. 442–470.

[5] K. Fan and R. M. Green. Monomials and Temperley-Lieb algebras. *J. Algebra*, **190** (1997) pp. 498–517.

[6] S. Fomin, W. Fulton, C. Li, and Y. Poon. Eigenvalues, singular values, and Littlewood-Richardson coefficients, 2003. Preprint math.AG/03013078 on ArXiv.

[7] S. Fomin and A. Zelevinsky. Total positivity: Tests and parametrizations. *Math. Intelligencer*, **22** (2000) pp. 23–33.

[8] I. Gessel and G. Viennot. Determinants and plane partitions, 1989. Preprint.

[9] F. M. Goodman, P. de la Harpe, and V. Jones. *Coxeter Graphs and Towers of Algebras*. Springer–Verlag, New York, 1989.

[10] I. P. Goulden and D. M. Jackson. Immanants of combinatorial matrices. *J. Algebra*, **148** (1992) pp. 305–324.

[11] C. Greene. Proof of a conjecture on immanants of the Jacobi-Trudi matrix. *Linear Algebra Appl.*, **171** (1992) pp. 65–79.

[12] M. Haiman. Hecke algebra characters and immanant conjectures. *J. Amer. Math. Soc.*, **6** (1993) pp. 569–595.

[13] P. Heyfron. Immanant dominance orderings for hook partitions. *Linear and Multilinear Algebra*, **24** (1988) pp. 65–78.

[14] D. Kazhdan and G. Lusztig. Representations of Coxeter groups and Hecke algebras. *Inv. Math.*, **53** (1979) pp. 165–184.

[15] G. Lusztig. Total positivity in reductive groups. In *Lie Theory and Geometry: in Honor of Bertram Kostant*, vol. 123 of *Progress in Mathematics*. Birkhäuser, Boston, 1994 pp. 531–568.

[16] B. Rhoades and M. Skandera. Temperley-Lieb immanants, 2004. Submitted.

[17] B. Sagan. *The Symmetric Group*. Springer, New York, 2001.

[18] M. Skandera. Inequalities in products of minors of totally nonnegative matrices. *J. Alg. Combin.*, **20** (2004).

[19] T. A. Springer. Quelques aplications de la cohomologie d'intersection. In *Séminaire Bourbaki, Vol. 1981/1982*, vol. 92 of *Astérisque*. Soc. Math. France, Paris, 1982 pp. 249–273.

[20] R. Stanley. *Enumerative Combinatorics*, vol. 2. Cambridge University Press, Cambridge, 1999.

[21] R. Stanley. Positivity problems and conjectures. In *Mathematics: Frontiers and Perspectives* (V. Arnold, M. Atiyah, P. Lax, and B. Mazur, eds.). American Mathematical Society, Providence, RI, 2000 pp. 295–319.

[22] R. Stanley and J. R. Stembridge. On immanants of Jacobi-Trudi matrices and permutations with restricted positions. *J. Combin. Theory Ser. A*, **62** (1993) pp. 261–279.

[23] J. Stembridge. Immanants of totally positive matrices are nonnegative. *Bull. London Math. Soc.*, **23** (1991) pp. 422–428.

[24] J. Stembridge. Some conjectures for immanants. *Can. J. Math.*, **44** (1992) pp. 1079–1099.

[25] J. Stembridge. Multiplicity-free products of Schur functions. *Ann. Comb.*, **5** (2001) pp. 113–121.

[26] J. Stembridge, 2004. Personal communication.

[27] B. W. Westbury. The representation theory of the Temperley-Lieb algebras. *Math. Z.*, **219** (1995) pp. 539–565.

[28] A. Zelevinsky. From Littlewood-Richardson coefficients to cluster algebras in three lectures. In *Symmetric Functions 2001: Surveys of Developments and Perspectives* (S. Fomin, ed.), vol. 74 of *NATO Science Series II: Mathematics, Physics, and Chemistry*. Kluwer, Dordrecht, 2002 pp. 253–273.

ASYMPTOTICS OF CHARACTERS OF SYMMETRIC GROUPS AND FREE PROBABILITY

PIOTR ŚNIADY

ABSTRACT. In order to answer the question *"what is the asymptotic theory of representations of S_n"* we will present two concrete problems. In both cases the solution requires a good understanding of the product (convolution) of conjugacy classes in the symmetric group and we will present a combinatorial setup for explicit calculation of such products. The asymptotic behavior of each summand in our expansion will depend on topology (genus) of a two-dimensional surface associated to some partitions and for this reason our method carries a strong resemblance to the genus expansion from the random matrix theory. In particular, non-crossing partitions and free probability play a special role.

1. WHAT IS THE ASYMPTOTIC THEORY OF THE REPRESENTATIONS OF THE SYMMETRIC GROUPS S_n?

1.1. **(Generalized) Young diagrams.** Irreducible representations ρ^λ of the symmetric group S_n are in a one-to-one correspondence with *Young diagrams λ having n boxes*. An example of a Young diagram is presented on Figure 1.1. This figure also explains the notion of a *profile* of a Young diagram.

For a Young diagram with n boxes the area of the shaded region is equal to $2n$. After we shrink the geometric representation of this Young diagram by by factor $\frac{1}{\sqrt{n}}$ we obtain a *generalized Young diagram* (cf Figure 1.2) for which the area of the shaded region is equal to 2. In the following we will compare the shapes of the Young diagrams only after such a rescaling.

Please note that the usual definition of a Young diagram λ says that it is a weakly decreasing sequence $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ of positive integers and the *generalized Young diagram* considered above do not fit into this category. Instead, any generalized Young diagram is by definition identified with its profile (i.e. a function $f : \mathbb{R} \to \mathbb{R}_+$ with some additional constraints which an interested Reader can easily guess) [Ker93].

1.2. **Example of a problem: characters of a large Young diagram.** Let (λ_n) be a sequence of Young diagrams such that λ_n has n boxes and that the shapes of the Young diagrams λ_n converge (after rescaling) to some generalized Young diagram.

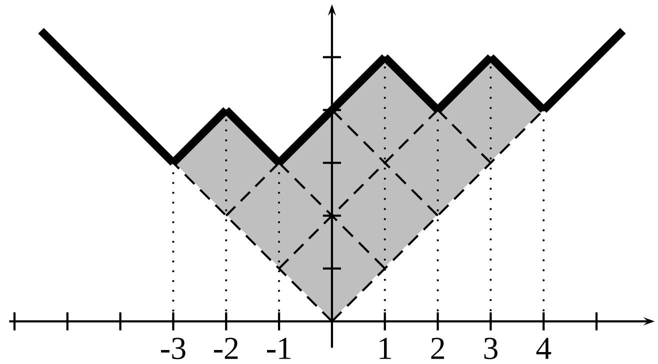


FIGURE 1.1. Graphical representation of a Young diagram $\lambda = (4, 3, 1)$ with 8 boxes. The thick ragged line is called the *profile* of a Young diagram.

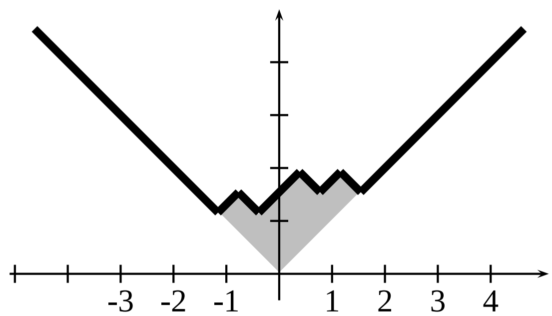


FIGURE 1.2. The Young diagram from Figure 1.1 after rescaling. The area of the shaded region is equal to 2.

What can we say about the characters of the corresponding irreducible representations

$$\chi^{\lambda_n}(\pi)$$

in the limit $n \to \infty$, where π is a fixed permutation?

1.3. **Example of a problem: what was the shape of the pile of stones?** For an integer $n \geq 1$ we consider a Young diagram ν with a shape of a $n \times n$ square. A *standard Young tableaux* is a filling of this Young diagram with numbers $1, \dots, n^2$ such that the numbers increase along the diagonals $\nearrow$, $\nwarrow$ from the bottom to the top, cf Figure 1.3 (left). We can think that a Young diagram is a *pile of stones* and the Young tableau is the order in which the stones are placed.

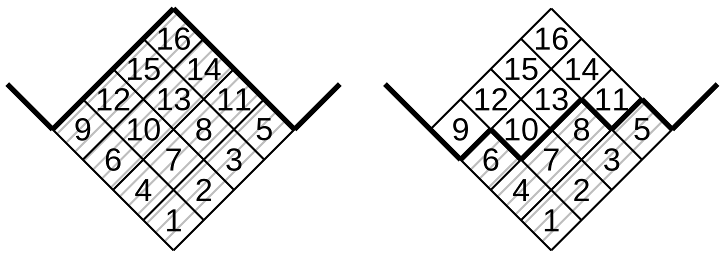


FIGURE 1.3. On the left: example of a standard Young tableaux of a square shape. On the right: the Young diagram resulting from this tableaux by removing the half of the boxes with the biggest numbers.

Let $0 < \alpha < 1$ be fixed; we remove from a randomly chosen standard Young tableaux all boxes with numbers bigger than αn^2, cf Figure 1.3 (right). What is the shape of the resulting Young diagram λ with αn^2 boxes, when $n \to \infty$? In other words: *What was the shape of this pile of stones in the past* [PR04]*?* This problem is equivalent to the study of *the restriction of representations*: the random Young diagram λ described above has the same distribution as a randomly chosen summand in the decomposition of $\rho = \rho^\nu \big\downarrow^{S_{n^2}}_{S_{\alpha n^2}}$ into irreducible components.

1.4. **Conclusions from the above examples.** In principle, for any question concerning representations of S_n there is a well-known answer given by some *combinatorial* algorithm. However, when $n \to \infty$ such *combinatorial* answers are too complicated to be useful. We need more analytic methods and—as we shall see—Kerov's transition measure is such an appropriate analytic tool.

The second problem (Section 1.3) indicates another phenomenon: in the asymptotic theory of representations some questions are of *statistical* flavor from the very beginning.

1.5. **Kerov's transition measure.** It was an idea of Kerov [Ker93] to associate to a Young diagram λ its *transition measure* μ_λ which is a certain probability measure on $\mathbb{R}$. The transition measure encodes the information about the shape of the Young diagram in a very compact and efficient way and it can be defined in (at least) four equivalent and interesting ways. Below we present just one of them (which is due to Biane [Bia98]).

We consider a matrix J, the entries of which belong to $\mathbb{C}(S_n)$, the symmetric group algebra:

$$J = \begin{bmatrix} 0 & (1,2) & \dots & (1,n) & 1 \\ (2,1) & 0 & \dots & (2,n) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (n,1) & (n,2) & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \in \mathcal{M}_{n+1}(\mathbb{C}) \otimes \mathbb{C}(S_n).$$

Except for the last row, the last column and the diagonal, the entry in the i-th row and the j-th column is equal to the transposition interchanging i and j. For a Young diagram λ we apply the irreducible representation ρ^λ to every entry of J and denote the outcome by $\rho^\lambda(J) \in \mathcal{M}_{n+1}(\mathbb{C}) \otimes \mathcal{M}_k(\mathbb{C}) = \mathcal{M}_{(n+1)k}(\mathbb{C})$.

The *transition measure* of λ (denoted μ^λ) is defined to be the spectral measure of the matrix $\rho^\lambda(J)$; in other words

$$\mu^\lambda = \frac{\delta_{\zeta_1} + \cdots + \delta_{\zeta_l}}{l}, \tag{1.1}$$

where $\zeta_1, \dots, \zeta_l$ are the eigenvalues of $\rho^\lambda(J)$ and δ_x denotes the Dirac measure at x. Since a compactly supported measure is determined by its moments this is equivalent with the requirement that

$$\int_{-\infty}^{\infty} x^k \, d\mu^\lambda(x) = \chi^\lambda(\operatorname{tr} J^k).$$

The element $\operatorname{tr} J^k \in \mathbb{C}(S_n)$ which appears in the right-hand side is called k*-th moment of the Jucys-Murphy element*.

Surprisingly, the definition of the transition measure can be naturally extended to generalized Young diagrams (which do not correspond to any irreducible representation). Furthermore, the transition measure behaves nicely with respect to the operation of rescaling.

2. How to deal with conjugacy classes?

2.1. **Where the difficulty is hidden?** The problem which we studied in Section 1.2 can be reformulated as follows: how to express a prescribed conjugacy class in $\mathbb{C}(S_n)$ (the calculation of the character on this class was our original goal) as a function of $(\operatorname{tr} J^k)_{k\geq 1}$ (the character $\chi^\lambda(\operatorname{tr} J^k)$ can be evaluated from (1.1))?

In Section 1.3 we studied the following problem: we know how to evaluate the characters of a reducible representation ρ (in our case: $\rho = \rho^\nu \big\downarrow^{S_{n^2}}_{S_{\alpha n^2}}$) on any conjugacy class and we ask statistical questions about the joint distribution of the random variables $\lambda \mapsto f_k(\lambda)$, where (f_k) is a family of some interesting functionals of the shape of a Young diagram and λ is a randomly chosen Young diagram contributing to ρ. Via Fourier transform we can view each f_k as a central element of $\mathbb{C}(S_n)$ and therefore this question can be reformulated as follows: how to express the products $f_{k_1} \cdots f_{k_l} \in \mathbb{C}(S_n)$ as a linear combination of conjugacy classes?

To summarize very briefly: *the problem is how to work efficiently with conjugacy classes and their products*.

2.2. **The main tool: partition–indexed conjugacy classes.** Let $p = (p_1, \dots, p_l)$ be a sequence with $p_1, \dots, p_l \in \{1, \dots, n+1\}$ and let π be a partition of the set $\{1, \dots, l\}$. We say that $p \sim \pi$ if for any $1 \leq i, j \leq l$ the equality $p_i = p_j$ holds if and only if i and j are connected by the partition π. We define

$$\Sigma_\pi = \frac{1}{n+1} \sum_{i \sim \pi} J_{p_1 p_2} J_{p_2 p_3} \cdots J_{p_{l-1} p_l} J_{p_l p_1} = \sum_{\substack{i \sim \pi \\ i(l) = n+1}} J_{p_1 p_2} J_{p_2 p_3} \cdots J_{p_{l-1} p_l} J_{p_l p_1} \in \mathbb{C}(S_n). \tag{2.1}$$

We can show [Śni03b] that indeed the second and the third expression in (2.1) are equal and that all non-zero summands which contribute to (2.1) are conjugate and for this reason we call the central element Σ_π *a partition–indexed conjugacy class*.

Our main idea in our recent series of papers [Śni03b, Śni03a, Śni05] is that *partition–indexed conjugacy classes are a very good tool for studying questions concerning symmetric groups*. We will outline some arguments for it in the following.

2.3. **The usual conjugacy classes.** The above definition of partition–indexed conjugacy classes might sound quite strange since usually such conjugacy classes are defined in the following way: for integers $k_1, \dots, k_m \geq 1$ the conjugacy class $\Sigma_{k_1, \dots, k_m} \in \mathbb{C}(S_n)$ is given by [KO94, Bia03]:

$$\Sigma_{k_1, \dots, k_m} = \sum_a (a_{1,1}, a_{1,2}, \dots, a_{1,k_1}) \cdots (a_{m,1}, a_{m,2}, \dots, a_{m,k_m}), \tag{2.2}$$

where the sum runs over all one–to–one functions

$$a : \big\{\{r, s\} : 1 \leq r \leq m, 1 \leq s \leq k_r\big\} \to \{1, \dots, n\}$$

and $(a_{1,1}, a_{1,2}, \dots, a_{1,k_1}) \cdots (a_{m,1}, a_{m,2}, \dots, a_{m,k_m})$ is a product of disjoint cycles.

In other words, we consider a Young diagram $(k_1, \dots, k_m)$ and all ways of filling it with the elements of the set $\{1, \dots, n\}$ in such a way that no element appears more than once. Each such a filling can be interpreted as a

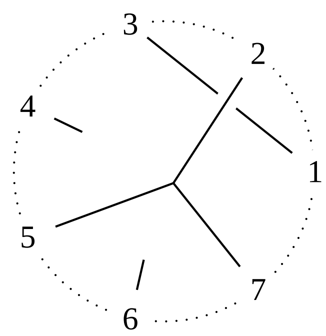


FIGURE 2.1. A graphical representation of a partition $\pi = \{\{1,3\},\{2,5,7\},\{4\},\{6\}\}$.

permutation when we treat rows of the Young tableau as disjoint cycles. It follows that each summand in (2.2) is a permutation with cycles of length $k_1, \dots, k_m$ and additionally with $n - (k_1 + \cdots + k_m)$ fix-points.

2.4. **A concrete form of the partition–indexed conjugacy classes.** Now a question arises how to relate the two kinds of conjugacy classes considered above and we will present a solution to this problem in this section.

It is convenient to represent partitions graphically, as it is shown on Figure 2.1. If we draw a partition with a very fat pen and take the boundary of this picture we obtain a fattened partition, as it is shown on the Figure 2.2 (left). We will use the convention that every vertex of the original partition is split into two *half-vertices*. The lines of this fat partition are equipped with the arrows which correspond to a counterclockwise orientation of the original blocks of the partition π and we added some extra lines at the boundary which connect elements 3 to $2'$, 2 to $1'$,.... In order to bring some order we ask a British policeman to organize a traffic circle so that every line must turn clockwise around the central disc, as it presented on Figure 2.2 (right). (French policemen are better in arranging such a *rond point* but unfortunately they prefer a counterclockwise traffic)

In this way we obtained a number of loops $L_1, \dots, L_r$: for a loop L we count the number V of vertices visited (attention! half-vertex is counted as $\frac{1}{2}$ of a vertex) and the winding number W of a loop around the British policemen in the center. In the example from Figure 2.2 (right) there are two loops: $1 \to 7' \to 2 \to 1' \to 3 \to 2' \to 5 \to 4' \to 4 \to 3' \to 1 \to \cdots$ ($V = 5$ vertices visited and $W = 3$ winds) and $7 \to 6' \to 6 \to 5' \to 7 \to \cdots$ ($V = 2$ vertices visited and $W = 1$ wind).

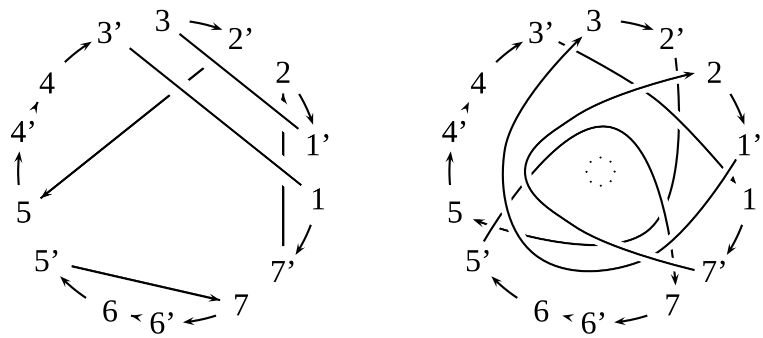


FIGURE 2.2. On the left: a fat partition corresponding to the partition π from Figure 2.1. On the right: a version of this figure in which all lines wind clockwise around the central disc.

We prove [Śni03b] that the relation between conjugacy classes Σ_π defined in 2.1 and the conjugacy classes $\Sigma_{k_1,\dots,k_m}$ defined in 2.2 is given explicitly by

$$\Sigma_\pi = \Sigma_{V_1 - W_1, \dots, V_r - W_r}.$$

2.5. **Products of conjugacy classes.** The first advantage of the partition–indexed conjugacy classes is that their product can be easily expressed combinatorially [Śni03b]. To be precise for any partition π_1 of a set $\{1,\dots,l_1\}$ and any partition π_2 of a set $\{l_1+1, l_1+2, \dots, l_1+l_2\}$

$$\Sigma_{l_1}\Sigma_{l_2} = \sum_\pi \Sigma_\pi, \tag{2.3}$$

where the sum runs over all partitions π of the set $\{1,\dots,l_1+l_2\}$ such that

- any elements $a, b \in \{1,\dots,l_1\}$ are connected by π if and only if they are connected by π_1,
- any elements $a, b \in \{l_1+1,\dots,l_1+l_2\}$ are connected by π if and only if they are connected by π_2,
- elements l_1 and l_1+l_2 are connected by π.

2.6. **Moments of Jucys-Murphy elements.** The second advantage of the partition–indexed conjugacy classes is that they can be used to express easily the moments $M_k = \operatorname{tr} J^k$ of Jucys-Murphy elements [Śni03b]:

$$M_k = \operatorname{tr} J^k = \sum_\pi \Sigma_\pi, \tag{2.4}$$

where the sum runs over all partitions π of the set $\{1,\dots,k\}$.

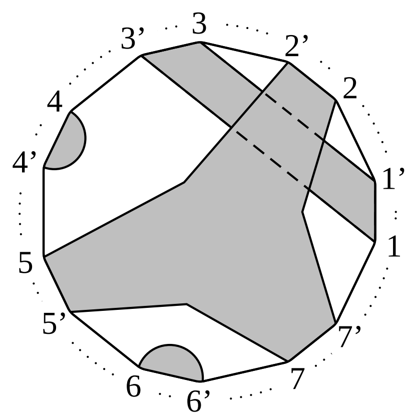


FIGURE 2.3. The first collection of discs for partition π from Figure 2.1.

2.7. **Genus expansion.** Equations (2.3) and (2.4) give us exact formulas, nevertheless in the asymptotic theory of representations of symmetric groups S_n we are rather interested in approximate formulas which hold asymptotically as $n \to \infty$. For this reason we define a filtration on the algebra of conjugacy classes [IK99, IO02, Śni04] by

$$\deg \Sigma_{k_1,\dots,k_l} = (k_1 + 1) + \cdots + (k_l + 1).$$

We consider a large sphere with a small circular hole. The boundary of this hole is the circle from Figure 2.1. Let us draw the blocks of the partition π with a fat pen; in this way each block becomes a disc glued to the boundary of the hole, cf Figure 2.3.

After gluing the first collection of discs, our sphere becomes a surface with a number of holes. The boundary of each hole is a circle and we shall glue this hole with another disc. Thus we obtained an orientable surface without a boundary. We call the genus of this surface the genus of the partition π and denote it by genus_π.

The following result was proved in our previous work [Śni03b]: for any partition π of an n–element set

$$\deg \Sigma_\pi = n - 2 \,\text{genus}_\pi \,. \tag{2.5}$$

In other words: the dominating contribution in the asymptotic problems will come from the partitions with the minimal possible genus. Similar formulas involving surfaces appear in the random matrix theory.

3. OK, SO WHERE IS FREE PROBABILITY THEORY?

Free probability of Voiculescu [VDN92] is a non-commutative probability theory in which the notion of independence was replaced by the notion of *freeness*. Free probability can be applied for example to describe asymptotic properties of some random matrices.

Speicher [Spe98] realized that the combinatorial structure behind freeness is the structure of *non-crossing partitions* [Kre72] and the corresponding *free cumulants*. We recall that a non-crossing partition is a partition with genus equal to 0. For a sequence (M_i), called the sequence of moments, the corresponding sequence (R_i) of free cumulants is given implicitly by equations

$$M_k = \sum_{\pi} R_{\pi}, \tag{3.1}$$

where the sum runs over all non-crossing partitions of the set $\{1, \dots, k\}$.

Please note the surprising similarity between equations (2.4) and (3.1). To the right-hand side of (3.1) contribute only non-crossing partitions and the first-order approximation of the right-hand side of (2.4) is given by non-crossing partitions. It follows that

$$\Sigma_k = R_{k+1} + \text{(lower degree terms)}. \tag{3.2}$$

To conclude: *free cumulants of the Jucys-Murphy element (or, equivalently, of the corresponding transition measure) are a very important tool for study of asymptotical theory of representations* and there are two reasons for it. Firstly, free cumulants are homogenous is a sense that they behave nicely with respect to rescaling of a generalized Young diagram and therefore it is easy to understand their asymptotic behavior. Secondly, the relation (3.2) between free cumulants and conjugacy classes is much simpler than the analogous relation between moments of the Jucys-Murphy element and conjugacy classes.

4. SOLUTIONS TO THE PROBLEMS

Methods presented in this article [Śni03b] can be used to attack the problems presented in Section 1 and we will present below the outcomes.

4.1. **Problem from Section 1.2 revisited.**

Finally, the problem from Section 1.2 can be formulated as follows: *what is the relation between the conjugacy classes* (Σ_k) *and the free cumulants* (R_k)*?* The first partial answer to this problem was given by Biane [Bia98] who proved (3.2) which can be regarded as a first-order approximation. In our recent article [Śni03a] we found explicitly the second-order asymptotics. Very recently Goulden and

Rattan [GR05] gave a complete solution to this problem (their solution uses very different methods).

4.2. **Problem from Section 1.3 revisited.** In our recent work [Śni05] we found a very large class of representations of S_n with a *approximate factorization of characters* and we proved that a shape of a randomly chosen Young diagram contributing to such representations concentrates around some limit shape (this part was already proved by Biane [Bia01]) and furthermore the fluctuations around this limit shape are Gaussian.

5. Acknowledgments

Research supported by State Committee for Scientific Research (Komitet Badań Naukowych) grant No. 2P03A00723; by EU Network "QP-APPLICATIONS", contract HPRN-CT-2002-00729; by KBN-DAAD project 36/2003/2004.

References

[Bia98] Philippe Biane. Representations of symmetric groups and free probability. *Adv. Math.*, 138(1):126–181, 1998.

[Bia01] Philippe Biane. Approximate factorization and concentration for characters of symmetric groups. *Internat. Math. Res. Notices*, (4):179–192, 2001.

[Bia03] Philippe Biane. Characters of symmetric groups and free cumulants. In Anatoly M. Vershik, editor, *Asymptotic Combinatorics with Applications to Mathematical Physics*, volume 1815 of *Lecture Notes in Mathematics*, pages 185–200. Springer, 2003.

[GR05] I.P. Goulden and A. Rattan. An explicit form for Kerov's character polynomials. Preprint, 2005.

[IK99] V. Ivanov and S. Kerov. The algebra of conjugacy classes in symmetric groups, and partial permutations. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 256(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 3):95–120, 265, 1999.

[IO02] Vladimir Ivanov and Grigori Olshanski. Kerov's central limit theorem for the Plancherel measure on Young diagrams. In S. Fomin, editor, *Symmetric Functions 2001: Surveys of Developments and Perspectives*, volume 74 of *NATO Science Series II. Mathematics, Physics and Chemistry*, pages 93–151. Kluwer, 2002.

[Ker93] S. V. Kerov. Transition probabilities of continual Young diagrams and the Markov moment problem. *Funktsional. Anal. i Prilozhen.*, 27(2):32–49, 96, 1993.

[KO94] Serguei Kerov and Grigori Olshanski. Polynomial functions on the set of Young diagrams. *C. R. Acad. Sci. Paris Sér. I Math.*, 319(2):121–126, 1994.

[Kre72] G. Kreweras. Sur les partitions non croisées d'un cycle. *Discrete Math.*, 1(4):333–350, 1972.

[PR04] Boris Pittel and Dan Romik. Limit shapes for random square Young tableaux and plane partitions. Preprint `arXiv:math.PR/0405190`, 2004.

[Śni03a] Piotr Śniady. Free probability and representations of large symmetric groups. Preprint `arXiv:math.CO/0304275`, 2003.

[Śni03b] Piotr Śniady. Symmetric groups and random matrices (the title of this preprint is tentative and might change). Preprint `arXiv:math.CO/0301299`, 2003.

[Śni04] Piotr Śniady. Asymptotics of characters of symmetric groups and genus expansion. Preprint `arXiv:math.CO/041164`, 2004.

[Śni05] Piotr Śniady. Gaussian fluctuations of characters of symmetric groups and of Young diagrams. Preprint `arXiv:math.CO/0501112`, 2005.

[Spe98] Roland Speicher. Combinatorial theory of the free product with amalgamation and operator-valued free probability theory. *Mem. Amer. Math. Soc.*, 132(627):x+88, 1998.

[VDN92] D. V. Voiculescu, K. J. Dykema, and A. Nica. *Free random variables*. American Mathematical Society, Providence, RI, 1992.

INSTITUTE OF MATHEMATICS, UNIVERSITY OF WROCLAW, PL. GRUNWALDZKI 2/4, 50-384 WROCLAW, POLAND

E-mail address: `Piotr.Sniady@math.uni.wroc.pl`

TABLEAUX ON PERIODIC SKEW DIAGRAMS AND IRREDUCIBLE REPRESENTATIONS OF DOUBLE AFFINE HECKE ALGEBRA OF TYPE A

TAKESHI SUZUKI AND MONICA VAZIRANI

Abstract. The irreducible representations of the symmetric group S_n are parameterized by combinatorial objects called *Young diagrams,* or shapes. A given irreducible representation has a basis indexed by *Young tableaux* of that shape. In fact, this basis consists of weight vectors (simultaneous eigenvectors) for a commutative subalgebra $\mathbb{F}[\mathfrak{X}]$ of the group algebra $\mathbb{F}S_n$.

The double affine Hecke algebra (DAHA) is a deformation of the group algebra of the *affine* symmetric group and it also contains a commutative subalgebra $\mathbb{F}[\mathfrak{X}]$.

Not every irreducible representation of the DAHA has a basis of weight vectors (and in fact it is quite difficult to parameterize all of its irreducible representations), but if we restrict our attention to those that do, these irreducible representations are parameterized by "affine shapes" and have a basis (of $\mathfrak{X}$-weight vectors) indexed by the "affine tableaux" of that shape. In this talk, we will construct these irreducible representations.

Introduction.

We introduce and study an affine analogue of skew Young diagrams and tableaux on them. The double affine Hecke algebra of type A acts on the space spanned by standard tableaux on each diagram. We show that the modules obtained this way are irreducible, and they exhaust all irreducible modules of a certain class over the double affine Hecke algebra. In particular, the classification of irreducible modules of this class, announced by Cherednik, is recovered.

As is well-known, Young diagrams consisting of n boxes parameterize isomorphism classes of finite dimensional irreducible representations of the symmetric group $\mathfrak{S}_n$, and moreover the structure of each irreducible representation is described in terms of tableaux on the corresponding Young diagram; namely, a basis of the representation is labeled by standard tableaux, on which the action of $\mathfrak{S}_n$ generators is explicitly described. This combinatorial description due to A. Young has played an essential role in the study of the representation theory of the symmetric group (or the affine Hecke algebra), and its generalization for the (degenerate) affine Hecke algebra $H_n(q)$ of GL_n has been given in [Ch1, Ra1, Ra2], where skew Young diagrams appear on combinatorial side.

The purpose of this paper is to introduce an "affine analogue" of skew Young diagrams and tableaux, which give a parameterization and a combinatorial description of a family of irreducible representations of the double affine Hecke algebra $\ddot{H}_n(q)$ of GL_n over a field $\mathbb{F}$, where $q \in \mathbb{F}$ is a parameter of the algebra.

The double affine Hecke algebra was introduced by I. Cherednik [Ch2, Ch3] and has since been used by him and by several authors to obtain important results about diagonal coinvariants, Macdonald polynomials, and certain Macdonald identities.

In this paper, we focus on the case where q is not a root of 1, and we consider representations of $\ddot{H}_n(q)$ that are $\mathfrak{X}$-semisimple; namely, we consider representations which have basis of simultaneous eigenvectors with respect to all elements in the commutative subalgebra $\mathbb{F}[\mathfrak{X}] = \mathbb{F}[x_1^{\pm 1}, x_2^{\pm 1}, \dots, x_n^{\pm 1}, \xi^{\pm 1}]$ of $\ddot{H}_n(q)$. (In [Ra1, Ra2], such representations for affine Hecke algebras are referred to as "calibrated.")

On combinatorial side, we introduce *periodic skew diagrams* as skew Young diagrams consisting of infinitely many boxes satisfying certain periodicity conditions. We define a tableau on a periodic skew diagram as a bijection from the diagram to $\mathbb{Z}$ which satisfies the condition reflecting the periodicity of the diagram.

Periodic skew diagrams are natural generalization of skew Young diagrams and have appeared in [Ch4] (or implicitly in [AST]), but the notion of tableaux on them seems new.

To connect the combinatorics with the representation theory of the double affine Hecke algebra $\ddot{H}_n(q)$, we construct, for each periodic skew diagram, an $\ddot{H}_n(q)$-module that has a basis of $\mathbb{F}[\mathfrak{X}]$-weight vectors labeled by standard tableaux on the diagram by giving the explicit action of the $\ddot{H}_n(q)$ generators.

Such modules are $\mathfrak{X}$-semisimple by definition. We show that they are irreducible, and that our construction gives a one-to-one correspondence between the set of periodic skew diagrams and the set of isomorphism classes of irreducible representations of the double affine Hecke algebra that are $\mathfrak{X}$-semisimple.

The classification results here recover those of Cherednik's in [Ch4] (see also [Ch5]), but in this paper we provide a detailed proof based on purely combinatorial arguments concerning standard tableaux on periodic skew diagrams.

Note that the corresponding results for the degenerate double affine Hecke algebra of GL_n easily follow from a parallel argument.

1. The affine root system and Weyl group

1.1. **The affine root system.** Let $n \in \mathbb{Z}_{\geq 2}$. Let $\tilde{\mathfrak{h}}$ be an $(n+2)$-dimensional vector space over $\mathbb{Q}$ with the basis $\{\epsilon_1^\vee, \epsilon_2^\vee, \dots, \epsilon_n^\vee, c, d\}$:

$$\tilde{\mathfrak{h}} = \left(\oplus_{i=1}^n \mathbb{Q}\epsilon_i^\vee\right) \oplus \mathbb{Q}c \oplus \mathbb{Q}d.$$

Introduce the non-degenerate symmetric bilinear form $(\,|\,)$ on $\tilde{\mathfrak{h}}$ by

$$(\epsilon_i^\vee|\epsilon_j^\vee) = \delta_{ij}, \quad (\epsilon_i^\vee|c) = (\epsilon_i^\vee|d) = 0, \quad (c|d) = 1, \quad (c|c) = (d|d) = 0.$$

Put $\mathfrak{h} = \oplus_{i=1}^n \mathbb{Q}\epsilon_i^\vee$ and $\dot{\mathfrak{h}} = \mathfrak{h} \oplus \mathbb{Q}c$. Let $\tilde{\mathfrak{h}}^* = (\oplus_{i=1}^n \mathbb{Q}\epsilon_i) \oplus \mathbb{Q}c^* \oplus \mathbb{Q}\delta$ be the dual space of $\tilde{\mathfrak{h}}$, where ϵ_i, c^* and δ are the dual vectors of $\epsilon_i^\vee$, c and d respectively. We identify the dual space $\dot{\mathfrak{h}}^*$ of $\dot{\mathfrak{h}}$ as a subspace of $\tilde{\mathfrak{h}}^*$ via the identification $\dot{\mathfrak{h}}^* = \tilde{\mathfrak{h}}^*/\mathbb{Q}\delta \cong \mathfrak{h}^* \oplus \mathbb{Q}c^*$.

The natural pairing is denoted by $\langle\,|\,\rangle : \tilde{\mathfrak{h}}^* \times \tilde{\mathfrak{h}} \to \mathbb{Q}$. There exists an isomorphism $\tilde{\mathfrak{h}}^* \to \tilde{\mathfrak{h}}$ such that $\epsilon_i \mapsto \epsilon_i^\vee$, $\delta \mapsto c$ and $c^* \mapsto d$. We denote by $\zeta^\vee \in \tilde{\mathfrak{h}}$ the image of $\zeta \in \tilde{\mathfrak{h}}^*$ under this isomorphism. Introduce the bilinear form $(\,|\,)$ on $\tilde{\mathfrak{h}}^*$ through this isomorphism. Note that

$$(\zeta \mid \eta) = \langle \zeta \mid \eta^\vee \rangle = (\zeta^\vee \mid \eta^\vee), \quad (\zeta, \eta \in \tilde{\mathfrak{h}}^*).$$

Put $\alpha_{ij} = \epsilon_i - \epsilon_j$ $(1 \leq i \neq j \leq n)$ and $\alpha_i = \alpha_{ii+1}$ $(1 \leq i \leq n-1)$. Then

$$R = \{\alpha_{ij} \mid i, j \in [1,n],\ i \neq j\},\ R^+ = \{\alpha_{ij} \mid i, j \in [1,n],\ i < j\},\ \Pi = \{\alpha_1, \alpha_2, \dots, \alpha_{n-1}\}$$

give the system of roots, positive roots and simple roots of type A_{n-1} respectively.

Put $\alpha_0 = -\alpha_{1n} + \delta$, and define the set $\dot{R}$ of (real) roots, $\dot{R}^+$ of positive roots and $\dot{\Pi}$ of simple roots of type $A^{(1)}_{n-1}$ by

$$\begin{aligned}
&\dot{R} = \{\alpha + k\delta \,|\, \alpha \in R,\, k \in \mathbb{Z}\}\,,\\
&\dot{R}^+ = \{\alpha + k\delta \mid \alpha \in R^+,\, k \in \mathbb{Z}_{\geq 0}\} \sqcup \{-\alpha + k\delta \mid \alpha \in R^+,\, k \in \mathbb{Z}_{\geq 1}\}\,,\\
&\dot{\Pi} = \{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}\,.
\end{aligned}$$

1.2. **Affine Weyl group.**

Definition 1.1. For $n \in \mathbb{Z}_{\geq 2}$, the *extended affine Weyl group* $\dot{W}_n$ of $\mathfrak{gl}_n$ is the group defined by the following generators and relations:

$$\begin{aligned}
&\textit{generators}\ : && s_0, s_1, \ldots, s_{n-1}, \pi^{\pm 1}.\\
&\textit{relations for } n \geq 3\ : && s_i^2 = 1\ (i \in [0, n-1]),\\
& && s_i s_j\ s_i = s_j s_i s_j\ (i - j \equiv \pm 1 \bmod n),\\
& && s_i s_j = s_j s_i\ (i - j \not\equiv \pm 1 \bmod n),\\
& && \pi s_i = s_{i+1}\pi,\ (i \in [0, n-2]), \quad \pi s_{n-1} = s_0 \pi,\\
& && \pi\pi^{-1} = \pi^{-1}\pi = 1.\\
&\textit{relations for } n = 2\ : && s_0^2 = s_1^2 = 1,\\
& && \pi s_0 = s_1 \pi, \quad \pi s_1 = s_0 \pi, \quad \pi\pi^{-1} = \pi^{-1}\pi = 1.
\end{aligned}$$

The subgroup W_n of $\dot{W}_n$ generated by the elements $s_1, s_2, \ldots, s_{n-1}$ is called the *Weyl group* of $\mathfrak{gl}_n$. The group W_n is isomorphic to the symmetric group of degree n.

In the following, we fix $n \in \mathbb{Z}_{\geq 2}$ and denote $\dot{W} = \dot{W}_n$ and $W = W_n$.

Put

$$P = \oplus_{i=1}^n \mathbb{Z}\epsilon_i.$$

Put $\tau_{\epsilon_1} = \pi s_{n-1} \cdots s_2 s_1$ and $\tau_{\epsilon_i} = \pi^{i-1}\tau_{\epsilon_1}\pi^{-i+1}$ $(i \in [2, n])$. Then there exists a group embedding $P \to \dot{W}$ such that $\epsilon_i \mapsto \tau_{\epsilon_i}$. By τ_η we denote the element in $\dot{W}$ corresponding to $\eta \in P$. It is well-known that the group $\dot{W}$ is isomorphic to the semidirect product $P \rtimes W$ with the relation $w\tau_\eta w^{-1} = \tau_{w(\eta)}$.

The group $\dot{W}$ acts on $\tilde{\mathfrak{h}}$ by

$$\begin{aligned}
&s_i(h) = h - \langle \alpha_i | h \rangle \alpha_i^\vee \quad \text{for } i \in [1, n-1],\ h \in \tilde{\mathfrak{h}},\\
&\tau_{\epsilon_i}(h) = h + \langle \delta | h \rangle \epsilon_i^\vee - \left(\langle \epsilon_i | h \rangle + \frac{1}{2}\langle \delta | h \rangle \right) c \quad \text{for } i \in [1, n],\ h \in \tilde{\mathfrak{h}}.
\end{aligned}$$

The dual action on $\tilde{\mathfrak{h}}^*$ is given by

$$\begin{aligned}
&s_i(\zeta) = \zeta - (\alpha_i | \zeta)\alpha_i \quad \text{for } i \in [1, n-1],\ \zeta \in \tilde{\mathfrak{h}}^*,\\
&\tau_{\epsilon_i}(\zeta) = \zeta + (\delta | \zeta)\epsilon_i - \left((\epsilon_i | \zeta) + \frac{1}{2}(\delta | \zeta) \right) \delta \quad \text{for } i \in [1, n],\ h \in \tilde{\mathfrak{h}}^*.
\end{aligned}$$

With respect to these actions, the inner products on $\tilde{\mathfrak{h}}$ and $\tilde{\mathfrak{h}}^*$ are $\dot{W}$-invariant. Note that the set $\dot{R}$ of roots is preserved by the dual action of $\dot{W}$ on $\tilde{\mathfrak{h}}^*$. For $\alpha \in \dot{R}$, there exists $i \in [0, n-1]$ and $w \in \dot{W}$ such that $w(\alpha_i) = \alpha$. We set $s_\alpha = w s_i w^{-1}$. Then s_α is independent of the choice of i and w, and we have

$$s_\alpha(h) = h - \langle \alpha \mid h \rangle \alpha^\vee$$

for $h \in \tilde{\mathfrak{h}}$. The element s_α is called the reflection corresponding to α. Note that $s_{\alpha_i} = s_i$. For $w \in \dot{W}$, set

$$R(w) = \dot{R}^+ \cap w^{-1}\dot{R}^-,$$

where $\dot{R}^- = \dot{R}\backslash\dot{R}^+$. The length $l(w)$ of $w \in \dot{W}$ is defined as the number $\sharp R(w)$ of elements in $R(w)$. For $w \in \dot{W}$, an expression $w = \pi^k s_{j_1} s_{j_2} \cdots s_{j_m}$ is called a reduced expression if $m = l(w)$. It can be seen that

$$R(w) = \{s_{j_m} \cdots s_{j_2}(\alpha_{j_1}), s_{j_m} \cdots s_{j_3}(\alpha_{j_2}), \ldots, \alpha_{j_m}\} \tag{1.1}$$

if $w = \pi^k s_{j_1} s_{j_2} \cdots s_{j_m}$ is a reduced expression.

Define the Bruhat order $\preceq$ in $\dot{W}$ by

$$x \preceq w \Leftrightarrow x \text{ is equal to a subexpression of a reduced expression of } w.$$

Let I be a subset of $[0, n-1]$. Put

$$\dot{\Pi}_I = \{\alpha_i \mid i \in I\} \subseteq \dot{\Pi}, \quad \dot{W}_I = \langle s_i \mid i \in I\rangle \subseteq \dot{W}, \quad \dot{R}_I^+ = \{\alpha \in \dot{R}^+ \mid s_\alpha \in \dot{W}_I\}.$$

The subgroup $\dot{W}_I$ is called the parabolic subgroup corresponding to $\dot{\Pi}_I$. Define

$$\dot{W}^I = \left\{ w \in \dot{W} \mid R(w) \cap \dot{R}_I^+ = \emptyset \right\}.$$

1.3. **Notation.** For any integer i, we introduce the following notation:

$$\epsilon_i = \epsilon_{\underline{i}} - k\delta \in \tilde{\mathfrak{h}}^*, \quad \epsilon_i^\vee = \epsilon_{\underline{i}}^\vee - kc \in \tilde{\mathfrak{h}}, \tag{1.2}$$

where $i = \underline{i} + kn$ with $\underline{i} \in [1, n]$ and $k \in \mathbb{Z}$.

Put $\alpha_{ij} = \epsilon_i - \epsilon_j$ and $\alpha_{ij}^\vee = \epsilon_i^\vee - \epsilon_j^\vee$ for any $i, j \in \mathbb{Z}$. Noting that $\epsilon_0 - \epsilon_1 = \delta + \epsilon_n - \epsilon_1 = \alpha_0$, we reset $\alpha_i = \epsilon_i - \epsilon_{i+1}$ and $\alpha_i^\vee = \epsilon_i^\vee - \epsilon_{i+1}^\vee$ for any $i \in \mathbb{Z}$.

Define the action of $\dot{W}$ on the set $\mathbb{Z}$ of integers by

$$\begin{aligned} &s_i(j) = j+1 \quad \text{for } j \equiv i \bmod n, && s_i(j) = j \quad \text{for } j \not\equiv i, i+1 \bmod n,\\ &s_i(j) = j-1 \quad \text{for } j \equiv i+1 \bmod n, && \pi(j) = j+1 \quad \text{for all } j. \end{aligned}$$

It is easy to see that the action of τ_{ϵ_i} ($i \in [1, n]$) is given by

$$\tau_{\epsilon_i}(j) = j + n \quad \text{for } j \equiv i \bmod n, \qquad \tau_{\epsilon_i}(j) = j \quad \text{for } j \not\equiv i \bmod n,$$

and that the following formula holds for any $w \in \dot{W}$:

$$w(j+n) = w(j) + n \quad \text{for all } j.$$

Lemma 1.2. *Let $w \in \dot{W}$.*

(i) $w(\epsilon_j) = \epsilon_{w(j)}$ *and* $w(\epsilon_j^\vee) = \epsilon_{w(j)}^\vee$ *for any* $j \in \mathbb{Z}$.

(ii) $w(\alpha_{ij}) = \alpha_{w(i)w(j)}$ *and* $w(\alpha_{ij}^\vee) = \alpha_{w(i)w(j)}^\vee$ *for any* $i, j \in \mathbb{Z}$.

2. Periodic skew diagrams and tableaux on them

Throughout this paper, we let $\mathbb{F}$ denote a field whose characteristic is not equal to 2.

2.1. **Periodic skew diagrams.** For $m \in \mathbb{Z}_{\geq 1}$ and $\ell \in \mathbb{Z}_{\geq 0}$, put

$$\widehat{\mathcal{P}}^{+}_{m,\ell} = \{\mu \in \mathbb{Z}^m \mid \mu_1 \geq \mu_2 \geq \cdots \geq \mu_m \text{ and } \ell \geq \mu_1 - \mu_m\}, \tag{2.1}$$

where μ_i denotes the i-th component of μ, i.e., $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$. Fix $n \in \mathbb{Z}_{\geq 2}$ and introduce the following subsets of $\mathbb{Z}^m \times \mathbb{Z}^m$:

$$\widehat{\mathcal{J}}^{n}_{m,\ell} = \left\{ (\lambda, \mu) \in \widehat{\mathcal{P}}^{+}_{m,\ell} \times \widehat{\mathcal{P}}^{+}_{m,\ell} \,\middle|\, \lambda_i \geq \mu_i \ (i \in [1,m]),\ \sum_{i=1}^{m} (\lambda_i - \mu_i) = n \right\},$$
$$\widehat{\mathcal{J}}^{*n}_{m,\ell} = \left\{ (\lambda, \mu) \in \widehat{\mathcal{P}}^{+}_{m,\ell} \times \widehat{\mathcal{P}}^{+}_{m,\ell} \,\middle|\, \lambda_i > \mu_i \ (i \in [1,m]),\ \sum_{i=1}^{m} (\lambda_i - \mu_i) = n \right\}.$$

For $(\lambda, \mu) \in \widehat{\mathcal{J}}^{n}_{m,\ell}$, define the subsets λ/μ and $\widehat{\lambda/\mu}_{(m,-\ell)}$ of $\mathbb{Z}^2$ by

$$\lambda/\mu = \left\{ (a,b) \in \mathbb{Z}^2 \mid a \in [1,m],\ b \in [\mu_a + 1, \lambda_a] \right\},$$
$$\widehat{\lambda/\mu}_{(m,-\ell)} = \left\{ (a + km, b - k\ell) \in \mathbb{Z}^2 \mid (a,b) \in \lambda/\mu,\ k \in \mathbb{Z} \right\}.$$

Let $\lambda/\mu[k] = \lambda/\mu + k(m, -\ell)$. Obviously we have

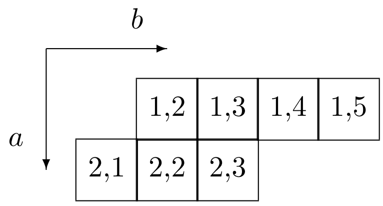
$$\widehat{\lambda/\mu}_{(m,-\ell)} = \bigsqcup_{k \in \mathbb{Z}} \lambda/\mu[k] = \bigsqcup_{k \in \mathbb{Z}} \left(\lambda/\mu + k(m, -\ell)\right).$$

The set λ/μ is the skew diagram (or skew Young diagram) associated with (λ, μ).
We call the set $\widehat{\lambda/\mu}_{(m,-\ell)}$ the *periodic skew diagram* associated with (λ, μ).
We will denote $\widehat{\lambda/\mu}_{(m,-\ell)}$ just by $\widehat{\lambda/\mu}$ when m and ℓ are fixed.

Example 2.1. Let $n = 7,\ m = 2$ and $\ell = 3$. Put $\lambda = (5,3),\ \mu = (1,0)$. Then $(\lambda, \mu) \in \widehat{\mathcal{J}}^{*n}_{m,\ell}$ and we have

$$\lambda/\mu = \{(1,2), (1,3), (1,4), (1,5), (2,1), (2,2), (2,3)\}.$$

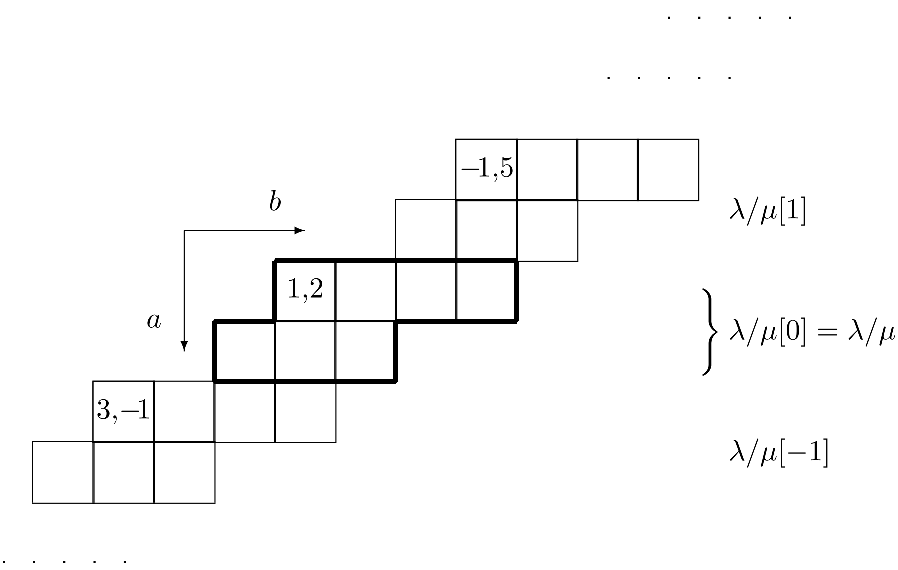
The set λ/μ is expressed by the following picture (usually, the coordinate in the boxes are omitted):



The periodic skew diagram

$$\widehat{\lambda/\mu}_{(2,-3)} = \bigsqcup_{k \in \mathbb{Z}} \lambda/\mu[k] = \bigsqcup_{k \in \mathbb{Z}} \left(\lambda/\mu + k(2, -3)\right)$$

is expressed by the following picture:



2.2. **Tableaux on periodic skew diagram.** Fix $n \in \mathbb{Z}_{\geq 2}$. Recall that a bijection from a skew Young diagram λ/μ of degree n to the set $[1, n]$ is called a tableau on λ/μ.

Definition 2.2. Given $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$, $\gamma = (m, -\ell)$, a bijection $T : \widehat{\lambda/\mu} \to \mathbb{Z}$ is said to be a γ-*tableau* on $\widehat{\lambda/\mu}$ if T satisfies

$$T(u + \gamma) = T(u) + n \quad \text{for all } u \in \widehat{\lambda/\mu}. \tag{2.2}$$

Let

$$\mathrm{Tab}(\widehat{\lambda/\mu}) = \mathrm{Tab}_{(m,-\ell)}(\widehat{\lambda/\mu})$$

denote the set of all γ-tableaux on $\widehat{\lambda/\mu}$.

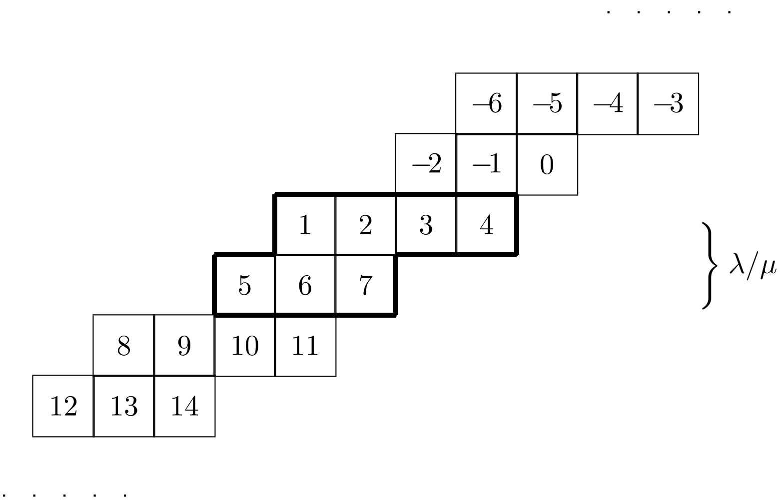
Remark 2.3. A tableau on $\widehat{\lambda/\mu}$ is determined uniquely from the values on a fundamental domain of $\widehat{\lambda/\mu}$ with respect to the action of $\mathbb{Z}\gamma$. It also holds that any bijection from a fundamental domain of $\mathbb{Z}\gamma$ to the set $[1, n]$ uniquely extends to a tableau on $\widehat{\lambda/\mu}$.

There exists a unique tableau $T_0^{\widehat{\lambda/\mu}} = T_0$ on $\widehat{\lambda/\mu}$ such that

$$T_0(i, \mu_i + j) = \sum_{k=1}^{i-1} (\lambda_k - \mu_k) + j \quad \text{for } i \in [1, m],\ j \in [1, \lambda_i - \mu_i]. \tag{2.3}$$

We call T_0 the *row reading tableau* on $\widehat{\lambda/\mu}$.

Example 2.4. Let $n = 7,\ m = 2,\ \ell = 3$ and $\lambda = (5, 3),\ \mu = (1, 0)$. The tableau T_0 on $\widehat{\lambda/\mu}$ given above is expressed as follows:



Proposition 2.5. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$. The group $\dot{W}$ acts on the set* $\mathrm{Tab}(\widehat{\lambda/\mu})$ *by*

$$(wT)(u) = w(T(u)) \tag{2.4}$$

for $w \in \dot{W}$, $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$ and $u \in \widehat{\lambda/\mu}$.

For each $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$, define the map $\psi_T : \dot{W} \to \mathrm{Tab}(\widehat{\lambda/\mu})$ by $\psi_T(w) = wT$ $(w \in \dot{W})$.

Proposition 2.6. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$. For any $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$, the correspondence ψ_T is a bijection.*

Lemma 2.7. *$T^{-1}(w^{-1}(i)) = (wT)^{-1}(i)$ for any $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$, $w \in \dot{W}$ and $i \in \mathbb{Z}$.*

2.3. **Content and weight.** Let C denote the map from $\mathbb{Z}^2$ to $\mathbb{Z}$ given by $C(a, b) = b - a$ for $(a, b) \in \mathbb{Z}^2$.

For a tableau $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$, define the map $C_T^{\widehat{\lambda/\mu}} : \mathbb{Z} \to \mathbb{Z}$ by

$$C_T^{\widehat{\lambda/\mu}}(i) = C(T^{-1}(i)) \quad (i \in \mathbb{Z}),$$

and call $C_T^{\widehat{\lambda/\mu}}$ the *content* of T. We simply denote $C_T^{\widehat{\lambda/\mu}}$ by C_T when (λ, μ) is fixed.

Lemma 2.8. *Let $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$. Then*
(i) *$C_T(i + n) = C_T(i) - (\ell + m)$ for all $i \in \mathbb{Z}$.*
(ii) *$C_{wT}(i) = C_T(w^{-1}(i))$ for all $w \in \dot{W}$ and $i \in \mathbb{Z}$.*

For $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$, we define $\zeta_T \in \dot{\mathfrak{h}}^*$ by

$$\zeta_T = \sum_{i=1}^{n} C_T(i)\epsilon_i + (\ell + m)c^*.$$

Then ζ_T belongs to the lattice $\dot{P} \stackrel{def}{=} P \oplus \mathbb{Z}c^* = (\oplus_{i=1}^n \mathbb{Z}\epsilon_i) \oplus \mathbb{Z}c^*$. Note that the action of $\dot{W}$ on $\dot{\mathfrak{h}}^*$ preserves $\dot{P}$. Lemma 2.8 immediately implies the following:

Lemma 2.9. *Let $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$. Then*
(i) $\langle \zeta_T \mid \epsilon_i^\vee \rangle = C_T(i)$ *for all* $i \in \mathbb{Z}$.
(ii) $w(\zeta_T) = \zeta_{wT}$ *for all* $w \in \dot{W}$.

2.4. **The affine Weyl group and row increasing tableaux.** Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$.

Definition 2.10. A tableau $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$ is said to be *row increasing* (resp. *column increasing*) if

$$(a,b),(a,b+1) \in \widehat{\lambda/\mu} \;\Rightarrow\; T(a,b) < T(a,b+1).$$
$$(\text{resp. } (a,b),(a+1,b) \in \widehat{\lambda/\mu} \;\Rightarrow\; T(a,b) < T(a+1,b).)$$

A tableau $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$ which is row increasing and column increasing is called *a standard tableau* (or *a row-column increasing tableau*).

Denote by $\mathrm{Tab}^{\mathrm{R}}(\widehat{\lambda/\mu})$ (resp. $\mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$) the set of all row increasing (resp. standard) tableaux on $\widehat{\lambda/\mu}$.

For $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$, put

$$I_{\lambda,\mu} = [1, n-1] \setminus \{n_1, n_2, \ldots, n_{m-1}\},$$

where $n_i = \sum_{j=1}^{i} (\lambda_j - \mu_j)$ for $i \in [1, m-1]$.

We write $\dot{R}^+_{\lambda-\mu} = \dot{R}^+_{I_{\lambda,\mu}}$, $\dot{W}_{\lambda-\mu} = \dot{W}_{I_{\lambda,\mu}}$ and $\dot{W}^{\lambda-\mu} = \dot{W}^{I_{\lambda,\mu}}$.

Note that $\dot{R}^+_{\lambda-\mu} \subseteq R^+$ and $\dot{W}_{\lambda-\mu} = W_{\lambda_1-\mu_1} \times W_{\lambda_2-\mu_2} \times \cdots \times W_{\lambda_m-\mu_m} \subseteq W$.

Recall that the correspondence $\psi_T : \dot{W} \to \mathrm{Tab}(\widehat{\lambda/\mu})$ given by $w \mapsto wT$ is bijective (Proposition 2.6) for any $T \in \mathrm{Tab}(\widehat{\lambda/\mu})$.

Proposition 2.11. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$. Then*

$$\psi_{T_0}^{-1}(\mathrm{Tab}^{\mathrm{R}}(\widehat{\lambda/\mu})) = \dot{W}^{\lambda-\mu},$$

or equivalently, $\mathrm{Tab}^{\mathrm{R}}(\widehat{\lambda/\mu}) = \dot{W}^{\lambda-\mu}T_0 = \{wT_0 \mid w \in \dot{W}^{\lambda-\mu}\}$.

2.5. **The set of standard tableaux.** The next lemma follows easily:

Lemma 2.12. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$ and $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$. If $(a,b) \in \widehat{\lambda/\mu}$ and $(a+1,b+1) \in \widehat{\lambda/\mu}$, then $T(a+1,b+1) - T(a,b) > 1$.*

Proposition 2.13. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$ and $T, S \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$. If $C_T = C_S$ then $T = S$.*

For $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$, put

$$\dot{Z}_T^{\widehat{\lambda/\mu}} = \left\{ w \in \dot{W} \,\middle|\, \langle \zeta_T \mid \alpha^\vee \rangle \notin \{-1, 1\} \text{ for all } \alpha \in R(w) \right\}. \tag{2.5}$$

Theorem 2.14. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$ and $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$. Then*

$$\psi_T^{-1}(\mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})) = \dot{Z}_T^{\widehat{\lambda/\mu}},$$

or equivalently, $\mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu}) = \dot{Z}_T^{\widehat{\lambda/\mu}}T$.

For $m \in \mathbb{Z}_{\geq 1}$, define an automorphism ω_m of $\mathbb{Z}^m$ by

$$\omega_m \cdot \lambda = (\lambda_m + \ell + 1, \lambda_1 + 1, \lambda_2 + 1, \ldots, \lambda_{m-1} + 1), \tag{2.6}$$

for $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{Z}^m$. Let $\langle \omega_m \rangle$ denote the free group generated by ω_m, and let $\langle \omega_m \rangle$ act on $\mathbb{Z}^m \times \mathbb{Z}^m$ by $\omega_m \cdot (\lambda, \mu) = (\omega_m \cdot \lambda, \omega_m \cdot \mu)$ for $(\lambda, \mu) \in \mathbb{Z}^m \times \mathbb{Z}^m$. Note that $\langle \omega_m \rangle$ preserves the subsets $\widehat{\mathcal{J}}^n_{m,\ell}$ and $\widehat{\mathcal{J}}^{*n}_{m,\ell}$ of $\mathbb{Z}^m \times \mathbb{Z}^m$.

Proposition 2.15. *Let $m, m' \in [1, n]$ and $\ell, \ell' \in \mathbb{Z}_{\geq 0}$. Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^{*n}_{m,\ell}$ and $(\eta, \nu) \in \widehat{\mathcal{J}}^{*n}_{m',\ell'}$. The following are equivalent*:

(a) $C_T^{\widehat{\lambda/\mu}} = C_S^{\widehat{\eta/\nu}}$ *for some* $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$ *and* $S \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\eta/\nu})$,

(b) $m = m'$, $\ell = \ell'$ *and* $\widehat{\lambda/\mu} = \widehat{\eta/\nu} + (r, r)$ *for some* $r \in \mathbb{Z}$.

(c) $m = m'$, $\ell = \ell'$ *and* $(\eta, \nu) = \omega_m^r \cdot (\lambda, \mu)$ *for some* $r \in \mathbb{Z}$.

3. Representations of the double affine Hecke algebra

Let $\mathbb{F}$ denote a field whose characteristic is not equal to 2.

3.1. Double affine Hecke algebra of type A. Let $q \in \mathbb{F}$.

The double affine Hecke algebra was introduced by Cherednik [Ch2, Ch3].

Definition 3.1. Let $n \in \mathbb{Z}_{\geq 2}$.

(i) The *double affine Hecke algebra* $\ddot{H}_n(q)$ of GL_n is the unital associative algebra over $\mathbb{F}$ defined by the following generators and relations:

$$\begin{aligned}
&\textit{generators}\ : && t_0, t_1, \ldots, t_{n-1}, \pi^{\pm 1}, x_1^{\pm 1}, x_2^{\pm 1}, \ldots, x_n^{\pm 1}, \xi^{\pm 1}.\\
&\textit{relations for } n \geq 3 : && (t_i - q)(t_i + 1) = 0 \ (i \in [0, n-1]),\\
& && t_i t_j t_i = t_j t_i t_j \ (j \equiv i \pm 1 \bmod n), \quad t_i t_j = t_j t_i \ (j \not\equiv i \pm 1 \bmod n),\\
& && \pi \pi^{-1} = \pi^{-1} \pi = 1,\\
& && \pi t_i \pi^{-1} = t_{i+1} \ (i \in [0, n-2]), \ \pi t_{n-1} \pi^{-1} = t_0,\\
& && x_i x_i^{-1} = x_i^{-1} x_i = 1 \ (i \in [1, n]), \quad x_i x_j = x_j x_i \ (i, j \in [1, n]),\\
& && t_i x_i t_i = q x_{i+1} \ (i \in [1, n-1]), \ t_0 x_n t_0 = \xi^{-1} q x_1\\
& && t_i x_j = x_j t_i \ (j \not\equiv i, i+1 \bmod n),\\
& && \pi x_i \pi^{-1} = x_{i+1} \ (i \in [1, n-1]), \ \pi x_n \pi^{-1} = \xi^{-1} x_1,\\
& && \xi \xi^{-1} = \xi^{-1} \xi = 1, \quad \xi^{\pm 1} h = h \xi^{\pm 1} \ (h \in \ddot{H}_n(q)).\\
&\textit{relations for } n = 2 : && (t_i - q)(t_i + 1) = 0 \ (i \in [0, 1]),\\
& && \pi \pi^{-1} = \pi^{-1} \pi = 1, \quad \pi t_0 \pi^{-1} = t_1, \ \pi t_1 \pi^{-1} = t_0,\\
& && x_i x_i^{-1} = x_i^{-1} x_i = 1 \ (i \in [1, 2]), \quad x_1 x_2 = x_2 x_1,\\
& && t_1 x_1 t_1 = q x_2, \quad t_0 x_2 t_0 = \xi^{-1} q x_1\\
& && \pi x_1 \pi^{-1} = x_2, \ \pi x_2 \pi^{-1} = \xi^{-1} x_1,\\
& && \xi \xi^{-1} = \xi^{-1} \xi = 1, \quad \xi^{\pm 1} h = h \xi^{\pm 1} \ (h \in \ddot{H}_2(q)).
\end{aligned}$$

(ii) Define the *affine Hecke algebra* $\dot{H}_n(q)$ of GL_n as the subalgebra of $\ddot{H}_n(q)$ generated by $\{t_0, t_1, \ldots, t_{n-1}, \pi^{\pm 1}\}$.

Remark 3.2. It is known that the subalgebra of $\ddot{H}_n(q)$ generated by

$$\{t_1, t_2, \ldots, t_{n-1}, x_1^{\pm 1}, x_2^{\pm 1}, \ldots, x_n^{\pm 1}\}$$

is also isomorphic to $\dot{H}_n(q)$.

For $\nu = \sum_{i=1}^{n} \nu_i \epsilon_i + \nu_c c^* \in \dot{P}$, put

$$x^\nu = x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n} \xi^{\nu_c}.$$

Let $\mathfrak{X}$ denote the commutative group $\{x^\nu \mid \nu \in \dot{P}\} \subseteq \ddot{H}_n(q)$. The group algebra $\mathbb{F}[\mathfrak{X}] = \mathbb{F}[x_1^{\pm 1}, x_2^{\pm 1}, \ldots, x_n^{\pm 1}, \xi^{\pm 1}]$ is a commutative subalgebra of $\ddot{H}_n(q)$.

For $w \in \dot{W}$ with a reduced expression $w = \pi^r s_{i_1} s_{i_2} \cdots s_{i_k}$, put

$$t_w = \pi^r t_{i_1} t_{i_2} \cdots t_{i_k}.$$

Then t_w does not depend on the choice of the reduced expression, and $\{t_w\}_{w \in \dot{W}}$ forms a basis of the affine Hecke algebra $\dot{H}_n(q) \subset \ddot{H}_n(q)$.

It is easy to see that $\{t_w x^\nu\}_{w \in \dot{W}, \nu \in \dot{P}}$ and $\{x^\nu t_w\}_{w \in \dot{W}, \nu \in \dot{P}}$ respectively form bases of $\ddot{H}_n(q)$.

Let $\mathfrak{X}^*$ denote the set of characters of $\mathfrak{X}$:

$$\mathfrak{X}^* = \mathrm{Hom}_{\,group}(\mathfrak{X}, \mathrm{GL}_1(\mathbb{F})).$$

Consider the correspondence $\dot{P} \to \mathfrak{X}^*$ which maps $\zeta \in \dot{P}$ to the character $q^\zeta \in \mathfrak{X}^*$ defined by

$$q^\zeta(x_i) = q^{\langle \zeta | \epsilon_i^\vee \rangle} \ (i \in [1, n]), \quad q^\zeta(\xi) = q^{\langle \zeta | c \rangle},$$

or equivalently, defined by $q^\zeta(x^\nu) = q^{\langle \zeta | \nu^\vee \rangle}$ $(\nu \in \dot{P})$. Through this correspondence, $\dot{P}$ is identified with the subset

$$\{\chi \in \mathfrak{X}^* \mid \chi(x^\nu) \in q^{\mathbb{Z}} \ (\forall \nu \in \dot{P})\}$$

of $\mathfrak{X}^*$, where $q^{\mathbb{Z}} = \{q^r \mid r \in \mathbb{Z}\}$.

For an $\ddot{H}_n(q)$-module M and $\zeta \in \dot{P}$, define the weight space M_ζ and the generalized weight space M_ζ^{gen} of weight ζ with respect to the action of $\mathbb{F}[\mathfrak{X}]$ by

$$M_\zeta = \left\{ v \in M \,\middle|\, (x^\nu - q^{\langle \zeta | \nu^\vee \rangle}) v = 0 \ \text{ for any } \nu \in \dot{P} \right\},$$
$$M_\zeta^{\mathrm{gen}} = \bigcup_{k \geq 1} \left\{ v \in M \,\middle|\, (x^\nu - q^{\langle \zeta | \nu^\vee \rangle})^k v = 0 \text{ for any } \nu \in \dot{P} \right\}.$$

For an $\ddot{H}_n(q)$-module M, an element $\zeta \in \dot{P}$ is called a weight of M if $M_\zeta \neq 0$, and an element $v \in M_\zeta$ (resp. M_ζ^{gen}) is called a weight vector (resp. generalized weight vector) of weight ζ.

For $\zeta \in \dot{P}$, put

$$\dot{\mathcal{Z}}_\zeta = \{w \in \dot{W} \mid \langle \zeta \mid \alpha^\vee \rangle \notin \{-1, 1\} \text{ for all } \alpha \in R(w)\}. \tag{3.1}$$

Note that $\dot{\mathcal{Z}}_{\zeta_T} = \dot{Z}_T^{\widehat{\lambda/\mu}}$ for $(\lambda, \mu) \in \widehat{\mathcal{J}}_{m,\ell}^n$ and $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$.

3.2. **$\mathfrak{X}$-semisimple modules.** Fix $n \in \mathbb{Z}_{\geq 2}$. Let $q \in \mathbb{F}$ and suppose that q is not a root of 1.

Fix $\kappa \in \mathbb{Z}$ and put $P_\kappa = P + \kappa c^* = \{\zeta \in \dot{P} \mid \langle \zeta \mid c \rangle = \kappa\}$.

Definition 3.3. Define $\mathcal{O}_\kappa^{ss}(\ddot{H}_n(q))$ as the set consisting of those $\ddot{H}_n(q)$-modules M which are finitely generated and admit a decomposition

$$M = \bigoplus_{\zeta \in P_\kappa} M_\zeta$$

with $\dim M_\zeta < \infty$ for all $\zeta \in P_\kappa$.

A module in $\mathcal{O}_\kappa^{ss}(\ddot{H}_n(q))$ is also called $\mathfrak{X}$-semisimple. We remark that the structure of all irreducible $\mathfrak{X}$-semisimple modules, without requiring the eigenvalues of the x_k to live in $\{q^i \mid i \in \mathbb{Z}\}$, is easily obtained once we understand the modules in $\mathcal{O}_\kappa^{ss}(\ddot{H}_n(q))$.

3.3. **Representations associated with periodic skew diagrams.** In the rest of this paper, we always assume that q is not a root of 1.

Let $n \in \mathbb{Z}_{\geq 2}$, $m \in \mathbb{Z}_{\geq 1}$ and $\ell \in \mathbb{Z}_{\geq 0}$.

For $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$, set

$$\ddot{V}(\lambda, \mu) = \bigoplus_{T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})} \mathbb{F} v_T. \tag{3.2}$$

Define linear operators $\tilde{x}_i$ $(i \in [1, n])$, $\tilde{\pi}$ and $\tilde{t}_i$ $(i \in [0, n-1])$ on $\ddot{V}(\lambda, \mu)$ by

$$\tilde{x}_i v_T = q^{C_T(i)} v_T, \tag{3.3}$$

$$\tilde{\pi} v_T = v_{\pi T}, \tag{3.4}$$

$$\tilde{t}_i v_T = \begin{cases} \frac{1-q^{1+\tau_i}}{1-q^{\tau_i}} v_{s_i T} - \frac{1-q}{1-q^{\tau_i}} v_T & \text{if } s_i T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu}), \\ -\frac{1-q}{1-q^{\tau_i}} v_T & \text{if } s_i T \notin \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu}), \end{cases} \tag{3.5}$$

where

$$\tau_i = C_T(i) - C_T(i+1) = \langle \zeta_T \mid \alpha_i^\vee \rangle \quad (i \in [0, n-1]).$$

The following lemma is easy and ensures that the operator $\tilde{t}_i$ is well-defined:

Lemma 3.4. *$C_T(i) - C_T(i+1) \neq 0$ for any $i \in [0, n-1]$ and $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$.*

Theorem 3.5. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$. There exists an algebra homomorphism $\theta_{\lambda,\mu} : \ddot{H}_n(q) \to \mathrm{End}_{\mathbb{F}}(\ddot{V}(\lambda, \mu))$ such that*

$$\theta_{\lambda,\mu}(t_i) = \tilde{t}_i \ (i \in [0, n-1]),\ \theta_{\lambda,\mu}(\pi) = \tilde{\pi},$$
$$\theta_{\lambda,\mu}(x_i) = \tilde{x}_i \ (i \in [1, n]), \quad \theta_{\lambda,\mu}(\xi) = q^{\ell+m}.$$

Theorem 3.6. *Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^n_{m,\ell}$.*

(i) $\ddot{V}(\lambda, \mu) = \bigoplus_{T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})} \ddot{V}(\lambda, \mu)_{\zeta_T}$, *and* $\ddot{V}(\lambda, \mu)_{\zeta_T} = \mathbb{F} v_T$ *for all* $T \in \mathrm{Tab}^{\mathrm{RC}}(\widehat{\lambda/\mu})$.

(ii) *The $\ddot{H}_n(q)$-module $\ddot{V}(\lambda, \mu)$ is irreducible.*

3.4. **Classification of $\mathfrak{X}$-semisimple modules.** Fix $n \in \mathbb{Z}_{\geq 2}$ and $\kappa \in \mathbb{Z}_{\geq 1}$. Let $q \in \mathbb{F}$ and suppose that q is not a root of 1.

Theorem 3.7. *Let $n \in \mathbb{Z}_{\geq 2}$ and $\kappa \in \mathbb{Z}_{\geq 1}$. Let L be an irreducible $\ddot{H}_n(q)$-module which belongs to $\mathcal{O}^{ss}_{\kappa}(\ddot{H}_n(q))$. Then there exist $m \in [1, \kappa]$ and $(\lambda, \mu) \in \widehat{\mathcal{J}}^{*n}_{m,\kappa-m}$ such that $L \cong \ddot{V}(\lambda, \mu)$.*

Theorem 3.8. *Let $m, m' \in \mathbb{Z}_{\geq 1}$ and $\ell, \ell' \in \mathbb{Z}_{\geq 0}$. Let $(\lambda, \mu) \in \widehat{\mathcal{J}}^{*n}_{m,\ell}$ and $(\eta, \nu) \in \widehat{\mathcal{J}}^{*n}_{m',\ell'}$. Then the following are equivalent*:

(a) $\ddot{V}(\lambda, \mu) \cong \ddot{V}(\eta, \nu)$.

(b) $m = m'$, $\ell = \ell'$ *and* $\widehat{\lambda/\mu} = \widehat{\eta/\nu} + (r, r)$ *for some* $r \in \mathbb{Z}$.

(c) $m = m'$, $\ell = \ell'$ *and* $(\eta, \nu) = \omega^r_m \cdot (\lambda, \mu)$ *for some* $r \in \mathbb{Z}$.

Remark 3.9. Combining Theorem 3.7 and Theorem 3.8, the classification we obtain agrees with that announced in [Ch4], where he also considers general q and ξ.

An alternative approach to prove these results is to use the result in [Va, Su], where the classification of irreducible modules over $\ddot{H}_n(q)$ of a more general class is obtained. Actually, it is easy to see that the $\ddot{H}_n(q)$-module $\ddot{V}(\lambda, \mu)$ coincides with the unique simple quotient $\ddot{L}(\lambda, \mu)$ of the induced module $\ddot{M}(\lambda, \mu)$ with the notation in [Su].

References

[AST] T. Arakawa, T. Suzuki and A. Tsuchiya, *Degenerate double affine Hecke algebras and conformal field theory*, in *Topological Field Theory, Primitive Forms and Related Topics* ; the proceedings of the 38^{th} Taniguchi symposium, Ed. M Kashiwara et al. (1998), Birkhäuser, 1-34.

[Ch1] I. V. Cherednik, *Special bases of irreducible representations of a degenerate affine Hecke algebra*, Funct. Anal. Appl. 20, No 1 (1986), 76-78.

[Ch2] I. V. Cherednik, *A unification of Knizhnik-Zamolodchikov and Dunkl operators via affine Hecke algebras.* Invent. Math. **106** (1991), 411-431.

[Ch3] I. Cherednik, *Double affine Hecke algebras, Knizhnik-Zamolodchikov equations, and Macdonald's operators*, Inter. Math. Res. Notices, 6 (1992), 171-179.

[Ch4] I. V. Cherednik, *Double affine Hecke algebras and differential Fourier transforms*, Invent. Math. **152**, No 2 (2003), 213-303.

[Ch5] I. V. Cherednik, *Introduction to double Hecke algebras*, Preprint (2004).

[Hu] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge University Press (1990).

[Ra1] A. Ram, *Skew shape representations are irreducible*, preprint (1998)

[Ra2] A. Ram, *Calibrated representations of affine Hecke algebras*, preprint (1998)

[Su] T. Suzuki, *Classification of simple modules over degenerate double affine Hecke algebras of type A*, Inter. Math. Res. Notices, **43** (2003), 2313–2339.

[Va] E. Vasserot, *Induced and simple modules of double affine Hecke algebras*, math.RT/0207127.

Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 606-8502, Japan
E-mail address: takeshi@kurims.kyoto-u.ac.jp

UC Davis, Department of Mathematics, One Shields Ave, Davis, CA 95616-8633
E-mail address: vazirani@math.ucdavis.edu

HALL–LITTLEWOOD FUNCTIONS AND THE A_2 ROGERS–RAMANUJAN IDENTITIES

S. OLE WARNAAR

Abstract. We prove an identity for Hall–Littlewood symmetric functions labelled by the Lie algebra A_2. Through specialization this yields a simple proof of the A_2 Rogers–Ramanujan identities of Andrews, Schilling and the author.

Nous démontrons une identité pour les functions symétriques de Hall–Littlewood associée à l'algèbre de Lie A_2. En spécialisant cette identité, nous obtenons une démonstration simple des identités du type Rogers–Ramanujan associées á A_2 d'Andrews, Schilling et l'auteur.

1. Introduction

The Rogers–Ramanujan identities, given by [10]

$$1+\sum_{n=1}^{\infty}\frac{q^{n^2}}{(1-q)(1-q^2)\cdots(1-q^n)}=\prod_{n=1}^{\infty}\frac{1}{(1-q^{5n-1})(1-q^{5n-4})} \tag{1.1a}$$

and

$$1+\sum_{n=1}^{\infty}\frac{q^{n(n+1)}}{(1-q)(1-q^2)\cdots(1-q^n)}=\prod_{n=1}^{\infty}\frac{1}{(1-q^{5n-2})(1-q^{5n-3})}, \tag{1.1b}$$

are two of the most famous q-series identities, with deep connections with number theory, representation theory, statistical mechanics and various other branches of mathematics.

Many different proofs of the Rogers–Ramanujan identities have been given in the literature, some bijective, some representation theoretic, but the vast majority basic hypergeometric. In 1990, J. Stembridge, building on work of I. Macdonald, found a proof of the Rogers–Ramanujan identities quite unlike any of the previously known proofs. In particular he discovered that Rogers–Ramanujan-type identities may be obtained by appropriately specializing identities for Hall–Littlewood polynomials. The Hall–Littlewood polynomials and, more generally, Hall–Littlewood functions are an important class of symmetric functions, generalizing the well-known Schur functions. Stembridge's Hall–Littlewood approach to Rogers–Ramanujan identities has been further generalized in recent work by Fulman [2], Ishikawa *et al.* [5] and Jouhet and Zeng [7].

Several years ago Andrews, Schilling and the present author generalized the two Rogers–Ramanujan identities to three identities labelled by the Lie algebra A_2 [1]. The simplest of these, which takes the place of (1.1a) when A_1 is replaced by A_2

Work supported by the Australian Research Council.

reads

$$\sum_{n_1,n_2=0}^{\infty} \frac{q^{n_1^2-n_1n_2+n_2^2}}{(q;q)_{n_1}(q;q)_{n_2}(q;q)_{n_1+n_2}} = \prod_{n=1}^{\infty} \frac{1}{(1-q^n)(1-q^{7n-1})^2(1-q^{7n-3})(1-q^{7n-4})(1-q^{7n-6})^2}, \tag{1.2}$$

where $(q;q)_0 = 1$ and $(q;q)_n = \prod_{i=1}^{n}(1-q^i)$ is a q-shifted factorial.

An important question is whether (1.2) and its companions can again be understood in terms of Hall–Littlewood functions. This question is especially relevant since the A_n analogues of the Rogers–Ramanujan identities have so far remained elusive, and an understanding of (1.2) in the context of symmetric functions might provide further insight into the structure of the full A_n generalization of (1.1).

In this paper we will show that the theory of Hall–Littlewood functions may indeed be applied to yield a proof of (1.2). In particular we will prove the following A_2-type identity for Hall–Littlewood functions.

Theorem 1.1. *Let $x = (x_1, x_2, \dots)$, $y = (y_1, y_2, \dots)$ and let $P_\lambda(x;q)$ and $P_\mu(y;q)$ be Hall–Littlewood functions indexed by the partitions λ and μ. Then*

$$\sum_{\lambda,\mu} q^{n(\lambda)+n(\mu)-(\lambda'|\mu')} P_\lambda(x;q) P_\mu(y;q) = \prod_{i\geq 1} \frac{1}{(1-x_i)(1-y_i)} \prod_{i,j\geq 1} \frac{1-x_iy_j}{1-q^{-1}x_iy_j}. \tag{1.3}$$

In the above λ' and μ' are the conjugates of λ and μ, $(\lambda|\mu) = \sum_{i\geq 1}\lambda_i\mu_i$, and $n(\lambda) = \sum_{i\geq 1}(i-1)\lambda_i$.

An appropriate specialization of Theorem 1.1 leads to a q-series identity of [1] which is the key-ingredient in proving (1.2).

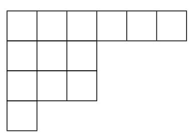
In the next section we give the necessary background material on Hall–Littlewood functions. Section 3 contains a proof of Theorem 1.1 and in Section 4 we present a proof of the A_2 Rogers–Ramanujan identities (1.2) based on Theorem 1.1.

2. Hall-Littlewood functions

We review some basic facts from the theory of Hall-Littlewood functions. For more details the reader may wish to consult Chapter III of Macdonald's book on symmetric functions [9].

Let $\lambda = (\lambda_1, \lambda_2, \dots)$ be a partition, i.e., $\lambda_1 \geq \lambda_2 \geq \dots$ with finitely many λ_i unequal to zero. The length and weight of λ, denoted by $\ell(\lambda)$ and $|\lambda|$, are the number and sum of the non-zero λ_i (called parts), respectively. The unique partition of weight zero is denoted by 0, and the multiplicity of the part i in the partition λ is denoted by $m_i(\lambda)$.

We identify a partition with its diagram or Ferrers graph in the usual way, and, for example, the diagram of $\lambda = (6,3,3,1)$ is given by



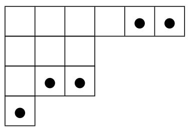
The conjugate λ' of λ is the partition obtained by reflecting the diagram of λ in the main diagonal. Hence $m_i(\lambda) = \lambda'_i - \lambda'_{i+1}$.

A standard statistic on partitions needed repeatedly is

$$n(\lambda) = \sum_{i\geq 1}(i-1)\lambda_i = \sum_{i\geq 1}\binom{\lambda'_i}{2}.$$

We also need the usual scalar product $(\lambda|\mu) = \sum_{i\geq 1}\lambda_i\mu_i$ (which in the notation of [9] would be $|\lambda\mu|$). We will occasionally use this for more general sequences of integers, not necessarily partitions.

If λ and μ are two partions then $\mu \subset \lambda$ iff $\lambda_i \geq \mu_i$ for all $i \geq 1$, i.e., the diagram of λ contains the diagram of μ. If $\mu \subset \lambda$ then the skew-diagram $\lambda - \mu$ denotes the set-theoretic difference between λ and μ, and $|\lambda - \mu| = |\lambda| - |\mu|$. For example, if $\lambda = (6, 3, 3, 1)$ and $\mu = (4, 3, 1)$ then the skew diagram $\lambda - \mu$ is given by the marked squares in



and $|\lambda - \mu| = 5$.

For $\theta = \lambda - \mu$ a skew diagram, its conjugate $\theta' = \lambda' - \mu'$ is the (skew) diagram obtained by reflecting θ in the main diagonal. Following [9] we define the components of θ and θ' by $\theta_i = \lambda_i - \mu_i$ and $\theta'_i = \lambda'_i - \mu'_i$. Quite often we only require knowledge of the sequence of components of a skew diagram θ, and by abuse of notation we will occasionally write $\theta = (\theta_1, \theta_2, \dots)$, even though the components θ_i alone do not fix θ.

A skew diagram θ is a horizontal strip if $\theta'_i \in \{0, 1\}$, i.e., if at most one square occurs in each column of θ. The skew diagram in the above example is a horizontal strip since $\theta' = (1, 1, 1, 0, 1, 1, 0, 0, \dots)$.

Let S_n be the symmetric group, $\Lambda_n = \mathbb{Z}[x_1, \dots, x_n]^{S_n}$ be the ring of symmetric polynomials in n independent variables and Λ the ring of symmetric functions in countably many independent variables.

For $x = (x_1, \dots, x_n)$ and λ a partition such that $\ell(\lambda) \leq n$ the Hall–Littlewood polynomials $P_\lambda(x; q)$ are defined by

$$P_\lambda(x;q) = \sum_{w\in S_n/S_n^\lambda} w\Big(x^\lambda \prod_{\lambda_i>\lambda_j}\frac{x_i - qx_j}{x_i - x_j}\Big). \tag{2.1}$$

Here S_n^λ is the subgroup of S_n consisting of the permutations that leave λ invariant, and $w(f(x)) = f(w(x))$. When $\ell(\lambda) > n$,

$$P_\lambda(x;q) = 0. \tag{2.2}$$

The Hall–Littlewood polynomials are symmetric polynomials in x, homogeneous of degree $|\lambda|$, with coefficients in $\mathbb{Z}[q]$, and form a $\mathbb{Z}[q]$ basis of $\Lambda_n[q]$. Thanks to the stability property $P_\lambda(x_1, \dots, x_n, 0; q) = P_\lambda(x_1, \dots, x_n; q)$ the Hall–Littlewood polynomials may be extended to the Hall–Littlewood functions in an infinite number of variables $x_1, x_2, \dots$ in the usual way, to form a $\mathbb{Z}[q]$ basis of $\Lambda[q]$. The indeterminate q in the Hall–Littlewood symmetric functions serves as a parameter interpolating between the Schur functions and monomial symmetric functions; $P_\lambda(x;0) = s_\lambda(x)$ and $P_\lambda(x;1) = m_\lambda(x)$.

We will also need the symmetric functions $Q_\lambda(x;q)$ (also referred to as Hall-Littlewood functions) defined by

$$Q_\lambda(x;q) = b_\lambda(q) P_\lambda(x;q), \tag{2.3}$$

where

$$b_\lambda(q) = \prod_{i=1}^{\lambda_1} (q;q)_{m_i(\lambda)}.$$

We already mentioned the homogeneity of the Hall–Littlewood functions;

$$P_\lambda(ax;q) = a^{|\lambda|} P_\lambda(x;q), \tag{2.4}$$

where $ax = (ax_1, ax_2, \dots)$. Another useful result is the specialization

$$P_\lambda(1, q, \dots, q^{n-1}; q) = \frac{q^{n(\lambda)} (q;q)_n}{(q;q)_{n-\ell(\lambda)} b_\lambda(q)}, \tag{2.5}$$

where $1/(q;q)_{-m} = 0$ for m a positive integer, so that $P_\lambda(1, q, \dots, q^{n-1}; q) = 0$ if $\ell(\lambda) > n$ in accordance with (2.2). By (2.3) this also implies the particularly simple

$$Q_\lambda(1, q, q^2, \dots; q) = q^{n(\lambda)}. \tag{2.6}$$

The skew Hall–Littlewood functions $P_{\lambda/\mu}$ and $Q_{\lambda/\mu}$ are defined by

$$P_\lambda(x, y; q) = \sum_\mu P_{\lambda/\mu}(x;q) P_\mu(y;q) \tag{2.7}$$

and

$$Q_\lambda(x, y; q) = \sum_\mu Q_{\lambda/\mu}(x;q) Q_\mu(y;q),$$

so that

$$Q_{\lambda/\mu}(x;q) = \frac{b_\lambda(q)}{b_\mu(q)} P_{\lambda/\mu}(x;q). \tag{2.8}$$

An important property is that $P_{\lambda/\mu}$ is zero if $\mu \not\subset \lambda$. Some trivial instances of the skew functions are given by $P_{\lambda/0} = P_\lambda$ and $P_{\lambda/\lambda} = 1$. By (2.8) similar statements apply to $Q_{\lambda/\mu}$.

The Cauchy identity for (skew) Hall–Littlewood functions is given by [11, Lemma 3.1]

$$\sum_\lambda P_{\lambda/\mu}(x;q) Q_{\lambda/\nu}(y;q) = \sum_\lambda P_{\nu/\lambda}(x;q) Q_{\mu/\lambda}(y;q) \prod_{i,j\ge1} \frac{1 - q x_i y_j}{1 - x_i y_j}. \tag{2.9}$$

We conclude our introduction of the Hall–Littlewood functions with the following two important definitions. Let $\lambda \supset \mu$ be partitions such that $\theta = \lambda - \mu$ is a horizontal strip, i.e., $\theta'_i \in \{0, 1\}$. Let I be the set of integers $i \ge 1$ such that $\theta'_i = 1$ and $\theta'_{i+1} = 0$. Then

$$\phi_{\lambda/\mu}(q) = \prod_{i \in I} (1 - q^{m_i(\lambda)}).$$

Similarly, let J be the set of integers $j \ge 1$ such that $\theta'_j = 0$ and $\theta'_{j+1} = 1$. Then

$$\psi_{\lambda/\mu}(q) = \prod_{j \in J} (1 - q^{m_j(\mu)}).$$

For example, if $\lambda = (5,3,2,2)$ and $\mu = (3,3,2)$ then θ is a horizontal strip and $\theta' = (1,1,0,1,1,0,0,\dots)$. Hence $I = \{2,5\}$ and $J = \{3\}$, leading to

$$\phi_{\lambda/\mu}(q) = (1-q^{m_2(\lambda)})(1-q^{m_5(\lambda)}) = (1-q^2)(1-q)$$

and

$$\psi_{\lambda/\mu}(q) = (1-q^{m_3(\mu)}) = (1-q^2).$$

The skew Hall–Littlewood functions $Q_{\lambda/\mu}(x;q)$ and $P_{\lambda/\mu}(x;q)$ can be expressed in terms of $\phi_{\lambda/\mu}(q)$ and $\psi_{\lambda/\mu}(q)$ [9, p. 229]. For our purposes we only require a special instance of this result corresponding to the case that x represents a single variable. Then

$$Q_{\lambda/\mu}(x;q) = \begin{cases} \phi_{\lambda/\mu}(q)x^{|\lambda-\mu|} & \text{if } \lambda-\mu \text{ is a horizontal strip,} \\ 0 & \text{otherwise} \end{cases} \tag{2.10a}$$

and

$$P_{\lambda/\mu}(x;q) = \begin{cases} \psi_{\lambda/\mu}(q)x^{|\lambda-\mu|} & \text{if } \lambda-\mu \text{ is a horizontal strip,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.10b}$$

3. Proof of Theorem 1.1

Throughout this section z represents a single variable.

To establish (1.3) it is enough to show its truth for $x = (x_1,\dots,x_n)$ and $y = (y_1,\dots,y_m)$, and by induction on m it then easily follows that we only need to prove

$$\sum_{\lambda,\mu} q^{n(\lambda)+n(\mu)-(\lambda'|\mu')} P_\lambda(x;q)P_\mu(y,z;q) \\ = \frac{1}{1-z}\prod_{i=1}^{n}\frac{1-zx_i}{1-q^{-1}zx_i}\sum_{\lambda,\mu} q^{n(\lambda)+n(\mu)-(\lambda'|\mu')} P_\lambda(x;q)P_\mu(y;q), \tag{3.1}$$

where we have replaced y_{m+1} by z.

If on the left we replace μ by ν and use (2.7) (with $\lambda \to \nu$ and $x \to z$) we get

$$\mathrm{LHS}(3.1) = \sum_{\lambda,\mu,\nu} q^{n(\lambda)+n(\nu)-(\lambda'|\nu')} P_\lambda(x;q)P_\mu(y;q)P_{\nu/\mu}(z;q).$$

From (2.9) with $\mu = 0$, $x = (x_1,\dots,x_n)$ and $y \to z/q$ it follows that

$$P_\nu(x;q)\prod_{i=1}^{n}\frac{1-zx_i}{1-q^{-1}zx_i} = \sum_\lambda Q_{\lambda/\nu}(z/q;q)P_\lambda(x;q).$$

Using this on the right of (3.1) with λ replaced by ν yields

$$\mathrm{RHS}(3.1) = \frac{1}{1-z}\sum_{\lambda,\mu,\nu} q^{n(\mu)+n(\nu)-(\mu'|\nu')} P_\lambda(x;q)P_\mu(y;q)Q_{\lambda/\nu}(z/q;q).$$

Therefore, by equating coefficients of $P_\lambda(x;q)P_\mu(y;q)$ we find that the problem of proving (1.3) boils down to showing that

$$\sum_\nu q^{n(\lambda)+n(\nu)-(\lambda'|\nu')} P_{\nu/\mu}(z;q) = \frac{1}{1-z}\sum_\nu q^{n(\mu)+n(\nu)-(\mu'|\nu')} Q_{\lambda/\nu}(z/q;q).$$

Next we use (2.10) to arrive at the equivalent but more combinatorial statement that

$$\sum_{\substack{\nu\supset\mu \\ \nu-\mu \text{ hor. strip}}} q^{n(\lambda)+n(\nu)-(\lambda'|\nu')} z^{|\nu-\mu|}\psi_{\nu/\mu}(q) = \frac{1}{1-z} \sum_{\substack{\nu\subset\lambda \\ \lambda-\nu \text{ hor. strip}}} q^{n(\mu)+n(\nu)-(\mu'|\nu')}(z/q)^{|\lambda-\nu|}\phi_{\lambda/\nu}(q). \tag{3.2}$$

To make further progress we need a lemma [12].

Lemma 3.1. *For k a positive integer let $\omega=(\omega_1,\dots,\omega_k)\in\{0,1\}^k$, and let $J=J(\omega)$ be the set of integers j such that $\omega_j=0$ and $\omega_{j+1}=1$. For $\lambda\supset\mu$ partitions let $\theta'=\lambda'-\mu'$ be a skew diagram. Then*

$$\sum_{\substack{\lambda\supset\mu \\ \lambda-\mu \text{ hor. strip} \\ \theta'_i=\omega_i,\ i\in\{1,\dots,k\}}} q^{n(\lambda)} z^{|\lambda-\mu|}\psi_{\lambda/\mu}(q) = \frac{q^{n(\mu)+(\mu'|\omega)}z^{|\omega|}}{1-z}(1-z(1-\omega_k)q^{\mu'_k})\prod_{j\in J}(1-q^{m_j(\mu)}).$$

The restriction $\theta'_i=\omega_i$ for $i\in\{1,\dots,k\}$ in the sum over λ on the left means that the first k parts of λ' are fixed. The remaining parts are free subject only to the condition that $\lambda-\mu$ is a horizontal strip, i.e., that $\lambda'_i-\mu'_i\in\{0,1\}$.

In view of Lemma 3.1 it is natural to rewrite the left side of (3.2) as

$$\mathrm{LHS}(3.2)=\sum_{\omega\in\{0,1\}^{\lambda_1}} \sum_{\substack{\nu\supset\mu \\ \nu-\mu \text{ hor. strip} \\ \theta'_i=\omega_i,\ i\in\{1,\dots,\lambda_1\}}} q^{n(\lambda)+n(\nu)-(\lambda'|\mu')-(\lambda'|\omega)} z^{|\nu-\mu|}\psi_{\nu/\mu}(q),$$

where $\theta=\nu-\mu$, and where we have used that $\theta'_i\in\{0,1\}$ as follows from the fact that $\nu-\mu$ is a horizontal strip.

Now the sum over ν can be performed by application of Lemma 3.1 with $\lambda\to\nu$ and $k\to\lambda_1$, resulting in

$$\begin{aligned}\mathrm{LHS}(3.2)=\frac{q^{n(\lambda)+n(\mu)-(\lambda'|\mu')}}{1-z} &\sum_{\omega\in\{0,1\}^{\lambda_1}} q^{(\mu'|\omega)-(\lambda'|\omega)}z^{|\omega|} \\ &\qquad\times(1-z(1-\omega_{\lambda_1})q^{\mu'_{\lambda_1}})\prod_{j\in J}(1-q^{m_j(\mu)})\end{aligned}$$

with $J=J(\omega)\subset\{1,\dots,\lambda_1-1\}$ the set of integers j such that $\omega_j<\omega_{j+1}$.

For the right-hand side of (3.2) we introduce the notation $\tau_i=\lambda'_i-\nu'_i$, so that the sum over ν can be rewritten as a sum over $\tau\in\{0,1\}^{\lambda_1}$. Using that

$$n(\nu)=\sum_{i=1}^{\lambda_1}\binom{\nu'_i}{2}=\sum_{i=1}^{\lambda_1}\binom{\lambda'_i-\tau_i}{2}=n(\lambda)-(\lambda'|\tau)+|\tau|$$

this yields

$$\mathrm{RHS}(3.2)=\frac{q^{n(\lambda)+n(\mu)-(\lambda'|\mu')}}{1-z}\sum_{\tau\in\{0,1\}^{\lambda_1}} q^{(\mu'|\tau)-(\lambda'|\tau)}z^{|\tau|}\prod_{i\in I}(1-q^{m_i(\lambda)}),$$

with $I=I(\tau)\subset\{1,\dots,\lambda_1\}$ the set of integers i such that $\tau_i>\tau_{i+1}$ (with the convention that $\lambda_1\in I$ if $\tau_{\lambda_1}=1$).

Equating the above two results for the respective sides of (3.2) gives

$$\sum_{\omega\in\{0,1\}^{\lambda_1}} q^{(\mu'|\omega)-(\lambda'|\omega)}z^{|\omega|}(1-z(1-\omega_{\lambda_1})q^{\mu'_{\lambda_1}})\prod_{j\in J}(1-q^{m_j(\mu)}) \\ =\sum_{\tau\in\{0,1\}^{\lambda_1}} q^{(\mu'|\tau)-(\lambda'|\tau)}z^{|\tau|}\prod_{i\in I}(1-q^{m_i(\lambda)}).$$

Using that $m_i(\lambda)=\lambda'_i-\lambda'_{i+1}$ it is not hard to see that this is the

$$k\to\lambda_1,\quad b_{k+1}\to 1,\quad a_i\to zq^{\mu'_i},\quad b_i\to q^{\lambda'_i},\quad i\in\{1,\dots,\lambda_1\}$$

specialization of the more general

$$\sum_{\omega\in\{0,1\}^k}(a/b)^\omega(1-(1-\omega_k)a_k/b_{k+1})\prod_{j\in J}(1-a_j/a_{j+1}) \\ =\sum_{\tau\in\{0,1\}^k}(a/b)^\tau\prod_{i\in I}(1-b_i/b_{i+1}),$$

where $(a/b)^\omega=\prod_{i=1}^k(a_i/b_i)^{\omega_i}$ and $(a/b)^\tau=\prod_{i=1}^k(a_i/b_i)^{\tau_i}$. Obviously, the set $J\subset\{1,\dots,k-1\}$ should now be defined as the set of integers j such that $\omega_j<\omega_{j+1}$ and the the set $I\subset\{1,\dots,k\}$ as the set of integers i such that $\tau_i>\tau_{i+1}$ (with the convention that $k\in I$ if $\tau_k=1$).

Next we split both sides into the sum of two terms as follows:

$$\Bigg(\sum_{\omega\in\{0,1\}^k}-(a_k/b_{k+1})\sum_{\substack{\omega\in\{0,1\}^k\\ \omega_k=0}}\Bigg)(a/b)^\omega\prod_{j\in J}(1-a_j/a_{j+1}) \\ =\Bigg(\sum_{\tau\in\{0,1\}^k}-(b_k/b_{k+1})\sum_{\substack{\tau\in\{0,1\}^k\\ \tau_k=1}}\Bigg)(a/b)^\tau\prod_{\substack{i\in I\\ i\neq k}}(1-b_i/b_{i+1}).$$

Equating the first sum on the left with the first sum on the right yields

$$\sum_{\omega\in\{0,1\}^k}(a/b)^\omega\prod_{j\in J}(1-a_j/a_{j+1})=\sum_{\tau\in\{0,1\}^k}(a/b)^\tau\prod_{\substack{i\in I\\ i\neq k}}(1-b_i/b_{i+1}). \tag{3.3}$$

If we equate the second sum on the left with the second sum on the right and use that $k-1\notin J(\omega)$ if $\omega_k=0$ and $k-1\notin I(\tau)$ if $\tau_k=1$, we obtain $(a_k/b_{k+1})((3.3)_{k\to k-1})$.

Slightly changing our earlier convention we thus need to prove that

$$\sum_{\omega\in\{0,1\}^k}(a/b)^\omega\prod_{j\in J}(1-a_j/a_{j+1})=\sum_{\tau\in\{0,1\}^k}(a/b)^\tau\prod_{i\in I}(1-b_i/b_{i+1}), \tag{3.4}$$

where from now on $I\subset\{1,\dots,k-1\}$ denotes the set of integers i such that $\tau_i>\tau_{i+1}$ (so that no longer $k\in I$ if $\tau_k=1$). It is not hard to see by multiplying out the respective products that boths sides yield $((1+\sqrt2)^{k+1}-(1-\sqrt2)^{k+1})/(2\sqrt2)$ terms. To see that the terms on the left and right are in one-to-one correspondence we again resort to induction. First, for $k=1$ it is readily checked that both sides yield $1+a_1/b_1$. For $k=2$ we on the left get

$$\underbrace{1}_{\omega=(0,0)} + \underbrace{(a_1/b_1)}_{\omega=(1,0)} + \underbrace{(a_2/b_2)(1-a_1/a_2)}_{\omega=(0,1)} + \underbrace{(a_1a_2/b_1b_2)}_{\omega=(1,1)}$$

and on the right

$$\underbrace{1}_{\tau=(0,0)} + \underbrace{(a_1/b_1)(1-b_1/b_2)}_{\tau=(1,0)} + \underbrace{(a_2/b_2)}_{\tau=(0,1)} + \underbrace{(a_1a_2/b_1b_2)}_{\tau=(1,1)}$$

which both give

$$1+a_1/b_1+a_2/b_2-a_1/b_2+a_1a_2/b_1b_2.$$

Let us now assume that (3.4) has been shown to be true for $1\le k\le K-1$ with $K\ge 3$ and prove the case $k=K$.

On the left of (3.4) we split the sum over ω according to

$$\sum_{\omega\in\{0,1\}^k} = \sum_{\substack{\omega\in\{0,1\}^k\\ \omega_1=1}} + \sum_{\substack{\omega\in\{0,1\}^k\\ \omega_1=\omega_2=0}} + \sum_{\substack{\omega\in\{0,1\}^k\\ \omega_1=0,\ \omega_2=1}}.$$

Defining $\bar\omega\in\{0,1\}^{k-1}$ and $\bar{\bar\omega}\in\{0,1\}^{k-2}$ by $\bar\omega=(\omega_2,\dots,\omega_k)$ and $\bar{\bar\omega}=(\omega_3,\dots,\omega_k)$, and also setting and $\bar a_j=a_{j+1}$, $\bar b_j=b_{j+1}$, and $\bar{\bar a}_j=a_{j+2}$, $\bar{\bar b}_j=b_{j+2}$, this leads to

$$\begin{aligned}
\mathrm{LHS}(3.4) &= (a_1/b_1)\sum_{\bar\omega\in\{0,1\}^{k-1}}(\bar a/\bar b)^{\bar\omega}\prod_{j\in J(\bar\omega)}(1-\bar a_j/\bar a_{j+1})\\
&\quad+\sum_{\substack{\bar\omega\in\{0,1\}^{k-1}\\ \bar\omega_1=0}}(\bar a/\bar b)^{\bar\omega}\prod_{j\in J(\bar\omega)}(1-\bar a_j/\bar a_{j+1})\\
&\quad+(1-a_1/a_2)\sum_{\substack{\bar\omega\in\{0,1\}^{k-1}\\ \bar\omega_1=1}}(\bar a/\bar b)^{\bar\omega}\prod_{j\in J(\bar\omega)}(1-\bar a_j/\bar a_{j+1})\\
&=(1+a_1/b_1)\sum_{\bar\omega\in\{0,1\}^{k-1}}(\bar a/\bar b)^{\bar\omega}\prod_{j\in J(\bar\omega)}(1-\bar a_j/\bar a_{j+1})\\
&\quad-(a_1/a_2)\sum_{\substack{\bar\omega\in\{0,1\}^{k-1}\\ \bar\omega_1=1}}(\bar a/\bar b)^{\bar\omega}\prod_{j\in J(\bar\omega)}(1-\bar a_j/\bar a_{j+1})\\
&=(1+a_1/b_1)\sum_{\bar\omega\in\{0,1\}^{k-1}}(\bar a/\bar b)^{\bar\omega}\prod_{j\in J(\bar\omega)}(1-\bar a_j/\bar a_{j+1})\\
&\quad-(a_1/b_2)\sum_{\bar{\bar\omega}\in\{0,1\}^{k-2}}(\bar{\bar a}/\bar{\bar b})^{\bar{\bar\omega}}\prod_{j\in J(\bar{\bar\omega})}(1-\bar{\bar a}_j/\bar{\bar a}_{j+1}).
\end{aligned}$$

On the right of (3.4) we split the sum over τ according to

$$\sum_{\tau\in\{0,1\}^k} = \sum_{\substack{\tau\in\{0,1\}^k\\ \tau_1=0}} + \sum_{\substack{\tau\in\{0,1\}^k\\ \tau_1=\tau_2=1}} + \sum_{\substack{\tau\in\{0,1\}^k\\ \tau_1=1,\ \tau_2=0}}.$$

Defining $\bar{\tau} \in \{0,1\}^{k-1}$ and $\bar{\bar{\tau}} \in \{0,1\}^{k-2}$ by $\bar{\tau} = (\tau_2, \dots, \tau_k)$ and $\bar{\bar{\tau}} = (\tau_3, \dots, \tau_k)$, this yields

$$
\begin{aligned}
\mathrm{RHS}(3.4) = & \sum_{\bar{\tau}\in\{0,1\}^{k-1}} (\bar{a}/\bar{b})^{\bar{\tau}} \prod_{j\in J(\bar{\tau})} (1-\bar{b}_j/\bar{b}_{j+1}) \\
& + (a_1/b_1) \sum_{\substack{\bar{\tau}\in\{0,1\}^{k-1}\\ \bar{\tau}_1=1}} (\bar{a}/\bar{b})^{\bar{\tau}} \prod_{j\in J(\bar{\tau})} (1-\bar{b}_j/\bar{b}_{j+1}) \\
& + (a_1/b_1)(1-b_1/b_2) \sum_{\substack{\bar{\tau}\in\{0,1\}^{k-1}\\ \bar{\tau}_1=0}} (\bar{a}/\bar{b})^{\bar{\tau}} \prod_{j\in J(\bar{\tau})} (1-\bar{b}_j/\bar{b}_{j+1}) \\
= & (1+a_1/b_1) \sum_{\bar{\tau}\in\{0,1\}^{k-1}} (\bar{a}/\bar{b})^{\bar{\tau}} \prod_{j\in J(\bar{\tau})} (1-\bar{b}_j/\bar{b}_{j+1}) \\
& - (a_1/b_2) \sum_{\substack{\bar{\tau}\in\{0,1\}^{k-1}\\ \bar{\tau}_1=0}} (\bar{a}/\bar{b})^{\bar{\tau}} \prod_{j\in J(\bar{\tau})} (1-\bar{b}_j/\bar{b}_{j+1}) \\
= & (1+a_1/b_1) \sum_{\bar{\tau}\in\{0,1\}^{k-1}} (\bar{a}/\bar{b})^{\bar{\tau}} \prod_{j\in J(\bar{\tau})} (1-\bar{b}_j/\bar{b}_{j+1}) \\
& - (a_1/b_2) \sum_{\bar{\bar{\tau}}\in\{0,1\}^{k-2}} (\bar{\bar{a}}/\bar{\bar{b}})^{\bar{\bar{\tau}}} \prod_{j\in J(\bar{\bar{\tau}})} (1-\bar{\bar{b}}_j/\bar{\bar{b}}_{j+1}).
\end{aligned}
$$

By our induction hypothesis this equates with the previous expression for the left-hand side of (3.4), completing the proof.

4. The A_2 Rogers–Ramanujan identities

Let $(a;q)_0 = 1$, $(a;q)_n = \prod_{i=1}^n (1-aq^{i-1})$ and $(a_1,\dots,a_k;q)_n = (a_1;q)_n \cdots (a_k;q)_n$.

Proposition 4.1. *There holds*

$$
\sum_{\lambda,\mu} \frac{a^{|\lambda|} b^{|\mu|} q^{(\lambda'|\lambda')+(\mu'|\mu')-(\lambda'|\mu')}}{(q;q)_{n-\ell(\lambda)}(q;q)_{m-\ell(\mu)} b_\lambda(q) b_\mu(q)} = \frac{(abq;q)_{n+m}}{(q,aq,abq;q)_n (q,bq,abq;q)_m}. \tag{4.1}
$$

Proof. In Theorem 1.1 set $x_i = aq^i$ for $1 \le i \le n$, $x_i = 0$ for $i > n$, $y_j = bq^j$ for $1 \le j \le m$ and $y_j = 0$ for $j > m$. Using the homogeneity (2.4) and specialization (2.5), and noting that $2n(\lambda) + |\lambda| = (\lambda'|\lambda')$, gives (4.1). □

We remark that (4.1) is a bounded version of the A_2 case of the following identity for the A_n root system due to Hua [4] (and corrected in [3]):

$$
\sum_{\lambda^{(1)},\dots,\lambda^{(n)}} \frac{q^{\frac{1}{2}\sum_{i,j=1}^n C_{ij}(\lambda^{(i)\prime}|\lambda^{(j)\prime})} \prod_{i=1}^n a_i^{|\lambda^{(i)}|}}{\prod_{i=1}^n b_{\lambda^{(i)}}(q)} = \prod_{\alpha\in\Delta_+} \frac{1}{(a^\alpha q;q)_\infty}. \tag{4.2}
$$

Here $C_{ij} = 2\delta_{i,i} - \delta_{i,j-1} - \delta_{i,j+1}$ is the (i,j) entry of the A_n Cartan matrix and Δ_+ is the set of positive roots of A_n, i.e., the set (of cardinality $\binom{n+1}{2}$) of roots of the form $\alpha_i + \alpha_{i+1} + \cdots + \alpha_j$ with $1 \le i \le j \le n$, where $\alpha_1, \dots, \alpha_n$ are the simple roots of A_n. Furthermore, if $\alpha = \alpha_i + \alpha_{i+1} + \cdots + \alpha_j$ then $a^\alpha = a_i a_{i+1} \cdots a_j$.

For $M=(M_1,\dots,M_n)$ with M_i a non-negative integer, we define the following bounded analogue of the sum in (4.2):

$$R_M(a_1,\dots,a_n;q)=\sum_{\lambda^{(1)},\dots,\lambda^{(n)}}\frac{q^{\frac{1}{2}\sum_{i,j=1}^n C_{ij}(\lambda^{(i)\prime}|\lambda^{(j)\prime})}\prod_{i=1}^n a_i^{|\lambda^{(i)}|}}{\prod_{i=1}^n (q;q)_{M_i-\ell(\lambda^{(i)})}b_{\lambda^{(i)}}(q)}.$$

By construction $R_M(a_1,\dots,a_n;q)$ satisfies the following invariance property.

Lemma 4.1. *We have*

$$\sum_{r_1=0}^{M_1}\cdots\sum_{r_n=0}^{M_n}\frac{q^{\frac{1}{2}\sum_{i,j=1}^n C_{ij}r_ir_j}\prod_{i=1}^n a_i^{r_i}}{\prod_{i=1}^n (q;q)_{M_i-r_i}}R_r(a_1,\dots,a_n;q)=R_M(a_1,\dots,a_n;q).$$

Proof. Take the definition of R_M given above and replace each of $\lambda^{(1)},\dots,\lambda^{(n)}$ by its conjugate. Then introduce the non-negative integer r_i and the partition $\mu^{(i)}$ with largest part not exceeding r_i through $\lambda^{(i)}=(r_i,\mu_1^{(i)},\mu_2^{(i)},\dots)$. Since $b_{\lambda'}(q)=(q;q)_{r-\mu_1}b_{\mu'}(q)$ for $\lambda=(r,\mu_1,\mu_2,\dots)$ this implies the identity of the lemma after again replacing each of $\mu^{(1)},\dots,\mu^{(n)}$ by its conjugate. $\square$

Next is the observation that the left-hand side of (4.1) corresponds to $R_{(n,m)}(a,b;q)$. Hence we may reformulate the A_2 instance of Lemma 4.1.

Theorem 4.1. *For M_1 and M_2 non-negative integers*

$$\sum_{r_1=0}^{M_1}\sum_{r_2=0}^{M_2}\frac{a^{r_1}b^{r_2}q^{r_1^2-r_1r_2+r_2^2}}{(q;q)_{M_1-r_1}(q;q)_{M_2-r_2}}\frac{(abq;q)_{r_1+r_2}}{(q,aq,abq;q)_{r_1}(q,bq,abq;q)_{r_2}}=\frac{(abq;q)_{M_1+M_2}}{(q,aq,abq;q)_{M_1}(q,bq,abq;q)_{M_2}}. \tag{4.3}$$

To see how this leads to the A_2 Rogers–Ramanujan identity (1.2) and its higher moduli generalizations, let k_1,k_2,k_3 be integers such that $k_1+k_2+k_3=0$. Making the substitutions

$$\begin{aligned} &r_1\to r_1-k_1-k_2, && a\to q^{k_2-k_3}, && M_1\to M_1-k_1-k_2,\\ &r_2\to r_2-k_1, && b\to q^{k_1-k_2}, && M_2\to M_2-k_1, \end{aligned}$$

in (4.3), we obtain

$$\sum_{r_1=0}^{M_1}\sum_{r_2=0}^{M_2}\frac{q^{r_1^2-r_1r_2+r_2^2}}{(q;q)_{M_1-r_1}(q;q)_{M_2-r_2}(q;q)_{r_1+r_2}^2}\begin{bmatrix}r_1+r_2\\r_1+k_1\end{bmatrix}\begin{bmatrix}r_1+r_2\\r_1+k_2\end{bmatrix}\begin{bmatrix}r_1+r_2\\r_1+k_3\end{bmatrix}=\frac{q^{\frac{1}{2}(k_1^2+k_2^2+k_3^2)}}{(q)_{M_1+M_2}^2}\begin{bmatrix}M_1+M_2\\M_1+k_1\end{bmatrix}\begin{bmatrix}M_1+M_2\\M_1+k_2\end{bmatrix}\begin{bmatrix}M_1+M_2\\M_1+k_3\end{bmatrix}, \tag{4.4}$$

where

$$\begin{bmatrix}n\\m\end{bmatrix}=\begin{bmatrix}n\\m\end{bmatrix}_q=\begin{cases}\dfrac{(q^{n-m+1};q)_m}{(q;q)_m} & \text{for } m\geq 0,\\ 0 & \text{otherwise}\end{cases}$$

is a q-binomial coefficient. The identity (4.4) which is equivalent to the type-II A_2 Bailey lemma of [1, Theorem 4.3].

The idea is now to apply (4.4) to the A_2 Euler identity [1, Equation (5.15)]

$$\sum_{k_1+k_2+k_3=0} q^{\frac{3}{2}(k_1^2+k_2^2+k_3^2)} \times \sum_{w\in S_3} \epsilon(w) \prod_{i=1}^{3} q^{\frac{1}{2}(3k_i-w_i+i)^2-w_ik_i} \begin{bmatrix} M_1+M_2 \\ M_1+3k_i-w_i+i \end{bmatrix} = \begin{bmatrix} M_1+M_2 \\ M_1 \end{bmatrix}, \tag{4.5}$$

where $w \in S_3$ is a permutation of $(1,2,3)$ and $\epsilon(w)$ denotes the signature of w.

Replacing M_1, M_2 by r_1, r_2 in (4.5), then multiplying both sides by

$$\frac{q^{r_1^2-r_1r_2+r_2^2}}{(q;q)_{M_1-r_1}(q;q)_{M_2-r_2}(q;q)_{r_1+r_2}^2},$$

and finally summing over r_1 and r_2 using (4.4) (with $k_i \to 3k_i - w_i + i$), yields

$$\sum_{k_1+k_2+k_3=0} q^{\frac{3}{2}(k_1^2+k_2^2+k_3^2)} \sum_{w\in S_3} \epsilon(w) \prod_{i=1}^{3} q^{(3k_i-w_i+i)^2-w_ik_i} \begin{bmatrix} M_1+M_2 \\ M_1+3k_i-w_i+i \end{bmatrix} = \sum_{r_1=0}^{M_1}\sum_{r_2=0}^{M_2} \frac{q^{r_1^2-r_1r_2+r_2^2}(q;q)_{M_1+M_2}^2}{(q;q)_{M_1-r_1}(q;q)_{M_2-r_2}(q;q)_{r_1}(q;q)_{r_2}(q;q)_{r_1+r_2}}. \tag{4.6}$$

Letting M_1 and M_2 tend to infinity, and using the Vandermonde determinant

$$\sum_{w\in S_3} \epsilon(w) \prod_{i=1}^{3} x_i^{i-w_i} = \prod_{1\le i<j\le 3} (1 - x_j x_i^{-1})$$

with $x_i \to q^{7k_i+2i}$, gives

$$\frac{1}{(q;q)_\infty^3} \sum_{k_1+k_2+k_3=0} q^{\frac{21}{2}(k_1^2+k_2^2+k_3^2)-k_1-2k_2-3k_3} \times (1-q^{7(k_2-k_1)+2})(1-q^{7(k_3-k_2)+2})(1-q^{7(k_3-k_1)+4}) = \sum_{r_1,r_2=0}^{\infty} \frac{q^{r_1^2-r_1r_2+r_2^2}}{(q;q)_{r_1}(q;q)_{r_2}(q;q)_{r_1+r_2}}.$$

Finally, by the A_2 Macdonald identity [8]

$$\sum_{k_1+k_2+k_3=0} \prod_{i=1}^{3} x_i^{3k_i} q^{\frac{3}{2}k_i^2-ik_i} \prod_{1\le i<j\le 3} (1 - x_j x_i^{-1} q^{k_j-k_i}) = (q;q)_\infty^2 \prod_{1\le i<j\le 3} (x_i^{-1}x_j, qx_ix_j^{-1}; q)_\infty$$

with $q \to q^7$ and $x_i \to q^{2i}$ this becomes

$$\sum_{r_1,r_2=0}^{\infty} \frac{q^{r_1^2-r_1r_2+r_2^2}}{(q;q)_{r_1}(q;q)_{r_2}(q;q)_{r_1+r_2}} = \frac{(q^2,q^2,q^3,q^4,q^5,q^5,q^7,q^7;q^7)_\infty}{(q;q)_\infty^3}.$$

This result is easily recognized as the A_2 Rogers–Ramanujan identity (1.2).

The identity (4.6) can be further iterated using (4.4). Doing so and repeating the above calculations (requiring the Vandermonde determinant with $x_i \to q^{(3n+1)k_i+ni}$

and the Macdonald identity with $q \to q^{3n+1}$ and $x_i \to q^{ni}$) yields the following A_2 Rogers–Ramanujan-type identity for modulus $3n+1$ [1, Theorem 5.1; $i=k$]:

$$\sum_{\substack{\lambda,\mu \\ \ell(\lambda),\ell(\mu)\le n-1}} \frac{q^{(\lambda|\lambda)+(\mu|\mu)-(\lambda|\mu)}}{b_{\lambda'}(q)b_{\mu'}(q)(q;q)_{\lambda_{n-1}+\mu_{n-1}}} = \frac{(q^n,q^n,q^{n+1},q^{2n},q^{2n+1},q^{2n+1},q^{3n+1},q^{3n+1};q^{3n+1})_\infty}{(q;q)_\infty^3}.$$

In the large n limit ones recovers the A_2 case of Hua's identity (4.2) with $a_1=a_2=1$.

References

1. G. E. Andrews, A. Schilling and S. O. Warnaar, *An A_2 Bailey lemma and Rogers–Ramanujan-type identities*, J. Amer. Math. Soc. **12** (1999), 677–702.
2. J. Fulman, *The Rogers–Ramanujan identities, the finite general linear groups, and the Hall–Littlewood polynomials*, Proc. Amer. Math. Soc. **128** (2000), 17–25.
3. J. Fulman, *A probabilistic proof of the Rogers–Ramanujan identities*, Bull. London Math. Soc. **33** (2001), 397–407.
4. J. Hua, *Counting representations of quivers over finite fields*, J. Algebra **226** (2000), 1011–1033.
5. M. Ishikawa, F. Jouhet and J. Zeng, *A generalization of Kawanaka's identity for Hall–Littlewood polynomials and applications*, preprint.
6. F. Jouhet and J. Zeng, *Some new identities for Schur functions*, Adv. Appl. Math. **27** (2001), 493–509.
7. F. Jouhet and J. Zeng, *New identities of Hall–Littlewood polynomials and applications*, The Ramanujan J., to appear.
8. I. G. Macdonald, *Affine root systems and Dedekind's η-function*, Inv. Math. **15** (1972), 91–143.
9. I. G. Macdonald, *Symmetric functions and Hall polynomials*, second edition, (Oxford University Press, New-York, 1995).
10. L. J. Rogers, *Second memoir on the expansion of certain infinite products*, Proc. London Math. Soc. **25** (1894), 318–343.
11. J. R. Stembridge, *Hall–Littlewood functions, plane partitions, and the Rogers–Ramanujan identities*, Trans. Amer. Math. Soc. **319** (1990), 469–498.
12. S. O. Warnaar, *Hall–Littlewood functions and the A_2 Rogers–Ramanujan identities*, Adv. Math. to appear, ArXiv:math.CO/0410592.

Department of Mathematics and Statistics, The University of Melbourne, VIC 3010, Australia

E-mail address: warnaar@ms.unimelb.edu.au

Solution Spaces of H-Systems and the Ore-Sato Theorem*

S. A. Abramov†
Russian Academy of Sciences
Dorodnicyn Computing Centre
Vavilova 40, 119991, Moscow GSP-1, Russia
`sabramov@ccas.ru`

M. Petkovšek‡
Faculty of Mathematics and Physics
University of Ljubljana,
Jadranska 19, SI-1000 Ljubljana, Slovenia
`marko.petkovsek@fmf.uni-lj.si`

Abstract

An H-*system* is a system of first-order linear homogeneous difference equations for a single unknown function T, with coefficients which are polynomials with complex coefficients. We consider solutions of H-systems which are of the form $T: \operatorname{dom}(T) \to \mathbb{C}$ where either $\operatorname{dom}(T) = \mathbb{Z}^d$, or $\operatorname{dom}(T) = \mathbb{Z}^d \setminus S$ and S is the set of integer singularities of the system. It is shown that any natural number is the dimension of the solution space of some H-system, and that in the case $d \geq 2$ there are H-systems whose solution space is infinite-dimensional. The relationships between dimensions of solution spaces in the two cases $\operatorname{dom}(T) = \mathbb{Z}^d$ and $\operatorname{dom}(T) = \mathbb{Z}^d \setminus S$ are investigated. Finally we give an appropriate formulation of the Ore-Sato theorem on possible forms of solutions of H-systems in this setting.

Résumé

Par un H-*système* nous désignons un système des équations aux differences linéaires homogènes pour une seule fonction inconnue T, à coefficients polynomiaux sur le corps des nombres complexes. Nous considérons les solutions des H-systèmes de la forme $T: \operatorname{dom}(T) \to \mathbb{C}$ où soit $\operatorname{dom}(T) = \mathbb{Z}^d$, soit $\operatorname{dom}(T) = \mathbb{Z}^d \setminus S$, et S est l'ensemble

*The work is partially supported by the ECO-NET program of the French Foreign Affairs Ministry.
†Partially supported by RFBR grant 04-01-00757.
‡Partially supported by MŠZŠ RS under grant P1-0294.

des singularités entières du système. Nous montrons que chaque nombre naturel est égal à la dimension de l'éspace des solutions d'un H-système, et que dans le cas $d \geq 2$ il y a des H-systèmes dont la dimension de l'éspace des solutions est infinie. Les rélations entre les dimensions des éspaces des solutions dans les cas $\mathrm{dom}(T) = \mathbb{Z}^d$ et $\mathrm{dom}(T) = \mathbb{Z}^d \setminus S$ sont recherchées. Enfin nous présentons une formulation propre du théorême d'Ore-Sato sur les formes possibles des solutions des H-systèmes.

1 Introduction

Linear homogeneous recurrence equations with polynomial coefficients and systems of such equations play a significant role in combinatorics and in the theory of hypergeometric functions; the question of the dimension of the space of solutions of such systems is of great importance for many problems.

Let $n_1, \ldots, n_d$ be variables ranging over the integers and E_{n_i} the corresponding shift operators, acting on functions (sequences) of $n_1, \ldots, n_d$ by $E_{n_i} f(n_1, \ldots, n_i) = f(n_1, \ldots, n_i + 1, \ldots, n_d)$, $i = 1, \ldots, d$. We consider H-systems, i.e., systems of equations of the form $f_i\, E_{n_i} T = g_i T$, where $f_i, g_i \in \mathbb{C}[n_1, \ldots, n_d] \setminus \{0\}$ for $i = 1, \ldots, d$. The notion of singular points (singularities) of such systems can be defined in the usual way. Such singularities make obstacles (sometimes insuperable) for continuation of partial solutions of the system on all of $\mathbb{Z}^d$.

In this paper we consider two spaces of solutions of H-systems: the space V_1 of solutions defined everywhere on $\mathbb{Z}^d$, and the space V_2 of solutions that are defined at all nonsingular points of $\mathbb{Z}^d$ (more precisely, if W is the set of all solutions of a given system that are defined at least at all non-singular elements of $\mathbb{Z}^d$, then V_2 contains the restrictions of all elements of W to the set of all non-singular elements of $\mathbb{Z}^d$). In Sections 3 and 4 we investigate the dimensions of the spaces V_1, V_2. It is well known [6] that if (in the case $d = 1$) one considers the germs of sequences at infinity (i.e., classes of sequences which agree from some point on), then the dimension of the solution space is 1. However, the situation is different with $\dim V_1$ and $\dim V_2$. In Section 3 we prove for the case $d = 1$ that if the equation has singularities then $1 \leq \dim V_1 < \dim V_2 < \infty$, and for any integers s, t such that $1 \leq s < t$ there exists an equation with $\dim V_1 = s$ and $\dim V_2 = t$ (the case where there is no singularity is trivial: $\dim V_1 = \dim V_2 = 1$). In turn, in Section 4 we show that in the case $d > 1$ the possibilities are even richer: for any $s, t \in \mathbb{Z}_+ \cup \{\infty\}$ there exists an H-system with $\dim V_1 = s$ and $\dim V_2 = t$.

In Section 5 we revisit the Sato-Ore theorem [5, 7] and show that, contrary to some interpretations in the literature (e.g., [3, 4]), this theorem does not imply that any solution of an H-system is of the form

$$R(n_1, \ldots, n_d) \frac{\prod_{i=1}^{p} \Gamma(a_{i,1} n_1 + \cdots + a_{i,d} n_d + \alpha_i)}{\prod_{j=1}^{q} \Gamma(b_{j,1} n_1 + \cdots + b_{j,d} n_d + \beta_j)}\, u_1^{n_1} \cdots u_d^{n_d}, \tag{1}$$

where $R \in \mathbb{C}(x_1, \ldots, x_d)$, $a_{ik}, b_{jk} \in \mathbb{Z}$, and $\alpha_i, \beta_j \in \mathbb{C}$ (for the case when the solution of the system is holonomic, and R is required to be a polynomial, we have already noted this in [2]). Finally we give an appropriate corollary of the Ore-Sato theorem on possible forms of solutions of systems under consideration.

We write $p \perp q$ if polynomials $p, q \in \mathbb{C}[x_1, \ldots, x_d]$ are relatively prime. We call a set $A \subseteq \mathbb{Z}^d$ *algebraic* if there is a polynomial $p \in \mathbb{C}[x_1, \ldots, x_d] \setminus \{0\}$ which vanishes on A.

Definition 1 *Let E denote the shift operator corresponding to x, so that $Ef(x) = f(x+1)$ for every $f \in \mathbb{C}(x)$. A rational function $u \in \mathbb{C}(x)$ is* shift-reduced *if there are $a, b \in \mathbb{C}[x]$ such that $u = a/b$ and $a \perp E^k b$ for all $k \in \mathbb{Z}$.*

Theorem 1 *For every rational function $F \in \mathbb{C}(x)$ there are rational functions $u, v \in \mathbb{C}(x)$ such that*

(i) $F = u \cdot \frac{Ev}{v}$,

(ii) *u is shift-reduced.*

Definition 2 *If u, v, F are as in Theorem 1, (u, v) is a* rational normal form *of F.*

Theorem 2 *Let (u, v) and (u_1, v_1) be two rational normal forms of $F \in \mathbb{C}(x) \setminus \{0\}$. Write $u = p/q$ and $u_1 = p_1/q_1$ where $p, q, p_1, q_1 \in \mathbb{C}[x]$, $p \perp q$, and $p_1 \perp q_1$. Then $\deg p = \deg p_1$ and $\deg q = \deg q_1$.*

For proofs of Theorems 1 and 2, see [1].

2 H-systems and their solution spaces

Definition 3 *An H-*system[1] *is a system of equations*

$$f_i(n_1, \ldots, n_d) T(n_1, \ldots, n_i + 1, \ldots, n_d) = g_i(n_1, \ldots, n_d) T(n_1, \ldots, n_i, \ldots, n_d), \quad \text{for } i = 1, 2, \ldots, d, \tag{2}$$

where $f_i, g_i \in \mathbb{C}[n_1, \ldots, n_d] \setminus \{0\}$ and $f_i \perp g_i$. We say that a d-variate sequence T (i.e., a function $T : \mathrm{dom}(T) \to \mathbb{C}$) is a solution *of (2) if (2) is satisfied for all $(n_1, \ldots, n_i, \ldots, n_d) \in \mathrm{dom}(T)$ such that $(n_1, \ldots, n_i + 1, \ldots, n_d) \in \mathrm{dom}(T)$ as well.*

Definition 4 *Let A be an H-system of the form (2).*

A d-tuple $(n_1, \ldots, n_d) \in \mathbb{Z}^d$ is a trailing integer singularity *of A if there exists i, $1 \le i \le d$, such that $g_i(n_1, \ldots, n_d) = 0$. A d-tuple $(n_1, \ldots, n_d) \in \mathbb{Z}^d$ is a* leading integer singularity *of A if there exists i, $1 \le i \le d$, such that $f_i(n_1, \ldots, n_{i-1}, n_i - 1, n_{i+1}, \ldots, n_d) = 0$. A d-tuple $(n_1, \ldots, n_d) \in \mathbb{Z}^d$ is an* integer singularity *of A if it is a leading or a trailing integer singularity of A.*

Let $S(A)$ denote the set of all integer singularities of A. Denote by $V_1(A)$ the $\mathbb{C}$-linear space of all solutions of A which are defined at all elements of $\mathbb{Z}^d$, and by $V_2(A)$ the $\mathbb{C}$-linear space of all solutions of A which are defined at all elements of $\mathbb{Z}^d \setminus S(A)$.

[1] The prefix "H" refers to Jakob Horn and to the adjective "hypergeometric" as well.

We consider only integer singularities here, therefore we will drop the adjective "integer" in the sequel. Sometimes we will also drop the name of the H-system, and will write V_1, V_2 instead of $V_1(A)$, $V_2(A)$.

Definition 5 *Call the two d-tuples $(n_1, \ldots, n_d), (n'_1, \ldots, n'_d) \in \mathbb{Z}^d$* adjacent *if $\sum_{i=1}^{d} |n_i - n'_i| = 1$. Call a finite sequence $t_1, \ldots, t_k \in \mathbb{Z}^d$* a path *from t_1 to t_k if t_i is adjacent to t_{i+1} for all $i = 1, \ldots, k-1$. Given an H-system A, we define* components *induced by A on $\mathbb{Z}^d$ as the equivalence classes of the following equivalence relation $\sim$ in $\mathbb{Z}^d$: $t' \sim t''$ iff there exists a path from t' to t'' which contains no singularity of A. If T is a solution of an H-system A, then its* constituent *is the sequence that is the restriction of T on a component induced by A.*

Definition 6 *Rational functions $F_1, \ldots, F_d \in \mathbb{C}(n_1, \ldots, n_d)$ are* compatible *if*

$$(E_{n_j} F_i) F_j = F_i (E_{n_i} F_j)$$

for all $1 \le i \le j \le d$.

Note that a single rational function (corresponding to the case $d = 1$) is always compatible.

Proposition 1 *Let A be an H-system of the form (2) where $g_1/f_1, \ldots, g_d/f_d$ are compatible rational functions. Then $\dim V_2$ is equal to the number of components induced by A.*

Proof: To each component C_i induced by A on $\mathbb{Z}^d$ we assign a solution T_i of (2) which is 1 at a selected point $p_i \in C_i$, and 0 on all the remaining components. The values of T_i on the remaining points of C_i are uniquely determined by (2). It is clear that the set of all T_i is a basis for V_2. □

3 Dimensions of solution spaces: The univariate case

When $d = 1$ the system (2) is of the form

$$f(n)T(n+1) = g(n)T(n) \tag{3}$$

where $f(n), g(n) \in \mathbb{C}[n] \setminus \{0\}$ and $f(n) \perp g(n)$.

Example 1 ($\dim V_1 = 1$, $\dim V_2 = k$) Consider the recurrence

$$T(n+1) = p_k(n)\ T(n) \tag{4}$$

where $k \ge 1$ and $p_k(n) = \prod_{i=0}^{k-2} (n - 2i + 1)$. Here we use the convention that a product is 1 if its lower limit exceeds its upper limit. Clearly the set of singularities of (4) is $\{2i-1;\ i = 0, 1, \ldots, k-2\}$, so $\dim V_2 = k$. To compute $\dim V_1$, note that any solution $T(n)$ of (4) defined for all $n \in \mathbb{Z}$ is a constant multiple of

$$F_k(n) = \begin{cases} (-1)^{(k-1)n} / \prod_{i=0}^{k-2} (2i - n - 1)!, & n < 0, \\ 0, & n \ge 0. \end{cases}$$

Therefore $\dim V_1 = 1$. ■

Example 2 ($\dim V_1 = m$, $\dim V_2 = m+1$) Now consider the recurrence

$$q_m(n+1)\ T(n+1) = q_m(n)\ T(n) \tag{5}$$

where $m \geq 1$ and $q_m(n) = \prod_{i=1}^{m}(n+2i+1)$. The set of singularities is $\{-(2i+1);\ i = 1, 2, \ldots, m\}$, so $\dim V_2 = m+1$. Let $T(n)$ be a solution of (5) defined for all $n \in \mathbb{Z}$. By substituting $n = -2(i+1)$ for $i = 1, 2, \ldots, m$ into (5), we see that $T(n) = 0$ for these values of n. Likewise, by substituting $n = -3$ into (5), we find that $T(-2) = 0$. Using (5) it follows by induction on n that $T(n) = 0$ for all $n \leq -2(m+1)$ and for all $n \geq -2$ as well. On the other hand, it is easy to check that

$$G_m^{(i)}(n) = \delta_{n,-(2i+1)}$$

(where δ is the Kronecker delta) is a solution of (5) for $i = 1, 2, \ldots, m$. Therefore $\dim V_1 = m$. ■

Before describing the general situation we need a definition and a lemma.

Definition 7 *Let A be an H-system of the form (3). An interval of integers*

$$I = \{k, k+1, \ldots, k+m\}, \quad m \geq 0, \tag{6}$$

is a segment of singularities *of A if $I \subseteq S(A)$ while $k-1, k+m+1 \notin S(A)$.*

Lemma 1 *Each segment of singularities (6) of equation (3) is of (at least) one of the following types:*

(i) all elements of the segment are trailing singularities;

(ii) all elements of the segment are leading singularities;

(iii) there exists j, $0 \leq j < m$, such that $k, k+1, \ldots, k+j$ are leading singularities, while $k+j+1, \ldots, k+m$ are trailing singularities.

Proof: If $u \in \mathbb{Z}$ is a trailing singularity and $u+1$ a leading singularity of (3) then $f(u) = g(u) = 0$, contrary to the assumption $f \perp g$. So any segment of singularities of (3) consists of a (possiby empty) interval of leading singularities followed by a (possiby empty) interval of trailing singularities. □

Theorem 3 *Let S denote the set of singularities of equation (3).*
a) If $S = \emptyset$ then $\dim V_1 = \dim V_2 = 1$.
b) If $S \neq \emptyset$ then $1 \leq \dim V_1 < \dim V_2 < \infty$.

Proof: a) This is clear.

b) There is only a finite set of components induced on $\mathbb{Z}$ by (3), therefore $\dim V_2 < \infty$.

Next we prove that $\dim V_1 < \dim V_2$. First we show that if (6) is a segment of singularities of (3), then the restriction of V_1 to

$$\hat{I} = \{k-1, k, , \ldots, k+m, k+m+1\}$$

has dimension ≤ 1, while the analogous restriction of V_2 obviously has dimension 2. Indeed, if u is a trailing singularity, then any sequence from V_1 vanishes at $u+1$; and if u is a leading singularity, then any sequence from V_1 vanishes at $u-1$. By Lemma 1 we have three possibilities (i), (ii), (iii) for (6). In case (i) we have $T(k+1) = T(k+2) = \cdots = T(k+m+1) = 0$, in case (ii) $T(k-1) = T(k) = \cdots = T(k+m-1) = 0$, in case (iii) $T(k-1) = T(k) = \ldots T(k+j-1) = 0$ and $T(k+j+2) = T(k+j+3) = \cdots = T(k+m+1) = 0$; in each case $T(n)$ can be nonzero at most in two points of $\hat{I}$, however the value at one of them is uniquely determined by the value at the other one. Therefore the dimension of the restricted V_1 is ≤ 1. The same holds for dimension of the restriction of V_1 to the set

$$\{k-v, k-v+1, \ldots, k, k+1, \ldots, k+m, k+m+1, \ldots, k+w\},$$

where $k, k+1, \ldots, k+m$ are singularities, while $k-v, \ldots, k-1$ and $k+m+1, \ldots, k+w$ are not. Gluing together two such restrictions with coinciding, say, $k+m+1, \ldots, k+w$, and non-intersecting singular parts, we get the dimension ≤ 2, while the dimension of the corresponding restriction of V_2 is 3 and so on. This proves that $\dim V_1 < \dim V_2$.

Finally we prove that $\dim V_1 \geq 1$. If there are leading singularities, let n_0 be the largest leading singularity. Set $T(n_0) = 1$ and $T(n) = 0$ for $n < n_0$. None of the points $n > n_0$ is a leading singularity, hence the value of T at $n > n_0$ is uniquely determined by the recurrence (3) and the initial condition $T(n_0) = 1$. If there are no leading singularities, let n_0 be the least trailing singularity. Set $T(n_0) = 1$ and $T(n) = 0$ for $n > n_0$. None of the points $n < n_0$ is a trailing singularity, hence the value of T at $n < n_0$ is uniquely determined by the recurrence (3) and the initial condition $T(n_0) = 1$. In either case V_1 contains a nonzero solution. □

Theorem 4 *For any integers s, t such that $1 \leq s < t$ there exists an equation of the form (3) such that $\dim V_1 = s$ and $\dim V_2 = t$.*

Proof: Consider the recurrence

$$q_m(n+1)\ T(n+1) = p_k(n) q_m(n)\ T(n) \tag{7}$$

where $k, m \geq 1$, $p_k(n)$ is as in Example 1, and $q_m(n)$ is as in Example 2. Here the set of singularities is $\{2i-1;\ i = 0, 1, \ldots, k-2\} \cup \{-(2i+1);\ i = 1, 2, \ldots, m\}$, so $\dim V_2 = k+m$. Let $T(n)$ be a solution of (7) defined for all $n \in \mathbb{Z}$. In exactly the same way as in Example 2 we can see that $T(n) = 0$ for $n = -2, -4, \ldots, -2(m+1)$, $n \leq -2(m+1)$ or $n \geq -2$, and that $G_m^{(i)}(n) = \delta_{n,-(2i+1)}$ is a solution of (7) for $i = 1, 2, \ldots, m$. Therefore $\dim V_1 = m$.

If $1 \leq s < t$, let $m = s$ and $k = t - s$. Then for equation (7), $\dim V_1 = m = s$ and $\dim V_2 = k + m = t$. □

We conclude this section by some remarks on computation of $\dim V_1$ and $\dim V_2$. Let A denote equation (3). According to Proposition 1, $\dim V_2(A)$ is the number of components induced on $\mathbb{Z}$ by A and is thus easy to compute. We claim that $\dim V_1(A)$ equals the dimension of the kernel of a bidiagonal matrix B defined as follows. Let α be the maximum and β the minimum of the integer roots of $f(x)g(x)$; if A has no integer singularities then we can take $\alpha = \beta = 1$. Let B be the $(\alpha - \beta + 1) \times (\alpha - \beta + 2)$ matrix with entries

$$b_{i,j} = \begin{cases} f(\alpha - i + 1), & j = i, \\ -g(\alpha - i + 1), & j = i + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $1 \leq i \leq \alpha - \beta + 1$ and $1 \leq j \leq \alpha - \beta + 2$. Indeed, any vector v such that $Bv = 0$ can be extended to a solution of A in a unique way. This mapping is an isomorphism between the kernel of B and $V_1(A)$.

Incidentally, this gives an alternative proof of the inequality $\dim V_1 \geq 1$: B has more columns than rows, hence its kernel is nontrivial.

4 Dimensions of solution spaces: The multivariate case

If $d \geq 2$ in (2) then the dimensions of V_1 and/or V_2 can be infinite as shown by the following examples.

Example 3 ($\dim V_1 = \infty$, $\dim V_2 = 1$) Let A be the system

$$\begin{aligned} (n_1 - 4n_2 + 1)T(n_1 + 1, n_2) &= (n_1 - 4n_2)T(n_1, n_2), \\ (n_1 - 4n_2 - 4)T(n_1, n_2 + 1) &= (n_1 - 4n_2)T(n_1, n_2). \end{aligned}$$

It is easy to check that

$$T_i(n_1, n_2) = \delta_{n_1,4i}\delta_{n_2,i}, \qquad \text{for } i \in \mathbb{Z},$$

are linearly independent solutions of A on all of $\mathbb{Z}^2$, hence $\dim V_1 = \infty$. On the other hand, $S(A) = \{(n_1, n_2);\ n_1 = 4n_2\}$, so A induces a single component on $\mathbb{Z}^2$, and $\dim V_2 = 1$. ■

Example 4 ($\dim V_1 = 1$, $\dim V_2 = \infty$) Let B be the system

$$\begin{aligned} (n_1 - 4n_2)T(n_1 + 1, n_2) &= (n_1 - 4n_2 + 1)T(n_1, n_2), \\ (n_1 - 4n_2)T(n_1, n_2 + 1) &= (n_1 - 4n_2 - 4)T(n_1, n_2). \end{aligned}$$

It can be shown that any solution of B defined on all $\mathbb{Z}^2$ is a constant multiple of $n_1 - 4n_2$, so $\dim V_1 = 1$. On the other hand, $S(B) = \{(n_1, n_2);\ n_1 - 4n_2 \in \{-4, -1, 1, 4\}\}$, so each of the points $(4i, i)$ for $i \in \mathbb{Z}$ is a separate component of $\mathbb{Z}^2$ induced by B, hence $\dim V_2 = \infty$. ■

Example 5 ($\dim V_1 = \dim V_2 = \infty$) Let C be the system

$$\begin{aligned} (n_1 - n_2 - 1)(n_1 - n_2 + 1)T(n_1 + 1, n_2) &= (n_1 - n_2)(n_1 - n_2 + 2)T(n_1, n_2), \\ (n_1 - n_2 - 1)(n_1 - n_2 + 1)T(n_1, n_2 + 1) &= (n_1 - n_2)(n_1 - n_2 - 2)T(n_1, n_2). \end{aligned}$$

It is easy to check that

$$T_i(n_1, n_2) = \delta_{n_1,i}\delta_{n_2,i}, \qquad \text{for } i \in \mathbb{Z}, \tag{8}$$

are linearly independent solutions of C on all of $\mathbb{Z}^2$, hence $\dim V_1 = \infty$. As $S(C) = \{(n_1, n_2);\ n_1 - n_2 \in \{-2, 0, 2\}\}$, each of the points $(i, i - 1)$ and $(i, i + 1)$ for $i \in \mathbb{Z}$ is a separate component of $\mathbb{Z}^2$ induced by C, so $\dim V_2 = \infty$ as well. ■

The following theorem describes the general situation.

Theorem 5 *Let $1 \le s, t \le \infty$. Then there exists an H-system such that* $\dim V_1 = s$ *and* $\dim V_2 = t$.

Proof: Let $t \ge 2$ and $p_t(n_1, n_2) = \prod_{i=0}^{t-2}(n_1 - n_2 + 3i)$. Then the set of singularities of

$$\begin{aligned} p_t(n_1+1, n_2)T(n_1+1, n_2) &= p_t(n_1, n_2)T(n_1, n_2), \\ p_t(n_1, n_2+1)T(n_1, n_2+1) &= p_t(n_1, n_2)T(n_1, n_2) \end{aligned}$$

is $S = \{(n_1, n_2);\ n_1 - n_2 \in \{-3i;\ 0 \le i \le t-2\}\}$. As in Example 5, the functions (8) are linearly independent solutions of this system on all of $\mathbb{Z}^2$, hence $\dim V_1 = \infty$. On the other hand, the number of components induced on $\mathbb{Z}^2$ is t, so $\dim V_2 = t$.

Let $s \ge 2$ and

$$q_s(n_1, n_2) = \prod_{i=1}^{s-1}((n_1 - 2i)^2 + n_2^2). \tag{9}$$

Then the set of singularities of

$$\begin{aligned} (n_1 - 4n_2)q_{s+1}(n_1+1, n_2)T(n_1+1, n_2) &= (n_1 - 4n_2 + 1)q_{s+1}(n_1, n_2)T(n_1, n_2), \\ (n_1 - 4n_2)q_{s+1}(n_1, n_2+1)T(n_1, n_2+1) &= (n_1 - 4n_2 - 4)q_{s+1}(n_1, n_2)T(n_1, n_2) \end{aligned}$$

is $S = \{(n_1, n_2);\ n_1 - 4n_2 \in \{-4, -1, 1, 4\}\} \cup \{(2i, 0);\ 1 \le i \le s\}$. Each of the points $(4i, i)$ for $i \in \mathbb{Z}$ is a separate component, so $\dim V_2 = \infty$. It can be shown that any solution $T(n_1, n_2)$ defined on all $\mathbb{Z}^2$ vanishes everywhere except at the points $(2i, 0)$ where $1 \le i \le s$, and that

$$T_i(n_1, n_2) = \delta_{n_1, 2i}\delta_{n_2, 0}, \tag{10}$$

for $i = 1, 2, \ldots, s$, are linearly independent solutions of this system defined on all $\mathbb{Z}^2$. Hence $\dim V_1 = \infty$.

Together with Examples 3 – 5 this proves the assertion in the case when at least one of s, t is infinite.

Now assume that s, t are natural numbers, and let $r_t(n_1, n_2) = \prod_{i=1}^{t-1}(n_1 + 2i + 1)$. Consider the system

$$\begin{aligned} q_s(n_1+1, n_2)T(n_1+1, n_2) &= q_s(n_1, n_2)r_t(n_1, n_2)T(n_1, n_2), \\ q_s(n_1, n_2+1)T(n_1, n_2+1) &= q_s(n_1, n_2)T(n_1, n_2), \end{aligned}$$

where q_s is as in (9). It can be shown that any solution $T(n_1, n_2)$ defined on all $\mathbb{Z}^2$ vanishes for all (n_1, n_2) such that $n_1 > -(2t-1)$ and (n_1, n_2) is not of the form $(2i, 0)$ with $1 \le i \le s-1$. Further, a basis of V_1 is given by the s functions $T_i(n_1, n_2)$ for $i = 0, 1, \ldots, s-1$ where

$$T_0(n_1, n_2) = \begin{cases} \frac{(-1)^{(t-1)n_1}}{\prod_{i=1}^{s-1}((n_1-2i)^2+n_2^2)\prod_{i=1}^{t-1}(-n_1-2i-1)!}, & n_1 \le -(2t-1), \\ 0, & \text{otherwise}, \end{cases}$$

and $T_i(n_1, n_2)$ are as in (10) for $i = 1, 2, \ldots, s-1$. It follows that $\dim V_1 = s$. The set of singularities of this system is $S = \{(2i, 0);\ 1 \le i \le s-1\} \cup \{(-(2i+1), j);\ 1 \le i \le t-1, j \in \mathbb{Z}\}$, and the number of components induced on $\mathbb{Z}^2$ is t, so $\dim V_2 = t$ as desired. $\square$

We considered the case $d = 2$ here. The corresponding H-systems for the case of an orbitrary $d > 1$ can be obtained by adding equations $E_{n_i}T = T$, $i = 3, \ldots, d$, to the systems with $d = 2$.

5 The Ore-Sato theorem and its consequences

The well-known Ore-Sato theorem (see [5], [7]) is commonly believed to imply that any solution of an H-system (2) is of the form (1). We show that this is not so, and give an appropriate corollary of the Ore-Sato theorem on possible forms of solutions of H-systems in our setting.

Definition 8 *Let T be a solution of (2). We write* $\operatorname{supp} T$ *for the* support *of T, i.e., for the set of points in $\mathbb{Z}^d$ where T is defined and does not vanish.*

If (2) has a solution with non-algebraic support, then the rational functions f_i/g_i, $i = 1, \ldots, d$, are compatible, and uniquely determined by this solution (see [2]).

Definition 9 *A polynomial $p \in \mathbb{C}[x_1, \ldots, x_d]$ is* integer-linear *if $p(x_1, \ldots, x_d) = a_0 + a_1x_1 + \cdots + a_dx_d$ where $a_1, \ldots, a_d \in \mathbb{Z}$.*

The Ore-Sato theorem states (in the case $d = 2$) that for any compatible rational functions $F_1(x, y)$ and $F_2(x, y)$ there are compatible rational functions $G_1(x, y)$ and $G_2(x, y)$ which factor into integer-linear factors, and a rational function $R(x, y)$ such that $F_1(x, y) = G_1(x, y)R(x + 1, y)/R(x, y)$ and $F_2(x, y) = G_2(x, y)R(x, y + 1)/R(x, y)$. The full statement gives a precise description of the integer-linear factors.

In the literature one often encounters the claim that as a corollary of this theorem, any solution of an H-system (2) is of the form (1). For example, in [3, p. 223] one can read: "*From Ore's result it can be deduced that the most general form of A_{mn} is of the form*

$$A_{mn} = R(m, n)\gamma_{mn}a^m b^n$$

where R is a fixed rational function of m and n, a and b are constants, and γ_{mn} is a gamma product (...) that is to say it is of the form

$$\gamma_{mn} = \prod_i \Gamma(a_i + u_i m + v_i n)/\Gamma(a_i)$$

where the a_i are arbitrary (real or complex) constants, and the u_i and v_i are arbitrary integers which may be positive, negative, or zero." A similar quote can be found in [4, p. 5].

It may be the case that in the literature referred to above the term A_{mn} is implicitly assumed to be nonzero for all $m, n \in \mathbb{Z}$. This possibility is supported by the fact that, e.g., in [3] the corresponding H-system is given in terms of the two quotients $A_{m+1,n}/A_{mn}$ and $A_{m,n+1}/A_{mn}$. But such a severe restriction would preclude many important functions from being hypergeometric, such as the binomial coefficient $T(n_1, n_2) = \binom{n_1}{n_2}$, and all polynomials with integer roots.

However if we do not adopt this restriction, then there are hypergeometric terms which cannot be written in the form (1), as illustrated by the following examples.

Example 6 Take the H-system

$$\begin{aligned} p(n_1,n_2)T(n_1+1,n_2) &= p(n_1+1,n_2)T(n_1,n_2), \\ p(n_1,n_2)T(n_1,n_2+1) &= p(n_1,n_2+1)T(n_1,n_2), \end{aligned} \tag{11}$$

where $p(n_1,n_2) = (n_1-n_2-1)(n_1-n_2+1)$. It can be checked that any sequence T which satisfies $T(n_1,n_2) = 0$ unless $n_1 = n_2$ is a solution of (11). In particular, the sequence

$$T(n_1,n_2) = \begin{cases} 2^{n_1^2}, & n_1 = n_2, \\ 0, & \text{otherwise}, \end{cases}$$

is a solution of (11), even though it does not have the form (1) because it grows too fast along the diagonal. ■

There are examples which look less artificial and where the solution has a non-algebraic support.

Example 7 In this example we show that $|n_1-n_2|$, although a hypergeometric term, cannot be written in the form (1).

Denote $D(n_1,n_2) = |n_1-n_2|$. Then

$$\begin{aligned} (n_1-n_2)D(n_1+1,n_2) &= (n_1-n_2+1)D(n_1,n_2), \\ (n_1-n_2)D(n_1,n_2+1) &= (n_1-n_2-1)D(n_1,n_2) \end{aligned} \tag{12}$$

for all $n_1,n_2 \in \mathbb{Z}$, so $D(n_1,n_2)$ is a hypergeometric term. Note that when restricted to $n_1,n_2 \geq 0$, it is also holonomic, because its generating function is rational:

$$\sum_{n_1,n_2\geq 0} |n_1-n_2| z_1^{n_1} z_2^{n_2} = \left(\frac{z_1}{(1-z_1)^2} + \frac{z_2}{(1-z_2)^2}\right)\frac{1}{1-z_1z_2}. \tag{13}$$

To compare, the generating function of the polynomial n_1-n_2 is

$$\sum_{n_1,n_2\geq 0} (n_1-n_2) z_1^{n_1} z_2^{n_2} = \left(\frac{z_1}{1-z_1} - \frac{z_2}{1-z_2}\right)\frac{1}{(1-z_1)(1-z_2)}.$$

Let $T(n_1,n_2)$ be a hypergeometric term of the form (1) with $d=2$, defined for all $n_1,n_2 \geq 0$. Pick $k_0 \in \mathbb{Z}$, $k_0 > 0$, and assume that $T(n,k_0) = |n-k_0|$ for all $n > k_0$. *Then we claim that $T(n,k_0) = n-k_0$ for all $n \geq 0$.* Hence $T(n,k_0)$ disagrees with $|n-k_0|$ for all n such that $0 \leq n < k_0$.

To prove the claim, define

$$t(n) := T(n,k_0) = R(n,k_0)\, u_1^n u_2^{k_0} \frac{\prod_{i=1}^p \Gamma(a_{i,1}n + a_{i,2}k_0 + \alpha_i)}{\prod_{j=1}^q \Gamma(b_{j,1}n + b_{j,2}k_0 + \beta_j)}, \quad \text{for all } n \geq 0. \tag{14}$$

We wish to rewrite the right-hand side of (14) in such a way that the coefficient of n in the arguments of the Gamma function will be 1. It is straightforward to verify that for $n \in \mathbb{Z}$, $a \in \mathbb{Z}\setminus\{0\}$ and $z \in \mathbb{C}$ such that $an+z$ is not a nonpositive integer,

$$\Gamma(an+z) = \begin{cases} C(a,z)\, a^{an} \prod_{m=0}^{a-1} \Gamma(n+(z+m)/a), & a > 0, \\ C(a,z)\, a^{an} / \prod_{m=1}^{|a|} \Gamma(n+(z-m)/a), & a < 0,\ z \notin \mathbb{Z}, \end{cases}$$

where $C(a,z) \in \mathbb{C}$ is independent of n. To be able to apply this to (14), we need to show that

(i) for $i = 1,\dots,p$ and for all $n \geq 0$, the number $a_{i,1}n + a_{i,2}k_0 + \alpha_i$ is not a nonpositive integer,

(ii) for $i = 1,\dots,p$ and for all $n \geq 0$, if $a_{i,1} < 0$ then $\alpha_i \notin \mathbb{Z}$,

(iii) for $j = 1,\dots,q$ and for all $n \geq 0$, the number $b_{j,1}n + b_{j,2}k_0 + \beta_j$ is not a nonpositive integer,

(iv) for $j = 1,\dots,q$ and for all $n \geq 0$, if $b_{j,1} < 0$ then $\beta_j \notin \mathbb{Z}$.

Assertion (i) is obvious, for otherwise $T(n,k_0)$ would be undefined at such n. Assertion (ii) holds for the same reason, since if $a_{i,1} < 0$ and $\alpha_i \in \mathbb{Z}$ for some i, then $a_{i,1}n + a_{i,2}k_0 + \alpha_i$ would be a nonpositive integer for all large enough n. If $b_{j,1}n + b_{j,2}k_0 + \beta_j$ is a nonpositive integer for some j and some $n \neq k_0$, then t vanishes at this n, contrary to the fact that $t(n) = |n - k_0| \neq 0$ for all $n \neq k_0$. If $b_{j,1}k_0 + b_{j,2}k_0 + \beta_j$ is a nonpositive integer then (depending on the sign of $b_{j,1}$) at least one of $b_{j,1}(k_0 - 1) + b_{j,2}k_0 + \beta_j$ and $b_{j,1}(k_0 + 1) + b_{j,2}k_0 + \beta_j$ is also a nonpositive integer, something we have just ruled out. This proves (iii). Assertion (iv) holds for the same reason, since if $b_{j,1} < 0$ and $\beta_j \in \mathbb{Z}$ for some j, then $b_{j,1}n + b_{j,2}k_0 + \beta_j$ would be a nonpositive integer for all large enough n.

Therefore the univariate term t can be written in the form

$$t(n) = r(n)\, v^n \frac{\prod_{i=1}^{p'} \Gamma(n + \gamma_i)}{\prod_{j=1}^{q'} \Gamma(n + \delta_j)}, \quad \text{for all } n \geq 0, \tag{15}$$

where $r \in \mathbb{C}(x)$ and $v, \gamma_i, \delta_j \in \mathbb{C}$. If $\gamma_i - \delta_j \in \mathbb{Z}$ then $\Gamma(n+\gamma_i)/\Gamma(n+\delta_j)$ is a rational function of n, hence (15) can be rewritten as

$$t(n) = s(n)\, v^n \frac{\prod_{i=1}^{p''} \Gamma(n + \varepsilon_i)}{\prod_{j=1}^{q''} \Gamma(n + \zeta_j)}, \quad \text{for all } n \geq 0, \tag{16}$$

where $s \in \mathbb{C}(x)$, $\varepsilon_i, \zeta_j \in \mathbb{C}$, and none of the differences $\varepsilon_i - \zeta_j$ is an integer. It follows that

$$g(x) \;:=\; v \frac{\prod_{i=1}^{p''} (x + \varepsilon_i)}{\prod_{j=1}^{q''} (x + \zeta_j)} \;\in\; \mathbb{C}(x)$$

is a shift-reduced rational function (see Definition 1). Let $f(x) := (x+1-k_0)/(x-k_0) \in \mathbb{C}(x)$. For $n > k_0$ we have

$$\frac{t(n+1)}{t(n)} \;=\; \frac{|n+1-k_0|}{|n-k_0|} \;=\; \frac{n+1-k_0}{n-k_0} \;=\; f(n)$$

and

$$\frac{t(n+1)}{t(n)} \;=\; g(n)\, \frac{s(n+1)}{s(n)}.$$

The two rational functions $f(x)$ and $g(x)\,s(x+1)/s(x)$ agree infinitely often, so they are equal. Since g is shift-reduced, both $(1, x-k_0)$ and $(g(x), s(x))$ are rational normal forms of f. Now Theorem 2 implies that $g(x) = 1$. Comparing this with the definition of $g(x)$, we see that $v = 1$ and $p'' = q'' = 0$. From (16) it follows that $s(n) = t(n)$ for all $n \geq 0$, therefore $s(n) = n - k_0$ for all $n > k_0$. As the two rational functions $s(x)$ and $x - k_0$ agree infinitely often, they are equal. But then $t(n) = n - k_0$ for all $n \geq 0$, which proves our claim.

We could have chosen a specific value for k_0 (such as $k_0 = 1$, say), but by doing so our result would be slightly weaker. In geometric language, we have shown that as soon as a hypergeometric term $T(n_1, n_2)$ agrees with $|n_1 - n_2|$ on *any* horizontal (or, by symmetry, vertical) line which contains integer points on both sides of the line $n_1 = n_2$, then it cannot have the form (1). It seems that, under some additional conditions, this could be generalized from $|n_1 - n_2|$ to $|R(n_1, n_2)|$ where $R \in \mathbb{C}(x_1, x_2)$, and to horizontal (or vertical) lines containing integer points (p_1, p_2) and (n_1, n_2) with $R(p_1, p_2) > 0$ and $R(n_1, n_2) < 0$.

In the theory of multivariate hypergeometric series, H-systems are used to specify coefficients for such series. The simple rational function on the right-hand side of (13) has series expansion whose coefficients satisfy the H-system (12), however are not of the form (1). ∎

The following statement is a corollary of the Ore-Sato theorem.

Corollary 1 *Any constituent (see Definition 5) of a solution with non-algebraic support of an H-system (2) is of the form (1).*

References

[1] S. A. Abramov and M. Petkovšek, On the structure of multivariate hypergeometric terms, *Adv. Appl. Math.* **29** (2002) 386–411.

[2] S. A. Abramov and M. Petkovšek, Rational normal forms and minimal decompositions of hypergeometric terms, *J. Symb. Comput.* **33** (2002) 521–543.

[3] H. Bateman, A. Erdelyi, *Higher Transcendental Functions, Vol. 1*, McGraw-Hill, New York–Toronto–London 1953.

[4] I. M. Gel'fand, M. I. Graev and V. S. Retakh, General hypergeometric systems of equations and series of hypergeometric type (Russian), *Uspekhi Mat. Nauk* **47** (1992) 3–82, 235; translation in *Russian Math. Surveys* **47** (1992) 1–88.

[5] O. Ore, Sur la forme des fonctions hypergéométriques de plusieurs variables, *J. Math. Pures Appl.* **9** (1930) 311–326.

[6] M. Petkovšek, H. S. Wilf, D. Zeilberger, *A = B*, A K Peters, Wellesley, Massachusetts, 1996.

[7] M. Sato, T. Shintani and M. Muro, Theory of prehomogeneous vector spaces (algebraic part), *Nagoya Math. J.* **120** (1990) 1–34.

THE HOPF ALGEBRA OF UNIFORM BLOCK PERMUTATIONS. EXTENDED ABSTRACT

MARCELO AGUIAR AND ROSA C. ORELLANA

Abstract. We introduce the Hopf algebra of uniform block permutations and show that it is self-dual, free, and cofree. These results are closely related to the fact that uniform block permutations form a factorizable inverse monoid. This Hopf algebra contains the Hopf algebra of permutations of Malvenuto and Reutenauer and the Hopf algebra of symmetric functions in non-commuting variables of Gebhard, Rosas, and Sagan.

Résumé. Nous présentons l'algèbre de Hopf des permutations de blocs uniformes est démontrons qu'elle est auto duale, libre et colibre. Ces résultats sont liés au fait que les permutations de blocs uniformes constituent un monoïde inverse factorisable. Cette algèbre de Hopf contient l'algèbre de Hopf des permutations de Malvenuto et Reutenauer et l'algèbre de Hopf des fonctions symétriques à variables non commutatives de Gebhard, Rosas, et Sagan.

1. Uniform block permutations

1.1. **Set partitions.** Let n be a non-negative integer and let $[n] := \{1, 2, \ldots, n\}$. A *set partition* of $[n]$ is a collection of non-empty disjoint subsets of $[n]$, called *blocks*, whose union is $[n]$. For example, $\mathcal{A} = \big\{\{2,5,7\}\{1,3\}\{6,8\}\{4\}\big\}$, is a set partition of [8] with 4 blocks. We often specify a set partition by listing the blocks from left to right so that the sequence formed by the minima of the blocks is increasing, and by listing the elements within each block in increasing order. For instance, the set partition above will be denoted $\mathcal{A} = \{1,3\}\{2,5,7\}\{4\}\{6,8\}$. We use $\mathcal{A} \vdash [n]$ to indicate that $\mathcal{A}$ is a set partition of $[n]$.

The *type* of a set partition $\mathcal{A}$ of $[n]$ is the partition of n formed by the sizes of the blocks of $\mathcal{A}$. The symmetric group S_n acts on the set of set partitions of $[n]$: given $\sigma \in S_n$ and $\mathcal{A} \vdash [n]$, $\sigma(\mathcal{A})$ is the set partition whose blocks are $\sigma(A)$ for $A \in \mathcal{A}$. The orbit of $\mathcal{A}$ consists of those set partitions of the same type as $\mathcal{A}$. The stabilizer of $\mathcal{A}$ consists of those permutations that preserve the blocks, or that permute blocks of the same size. Therefore, the number of set partitions of type $1^{m_1}2^{m_2}\ldots n^{m_n}$ (m_i blocks of size i) is

$$\frac{n!}{m_1!\cdots m_n!(1!)^{m_1}\cdots(n!)^{m_n}}\,. \tag{1}$$

1.2. **The monoid of uniform block permutations.** The monoid (and the monoid algebra) of uniform block permutations has been studied by FitzGerald [9] and Kosuda [13, 14] in analogy to the partition algebra of Jones and Martin [12, 17].

A *block permutation* of $[n]$ consists of two set partitions $\mathcal{A}$ and $\mathcal{B}$ of $[n]$ with the same number of blocks and a bijection $f : \mathcal{A} \to \mathcal{B}$. For example, if $n = 3$, $f(\{1,3\}) = \{3\}$ and $f(\{2\}) = \{1,2\}$ then f is a block permutation. A block permutation is called *uniform* if it maps each block of $\mathcal{A}$ to a block of $\mathcal{B}$ of the same cardinality. For example, $f(\{1,3\}) = \{1,2\}$, $f(\{2\}) = \{3\}$ is uniform. Each permutation may be viewed as

Date: November 17, 2004.
2000 *Mathematics Subject Classification.* Primary: 05E99, 16W30; Secondary: 16G99, 20C30.
Key words and phrases. Hopf algebra, uniform block permutation, set partition, symmetric functions, Schur-Weyl duality.
Aguiar supported in part by NSF grant DMS-0302423.
Orellana supported in part by the Wilson Foundation.

a uniform block permutation for which all blocks have cardinality 1. In this paper we only consider block permutations that are uniform.

To specify a uniform block permutation $f : \mathcal{A} \to \mathcal{B}$ we must choose two set partitions $\mathcal{A}$ and $\mathcal{B}$ of the same type $1^{m_1} \dots n^{m_n}$ and for each i a bijection between the m_i blocks of size i of $\mathcal{A}$ and those of $\mathcal{B}$. We deduce from (1) that the total number of uniform block permutations of $[n]$ is

$$u_n := \sum_{1^{m_1}\dots n^{m_n} \vdash n} \left(\frac{n!}{(1!)^{m_1} \cdots (n!)^{m_n}} \right)^2 \frac{1}{m_1! \cdots m_n!} \tag{2}$$

where the sum runs over all partitions of n. Starting at $n = 0$, the first values are

$$1, 1, 3, 16, 131, 1496, 22482, \dots$$

This is sequence A023998 in [20]. These numbers and generalizations are studied in [19]; in particular, the following recursion is given in [19, equation (11)]:

$$u_{n+1} = \sum_{k=0}^{n} \binom{n}{k} \binom{n+1}{k} u_k \,, \quad u_0 = 1 \,.$$

We represent uniform block permutations by means of graphs. For instance, either one of the two graphs in Figure 1 represents the uniform block permutation f given by

$$\{1,3,4\} \to \{3,5,6\},\ \{2\} \to \{4\},\ \{5,7\} \to \{1,2\},\ \{6\} \to \{8\}, \ \text{ and } \{8\} \to \{7\} \,.$$

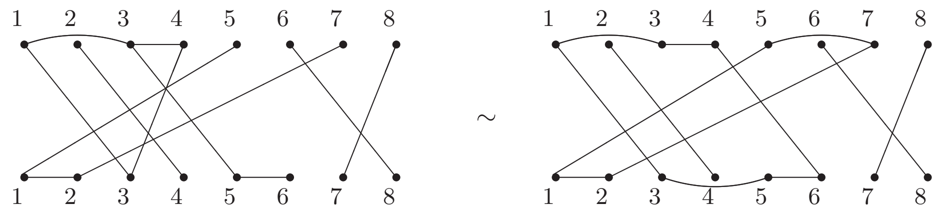


FIGURE 1. Two graphs representing the same uniform block permutation

Different graphs may represent the same uniform block permutation. For a graph to represent a uniform block permutation $f : \mathcal{A} \to \mathcal{B}$ of $[n]$ the vertex set must consist of two copies of $[n]$ (top and bottom) and each connected component must contain the same number of vertices on the top as on the bottom. The set partition $\mathcal{A}$ is read off from the adjacencies on the top, $\mathcal{B}$ from those on the bottom, and f from those in between.

The *diagram* of f is the unique representing graph in which all connected components are cycles and the elements in each cycle are joined in order, as in the second graph of Figure 1.

The set P_n of block permutations of $[n]$ is a monoid. The product $g \cdot f$ of two uniform block permutations f and g of $[n]$ is obtained by gluing the bottom of a graph representing f to the top of a graph representing g. The resulting graph represents a uniform block permutation which does not depend on the graphs chosen. An example is given in Figure 2. Note that gluing the diagram of f to the diagram of g may not result in the diagram of $g \cdot f$.

The identity is the uniform block permutation that maps $\{i\}$ to $\{i\}$ for all i. Viewing permutations as uniform block permutations as above, we get that the symmetric group S_n is a submonoid of P_n.

We recall a presentation of the monoid P_n given in [9, 13, 14]. Consider the uniform block permutations b_i and s_i with diagrams

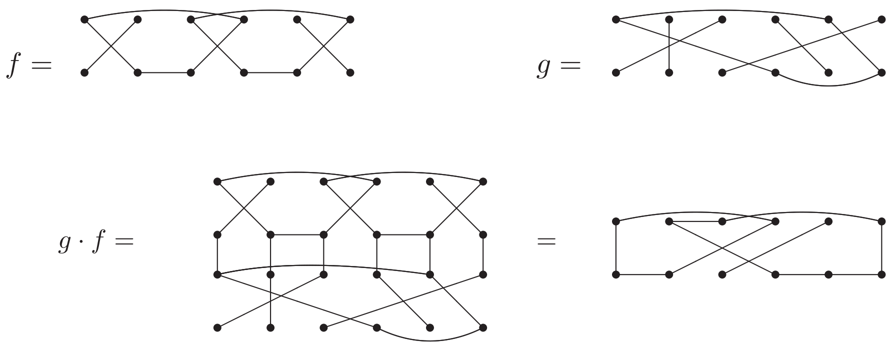
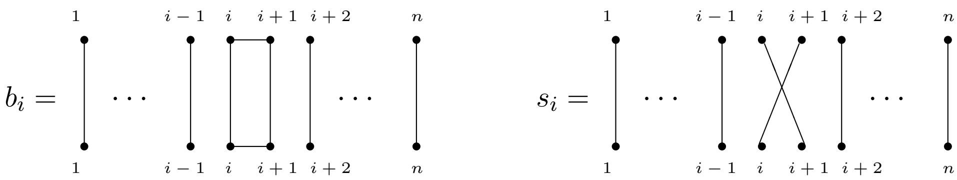


FIGURE 2. Product of uniform block permutations



The monoid P_n is generated by the elements $\{b_i, s_i \,|\, 1 \le i \le n-1\}$ subject to the following relations:

(1) $s_i^2 = 1, \qquad b_i^2 = b_i, \qquad 1 \le i \le n-1;$
(2) $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}, \qquad s_i b_{i+1} s_i = s_{i+1} b_i s_{i+1}, \qquad 1 \le i \le n-2;$
(3) $s_i s_j = s_j s_i, \qquad b_i s_j = s_j b_i, \qquad |i-j| > 1;$
(4) $b_i s_i = s_i b_i = b_i, \qquad 1 \le i \le n-1;$
(5) $b_i b_j = b_j b_i, \qquad 1 \le i, j \le n-1.$

The submonoid generated by the elements s_i, $1 \le i \le n-1$ is the symmetric group S_n, viewed as a submonoid of P_n as above.

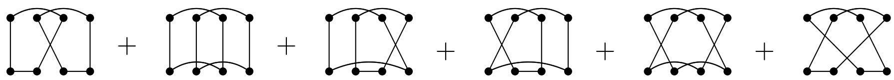
We will see in Sections 2.3 and 2.4 that P_n is a factorizable inverse monoid. Therefore, a presentation for P_n may also be derived from the results of [8].

1.3. **An ideal indexed by set partitions.** Let $\Bbbk P_n$ be the monoid algebra of P_n over a commutative ring $\Bbbk$.

Given a set partition $\mathcal{A} \vdash [n]$, let $Z_{\mathcal{A}} \in \Bbbk P_n$ denote the sum of all uniform block permutations $f : \mathcal{A} \to \mathcal{B}$, where $\mathcal{B}$ varies:

$$Z_{\mathcal{A}} := \sum_{f:\mathcal{A}\to\mathcal{B}} f\,.$$

For instance,

$Z_{\{1,3\}\{2,4\}} =$ 

Lemma 1.1. *Let $\mathcal{A}$ be a set partition of $[n]$ and σ a permutation of $[n]$. Then*

$$\sigma \cdot Z_{\mathcal{A}} = Z_{\mathcal{A}} \quad \textit{and} \quad Z_{\mathcal{A}} \cdot \sigma = Z_{\sigma^{-1}(\mathcal{A})}\,.$$

In addition,

$$Z_{\mathcal{A}} \cdot b_i = \begin{cases} Z_{\mathcal{A}} & \text{if } i \text{ and } i+1 \text{ belong to the same block of } \mathcal{A} \\ \binom{|A|+|A'|}{|A|} Z_{\mathcal{B}} & \text{if } i \text{ and } i+1 \text{ belong to different blocks } A \text{ and } A' \text{ of } \mathcal{A} \end{cases}$$

where the set partition $\mathcal{B}$ is obtained by merging the blocks A and A' of $\mathcal{A}$ and keeping the others.

Let $\mathcal{Z}_n$ denote the subspace of $\Bbbk P_n$ linearly spanned by the elements $Z_{\mathcal{A}}$ as $\mathcal{A}$ runs over all set partitions of $[n]$.

Corollary 1.2. *$\mathcal{Z}_n$ is a right ideal of the monoid algebra $\Bbbk P_n$.*

2. The Hopf algebra of uniform block permutations

In this section we define the Hopf algebra of uniform block permutations. It contains the Hopf algebra of permutations of Malvenuto and Reutenauer as a Hopf subalgebra.

2.1. **Schur-Weyl duality for uniform block permutations.** Let r and m be positive integers. Consider the complex reflection group

$$G(r,1,m) := \mathbb{Z}_r \wr S_m \,.$$

Let t denote the generator of the cyclic group $\mathbb{Z}_r$, $t^r = 1$.

Let V be the *monomial* representation of $G(r,1,m)$. Thus, V is an m-dimensional vector space with a basis $\{e_1, e_2, \dots, e_m\}$ on which $G(r,1,m)$ acts as follows:

$$t \cdot e_1 = e^{2\pi i/r} e_1 \,, \quad t \cdot e_i = e_i \text{ for } i > 1, \text{ and } \ \sigma \cdot e_i = e_{\sigma(i)} \text{ for } \sigma \in S_m.$$

Consider now the diagonal action of $G(r,1,m)$ on the tensor powers $V^{\otimes n}$,

$$g \cdot (e_{i_1} e_{i_2} \cdots e_{i_n}) = (g \cdot e_{i_1})(g \cdot e_{i_2}) \cdots (g \cdot e_{i_n}) \,.$$

The centralizer of this representation has been calculated by Tanabe.

Proposition 2.1. [21] *There is a right action of the monoid P_n on $V^{\otimes n}$ determined by*

$$(e_{i_1} \cdots e_{i_n}) \cdot b_j = \delta(i_j, i_{j+1}) e_{i_1} \cdots e_{i_n} \quad \text{and} \quad (e_{i_1} \cdots e_{i_n}) \cdot \sigma = e_{i_{\sigma(1)}} \cdots e_{i_{\sigma(n)}}$$

for $1 \le i \le n-1$ and $\sigma \in S_n$. This action commutes with the left action of $G(r,1,m)$ on $V^{\otimes n}$. Moreover, if $m \ge 2n$ and $r > n$ then the resulting map

$$\mathbb{C}P_n \to \operatorname{End}_{G(r,1,m)}(V^{\otimes n}) \tag{3}$$

is an isomorphism of algebras.

Classical Schur-Weyl duality states that the symmetric group algebra can be similarly recovered from the diagonal action of $GL(V)$ on $V^{\otimes n}$: if $\dim V \ge n$ then

$$\mathbb{C}S_n \cong \operatorname{End}_{GL(V)}(V^{\otimes n}) \,. \tag{4}$$

Malvenuto and Reutenauer [16] deduce from here the existence of a multiplication among permutations as follows. Given $\sigma \in S_p$ and $\tau \in S_q$, view them as linear endomorphisms of the tensor algebra

$$T(V) := \bigoplus_{n \ge 0} V^{\otimes n}$$

by means of (4) (σ acts as 0 on $V^{\otimes n}$ if $n \ne p$, similarly for τ). The tensor algebra is a Hopf algebra, so we can form the *convolution* product of any two linear endomorphisms:

$$T(V) \xrightarrow{\Delta} T(V) \otimes T(V) \xrightarrow{\sigma \otimes \tau} T(V) \otimes T(V) \xrightarrow{m} T(V) \,,$$

where Δ and m are the coproduct and product of the tensor algebra. Since these two maps commute with the action of $GL(V)$, the convolution of σ and τ belongs to $\mathrm{End}_{GL(V)}(V^{\otimes n})$, where $n = p + q$. Therefore, there exist an element $\sigma * \tau \in \mathbb{C}S_n$ whose right action equals the convolution of σ and τ. This is the product of Malvenuto and Reutenauer.

The same argument applies to uniform block permutations, in view of Proposition 2.1. We proceed to describe the resulting operation in explicit terms. As for permutations, this structure can be enlarged to that of a graded Hopf algebra.

2.2. **Product and coproduct of uniform block permutations.** Consider the graded vector space

$$\mathcal{P} := \bigoplus_{n\geq 0} \Bbbk P_n\,.$$

P_0 consists of the unique uniform block permutation of $[n]$, represented by the empty diagram, which we denote by $\emptyset$.

Let f and g be uniform block permutations of $[n]$ and $[m]$ respectively. Adding n to every entry in the diagram of g and placing it to the right of the diagram of f we obtain the diagram of a uniform block permutation of $[n+m]$, called the concatenation of f and g and denoted $f \times g$. Figure 3 shows an example.

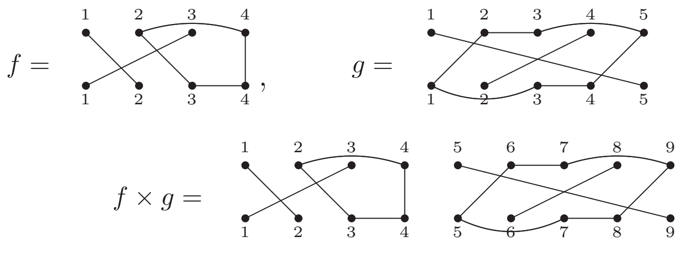


FIGURE 3. Concatenation of diagrams

Let $\mathrm{Sh}(n,m)$ denote the set of (n,m)-shuffles, that is, those permutations $\xi \in S_{n+m}$ such that

$$\xi(1) < \xi(2) < \cdots < \xi(n) \quad \text{and} \quad \xi(n+1) < \xi(n+2) < \cdots < \xi(n+m)\,.$$

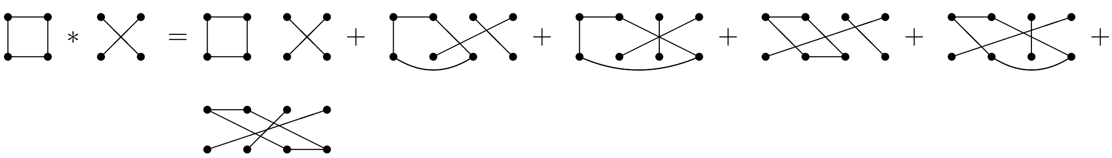
Let $sh_{n,m} \in \Bbbk S_{n+m}$ denote the sum of all (n,m)-shuffles.

The product $*$ on $\mathcal{P}$ is defined by

$$f * g := sh_{n,m} \cdot (f \times g) \in \Bbbk P_{n+m}$$

for all $f \in P_n$ and $g \in P_m$, and extended by linearity. It is easy to see that this product corresponds to convolution of endomorphisms of the tensor algebra via the map (3), when $\Bbbk = \mathbb{C}$.

For example,

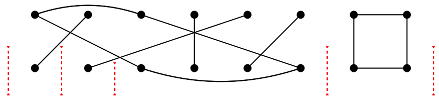


A *breaking point* of a set partition $\mathcal{B}$ is an integer $i \in \{0, 1, \ldots, n\}$ for which there exists a subset $S \subseteq \mathcal{B}$ such that

$$\bigcup_{B \in S} B = \{1, \ldots, i\} \quad \text{(and hence)} \quad \bigcup_{B \in \mathcal{B} \setminus S} B = \{i+1, \ldots, n\}\,.$$

Given a uniform block permutation $f : \mathcal{A} \to \mathcal{B}$, let $B(f)$ denote the set of breaking points of $\mathcal{B}$. Note that $i = 0$ and $i = n$ are breaking points of any f. If f is a permutation, that is if all blocks of f are of size 1, then $B(f) = \{0, 1, \ldots, n\}$.

In terms of the diagram of a uniform block permutation, if it is possible to put a vertical line between the first i and the last $n - i$ vertices in the bottom row without intersecting an edge between the two sets of vertices, then i is a breaking point.

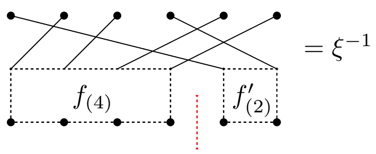
$f =$  $\Rightarrow \quad B(f) = \{0, 1, 2, 6, 8\}$.

Lemma 2.2. *If i is a breaking point of f, then there exists a unique $(i, n-i)$-shuffle $\xi \in S_n$ and unique uniform block permutations $f_{(i)} \in P_i$ and $f'_{(n-i)} \in P_{n-i}$ such that*

$$f = (f_{(i)} \times f'_{(n-i)}) \cdot \xi^{-1}\,.$$

Conversely, if such a decomposition exists, i is a breaking point of f.

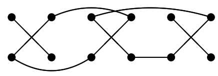
We illustrate this statement with an example where $i = 4$ and $\xi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 5 & 6 & 1 & 4 \end{pmatrix}$:

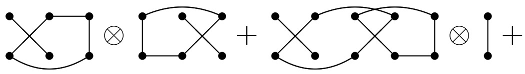
$f =$ 

We are now ready to define the coproduct on $\mathcal{P}$. Given $f \in P_n$ set

$$\Delta(f) := \sum_{i \in B(f)} f_{(i)} \otimes f'_{(n-i)},$$

where $f_{(i)}$ and $f'_{(n-i)}$ are as in Lemma 2.2. An example follows.

$f =$ 

$\Delta(f) = \quad f \otimes \emptyset +$  $+ \quad \emptyset \otimes f.$

Recall that an element $x \in \mathcal{P}$ is called primitive if $\Delta(x) = x \otimes \emptyset + \emptyset \otimes x$. Every uniform block permutation with breaking set $\{0, n\}$ is primitive, but there other primitive elements in $\mathcal{P}$. For example, the following element of $\Bbbk P_3$ is primitive:

Recall that $\emptyset$ denotes the empty uniform block permutation. Let $\varepsilon : \mathcal{P} \to \Bbbk$ be

$$\varepsilon(f) = \begin{cases} 1 & \text{if } f = \emptyset \in P_0, \\ 0 & \text{if } f \in P_n,\ n \geq 1. \end{cases}$$

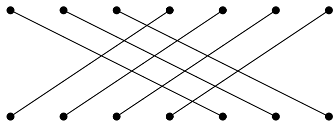
Theorem 2.3. *The graded vector space $\mathcal{P}$, equipped with the product $*$, coproduct Δ, unit $\emptyset$ and counit ε, is a graded connected Hopf algebra.*

Associativity and coassociativity follow from basic properties of shuffles (for the product one may also appeal to (3) and associativity of the convolution product). The existence of the antipode is guaranteed in any graded connected bialgebra. Compatibility between Δ and $*$ requires a special argument. We sketch part of it.

Let $\beta_{n,m}$ be the (n, m)-shuffle such that

$$\beta_{n,m}(i) = \begin{cases} m + i & \text{if } 1 \leq i \leq n, \\ i - n & \text{if } n + 1 \leq i \leq n + m. \end{cases}$$

The diagram of $\beta_{3,4}$ is shown below.



The inverse of $\beta_{n,m}$ is $\beta_{m,n}$.

Let $f \in P_n$, $g \in P_m$. A summand in $\Delta(f) * \Delta(g)$ is of the form

$$\xi_1 \cdot (f' \times g') \otimes \xi_2 \cdot (f'' \times g'')$$

where $p \in B(f)$, $f' \in P_p$, $f'' \in P_{n-p}$, $q \in B(g)$, $g' \in P_q$, $g'' \in P_{m-q}$, $\xi_1 \in \mathrm{Sh}(p, q)$, $\xi_2 \in \mathrm{Sh}(n-p, m-q)$, and there exist unique $\eta_1 \in \mathrm{Sh}(p, n-p)$ and $\eta_2 \in \mathrm{Sh}(q, m-q)$ such that $f \cdot \eta_1 = f' \times f''$ and $g \cdot \eta_2 = g' \times g''$.

Let $\beta := 1_p \times \beta_{n-p,q} \times 1_{m-q}$. Then

$$\begin{aligned} \xi_1 \cdot (f' \times g') \times \xi_2 \cdot (f'' \times g'') &= (\xi_1 \times \xi_2) \cdot ((f' \times g') \times (f'' \times g'')) \\ &= (\xi_1 \times \xi_2) \cdot \beta \cdot ((f' \times f'') \times (g' \times g'')) \cdot \beta^{-1} \\ &= (\xi_1 \times \xi_2) \cdot \beta \cdot (f \cdot \eta_1 \times g \cdot \eta_2) \cdot \beta^{-1} \\ &= (\xi_1 \times \xi_2) \cdot \beta \cdot (f \times g) \cdot (\eta_1 \times \eta_2) \cdot \beta^{-1}. \end{aligned}$$

Let $\xi := (\xi_1 \times \xi_2) \cdot \beta$ and $\eta := (\eta_1 \times \eta_2) \cdot \beta^{-1}$. One verifies that $\xi \in \mathrm{Sh}(n, m)$ and $\eta \in \mathrm{Sh}(p+q, n+m-p-q)$. Therefore,

$$\xi_1 \cdot (f' \times g') \times \xi_2 \cdot (f'' \times g'') = \xi \cdot (f \times g) \cdot \eta$$

is a summand in $\Delta(f * g)$. $\square$

Consider the following graded subspace of $\mathcal{P}$:

$$\mathcal{S} := \bigoplus_{n \geq 0} \Bbbk S_n .$$

Proposition 2.4. *$\mathcal{S}$ is a Hopf subalgebra of $\mathcal{P}$.*

$\mathcal{S}$ is the Hopf algebra of permutations of Malvenuto and Reutenauer [16]. Let σ be a permutation. In the notation of [2], the element $\sigma \in \mathcal{S}$ corresponds to the basis element F_σ^* of $\mathcal{S}Sym^*$, or equivalently the element $F_{\sigma^{-1}}$ of $\mathcal{S}Sym$.

2.3. **Inverse monoid structure and self-duality.** As $\mathcal{S}$, the Hopf algebra $\mathcal{P}$ is self-dual. To see this, recall that a block permutation is a bijection $f : \mathcal{A} \to \mathcal{B}$ between two set partitions of $[n]$. Let $\tilde{f} : \mathcal{B} \to \mathcal{A}$ denote the inverse bijection. If f is uniform then so is $\tilde{f}$. The diagram of $\tilde{f} \in P_n$ is obtained by reflecting the diagram of f across a horizontal line. Note that for $\sigma \in S_n \subseteq P_n$ we have $\tilde{\sigma} = \sigma^{-1}$.

Let $\mathcal{P}^*$ be the graded dual space of $\mathcal{P}$:

$$\mathcal{P}^* = \bigoplus_{n \geq 0} (\Bbbk P_n)^* .$$

Let $\{f^* \mid f \in P_n\}$ be the basis of $(\Bbbk P_n)^*$ dual to the basis P_n of $\Bbbk P_n$.

Proposition 2.5. *The map $\mathcal{P}^* \to \mathcal{P}$, $f^* \mapsto \tilde{f}$, is an isomorphism of graded Hopf algebras.*

The operation $f \mapsto \tilde{f}$ is also relevant to the monoid structure of P_n. Indeed, the following properties are satisfied

$$f = f\tilde{f}f \quad \text{and} \quad \tilde{f} = \tilde{f}f\tilde{f} .$$

Together with (5) below, these properties imply that P_n is an *inverse monoid* [6, Theorem 1.17]. The following properties are consequences of this fact [6, Lemma 1.18]:

$$\widetilde{fg} = \tilde{g}\tilde{f}, \quad \tilde{\tilde{f}} = f$$

(they can also be verified directly).

2.4. **Factorizable monoid structure and the weak order.** Let E_n denote the poset of set partitions of $[n]$: we say that $\mathcal{A} \leq \mathcal{B}$ if every bock of $\mathcal{B}$ is contained in a block of $\mathcal{A}$. This poset is a lattice, and this structure is related to the monoid structure of uniform block permutations as follows. If $id_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}$ denotes the uniform block permutation which is the identity map on the set of blocks of $\mathcal{A}$, then

$$id_{\mathcal{A}} \cdot id_{\mathcal{B}} = id_{\mathcal{A} \wedge \mathcal{B}} . \tag{5}$$

In other words, viewing E_n as a monoid under the meet operation $\wedge$, the map

$$E_n \to P_n , \quad \mathcal{A} \mapsto id_{\mathcal{A}} ,$$

is a morphism of monoids.

Any uniform block permutation $f \in P_n$ decomposes (non-uniquely) as

$$f = \sigma \cdot id_{\mathcal{A}} \tag{6}$$

for some $\sigma \in S_n$ and $\mathcal{A} \in E_n$. Note that σ is invertible and $id_{\mathcal{A}}$ is idempotent, by (5). It follows that P_n is a *factorizable inverse monoid* [5, Section 2], [15, Chapter 2.2]. Moreover, by Lemma 2.1 in [5], any invertible element in P_n belongs to S_n and any idempotent element in P_n belongs to (the image of) E_n. This lemma also guarantees that in (6), the idempotent $id_{\mathcal{A}}$ is uniquely determined by f (which is clear since $\mathcal{A}$ is the domain of f). On the other hand, σ is not unique, and we will make a suitable choice of this factor to define a partial order on P_n.

Consider the action of S_n on P_n by left multiplication. Given $\mathcal{A} \in E_n$, the orbit of $id_{\mathcal{A}}$ consists of all uniform block permutations $f : \mathcal{A} \to \mathcal{B}$ with domain $\mathcal{A}$, and the stabilizer is the *parabolic* subgroup

$$S_{\mathcal{A}} := \{\sigma \in S_n \mid \sigma(A) = A \ \forall A \in \mathcal{A}\} .$$

Consider the set of $\mathcal{A}$-*shuffles*:

$$\mathrm{Sh}(\mathcal{A}) := \{\xi \in S_n \mid \text{ if } i < j \text{ are in the same block of } \mathcal{A} \text{ then } \xi(i) < \xi(j)\} .$$

It is well-known that these permutations form a set of representatives for the left cosets of the subgroup $S_{\mathcal{A}}$. Therefore, given a uniform block permutation $f : \mathcal{A} \to \mathcal{B}$ there is a unique $\mathcal{A}$-shuffle ξ_f such that

$$f = \xi_f \cdot id_{\mathcal{A}}\,.$$

We use this decomposition to define a partial order on P_n as follows:

$$f \leq g \iff \xi_f \leq \xi_g\,,$$

where the partial order on the right hand side is the left weak order on S_n (see for instance [2]). We refer to this partial order as the *weak order* on P_n. Thus, P_n is the disjoint union of certain subposets of the weak order on S_n:

$$P_n \cong \bigsqcup_{\mathcal{A}\vdash[n]} \mathrm{Sh}(\mathcal{A})$$

(in fact, each $\mathrm{Sh}(\mathcal{A})$ is a lower order ideal S_n). Figures 4-8 show 5 of the 15 components of P_4. Note that even when $\mathcal{A}$ and $\mathcal{B}$ are set partitions of the same type the posets $\mathrm{Sh}(\mathcal{A})$ and $\mathrm{Sh}(\mathcal{B})$ need not be isomorphic.

The partial order we have defined on P_n should not be confused with the *natural partial order* which is defined on any inverse semigroup [7, Chapter 7.1], [15, Chapter 1.4].

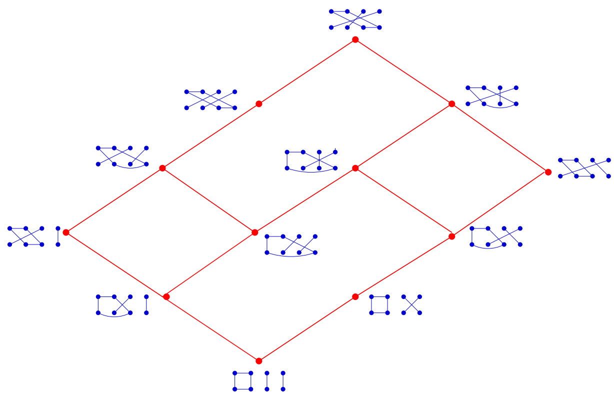


FIGURE 4. The component of P_4 corresponding to $\mathcal{A} = \{1,2\}\{3\}\{4\}$

Remark 2.6. As observed by Sloane [20], there is a connection between uniform block permutations and the *patience games* of Aldous and Diaconis [3]. Starting from a deck of cards a patience game produces a number of card piles according to certain simple rules (the output is not unique). If the cards are numbered $1,\ldots,n$, the initial deck is a permutation of $[n]$ and the resulting piles form a set partition of $[n]$. Suppose $\sigma \in S_n$. The set partitions $\mathcal{A}$ such that $\sigma \in \mathrm{Sh}(\mathcal{A})$ are precisely the possible outputs of patience games played from a deck of cards with $\sigma^{-1}(1)$ in the bottom, followed by $\sigma^{-1}(2)$, up to $\sigma^{-1}(n)$ on the top. Thus, uniform block permutations are in bijection with the pairs consisting of the input and the output of a patience game via $(\sigma,\mathcal{A}) \leftrightarrow \sigma \cdot id_{\mathcal{A}}$.

2.5. **The second basis and the Hopf algebra structure.** Following the ideas of [2], we use the weak order on P_n to define a new linear basis of the spaces $\Bbbk P_n$, on which the algebra structure of $\mathcal{P}$ is simple.

For each element $g \in P_n$ let

$$X_g := \sum_{f\leq g} f\,.$$

By Möbius inversion, the set $\{X_g \mid g \in P_n\}$ is a linear basis of $\mathcal{P}_n$.

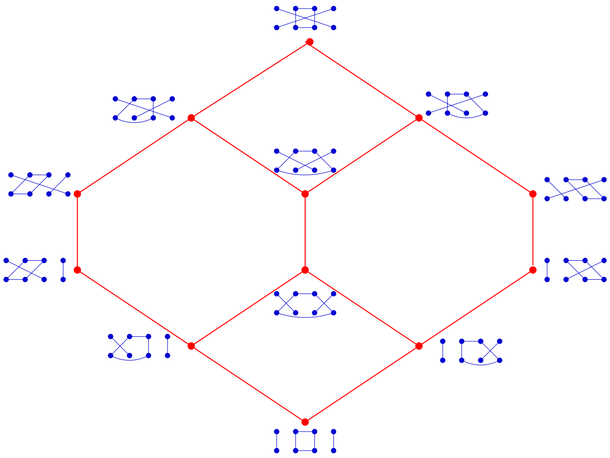


FIGURE 5. The component of P_4 corresponding to $\mathcal{A} = \{1\}\{2,3\}\{4\}$

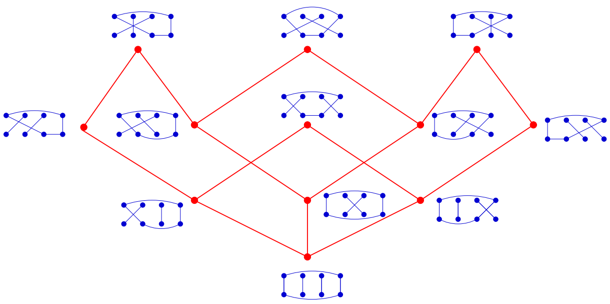


FIGURE 6. The component of P_4 corresponding to $\mathcal{A} = \{1,4\}\{2\}\{3\}$

Given $p, q \geq 0$, let $\xi_{p,q} \in S_{p+q}$ be the permutation

$$\xi_{p,q} := \begin{pmatrix} 1 & 2 & \dots & p & p+1 & p+2 & \dots & p+q \\ q+1 & q+2 & \dots & q+p & 1 & 2 & \dots & q \end{pmatrix} .$$

This is the maximum element of $\mathrm{Sh}(p,q)$ under the weak order. The product of $\mathcal{P}$ takes the following simple form on the X-basis.

Proposition 2.7. *Let $g_1 \in P_p$ and $g_2 \in P_q$ be uniform block permutations. Then*

$$X_{g_1} * X_{g_2} = X_{\xi_{p,q}\cdot(g_1 \times g_2)} .$$

Corollary 2.8. *The Hopf algebra $\mathcal{P}$ is free as an algebra and cofree as a graded coalgebra.*

Let V denote the space of primitive elements of $\mathcal{P}$. It follows that the generating series of $\mathcal{P}$ and V are related by

$$\mathcal{P}(x) = \frac{1}{1 - V(x)} .$$

Since

$$\mathcal{P}(x) = 1 + x + 3x^2 + 16x^3 + 131x^4 + 1496x^5 + 22482x^6 + \cdots$$

we deduce that

$$V(x) = x + 2x^2 + 11x^3 + 98x^4 + 1202x^5 + 19052x^6 + \cdots .$$

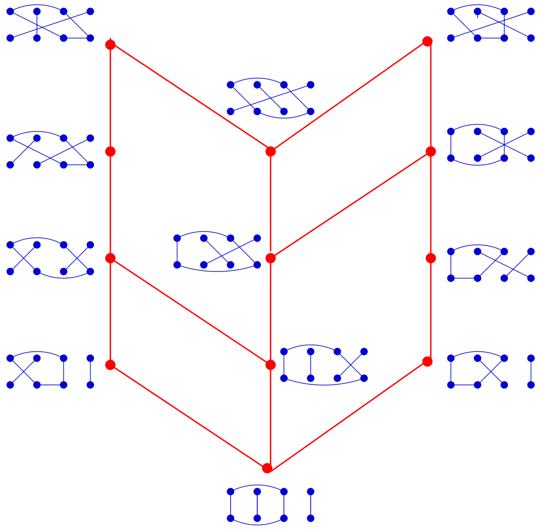


FIGURE 7. The component of P_4 corresponding to $\mathcal{A} = \{1,3\}\{2\}\{4\}$

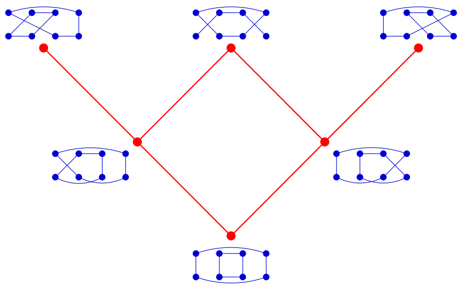


FIGURE 8. The component of P_4 corresponding to $\mathcal{A} = \{1,4\}\{2,3\}$

Remark 2.9. The same conclusion may be derived by introducing another basis

$$Z_g := \sum_{f \geq g} f .$$

This has the property that

$$Z_{g_1} * Z_{g_2} = Z_{g_1 \times g_2} .$$

Note that $Z_{id_{\mathcal{A}}}$ is the element denoted $Z_{\mathcal{A}}$ in Section 1.3.

3. The Hopf algebra of symmetric functions in non-commuting variables

Let X be a countable set, the *alphabet*. A *word of length* n is a function $w : [n] \to X$. Let $\Bbbk\langle\langle X \rangle\rangle$ be the algebra of non-commutative power series on the set of variables X. Its elements are infinite linear combinations of words, finitely many of each length, and the product is concatenation of words.

The *kernel* of a word w of length n is the set partition $\mathcal{K}(w)$ of $[n]$ whose blocks are the non-empty fibers of w. Order the set of set partitions of $[n]$ by refinement, as in Section 2.4. For each set partition $\mathcal{A}$ of $[n]$, let

$$p_A := \sum_{\mathcal{K}(w) \leq A} w \in \Bbbk\langle\langle X \rangle\rangle\,.$$

This is the sum of all words w such that if i and j are in the same block of $\mathcal{A}$ then $w(i) = w(j)$. For instance

$$p_{\{1,3\}\{2,4\}} = xyxy + xzxz + yxyx + \cdots + x^4 + y^4 + z^4 + \cdots\,.$$

The subspace of $\Bbbk\langle\langle X \rangle\rangle$ linearly spanned by the elements $p_{\mathcal{A}}$ as $\mathcal{A}$ runs over all set partitions of $[n]$, $n \geq 0$, is a subalgebra Π of $\Bbbk\langle\langle X \rangle\rangle$, graded by length. The elements of Π can be characterized as those power series of finite degree that are invariant under any permutation of the variables. Π is the algebra of symmetric functions in non-commuting variables introduced by Wolf [22] and recently studied by Gebhard, Rosas, and Sagan [10, 11, 18] in connection to Stanley's chromatic symmetric function.

Π is in fact a graded Hopf algebra [4, 1]. The coproduct is defined via evaluation of symmetric functions on two copies of the alphabet X. In order to describe the product and coproduct of Π on the basis elements $p_{\mathcal{A}}$ we introduce some notation.

Given set partitions $\mathcal{A} \vdash [n]$ and $\mathcal{B} \vdash [m]$ let $\mathcal{A} \times \mathcal{B}$ be the set partition of $[n+m]$ whose blocks are the blocks of $\mathcal{A}$ and the sets $\{b + n \,|\, b \in B\}$ where B is a block of $\mathcal{B}$. This corresponds to the operation $\times$ on uniform block permutations in the sense that $id_{\mathcal{A}} \times id_{\mathcal{B}} = id_{\mathcal{A}\times\mathcal{B}}$. For example, if $\mathcal{A} = \{1,3,4\}\{2,5\}\{6\} \vdash [6]$ and $\mathcal{B} = \{1,4\}\{2\}\{3,5\} \vdash [5]$, then $\mathcal{A} \times \mathcal{B} = \{1,3,4\}\{2,5\}\{6\}\{7,10\}\{8\}\{9,11\} \vdash [11]$.

To a set partition $\mathcal{A} \vdash [n]$ and a subset $S \subseteq \mathcal{A}$ we associate a new set partition $\mathcal{A}_S$ as follows. Write

$$\bigcup_{A \in S} A = \{j_1, \cdots, j_m\} \subseteq [n]$$

with $j_1 < j_2 < \cdots < j_m$. $\mathcal{A}_S$ is the set partition of $[m]$ whose blocks are obtained from the blocks $A \in S$ by replacing each j_t by t, for $1 \leq t \leq m$. For instance, if $S = \{1,5\}\{2,7\}$ then $\mathcal{A}_S = \{1,3\}\{2,4\} \vdash [4]$.

The product and coproduct of Π are given by

$$p_{\mathcal{A}} p_{\mathcal{B}} = p_{\mathcal{A}\times\mathcal{B}}\,, \tag{7}$$

$$\Delta(p_{\mathcal{A}}) = \sum_{S \sqcup T = \mathcal{A}} p_{\mathcal{A}_S} \otimes p_{\mathcal{A}_T}\,, \tag{8}$$

the sum over all decompositions of $\mathcal{A}$ into disjoint sets of blocks S and T. For example, if $\mathcal{A} = \{1,2,6\}\{3,5\}\{4\}$, then

$$\begin{aligned}\Delta(p_{\mathcal{A}}) \;=\;& p_{\mathcal{A}} \otimes 1 + p_{\{1,2,5\}\{3,4\}} \otimes p_{\{1\}} + p_{\{1,2,4\}\{3\}} \otimes p_{\{1,2\}} + p_{\{1,3\}\{2\}} \otimes p_{\{1,2,3\}} + \\ & p_{\{1,2,3\}} \otimes p_{\{1,3\}\{2\}} + p_{\{1,2\}} \otimes p_{\{1,2,4\}\{3\}} + p_{\{1\}} \otimes p_{\{1,2,5\}\{3,4\}} + 1 \otimes p_{\mathcal{A}}\,.\end{aligned}$$

Consider now the direct sum of the subspaces $\mathcal{Z}_n$ of $\Bbbk P_n$ introduced in Section 1.3:

$$\mathcal{Z} := \bigoplus_{n \geq 0} \mathcal{Z}_n \subset \mathcal{P}\,.$$

Theorem 3.1. *$\mathcal{Z}$ is a Hopf subalgebra of $\mathcal{P}$. Moreover, the map*

$$\mathcal{Z} \to \Pi, \quad Z_{\mathcal{A}} \mapsto p_{\mathcal{A}}$$

is an isomorphism of graded Hopf algebras.

Thus the Hopf algebra of uniform block permutations $\mathcal{P}$ contains the Hopf algebra Π of symmetric functions in non-commuting variables. Note also that this reveals the existence of a second operation on Π: according to Corollary 1.2, each homogeneous component Π_n carries an associative non-unital product that turns it into a right ideal of the monoid algebra $\Bbbk P_n$. Connections between Π and other combinatorial Hopf algebras are studied in [1].

References

[1] Marcelo Aguiar and Swapneel Mahajan, *Species and equivariant Hopf algebras*, in preparation (2004).

[2] Marcelo Aguiar and Frank Sottile, *Structure of the Malvenuto-Reutenauer Hopf algebra of Permutations*, Advances in Mathematics **191** n2 (2005) 225–275.

[3] David Aldous and Persi Diaconis, *Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem*, Bull. Amer. Math. Soc. **36** (1999) 413–432.

[4] Nantel Bergeron , Christophe Reutenauer, Mercedes Rosas, and Mike Zabrocki, *The Hopf algebra of symmetric functions in noncommutative variables*, in preparation (2004).

[5] S. Y. Chen and S. C. Hsieh, *Factorizable Inverse Semigroups*, Semigroup Forum **8** (1974) n4 283–297.

[6] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups. Vol. I*, Mathematical Surveys, No. 7 American Mathematical Society, Providence, R.I., 1961 xv+224 pp.

[7] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups. Vol. II*, Mathematical Surveys, No. 7 American Mathematical Society, Providence, R.I., 1967 xv+350 pp.

[8] David Easdown, James East, and D. G. FitzGerlad, *Presentations of factorizable inverse monoids*, 2004.

[9] Desmond G. FitzGerald, *A presentation for the monoid of uniform block permutations*, Bull. Austral. Math. Soc., **68** (2003) 317–324.

[10] David D. Gebhard, Bruce E. Sagan, *Sinks in Acyclic Orientations of Graphs*, J. Combin. Theory (B) **80** (2000) 130–146.

[11] David D. Gebhard, Bruce E. Sagan, *A chromatic symmetric function in noncommuting variables*, J. Alg. Combin. **13** (2001) 227–255.

[12] Vaughan F. R. Jones, *The Potts Model and the symmetric group*, in Subfactors: Proceedings of the Tanaguchi Symposium on Operator Algebras, Kyuzeso, 1993, pp. 259-267, World Scientific, River Edge, NJ 1994.

[13] Masashi Kosuda, *Characterization for the party algebra*, Ryukyu Math. J. **13** (2000) 7–22.

[14] Masashi Kosuda, *Party algebra and construction of its irreducible representations*, paper presented at Formal Power Series and Algebraic Combinatorics (FPSAC01), Tempe, Arizona (USA), May 20-26, 2001.

[15] Mark V. Lawson, *Inverse semigroups. The theory of partial symmetries*, World Scientific, River Edge, NJ, 1998. xiv+411 pp.

[16] Claudia Malvenuto and Christophe Reutenauer, *Duality between quasi-symmetric functions and the Solomon descent algebra*, J. Algebra **177** n3 (1995), 967–982.

[17] Paul P. Martin, *Temperley-Lieb algebras for non-planar statistical mechanics - The partition algebra construction*, J. Knot Theory and its Applicatitions, **3** n1 (1994) 51–82.

[18] Mercedes Rosas and Bruce Sagan, *Symmetric functions in non-commuting variables*, to appear in Trans. Amer. Math. Soc.

[19] J.-M. Sixdeniers, K. A. Penson, and A. I. Solomon, *Extended Bell and Stirling numbers from hypergeometric exponentiation*, J. Integer Seq., **4** (2001), Article 01.1.4.

[20] Neil J. A. Sloane, *An on-line version of the encyclopedia of integer sequences*, Electron. J. Combin. **1** (1994), Feature 1, approx. 5 pp. (electronic), `http://akpublic.research.att.com/~njas/sequences/ol.html`.

[21] Kenichiro Tanabe, *On the centralizer algebra of the unitary reflection group $G(m,p,n)$*, Nagoya Math. J. **148** (1997) 113–126.

[22] M. C. Wolf, *Symmetric functions of non-commuting elements*, Duke Math. J. **2** (1936) 626–637.

Department of Mathematics, Texas A&M University, College Station, TX 77843, USA
E-mail address: `maguiar@math.tamu.edu`
URL: `http://www.math.tamu.edu/∼maguiar`

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA
E-mail address: `Rosa.C.Orellana@Dartmouth.EDU`
URL: `http://www.math.dartmouth.edu/∼orellana/`

ENUMERATION OF CONNECTED UNIFORM HYPERGRAPHS

TSIRY ANDRIAMAMPIANINA AND VLADY RAVELOMANANA

Abstract. In this paper, we are concerned in counting exactly and asymptotically connected labeled b-uniform hypergraphs ($b \geq 3$). Enumerative results on connected graphs are generalized here to connected uniform hypergraphs. For this purpose, these structures are counted according to the number of vertices and hyperedges. First, we show how to compute step by step the associated exponential generating functions (EGFs) by means of differential equations and provide combinatorial interpretations of the obtained results. Next, we turn on asymptotic enumeration. We establish Wright-like inequalities for hypergraphs and by means of complex analysis, we obtain the asymptotic number of connected b-uniform hypergraphs with n vertices and $(n+\ell)/(b-1)$ hyperedges whenever $\ell = o(n^{1/3}/b^{1/3})$. This latter result confirms a conjecture made by Karoński and Łuczak in [20] about the validity of their formula for excesses in the 'Wright's range'.

Résumé. Dans cet article, nous nous intéressons à l'énumération exacte puis asymptotique des hypergraphes b-uniformes $(b \geq 3)$. Des résultats énumératifs sur les graphes sont ici généralisés pour les hypergraphes b-uniformes. Dans cette optique, ces structures sont énumérées suivant le nombre de sommets et le nombre d'hyperarêtes. Premièrement, nous montrons comment obtenir récursivement leurs fonctions génératrices exponentielles et nous justifions alors les formes des résultats ainsi obtenus par des arguments combinatoires. Ensuite, nous faisons l'énumération asymptotique. Nous établissons des inégalités similaires à celles de Wright pour les hypergraphes, en passant par de l'analyse complexe, nous obtenons l'asymtotique du nombre d'hypergraphes connexes b-uniformes ayant n sommets et $(n+\ell)/(b-1)$ hyperarêtes quand $\ell = o(n^{1/3}/b^{1/3})$. Ce dernier résultat confirme une conjecture de Karoǹski et Łuczak dans [20] pour avoir une formule valide avec des excès dans 'l'intervalle de Wright'.

1. Introduction

In this paper we are concerned with counting exactly and asymptotically members of families of labeled connected b-uniform hypergraphs with a given number of vertices and hyperedges and without multiple hyperedges. A labeled *hypergraph* $\mathcal{H} = (V, E)$ is given by a set V of n vertices with a family E of subsets of V of cardinal ≥ 2 (see Berge [4]). A member of E is called *hyperedge* and $\mathcal{H} = (V, E)$ is said b-*uniform* ($b \geq 2$) iff each member of E contains *exactly* b vertices. Therefore, 2-uniform hypergraphs are simply graphs. Let $\mathcal{H} = (V, E)$ be a hypergraph, uniform or not, then its *excess* is defined as (see [20]):

$$\text{excess}(\mathcal{H}) = \sum_{e \in E} (|e| - 1) - |V| \,. \tag{1}$$

Therefore, if $\mathcal{H} = (V, E)$ is a b-uniform hypergraph then its excess is given by the expression

$$\text{excess}(\mathcal{H}) = |E| \times (b-1) - |V| \,. \tag{2}$$

The notion of excess was first used in [27] where the author obtained substantial enumerative results in the study of connected graphs according to their number of vertices and edges. Wright's results appeared to be very important in the study of random graphs [6, 18, 19]. Later, Bender, Canfield and McKay [2] but also Pittel and Wormald [24] generalized Wright's results and obtained the asymptotic number of connected graphs of any given number of vertices and edges.

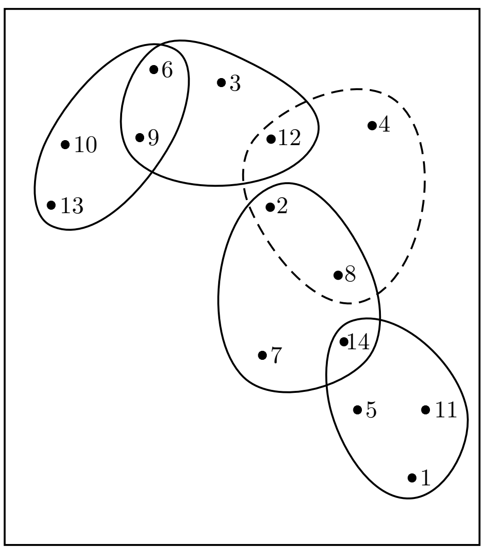
Key words and phrases. b-uniform hypergraphs; enumerative and analytic combinatorics; saddle-point method; generalized Wright's coefficients; random hypergraphs.

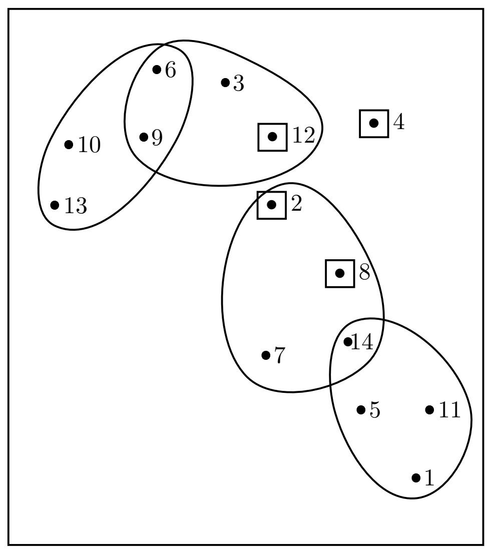
In contrary, much less is known about the number of hypergraphs of a given size. As far as we know, the most important results in these directions are those of Karoński and Łuczak in [20, 21]. In [20], the two authors used 'purely combinatorial arguments' to obtain their results. In this paper, our aim is to obtain analogous results to that of Wright [27, 29]. To do so, our approach is based on generating functions. Following the previously cited works, namely [18], connected hypergraphs with excess -1 are called *hypertrees*, connected hypergraphs with excess 0 are called *unicyclic components* or *unicycles*. Since these structures are labeled, we will use exponential generating functions (EGFs, for short) [15] to encode their number. Then, denote by H_ℓ the EGF of labeled connected b-uniform hypergraphs with excess ℓ. The purpose of this work is to compute the sequence of EGFs $(H_\ell)_{\ell \geq -1}$.

The outline of this paper is as follows. In the second section, we establish the differential equation satisfied by the EGFs H_ℓ ($\ell \geq -1$). We show how these EGFs can be computed exactly from this combinatorial equation and we retrieve some results that appeared in [25, 20]. In the third section, we give the forms of the expression of H_ℓ. We show that for every $\ell \geq -1$, H_ℓ can be expressed in terms of the EGF of rooted hypertrees and we give combinatorial interpretations of the forms of these EGFs. The next section is devoted to the asymptotic enumeration of uniform hypergraphs. First, we establish Wright-like inequalities for hypergraphs. Next, these inequalities are combined with methods from complex analysis and lead us to the asymptotic number of connected hypergraphs with n vertices and $(n + o(n^{1/3}))/(b-1)$ hyperedges.

2. Combinatorial equations satisfied by H_ℓ

2.1. **Hypergraphs surgery.** Let us start with a figure that illustrates, with 4-uniform hypergraph, the main idea from which we deduce the enumeration.





The figure on the left side is a connected 4-uniform hypergraph with 14 vertices and 5 hyperedges one of which is distinguished, namely the dashed hyperedge {2, 4, 8, 12}. The figure on the right side is a 4-uniform hypergraph with also 14 vertices but with only 4 hyperedges. This latter hypergraph is not connected but contains 3 components in which one or more vertices are distinguished (resp. {2, 8}, {4} and {12}). The above figures reflect combinatorial relations between families of connected hypergraphs with on first hand a distinguished hyperedge and on the other hand marked vertices. For instance, we refer the reader to Bergeron, Labelle and Leroux [5] for the use of distinguishing/marking and pointing in combinatorial species. The following lemma describes the relationships between number of edges and excesses in b-uniform connected components with distinguished hyperedge and marked vertices:

Lemma 2.1. *Consider a set $\mathcal{M}$ of connected b-uniform hypergraphs with one or more distinguished vertices. For any couple $(j,\, k)$, denote by m_{jk} the number of connected components, with excess j and with k marked vertices, in $\mathcal{M}$. Then, the hypergraph obtained when creating a (new) hyperedge connecting all the distinguished vertices of all the components in $\mathcal{M}$ is* (**i**) *connected,* (**ii**) *b-uniform and* (**iii**) *has excess ℓ if and only if*

$$\sum_{j,k} k\, m_{jk} = b \qquad \text{and} \qquad \sum_{j,k} (j+k)\, m_{jk} = \ell + 1\,. \tag{3}$$

Proof. It is clear that the created hypergraph is connected and is b-uniform if and only if the total number $(\sum_{j,k} k\, m_{jk})$ of distinguished vertices in the set is equal to b. Let N be the number of (connected) hypergraphs in the considered set $\mathcal{M}$ and let n be the total number of vertices in this set. Let us assign an arbitrary order to the members of the set and let n_i, s_i and k_i be respectively the number of vertices, hyperedges and distinguished vertices in the i-th hypergraph. The excess of the newly created hypergraph is then equals to ℓ if and only if $\sum_{i=1}^{N} s_i(b-1) + (b-1) - n = \ell$. We get $\sum_{i=1}^{N} s_i(b-1) + \sum_{i=1}^{N} k_i - \sum_{i=1}^{N} n_i = \ell + 1$ and $\sum_{i=1}^{N} \big(\{s_i(b-1) - n_i\} + k_i\big) = \ell + 1$. Therefore, $\sum_{j,k} m_{jk}(j+k) = \ell + 1$. $\square$

2.2. **Combinatorial equations.** In this paragraph, the previous correspondences are expressed in terms of EGFs. Let us consider the bivariate EGF H_ℓ. We have

$$\begin{aligned} H_\ell(w,z) &= \sum_{s=0}^{\infty}\sum_{n=0}^{\infty} h_\ell(s,n)\, w^s\, \frac{z^n}{n!} \\ &= \sum_{n=0}^{\infty} h_\ell\Big((n+\ell)/(b-1), n\Big)\, w^{(n+\ell)/(b-1)}\, \frac{z^n}{n!}\,, \end{aligned} \tag{4}$$

where w (resp. z) is the variable related to the number of hyperedges (resp. labeled vertices). In (4), $h_\ell(s,n)$ denotes the number of connected b-uniform hypergraphs with excess ℓ with s hyperedges and n vertices. Using (2), we note that $h_\ell(s,n) \neq 0$ iff $(n+\ell)/(b-1) \in \mathbb{N}$. The following theorem is inspired by the observations of paragraph 2.1 and gives recursive relation between the EGFs H_ℓ.

Theorem 2.2. *The bivariate EGFs $(H_\ell)_{\ell \geq -1}$ of labeled connected b-uniform hypergraphs satisfy*

$$w\frac{\partial}{\partial w} H_\ell(w,z) = w \sum_{(m_{jk}) \in S_\ell} \left\{ \prod_{j,k} \frac{1}{m_{jk}!} \left(\frac{z^k}{k!} \frac{\partial^k}{\partial z^k} H_j(w,z) \right)^{m_{jk}} \right\} - w\frac{\partial}{\partial w} H_{\ell-b+1}(w,z) \tag{5}$$

where S_ℓ is the following set of matrix:

$$S_\ell = \left\{ (m_{jk})_{\substack{-1 \leq j \leq \ell \\ 1 \leq k \leq b}} \text{ with } (m_{jk} \in \mathbb{N}) \text{ such that } \sum_{j,k} m_{jk}(j+k) = \ell + 1 \text{ and } \sum_{j,k} k m_{jk} = b \right\} \tag{6}$$

and $H_j \equiv 0$ if $j \leq -2$.

Proof. This equation relates, in terms of generating functions, the bijection between two sets of objects described by **a**) and **b**) as follows. **a**) In the left-hand side of (5), we have the EGF of the set of connected hypergraphs with excess ℓ and with a marked hyperedge. **b**) In the right-hand side, there are union of sets of components with one or more distinguished vertices that can be obtained from the removal, in a connected hypergraph of excess ℓ, of a hyperedge. After such removal, in each newly created component, the vertices which belonged to the removed hyperedge are marked. If there is k such distinguished vertices, in terms of EGFs, we then have $\frac{z^k}{k!}\frac{\partial^k}{\partial z^k} H_j(w,z)$. The second member of our equation can be interpreted as the creation of a (future) hyperedge with a total of b distinguished vertices in order to reconnect a set of hypergraphs. In the case where there is only one component, necessarily its excess is $\ell - b - 1$ and there are b of its vertices that are distinguished. These b vertices must not form an already existing hyperedge because we consider here hypergraphs without multiple hyperedges. It is the reason why we have to subtract the term $w\frac{\partial}{\partial w} H_{\ell-b+1}(w,z)$

in the RHS of (5). Observe that by the previous lemma, the definition of the set S_ℓ, *viz.* (6), ensures that the hypergraph obtained by the creation of a hyperedge connecting the marked vertices in the RHS is with excess ℓ and that the hyperedge which is created is formed with b vertices. $\square$

Remark 2.3. *We note that it is sufficient to determine the univariate EGFs since the corresponding bivariate EGFs can be deduced from the univariate ones simply using the relation*

$$H_\ell(w,z) = w^{\ell/(b-1)} H_\ell(w^{1/(b-1)} z)\,. \tag{7}$$

Saving the justification of its use for later, let us denote by $T(z)$ the (univariate) EGF corresponding to rooted hypertrees. Since a rooted hypertree is either a root or a root with a non-empty set of rooted (sub)hypertrees, borrowing methods from symbolic combinatorics (cf. [11]), we get

$$T(z) = z\ \exp\left(\frac{T(z)^{(b-1)}}{(b-1)!}\right). \tag{8}$$

Remark 2.4. *Throughout this paper, we use the notation H_ℓ followed by the couple of variables (w,z) to express the bivariate EGF, the notation H_ℓ followed by the variable (z) to express the univariate EGF. Whenever the variable are intentionally omitted, the EGF in used is*
$H_\ell \equiv H_\ell(T(z))$ where $T(z)$ is the EGF of rooted (b-uniform) hypertrees implicitly given by (8).

The EGFs $H_\ell \equiv H_\ell(T(z))$ satisfies the following.

Corollary 2.5. *For excess $\ell = -1$*

$$H_{-1} = T - \frac{(b-1)T^b}{b!}\,, \qquad T \equiv T(z) \tag{9}$$

and for $\ell \geq 0$

$$\begin{aligned}\frac{1}{b-1}\left(\ell H_\ell + T\frac{\mathrm{d}}{\mathrm{d}T}H_\ell\right) \quad &= \sum_{(m_{jk})\in S_\ell{}^*}\left\{\prod_{j,k}\frac{1}{m_{jk}!}\left(\frac{z^k}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}H_j(z)\right)^{m_{jk}}\right\}\\ &\quad -\frac{1}{b-1}\left((\ell-b+1)H_{\ell-b+1}(z) + z\frac{\mathrm{d}}{\mathrm{d}z}H_{\ell-b+1}(z)\right)\end{aligned} \tag{10}$$

where $S_\ell{}^*$ is the same as S_ℓ (see (6)) without the matrix where all coefficients equal zero except for the coefficients $m_{1,1} = b-1$ and $m_{\ell,1} = 1$.

Sketch of proof. Use the fact that

$$w\frac{\partial}{\partial w}H_j(w,z) = \frac{1}{b-1}\left(jH_j(w,z) + z\frac{\partial}{\partial z}H_j(w,z)\right)\,,$$

with (5) and (7) and set $w=1$. For $\ell=-1$, we have $S_{-1} = \{(m_{-11}, m_{-12}, \ldots, m_{-1,b}) = (b,0,0,\ldots,0)\}$. Therefore, we obtain

$$\frac{1}{b-1}\left(-H_{-1} + z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z)\right) \quad = \quad \frac{\left(z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z)\right)^b}{b!}\,,$$

and using the fact that $z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z) = T(z)$, it yields (9). To prove (10), first we note that for $\ell \geq 0$ the range of the matrix in $S_\ell{}^*$ can be rearranged so that the line index ranges from -1 to $\ell-1$ and the column index ranges from 1 to b. After some algebra, we get

$$\frac{1}{b-1}\left(\ell H_\ell + z\frac{\mathrm{d}}{\mathrm{d}z}H_\ell(z)\right) \quad = \quad J_\ell + \frac{\left(z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z)\right)^{b-1}}{(b-1)!}\left(z\frac{\mathrm{d}}{\mathrm{d}z}H_\ell(z)\right)$$

where J_ℓ is the RHS of (10). Again using $z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z) = T$, we obtain

$$\frac{1}{b-1}\left(\ell H_\ell + \left(z\frac{\mathrm{d}}{\mathrm{d}z}H_\ell(z)\right)\left(1 - \frac{T^{b-1}}{(b-2)!}\right)\right) = J_\ell\,.$$

From (8), we have

$$\frac{\mathrm{d}T}{\mathrm{d}z} = \frac{T}{z\left(1 - \frac{T^{b-1}}{(b-2)!}\right)} \tag{11}$$

and by the chain rule for differentiation we get the desired result. Note also that $\frac{z^k}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}H_j(z)$ can be expressed in terms of T so that (10) is completely a differential equation w.r.t. T. ☐

2.3. **Analytical resolution.** In this section we show how to compute the expression of H_ℓ, $\ell \geq 0$, in terms of the EGF T of rooted hypertrees. We note that the equation (10) for $\ell \geq 0$ allows us to compute recursively the expression of H_j for successive values of j. Thus, for each step, we have to solve a differential equation of order one in the variable T to get the expression of H_j which verifies the condition that ${H_j}_{|T=0} = 0$.

Lemma 2.6. *Let us define θ as*

$$\theta = 1 - \frac{T^{b-1}}{(b-2)!}\,. \tag{12}$$

For all $j \geq -1$ and for all $k \geq 0$, there is a function f_{jk} such that

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k}H_j(z) = \frac{f_{jk}(\theta)}{z^k T^j}\,. \tag{13}$$

Denoting $f_j \equiv f_{j0}$, in particular we have

$$H_j = \frac{f_j(\theta)}{T^j}\,. \tag{14}$$

Proof. From (11) and by the chain rule for differential we deduce (13):

$$f_{j,k+1}(\theta) = -(b-1)\frac{{f_{jk}}'(\theta)}{\theta} + (b-1){f_{jk}}'(\theta) - j\frac{f_{jk}(\theta)}{\theta} - kf_{jk}(\theta)\,. \tag{15}$$

The change of variable given by (12) allow us to deduce that $f_\ell(\theta)$ satisfies:

$$\frac{\mathrm{d}\left(f_\ell(\theta)\right)}{-(b-2)!} = \left(\sum_{(m_{jk})\in S_\ell{}^*}\prod_{j,k}\frac{1}{m_{jk}!}\left(\frac{f_{jk}(\theta)}{k!}\right)^{m_{jk}} - \frac{1}{b-1}\left((\ell-b+1)f_{\ell-b+1,0}(\theta) + f_{\ell-b+1,1}(\theta)\right)\right)\mathrm{d}\theta \tag{16}$$

☐

3. On the form of the EGFs H_ℓ

In order to establish the forms of the EGFs H_ℓ, we introduce some definitions.

Definitions. The *degree* of a vertex v is the number of the hyperedges that contain v.
A *special hyperedge* is one that contains 3 or more vertices of degree at least 2.
A *special vertex* is either a vertex that belongs to a special hyperedge or a vertex of degree ≥ 3.
A *pendant hyperedge* is one where there are $(b-1)$ vertices of degree 1. In the following, we call *path* a sequence of hyperedges. A path is also characterized by a starting vertex that belongs to the first hyperedge and by an ending vertex that belongs to the last hyperedge, and the sequence of hyperedges defining a path is such that each hyperedge contains exactly $(b-2)$ vertices of degree 1 in the hypergraph and where any pair of successive hyperedges share exactly one vertex that is not the starting nor the ending vertex of the path. We distinguish four kind of paths:

- *α-path*: a path that starts from and ends to the same special vertex, there are at least 2 hyperedges in an α-path and if there are exactly 2 hyperedges in an α-path then it is said to be *minimal*.
- *β-path*: a path that connects 2 special vertices such that if any hyperedge in the path is broken, these 2 special vertices become disconnected, there is at least 1 hyperedge in such a path; a single

hyperedge β-path is said to be *minimal*.
• γ*-path*: a path that joins 2 special vertices such that these vertices remain connected even if this path is broken, there is at least 1 hyperedge in such a path; a single hyperedge γ-path is said to be *minimal*.

A *basic hypergraph* is an unlabeled hypergraph that can be obtained from a labeled hypergraph by the following procedure:

- Discard the labels.
- Remove recursively pendant hyperedges.
- Shrink paths to a minimal special path of the same kind.

Thus, basic hypergraph has the same excess as any hypergraph from which it may be obtained: in the procedure we have described each time a hyperedge is removed (this happens when shrinking a path), it is just as if we have removed $(b-1)$ vertices. Furthermore, basic hypergraph has, for a given kind of paths, as much number of this kind of paths as in any (original) labeled hypergraph from which it may be obtained.

Let us enumerate the number of hypergraphs from which a fixed basic hypergraph with excess ℓ can be obtained. Let J be the EGF of such hypergraphs. Let m be the number of vertices in the basic hypergraph and respectively c_α, c_β and c_γ be the number of α-, β- and γ-paths, then

$$J = \frac{1}{g}\frac{T^m}{\theta^p} \tag{17}$$

where g is the number of automorphisms (e.g. [14]) of the basic hypergraph and $p = c_\alpha + c_\beta + c_\gamma$ is the number of α-, β- or γ-paths. The proof of this relation is immediate since "original" hypergraphs are obtained by rooting m rooted hypertrees in the basic hypergraph and by re-inserting p "chains" of eventually zero length in α- β- and γ-paths of the basic hypergraph. Thus, each hypergraph is obtained g times because of the number of choices where the m rooted hypertrees can be fixed. Furthermore, with s denoting the number of hyperedges in the considered basic hypergraph, there is a positive rational λ such that:

$$J = \frac{1}{g}\frac{T^{s(b-1)-\ell}}{\theta^p} = \lambda\frac{(1-\theta)^s}{T^\ell\theta^p}\,, \tag{18}$$

and necessary $s \geq \left\lfloor \frac{\ell+1}{b-1} + 1 \right\rfloor$.

Lemma 3.1. *For any basic hypergraph with excess ℓ, the total number of α-, β- and γ-paths verifies*

$$c_\alpha + c_\beta + c_\gamma \leq 3\ell$$

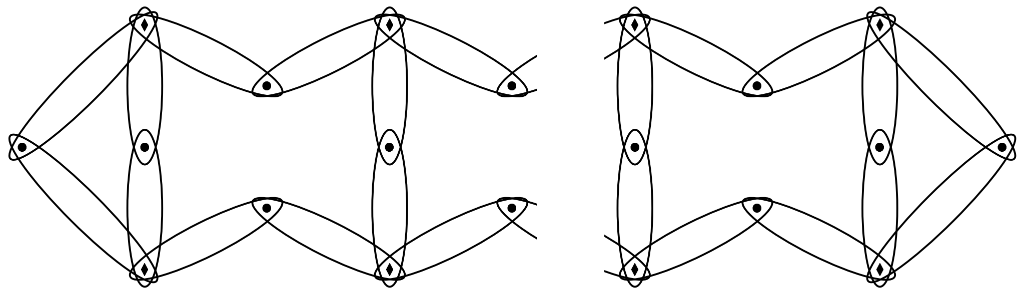
Proof. Let B_0 be the hypergraph induced by the special vertices in the basic hypergraph, let m_0 be the number of special vertices and r_0 be the number of special hyperedges then

$$m_0 + c_\alpha + 2(b-2)c_\alpha + (b-2)c_\beta + (b-2)c_\gamma + \ell = (b-1)\left(r_0 + 2c_\alpha + c_\beta + c_\gamma\right)\,. \tag{19}$$

Thus, $m_0 - r_0(b-1) + \ell = c_\alpha + c_\beta + c_\gamma$ and

$$-\text{excess}(B_0) + \ell = p\,. \tag{20}$$

Therefore, to determine the maximum of the number p over the basic hypergraph, it is sufficient to determine the minimal value of excess(B_0). excess(B_0) has minimal value if B_0 is a forest where hypertrees are either a single vertex (of degree 3 in the basic hypergraph) or a hyperedge (with exactly 3 vertices of degree 2 in the basic hypergraph). Therefore, if there is a hypergraph with hypergraph induced by the special vertices satisfying the above condition, we can deduce the maximum of the number p. The construction of such hypergraph is depicted in the following figure:



where only the vertices of degree at least 2 are represented. The basic hypergraph that can be obtained from the hypergraph in the figure above is only with special vertices of degree 3 and the hypergraph induced by these vertices consists of exactly $2 \times \ell$ isolated vertices (because the excess of the hypergraph is ℓ). Thus, $p \leq 3\ell$. ☐

Lemma 3.2. *$H_\ell = \frac{f_\ell(\theta)}{T^\ell}$ with f_ℓ a polynomial of maximum degree $\left\lfloor \frac{\ell+1}{b-1} + 1 \right\rfloor$.*

Proof. A matrix in ${S_\ell}^*$ corresponds to constructions as the one we have described in lemma 2.1. After having assigned an arbitrary order to the marked hypergraphs used in a such construction, let:

- the i-th hypergraph be of excess ℓ_i such that $\ell_i + 1 = q_i(b-1) + r_i$
- $\ell + 1 = q(b-1) + r$

with q, $q_i \geq 0$ and r, $r_i < b-1$.
As in the proof of lemma 2.1, we get here $\ell + 1 = q(b-1) + r = \sum_{i=1}^{N}(q_i(b-1) + r_i - 1 + k_i)$ since $\sum_{i=1}^{N}(r_i - 1 + k_i) \geq 0$, we deduce that $q \geq \sum_{i=1}^{N} q_i$. So,

$$q + 1 \geq \sum_{i=1}^{N} q_i + 1\,.$$

The lemma follows, since the summation in the right-hand side maximizes the degree of θ in $f_\ell(\theta)$. ☐

Using lemmas (3.1) and (3.2) with combinatorial identities , we obtain the following theorem and its corollary about the forms of the EGFs H_ℓ.

Theorem 3.3. *The EGF of connected b-uniform hypergraphs with excess ℓ can be put into the form*

$$H_\ell = \frac{(1-\theta)^{\left\lfloor \frac{\ell+1}{b-1}+1 \right\rfloor}}{T^\ell} \sum_{p=0}^{3\,\ell} A_{\ell\,p} \left(\frac{1-\theta}{\theta}\right)^p$$

with the coefficients $A_{\ell\,p}$ being rational.

Corollary 3.4. *The EGF of connected b-uniform hypergraphs with excess $\ell \geq 1$ can be rewritten as*

$$H_\ell = \frac{1}{T^\ell} \sum_{j=-3\,\ell}^{\left\lfloor \frac{\ell+1}{b-1}+1 \right\rfloor} c_j(\ell, b)\,\theta^j\,,$$

where $c_j(\ell, b) \in \mathbb{Q}$.

The proofs of theorem 3.3 and corollary 3.4 are omitted in this extended abstract.

4. Asymptotic results

4.1. **Wright-like inequalities for hypergraphs.** In order to compute the asymptotic number of connected ℓ-excess hypergraphs of a given size, we need the following result which gives the first two terms of H_ℓ. Let us recall that $\theta = 1 - T^{b-1}/(b-2)!$.

Lemma 4.1. *Developing the two first coefficients of the partial fraction form of H_ℓ, we get for $\ell \geq 1$*

$$T(z)^\ell H_\ell(z) = \frac{\lambda_\ell (b-1)^{2\ell}}{3\,\ell\,\theta(z)^{3\ell}} - \frac{(\kappa_\ell - \nu_\ell(b-2))(b-1)^{2\ell-1}}{(3\,\ell - 1)\,\theta(z)^{3\ell-1}} + \sum_{j=-3\ell+2}^{\left\lfloor \frac{\ell+1}{b-1}+1 \right\rfloor} c_j(\ell, b)\,\theta(z)^j \,. \tag{21}$$

In (21), $(\lambda_\ell)_{\ell\in\mathbb{N}}$ is defined recursively by $\lambda_0 = \frac{1}{2}$ and

$$\lambda_\ell \;=\; \frac{1}{2}\lambda_{\ell-1}(3\ell-1) \;+ \frac{1}{2}\sum_{t=0}^{\ell-1} \lambda_t \lambda_{\ell-1-t}\,, \qquad (\ell \geq 1)\,. \tag{22}$$

Similarly, define $(\nu_\ell)_{\ell\geq 1}$, $(\mu_\ell)_{\ell\geq 0}$ and $(\kappa_\ell)_{\ell\geq 1}$ as follows: $\nu_1 = \frac{5}{12}$ and

$$\begin{aligned} \nu_\ell \;&=\; \frac{1}{2}\lambda_{\ell-1} + \frac{1}{6}(3\ell-4)(3\ell-2)\lambda_{\ell-2} + \frac{1}{2}\sum_{t=0}^{\ell-2}(3t+2)\lambda_t\lambda_{\ell-2-t} \\ &+\; \frac{1}{6}\sum_{s=0}^{\ell-2}\sum_{t=0}^{\ell-2-s} \lambda_s\lambda_t\lambda_{\ell-2-s-t}\,(\ell\geq 2)\,. \end{aligned} \tag{23}$$

$$\kappa_\ell \;=\; \frac{1}{2}\left((3\ell-2)\mu_{\ell-1} + (3b\ell - b - 2\ell)\,\lambda_{\ell-1}\right) + \sum_{t=0}^{\ell-1}\mu_t\lambda_{\ell-1-t}\,. \tag{24}$$

$\mu_0 = b-1$ and for $\ell \geq 1$, μ_ℓ is given by

$$\mu_\ell \;=\; \kappa_\ell - \nu_\ell(b-2) + \lambda_\ell(b - \frac{2}{3})\,, \qquad (\ell \geq 1)\,. \tag{25}$$

Sketch of proof. Use the differential equation (10) given in corollary 2.5 with corollary 3.4, mainly focusing on the 'first two terms' of H_ℓ after a bit of standard algebra we get (21).

We are now ready to state similar inequalities such as those obtained by Wright in [29]. If A and B are two formal power series such that for all $n \geq 0$ we have $[z^n]\,A(z) \leq [z^n]\,B(z)$ then we denote this relation $A \preceq B$ (or $A(z) \preceq B(z)$).

Lemma 4.2. *For any $\ell \geq 1$, H_ℓ satisfies*

$$\frac{\lambda_\ell(b-1)^{2\ell}}{3\,\ell\,T(z)^\ell\,\theta(z)^{3\ell}} - \frac{(\kappa_\ell - \nu_\ell(b-2))(b-1)^{2\ell-1}}{(3\,\ell-1)\,T(z)^\ell\,\theta(z)^{3\ell-1}} \preceq H_\ell(z) \preceq \frac{\lambda_\ell(b-1)^{2\ell}}{3\,\ell\,T(z)^\ell\,\theta(z)^{3\ell}}\,, \tag{26}$$

where $(\lambda_\ell)_{\ell\in\mathbb{N}}$, $(\kappa_\ell)_{\ell\in\mathbb{N}^\star}$ and $(\nu_\ell)_{\ell\in\mathbb{N}^\star}$ are defined as in lemma 4.1.

The proof of this lemma will be provided in the full paper.

The following lemma gives the order of magnitude of the two first coefficients of the partial fraction form of H_ℓ.

Lemma 4.3. *We have*

$$\lambda_\ell = 3\left(\frac{3}{2}\right)^\ell \frac{\ell!}{2\pi}\left(1 + O\left(\frac{1}{\ell}\right)\right)\,, \tag{27}$$

$$\Big|\kappa_\ell - \nu_\ell(b-2)\Big| = O\left(\ell\lambda_\ell\right)\,. \tag{28}$$

Proof. To prove (27), it suffices to remark that $\lambda_\ell = 3\,\ell b_\ell$ where the sequence (b_ℓ) corresponds to the Wright's coefficients defined in [27, eq. (3.2)]. Therefore, by the proof of Lambert Meertens reported in [2] (see also Voblyĭ [26]), (27) holds. The remaining proof of (28) is technical and is omitted in this extended abstract. □

4.2. **A lemma from contour integration.** In order to get rid of the asymptotic behavior of the coefficients of $H_\ell(z)$, we need a last intermediate step. Define $h_n(a,\beta)$ as follows

$$\frac{1}{T(z)^a\left(1-\frac{T(z)^{b-1}}{(b-2)!}\right)^{3a+\beta}} = \sum_{n\geq 0} h_n(a,\beta)\frac{z^n}{n!}\,. \tag{29}$$

The following lemma is an application of the saddle point method [8, 11] which is well suited to cope with our analysis :

Lemma 4.4. *Let $a \equiv a(n)$ be such that $a(b-1)\to 0$ but $\frac{a(b-1)n}{\ln n^2}\to\infty$ and let β be a fixed number. Then $h_n(an,\beta)$ defined in (29) satisfies*

$$\begin{aligned} h_n(an,\beta) &= \frac{n!}{\sqrt{2\pi n\Big(b-1\Big)}\,\Big((b-1)!\Big)^{\frac{an+n}{b-1}}}\Big(1-(b-1)u_0\Big)^{(1-\beta)} \\ &\times\ \exp\left(n\Phi(u_0)\right)\left(1+O\Big(\sqrt{a(b-1)}\Big)+O\Big(\frac{1}{\sqrt{a(b-1)n}}\Big)\right)\,, \end{aligned} \tag{30}$$

where

$$\begin{aligned} \Phi(u) &= u-\left(\frac{a+1}{b-1}\right)\ln u - 3\,a\ln\left(1-(b-1)u\right) \\ u_0 &= \frac{3/2\,ab-a+1-1/2\,\sqrt{\Delta}}{b-1} \qquad \textit{with}\ \ \Delta = 9\,a^2b^2-12\,a^2b+12\,ab+4\,a^2-12\,a. \end{aligned} \tag{31}$$

Proof. Cauchy's integral formula gives

$$h_n(an,\beta) = n!\,[z^n]\,\frac{1}{T(z)^{an}\left(1-\frac{T(z)^{b-1}}{(b-2)!}\right)^{3a\,n+\beta}} = \frac{n!}{2\pi i}\oint\frac{1}{(T(z))^{a\,n}\left(1-\frac{T(z)^{(b-1)}}{(b-2)!}\right)^{(3a\,n+\beta)}}\frac{dz}{z^{n+1}}\,. \tag{32}$$

Note that the radius of convergence of the series $T(z)$ is given by $\sqrt[(b-1)]{(b-2)!}\,\exp\left(-1/(b-1)\right)$. We make the substitution $u = T(z)^{(b-1)}/(b-1)!$ and get successively

$$\begin{aligned} T(z) &= \sqrt[(b-1)]{(b-1)!\,u}\,, \quad z = \sqrt[(b-1)]{(b-1)!\,u}\,e^{-u} \quad \text{and} \\ dz &= \left(\frac{1}{(b-1)\,u}-1\right)\Big((b-1)!\,u\Big)^{\frac{1}{(b-1)}}e^{-u}\,du\,. \end{aligned} \tag{33}$$

From (32), we then obtain

$$h_n(a\,n,\beta) = \frac{n!}{2\pi i\Big((b-1)!\Big)^{(a\,n+n)/(b-1)}}\oint\frac{\left(1-(b-1)u\right)^{1-\beta}}{(b-1)\,u}\exp\left(n\Phi(u)\right)du\,, \tag{34}$$

where $\Phi(u) = u-\left(\frac{a+1}{b-1}\right)\ln u - 3\,a\ln\left(1-(b-1)u\right)$. The big power in the integrand, *viz.* $\exp\left(n\Phi(u)\right)$, suggests us to use the saddle point method. Investigating the roots of $\Phi'(u)=0$, we find two saddle points, $u_0 = \frac{3/2\,ab-a+1-1/2\,\sqrt{\Delta}}{b-1}$ and $u_1 = \frac{3/2\,ab-a+1+1/2\,\sqrt{\Delta}}{b-1}$ with $\Delta = 9\,a^2b^2-12\,a^2b+12\,ab+4\,a^2-12\,a$ Moreover, we have $\Phi''(u) = \frac{a+1}{(b-1)u^2}+3\,\frac{a(-b+1)^2}{(1-(b-1)u)^2}$ so that for $u\notin\{0,1/(b-1)\}$, $\Phi''(u)>0$. The main point of the application of the saddle point method here is that $\Phi^{'}(u_0)=0$ and $\Phi^{''}(u_0)>0$, hence $n\Phi(u_0\exp(i\tau))$ is well approximated by $n\Phi(u_0)-n{u_0}^2\Phi^{''}(u_0)\frac{\tau^2}{2}$ in the vicinity of $\tau=0$. If we integrate (34) around a circle passing vertically through $u=u_0$ in the z-plane, we obtain

$$h_n(an,\beta) = \frac{n!}{2\pi\left((b-1)!\right)^{(an+n)/(b-1)}}\int_{-\pi}^{\pi}\frac{\left(1-(b-1)u_0e^{i\tau}\right)^{1-\beta}}{(b-1)}\exp\left(n\Phi(u_0e^{i\tau})\right)d\tau \tag{35}$$

where

$$\Phi(u_0 e^{i\tau}) = u_0\cos\tau + iu_0\sin\tau - \frac{a+1}{b-1}\ln u_0 - i\frac{a+1}{b-1}\tau - 3a\ln(1-(b-1)u_0e^{i\tau})\ . \tag{36}$$

Denoting by $\mathfrak{Re}(z)$ the real part of z, if $f(\tau) = \mathfrak{Re}(\Phi(u_0e^{i\tau}))$ we have

$$f(\tau) = u_0\cos\tau - \frac{a+1}{b-1}\ln u_0 - 3a\ln u_0 - 3a\ln(b-1) - \frac{3a}{2}\ln\Big(1+\frac{1}{(b-1)^2u_0^2} - \frac{2\cos\tau}{(b-1)u_0}\Big)\,. \tag{37}$$

It comes

$$f'(\tau) = \frac{d}{d\tau}\mathfrak{Re}(h(u_0e^{i\tau})) = -u_0\sin\tau - \frac{3a\sin\tau}{u_0(b-1)+\frac{1}{(b-1)u_0} - 2\cos\tau}\,. \tag{38}$$

Therefore, if $\tau = 0$ $f^{'}(\tau) = 0$. Also, $f(\tau)$ is a symmetric function of τ and in $[-\pi, -\tau_0]\cup[\tau_0,\pi]$, for any given $\tau_0 \in (0,\pi)$, and $f(\tau)$ takes its maximum value for $\tau = \tau_0$. Since $|\exp(\Phi(u))| = \exp(\mathfrak{Re}(\Phi(u)))$, when splitting the integral in (35) into three parts, viz. "$\int_{-\pi}^{-\tau_0} + \int_{-\tau_0}^{\tau_0} + \int_{\tau_0}^{\pi}$", we know that it suffices to integrate from $-\tau_0$ to τ_0, for a convenient value of τ_0, because the others can be bounded by the magnitude of the integrand at τ_0. In fact, we have

$$\Phi(u_0e^{i\theta}) = \Phi(u_0) + \sum_{p\geq 2}\phi_p(e^{i\theta}-1)^p \tag{39}$$

where $\phi_p = \frac{{u_0}^p}{p!}\Phi^{(p)}(u_0)$. We easily compute $\Phi^{(p)}(u_0) = (-1)^p(p-1)!\Big(\frac{a+1}{(b-1){u_0}^p} + \frac{3a(1-b)^p}{(1-(b-1)u_0)^p}\Big)$, for $p \geq 2$. Whenever $ab \to 0$, we have

$$(b-1)u_0 = 1 - \sqrt{3\,(b-1)\,a} + (3/2\,b - 1)\,a + O\left(b^{3/2}a^{3/2}\right)\,. \tag{40}$$

Therefore, we obtain after a bit of algebra

$$|\phi_p| \leq O\left(\frac{2^p}{a^{\frac{p}{2}-1}(b-1)^{\frac{p}{2}}}\right)\,, \qquad \text{as } a(b-1)\to 0\,. \tag{41}$$

On the other hand,

$$|e^{i\tau}-1| = \sqrt{2(1-\cos\tau)} < \tau\,, \qquad \tau > 0\,. \tag{42}$$

Thus, the summation in (39) can be bounded for values of τ and a such that $\tau \to 0$, $ab \to 0$ ($a \to 0$) but $\frac{\tau}{\sqrt{a}} \to 0$ and we have

$$\Big|\sum_{p\geq 4}\phi_p(e^{i\tau}-1)^p\Big| \leq \sum_{p\geq 4}|\phi_p\tau^p| \leq \sum_{p\geq 4}O\Big(\frac{2^p\tau^p}{a^{\frac{p}{2}-1}(b-1)^{\frac{p}{2}}}\Big) = O\Big(\frac{\tau^4}{a(b-1)}\Big)\,. \tag{43}$$

It follows that for $\tau \to 0$, $a(b-1) \to 0$ and $\frac{\tau}{\sqrt{a(b-1)}} \to 0$, $\Phi(u_0e^{i\tau})$ can be rewritten as

$$\begin{aligned}\Phi(u_0e^{i\tau}) &= \Phi(u_0) - \frac{1}{(b-1)}\left(1 - \frac{\sqrt{a}}{\sqrt{3(b-1)}}\,\frac{3b-4}{2} + \frac{(9b^2-12b+4)}{12(b-1)}a\right)\tau^2 \\ &- \frac{i}{(b-1)}\left(1 - \frac{(3b-4)\sqrt{a}}{2\sqrt{3(b-1)}} + \frac{(9b^2-12b+4)}{12(b-1)}\,a\right)\tau^3 + O\left(\frac{\tau^4}{a(b-1)}\right)\,.\end{aligned} \tag{44}$$

Therefore, if $a(b-1) \to 0$ but $\frac{a(b-1)n}{(\ln n)^2} \to \infty$, if we let $\tau_0 = \frac{\ln n}{\sqrt{n\,u_0^2\Phi''(u_0)}}$ (with $u_0^2\Phi''(u_0) = \frac{2}{b-1} + O(\sqrt{a(b-1)})$) we can remark (as already said) that it suffices to integrate (35) from $-\tau_0$ to τ_0, using the magnitude of the integrand at τ_0 to bound the resulting error. In fact,

$$\begin{aligned}&\Big|(1-(b-1)u_0e^{i\tau_0})^{(1-\beta)}\exp\left(n\Phi(u_0e^{i\tau_0}) - nu_0 + \frac{n(a+1)}{(b-1)}\ln u_0 + 3an\ln\Big(1-(b-1)u_0\Big)\right)\Big| = \\ &\Big|1-(b-1)u_0e^{i\tau_0}\Big|^{(1-\beta)}\exp\Big(-\frac{n}{2}u_0^2\Phi''(u_0)\,{\tau_0}^2 + O\Big(n\frac{{\tau_0}^4}{a(b-1)}\Big)\Big) = O\Big(e^{-\frac{(\ln n)^2}{2}}\Big)\,.\end{aligned} \tag{45}$$

The rest of the proof is now standard application of the saddle point method (see for instance De Bruijn [8, Chapters 5 & 6]) and is omitted in this extended abstract. After a bit of algebra, one gets the formula (30). □

4.3. **Asymptotic number of connected hypergraphs.** We are now ready to state the main result of this section.

Theorem 4.5. *For $\ell \equiv \ell(n)$ such that $\ell = o\left(\sqrt[3]{\frac{n}{b}}\right)$ as $n \to \infty$, the number of connected b-uniform hypergraphs built with n vertices and having excess ℓ satisfies*

$$\sqrt{\frac{3}{2\pi}} \quad \frac{\left(b-1\right)^{\frac{\ell}{2}} \quad e^{\frac{\ell}{2}} \quad n^{n+\frac{3\ell}{2}-\frac{1}{2}}}{12^{\frac{\ell}{2}}\,\ell^{\frac{\ell}{2}}\left((b-2)!\right)^{\frac{n+\ell}{b-1}}} \exp\left(\frac{n}{b-1}-n\right)\left(1+O\left(\frac{1}{\sqrt{\ell}}\right)+O\left(\sqrt{\frac{b\,\ell^3}{n}}\right)\right). \tag{46}$$

We urge the reader to compare the methods and results obtained by Karoński and Łuczak in [20] with ours. In particular, the authors of [20] obtained results concerning various kinds of hypergraphs (smooth hypergraphs, clean hypergraphs, etc.). Unlike their results, where the excesses are of order $o(\log n/\log\log n)$, the theorem above states that the three variables n, ℓ and b can tend together to infinity but (46) remains valid whenever $\ell = o\left(\sqrt[3]{\frac{n}{b}}\right)$. Note also that by setting $b = 2$ in (46), we retrieve Wright's formula for graphs [29]. We remark also that the powerful methods developed in [2] and in [24] can be used to extend the validity of our asymptotic result.

Acknowledgments

The authors would like to thank warmly Gerard Duchamp (LIPN – University of Paris-Nord) and Jean-Gabriel Luque (IGM – Marne la Vallée) for providing proofs of the identities used to prove theorem 3.3 and its corollary. They would like to thank also Michal Karoński for useful discussions about [20] during the AofA meeting in Gdansk (2000).

References

[1] Bender E. A. (1974). Asymptotic Methods in Enumeration, *SIAM Review*, 16:485–515.

[2] Bender, E. A., Canfield, E. R. and McKay B. D. (1990). The asymptotic number of labeled connected graphs with a given number of vertices and edges. *Random Structures and Algorithms*, 1:127–169.

[3] Bender, E. A., Canfield, E. R. and McKay B. D. (1992). Asymptotic properties of labeled connected graphs. *Random Structures and Algorithms*, 3:183–202.

[4] Berge, C. (1976). *Graphs and hypergraphs* North-Holland Mathematical Library.

[5] Bergeron, F., Labelle, G. and Leroux, P. (1998). *Combinatorial Species and Tree-like Structures* Encyclopedia of Mathematics and its Applications, Vol. 67, Cambridge Univ. Press.

[6] Bollobás, B. (1985) *Random Graphs* Academic Press, London.

[7] Bouwkamp and De Bruijn *Hypergraphs and a functional equation* Journal of Combinatorial Theory, Ser.

[8] De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis.* Dover, New-York.

[9] Cayley, A. (1889). A Theorem on Trees. *Quart. J. Math. Oxford Ser.*, 23:376–378.

[10] Flajolet, P., Knuth, D. E. and Pittel B. (1989). The first cycles in an evolving graph. *Discrete Math.*, 75:167–215.

[11] Flajolet, P. and Sedgewick, R. *Analytic Combinatorics.* To appear (chapters are available as Inria research reports). See *http : //algo.inria.fr/flajolet/Publications/books.html*.

[12] Flajolet, P. and Sedgewick, R. *The average case analysis of algorithms: Counting and generating functions*

[13] Goulden, I. P. and Jackson, D. M. (1983). *Combinatorial Enumeration.* Wiley, New York.

[14] Harary, F. and Palmer, E. (1973). *Graphical Enumeration.* Academic Press, New-York and London.

[15] Herbert S. Wilf *generatingfunctionology*

[16] Janson, S. (1993). Multicyclic components in random graphs process. *Random Structures and Algorithms*, 4:71–84.

[17] Janson, S. (2000). Growth of components in random graphs. *Random Structures and Algorithms*, 17:343-356.

[18] Janson, S., Knuth, D. E., Łuczak, T. and Pittel B. (1993). The birth of the giant component. *Random Structures and Algorithms*, 4:233–358.

[19] Janson, S., Łuczak, T. and Rucinski A. (2000). *Random Graphs.* John Wiley, New York.

[20] Karoński, M. and Łuczak, T. (1997). The number of sparsely edges connected uniform hypergraphs. *Discrete Math.* 171:153–168.

[21] Karoński, M. and Łuczak, T. and (2002). The phase transition in a random hypergraph. *J. Comput. Appl. Math.* 142:125–135.

[22] Knuth, D. E. and Pittel, B. (1989). A recurrence related to trees. *Proc. Am. Math. Soc.*, 105:335–349.

[23] Rényi, A. (1959). On connected graphs I. *Publ. Math. Inst. Hungarian Acad. Sci.* 4:385–388.

[24] Pittel, B. and Wormald N. (2003). Counting connected graphs inside-out. *J. Comb. Th. Series B*, to appear.

[25] Selivanov, B. I. (1972). Perechislenie odnorodnykh hipergrafov c protstoĭ ciklovoĭ strukturoĭ, Kombinatornyĭ Analiz 2, pages 60–67.

[26] Voblyĭ, V. A. O koeffitsientakh Raĭta i Stepanova-Raĭta *Matematicheskie Zametki* 42 (1987), 854–862. English translation, V. A. Voblyi, "Wright and Stepanov-Wright coefficients," *Mathematical Notes* **42** (1987), 969–974.

[27] Wright, E. M. (1977). The Number of Connected Sparsely Edged Graphs. *Journal of Graph Theory*, 1:317–330.

[28] Wright, E. M. (1978). The Number of Connected Sparsely Edged Graphs. II. Smooth graphs and blocks *Journal of Graph Theory*, 2:299–305.

[29] Wright, E. M. (1980). The Number of Connected Sparsely Edged Graphs. III. Asymptotic results *Journal of Graph Theory*, 4:393–407.

E-mail address: {tsr,vlad}@lipn.univ-paris13.fr

Tsiry ANDRIAMAMPIANINA, LIPN – UMR CNRS 7030, Université de Paris-Nord, F 93430 Villetaneuse, France.

Vlady RAVELOMANANA, LIPN – UMR CNRS 7030, Université de Paris-Nord, F 93430 Villetaneuse, France.

Combinatorics and Representations of Complex Reflection Groups $G(r,p,n)$ (Extended Abstract)

Eli Bagno and Riccardo Biagioli

June 20, 2005

Abstract

For every $r, n, p|r$ there is a complex reflection group, denoted $G(r,p,n)$, consisting of all monomial $n \times n$ matrices such that all the nonzero entries are r^{th} roots of the unity and the r/p^{th} power of the product of the nonzero entries is 1. By considering these groups as subgroups of the colored permutation groups, $\mathbb{Z}_r \wr S_n$, we use Clifford theory to define on $G(r,p,n)$ combinatorial parameters and descent representations previously defined on Classical Weyl groups. One of these parameters is the major index which also has an important role in the decomposition of descent representations into irreducibles. We present also a Carlitz identity for these complex reflection groups.

1 Introduction

Let V be a complex vector space of dimension n. A *pseudo-reflection* on V is a linear transformation on V of finite order which fixes a hyperplane in V pointwise. A *complex reflection group* on V is a finite subgroup $W \leq \mathrm{GL}(V)$ generated by pseudo-reflections. Such groups are characterized by the structure of their invariant ring. More precisely, let $\mathbb{C}[V]$ be the symmetric algebra of V and let us denote by $\mathbb{C}[V]^W$ the algebra of invariants of W. Then Shephard-Todd [26] and Chevalley [13] proved that W is generated by pseudo-reflections if and only if $\mathbb{C}[V]^W$ is a polynomial ring.

Irreducible finite complex reflection groups have been classified by Shephard-Todd [26]. In particular there is a single infinite family of groups and exactly 34 other "exceptional" complex reflection groups. The infinite family $G(r,p,n)$ where r,p,n are positive integers numbers with $p|r$, consists of the groups of $n \times n$ matrices such that

1) the entries are either 0 or r^{th} roots of unity;
2) there is exactly one nonzero entry in each row and each column;
3) the $(r/p)^{\text{th}}$ power of the product of the nonzero entries is 1.

In particular the classical Weyl groups appear as special cases: $G(1,1,n) = S_n$ the symmetric group, $G(2,1,n) = B_n$, the Weyl group of type B, and $G(2,2,n) = D_n$ the Weyl group of type D.

Throughout research on complex reflection groups and their braid groups and Hecke algebras, the fact that they behave like Weyl groups has become more and more clear. In particular, it has recently discoverd that complex reflection groups (and not only Weyl groups) play a key role in the structure as well as in the representation theory of finite reductive groups. For more information on these results the reader is advised to consult the survey article of Broué [9], and the handbook of Geck and Malle [11].

One of the aims of this paper is to show that complex reflection groups continue to behave like Weyl groups also from the combinatorial point of view. In a way similar to Coxeter groups, they have presentations in terms of generators and relations, that can be visualized by Dynkin type diagrams (see e.g., [10]). Moreover, their elements can be represented as colored permutations. In fact, the complex reflection group $G(r,p,n)$ can be naturally identified as a normal subgroups of index p of the wreath product $G(r,n) := \mathbb{Z}_r \wr S_n$, where $\mathbb{Z}_r$ is the cyclic group of order r. This makes it possible to handle complex reflection groups by purely combinatorial methods. In Sections 2 and 8 we follow this approach. In particular, we introduce the concept of major index and descent number for complex reflection groups. Their joint distribution over the group is computed, giving rise to a nice identity that relates the two new statistics with the degrees of $G(r,p,n)$.

Then our investigation continues by showing the interplay between these new combinatorial objects and the representation theory of the group. More precisely, if we set $\mathbf{x} = x_1, \ldots, x_n$ as a basis for V, then $\mathbb{C}[V]$ can be identified with the ring of polynomials $\mathbb{C}[\mathbf{x}]$. The ring of invariants $\mathbb{C}[\mathbf{x}]^W$ is then generated by 1 and by a set of n algebraically independent homogeneous polynomials $\{f_1, \ldots, f_n\}$ which are called *basic invariants.* Although these polynomials are not uniquely determined, their degrees $d_1, \ldots, d_n$ are basic numerical invariants of the group, and are called the *degrees of* W. Let us denote by $\mathcal{I}_W$ the ideal generated by the invariants of strictly positive degree. The *module of coinvariants* of W is defined by

$$\mathbb{C}[\mathbf{x}]_W := \mathbb{C}[\mathbf{x}]/\mathcal{I}_W.$$

Since $\mathcal{I}_W$ is W-invariant, the group W acts naturally on $\mathbb{C}[\mathbf{x}]_W$. In fact, it is well known that $\mathbb{C}[\mathbf{x}]_W$ is isomorphic to the *left regular representation* of W. It follows that the dimension of $\mathbb{C}[\mathbf{x}]_W$ as a $\mathbb{C}$-module is equal to the order of the group W. In section 3, by using the combinatorial tools previously introduced, an explicit monomial basis for the module of coinvariants, called colored-descent basis, is provided.

Recently, another basis for $\mathbb{C}[\mathbf{x}]_W$ has been given by Allen [4]. Although both our and Allen's basis coincide with the Garsia-Stanton basis in the case of S_n, in general they are different as can be checked already in the small case of $G(2,2,2)$. It would be interesting to see if Allen's basis leads to an analogous definition of descent representations.

All this machinery leads to a natural definition of a new set of $G(r,p,n)$-modules, that we call *colored-descent representations*. They are generalizations of the descent representations introduced by Adin, Brenti, and Roichman in [3] for the symmetric and hyperoctahedral group, and which refine the descent representations of Solomon [25]. The decomposition into irreducibles of the colored-descent representations is provided. Moreover it turns out that the multiplicity of any irreducible representations is counted by the cardinality of a particular class of standard Young tableaux.

2 Complex Reflection Groups

For our exposition it will be much more convenient to consider wreath products not as groups of complex matrices, but as groups of colored permutations.

For any $n \in \mathbb{P} := \{1, 2, \ldots\}$ we let $[n] := \{1, 2, \ldots, n\}$, and for any $a, b \in \mathbb{N}$ we let $[a, b] := \{a, a+1, \ldots, b\}$. Let S_n be the symmetric group on $[n]$. A permutation $\sigma \in S_n$ will be denoted by $\sigma = \sigma(1) \cdots \sigma(n)$.

Let $r, n \in \mathbb{P}$. Define:

$$G(r,n) := \{((c_1, \ldots, c_n), \sigma) \mid c_i \in [0, r-1], \sigma \in S_n\}. \tag{1}$$

Any c_i can be considered as the color of the corresponding entry $\sigma(i)$. This explains the fact that this group is also called the *group of r-colored permutations*. Sometimes we will represent its elements in window notation as

$$g = g(1) \cdots g(n) = \sigma(1)^{c_1} \cdots \sigma(n)^{c_n}.$$

When it is not clear from the context, we will denote c_i by $c_i(g)$. Moreover, if $c_i = 0$, it will be omitted in the window notation of g. We denote by

$$\mathrm{Col}(g) := (c_1, \ldots, c_n) \quad \text{and} \quad \mathrm{col}(g) := \sum_{i=1}^{n} c_i,$$

the *color vector* and the *color weight* of any $\gamma := ((c_1, \ldots, c_n), \sigma) \in G(r,n)$.
For example, for $g = 4^1 3 2^4 1^2 \in G(5,4)$ we have $\mathrm{Col}(g) = (1,0,4,2)$ and $\mathrm{col}(g) = 7$.

Now let $p \in \mathbb{P}$ such that $p|r$. The *complex reflection group* $G(r,p,n)$ is the subgroup of $G(r,n)$ defined by

$$G(r,p,n) := \{g \in G(r,n) : \mathrm{col}(g) \equiv 0 \bmod p\}. \tag{2}$$

In particular we have the following chain of inclusions

$$G(r,r,n) \trianglelefteq G(r,p,n) \trianglelefteq G(r,1,n) = G(r,n),$$

where $\trianglelefteq$ stands for normal subgroup.

3 Colored Descent Basis

In order to lighten the notation, we let $G := G(r,n)$, $H := G(r,p,n)$, and $d := r/p$. The wreath product G acts on the ring of polynomials $\mathbb{C}[\mathbf{x}]$ as follows

$$\sigma(1)^{c_1}\cdots\sigma(n)^{c_n}\cdot P(x_1,\ldots,x_n) = P(\zeta^{c_{\sigma(1)}}x_{\sigma(1)},\ldots,\zeta^{c_{\sigma(n)}}x_{\sigma(n)}),$$

where ζ denotes a primitive r^{th} root of unity. A set of fundamental invariants under this actions is given by the *elementary symmetric functions* $e_j(x_1^r,\ldots,x_n^r)$, $1 \le j \le n$. Now, consider the restriction of the previous action on $\mathbb{C}[\mathbf{x}]$ to H. A set of fundamental invariants is given by

$$f_j(x_1,\ldots,x_n) := \begin{cases} e_j(x_1^r,\ldots,x_n^r) & \text{for } j = 1,\ldots,n-1 \\ x_1^d\cdots x_n^d & \text{for } j = n. \end{cases}$$

It follows that the degrees of H are $r, 2r, \ldots, (n-1)r, nd$.

Let $\mathcal{I}_H := (f_1,\ldots,f_n)$, the module of coinvariants $\mathbb{C}[\mathbf{x}]_H := \mathbb{C}[\mathbf{x}]/\mathcal{I}_H$ has dimension equal to $|H|$, that is $\frac{n!r^n}{p}$. In what follows we will associate to any element $h \in H$ an ad-hoc monomial in $\mathbb{C}[\mathbf{x}]$. Those monomials will form a linear basis for the module of coinvariants. In order to do this, we need to introduce various statistics on complex reflection groups.

For any $r,p,n \in \mathbb{P}$, with $p|r$ and $d := r/p$ we define the following subset of $G(r,n)$,

$$\Gamma(r,p,n) = \{\gamma = ((c_1,\ldots,c_n),\sigma) \in G(r,n) \mid c_n < d\}. \tag{3}$$

Note that $\Gamma := \Gamma(r,p,n)$ it is not a subgroup of G. Clearly, $|\Gamma| = n!r^{n-1}d$ and so it is in bijection with H. Although it is not fundamental for our purposes, we specify a bijection, in such a way that some of the definitions we will introduce, coincide with the usual ones, once we specialize H to any classical Weyl group. Indeed, one can easily check that the mapping

$$((c_1,\ldots,c_n),\sigma) \mapsto ((c_1,\ldots,\lfloor\frac{c_n}{p}\rfloor),\sigma) \tag{4}$$

is a bijection between H and Γ. As usual for any $a \in \mathbb{Q}$, $\lfloor a \rfloor$ denotes the greatest integer $\le a$. In order to make our definitions more natural and clear, from now on, we will work with Γ instead of H. Clearly, via the above bijection every function on Γ can be considered as a function on H and viceversa.

We fix the following order $\prec$ on colored integer numbers

$$1^{r-1} \prec 2^{r-1} \prec \ldots \prec n^{r-1} \prec \ldots \prec 1^1 \prec 2^1 \prec \ldots \prec n^1 \prec 1 \prec 2 \prec \ldots \prec n. \tag{5}$$

The *descent set* of an colored integer sequence $\gamma \in \Gamma$ is defined by $\mathrm{Des}(\gamma) := \{i \in [n-1] : \gamma_i \succ \gamma_{i+1}\}$. Moreover for any $\gamma = ((c_1,\ldots,c_n),\sigma) \in \Gamma$ we let

$$d_i(\gamma) := |\{j \in \mathrm{Des}(\gamma) : j \succeq i\}| \text{ and } m_i(\gamma) := r\cdot d_i(\gamma) + c_i(\gamma). \tag{6}$$

For every $\gamma \in \Gamma$ we define the $G(r,p,n)$-*major index* of γ by

$$\mathrm{m}(\gamma) := \sum_{i=1}^{n} m_i(\gamma). \tag{7}$$

For example, let $\gamma = 62^5 4^4 3^1 1^6 5^3 \in \Gamma(8,2,6)$. We have $(d_1(\gamma), \ldots, d_n(\gamma)) = (2,1,1,0,0,0)$, $(m_1(\gamma), \ldots, m_n(\gamma)) = (16,13,12,9,6,3)$ and $\mathrm{m}(\gamma) = 59$.

We are ready to associate to every element of Γ a monomial in $\mathbb{C}[\mathbf{x}]$. Let $\gamma = ((c_1, \ldots, c_n), \sigma) \in \Gamma$. We define

$$\mathbf{x}_\gamma := \prod_{i=1}^{n} x_{\sigma(i)}^{m_i(\gamma)}. \tag{8}$$

It is clear that $m_n(\gamma) < d$, hence $\mathbf{x}_\gamma$ is nonzero in $\mathbb{C}[\mathbf{x}]_H$.
For example, if $\gamma = 62^5 4^4 3^1 1^6 5^3 \in \Gamma(8,2,6)$ then $\mathbf{x}_\gamma = x_1^6 x_2^{13} x_3^9 x_4^{12} x_5^3 x_6^{16}$.

We restrict our attention to the quotient $S := \mathbb{C}[\mathbf{x}]/(f_n)$. Hence we consider nonzero monomials $M = \prod_{i=1}^{n} x_i^{a_i}$ such that $a_i < d$ for at least one $i \in [n]$. We associate to M the element $\gamma(M) = ((c_1, \ldots, c_n), \sigma) \in \Gamma$ such that for all $i \in [n]$

i) $a_{\sigma(i)} \geq a_{\sigma(i+1)}$;
ii) $a_{\sigma(i)} = a_{\sigma(i+1)} \Longrightarrow \sigma(i) < \sigma(i+1)$,
iii) $c_i \equiv a_{\sigma(i)} \pmod r$.

We denote by $\lambda(M) := (a_{\sigma(1)}, \ldots, a_{\sigma(n)})$ the *exponent partition* of M, and we call $\gamma(M) \in \Gamma$ the *colored index permutation*.

Now, let $M = \prod_{i=1}^{n} x_i^{a_i}$ be a nonzero monomial in S, and let $\gamma := \gamma(M)$ be its colored index permutation. Consider now the monomial $\mathbf{x}_\gamma$ associated to γ.

We associate to M another partition, $\mu(M)$, defined by

$$\mu'(M) := \left(\frac{a_{\sigma(i)} - m_i(\gamma)}{r} \right)_{i=1}^{n-1}, \tag{9}$$

where, as usual, μ' denotes the conjugate partition of μ.

Example 3.1. Let $r = 8$, $p = 2$, and $n = 6$ and consider the monomial $M = x_1^6 x_2^{21} x_3^{17} x_4^{20} x_5^3 x_6^{32} \in \mathbb{C}[x_1, \ldots, x_6]/(f_6)$. The exponent partition $\lambda(M) = (32,21,20,17,6,3)$ is obtained by reordering the power of x_i's following the colored index permutation $\gamma(M) = 62^5 4^4 3^1 1^6 5^3 \in \Gamma(8,2,6)$. We have already computed the monomial $\mathbf{x}_{\gamma(M)} = x_1^6 x_2^{13} x_3^9 x_4^{12} x_5^3 x_6^{16}$. It follows that $\mu(M) = (4,1)$.

We now define a partial order on the monomials of the same total degree in S. Let M and M' be nonzero monomials in S with the same total degree and such that the exponents of x_i in M and M' have the same parity (mod r) for every $i \in [n]$. Then we write $M' < M$ if one of the following holds:

1) $\lambda(M') \lhd \lambda(M)$, or

2) $\lambda(M') = \lambda(M)$ and $\mathrm{inv}(\gamma(M')) > \mathrm{inv}(\gamma(M))$.

Here, $\mathrm{inv}(\gamma) := |\{(i,j) \mid i < j \text{ and } \gamma(i) \succ \gamma(j)\}|$, and $\lhd$ denotes the *dominance order* defined on the set partitions of a fixed nonnegative integer n by: $\mu \unlhd \lambda$ if for all $i \geq 1$

$$\mu_1 + \mu_2 + \cdots + \mu_i \leq \lambda_1 + \lambda_2 + \cdots + \lambda_i.$$

Theorem 3.2. *The set*

$$\{\mathbf{x}_\gamma + \mathcal{I}_H : \gamma \in \Gamma\}$$

is a basis for $\mathbb{C}[\mathbf{x}]_H$.

Example 3.3. The elements of $\Gamma(6,3,2)$, $d = 2$, are

$$\begin{array}{llllll} 12 & 1^12 & 1^22 & 1^32 & 1^42 & 1^52 \\ 12^1 & 1^12^1 & 1^22^1 & 1^32^1 & 1^42^1 & 1^52^1 \\ 21 & 2^11 & 2^21 & 2^32 & 2^41 & 2^51 \\ 21^1 & 2^11^1 & 2^21^1 & 2^32^1 & 2^41^1 & 2^51^1. \end{array}$$

The corresponding monomials are

$$\begin{array}{llllll} 1 & x_1 & x_1^2 & x_1^3 & x_1^4 & x_1^5 \\ x_1^6x_2 & x_1x_2 & x_1^2x_2 & x_1^3x_2 & x_1^4x_2 & x_1^5x_2 \\ x_2^6 & x_2 & x_2^2 & x_2^3 & x_2^4 & x_2^5 \\ x_2^6x_1 & x_2^7x_1 & x_2^2x_1 & x_2^3x_1 & x_2^4x_1 & x_2^5x_1. \end{array}$$

It is easy to check that they form a basis for $\mathbb{C}[x_1, x_2]/(x_1^6 + x_2^6, x_1^2x_2^2)$.

4 The Representation Theory of $G(r, p, n)$

In this section we present the representation theory of the group $H := G(r, p, n)$. We follow the exposition of [31], (see also [20]). Since the irreducible representations of H are related to the irreducible representations of G via Clifford Theory, we start this section by presenting the representation theory of G.

Let $g = \sigma(1)^{c_1} \cdots \sigma(n)^{c_n} \in G$. First divide $\sigma \in S_n$ into cycles, and then provide the entries with their original color c_i by obtaining colored cycles. The color of a cycle is simply the sum of all the colors of its entries. For every $i \in [0, r-1]$, let α^i be the partition formed by the lengths of the cycles of g having color i. We may thus associate g with the r-partition $\vec{\alpha} = (\alpha^0, \ldots, \alpha^{r-1})$. Note that $\sum\limits_{i=0}^{r-1} |\alpha^i| = n$. We refer to $\vec{\alpha}$ as the *type* of g. One can prove that two elements of G are conjugate if and only if the have the same type.

It is well known that irreducible representations of G are also indexed by r-tuple of partitions $\vec{\lambda} := (\lambda^0, \ldots, \lambda^{r-1})$ with $\sum\limits_{i=0}^{r-1} |\lambda^i| = n$. We denote this set by $\mathcal{P}_{r,n}$.

As mentioned above, the passage to the representation theory of $G(r,p,n)$, is by Clifford theory. The group G/H can be identified with the cyclic group C of order p of the characters δ of G satisfying $H \subset Ker(\delta)$. More precisely, define the linear character δ_0 of G by $\delta_0((c_1,\ldots,c_n),\sigma) := \zeta^{c_1+\ldots+c_n}$, so that $C =< \delta_0^d >\simeq \mathbb{Z}_p$. The group C acts on the set of irreducible representations of G by

$$V(\vec{\lambda}) \mapsto \delta \otimes V(\vec{\lambda})$$

where $V(\vec{\lambda})$ is the irreducible representation of G indexed by $\vec{\lambda}$, and $\delta \in C$. This action can be explicitly described as follows. Let $\vec{\lambda} = (\lambda^0,\ldots,\lambda^{r-1}) \in \mathcal{P}_{r,n}$, we define a *1-shift* of $\vec{\lambda}$ by

$$(\vec{\lambda})^{\circlearrowleft 1} := (\lambda^{r-1}, \lambda^0, \ldots, \lambda^{r-2}). \tag{10}$$

By applying i-times the *shift operator* we get $(\vec{\lambda})^{\circlearrowleft i}$. Then once can show (see [20, Section 4]) that

$$\delta^i \otimes V(\vec{\lambda}) \simeq V((\vec{\lambda})^{\circlearrowleft i}), \tag{11}$$

for every $\delta \in C$.

Now let us denote by $[\vec{\lambda}]$ a C-orbit of the representation $V(\vec{\lambda})$. From (11) we obtain that $[\vec{\lambda}] = \{V(\vec{\mu}) \ : \ \vec{\mu} \sim \vec{\lambda}\}$, where the equivalence relation is defined by

$$\vec{\lambda} \sim \vec{\mu} \ \text{ if and only } \ \vec{\mu} = (\vec{\lambda})^{\circlearrowleft i\cdot d} \ \text{ for some } \ i \in [0, p-1]. \tag{12}$$

Let us denote $b(\vec{\lambda}) := |[\vec{\lambda}]|$, and set $u(\vec{\lambda}) := \frac{p}{b(\vec{\lambda})}$. Consider the stabilizer of $\vec{\lambda}$, $C_{\vec{\lambda}}$, that is:

$$C_{\vec{\lambda}} := \{\delta \in C \mid V(\vec{\lambda}) = \delta \otimes V(\vec{\lambda})\}.$$

Clearly, $C_{\vec{\lambda}}$ is a subgroup of C generated by $\delta_0^{b(\vec{\lambda})\cdot d}$ and so $|C_{\vec{\lambda}}| = u(\vec{\lambda})$.

It can be proven that the restriction of the irreducible representation $V(\vec{\lambda})$ of G to H decomposes into $u(\vec{\lambda}) = |C_{\vec{\lambda}}|$ non-isomorphic irreducible H modules. On the other hand, any other G-module in the same orbit $[\vec{\lambda}]$ will give us the same result. Actually, one can prove even more:

Theorem 4.1. *(See [31])*

There is a one to one correspondence between the irreducible representations of H and the ordered pairs $([\vec{\lambda}], \delta)$ where $[\vec{\lambda}]$ is the orbit of the irreducible representation $V(\vec{\lambda})$ of G and $\delta \in C_{\vec{\lambda}}$. Moreover if $\chi^{\vec{\lambda}}$ denotes the character of $V(\vec{\lambda})$ then

i) $\chi^{\vec{\lambda}} = \chi^{\vec{\mu}}$ for all $\vec{\mu} \in [\vec{\lambda}]$, and

ii) $\chi^{\vec{\lambda}} = \sum\limits_{\delta \in C_{\vec{\lambda}}} \chi^{([\vec{\lambda}],\delta)}$.

Here is a simple but important example. The irreducible representations of B_n ($G(2,1,n)$ in our notation) are indexed by bi-partitions of n. The Coxeter group D_n ($G(2,2,n)$ in our notation), is a subgroup of B_n of index 2. Thus the stabilizer of the action of $B_n/D_n \cong \mathbb{Z}_2$ on a pair of Young diagrams $\vec{\lambda} = (\lambda^1, \lambda^2)$ is either $\mathbb{Z}_2$ if $\lambda^1 = \lambda^2$, or $\{\mathrm{id}\}$ if $\lambda^1 \neq \lambda^2$. In the first case, the irreducible representation of B_n corresponding to $\vec{\lambda}$, when restricted to D_n, splits into two non-isomorphic irreducible representations of D_n. In the second case $\vec{\lambda} = (\lambda^1, \lambda^2)$ and $\vec{\lambda}^T = (\lambda^2, \lambda^1)$ correspond to two isomorphic irreducible representations of D_n.

5 n-Orbital Standard Tableaux

In this section we introduce a new class of standard Young r-tableaux that will be useful for our purposes later on.

Let $\vec{\lambda} = (\lambda^0, \ldots, \lambda^{r-1}) \in \mathcal{P}_{r,n}$ be an r-partition of n. A *Ferrers diagram* of shape $\vec{\lambda}$ is obtained by the union of the Ferrers diagrams of shapes $\lambda^0, \ldots, \lambda^{r-1}$, where the $(i+1)^{\text{th}}$ diagram lies south west of the i^{th}. A *standard Young r-tableau* $T := (T^0, \ldots, T^{r-1})$ of shape $\vec{\lambda}$ is obtained by inserting the integers $1, 2, \ldots, n$ as *entries* in the corresponding Ferrers diagram increasing along rows and down columns of each diagram separately. We denote by $\mathrm{SYT}(\vec{\lambda})$ the set of all r-standard Young tableaux of shape $\vec{\lambda}$. Any entry in the i component T^i of $T \in \mathrm{SYT}(\vec{\lambda})$ will be considered of color i.

A *descent* in an r-standard Young tableau T is an entry i such that $i+1$ is strictly below i. We denote the set of descents in T by $\mathrm{Des}(T)$.

Let

$$
\begin{aligned}
d_i(T) &:= |\{j \geq i \ : \ j \in \mathrm{Des}(T)\}, \quad c_i = c_i(T) := k \text{ if } i \in T^k; \\
f_i(T) &:= r \cdot d_i(T) + c_i(T), \quad f(T) := (f_1(T), \ldots, f_n(T)) \\
\mathrm{col}(T) &:= c_1 + \ldots + c_n, \\
\mathrm{maj}(T) &:= \sum_{i \in \mathrm{Des}(T)} i, \quad \text{and} \quad \mathrm{fmaj}(T) := r \cdot \mathrm{maj}(T) + \mathrm{col}(T).
\end{aligned}
$$

For example, the tableau T_1 in Figure 1 belongs to $\mathrm{SYT}((1),(2),(2,1),(1,1),(3,1),(2))$. We have that $\mathrm{Des}(T) = \{1,3,5,8,11,12\}$, $\mathrm{maj}(T) = 40$, $\mathrm{col}(T) = 1\cdot 2 + 2 \cdot 3 + 3 \cdot 2 + 4 \cdot 4 + 5 \cdot 2 = 40$, and so $\mathrm{fmaj}(T) = 280$.

Let $\vec{\lambda} = (\lambda^0, \ldots, \lambda^{r-1}) \in \mathcal{P}_{r,n}$. Let $[\vec{\lambda}] = \{\vec{\mu} \in \mathcal{P}_{r,n} \mid \vec{\mu} \sim \vec{\lambda}\}$ be the orbit of $\vec{\lambda}$ under the equivalence relation $\sim$ defined in (12). An *orbital standard Young tableau* $T = (T^0, \ldots, T^{r-1})$ *of type* $[\vec{\lambda}]$ is a standard Young r-tableau having one of the shapes in $[\vec{\lambda}]$. The following definition is fundamental in our work. An *n-orbital standard Young tableau* of type $[\vec{\lambda}]$ is an orbital tableau of type $[\vec{\lambda}]$ such that $n \in T^0 \cup \cdots \cup T^{d-1}$. We denote by $\mathrm{OSYT}_n[\vec{\lambda}]$ the set of all n-orbital r-tableaux of type $\vec{\lambda}$.

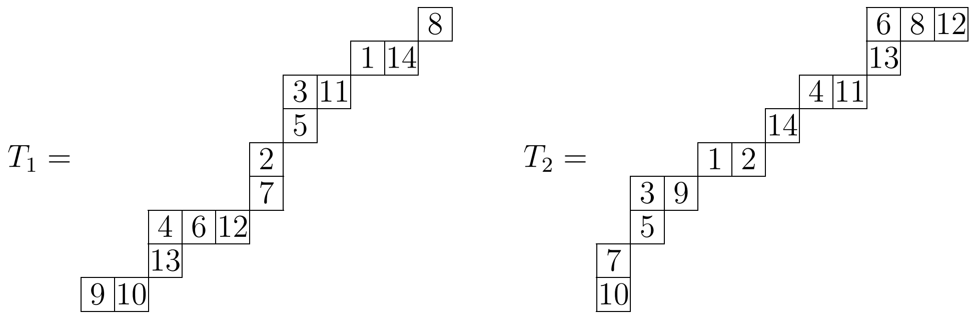


Figure 1: Two 6-standard Young tableaux

More precisely, let T be a standard Young r-tableau of shape $\vec{\lambda}$. From (12), it follows that all the possible orbital tableaux of type $[\vec{\lambda}]$, have shapes obtained from that of $\vec{\lambda}$ by applying $i \cdot d$-times the shift operator (10) , for $i = 0, \ldots, p-1$.

Example 5.1. Let $r = 6$ and $n = 17$. If $p = 3$, and so $d = 2$, then the two tableaux T_1 and T_2 in Figure 1 are of the same type $[\vec{\lambda}] = [(1), (2), (2, 1), (1, 1), (3, 1), (2)]$: T_1 is n-orbital, while T_2 is not. Differently, for $p = 2$, and $d = 3$ the two tableaux T_1 and T_2 are not in the same orbit $[\vec{\lambda}]$. Nervertheless, T_2 is an n-orbital tableau of type $[(3, 1), (2), (1), (2), (2, 1), (1, 1)]$.

6 Colored-descent representations of $G(r, p, n)$

The module of coinvariants $\mathbb{C}[\mathbf{x}]_H$ has a natural grading induced from that of $\mathbb{C}[\mathbf{x}]$. If we denote by R_k its k^{th} homogeneous component, we have

$$\mathbb{C}[\mathbf{x}]_H = \bigoplus_{k \geq 0} R_k.$$

In this section we introduce a set of $G(r, p, n)$-modules $R_{\mathcal{D},\mathcal{C}}$ which decompose R_k. These representations, called colored-descent representations, generalize the descent representations introduced for S_n and B_n by Adin, Brenti and Roichman in [3]. See also [8] for the case of D_n.

If $|\lambda| = k$ then one has:

$$\begin{aligned} J_{\lambda}^{\trianglelefteq} &:= \operatorname{span}_{\mathbb{C}}\{\mathbf{x}_{\gamma} + \mathcal{I}_H \mid \gamma \in \Gamma,\ \lambda(\mathbf{x}_{\gamma}) \trianglelefteq \lambda\} \text{ and} \\ J_{\lambda}^{\triangleleft} &:= \operatorname{span}_{\mathbb{C}}\{\mathbf{x}_{\gamma} + \mathcal{I}_H \mid \gamma \in \Gamma,\ \lambda(\mathbf{x}_{\gamma}) \triangleleft \lambda\} \end{aligned}$$

are two submodules of R_k. Their quotient is still an H-module, denoted by

$$R_{\lambda} := \frac{J_{\lambda}^{\trianglelefteq}}{J_{\lambda}^{\triangleleft}}.$$

For any $\mathcal{D} \subseteq [n-1]$ we define the partition $\lambda_{\mathcal{D}} := (\lambda_1, \ldots, \lambda_{n-1}, 0)$, where $\lambda_i := |\mathcal{D} \cap [i, n-1]|$. For $\mathcal{D} \subseteq [n-1]$ and $\mathcal{C} \in [0, r-1]^{n-1} \times \{0, ..., d-1\}$, we define the vector

$$\lambda_{\mathcal{D},\mathcal{C}} := r \cdot \lambda_{\mathcal{D}} + \mathcal{C},$$

where sum stands for sum of vectors.

From now on we denote $R_{\mathcal{D},\mathcal{C}} := R_{\lambda_{\mathcal{D},\mathcal{C}}}$, and by $\bar{\mathbf{x}}_\gamma$ the image of the colored-descent basis element $\mathbf{x}_\gamma \in J^{\trianglelefteq}_{\lambda_{\mathcal{D},\mathcal{C}}}$ in the quotient $R_{\mathcal{D},\mathcal{C}}$.

Proposition 6.1. *For any $\mathcal{D} \subseteq [n-1]$ and $\mathcal{C} \in [0, r-1]^{n-1} \times \{0, ..., d-1\}$, the set*

$$\{\bar{\mathbf{x}}_\gamma \ : \ \gamma \in \Gamma, \ \mathrm{Des}(\gamma) = \mathcal{D} \text{ and } \mathrm{Col}(\gamma) = \mathcal{C}\}$$

is a basis of $R_{\mathcal{D},\mathcal{C}}$. □

The H-modules $R_{\mathcal{D},\mathcal{C}}$ are called *colored-descent representation* in analogy with [3, Section 3.5]. They decompose the k^{th} component of $\mathbb{C}[\mathbf{x}]_H$ as follows.

Theorem 6.2. *For every $0 \leq k \leq r\binom{n}{2} + n(d-1)$,*

$$R_k \cong \bigoplus_{\mathcal{D},\mathcal{C}} R_{\mathcal{D},\mathcal{C}}$$

as H-modules, where the sum is over all $\mathcal{D} \subseteq [n-1], \mathcal{C} \in [0, r-1]^{n-1} \times \{0, ..., d-1\}$ such that

$$r \cdot \sum_{i \in \mathcal{D}} i + \sum_{j \in \mathcal{C}} j = k.$$

7 Decomposition of $R_{\mathcal{D},\mathcal{C}}$

In this section we prove a simple combinatorial description of the multiplicities of the irreducible representations of H in $R_{\mathcal{D},\mathcal{C}}$.

Theorem 7.1. *For every $\mathcal{D} \subseteq [n-1]$ and $\mathcal{C} \subseteq [0, r-1]^{n-1} \times \{0, ..., d-1\}$, $\vec{\lambda} \in \mathcal{P}_{r,n}$ and $\delta \in C_{\vec{\lambda}}$, the multiplicity of the irreducible representation of $G(r, p, n)$ corresponding to the pair $([\vec{\lambda}], \delta)$ in $R_{\mathcal{D},\mathcal{C}}$ is*

$$|\{T \in \mathrm{OSYT}[\vec{\lambda}] \mid \mathrm{Des}(T) = \mathcal{D}, \mathrm{Col}(T) = \mathcal{C}\}|.$$

As a corollary of this and of Theorem 6.2 we obtain the following result that is a generalization of a well known theorem on the decomposition of the coinvariant algebra of the symmetric group, (see e.g., [18] and [29]).

Theorem 7.2. *For $0 \leq k \leq r\binom{n}{2} + n(d-1)$, the representation R_k is isomorphic to the direct sum $\oplus m_{k,(\lambda,\delta)} V^{([\vec{\lambda}],\delta)}$, where $V^{([\vec{\lambda}],\delta)}$ is the irreducible representation of H labeled by $([\vec{\lambda}], \delta)$, and*

$$m_{k,([\vec{\lambda}],\delta)} := \mid \{T \in \mathrm{OSYT}[\vec{\lambda}] \ : \ \mathrm{m}(T) = k\} \mid .$$

8 Combinatorial Identities

In the case of classical Weyl groups and wreath products, any major statistic is associated with a descent statistic and their joint distribution is given by a nice closed formula, called *Carlitz identity*. In this last section we show that this is the case also for the complex reflection groups $G(r,p,n)$.

The following theorem presents the joint distribution of fdes and fmaj over $G(r,n)$.

Theorem 8.1 (Carlitz identity for G). *Let $n \in \mathbb{N}$. Then*

$$\sum_{k\geq 0}[k+1]_q^n t^k = \frac{\sum_{g\in G(r,n)} t^{\mathrm{fdes}(g)} q^{\mathrm{fmaj}(g)}}{(1-t)(1-t^r q^r)(1-t^r q^{2r})\cdots(1-t^r q^{nr})}.$$

Using the above theorem and a specific decomposition of $G(r,n)$ into subsets which are in a bijection with $G(r,p,n)$, we get the following:

Theorem 8.2 (Carlitz identity for H). *Let $n \in \mathbb{N}$. Then*

$$\sum_{k\geq 0}[k+1]_q^n t^k = \frac{\sum_{h\in G(r,n,p)} t^{\mathrm{d}(h)} q^{\mathrm{m}(h)}}{(1-t)(1-t^r q^r)(1-t^r q^{2r})\cdots(1-t^r q^{(n-1)r})(1-t^d q^{nd})}.$$

We refer to Theorem 8.1 and 8.2 as the Carlitz identities for G and H, respectively. It is worth to note that the powers of the q's in the denominators of the two formulas, $r, 2r, \ldots, nr$, and $r, 2r, \ldots, (n-1)r, nd$ are actually the degrees of $G(r,n)$ and $G(r,p,n)$.

References

[1] R. M. Adin and Y. Roichman, *The flag major index and group actions on polynomial rings*, European J. Combin., **22** (2001), 431-446

[2] R. M. Adin, F. Brenti and Y. Roichman, *Descent numbers and major indices for the hyperoctahedral Group*, Adv. in Appl. Math., **27** (2001), 210-224.

[3] R. M. Adin, F. Brenti and Y. Roichman, *Descent representations and multivariate statistics*, Trans. Amer. Math. Soc., to appear.

[4] E. E. Allen, *Descent monomials, P-partitions and dense Garsia-Haiman modules,* J. Alg. Combin., **20** (2004), 173-193.

[5] E. Bagno, *Euler-Mahonian parameter on colored permutation groups*, Sém. Loth. Combin., B51f (2004), 16 pp.

[6] P. Baumann and C. Hohlweg. *A Solomon descent theory for the wreath products* $G \wr \mathfrak{S}_n$, preprint `arXiv : math.CO/0503011`

[7] R. Biagioli and F. Caselli, *Invariant algebras and major indices for classical Weyl groups*, Proc. London Math. Soc., **88** (2004), 603-631.

[8] R. Biagioli and F. Caselli, *A descent basis for the coivariant algebra of type D*, J. Algebra, **275** (2004), 517-539.

[9] M. Broué, *Reflection groups, braid groups, Hecke algebras, finite reductive groups*, Current developments in mathematics, Int. Press, Somerville, (2001), 1-107.

[10] M. Broué, G. Malle and R. Rouquier, *Complex reflection groups, braid groups, Hecke algebras*, J. Reine Angew. Math., **500** (1998), 127-190.

[11] M. Geck and G. Malle, *Reflection groups*, Handbook of Algebra, vol. 4, North-Holland, Amsterdam, to appear.

[12] L. Carlitz, *A combinatorial property of q-Eulerian numbers*, Amer. Math. Montly, **82** (1975), 51-54.

[13] C. Chevalley, *Invariants of finite groups generated by reflections* Amer. J. Math., **77** (1955), 778-782.

[14] D. Foata, *On the Netto inversion number of a sequence*, Proc. Amer. Math. Soc., **19** (1968), 236-240.

[15] A. Garsia and I. Gessel, *Permutation statistics and partitions*, Adv. Math., **31** (1979), 288-305.

[16] A. Garsia and D. Stanton, *Group actions of Stanley-Reisner rings and invariant of permutation groups*, Adv. Math., **51** (1984), 107-201.

[17] I. M. Gessel, *Generating functions and enumeration of sequences*, Ph.D. Thesis, M.I.T., 1977.

[18] W. Kraskiewicz and Weymann, *Algebra of coinvariants and the action of Coxeter elements*, Bayreuth. Math. Schr., **63** (2001), 265-284.

[19] I. G. Macdonald, *Symmetric functions and Hall polynomials*, second edition, Oxford Math. Monographs, Oxford Univ. Press, Oxford 1995.

[20] H. Morita and H. F. Yamada, *Higher Specht polynomials for the complex reflection groups*, Hokkaido Math. J. **27** (1998), 505-515.

[21] P. A. MacMahon, *Combinatory analysis*, vol. 1, Cambridge Univ. Press, London, 1915.

[22] R. Mantaci and C. Reutenauer, *A generalization of Solomon's algebra for hyperoctahedral groups and other wreath products*, Comm. Algebra, **23** (1995), 27-56.

[23] V. Reiner, *Signed permutation statistics*, European J. Combin., **14** (1993), 553-567.

[24] C. Reutenauer, *Free Lie algebras,* London Mathematical Soc. Monographs, New series 7, Oxford Univ. Press, 1993.

[25] L. Solomon, *A Mackey formula in the group ring of a Coxeter group*, J. Algebra, **41** (1976), 255-264.

[26] G, C. Shephard and J. A. Todd, *Finite unitary reflection groups*, Canad. J. Math., **6** (1954), 274-304.

[27] T. Shoji, *Green functions associated to complex reflection groups*, J. Algebra, **245**, (2001), 650-694.

[28] R. P. Stanley, *Enumerative combinatorics*, vol. 2, Cambridge Stud. Adv. Math., no. 62, Cambridge Univ. Press, Cambridge, 1999.

[29] R. P. Stanley, *Recent progress in algebraic combinatorics*, Bull. Amer. Math. Soc. (new series), **40** (2003), 55-68.

[30] E. Steingrimsson, *Permutation statistics of indexed permutations*, European J. Combin., **15** (1994), 187-205.

[31] J. Stembridge, *On the eigenvalues of representations of reflection groups and wreath products,* Pacific J. Math., **140** (1989), 359-396.

THE EXCEDANCE NUMBER OF SOME COLORED PERMUTATION GROUPS (EXTENDED ABSTRACT)

ELI BAGNO AND DAVID GARBER

Abstract. We generalize the results of Ksavrelof and Zeng about the multidistribution of the excedance number of S_n with some natural parameters to the *colored permutation group* and to the Coxeter group of type D. We define two different orders on these groups which induce two different excedance numbers. Surprisingly, in the case of the *colored permutation group* we get the same generalized formulas for both orders.

1. Introduction

Let r and n be two positive integers. The *colored permutation group* $G_{r,n}$ consists of all permutations of the set

$$\Sigma = \{1, \dots, n, \bar{1}, \dots, \bar{n}, \dots, 1^{[r-1]}, \dots, n^{[r-1]}\}$$

satisfying $\pi(\bar{i}) = \overline{\pi(i)}$.

The symmetric group S_n is a special case of $G_{r,n}$ for $r = 1$ while for $r = 2$ we get the Weyl group of type B: B_n. In S_n one can define the following well-known parameters: Given $\sigma \in S_n$, $i \in [n]$ is *an excedance of* σ if and only if $\sigma(i) > i$. The number of excedances is denoted by $\text{exc}(\sigma)$. Two other natural parameters on S_n are the number of fixed points and the number of cycles of σ, denoted by $\text{fix}(\sigma)$ and $\text{cyc}(\sigma)$ respectively.

Consider the following generating function over S_n:

$$P_n(q, t, s) = \sum_{\sigma \in S_n} q^{\text{exc}(\sigma)} t^{\text{fix}(\sigma)} s^{\text{cyc}(\sigma)}.$$

$P_n(q, 1, 1)$ is the classical Eulerian polynomial, while $P_n(q, 0, 1)$ is the counter part for the derangements, i.e. the permutations without fixed points, see [**?**].

In the case $s = -1$, the two polynomials $P_n(q, 1, -1)$ and $P_n(q, 0, -1)$ have simple closed formulas:

Date: June 9, 2005.

$$P_n(q,1,-1) = -(q-1)^{n-1} \tag{1}$$

$$P_n(q,0,-1) = -q[n-1]_q \tag{2}$$

Recently, Ksavrelof and Zeng [?] proved some new recursive formulas which induce Equations (**??**) and (**??**). A natural problem is to generalize the results of [?] to other groups. The main challenge here is to choose a suitable order on the alphabet Σ of $G_{r,n}$ and define the parameters properly.

In this paper we cope with this challenge. We define two different orders on Σ, one of them, the *absolute order* 'forgets' the colors, while the other is much more natural, since it takes into account the colorful structure of $G_{r,n}$. This order is called the *color order*. The parameter exc will be defined according to both orders in two different ways. The interesting point is that for the group $G_{r,n}$ we get the same recursive formulas for both cases.

Define

$$P^{\mathbf{ord}}_{G_{r,n}}(q,t,s) = \sum_{\pi \in G_{r,n}} q^{\mathrm{exc}^{\mathbf{ord}}(\pi)} t^{\mathrm{fix}(\pi)} s^{\mathrm{cyc}(\pi)}$$

where **ord** can be either the absolute order or the color order.

For $G_{r,n}$, we prove the following two main results:

Theorem 1.1.

$$P^{\mathrm{Abs}}_{G_{r,n}}(q,1,-1) = P^{\mathrm{Clr}}_{G_{r,n}}(q,1,-1) = (q^r-1)P_{G_{r,n-1}}(q,1,-1).$$

Hence,

$$P^{\mathrm{Abs}}_{G_{r,n}}(q,1,-1) = P^{\mathrm{Clr}}_{G_{r,n}}(q,1,-1) = -\frac{(q^r-1)^n}{q-1}.$$

Theorem 1.2.

$$P^{\mathrm{Abs}}_{G_{r,n}}(q,0,-1) = P^{\mathrm{Clr}}_{G_{r,n}}(q,0,-1) = [r]_q(P_{G_{r,n-1}}(q,0,-1) - q^{n-1}[r]_q^{n-1}).$$

Hence,

$$P^{\mathrm{Abs}}_{G_{r,n}}(q,0,-1) = P^{\mathrm{Clr}}_{G_{r,n}}(q,0,-1) = -q[r]_q^n[n-1]_q,$$

where $[r]_q = 1 + \cdots + q^{r-1}$.

One can easily check that the formulas appeared in Theorem **??** and Theorem **??** indeed generalize the formulas of Ksavrelof and Zeng (for $r = 1$).

As mentioned above, when $r = 2$ we get the group B_n. This group has a well known normal subgroup called D_n consisting of the even signed permutations, i.e., permutations with an even number of minus signs. This group is also known as the Coxeter group of type D. With respect to D_n, we prove:

Theorem 1.3.

$$P_{D_n}^{\mathrm{Clr}}(q,1,-1) = (q^2-1)P_{D_{n-1}}(q,1,-1).$$

Hence,

$$P_{D_n}^{\mathrm{Clr}}(q,1,-1) = -(q^2-1)^{n-1}.$$

Theorem 1.4.

$$P_{D_n}^{\mathrm{Abs}}(q,1,-1) = -\frac{1}{2}(q-1)^{n-1}((1+q)^n+(1-q)^n).$$

2. Preliminaries

2.1. **Notations.** For $n \in \mathbb{N}$, let $[n] := \{1,2,\ldots,n\}$ (where $[0] := \emptyset$).
Also, let:

$$[n]_q := \frac{1-q^n}{1-q} = 1+q+\cdots+q^{n-1},$$

(so $[0]_q = 0$).

2.2. **The group of colored permutations.**

Definition 2.1. Let r and n be two positive integers. *The group of colored permutations of n digits with r colors* is the wreath product $G_{r,n} = \mathbb{Z}_r \wr S_n = \mathbb{Z}_r^n \rtimes S_n$, consisting of all the pairs (z,τ) where z is an n-tuple of integers between 0 and $r-1$ and $\tau \in S_n$. The multiplication is defined by the following rule: For $z = (z_1,...,z_n)$ and $z' = (z'_1,...,z'_n)$

$$(z,\tau)\cdot(z',\tau') = ((z_1+z'_{\tau(1)},...,z_n+z'_{\tau(n)}),\tau\circ\tau')$$

(here $+$ is taken modulo r).

In particular, $G_{1,n} = C_1 \wr S_n$ is the symmetric group S_n while $G_{2,n} = C_2 \wr S_n$ is the group of signed permutations B_n, also known as the hyperoctahedral group, or the classical Weyl group of type B.

A natural way to present $G_{r,n}$, which justifies its name, is the following: Consider the alphabet $\Sigma = \{1,\ldots,n,\bar{1},\ldots,\bar{n},\ldots,1^{[r-1]},\ldots,n^{[r-1]}\}$ as the set $[n]$ colored by the colors $0,\ldots,r-1$. Then, an element of $G_{r,n}$ is a *colored permutation*, i.e. a bijection $\pi : \Sigma \to \Sigma$ such that $\pi(\bar{i}) = \overline{\pi(i)}$.

Here are some conventions we use along this paper: For an element $\pi = (z,\tau) \in G_{r,n}$ with $z = (z_1,...,z_n)$ we write $z_i(\pi) = z_i$. For $\pi = (z,\tau)$, we denote $|\pi| = (0,\tau), (0 \in \mathbb{Z}_r^n)$. The element $(z,\tau) = ((1,0,3,2),(2,1,4,3)) \in G_{3,4}$ will be written as $(\bar{2}1\bar{\bar{\bar{4}}}\bar{\bar{3}})$.

2.3. **The Coxeter group of type D.** We define here the following normal subgroup of B_n of index 2, called the *Coxeter group of type D*:

$$D_n = \{\pi \in B_n \mid \sum_{i=1}^{n} z_i(\pi) \equiv 0 \pmod 2\}.$$

3. Statistics on $G_{r,n}$

Given any ordered alphabet Σ', we recall the definition of the *excedance set* of a permutation π on Σ' by :

$$\mathrm{Exc}(\pi) = \{i \in \Sigma' \mid \pi(i) > i\}$$

and the *excedance number* to be $\mathrm{exc}(\pi) = |\mathrm{Exc}(\pi)|$.

Example 3.1. *Given the order:* $\bar{\bar{1}} < \bar{\bar{2}} < \bar{\bar{3}} < \bar{1} < \bar{2} < \bar{3} < 1 < 2 < 3$, *we write* $\sigma = (3\bar{1}\bar{\bar{2}}) \in G_{3,3}$ *in an extended form:*

$$\begin{pmatrix} \bar{\bar{1}} & \bar{\bar{2}} & \bar{\bar{3}} & \bar{1} & \bar{2} & \bar{3} & 1 & 2 & 3 \\ \bar{\bar{3}} & 1 & \bar{2} & \bar{3} & \bar{\bar{1}} & 2 & 3 & \bar{1} & \bar{\bar{2}} \end{pmatrix}$$

and calculate: $\mathrm{Exc}(\sigma) = \{\bar{\bar{1}}, \bar{\bar{2}}, \bar{\bar{3}}, \bar{1}, \bar{3}, 1\}$ *and* $\mathrm{exc}(\sigma) = 6$.

We start by defining two orders on the set $\Sigma = \{1, \dots, n, \bar{1}, \dots, \bar{n}, \dots, 1^{[r-1]}, \dots, n^{[r-1]}\}$.

Definition 3.2. Define the *absolute order* on Σ to be:

$1^{[r-1]} < \cdots < \bar{1} < 1 < 2^{[r-1]} < \cdots < \bar{2} < 2 < \cdots < n^{[r-1]} < \cdots < \bar{n} < n$

and the *color order* on Σ by:

$$1^{[r-1]} < \cdots < n^{[r-1]} < \cdots < 1^{[1]} < \cdots < n^{[1]} < 1 < \cdots < n$$

Before defining the excedance number with respect to both orders, we have to introduce some notations.

Let $\sigma \in G_{r,n}$. We define:

$$\mathrm{csum}(\sigma) = \sum_{i=1}^{n} z_i(\sigma)$$

$$\mathrm{Exc}_A(\sigma) = \{i \in [n-1] \mid \sigma(i) > i\},$$

where the comparison is with respect to the color order.

$$\mathrm{exc}_A(\sigma) = |\mathrm{Exc}_A(\sigma)|$$

Let $\sigma \in G_{r,n}$. Recall that for $\sigma = (z, \tau) \in G_{r,n}$, $|\sigma|$ is the permutation of $[n]$ satisfying $|\sigma|(i) = \tau(i)$. For example, if $\sigma = (\bar{2}\bar{\bar{3}}1\bar{4})$ then $|\sigma| = (2314)$.

Now we can define the excedance numbers.

Definition 3.3. Define:

$$\mathrm{exc}^{\mathrm{Abs}}(\sigma) = \mathrm{exc}(|\sigma|) + \mathrm{csum}(\sigma)$$

$$\mathrm{exc}^{\mathrm{Clr}}(\sigma) = r \cdot \mathrm{exc}_A(\sigma) + \mathrm{csum}(\sigma)$$

Example 3.4. *Take* $\sigma = (\bar{1}\bar{\bar{3}}4\bar{2}) \in G_{3,4}$*. Then* $\mathrm{csum}(\sigma) = 4$, $\mathrm{Exc}_{\mathrm{A}}(\sigma) = \{3\}$, $\mathrm{Exc}(|\sigma|) = \{2,3\}$ *and thus* $\mathrm{exc}^{\mathrm{Abs}}(\sigma) = 6$ *and* $\mathrm{exc}^{\mathrm{Clr}}(\sigma) = 7$*.*

Recall that any permutation of S_n can be decomposed into a product of disjoint cycles. This notion can be easily generalized to the group $G_{r,n}$ as follows. Given any $\pi \in G_{r,n}$ we define the cycle number of $\pi = (z, \tau)$ to be the number of cycles in τ.

We say that $i \in [n]$ is an *absolute fixed point* of $\sigma \in G_{r,n}$ if $|\sigma(i)| = i$.

4. Proof of Theorem ?? for the color order

In this section we prove Theorem **??** for the color order. The idea of proving such identities is constructing a subset S of $G_{r,n}$ whose contribution to the generating function is exactly the right side of the identity and a killing involution on $G_{r,n} - S$, i.e., an involution on $G_{r,n} - S$ which preserves the number of excedances but changes the sign of each element of $G_{r,n} - S$ and hence shows that $G_{r,n} - S$ contributes nothing to the generating function.

Therefore, we divide $G_{r,n}$ into $2r + 1$ disjoint subsets as follows:

$$K_{r,n} = \{\sigma \in G_{r,n} \mid |\sigma(n)| \neq n, |\sigma(n-1)| \neq n\}$$

$$T^i_{r,n} = \{\sigma \in G_{r,n} \mid \sigma(n) = n^{[i]}\}, \qquad (0 \leq i \leq r-1)$$

$$R^i_{r,n} = \{\sigma \in G_{r,n} \mid \sigma(n-1) = n^{[i]}\}, \qquad (0 \leq i \leq r-1)$$

We first construct a killing involution on the set $K_{r,n}$. Define $\varphi : K_{r,n} \to K_{r,n}$ by

$$\sigma' = \varphi(\sigma) = (\sigma(n-1), \sigma(n))\sigma, \qquad \sigma \in K_{r,n}$$

Note that φ exchanges $\sigma(n-1)$ with $\sigma(n)$. It is obvious that φ is indeed an involution.

We will show that $\mathrm{exc}^{\mathrm{Clr}}(\sigma) = \mathrm{exc}^{\mathrm{Clr}}(\sigma')$. First, for $i < n-1$, it is clear that $i \in \mathrm{Exc}_{\mathrm{A}}(\sigma)$ if and only if $i \in \mathrm{Exc}_{\mathrm{A}}(\sigma')$. Now, as $\sigma(n-1) \neq n$, $n-1 \notin \mathrm{Exc}_{\mathrm{A}}(\sigma)$ and thus $n \notin \mathrm{Exc}_{\mathrm{A}}(\sigma')$. Finally, $|\sigma(n)| \neq n$ implies that $n-1 \notin \mathrm{Exc}_{\mathrm{A}}(\sigma')$. Now since it is obvious that $\mathrm{csum}(\sigma) = \mathrm{csum}(\sigma')$, we have that $\mathrm{exc}^{\mathrm{Clr}}(\sigma) = \mathrm{exc}^{\mathrm{Clr}}(\sigma')$.

On the other hand, $\mathrm{cyc}(\sigma)$ and $\mathrm{cyc}(\sigma')$ have different parities due to a multiplication by a transposition. Hence, φ is indeed a killing involution on $K_{r,n}$.

We turn now to the sets $T^i_{r,n}$. Note that there is a natural bijection between $T^i_{r,n}$ and $G_{r,n-1}$ defined by ignoring the last digit. Denote the image of $\sigma \in T^i_{r,n}$ under this bijection by σ'. Since $n \notin \mathrm{Exc_A}(\sigma)$, we have $\mathrm{exc_A}(\sigma) = \mathrm{exc_A}(\sigma')$. Now, since $z_n(\sigma) = i$ we have $\mathrm{csum}(\sigma') = \mathrm{csum}(\sigma) - \mathrm{i}$ and we get:

$$\mathrm{exc}^{\mathrm{Clr}}(\sigma) = \mathrm{exc}^{\mathrm{Clr}}(\sigma') + \mathrm{i}.$$

Now, since n is an absolute fixed point of σ, $\mathrm{cyc}(\sigma') = \mathrm{cyc}(\sigma) - 1$.

To summarize, we get that the total contribution of the elements in $T^i_{r,n}$ is:

$$P^{Clr}_{T^i_{r,n}} = -q^i P^{\mathrm{Clr}}_{G_{r,n-1}}(q, 1, -1)$$

for $0 \leq i \leq r-1$.

Now, we treat the sets $R^i_{r,n}$. There is a bijection between $R^i_{r,n}$ and $T^i_{r,n}$ using the same function φ we used above. Define $\varphi : R^i_{r,n} \to T^i_{r,n}$ by

$$\sigma' = \varphi(\sigma) = (\sigma(n-1), \sigma(n))\sigma.$$

When we compute the change in the excedance, we split our treatment into two cases: $i = 0$ and $i > 0$.

We start with the case $i = 0$. Note that $n - 1 \in \mathrm{Exc_A}(\sigma)$. On the other hand, in σ', $n-1, n \notin \mathrm{Exc_A}(\sigma')$. Hence, $\mathrm{exc_A}(\sigma) - 1 = \mathrm{exc_A}(\sigma')$.

Now, for the case $i > 0 : n-1, n \notin \mathrm{Exc_A}(\sigma)$ (since $\sigma(n-1) = n^{[i]}$ is not an excedance with respect to the color order). We also have: $n-1, n \notin \mathrm{Exc_A}(\sigma')$ and thus $\mathrm{Exc_A}(\sigma) = \mathrm{Exc_A}(\sigma')$ for $\sigma \in R^i_{r,n}$ where $i > 0$.

In both cases, we have that $\mathrm{csum}(\sigma) = \mathrm{csum}(\sigma')$. Hence, $\mathrm{exc}^{\mathrm{Clr}}(\sigma) - \mathrm{r} = \mathrm{exc}^{\mathrm{Clr}}(\sigma')$ for $i = 0$ and $\mathrm{exc}^{\mathrm{Clr}}(\sigma) = \mathrm{exc}(\sigma')$ for $i > 0$.

As before, the number of cycles changes its parity due to the multiplication by a transposition, and hence: $(-1)^{\mathrm{cyc}(\sigma)} = -(-1)^{\mathrm{cyc}(\sigma')}$.

Hence, the total contribution of elements in $R^i_{r,n}$ is

$$q^r P^{\mathrm{Clr}}_{G_{r,n-1}}(q, 1, -1)$$

for $i = 0$, and

$$q^i P^{\mathrm{Clr}}_{G_{r,n-1}}(q, 1, -1)$$

for $i > 0$.

Now, if we sum up all the parts, we get:

$$P^{\mathrm{Clr}}_{G_{r,n}}(q,1,-1)=\sum_{i=0}^{r-1}(-q^iP^{\mathrm{Clr}}_{G_{r,n-1}}(q,1,-1))+q^rP^{\mathrm{Clr}}_{G_{r,n-1}}(q,1,-1)+$$

$$\sum_{i=1}^{r-1}q^iP^{\mathrm{Clr}}_{G_{r,n-1}}(q,1,-1)=(q^r-1)P^{\mathrm{Clr}}_{G_{r,n-1}}(q,1,-1)$$

as needed.

Now, for $n=1$, $G_{r,1}$ is the cyclic group of order r and thus

$$P^{\mathrm{Clr}}_{G_{r,1}}(q,1,-1)=-(1+q+\cdots+q^{r-1})=-\frac{q^r-1}{q-1}$$

so we have

$$P^{\mathrm{Clr}}_{G_{r,n}}(q,1,-1)=-\frac{(q^r-1)^n}{q-1}$$

5. Proof of Theorem ?? for the color order

We recall that D_n is the subgroup of B_n consisting of the even signed permutations, i.e., permutations with an even number of minus signs. We divide D_n into 5 subsets:

$$K_n=\{\sigma\in D_n \mid |\sigma(n)|\neq n, |\sigma(n-1)|\neq n\}.$$

$$T_n^0=\{\sigma\in D_n \mid \sigma(n)=n\}.$$
$$T_n^1=\{\sigma\in D_n \mid \sigma(n)=\bar{n}\}.$$
$$R_n^0=\{\sigma\in D_n \mid \sigma(n-1)=n\}.$$
$$R_n^1=\{\sigma\in D_n \mid \sigma(n-1)=\bar{n}\}.$$

Now we denote:

$$a_n=P^{\mathrm{Clr}}_{D_n}(q,1,-1),$$
$$b_n=P^{\mathrm{Clr}}_{D_n^c}(q,1,-1),$$

where D_n^c is the complement of D_n in B_n.

Define $\varphi: K_n\to K_n$ by

$$\sigma'=\varphi(\sigma)=(\sigma(n-1),\sigma(n))\sigma.$$

Note that φ exchanges $\sigma(n-1)$ with $\sigma(n)$. It is easy to see that φ is a killing involution on K_n.

We turn now to the set T_n^0. Note that there is a natural bijection between T_n^0 and D_{n-1} defined by ignoring the last digit. Let $\sigma\in T_n^0$. Denote the image of $\sigma\in T_n^0$ under this bijection by σ'. Note that $\mathrm{csum}(\sigma')=\mathrm{csum}(\sigma)$, $\mathrm{Exc}_{\mathrm{A}}(\sigma')=\mathrm{Exc}_{\mathrm{A}}(\sigma)$ and $\mathrm{Exc}^{\mathrm{Clr}}(\sigma')=\mathrm{Exc}^{\mathrm{Clr}}(\sigma)$.

On the other hand, $\mathrm{cyc}(\sigma') = \mathrm{cyc}(\sigma) - 1$ and thus the restriction of a_n to T_n^0 is just $-a_{n-1}$.

For the contribution of the set T_n^1 note that the function φ defined above gives us a bijection between T_n^1 and D_{n-1}^c. In this case, $\mathrm{csum}(\sigma') = \mathrm{csum}(\sigma) - 1$, $\mathrm{exc}_{\mathrm{A}}(\sigma') = \mathrm{exc}_{\mathrm{A}}(\sigma)$ and $\mathrm{exc}(\sigma')^{\mathrm{Clr}} = \mathrm{exc}^{\mathrm{Clr}}(\sigma)$. On the other hand, $\mathrm{cyc}(\sigma') = \mathrm{cyc}(\sigma) - 1$ as before. Hence, the restriction of a_n to T_n^1 is $-qb_{n-1}$.

Now, for the set R_n^0, we have the following bijection between R_n^0 and D_{n-1}: for $\sigma \in R_n^0$, exchange the last two digits, and then ignore the last digit. If we denote the image of σ by σ', we have: $\mathrm{csum}(\sigma') = \mathrm{csum}(\sigma)$, $\mathrm{exc}_{\mathrm{A}}(\sigma') = \mathrm{exc}_{\mathrm{A}}(\sigma) - 1$, $\mathrm{exc}^{\mathrm{Clr}}(\sigma') = \mathrm{exc}^{\mathrm{Clr}}(\sigma) - 2$ and $\mathrm{cyc}(\sigma') \equiv \mathrm{cyc}(\sigma) \pmod 2$. Hence, the restriction of a_n to R_n^0 is $q^2 a_{n-1}$.

Similarly, for the set R_n^1, we have a bijection between R_n^1 and D_{n-1}^c: for $\sigma \in R_n^1$, exchange the last two digits, and then ignore the last digit. Denoting the image of σ by σ', we have $\mathrm{csum}(\sigma') = \mathrm{csum}(\sigma) - 1$, $\mathrm{exc}_{\mathrm{A}}(\sigma') = \mathrm{exc}_{\mathrm{A}}(\sigma)$, and hence $\mathrm{exc}^{\mathrm{Clr}}(\sigma') = \mathrm{exc}^{\mathrm{Clr}}(\sigma) - 1$. Also, we have $\mathrm{cyc}(\sigma') \equiv \mathrm{cyc}(\sigma) \pmod 2$. Hence, the restriction of a_n to R_n^1 is qb_{n-1}.

We summarize all the contributions over all the four subsets, and we have:

$$a_n = -a_{n-1} - qb_{n-1} + q^2 a_{n-1} + qb_{n-1} = (q^2 - 1)a_{n-1}.$$

For computing a_1, note that $D_1 = \{1\}$ and thus $a_1 = -1$.

Therefore, we have:

$$P_{D_n}^{\mathrm{Clr}}(q, 1, -1) = a_n = -(q^2 - 1)^{n-1},$$

and we are done. □

Acknowledgements

We wish to thank Alex Lubotzky and Ron Livne. We also wish to thank the Einstein Institute of Mathematics at the Hebrew University for hosting their stays.

References

[1] G. Ksavrelof and J. Zeng, *Two involutions for signed excedance numbers*, Semi. Loth. Comb. **49** (2002/04), Art. B49e, 8 pp. (electronic).

[2] R. P. Stanley, *Enumerative combinatorics*, Vol 1 and 2, Cambridge University Press, 1997.

Einstein institute of Mathematics, The Hebrew University, Givat Ram, 91904 Jerusalem, Israel

E-mail address: `bagnoe@math.huji.ac.il`

Einstein institute of Mathematics, The Hebrew University, Givat Ram, 91904 Jerusalem, Israel and, Department of Sciences, Holon Academic Institute of Technology, PO Box 305, 58102 Holon, Israel

E-mail address: `garber@math.huji.ac.il,garber@hait.ac.il`

On the Kronecker Product $s_{(n-p,p)} * s_\lambda$

C.M. Ballantine
College of the Holy Cross
Worcester, MA
cballant@holycross.edu

R.C. Orellana
Dartmouth College
Hanover, NH 03755
Rosa.C.Orellana@Dartmouth.edu

Abstract

The Kronecker product of two Schur functions s_λ and s_μ, denoted $s_\lambda * s_\mu$, is defined as the Frobenius characteristic of the tensor product of the irreducible representations of the symmetric group indexed by partitions of n, λ and μ, respectively. The coefficient, $g_{\lambda,\mu,\nu}$, of s_ν in $s_\lambda * s_\mu$ is equal to the multiplicity of the irreducible representation indexed by ν in the tensor product. In this paper we give an algorithm for expanding the Kronecker product $s_{(n-p,p)} * s_\lambda$ whenever $\lambda_1 - \lambda_2 \geq 2p$. As a consequence of this algorithm we obtain a formula for the coefficients $g_{\lambda,\mu,\nu}$ in terms of Littlewood-Richardson coefficients which does not involve cancellations. We also show that the coefficients in the expansion of $s_{(n-p,p)} * s_\lambda$ are stable. Moreover, we obtain a simple combinatorial interpretation for $g_{\lambda,(n-p,p),\nu}$ if λ is not a partition inside the $2(p-1) \times 2(p-1)$ square.

Introduction

Let χ^λ and χ^μ be the irreducible characters of S_n (the symmetric group on n letters) indexed by the partitions λ and μ of n. The *Kronecker product* $\chi^\lambda\chi^\mu$ is defined by $(\chi^\lambda\chi^\mu)(w) = \chi^\lambda(w)\chi^\mu(w)$ for all $w \in S_n$. Hence, $\chi^\lambda\chi^\mu$ is the character that corresponds to the diagonal action of S_n on the tensor product of the irreducible representations indexed by λ and μ. Then we have

$$\chi^\lambda\chi^\mu = \sum_{\nu\vdash n} g_{\lambda,\mu,\nu}\chi^\nu,$$

where $g_{\lambda,\mu,\nu}$ is the multiplicity of χ^ν in $\chi^\lambda\chi^\mu$. Hence the $g_{\lambda,\mu,\nu}$ are non-negative integers.

By means of the Frobenius map one can define the Kronecker (internal) product on the Schur symmetric functions by

$$s_\lambda * s_\mu = \sum_{\nu\vdash n} g_{\lambda,\mu,\nu}s_\nu.$$

A reasonable formula for decomposing the Kronecker product is unavailable, although the problem has been studied since the early twentieth century. In recent years Lascoux [La], Remmel [R], Remmel and Whitehead [RWd] and Rosas [Ro] derived closed formulas for

Kronecker products of Schur functions indexed by two row shapes or hook shapes. Gessel [Ge] obtained a combinatorial interpretation for zigzag partitions.

More general results include a formula of Garsia and Remmel [GR-1] which decomposes the product of homogeneous symmetric functions with a Schur function. Dvir [D] and Clausen and Meier [CM] have found bounds for the largest part and the maximal number of parts in a constituent of a product. Bessenrodt and Kleshchev [BK] have looked at the problem of determining when the decomposition of the Kronecker product has one or two constituents.

In 1937 Murnaghan [M] noticed that for large n the Kronecker product did not depend on the first part of the partitions λ and μ. That is, if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$ is a partition of n (written $\lambda \vdash n$) and $\bar{\lambda} = (\lambda_2, \ldots, \lambda_{\ell(\lambda)})$ denotes the partition obtained by removing the first part of λ, then there exists an n such that $g_{(n-|\bar{\lambda}|,\bar{\lambda}),(n-|\bar{\mu}|,\bar{\mu}),(n-|\bar{\nu}|,\bar{\nu})} = g_{(m-|\bar{\lambda}|,\bar{\lambda}),(m-|\bar{\mu}|,\bar{\mu}),(m-|\bar{\nu}|,\bar{\nu})}$ for all $m \geq n$. In this case we say that $g_{\lambda,\mu,\nu}$ is *stable*. Vallejo [V] has recently found a bound for n for the stability of $g_{\lambda,\mu,\nu}$. In this paper we show that $g_{(n-p,p),\lambda,\nu}$ is stable for all ν if $\lambda_1 - \lambda_2 \geq 2p$.

There is a simple algorithm for the decomposition of $s_{(n-1,1)} * s_\lambda$ whenever $\lambda_1 - \lambda_2 \geq 2$.

First Step: Everywhere possible delete zero or one box from $\bar{\lambda}$ such that the resulting diagram corresponds to a partition.

Second step: To each diagram $\beta \neq \bar{\lambda}$ obtained in the first step, everywhere possible add zero or one box so that the resulting diagram corresponds to a partition. And to $\beta = \bar{\lambda}$ add everywhere possible one box.

Finally, we complete the resulting diagrams $\bar{\nu}$ obtained in the second step such that $\nu = (n - |\bar{\nu}|, \bar{\nu})$ is a partition of n. Then $s_{(n-1,1)} * s_\lambda$ is equal to the sum of the Schur functions corresponding to all diagrams ν obtained via the remove/add steps above.

We generalize this algorithm for the Kronecker product $s_{(n-p,p)} * s_\lambda$ whenever $\lambda_1 - \lambda_2 \geq 2p$. We use the algorithm to obtain a close formula for $g_{\lambda,\mu,\nu}$ as well as bounds for the size of ν_1 and ν_2. Our main tools are the Garsia-Remmel identity [GR-1, Lemma 6.3] and the Remmel-Whitney algorithm for multiplying Schur functions [RWy].

We also give a combinatorial interpretation for the coefficient of s_ν in $s_{(n-p,p)} * s_\lambda$, if $\lambda_1 \geq 2p-1$ or $\ell(\lambda) \geq 2p-1$, in terms of what we call *Kronecker Tableaux*. In particular, our combinatorial interpretation holds for all λ if $n > (2p-2)^2$. Our analysis involves studying the Schur positivity of the symmetric function $s_{\lambda/\alpha} s_\alpha - s_{\lambda/\beta} s_\beta$, where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{\ell(\alpha)})$ with $\alpha_1 > \alpha_2$ and $\beta = (\alpha_1 - 1, \alpha_2, \ldots, \alpha_{\ell(\alpha)})$. We prove that this symmetric function is Schur positive if and only if $\lambda_1 \geq 2\alpha_1 - 1$. This result is then used to give a combinatorial interpretation for $g_{(n-p,p),\lambda,\nu}$ whenever λ is not a partition that fits in the $(2p-2) \times (2p-2)$ square.

Summary of results

1) The (modified) Remmel-Whitney algorithms.

The *reverse lexicographic filling* of μ, $rl(\mu)$, is a filling of the Young diagram μ with the numbers $1, 2, \ldots, |\mu|$ so that the numbers are entered in order from right to left and top to bottom.

Definition: A tableau T is (λ, μ)-*compatible* if it contains $|\lambda|$ unlabelled boxes and $|\mu|$ labelled boxes (with labels $1, 2 \ldots, |\mu|$) and all of the following conditions are satisfied:

(a) T contains $|\lambda|$ unlabelled boxes in the shape λ. They are positioned in the upper-left corner of T.

(b) The labelled boxes in T are in increasing order in each row from left to right and in each column from top to bottom. If one box of T is labelled, so are all the boxes in the same row that are to the right of it.

(c) If a box labelled $i+1$ occurs immediately to the left of the box labelled i in $rl(\mu)$, then in T the label $i+1$ occurs weakly above and strictly to the right of i.

(d) If the box labelled y occurs immediately below the box labelled x in $rl(\mu)$, then in T the label y occurs strictly below and weakly to the left of x.

Remmel and Whitney showed that $c^{\nu}_{\lambda\mu}$ is the number of (λ, μ)-compatible tableaux of shape ν [RWy].

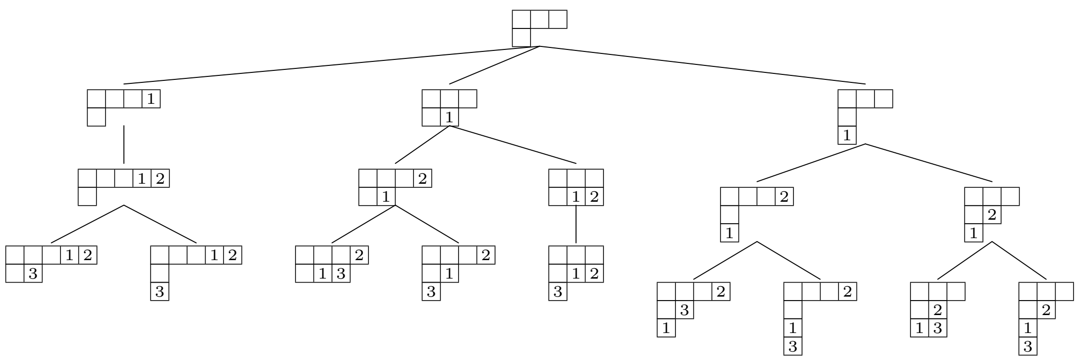
Multiplication: $s_\lambda s_\mu$ - Add$[\mu]$ to λ. Computing $s_\lambda s_\mu = \sum\limits_{|\nu|=|\lambda|+|\mu|} c^{\nu}_{\lambda\mu} s_\nu$:

(1) To the Young diagram λ add a box labelled 1 everywhere possible so that the rows are weakly increasing in size.

(2) We add each subsequent number so that, at each step, the conditions of the definition of (λ, μ)-compatible tableau are satisfied.

In this way we obtain a tree. The leaves of this tree are the elements of the multi-set Add$[\mu]$ to λ. They are the summands in the decomposition of $s_\lambda s_\mu$.

Example: The decomposition of $s_\lambda s_\mu$, where $\lambda = (3,1)$, $\mu = (2,1)$: $\lambda = \begin{array}{|c|c|c|}\hline \, & \, & \, \\ \hline \, & \multicolumn{2}{c}{} \\ \cline{1-1} \end{array}$ and $rl(\mu) = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}$.



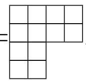
Hence $s_\lambda s_\mu = s_{(5,2)} + s_{(5,1,1)} + s_{(4,3)} + 2s_{(4,2,1)} + s_{(3,3,1)} + s_{(4,1,1,1)} + s_{(3,2,2)} + s_{(3,2,1,1)}$.

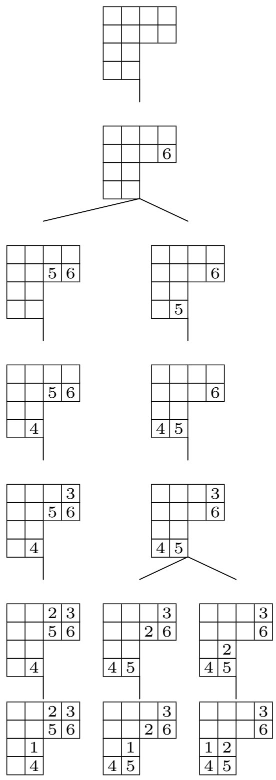
Skew: $s_{\lambda/\mu}$ **- Delete**$[\mu]$ **from** λ. Computing $s_{\lambda/\mu} = \sum_{|\nu|=|\lambda|-|\mu|} c^{\lambda}_{\mu\nu} s_\nu$:

(1) Form the reverse lexicographic filling of μ.

(2) Starting with the Young diagram λ we will label its outermost boxes with the numbers $1, 2, \ldots, |\mu|$ in decreasing order, starting with $|\mu|$, in the following way. At every step, the diagram obtained from λ by deleting the labelled boxes must be a Young diagram. Suppose the position (i, j) in $rl(\mu)$ is labelled x. If $j > 1$, let x^- be the label in position $(i, j-1)$ in $rl(\mu)$. If $i < \ell(\mu)$, let x^+ be the label in position $(i+1, j)$ in $rl(\mu)$. In λ, x will be placed to the left and weakly below (to the SW) of x^- and above and weakly to the right (to the NE) of x^+.

From each of the diagrams obtained (with $|\mu|$ labelled boxes) we remove all labelled boxes. The resulting diagrams are the elements in the multi-set Delete$[\mu]$ from λ. They are the summands in the decomposition of $s_{\lambda/\mu}$.

Example: The decomposition of $s_{\lambda/\mu}$, $\lambda = (4, 4, 2, 2)$, $\mu = (3, 3)$: $\lambda =$ , $rl(\mu) = \begin{array}{|c|c|c|}\hline 3&2&1\\ \hline 6&5&4\\ \hline\end{array}$.



Hence $s_{\lambda/\mu} = s_{(2,2,1,1)} + s_{(3,2,1)} + s_{(3,3)}$.

2) Algorithm for computing $s_{(n-p,p)} * s_\lambda$

If $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$, we denote by $\bar{\mu}$ the partition $\bar{\mu} = (\mu_2, \ldots, \mu_k)$. We will follow the philosophy of [M], and attempt to work with the partition $\bar{\mu}$ instead of μ whenever possible. Knowing that $\mu \vdash n$, μ_1 is completely determined by $\bar{\mu}$.

Let p be a positive integer and λ a partition of n such that $\lambda_1 - \lambda_2 \geq 2p$. We consider the subset of partitions of p contained in λ: $S_\lambda = \{\alpha \vdash p \,|\, \alpha \subseteq \lambda\}$.

Algorithm: For every $\alpha \in S_\lambda$ form the following set of Young diagrams:

$$Q(\alpha) = \bigcup_{j=0}^{\alpha_1} \{\nu |\, \nu \text{ is obtained by removing a horizontal strip with } j \text{ boxes from } \alpha\}$$

$$= \bigcup_{j=0}^{\alpha_1} \text{ Delete } [(j)] \text{ from } \alpha$$

For each $\alpha \in S_\lambda$ perform the following two steps:

(1) Remove$[\alpha]$**:** For each $\delta \in Q(\alpha)$ perform *Delete[δ] from* $\bar{\lambda}$. Record all diagrams obtained from *Delete[δ] from* $\bar{\lambda}$, with multiplicity, in the multi-set $D(\alpha)$. Denote by $d_{\alpha\lambda\beta}$ the multiplicity of β in $D(\alpha)$. If $\alpha_1 > \alpha_2$, let $D'(\alpha)$ be the submulti-set of $D(\alpha)$ of diagrams obtained by performing *Delete[δ] from* $\bar{\lambda}$ whenever $\delta_1 = \alpha_1$. Denote the multiplicity of $\beta \in D'(\alpha)$ by $d'_{\alpha\lambda\beta}$. If $\alpha_1 = \alpha_2$, set $d'_{\alpha\lambda\beta} = 0$.

(2) Add$[\alpha]$**:** For each (distinct) $\beta \in D(\alpha)$,

(a) If $d'_{\alpha\lambda\beta} = 0$, then for each $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ perform *Add[γ] to* β. The multiplicity of each resulting diagram is multiplied by $d_{\alpha\lambda\beta}$.

(b) If $0 < d'_{\alpha\lambda\beta} = d_{\alpha\lambda\beta}$, then for each $\gamma \in Q(\alpha)$ perform *Add[γ] to* β. The multiplicity of each resulting diagram is multiplied by $d_{\alpha\lambda\beta}$.

(c) If $0 < d'_{\alpha\lambda\beta} < d_{\alpha\lambda\beta}$, then for each $\gamma \in Q(\alpha)$ perform *Add[γ] to* β. For each $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ the multiplicity of each resulting diagram is multiplied by $d_{\alpha\lambda\beta}$. And for each γ such that $\gamma_1 < \alpha_1$ the multiplicity of each resulting diagram is multiplied by $d'_{\alpha\lambda\beta}$.

Finally, we record all diagrams obtained in step (2), for every β, in a multi-set R_α.

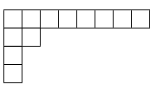
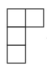
Note: Whenever we perform *Delete[η] from* η, the empty diagram, denoted ϵ, will be recorded. Thus, if $\alpha = (p)$, then $\epsilon \in Q(\alpha)$. Similarly, in the **Remove**$[\alpha]$ step, if $\delta = \bar{\lambda} \in Q(\alpha)$, then $\epsilon \in D(\alpha)$.

If $\eta = (\eta_1, \ldots, \eta_{\ell(\eta)}) \in R_\alpha$, let $\tilde{\eta} = (\eta_0, \eta_1, \ldots, \eta_{\ell(\eta)})$, where $\eta_0 = n - |\eta|$. Thus $\tilde{\eta} \vdash n$.

Theorem 1: Let p be a positive integer and λ a partition of n such that $\lambda_1 - \lambda_2 \geq 2p$. Then

$$s_{(n-p,p)} * s_\lambda = \sum_{\alpha \in S_\lambda} \sum_{\eta \in R_\alpha} s_{\tilde{\eta}}.$$

Example: We will perform the algorithm for $s_{(n-p,p)} * s_\lambda$ in the case when $n = 12$, $p = 3$ and $\lambda = (8, 2, 1, 1)$. Since $\lambda_1 - \lambda_2 = 8 - 2 = 6 \geq 2p$, the condition of the algorithm is satisfied. The Young diagrams for λ and $\bar{\lambda}$ are

$\lambda =$  and $\bar{\lambda} =$ .

We have $S_\lambda = \{\alpha \vdash 3 \mid \alpha \leq \lambda\} = \left\{ \ydiagram{3}\,, \ydiagram{2,1}\,, \ydiagram{1,1,1} \right\}$

$\boldsymbol{\alpha} = \ydiagram{3}$: From α remove j boxes, $0 \leq j \leq 3$, no two in the same column.

$$Q(\alpha) = \{ \ydiagram{3}\,, \ydiagram{2}\,, \ydiagram{1}\,, \epsilon\}$$

(1) Remove$[\alpha]$**:** For each $\delta \in Q(\alpha)$ perform *Delete[δ] from* $\bar{\lambda}$.

$Delete[\young(321)]$, $Delete[\young(21)]$, $Delete[\young(1)]$, and $Delete[\epsilon]\ from\ \ydiagram{2,1,1}$. Then we have

$$D(\alpha) = \left\{ \ydiagram{1,1}\,, \ydiagram{1,1,1}\,, \ydiagram{2,1}\,, \ydiagram{2,1,1} \right\} \quad \text{and} \quad D'(\alpha) = \emptyset.$$

(2) Add$[\alpha]$**:** Since $D'(\alpha) = \emptyset$, we have $d'_{\alpha\lambda\beta} = 0$ for all $\beta \in D(\alpha)$. We are in case (a). The only $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ is $\gamma = \ydiagram{3}$. For every $\beta \in D(\alpha)$ we perform $Add[\ydiagram{3}]\ to\ \beta$.

$Add[\young(321)]\ to\ \ydiagram{1,1} = \{(4,1), (3,1,1)\}$;

$Add[\young(321)]\ to\ \ydiagram{1,1,1} = \{(4,1,1), (3,1,1,1)\}$;

$Add[\young(321)]\ to\ \ydiagram{2,1} = \{(5,1), (4,2), (4,1,1), (3,2,1)\}$;

$Add[\young(321)]\ to\ \ydiagram{2,1,1} = \{(5,1,1), (4,2,1), (4,1,1,1), (3,2,1,1)\}$.

We take the union of these four multi-sets to get:

$R_{\ydiagram{3}} = \{(4,1), (3,1,1), 2(4,1,1), (3,1,1,1), (5,1), (4,2), (3,2,1), (5,1,1), (4,2,1), (4,1,1,1), (3,2,1,1)\}$

$\boldsymbol{\alpha} = \ydiagram{2,1}$: From α remove j boxes, $0 \leq j \leq 2$, no two in the same column.

$$Q(\alpha) = \left\{ \ydiagram{2,1}\,, \ydiagram{1,1}\,, \ydiagram{2}\,, \ydiagram{1} \right\}$$

(1) Remove$[\alpha]$**:** For each $\delta \in Q(\alpha)$ perform *Delete[δ] from* $\bar{\lambda}$.

$Delete[\young(21,3)]\ from\ \ydiagram{2,1,1} = \{(1)\}$; $\quad Delete[\young(1,2)]\ from\ \ydiagram{2,1,1} = \{(1,1), (2)\}$;

$Delete[\young(21)]\ from\ \ydiagram{2,1,1} = \{(1,1)\}$; $\quad Delete[\young(1)]\ from\ \ydiagram{2,1,1} = \{(2,1), (1,1,1)\}$.

This yields:

$$D(\alpha) = \left\{ \ydiagram{1}\,, 2\,\ydiagram{1,1}\,, \ydiagram{2}\,, \ydiagram{2,1}\,, \ydiagram{1,1,1} \right\} \quad \text{and} \quad D'(\alpha) = \left\{ \ydiagram{1}\,, \ydiagram{1,1} \right\}.$$

(2) Add$[\alpha]$**:** If $\beta = \ydiagram{1}$, then we have $d'_{\alpha\lambda\beta} = 1 = d_{\alpha\lambda\beta}$ and we are in case (b). For each $\gamma \in Q(\alpha)$ we perform *Add[*γ*] to* $\ydiagram{1}$.

$Add[\ytableaushort{21,3}]\ to\ \ydiagram{1} = \{(3,1),(2,2),(2,1,1)\}$; $Add[\ytableaushort{1,2}]\ to\ \ydiagram{1} = \{(2,1),(1,1,1)\}$;

$Add[\ytableaushort{21}]\ to\ \ydiagram{1} = \{(3),(2,1)\}$; $Add[\ytableaushort{1}]\ to\ \ydiagram{1} = \{(2),(1,1)\}$.

If $\beta = \ydiagram{1,1}$, then $d'_{\alpha\lambda\beta} = 1$ and $d_{\alpha\lambda\beta} = 2$. Thus we are in case (c).

For each $\gamma \in Q(\alpha)$ we perform *Add[*γ*] to* $\ydiagram{1,1}$ and if $\gamma_1 = \alpha_1$ count the resulting diagrams with multiplicity $d_{\alpha\lambda\beta} = 2$.

$2 \times Add[\ytableaushort{21,3}]\ to\ \ydiagram{1,1} = \{2(3,2),2(3,1,1),2(2,2,1),2(2,1,1,1)\}$;

$Add[\ytableaushort{1,2}]\ to\ \ydiagram{1,1} = \{(2,2),(2,1,1),(1,1,1,1)\}$;

$2 \times Add[\ytableaushort{21}]\ to\ \ydiagram{1,1} = \{2(3,1),2(2,1,1)\}$;

$Add[\ytableaushort{1}]\ to\ \ydiagram{1,1} = \{(2,1),(1,1,1)\}$.

If $\beta = \ydiagram{2}$, then $d'_{\alpha\lambda\beta} = 0$. We are in case (a). The only $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ are $\gamma = \ydiagram{2,1}$ and $\gamma = \ydiagram{2}$.

$Add[\ytableaushort{21,3}]\ to\ \ydiagram{2} = \{(4,1),(3,2),(3,1,1),(2,2,1)\}$;

$Add[\ytableaushort{21}]\ to\ \ydiagram{2} = \{(4),(3,1),(2,2)\}$;

If $\beta = \ydiagram{2,1}$, then $d'_{\alpha\lambda\beta} = 0$. We are in case (a). As before, the only $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ are $\gamma = \ydiagram{2,1}$ and $\gamma = \ydiagram{2}$.

$Add[\ytableaushort{21,3}]\ to\ \ydiagram{2,1} = \{(4,2),(4,1,1),(3,3),2(3,2,1),(3,1,1,1),(2,2,2),(2,2,1,1)\}$;

$Add[\ytableaushort{21}]\ to\ \ydiagram{2,1} = \{(4,1),(3,2),(3,1,1),(2,2,1)\}$.

If $\beta = \ydiagram{1,1,1}$, then $d'_{\alpha\lambda\beta} = 0$. We are in case (a). As before, the only $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ are $\gamma = \ydiagram{2,1}$ and $\gamma = \ydiagram{2}$.

$Add[\ytableaushort{21,3}]\ to\ \ydiagram{1,1,1} = \{(3,2,1),(3,1,1,1),(2,2,1,1),(2,1,1,1,1)\}$;

$Add[\ytableaushort{21}]\ to\ \ydiagram{1,1,1} = \{(3,1,1),(2,1,1,1)\}$.

We take the union of all the multi-sets above (from the Add step):

$$\begin{aligned} R_{\ydiagram{2,1}} = \{&4(3,1),3(2,2),4(2,1,1),3(2,1),2(1,1,1),(3),(2),(1,1),4(3,2),\\ &5(3,1,1),4(2,2,1),3(2,1,1,1),(1,1,1,1),2(4,1),(4),(4,2),(4,1,1),\\ &(3,3),3(3,2,1),2(3,1,1,1),(2,2,2),2(2,2,1,1),(2,1,1,1,1)\} \end{aligned}$$

$\boldsymbol{\alpha} = \ydiagram{1,1,1}$: From α remove j boxes, $0 \leq j \leq 1$, no two in the same column.

$$Q(\alpha) = \left\{ \ydiagram{1,1,1}\, , \ydiagram{1,1} \right\}$$

(1) Remove$[\alpha]$**:** For each $\delta \in Q(\alpha)$ perform *Delete[δ] from* $\bar{\lambda}$.

$Delete[\young(1,2,3)]\ from\ \ydiagram{2,1} = \{(1)\}$; $\qquad Delete[\young(1,2)]\ from\ \ydiagram{2,1} = \{(2), (1,1)\}$.

This yields:

$$D(\alpha) = \left\{ \ydiagram{1}\, , \ydiagram{1,1}\, , \ydiagram{2} \right\}.$$

(2) Add$[\alpha]$**:** Since $\alpha_1 = \alpha_2$, $d'_{\alpha\lambda\beta} = 0$ for all $\beta \in D(\alpha)$. We are in case (a). For $\alpha = (1,1,1)$, all $\gamma \in Q(\alpha)$ satisfy $\gamma_1 = \alpha_1$. We perform *Add[γ] to β* for all $\gamma \in Q(\alpha)$ and all $\beta \in D(\alpha)$.

$Add[\young(1,2,3)]\ to\ \ydiagram{1} = \{(2,1,1), (1,1,1,1)\}$; $\qquad Add[\young(1,2)]\ to\ \ydiagram{1} = \{(2,1), (1,1,1)\}$;

$Add[\young(1,2,3)]\ to\ \ydiagram{1,1} = \{(2,2,1), (2,1,1,1), (1,1,1,1,1)\}$; $\quad Add[\young(1,2)]\ to\ \ydiagram{1,1} = \{(2,2), (2,1,1), (1,1,1,1)\}$;

$Add[\young(1,2,3)]\ to\ \ydiagram{2} = \{(3,1,1), (2,1,1,1)\}$; $\qquad Add[\young(1,2)]\ to\ \ydiagram{2} = \{(3,1), (2,1,1)\}$.

We take the union of all the multi-sets above:

$$R_{\ydiagram{1,1,1}} = \{3(2,1,1), 2(1,1,1,1), (2,1), (1,1,1), (2,2,1),$$
$$2(2,1,1,1), (1,1,1,1,1), (2,2), (3,1,1), (3,1)\}$$

Finally, we use Theorem 1 to obtain the decomposition of $s_{(9,3)} * s_{(8,2,1,1)}$. Consider the union of the multi-sets R_α, for all $\alpha \in S_{(8,2,1,1)}$, and "complete" each shape to size 12.

Thus
$s_{(9,3)} * s_{(8,2,1,1)} = 3s_{(7,4,1)} + 7s_{(7,3,1,1)} + 3s_{(6,4,1,1)} + 3s_{(6,3,1,1,1)} + s_{(6,5,1)} + 2s_{(6,4,2)} + 4s_{(6,3,2,1)} + s_{(5,5,1,1)} + s_{(5,4,2,1)} + s_{(5,4,1,1,1)} + s_{(5,3,2,1,1)} + 5s_{(8,3,1)} + 4s_{(8,2,2)} + 7s_{(8,2,1,1)} + 4s_{(9,2,1)} + 3s_{(9,1,1,1)} + s_{(9,3)} + s_{(10,2)} + s_{(10,1,1)} + 4s_{(7,3,2)} + 5s_{(7,2,2,1)} + 5s_{(7,2,1,1,1)} + 3s_{(8,1,1,1,1)} + s_{(8,4)} + s_{(6,3,3)} + s_{(6,2,2,2)} + 2s_{(6,2,2,1,1)} + s_{(6,2,1,1,1,1)} + s_{(7,1,1,1,1,1)}$.

3) Multiplicities in the Kronecker Product

Denote by $c^{\mu}_{\nu\,\eta}$ the Littlewood-Richardson coefficient. If we denote by $T^{\eta}_{\mu/\nu}$ the set of the semistandard Young tableaux of shape μ/ν and type η whose reverse reading word is a lattice permutation, then the cardinality of $T^{\eta}_{\mu/\nu}$ is equal to $c^{\mu}_{\nu\,\eta}$. Let $T^{\eta}_{\mu/\nu}(i,j)$ be the subset of $T^{\eta}_{\mu/\nu}$ of SSYTs of shape μ/ν and type η with label 1 in position (i,j). Note that this multi-subset could be empty. Define

$$a^{\mu}_{\nu\,\eta} := \begin{cases} |T^{\eta}_{\mu/\nu}(2,\nu_1)|, & \text{if } \mu_2 \geq \nu_1 \text{ and } \nu_1 > \nu_2, \\ 0 & \text{otherwise.} \end{cases}$$

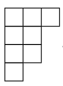
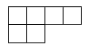
If $\beta = (\beta_1, \beta_2, \ldots, \beta_{\ell(\beta)}) \vdash m < n-p$, let $\hat{\beta} = (n-p-|\beta|, \beta_1, \beta_2, \ldots, \beta_{\ell(\beta)})$ be the partition of $n-p$ obtained from β by adding a first row of the correct size.

Theorem 2: Let n and p be positive integers such that $n \geq 2p$ and let λ be a partition of n with $\lambda_1 - \lambda_2 \geq 2p$. The multiplicity of s_ν in $s_{(n-p,p)} * s_\lambda$ is equal to

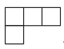
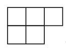
$$\sum_{\substack{\beta\subseteq\bar{\lambda},\,\beta\subseteq\bar{\nu} \\ |\beta|\geq n-\lambda_1-p}} \sum_{\substack{\alpha\vdash p \\ \alpha\subseteq\lambda}} \left(\sum_{\substack{\gamma\in Q(\alpha) \\ \gamma_1=\alpha_1,\,\gamma\subseteq\bar{\nu} \\ |\gamma|=|\bar{\nu}|-|\beta|}} c^{\lambda}_{\alpha\hat{\beta}} c^{\bar{\nu}}_{\beta\gamma} + \sum_{\substack{\gamma\in Q(\alpha) \\ \gamma_1<\alpha_1,\,\gamma\subseteq\bar{\nu} \\ |\gamma|=|\bar{\nu}|-|\beta|}} a^{\lambda}_{\alpha\hat{\beta}} c^{\bar{\nu}}_{\beta\gamma} \right).$$

Example: We use the above theorem to determine the multiplicity of $s_{(13,4,2)}$ in the Kronecker product $s_{(15,4)} * s_{(11,3,2,2,1)}$.

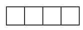
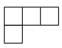
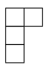
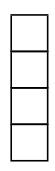
We have $n = 19$, $p = 4$, $\bar{\lambda} = (3,2,2,1)$ and $\bar{\nu} = (4,2)$, i.e

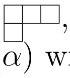
$\bar{\lambda} =$ , $\bar{\nu} =$ .

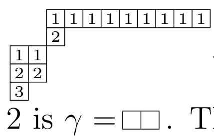
Since $n - \lambda_1 - p = 19 - 11 - 4 = 4$, the first summation in the formula of Theorem 2 runs over all Young diagrams β such that $|\beta| \geq 4$, $\beta \subseteq \bar{\lambda}$ and $\beta \subseteq \bar{\nu}$. Thus β has at most two rows: $\beta = (\beta_1, \beta_2)$ with $\beta_1 \leq 3$ and $\beta_2 \leq 2$. The possible β's in the first summation are

, , .

The second summation runs over all Young diagrams α of size $p = 4$ with $\alpha \subseteq \lambda$. They are the elements of

$S_\lambda = \{$ , , , ,  $\}$.

(1) If $\beta =$ , then $\hat{\beta} = (11,3,1) \vdash n-p = 15$. For each α, the inner sums will run over all $\gamma \in Q(\alpha)$ with $|\gamma| = |\bar{\nu}| - |\beta| = 6 - 4 = 2$.

If $\alpha =$, then the only SSYT of shape λ/α and type $\hat{\beta} = (11,3,1)$ is . Thus $c^{\lambda}_{\alpha\,\hat{\beta}} = 1$ and, since $\alpha_1 = \alpha_2$, $a^{\lambda}_{\alpha\,\hat{\beta}} = 0$. The only $\gamma \in Q(\alpha)$ with $|\gamma| = 2$ is $\gamma =$ . There is one SSYT of shape $\bar{\nu}/\beta$ and type $\gamma = (2)$: $\ytableaushort{\none\none 1, 1}$. Therefore $c^{\bar{\nu}}_{\beta\,\gamma} = 1$. Hence, $c^{\lambda}_{\alpha\,\hat{\beta}} c^{\bar{\nu}}_{\beta\,(2)} = 1$ *This contributes* 1 *to the multiplicity.*

If $\alpha = \ydiagram{2,1,1}$, then the only SSYT of shape λ/α and type $\hat{\beta} = (11, 3, 1)$ is $\ytableaushort{\none\none 111111111, \none 12, \none 2, 13, 2}$. Thus $c^{\lambda}_{\alpha\,\hat{\beta}} = a^{\lambda}_{\alpha\,\hat{\beta}} = 1$. The only $\gamma \in Q(\alpha)$ with $|\gamma| = 2$ is $\gamma = \ydiagram{1,1}$. There is one SSYT of shape $\bar{\nu}/\beta$ and type $\gamma = (1, 1)$: $\ytableaushort{\none\none 1, 2}$. Therefore $c^{\bar{\nu}}_{\beta\,\gamma} = 1$. Hence, $c^{\lambda}_{\alpha\,\hat{\beta}} c^{\bar{\nu}}_{\beta\,(1,1)} = 1$. *This contributes* 1 *to the multiplicity.*

For all other $\alpha \in S_\lambda$ we have $c^{\lambda}_{\alpha\,\hat{\beta}} = a^{\lambda}_{\alpha\,\hat{\beta}} = 0$. Hence, they do not contribute to the multiplicity.

(2) If $\beta = \ydiagram{3,2}$, then $\hat{\beta} = (10, 3, 2) \vdash n - p = 15$. For each α, the inner sums will run over all $\gamma \in Q(\alpha)$ with $|\gamma| = |\bar{\nu}| - |\beta| = 6 - 5 = 1$.

If $\alpha = \ydiagram{4}$ then $c^{\lambda}_{\alpha\,\hat{\beta}} = a^{\lambda}_{\alpha\,\hat{\beta}} = 0$.

If $\alpha = \ydiagram{3,1}$, $\ytableaushort{\none\none\none 11111111, \none 12, 12, 23, 3}$ is the only SSYT of shape λ/α and type $\hat{\beta} = (10, 3, 2)$. Thus $c^{\lambda}_{\alpha\,\hat{\beta}} = 1$ and $a^{\lambda}_{\alpha\,\hat{\beta}} = 0$. Since $\alpha_1 = 3$, there is no $\gamma \in Q(\alpha)$ with $\gamma_1 = \alpha_1$ and $|\gamma| = 1$.

If $\alpha = \ydiagram{2,2}$, $\alpha = \ydiagram{2,1,1}$ or $\alpha = \ydiagram{1,1,1,1}$, there is no $\gamma \in Q(\alpha)$ with $|\gamma| = 1$.

(3) Finally, if $\beta = \ydiagram{2,2}$, then $\hat{\beta} = (11, 2, 2) \vdash n - p = 15$. For each α, the inner sums will run over all $\gamma \in Q(\alpha)$ with $|\gamma| = |\bar{\nu}| - |\beta| = 6 - 4 = 2$.

If $\alpha = \ydiagram{3,1}$, $\ytableaushort{\none\none\none 11111111, \none 11, 12, 23, 3}$ is the only SSYT of shape λ/α and type $\hat{\beta} = (11, 2, 2)$. Thus $c^{\lambda}_{\alpha\,\hat{\beta}} = a^{\lambda}_{\alpha\,\hat{\beta}} = 1$. The shapes $\gamma \in Q(\alpha)$ with $|\gamma| = 2$ are $\gamma = \ydiagram{2}$ and $\gamma = \ydiagram{1,1}$. There is exactly one SSYT of shape $\bar{\nu}/\beta$ and type $\gamma = (2)$. Thus, for $\gamma = (2)$, $c^{\bar{\nu}}_{\beta\,\gamma} = 1$. Hence, $c^{\lambda}_{\alpha\,\hat{\beta}} c^{\bar{\nu}}_{\beta\,(2)} = 1$. *This contributes* 1 *to the multiplicity.* We also have $c^{\bar{\nu}}_{\beta\,(1,1)} = 0$

If $\alpha = \ydiagram{2,2}$, then $\ytableaushort{\none\none 111111111, \none\none 2, 11, 23, 3}$ is the only SSYT of shape λ/α and type $\hat{\beta} = (11, 2, 2)$. Thus $c^{\lambda}_{\alpha\,\hat{\beta}} = 1$ and, since $\alpha_1 = \alpha_2$, $a^{\lambda}_{\alpha\,\hat{\beta}} = 0$. The only $\gamma \in Q(\alpha)$ with $|\gamma| = 2$ (and $\gamma_1 = \alpha_1$) is $\gamma = \ydiagram{2}$. As before, there is one SSYT of shape $\bar{\nu}/\beta$ and type $\gamma = (2)$. Therefore $c^{\bar{\nu}}_{\beta\,(2)} = 1$. Hence, $c^{\lambda}_{\alpha\,\hat{\beta}} c^{\bar{\nu}}_{\beta\,(2)} = 1$. *This contributes* 1 *to the multiplicity.*

If $\alpha = \ydiagram{2,1,1}$, then $\ytableaushort{\none\none 111111111, \none 12, \none 2, 13, 3}$ is the only SSYT of shape λ/α and type $\hat{\beta} = (11, 2, 2)$. Thus $c^{\lambda}_{\alpha\,\hat{\beta}} = a^{\lambda}_{\alpha\,\hat{\beta}} = 1$. The only $\gamma \in Q(\alpha)$ with $|\gamma| = 2$ is $\gamma = \ydiagram{1,1}$. However, $c^{\bar{\nu}}_{\beta\,(1,1)} = 0$.

For all other $\alpha \in S_\lambda$ we have $c^{\lambda}_{\alpha\,\hat{\beta}} = a^{\lambda}_{\alpha\,\hat{\beta}} = 0$.

Therefore the multiplicity of $s_{(13,4,2)}$ in $s_{(15,4)} * s_{(11,3,2,2,1)}$ equals 4.

Proposition 3: Let n and p be positive integers with $n \geq 2p$ and let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$ be a partition of n with $\lambda_1 - \lambda_2 \geq 2p$. Consider the partition $\nu = (\nu_1, \nu_2, \ldots, \nu_{\ell(\nu)})$ of n. If the multiplicity $g_{(n-p,p),\lambda,\nu}$ of s_ν in $s_{(n-p,p)} * s_\lambda$ is non-zero, then $\lambda_1 - p \leq \nu_1 \leq \lambda_1 + p$. Moreover, if $\lambda_2 < p$ and $g_{(n-p,p),\lambda,\nu} \neq 0$, then $\lambda_1 - p \leq \nu_1 \leq \lambda_1 + \lambda_2$.

Proposition 4: Let n and p and $\lambda \vdash n$ be as in the previous proposition, i.e. $\lambda_1 - \lambda_2 \geq 2p$. Consider the partition $\nu = (\nu_1, \nu_2, \ldots, \nu_{\ell(\nu)})$ of n. If $\nu_2 > \lambda_2 + p$, then the multiplicity $g_{(n-p,p),\lambda,\nu}$ of s_ν in $s_{(n-p,p)} * s_\lambda$ is equal to zero. Moreover, if $\nu = (\lambda_1 - p, \lambda_2 + p, \lambda_3, \ldots, \lambda_{\ell(\lambda)})$, then $g_{(n-p,p),\lambda,\nu} = 1$.

If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}) \vdash n$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}) \vdash m$, we say that λ is less than μ *in lexicographic order*, and write $\lambda <_l \mu$, if there is a non-negative integer k such that $\lambda_i = \mu_i$ for all $i = 1, 2, \ldots, k$ and $\lambda_{k+1} < \mu_{k+1}$. Note that the lexicographic order is a total order on the set of all partitions.

Corollary 5: Let n and p be positive integers such that $n \geq 2p$ and let $\lambda \vdash n$ such that $\lambda_1 - \lambda_2 \geq 2p$. The smallest partition in lexicographic order $\nu \vdash n$ such that s_ν appears in the decomposition of $s_{(n-p,p)} * s_\lambda$ is the partition whose parts are $\lambda_1 - p, \lambda_2, \ldots, \lambda_{\ell(\lambda)}, p$, reordered to form a partition. Moreover, this s_ν appears with multiplicity 1.

4) Stability of Kronecker coefficients

Theorem 6: Given an arbitrary partition $\bar{\lambda} = (\lambda_2, \lambda_3, \ldots, \lambda_{\ell(\lambda)})$, let n be a positive integer such that $n \geq 2p + |\bar{\lambda}| + \lambda_2$. Then $g_{(n-p,p),(n-|\bar{\lambda}|,\bar{\lambda}),(n-|\bar{\nu}|,\bar{\nu})} = g_{(m-p,p),(m-|\bar{\lambda}|,\bar{\lambda}),(m-|\bar{\nu}|,\bar{\nu})}$ for all $m \geq n$ and all partitions $\nu \vdash n$.

5) Combinatorial interpretation of the Kronecker coefficients

A SSYT T of shape λ/α and type $\nu - \alpha$ whose reverse reading word is an α-lattice permutation (i.e. in any initial factor $a_1 a_2 \cdots a_j$, $1 \leq j \leq n$, the number of $i's + \alpha_i \geq$ the number of $(i+1)'s + \alpha_{i+1}$) is called a *Kronecker Tableau* of shape λ/α and type $(\nu - \alpha)$ if

(I) $\alpha_1 = \alpha_2$ or

(II) $\alpha_1 > \alpha_2$ and any one of the following two conditions is satisfied:

- (i) The number of 1's in the second row of λ/α is exactly $\alpha_1 - \alpha_2$.
- (ii) The number of 2's in the first row of λ/α is exactly $\alpha_1 - \alpha_2$.

Denote by $k^\lambda_{\alpha\nu}$ the number of Kronecker tableaux of shape λ/α and type $\nu - \alpha$.

Theorem 7: Let n and p be positive integers such that $n \geq 2p-1$. Let $\lambda = (\lambda_1, \ldots, \lambda_{\ell(\lambda)}) \vdash n$

such that $\lambda_1 \geq 2p-1$. If ν is a partition of n, the multiplicity of s_ν in $s_{(n-p,p)} * s_\lambda$ equals

$$\sum_{\substack{\alpha \vdash p \\ \alpha \subseteq \lambda}} k^{\lambda}_{\alpha\nu},$$

where $\alpha \subseteq \lambda$ means $\ell(\alpha) \leq \ell(\lambda)$ and $\alpha_i \leq \lambda_i$ for all $1 \leq i \leq \ell(\alpha)$.

References

[BK] Bessenrodt, C., Keleshchev, A.; "On Kronecker products of Complex Representations of the Symmetric and Alternating groups", Pacific J. of Math. Vol **190** 2, 1999.

[CM] Clausen, M., Meier, H.; "Extreme irreduzible Konstituenten in Tensordarstelhungen symmetrischer Gruppen", *Bayreuther Math. Schriften.* **45** (1993), 1-17.

[D] Dvir, Y.; "On the Kronecker product of S_n characters", *J. Algebra* **154** (1993), 125-140.

[GR-1] Garsia, A.M., Remmel, J., "Shuffles of Permutations and the Kronecker Product", *Graphs and Combinatorics* **1**, 217-263 (1985).

[Ge] Gessel, I.M; "Multipartite P-partitions and inner products of Schur functions", *Contemp. Math.* 1984, pp. 289-302.

[La] Lascoux, A.; "Produit de Kronecker des representations du group symmetrique", *Lecture Notes in Mathematics* 1980, **795**, Springer Verlag pp. 319-329.

[M] Murnaghan, F.D., "The Analysis of the Kronecker Product of Irreducible Representation of the Symmetric Group", *American Journal of Mathematics*, Vol. 60. No. 3, 761-784 (1938).

[R] Remmel, J., "A Formula for the Kronecker Products of Schur Functions of Hook Shapes", *Jornal of Algebra* **120**, 100-118 (1989).

[RWd] Remmel, J.,Whitehead, T.; "On the Kronecker product of Schur functions of two row shapes", *Bull. Belg. Math. Soc.* 1, 1994, pp. 649-683.

[RWy] Remmel, J., Whitney, R. "Multiplying Schur functions", *J. Algorithms*, **5**, 471-487, (1984).

[Ro] Rosas, M. H.; "The Kronecker product of Schur functions indexed by two-row shapes or hook shapes", *J. Algebraic Combin.* 14 (2001), no. 2, 153–173.

[V] Vallejo, E.; "Stability of the Kronecker products or irreducible characters of the symmetric group" *The Elec. J. of Comb.* **6** (1999).

A distributive lattice structure on noncrossing partitions*

Elena Barcucci† Antonio Bernini† Luca Ferrari‡ Maddalena Poneti†

Abstract

In [FP2] a natural order on Dyck paths of any fixed length inducing a distributive lattice structure is defined. We transfer this order on noncrossing partitions along a well-known bijection [S], thus showing that noncrossing partitions can be endowed with a distributive lattice structure having some combinatorial relevance. Finally we prove that our lattices are isomorphic to the posets of 312-avoiding permutations with the order induced by the strong Bruhat order of the symmetric group.

1 Introduction

Every paper dealing with Catalan numbers contains a sentence somehow like the following: "In [S2] Stanley gives 66 different combinatorial interpretations of Catalan numbers". Indeed, exercise 6.19 is maybe the best source of information on the Catalan family, at least from a purely enumerative point of view. A further step should be to consider some interesting order structures on the objects of the Catalan family and try to compare them. What we would like to do in the present paper is a first instance of this program.

We start by considering noncrossing partitions. They can be endowed with the refinement order, so to obtain the well-known noncrossing partition lattices, first studied by Kreweras [Kre], which have been proved very useful in several, different contexts. These lattices possess many interesting properties, however they are not distributive (actually not even modular). Is there the possibility of defining some *interesting* distributive lattice structure on noncrossing partitions? We claim that the answer is affirmative by explicitly finding an order on noncrossing partitions which is isomorphic to at least two combinatorially meaningful distributive lattices.

We first consider Dyck paths and define an order on them as follows: given two Dyck paths P, Q of the same length, we say that $P \leq Q$ when P entirely lies below Q (possibly coinciding with Q in some points). It is possible to show [FP2] that the set of Dyck paths of any given length endowed with this order is a distributive lattice. These *Dyck lattices* are not so well known; they have been studied first in [FP2] (following some general ideas of Narayana [N]), and in [CJ] the authors show their importance in the study of some matters related with Temperley-Lieb algebras. Our idea is to transport such a structure on noncrossing partitions along a famous bijection (see [S]). We have called *Bruhat noncrossing partition lattices* the distributive lattices of noncrossing partitions arising in this way; section 3 is devoted to the study of some properties of these lattices. Moreover,

*This work was partially supported by MIUR project: *Linguaggi formali e automi: metodi, modelli e applicazioni.*

†Dipartimento di Sistemi e Informatica, viale Morgagni 65, 50134 Firenze, Italy barcucci@dsi.unifi.it bernini@dsi.unifi.it poneti@dsi.unifi.it

‡Dipartimento di Scienze Matematiche ed Informatiche, Pian dei Mantellini, 44, 53100, Siena, Italy ferrari@math.unifi.it

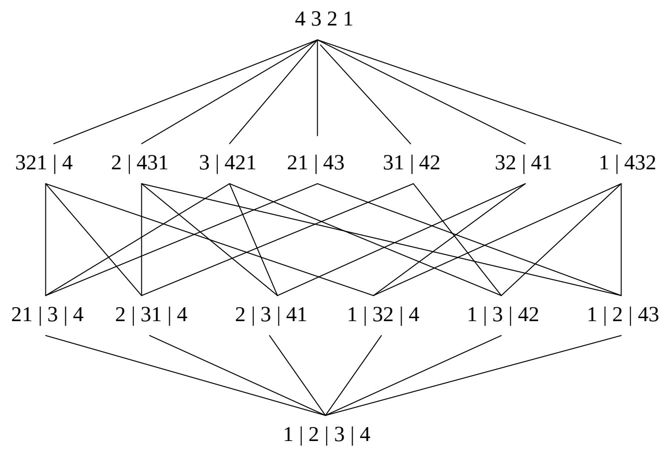


Figure 1: $\Pi(4)$.

Bruhat noncrossing partition lattices turn out to be isomorphic to an even more interesting class of lattices. It is not difficult to explicitly find a trivial bijection between noncrossing partitions and 312-avoiding permutations. More precisely, we show that such a bijection is an order-isomorphism between the Bruhat lattice of noncrossing partitions of an n set and the class $S_n(312)$ of 312-avoiding permutations of an n set endowed with the (strong) Bruhat order. As a byproduct, we have that $S_n(312)$ is a distributive sublattice of the symmetric group of order n with the Bruhat order. These results are contained in section 4, where we also find a criterion to determine the meet and the join of two 312-avoiding permutations in $S_n(312)$. To the best of our knowledge, the only paper dealing with this kind of matters is [P], where the author determines the Bruhat posets (arising from Weyl groups) which are lattices. However, the language and the aims of [P] are totally different from the ones of our approach. It would be interesting to compare our results with those of Proctor. However, it seems to us that our result is the first one concerning the order structure induced by the Bruhat order on a class of pattern-avoiding permutations.

The final part of this introduction is devoted to the explanation of the main notations we use through the paper and to the presentation of the basics of some general theories we refer to in the next pages.

The set (and the lattice) of partitions of $[n] = \{1, 2, \ldots, n\}$ will be denoted by $\Pi(n)$. If $\pi \in \Pi(n)$, we will always use the notation $\pi = B_1|B_2|\ldots|B_k$, where the B_i's are the blocks of π, the elements inside each block are in decreasing order and $\max B_i < \max B_j$, for $i < j$. Given $\pi, \rho \in \Pi(n)$, define $\pi \leq \rho$ when every block of π is contained into some block of ρ. The many properties of this classical order can be found in several textbooks, such as [S1, A]. Here we only mention that $\Pi(n)$ endowed with this *refinement order* is a lattice which is neither distributive nor modular. Nevertheless, it possesses a rank function: the rank of $\pi = B_1|B_2|\ldots|B_k$ is $n - k$. The Whitney numbers of the partition lattices are the well-known Stirling numbers of the second kind. The Hasse diagram of $\Pi(4)$ is shown in Figure 1.

We will often deal with Dyck paths and, depending on the context, we will find convenient to describe them in several different ways. Therefore a Dyck path will be alternatively described as a particular lattice path in the discrete plane $\mathbf{N} \times \mathbf{N}$ (and denoted by capital letters like $P, Q, R, \ldots$)

or as a function $f : \mathbf{N} \longrightarrow \mathbf{N}$ satisfying certain properties (and denoted by lowercase letters like $f, g, h, \ldots$) or else as a particular word of the two-letter alphabet $\{U, D\}$ (and denoted by Greek letters such as $\omega(U, D), \psi(U, D), \ldots$). We leave to the reader the details of the descriptions of Dyck paths we have sketched in the previous sentence.

In section 4 we make use of the concept of (generalized) pattern-avoiding permutation. Let $\pi = \pi_1\pi_2 \cdots \pi_n$ and $\sigma = \sigma_1\sigma_2 \cdots \sigma_k$ be two permutations of $[n]$ and $[k]$, respectively, with $k \leq n$. The permutation π *avoids the pattern* σ if there exist no indexes $i_1 < i_2 < \cdots < i_k$ such that $\pi_{i_1}\pi_{i_2} \ldots \pi_{i_k}$ is in the same relative order as $\sigma_1\sigma_2 \ldots \sigma_k$. The permutation π is called σ*-avoiding* and the subset of σ-avoiding permutations of S_n is denoted $S_n(\sigma)$. A huge amount of papers can be found dealing with pattern avoidance, see for instance [SS, F, Kra]. In [BS] the authors introduced *generalized patterns* for the study of Mahonian statistics on permutations. A generalized pattern is a permutation $\sigma \in S_k$ equipped with a dash between two of its elements (e.g. $1-32$ and $23-1$ are generalized patterns of length 3) and a permutation π contains a generalized pattern when adjacent elements in the generalized pattern correspond to adjacent elements in π. Classes of generalized pattern avoiding permutations has been widely studied in recent years (see [BS, BFP, C, CM], to cite a very few).

2 Noncrossing partitions and Dyck paths

A partition of $1, 2, \ldots, n$ is *noncrossing* when, given four elements, $1 \leq a < b < c < d \leq n$, such that a, c are in the same block and b, d are in the same block, then the two blocks coincide. The set of all noncrossing partitions of an n-set will be denoted $NC(n)$. We refer the reader to the fairly complete survey [S] and to the references therein for the plentiful applications of this notion.

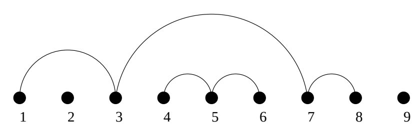


Figure 2: The noncrossing partition $2|654|8731|9 \in NC(9)$.

The refinement order can be restricted to noncrossing partitions: what we obtain is again a lattice, which is usually referred to as the *noncrossing partition lattice*. Among the main features of these lattices we recall here that they are not distributive and the lattice operations are different from those of the partition lattices (the join of two noncrossing partitions needs not be noncrossing within the full partition lattice).

Noncrossing partitions are enumerated by Catalan numbers, so, as it often happens, it is possible to find a bijection with Dyck paths. The nice bijection we are going to describe can also be found, for instance, in [D, S]. Fix a Dyck path and label its up steps by enumerating them from left to right (so that the k-th up step is labelled k). Next assign to each down step the same label of the up step it is matched with. Now consider the partition whose blocks are constituted by the labels of each sequence of consecutive down steps. Such a partition is easily seen to be noncrossing. In Figure 3 we have illustrated this bijection on a concrete example; the bold labels next to the down steps are the elements of the corresponding noncrossing partition, whereas the up steps are simply labelled in increasing order.

Now denote with $\mathcal{D}_n$ the set of Dyck paths of length $2n$. It is possible to define a natural order on $\mathcal{D}_n$ by setting $f \leq g$ whenever $f(n) \leq g(n)$, for every $n \in \mathbf{N}$. This means that $f \leq g$ when f "lies

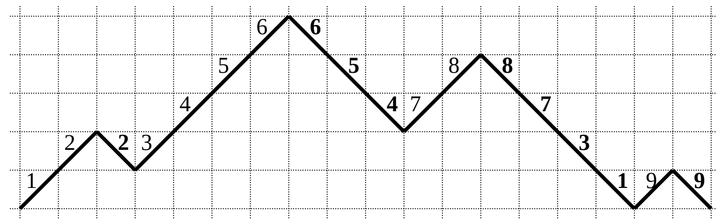


Figure 3: The Dyck path associated with 2|654|8731|9.

weakly" below g. The set $\mathcal{D}_n$, endowed with such an order, turns out to be a distributive lattice, which has been studied in some detail in [FP2] under the name of *Dyck lattice (of order n)*. We point out that Dyck lattices have also been considered in [CJ], where the authors speak of *geometric inclusion* of paths.

Our idea is to transfer the order structure of Dyck lattices along the above described bijection. In this way we define a new order on noncrossing partitions. The distributive lattices so obtained will be called *Bruhat noncrossing partition lattices.* The reason of this name, which is at present rather obscure, will become clear in the last section. Our main goal is to give a satisfactory description of such lattices.

3 The Bruhat noncrossing partition lattice

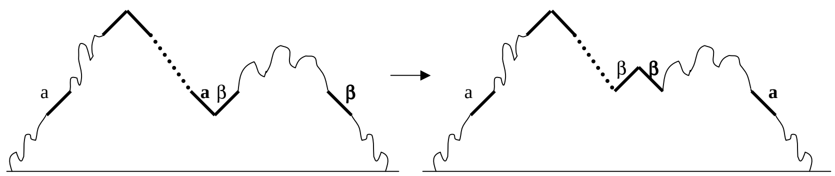
In the rest of the paper it is tacitly assumed that noncrossing partitions are endowed with the Bruhat order.

Given two noncrossing partitions π, ρ we look for some condition to recognize if $\pi \prec \rho$ or not. The following theorem gives a precise answer to this problem.

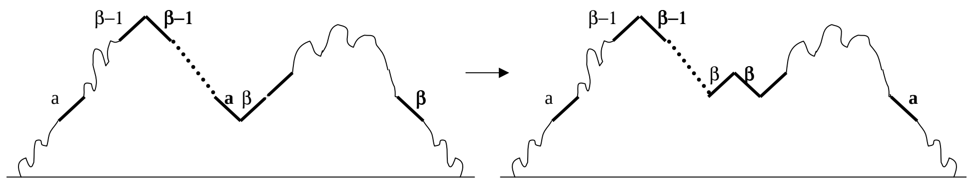
Theorem 3.1 *(Characterization of coverings) Given two noncrossing partitions $\pi, \rho \in NC(n)$, we have $\pi \prec \rho$ if and only if ρ is obtained from π by moving the minimum of some block B of π into the block $\tilde{B}$ containing the element $\beta = \max B + 1$ and either*

1. *keeping β inside $\tilde{B}$, if $\beta = \max \tilde{B}$, or*
2. *adding a new block $\overline{B} = \{\beta\}$, if $\beta \neq \max \tilde{B}$.*

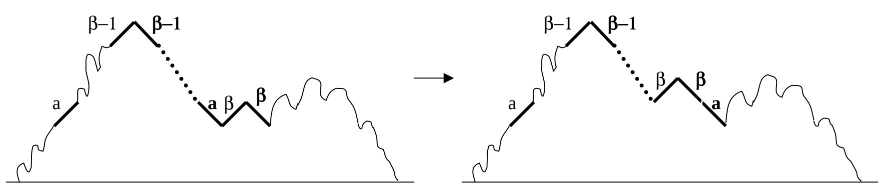
Proof. Suppose that P_π, P_ρ are the Dyck paths associated with π, ρ, respectively. The fact that $P_\pi \prec P_\rho$ in $\mathcal{D}_n$ means that P_ρ is obtained from P_π by replacing a valley with a peak. In the context of noncrossing partitions this amounts to moving the minimum a of a block, since the down step of a valley is the last step of a descent. The element a is moved into the block containing the element corresponding to the down step matched with the up step of the valley. It follows directly from the above bijection that such a down step has label equal to $\beta = \max B + 1$, where B is the block containing a in π. The following figure illustrates these facts.



Now, what happens with the element β? There are essentially two different cases. If the up step of the valley in P_π is followed by another up step, then β is not the maximum of its block in π, and it is easy to check that in ρ it becomes a singleton block (since in P_ρ the corresponding step is preceded and followed by up steps).



If the up step of the valley is followed by a down step, then β is the maximum of its block in π, and it remains in the same block also in ρ, as illustrated in the next figure.



■

Example. Given the partition $2|54|631 \in NC(6)$, there are precisely two partitions covering it, which are $3|54|621$ (2 is moved and 3 is not the maximum of its block) and $2|5|6431$ (4 is moved and 6 is the maximum of its block).

It is interesting to observe that the two "instructions" 1. and 2. in the previous theorem have a striking analogy with the definition of a *filler point* given in [DS]. Indeed, a filler point is produced whenever a valley preceded by an up step is changed into a peak in the associated Dyck path. Thus a filler point in a noncrossing partition corresponds to a down step preceded by a long ascent in the associated Dyck path (where a long ascent is a sequence of two or more consecutive up steps). Therefore, the number of noncrossing partitions of an n-set having k filler points coincides with the number $T_{n,k}$ of Dyck paths of length $2n$ having k long ascents, namely (see [Sl]):

$$T_{n,k} = \frac{1}{n+1}\binom{n+1}{k}\sum_{j=0}^{n-2k}\binom{k+j-1}{k-1}\binom{n+1-k}{n-2k-j}.$$

Our next result is a criterion to compare two given noncrossing partitions. In order to properly state it, we need to introduce a technical definition. Consider a noncrossing partition $\pi \in NC(n)$. We define the *max-vector* of π to be the vector $\max(\pi) = (\mu_1, \ldots, \mu_n)$ such that μ_i is the maximum of the first i elements of π. So, for instance, if $\pi = 2|31|54$, then $\max(\pi) = (2, 3, 3, 5, 5)$. We invite the reader to check that the max-vector uniquely determines its associated noncrossing partition. This fact will be very important in the sequel.

Theorem 3.2 *(Characterization of the Bruhat order of NC) Let $\pi, \rho \in NC(n)$. Then $\pi \leq \rho$ if and only if $\max(\pi) \leq \max(\rho)$ in the coordinatewise order.*

Proof. Let $\omega_1 = \omega_1(U, D)$ and $\omega_2 = \omega_2(U, D)$ be the two Dyck paths corresponding to π and ρ, respectively. Then it is clear that $\omega_1 \leq \omega_2$ if and only if every prefix of ω_1 contains at least as many D's as the corresponding prefix of ω_2. This can be translated on partitions using max-vectors. Indeed, if $\max(\pi) = (\mu_1, \ldots, \mu_n)$ and $\max(\rho) = (\nu_1, \ldots, \nu_n)$, consider the two vectors $(\overline{\mu_1}, \ldots, \overline{\mu_n})$ and $(\overline{\nu_1}, \ldots, \overline{\nu_n})$, where $\overline{\mu_i} = \mu_i + i$ and $\overline{\nu_i} = \nu_i + i$. Then, it is not difficult to observe that $\overline{\mu_i}$ and $\overline{\nu_i}$ encode the position of the i-th D in the corresponding Dyck path. From the hypotheses, we have that the i-th D of ω_1 occurs before the i-th D of ω_2, and so $\overline{\mu_i} \leq \overline{\nu_i}$. Since this holds for every $i \leq n$, the thesis follows. ■

Example. Let $\pi = 2|43|51|6$, $\rho = 43|52|61 \in NC(6)$. We easily find $\max(\pi) = (2, 4, 4, 5, 5, 6)$ and $\max(\rho) = (4, 4, 5, 5, 6, 6)$. It is immediate to see that $\max(\pi) \leq \max(\rho)$, whence $\pi \leq \rho$.

Remark. Observe that, if $\pi \prec \rho$, then $\max(\pi)$ and $\max(\rho)$ differ precisely in one position.

It is known [FP2] that Dyck lattices possess a rank function (simply because they are distributive lattices) which is essentially given by the *area* bounded by a Dyck path and the x-axis. More precisely, if $A(P)$ is the area of a Dyck path P of length n, then the rank of P inside its Dyck lattice is given by $r(P) = \frac{A(P)-n}{2}$. Our next goal is to translate the parameter "area under Dyck paths" into a parameter on noncrossing partitions, in order to define a rank on the Bruhat noncrossing partition lattices.

Our first result is a formula for the area of Dyck paths in terms of its peaks and valleys. Since we have not found such a formula in the literature, we also propose a proof for the reader's convenience.

Lemma 3.1 *Let P be a Dyck path. Let p_i and v_j denote the height of the i-th peak and the j-th valley of P, respectively. Then*

$$A(P) = \sum_i (p_i^2 - v_i^2). \tag{1}$$

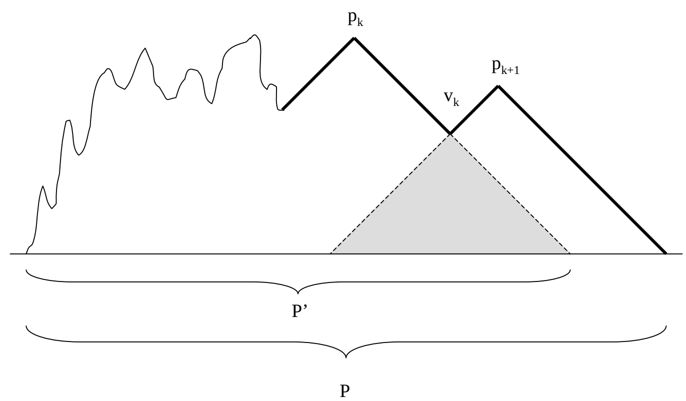


Figure 4: How P' is obtained from P.

Proof. We proceed by induction on the number of peaks. If a Dyck path P has only one peak, then it is the maximum of its Dyck lattice, and the formula immediately follows. Now suppose that P has $k+1$ peaks. Consider the path P' obtained by P by removing the last peak, i.e. coinciding with P up to the k-th peak and then ending with a sequence of down steps (see Figure 4).

It is now easy to see that

$$A(P) = A(P') + p_{k+1}^2 - v_k^2,$$

whence, thanks to the induction hypothesis:

$$A(P) = \sum_i (p_i^2 - v_i^2). \quad \blacksquare$$

Now we are ready to find a formula to express the rank of a partition in the Bruhat noncrossing partition lattice. The proof of the next theorem is left to the reader.

Theorem 3.3 *$NC(n)$ is a distributive lattice, and therefore it is ranked. More precisely, if $\pi = B_1|\dots|B_k \in NC(n)$, then its rank is given by:*

$$r_n(\pi) = \frac{A(\pi) - n}{2}, \tag{2}$$

where

$$A(\pi) = \sum_{i=1}^{k} \left(|B_i| \left(2b_i - 2\sum_{j=1}^{i-1} |B_j| - |B_i| \right) \right) \tag{3}$$

(here $b_i = \max B_i$).

4 Relationship with the strong Bruhat order on permutations

The last formula given for the rank of a noncrossing partition inside its Bruhat lattice is not as easy to understand as the rank function for Dyck paths. In order to find a better way to express this parameter, we make use of the concept of (generalized) pattern avoiding permutation. What we obtain is yet another description of Bruhat noncrossing partition lattices which provides some important information on the (strong) Bruhat order of the symmetric groups.

Proposition 4.1 *Removing the bars in noncrossing partitions defines a bijection between $NC(n)$ and the set $S_n(312)$ of 312-avoiding permutations of $[n]$, for any $n \in \mathbf{N}$.*

Proof. First observe that, for any $n \in \mathbf{N}$, $S_n(312) = S_n(31-2)$, since it is known that these two finite sets are both enumerated by Catalan numbers and obviously $S_n(312) \subseteq S_n(31-2)$. Now, if a pattern 31-2 appears in a noncrossing partition, then, denoting by $b < c < a$ the three elements corresponding to such a pattern, a and b must belong to the same block, and the maximum d of the block containing c must be larger than a (since the maximum of a block in a noncrossing partition is larger than every element preceding it). Thus, the four elements a, b, c, d would constitute a crossing, against the hypothesis. ■

Remark. In the rest of this section we will make an extensive use of the above described canonical bijection. In particular, we will freely switch from a noncrossing partition to its associated 312-avoiding permutation without stating it explicitly. Moreover, we will always use the same Greek letters $(\pi, \rho, \sigma, \ldots)$ to denote both a noncrossing partition and its associated 312-avoiding permutation. Finally, observe that each maximum of a block of a noncrossing partition corresponds to a left-to-right maximum in the corresponding permutation, that is an element which is greater than every other element on its left.

Observe that the composition of the bijection between Dyck paths and noncrossing partitions with the above one between noncrossing partitions and 312-avoiding permutations is precisely the bijection considered in [BK] and in [F]. More specifically, in [BK] the authors show that the area of a Dyck path corresponds to the inversion number of the associated permutation. Since the rank function of the strong Bruhat order on permutations is precisely the inversion number, we are led to conjecture a close relation between our noncrossing partition lattices and the subposets induced by the Bruhat order on 312-avoiding permutations.

Theorem 4.1 *Let $(S_n(312); \leq)$ be the poset obtained by transferring the structure of the Bruhat noncrossing partition lattice $NC(n)$ along the previous bijection. This is precisely the subposet induced on $S_n(312)$ by the strong Bruhat order of the symmetric group S_n. Therefore $S_n(312)$ is a distributive sublattice of S_n endowed with the strong Bruhat order.*

Proof. What we have to show is that the Hasse diagram of the Bruhat noncrossing partition lattice is isomorphic to that of $S_n(312)$ with the induced strong Bruhat order. To do this, it is enough to prove that the sets of elements covering a noncrossing partition and its associated 312-avoiding permutation coincide, via the bar-removing bijection.

Let π, ρ be noncrossing partitions, and suppose that $\pi \prec \rho$ in the Bruhat noncrossing partition lattice. This means that ρ is obtained by π using one of the two rules described in Theorem 3.1. In both cases, ρ is obtained from π by interchanging the minimum a of a block B with $\beta = \max B + 1$. On permutations this means that the inversion number of ρ is larger than that of π (since $a < \beta$). Now to conclude that $\pi \prec \rho$ in $S_n(312)$ it remains only to show that the above transposition does not generate other inversions, or, equivalently, that all the entries between a and β in π are either smaller than a or larger than β. Indeed, $\beta - 1$ is the maximum of B, so it appears before a in π. Hence, if there is an element x such that $a < x < \beta$ and x is between a and β in π, then we would have a pattern 312, which is excluded. Therefore we have shown that, if $\pi \prec \rho$ in $NC(n)$, then also $\pi \prec \rho$ in $S_n(312)$.

To conclude the proof we will show that, if $\pi \prec \rho$ in $S_n(312)$, then necessarily ρ is obtained by π as in Theorem 3.1. From the hypothesis it follows that ρ differs from π by a transposition of a pair of elements a and β. Suppose that $a < \beta$ and so a appears before β in π. If a was not a minimum in the noncrossing partition associated with π, then there would be an entry $x < a$ appearing after a, and so in ρ the elements β, x, a would show a pattern 312. Therefore a must be the minimum of its block B in the noncrossing partition π. Now set $b = \max B$. We claim that $\beta = b+1$. Indeed, if it is not, then $\beta - 1$ could not appear between a and β in π (since otherwise ρ would contain too many inversions). Clearly $\beta - 1$ can not appear before b too, since every entry before b must be smaller than b. Thus $\beta - 1$ lies necessarily on the right of β in π. But in this case the permutation ρ would contain a pattern 312 in the entries $\beta, a, \beta-1$, a contradiction. Therefore $\beta = b+1$, and the theorem is finally proved. ■

At this stage it is worth mentioning the following, remarkable corollary.

Corollary 4.1 *For any $n \in \mathbf{N}$, the Dyck lattice $\mathcal{D}_n$ is isomorphic to the lattice $S_n(312)$ with the strong Bruhat order.*

Our next goal is to find a synthetic description of the meet and join operations in the Bruhat lattices of 312-avoiding permutations.

Let $\pi = \pi_1 \cdots \pi_n$, $\rho = \rho_1 \cdots \rho_n \in S_n(312)$. Define the permutation $\pi \vee \rho = \sigma_1 \cdots \sigma_n$ by setting σ_i equal to the largest element among those smaller than or equal to $\max\{\pi_1, \ldots, \pi_i, \rho_1, \ldots, \rho_i\}$ not yet appeared in some previous positions. Analogously, the permutation $\pi \wedge \rho = \tau_1 \cdots \tau_n$ is defined by setting τ_i equal to the smallest element among those larger than or equal to $\min\{\pi_1, \ldots, \pi_i, \rho_1, \ldots, \rho_i\}$ not yet appeared in some previous positions. For instance, given $\pi = 32657481, \rho = 24378651$ we get $\pi \vee \rho = 34678521$ and $\pi \wedge \rho = 23457681$. In the following proposition we show that the above defined operations actually coincide with the join and meet operations in $S_n(312)$.

Proposition 4.2 *For any $\pi, \rho \in S_n(312)$, the permutations $\pi \vee \rho$ and $\pi \wedge \rho$ are respectively the join and the meet of π and ρ in the Bruhat lattice $S_n(312)$.*

Proof. Let $\max(\pi)$ and $\max(\rho)$ be the max-vectors of the noncrossing partitions associated with π and ρ, respectively. The join of the two Dyck paths associated with π and ρ corresponds to the Dyck path determined by the coordinatewise join of $\max(\pi)$ and $\max(\rho)$, say $\max(\pi) \vee \max(\rho)$, which is then the max-vector of the join of π and ρ in $S_n(312)$. There is a unique 312-avoiding permutation associated with $\max(\pi) \vee \max(\rho)$, which can be obtained as follows: the i-th entry of the permutation is the largest element among those smaller than or equal to the i-th component of the max-vector not yet appeared in the permutation. This corresponds precisely to our definition of $\pi \vee \rho$. The argument for the meet is completely analogous, and so the proof is complete. ■

References

[A] M. Aigner, *Combinatorial Theory*, Springer-Verlag, Berlin-New York, 1979.

[BS] E. Babson, E. Steingrímsson, *Generalized permutation patterns and a classification of the Mahonian statistics*, Sém. Lothar. Combin. 44 (2000), Art. B44b, 18 pp.

[BK] J. Bandlow, K. Killpatrick, *An area-to-inv bijection between Dyck paths and 312-avoiding permutations*, Electron. J. Combin. 8 (2001) #R40 (16 pp.).

[BFP] A. Bernini, L. Ferrari, R. Pinzani, *Enumerating permutations avoiding three Babson-Steingrímsson patterns*, Ann. Comb. (to appear).

[CJ] S. Cautis, D. M. Jackson, *The matrix of chromatic joins and the Temperley-Lieb algebra*, J. Combin. Theory Ser. B 89 (2003) 109-155.

[C] A. Claesson, *Generalized pattern avoidance*, European J. Combin., 22 (2001) 961-971.

[CM] A. Claesson, T. Mansour, *Enumerating permutations avoiding a pair of Babson-Steingrímsson patterns*, Ars Combinatoria (to appear).

[DS] A. Denise, R. Simion, *Two combinatorial statistics on Dyck paths*, Discrete Math. 137 (1995) 155-176.

[D] E. Deutsch, *Dyck path enumeration*, Discrete Math. 204 (1999) 167-202.

[F] M. Fulmek, *Enumeration of permutations containing a prescribed number of occurrences of a pattern of length three*, Adv. in Appl. Math. 30 (2003) 607-632.

[FP2] L. Ferrari, R. Pinzani, *Lattices of lattice paths*, J. Statist. Plann. Inference (to appear), available at *arXiv:math.CO/0411427*.

[Kra] C. Krattenthaler, *Permutations with restricted patterns and Dyck paths*, Adv. in Appl. Math. 27 (2001) 510-530.

[Kre] G. Kreweras, *Sur les partitions non croises d'un cycle*, Discrete Math. 1 (1972) 333-350.

[N] T. V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, University of Toronto Press, Toronto, 1979.

[P] R. A. Proctor, *Bruhat lattices, plane partitions generating functions, and minuscule representations*, European J. Combin., 5 (1984) 331-350.

[S] R. Simion, *Noncrossing partitions*, Discrete Math. 217 (2000) 367-409.

[SS] R. Simion, F. W. Schmidt, *Restricted Permutations*, European J. Combin., 6 (1985) 383-406.

[Sl] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, at http://www.research.att.com/ ~njas/sequences/index.html.

[S1] R. P. Stanley, *Enumerative Combinatorics 1*, Cambridge University Press, New York, 1997.

[S2] R. P. Stanley, *Enumerative Combinatorics 2*, Cambridge University Press, Cambridge, 1999.

ON TRIANGULATIONS WITH HIGH VERTEX DEGREE

OLIVIER BERNARDI

Abstract. We solve three enumerative problems concerning families of planar maps. More precisely, we establish algebraic equations for the generating function of non-separable triangulations in which all vertices have degree at least d, for a certain value d chosen in $\{3,\ 4,\ 5\}$.
The originality of the problem lies in the fact that degree restrictions are placed both on vertices and faces. Our proofs first follow Tutte's classical approach: we decompose maps by deleting the root and translate the decomposition into an equation satisfied by the generating function of the maps under consideration. Then we proceed to solve the equation obtained using a recent technique that extends the so-called *quadratic method*.

Résumé: Nous énumérons trois familles de cartes planaires. Plus précisément, nous démontrons des résultats d'algébricité pour les familles de triangulations non-séparables dont le degré des sommets est au moins égal à une certaine valeur d choisie parmi $\{3,\ 4,\ 5\}$.
L'originalité de nos résultats tiens au fait que les restrictions de degré portent simultanément sur les faces et les sommets. Nous adoptons, dans un premier temps, la démarche classique de Tutte : nous décomposons nos cartes par suppression de la racine et traduisons cette décomposition en une équation portant sur la série génératrice correspondante. Nous résolvons ensuite l'équation obtenue en utilisant des techniques récentes qui généralisent la *méthode quadratique*.

1. Introduction

The enumeration of planar maps (or *maps* for short) has received a lot of attention in the combinatorists community for nearly fifty years. Originally motivated by the four-color problem, W.Tutte introduced the concept of map in the late fifties and considered a great number of map families corresponding to various constraints on face or vertex degrees. These seminal works, based on basic decomposition techniques allied to a generating function approach, gave rise to many explicit results ([9] - [11]). Fifteen years later, some physicists became interested in the subject and developed their own tools [5] for tackling the problem. Their techniques based on matrix integrals (see [15] for an introduction) proved very powerful [4]. More recently, a more bijective approach based on conjugacy classes of trees emerged providing new insights on the subject ([8], [3], [1], [7]).
However, when one considers a map family defined by both face and vertex constraints, each of the mentioned methods seems relatively ineffective and very few enumerative results are known. There are however two major exceptions. The enumeration of bipartite (i.e. faces have even degree) cubic (i.e. vertices have degree 3) maps was first performed by Tutte by a classic generating function approach ([10],[13]). And, more recently, the enumeration of all bipartite maps according to the degree distribution of their vertices was accomplished using conjugacy classes of trees [1].

Date: 6th May 2005.

In this paper, we shall consider triangulations (i.e. faces have degree 3) constrained by vertex degree conditions. We first follow Tutte's classical approach, which consists in trying to translate the decomposition obtained by deletion of the root into a functional equation satisfied by the generating function. It is not clear at first sight why this approach should work, but it does up to the condition of relaxing some of the constraints at this stage of the resolution. This process requires to take into account in our generating function, beside the size of the map, the degree of its root-face. We end up with a polynomial equation for the (bivariate) generating function in which the variable counting the degree of the root-face cannot be trivially eliminated. We then use a recent generalization of the quadratic method [2] to get rid of this extra variable and compute an algebraic equation characterizing the univariate generating function.

We begin by some vocabulary on maps. A map is a proper embedding of a connected graph into the two-dimensional Riemann sphere, considered up to continuous deformations. A map is *rooted* if one of its edges is distinguished as the *root* and attributed an orientation. Unless otherwise specified, all maps under consideration in this paper are rooted. The face at the right of the root is called the *root-face* and the other faces are said *internal*. Similarly, the vertices incident to the root-face are said *external* and the others are said *internal*. Graphically, the root-face is usually represented as the infinite face when the map is projected on the plane (see Figure 1). The endpoints of the root are distinguished as its *origin* and *end* according to the orientation of the root. A map is *separable* if it can be decomposed into two parts (not reduced to a vertex) whose intersection is reduced to a vertex. It is *non-separable* otherwise. For instance, the map in Figure 1 is non-separable. Lastly, a map is a *triangulation* (resp. *near-triangulation*) if all its faces (resp. all its internal faces) have degree 3. For instance, the map of Figure 1 is a near-triangulation with root-face of degree 5.

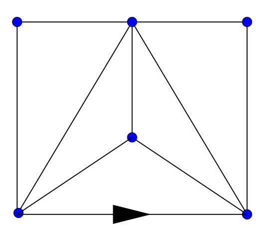


FIGURE 1. A non-separable near-triangulation.

In the sequel, we shall enumerate 3 families of non-separable triangulations. We recall some basic facts about these maps. Observe that a non-separable map (not reduced to an edge) cannot have loops nor isthmuses. Non-separable maps cannot have vertices of degree 1 either. Moreover, it is well known that planar graphs have at least one vertex of degree less than 6. We prove a stronger property: any triangulation has a vertex not incident to the root of degree less than 6. Indeed, consider a triangulation with f faces, e edges and v vertices. The incidence relation between faces and edges shows that $2e = 3f$ and reporting this identity in the Euler relation gives $e = 3v - 6$. If all vertices not incident to the root have degree at least 6 we have the inequality $2e \geq 6(v-2) + 2$ which contradicts the previous identity.

Let $\mathbf{S}$ be the set of non-separable rooted near-triangulations. As observed above, the vertices of maps in $\mathbf{S}$ have degree at least 2. We consider three sub-families $\mathbf{T}$, $\mathbf{U}$, $\mathbf{V}$ of $\mathbf{S}$. The set $\mathbf{T}$ (resp. $\mathbf{U}$, $\mathbf{V}$) is the subset of non-separable near-triangulations in which any internal vertex has degree at least 3 (resp. 4, 5). For each of the families $\mathbf{R} = \mathbf{S}$, $\mathbf{T}$, $\mathbf{U}$, $\mathbf{V}$, we consider the bivariate generating function $\mathbb{R}(x, z)$, where z counts the size (the number of edges) and x the degree of the root-face minus 2. That is to say, $\mathbb{R}(x) \equiv \mathbb{R}(x, z) = \sum_{n,d} a_{n,d} x^d z^n$ where $a_{n,d}$ is the number of maps in $\mathbf{R}$ with size n and root-face of degree $d + 2$. Note that the degree of the root-face is less than the number of edges. Therefore $\mathbb{R}(x, z)$ is a power series in the main variable z with polynomial coefficients in the secondary variable x. For each family $\mathbf{R} = \mathbf{S}$, $\mathbf{T}$, $\mathbf{U}$, $\mathbf{V}$, we will characterize the generating function $\mathbb{R}(x)$ as the unique power series solution of a functional equation (see Equations (2),(11),(13),(14)).
We also consider the set $\mathbf{F}$ of non-separable rooted triangulations and three subsets $\mathbf{G}$,$\mathbf{H}$,$\mathbf{K}$. The set $\mathbf{G}$ (resp. $\mathbf{H}$, $\mathbf{K}$) is the subset of non-separable triangulations in which any vertex not incident to the root has degree at least 3 (resp. 4, 5). As observed above, the subset of non-separable triangulations in which any vertex not incident to the root has degree at least 6 is empty. Note that, given the incidence relation between faces and edges, any triangulation has a size (number of edges) multiple of 3. To each of the families $\mathbf{L} = \mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, $\mathbf{K}$, we associate the univariate generating function $\mathbb{L}(t) = \sum_n a_n t^n$ where a_n is the number of maps in $\mathbf{L}$ of size $3n$. For each family $\mathbf{L} = \mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, $\mathbf{K}$ we will give an algebraic characterization of $\mathbb{L}(t)$ (see Equations (4),(17),(18) and Proposition 6).

There is a simple connection between the generating functions $\mathbb{F}(t)$ (resp. $\mathbb{G}(t)$, $\mathbb{H}(t)$, $\mathbb{K}(t)$) and $\mathbb{S}(x) \equiv S(x, z)$ (resp. $\mathbb{T}(x)$, $\mathbb{U}(x)$, $\mathbb{V}(x)$). This connection relies on a very simple bijection between non-separable triangulations and non-separable near-triangulations rooted on a digon (i.e. the root-face has degree 2). The bijection consists in deleting the external edge which is not the root in a near-triangulation rooted on a digon (see Figure 2). This bijection establishes a one-to-one correspondence between the set of triangulations $\mathbf{F}$ (resp. $\mathbf{G}$,$\mathbf{H}$,$\mathbf{K}$) and the set of near-triangulations in $\mathbf{S}$ (resp. $\mathbf{T}$, $\mathbf{U}$, $\mathbf{V}$) rooted on a digon.

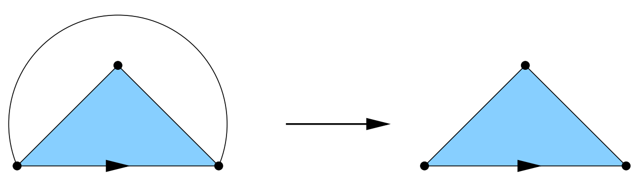


FIGURE 2. Near-triangulations rooted on a digon and triangulations.

Observe that for $\mathbf{R} \in \{\mathbf{S}, \mathbf{T}, \mathbf{U}, \mathbf{V}\}$, the power series $\mathbb{R}(0) = \mathbb{R}(0, z)$ is the generating function of near-triangulations in $\mathbf{R}$ rooted on a digon. Thus, we have the relations:

$$\mathbb{S}(0) = z\mathbb{F}(z^3), \qquad \mathbb{T}(0) = z\mathbb{G}(z^3), \qquad \mathbb{U}(0) = z\mathbb{H}(z^3), \qquad \mathbb{V}(0) = z\mathbb{K}(z^3). \tag{1}$$

This paper is organized as follows. In Section 2, we recall the decomposition scheme on maps due to W.T. Tutte. We apply it to the set $\mathbf{S}$ of unconstrained

non-separable near-triangulation and recall some known results about the generating functions $\mathbb{S}(x)$ and $\mathbb{F}(t)$. In Section 3, we apply the same decomposition scheme to the sets of near-triangulations **T**, **U** and **V** . This allows to write functional equations concerning the generating functions $\mathbb{T}(x)$, $\mathbb{U}(x)$, $\mathbb{V}(x)$. In the equations obtained, the variable x appears and cannot be trivially eliminated. This is precisely why we introduced this variable in our generating functions: it allows us to write the functionnal equations. In Section 4, we use techniques generalizing the *quadratic method* in order to get rid of the variable x. We obtain algebraic equations characterizing the generating function of the subsets of triangulations **G** , **H** , **K**. At this point, only the degree of the endpoints of the root remains unconstrained. That is, we have an algebraic characterization of triangulations for which any vertex *not incident to the root* has degree at least 3,4 or 5. In Section 5, we show that we can constrain, *a posteriori*, the degree of the endpoints of the root in the two first cases. This provides an algebraic characterization for triangulations in which *any* vertex has degree at least 3 or 4. However, no similar result is found for the set of triangulations in which *any* vertex has degree at least 5. We also study the singularities of our series and deduce the asymptotic behavior of the number of maps in each family.
Some of the results concerning triangulations in which any vertex has degree at least 3 were already proved in [6] via a compositional approach. We give here an alternative proof.

2. The decomposition principle

In the following, we adopt Tutte's classical approach for enumerating maps. That is, we decompose maps by deleting their root and translate this combinatorial decomposition into an equation satisfied by the corresponding generating function. Let us illustrate this classic approach on the problem of enumerating unconstrained non-separable triangulations (this was first done in [12]). We recall that **S** denotes the set of non-separable near-triangulations and $\mathbb{S}(x) = \mathbb{S}(x, z)$ the corresponding generation function. By convention, we exclude the map reduced to a vertex from **S**. Thus, the smallest map in **S** is the map reduced to a straight edge (see Figure 3). This map is called the *link-map* and denoted L. Its contribution to the generating function is z, hence $\mathbb{S}(x) = z + o(z)$.

FIGURE 3. The link-map L.

We decompose maps distinct from L by deleting the root. Note that, if M is a non-separable triangulation distinct from L, the face at the left of the root is internal (otherwise M would be separable) thus it has degree 3. Moreover, since M has no loop, the three vertices incident to this face are distinct. Let v be the vertex not incident to the root. When examining what can happen to M when deleting its root, one is led to distinguish two cases (see Figure 4).

Either the vertex v was incident to the root-face, in which case the map obtained by deletion of the root is separable (see Figure 5). Or v was not incident to the root-face and the map obtained by deletion of the root is a non-separable near-triangulation (see Figure 6). In the first case, the map obtained is in correspondence with an ordered pair of non-separable near-triangulations. This correspondence is bijective, that is, any ordered pair is the image of exactly one near-triangulation. In the

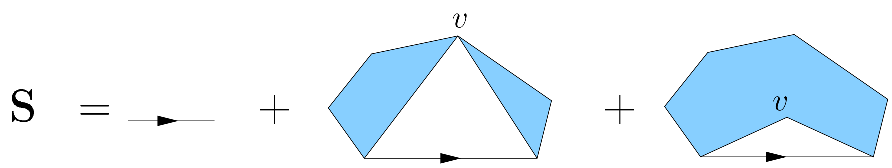


FIGURE 4. Decomposition of non-separable near-triangulations.

second case the degree of the root-face is increased by one. Hence the root-face of the near-triangulation obtained has degree at least 3. Here again, any near-triangulation for which the root-face has degree at least 3 is the image of exactly one near-triangulation.

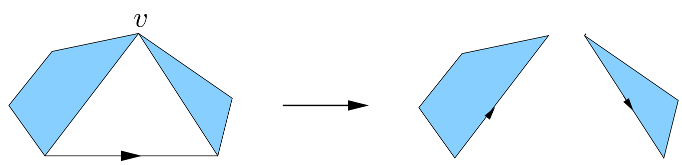


FIGURE 5. Case 1. The vertex v was incident to the root-face.

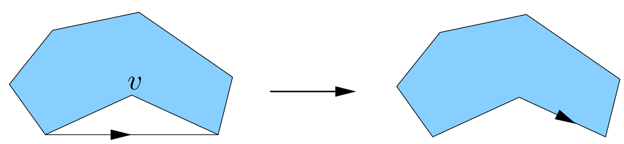


FIGURE 6. Case 2. The vertex v was not incident to the root-face.

We want to translate this analysis into a functional equation. Observe that the degree of the root-face appears in this analysis. This is why we are *forced* to introduce the variable x counting this parameter in our generating function $\mathbb{S}(x,z)$. For this reason, following Zeilberger's terminology [14], the secondary variable x is said to be *catalytic*: we need it to write the functional equation, but we shall try to get rid of it later.

In our case, the decomposition translates into the following equation:

$$\mathbb{S}(x,z) = z + xz\mathbb{S}(x,z)^2 + \frac{z}{x}\left(\mathbb{S}(x,z) - \mathbb{S}(0,z)\right) \ . \tag{2}$$

We shall explain this equation later. The first summand in the right-hand side accounts for the link map, the second summand corresponds to the case in which the vertex v is incident to the root-face, and the third summand corresponds to the case in which v is not incident to the root-face.

It is an easy exercise to check that this equation defines the series $\mathbb{S}(x,z)$ uniquely as a power series in z. By resolutions techniques presented in Section 4, we can derive from equation (2) a polynomial equation satisfied by the series $\mathbb{S}(0,z)$ where the extra variable x does not appear anymore. This equation reads

$$\mathbb{S}(0,z) \;=\; z - 27z^4 + 36z^3\mathbb{S}(0,z) - 8z^2\mathbb{S}(0,z)^2 - 16z^4\mathbb{S}(0,z)^3 \ . \tag{3}$$

Knowing that $\mathbb{S}(0,z) = z\mathbb{F}(z^3)$, we deduce the algebraic equation

$$\mathbb{F}(t) \ = 1 - 27t + 36t\mathbb{F}(t) - 8t\mathbb{F}(t)^2 - 16t^2\mathbb{F}(t)^3 \ , \tag{4}$$

characterizing $\mathbb{F}(t)$ uniquely as a power series in t. From this equation we can deduce the asymptotic behavior of the coefficients of $\mathbb{F}(t)$, that is, the number of non-separable triangulations of a given size.

3. Functional equations

In this section, we apply the decomposition principle presented in the previous section to the families $\mathbf{T},\ \mathbf{U},\ \mathbf{V}$ of non-separable near-triangulations in which all internal vertices have degree at least 3, 4, 5. We obtain functional equations satisfied by the corresponding generating functions $\mathbb{T}(x),\ \mathbb{U}(x),\ \mathbb{V}(x)$.
When we delete the root of a map, the degree of its endpoints is lowered by one. In order to translate the decomposition of maps into equations, we are forced to relax the constraints on the degree of external vertices and to control some parameters. Let $\mathbf{R}$ be one of the set $\mathbf{S},\ \mathbf{T},\ \mathbf{U},\ \mathbf{V}$. We define $\mathbf{R}_k$ as the set of maps in $\mathbf{R}$ such that the root-face has degree at least 3 and the origin of the root has degree k. We also define the set $\mathbf{R}_\infty$ as the image of $\mathbf{R}^2$ by the mapping ϕ represented in Figure 7. The mapping ϕ takes an ordered pair of maps and glues the end of the root of the first map to the origin of the second. The new root is chosen to be the root of the second map. Lastly, we write $\mathbf{R}_{\geq k} \triangleq \mathbf{R}_\infty \cup \bigcup\limits_{j\geq k} \mathbf{R}_j$.

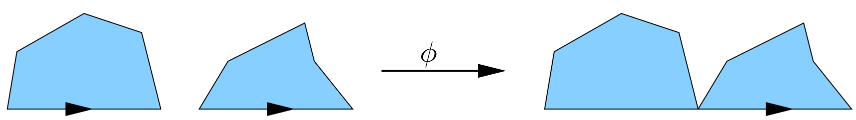


FIGURE 7. The mapping ϕ.

We shall use the symbols $\mathbb{R}_k(x,z)$, $\mathbb{R}_\infty(x,z)$ and $\mathbb{R}_{\geq k}(x,z)$ for the bivariate generating functions corresponding to the sets $\mathbf{R}_k$, $\mathbf{R}_\infty$ and $\mathbf{R}_{\geq k}$ respectively. In these series, as in $\mathbb{R}(x,z)$, the contribution of a map of size n and root-face degree d is $x^{d-2}z^n$.

We are now ready to prove our first results. Let us write L for the link-map (see Figure 3) and consider a near-triangulation M in $\mathbf{R} - \{L\}$. As observed before, the face at the left of the root is an internal face incident to three distinct vertices. Let v be the vertex not incident to the root. If v is external, the deletion of the root produces a map in $\mathbf{R}_\infty$. If v is internal and M is in $\mathbf{S}$ (resp. $\mathbf{T},\ \mathbf{U},\ \mathbf{V}$) then v has degree at least 2 (resp. 3, 4, 5) and the map obtained by deletion of the root is in $\bigcup_{k\geq 2}\mathbf{S}_k$ (resp. $\bigcup_{k\geq 3}\mathbf{T}_k$, $\bigcup_{k\geq 4}\mathbf{U}_k$, $\bigcup_{k\geq 5}\mathbf{V}_k$). Therefore, the mapping goes from $\mathbf{S}-\{L\}$ (resp. $\mathbf{T}-\{L\}$, $\mathbf{U}-\{L\}$, $\mathbf{V}-\{L\}$) to $\mathbf{S}_{\geq 2}$ (resp. $\mathbf{T}_{\geq 3}$, $\mathbf{U}_{\geq 4}$, $\mathbf{V}_{\geq 5}$). And this mapping is clearly bijective. Moreover, the map obtained after deletion of the root has size lowered by one and root-face degree increased by one. This analysis translates into the following equations:

$$\begin{aligned}\mathbb{S}(x) &= z + \frac{z}{x}\mathbb{S}_{\geq 2}(x) \ , &(5)\\ \mathbb{T}(x) &= z + \frac{z}{x}\mathbb{T}_{\geq 3}(x) \ , &(6)\\ \mathbb{U}(x) &= z + \frac{z}{x}\mathbb{U}_{\geq 4}(x) \ , &(7)\\ \mathbb{V}(x) &= z + \frac{z}{x}\mathbb{V}_{\geq 5}(x) \ . &(8)\end{aligned}$$

In view of Equation (5), we will obtain a non-trivial equation for $\mathbb{S}(x)$ if we can express $\mathbb{S}_{\geq 2}(x)$ in terms of $\mathbb{S}(x)$. Similarly, we will obtain a non-trivial equation for $\mathbb{T}(x)$ if we can express $\mathbb{T}_{\geq 2}(x)$ and $\mathbb{T}_2(x)$ in terms of $\mathbb{T}(x)$. For $\mathbb{U}(x)$ (resp. $\mathbb{V}(x)$) we need to express $\mathbb{U}_{\geq 2}(x)$, $\mathbb{U}_2(x)$ and $\mathbb{U}_3(x)$ (resp. $\mathbb{V}_{\geq 2}(x)$, $\mathbb{V}_2(x)$, $\mathbb{V}_3(x)$ and $\mathbb{V}_4(x)$). Our first task will thus be to evaluate $\mathbb{R}_{\geq 2}(x)$ for $\mathbb{R}$ in $\{\mathbb{S},\ \mathbb{T},\ \mathbb{U},\ \mathbb{V}\}$.
Let the set $\mathbf{R}$ be one of $\{\mathbf{S},\ \mathbf{T},\ \mathbf{U},\ \mathbf{V}\}$. By definition, $\mathbf{R}_\infty$ is in bijection with $\mathbf{R}^2$. This bijection translates into the following functional equation

$$\mathbb{R}_\infty(x) \ = \ x^2\mathbb{R}(x)^2 \ .$$

Observe that $\bigcup_{k\geq 2}\mathbf{R}_k$ is the set of maps in $\mathbf{R}$ for which the root-face has degree at least 3. We thus have the set identity $\bigcup_{k\geq 2}\mathbf{R}_k = \mathbf{R} - \{M \in \mathbf{R} \ / \ d(M) = 2\}$. And from this, we deduce

$$\sum_{k\geq 2}\mathbb{R}_k(x) \ = \ \mathbb{R}(x) - \mathbb{R}(0) \ ,$$

since $\mathbb{R}(0)$ is the generating function of maps in $\mathbf{R}$ rooted on a digon.
Now, since

$$\mathbf{R}_{\geq 2} = \mathbf{R}_\infty \cup \bigcup_{k\geq 2}\mathbf{R}_k \ ,$$

we have

$$\mathbb{R}_{\geq 2}(x) \ = \ x^2\mathbb{R}(x)^2 \ + \ (\mathbb{R}(x) - \mathbb{R}(0)) \qquad \text{for } \mathbb{R} \text{ in } \{\mathbb{S}(x),\ \mathbb{T}(x),\ \mathbb{U}(x),\ \mathbb{V}(x)\} \ . \quad (9)$$

Equations (5) and (9) already allow us to recover Equation (2) announced in Section 2:

$$\mathbb{S}(x) = z + xz\mathbb{S}(x)^2 + z\left(\frac{\mathbb{S}(x) - \mathbb{S}(0)}{x}\right) \ .$$

In order to go further, we will need to express $\mathbb{T}_2(x)$, $\mathbb{U}_2(x)$, $\mathbb{U}_3(x)$, $\mathbb{V}_2(x)$, $\mathbb{V}_3(x)$ and $\mathbb{V}_4(x)$ (see Equations (5-8)). We begin by the equation concerning $\mathbb{R}_2(x)$ for $\mathbb{R}$ in $\{\mathbb{S},\ \mathbb{T},\ \mathbb{U},\ \mathbb{V}\}$.
Observe that if the set $\mathbf{R}$ is one of $\{\mathbf{S},\ \mathbf{T},\ \mathbf{U},\ \mathbf{V}\}$, then the set $\mathbf{R}_2$ is in bijection with $\mathbf{R}$ by the mapping illustrated in Figure 8. Consequently we can write

$$\mathbb{R}_2(x) \ = \ xz^2\mathbb{R}(x) \qquad \text{for } \mathbb{R} \text{ in } \{\mathbb{S}(x),\ \mathbb{T}(x),\ \mathbb{U}(x),\ \mathbb{V}(x)\} \ . \quad (10)$$

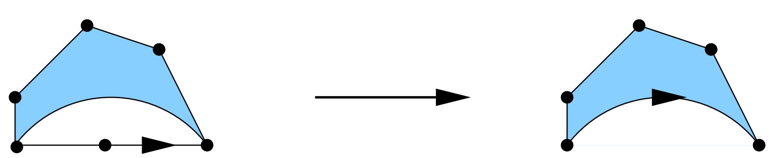


FIGURE 8. A bijection between $\mathbf{R}_2$ and $\mathbf{R}$.

We put immediately this equation to contribution in order to write an equation for $\mathbf{T}$:

$$\begin{aligned}\mathbb{T}(x) &= z + \frac{z}{x}\mathbb{T}_{\geq 3}(x) && \text{by equation (6)}\\ &= z + \frac{z}{x}\left(\mathbb{T}_{\geq 2}(x) - \mathbb{T}_2(x)\right) \\ &= z + \frac{z}{x}\left(x^2\mathbb{T}(x)^2 + (\mathbb{T}(x) - \mathbb{T}(0)) - xz^2\mathbb{T}(x)\right) && \text{by equations (9) and (10).}\end{aligned}$$

Proposition 1. *The generating function $\mathbb{T}(x)$ of non-separable near-triangulations in which all internal vertices have degree at least 3 satisfies:*

$$\mathbb{T}(x) \;=\; z \;+\; xz\mathbb{T}(x)^2 \;+\; z\left(\frac{\mathbb{T}(x) - \mathbb{T}(0)}{x}\right) \;-\; z^3\mathbb{T}(x)\;. \tag{11}$$

We go a step further to find an equation concerning the set $\mathbf{U}$. We have to express $\mathbb{U}_3(x)$ in terms of $\mathbb{U}(x)$. Let M be a map in $\mathbf{U}_3$ and n the origin of its root. By definition, n has degree 3 and the root-face of M has degree at least 3. Observe that, since the map is non-separable, the vertex a preceding n on the root-face is distinct from the vertex b following n (see Figure 9). Let us denote by v the third vertex adjacent to n. Since there can be no loop, it is clear that v is distinct from a and b. With these considerations, it is clear that maps in $\mathbf{U}_3$ are in bijection with maps in $\mathbf{U}_{\geq 3}$ by the mapping illustrated in Figure 9. Indeed, the vertex v must be either incident to the root-face (in which case the result is in $\mathbf{U}_\infty$) or of degree $d \geq 4$ (in which case the result is in $\mathbf{U}_{d-1}$).
The bijection translates into the identity:

$$\mathbb{U}_3(x) = z^3\mathbb{U}_{\geq 3}(x) = z^3(\mathbb{U}_{\geq 2}(x) - \mathbb{U}_2(x))\;. \tag{12}$$

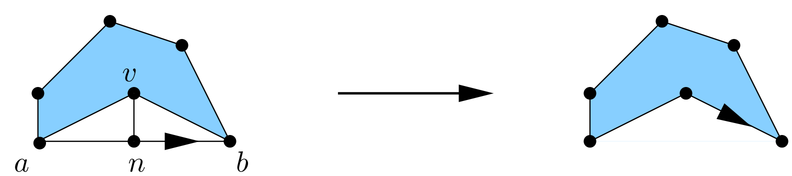


FIGURE 9. A bijection between $\mathbf{U}_3$ and $\mathbf{U}_{\geq 3}$.

We are now ready to establish an equation for $\mathbf{U}$:

$$\begin{aligned}\mathbb{U}(x) &= z + \frac{z}{x}\mathbb{U}_{\geq 4}(x) && \text{by equation (7)}\\ &= z + \frac{z}{x}\left(\mathbb{U}_{\geq 2}(x) - \mathbb{U}_2(x) - \mathbb{U}_3(x)\right) \\ &= z + \frac{z(1-z^3)}{x}\left(\mathbb{U}_{\geq 2}(x) - \mathbb{U}_2(x)\right) && \text{by equation (12)}\\ &= z + \frac{z(1-z^3)}{x}\left(x^2\mathbb{U}(x)^2 + (\mathbb{U}(x) - \mathbb{U}(0)) - xz^2\mathbb{U}(x)\right) && \text{by (9) and (10).}\end{aligned}$$

Proposition 2. *The generating function $\mathbb{U}(x)$ of non-separable near-triangulations in which all internal vertices have degree at least 4 satisfies:*

$$\mathbb{U}(x) = z + xz(1-z^3)\mathbb{U}(x)^2 + z(1-z^3)\left(\frac{\mathbb{U}(x) - \mathbb{U}(0)}{x}\right) - z^3(1-z^3)\mathbb{U}(x)\;. \tag{13}$$

It is possible to use the same approach to establish an equation concerning the set $\mathbf{V}$. Unfortunately, the decomposition happens to be considerably more entangled and the proof quite heavy. We spare the reader the details of these calculations and simply state the following proposition.

Proposition 3. *The generating function $\mathbb{V}(x) = VF(x,z)$ of non-separable near-triangulations in which all internal vertices have degree at least 5 satisfies:*

$$\begin{aligned} & x^2z^5(z^3-1)\mathbb{V}(x)^3 - z(-x+2z^4+xz^9-2xz^6-2z^7+xz^{12})\mathbb{V}(x)^2 \\ & +\frac{(2x^2z^6+xz-2x^2z^3-z^5+z^8-xz^{10}-xz^{13}-x^2z^{12}+x^2z^{15}-2x^2z^9+2xz^7-x^2)}{x^2}\mathbb{V}(x) \\ & \quad +\frac{z(-x+z^4+xz^9-2xz^6-z^7+xz^{12})}{x^2}\mathbb{V}(0) - \frac{z^5(z^3-1)}{x}[x]\mathbb{V}(x) + z(z^3+1) \quad = \quad 0, \end{aligned} \tag{14}$$

where $[x]V(x)$ denotes the coefficient of x in $\mathbb{V}(x)$.

4. Algebraic equations for triangulations with high degree

In the previous section, we have exhibited functional equations concerning the families of near-triangulations $\mathbf{T}, \mathbf{U}, \mathbf{V}$. We now solve these equations and establish algebraic characterization for the families $\mathbf{F}, \mathbf{G}, \mathbf{H}$ of triangulations in which vertices not incident to the root have degree at least $3, 4, 5$ respectively. As observed in the introduction, the generating functions $\mathbb{F}(t), \mathbb{G}(t), \mathbb{H}(t)$ are related to the series $\mathbb{T}(0), \mathbb{U}(0), \mathbb{V}(0)$ (see Equation (1)).

Let us look at Equations (11), (13) and (14) satisfied by the series $\mathbb{T}(x)$, $\mathbb{U}(x)$ and $\mathbb{V}(x)$ respectively. We begin with Equation (11). Multiplying each side of this equation by x, we obtain a polynomial equation in the two variables x and z and the two unknown series $\mathbb{T}(x)$ and $\mathbb{T}(0)$. It is easily seen that this equation allows us to compute the coefficients of $\mathbb{T}(x)$ (hence those of $\mathbb{T}(0)$) iteratively. In particular, Equation (11) defines the series $\mathbb{T}(0)$ uniquely as a power series in z. The same property holds for Equations (13) and (14): reducing both sides of the equation to the same denominator we obtain a polynomial equation. In the case of Equation (14) there is a third unknown series, $[x]\mathbb{V}(x)$ appearing. But in both cases, it is easily seen that the equation defines the unknown series $\mathbb{U}(0)$, $\mathbb{V}(0)$ uniquely as a power series in z.
So, in a sense, these equations solve the enumeration problems. However we want to find algebraic equations for the series $\mathbb{T}(0)$, $\mathbb{U}(0)$ and $\mathbb{V}(0)$ in which the extra variable x do not appear anymore. Techniques for performing such manipulations appear many times in the literature. In the cases of Equation (11) and (13) we can routinely apply the so called *quadratic method*. This method allows one to solve any quadratic equation in which there is one unknown bivariate series and one unknown univariate series. However, Equation (14) has two unknown univariate series and is moreover cubic in the bivariate series. Very recently, a new formalism due to Bousquet-Mélou and Jehanne has emerged allowing one to solve this kind of equation [2]. We adopt this formalism.

Let us begin with Equation (11) concerning $\mathbb{T}(0)$. We define the polynomial

$$P(T, T_0, X, Z) = XZ + X^2ZT^2 + ZT - ZT_0 - XZ^3T - XT .$$

Equation (11) can be written as

$$P(\mathbb{T}(x), \mathbb{T}(0), x, z) = 0 . \tag{15}$$

Let us consider the equation $P_1'(\mathbb{T}(x), \mathbb{T}(0), x, z) = 0$, where P_1' denotes the first derivative of P with respect to its first variable. This equation can be written as

$$x = z + 2x^2z\mathbb{T}(x) - xz^3 .$$

This equation is not satisfied for a generic x. However, considered as an equation in x, it is straightforward to show that it admits a unique power series solution $X(z)$. Taking the derivative of Equation (15) with respect to x one obtains

$$\frac{\partial \mathbb{T}(x)}{\partial x} \cdot P'_1(\mathbb{T}(x), \mathbb{T}(0), x, z) + P'_3(\mathbb{T}(x), \mathbb{T}(0), x, z) = 0$$

where P'_3 denotes the first derivative of P with respect to its third variable. Substituting the series $X(z)$ for x in that equation, we see that the series $X(z)$ is also solution of equation $P'_3(\mathbb{T}(x), \mathbb{T}(0), x, z) = 0$. Hence, we have a system of three equations

$$\begin{array}{rcl} P(\mathbb{T}(X(z)), \mathbb{T}(0), X(z), z) & = & 0\ , \\ P'_1(\mathbb{T}(X(z)), \mathbb{T}(0), X(z), z) & = & 0\ , \\ P'_3(\mathbb{T}(X(z)), \mathbb{T}(0), X(z), z) & = & 0\ , \end{array} \tag{16}$$

for the three unknown series $\mathbb{T}(X(z))$, $\mathbb{T}(0)$ and $X(z)$. This polynomial system can be solved by elimination techniques using either resultant calculations or Gröbner bases. Performing these eliminations one obtains an algebraic equation for $\mathbb{T}(0)$:

$$\mathbb{T}(0) = z - 24z^4 + 3z^7 + z^{10} + (32z^3 + 30z^6 - 4z^9 - z^{12})\mathbb{T}(0) - 8z^2(1+z^3)^2\mathbb{T}(0)^2 - 16z^4\mathbb{T}(0)^3.$$

Using the fact that $\mathbb{T}(0) = z\mathbb{G}(z^3)$ we get the following theorem.

Theorem 4. *Let **G** be the set of non-separable triangulations for which any vertex not incident to the root has degree at least 3 and let $\mathbb{G}(t)$ be its generating function. The series $\mathbb{G}(t)$ is uniquely defined as a power series in t by the algebraic equation:*

$$\begin{array}{rl} 1 - 24t + 3t^2 + t^3 \ - & (1+t)(1 - 33t + 3t^2 + t^3)\mathbb{G}(t) \\ & -\ 8t(1+t)^2\mathbb{G}(t)^2 \ - \ 16t^2\mathbb{G}(t)^3 \ = \ 0\ . \end{array} \tag{17}$$

The same manipulations lead to a similar result concerning the set **H**.

Theorem 5. *Let **H** be the set of non-separable triangulations for which any vertex not incident to the root has degree at least 4 and let $\mathbb{H}(t)$ be its generating function. The series $\mathbb{H}(t)$ is uniquely defined as a power series in t by the algebraic equation:*

$$\begin{array}{c} 1 - 24t + 54t^2 - 32t^3 + 3t^5 - t^6 + 8(1-t)^2(1+t+t^2)^2 t\mathbb{H}(t)^2 \\ -(1+t-t^2)(1 - 33t + 72t^2 - 41t^3 + 3t^5 - t^6)\mathbb{H}(t) - 16t^2(t-1)^4\mathbb{H}(t)^3 = 0. \end{array} \tag{18}$$

For the Equation (14) concerning $\mathbb{V}(0)$ the method is almost identical. We see that there is a polynomial $Q(V, V_0, V_1, x, z)$ such that Equation (14) can be written as $Q(\mathbb{V}(x), \mathbb{V}(0), [x]\mathbb{V}(x), x, z) = 0$. But we can show that there are exactly *two* series $X_1(z)$, $X_2(z)$ such that $Q'_1(\mathbb{V}(X(z)), \mathbb{V}(0), [x]\mathbb{V}(x), X(z), z) = 0$. Thus, we obtain a system of 6 equations

$$\begin{array}{rcl} Q(\mathbb{V}(X_i(z)), \mathbb{V}(0), [x]\mathbb{V}(x), X_i(z), z) & = & 0 \\ Q'_1(\mathbb{V}(X_i(z)), \mathbb{V}(0), [x]\mathbb{V}(x), X_i(z), z) & = & 0 \qquad\qquad i = 1,\ 2 \\ Q'_3(\mathbb{V}(X_i(z)), \mathbb{V}(0), [x]\mathbb{V}(x), X_i(z), z) & = & 0 \end{array} \tag{19}$$

for the 6 unknown series $\mathbb{V}(X_1(z))$, $\mathbb{V}(X_2(z))$, $X_1(z)$, $X_2(z)$, $\mathbb{V}(0)$ and $[x]\mathbb{V}(x)$. This system can be solved via elimination techniques. The calculus involved are really heavy, and the result is too big to fit in here. However, we have the following theorem.

Theorem 6. *Let **K** be the set of non-separable triangulations for which any vertex not incident to the root has degree at least 5 and let $\mathbb{K}(t)$ be its generating function. The series $\mathbb{K}(t)$ is algebraic of degree 6.*

5. Constraining the vertices incident to the root

We have established algebraic equations for triangulations in which any vertex *not incident to the root* has degree at least 3 (resp. 4). We will now establish equations for triangulations in which *any* vertex has degree at least 3 (resp. 4). This can be done by expressing the generating function of triangulations in which any vertex has degree at least 3, 4 in terms of the series $\mathbb{G}(t)$, $\mathbb{H}(t)$.

Theorem 7. *Let* $\boldsymbol{G}^*$ *be the set of non-separable triangulations for which* any *vertex has degree at least 3 and let* $\mathbb{G}^*(t)$ *be its generating function. The series* $\mathbb{G}$ *and* $\mathbb{G}^*$ *are related by the identity*

$$\mathbb{G}^*(t) = (1-2t)\mathbb{G}(t)\ . \tag{20}$$

Theorem 8. *Let* $\boldsymbol{H}^*$ *be the set of non-separable triangulations for which* any *vertex has degree at least 4 and let* $\mathbb{H}^*(t)$ *be its generating function. The series* $\mathbb{H}$ *and* $\mathbb{H}^*$ *are related by the identity*

$$\mathbb{H}^*(t) = \frac{1-5t+5t^2-3t^3}{1-t}\mathbb{H}(t)\ . \tag{21}$$

The proofs of Theorems 7 and 8 are reminiscent of the manipulations practiced in Section 3. We do not detail them here.

Plugging Equation (20) (resp. (21)) in Equation (17) (resp. (18)) one obtains algebraic characterizations of the series $\mathbb{G}^*$ (resp. $\mathbb{H}^*$). In particular, it is possible to compute the first coefficients of these series. We find:

$$\mathbb{G}^*(t) = t^2 + 3t^3 + 19t^4 + 128t^5 + 909t^6 + 6737t^7 + 51683t^8 + o(t^8),$$
$$\mathbb{H}^*(t) = t^4 + 3t^5 + 12t^6 + 59t^7 + 325t^8 + 1875t^9 + 11029t^{10} + o(t^{10}).$$

In the expansion of $\mathbb{G}^*(t)$, the smallest non-zero coefficient t^2 corresponds to the tetrahedron. In the expansion of $\mathbb{H}^*(t)$, the smallest non-zero coefficient t^4 corresponds to the octahedron (see Figure 10).
We were unable to find an equation that would permit to count non-separable triangulations in which *any* vertex has degree at least 5. However, we can use the algebraic equation satisfied by the series $\mathbb{K}(t)$ and discover that the first non-zero coefficient corresponds to the icosahedron (see Figure 10).

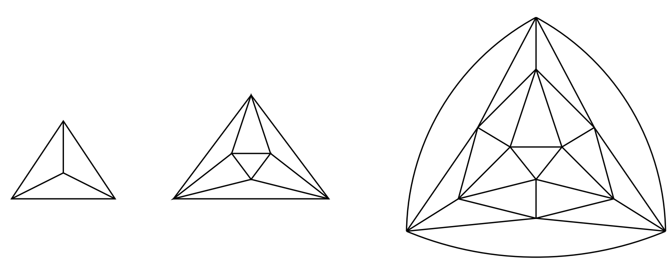


FIGURE 10. The platonic solids: tetrahedron, octahedron, icosahedron.

The algebraic equations for $\mathbb{G}^*(t)$, $\mathbb{H}^*(t)$ allow one to study their dominant singularity and deduce the asymptotic behavior of the number of triangulations in which any vertex has degree at least 3, 4. It can be shown that the number of triangulations in which any vertex has degree at least 5 has the same asymptotic behavior (up to a constant factor) as the number of triangulations in which any vertex *not incident to the root* has degree at least 5. Therefore, this behavior can be deduced from the algebraic equation concerning $\mathbb{K}(t)$. As expected, the number f_n, g_n, h_n, k_n of triangulations in which any vertex has degree at least 2,3,4,5 respectively have the generic form: $f_n, g_n, h_n, k_n \sim \alpha n^{-5/2}\mu^n$. And the exponential factors are approximately equal to $\mu_F = 13.5,\ \mu_G = 10.20,\ \mu_H = 7.03,\ \mu_K = 4.06$.

6. Concluding remarks

We studied three families of triangulations. We were able to establish algebraic equations for the generating functions of non-separable triangulations in which vertices not incident to the root have degree at least 3, 4, 5. It was then possible to obtain algebraic equations for non-separable triangulations in which any vertex has degree at least 3, 4. However, no similar result was found for degree 5.

The algebraic equations can be converted into differential equations (using for instance the *algeqtodiffeq* Maple command) from which one is able to compute the coefficients of the series in linear time. Thus our equations allow to compute efficiently the number of maps of a given size for each of the mentioned families. Moreover, asymptotic results for the number of such maps can also be found routinely from the algebraic equations.

We proved our results using basic decomposition techniques allied with a generating function approach. Alternatively, it is possible to obtain some of these results by a compositional approach. The equation concerning non-separable triangulations in which any vertex has degree at least 3 was proved by this method in [6]. It is also possible to recover the equation concerning non-separable triangulations in which any vertex has degree at least 4. However, the results concerning non-separable triangulations in which vertices not incident to the root have degree at least 5 seem hard to obtain by this method.

The approach adopted in this paper is classic except for some manipulations of equations relying on very recent techniques. Still, it is quite surprising that our approach works since the maps under consideration were constrained by degree conditions placed both on vertices and faces. This relies on the possibility to relax some of the constraints (the degree of the root-face and external vertices) until the last steps of the resolution.

In this paper, we have focused on non-separable triangulations but it is possible to practice the same kind of manipulations for all (that is to say *possibly separable*) triangulations. The method should also apply to some other families of maps, in particular to quadrangulations. Thus, a whole new class of map families is expected to have algebraic generating functions.

References

[1] M. Bousquet-Mélou and G.Schaeffer. The degree distribution in bipartite planar maps: application to the Ising model. In *FPSAC*, 2003. See also ArXiv: math.CO/0211070v2.

[2] M. Bousquet-Mélou and A. Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *ArXiv: math.CO/0504018*, 2005.

[3] M. Bousquet-Mélou and G. Schaeffer. Enumeration of planar constellations. *Adv. in Appl. Math.*, 24(4):337–368, 2000.

[4] J. Bouttier, P. Di Francesco, and E. Guitter. Census of planar maps: from the one-matrix model solution to a combinatorial proof. *Nucl. Phys. B*, 645:477–499, 2002.

[5] E. Brezin, C. Itzykson, G. Parisi, and J-B. Zuber. Planar diagrams. *Commun. Math. Phys.*, 49:35–51, 1978.

[6] Z. Gao and C. Wormald. Enumeration of rooted cubic planar maps. *Annals Comb.*, 6:313 – 325, 2002.

[7] 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.

[8] G. Schaeffer. *Conjugaison d'arbres et cartes combinatoires aléatoires.* PhD thesis, Université Bordeaux I, Bordeaux, 1998.

[9] W.T. Tutte. A census of planar triangulations. *Canad. J. Math.*, 14:21–38, 1962.

[10] W.T. Tutte. A census of slicings. *Canad. J. Math.*, 14:708–722, 1962.

[11] W.T. Tutte. A census of planar maps. *Canad. J. Math.*, 15:249–271, 1963.

[12] W.T. Tutte. Chromatic sum for rooted planar triangulations : the cases λ=1 and λ=2. *Canad. J. Math.*, 25(2):426–447, 1973.

[13] W.T. Tutte. Chromatic sum for rooted planar triangulations III: the case λ=3. *Canad. J. Math.*, 25(4):780–790, 1973.

[14] D. Zeilberger. The umbral transfer-matrix method: I. Foundations. *J. Combin. Theory, Ser. A*, 91:451–463, 2000.

[15] A. Zvonkin. Matrix integrals and map enumeration : an accessible introduction. *Math. Comput. Model.*, 26:281–304, 1997.

LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France

E-mail address: `bernardi@labri.fr`

A DIRECT DECOMPOSITION OF 3-CONNECTED PLANAR GRAPHS

MANUEL BODIRSKY, CLEMENS GRÖPL, DANIEL JOHANNSEN, AND MIHYUN KANG

ABSTRACT. We present a decomposition strategy for c-nets, i. e., rooted 3-connected planar maps. The decomposition yields an algebraic equation for the number of c-nets with a given number of vertices and a given size of the outer face. The decomposition also leads to a deterministic and polynomial time algorithm to sample c-nets *uniformly at random*. Using rejection sampling, we can also sample isomorphism types of convex polyhedra, i.e., 3-connected planar graphs, uniformly at random.

RÉSUMÉ. Nous proposons une stratégie de décomposition pour les cartes pointées 3-connexes (c-réseaux). Cette décomposition permet d'obtenir une équation algébrique pour le nombre de c-réseaux suivant le nombre de sommets et la taille de la face extèrieure. On en déduit un algorithme de complexité en temps polynomiale pour le tirage aléatoire uniforme des c-réseaux. En utilisant une méthode à rejet, nous obtenons aussi un algorithme de tirage aléatoire uniforme pour les graphes planaires 3-connexes.

1. INTRODUCTION

Three-connected planar graphs are in a one-to-one relationship to the edge-graphs of convex polyhedra [23]. The enumeration of such graphs has a long history. Already Euler attempted to find an exact formula for the number of isomorphism types of convex polyhedra [10], which is still unknown. However, since almost all such graphs have a trivial automorphism group [3, 26], and since all embeddings of such a graph are equivalent (due to Whitney; see e.g. [9]), the *asymptotic* behavior of these numbers is the same as for the number of *c-nets*, i.e., three-connected planar maps with a distinguished directed edge at the outer face. The exact and the asymptotic number of c-nets for a given number of edges was first computed by Tutte [25]. Mullin and Schellenberg [18] found exact formulas in terms of vertices and faces. The algebraic equation derived there was analyzed by Bender and Richmond in [2], who showed that the growth constant for the number of c-nets depending on the number of vertices is $16/27(17+7\sqrt{7}) \doteq 21.049042$.

Other motivations to study c-nets come from *random sampling* in theoretical computer science[1]. The only known algorithm to sample labeled planar graphs uniformly at random in polynomial time requires a sampling procedure for c-nets in its "inner loop" [4]. A sampling procedure from [1, 21, 22] for planar maps with given numbers of vertices and edges was applied for that step in [4], and the analysis shows that this is the bottleneck for the performance. Recently, the sampling procedure for c-nets was improved [13]. But still this approach applies rejection sampling, and therefore can only lead to expected polynomial time sampling procedures.

In this paper, we present a new decomposition strategy for the number of c-nets with a given number of vertices and a given size of the outer face. We will formulate the decomposition using the generating function for the number of c-nets. The resulting equations

Key words and phrases. Random sampling, planar graphs, algorithms.

[1] In the literature often the word "generating" is used instead of "sampling". We prefer "sampling" because it is more specific.

can be solved with the quadratic method [6,12], and the generating function for the number of c-nets is algebraic of degree four, and therefore has an explicit description with radicals. Using the computer algebra package GFUN [20], we compute a linear differential equation with polynomial coefficients that describes the generating function. From that we get a single-parameter recurrence for its coefficients that allows to compute the number of c-nets with more than 100000 vertices within reasonable time. Following the discussion in the forthcoming book of Flajolet and Sedgewick [12] we compute the mentioned growth constant.

With the decomposition strategy we obtain the first *deterministic* polynomial time sampling procedure for c-nets. Together with the results in [4] we obtain the first *deterministic* polynomial time sampling procedure for labeled planar graphs. Since almost all 3-connected graphs have a trivial automorphism group [3], this can also be used in a rejection sampling procedure to sample 3-connected planar graphs in *expected* polynomial time. The algorithm uses a recursive formula for c-nets on n vertices with a specified size of the outer face. Our decomposition strategy is flexible enough to also control other parameters of c-nets, for instance the total number of edges, faces, or the degrees of root vertices, if needed. From a methodological point of view, the decomposition is interesting, since it generalizes the well-known and classical approach of Tutte to count triangulations [24]. This direct technique was never carried out for c-nets – yet it is particularly suited for the *recursive method for sampling* (an early reference is [19]; see [8, 11] for recent developments).

The fact that we can control the size of the outer face opens new applications for counting *unlabeled* planar graphs. The only approach in question to enumerate unlabeled planar graphs exploits the connectivity structure, and was already proposed in [27]. As a first step, we can use the result of the present paper to compute the number of *unlabeled rooted 2-connected* planar graphs on a given number of edges. Moreover, using the sampling procedure for c-nets with a specified size of the outer face, we obtain the first expected polynomial time sampling procedure for 2-connected planar graphs [5]. With the sampling procedures for c-nets from [13] this is not possible.

Outline. The paper is organized as follows: We first introduce c-nets, and mention previous enumerative results. In Section 3, we describe the unique decomposition strategy for c-nets, which directly translates into equations for the generating function for the number of c-nets. In Section 4 we apply the quadratic method to derive a single algebraic equation of degree four that defines this generating function, and to derive a single parameter recurrence. Section 5 uses the decomposition to sample c-nets uniformly at random.

2. Planar Structures and c-nets

A *map* is a graph embedded in the plane. A *planar graph* is a graph that has an embedding in the plane. A graph is k*-connected* if the graph stays connected after deleting any k vertices. By Whitney's theorem (see e.g. [9]), all embeddings of 3-connected planar graphs are equivalent. A *rooted map* is a map with a distinguished directed edge st on the outer face. If we count rooted maps, we count them up to isomorphisms that map the outer face to the outer face and the root edge to the root edge.

A *c-net* is a rooted and 3-connected map on at least three vertices. We distinguish between *outer* vertices, which lie on the outer face and *inner* vertices, which do not lie on the outer face. The outer vertices include the vertices of the root st and are labeled $s, t, u_1, \ldots, u_k$ in clockwise order starting with the root; see Figure 1.

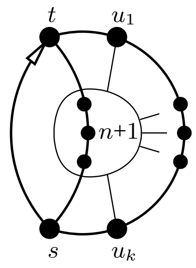


FIGURE 1. A c-net on $n + k + 3$ vertices

Starting with Tutte's pioneering work [25], many classes of planar maps were enumerated. It is possible to compute the number of unrooted planar maps on m edges [17,28,29]. For *rooted* maps, the enumeration is easier. The formulas for 3-connected, 2-connected, connected, and all rooted planar maps are related via a connectivity decomposition [25]. Mulling and Schellenberg [18] used a bijection between 3-connected rooted maps, i.e., c-nets, and *quadrangulations*, which can be further decomposed, to enumerate c-nets in terms of edges and faces (by Euler's formula, one can then also control the number of vertices). The evaluation of their formula, however, involves the evaluation of a double summation. In this paper, we present a single parameter recurrence that can be computed much faster. Since the generating function is algebraic, it is straightforward to use singularity analysis (an excellent exposition of which can be found in the forthcoming book of Flajolet and Sedgewick) to reproduce the asymptotic results of Bender and Richmond [2].

3. DECOMPOSITION

In this section we present a unique decomposition strategy for c-nets. Informally, the idea is to *remove* the root edge, and to describe the remaining graph in terms of smaller c-nets. Tutte [24] applied this technique successfully to *near-triangulations*, which generalize triangulations. The decomposition proposed by Tutte is simple: Either the graph without the root edge is 3-connected, or it is decomposed at its 2-cuts into 3-connected components. In either case the decomposition yields one or more smaller near-triangulations. The uniqueness of the decomposition is ensured by an important property of the simple structure of near-triangulations: The components of a decomposition at a 2-cut are independent, i.e., an arbitrary combination of near-triangulations can be composed to obtain a near-triangulation.

The generalization of this decomposition for c-nets faces mainly two problems. First, the objects resulting from the decomposition are not necessarily c-nets. Second, the components induced by a 2-cut are in general not independent as described before. We solve these problems by assigning distinct generating functions to each type of component and by introducing a third case for the decomposition into dependent components. This leads us to the notions of d-nets (one 3-connected component), e-nets (there is a 2-cut that yields two dependent components) and f-nets (there is a 2-cut that yields two independent components), which are depicted in Fig. 3.

In figures, we draw the root edge st as a directed edge. Edges that are added to the graph are indicated as dotted lines. If a pair of vertices forms a 2-cut, we draw a dashed

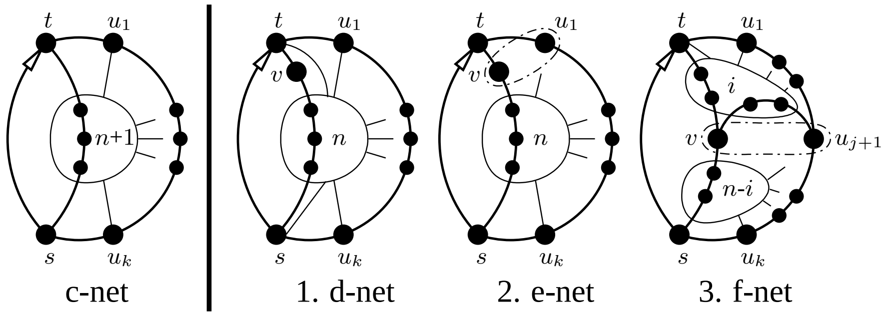


FIGURE 2. The basic case distinction: Every c-net (except K_3) is either a d-net, and e-net, or an f-net.

circle around the two vertices. The set of inner vertices is represented by a closed line with its size noted inside.

We formulate the decomposition in terms of generating functions. Let $c(n,k)$ be the number of all c-nets on $n+1$ inner vertices and $k+2$ outer vertices. For technical reasons, we also define $c(n,k)$ for $k=0$: This case corresponds to graphs where the root st is a double edge which bounds the outer face. Every c-net with a double root edge can be identified with a simple c-net by removing one of the parallel edges, hence $c(n) := c(n,0) = \sum_{k=1}^{n} c(n-k,k)$. Let $C(t,u) := \sum_{n\geq 0}\sum_{k\geq 0} c(n,k)t^n u^k$ be the two variable ordinary generating function for the number of c-nets, and let $C(t) := \sum_{n\geq 0} c(n)t^n$.

Decomposition of c-nets. If a c-net has only three vertices (s, t and an inner vertex) then it is the K_3 and represents the only initial case of the whole decomposition. (The decomposition terminates trivially for negative values of n or k.) Now consider c-nets on at least four vertices. We distinguish three disjoint cases; they are depicted in Fig. 2.

(1) After removing the root edge, the remaining graph is still three-connected.
(2) There is a 2-cut in the graph without the root edge, and vertex t is of degree three. (The two neighbors of t besides vertex s necessarily form a 2-cut in the graph without the root.)
(3) There is a 2-cut in the graph without the root edge, and vertex t is at least of degree four.

Now let $D(n,k)$, $E(n,k)$ and $F(n,k)$ be the generating functions representing the c-nets of the first, second and third case, with coefficients $d(n,k)$, $e(n,k)$ and $f(n,k)$. For convenience we call these three different kind of c-nets *d-nets*, *e-nets* and *f-nets*. Then the basic case distinction can be formulated as follows.

$$C(t,u) = 1 + D(t,u) + E(t,u) + F(t,u)\,. \tag{1}$$

Decomposition of d-nets. Let G be a d-net, i.e., G is a c-net which is 3-connected after removing the root st. The decomposition of d-nets is easy. Let v be the neighbor of t (different from s) on the inner face. There are two distinct cases, depicted in Fig. 3.

(1) The vertex v is the only vertex on the inner face of st except s and t.
Decomposition: Remove st and choose sv as new root edge.
Result: A c-net with one inner vertex less and one outer vertex more than G.

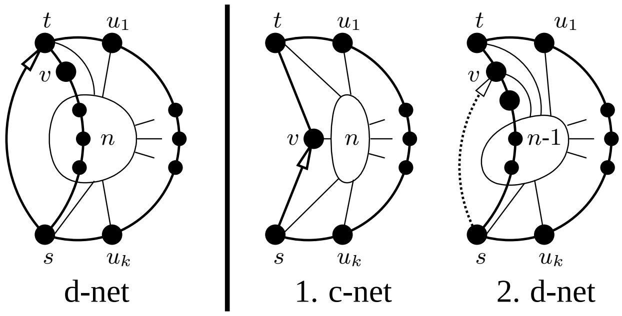


FIGURE 3. The decomposition of d-nets.

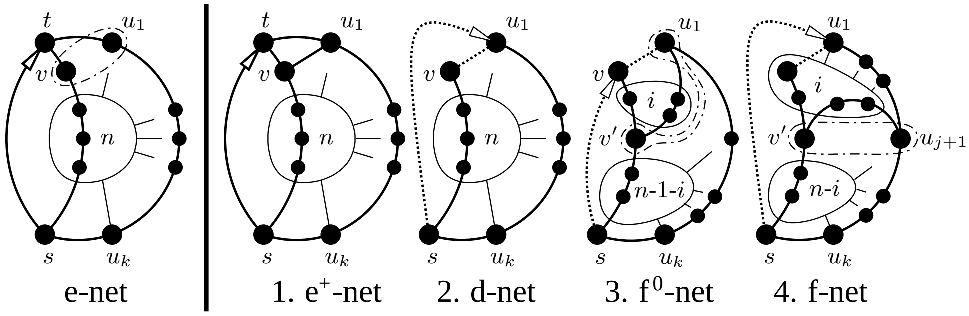


FIGURE 4. The decomposition of e-nets.

(2) There is at least one other vertex than v on the inner face of st except s and t.
Decomposition: Remove st and insert sv as new root edge.
Result: A d-net with one inner vertex less and one outer vertex more than G.

With exception of the initial case every c-net with a double edges root is a d-net. Hence

$$D(t,u) = \frac{t}{u}(C(t,u) - C(t,0)) + \frac{t}{u}(D(t,u) - D(t,0)) \tag{2}$$

$$D(t,0) = C(t,0) - 1\,. \tag{3}$$

Decomposition of e-nets. Let G be an e-net, i.e., G is a c-net and t is of degree 3. The two neighbors of t apart from s are u_1 on the outer and v on the inner face and $\{v, u_1\}$ forms a 2-cut on G without st. We now introduce the last two kinds of c-nets that appear in the decomposition. *e$^+$-nets* (represented by $E^+(t,u)$) are defined as e-nets where the two neighbors (other than s) of t are connected by an edge, whereas *f^0-nets* (represented by $F^0(t,u)$) are defined as f-nets where u_1 has to be one of the cut vertices. In the decomposition of d-nets there are four distinct cases; they are depicted in Fig. 4.

(1) There is an edge vu_1 in G.
Result: An e$^+$-net with the same number of vertices like G.
(2) There is no edge vu_1 and G without t is 3-connected.
Decomposition: Remove t, insert the edge vu_1 and insert su_1 as new root edge.
Result: A d-net with one outer vertex less than G.

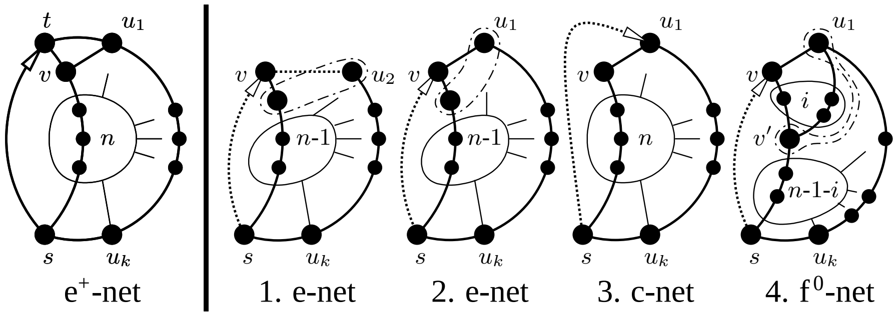


FIGURE 5. The decomposition of e⁺-nets.

(3) There is no edge vu_1 and G without t has a 2-cut including u_1.
Decomposition: Remove t, insert vu_1 and insert sv as new root edge.
Result: An f^0-net with one inner vertex less than G.

(4) There is no edge vu_1 and G without t has a 2-cut, where u_1 is no cut vertex.
Decomposition: Remove t, insert vu_1 and insert su_1 as new root edge.
Result: An f-net with one outer vertex less than G.

Hence

$$E(t,u) = E^{+}(t,u) + u\,D(t,u) + t\,F^0(t,u) + u\,F(t,u)\,. \tag{4}$$

Decomposition of e⁺-nets. Next, let G be an e⁺-net, i.e., an e-net with an edge vu_1. Again, there are four distinct cases; they are depicted in Fig. 5.

(1) The degrees of v and u_1 in G are both three.
Decomposition: Remove t and u_1, insert the edge vu_2 (which cannot exist in G) and insert sv as new root edge.
Result: An e-net with one inner and one outer vertex less than G.

(2) The degree of v in G is three and the degree of u_1 in G is at least four.
Decomposition: Remove t and insert sv as new root edge.
Result: An e-net with one inner vertex less than G.

(3) The degree of v in G is at least four, and u_1 is not a cut-vertex of any 2-cut in G without t.
Decomposition: Remove t and insert su_1 as new root edge.
Result: A c-net with one outer vertex less than G.

(4) The degree of v in G is at least four, and u_1 is a cut-vertex of a 2-cut in G without t.
Decomposition: Remove t and insert sv as new root edge.
Result: An f^0-net with one inner vertex less than G.

Hence

$$E^{+}(t,u) = tu\,E(t,u) + t\,E(t,u) + u\,C(t,u) + t\,F^0(t,u)\,. \tag{5}$$

Decomposition of f-nets and f^0-nets. Let G be an f-net, i.e., G is a c-net where the degree of t is at least four and which has a 2-cut after removing st. Because of planarity there exists a unique 2-cut vu_{j+1} $(0 \le j \le k-1)$ that is closest to t (see Figure 6). As introduced above, G is an f^0-net if $j = 0$. G without v and u_{j+1} has two components, one of which includes t and i inner vertices and the other includes s and $n-i$ inner vertices. Let G_t be the subgraph induced by v, u_{j+1} and the component containing t, and let G_s be the

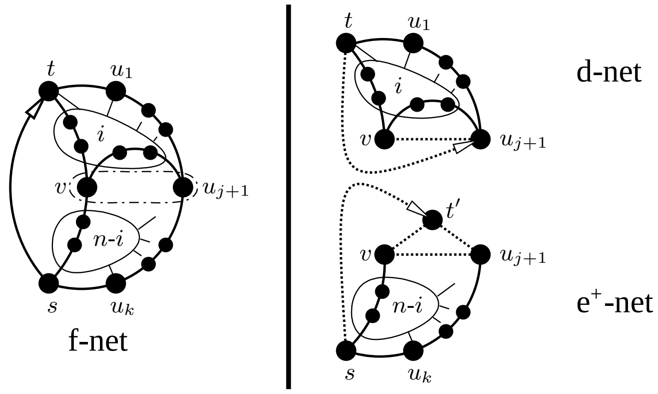


FIGURE 6. The decomposition of f-nets and f^0-nets.

subgraph induced by v, u_{j+1} and the component containing s. Note that the edge vu_{j+1} might or might not be present in G.
Decomposition: If vu_{j+1} is not an edge of G, then insert it into both G_t and G_s. Insert tu_{j+1} as root edge into G_t. Add a new vertex t' to G_s, insert the edges st', $t'v$ and $t'u_{j+1}$, and choose st' to be the root edge of G_s.
Result: G_t is an d-net with i inner and j outer vertices. G_s is an e$^+$-net with $n-i$ inner and $k-j$ outer vertices.

For given parameters i and j the choice whether vu_{j+1} is an edge of G, the choice of G_t and the choice of G_s are all independent, i.e., changing any of these choices in an f-net yields a different f-net with the same parameters. The decomposition for f^0-nets is the same, and since an f^0-net is an f-net with $j=0$, we have

$$F(t,u) = 2\, D(t,u)E^+(t,u) \tag{6}$$

$$F^0(t,u) = 2\, D(t,0)E^+(t,u)\,. \tag{7}$$

4. GENERATING FUNCTIONS

We now use the equations (1) - (7) from the decomposition to derive an algebraic equation and an explicit description for $C(t,u)$ and for $C(t) = C(t,0)$. First, we eliminate the auxiliary generating functions $D(t,u)$, $D(t,0)$, $E(t,u)$, $E^+(t,u)$, $F(t,u)$ and $F^0(t,u)$ within (1) - (7), which yields an equation in $C(t,u)$, $C(t)$, t and u:

$$0 = \big(\, g_1(t,u)\, C(t,u) + g_2(t,u)\,\big)^2 - g_3(t,u)\,, \qquad \text{where} \tag{8}$$

$$\begin{aligned}
g_1(t,u) &:= 4tu(t{+}1)(u{+}1)\,,\\
g_2(t,u) &:= 2t{+}2t^2{+}4t^3{-}u{+}tu{+}4t^3u{+}u^2{+}tu^2{-}2t^2u^2 - 2t(t{+}1)(u{+}1)(2t{+}u)\,C(t)\,,\\
g_3(t,u) &:= 4t^4(u{+}1)^2(4t^2{-}4tu{+}u^2{+}4t{-}4u{+}5){+}2tu(u^3{-}4u^2{-}3u{-}2){+}u^2(u{-}1)^2\\
&\quad +4t^3(u^4{-}5u^3{-}9u^2{-}u{+}2){+}t^2(5u^4{-}10u^3{-}15u^2{+}4)\\
&\quad + 4t^2(t{+}1)^2(u{+}1)^2(2t{-}u)^2\,C(t)^2 - 4t(t{+}1)(u{+}1)(4t^2{+}4t^3{+}8t^4\\
&\quad -4tu{-}4t^3u{+}8t^4u{-}u^2{-}5tu^2{-}2t^2u^2{-}8t^3u^2{+}u^3{+}tu^3{+}2t^2u^3)\,C(t)\,.
\end{aligned}$$

Both $C(t,u)$ and $C(t)$ appear in (8), and we cannot solve directly for one of these functions in t and u only. Setting $u=0$ we only yield the trivial equation $0=0$. Instead, we apply the quadratic method due to Tutte [24], and follow the presentation in [15]. We assume that there exists a function $u_t := u(t)$ such that $g_3(t,u_t) = 0$. Equation (8) directly yields $0 = g_3(t,u_t) = (g_1\ C + g_2)^2(t,u_t)$, hence $0 = (g_1\ C + g_2)(t,u_t)$ and then $(\frac{\partial}{\partial u} g_3)(t,u_t) = \frac{\partial}{\partial u}(g_1 C + g_2)^2(t,u_t) = \big(2(g_1 C + g_2)\frac{\partial}{\partial u}(g_1 C + g_2)\big)(t,u_t) = 0$ holds as well. We now have the following pair of simultaneous equations: $0 = g_3(t,u_t)$ and $0 = (\frac{\partial}{\partial u} g_3)(t,u_t)$, depending on $C(t)$, t and u. We eliminate u by calculating the resultant, i.e., the Sylvester determinant, of $g_3(t,u_t)$ and $(\frac{\partial}{\partial u} g_3)(t,u_t)$ with respect to u, and obtain one polynomial in $C := C(t)$ and t, the roots of which include the common roots of $g_3(t,u_t)$ and $(\frac{\partial}{\partial u} g_3)(t,u_t)$. The resultant has several nontrivial factors, but only the following factor $p(C,t)$ will be relevant for us.

$$
\begin{aligned}
p(C,t) = {}& (8t^3{+}72t^4{+}264t^5{+}504t^6{+}528t^7{+}288t^8{+}64t^9)\ C^4 \\
&+ (12t^2{-}228t^3{-}988t^4{-}1756t^5{-}2032t^6{-}1792t^7{-}1024t^8{-}256t^9)\ C^3 \\
&+ (6t{+}218t^2{+}894t^3{+}2190t^4{+}3284t^5{+}3120t^6{+}2304t^7{+}1344t^8{+}384t^9)\ C^2 \\
&+ (1{-}43t{-}337t^2{-}1021t^3{-}1828t^4{-}2404t^5{-}2128t^6{-}1344t^7{-}768t^8{-}256t^9)\ C \\
&+ ({-}1{+}36t{+}131t^2{+}350t^3{+}540t^4{+}616t^5{+}536t^6{+}304t^7{+}160t^8{+}64t^9)\,.
\end{aligned}
$$

As the order of $p(C,t)$ as a polynomial in C is four, and $p(C,t) = 0$ yields four algebraic solutions for C. Comparing initial coefficients, we find that the following is the explicit form of the generating function $C(t)$.

$$
\begin{aligned}
a =& -729 - 49113t - 61936t^2 - 137856t^3 + 6144t^4 + 8192t^5 \\
b =& (t-1)(-\frac{3}{2}(32t + 17 - 7\sqrt{7})(32t + 17 + 7\sqrt{7}))^{\frac{3}{2}} \\
s =& -3 + 2126t - 1571t^2 - 11800t^3 - 9392t^4 - 256t^5 + 1024t^6 \\
y =& -3(2^{\frac{2}{3}} 4t^{\frac{1}{3}}(1+t)(1+2t)^3((a+b)^{\frac{1}{3}} - (-1)^{\frac{1}{3}}(a-b)^{\frac{1}{3}}) + s) \\
C(t) =& \Big(3(-3 + 63t + 124t^2 + 128t^3 + 128t^4 + 64t^5) + \sqrt{y} \\
&+ \big(-9s - y + 54(1 + 2681t - 46609t^2 - 96397t^3 + 48468t^4 + 188304t^5 \\
&+ 62016t^6 - 63488t^7 - 32768t^8)/\sqrt{y}\big)^{\frac{1}{2}}\Big)/(24t(1+t)(1+2t)^3)\,.
\end{aligned}
$$

An explicit form for $C(t,u)$ can be obtained by solving equation (8) for $C(t,u)$, and substituting $C(t)$ by its explicit form.

Having the algebraic equation at hand, we can apply singularity analysis: The dominant singularity lies in the exceptional set of the algebraic curve, and can be computed by evaluating the resultant R of $p(C,t)$ and $\frac{\partial}{\partial C}p(C,t)$ with respect to C. The solutions for t in the equation $R = 0$ can be computed symbolically with Mathematica, and the smallest real solution t_0, where additionally the equations $p(C,t_0) = 0$ and $\frac{\partial}{\partial C}p(C,t_0) = 0$ have a simultaneous solution, is a dominant singularity of $C(t)$. In this way, it is easy to compute the dominant singularity of $C(t)$, which is at $t_0 = 1/32(7\sqrt{7} - 17) \doteq 0.047508$ (that was computed before from the equations of Mullin and Schellenberg; see [2]), and proves the following.

Theorem 1 (essentially from [2]). *The number of c-nets $c(n)$ is in $(1/t_0)^{n+o(n)}$), where $1/t_0 = 16/27(17 + 7\sqrt{7}) \doteq 21.049042$.*

Using the Maple package GFUN [20], the algebraic equation $p(C,t)$ can be transformed automatically into a linear differential equation with polynomial coefficients, which in turn translates to a one parameter recurrence formula for c_n. Using Horner's method and this formula we computed the value of $c(100000)$ in 100 seconds on a PC.

Theorem 2. *For the coefficients $c(n)$ of $C(t)$ the following recursion holds.*

$$
\begin{aligned}
c(0) &= 0,\ c(1) = 1,\ c(2) = 7,\ c(3) = 73,\ c(4) = 879,\ c(5) = 11713,\\
&\quad c(6) = 167423,\ c(7) = 2519937,\ \textit{and for } n \geq 8\,,\\
c(n) &= \Big((42147840 + 49975296(n-7) + 19267584(n-7)^2 + 2408448(n-7)^3)\ c(n-7)\\
&\quad + (291529728 + 269461504(n-7) + 83615232(n-7)^2 + 8692736(n-7)^3)\ c(n-6)\\
&\quad + (533308032 + 435701440(n-7) + 119431200(n-7)^2 + 11026784(n-7)^3)\ c(n-5)\\
&\quad + (259749888 + 220560168(n-7) + 59988636(n-7)^2 + 5361276(n-7)^3)\ c(n-4)\\
&\quad + (-45552288 - 9821452(n-7) + 1941468(n-7)^2 + 418816(n-7)^3)\ c(n-3)\\
&\quad + (-16057320 - 11696062(n-7) - 2582841(n-7)^2 - 180467(n-7)^3)\ c(n-2)\\
&\quad + (5063688 + 2370408(n-7) + 367734(n-7)^2 + 18930(n-7)^3)\ c(n-1) \Big)\\
&\quad \Big/\ (255024 + 99918(n-7) + 13041(n-7)^2 + 567(n-7)^3)\,.
\end{aligned}
$$

5. SAMPLING

We now discuss how to use the presented decomposition to sample c-nets uniformly at random. (As usual, $\tilde{O}(\cdot)$ denotes growth up to logarithmic factors.) Note that the analysis of [13] applies to expected running time, whereas our bound is deterministic. Moreover, they have parameters for vertices and faces, whereas we have parameters for the number of vertices and the size of the outer face. Thus the results are not directly comparable. Their upper bound is $O(n^4)$ for n vertices, and reduces to $O(n)$ if the ratio of vertex number to face number is fixed to a constant. The worst case is attained for triangulations.

Theorem 3. *There exists a deterministic polynomial time algorithm to sample c-nets on a given number of vertices and a given number of vertices on the outer face uniformly at random. The algorithm runs in $\tilde{O}(n^5)$ time and $O(n^3)$ space. If we allow a pre-computation, the algorithm can sample a c-net in $\tilde{O}(n^2)$ time and $O(n^5)$ space.*

Proof. The decomposition yields recursive counting functions for c-nets, d-nets, e-nets, e$^+$-nets, f-nets, and f^0-nets. For all $n, k \geq 0$:

$$
\begin{aligned}
c(n,k) &= \begin{cases} 1 & \text{if } n = k = 0\,,\\ d(n,k) + e(n,k) + f(n,k) & \text{else.} \end{cases}\\
d(n,k) &= c(n-1,k+1) + d(n-1,k+1)\,.\\
e(n,k) &= e^+(n,k) + d(n,k-1) + f^0(n-1,k) + f(n,k-1)\,.\\
e^+(n,k) &= e(n-1,k-1) + e(n-1,k) + c(n,k-1) + f^0(n-1,k)\,.\\
f(n,k) &= 2\ \sum_{i=0}^{n} \sum_{j=0}^{k} d(i,j) e^+(n-i,k-j)\,.\\
f^0(n,k) &= 2\ \sum_{i=0}^{n} d(i,0) e^+(n-i,k)\,.
\end{aligned}
$$

By induction on the lexicographically ordered pair (n, k), the decomposition reduces to the initial case within $O(nk)$ steps of recursion. Hence we can evaluate the functions using dynamic programming. The representation size of all computed numbers is linear, because it is bounded by the logarithm of the number of unlabeled c-nets, which is $O(2^{O(n)})$ according to Theorem 1. Note that the functions d, e, e^+, f, and f^0 are at most as large as c according to their definitions. Since we employ a constant number of two-dimensional tables, the total space requirement is $O(n^3)$. Concerning the running time, each summation runs over at most two indices, and for each summand we have to perform one multiplication with $O(n)$ bit numbers. We assume an $O(n \log n \log \log n)$ multiplication algorithm (see e.g. [7]). Thus the running time for the computation of the values is within $\tilde{O}(n^5)$.

The values in the dynamic programming tables can be used to make the correct probabilistic decisions in a recursive construction of c-nets, which is essentially the inversion of the presented decomposition – this method is standard and known as the *recursive method* for sampling [8,11,19]. For each entry, we scan over all the entries from which it was computed (there are at most n^2 of them). We compute partial sums in another pass over these entries and build a balanced binary tree, where in each internal node the maximum over its left-hand siblings is stored. This will take $O(n^5)$ time in total, since we have $O(n^2)$ table entries, each with $O(n^2)$ dependencies, and each tree node stores an $O(n)$ bit number. After that, when given a random number between 1 and the maximum (i. e., the value of the entry for which the tree was built), we can find the corresponding table entry in one pass through the tree, while reading each bit of the random number only a constant number of times, and hence in $O(n)$ time. Then the procedure calls itself recursively. In the case of f and f^0, we have to trace back two separate lines, as the random sibling corresponds to a choice of the summation indices i (for f), respectively (i, j) (for f^0) and the actual summand is a product of two entries (e.g. $d(i, j)$ and $e^+(n - i, k - j)$ for $f(n, k)$ and (i, j)). Note that the sum of the bit lengths of both factors is linear in the bit length of the entry. It follows that the total running time for generating the decomposition tree is $\tilde{O}(n^2)$ (details omitted due to lack of space). If the decomposition tree is stored appropriately, we can output the sampled random graph in $O(n)$ time.

It is not necessary to create the binary trees physically for each entry of the tables. Instead, we can just redo the computations from the preprocessing and stop if the partial sum exceeds the random number. This way, the algorithm uses $\tilde{O}(n^5)$ time and $O(n^3)$ space. $\square$

To sample unlabeled, unrooted 3-connected planar graphs uniformly at random, we apply *rejection sampling*. That is, we generate a c-net uniformly at random, but the resulting graph is accepted only with a probability that is inverse proportional to the size of the orbit of the root edge together with an incident face in the automorphism group of the graph. (It is well-known that the automorphism group of a planar graph can be computed efficiently, see e.g. [16].) If we do not output the graph, we restart the algorithm. Clearly, the output of this procedure are uniform random samples from the class of all 3-connected planar graphs. Since a 3-connected planar graph has with high probability a trivial automorphism group [3], the expected number of restarts is constant.

Corollary 1. *Using rejection sampling, we can sample 3-connected planar graphs using the algorithm of Theorem 3 in an expected constant number of rounds.*

6. Conclusion

Our main structural result is a new decomposition of rooted 3-connected planar graphs, which can easily be expressed in terms of recursive counting formulas, or equations for

$c(n,k)$	0	1	2	3	4	5	$k=6$
0	1	6	56	640	8256	115456	1710592
1	1	16	208	2848	41216	624384	9812992
2	1	30	560	9440	156592	2613664	44169600
3	1	48	1240	25864	496944	9234368	169378560
4	1	70	2408	61712	1377600	28663040	574139904
5	1	96	4256	132480	3430528	80104448	1758695424
6	1	126	7008	261648	7826544	205083936	4944057984
7	1	160	10920	483080	16600944	487362496	12906193920
8	1	198	16280	843744	33111232	1086226944	31579350528
9	1	240	23408	1406752	62659200	2289692416	72985375744
$n=10$	1	286	32656	2254720	113313200	4596347808	160355238784

$c(n,k)$	7	8	9	$k=10$
0	26468352	423641088	6966960128	117148778496
1	158883840	2636197888	44640468992	769058340864
2	756712960	13136471040	230851792896	4102116843520
3	3095526912	56624998400	1039080697856	19147850612736
4	11259283200	218198045184	4201424145408	80643838062592
5	37158281984	765948707328	15534537453568	311681600004096
6	112834665216	2481031718144	53154302311936	1117907385569280
7	318621198720	7487670554880	169818439763968	3751908804540416
8	843790483712	21217661003264	510172604564480	11860405982539776
9	2110406347008	56815355557376	1449735177678848	35506327812194304
$n=10$	5014608178944	144547875949568	3916271978577920	101129913041264640

FIGURE 7. A table of $c(n,k)$ for small c-nets on up to 23 vertices. The number of vertices on the outer face is $k+2$. The total number of vertices is $n+k+3$.

their generating functions. We use these equations to derive an algebraic equation of degree four that determines the generating function for the number of rooted 3-connected planar graphs on n vertices. Here we apply computer algebra systems, and also derive a single parameter recurrence formula, which allows to compute these numbers for much larger n than the previously known formulas of Mullin and Schellenberg [18].

The main algorithmic result is the first deterministic polynomial time algorithm to sample c-nets with a given number of vertices and a given size of the outer face uniformly at random. Since the recurrences of the decomposition do not involve any subtractions, the decomposition immediately translates into a sampling algorithm that produces a rooted 3-connected planar graph uniformly at random. The recursive counting formulas were implemented by top-down dynamic programming in C++ using the GMP library for exact arithmetic [14]. A table for small values of n and k is given in Figure 7.

It is fairly straightforward to see that the decomposition can be refined to control more parameters of the graph, e. g., the number of edges, or the degree of a root vertex. Each parameter comes at the cost of another dimension in the tables and hence increases the precomputation time by a quadratic factor. The recursive counting formulas with an additional parameter for the number for edges were also implemented, and we used the numbers of Mullin and Schellenberg [18] to check both implementations.

The algorithm can be used to obtain a faster and now fully deterministic polynomial time sampler for labeled planar graphs [4]. Also, using the rejection sampling method, we obtain an expected polynomial time algorithm for 3-connected planar graphs (isomorphism types of convex polyhedra). In forthcoming work, we apply the n,k,m-recurrence and rejection sampling to generate 3-connected planar graphs with a sense-reversing automorphism, and unlabeled 2-connected planar graphs [5].

REFERENCES

1. Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michèle Soria, *Random maps, coalescing saddles, singularity analysis, and Airy phenomena*, Random Structures and Algorithms **19** (2001), 194–246.
2. E.A. Bender and L.B.Richmond, *The asymptotic enumeration of rooted convex polyhedra*, Journal of Combinatorial Theory (1984), no. 36, 276–283.
3. Edward A Bender and N.C. Wormald, *Almost all convex polyhedra are asymmetric*, Can. J. Math. **27** (1985), no. 5, 854–871.
4. M. Bodirsky, C. Gröpl, and M. Kang, *Generating labeled planar graphs uniformly at random*, Proceedings of Thirtieth International Colloquium on Automata, Languages and Programming (ICALP'03), 2003, pp. 1095–1107.
5. ______, *Sampling unlabeled 2-connected planar graphs uniformly at random*, Preprint, 2004.
6. Mireille Bousquet-Mélou, *On (some) functional equations arising in enumerative combinatorics*, extended abstract for FPSAC'01, 2001.
7. P. Bürgisser, M. Clausen, and M.A. Shokrollahi, *Algebraic complexity theory*, Grundlehren der mathematischen Wissenschaften, no. 315, Springer Verlag, 1997.
8. Alain Denise and Paul Zimmermann, *Uniform random generation of decomposable structures using floating-point arithmetic*, Theoretical Computer Science **218** (1999), 233–248.
9. Reinhard Diestel, *Graph theory*, Springer–Verlag, New York, 1997.
10. P.J. Federico, *The number of polyhedra*, Philips Res. Repts **30** (1975), 220–231.
11. P. Flajolet, P. Zimmerman, and B. Van Cutsem, *A calculus for the random generation of labelled combinatorial structures*, Theoretical Computer Science **132** (1994), no. 1-2, 1–35.
12. Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics — symbolic combinatorics*, Book in preparation (Oct 2004), and Rapport de recherche 4103, INRIA, 2001.
13. Eric Fusy, Dominique Poulalhon, and Gilles Schaeffer, *Dissections and trees: applications to optimal mesh encoding and random sampling*, Accepted for publication in the proceedings of SODA'05, 2005.
14. *The GNU multiple precision arithmetic library, version 4.1.2.*, http://swox.com/gmp/.
15. Ian P. Goulden and David M. Jackson, *Combinatorial enumeration*, John Wiley, New York, 1983.
16. J.E. Hopcroft and R.E. Tarjan, *A v^2 algorithm for determining isomorphism of planar graphs*, Information Processing Letters **1** (1971), no. 1, 32–34.
17. V. A. Liskovets and T. R. Walsh, *Counting non-isomorphic 2-connected planar maps*, Graph Theory Newsletter **9** (1980), no. 6, 3.
18. R.C. Mullin and P.J. Schellenberg, *The enumeration of c-nets via quadrangulations*, Journal of Combinatorial Theory **4** (1968), 259–276.
19. A. Nijenhuis and H.S. Wilf, *Combinatorial algorithms*, Academic Press Inc., 1979.
20. Bruno Salvy and Paul Zimmermann, *GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable*, ACM Transactions on Mathematical Software **20** (1994), no. 2, 163–177.
21. Gilles Schaeffer, *Conjugaison d'arbres et cartes combinatoires aléatoires*, Ph.D. thesis, Université Bordeaux I, 1998.
22. Gilles Schaeffer, *Random sampling of large planar maps and convex polyhedra*, Proc. of the thirty-first annual ACM symposium on theory of computing (STOC'99) (Atlanta, Georgia), ACM press, May 1999, pp. 760–769.
23. E. Steinitz, *Polyeder und Raumeinteilungen*, Encyclopädie der mathematischen Wissenschaften **3** (1922), no. 9.
24. W. T. Tutte, *A census of planar triangulations*, Canad. J. Math. **14** (1962), 21–38.
25. ______, *A census of planar maps*, Canad. J. Math. **15** (1963), 249–271.
26. ______, *Graph theory*, Cambridge University Press, 1984.
27. T. Walsh and V. A. Liskovets, *Ten steps to counting planar graphs*, Proc. of Eighteenth Southeastern International Conference on Combinatoris, Graph Theory, and Computing, Congr. Numer., 1987, pp. 269–277.
28. Timothy Walsh, *Counting non-isomorphic three-connected planar maps*, J. Combin. Theory **32** (1982), 33–44.
29. Nicholas C. Wormald, *On the number of planar maps*, Canad. J. Math. **33** (1981), no. 1, 1–11.

MANUEL BODIRSKY, HUMBOLDT-UNIVERSITÄT ZU BERLIN, GERMANY
E-mail address: `bodirsky@informatik.hu-berlin.de`

CLEMENS GRÖPL, FREIE UNIVERSITÄT BERLIN, GERMANY
E-mail address: `groepl@inf.fu-berlin.de`

DANIEL JOHANNSEN, HUMBOLDT-UNIVERSITÄT ZU BERLIN, GERMANY
E-mail address: `johannse@informatik.hu-berlin.de`

MIHYUN KANG, HUMBOLDT-UNIVERSITÄT ZU BERLIN, GERMANY
E-mail address: `kang@informatik.hu-berlin.de`

The limit of a Stanley-Wilf sequence is not always rational, and layered patterns beat monotone patterns

Miklós Bóna *

Abstract

We show the first known example for a pattern q for which $L(q) = \lim_{n\to\infty} \sqrt[n]{S_n(q)}$ is not an integer, where $S_n(q)$ denotes the number of permutations of length n avoiding the pattern q. We find the exact value of the limit and show that it is irrational, but algebraic. Then we generalize our results to an infinite sequence of patterns. We provide further generalizations that start explaining why certain patterns are easier to avoid than others. Finally, we show that if q is a layered pattern of length k, then $L(q) \geq (k-1)^2$ holds.

1 Introduction

Let $S_n(q)$ be the number of permutations of length n (or, in what follows, n-permutations) that avoid the pattern q. For a brief introduction to the area of pattern avoidance, see [5]; for a more detailed introduction, see [6]. A recent spectacular result of Marcus and Tardos [9] shows that for any pattern q, there exists a constant c_q so that $S_n(q) < c_q^n$ holds for all n. As pointed out by Arratia in [2], this is equivalent to the statement that $L(q) = \lim_{n\to\infty} \sqrt[n]{S_n(q)}$ exists. Let us call the sequence $\sqrt[n]{S_n(q)}$ a *Stanley-Wilf* sequence. It is a natural and intriguing question to ask what the limit $L(q)$ of a Stanley-Wilf sequence can be, for various patterns q.

The main reason this question has been so intriguing is that in all cases where $L(q)$ has been known, it has been known to be an *integer*. Indeed, the results previously known are listed below.

1. When q is of length three, then $L(q) = 4$. This follows from the well-known fact [11] that in this case, $S_n(q) = \binom{2n}{n}/(n+1)$.
2. When $q = 123\cdots k$, or when q is such that $S_n(q) = S_n(12\cdots k)$, then $L(q) = (k-1)^2$. This follows from an asymptotic formula of Regev [10].
3. When $q = 1342$, or when q is such that $S_n(q) = S_n(1342)$, then $L(q) = 8$. See [4] for this result and an exact formula for the numbers $S_n(1342)$.

In this paper, we show that $L(q)$ is *not* always an integer. We achieve this by proving that $14 < L(12453) < 15$. Then we compute the exact value of this limit, and see that it is not even rational; it is the number $9+4\sqrt{2}$. We compute the limit of the Stanley-Wilf sequence for an infinite sequence of patterns, and see that as the length k of these patterns grows, $L(q)$ will fall further and further below $L(12\cdots k) =, (k-1)^2$. Finally, we show

*University of Florida, Gainesville FL 32611-8105. Partially supported by an NSA Young Investigator Award. Email: bona@math.ufl.edu.

that while for certain patterns, our methods provide the exact value of the limit of the Stanley-Wilf sequence, for certain others they only provide a *lower bound* on this limit. This starts explaining why certain patterns are easier to avoid than others. Among other results, we will confirm a seven-year old conjecture by proving that in the sense of logarithmic asymptotics, a layered pattern q is always easier to avoid than the monotone pattern of the same length.

2 Proving an upper bound

Let $p = p_1 p_2 \cdots p_n$ be a permutation. Recall that p_i is called a *left-to-right minimum* of p if $p_j > p_i$ for all $j < i$. In other words, a left-to-right minimum is an entry that is smaller than everything on its left. Note that p_1 is always a left-to-right minimum, and so is the entry 1 of p. Also note that the left-to-right minima of p always form a decreasing sequence. For the rest of this paper, entries that are not left-to-right minima are called *remaining entries*.

Now we are in a position to prove our promised upper bound for the numbers $S_n(12453)$.

Lemma 2.1 *For all positive integers n, we have*

$$S_n(12453) < (9 + 4\sqrt{2})^n < 14.66^n.$$

Proof: Let p be a permutation counted by $S_n(12453)$, and let p have k left-to-right minima. Then we have at most $\binom{n}{k}$ choices for the set of these left-to-right minima, and we have at most $\binom{n}{k}$ choices for their positions. The string of the remaining entries has to form a 1342-avoiding permutation of length $n-k$. Indeed, if there was a copy $acdb$ of 1342 among the entries that are not left-to-right minima, then we could complete it to a 12453 pattern by simply prepending it by the closest left-to-right minimum that is on the left of a. The number of 1342-avoiding permutations on $n-k$ elements is less than 8^{n-k} as we know from [4]. This shows that

$$\begin{aligned} S_n(12453) &< \sum_{k=1}^{n} \binom{n}{k}^2 \cdot 8^{n-k} \\ &< \sum_{k=1}^{n} \left(\binom{n}{k} \cdot \sqrt{8}^{n-k} \right)^2 \leq \left(\sum_{k=1}^{n} \binom{n}{k} \cdot \sqrt{8}^{n-k} \right)^2 \\ &< (1+\sqrt{8})^{2n} = (9+4\sqrt{2})^n, \end{aligned}$$

and the proof is complete. $\diamond$

Corollary 2.2 *We have*

$$L(12453) \leq 9 + 4\sqrt{2} < 14.66.$$

3 Proving a lower bound

We have seen in Corollary 2.2 that $L(12453) \leq 9 + 4\sqrt{2} < 14.66$. In order to prove that this limit is not an integer, it suffices to show that it is larger than 14. In what follows, we are going to work toward a good lower bound for the numbers $S_n(12453)$, and thus the number $L(12453)$.

Where is the waste in the proof of the upper bound in the previous section? The waste is that there are some choices for the left-to-right minima that are incompatible with some choices for the 1342-avoiding permutation of the remaining entries. This is a crucial concept of the upcoming proof, so we will make it more precise.

We have mentioned in the previous section, that determining the left-to-right minima of a permutation p means to determine the set T of positions where these minima will be, and to determine the set Z of entries that are the left-to-right minima. In other words, the ordered pair (T, Z) of equal-sized subsets of $[n] = \{1, 2, \cdots, n\}$ describes the left-to-right minima of p.

Definition 3.1 *Let n be a positive integer, and let $m \leq n$ be a positive integer. Let T and Z be two m-element subsets of $[n]$. Finally, let S be a permutation of the elements of the set $[n] - Z$. If there exists an n-permutation p so that its left-to-right minima are precisely the elements of Z, they are located in positions belonging to T, and its string of remaining entries is S, then we say that the triple (T, Z, S) is* compatible. *Otherwise, we say that the triple (T, Z, S) is* incompatible.

Clearly, if (T, Z, S) is *compatible*, then there is exactly one permutation p satisfying all criteria specified by (T, Z, S).

Example 3.2 *If $n = 4$, and $T = \{1, 3\}$, $Z = \{1, 2\}$, and $S = 43$, then (T, Z, S) is* compatible *as shown by the permutation 2413.*

Example 3.3 *If $n = 4$, and $T = \{1, 3\}$, $Z = \{1, 3\}$, and $S = 24$, then (T, Z, S) is* incompatible. *Indeed, the only permutation allowed by T and S is 3214, but for this permutation $Z = \{1, 2, 3\}$, not $\{1, 3\}$.*

Returning to the method by which we proved our upper bound for $L(12453)$, we will show that in a sufficient number of cases, our triples (T, Z, S) are compatible. This will show that the upper bound is quite close to the precise value of $L(12453)$.

What is a good way to check that a particular choice (T, Z) of left-to-right minima is compatible with a particular choice of S? For shortness, let us call the procedure of putting together S and a string (T, Z) of left-to-right minima *merging*. One has to check that in the permutation obtained by merging our left-to-right minima with S, the left-to-right minima are indeed the entries in Z. That is, there are no additional left-to-right minima, and the entries in Z are indeed all left-to-right minima. This is achieved exactly when any remaining entry is larger than the closest left-to-right minimum on its left.

In our efforts to find a good lower bound on $L(12453)$, we will only consider a special kind of permutations. Let N be a positive integer so that $S_n(1342) > 7.99^n$ for all $n \geq N$. (We know from [4] that such an N exists as $L(1342) = 8$.)

Assume first that the string S of remaining entries of our permutations has length s, where s is divisible by N. Consider permutations having the following additional property. If we cut S into s/N blocks of consecutive entries of length N each, then the entries of any given block B are all smaller than the entries of any block on the left of B, and larger than the entries of any block on the right of B. Let us call these strings S *block-structured.* See Fig. 1 for the generic diagram of a block-structured string in the (unrealistic) case of $N = 2$.

If s is not divisible by N, that is, when $s = Nt + r$ for some $r \in [1, N - 1]$, then we call S block-structured if its last r entries are its smallest entries, and they are in decreasing order, and its first $s - r$ entries have the block-structured property in the above sense. For instance, for $N = 3$ and $s = 8$, the string 798|645|21 is block structured. As the last r entries must be in decreasing order, we will not call their string a block.

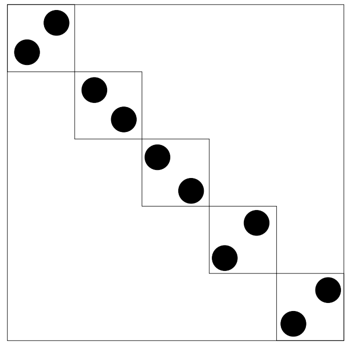


Figure 1: A block-structured string.

Let S be a block-structured string in which each block is a 1342-avoiding substring. It is then clear that S itself is 1342-avoiding as a 1342-pattern cannot start in a block and end in another one. The definition of N implies that we have more than 7.99^N choices for the substring of each block. Therefore, we have at least 7.99^{s-r} block-structured strings S of length s that avoid 1342. (Recall that r is the remainder of s modulo N). As $r < N$, this implies that the number of block-structured strings of length s is always more than $\frac{1}{7.99^N} \cdot 7.99^s = c \cdot 7.99^s$, for an absolute constant c. (The constant c will become insignificant when we take nth roots.)

We claim that a sufficient number of these strings S will be compatible with a sufficient number of the choices (T, Z) of left-to-right minima.

First, look at the very special case when S is *decreasing*. In this case, we will write S^{dec} instead of S. Now our permutation p consists of two decreasing sequences (so it is 123-avoiding), namely the left-to-right minima and S^{dec}. The following Proposition is very well-known.

Proposition 3.4 *Let $1 \leq m \leq n$. Then the number of 123-avoiding n-permutations having exactly m left-to-right minima is*

$$A(n,m) = \frac{1}{n}\binom{n}{m}\binom{n}{m-1}, \tag{1}$$

a Narayana *number.*

For a proof, see [12] or [6].

The significance of this result for us is the following. If we just wanted to merge (T, Z) and S^{dec} together, with no regard to the existing constraints, the total number of ways to do that would be of course at most $\binom{n}{m} \cdot \binom{n}{m}$. The above formula shows that roughly $\frac{1}{n}$ of these mergings will actually be good, that is, they will not violate any constraints, they will lead to compatible triples (T, Z, S). The factor $\frac{1}{n}$ is not a significant loss from our point of view, since $\lim_{n\to\infty} \sqrt[n]{1/n} = 1$.

Now let us return to the general case of block-structured strings S. In other words, take a 123-avoiding n-permutation (T, Z, S^{dec}), and replace its string S^{dec} by a block-structured string S *taken on the entries that belong to S^{dec}*. We claim that after this replacement, a sufficient number of triples (T, Z, S) will be compatible.

Here is the outline of the proof of that claim. Because of the definition of a block-structured S, it is true that every entry in S is at most N positions away from the position it was in S^{dec}. (We will take the left-to-right minima into account next.) Therefore, if we merge (T, Z) and S^{dec} together so that each left-to-right minimum

y is not only smaller than all entries on its left, and smaller than all remaining entries located between y and the closest left-to-right minimum y' on the right of y, but also smaller than the N closest remaining entries on the right of y', then we will be done. Indeed, in this case replacing S^{dec} by any block-structured string S will not violate any constraints since no remaining entry moves up by more than N slots in the string of remaining entries.

For example, set $N = 3$ (which is unrealistic because in reality, N needs to be much larger). Then the permutation 592871643 has the desired property. Indeed, each left-to-right minimum y of this permutation is smaller than the three remaining entries immediately following the left-to-right minimum y' that comes after y. That is, 5 is smaller than 8, 7 and 6, and 2 is smaller than 6, 4, and 3. (The condition always vacuously holds for 1.) Therefore, if we rearrange the string 987643 so that no entry moves up by more than three slots within this string, then no constraints will be violated, that is, the obtained permutation will still have left-to-right minima 5, 2, and 1.

Therefore, we will have a lower bound for the number of compatible triples (T, Z, S) if we find a lower bound for the number of compatible triples (T, Z, S^{dec}) in which each left-to-right minimum has the mentioned stronger property.

In order to find such a lower bound, take a 123-avoiding permutation p' which is of length $n - N$. Let p' have m left-to-right minima. Denote (T', Z') the string of the left-to-right minima of p', and let $S^{dec'}$ denote the decreasing string of remaining entries of p'. Now prepend p' with the decreasing string taken on the N-element set $\{n - N + 1, n - N + 2, \cdots, n\}$, to get an n-permutation. In this n-permutation, move each of the original m left-to-right minima of p' to the left by N positions. Let us call the obtained n-permutation p''.

For example, with $n = 9$, $N = 3$, and $p' = 456123$, we first prepend p' with the string 987, to get 987456123, then move the original two left-to-right minima of p', that is, the entries 4 and 1, to the left by three positions, to get $p'' = 498715623$. Note that means that the set of positions of the left-to-right minima of p'' is actually still T'.

It is then clear that the left-to-right minima of p'' are the same as the left-to-right minima of p'. Furthermore, because of the translation we used to create our new permutation, p'' has the property that if y and y' are two left-to-right minima so that y' is the closest left-to-right minimum on the right of y, then y is smaller than the N remaining entries immediately on the right of y'. Indeed, these N remaining entries were on the right of y' in the original permutation p'.

Now we can use the argument that we outlined five paragraphs ago. For easy reference, we sketch that argument again. If S^{dec} is replaced by any block-structured permutation of the same size taken on the same set of elements, (resulting in the n-permutation $p*$) then each remaining entry x will move within its block only, that is, x will move at most N positions from its original position in the string S of remaining entries. Therefore, x will still be larger than the left-to-right minimum closest to it and preceding it.

This shows that if p' and (T', Z') lead to a compatible triple, then so too will $p*$ and (T, Z), where (T, Z) describes the left-to-right minima of $p*$. Proposition 3.4 implies that the number of compatible triples (T', Z', p') is $\frac{1}{n-N}\binom{n-N}{m}\binom{n-N}{m-1}$. As N is a constant, we have

$$\lim_{n\to\infty} \sqrt[n]{\frac{1}{n-N}\binom{n-N}{m}\binom{n-N}{m-1}} = \lim_{n\to\infty} \sqrt[n]{\binom{n}{m}\binom{n}{m}}. \tag{2}$$

Now restrict our attention to the particular case when $m = \lfloor n/3 \rfloor$. We claim that permutations of this particular type are sufficiently numerous to provide the lower bound we need. Using Stirling's formula, a routine computation yields that in this case, we have

$$\lim_{n\to\infty} \sqrt[n]{\binom{n}{m}\binom{n}{m}} = \lim_{n\to\infty} \sqrt[n]{\left(\frac{3^n}{2^{2n/3}}\right)^2} \geq 1.88^2.$$

Besides, we have more than $c \cdot 7.99^{2n/3}$ choices for the block-structured string S by which we replace S^{dec}. Therefore, we have proved the following lower bound.

Lemma 3.5 *For n sufficiently large, the number of n-permutations of length n that avoid the pattern 12453 is larger than*

$$1.88^{2n} \cdot c \cdot 7.99^{2n/3} \geq c \cdot 14.12^n,$$

where $c = 7.99^{-N}$.

Lemma 3.5 and Corollary 2.2 together immediately yield the following.

Theorem 3.6 *We have*

$$14.12 \leq L(12453) \leq 14.66.$$

In particular, $L(12453) = \lim_{n\to\infty} \sqrt[n]{S_n(12453)}$ is not an integer.

4 The exact value of $L(12453)$

If we are a little bit more careful with our choice of m in the argument of the previous section, we can find the exact value of $L(12453)$. It turns out to be the upper bound proved in Corollary 2.2.

Theorem 4.1 *We have $L(12453) = (1+\sqrt{8})^2 = 9 + 4\sqrt{2}$.*

Proof: The above argument works for any $1 \leq m \leq n-N$ instead of $m = \lfloor n/3 \rfloor$, and for any positive real number $8 - \epsilon < 8$ instead of 7.99. The inequality generalizing Lemma 3.5 we get is

$$S_n(12453) \geq c \cdot \frac{1}{n-N}\binom{n-N}{m}\binom{n-N}{m-1}(8-\epsilon)^{n-m}, \tag{3}$$

where $c = (8-\epsilon)^{-N}$.

Taking (3) for all $m \in [1, n-N]$, then summing all the obtained inequalities, we get

$$(n-N)S_n(12453) \geq \frac{c}{n-N}\sum_{m=1}^{n-N}\binom{n-N}{m}\binom{n-N}{m-1}(8-\epsilon)^{n-m}. \tag{4}$$

By a routine computation, we see that

$$\binom{n-N}{m-1} \geq \frac{1}{n-N}\binom{n-N}{m}.$$

The last inequality and (4) together yield

$$(n-N)S_n(12453) \geq \frac{c}{(n-N)^2}\sum_{m=1}^{n-N}\binom{n-N}{m}^2(8-\epsilon)^{n-m}.$$

Finally, if $n > N$, then clearly, $S_n(12453) \geq S_n(1342) \geq (8-\epsilon)^n$. Comparing this to the last inequality, we get

$$(n-N+1)S_n(12453) \geq c \cdot \frac{(8-\epsilon)^N}{(n-N)^2} \sum_{m=0}^{n-N} \binom{n-N}{m}^2 (8-\epsilon)^{n-N-m}. \tag{5}$$

Let us now resort to the well-known Cauchy-Schwarz inequality stating that if $a_1, a_2, \cdots, a_d$ are positive real numbers, then

$$\frac{1}{d}(a_1 + a_2 + \cdots + a_d)^2 \leq a_1^2 + a_2^2 + \cdots + a_d^2. \tag{6}$$

The right-hand side of (5) can be viewed as the sum of $(n-N+1)$ squares, namely the squares of the positive real numbers $\binom{n-N}{m}\sqrt{(8-\epsilon)^{n-N-m}}$. Therefore, setting $d = n-N+1$, we can apply (6) to the sum on the right-hand side of (5), which then leads us to the inequality

$$\begin{aligned} \frac{1}{n-N+1}(1+\sqrt{8-\epsilon})^{2(n-N)} &= \frac{1}{n-N+1}\left(\sum_{m=0}^{n-N}\binom{n-N}{m}\sqrt{8-\epsilon}^{\,n-N-m}\right)^2 \\ &\leq \sum_{m=0}^{n-N}\binom{n-N}{m}^2 (8-\epsilon)^{n-N-m}. \end{aligned}$$

Comparing this with (5), we see that

$$S_n(12453) \geq c \cdot \frac{(8-\epsilon)^N}{(n-N+1)^4}(1+\sqrt{8-\epsilon})^{2(n-N)}.$$

Taking nth roots, then taking limits as n goes to infinity, we see that

$$L(12453) \geq (1+\sqrt{8-\epsilon})^2$$

for any positive ϵ, proving our claim. ◇

5 Some generalizations

In this Section, we will provide some interesting generalizations of our results. We will need the following simple recursive properties of pattern avoiding permutations.

Proposition 5.1 *Let q be a pattern of length k, and let q' be the pattern of length $k+1$ that is obtained from q by adding 1 to each entry of q and prepending it with 1. Let p be a permutation whose string of remaining entries is S. Then the following hold.*

1. *If S avoids q, then p avoids q'.*
2. *If q itself starts with 1, then p avoids q' if and only if S avoids q.*

Iteratively applying part 2 of Proposition 5.1, and the method explained in the previous sections, we get the following theorem.

Theorem 5.2 *Let $k \geq 4$, and let q_k be the pattern $12\cdots(k-3)(k-1)k(k-2)$. So $q_4 = 1342$, $q_5 = 12453$, and so on. Then we have*

$$L(q_k) = (k-4+\sqrt{8})^2.$$

Proof: Induction on k. For $k = 4$, the result is proved in [4], and for $k = 5$, we have just proved it in the previous section. Assuming that the statement is true for k, we can prove the statement for $k+1$ the very same way we proved it for $k = 5$, using the result for $k = 4$, and part 2 of Proposition 5.1. $\diamond$

The method we used to prove Lemma 2.1 can also be used to prove the following recursive result.

Lemma 5.3 *Let q be a pattern of length k that starts with 1, and let q' be the pattern of length $k+1$ that is obtained from q by adding 1 to each entry of q and prepending it with 1. Let c be a constant so that $S_n(q) < c^n$ for all n. Then we have*

$$S_n(q') < (1+\sqrt{c})^{2n} = (1+c+2\sqrt{c})^n.$$

This is an improvement of the previous best result [7], that only showed $S_n(q') < (4c)^n$.

The following generalization of Theorem 4.1 can be proved just as that Theorem is.

Theorem 5.4 *Let q and q' be as in Lemma 5.3. Then we have*

$$L(q') = 1 + L(q) + 2\sqrt{L(q)}.$$

In a sense, this result generalizes Regev's result [10] that showed that $L(12\cdots k) = (k-1)^2$. Our result shows that this particular growth rate, that is, that $\sqrt{L(q)}$ grows by one as the pattern grows by one, is not limited to monotone patterns.

An interesting consequence of this Theorem is that if q is as above, and $L(q) < (k-1)^2$, in other words, q is harder (or easier, for that matter) to avoid than the monotonic pattern of the same length, then repeatedly prepending q with 1 will not change this. That is, the obtained new patterns will still be more difficult to avoid than the monotonic pattern of the same length.

Are the methods presented in this paper useful at all if the pattern q does not start in the entry 1? We will show that for most patterns q, the answer is in the affirmative, as far as a *lower* bound is concerned. Let us say that the pattern q is *indecomposable* if it cannot be cut into two parts so that all entries on the left of the cut are larger than all entries on the right of the cut. For instance, 1423 and 3142 are indecomposable, but 3412 is not as we could cut it after two entries. Therefore, we call 3412 *decomposable.* It is routine to verify that as k grows, the ratio of indecomposable patterns among all $k!$ patterns of length k goes to 1.

Theorem 5.5 *Let q be an* indecomposable *pattern of length k, and let $L = \lim_{n\to\infty} \sqrt[n]{S_n(q)}$. Let q' be defined as in Lemma 5.3. Then we have*

$$\lim_{n\to\infty} \sqrt[n]{S_n(q')} \geq 1 + L + 2\sqrt{L}.$$

Proof: This Theorem can be proved as Lemma 3.5, and Theorem 4.1 are. Indeed, as q is indecomposable, any block-structured string S will avoid q if each block does. Now apply part 1 of Proposition 5.1 to see that our argument will still provide the required lower bound. $\diamond$

Note that the fact that the reverse complement of an indecomposable pattern is also an indecomposable pattern makes it possible to prove an analogous version of Theorem 5.5, in which instead of prepending the indecomposable pattern q by a minimal entry, we affix a maximal entry to the end of q.

Our methods will not provide an upper bound for $\lim_{n\to\infty} \sqrt[n]{S_n(q')}$ because the string S of remaining entries of a q'-avoiding permutation does not have to be q-avoiding. (Only part 1 and not part 2 of Proposition 5.1 applies.) That condition is simply sufficient, but not necessary, in this general case. Nevertheless, Theorem 5.5 is interesting. It shows that for almost all patterns q, if we prepend q by the entry 1, the limit of the corresponding Stanley-Wilf sequence will grow at least as fast as for monotone q. If q started in 1, then this growth will be the same as for monotone q.

Now it is a little easier to understand why, in the case of length 4, the patterns that are the hardest to avoid, are along with certain equivalent ones, 1423 and 1342. Indeed, removing the starting 1 from them, we get the *decomposable* patterns 423 and 342. As these patterns are decomposable, Theorem 5.5 does not hold for them, so the limit of the Stanley-Wilf sequence for the patterns 1423 or 1342 does not have to be at least $1+4+4=9$, and in fact it is not.

A particularly interesting application of Theorem 5.4 is as follows. Recall that a *layered* pattern is a pattern that consists of decreasing subsequences (the layers) so that the entries increase among the layers. For instance, 3217654 is a layered pattern. In 1997, several people (including present author) have observed, using numerical evidence computed in [13], that if q is a layered pattern of length k, then for small n, the inequality $S_n(12\cdots k) \leq S_n(q)$ seems to hold. We will now show that this is indeed true in the sense of logarithmic asymptotics.

Theorem 5.6 *Let q be a layered pattern of length k. Then we have*

$$L(q) \geq (k-1)^2.$$

Equivalently, $L(q) \geq L(12\cdots k)$.

In order to prove Theorem 5.6, we need the following powerful Lemma, due to Backelin, West, and Xin.

Lemma 5.7 *[3] Let $r < k$, and let v be any pattern of length $k-r$ taken on the set $\{r+1, r+2, \cdots, k\}$. Then for all positive integers n, we have*

$$S_n(12\cdots rv) = S_n(r(r-1)\cdots 21v).$$

Now we are in position to prove Theorem 5.6.

Proof: (of Theorem 5.6.) Induction on k. If q has only one layer, then q is the decreasing pattern, and the statement is obvious. Now assume q has at least two layers, and that we know the statement for all layered patterns of length $k-1$. As q is layered, it is of the form $r(r-1)\cdots 21v$ for some r, and some layered pattern v. Therefore, Lemma 5.7 applies, and we have $S_n(q) = S_n(12\cdots rv)$. If this last pattern is denoted by q^*, then we obviously also have $L(q) = L(q^*)$. We further denote by q^{*-} the pattern obtained from q^* by removing its first entry. Note that q^{*-} is still a layered pattern, just its first several layers may have length 1.

Case 1 Assume first that $r > 1$. Then note that q^{*-} starts with its smallest entry. Therefore, Theorem 5.4 applies, and by the induction hypothesis we have

$$L(q) = L(q^*) = 1 + L(q^{*-}) + 2\sqrt{L(q^{*-})} \geq 1 + (k-2)^2 + 2(k-2) = (k-1)^2,$$

which was to be proved.

Case 2 Now assume that $r = 1$. Then q is a layered pattern that starts with a layer of length 1. Therefore, instead of applying Theorem 5.4, we need to, and almost always can, apply Theorem 5.5 for the pattern q^{*-}. Indeed, q^{*-} is a layered pattern, and as such, is indecomposable, except when it has only one layer, that is, it is the decreasing permutation.

Subcase 2a First look at the case when q^{*-} has more than one layers. That implies that q^{*-} is indecomposable. Therefore, we can apply Theorem 5.5 to get

$$L(q) \geq 1 + L(q^{*-}) + 2\sqrt{L(q^{*-})} \geq 1 + (k-2)^2 + 2(k-2) = (k-1)^2.$$

Subcase 2b Finally, if q^{*-} has only one layer, then by the definition of layered patterns, q^{*-} must be the decreasing pattern. As we are in the case when $r = 1$, we simply have $q = 1k(k-1)\cdots 2$. Then we have

$$S_n(q) = S_n(k-1\cdots 21k) = S_n(12\cdots k),$$

where the first equality follows by taking reverse complements, and the second one is a special case of Lemma 5.7. Indeed, simply set $r = k-1$ in that Lemma, and let v be the one-element pattern.

As we have covered all possible cases, the proof is complete. $\diamond$

We point out that while Case 1 could have been treated the same way as Subcase 2a, that would have been less elucidating. Indeed, in Case 1, we prove an *equality*, while in Subcase 2a, we only prove an *inequality*. This lends some further support to the conjecture, supported by numerical evidence, that among all layered patterns q of length k, the one for which $S_n(q)$ is maximal for large n is $q = 1325476\cdots$. As this pattern has as many non-singleton layers as possible (without being equivalent to the monotone pattern), for this pattern our inductive proof will go to Subcase 2a as many times as possible.

Here is another way in which our results start explaining why certain patterns are easier to avoid than others. We formulate our observations in the following Corollary.

Corollary 5.8 *Let q_1 and q_2 be patterns so that $L(q_1) \leq L(q_2)$. Let q_i' be the pattern obtained from q_i by prepending q_i by a 1. Furthermore, let q_1 start with the entry 1, and let q_2 be indecomposable. Then we have*

$$L(q_1') = 1 + L(q_1) + 2\sqrt{L(q_1)} \leq 1 + L(q_2) + 2\sqrt{L(q_2)} \leq L(q_2').$$

For instance, if we set $q_1 = 123$ and $q_2 = 213$, we get the well-known statement weakly comparing the limits of the Stanley-Wilf sequences of 1234 and 1324, first proved in [7].

6 Further Directions

Our results raise two interesting kinds of questions. We have seen that the limit of a Stanley-Wilf sequence is not simply not always an integer, but also not always rational. Is it always an *algebraic* number? If yes, can its degree be arbitrarily high? Can it be more than two? Is it always an *algebraic integer*, that is, the root of a *monic* polynomial with integer coefficients? The results so far leave that possibility open.

The second question is related to the size of the limit of $\sqrt[n]{S_n(q)}$ if q is of length k. The largest value that this limit is known to take is $(k-1)^2$, attained by the monotonic pattern. Before present paper, the smallest known value, in terms of k, for this limit was $(k-1)^2 - 1 = 8$, attained by $q = 1342$. As Theorem 5.2 shows, the value $(k-4+\sqrt{8})^2$ is also possible. As k goes to infinity, the *difference* of the assumed maximum $(k-1)^2$ and this value also goes to infinity, while their ratio goes to 1. Is it possible to find a series of patterns q_k so that this ratio does not converge to 1? We point out that it follows from a result of P. Valtr (published in [8]) that for any pattern q of length k, we have $\lim_{n\to\infty} \sqrt[n]{S_n(q)} \geq e^{-3}k^2$, so the mentioned ratio cannot be more then e^3.

Finally, now that the Stanley-Wilf conjecture has been proved, and we know that the limit of a Stanley-Wilf sequence always exists, we can ask what the largest possible value of this limit is, in terms of k. In [2], R. Arratia conjectured that this limit is at most $(k-1)^2$, and, following the footsteps of Erdős, he offered 100 dollars for a proof or disproof of the conjecture $S_n(q) \leq (k-1)^{2n}$, for all n and q. However, this conjecture was recently disproved in [1], where the authors showed that $L(1324) \geq 9.35$. Using our methods, it is straightforward to extend this result to the inequality

$$L(132456 \cdots k) \geq (\sqrt{9.35} + k - 4)^2,$$

for $k \geq 5$. This shows that if k is large enough, then the difference $L(q) - (k-1)^2$ can be arbitrarily large. We hope to extend this line of thinking in a subsequent paper.

References

[1] M. Albert, M. Elder, A. Rechnitzer, P. Westcott, and M. Zabrocki, A lower bound on the growth rate of the class of 4231 avoiding permutations, *Adv. Appl. Math, submitted.*

[2] R. Arratia, On the Stanley-Wilf Conjecture for the Number of Permutations Avoiding a Given Pattern. *Electronic J. Combin.*, **6** (1999), no. 1, N1.

[3] J. Backelin, J. West, G. Xin, Wilf equivalence for singleton classes, *Proceedings of the 13th Conference on Formal Power Series and Algebraic Combinatorics,* Tempe, AZ, 2001.

[4] M. Bóna, Exact enumeration of 1342-avoiding permutations; A close link with labeled trees and planar maps. *J. Combin. Theory A*, **80** (1997), 257–272.

[5] M. Bóna, **A Walk Through Combinatorics**, World Scientific, 2002.

[6] M. Bóna, **Combinatorics of Permutations**, CRC Press, 2004.

[7] M. Bóna, Permutations avoiding certain patterns; The case of length 4 and generalizations, *Discrete Math.,* **175** (1997), no. 1-3, 55–67.

[8] T. Kaiser, M. Klazar, On growth rates of hereditary permutation classes, *Electr. J. Combinatorics* **9** (2003), vol. 2, R10.

[9] A. Marcus, G. Tardos, Excluded Permutation Matrices and the Stanley-Wilf conjecture, *preprint, 2003.*

[10] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Advances in Mathematics*, **41** (1981), 115–136.

[11] R. Simion, F. W. Schmidt, Restricted Permutations. *European Journal of Combinatorics*, **6** (1985), 383–406.

[12] R. Stanley, **Enumerative Combinatorics**, volume 2, Cambridge University Press, 1997.

[13] J. West, Permutations with forbidden subsequences; and, Stack sortable permutations. *PHD-thesis,* Massachusetts Institute of Technology, 1990.

A COMBINATORIAL PROOF OF THE ROGERS-RAMANUJAN AND SCHUR IDENTITIES

CILANNE BOULET* AND IGOR PAK*

Abstract/Résumé. We give a combinatorial proof of the first Rogers-Ramanujan identity by using two symmetries of a new generalization of Dyson's rank. These symmetries are established by direct bijections.

Nous donnons une preuve combinatoire de la première identité de Rogers-Ramanujan en utilisant deux nouvelles symétries obtenues grâce à une généralisation de la notion de rang de Dyson. Ces symétries sont démontrées bijectivement.

Introduction

The Roger-Ramanujan identities are perhaps the most mysterious and celebrated results in partition theory. They have a remarkable tenacity to appear in areas as distinct as enumerative combinatorics, number theory, representation theory, group theory, statistical physics, probability and complex analysis [4, 6]. The identities were discovered independently by Rogers, Schur, and Ramanujan (in this order), but were named and publicized by Hardy [20]. Since then, the identities have been greatly romanticized and have achieved nearly royal status in the field. By now there are dozens of proofs known, of various degree of difficulty and depth. Still, it seems that Hardy's famous comment remain valid: *"None of the proofs of* [the Rogers-Ramanujan identities] *can be called "simple" and "straightforward"* [...]; *and no doubt it would be unreasonable to expect a really easy proof"* [20].

In this paper we propose a new combinatorial proof with a minimum amount of algebraic manipulation. Almost completely bijective, our proof would not satisfy Hardy as it is neither "simple" nor "straightforward". On the other hand, the heart of the proof is the analysis of two bijections and their properties, each of them elementary and approachable. In fact, our proof gives new generating function formulas (see (✠) in section 1) and is amenable to advanced generalizations which will appear elsewhere (see [8]).

We should mention that on the one hand, our proof is heavily influenced by the works of Bressoud and Zeilberger [10, 11, 12, 13], and on the other hand by Dyson's papers [14, 15], which were further extended by Berkovich and Garvan [7] (see also [19, 21]). In fact, the basic idea to use a generalization of Dyson's rank was explicit in [7, 19]. We postpone historical and other comments until section 3.

Let us say a few words about the structure of the paper. We split the proof of the Rogers-Ramanujan identities into two virtually independent parts. In the first, *the algebraic part*, we use the Jacobi triple product identity and some additional elementary algebraic manipulations to derive the equations. This proof is based on two symmetry

Date: November 21, 2004.

Key words and phrases. Rogers-Ramanujan identity, Schur's identity, bijection, Dyson's rank.

*Department of Mathematics, MIT, Cambridge, MA, 02139; Email: {cilanne,pak}@math.mit.edu.

equations whose proofs are given in *the combinatorial part* by direct bijections. Our presentation is elementary and completely self contained, except for the use of the classical Jacobi triple product identity.

1. The algebraic part

We consider the first *Rogers-Ramanujan identity*:

$$(\blacklozenge)\qquad 1+\sum_{k=1}^{\infty}\frac{t^{k^2}}{(1-t)(1-t^2)\cdots(1-t^k)} \;=\; \prod_{i=0}^{\infty}\frac{1}{(1-t^{5i+1})(1-t^{5i+4})}\,.$$

Our first step is standard. Recall the *Jacobi triple product identity* (see e.g. [4]):

$$\sum_{k=-\infty}^{\infty} z^k q^{\frac{k(k+1)}{2}} \;=\; \prod_{i=1}^{\infty}(1+zq^i)\,\prod_{j=0}^{\infty}(1+z^{-1}q^j)\,\prod_{i=1}^{\infty}(1-q^i).$$

Set $q\leftarrow t^5$, $z\leftarrow(-t^{-2})$, and rewrite the r.h.s. of $(\blacklozenge)$ as follows:

$$\prod_{r=0}^{\infty}\frac{1}{(1-t^{5r+1})(1-t^{5r+4})} \;=\; \sum_{m=-\infty}^{\infty}(-1)^m\, t^{\frac{m(5m-1)}{2}}\,\prod_{i=1}^{\infty}\frac{1}{(1-t^i)}\,.$$

This gives us *Schur's identity*, which is equivalent to $(\blacklozenge)$:

$$(\lozenge)\qquad \left(1+\sum_{k=1}^{\infty}\frac{t^{k^2}}{(1-t)(1-t^2)\cdots(1-t^k)}\right) \;=\; \prod_{i=1}^{\infty}\frac{1}{(1-t^i)}\,\sum_{m=-\infty}^{\infty}(-1)^m\, t^{\frac{m(5m-1)}{2}}\,.$$

To prove Schur's identity we need several combinatorial definitions. Denote by $\mathcal{P}_n$ the set of all partitions λ of n, and let $\mathcal{P}=\cup_n\mathcal{P}_n$, $p(n)=|\mathcal{P}_n|$. Denote by $\ell(\lambda)$ and $e(\lambda)$ the number of parts and the smallest part of the partition, respectively. By definition, $e(\lambda)=\lambda_{\ell(\lambda)}$. We say that λ is a *Rogers-Ramanujan partition* if $e(\lambda)\geq\ell(\lambda)$. Denote by $\mathcal{Q}_n$ the set of Rogers-Ramanujan partitions, and let $\mathcal{Q}=\cup_n\mathcal{Q}_n$, $q(n)=|\mathcal{Q}_n|$. Recall that

$$P(t)\;:=\;1\;+\;\sum_{n=1}^{\infty}p(n)\,t^n\;=\;\prod_{i=1}^{n}\frac{1}{1-t^i}\,,$$

and

$$Q(t)\;:=\;1\;+\;\sum_{n=1}^{\infty}q(n)\,t^n\;=\;1\;+\;\sum_{k=1}^{\infty}\frac{t^{k^2}}{(1-t)(1-t^2)\cdots(1-t^k)}\,.$$

We consider a statistic on $\mathcal{P}\smallsetminus\mathcal{Q}$ which we call $(2,0)$-*rank* of a partition, and denote by $r_{2,0}(\lambda)$, for $\lambda\in\mathcal{P}\smallsetminus\mathcal{Q}$. Similarly, for $m\geq1$ we consider a statistic on $\mathcal{P}$ which we call $(2,m)$-*rank* of a partition, and denote by $r_{2,m}(\lambda)$, for $\lambda\in\mathcal{P}$. We formally define and study these statistics in the next section. Denote by $h(n,m,r)$ the number of partitions λ of n with $r_{2,m}(\lambda)=r$. Similarly, let $h(n,m,\leq r)$ and $h(n,m,\geq r)$ be the number of partitions with the $(2,m)$-rank $\leq r$ and $\geq r$, respectively. The following is apparent from the definitions:

$$(\circledast)\qquad \begin{aligned} h(n,m,\leq r)\;+\;h(n,m,\geq r+1)\;&=\;p(n),\;\text{ for }\;m>0,\\ h(n,0,\leq r)\;+\;h(n,0,\geq r+1)\;&=\;p(n)-q(n),\end{aligned}$$

for all $r\in\mathbb{Z}$ and $n\geq1$. The following two equations are the main ingredients of the proof. For all $m,r\geq0$ we have:

(first symmetry) $\qquad h(n,0,r)\;=\;h(n,0,-r),$

(second symmetry) $\qquad h(n,m,\leq -r) \,=\, h(n-r-2m-2,m+2,\geq -r).$

Both symmetry equations will be proved in the next section. For now, let us continue to prove Schur's identity. For every $j \geq 0$ let

$$\begin{aligned} a_j \,&=\, h\left(n-jr-2jm-j(5j-1)/2, m+2j, \leq -r-j\right), \\ b_j \,&=\, h\left(n-jr-2jm-j(5j-1)/2, m+2j, \geq -r-j+1\right). \end{aligned}$$

The equation $(\circledast)$ gives us $a_j + b_j = p(n-jr-2jm-j(5j-1)/2)$, for all $r,j>0$. The second symmetry equation gives us $a_j = b_{j+1}$. Applying these multiple times we get:

$$\begin{aligned} h(n,m,\leq -r) \,&=\, a_0 \,=\, b_1 \\ &=\, b_1 + (a_1-b_2) - (a_2-b_3) + (a_3-b_4) - \ldots \\ &=\, (b_1+a_1) - (b_2+a_2) + (b_3+a_3) - (b_4+a_4) + \ldots \\ &=\, p(n-r-2m-2) - p(n-2r-4m-9) + p(n-3r-6m-42) - \ldots \\ &=\, \sum_{j=1}^{\infty} (-1)^{j-1} p(n-jr-2jm-j(5j-1)/2)\,. \end{aligned}$$

In terms of the generating functions

$$H_{m,\leq r}(t) \,:=\, \sum_{n=1}^{\infty} h(n,m,\leq r)\, t^n\,,$$

this gives (for $m,r \geq 0$)

$$(\maltese) \qquad H_{m,\leq r}(t) \,=\, \prod_{n=1}^{\infty} \frac{1}{(1-t^n)} \, \sum_{j=1}^{\infty} (-1)^{j-1}\, t^{jr+2jm+j(5j-1)/2}\,.$$

In particular, we have:

$$\begin{aligned} H_{0,\leq 0}(t) \,&=\, \prod_{i=1}^{\infty} \frac{1}{(1-t^i)} \, \sum_{j=1}^{\infty} (-1)^{j-1}\, t^{\frac{j(5j-1)}{2}}\,, \\ H_{0,\leq -1}(t) \,&=\, \prod_{i=1}^{\infty} \frac{1}{(1-t^i)} \, \sum_{j=1}^{\infty} (-1)^{j-1}\, t^{\frac{j(5j+1)}{2}}\,. \end{aligned}$$

From the first symmetry equation and $(\circledast)$ we have:

$$H_{0,\leq 0}(t) + H_{0,\leq -1}(t) \,=\, H_{0,\leq 0}(t) + H_{0,\geq 1}(t) \,=\, P(t) - Q(t).$$

We conclude:

$$\begin{aligned} &\prod_{n=1}^{\infty} \frac{1}{(1-t^i)} \left(\sum_{j=1}^{\infty} (-1)^{j-1}\, t^{\frac{j(5j-1)}{2}} + \sum_{j=1}^{\infty} (-1)^{j-1}\, t^{\frac{j(5j+1)}{2}} \right) \\ &\qquad = \prod_{i=1}^{\infty} \frac{1}{(1-t^i)} \, - \left(1 + \sum_{k=1}^{\infty} \frac{t^{k^2}}{(1-t)(1-t^2)\ldots(1-t^k)} \right), \end{aligned}$$

which implies $(\Diamond)$ and completes the proof of $(\blacklozenge)$.

2. The combinatorial part

2.1. **Definitions.** Let $\lambda = (\lambda_1, \ldots, \lambda_{\ell(\lambda)})$, $\lambda_1 \geq \ldots \geq \lambda_{\ell(\lambda)} > 0$, be an integer partition of $n = \lambda_1 + \ldots + \lambda_{\ell(\lambda)}$. We will say that $\lambda_j = 0$ for $j > \ell(\lambda)$. We graphically represent partition λ by a Young diagram $[\lambda]$ as in Figure 1. Denote by λ' a *conjugate partition* of λ obtained by reflection upon main diagonal (see Figure 1).

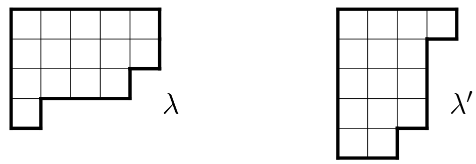


Figure 1. Partition $\lambda = (5, 5, 4, 1)$ and conjugate partition $\lambda' = (4, 3, 3, 3, 2)$.

For $m \geq 0$, define an *m-rectangle* to be a rectangle whose height minus its width is m. Define the *first m-Durfee rectangle* to be the largest m-rectangle which fits in diagram $[\lambda]$. Denote by $s_m(\lambda)$ the height of the first m-Durfee rectangle. Define the *second m-Durfee rectangle* to be the largest m-rectangle which fits in diagram $[\lambda]$ below the first m-Durfee rectangle, and let $t_m(\lambda)$ be its height. We will allow an m-Durfee rectangle to have width 0 but never height 0. Finally, denote by α, β, and γ the three partitions to the right of, in the middle of and below the m-Durfee rectangles (see Figure 2).

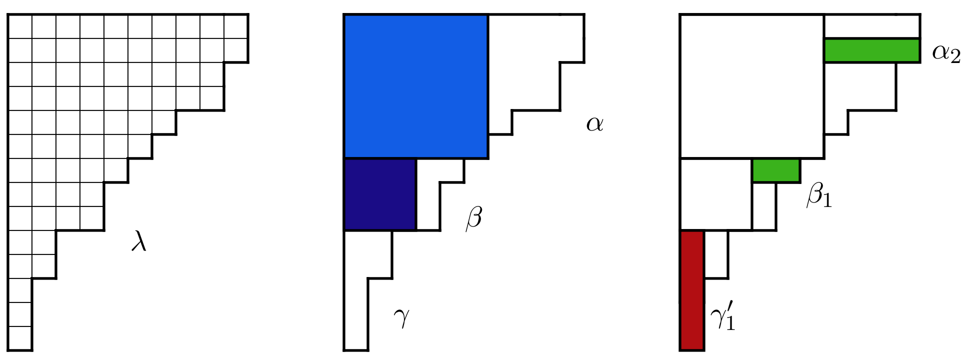


Figure 2. Partition $\lambda = (10, 10, 9, 9, 7, 6, 5, 4, 4, 2, 2, 1, 1, 1)$, the first Durfee square of height $s_0(\lambda) = 6$, and the second Durfee square of height $t_0(\lambda) = 3$. Here the remaining partitions are $\alpha = (4, 4, 3, 3, 1)$, $\beta = (2, 1, 1)$, and $\gamma = (2, 2, 1, 1, 1)$. In this case, the $(2, 0)$-rank $r_{2,0}(\lambda) = \beta_1 + \alpha_2 - \gamma'_1 = 1$.

We define *(2, m)-rank*, $r_{2,m}(\lambda)$, of a partition λ by the formula:

$$r_{2,m}(\lambda) := \beta_1 + \alpha_{s_m(\lambda)-t_m(\lambda)-\beta_1+1} - (\ell(\lambda) - s_m(\lambda) - t_m(\lambda)) .$$

From this definition, it is easy to see that

$$r_{2,m}(\lambda) = \begin{cases} \beta_1 + \alpha_{s_m(\lambda)-t_m(\lambda)-\beta_1+1} - \gamma'_1 & \text{if } s_m(\lambda),\ t_m(\lambda) > m, \\ \beta_1 + \alpha_{s_m(\lambda)-t_m(\lambda)-\beta_1+1} - (\beta'_1 - m) & \text{if } s_m(\lambda) > t_m(\lambda) = m, \\ \alpha_1 - (\alpha'_1 - 2m) & \text{if } s_m(\lambda) = t_m(\lambda) = m . \end{cases}$$

For example, $r_{2,0}(\lambda) = \beta_1 + \alpha_2 - \gamma_1' = 2 + 4 - 5 = 1$ for λ as in Figure 2.

Note that $(2,0)$-rank is only defined for non-Rogers-Ramanujan partitions because otherwise β_1 does not exist, while $(2,m)$-rank is defined for all partitions for all $m > 0$.

Let $\mathcal{H}_{n,m,r}$ to be the set of partitions of n with $(2,m)$-rank r. In the notation above, $h(n,s,r) = \big|\mathcal{H}_{n,s,r}\big|$.

2.2. **Proof of the first symmetry.** In order to prove the first symmetry we present an involution φ on $\mathcal{P} \smallsetminus \mathcal{Q}$ which preserves the size of partitions as well as their Durfee squares, but changes the sign of the rank:

$$\varphi : \mathcal{H}_{n,0,r} \to \mathcal{H}_{n,0,-r}\,.$$

Let λ be a partition with two Durfee square and partitions α, β, and γ to the right of, in the middle of, and below the Durfee squares. This map φ will preserve the Durfee squares of λ whose sizes we denote by

$$s = s_0(\lambda) \qquad \text{and} \qquad t = t_0(\lambda)\,.$$

We will describe the action of $\varphi : \lambda \mapsto \widehat{\lambda}$ by first mapping (α, β, γ) to a 5-tuple of partitions $(\mu, \nu, \pi, \rho, \sigma)$, and subsequently mapping that 5-tuple to different triple $(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma})$ which goes to the right of, in the middle of, and below of the Durfee squares in $\widehat{\lambda}$.

(1) First, let $\mu = \beta$.
Second, remove the following parts from α: $\alpha_{s-t-\beta_j+j}$ for $1 \le j \le t$. Let ν be the partition comprising of parts removed from α and π be the partitions comprising of the parts which were not removed.
Third, for $1 \le j \le t$, let
$$k_j = \max\{k \le s - t \mid \gamma_j' - k \ge \pi_{s-t-k+1}\}\,.$$
Let ρ be the partition with parts $\rho_j = k_j$ and σ be the partition with parts $\sigma_j = \gamma_j' - k_j$.

(2) First, let $\widehat{\gamma}' = \nu + \mu$ be the *sum* of partitions, defined to have parts $\widehat{\gamma}_j' = \nu_j + \mu_j$.
Second, let $\widehat{\alpha} = \sigma \cup \pi$ be the *union* of partitions, defined as a union of parts in σ and π.[1]
Third, let $\widehat{\beta} = \rho$.

Figure 3 shows an example of φ and the relation between these two steps.

Remark 2.1. The key to understanding the map φ is the definition of k_j. By considering $k = 0$, we see that k_j is defined for all $1 \le j \le t$. Moreover, one can check that k_j is the *unique* integer k which satisfies

$$(\dagger) \qquad \pi_{s-t-k+1} \le \gamma_j' - k \le \pi_{s-t-k}\,.$$

(We do not consider the upper bound for $k = s - t$.) This characterization of k_j can also be taken as its definition. Equation $(\dagger)$ is used repeatedly in our proof of the next lemma.

Lemma 2.2. *The map φ defined above is an involution.*

[1]Alternatively, the union can be defined via the sum: $\sigma \cup \pi = (\sigma' + \pi')'$.

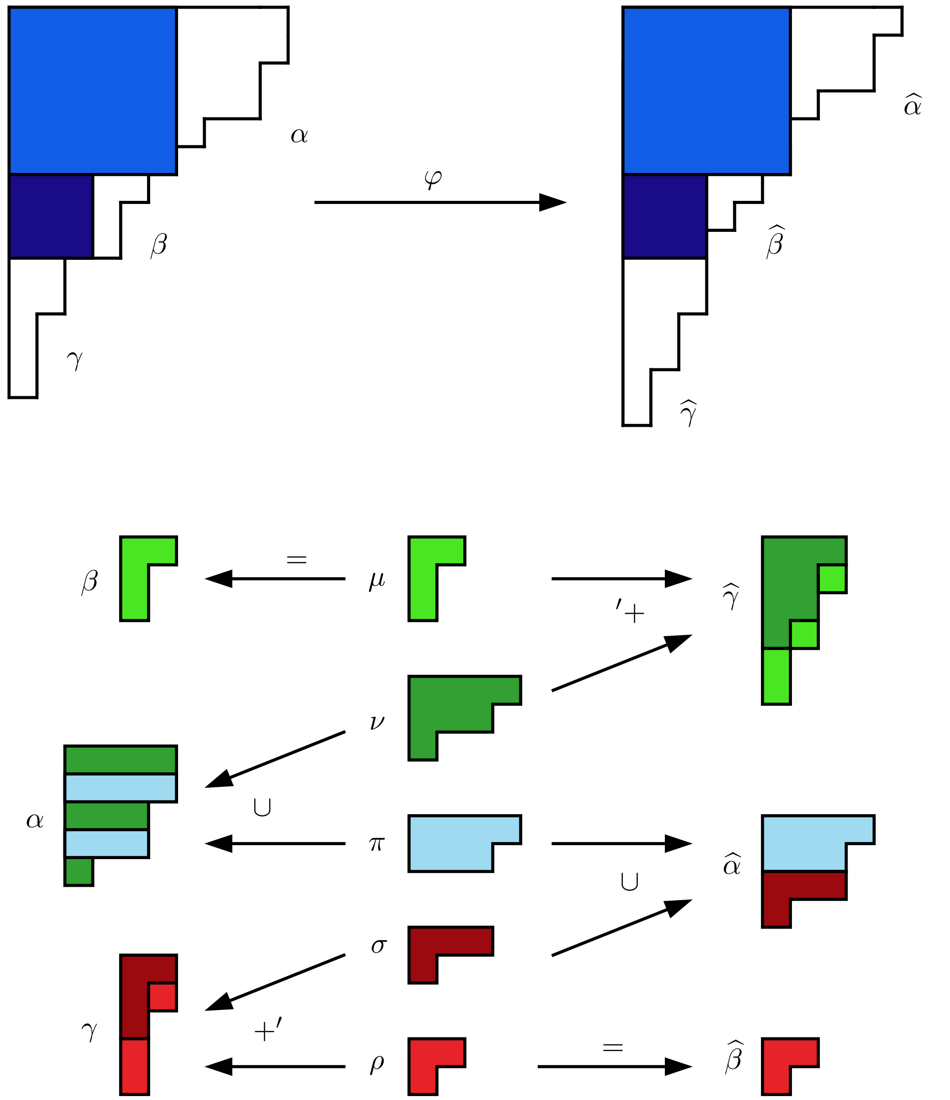


FIGURE 3. An example of the first symmetry involution $\varphi : \lambda \mapsto \widehat{\lambda}$, where $\lambda \in \mathcal{H}_{n,0,r}$ and $\widehat{\lambda} \in \mathcal{H}_{n,0,-r}$ for $n = 71$, and $r = 1$. The maps are defined by the following rules: $\beta = \mu$, $\alpha = \nu \cup \pi$, $\gamma' = \sigma + \rho$, while $\widehat{\beta} = \rho$, $\widehat{\alpha} = \pi \cup \sigma$, $\widehat{\gamma}' = \mu + \nu$. Also, $\lambda = (10, 10, 9, 9, 7, 6, 5, 4, 4, 2, 2, 1, 1, 1)$ and $\widehat{\lambda} = (10, 9, 9, 7, 6, 6, 5, 4, 3, 3, 3, 2, 2, 1, 1)$.

Proof. Our proof is divided into five parts; we prove that

(1) ρ is a partition, **(2)** σ is a partition, **(3)** $\widehat{\lambda} = \varphi(\lambda)$ is a partition,

(4) φ^2 is the identity map, and **(5)** $r_{2,0}(\widehat{\lambda}) = -r_{2,0}(\lambda)$.

(1) Considering the bounds (†) for j and $j+1$, we note that, if $k_j \le k_{j+1}$, then

$$\pi_{s-t-k_j+1} + k_j \;\le\; \pi_{s-t-k_{j+1}+1} + k_{j+1} \;\le\; \gamma'_{j+1} \le \gamma'_j \;\le \pi_{s-t-k_j} + k_j\,.$$

This gives us

$$\pi_{s-t-k_j+1} \leq \gamma'_{j+1} - k_j \leq \pi_{s-t-k_j}$$

and uniqueness therefore implies that $k_j = k_{j+1}$. We conclude that $k_j \geq k_{j+1}$ and that ρ is a partition.

(2) If $k_j > k_{j+1}$, then we have $s - t - k_j + 1 \leq s - t - k_{j+1}$ and therefore

$$\pi_{s-t-k_{j+1}} \leq \pi_{s-t-k_j+1} .$$

Again, by considering (†) for j and $j+1$, we conclude that

$$\gamma'_j - k_j \geq \gamma'_{j+1} - k_{j+1} .$$

If $k_j = k_{j+1}$, then we simply need to recall that γ' is a partition to see that

$$\gamma'_j - k_j \geq \gamma'_{j+1} - k_{j+1} .$$

This implies that σ is a partition.

(3) By their definitions, it is clear that μ, ν, and π are partitions. Since we just showed that ρ and σ are all partition, it follows that $\widehat{\alpha}$, $\widehat{\beta}$, and $\widehat{\gamma}$ are also partitions. Moreover, by their definitions, we see that μ, ν, and σ have at most t parts, π has at most $s-t$, and ρ has at most t parts each of which is less than or equal to $s-t$. This implies that $\widehat{\alpha}$ has at most s parts, $\widehat{\beta}$ has at most t parts each of which is less than or equal to $s-t$, and $\widehat{\gamma}'$ has parts at most t. Therefore, $\widehat{\alpha}$, $\widehat{\beta}$, and $\widehat{\gamma}$ fit to the right of, in the middle of, and below Durfee squares of sizes s and t and so $\varphi(\lambda)$ is a partition.

(4) We will apply φ twice to a Rogers-Ramanujan partition λ with α, β, and γ to the right of, in the middle of, and below its two Durfee squares. As usual, let μ, ν, π, ρ, σ be the partitions occurring in the intermediate stage of the first application of φ to λ and let $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\gamma}$ be the partitions to the right of, in the middle of, and below the Durfee squares of $\widehat{\lambda} = \varphi(\lambda)$. Similarly, let $\widehat{\mu}$, $\widehat{\nu}$, $\widehat{\pi}$, $\widehat{\rho}$, $\widehat{\sigma}$ be the partitions occurring in the intermediate stage of the second application of φ and let α^*, β^*, and γ^* be the partitions to the right of, in the middle of and below the Durfee squares of $\varphi^2(\lambda) = \varphi(\widehat{\lambda})$.

We need several observations. First, note that $\widehat{\mu} = \widehat{\beta} = \rho$. Second, by (†) we have:

$$\pi_{s-t-k_j+1} \leq \gamma'_j - k_j = \sigma_j \leq \pi_{s-t-k_j} .$$

Since σ is a partition, this implies that $\widehat{\alpha}_{s-t-k_j+j} = \sigma_j$. On the other hand, since $\widehat{\beta}_j = \rho_j = k_j$, the map φ removes the rows $\widehat{\alpha}_{s-t-k_j+j} = \sigma_j$ from $\widehat{\alpha}$. From here we conclude that $\widehat{\nu} = \sigma$ and $\widehat{\pi} = \pi$. Third, define

$$\widehat{k}_j = \max\{\widehat{k} \leq s-t \mid \gamma'_j - \widehat{k} \geq \pi_{s-t-\widehat{k}+1}\} .$$

By remark 2.1, we know that $\widehat{k}_j$ as above is the unique integer $\widehat{k}$ which satisfies:

$$\widehat{\pi}_{s-t-\widehat{k}+1} \leq \widehat{\gamma}'_j - \widehat{k} \leq \widehat{\pi}_{s-t-\widehat{k}} .$$

On the other hand, recall that $\widehat{\gamma}'_j = \mu_j + \nu_j$ and $\beta_j = \mu_j$. This implies $\widehat{\gamma}'_j - \beta_j = \nu_j$. Also, by the definition of ν, we have $\nu_j = \alpha_{s-t-\beta_j+j}$. Therefore, by the definition of π, we have:

$$\pi_{s-t-\beta_j+1} \leq \alpha_{s-t-\beta_j+j} = \nu_j = \widehat{\gamma}'_j - \beta_j \leq \pi_{s-t-\beta_j} .$$

Since, $\widehat{\pi} = \pi$, by the uniqueness in remark 2.1 we have $\widehat{k}_j = \beta_j = \mu_j$. This implies that $\widehat{\rho} = \mu$ and $\widehat{\sigma} = \nu$.

Finally, the second step of our bijection gives $\alpha^* = \nu \cup \pi = \alpha$, $\beta^* = \mu = \beta$, and $(\gamma^*)' = \rho + \sigma = \gamma'$. This implies that φ^2 is the identity map.

(5) Using the results from **(4)**, we have:

$$r_{2,0}(\lambda) = \beta_1 + \alpha_{s-t-\beta_1+1} - \gamma_1' = \mu_1 + \nu_1 - \rho_1 - \sigma_1 .$$

On the other hand,

$$r_{2,0}(\widehat{\lambda}) = \widehat{\beta}_1 + \widehat{\alpha}_{s-t-\widehat{\beta}_1+1} - \widehat{\gamma}_1' = \rho_1 + \sigma_1 - \mu_1 - \nu_1 .$$

We conclude that $r_{2,0}(\widehat{\lambda}) = -r_{2,0}(\lambda)$. $\square$

2.3. **Proof of the second symmetry.** In order to prove the second symmetry we present a bijection

$$\psi_{m,r} : \mathcal{H}_{n,m,\leq -r} \to \mathcal{H}_{n-r-2m-2,m+2,\geq -r} .$$

This map will only be defined for $r \geq 0$, in which case the first and second m-Durfee rectangles of a partition $\lambda \in \mathcal{H}_{n,m,\leq -r}$ have non-zero width because

$$\beta_1 + \alpha_{s_m(\lambda)-t_m(\lambda)-\beta_1+1} - (\ell(\lambda) - s_m(\lambda) - t_m(\lambda)) \leq -r \leq 0$$

implies that $\ell(\lambda) \geq s_m(\lambda) + t_m(\lambda)$. As a consequence,

$$r_{2,m}(\lambda) = \beta_1 + \alpha_{s_m(\lambda)-t_m(\lambda)-\beta_1+1} - \gamma_1' .$$

We describe the action of $\psi := \psi_{m,r}$ by giving the sizes of the Durfee rectangles of $\widehat{\lambda} := \psi_{m,r}(\lambda) = \psi(\lambda)$ and the partitions $\widehat{\alpha}$, $\widehat{\beta}$, and $\widehat{\gamma}$ which go to the right of, in the middle of, and below those Durfee rectangles in $\widehat{\lambda}$.

(1) If λ has two m-Durfee rectangle of height

$$s := s_m(\lambda) \qquad \text{and} \qquad t := t_m(\lambda)$$

then μ has two $(m+2)$-Durfee rectangle of height

$$s' := s_{m+2}(\widehat{\lambda}) = s + 1 \qquad \text{and} \qquad t' := t_{m+2}(\widehat{\lambda}) = t + 1 .$$

(2) Let

$$k_1 = \max\{k \leq s - t \mid \gamma_1' - r - k \geq \alpha_{s-t-k+1}\} .$$

Obtain $\widehat{\alpha}$ from α by adding a new part of size $\gamma_1' - r - k_1$, $\widehat{\beta}$ from β by adding a new part of size k_1, and $\widehat{\gamma}$ from γ by removing its first column.

Figure 3 shows an example of the bijection $\psi = \psi_{m,r}$.

Remark 2.3. As in remark 2.1, by considering $k = \beta_1$ we see that k_1 is defined and indeed we have $k_1 \geq \beta_1$. Moreover, it follows from its definition that k_1 is the *unique* k such that

$$(\ddagger) \qquad \alpha_{s-t-k+1} \leq \gamma_1' - r - k \leq \alpha_{s-t-k} .$$

(If $k = s - t$ we do not consider the upper bound.)

Lemma 2.4. *The map $\psi = \psi_{m,r}$ defined above is a bijection.*

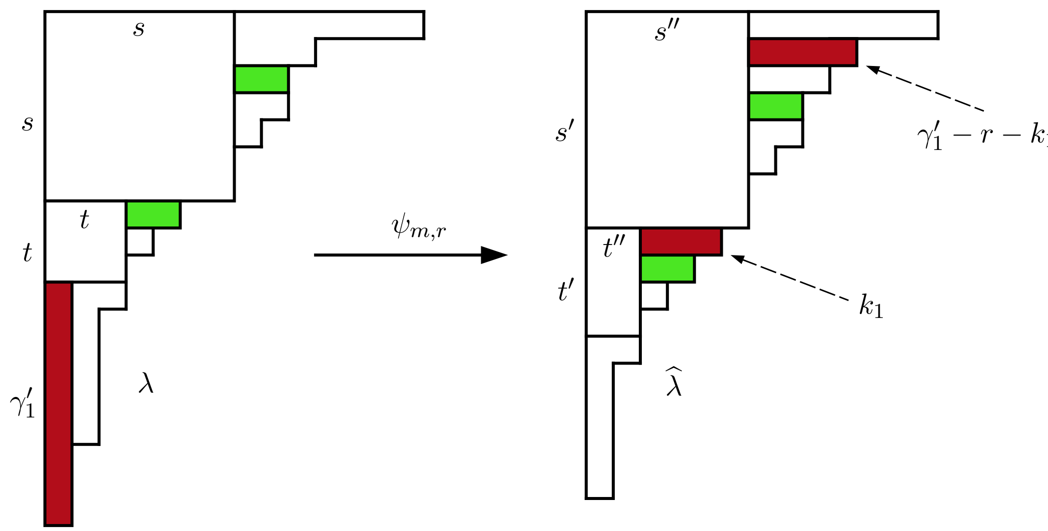


FIGURE 4. An example of the second symmetry bijection $\psi_{m,r} : \lambda \mapsto \widehat{\lambda}$, where $\lambda \in \mathcal{H}_{n,m,\le -r}$, $\widehat{\lambda} \in \mathcal{H}_{n',m+2,\ge -r}$, for $m = 0$, $r = 2$, $n = 92$, and $n' = n - r - 2m - 2 = 88$. Here $r_{2,0}(\lambda) = 2+2-9 = -5 \le -2$ and $r_{2,2}(\widehat{\lambda}) = 3 + 4 - 6 = 1 \ge -2$, where $\lambda = (14, 10, 9, 9, 8, 7, 7, 5, 4, 3, 3, 2, 2, 2, 2, 2, 1, 1, 1)$ and $\widehat{\lambda} = (13, 10, 9, 8, 8, 7, 6, 6, 5, 4, 3, 2, 2, 1, 1, 1, 1, 1)$. Also, $s = 7$, $s' = s + 1 = 8$, $s'' = s' - m - 2 = 6$, $t = 3$, $t' = 4$, $t'' = 2$, $\gamma'_1 = 9$, $k_1 = 3$, and $\gamma'_1 - r - k_1 = 4$.

Proof. Our proof has three parts:

(1) we prove that the size of $\widehat{\lambda} = \psi(\lambda)$ is $n - r - 2m - 2$,
(2) we prove that $r_{2,m+2}(\widehat{\lambda}) \ge -r$, and
(3) we present the inverse map ψ^{-1}.

(1) To prove that the above construction gives a partition $\widehat{\lambda}$ of size $n - r - 2m - 2$ note that together the rows added to α and β have size r less than the size of the column removed from γ. Also, both the first and second $(m + 2)$-Durfee rectangles of $\widehat{\lambda}$ have size $m + 1$ less than the size of the corresponding m-Durfee rectangle of λ. This implies the claim.

(2) By remark 2.3, the part we inserted into β will be the largest part of the resulting partition, i.e. $\widehat{\beta}_1 = k_1$. By equation (‡) we have:

$$\alpha_{s-t-k_1+1} \le \gamma'_1 - r - k_1 \le \alpha_{s-t-k_1} .$$

Therefore, we must have:

$$\widehat{\alpha}_{s'-t'-\widehat{\beta}_1+1} = \widehat{\alpha}_{s-t-k_1+1} = \gamma'_1 - r - k .$$

Indeed, we have chosen k_1 in the *unique* way so that the rows we insert into α and β are $\widehat{\alpha}_{s'-t'-\widehat{\beta}_1+1}$ and $\widehat{\beta}_1$ respectively.

The above two equations now allow us to bound the $(2, m + 2)$-rank of $\widehat{\lambda}$:

$$\widehat{\alpha}_{s'-t'-\widehat{\beta}_1+1} + \widehat{\beta}_1 - (\ell(\widehat{\lambda}) - s' - t') \;=\; (\gamma'_1 - r - k_1) + k_1 - (\ell(\widehat{\lambda}) - s' - t') \;\ge\; -r ,$$

where the last inequality follows from

$$\gamma'_1 \geq \gamma'_2 \geq \ell(\widehat{\lambda}) - s' - t' \, .$$

(3) The above characterization of k_1 also shows us that to recover α, β, and γ from $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\gamma}$, we remove part $\widehat{\alpha}_{s'-t'-\widehat{\beta}_1+1}$ from $\widehat{\alpha}$, remove part $\widehat{\beta}_1$ from $\widehat{\beta}$, and add a column of height $\widehat{\alpha}_{s'-t'-\widehat{\beta}_1+1} + \widehat{\beta}_1 + r$ to $\widehat{\gamma}$. Since we can also easily recover the sizes of the previous m-Durfee rectangles, we conclude that ψ is a bijection between the desired sets. □

3. Final remarks

3.1. Of the many proofs of Rogers-Ramanujan identities only a few can be honestly called "combinatorial". We would like to single out [3] as an interesting example. Perhaps, the most important combinatorial proof was given by Schur in [24] where he deduced his identity by a direct involutive argument. The celebrated bijection of Garsia and Milne [18] is based on this proof and the involution principle. In [11], a different involution principle proof was obtained (see also [13]) based on a short proof of Bressoud [10]. We refer to [22] for further references, historical information, and combinatorial proofs of other partition identities.

3.2. Dyson's rank $r_1(\lambda) = \lambda_1 - \lambda'_1$ was defined in [14] for the purposes of finding a combinatorial interpretation of Ramanujan's congruences. Dyson used the rank to obtain a simple combinatorial proof of Euler's pentagonal theorem in [15] (see also [16, 21]). It was shown in [21] that this proof can be converted into a direct involutive proof, and such a proof in fact coincides with the involution obtained by Bressoud and Zeilberger [12].

Roughly speaking, our proof of Schur's identity is a Dyson-style proof with a modified Dyson's rank, where the definition of the latter was inspired by [11, 12, 13]. Unfortunately, reverse engineering the proofs in [13] is not straightforward due to the complexity of that paper. Therefore, rather than giving a formal connection, we will only say that our map $\psi_{m,r}$ is related to the maps φ in [11] and Φ in [13].

It would be interesting to extend our Dyson-style proof to the generalization of Schur's identity found in [17]. This would give a new combinatorial proof of the generalizations of the Rogers-Ramanujan identities found in that paper and, in a special case, provide a new combinatorial proof of the second Rogers-Ramanujan identity (see e.g. [4, 6, 20, 22]).

3.3. The idea of using iterated Durfee squares to study the Rogers-Ramanujan identities and their generalizations is due to Andrews [5]. The $(2, m)$-rank of a partition is a special case of a general (but more involved) notion of (k, m)-rank which will be presented in [8]. It leads to combinatorial proofs of some of Andrews' generalizations of Rogers-Ramanujan identities mentioned above.

Garvan [19] defined a generalized notion of a rank to partitions with iterated Durfee squares, that is different from ours, but still satisfies equation (✠) (for $m = 0$). In [7], Berkovich and Garvan asked for a Dyson-style proof of (✠) but unfortunately, they were unable to carry out their program in full as the combinatorial symmetry they obtain seem to be hard to establish bijectively. (This symmetry is somewhat different from our second symmetry.) Most recently, the first author was able to relate the two generalizations of rank by a combinatorial, but not completely bijective, argument. This will also appear in [8].

3.4. Yet another generalization of Dyson's rank was kindly brought to our attention by George Andrews. The notion of successive rank can also be used to give a combinatorial proof of the Rogers-Ramanujan identities and their generalizations by a sieve argument (see [2, 9]). However, this proof involves a different combinatorial description of the partitions on the left hand side of the Rogers-Ramanujan identities than the proof presented here.

3.5. Finally, let us note that the Jacobi triple product identity has a combinatorial proof due to Sylvester (see [22, 25]). We refer to [1] for an elementary algebraic proof.

Also, while our proof is mostly combinatorial it is by no means a direct bijection. The quest for a direct bijective proof is still under way, and as recently as this year Zeilberger lamented on the lack of such proof [26]. The results in [23] seem to discourage any future work in this direction.

Acknowledgments. The authors are grateful to George Andrews and Richard Stanley for their support and encouragement of our studies of partition identities. The first author was supported by NSERC(Canada) and the second author by the NSA and the NSF.

References

[1] G. E. Andrews, A simple proof of Jacobi's triple product identity, *Proc. Amer. Math. Soc.* **16** (1965), 333–334.
[2] G. E. Andrews, Sieves in the theory of partitions, *Amer. J. Math.* **94** (1972), 1214–1230.
[3] G. E. Andrews, Partially ordered sets and the Rogers-Ramanujan identities, *Aequationes Math.* **12** (1975), 94–107.
[4] G. E. Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, MA, 1976.
[5] G. E. Andrews, Partitions and Durfee dissection, *Amer. J. Math.* **101** (1979), 735–742.
[6] G. E. Andrews, *q-series: their development and application in analysis, number theory, combinatorics, physics, and computer algebra*, CBMS Regional Conference Series in Mathematics **66**, AMS, Providence, RI, 1986.
[7] A. Berkovich and F. G. Garvan, Some Observations on Dyson's New Symmetries of Partitions, *J. Combin. Theory, Ser. A* **100** (2002), 61–93.
[8] C. Boulet, MIT Ph.D. thesis, in preparation.
[9] D. M. Bressoud, Extensions of the partition sieve, *J. Number Theory* **12** (1980), 87–100.
[10] D. M. Bressoud, An easy proof of the Rogers-Ramanujan identities, *J. Number Theory* **16** (1983), 235–241.
[11] D. M. Bressoud and D. Zeilberger, A short Rogers-Ramanujan bijection, *Discrete Math.* **38** (1982), 313–315.
[12] D. M. Bressoud and D. Zeilberger, Bijecting Euler's partitions-recurrence, *Amer. Math. Monthly* **92** (1985), 54–55.
[13] D. M. Bressoud and D. Zeilberger, Generalized Rogers-Ramanujan bijections, *Adv. Math.* **78** (1989), 42–75.
[14] F. J. Dyson, Some Guesses in The Theory of Partitions, *Eureka* (Cambridge) **8** (1944), 10–15.
[15] F. J. Dyson, A new symmetry of partitions, *J. Combin. Theory* **7** (1969), 56–61.
[16] F. J. Dyson, A walk through Ramanujan's garden, in *"Ramanujan revisited"*, Academic Press, Boston, 1988, 7–28.
[17] K. Garrett, M. E. H. Ismail, and D. Stanton, Variants of the Rogers-Ramanujan Identities, *Adv. in Appl. Math.* **23** (1999), 274–299.
[18] A. M. Garsia and S. C. Milne, A Rogers-Ramanujan bijection, *J. Combin. Theory, Ser. A* **31** (1981), 289–339.
[19] F. G. Garvan, Generalizations of Dyson's rank and non-Rogers-Ramanujan partitions, *Manuscripta Math.* **84** (1994), 343–359.

[20] G. H. Hardy, *Ramanujan. Twelve lectures on subjects suggested by his life and work*, Cambridge University Press, Cambridge, England, 1940.

[21] I. Pak, On Fine's partition theorems, Dyson, Andrews, and missed opportunities, *Math. Intelligencer* **25**, No. 1 (2003), 10–16.

[22] I. Pak, Partition bijections. A survey, to appear in *Ramanujan Journal*, 69 pp., available at `http://www-math.mit.edu/~pak/research.html`

[23] I. Pak, Rogers-Ramanujan bijections are not geometric, in preparation.

[24] I. Schur, Ein Beitrag zur Additiven Zahlentheorie und zur Theorie der Kettenbrüche, *S.-B. Preuss. Akad. Wiss. Phys. Math. Klasse* (1917), 302–321.

[25] E. M. Wright, An enumerative proof of an identity of Jacobi, *J. London Math. Soc.* **40** (1965), 55–57.

[26] D. Zeilberger, Enumerative and Algebraic Combinatorics, to appear in *"Princeton Companion of Mathematics"* (T. Gowers, ed.), Princeton University Press, Princeton, NJ.

COMBINATORICS OF PATIENCE SORTING PILES

ALEXANDER BURSTEIN AND ISAIAH LANKHAM

Abstract. Despite having been introduced in 1962 by C.L. Mallows, the combinatorial algorithm *Patience Sorting* is only now beginning to receive significant attention due to such recent deep results as the Baik-Deift-Johansson Theorem that connect it to fields including Probabilistic Combinatorics and Random Matrix Theory.

The aim of this work is to develop some of the more basic combinatorics of the Patience Sorting Algorithm. In particular, we exploit the similarities between Patience Sorting and the Schensted Insertion Algorithm in order to do things that include defining an analog of the Knuth relations and extending Patience Sorting to a bijection between permutations and certain pairs of set partitions. As an application of these constructions we characterize and enumerate the set $S_n(3\text{-}\bar{1}\text{-}42)$ of permutations that avoid the generalized permutation pattern 2-31 unless it is part of the generalized pattern 3-1-42.

Résumé. En dépit de la introduction en 1962 par C.L. Mallows, combinatoire d'algorithme *Patience Sorting* commence seulement maintenant à susciter l'attention significative dû à des résultats profonds récents tels que le théorème de Baik-Deift-Johansson qui le relient à la combinatoire probabiliste et à la théorie des matrices aléatoires.

On développe une partie plus fondamentale de la combinatoire de l'algorithme de *Patience Sorting*. En particulier, on utilise les similitudes entre *Patience Sorting* et la correspondence de Schensted pour définir un analogue des relations de Knuth et pour généraliser *Patience Sorting* en une bijection entre les permutations et certaines paires de partitions d'ensemble. Comme application de ces constructions on caractérise et énumére l'ensemble $S_n(3\text{-}\bar{1}\text{-}42)$ de permutations qui évitent le motif généralisé 2-31 de permutation à moins qu'il soit partie du motif généralisé 3-1-42.

1. Introduction

The term *Patience Sorting* was introduced in 1962 by C.L. Mallows [15, 16] as the name of a card sorting algorithm invented by A.S.C. Ross. This algorithm works by first partitioning a shuffled deck of cards (which we take to be a permutation $\sigma \in \mathfrak{S}_n$) into sorted subsequences called *piles* using what Mallows referred to as a "patience sorting procedure":

Algorithm 1.1 (Mallows' Patience Sorting Procedure)**.** Given a shuffled deck of cards $\sigma = c_1 c_2 \cdots c_n$, inductively build the set of piles $R = R(\sigma) = \{r_1, r_2, \ldots, r_m\}$ 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.

Date: May 1, 2005.
2000 *Mathematics Subject Classification.* Primary: 05A05, 05A18; Secondary: 05A15.
Key words and phrases. Patience Sorting, Set Partitions, Bell Numbers, Generalized Permutation Patterns, Left-to-right Minima Subsequences, Basic Subsequences, Shadow Diagrams.
The work of the second author was supported in part by the U.S. National Science Foundation under Grants DMS-0135345 and DMS-0304414.

We call the collection of piles $R(\sigma)$ the *pile configuration* associated to the deck of cards σ and illustrate their formation via an extended version of Algorithm 1.1 in Section 3.1 below.

Since each card c_i is either larger than the top card of every pile or is placed on top of the left-most top card d_j larger than it, the cards $d_1, d_2, \ldots, d_k$ atop the piles will be in increasing order from left to right at each step of the algorithm. Thus, Algorithm 1.1 resembles repeated application of the Schensted Insertion Algorithm (see [10]) for interposing a value into the sequence $d_1, d_2, \ldots, d_k$ as if it were the top row of a Young tableau. The distinction is that cards remain in place and have other cards placed on top of them instead of being actively "bumped" from the row so that the Schensted Insertion Algorithm can be recursively applied to the "bumped" value and the next lower row in the Young tableau. In this sense, Patience Sorting can be viewed as a non-recursive analog of the remarkable Robinson-Schensted-Knuth (or RSK) Algorithm due to G. Robinson [19] for permutations in 1938, C. Schensted [21] for words in 1961, and Knuth [12] for so-called $\mathbb{N}$-matrices in 1970. (See Fulton [10] for a detailed account of the differences between these algorithms.)

Recall that the RSK Algorithm bijectively associates an ordered pair of standard Young tableaux $(P(\sigma), Q(\sigma))$ to each permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n \in \mathfrak{S}_n$ by first building a so-called "insertion tableau" $P(\sigma)$ through repeated Schensted Insertion of the components $\sigma_1, \sigma_2, \ldots, \sigma_n$ into an initially empty tableau. It also simultaneously constructs the "recording tableau" $Q(\sigma)$ by literally *recording* how $P(\sigma)$ is formed. These tableaux have the same shape (a partition λ of n, denoted $\lambda \vdash n$), and this correspondence has many interesting properties. E.g., RSK applied to a permutation is symmetric in the sense that if $\sigma \in \mathfrak{S}_n$ corresponds to the ordered pair of tableaux $(P(\sigma), Q(\sigma))$, then $(Q(\sigma), P(\sigma))$ corresponds to the inverse permutation σ^{-1}. As a result, there is a bijection between the set of involutions $\mathfrak{I}_n \subset \mathfrak{S}_n$ and the set $\mathfrak{T}_n$ of all standard Young tableaux with entries $1, 2, \ldots, n$.

In this paper we develop a bijection extension of Algorithm 1.1 and then study analogues for such properties of RSK. To facilitate this, we first characterize in Section 2 when two permutations have the same pile configurations under Algorithm 1.1. This yields an equivalence relation $\overset{PS}{\sim}$ on $\mathfrak{S}_n$ that is analogous to the Knuth relation $213 \overset{RSK}{\sim} 231$. (Recall that the Knuth relations describe when two permutations have the same "insertion tableau" P under RSK; see Sagan [20].)

In Section 3 we then explicitly describe a bijection between $\mathfrak{S}_n$ and certain pairs of pile configurations having the same shape (a composition γ of n, denoted $\gamma \multimapinv n$). Since there are many more possible pile configurations than standard Young tableaux, it is necessary to specify which pairs are possible; this turns out to be related to the other Knuth relation $312 \overset{RSK}{\sim} 132$. Moreover, this bijection shares the same symmetry property as RSK, and so we can immediately characterize a certain collection of pile configurations that are in bijection with the set of involutions $\mathfrak{I}_n$ (as well as with the set $\mathfrak{T}_n$).

In Section 4 we conclude by using the equivalence relation $\overset{PS}{\sim}$ to characterize and enumerate the set $S_n(3\text{-}\bar{1}\text{-}42)$ of permutations avoiding the generalized barred permutation pattern 3-$\bar{1}$-42. Such permutations avoid the pattern 2-31 unless it is contained in a 3-1-42 pattern.

Another interesting property of RSK is that, given $\sigma \in \mathfrak{S}_n$, the number of boxes in the top row of the "insertion tableau" $P(\sigma)$ is exactly the length of the longest increasing subsequence in σ. (This was first proven by Schensted [21] but is now a special case of Greene's Theorem [11]). Due to the similarity between the Schensted Insertion Algorithm and Algorithm 1.1, it is clear that the cards atop the piles when Patience Sorting terminates will be exactly the elements in the top row of $P(\sigma)$. Thus, the number of piles formed under Patience Sorting is also equal to the length of the longest increasing subsequence in σ, and so one can apply the recent but now highly celebrated Baik-Deift-Johansson Theorem [3] in order to get the asymptotic distribution for the number of piles (up to rescaling). Due to this deep connection between Patience Sorting and Probabilistic Combinatorics, it has been suggested (see, e.g., [13], [14] and [18]; cf. [7]) that studying generalizations of Patience

Sorting might be the key to tackling certain open problems that can be viewed from the standpoint of Random Matrix Theory—the most notable being the Riemann Hypothesis.

At the same time, there is a lot more to Patience Sorting than just resembling the RSK Algorithm for permutations. E.g., after applying Algorithm 1.1 to a deck of cards, it is easy to recollect each card in ascending order from amongst the current top cards of the piles (and thus complete A.S.C. Ross' card sorting algorithm). While this is not necessarily the fastest sorting algorithm one can apply to a deck of cards, the *patience* in *Patience Sorting* is not intended to describe a prerequisites for its use. Instead it refers to how pile formation in Algorithm 1.1 resembles the way in which one places cards into piles when playing the popular single-person card game *Klondike Solitaire*, which is often called *Patience* in the UK. This is more than a coincidence, though, as Algorithm 1.1 also happens to be an optimal strategy (in the sense of forming as few piles as possible; see [1] for a proof) when playing an idealized model of Klondike Solitaire known as *Floyd's Game*:

Game 1.2 (Floyd's Game)**.** Given a shuffled deck of cards $c_1, c_2, \ldots, c_n$,

- Place the first card c_1 from the deck into a pile r_1 by itself.
- Then for each card c_i ($i = 2, \ldots, n$), either
  - put c_i into a new pile by itself or
  - play c_i on top of any pile whose current top card is larger than c_i.
- The object of the game is to end with as few piles as possible.

In other words, the cards are played one at a time according to the order they appear in the deck so that piles are created in much the same way they are formed under Patience Sorting. According to [1], Floyd's Game was developed independently of Mallow's work and originated in unpublished correspondence between Computer Scientists Bob Floyd and Donald Knuth during 1964.

Note that unlike Klondike Solitaire, there is a known strategy (Algorithm 1.1) for Floyd's Game under which one will always win. In fact, Klondike Solitaire—though so popular that it has come pre-installed on the vast majority of personal computers shipped since 1989—is very poorly understood mathematically. (Recent progress, however, has been made in developing an optimal strategy for a version called *thoughtful solitaire* [25].) As such, Persi Diaconis ([1] and private communication with the second author) has suggested that a deeper understanding of Patience Sorting and its generalization would undoubtedly help in developing a better mathematical model for analyzing Klondike Solitaire.

2. Pile Configurations Resulting from Patience Sorting

2.1. Pile Configurations, Shadow Diagrams, and Reverse Patience Words.

We begin by explicitly characterizing the pile configurations that result from applying Patience Sorting (Algorithm 1.1) to a permutation:

Lemma 2.1. *Let $\sigma \in \mathfrak{S}_n$ be a permutation and $R(\sigma) = \{r_1, r_2, \ldots, r_k\}$ be the pile configuration associated to σ. Then $R(\sigma)$ is a partition of $[n] = \{1, 2, \ldots, n\}$ such that denoting $r_j = \{r_{j1} > r_{j2} > \cdots > r_{js_j}\}$,*

$$r_{js_j} < r_{is_i} \quad \text{if} \quad j < i \tag{2.1}$$

Moreover, for every set partition $S = \{s_1, s_2, \ldots, s_k\}$ satisfying Equation (2.1), there is a permutation $\sigma \in \mathfrak{S}_n$ such that $R(\sigma) = S$.

Proof. Omitted. □

We will often express a pile configuration R with its constituent piles $r_1, r_2, \ldots, r_k$ written vertically and bottom-justified with respect to the largest value r_{j1} in each pile r_j. This motivate the following definition:

Definition 2.2. The *reverse patience word* $RPW(R)$ for a pile configuration R is the permutation formed by concatenating the piles $r_1, r_2, \ldots, r_k$ together with each pile r_j

written in decreasing order (i.e., read from bottom to top in order from left to right). In the notation of Lemma 2.1,

$$RPW(R) = r_{11}r_{12}\cdots r_{1s_1}\ r_{21}r_{22}\cdots r_{2s_2}\ \ \cdots\ \ r_{k1}r_{k2}\cdots r_{ks_k}$$

Example 2.3. The pile configuration $R = \{\{6 > 4 > 1\}, \{5 > 2\}, \{8 > 7 > 3\}\}$ is represented by the piles

$$\begin{array}{ccc} 1 & & 3 \\ 4 & 2 & 7 \\ 6 & 5 & 8 \end{array}$$

and has the reverse patience word $RPW(R) = 64152873$.

The following Lemma should be clear from the above definitions and example:

Lemma 2.4. *Given a permutation $\sigma \in \mathfrak{S}_n$, $R(RPW(R(\sigma))) = R(\sigma)$.*

Proof. Omitted. □

At the same time, it is also clear that in general there will be many permutations $\tau \in \mathfrak{S}_n$ for which $R(\sigma) = R(\tau)$. In Section 2.2 below we characterize when two permutations have the same pile configuration, and we will denote this equivalence relation by $\sigma \overset{PS}{\sim} \tau$. Moreover, we will also see that the reverse patience word $RPW(R(\sigma))$ is the most natural representative for the equivalence class generated by σ.

We close this section by giving an alternate characterization for pile configurations in terms of the so-called *shadow diagram* construction that G. Viennot [23] introduced in the context of studying the RSK Algorithm for permutations.

Definition 2.5. Given a lattice point $(m, n) \in \mathbb{Z}^2$, we define the *(northeast) shadow* of (m, n) to be the quarter space $S(m, n) = \{(x, y) \in \mathbb{R}^2 \mid x \geq m,\ y \geq n\}$.

See Figure 2.1(a) for an example of a point's shadow.

The most important use of shadows is in building shadowlines:

Definition 2.6. Given lattice points $(m_1, n_1), (m_2, n_2), \ldots, (m_k, n_k) \in \mathbb{Z}^2$, we define their *(northeast) shadowline* to be the boundary of the quarter space formed by taking the union of the shadows $S(m_1, n_1), S(m_2, n_2), \ldots, S(m_k, n_k)$.

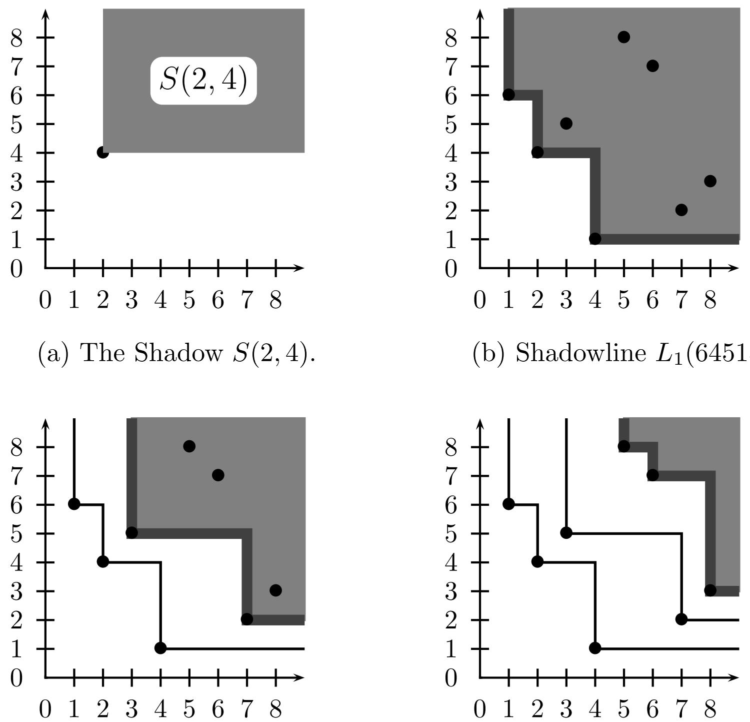
In particular, we wish to associate to each permutation a certain collection of shadowlines (as illustrated in Figure 2.1(b)–(d)):

Definition 2.7. Given a permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n \in \mathfrak{S}_n$, the *(northeast) shadow diagram* of σ consists of the shadowlines $L_1(\sigma), L_2(\sigma), \ldots, L_k(\sigma)$ formed as follows:

- $L_1(\sigma)$ is the shadowline for the lattice points $\{(1, \sigma_1), (2, \sigma_2), \ldots, (n, \sigma_n)\}$.
- While at least one of the points $(1, \sigma_1), (2, \sigma_2), \ldots, (n, \sigma_n)$ is not contained in the the shadowlines $L_1(\sigma), L_2(\sigma), \ldots, L_j(\sigma)$, define $L_{j+1}(\sigma)$ to be the shadowline for the points
$$\{(i, \sigma_i) \mid (i, \sigma_i) \notin \bigcup_{k=1}^{j} L_k(\sigma)\}.$$

In other words, we define the shadow diagram inductively by taking $L_1(\sigma)$ to be the shadowline for the diagram $\{(1, \sigma_1), (2, \sigma_2), \ldots, (n, \sigma_n)\}$ of the permutation. Then we ignore the points whose shadows were actually used in building $L_1(\sigma)$ and define $L_2(\sigma)$ to be the shadowline of the resulting subset of the permutation's diagram. We then build $L_3(\sigma)$ as the shadowline for the points not yet used in constructing both $L_1(\sigma)$ and $L_2(\sigma)$, and this process continues until all points in the permutation diagram are exhausted.

One of the most basic properties of the shadow diagram for a permutation σ is that it encodes the top row of the insertion tableau $P(\sigma)$ (resp. recording tableau $Q(\sigma)$) as



(a) The Shadow $S(2,4)$.

(b) Shadowline $L_1(64518723)$.

(c) Shadowline $L_2(64518723)$.

(d) Shadowline $L_3(64518723)$.

FIGURE 2.1. Examples of Shadow and Shadowline Constructions

the smallest ordinates (resp. smallest abscissae) of all points belonging to the constituent shadowlines $L_1(\sigma), L_2(\sigma), \ldots, L_k(\sigma)$. (A proof of this can be found in Sagan [20].) In particular, this means that if σ has pile configuration $R(\sigma) = \{r_1, r_2, \ldots, r_m\}$, then $m = k$ since the number of piles is equal to the length of the top row of $P(\sigma)$. We can say even more about the relationship between $L_1(\sigma), L_2(\sigma), \ldots, L_k(\sigma)$ and $R(\sigma)$ when both are viewed in terms of left-to-right minima subsequences (a.k.a. basic subsequences or records):

Definition 2.8. 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 subsequence* of π consists of those $\pi_j = \min\{\pi_i \mid 1 \leq i \leq j\}$.

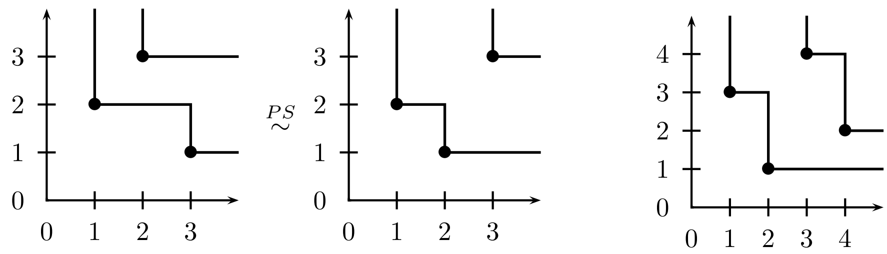
We then inductively define the left-to-right minima subsequences $s_1, s_2, \ldots, s_k$ of a permutation σ by taking s_1 to be the left-to-right minima subsequence for σ itself and then s_i to be the left-to-right minima subsequence for the partial permutation obtained by removing the elements of $s_1, s_2, \ldots, s_{i-1}$ from σ.

Lemma 2.9. *Suppose that $\sigma \in \mathfrak{S}_n$ has shadow diagram $L_1(\sigma), L_2(\sigma), \ldots, L_k(\sigma)$. Then the ordinates of the southwest corners of L_j are exactly the cards in the j^{th} pile $r_j \in R(\sigma)$ formed by applying Patience Sorting to σ.*

Proof. The left-to-right minima subsequence s_i of σ consists of those elements σ_t that appears at the end of an increasing subsequence of length i but not at the end of an increasing subsequence of length $i + 1$. Thus, since each element added to a pile must be smaller than all other elements already in the pile, $s_1 = r_1$. It then follows by induction that $s_i = r_i$ for $i = 2, \ldots, k$.

The proof that the ordinates of the southwest corners of the L_i are also the elements of the left-to-right minima subsequences s_i is similar. ☐

Lemma 2.9 gives a particularly nice correspondence between the piles formed under Patience Sorting and the shadowlines forming the shadow diagram of a permutation. In



(a) "Stretching" shadowlines effects $231 \overset{PS}{\sim} 213$. Thus, $\underset{\sim}{231} = \{231, 213\}$.

(b) No "stretching" can interchange "4" and "2".

FIGURE 2.2. Examples of patience sorting equivalence and non-equivalence

particular, we have that forming $RPW(R(\sigma))$ essentially amounts to sorting σ into left-to-right minima subsequences.

We will rely heavily upon this correspondence in the sections below.

2.2. **Permutations Having Equivalent Pile Configurations.** In this section we characterize the following equivalence relation:

Definition 2.10. Two permutations $\sigma, \tau \in \mathfrak{S}_n$ are said to be *patience sorting equivalent*, written $\sigma \overset{PS}{\sim} \tau$, if they have the same pile configuration $R(\sigma) = R(\tau)$ under Algorithm 1.1. We denote the equivalence class generated by σ as $\underset{\sim}{\sigma}$.

By Lemma 2.9 in Section 2.1 above, the pile configurations $R(\sigma)$ and $R(\tau)$ correspond to certain shadow diagrams. Thus, it should be intuitive clear that preserving a given pile configuration is equivalent to preserving the ordinates for the southwest corners of the shadowlines. In particular, this means that we are limited to horizontally "stretching" shadowlines up to the point of not allowing them to cross as is illustrated in Figure 2.2 and the following examples.

Example 2.11. The only non-singleton patience sorting equivalence class for $\mathfrak{S}_3$ consists of $\underset{\sim}{231} = \{231, 213\}$. We illustrate $231 \overset{PS}{\sim} 213$ in Figure 2.2(a).

Notice that the actual values of the elements interchanged in Example 2.11 are immaterial so long as they have the same relative magnitudes as the literal values in the word 231. (I.e., they have to be order-isomorphic.) Moreover, it should also be clear that any value greater than the element playing the role of "1" can be inserted between the elements playing the roles of "2" and "3" without affecting the ability to interchange the "1" and "3" elements. Problems with this interchange only start to arise when a value smaller than the element playing the role of "1" is inserted between the elements playing the roles of "2" and "3". We can formally describe this idea using the language of generalized permutation patterns (as was recently defined in [2]; cf. [4]).

Definition 2.12. Let $\sigma = \sigma_1\sigma_2 \cdots \sigma_n \in \mathfrak{S}_n$ and $\tau \in \mathfrak{S}_m$ for $m \leq n$. Then we say that σ *contains the (classical) pattern* τ 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 τ.

If σ does not contain τ, then we say that σ *avoids* the pattern τ, and we denote by $S_n(\tau)$ the subset of the symmetric group $\mathfrak{S}_n$ that avoids τ.

Note that the elements in the subsequence $\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_m}$ are not required to be contiguous in σ. In a *generalized pattern* one assumes that every element in the subsequence must be taken contiguously unless a dash is inserted in the pattern τ between elements that are not required to be contiguous in σ. (A generalized patterns with no dashes is sometimes called a *segment* or a *consecutive pattern*.)

Example 2.13.

(1) Notice that 2431 contains a 2-31 pattern as the bold underlined subsequence **<u>2</u>4<u>31</u>**. Moreover, it is clear that 2431 $\overset{PS}{\sim}$ 2413.

(2) Even though 3142 contains a 2-31 pattern (as the subsequence **<u>3</u>1<u>42</u>**), we cannot interchange "4" and "2", and so $R(3142) \neq R(3124)$. As illustrated in Figure 2.2(b), this is because "4" and "2" are on the same shadowline.

We can now state our main result on patience sorting equivalence:

Theorem 2.14. *Let $\sigma, \tau \in \mathfrak{S}_n$. Then σ and τ have the same pile configurations under Algorithm 1.1 (so that $\sigma \overset{PS}{\sim} \tau$) if and only if there exists a sequence of* 2-31 *to* 2-13 *interchanges (with no* 2-31 *pattern contained in a* 3-1-42 *pattern) that transform σ into τ.*

In other words, $\overset{PS}{\sim}$ is the transitive closure of such interchanges.

Proof. (Sketch) By Lemma 2.9 it suffices to show that 2-31 to 2-13 interchanges (with no 2-31 pattern contained in a 3-1-42 pattern), preserve the left-to-right minima subsequences $s_1, s_2, \ldots, s_k$ of σ. This amounts to showing by induction that such interchanges suffice to transform σ into $RPW(R(\sigma))$ via the sequence of pattern interchanges

$$\sigma = \sigma_0 \rightsquigarrow \sigma_1 \rightsquigarrow \sigma_2 \rightsquigarrow \cdots \rightsquigarrow \sigma_l = RPW(R(\sigma))$$

where each $\sigma_i \overset{PS}{\sim} \sigma_{i+1}$. □

Remark 2.15. It follows from Theorem 2.14 that Examples 2.11 and 2.13(2) sufficiently characterize when two permutations yield the same pile configurations under Patience Sorting. However, it is worth pointing out that these examples also begin to illustrate how one can build an infinite sequence of generalized permutation patterns (all of them containing either 2-13 or 2-31) with the following property: an interchange of the pattern 2-13 with the pattern 2-31 is allowed within an odd-length pattern in this sequence unless the elements used to form the odd-length pattern can also be used as part of a longer even-length pattern in this sequence.

Example 2.16. Even though the permutation 34152 contains a 3-1-42 pattern in the suffix "4152", one can still directly interchange the "5" and the "2" because of the "3" prefix (or via the following sequence of interchanges: 34152 $\rightsquigarrow$ 31452 $\rightsquigarrow$ 31425 $\rightsquigarrow$ 34125).

3. Bijectively Extending Patience Sorting to "Stable Pairs"

3.1. **The Extended Patience Sorting Algorithm.** Recall from Section 1 that Patience Sorting (Algorithm 1.1) can be viewed as an iterated, non-recursive form of the Schensted Insertion Algorithm for inserting a value into the top row of a Young Tableau. In this section we extend the Patience Sorting construction so that it becomes a full non-recursive analog of the RSK Algorithm for permutations. In particular, we mimic the RSK recording tableau construction so that "recording piles" are formed while assembling the usual pile configuration (which we will similarly now call "insertion piles") under Patience Sorting:

Algorithm 3.1 (Extended Patience Sorting Algorithm)**.** Given a shuffled deck of cards $\sigma = c_1 c_2 \cdots c_n$, inductively build *insertion piles* $R = R(\sigma) = \{r_1, r_2, \ldots, r_m\}$ and *recording piles* $S = S(\sigma) = \{s_1, s_2, \ldots, s_m\}$ 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.

We call the pile configuration pairs that result from Algorithm 3.1 *stable pairs* and give a characterization for them in Section 3.2 below. Note that the pile configurations that comprise a resulting stable pair must have the same "shape", which we define as follows.

Definition 3.2. 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 $\mathrm{sh}(R) = \gamma \models n$.

Example 3.3. Let $\sigma = 64518723 \in \mathfrak{S}_8$. Then according to Algorithm 3.1 we simultaneously form the following pile configurations with shape $\mathrm{sh}(R(\sigma)) = \mathrm{sh}(S(\sigma)) = (3, 2, 3)$:

	insertion piles	recording piles		insertion piles	recording piles
Form a new pile with **6**:	$\begin{array}{l} \mathbf{6} \end{array}$	$\begin{array}{l} \mathbf{1} \end{array}$	Then play the **4** on it:	$\begin{array}{l} \mathbf{4} \\ 6 \end{array}$	$\begin{array}{l} 1 \\ \mathbf{2} \end{array}$
Form a new pile with **5**:	$\begin{array}{ll} 4 & \\ 6 & \mathbf{5} \end{array}$	$\begin{array}{ll} 1 & \\ 2 & \mathbf{3} \end{array}$	Add the **1** to left pile:	$\begin{array}{ll} \mathbf{1} & \\ 4 & \\ 6 & 5 \end{array}$	$\begin{array}{ll} 1 & \\ 2 & \\ \mathbf{4} & 3 \end{array}$
Form a new pile with **8**:	$\begin{array}{lll} 1 & & \\ 4 & & \\ 6 & 5 & \mathbf{8} \end{array}$	$\begin{array}{lll} 1 & & \\ 2 & & \\ 4 & 3 & \mathbf{5} \end{array}$	Then play the **7** on it:	$\begin{array}{lll} 1 & & \\ 4 & & \mathbf{7} \\ 6 & 5 & 8 \end{array}$	$\begin{array}{lll} 1 & & \\ 2 & & 5 \\ 4 & 3 & \mathbf{6} \end{array}$
Add the **2** to a pile:	$\begin{array}{lll} 1 & & \\ 4 & \mathbf{2} & 7 \\ 6 & 5 & 8 \end{array}$	$\begin{array}{lll} 1 & & \\ 2 & 3 & 5 \\ 4 & \mathbf{7} & 6 \end{array}$	Add the **3** to a pile:	$\begin{array}{lll} 1 & & \mathbf{3} \\ 4 & 2 & 7 \\ 6 & 5 & 8 \end{array}$	$\begin{array}{lll} 1 & & 5 \\ 2 & 3 & 6 \\ 4 & 7 & \mathbf{8} \end{array}$

The idea behind Algorithm 3.1 is that we are using the recording piles $S(\sigma)$ to implicitly label the order in which the elements of the permutation σ are added to the insertion piles $R(\sigma)$. It is clear that this information then allows us to uniquely reconstruct σ by reversing the order in which the cards were played. However, even though reversing the Extended Patience Sorting Algorithm is much easier than reversing the RSK Algorithm through recursive "reverse row bumping", the trade-off is that the stable pairs that result from the former are not independent whereas the tableau pairs generated by RSK are completely independent (up to shape).

That $S(\sigma) = \{s_1, s_2, \ldots, s_m\}$ records the order of the cards being added to the insertion piles is made clear if we alternatively add cards to the tops of new piles s'_j in Algorithm 3.1 instead of to the bottoms of the piles s_j. This yields modified recording piles $S'(\sigma)$ from which each original recording pile $s_j \in S(\sigma)$ can be recovered by simply reflecting the corresponding pile s'_j vertically.

Example 3.4. As in Example 3.3 above, let $\sigma = 64518723 \in \mathfrak{S}_8$. Then $R(\sigma)$ is formed as before and

$$S'(\sigma) \;=\; \begin{array}{ccc} 4 & & 8 \\ 2 & 7 & 6 \\ 1 & 3 & 5 \end{array} \;\overset{reflect}{\dashrightarrow}\; \begin{array}{ccc} 1 & & 5 \\ 2 & 3 & 6 \\ 4 & 7 & 8 \end{array} \;=\; S(\sigma)$$

We are now in a position to prove that the Extended Patience Sorting Algorithm has the same form of symmetry as the RSK Algorithm has for permutations.

Proposition 3.5. *Let $(R(\sigma), S(\sigma))$ be the insertion and recording piles, respectively, formed by applying Algorithm 3.1 to $\sigma \in \mathfrak{S}_n$. Then reversing Algorithm 3.1 for $(S(\sigma), R(\sigma))$ yields the inverse permutation σ^{-1}.*

Proof. Construct $S'(\sigma)$ from $S(\sigma)$ as discussed above, and form the n ordered pairs (r_{ij}, s'_{ij}) where i indexes the individual piles and j the cards in the i^{th} piles. Then these n points correspond to the diagram of a permutation $\tau \in \mathfrak{S}_n$. However, since reflecting these points through the line $y = x$ yields the diagram for σ, it follows that $\tau = \sigma^{-1}$. □

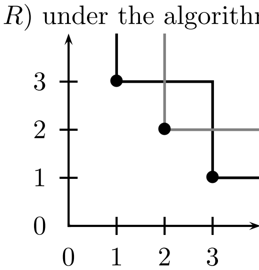
Proposition 3.5 suggests that Algorithm 3.1 is the right generalization of Algorithm 1.1 since we obtain the same symmetry property as for RSK. At the same time, though, since there are many more possible pile configurations than standard Young Tableau (as we'll show in Section 4 below), not every ordered pair of pile configurations with the same shape will result from Algorithm 3.1. Thus, it is necessary to first characterize the "stable pairs" that result from applying Extended Patience Sorting to a permutation.

3.2. **Characterizing "Stable Pairs" and Pile Configurations for Involutions.** Based upon Proposition 3.5 above, there is a bijection between involutions and certain pile configurations. We will describe this bijection as a corollary to the more general construction for the "stable pairs" of pile configurations that can result from apply the Extended Patience Sorting Algorithm to a permutation.

The following example, though very small, illustrates the most generic behavior that must be avoided in constructing stable pairs. As in section 3.1 above, we denote by S' the "reverse pile configuration" of S (which has all piles listed in reverse order).

Example 3.6. Even though the pile configuration $R = \{\{3 > 1\}, \{2\}\}$ cannot result as the insertion piles given by an involution under the Extended Patience Sorting Algorithm , we can still try to look at the pre-image of the pair (R, R) under the algorithm:

$$R \quad = \quad \begin{array}{cc} 1 & \\ 3 & 2 \end{array} \quad \text{and} \quad S' \quad = \quad \begin{array}{cc} 3 & \\ 1 & 2 \end{array} \quad \Longrightarrow$$



Note that there are two competing constructions here. On the one hand we have the diagram $\{(1,3), (2,2), (3,1)\}$ of a permutation given by the entries in the pile configurations. (In particular, the values in R specify the ordinates and the values in the corresponding boxes of S' the abscissae.) On the other hand, the piles in R also specify the shadowlines for this permutation diagram. Here the pair (R, S) of pile configurations is "unstable" because their combination yields crossing shadowlines—which is clearly not allowed.

We can now make the following important definitions:

Definition 3.7. Given a composition γ of n (denoted $\gamma \multimap n$), we define the set $\mathfrak{P}_\gamma(n)$ to be all pile configurations R such that $\operatorname{sh}(R) = \gamma$ and set

$$\mathfrak{P}(n) = \bigcup_{\gamma \multimap n} \mathfrak{P}_\gamma(n).$$

Definition 3.8. Define the set $\Sigma(n) \subset \mathfrak{P}(n) \times \mathfrak{P}(n)$ to consist of all ordered pairs (R, S) with $\operatorname{sh}(R) = \operatorname{sh}(S)$ such that if $RPW(R)$ contains a 31-2 pattern as a subword ω, then $RPW(S')$ avoids a 13-2 patterns in the subword whose elements have the same positions in $RPW(S')$ as ω does in $RPW(R)$.

In other words, Definition 3.8 characterizes "stable pairs" of pile configurations (R, S) by forcing R and S to avoid certain sub-pile pattern pairs. As in Example 3.6, we are characterizing when the induced shadowlines cross.

Theorem 3.9. *Extended Patience Sorting (Algorithm 3.1) gives a bijection between the symmetric group $\mathfrak{S}_n$ and the "stable pairs" set $\Sigma(n)$ given in Definition 3.8 above.*

Proof. Omitted. □

We illustrate this general form for these "forbidden sub-pile patterns" in the following example:

Example 3.10. For $x_1 < x_2 < x_3$ and $y_1 < y_2 < y_3$, we forbid the following simultaneous sub-pile patterns:

$$\begin{array}{cc} \uparrow & \\ (i\text{ cards}) & \uparrow \\ \downarrow & (j\text{ cards}) \\ y_1 & \downarrow \\ y_3 & y_2 \\ \uparrow & \uparrow \\ (k\text{ cards}) & (l\text{ cards}) \\ \downarrow & \downarrow \\ \hline \end{array} \quad \subset \quad R \qquad \text{and} \qquad \begin{array}{cc} \uparrow & \\ (k\text{ cards}) & \uparrow \\ \downarrow & (l\text{ cards}) \\ x_1 & \downarrow \\ x_3 & x_2 \\ \uparrow & \uparrow \\ (i\text{ cards}) & (j\text{ cards}) \\ \downarrow & \downarrow \\ \hline \end{array} \quad \subset \quad S$$

The reason we disallow these sub-pile patterns is clear from the diagram given in Example 3.6 above: these patterns cause the partial shadowlines dictated by the sub-pile pattern in R to necessarily cross when applied to the lattice points $(x_1, y_3), (x_2, y_2), (x_3, y_1)$ given by the sub-pile patterns in both R and S.

Based upon the characterization of stable pairs given in Theorem 3.9 and the Symmetry Property proven in Proposition 3.5, we can immediately describe a bijection between involutions and certain pile configurations. In particular, these pile configurations must avoid simultaneously containing the symmetric sub-pile patterns as given in Example 3.10.

This corresponds to the reverse patience word for a pile configuration simultaneously avoiding a symmetric pair of the generalized patterns 31-2 and 32-1. As such it is interesting to compare this construction to two results recently obtained by Claesson and Mansour [6]:

(1) The size of $S_n(3\text{-}12, 3\text{-}21)$ is equal to the number of involutions $|\mathfrak{I}_n|$ in $\mathfrak{S}_n$.

(2) The size of $S_n(31\text{-}2, 32\text{-}1)$ is 2^{n-1}.

The first result suggests that there should be a way to relate the result in Theorem 3.9 to simultaneous avoidance of the very similar patterns 3-12 and 3-21. The second result suggests that restricting to complete avoidance of all simultaneous occurrences of 31-2 and 32-1 will yield a natural bijection between $S_n(31\text{-}2, 32\text{-}1)$ and a subset $\mathfrak{N} \subset \mathfrak{P}(n)$ such that $\mathfrak{N} \cap \mathfrak{P}_\gamma(n)$ contains exactly one pile configuration of each shape γ. A natural family for this collection of pile configurations consists of what we call *non-crossing pile configurations*; namely, for the composition $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_k) \multimap n$,

$$\mathfrak{N} \cap \mathfrak{P}_\gamma(n) = \{\{\gamma_1 > \cdots > 1\}, \{\gamma_1 + \gamma_2 > \cdots > \gamma_1 + 1\}, \ldots, \{n > \cdots > n - \gamma_{k-1}\}\}$$

so that there are exactly 2^{n-1} such pile configurations. One can also show that $\mathfrak{N}$ is the image $R(S_n(3\text{-}1\text{-}2))$ of all permutations avoiding the classical pattern 3-1-2 under the Patience Sorting Algorithm.

4. Enumerating $S_n(3\text{-}\bar{1}\text{-}42)$

In this section we use the results from Section 2 to both enumerate and characterize the permutations that avoid the generalized permutation pattern 2-31 unless it's part of the generalized pattern 3-1-42. We call this restricted form of the generalized pattern 2-31 a *(generalized) barred permutation pattern* and denote it by $3\text{-}\bar{1}\text{-}42$. (This notation is due to J. West, et al., and first appeared in the study of two-stack sortable permutations [8, 9, 24].)

Theorem 4.1.

(1) *The set of permutations* $S_n(3\text{-}\bar{1}\text{-}42)$ *that avoid the pattern* $3\text{-}\bar{1}\text{-}42$ *is exactly the set* $RPW(R(\mathfrak{S}_n))$ *of reverse patience words obtainable from the symmetric group* $\mathfrak{S}_n$.

(2) *The size of* $S_n(3\text{-}\bar{1}\text{-}42)$ *is given by the* n^{th} *Bell number* B_n.

Proof.

(1) Let $\sigma \in S_n(3\text{-}\bar{1}\text{-}42)$. Then for $i = 1, 2, \ldots, n-1$, define $\sigma_{m_i} = \min\{\sigma_j \mid i \leq j \leq n\}$. Since σ avoids $3\text{-}\bar{1}\text{-}42$, the subpermutation $\sigma_i \sigma_{i+1} \cdots \sigma_{m_i}$ must be a decreasing subsequence of σ. (Otherwise σ would necessarily contain a 2-31 pattern that is not part of a 3-1-42 pattern.) It follows that the left-to-right minima subsequences $s_1, s_2, \ldots, s_k$ of σ must be disjoint and satisfy Equation (2.1) so that the result follows by Lemmas 2.1 and 2.9.

(2) Recall that the Bell number B_n enumerates the set partitions of $[n] = \{1, 2, \ldots, n\}$. From Part (1), the elements of $S_n(3\text{-}\bar{1}\text{-}42)$ are in bijection with pile configurations. Thus, since pile configurations are themselves set partitions, we need only show that every set partition is also a pile configuration. This follows by ordering the components of a given set partition by their smallest element so that Equation (2.1) is satisfied.

□

Remark 4.2. We conclude by remarking that even though the set $S_n(3\text{-}\bar{1}\text{-}42)$ is enumerated by the very well known Bell numbers, it cannot be described in a simpler way using classical pattern avoidance. This means that there does not exist a countable set of non-generalized (a.k.a. classical) permutation patterns $\tau_1, \tau_2, \ldots$ such that

$$S_n(3\text{-}\bar{1}\text{-}42) = S_n(\tau_1, \tau_2, \ldots) = \bigcap_{i \geq 1} S_n(\tau_i).$$

There are two very important reasons that this cannot happen:

First of all, the Bell numbers satisfy $\log B_n = n(\log n - \log\log n + O(1))$ and so exhibit superexponential growth. However, in light of the Stanley-Wilf ex-Conjecture (which was recently proven by Marcus and Tardos [17]), the set of permutations $S_n(\tau)$ avoiding any classical pattern τ can only grow at most exponentially in n.

On the other hand, the class of permutations

$$S(3\text{-}\bar{1}\text{-}42) = \bigcup_{n \geq 4} S_n(3\text{-}\bar{1}\text{-}42)$$

is not closed under taking order-isomorphic subpermutations, whereas it is easy to see that classes of permutations defined by classical pattern avoidance must be closed. (See Bóna [4], Chap. 5.) In particular, the permutation $3142 \in S(3\text{-}\bar{1}\text{-}42)$ but $231 \notin S(3\text{-}\bar{1}\text{-}42)$.

At the same time, Theorem 4.1(2) implies that $3\text{-}\bar{1}\text{-}42$ belongs to the so-called *Wilf Equivalence class* for the generalized pattern 1-23. That is, if

$$\tau \in \{1\text{-}23, 3\text{-}21, 12\text{-}3, 32\text{-}1, 1\text{-}32, 3\text{-}12, 21\text{-}3, 23\text{-}1\}$$

then the size of the avoidance class $S_n(\tau)$ is also given by the n^{th} Bell number B_n. In particular, Claesson [5] showed that $|S_n(23\text{-}1)| = B_n$ via direct bijection between permutations avoiding 23-1 and set partitions. Furthermore, in any permutation $\sigma \in S_n(3\text{-}\bar{1}\text{-}42)$ each segment between consecutive right-to-left minima must be a single decreasing run (when from read left to right), so it is easy to see that $S_n(3\text{-}\bar{1}\text{-}42) = S_n(23\text{-}1)$. Thus, the barred pattern $3\text{-}\bar{1}\text{-}42$ and the generalized pattern 23-1 are not just in the same Wilf equivalence class but also have identical avoidance classes.

Still, even though $S_n(3\text{-}\bar{1}\text{-}42) = S_n(23\text{-}1)$, it is more natural to use avoidance of 3-$\bar{1}$-42 when studying Patience Sorting. Fundamentally, this lets us look at $S_n(3\text{-}\bar{1}\text{-}42)$ as the set of equivalence classes in $\mathfrak{S}_n$ modulo 3-$\bar{1}$-42 $\overset{PS}{\sim}$ 3-$\bar{1}$-24, where each equivalence class corresponds to a unique pile configuration. The same equivalence relation is not easy to describe when starting with an occurrence of 23-1. (Note that 23-1 $\sim$ 2-13 or 23-1 $\sim$ 21-3 is wrong since we would incorrectly get 2431 $\sim$ 2314 or 2431 $\sim$ 2134 instead of the correct 2431 $\sim$ 2413).

This suggests that there is even more information about pattern avoidance to be gotten from such a simple algorithm as Patience Sorting.

References

[1] 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.

[2] E. Babson and E. Steingrímsson. "Generalized Permutation Patterns and a Classification of the Mahonian Statistics", *Séminaire Lotharingien de Combinatoire*, **B44b** (2000), 18 pp.

[3] J. Baik, P. Deift, and K. Johansson. "On the Distribution of the Length of the Longest Increasing Subsequence in a Random Permutation", *J. Amer. Math. Soc.* **12** (1999), 1119–1178.

[4] M. Bóna. *Combinatorics of Permutations.* Chapman & Hall/CRC Press, 2004.

[5] A. Claesson. "Generalized Pattern Avoidance". *Europ. J. Combin.* **22** (2001), 961–971.

[6] A. Claesson and T. Mansour. "Enumerating Permutations Avoiding a Pair of Babson-Steingrímsson Patterns". *Ars Combinatoria*, to appear.

[7] J.B. Conrey. "The Riemann Hypothesis". *Not. Amer. Math. Soc.* **50** (2003), 341–353.

[8] S. Dulucq, S. Gire, and O. Guibert. "A Combinatorial Proof of J. West's Conjecture". *Discrete Math.* **187** (1998), 71–96.

[9] S. Dulucq, S. Gire, and J. West. "Permutations with Forbidden Subsequences and Nonseparable Maps". *Discrete Math.* **153** (1996), 85–103.

[10] W. Fulton. *Young Tableaux.* LMS Student Texts 35. Cambridge University Press, 1997.

[11] C. Greene. "An Extension of Schensted's Theorem". *Adv. in Math.*, **14** (1974), 254–265.

[12] D. E. Knuth. "Permutations, Matrices, and Generalized Young Tableaux", *Pac. J. Math.* **34** (1970), 709–727.

[13] C. Kuykendall. "Analyzing Solitaire". *Science*, **283** (1999), 794–795.

[14] D. Mackenzie. "From Solitaire, a Clue to the World of Prime Numbers". *Science* **282** (1998), 1631–1633.

[15] C. L. Mallows. "Problem 62-2, Patience Sorting". *SIAM Review* **4** (1962), 148–149.

[16] C. L. Mallows. "Problem 62-2". *SIAM Review* **5** (1963), 375–376.

[17] A. Marcus and G. Tardos. "Excluded Permutation Matrices and the Stanley-Wilf Conjecture." *J. Combin. Th. Ser. A* **107** (2004), 153–160.

[18] I. Peterson. "Solitaire-y Sequences". *MathTrek.*
Available online at `http://www.maa.org/mathland/mathtrek_7_5_99.html`

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

[20] B. Sagan. *The Symmetric Group, Second Edition.* Graduate Texts in Mathematics 203. Springer-Verlag, 2000.

[21] C. Schensted. "Longest Increasing and Decreasing Subsequences", *Can. J. Math.* **13** (1961), 179–191.

[22] R. Stanley. *Enumerative Combinatorics, Volume 2.* Studies in Advanced Mathematics 62. Cambridge University Press, 1999.

[23] 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.

[24] J. West. Permutations with Forbidden Subsequences and Stack-sortable Permutations. Ph.D. thesis, M.I.T., 1990.

[25] X. Yan, P. Diaconis, P. Rusmevichientong, and B. Van Roy. "Solitaire: Man Versus Machine," in *Advances in Neural Information Processing Systems 17.* L.K. Saul, Y. Weiss, and L. Bottou, eds. MIT Press, 2005, to appear.

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/`

Some Symmetry and Unimodality Properties of the q, x, y-Hit Numbers

Fred Butler
Department of Mathematics
West Virginia University
P.O. Box 6310
Morgantown, WV 26506-6310
fbutler@math.wvu.edu

Abstract

We prove symmetry and in some cases symmetry and unimodality of polynomials related to the q, x, y-hit numbers introduced by Haglund. These results generalize theorems proven by Haglund for the q-hit numbers. We also apply one of these results to obtain a corollary concerning a generalization of the Eulerian numbers.

Résumé

Nous prouvons la symétrie et dans certains cas la symétrie et l'unimodalité de polynômes relatifs aux q, x, y nombres de contacts introduits par Haglund, généralisant ainsi certains théorèmes. Un de ces résultats nous permet d'obtenir un corollaire à propos d'une généralisation des nombres Eulériens.

1 Introduction

1.1 Preliminaries

We will use the notation SQ_n to denote the $n \times n$ square chess board. We will number the columns of SQ_n with 1 through n going from left to right across the bottom, and the rows of SQ_n with 1 through n going from bottom to top. We will label a square on SQ_n in column i row j with (i, j).

More generally, a *board* will be any subset of SQ_n for some $n \in \mathbb{N}$. A *Ferrers board* is a board with non-decreasing column heights from left to right, or more precisely a board of the form $\{(i, j) \in SQ_n | \ 1 \le j \le b_i, \ 1 \le i \le n\}$ where $b_1 \le b_2 \le \cdots \le b_n$. We will denote the Ferrers board with

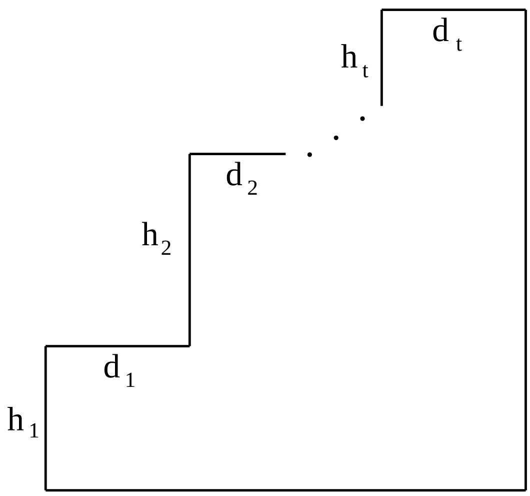


Figure 1: The Ferrers board $B(h_1, d_1; \ldots; h_t, d_t)$.

column heights $b_1, b_2, \ldots b_n$ by $B(b_1, \ldots, b_n)$. We will also specify a Ferrers board by its step heights and depths. The Ferrers board $B(h_1, d_1; \ldots; h_t, d_t)$ is shown in Figure 1. We will call $B = B(b_1, \ldots, b_n) = B(h_1, d_1; \ldots; h_t, d_t)$ a *regular Ferrers board* if $b_i \geq i$ for $1 \leq i \leq n$, or equivalently if $h_1 + \cdots + h_i \geq d_1 + \cdots + d_i$ for $1 \leq i \leq t$ as was defined in [9]. In this paper we will focus on regular Ferrers boards.

A *rook placement* on a board $B \subseteq SQ_n$ is a subset of squares of B such that no two of these squares lie in the same row or the same column. As the name suggests, these squares represent positions on an $n \times n$ chess board where non-attacking rooks can be placed. Let $r_k(B)$ denote the number of k rook placements on B, and let $h_{n,k}(B)$ denote the number of n rook placements on SQ_n such that exactly k rooks lie on B. These are known as the *kth rook number* and the *kth hit number*, respectively, of the board B. Classical rook theory is concerned with studying the relationships between these two numbers.

1.2 Cycle-counting q-rook theory

The cycle-counting q-rook numbers were first introduced in the unpublished work of Ehrenborg, Haglund, and Readdy [4], defined only for Ferrers boards. These rook numbers generalize both the q-rook numbers $R_k(q, B)$ of Garsia and Remmel [5], and the cycle-counting rook numbers $r_k(y, B)$ of Chung and Graham [2]. In order to describe them, we need to define the following three statistics.

The first statistic is denoted inv_B, a generalization of the number of inversions of a permutation. Given a placement P of rooks on a Ferrers board $B \subseteq SQ_n$, let each rook cancel all squares to the right in its row and below in its column. We can then define $inv_B(P)$ to be the number of squares of B which neither contain a rook from P nor are cancelled.

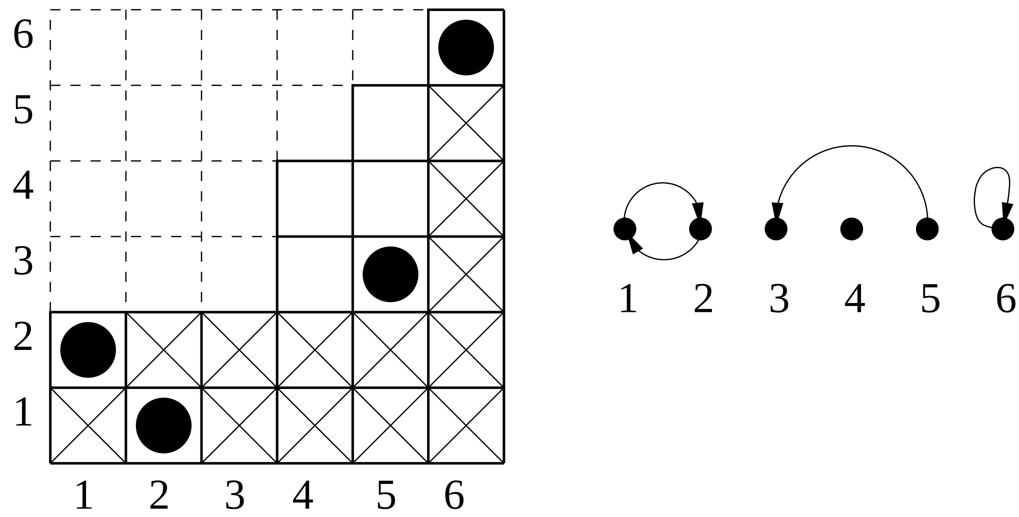


Figure 2: The placement P on B and the associated digraph G_P.

The second statistic is denoted *cyc*, and is a generalization of the number of cycles of a permutation. Given a rook placement P on a board $B \subseteq SQ_n$, it is possible to associate to P a simple directed graph G_P on n vertices. This fact was first noted in [6] (see also [2] and [3]). There is an edge from i to j in G_P if and only if there is a rook from P on the square (i, j). We can then define $cyc(P)$ to be the number of cycles in G_P.

The third statistic, denoted E, depends on the following fact. Given any placement P of j non-attacking rooks in columns 1 through $i-1$ of a Ferrers board B (where $j \leq i-1$), it is an easy exercise to see that if $b_i \geq i$ then there is exactly one square in column i where placement of a rook will complete a new cycle in the digraph G_P. If $b_i < i$ then there is no square where placing a rook will complete a new cycle. Note that a regular Ferrers board will have such a square in each of its columns (since $b_i \geq i$ for all $1 \leq i \leq n$). Now for i with $b_i \geq i$ we can define $s_i(P)$ to be the unique square which, considering only rooks from P in columns 1 through $i-1$ of P, completes a new cycle. Then let $E(P)$ be the number of i such that $b_i \geq i$ and there is no rook from P in column i on or above square $s_i(P)$.

For the rook placement P pictured in Figure 2, we see that $inv_B(P) = 4$, $cyc(P) = 2$, and $E(P) = 2$ (corresponding to $i = 4$ and $i = 5$). We will use the common notation of

$$[x] = \frac{1-q^x}{1-q}$$

to denote the *q-analog* of the real number x. Note that when $x = n \in \mathbb{N}$,

$$[n] = \frac{1-q^n}{1-q} = 1 + q + \cdots + q^{n-1}$$

is a polynomial in q. We use the notation $[n]!$ to denote the q-analog of $n!$, the product $[n][n-1]\cdots[2][1]$. Finally, for $n, k \in \mathbb{N}$ we denote by $\begin{bmatrix} n \\ k \end{bmatrix}$ the

q-analog of the binomial coefficient $\binom{n}{k}$, equal to

$$\frac{[n]!}{[k]![n-k]!} = \frac{[n][n-1]\cdots[n-k+1]}{[k]!}$$

for $k \leq n$ and equal to 0 for $k > n$. The fact that $\begin{bmatrix} n \\ k \end{bmatrix}$ is also a polynomial in q is proven in [10]. More generally for $z \in \mathbb{C}$ we will write $\begin{bmatrix} z \\ k \end{bmatrix}$ for $[z][z-1]\cdots[z-k+1]/[k]!$.

As in [4], we now define the *kth cycle-counting q-rook number* of a Ferrers board B by the equation

$$R_k(y,q,B) = \sum_{P \ k \text{ rooks on } B} [y]^{cyc(P)} q^{inv_B(P)+(y-1)E(P)}. \tag{1}$$

Letting $y = 1$ in (1) yields the q-rook numbers of [5], and letting $q \to 1$ gives the cycle-counting rook numbers of [2]. The $R_k(y,q,B)$ satisfy the useful equation

$$\sum_{k=0}^{n} R_{n-k}(y,q,B)[z][z-1]\cdots[z-k+1] =$$
$$\prod_{i \text{ with } b_i \geq i} [z+b_i-i+y] \prod_{i \text{ with } b_i < i} [z+b_i-i+1], \tag{2}$$

a version of the well-known factorization theorems proven for the $r_k(B)$ [7], $R_k(q,B)$ [5], and $r_k(y,B)$ [2].

Haglund [9] further extended this model by defining the *q,x,y-hit numbers* algebraically by the equation

$$\sum_{k=0}^{n} A_{n,k}(x,y,q,B)z^k = \tag{3}$$

$$\sum_{k=0}^{n} R_{n-k}(y,q,B)[x][x+1]\cdots[x+k-1]z^k \prod_{i=k+1}^{n} (1-zq^{x+i-1}),$$

which generalize the $a_{n,k}(x,y,B)$ also discussed in [9]. The case $x = y$ is studied in [1], where a combinatorial interpretation for $A_{n,k}(y,y,q,B)$ is given. In addition to generalizing the q-hit numbers of Garsia and Remmel [5], the $A_{n,k}(y,y,q,B)$ also generalize the cycle-counting hit numbers in the model of Chung and Graham [2].

In Section 2 we prove symmetry and unimodality of $A_{n,k}(a,b,q,B)$ for $a,b \in \mathbb{N}$. We then apply this theorem to prove a symmetry and unimodality property of the *cycle-counting q-Eulerian numbers* introduced in [1]. In Section 3, we prove symmetry of the polynomial $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$ for any regular Ferrers board $B = B(h_1,d_1;\ldots;h_t,d_t)$. Finally in Section 4, we prove unimodality of $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$ for a certain class of regular Ferrers boards.

2 Symmetry and Unimodality of $A_{n,k}(a, b, q, B)$

If $B = B(h_1, d_1; \ldots; h_t, d_t) \subseteq SQ_n$ is a Ferrers board, let us denote by $B - h_p - d_p$ the Ferrers board $B(h_1, d_1; \ldots; h_p - 1, d_p - 1; \ldots h_t, d_t) \subseteq SQ_{n-1}$, obtained from B by decreasing the pth step by 1. We will call the number of squares in the board B the $Area(B)$.

Suppose

$$f(q) = \sum_{i=M}^{N} a_i q^i,$$

is a polynomial in q with $a_M, a_N \neq 0$. We call $M + N$ the *virtual degree* of f. We will say the polynomial $f(q)$ is $zsu(d)$ if either

1. $f(q)$ is identically zero, or

2. $f(q)$ is in $\mathbb{N}[q]$, symmetric, and unimodal with virtual degree d.

Note that for $s \in \mathbb{N}$, q^s is $zsu(2s)$ and $[s]$ is $zsu(s-1)$. We have the following lemmas. The proof of Lemma 2.1 is trivial, and a proof of Lemma 2.2 can be found in [11].

Lemma 2.1. *If f and g are polynomials which are both $zsu(d)$, then $f + g$ is $zsu(d)$.*

Lemma 2.2. *If f is $zsu(d)$ and g is $zsu(e)$, then fg is $zsu(d + e)$.*

Combining Lemmas 2.1 and 2.2 with (2) and (3), we can easily prove the following.

Lemma 2.3. *Let $a, b \in \mathbb{N}$. For any regular Ferrers board $B \subseteq SQ_n$, $A_{n,0}(a, b, q, B)$ is $zsu(Area(B) + n(b-1) - \binom{n+1}{2})$.*

Lemma 2.4. *Let $B = B(h_1, d_1; \ldots; h_t, d_t) \subseteq SQ_n$ be a regular Ferrers board, $B - h_t - d_t \subseteq SQ_{n-1}$ as described earlier. Then*

$$A_{n,k}(x, y, q, B) = [k + y + d_t - 1]A_{n-1,k}(x, y, q, B - h_t - d_t) +$$

$$q^{k+y+d_t-2}[n + x - y - d_t - k + 1]A_{n-1,k-1}(x, y, q, B - h_t - d_t)$$

for any $1 \leq k \leq n$.

Proof. Let $p = t$ in Lemma 5.7 of [9]. □

The following is now a simple corollary of the above lemmas.

Corollary 2.5. *Let $B = B(h_1, d_1; \ldots; h_t, d_t) \subseteq SQ_n$ be a regular Ferrers board, $a, b \in \mathbb{N}$. If $n+a+1 \geq b+d_t+k$, then $A_{n,k}(a, b, q, B)$ is $zsu(Area(B)+ n(b + k - 1) + k(a - 1) - \binom{n+1}{2})$ for $0 \leq k \leq n$.*

Proof. The proof is by induction on $Area(B)$. When $Area(B) = 1$ the only regular Ferrers board is the 1×1 square SQ_1. A quick calculation from the definition shows that $A_{1,0}(x, y, q, SQ_1) = [y]$, $A_{1,1}(x, y, q, SQ_1) = q^y[x - y]$, and $A_{1,k}(x, y, q, SQ_1) = 0$ for $k > 1$. Thus $A_{1,0}(a, b, q, B) = [b]$ and $A_{1,1}(a, b, q, B) = q^b[a - b]$ are $zsu(b - 1)$ and $zsu(a + b - 1)$ respectively, and the result holds for the case $Area(B) = 1$.

Now assume the result holds for all regular Ferrers boards with $Area < A$, and let B be such a board with $Area(B) = A$. We know by Lemma 2.3 that the result holds for $A_{n,0}(a, b, q, B)$, so assume $k > 0$. Then by Lemma 2.4 with $x = a$ and $y = b$, we have that

$$A_{n,k}(a, b, q, B) = [k + b + d_t - 1]A_{n-1,k}(a, b, q, B - h_t - d_t)+ \\ q^{k+b+d_t-2}[n + a - b - d_t - k + 1]A_{n-1,k-1}(a, b, q, B - h_t - d_t). \qquad (4)$$

Now we know that $[k + b + d_t - 1]$ is $zsu(k + b + d_t - 2)$, and by the induction hypothesis, $A_{n-1,k}(a, b, q, B - h_t - d_t)$ is $zsu(Area(B - h_t - d_t) + (n - 1)(b + k - 1) + k(a - 1) - \binom{n}{2})$. Note here that $Area(B - h_t - d_t) = Area(B) - n - d_t + 1$. Then by Lemma 2.2, the first term on the right side of (4) is $zsu(Area(B) + n(b + k - 1) + k(a - 1) - \binom{n+1}{2})$.

For the second term on the right side of (4), we know that q^{k+b+d_t-2} is $zsu(2k+2b+2d_t-4)$, $[n+a-b-d_t-k+1]$ is $zsu(n+a-b-d_t-k)$ (since we have assumed $n + a + 1 \geq b + d_t + k$), and by the induction hypothesis $A_{n-1,k-1}(a, b, q, B - h_t - d_t)$ is $zsu(Area(B - h_t - d_t) + (n - 1)(b + k - 2) + (k - 1)(a - 1) - \binom{n}{2})$. Finally, applying Lemma 2.2 one last time we get that the second term on the right side of (4) is $zsu(Area(B)+n(b+k-1)+k(a-1) - \binom{n+1}{2})$, and Lemma 2.1 gives us the result for $A_{n,k}(a, b, q, B)$ as well. □

A cycle-counting version of the Eulerian numbers is given in [1], defined by the equation

$$\tilde{E}_{n,k}(y, q) = \sum_{\sigma \in S_n,\ des(\sigma)=k-1} [y]^{\ell rmin(\sigma)} q^{(n-\ell rmin(\sigma))(y-1)+maj(\sigma)}. \qquad (5)$$

Here $\ell rmin(\sigma)$ denotes the number of *left-to-right minima* of the permutation σ, computed by the following algorithm. For $\sigma = \sigma_1\sigma_2 \cdots \sigma_n \in S_n$, if $\sigma_{j_1} = 1$ then let y_1 be the cycle $(\sigma_1 \cdots \sigma_{j_1})$. If α is the smallest integer not contained in y_1, and $\sigma_{j_2} = \alpha$, let y_2 be the cycle $(\sigma_{j_1+1} \cdots \sigma_{j_2})$, etc. If the result of the above procedure is the product of cycles $y_1y_2 \cdots y_p$, we will let $p = \ell rmin(\sigma)$.

It was proven in [1] that

$$\tilde{E}_{n,k}(y, q) = A_{n,k-1}(y, y, q, \mathbb{T}_n), \qquad (6)$$

where $\mathbb{T}_n = B(1, 2, \ldots, n)$ denotes the triangular Ferrers board. In light of (6) and Corollary 2.5, the following can be easily proven.

Corollary 2.6. *For $m \in \mathbb{N}$, the polynomial*

$$\sum_{\sigma \in S_n,\ des(\sigma)=k-1} [m]^{\ell rmin(\sigma)} q^{(n-\ell rmin(\sigma))(m-1)+maj(\sigma)}$$

is symmetric and unimodal with virtual degree $n(m+k-2)+(k-1)(m-1)$.

3 Symmetry of $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$

In this section we prove a more general symmetry result for all regular Ferrers boards, namely the symmetry of the polynomial

$$\frac{A_{n,k}(a,b,q,B)}{\prod_{i=1}^{t}[d_i]!},$$

where $B = B(h_1, d_1; \ldots ; h_t, d_t)$. Throughout the rest of the paper we will use the notation H_i for the partial sum $h_1 + \cdots + h_i$, and D_i for $d_1 + \cdots + d_i$. We have the following lemmas.

Lemma 3.1. *Let $B = B(b_1, \ldots, b_n) = B(h_1, d_1; \ldots ; h_t, d_t)$ be a regular Ferrers board, $j \in \mathbb{N}$. Then*

$$\frac{\prod_{i=1}^{n}[j + b_i - i + y]}{\prod_{i=1}^{t}[d_i]!} = \prod_{i=1}^{t}\begin{bmatrix} j + H_i - D_{i-1} + y - 1 \\ d_i \end{bmatrix}.$$

Proof. We see that

$$\prod_{i=1}^{n}[j + b_i - i + y] =$$

$$\prod_{i=1}^{t}[j+H_i-D_{i-1}+y-1][(j+H_i-D_{i-1}+y-1)-1]\cdots[(j+H_i-D_{i-1}+y-1)-d_i+1].$$

Thus

$$\frac{\prod_{i=1}^{n}[j + b_i - i + y]}{\prod_{i=1}^{t}[d_i]!} =$$

$$\prod_{i=1}^{t}\frac{[j + H_i - D_{i-1} + y - 1]\cdots[(j + H_i - D_{i-1} + y - 1) - d_i + 1]}{[d_i]!},$$

which is

$$\prod_{i=1}^{t}\begin{bmatrix} j + H_i - D_{i-1} + y - 1 \\ d_i \end{bmatrix}$$

by definition. □

Lemma 3.2. *Let $B = B(b_1, \ldots, b_n) = B(h_1, d_1; \ldots ; h_t, d_t) \subseteq SQ_n$ be a regular Ferrers board. Then $A_{n,k}(x, y, q, B)/\prod_{i=1}^{t}[d_i]! =$*

$$\sum_{j=0}^{k}\begin{bmatrix} n + x \\ k - j \end{bmatrix}\begin{bmatrix} x + j - 1 \\ j \end{bmatrix}(-1)^{k-j}q^{\binom{k-j}{2}}\prod_{i=1}^{t}\begin{bmatrix} j + H_i - D_{i-1} + y - 1 \\ d_i \end{bmatrix}.$$

Proof. By Lemma 5.1 of [9], we have

$$A_{n,k}(x,y,q,B)=\sum_{j=0}^{k}\begin{bmatrix}n+x\\k-j\end{bmatrix}\begin{bmatrix}x+j-1\\j\end{bmatrix}(-1)^{k-j}q^{\binom{k-j}{2}}\prod_{i=1}^{n}[j+b_i-i+y].$$

The lemma now follows trivially from Lemma 3.1. □

We can now prove the following.

Theorem 3.3. *Let $B=B(h_1,d_1;\dots;h_t,d_t)$ be a regular Ferrers board (so $H_i\geq D_i$ for $1\leq i\leq t$). Let $a,b\in\mathbb{N}$ with $a\geq b\geq 1$, and set*

$$L_k^{a,b}(B)=Area(B)+n(b-1)+k(n+a-1)-\sum_{i=1}^{t}d_iD_i.$$

Then $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$ is either zero or symmetric with virtual degree $L_k^{a,b}(B)$.

Proof. By Lemma 3.2, $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!=$

$$\sum_{j=0}^{k}\begin{bmatrix}n+a\\k-j\end{bmatrix}\begin{bmatrix}a+j-1\\j\end{bmatrix}(-1)^{k-j}q^{\binom{k-j}{2}}\prod_{i=1}^{t}\begin{bmatrix}j+H_i-D_{i-1}+b-1\\d_i\end{bmatrix},$$

which is a polynomial in q (the first two q-binomial coefficients in each summand are clearly polynomials, and the third is since $H_i\geq D_i\geq D_{i-1}$ and $b\geq 1$). Using the fact that $\begin{bmatrix}r\\s\end{bmatrix}$ is $zsu(s(r-s))$ (see [8, 12] for a proof) and Lemma 2.2, each term on the right side above has virtual degree $(k-j)(n+a-k+j)+j(a-1)+(k-j)(k-j-1)+\sum_{i=1}^{t}d_i(j+H_i-D_i+b-1)$, which is exactly $L_k^{a,b}(B)$. Since the sign alternates, we can only conclude that $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$ is symmetric with virtual degree $L_k^{a,b}(B)$. □

4 Unimodality of $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$

In this section we give some sufficient conditions on the regular Ferrers board B for the polynomial of the previous section to also be unimodal. Let us first define some more notation.

Suppose we have integers $h_1,\dots,h_t$, $d_1,\dots,d_t$, and $e_1,\dots,e_t$ with $d_i\in\mathbb{P}$, $h_i\in\mathbb{N}$, and $0\leq e_i\leq d_i$. We will denote the vector $(e_1,e_2,\dots,e_t)$ by $\vec{e}$. We will continue to denote the partial sum $h_1+\cdots+h_i$ by H_i, $d_1+\cdots+d_i$ by D_i, and we will also let $E_i=e_1+\cdots+e_i$. We make the convention that $H_0=D_0=E_0=0$. For fixed $h_1,\dots,h_t$ and $d_1,\dots,d_t$ we can define

$$P(\vec{e},x,y)=\prod_{i=1}^{t}\begin{bmatrix}H_i-D_{i-1}+E_{i-1}+y-1\\d_i-e_i\end{bmatrix}\begin{bmatrix}D_i+D_{i-1}-H_i-E_{i-1}+x-y\\e_i\end{bmatrix}$$

and prove the following lemmas.

Lemma 4.1. *Let $B = B(h_1, d_1; \dots h_{t-1}, d_{t-1}; h_t, d_t) \subseteq SQ_n$ be a regular Ferrers board, $B' = B(h_1, d_1; \dots; h_{t-1}, d_{t-1}) \subseteq SQ_{H_{t-1}}$. Then*

$$A_{n,k}(x, y, q, B) = [d_t]! \sum_{s=k-d_t}^{k} A_{H_{t-1},s}(x, y, q, B') \begin{bmatrix} y + d_t + s - 1 \\ d_t - k + s \end{bmatrix}$$

$$\times \begin{bmatrix} n - y - d_t + x - s \\ k - s \end{bmatrix} q^{(k-s)(y+k-1)}.$$

Proof. Let $p = t$ in Corollary 5.10 of [9] and note that because B is a regular Ferrers board, $H_t = D_t = n$. □

Lemma 4.2. *Let $B = B(h_1, d_1; \dots; h_t, d_t)$ be a regular Ferrers board. Then*

$$A_{n,k}(x, y, q, B) =$$

$$\prod_{i=1}^{t} [d_i]! \sum_{e_1+\dots+e_t=k,\ 0 \le e_i \le d_i} P(\vec{e}, x, y) \prod_{i=1}^{t} q^{e_i(H_i - D_i + E_i + y - 1)}. \tag{7}$$

Proof. By induction on t. When $t = 1$ we have that $d_1 = n$, and Lemma 4.1 gives us

$$A_{n,k}(x, y, q, B) = [d_1]! \sum_{s=k-n}^{k} A_{0,s}(x, y, q, \emptyset) \begin{bmatrix} y + n + s - 1 \\ d_1 - k + s \end{bmatrix} \tag{8}$$

$$\times \begin{bmatrix} n - y - n + x - s \\ k - s \end{bmatrix} \times q^{(k-s)(y+k-1)}.$$

In this case we have that $H_1 = D_1 = d_1 = n$ and $D_0 = H_0 = 0$, so we get that the $s = 0$ term in (8) is equal to

$$[d_1]! \begin{bmatrix} H_1 - D_0 + y - 1 \\ d_1 - k \end{bmatrix} \begin{bmatrix} D_1 + D_0 - H_1 + x - y \\ k \end{bmatrix} \times q^{k(H_1 - D_1 + k + y - 1)}. \tag{9}$$

Note that by definition

$$A_{0,s}(x, y, q, \emptyset) = \delta_{s,0},$$

so the only nonzero summand in (8) occurs when $s = 0$ and hence (9) is actually equal to (8). Finally if we recall that $E_1 = e_1$ and $E_0 = 0$, we can rewrite (9) as

$$[d_1]! \sum_{e_1=k,\ 0 \le e_1 \le d_1} \begin{bmatrix} H_1 - D_0 + E_0 + y - 1 \\ d_1 - e_1 \end{bmatrix}$$

$$\times \begin{bmatrix} D_1 + D_0 - H_1 - E_0 + x - y \\ e_1 \end{bmatrix} \times q^{e_1(H_1 - D_1 + E_1 + y - 1)},$$

which is exactly of the form of (7).

For $t > 1$, Lemma 4.1 gives that

$$A_{n,k}(x,y,q,B) = [d_t]! \sum_{E_{t-1}=E_t-d_t}^{E_t} A_{H_{t-1},E_{t-1}}(x,y,q,B') \begin{bmatrix} y+d_t+E_{t-1}-1 \\ d_t-e_t \end{bmatrix}$$

$$\times \begin{bmatrix} n-y-d_t+x-E_{t-1} \\ e_t \end{bmatrix} \times q^{e_t(y+E_t-1)}. \tag{10}$$

Here we are letting $E_{t-1} = s$ and defining $e_t = k-s$ and $E_t = E_{t-1}+e_t = k$. Since B is regular $H_t = D_t = n$, so $H_t - D_{t-1} = D_t - D_{t-1} = d_t$ and (10) can be rewritten as

$$A_{n,k}(x,y,q,B) = [d_t]! \sum_{e_t=0}^{d_t} A_{H_{t-1},E_{t-1}}(x,y,q,B') \begin{bmatrix} H_t-D_{t-1}+E_{t-1}+y-1 \\ d_t-e_t \end{bmatrix}$$

$$\times \begin{bmatrix} D_t+D_{t-1}-H_t-E_{t-1}+x-y \\ e_t \end{bmatrix} \times q^{e_t(H_t-D_t+E_t+y-1)}.$$

By the inductive hypothesis, the above is equal to

$$[d_t]! \sum_{e_t=0}^{d_t} \left\{ \prod_{i=1}^{t-1} [d_i]! \sum_{e_1+\cdots+e_{t-1}=E_{t-1},\ 0\le e_i \le d_i} \prod_{i=1}^{t-1} \begin{bmatrix} H_i-D_{i-1}+E_{i-1}+y-1 \\ d_i-e_i \end{bmatrix} \right.$$

$$\left. \times \begin{bmatrix} D_i+D_{i-1}-H_i-E_{i-1}+x-y \\ e_i \end{bmatrix} q^{e_i(H_i-D_i+E_i+y-1)} \right\}$$

$$\times \begin{bmatrix} H_t-D_{t-1}+E_{t-1}+y-1 \\ d_t-e_t \end{bmatrix} \begin{bmatrix} D_t+D_{t-1}-H_t-E_{t-1}+x-y \\ e_t \end{bmatrix} q^{e_t(H_t-D_t+E_t+y-1)}$$

which is

$$\prod_{i=1}^{t} [d_i]! \sum_{e_1+\cdots+e_t=k,\ 0\le e_i\le d_t} P(\vec{e},x,y) \prod_{i+1}^{t} q^{e_i(H_i-D_i+E_i+y-1)}$$

as desired. □

Lemma 4.3. *Let $B = B(h_1,d_1;\ldots;h_t,d_t)$ be a regular Ferrers board, $a,b \in \mathbb{N}$ with $a \ge b \ge 1$. Let e_i, d_i, h_i, E_i, D_i, and H_i be as in the definition of $P(\vec{e},x,y)$. Assume that B is such that $d_{i-1}+d_i \ge h_i$ for $1 \le i \le t$ (where $d_0 := 0$). If any of the numerators of the q-binomial coefficients in*

$$P(\vec{e},a,b) = \prod_{i=1}^{t} \begin{bmatrix} H_i-D_{i-1}+E_{i-1}+b-1 \\ d_i-e_i \end{bmatrix} \begin{bmatrix} D_i+D_{i-1}-H_i-E_{i-1}+a-b \\ e_i \end{bmatrix}$$

are negative, then $P(\vec{e},a,b) = 0$.

Proof. First note that $H_i - D_{i-1} + E_{i-1} + b - 1 \geq 0$ for $1 \leq i \leq t$, since $H_i \geq D_i \geq D_{i-1}$ and $b \geq 1$, so none of the numerators in the first q-binomial coefficient of the product are ever negative.

Now suppose that $D_k + D_{k-1} - H_k - E_{k-1} + a - b < 0$ for some k with $0 \leq k \leq t$. Note $D_1 + D_0 - H_1 - E_0 + a - b = d_1 - h_1 + a - b$, and since we assumed $d_{i-1} + d_i \geq h_i$ (and in particular $d_1 \geq h_1$) and $a \geq b$, we have that $d_1 - h_1 + a - b \geq 0$. Thus we see that such a k must be greater than 2.

Now choose j such that $D_i + D_{i-1} - H_i - E_{i-1} + a - b \geq 0$ for $1 \leq i < j$, but $D_j + D_{j-1} - H_j - E_{j-1} + a - b < 0$ (such a j exists because of the remarks in the previous paragraph). Then $D_j + D_{j-1} - H_j - E_{j-1} + a - b < 0$ implies $D_j + D_{j-1} - H_j - E_{j-2} + a - b < e_{j-1}$, which is equivalent to $d_j + d_{j-1} - h_j + D_{j-1} + D_{j-2} - H_{j-1} - E_{j-2} + a - b < e_{j-1}$, which implies $D_{j-1} + D_{j-2} - H_{j-1} - E_{j-2} + a - b < e_{j-1}$ (since $d_j + d_{j-1} \geq h_j$). Hence

$$\begin{bmatrix} D_{j-1} + D_{j-2} - H_{j-1} - E_{j-2} + a - b \\ e_{j-1} \end{bmatrix} = 0$$

since the numerator is non-negative by definition of j but less than the denominator, so the product $P(\vec{e}, a, b) = 0$ as well. □

Theorem 4.4. *Let $B = B(h_1, d_1; \ldots ; h_t, d_t)$ be a regular Ferrers board such that $d_{i-1} + d_i \geq h_i$ for $1 \leq i \leq t$. Let $a, b \in \mathbb{N}$ with $a \geq b \geq 1$, and set*

$$L_k^{a,b}(B) = Area(B) + n(b-1) + k(n+a-1) - \sum_{i=1}^{t} d_i D_i$$

as before. Then $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$ is $zsu(L_k^{a,b}(B))$.

Proof. We apply Lemma 4.2, which says that

$$\frac{A_{n,k}(a,b,q,B)}{\prod_{i=1}^{t}[d_i]!} = \sum_{e_1+\cdots+e_t=k,\ 0\leq e_i \leq d_i} P(\vec{e}, a, b) \prod_{i=1}^{t} q^{e_i(H_i - D_i + E_i + b - 1)},$$

and all of the terms on the right hand side above are in $\mathbb{N}[q]$ by Lemma 4.3. Each term is $zsu(\sum_{i=1}^{t}\{(d_i - e_i)(H_i - D_i + E_i + b - 1) + e_i(D_i + D_{i-1} - H_i - E_i + a - b) + 2e_i(H_i - D_i + E_i + b - 1)\})$, which a simple calculation shows is the same $zsu(L_k^{a,b}(B))$. Thus by Lemma 2.1, $A_{n,k}(a,b,q,B)/\prod_{i=1}^{t}[d_i]!$ is $zsu(L_k^{a,b}(B))$ as well. □

References

[1] F. Butler. Rook theory and cycle-counting permutation statistics. *Adv. in Apl. Math.*, 33:655–675, 2004.

[2] F. Chung and R. Graham. On the cover polynomial of a digraph. *J. Combin. Theory, Ser. B*, 65:273–290, 1995.

[3] M. Dworkin. Factorization of the cover polynomial. *J. Combin. Theory, Ser. B*, 71:17–53, 1997.

[4] R. Ehrenborg, J. Haglund, and M. Readdy. Colored juggling patterns and weighted rook placements. Unpublished manuscript.

[5] A. Garsia and J. Remmel. q-Counting rook configurations and a formula of Frobenius. *J. Combin. Theory, Ser. A*, 41:246–275, 1986.

[6] I. Gessel. Generalized rook polynomials and orthogonal polynomials. In Dennis Stanton, editor, *q-Series and Partitions*, IMA Volumes in Mathematics and Its Applications, pages 159–167. Springer Verlag, 1989.

[7] J. Goldman, J. Joichi, and D. White. Rook theory I: Rook equivalence of Ferrers boards. *Proc. Amer. Math. Soc.*, 52:485–492, 1975.

[8] F. Goodman and K. O'Hara. On the Gaussian polynomials. In Dennis Stanton, editor, *q-Series and Partitions*, IMA Volumes in Mathematics and Its Applications, pages 57–66. Springer Verlag, 1989.

[9] J. Haglund. Rook theory and hypergeometric series. *Adv. in Appl. Math.*, 17:408–459, 1996.

[10] R. Stanely. *Enumerative Combinatorics Voulme I.* Cambridge University Press, 1997.

[11] R. Stanley. Unimodal and log-concave sequences in algebra, combinatorics, and geometry. In *Graph Theory and Its Applications: East and West*, volume 576 of *Ann. New York Acad. Sci.*, pages 500–535. 1989.

[12] D. Zeilberger. A one-line high school algebra proof of the unimodality of the Gaussian polynomials $\begin{bmatrix} n \\ k \end{bmatrix}$ for $k < 20$. In Dennis Stanton, editor, *q-Series and Partitions*, IMA Volumes in Mathematics and Its Applications, pages 67–72. Springer Verlag, 1989.

DECIDING THE COHEN-MACAULAY PROPERTY FOR BIPARTITE GRAPHS

GIUSEPPA CARRA' FERRO
DANIELA FERRARELLO

Abstract. An algorithm is provided in order to decide whether a bipartite graph is Cohen-Macaulay. It works by appropriately deleting vertices from the given graph and by applying known properties on the obtained subgraphs.

1. Introduction

The Cohen-Macaulay property of a graph (CM for short) is worth investigating, since it comes from an algebraic concept and the combinatorial meaning is not so evident. In fact it is difficult to recognize a CM graph just by looking at it. So, it is interesting to find necessary and sufficient conditions for a graph in order to be Cohen Macaulay and it would be very useful to find a decision procedure.

Cohen-Macaulay graphs are investigated in several works, see for example [6], where one can find constructions of CM graphs and properties about bipartite CM graphs. The latter ones are characterized in [4]. It is also known that a chordal graph is CM if and only if it is unmixed (see [5]) and that the complement of a d-tree is CM (see [2]). Actually there is no decision procedure for CM graphs. In this paper we show an algorithm for checking Cohen-Macauly property of a bipartite graph. Such algorithm uses some results about CM graphs in [6] and it is based on the decision procedure for bipartite graphs and vertex covers in [1].

2. Cohen-Macaulay Graphs

Here we will introduce the concept of Cohen-Macaulay graph and all definitions and properties, that we will use as tools for studying such graphs.

DEFINITION 2.1. *The ascending chain condition, commonly abbreviated "A.C.C.," for a partially ordered set X requires that all increasing sequences in X become stationary.*

DEFINITION 2.2. *A ring is called* Noetherian *if it does not contain an infinite ascending chain of ideals.*

REMARK 2.1. *If R is Noetherian, it satisfies the ascending chain condition on ideals.*

PROPOSITION 2.1. *The following properties are equivalent.*

(1) *R satisfies the ascending chain condition on ideals.*
(2) *Every ideal of R is finitely generated.*
(3) *Every set of ideals contains a maximal element.*

Let M be a module over a ring R. We say that $x \in R$ is a *M-regular element* if it is not a zero-divisor on M.

DEFINITION 2.3. *A sequence $\mathbf{x} = x_1, \ldots, x_n$ of elements of R is called a* M-regular sequence *if*

- *(i) x_i is a $M/(x_1, \ldots, x_{i-1})$ M-regular element for $i = 1, \ldots, n$;*
- *(ii) $M/\mathbf{x}M \neq 0$.*

EXAMPLE 2.1. *The typical example of regular sequence is the sequence $x_1, \ldots, x_n$ of indeterminates in a polynomial ring $R = S[x_1, \ldots, x_n]$.*

Let R be a Noetherian ring and let M be a R-module. If $\mathbf{x} = x_1, \ldots, x_n$ is a M-sequence, then the sequence of ideals $(x_1) \subset (x_1, x_2) \subset \ldots \subset (x_1, \ldots, x_n)$ is strictly ascending. Therefore a M-sequence can be extended to a maximal sequence in the following way: a M-sequence $\mathbf{x}$ in an ideal I is *maximal* in I if $x_1, \ldots, x_{n+1}$ is not a M-sequence for any $x_{n+1} \in I$.

THEOREM 2.1. *(Rees)*
Let R be a Noetherian ring. Let M be a finite $R-$module and let I be an ideal, such that $IM \neq M$. Then all maximal M-sequences in I have the same length n, that is called grade *of I on M, and it is denoted by $grade(I, M)$.*

DEFINITION 2.4. *A ring R is called* local*, if it has a unique maximal ideal* $\mathfrak{m}$ *and it is denoted by* $(R, \mathfrak{m})$.

The notions of grade and Noetherian local ring imply the definition of depth of a local ring R.

DEFINITION 2.5. *Let* $(R, \mathfrak{m})$ *be a Noetherian local ring and let M be a finite R-module. Then the grade of* $(R, \mathfrak{m})$ *on M is called the* depth *of M and it is denoted by $depthM$.*

We introduce the notion of height of an ideal $\mathfrak{p}$ in R, in order to define the dimension of a commutative ring R. *height* $\mathfrak{p}$, is the supremum of the lengths t of strictly descending chains $\mathfrak{p} = \mathfrak{p}_0 \supset \mathfrak{p}_1 \supset \ldots \supset \mathfrak{p}_t$ of prime ideals.

DEFINITION 2.6. *Let* $(R, \mathfrak{m})$ *be a local ring. The* dimension *of R is the height of* $\mathfrak{m}$ *and it is denoted by $dimR$.*

In general $depthR \leq dimR$.
Finally we define a Cohen-Macaulay ring

DEFINITION 2.7. *Let R be a Noetherian local ring. A finite R-module $M \neq 0$ is a* Cohen-Macaulay *module if depth M=dim M. If R itself is a Cohen-Macaulay module, then it is called a* Cohen Macaulay *ring.*

DEFINITION 2.8. *A noetherian ring R is said to be a* Cohen-Macaulay *ring if $R_{\mathfrak{m}}$ is a* Cohen-Macaulay *ring for every maximal ideal* $\mathfrak{m}$ *of R.*

To every undirected graph G with the vertex set $V(G) = \{v_1, \ldots, v_n\}$ and the edge set $E(G) = \{e_1, \ldots, e_m\}$ it is possible to associate a monomial ideal $I(G)$, that is generated by all square free monomials $v_i v_j$, such that $\{v_i, v_j\} = e_h$ is an edge of G. Such an ideal is usually called the monomial *edge ideal*.

DEFINITION 2.9. *G is said* Cohen-Macaulay *(CM for short) with respect to the field K, if the quotient ring $K[v_1, \ldots, v_n]/I(G)$ is Cohen-Macaulay.*

DEFINITION 2.10. *A* vertex cover V' *of a graph* G *is a subset of vertices of* G*, such that at least one vertex of every edge of* G *is in* V'*. A vertex cover* V' *is said to be* minimal *if no subsets of* V' *is itself a vertex cover.*

Of course every graph has vertex covers (it is enough to take the whole set of vertices). By using the following proposition of Villarreal it is possible to find them by looking at the primary decomposition of the monomial edge ideal.

PROPOSITION 2.2. *(See* [6]*, chapter 6 proposition 1.16)*
Let $K[v] = K[v_1, ..., v_n]$ *be a polynomial ring over a field* K *and let* G *be an undirected graph. If* P *is the ideal of* $K[v]$ *generated by* $A = \{v_{i1}, \dots, v_{ir}\}$*, then* P *is a minimal prime over the edge ideal* $I(G)$ *if and only if* A *is a minimal vertex cover of* G*.*

As a corollary of the previous proposition we obtain a way to compute the height of an edge ideal.

COROLLARY 2.1. *(See* [6]*, chapter 6 corollary 1.18)*
If G *is a graph and* $I(G)$ *its monomial edge ideal, then the height of* $I(G)$ *is equal to the vertex covering number* $\alpha_0(G)$*, that is the smallest number of vertices in a minimal vertex cover.*

DEFINITION 2.11. *A graph is said* unmixed *if all minimal vertex covers have the same cardinality.*

REMARK 2.2. *A Cohen-Macaulay graph is unmixed. (See, for instance,* [4]*)*

Finally it is useful to introduce the definition of bipartite graph.

DEFINITION 2.12. *A graph* G *is* bipartite*, if its vertices can be divided in two sets, such that no edge connects vertices in the same set. Here we will call these two sets* partition sets*. Equivalently* G *is* bipartite *iff all cycles in* G *are even.*

2.1. **Construction of Cohen-Macaulay Graphs.** The main part of the results in this subsection can be found in [6], chapter 6 section 2.

The degree of a vertex v, $deg(v)$, is the number of edges incident in v and the set of neighbors of v, $N(v)$ is the set of vertices connected with v.
Of course $\mid deg(v) \mid = \mid N(v) \mid, \forall v \in V(G)$.

First construction

Let G be a graph on the vertex set $V = \{v_1, \dots, v_r, z, w\}$ with $deg(w) = 1$, $N(w) = \{z\}$, $deg(z) = k + 1$, $N(z) = \{w, v_1, \dots, v_k\}$.
Let G_1 be the graph obtained by deleting the vertices w and z in G, and let F_1 be the graph obtained by deleting the vertices $\{v_1, \dots, v_k\}$ in G_1 .[1]
Then the following propositions hold:

(1) If G is CM, then both G_1 and F_1 are CM
(2) If G_1 and F_1 are CM and $\{v_1, \dots, v_k\}$ form a part of a minimal vertex cover for G, then G is CM
(3) If G_1 is CM and $\{v_1, \dots, v_k\}$ is a minimal vertex cover for G_1, then G is CM
(4) Every bipartite CM graph has a vertex of degree 1.

Second construction

Let G be a graph on the vertex set $V = \{v_1, \dots, v_n, z\}$ with $deg(z) \geq 2$, $N(z) = \{v_1, \dots, v_k\}$, and $deg(v_i) \geq 2$ for all $i = 1, \dots, k$.
Let G_1 be the graph obtained by deleting z in G and let F_1 be the graph obtained from by deleting $v_1, \dots, v_k$ in G_1.
Let I be the edge ideal of G_1.
Then the following propositions hold:

(1') If G is CM, then F_1 is CM
(2') Suppose that $\{v_1, \dots, v_k\}$ do not form a part of a minimal vertex cover for G_1 and $height(I, v_1, \dots, v_k)$=$height(I) + 1$. If F_1 and G_1 are CM, then G is CM
(3') If G_1 is CM and $\{v_1, \dots, v_{k-1}\}$ is a minimal vertex cover for G_1, then G is CM

[1] When a vertex is removed from a graph, then all edges incident in the vertex are also removed.

(4') For every CM graph two vertices of degree 1 cannot have a vertex in common.

2.2. **Characterization of bipartite CM graphs.** The main results in this subsection can be found in [4].

DEFINITION 2.13. *A* simple *graph is an unweighted, undirected graph and with no self-loops.*

Let G be a simple finite bipartite graph and let us suppose G unmixed. Let us call W and W' the two bipartition sets.

REMARK 2.3. *If G is a bipartite graph without isolated vertices, then the partition sets are minimal vertex covers of G. In fact if W and W' are the two partition sets, then W (resp W') is a vertex cover, because every edge has a vertex in W (resp W'). Moreover W (resp.W') is minimal, because there are no edges connecting two vertices in W' (resp. W).*

So, if G is bipartite and unmixed, then W and W' have the same cardinality n.
Now $(W \backslash U) \cup N(U)$ is a vertex cover of G for every subset U of W. In fact every edge incident in a vertex of U is covered by a vertex in $N(U)$ and every vertex not incident in a vertex of U is covered by a vertex in $W \backslash U$.
So $|(W \backslash U) \cup N(U)| \geq |W|$ and then $|U| \leq |N(U)|$. By marriage theorem every vertex in W is connected with a vertex in W'. This means we can relabel the names of the vertices in the following way: $W = \{x_1, \ldots, x_n\}$ and $W' = \{y_1, \ldots, y_n\}$, such that **(a)** $\{x_i, y_i\}$ is an edge of G for all $1 \leq i \leq n$.

LEMMA 2.1. *(see* [4]*, lemma 3.3)*
*With the above notation, let us suppose that G is a simple bipartite Cohen-Macaulay graph. Then G satisfies condition **(a)** and, furthermore, it satisfies also the condition **(b)** if $\{x_i, y_j\}$ is an edge of G, then $i \leq j$.*

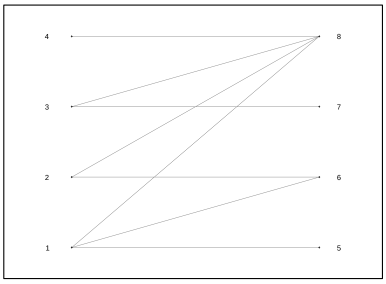
THEOREM 2.2. *(see* [4]*, theorem 3.4)*
Let G be a simple bipartite graph without loops on the vertex set $W \bigcup W'$, with $W = \{x_1, \ldots, x_n\}$, $W' = \{y_1, \ldots, y_n\}$ such that

(a): $\{x_i, y_i\}$ *is an edge for all* $1 \leqslant i \leqslant n$*;*
(b): *if* $\{x_i, y_j\}$ *is an edge then* $i \leqslant j$*;*

then G is CM iff

(c): *whenever* $\{x_i, y_j\}$ *and* $\{x_j, y_k\}$ *are edges, then* $\{x_i, y_k\}$ *is an edge.*

The previous theorem allows to know how a bipartite CM graph looks like. See the picture below. This fact is not trivial, because it is not clear just by looking only at the definition.



3. A Decision Algorithm for CM Graphs

Actually there is no algorithm for checking whether a graph is CM. Here we found a decision procedure for graphs, when the graph is bipartite. The main strategy is given by removing vertices from the initial graph and by checking some properties for the corresponding subgraphs. So if the graph is CM, then at the end of the algorithm we will find either a vertex or an edge, that are trivially Cohen-Macaulay. In order to write the algorithm we need the following results. (See [1]).

DEFINITION 3.1. *Let G=$(V(G), E(G))$ be a finite undirected graph. The binomial* extended edge ideal of *G is the ideal $I(G, E(G))$= ($e_h - v_i v_j$: $e_h = \{v_i, v_j\}$ is in $E(G)$).*
The ideal $I(G)_{E(G)} = I(G, E(G)) \cap K[e_1, \ldots, e_m]$ is the binomial edge ideal of G.

REMARK 3.1. *$I(G)_{E(G)}$ is the toric ideal of the incidence matrix $IM(G) = (a_{ih})_{i=1,\dots,n,h=1,\dots,m}$ of G defined by $a_{ih} = 1$ if $v_i \in e_h$ and $a_{ih} = 0$ if $v_i \notin e_h$ for every $v_i \in V(E)$ and $e_h \in E(G)$.*

DEFINITION 3.2. *Let G=$(V(G), E(G))$ be a finite undirected graph. The ideal $I(G, V(G))$=($v_i - \prod e_h$: v_i belongs to the edge e_h in G) is the binomial* extended vertex ideal of G. $I_{V(G)}$=$I(G, V(G)) \cap K[v_1, \dots, v_n]$. *is the binomial* vertex ideal of G.

$I_{V(G)}$ is the toric ideal of the transpose of the incidence matrix $IM(G)$ of G.

THEOREM 3.1. *Let G=$(V(G), E(G)$ be a finite undirected graph with $V(G)$= $\{v_1, \dots, v_n\}$ and $E(G)$=$\{e_1, \dots, e_m\}$. The odd cycle C=$(e_{i_1} = \{v_{i_1}, v_{i_2}\}, e_{i_2} = \{v_{i_2}, v_{i_3}\}, \dots, e_{i_{2q-2}} = \{v_{i_{2q-2}}, v_{i_{2q-1}}\}, e_{i_{2q-1}} = \{v_{i_{2q-1}}, v_{i_1}\})$ is in G iff the binomial $f_C = \prod_{k=1,\dots,q-1} e_{i_{2k}} v_{i_1}^2 - \prod_{k=1,\dots,q} e_{i_{2k-1}} \in I(G, E(G))$.*
Let σ be a lexicographic term ordering on the set of the power products in $\{e_1, \dots, e_m, v_1, \dots, v_n\}$ with $v_i > e_j$ for all i and j and $v_{i_{2q-1}} >_\sigma v_{i_{2q-2}} >_\sigma \dots >_\sigma v_{i_1}$. If C is minimal, then the binomial f_C is in the Gröbner basis of $I(G, E(G))$ with respect to σ.

THEOREM 3.2. *Let G=$(V(G), E(G))$ be a simple undirected connected graph without isolated vertices. If $I_{V(G)}$ contains an irreducible polynomial p of the form p=$\prod_{j\in J} v_j - \prod_{k\in K} v_k$, then G is bipartite and the partition sets are V'=$\{v_j : j \in J\}$ and V''=$\{v_k : k \in K\}$.*

3.1. **The Decision Algorithm.** Now we can show our decision procedure for bipartite CM graphs. First of all we have to check if our graph is bipartite. This can be done in the following way:

- We observe that a graph is CM if and only if every connected components is CM, and an isolated vertex is trivially CM; so we can apply the following step to the graph, that it is obtained by deleting the isolated vertices of G;

- Given a graph G, check if $\mid V(G) \mid$ is even. If $\mid V(G) \mid$ is odd, then G cannot be a bipartite unmixed graph and then it is not CM.
- Given a graph G, by using the package networks of Maple we can get the incidence matrix I_G of G;
- Given the matrix I_G, we can get the ideal $I(G, E(G))$ in the ring $K[e_1, \ldots, e_m, v_1, \ldots, v_n]$;
- Given the ideal $I(G, E(G))$, by using the package Groebner of Maple we can get a Gröbner basis of it with respect to a lexicographic term ordering σ with $v_i >_\sigma e_j$ for all i and j and then we can get the ideal $I(G)_{E(G)}$;
- By using the theorem 3.1 and the property of Gröbner bases about the decidability of the membership problem for polynomials, if G has no odd cycle, then G is bipartite;
- If G is not bipartite, we cannot conclude anything about the Cohen-Macaulay property.

Once we know that G is bipartite we can start the following algorithm

- We can get the transpose of the incidence matrix I_G^T of G;
- Given I_G^T we can get the ideal $I(G, V(G))$ in the ring $K[v_1, \ldots, v_n, e_1, \ldots, e_m]$;
- We can find the two partition sets by computing a Gröbner basis of the ideal $I(G, V(G))$ with respect to a lexicographic term ordering τ with $e_j >_\tau v_i$ for all i and j and then we can get the ideal $I(G)_{V(G)}$. The monomials appearing in the binomial of the basis represent the two sets, as in theorem 3.2;
- If the partition sets have different cardinality, then we can conclude that G is not CM. (In fact, the graph is not unmixed by remark as above);
- Given the bipartite graph G we can get the monomial edge ideal $I(G)$ and then we can find its minimal primes, that represent the minimal vertex covers of G, according to proposition 2.2
- Given the minimal vertex covers, we can decide if G is unmixed by looking at the cardinality of its minimal vertex covers;

- If G is not unmixed, then we can conclude that it is not CM;
- If G is unmixed, we can look at the degrees of its vertices; if there is not a vertex of degree one, then we can conclude that G is not CM, according to condition 4 in section 2.1;
- If G has a vertex of degree one, then we can consider the two partition sets being the sets W and W', with $W = \{x_1, \ldots, x_n\}$, $W' = \{y_1, \ldots, y_n\}$. W and W' have the same cardinality, and G is unmixed. We can order the vertices in such a way as the vertex of degree one is y_1 (it belongs just to the edge x_1, y_1 by condition **(b)** in theorem 2.1). $G_1 = G \backslash \{x_1, y_1\}$. If the set of neighbors of x_1 is the entire set $\{y_2, \ldots, y_n\}$, that is a minimal vertex cover for G_1, then by condition 3 G is CM if and only if G_1 is CM. So we can apply the algorithm to G_1;
- If the set of neighbors of x_1 is a proper subset of $\{y_2, \ldots, y_n\}$, [2], then by condition 2 in 2.1 G is CM if and only if both G_1 and F_1 are CM. So we start with F_1 (the smallest subgraph) as input of the algorithm. If F_1 is not CM, we can conclude that G is not CM. Else we put G_1 as input of the algorithm. If G_1 is not CM, then G is not CM;
- If in the previous steps we did not conclude that G is not CM, then we can conclude that G is CM.

Finiteness and correctness

Note that the algorithm finishes, because at each recursive step the input is a subgraph of the given one. So, in the worst case we apply the algorithm until the input is either an isolated vertex or an edge, that are trivially Cohen-Macaulay.

Moreover of course by deleting vertices from a bipartite graph we get a graph, that is still bipartite, since there are no new edges.

[2] It is clear that every set of x_i's or y_i's is at least a part of a minimal vertex cover for G

4. Example

Our algorithm is implemented in Maple 9.5 by using the packages *networks*, for graphs, and *Groebner*, for Gröbner bases, and it is called *isbipCM*. Here we will show an example.

EXAMPLE 4.1. *Let us consider the following graph G with eight vertices*

$> with(networks):$

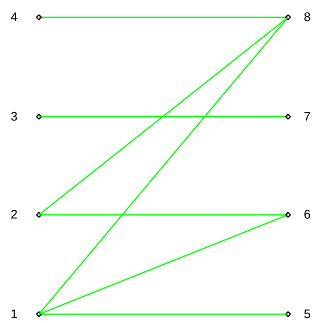
$> with(Groebner):$

$> G := void(8):$

$> addedge([\{1,5\},\{2,6\},\{3,7\},\{4,8\},\{1,6\},\{1,8\},\{2,8\}], G);$

$$e1, e2, e3, e4, e5, e6, e7$$

$> draw(Linear([1,2,3,4],[5,6,7,8]), G);$



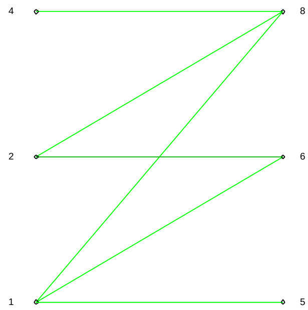
$> isbipCM(G);$

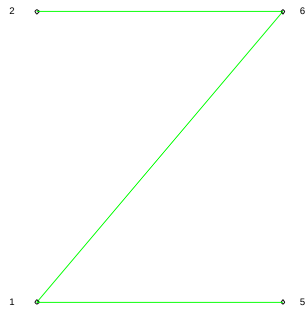
$$true$$

At each step the algorithm chooses a vertex of degree 1 and it works on the obtained subgraphs, according to the construction of CM graphs with the following choices:

$w = 3$, $z = 7$, $N(z)\backslash\{w\} = \{\}$, G_1 is the graph on the left of the following picture and F_1 is the same.

and then $w = 4$, $z = 8$, $N(z)\backslash\{w\} = \{1, 2\}$, G_1 is the graph on the right and since $\{1, 2\}$ is the entire set of bipartition of G_1 we do not need to construct F_1.





At the third step $w = 5$, $z = 1$, $N(z)\backslash\{w\} = \{6\}$, G_1 is just the edge $\{5, 6\}$ and we do not need to construct F_1.

An edge is trivially Cohen-Macaulay and the algorithm returns *true*.

References

[1] Carra' Ferro G. and Ferrarello D. *"Ideals and graphs" (2004) preprint*

[2] Ferrarello D. *"The complement of a d-tree is Cohen-Macaulay" (2004) preprint*

[3] Fröberg R. *"On Stanley-Reisner rings". Topics in Algebra vol. 26, Part 2 (1990)*

[4] Herzog J. and Hibi T. *"Distributive lattices, bipartite graphs and Alexander duality" (2003) preprint*

[5] Herzog J., Hibi T. and Zheng X. *"Cohen Macaulay chordal graphs" (2004) preprint*

[6] Villarreal R.H. *"Monomial algebras" (2001) Marcel Dekker Inc., Pure and Applied Mathematics 238*

Department of Mathematics and Computer Science, University of Catania, Viale Andrea Doria 6, 95125 Catania, Italy

E-mail address: carra@dmi.unict.it, ferrarello@dmi.unict.it

Enumeration of L-convex polyominoes. Bijection and area.

G. Castiglione * A. Frosini † E. Munarini ‡ A. Restivo * S. Rinaldi †

Abstract

We consider the class of L-convex polyominoes, i.e. (convex) polyominoes in which any two cells can be connected by a path of cells in the polyomino that switches direction between the vertical and the horizontal at most once - such paths with one change of direction look like the letter L in one of its four cyclic orientations, hence the name. In this paper we prove that the number f_n of L-convex polyominoes with perimeter $2(n+2)$ satisfies the linear recurrence relation $f_{n+2} = 4f_{n+1} - 2f_n$, by determining a coding of such polyominoes in terms of words of a regular language over four letters, thus giving a bijection with the class of 2-compositions (a simple generalization of the ordinary compositions) with sum equal to n. Moreover we study some combinatorial properties of 2-compositions. In the last section we determine the area generating function of L-convex polyominoes.

1 L-convex polyominoes: basic definitions

A *polyomino* is a finite union of elementary cells of the lattice $\mathbb{Z} \times \mathbb{Z}$, whose interior is connected (see Fig. 1 (a)). A polyomino is *h-convex* (resp. *v-convex*) if every row (resp. column) is connected. A polyomino is *hv-convex*, or simply *convex*, if it is both h-convex and v-convex (see Fig. 1 (b)). In a polyomino the *semi-perimeter* is half the length of the border, while the *area* is the number of its cells.

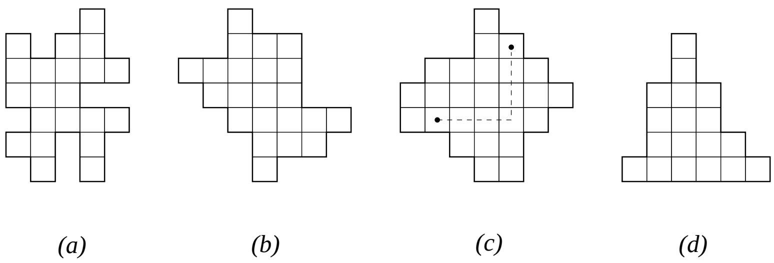


Figure 1: (a) a polyomino; (b) a convex polyomino; (c) a L-convex polyomino; (d) a stack polyomino.

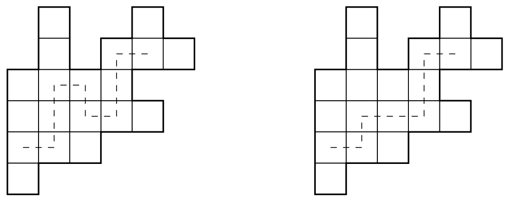
In a polyomino we will define a *path* as a self-avoiding sequence of unitary steps of four types: *north* $(0,1)$, *south* $(0,-1)$, *east* $(1,0)$, and *west* $(-1,0)$. A path connecting two distinct cells A and B, starts from the center of A, and ends in the center of B (see Fig. 2 (a)). We say that a path is *monotone* if it is constituted only of steps of at most two types (see Fig. 2 (b)). Given a path $w = u_1 \dots u_k$, 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*.

In [4] the authors observe that a polyomino P is convex if and only if every pair of cells is connected by a monotone path. Hence, taking into account the minimum number of changes of

*Università di Palermo, Dipartimento di Matematica e Applicazioni, Via Archirafi, 34, 90123, Palermo, Italy [giusi, restivo]@math.unipa.it.

†Università di Siena, Dipartimento di Matematica, Pian dei Mantellini, 44, 53100, Siena, Italy [frosini, rinaldi]@unisi.it.

‡Politecnico di Milano, Dipartimento di Matematica, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy munarini@mate.polimi.it.



(a) *(b)*

Figure 2: (a) a path between two cells in a polyomino; (b) a monotone path made only of north and east steps.

direction in their monotone paths, they give a classification of convex polyominoes. In particular, they call *k-convex* a convex polyomino such that every pair of cells can be connected by a monotone path with at most k changes of direction. For $k = 1$ we have the *L-convex polyominoes* i.e. the class of polyominoes such that each two cells can be connected by a path with at most one change of direction (see Fig. 1 (c)). In an L-convex polyomino the *horizontal basis* (resp. *vertical basis*) is the set of rows (resp. columns) having maximal length; by definition, both the horizontal and the vertical basis are rectangles.

In this paper we will also deal with the well-known class of *stack polyominoes* [8] [12, p. 76] [13] (see Fig. 1 (d)).

Let us denote by L_n the set of L-convex polyominoes having semi-perimeter $n+2$. In [3], using the ECO method, it was proved that the numbers $f_n = |L_n|$ satisfy the recurrence relation:

$$f_{n+2} = 4f_{n+1} - 2f_n \qquad (n \geq 1) \tag{1}$$

with $f_0 = 1$, $f_1 = 2$, $f_2 = 7$, giving the sequence 1,2,7,24,82,280,956,3264, ... (sequence A003480 in [11]).

The main results of the paper are the following:

1. we prove that the class of 2-compositions (a natural extension of the ordinary compositions) is enumerated by the sequence $(f_n)_{n\geq 0}$, and then we obtain several other properties of such a sequence;
2. we determine a bijection between 2-compositions and L-convex polyominoes, thus giving a combinatorial explanation that L-convex polyominoes satisfy the recurrence in (1);
3. finally we find the generating function for L-convex polyominoes according to the area.

2 2-compositions

A *composition* of a natural number n is an ordered partition of n, that is a k-tuple $(x_1, \ldots, x_k)$ of positive integers such that $x_1 + \cdots + x_k = n$ (see [5]).

We now extend the definition of composition to the 2-dimensional case. For any positive integer k, a *2-composition* of *length* k is a $2 \times k$ matrix whose entries are nonnegative integers, such that each column has at least one non null element; the sum of the elements in a 2-composition M is called the *sum* of M. Let U_n be the class of 2-compositions with sum equal to n and let $u_n = |U_n|$. For instance

$$U_1 = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \qquad U_2 = \left\{ \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right\}$$

and $u_1 = 2$, $u_2 = 7$. In particular U_0 contains only the empty 2-composition, with length 0, and $u_0 = 1$.

In what follows we will study 2-compositions. Some of their properties are easy to prove and for brevity they will only be stated.

Proposition 1 *The numbers* u_n *satisfy the recurrence* $u_{n+2} = 4u_{n+1} - 2u_n$ *for* $n \geq 1$ *, with the initial values* $u_0 = 1$ *,* $u_1 = 2$ *,* $u_2 = 7$ *.*

Proof. Let $n \geq 1$. The 2-compositions in U_{n+2} can be all obtained by performing the following operations on each 2-composition $M \in U_{n+1}$:

1. add a column $\begin{bmatrix}1\\0\end{bmatrix}$ on the left of M;
2. add a column $\begin{bmatrix}0\\1\end{bmatrix}$ on the left of M;
3. increase by one the first element on the first row of M;
4. increase by one the first element on the second row of M.

By performing the four operations on the 2-compositions of U_{n+1} we obtain a set of $4u_{n+1}$ elements of U_{n+2}. However, some 2-compositions are obtained twice, and they are precisely those containing no null elements in the first column, that is:

1. those whose first column is $\begin{bmatrix}1\\1\end{bmatrix}$;
2. those whose first column is $\begin{bmatrix}x+1\\y+1\end{bmatrix}$, with x , $y \geq 0$ and $(x, y) \neq (0, 0)$.

Since the number of elements in each class is clearly given by u_n it follows that $u_{n+2} = 4u_{n+1} - 2u_n$. Finally the initial values have been already determined in the initial examples. □

We have then the remarkable fact that the number of the L-convex polyominoes with semi-perimeter $n+2$ is equal to the number of the 2-compositions of n . In Section 3 we will determine a simple bijection between these two classes.

Let $u_{n,k}$ be the number of the elements of U_n having length k . The first terms of $u_{n,k}$ are presented in the table (a) of Fig. 3.

Proposition 2 *The numbers* $u_{n,k}$ *satisfy the recurrence relations*

$$u_{n+2,k+1} = 2u_{n+1,k+1} + 2u_{n+1,k} - u_{n,k+1} - u_{n,k}$$

$$u_{n+1,k+1} = u_{n,k+1} + 2u_{n,k} + u_{n-1,k} + \ldots + u_{0,k} \, .$$

In particular the infinite lower triangular matrix $[u_{n,k}]_{n,k\geq 0}$ *is a Riordan matrix with spectrum*

$$\left(1, \; \frac{2x - x^2}{(1-x)^2}\right) \, .$$

See [10] for the theory of Riordan matrices.

Proposition 3 *The numbers* u_n *have the Pisot property:*

$$u_{n+1}^2 - u_{n+2}\, u_n = 2^{n-1} \qquad (n \geq 1) \tag{2}$$

(which reassembles the well-known Cassini's identity for Fibonacci numbers [7]).

Since every 2-composition can be viewed as the concatenation of its columns, it follows that the set $\mathbb{U}$ of all 2-compositions is the language $\mathcal{A}^*$ on the infinite alphabet

$$\mathcal{A} = \{ a_{10}, \; a_{01}, \; a_{20}, \; a_{11}, \; a_{02}, \; \ldots \},$$

where the letter a_{ij} corresponds to the column $\begin{bmatrix}i\\j\end{bmatrix}$. Then the generating series of $\mathbb{U}$ is

$$u(x_{10}, x_{01}, x_{20}, x_{11}, x_{02}, \ldots) = \frac{1}{1 - x_{10} - x_{01} - x_{20} - x_{11} - x_{02} - \cdots}$$

n/k	0	1	2	3	4	5	6
0	1						
1	0	2					
2	0	3	4				
3	0	4	12	8			
4	0	5	25	36	16		
5	0	6	44	102	96	32	
6	0	7	70	231	344	240	64

n/k	0	1	2	3	4	5	6
0	1						
1	1	1					
2	2	3	2				
3	4	8	8	4			
4	8	20	26	20	8		
5	16	48	76	76	48	16	
6	32	112	208	252	208	112	32

Figure 3: (a) Table of the numbers $u_{n,k}$; (b) table of the numbers $v_{n,k}$.

(see [9]). In particular, for $x_{ij} = x^{i+j}y$ we obtain the generating series

$$u(x,y) = \sum_{n,k\geq 0} u_{n,k}\, x^n y^k = \frac{1}{1 - xh(x)\,y}, \tag{3}$$

where

$$h(x) = \sum_{n\geq 0}(n+2)x^n = \frac{2-x}{(1-x)^2}\,.$$

This also proves the second part of Proposition 2. From (3) it follows the identity $u(x,y) = 1 + xy\,h(x)\,u(x,y)$ which implies the recurrence

$$u_{n+1,k+1} = \sum_{i=0}^{n}(i+2)u_{n-i,k}\,.$$

Finally, expanding (3) we can obtain the following explicit formula (for n, $k \geq 1$):

$$u_{n,k} = \sum_{j=0}^{k}\binom{k}{j}\binom{n+k-j-1}{2k-1}(-1)^j\, 2^{k-j}\,.$$

For $y = 1$ in (3), we reobtain the generating series $u(x)$ for the numbers u_n. We also retrieve that $u(x)$ is the *quasi-inversion* of the series $xh(x)$ as pointed out in [2], in a completely different study. Moreover, since $u(x) = 1 + xh(x)$, it follows that

$$u_{n+1} = \sum_{k=0}^{n}(k+2)u_{n-k}\,.$$

Another interesting statistic can be obtained in the following way. For any $n \geq 1$, the *projection* (here the term is used in the sense of the discrete tomography [6]) of the 2-composition

$$M = \begin{bmatrix} x_1 & x_2 & \dots & \dots & x_k \\ y_1 & y_2 & \dots & \dots & y_k \end{bmatrix} \in U_n$$

is the 2-composition

$$\pi(M) = \begin{bmatrix} x_1 + x_2 + \ldots + x_k \\ y_1 + y_2 + \ldots + y_k \end{bmatrix}.$$

Clearly $\pi(M)$ is still an element of U_n. Moreover, for any $M \in U_n$ let us define

$$[M] = \{Q \in U_n : \pi(Q) = \pi(M)\}\,.$$

For instance:

$$\left[\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right] = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}.$$

One can easily observe that for any $n \geq 0$, there are $n+1$ distinct classes $[M]$ in U_n. For $0 \leq k \leq n$, let $v_{n,k}$ be the number of elements of U_n whose projection is equal to $\begin{bmatrix} n-k \\ k \end{bmatrix}$. The first terms of $v_{n,k}$ are presented in the table in Fig. 3 (b) (sequence A059576 in [11]).

Proposition 4 *The generating series for the numbers* $v_{n,k}$ *is*

$$v(x,y) = \sum_{n,k\geq 0} v_{n,k}\, x^n y^k = \frac{1-x-xy+x^2y}{1-2x-2xy+2x^2y}\,.$$

In particular

$$\sum_{n\geq 0} v_{n,0}\, x^n = \frac{1-x}{1-2x}\,, \qquad \sum_{n\geq k} v_{n,k}\, x^n = \frac{2^{k-1}(x-x^2)^k}{(1-2x)^{k+1}} \quad (k\geq 1)\,.$$

Moreover the numbers $v_{n,k}$ *satisfy the recurrence*

$$v_{n+2,k+1} = 2v_{n+1,k+1} + 2v_{n+1,k} - 2v_{n,k}$$

and (for $(n,k)\neq(0,0)$ *)*

$$v_{n,k} = \sum_{j=0}^{\min(k,n-k)} \binom{k}{j}\binom{n-j}{k}(-1)^j\, 2^{n-j-1}\,.$$

3 A bijection between U_n and L_n

In this section we will present a bijection between L-convex polyominoes with semi-perimeter equal to $n+2$ and 2-compositions with sum n. In order to do this, we need first to represent L-convex polyominoes in terms of *2-colored stacks*. A stack polyomino is *2-colored* when its rows are colored black or white and satisfy the following priority conditions:

1. if a row is white then all the other rows of the same length above it (if any) have the same color;
2. the rows having maximal length are colored white.

Starting from an L-convex polyomino, we give the black color to the rows placed below the horizontal basis, and then vertically translate them above the basis respecting condition 1. (see Fig. 4 (b)). We observe that by the definition of L-convexity, the obtained polyomino is actually a 2-colored stack polyomino. Conversely, to each 2-colored stack polyomino there corresponds a unique L-convex polyomino.

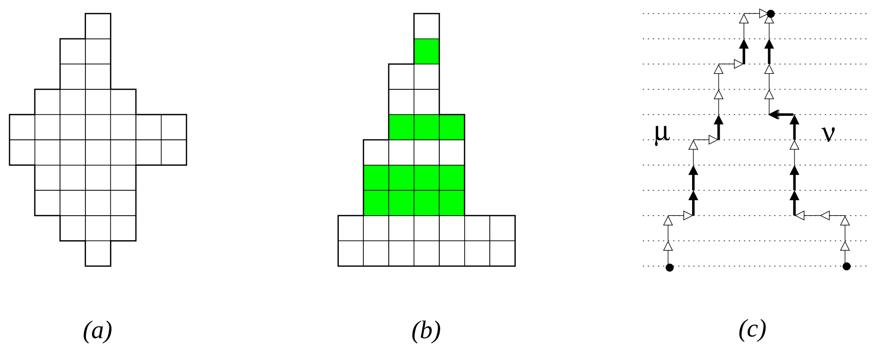


Figure 4: (a) an L-convex polyomino; (b) the corresponding 2-colored stack polyomino; (c) the paths μ and ν; for simplicity we represent the north, east and west steps by means of 2-colored arrows.

The boundary of a 2-colored stack polyomino is uniquely determined by two non-intersecting (except at the end points) lattice paths μ and ν (see Fig. 4 (c)):

1. μ runs from the leftmost point having minimal ordinate to the rightmost point having maximal ordinate in the polyomino, and uses 2-colored north and east unitary steps, that are black (resp. white) when it meets a black (resp. white) cell;

2. ν runs from the rightmost point having minimal ordinate to the rightmost point having maximal ordinate in the polyomino, and uses 2-colored north and west unitary steps, that are black (resp. white) when it meets a black (resp. white) cell.

By definition, both μ and ν start with a white north step, and μ ends with an east step.

Now we give a coding of 2-colored stacks in terms of words of a regular language over the alphabet $\{a, b, c, d\}$. The word representation of the polyomino is obtained by following the two paths, μ and ν, level by level from the bottom to the top of the polyomino. At each level one can meet:

1. a pair of north steps, one in μ, and the other in ν; in this case we write a (resp. d) if the steps are white (resp. black);
2. a sequence of east steps in μ, and, on the same horizontal line, a sequence of west steps in ν; in this case we write a b for each east step, and a c for each west step. By convention, we assume that, at the same level, we read east steps before west steps.

Using such a coding we have that any L-convex polyomino having semi-perimeter $n+2$ can be represented as a word in the alphabet $\{a, b, c, d\}$, having the same length. The language of all such words will be referred to as $\mathcal{K}$. For example, the word corresponding to the polyomino in Fig. 4 (a) is $aabccddabdcaabdab$. The number of rows (resp. columns) of the polyomino is given by the number of a plus the number of d (resp. the number of b plus the number of c) in the corresponding word of $\mathcal{K}$.

The words of $\mathcal{K}$ are characterized by the property that they begin with an a and end with a b, and contain neither the factor ad nor the factor cb. These simple observations lead us to state that $\mathcal{K}$ is a regular language, whose regular expression is:

$$a\left(a+b+c^{+}a+bd^{+}+c^{+}d^{+}\right)^{*}b\,. \tag{4}$$

Notice that using the same coding we can represent stack polyominoes in terms of words on the alphabet $\{a, b, c\}$, beginning with an a and ending with a b, and not containing the factor cb. A coding of L-convex polyominoes in terms of a regular language has been also considered in [1] in order to investigate about ordering properties of polyominoes.

Let $l_{i,j}$ be the number of L-convex polyominoes with $i+1$ rows and $j+1$ columns, as shown in the table of Fig.5. From (4), removing the first and the last letter, we can obtain the generating function for these numbers, as described in [9], after setting $a=d=x$ and $b=c=y$:

$$l(x,y)=\sum_{i,j\geq 0} l_{i,j}\,x^{i}y^{j}=\frac{1}{1-x-y-\dfrac{xy}{1-x}-\dfrac{xy}{1-y}-\dfrac{xy}{(1-x)(1-y)}}$$

that is

$$l(x,y)=\frac{(1-x)(1-y)}{1-2x-2y+x^{2}+y^{2}}\,. \tag{5}$$

Hence, it follows that the numbers $l_{i,j}$ satisfy the recurrence

$$l_{i+2,j+2}=2l_{i+1,j+2}+2l_{i+2,j+1}-l_{i,j+2}-l_{i+2,j}\,.$$

Letting $x=y$ in (5) we reobtain the generating function $f(x)=l(x,x)$ of $\mathcal{K}$ i.e. the generating function of L-convex polyominoes according to the semi-perimeter.

To conclude our bijection, we now give a representation of the words of $\mathcal{K}$ of length $n+2$ in terms of 2-compositions of U_n. First we observe that each word of $\mathcal{K}$ can be uniquely factorized into the factors:

$$c^{h}a, \quad bd^{j}, \quad c^{r}d^{s}, \qquad\qquad h,j\geq 0, \quad r,s\geq 1.$$

Let w be the word of $\mathcal{K}$ corresponding to a polyomino $P\in L_n$, from which we have removed the first and the last symbol; we use the following coding:

$$c^{h}a\to\begin{bmatrix}h+1\\0\end{bmatrix}, \qquad bd^{k}\to\begin{bmatrix}0\\k+1\end{bmatrix}\quad(h,k\geq 0), \qquad c^{r}d^{s}\to\begin{bmatrix}r\\s\end{bmatrix}\quad(r,s\geq 1),$$

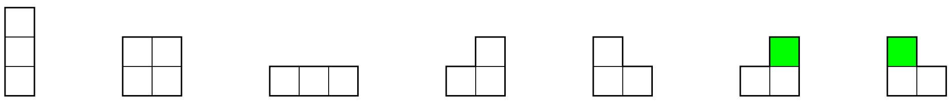
i/j	0	1	2	3	4	5	6
0	1	1	1	1	1	1	1
1	1	5	11	19	29	41	55
2	1	11	42	110	235	441	756
3	1	19	110	402	1135	2709	5740
4	1	29	235	1135	4070	11982	30618
5	1	41	441	2709	11982	42510	128534
6	1	55	756	5740	30618	128534	452900

Figure 5: Table of the numbers $l_{i,j}$ of L-convex polyominoes with $i+1$ rows and $j+1$ columns.

thus obtaining a 2-composition. For instance, the word $aabccddabdcaabdab$ (corresponding to the polyomino in Fig. 4 (a)) is translated into the 2-composition

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 0 & 2 & 0 & 0 & 2 & 0 \end{bmatrix}.$$

The reader can easily observe that using the above coding is also easy to pass from a 2-composition to a word of $\mathcal{K}$, and then to an L-convex polyomino, which completes the bijection. Fig. 6 shows the bijection between L_2 and U_2.



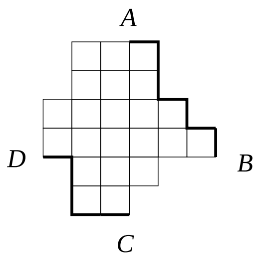
aaab *aabb* *abbb* *abab* *acab* *abdb* *acdb*

$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Figure 6: The bijection between L_2 and U_2.

Using the previously defined bijection, one natural question is how to interpret in terms of L-convex polyominoes the various properties determined in Section 2.

For instance, let us now consider the statistic in the table (b), Fig. 3. In terms of the word representation of a 2-composition M, the two entries of $\pi(M)$ are given by the number of a plus the number of c, and by the number of b plus the number of d, respectively.



acabcdaacdbb

$$\begin{pmatrix} 2 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Figure 7: A polyomino P, the corresponding 2-composition M, with projection $\pi(M) = \binom{6}{4}$; the two entries of $\pi(M)$ are given by the lengths of the paths AB and CD minus one, respectively.

It is also possible to read the previous statistic in terms of L-convex polyominoes. Let P be an L-convex polyomino, and $M(P)$ (briefly, M) the corresponding 2-composition. Let us consider the following discrete points on P (see Fig. 7):

i) A is rightmost point having maximal ordinate of the vertical basis;

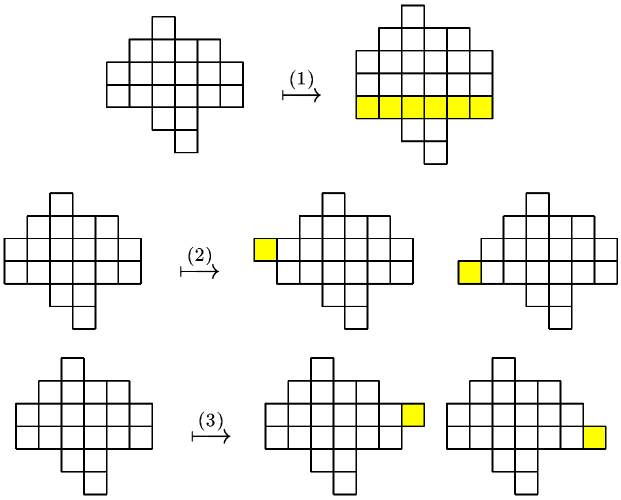


Figure 8: Generation of L-convex polyominoes.

ii) B is the rightmost point having minimal ordinate of the horizontal basis;

iii) C is the rightmost point having minimal ordinate of the vertical basis;

iv) D is the leftmost point having minimal ordinate of the horizontal basis.

Now, the element in the first row of $\pi(M)$ is given by the number of steps in the path connecting A to B, minus one. Analogously, the element in the second row of $\pi(M)$ is the number of steps in the path connecting C to D, minus one.

It would be also worth studying the interpretation of some parameters defined on a 2-composition (for instance the length of the composition), with respect to the correspondent L-convex polyomino through the bijection we have described above. On the other side, it would be interesting to investigate how some parameters defined on L-convex polyominoes (for example, the number of rows and columns, the area) can be interpreted on the corresponding 2-compositions.

4 Enumeration according to the area

In this section we determine the generating function of L-convex polyominoes according to the area, solving a problem posed in [3]. In order to do this, we first observe that every L-convex polyomino can be obtained with a sequence of the following operations starting from the polyomino formed by one cell:

1. add a new row of maximal length;
2. add a new cell on the left of a row of maximal length;
3. add a new cell on the right of a row of maximal length.

(See Fig. (8) for an example.) In this way, however, every polyomino with exactly one row of maximal length with a cell protruding on the left and a cell protruding on the right can be obtained two times applying operations 2. and 3. first in this order and then in the inverse order.

Let $L_{n,i,j}$ be the set of all L-convex polyominoes with semi-perimeter $n+2$ and $i+1$ rows of maximal length $j+1$ and let

$$a_{n,i,j}(q) = \sum_{P \in L_{n,i,j}} q^{a(\pi)}$$

where $a(P)$ is the area (i.e. the number of cells) of P. It follows that

$$a_{n+1,i+1,j}(q) = q^{j+1} a_{n,i,j}(q) \tag{6}$$

$$a_{n+2,0,j+2}(q) = \sum_{k=0}^{n+1} (k+1)(2q a_{n+1,k,j+1}(q) - q^2 a_{n,k,j}(q)). \tag{7}$$

Consider now the generating series

$$a_{i,j}(q;x) = \sum_{n \geq 0} a_{n,i,j}(q)\, x^n\,, \qquad a_j(q;x,y) = \sum_{n,i \geq 0} a_{n,i,j}(q)\, x^n y^i\,.$$

Let us consider L-convex polyomino whose maximal rows have length $j+1$. It may be reduced to an L-convex polyomino having a single row with that length. Otherwise, if it has several such rows, by removing one of them we can obtain a new L-convex polyomino of the same type. Fig. 9 depicts this decomposition.

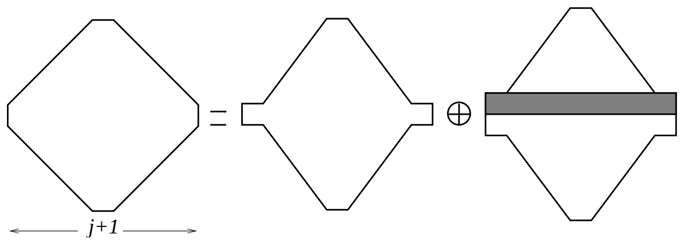


Figure 9: The decomposition of L-convex polyominoes whose maximal rows have length $j+1$

It follows that

$$a_j(q;x,y) = \frac{b_j(q;x)}{1 - q^{j+1}\, x\, y} \tag{8}$$

where $b_j(q;x) = a_{0,j}(q;x)$.

From equation (7) it follows

$$\mathcal{R}_x^2 b_{j+2}(q;x) = 2q\,[\,\mathcal{R}_x(\theta_y + 1)a_{j+1}(q;x,y)\,]_{y=1} - q^2\,[\,(\theta_y+1)a_j(q;x,y)\,]_{y=1} \tag{9}$$

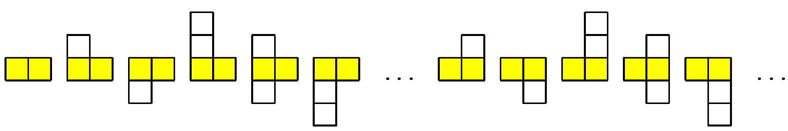
where $\mathcal{R}_x$ is the operator defined by $\mathcal{R}_x f(x) = (f(x) - f(0))/x$ and $\theta_y = y\frac{\mathrm{d}}{\mathrm{d}y}$. From equation (8) it follows that

$$[\theta_y a_j(q;x,y)]_{y=1} = \frac{q^{j+1} x b_j(q;x)}{(1-q^{j+1}x)^2}\,.$$

Since $b_{0,j+2}(q) = a_{0,0,j+2}(q) = 0$ and $b_{1,j+2}(q) = a_{1,0,j+2}(q) = 0$ for every j, equation (9) becomes

$$b_{j+2}(q;x) = \frac{2qx}{(1-q^{j+2}x)^2}\, b_{j+1}(q;x) - \frac{q^2x^2}{(1-q^{j+1}x)^2}\, b_j(q;x)\,. \tag{10}$$

Finally we need the initial values. For $j=0$ there is only the L-convex polyomino $\square$ and hence $b_0(q;x) = q$. For $j=1$ we have all the following polyominoes



and then

$$b_1(q;x) = 2\sum_{h,k \geq 0} q^{h+k+2} x^{h+k+1} - q^2 x = \frac{q^2 x(1 + 2qx - q^2x^2)}{(1-qx)^2}\,.$$

Recurrence (10), with the given initial values, completely determines the sequence $b_j(q;x)$ and easily allow to find that

$$b_j(q;x) = \frac{q^{j+1}x^j f_j(q;x)}{(1-qx)^2(1-q^2x)^2\cdots(1-q^jx)^2} \tag{11}$$

for suitable polynomials $f_j(q;x)$. Substituting the expression of $b_j(q;x)$ given by (11) in (10), it follows that the polynomials $f_j(q;x)$ satisfy the recurrence

$$f_{j+2}(q;x) = 2f_{j+1}(q;x) - (1-q^{j+2}x)^2 f_j(q;x) \tag{12}$$

with the initial conditions $f_0(q;x) = 1$ and $f_1(q;x) = 1 + 2qx - q^2x^2$. This completely defines the polynomials $f_k(q;x)$.

Then we have that

$$a_j(q;x,y) = \frac{q^{j+1}x^j f_j(q;x)}{(1-qx)^2(1-q^2x)^2\cdots(1-q^jx)^2(1-q^{j+1}xy)}$$

and finally

$$a(q;x,y,z) = \sum_{n,i,j\geq 0} a_{n,i,j}(q)x^n y^i z^j = \sum_{k\geq 0} \frac{q^{k+1}x^k z^k f_k(q;x)}{(1-qx)^2(1-q^2x)^2\cdots(1-q^kx)^2(1-q^{k+1}xy)} \tag{13}$$

From this series we can deduce several other generating series. First of all for, $q = 1$, equation (12) becomes

$$f_{j+2}(1;x) = 2f_{j+1}(1;x) - (1-x)^2 f_j(1;x)$$

with $f_0(1;x) = 1$ and $f_1(1;x) = 1 + 2x - x^2$. Then the generating series for these polynomials is

$$f(1;x,t) = \frac{1-(1-x)^2t}{1-2t+(1-x)^2t^2}. \tag{14}$$

Hence, for $q = 1$ the series (13) becomes

$$a(1;x,y,z) = \frac{1}{1-xy} f\left(1;x,\frac{xz}{(1-x)^2}\right) = \frac{(1-x)^2(1-xz)}{(1-xy)(1-2x+x^2-2xz+x^2z^2)}.$$

In particular, for $y = z = 1$, we obtain another time the generating series

$$a(1;x,1,1) = \frac{(1-x)^2}{1-4x+2x^2}$$

for the number of L-convex polyominoes according to semi-perimeter.

Finally, let a_n be the number of all L-convex polyominoes with area n. The first terms of this sequence are: 1, 1, 2, 6, 15, 35, 76, 156, 310, 590, 1098, 1984, 3515, 6094, 10398, 17434, 28837, 47038, 75820. This sequence is not in the *Encyclopedia of Integer Sequences* [11]. From (13) follows the main proposition:

Proposition 5 *The generating series of the sequence* a_n *is*

$$a(q) = \sum_{n\geq 0} a_n\, q^n = 1 + \sum_{k\geq 0} \frac{q^{k+1} f_k(q;1)}{(1-q)^2(1-q^2)^2\cdots(1-q^k)^2(1-q^{k+1})}.$$

This series is very similar to the generating series [13]

$$s(q) = \sum_{n\geq 0} s_n\, q^n = 1 + \sum_{k\geq 0} \frac{q^{k+1}}{(1-q)^2(1-q^2)^2\cdots(1-q^k)^2(1-q^{k+1})}.$$

for the numbers s_n that count stacks with area n (sequence A001523 in [11]). They only differ for the presence of the polynomials $f_k(q;1)$. It could be interesting a deeper investigation about the (combinatorial and formal) structure of such polynomials.

References

[1] F. Burderi, G. Castiglione, A. Restivo, *Highmann's theorem on Discrete Sets*, preprint.

[2] P. Cameron, *Some sequences of integers*, Discrete Math. **75** (1989), 89–102.

[3] G. Castiglione, A. Frosini, A. Restivo, S. Rinaldi, *Enumeration of L-convex polyominoes by rows and columns*, to appear in Theoret.Comput. Sci.

[4] G. Castiglione, A. Restivo, *Reconstruction of L-convex Polyominoes*, Electronic Notes in Discrete Mathematics **12**, Elsevier Science (2003).

[5] L. Comtet, *"Advanced Combinatorics"*, Reidel, Dordrecht-Holland, Boston, 1974.

[6] A. Kuba, G. Herman, *"Discrete Tomography: Foundations, Algorithms and Applications"*, Birkhauser, Boston, USA, (1999).

[7] R.L. Graham, D.E. Knuth, O. Patashnik, *"Concrete Mathematics"*, Addison-Wesley, Reading, MA, 1989.

[8] S. Rinaldi, D. G. Rogers, *How the odd terms in the Fibonacci sequence stack up*, Math. Gaz., to appear.

[9] A. Salomaa and M. Soittola, *"Automata-theoretic aspects of formal power series"*, Springer-Verlag, New York, 1978.

[10] L. W. Shapiro, S. Getu, W. J. Woan, L. C. Woodson, *The Riordan group*, Discrete Appl. Math. **34** (1991), 229–239.

[11] N.J.A. Sloane, *"On-Line Encyclopedia of Integer Sequences"*, published electonically at `http://www.research.att.com/~njas/sequences/`.

[12] R.P. Stanley, *"Enumerative Combinatorics"*, Volume 1, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, Cambridge, 1997.

[13] H. N. V. Temperley, *Statistical mechanics and the partitions of numbers. II. The form of crystal surphaces*, Proc. Camb. Phil. Soc. 48 (1952), 683–697.

PERMUTATION REPRESENTATIONS ON INVERTIBLE MATRICES

YONA CHERNIAVSKY , ELI BAGNO

(Extended abstract)

1. Introduction

The (0,1)-matrices have a wide variety of applications in combinatorics as well as in computer science. A lot of research had been devoted to this area. By considering the set of $n \times n$ (0,1)-matrices as a boolean monoid and relating them to posets, one can get interesting representations of S_n. Note that $S_n \times S_n$ acts on matrices by permuting rows and columns. Some aspects of the corresponding equivalence relation are treated in [I] and [Li]. A simultaneous lexicographic ordering of the rows and the columns using this action is shown in [MM].

The above action of $S_n \times S_n$ gives rise to a permutation representation of $S_n \times S_n$ on (0,1)-matrices. If we diagonally embed S_n in $S_n \times S_n$ we get a generalization of the conjugacy representation of S_n.

Adin and Frumkin [AF] showed that the conjugacy character of the symmetric group is close, in some sense, to the regular character of S_n. More precisely, the quotient of the norms of the regular character and the conjugacy character as well as the cosine of the angle between them tend to 1 when n tends to infinity. This implies that these representations have essentially the same decompositions.

Roichman [R] further points out a wide family of irreducible representations of S_n whose multiplicity in the conjugacy representation is asymptotically equal to their dimension, i.e. their multiplicity in the regular representation.

In this paper we use the action of $S_n \times S_n$ on the (0,1)- matrices to define two families of representations on a family of orbits of this action. The first family forms an interpolation between the regular representation of $S_n \times S_n$ and the 'diagonal sum' of the irreducible representations of S_n: $\bigoplus_{\lambda \vdash n} S^\lambda \otimes S^\lambda$. The other family is a generalization of the conjugacy representation of S_n. In both cases we calculate characters and present the decomposition of these representations into irreducibles. The second family of representations can be seen as an extension of the results of [AF] and [R].

2. Preliminaries

2.1. **Symmetric Groups.** S_n is the group of all bijections from the set $\{1...n\}$ to itself. Every $\pi \in S_n$ may be written in *disjoint cycle form* usually omitting the 1-cycles of π. For example, $\pi = 365492187$ may also be written as $\pi = (9,7,1,3,5)(2,6)$. Given $\pi, \tau \in S_n$ let $\pi\tau := \pi \circ \tau$ (composition of functions) so that, for example, $(1,2)(2,3) = (1,2,3)$. Note that two permutations are conjugate in S_n if and only if they have the same cycle structure. In this paper we write $\pi \sim \sigma$ if the permutations π and σ are conjugate in S_n. We denote by $\hat{S}_n$

the set of conjugacy classes of S_n and by $C_\pi \leq S_n$ the centralizer subgroup of the element $\pi \in S_n$. Let $C(\pi) \subseteq S_n$ denote the conjugacy class of the element $\pi \in S_n$. By $supp(\pi)$ we mean the set of digits which are not fixed by π. An element $\pi \in S_n$ with $|supp(\pi)| = t$ can be considered as an element of S_t and then C_π^t denotes the centralizer subgroup of the element π in S_t while $C^t(\pi)$ denotes the conjugacy class of the element π in S_t. $\pi_k\pi_{n-k}$ denotes an element of $S_k \times S_{n-k}$ where $\pi_k \in S_k$ and $\pi_{n-k} \in S_{n-k}$.
$C^{k\times(n-k)}(\pi_k\pi_{n-k})$ denotes the conjugacy class of the element $\pi_k\pi_{n-k}$ in $S_k \times S_{n-k}$.

There is an obvious embedding of S_n in $GL_n(\mathbb{F})$ where is $\mathbb{F}$ is any field. Just think about 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:

$$[\pi]_{i,j} = \begin{cases} 1 & i = \pi(j) \\ 0 & \text{otherwise} \end{cases}$$

Further we identify a permutation with the corresponding permutation matrix.

2.2. **Color permutation groups.** For later use, we define here the color permutation groups. For $r, n \in \mathbb{N}$, let $G_{r,n}$ denote the group of all n by n monomial matrices whose non-zero entries are complex r-th roots of unity. This group can also be described as the wreath product $C_r \wr S_n$ which is the semi-direct product $C_r^n \rtimes S_n$, where C_r^n is taken as the subgroup of all diagonal matrices in $G_{r,n}$. For $r = 1$, $G_{r,n}$ is just S_n while for $r = 2$, $G_{r,n} = B_n$, the Weyl group of type B.

2.3. **Representations.**

2.3.1. *Permutation representations.* In this work we deal mainly with permutation representations. Given an action of a group G on a set M, the appropriate representation space is the space spanned by the elements of M on which G acts by linear extension. We list two well known facts about permutation representations.

Fact 2.1. *The character of the permutation representation calculated at some $g \in G$ equals to the number of fixed points under g.*

Fact 2.2. *The multiplicity of the trivial representation in a given permutation representation is equal to the number of orbits under the corresponding action.*

An important example we will use extensively in this work is the conjugacy representation which is the permutation representation obtained by the action of the group on itself by conjugation.

2.3.2. *Representations of S_n.* Let n be a nonnegative integer. A *partition* of n is an infinite sequence of nonnegative integers with finitely many nonzero terms $\lambda = (\lambda_1, \lambda_2, \ldots)$, where $\lambda_1 \geq \lambda_2 \geq \ldots$ and $\sum\limits_{i=1}^{\infty} \lambda_i = n$.

The sum $\sum \lambda_i = n$ is called the *size* of λ, denoted $|\lambda|$; write also $\lambda \vdash n$. The number of parts of λ, $\ell(\lambda)$, is the maximal j for which $\lambda_j > 0$. The unique partition of $n = 0$ is the *empty partition* $\emptyset = (0, 0, \ldots)$, which has length $\ell(\emptyset) := 0$. For a partition $\lambda = (\lambda_1, \ldots, \lambda_k, \ldots)$ define the *conjugate partition* $\lambda' = (\lambda'_1, \ldots, \lambda'_i, \ldots)$ by letting λ'_i be the number of parts of λ that are $\geq i$ $(\forall i \geq 1)$.

A partition $\lambda = (\lambda_1, ..., \lambda_k)$ may be viewed as the subset

$$\{(i,j) \mid 1 \le i \le k, 1 \le j \le \lambda_i\} \subseteq \mathbb{Z}^2,$$

the corresponding *Young diagram*. Using this interpretation we may speak of the intersection $\lambda \cap \mu$, the set difference $\lambda \setminus \mu$ and the symmetric set difference $\lambda \triangle \mu$ of any two partitions. Note that $|\lambda \triangle \mu| = \sum_{k=1}^{\infty} |\lambda_k - \mu_k|$.

It is well known that the irreducible representations of S_n are indexed by partitions of n (See for example [Sa]) and the representations of $S_n \times S_n$ are indexed by pairs of partitions (λ, μ) where $\lambda, \mu \vdash n$. For every two representations of S_n, λ and ρ, we denote by $m(\lambda, \rho)$ the multiplicity of λ in ρ. If we denote by $\langle \, , \rangle$ the standard scalar product of characters of a finite group G i.e.

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|} \sum_{\pi \in G} \chi_1(\pi) \overline{\chi_2(\pi)}$$

then $m(\lambda, \rho) = \langle \chi_\lambda, \chi_\rho \rangle$.

Similarly, $m\big((\lambda, \mu), \varphi\big)$ denotes the multiplicity of the representation of $S_n \times S_n$ corresponding to the pair of partitions (λ, μ) , $\lambda \vdash n$, $\mu \vdash n$ in the decomposition of φ, where φ is any representation of $S_n \times S_n$.

We cite here for later use the branching rule for the representations of S_n. We start with a definition needed to state the branching rule.

Definition 2.3. *Let $\lambda \vdash n$ be a Young diagram. Then a corner of λ is a cell $(i,j) \in \lambda$ such whose removal leaves leaves the Young diagram of a partition. Any partition obtained by such a removal is denoted by λ^-.*

Proposition 2.4. [Sa] *If $\lambda \vdash n$ then*

$$S^\lambda \downarrow_{S_{n-1}}^{S_n} \cong \bigoplus_{\lambda^-} S^{\lambda^-}.$$

3. The action of $S_n \times S_n$ on invertible matrices

Definition 3.1. *Let G be a subgroup of $S_n \times S_n$ and let $\mathbb{F}$ be any field. We define an action of G on the group $GL_n(\mathbb{F})$ by*

$$(\pi, \sigma) \bullet A = \pi A \sigma^{-1} \text{ where } (\pi, \sigma) \in G \text{ and } A \in GL_n(\mathbb{F}) \qquad (1)$$

It is easy to see that this really defines a group action.

In this work we deal only with the cases: $G = S_n \times S_n$ and $G = (S_k \times S_{n-k}) \times (S_k \times S_{n-k})$.

Definition 3.2. *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 G obtained from the action* (1) . *In the sequel we identify the action* (1) *with the permutation representation α_M associated with it.*

3.1. **A generalization of the conjugacy representation of S_n.**

In this section we present a conjugacy representation of S_n on a subset M of $GL_n(\mathbb{F})$.

Definition 3.3. *Denote by β the permutation representation of S_n obtained by the following action on M.*

$$\pi \circ A = (\pi, \pi) \bullet A = \pi A \pi^{-1} \qquad (2)$$

The connection between α_M and β_M is given by the following easily seen claim:

Claim 3.4. *Consider the diagonal embedding of S_n into $S_n \times S_n$. Then*

$$\beta_M = \alpha_M \downarrow_{S_n}^{S_n \times S_n} . \qquad \square$$

Theorem 3.5. *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. See Theorem 4.5 in [CS]. $\square$

4. The action of $S_n \times S_n$ on (0,1)-matrices

In this section we specialize the action (1) of $S_n \times S_n$ defined in Section 3 to (0,1)-matrices. Consider the group $G = GL_n(\mathbb{Z}_2)$. For every $A \in G$ denote by $o(A)$ the number of nonzero entries in A. One can associate with A a pair of partitions of $o(A)$ with n parts $(\eta(A), \theta(A))$ where $\eta(A)$ describes the distribution of nonzero entries in the rows of A and $\theta(A)$ describes the same distribution for columns. For example, if:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

then $\eta(A) = (4,3,1,1) \vdash 9$ and $\theta(A) = (3,3,2,1) \vdash 9$.

If we fix a pair of partitions (η, θ) then the set of matrices corresponding to (η, θ) is closed under the action (1), but this action is not necessarily transitive on such a set, i.e. it can be decomposed into a union of several orbits.

We present now a family of subsets of $GL_n(\mathbb{Z}_2)$ which will be proven shortly to be orbits of our action:

Definition 4.1.

$$H_n^0 = \{A \in G \mid \eta(A) = \theta(A) = (1,1,1,\ldots,1) = 1^n\}$$

$$H_n^1 = \{A \in G \mid \eta(A) = (n,1,1,\ldots,1), \theta(A) = (2,2,\ldots,2,1) = 2^{n-1}\,1\}$$

$$H_n^2 = \{A \in G \mid \eta(A) = (n,n-1,1,\ldots,1), \theta(A) = (3,3,\ldots,3,2,1) = 3^{n-2}\,2\,1\}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\begin{aligned} H_n^k = \{A \in G \mid \eta(A) &= (n, n-1, \ldots, n-(k-1), 1, \ldots, 1), \\ \theta(A) &= (k+1, k+1, \ldots, k+1, k, k-1, \ldots, 2, 1) = (k+1)^{n-k}\, k\,(k-1)\ \ldots\ 2\,1\} \end{aligned}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$H_n^n = \{A \in \ \mid \eta(A) = \theta(A) = (n, n-1, n-2, \ldots, n-(k-1), \ldots, 3, 2, 1)\}$$

Note that in the above example $A \in H_4^2$.
A few remarks on the sets H_n^k are in order: First, note that $|H_n^k| = n!(n)_k$. Secondly, note that H_n^0 is S_n, embedded as permutation matrices. Also note that the set $H_n^0 \cup H_n^1$ is closed under matrix multiplication and matrix inversion and is actually isomorphic to the group S_{n+1}. Another simple observation is that $H_n^n = H_n^{n-1}$.

In order to prove that the sets H_n^k are transitive under the action we need the following definition:

Definition 4.2. *Denote by $U_{n,k}$ the following binary $n \times n$ matrix : the upper left $k \times k$ block is upper triangular with the upper triangle filled by ones, the upper right $k \times (n-k)$ block is filled by ones, the lower left $(n-k) \times k$ block is the zero matrix and the lower right $(n-k) \times (n-k)$ block is the identity matrix I_{n-k}.*

Proposition 4.3. *Each set H_n^k is transitive under the action α of $S_n \times S_n$. More explicitly, $H_n^k = \{\pi U_{n,k} \sigma \,|\, \pi, \sigma \in S_n\}$*

For the case $k = n$ the permutation representation $\alpha_{H_n^k}$ can be easily described:

Proposition 4.4. *The representation $\alpha_{H_n^n}$ is isomorphic to the regular representation of $S_n \times S_n$.*

4.1. **A natural mapping from H_n^k onto S_n.** In this section we present an epimorphism between the representation of $S_n \times S_n$ on H_n^k to the representation of $S_n \times S_n$ on S_n. We will use this mapping later when we decompose the permutation representation α into irreducibles representations.

Definition 4.5. *Define the mapping $T_{n,k} : H_n^k \longrightarrow S_n$ by $T_{n,k}(\pi U_{n,k} \sigma) = \pi\sigma$.*

Proposition 4.6. *The mapping $T_{n,k}$ preserves the action α of $S_n \times S_n$ on H_n^k, i.e.*

$$T_{n,k}(\pi A \sigma) = \pi T_{n,k}(A)\sigma \text{ for any } A \in H_n^k.$$

It is also clear from the definition that $T_{n,k}$ is onto and it is easy to see that $|T_{n,k}^{-1}(\pi)| = k!\binom{n}{k} = (n)_k$.

5. The representation β_M for $M = H_n^k$.

In [F] it was proven that the conjugacy representation of S_n contains every irreducible representation of S_n as a constituent. The representation β defined in Section 3.1 is a type of a conjugacy representation of S_n on H_n^k.

Proposition 5.1. *Denote the conjugacy representation of S_n by ψ. Then every irreducible representation of S_n is a constituent in $\beta_{H_n^k}$. In other words*

$$m\left(\lambda, \beta_{H_n^k}\right) > 0 \qquad \text{for any} \quad \lambda \vdash n.$$

where $m\big(\lambda, \beta_{H_n^k}\big)$ denotes the multiplicity of the irreducible representation corresponding to λ in $\beta_{H_n^k}$.

We turn now to the calculation of the character of $\beta_{H_n^k}$. By the definition, we have:

$$\chi_{\beta_{H_n^k}}(\pi)\,(= \chi_{\alpha_{H_n^k}}(\pi,\pi)) \,= \#\{A \in H_n^k \,|\, \pi A = A\pi\}$$

but we can achieve much more than that:

Proposition 5.2.

$$\chi_{\beta_{H_n^k}}(\pi) = |C_\pi|(n - |supp(\pi)|)_k = (n - |supp(\pi)|)_k \chi_{Conj}(\pi)$$

where χ_{Conj} is the conjugacy character of S_n.

We turn now to the calculation of the multiplicity of every irreducible representation of S_n in $\beta_{H_n^k}$.

Proposition 5.3. *Let $\lambda \vdash n$.*

$$m\big(\lambda, \beta_{H_n^k}\big) = \sum_{C \in \hat{S}_n} \chi_\lambda(C)(n - |supp(C)|)_k$$

where $\hat{S}_n$ denotes the set of conjugacy classes of S_n. □

6. Asymptotic behavior of the representation $\beta_{H_n^k}$.

In this section we generalize the results of Roichman [R], Adin, and Frumkin [AF] concerning the asymptotic behavior of the conjugacy representation of S_n. These two results imply that the conjugacy representation and the regular representation of S_n have essentially the same decomposition. In our case, as we prove in this section, the representation $\beta_{H_n^k}$ is essentially $(n)_k$ times the regular representation of S_n. We start by citing the result from [R].

Theorem R1 Let $m(\lambda)$ be the multiplicity of the irreducible representation S^λ in the conjugacy representation of S_n, and let f^λ be the multiplicity of S^λ in the regular representation of S_n. Then for any $0 < \varepsilon < 1$ there exist $0 < \delta(\varepsilon)$ and $N(\varepsilon)$ such that, for any partition λ of $n > N(\varepsilon)$ with $\max\{\frac{\lambda_1}{n}, \frac{\lambda'_1}{n}\} \le \delta(\varepsilon)$,

$$1 - \varepsilon < \frac{m(\lambda)}{f^\lambda} < 1 + \varepsilon.$$

The following generalization of this theorem is straightforward:

Proposition 6.1. *For any $0 < \varepsilon < 1$ there exist $0 < \delta(\varepsilon)$ and $N(\varepsilon)$ such that, for any partition λ of $n > N(\varepsilon)$ with max$\{\frac{\lambda_1}{n}, \frac{\lambda'_1}{n}\} \le \delta(\varepsilon)$,and for any $k \le n$*

$$1 - \varepsilon < \frac{m(\lambda, \beta_{H_n^k})}{(n)_k f^\lambda} < 1 + \varepsilon.$$

The following asymptotic result from [AF] can also be generalized for the characters $\chi_{\beta_{H_n^k}}$.

Theorem AF Let $\chi_R^{(n)}$ and $\chi_{Conj}^{(n)}$ be the regular and the conjugacy characters of S_n respectively. Then

$$\lim_{n\to\infty} \frac{\|\chi_R^{(n)}\|}{\|\chi_{Conj}^{(n)}\|} = 1\,,$$

$$\lim_{n\to\infty} \frac{\langle \chi_R^{(n)}, \chi_{Conj}^{(n)} \rangle}{\|\chi_R^{(n)}\| \cdot \|\chi_{Conj}^{(n)}\|} = 1$$

where $\| * \|$ denotes the norm with respect to the standard scalar product of characters.

Our generalization looks as follows:

Proposition 6.2. *In the notations of Theorem AF*

$$\lim_{n\to\infty} \frac{\|(n)_k\chi_R^{(n)}\|}{\|\chi_{\beta_{H_n^k}}\|} = 1\,,$$

$$\lim_{n\to\infty} \frac{\langle (n)_k\chi_R^{(n)}, \chi_{\beta_{H_n^k}}\rangle}{\|(n)_k\chi_R^{(n)}\|\cdot\|\chi_{\beta_{H_n^k}}\|} = \lim_{n\to\infty} \frac{\langle \chi_R^{(n)}, \chi_{\beta_{H_n^k}}\rangle}{\|\chi_R^{(n)}\|\cdot\|\chi_{\beta_{H_n^k}}\|} = 1\,,$$

where k is bounded or tends to infinity remaining less than n.

7. The representations α_M for $M = H_n^k$

In this section we deal with the representations $\alpha_{H_n^k}$ defined in Section 3. We use the branching rule and the Frobenious reciprocity to decompose these representations into irreducible representations of $S_n \times S_n$. As we have already seen in example **??**, $\alpha_{H_n^0} \cong \bigoplus_{\lambda\vdash n} S^\lambda \otimes S^\lambda$ while $\alpha_{H_n^n}$ is the regular representation of $S_n\times S_n \cong \bigoplus_{\lambda,\rho\vdash n} f^\lambda f^\rho S^\lambda \otimes S^\rho$ and thus $\alpha_{H_n^k}$ can be seen as a type of an interpolation between these two representations.

First, concerning the character of $\alpha_{H_n^k}$, by combining Proposition 5.2 and Theorem 3.5 together we get:

$$\chi_{\alpha_{H_n^k}}(\pi,\sigma) = \begin{cases} |C_\pi|(n-|supp(\pi)|)_k\;, & \pi \text{ and } \sigma \text{ are conjugate in } S_n \\ 0\;, & \text{otherwise} \end{cases}$$

We turn now to the lation of the multiplicity of an irreducible representation of $S_n \times S_n$ in $\alpha_{H_n^k}$.

Proposition 7.1. *For any n and any $0 \le k \le n$*

$$m\left((\lambda,\mu),\alpha_{H_n^k}\right) = \frac{1}{n!}\sum_{\pi\in S_n} \chi_\lambda(\pi)\chi_\mu(\pi)(n-|supp(\pi)|)_k \qquad \square$$

The boundary cases $k=0$ and $k=n$ are discussed in Example **??** and Proposition 4.4 respectively.

7.1. A combinatorial view of $\alpha_{H_n^k}$.

In this section we present another approach to the representation $\alpha_{H_n^k}$. This approach will give us a combinatorial view on the multiplicity formulas we calculated in the last section.

Definition 7.2. *Define the following subset of H_n^k:*

$$W_n^k = \{\pi_k\pi_{n-k}U_{n,k}\sigma_k\sigma_{n-k} \mid \pi_k,\sigma_k \in S_k \text{ and } \pi_{n-k},\sigma_{n-k}\in S_{n-k}\}\,.$$

The set W_n^k is the orbit of the matrix $U_{n,k}$ under the action α restricted to the subgroup $(S_k\times S_{n-k})\times(S_k\times S_{n-k})$.

Definition 7.3. *Denote by $\omega_{n,k}$ the permutation representation of the group $(S_k\times S_{n-k})\times(S_k\times S_{n-k})$ on W_n^k corresponding to the action α.*

Claim 7.4.

$$\omega_{n,k} \cong R_k \otimes \left(\bigoplus_{\rho\vdash n-k} S^\rho\otimes S^\rho\right)$$

where R_k is the regular representation of $S_k\times S_k$.

This implies the following:

Claim 7.5.

$$\chi_{\omega_{n,k}}(\pi_k\pi_{n-k},\sigma_k\sigma_{n-k}) = \begin{cases} 0 & \textit{when } \pi_k \neq e \textit{ or } \sigma_k \neq e \\ 0 & \textit{when } \pi_{n-k} \textit{ is not conjugate to } \sigma_{n-k} \textit{ in } S_{n-k} \\ (k!)^2|C^{n-k}_{\pi_{n-k}}| & \textit{when } \pi_k = \sigma_k = e \textit{ and } \pi_{n-k} \sim \sigma_{n-k} \textit{ in } S_{n-k} \end{cases}$$

□

We can use $\omega_{n,k}$ to get information of $\alpha_{n,k}$.

Proposition 7.6.

$$\alpha_{H_n^k} = \omega_{n,k} \uparrow^{S_n \times S_n}_{(S_k \times S_{n-k}) \times (S_k \times S_{n-k})}$$

We use now the Frobenius reciprocity to obtain the multiplicity of any irreducible representation of $S_n \times S_n$ in $\alpha_{H_n^k}$.

Proposition 7.7. *Let $0 \leq k \leq n$ and let λ, μ be partitions of n. Then*

$$m\left((\lambda,\mu), \alpha_{H_n^k}\right) = \langle \chi_\lambda \downarrow^{S_n}_{S_{n-k}}, \chi_\mu \downarrow^{S_n}_{S_{n-k}} \rangle$$

or in other words:

$$\alpha_{H_n^k} = \bigoplus_{\lambda,\mu \vdash n} \langle \chi_\lambda \downarrow^{S_n}_{S_{n-k}}, \chi_\mu \downarrow^{S_n}_{S_{n-k}} \rangle S^\lambda \otimes S^\mu \quad .$$

The number $\langle \chi_\lambda \downarrow^{S_n}_{S_{n-k}}, \chi_\mu \downarrow^{S_n}_{S_{n-k}} \rangle$ has a very nice combinatorial interpretation. It follows from the branching rule that this is just the number of ways to delete k boundary cells from the diagrams corresponding to the partitions λ and μ to get the same Young diagram of $n-k$ cells. By the branching rule (see Proposition 2.4) we have thus:

Claim 7.8.

$$\langle \chi_\lambda \downarrow^{S_n}_{S_{n-k}}, \chi_\mu \downarrow^{S_n}_{S_{n-k}} \rangle = 0 \quad \textit{when} \quad |\lambda \bigtriangleup \mu| > 2k$$

and it does not vanish otherwise.

Corollary 7.9.

$$m\left((\lambda,\mu), \alpha_{H_n^k}\right) = 0 \quad \textit{when} \quad |\lambda \bigtriangleup \mu| > 2k$$

and

$$m\left((\lambda,\mu), \alpha_{H_n^k}\right) \neq 0 \quad \textit{when} \quad |\lambda \bigtriangleup \mu| \leq 2k.$$

□

8. The actions α and β on colored permutations

In this section we introduce actions of S_n and $S_n \times S_n$ on another family of sets, namely the colored permutation groups. We start with the actions on $B_n = C_2 \wr S_n$.

8.1. The action α of $S_n \times S_n$ on signed permutations.

Consider the action α of $S_n \times S_n$ on B_n. We start by describing the orbits of this action.

Definition 8.1. *For every $0 \leq k \leq n$ define*

$$X_n^k = \{A \in B_n \mid A \textit{ has exactly } k \textit{ minuses}\}.$$

For example

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \in X_4^2.$$

It is easy to see that the sets X_n^k form a partition of B_n. Also, note that $|X_n^k| = n!\binom{n}{k}$.

Claim 8.2. *Each set X_n^k is an orbit under the action α of $S_n \times S_n$ on B_n, i.e.*

$$X_n^k = \{\pi \tilde{U}_{n,k} \sigma \mid \pi, \sigma \in S_n\},$$

where

$$\tilde{U}_{n,k} = \begin{pmatrix} -I_k & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & I_{n-k} \end{pmatrix}$$

and I_t is the identity $t \times t$ matrix.

We decompose now the representations $\alpha_{X_n^k}$ into irreducible representations just as we did in the previous section.

Definition 8.3. *Define the following subset of X_n^k:*

$$\tilde{W}_n^k = \left\{ \pi_k \pi_{n-k} \tilde{U}_{n,k} \sigma_k \sigma_{n-k} \mid \pi_k, \sigma_k \in S_k \text{ and } \pi_{n-k}, \sigma_{n-k} \in S_{n-k} \right\}.$$

The set $\tilde{W}_n^k$ is the orbit of the matrix $\tilde{U}_{n,k}$ under the action α by the group $(S_k \times S_{n-k}) \times (S_k \times S_{n-k})$.

Definition 8.4. *Denote $\tilde{\omega}_{n,k}$ the permutation representation of the group $(S_k \times S_{n-k}) \times (S_k \times S_{n-k})$ which is obtained from the action α of this group on the set $\tilde{W}_n^k$.*

Claim 8.5.

$$\tilde{\omega}_{n,k} \cong \left(\bigoplus_{\rho \vdash k} S^\rho \otimes S^\rho \right) \otimes \left(\bigoplus_{\rho \vdash n-k} S^\rho \otimes S^\rho \right)$$

$$\chi_{\tilde{\omega}_{n,k}}(\pi_k \pi_{n-k}, \sigma_k \sigma_{n-k}) = \begin{cases} |C^k_{\pi_k}||C^{n-k}_{\pi_{n-k}}| & \textit{when } \pi_k \pi_{n-k} \sim \sigma_k \sigma_{n-k} \textit{ in } S_k \times S_{n-k} \\ 0 & \textit{otherwise} \end{cases}$$

□

Proposition 8.6.

$$\alpha_{X_n^k} = \tilde{\omega}_{n,k} \uparrow^{S_n \times S_n}_{(S_k \times S_{n-k}) \times (S_k \times S_{n-k})}$$

Recall from [Sa] the definition of $c^\lambda_{\rho\nu}$ – the Littlewood-Richardson coefficients defined by the following formula:

$$(S^\rho \otimes S^\nu) \uparrow^{S_n}_{S_k \times S_{n-k}} = \bigoplus_{\lambda \vdash n} c^\lambda_{\rho\nu} S^\lambda,$$

where $\rho \vdash k$ and $\nu \vdash n-k$. Using the Frobenius reciprocity formula we have for every $\lambda \vdash n$:

Claim 8.7.

$$S^{\lambda}\downarrow^{S_n}_{S_k\times S_{n-k}}=\bigoplus_{\rho\vdash k,\nu\vdash n-k}c^{\lambda}_{\rho\nu}\left(S^{\rho}\otimes S^{\nu}\right).$$

$$\chi_{\lambda}\downarrow^{S_n}_{S_k\times S_{n-k}}=\sum_{\rho\vdash k,\nu\vdash n-k}c^{\lambda}_{\rho\nu}\chi_{(\rho,\nu)}.$$

□

We use now the Frobenius reciprocity to obtain the multiplicity of any irreducible representation of $S_n\times S_n$ in $\alpha_{X^k_n}$.

Proposition 8.8. *Let* $0\le k\le n$ *and* $\lambda,\mu\vdash n$.

$$m\big((\lambda,\mu),\alpha_{X^k_n}\big)=$$
$$=\langle\chi_{\lambda}\downarrow^{S_n}_{S_k\times S_{n-k}},\chi_{\mu}\downarrow^{S_n}_{S_k\times S_{n-k}}\rangle=$$
$$=\sum_{\rho\vdash k,\nu\vdash n-k}c^{\lambda}_{\rho\nu}c^{\mu}_{\rho\nu}$$

By the definition of X^k_n we have $\alpha_{B_n}=\bigoplus_{k=0}^{n}\alpha_{X^k_n}$ and thus:

Corollary 8.9.

$$m\left((\lambda,\mu),\alpha_{B_n}\right)=\sum_{k=0}^{n}\sum_{\rho\vdash k,\nu\vdash n-k}c^{\lambda}_{\rho\nu}c^{\mu}_{\rho\nu}.$$

□

There is a natural mapping between the sets H^k_n and X^k_n defined by:

$$H^k_n\ni\pi U_{n,k}\sigma\overset{\tilde{T}_{n,k}}{\longmapsto}\pi\tilde{U}_{n,k}\sigma\in X^k_n$$

One can verify that $\tilde{T}_{n,k}$ is well defined. Moreover, $\tilde{T}_{n,k}$ commutes with the action α of $S_n\times S_n$ on X^k_n, i.e.:

$$\tilde{T}_{n,k}(\pi A\sigma)=\pi\tilde{T}_{n,k}(A)\sigma\text{ for any }A\in X^n_k.$$

It is easy to see that $\tilde{T}_{n,k}$ is also surjective and thus it induces epimorphisms of modules from the $S_n\times S_n$-module $\alpha_{H^k_n}$ to the $S_n\times S_n$- module $\alpha_{X^k_n}$ and from the S_n-module $\beta_{H^k_n}$ to the S_n- module $\beta_{X^k_n}$. Note also that for $k=0$ this mapping is the identity mapping since $H^0_n=X^0_n=S_n$ and for $k=1$ this mapping is bijective. We conclude:

Claim 8.10.

$$m\left((\lambda,\mu),\alpha_{H^k_n}\right)\ge m\left((\lambda,\mu),\alpha_{X^k_n}\right)$$

□

This implies that if

$$\sum_{\rho\vdash k,\nu\vdash n-k}c^{\lambda}_{\rho\nu}c^{\mu}_{\rho\nu}\neq 0$$

then $|\lambda\triangle\mu|\le 2k$. This can also be seen by the combinatorial interpretation of the Littlewood-Richardson coefficients.

8.2. **The action β on colored permutations.** Recall that every matrix $B \in B_n$ can be written uniquely in the form $B = Z\pi$ for some $\pi \in S_n$ and some $Z \in C_2^n$. There exists a natural epimorphism from B_n onto S_n defined by omitting the minuses:

$$p : B_n \longrightarrow S_n \qquad\qquad p(Z\pi) = \pi.$$

If we restrict p to X_n^k we obtain a surjective mapping from X_n^k onto S_n which commutes with the action α of $S_n \times S_n$ on X_n^k (and clearly also commutes with the action β of S_n on X_n^k by conjugation). It gives us a surjective homomorphism from the representation $\beta_{X_n^k}$ onto the conjugacy representation representation of S_n denoted by ψ. Therefore, using the result of [F], we have

$$m\left(\lambda, \beta_{X_n^k}\right) \geq m\left(\lambda, \psi\right) > 0 \quad \text{for any} \quad \lambda \vdash n,$$

$$m\left(\lambda, \beta_{B_n}\right) = \sum_{k=0}^{n} m\left(\lambda, \beta_{X_n^k}\right) > 0 \quad \text{for any} \quad \lambda \vdash n.$$

Although the calculation of $\chi_{\beta_{X_n^k}}$ is rather involved, the asymptotic results of [R] and [AF] can be generalized for the representations $\beta_{X_n^k}$ and β_{B_n}. We start by presenting the generalization of Theorem R1:

Proposition 8.11. *For any* $0 < \varepsilon < 1$ *there exist* $0 < \delta(\varepsilon)$ *and* $N(\varepsilon)$ *such that, for any partition* λ *of* $n > N(\varepsilon)$ *with max*$\{\frac{\lambda_1}{n}, \frac{\lambda_1'}{n}\} \leq \delta(\varepsilon)$,

$$1 - \varepsilon < \frac{m(\lambda, \beta_{X_n^k})}{\binom{n}{k} f^\lambda} < 1 + \varepsilon\ ,$$

$$1 - \varepsilon < \frac{m(\lambda, \beta_{B_n})}{2^n f^\lambda} < 1 + \varepsilon\ .$$

The generalization of Theorem [AF] is as follows and can be proved by using the inequality $\chi_{\beta_{X_n^k}}(\pi) \leq \binom{n}{k}|C_\pi|$:

Proposition 8.12. *In the notations of Theorem* [AF]

$$\lim_{n\to\infty} \frac{\|\binom{n}{k}\chi_R^{(n)}\|}{\|\chi_{\beta_{X_n^k}}\|} = \lim_{n\to\infty} \frac{\|2^n\chi_R^{(n)}\|}{\|\chi_{\beta_{B_n}}\|} = 1\,,$$

$$\lim_{n\to\infty} \frac{\langle \chi_R^{(n)}, \chi_{\beta_{X_n^k}}\rangle}{\|\chi_R^{(n)}\| \cdot \|\chi_{\beta_{X_n^k}}\|} = \lim_{n\to\infty} \frac{\langle \chi_R^{(n)}, \chi_{\beta_{B_n}}\rangle}{\|\chi_R^{(n)}\| \cdot \|\chi_{\beta_{B_n}}\|} = 1\,,$$

where k is bounded or tends to infinity remaining less than n.

These asymptotic results can be also obtained for the action β (conjugation by permutations) on the group $C_r \wr S_n$. Similarly to $X_n^k \subset B_n$ define the sets $Y_n^k \subset C_r \wr S_n$:

Definition 8.13.

$$Y_n^k = \{A \in C_r \wr S_n \mid A \textit{ has exactly } k \textit{ entries} \neq 0\,,\, 1\}.$$

Note that the sets Y_n^k form a partition of $C_r \wr S_n$ and $Y_n^k = n!\binom{n}{k}(r-1)^k$. The sets Y_n^k are closed under the action α of $S_n \times S_n$ but they are not transitive under this action.

Consider C_r^n as the group of diagonal matrices with the entries of the form ω^ℓ (where $\omega = \exp\frac{2\pi i}{r}$ – the primitive r-th root of unity and $0 \leq \ell < r$) on the diagonal. Then each matrix $A \in C_r \wr S_n$ can be uniquely written as $A = Z\sigma$ for some $\sigma \in S_n$ and some $Z \in C_r^n$. Just as in the case of B_n, we consider the epimorphism $p : B_n \longrightarrow S_n$ defined by: $p(Z\sigma) = \sigma$.

p induces an epimorphism of modules between $\beta_{Y_n^k}$ and the conjugacy representation of S_n.

We conclude, using the result of [F]:

$$m\left(\lambda, \beta_{Y_n^k}\right) \geq m\left(\lambda, \psi\right) > 0 \quad \text{for any} \quad \lambda \vdash n,$$

$$m\left(\lambda, \beta_{C_r \wr S_n}\right) = \sum_{k=0}^{n} m\left(\lambda, \beta_{Y_n^k}\right) > 0 \quad \text{for any} \quad \lambda \vdash n.$$

The Theorems R1 and AF are obtained in a way similar to the one we used for B_n:

Proposition 8.14. *In the conditions and notations of Theorem R1*

$$1 - \varepsilon < \frac{m(\lambda, \beta_{Y_n^k})}{\binom{n}{k}(r-1)^k f^\lambda} < 1 + \varepsilon,$$

$$1 - \varepsilon < \frac{m(\lambda, \beta_{C_r \wr S_n})}{r^n f^\lambda} < 1 + \varepsilon.$$

In the notations of Theorem AF

$$\lim_{n\to\infty} \frac{\|\binom{n}{k}(r-1)^k \chi_R^{(n)}\|}{\|\chi_{\beta_{Y_n^k}}\|} = \lim_{n\to\infty} \frac{\|r^n \chi_R^{(n)}\|}{\|\chi_{\beta_{C_r \wr S_n}}\|} = 1\,,$$

$$\lim_{n\to\infty} \frac{\langle \chi_R^{(n)}, \chi_{\beta_{Y_n^k}} \rangle}{\|\chi_R^{(n)}\| \cdot \|\chi_{\beta_{Y_n^k}}\|} = \lim_{n\to\infty} \frac{\langle \chi_R^{(n)}, \chi_{\beta_{C_r \wr S_n}} \rangle}{\|\chi_R^{(n)}\| \cdot \|\chi_{\beta_{C_r \wr S_n}}\|} = 1\,,$$

where k is bounded or tends to infinity remaining less than n.

References

[AF] R. M. Adin and A. Frumkin, *The conjugacy character tends to be regular*, Israel J. Math. **59** (1987), 234–240.

[CS] Y. Cherniavsky and M. Sklarz, *On conjugation action of S_n on invertible matrices*, Preprint, available on the website: www.math.biu.ac.il/˜cherniy

[F] A. Frumkin, *Theorem about the conjugacy representation of S_n*, Israel J. Math. **55** (1986), 121–128.

[I] Iwahori N.*On an equivalence relatioion among (0,1)-matrices*, Sci.Papers College Gen.Ed.Univ.Tokyo **21**(1971), 1-9.

[Li] Li J.S.*Permutation equivalence of (0,1)- square matrices*, Acta Math.Sinica **26** (1983), no.5, 586-596.

[Md] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, second edition, Oxford Math. Monographs, Oxford Univ. Press, Oxford, 1995.

[MM] Mader A., Mutzbauer O.,*Double orderings of (0,1)-matrices*, Ars Combin. **61**(2001),81-95.

[R] Y. Roichman, *Decomposition of the conjugacy representation of the symmetric groups*, Israel J. Math. **97** (1997), 305–316.

[Sa] B. Sagan, *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, Springer Verlag (2001).

Department of Mathematics and Statistics Bar-Ilan University Ramat-Gan, Israel 52900

E-mail address: `cherniy@math.biu.ac.il , bagnoe@math.biu.ac.il`

A COMBINATORIAL DERIVATION OF THE PASEP STATIONARY STATE

SYLVIE CORTEEL, RICHARD BRAK, ANDREW RECHNITZER, AND JOHN ESSAM

Abstract. We give a combinatorial derivation and interpretation of the algebra associated with the stationary distribution of the partially asymmetric exclusion process. Our derivation works by constructing a larger Markov chain on a larger set of generalised configurations. A bijection on this new set of configurations allows us to find the stationary distribution of the new chain. We then show that a *subset* of the generalised configurations is equivalent to the original chain and that the stationary distribution on this subset is simply related to that of the original chain. This derivation allows us to find expressions for the normalisation using both recurrences and path models. These results exhibit classical combinatorial numbers such as $n!$, 2^n and the Catalan numbers.

RESUME. Nous donnons une interprétation combinatoire de l'algèbre associé à la distribution stationnaire du processus d'exclusion partiellement symétrique. Pour cela nous construisons une chaine de Markov qui agit sur les configurations et les configurations marquées. Nous utilisons une bijection sur les configurations marquées pour trouver la distribution stationnaire de la nouvelle chaine qui nous donne aussi la distribution stationnaire de la chaine originelle. Grace à cette construction, nous pouvons calculer dans plusieurs cas particuliers le coefficient de normalisation de la chaine en utilisant des récurrences et/ou des chemins. De nombreux nombres classiques apparaissent: Catalan, factoriel, 2^n...

1. Introduction

The PASEP (Partially asymmetric exclusion process) is a generalisation of the TASEP model presented last year at FPSAC'04 [13]. This model was introduced by physicists [2, 9, 10, 11, 12, 16]. The TASEP model consists of black particles entering a row of n cells, each of which is occupied by a black particle or vacant. A particle may enter the system from the left hand side, hop to the right and leave the system from the right hand side, with the constraint that a cell contains at most one particle. The particles in the PASEP move in the same way as those in the TASEP, but in addition may enter the system from the right hand side, hop to the left and leave the system from the left hand side. An example is given in Figure 1.

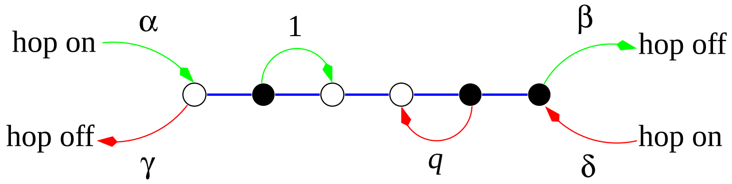


Figure 1. The PASEP model. Note that we also use the variables $c = 1/\alpha - 1$ and $d = 1/\beta - 1$.

¿From now on we will say that the empty cells are filled with white particles $\circ$. A basic configuration is a row of n cells, each containing either a black or a white particle. Let $\mathcal{B}_n$ be the set of basic configurations of n particles. We write these configurations as though they are words of length n in the language $\{\circ, \bullet\}^*$, so that "$\bullet^k$" denotes a string of k black particles and "AB" denotes the configuration made up of the word "A" followed by the word "B". We denote the length of the word A by $|A|$.

The PASEP defines a Markov chain P defined on $\mathcal{B}_n$ with the transition probabilities α, β, γ, δ, η and q. The probability $P_{X,Y}$, of finding the system in state Y at time $t+1$ given that the system is in state X at time t is defined by:

∗ If $X = A \bullet \circ B$ and $Y = A \circ \bullet B$ then

(1a) $$P_{X,Y} = \eta/(n+1); \qquad P_{Y,X} = q/(n+1)$$

∗ If $X = \circ B$ and $Y = \bullet B$ then

$$P_{X,Y} = \alpha/(n+1); \qquad P_{Y,X} = \gamma/(n+1) \tag{1b}$$

∗ If $X = B\bullet$ and $Y = B\circ$ then

$$P_{X,Y} = \beta/(n+1); \qquad P_{Y,X} = \delta/(n+1) \tag{1c}$$

∗ Otherwise $P_{X,Y} = 0$ for $Y \neq X$ and $P_{X,X} = 1 - \sum_{X \neq Y} P_{X,Y}$.

See an example for $n = 2$ in Figure 2.

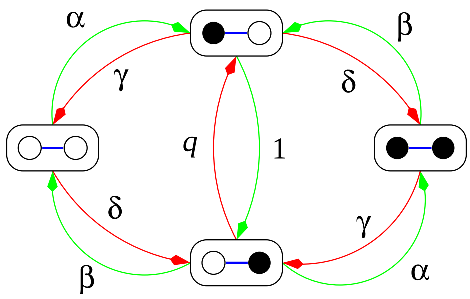


FIGURE 2. The chain P for $n = 2$.

There are many results for the PASEP model. One central question is the computation of the stationary distribution of the chain. This has been most successfully analysed using a matrix product Ansatz [9, 11, 16].

In this paper we give a combinatorial derivation and interpretation of the stationary distribution of the PASEP model. To our knowledge [4, 15] are the only purely combinatorial derivations of the solution to the TASEP model (which is a special case of PASEP). Our derivation works by

[i] constructing a larger Markov chain on both basic configurations and marked basic configurations which we call *marked configurations*.

[ii] using a bijection between marked configurations to find the stationary distribution of the larger chain.

[iii] showing that a *subset* of the configurations is equivalent to the original chain and that the stationary state on this subset is simply related to that of the original chain.

We note that this is similar to the work done in [15] where the authors studied the case $\delta = \gamma = q = 0$. Here we study the stationary distribution of the full model.

The main result of this paper is to give a combinatorial derivation of the following theorem:

Theorem 1. *Let the stationary distribution be denoted*

$$\mathrm{P}_\infty(X) = \frac{W(X)}{Z_n} \tag{2}$$

where $W(X)$ is the weight *of a basic configuration X and Z_n is the normalisation:*

$$Z_n = \sum_{X \in \mathcal{B}_n} W(X). \tag{3}$$

The weight of a basic configuration X satisfies the following algebra :

$$W(X) = 1 \quad \text{if } X \in \mathcal{B}_0 \tag{4a}$$

$$\alpha W(\circ X) = W(X) + \gamma W(\bullet X) \tag{4b}$$

$$\beta W(X\bullet) = W(X) + \delta W(X\circ) \tag{4c}$$

$$\eta W(A \bullet \circ B) = W(A \circ B) + W(A \bullet B) + q W(A \circ \bullet B) \tag{4d}$$

where A and B denote configurations of particles.

We may use this algebra recursively to find expressions for Z_n. Unfortunately we have not yet found a combinatorial derivation of Sasamoto's full five parameter expression for Z_n [16] (one of the six parameters can be set to one without loss of generality), but we are able to find many different specialisations and for example the result of [2]. See Table 1 for a summary of these results.

We also find simple expressions for the stationary distribution of certain configurations. For example:

Proposition 1. *If $\gamma = \delta = 0$ and $q = \alpha = \beta = 1$ then*

$$\mathrm{P}_\infty(X = \bullet^k A) = \frac{(n-k+1)^k (n-k+1)!}{(n+1)!} \tag{5}$$

where A is any configuration in $\mathcal{B}_{n-k}$.

Another special case is:

Proposition 2. *If $\gamma = \delta = q = 0$ and $\alpha = \beta = 1/2$ then*

$$\mathrm{P}_\infty(X) = \frac{1}{2^n}, \tag{6}$$

independent of the configuration X.

In Section 2 we use the weight algebra to find recurrences for the normalisations and also describe a path model interpretation of the stationary states and normalisations. In Section 3 we define a larger Markov chain whose stationary distribution is related to that of the PASEP chain. In Sections 4 and 5 we give proofs of some of the propositions. A longer version of this paper which includes all the proofs is in preparation [5].

2. Algorithms, numbers and paths

In this section we set[1] $\eta = 1$ and study the normalisation using the weight algebra. By considering the position of first (leftmost) $\circ$ particle the algebra may be used to obtain a recursion to compute Z_n when either γ or δ are zero. Here we consider $\delta = 0$, and we note that similar results may be obtained for $\gamma = 0$.

2.1. **Recursions for Z_n.** Let $W_{n,k}$ be the sum of the weights of configurations in $\mathcal{B}_n$ that start with exactly k $\bullet$s and then at least one $\circ$ or are all black. Similarly let $W_{n,k,j}$ the sum of the weights of basic configurations in $\mathcal{B}_n$ that start with exactly k $\bullet$s then a single $\circ$ and then at least j $\bullet$s. Finally let $Z_{n,k}$ the sum of the weights of basic configurations in $\mathcal{B}_n$ that start with *at least* k $\bullet$s.

Proposition 3. *If $\delta = 0$ then $Z_{n,k}$, $W_{n,k}$ and $W_{n,k,j}$ satisfy the following equations*

$$W_{n,k} = \begin{cases} (Z_{n-1,0} + \gamma Z_{n,1})/\alpha & \text{if } k = 0 \\ W_{n-1,k-1} + Z_{n-1,k} + qW_{n,k-1,1} & \text{if } k \in [1, n-1] \\ W_{n-1,n-1}/\beta & \text{if } k = n \\ 0 & \text{if } k > n \end{cases} \tag{7a}$$

$$W_{n,k,j} = \begin{cases} W_{n,k} & \text{if } j = 0 \\ (Z_{n-1,j} + \gamma Z_{n,j})/\alpha & \text{if } k = 0 \\ 0 & \text{if } k + j > n \\ Z_{n-1,k+j} + W_{n-1,k-1,j} + qW_{n,k+1,j-1} & \text{otherwise} \end{cases} \tag{7b}$$

$$Z_{n,k} = \begin{cases} W_{n,k} + Z_{n,k+1} & \text{if } k \in [0, n-1] \\ W_{n,n} & \text{if } k = n \\ 0 & \text{if } k > n \end{cases} \tag{7c}$$

These follow from application of the relations of the weights equations to basic configurations where the first $\bullet\circ$ pair occurs at position k for $k \in [1, n-1]$. The case $k = n$ corresponds to an all black particle configuration.

Using these recurrences we can compute $Z_n = Z_{n,0}$ for any n. We used this data to guess some of the results for specific α, β, γ, q, and then proved the conjectured forms. In Table 1 we give a list of some results like those in Proposition 1. Note we have used $\alpha = 1/(1+c)$ and $\beta = 1/(1+d)$. The result $q = c = d = 0$ appeared in [4, 13, 11]. The proofs will appear in [5].

---

[1] One can show that setting $\eta = 1$ is equivalent to rescaling the other parameters in the stationary state distribution.

set to 0	set to 1	Z_n	$Z_{n,k}$
q, γ	c, d	2^{2n}	2^{2n-k}
q, γ, d, c		$\dfrac{1}{n+2}\dbinom{2n+2}{n+1}$	$\dfrac{k+2}{n+2}\dbinom{2n-k+1}{n-k}$
q, γ, d	c	$\dbinom{2n}{n}$	$\dbinom{2n-k+1}{n-k}$
γ	q	$\prod_{k=1}^{n}(c+d+k+1)$	$(n-k+d+1)^k Z_{n-k}$
	γ, q	$\prod_{i=1}^{n}((c+2)(i+d+1)-1)$	$(n-k+d+1)^k Z_{n-k}$

TABLE 1. Some results for Z_n when $\delta = 0$. Note that $c, d = 0$ implies that $\alpha, \beta = 1$ (respectively) and $c, d = 1$ implies that $\alpha, \beta = 1/2$ (respectively).

2.2. **Path models for $\gamma = \delta = 0$.** Another way to compute Z_n is to give a bijection between basic configurations counted by their weights and a family of weighted lattice paths. A similar result was given in [4] for $\gamma = \delta = q = 0$. In [14] the complete configurations can also be interpreted as paths when $\gamma = \delta = q = 0$. Here we generalise one of the approaches of [4] to get the result for general q.

Definition 1. *A* Motzkin path [1] *of length n is a sequence of vertices $p = (v_0, v_1, \ldots, v_n)$, with $v_i \in \mathbb{N}^2$ (where $\mathbb{N} = \{0, 1, \ldots\}$), with steps $v_{i+1} - v_i \in \{(1,1), (1,-1), (1,0)\}$ and $v_0 = (0,0)$ and $v_n = (n, 0)$. A* bicoloured Motzkin path *is a Motzkin path in which each step $(1,0)$ is labelled by one of two colours.*

These paths can be mapped to words in the language $\{N, S, \overset{\circ}{E}, \overset{\bullet}{E}\}$ by mapping the steps $(1,1), (1,-1)$ to N and S, respectively and mapping the two different coloured horizontal steps to $\overset{\circ}{E}$ and $\overset{\bullet}{E}$. The height of a step $v_{i+1} - v_i$ is the y-coordinate of the vertex v_i. These heights imply the weights of the paths.

Definition 2. *Let $\mathcal{P}(n)$ be the set of bicoloured Motzkin paths of length n. The weight of the path in $\mathcal{P}(0)$ is 1 and the weight of any other path is the product of the weights of its steps. The weight of a step, p_k, starting at height h is given by:*

$$\text{if } p_k = N \quad \text{then} \quad w(p_k) = N_h = [h+1]_q \tag{8a}$$

$$\text{if } p_k = \overset{\circ}{E} \quad \text{then} \quad w(p_k) = \overset{\circ}{E}_h = [h+1]_q + q^h c \tag{8b}$$

$$\text{if } p_k = \overset{\bullet}{E} \quad \text{then} \quad w(p_k) = \overset{\bullet}{E}_h = [h+1]_q + q^h d \tag{8c}$$

$$\text{if } p_k = S \quad \text{then} \quad w(p_k) = S_h = [h]_q + q^h(1+c)(1+d) - q^{h-1}cd \tag{8d}$$

where $[h]_q = 1 + q + \ldots + q^{h-1}$ for $h > 0$ and $[h]_q = 0$ for $h \leq 0$.

An example of a path in $\mathcal{P}(11)$ is given in Figure 3.

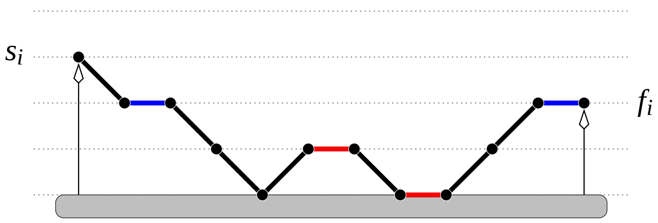


FIGURE 3. A path in $\mathcal{P}(11)$ which corresponds to the word $NN\overset{\circ}{E}SSN\overset{\bullet}{E}S\overset{\bullet}{E}NS$.

Let us define a mapping $\theta : \mathcal{P}(n) \mapsto \mathcal{B}_n$ where each bicoloured Motzkin path is mapped to a basic configuration such that each step S or $\overset{\circ}{E}$ is mapped to a white particle and each step N or $\overset{\bullet}{E}$ is mapped to a black particle. This mapping is many-to-one and we denote $\theta^{-1}(X)$ as the set of all paths that map to X.

Theorem 2. *When $\gamma = \delta = 0$ the weight of a basic configuration, X, is given by*

$$W(X) = \sum_{p \in \theta^{-1}(X)} w(p) \tag{9}$$

and

$$Z_n = \sum_{p \in \mathcal{P}(n)} w(p) \tag{10}$$

This Theorem gives a combinatorial derivation of the stationary distribution that does not make use of the matrix product Ansatz which was used to obtain the results in [11, 16]. The proof is given in Section 5 and works by showing that the the weight of the paths obeys the same equations as the weight of the basic configurations.

We may specialise the above result to get the corollary:

Corollary 3. *If $c = d = 0$ ($\alpha = \beta = 1$) the paths $\mathcal{P}$ correspond to bicoloured Motzkin Paths of length n where the weight of a step starting at height i is $[i+1]_q$.*

- *If $q = 0$ then $Z_n = C_{n+1}$, where C_n is the n^{th} Catalan number.*
- *If $q = 1$ then $Z_n = (n+1)!$ (see [3, 17]).*

This can be linked to well-known results on the q-enumeration of permutations [7, 8]. Also if $c = d = 1$ (i.e. $\alpha = \beta = 1/2$) and $q = 0$, only the paths that are made of east steps have non zero weight. Each such path has weight 2^n. Therefore $W(X) = 2^n$ for any $X \in \mathcal{B}_n$ and $Z_n = 4^n$ in that case — this is Proposition 2.

3. Stationarity and Marked configurations

In this section we define a larger Markov chain, the M-PASEP chain, we which use to study the stationary distribution of the original PASEP chain. In particular we show that the stationary distributions of the M-PASEP and the PASEP chains are simply related.

3.1. Marked configurations.

We enlarge the state space of the original chain by adding "embellished" configurations which we call "marked configurations" (hence the "M" in M-PASEP).

Definition 3. *A marked configuration (X, i, D) of size n consists of a basic configuration $X \in \mathcal{B}_n$, an integer $i \in [0, n]$ and a "direction" $D \in \{L, R, S, N\}$. The directions are L for "left", R for "right", S for "stable" and N for "nothing". The possible values of D depend on the values of X and i. All triples satisfying the following conditions occur (depending on the structure of X):*

- *for all X and all $i \in [0, n]$ $D = N$.*
- *if $X = \circ A$ then $i = 0$ and $D \in \{R, S\}$.*
- *if $X = \bullet A$ then $i = 0$ and $D \in \{S\}$.*
- *if $X = A\circ$ then $i = n$ and $D \in \{S\}$.*
- *if $X = A\bullet$ then $i = n$ and $D \in \{L, S\}$.*
- *if $X = A \bullet \circ B$ and $|A| \in [0, n-2]$ then $i = |A| + 1$ and $D \in \{L, R, S\}$.*
- *if $X = A \circ \bullet B$ and $|A| \in [0, n-2]$ then $i = |A| + 1$ and $D \in \{S\}$.*

We define a projection $U(M) = X$ from a marked configuration $M = (X, i, D)$ to the corresponding unmarked configuration, X. We denote the set of all marked configurations of size n by $\mathcal{M}_n$.

Here are all the marked configurations for $n = 2$:

$$\left\{ \begin{array}{llllllll} (\circ\circ, 0, R), & (\circ\circ, 0, S), & (\circ\circ, 0, N), & (\circ\circ, 1, N), & (\circ\circ, 2, S), & (\circ\circ, 2, N), & & \\ (\circ\bullet, 0, R), & (\circ\bullet, 0, S), & (\circ\bullet, 0, N), & (\circ\bullet, 1, S), & (\circ\bullet, 1, N), & (\circ\bullet, 2, L), & (\circ\bullet, 2, S), & (\circ\bullet, 2, N), \\ (\bullet\circ, 0, S), & (\bullet\circ, 0, N), & (\bullet\circ, 1, L), & (\bullet\circ, 1, R), & (\bullet\circ, 1, S), & (\bullet\circ, 1, N), & (\bullet\circ, 2, S), & (\bullet\circ, 2, N), \\ (\bullet\bullet, 0, S), & (\bullet\bullet, 0, N), & (\bullet\bullet, 1, N), & (\bullet\bullet, 2, L), & (\bullet\bullet, 2, S), & (\bullet\bullet, 2, N) & & \end{array} \right\}$$

3.2. **The M-PASEP chain.** We define the M-PASEP chain, C, whose states are both basic and marked configurations. For any $X \in \mathcal{B}_n$ and $M \in \mathcal{M}_n$ the transition probabilities between states in the chain are given by

$$\text{if} \quad U(M) = X \quad \text{then} \quad C_{X,M} = \frac{W(M)}{(n+1)W(X)} \tag{11a}$$

$$\text{if} \quad U(T(M)) = X \quad \text{then} \quad C_{M,X} = 1 \tag{11b}$$

$$\text{otherwise} \qquad C_{M,X} = C_{X,X} = C_{M,M} = 0, \tag{11c}$$

where $W(M)$ is the weight of a marked configuration which we define below and $T : \mathcal{M}_n \mapsto \mathcal{M}_n$ is a weight preserving bijection given in Definition 5 below.

Definition 4. *The weight of a marked configuration M is defined in terms of the weights of unmarked configurations as follows:*

$$W(\circ A, 0, R) = W(A) \tag{12a}$$
$$W(\circ A, 0, S) = W(\bullet A, 0, S) = \gamma W(\bullet A) \tag{12b}$$
$$W(\circ A, 0, N) = (1-\alpha)W(\circ A) \tag{12c}$$
$$W(A\bullet, n, L) = W(A) \tag{12d}$$
$$W(A\bullet, n, S) = W(A\circ, n, S) = \delta W(A\circ) \tag{12e}$$
$$W(A\bullet, n, N) = (1-\beta)W(A\bullet) \tag{12f}$$
$$W(A\bullet\circ B, i, R) = W(A \circ B) \quad \textit{if} \quad A \in \mathcal{B}_{i-1} \tag{12g}$$
$$W(A\bullet\circ B, i, L) = W(A \bullet B) \quad \textit{if} \quad A \in \mathcal{B}_{i-1} \tag{12h}$$
$$W(A\bullet\circ B, i, S) = W(A\circ\bullet B, i, S) = qW(A\circ\bullet B) \quad \textit{if} \quad A \in \mathcal{B}_{i-1} \tag{12i}$$
$$W(A\bullet\circ B, i, N) = (1-\eta)W(A\bullet\circ B) \quad \textit{if} \quad A \in \mathcal{B}_{i-1} \tag{12j}$$
$$\textit{Otherwise} \quad W(X, i, N) = W(X). \tag{12k}$$

Note that the chain alternates between marked and basic configurations. The state graph of the chain C for $n = 2$ is shown in Figure 4. The weights of marked and unmarked configurations are simply related and from Definition 4 and Theorem 1, we get:

Lemma 4. *For all $X \in \mathcal{B}_n$ and $i \in [0, n]$:*

$$\sum_D W(X, i, D) = W(X) \tag{13}$$

and

$$\sum_{M \,:\, U(M)=X} W(M) = (n+1)W(X). \tag{14}$$

The stationary distribution of the PASEP chain is simply related to that of the new chain C:

Proposition 4. *The conditional stationary probability of finding the M-PASEP chain, C, in a state Y given that it is in an unmarked state is related to the stationary distribution of PASEP chain by*

$$\mathrm{P}^C_\infty(Y \textit{ given that } Y \in \mathcal{B}_n) = \mathrm{P}_\infty(Y). \tag{15}$$

This proposition relates the stationary distribution of the two chains, but does not tell us what the distributions are. We prove the above proposition in the next section and also expand it to give the proof of Theorem 1.

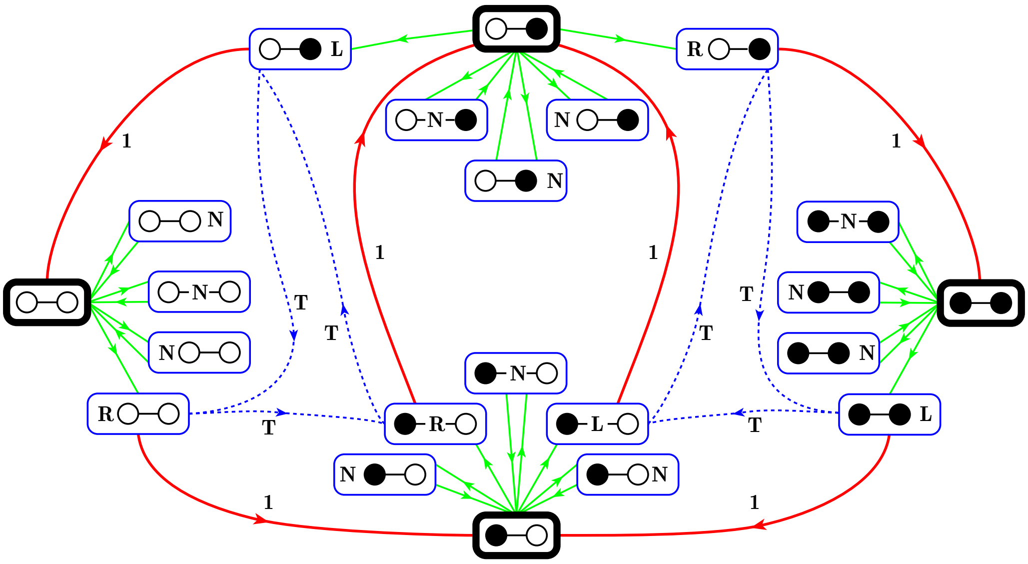


FIGURE 4. The M-PASEP chain C for $n = 2$. Each marked state (X, i, D) is written as the configuration X with its direction D in position i. The short thin arcs have probabilities given by equations (11). The dashed lines show the action of the bijection T.

4. PROOF OF STATIONARITY

To prove stationarity we need two major ingredients. The first is a bijection between states on the M-PASEP chain and the second is lemma which gives conditions under which a Markov chain may be altered while leaving its stationary distribution essentially unchanged.

4.1. The bijection.

Definition 5. *We define the mapping $T : \mathcal{M}_n \to \mathcal{M}_n$. Let $M = (X, i, D) \in \mathcal{M}_n$. We first note that if $D = N$ then $T(M) = M$. Otherwise we define the bijection by the following algorithm:*

- $*$ *if $i = 0$ then the colour of the first particle is changed and*
  - $\diamond$ *if $D = S$ then i and D are unchanged, or*
  - $\diamond$ *if $D = R$ then* "shuffle M right"*.*
  - $\diamond$ *Note that $M = (X, 0, L)$ cannot occur.*
- $*$ *if $i \in [1, n-1]$ then swap the i^{th} and $i+1^{th}$ particles and*
  - $\diamond$ *if $D = S$ then i and D are unchanged, or*
  - $\diamond$ *if $D = L$ then* "shuffle M left"*.*
  - $\diamond$ *if $D = R$ then* "shuffle M right"*.*
- $*$ *if $i = n$ then the colour of the last particle is changed and*
  - $\diamond$ *if $D = S$ then i and D are unchanged, or*
  - $\diamond$ *if $D = L$ then* "shuffle M left"*.*
  - $\diamond$ *Note that $M = (X, n, R)$ cannot occur.*

where "shuffle M right" *means*

- $*$ *choose the minimal $j \in (i, n)$ such that the j^{th} particle is black and the $(j+1)^{th}$ particle is white.*
- $*$ *if such a j exists then set $i = j$ and $D = R$*
- $*$ *otherwise set $i = n$ and $D = L$.*

and "shuffle M left" *means*

- $*$ *choose the maximal $j \in [0, i)$ such that the j^{th} particle is black and the $(j+1)^{th}$ particle is white.*
- $*$ *if such a j exists then set $i = j$ and $D = L$*
- $*$ *otherwise set $i = 0$ and $D = R$.*

Some examples of this mapping are given below

$$\begin{aligned} T(\bullet\circ\bullet\bullet\circ\bullet,0,S) &= (\circ\circ\bullet\bullet\circ\bullet,0,S)\\ T(\bullet\circ\bullet\bullet\circ\bullet,4,R) &= (\bullet\circ\bullet\circ\bullet\bullet,6,L)\\ T(\bullet\circ\bullet\bullet\circ\bullet,4,L) &= (\bullet\circ\bullet\circ\bullet\bullet,1,L). \end{aligned}$$

The bijection is weight preserving:

Proposition 5. *The mapping T defined in Definition 5 is a bijection from $\mathcal{M}_n$ to itself and $\forall M \in \mathcal{M}_n$*

$$\text{(16)} \qquad \textit{either} \quad T(M) = M \quad \textit{or} \quad T^2(M) = M \quad \textit{or} \quad T^{n+1}(M) = M.$$

The bijection is also weight-preserving: $W(T(M)) = W(M)$.

Sketch of Proof. It follows directly from the definition of the bijection that if $D = N$ then $T(M) = M$ and that if $D = S$ then $T^2(M) = M$. The weight is invariant in both cases. To show that if $D = L, R$ then $T^{n+1}(M) = M$, we use a one-to-n+1 weight preserving correspondence between the basic configurations in $\mathcal{B}_{n-1}$ and marked configurations in $\mathcal{M}_n$ with $D = L, R$.

Start with a configuration X in $\mathcal{B}_{n-1}$. Suppose that the white particles are located at positions $j_1 < j_2 < \ldots < j_i$ and that the black particles are located at positions $j_{i+2} > j_{i+3} > \ldots > j_n$. Now create $n+1$ marked configuration $M_0, M_1, \ldots, M_n$ as follows :

* $M_0 = (\circ X, 0, R)$.
* $M_l = (X_l, j_l, R)$ where X_l is obtained from X by replacing the $\circ$ located at j_l by $\bullet\circ$ with $1 \leq l \leq i$.
* $M_{i+1} = (X\bullet, n, L)$
* $M_l = (X_l, j_l, L)$ where X_l is obtained from X by replacing the $\bullet$ located at j_l is replaced by $\bullet\circ$ with $i+2 \leq l \leq n$.

One can check that $T(M_i) = M_{i+1}$, $0 \leq i \leq n-1$ and $T(M_n) = M_0$. Moreover, the definition of the weight of marked configurations implies that $W(M_i) = W(X)$ for $i = 0 \ldots n$ and so the weight is invariant under T.

To illustrate this consider $X = \circ\bullet$. We obtain the configurations $\{(\circ\circ\bullet, 0, R), (\bullet\circ\bullet, 1, R), (\circ\bullet\bullet, L, 3), (\circ\bullet\circ, L, 2)\}$. One may verify that these form a 4-cycle under T and that all the configurations have weight equal to that of the configuration $\circ\bullet$. $\square$

4.2. **Enriching a Markov chain.** In order to show that the stationary state of M-PASEP chain C is simply related to that of the PASEP we use the following lemma:

Lemma 5. *Consider a Markov chain C_1 with a transition from a state a to a state b with probability r. We replace the arc $\overrightarrow{ab}$ by the subgraph $\mathcal{H}$ as shown in Figure 5 to create a new chain C_2. Let H be the set of vertices in $\mathcal{H} \setminus \{a, b\}$. If:*

* *the weighted sum over all directed spanning trees in $\mathcal{H}$ rooted at b is equal to r, and*
* *the weighted sum over all directed spanning forests of $\mathcal{H}$ that contain 2 components (one rooted at a and one rooted at b) is equal to 1,*

then $\Pr_\infty^{C_2}(x$ given that $x \notin H) = \Pr_\infty^{C_1}(x)$, where $\Pr_\infty^{C_i}(x)$ is the stationary state probability of finding chain C_i in state x.

This follows from the Markov-Tree Theorem [6] and can be proved by applying the Matrix-tree theorem to a transition matrix, though it can also be proved combinatorially (see [5]).

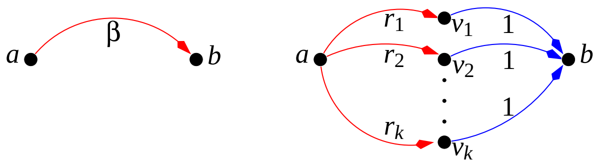


FIGURE 5. Replace the arc from a to b by a subgraph $\mathcal{H}$.

4.3. **The stationary distribution.** Using Lemma 5 and the definitions of the chains C and P, it is then easy to check that Proposition 4 holds. This shows that the two stationary distributions are simply related but does not tell us what they are. The following proposition does this and together with Proposition 4 implies Theorem 1.

Proposition 6. *Let Y be a state in the M-PASEP chain C, then:*

$$\mathrm{P}_\infty^C(Y \text{ given that } Y \in \mathcal{B}_n) = \frac{W(Y)}{Z_n}, \text{ and} \tag{17a}$$

$$\mathrm{P}_\infty^C(Y \text{ given that } Y \in \mathcal{M}_n) = \frac{W(Y)}{(n+1)Z_n}. \tag{17b}$$

where $W(Y)$ is the weight defined by the equations in Theorem 1 and Definition 4.

Proof. To prove stationarity we need to show that the above probabilities are unchanged under the action of the transitions. Denote the probability of finding the chain C in state Y at time t by $\Pr(C(t) = Y)$.

Now suppose that

$$\Pr(C(t) = Y \text{ given that } Y \in \mathcal{B}_n) = \frac{W(Y)}{Z_n}, \text{ and} \tag{18a}$$

$$\Pr(C(t) = Y \text{ given that } Y \in \mathcal{M}_n) = \frac{W(Y)}{(n+1)Z_n} \tag{18b}$$

Then because the chain alternates between basic and marked states we have

$$\begin{aligned}
\Pr(C(t+1) = Y \in \mathcal{B}_n) &= \sum_{M\,:\,U(T(M))=Y} \Pr(C(t) = M) \\
&= \sum_{M\,:\,U(T(M))=Y} \frac{1}{n+1}\frac{W(M)}{Z_n} \\
&= \sum_{M\,:\,U(T(M))=Y} \frac{1}{n+1}\frac{W(T(M))}{Z_n} \\
&= \sum_{M\,:\,U(M)=Y} \frac{1}{n+1}\frac{W(M)}{Z_n} \\
&= \frac{W(Y)}{Z_n}
\end{aligned} \tag{19}$$

Above we have used Lemma 4 and the fact that the bijection T is weight preserving, Proposition 5. Similarly we get:

$$\begin{aligned}
\Pr(C(t+1) = Y \in \mathcal{M}_n) &= \frac{W(M)}{(n+1)W(U(Y))}\Pr(C(t) = U(Y)) \\
&= \frac{W(Y)}{(n+1)W(U(Y))}\frac{W(U(Y))}{Z_n} = \frac{W(Y)}{(n+1)Z_n}.
\end{aligned} \tag{20}$$

This completes the proof. □

5. Weight-equations and lattice paths

In this section we prove Theorem 2 by showing that the weight of the paths in $\mathcal{P}(n)$ and the weight of the basic configurations satisfy the same equations. Let

$$Q(X) = \sum_{p\in\theta^{-1}(X)} w(p) \tag{21}$$

We work by showing that $Q(X)$ satisfies the same equations as $W(X)$.

Recall that given a path p, $\theta(p)$ is the basic configuration such that each $\overset{\circ}{E}$ and S step is changed to $\circ$ and each $\overset{\bullet}{E}$ and N step is changed to $\bullet$. Also, recall that N_i (resp. S_i, $\overset{\circ}{E}_i$, $\overset{\bullet}{E}_i$) is the weight of a North-East (resp. South East, East $\overset{\circ}{E}$, East $\overset{\bullet}{E}$) step starting at height i.

* If $X \in \mathcal{B}_0$ then $W(X) = 1$ and it follows that $Q(X) = 1$.
* If $X = \circ A$ then $W(X) = (1+c)W(A)$ by Theorem 1. The first step of a path $p \in \theta^{-1}(X)$, must be $\overset{\circ}{E}$ (it cannot start with a South-East step) and its weight is $1+c$. Removing this step gives:

$$Q(\circ A) = \sum_{\substack{p\in\mathcal{P}(n)\\ \theta(p)=\circ A}} w(p) = (1+c) \sum_{\substack{p'\in\mathcal{P}(n-1)\\ \theta(p')=A}} w(p') = (1+c)Q(A). \tag{22}$$

* Let $X = A\bullet$ then $W(X) = (1+d)W(A)$. The last step of a path $p \in \theta^{-1}(X)$, must be $\overset{\bullet}{E}$ (it cannot end with a North-East step) and its weight is $1+d$. Removing this step gives:

$$Q(A\bullet) = \sum_{\substack{p\in\mathcal{P}(n)\\ \theta(p)=A\bullet}} w(p) = (1+d) \sum_{\substack{p'\in\mathcal{P}(n-1)\\ \theta(p')=A}} w(p'). = (1+d)Q(A). \tag{23}$$

* Let $X = A \bullet \circ B$ where $A \in \mathcal{B}_{k-1}$, then $W(X) = W(A \bullet B) + W(A \circ B) + qW(A \circ \bullet B)$. There are four possibilities that we must consider — let p be a path such that $\theta(p) = X$ such that $(p_k, p_{k+1}) = (N, S)$, $(\overset{\bullet}{E}, \overset{\circ}{E})$, $(\overset{\bullet}{E}, S)$ or $(N, \overset{\circ}{E})$. Let $U = (p_1, \ldots p_{k-1})$ and $V = (p_{k+2}, \ldots, p_n)$.
  [i] Consider $\hat{p} = UNSV$ and $\bar{p} = U\overset{\bullet}{E}\overset{\circ}{E}V$. Form three new paths $p' = U\overset{\circ}{E}V$, $p'' = U\overset{\bullet}{E}V$ and $p''' = U\overset{\circ}{E}\overset{\bullet}{E}V$. Note that $\theta(p') = A \circ B$, $\theta(p'') = A \bullet B$ and $\theta(p''') = A \circ \bullet B$. The heights of vertices in U and V are the same in all of these paths and hence the weight contribution from U and V is the same in all of these paths and can be factored out. Hence the equation

$$w(\hat{p}) + w(\bar{p}) = w(p') + w(p'') + qw(p''') \tag{24}$$

  is true providing

$$N_iS_{i+1} + \overset{\bullet}{E}_i\overset{\circ}{E}_i = \overset{\circ}{E}_i + \overset{\bullet}{E}_i + q(S_iN_{i-1} + \overset{\circ}{E}_i\overset{\bullet}{E}_i) \tag{25}$$

  is true. The latter can be verified directly from the definition of the weights.
  [ii] Consider $p = UN\overset{\circ}{E}V$ and form $p' = UNV$ and $p'' = UENV$. Note that $\theta(p') = A \bullet B$ and $\theta(p'') = A \circ \bullet B$. Again the heights of vertices in U and V are the same in all of these paths and so the contribution of U and V to the weights of these paths may be factored out. Hence the equation

$$w(p) = w(p') + qw(p'') \tag{26}$$

  is true providing

$$N_i\overset{\circ}{E}_i + 1 = N_i + q\overset{\circ}{E}_iN_i \tag{27}$$

  is true, which can be verified from the definition of the weights.
  [iii] Consider $p = U\overset{\bullet}{E}SV$. This case follows similarly to the previous case.

Taking the above and summing over all A and B gives the following

$$\sum_{\substack{p\in\mathcal{P}(n)\\ \theta(p)=A\bullet\circ B}} w(p) = \sum_{\substack{p'\in\mathcal{P}(n-1)\\ \theta(p')=A\bullet B}} w(p') + \sum_{\substack{p'\in\mathcal{P}(n-1)\\ \theta(p')=A\circ B}} w(p') + q \sum_{\substack{p'\in\mathcal{P}(n)\\ \theta(p')=A\circ\bullet B}} w(p'). \tag{28}$$

Which is $Q(A \bullet \circ B) = Q(A \circ B) + Q(A \bullet B) + qQ(A \circ \bullet B)$.

□

References

[1] P. Flajolet, *Combinatorial aspects of continued fractions*, Disc. Math. 41 (1982), 145-153.

[2] R. A. Blythe, M. R. Evans, F. Colaiori, F. H. L. Essler, *Exact solution of a partially asymmetric exclusion model using a deformed oscillator algebra*, J. Phys. A: Math. Gen. 33, 2313–2332 (2000). `arxiv cond-mat/9910242`

[3] P. Biane, Permutations suivant le type d'excdance et le nombre d'inversions et interprtation combinatoire d'une fraction continue de Heine. European J. Combin. 14, no. 4, 277–284 (1993).

[4] R. Brak, J. Essam, *Asymmetric exclusion model and weighted lattice paths*, J. Phys. A 37 no. 14, 4183–4217 (2004). `arxiv: cond-mat/0311153`.

[5] R. Brak, S. Corteel, J. Essam, A. Rechnitzer, *A combinatorial derivation of the PASEP algebra*, long version , in preparation.

[6] S. Chaiken, *A combinatorial proof of the all minors matrix tree theorem*, SIAM J. Alg. Disc. Meth. 3 (1982), 319–329.

[7] A. Claesson and T. Mansour, *Counting occurences of a pattern of type (1,2) or (2,1) in permutations*, Adv in Appl. Math, 29:293-310 (2002).

[8] R. Clarke, E. Steingrimsson, J. Zeng, *New Euler-Mahonian statistics on permutations and words*, Adv in Appl. Math, 18(3):237-270 (1997).

[9] B. Derrida, E. Domany, D. Mukamel, *An exact solution of the one dimensional asymmetric exclusion model with open boundaries* J. Stat. Phys. 69, 667-687 (1992).

[10] B. Derrida, M.R. Evans, *Exact correlation functions in an asymmetric exclusion model with open boundaries* J. Phys. I France 3, 311-322 (1993).

[11] B. Derrida, M.R. Evans, V. Hakim, V. Pasquier, *Exact solution of a 1d asymmetric exclusion model using a matrix formulation* J. Phys. A26, 1493-1517 (1993).

[12] B. Derrida, C. Enaud, J.L. Lebowitz, *The asymmetric exclusion process and Brownian excursions* J. Stat. Phys. 115, 365-382 (2004)

[13] E. Duchi and G. Schaeffer, *Jumping particles I: maximal flow regime*, 16th International Conference on Formal Power Series and Algebraic Combinatorics, (FPSAC'04), Vancouver (Canada).

[14] E. Duchi, G. Schaeffer, *Jumping particles II: general boundary conditions*, International Colloquium on Mathematics and Computer Science: trees, algorithms and probability (MathInfo'2004), Vienna (Austria).

[15] E. Duchi, G. Schaeffer, *A combinatorial approach to jumping particles*, `arxiv math.CO/0407519`.

[16] M. Uchiyama, T. Sasamoto, M. Wadati, *Asymmetric Simple Exclusion Process with Open Boundaries and Askey-Wilson Polynomials*, J. Phys. A 37 (2004), no. 18, 4985–5002. `arxiv cond-mat/0312457`.

[17] G. Viennot, *Une théorie combinatoire des polynômes orthogonaux généraux*, Notes de Cours, UQAM, 1983.

Enumeration of Sequences Constrained by the Ratio of Consecutive Parts

Sylvie Corteel [*] Sunyong Lee [†] Carla D. Savage [‡]

November 13, 2004; Revised March 3, 2005

Abstract

Recurrences are developed to enumerate any family of nonnegative integer sequences $\lambda = (\lambda_1, \ldots, \lambda_n)$ satisfying the constraints:

$$\frac{\lambda_1}{a_1} \geq \frac{\lambda_2}{a_2} \geq \cdots \geq \frac{\lambda_{n-1}}{a_{n-1}} \geq \frac{\lambda_n}{a_n} \geq 0,$$

for a given constraint sequence $\mathbf{a} = [a_1, \ldots, a_n]$ of positive integers. They are applied to derive new counting formulas, to reveal new relationships between families, and to give simple proofs of the truncated lecture hall and anti-lecture hall theorems.

Nous développons des récurrences pour énumérer des familles de suites d'entiers $\lambda = (\lambda_1, \ldots, \lambda_n)$ satisfaisant les contraintes

$$\frac{\lambda_1}{a_1} \geq \frac{\lambda_2}{a_2} \geq \cdots \geq \frac{\lambda_{n-1}}{a_{n-1}} \geq \frac{\lambda_n}{a_n} \geq 0,$$

pour une suite d'entiers positifs donnée $a = [a_1, \ldots, a_n]$. Ces récurrences permettent de dériver de nouvelles formules dénumération, de révéler de nouvelles relations entre certaines familles, et de donner des preuves simples des théorèmes des partitions Lecture Hall tronquées et des compositions Lecture Hall tronquées.

1 Introduction

We consider the problem of enumerating nonnegative integer sequences $\lambda = (\lambda_1, \ldots, \lambda_n)$ satisfying the constraints:

$$\frac{\lambda_1}{a_1} \geq \frac{\lambda_2}{a_2} \geq \cdots \geq \frac{\lambda_{n-1}}{a_{n-1}} \geq \frac{\lambda_n}{a_n} \geq 0, \tag{1}$$

*CNRS PRiSM, UVSQ, 45 Avenue des Etats-Unis, 78035 Versailles, France. (syl@prism.uvsq.fr)
†Computer Science, North Carolina State University, Box 8206, Raleigh, NC 27695, USA. (slee7@unity.ncsu.edu) Research supported in part by NSF grant DMS-0300034.
‡Computer Science, North Carolina State University, Box 8206, Raleigh, NC 27695, USA. (savage@csc.ncsu.edu) Research supported in part by NSF grants DMS-0300034 and INT-0230800.

for a given constraint sequence $\mathbf{a} = [a_1, \ldots, a_n]$ of positive integers.

We refer to a sequence $\lambda = (\lambda_1, \ldots, \lambda_n)$ of nonnegative integers as a *composition* into n nonnegative parts. If the parts of λ are nonincreasing, then λ is a *partition.* Partitions and compositions are commonly defined by the set of parts allowed, the number of occurrences of a part, or the difference between consecutive parts. In contrast, the compositions satisfying (1) are constrained by the ratio of consecutive parts and we refer to them as *ratio compositions.* Generating functions are known for ratio compositions only for some special constraint sequences, $\mathbf{a}$, including:

$\mathbf{a} = [1, 1, \ldots, 1]$: ordinary partitions [1];

$\mathbf{a} = [1, 2, 4, \ldots, 2^{n-1}]$: Cayley compositions [8, 2, 14, 4];

$\mathbf{a} = [r^{n-1}, r^{n-2}, \ldots, r, 1]$: Hickerson partitions [13];

$\mathbf{a} = [n, n-1, \ldots, 1]$: lecture hall partitions [5, 6, 7, 15, 16, 3];

$\mathbf{a} = [1, 2, \ldots, n]$: anti-lecture hall compositions [9];

$\mathbf{a} = [1, 2, 1, 2, ..., 2-(-1)^n]$: one-two compositions [11];

$\mathbf{a} = [n, n-1 \ldots, n-t+1]$: truncated lecture hall partitions [10];

$\mathbf{a} = [n-t+1, n-t, \ldots, n]$: truncated anti-lecture hall compositions [10].

In this paper, we introduce a common approach for the enumeration of ratio compositions by using as a statistic a bound on the size of the first part. This generalizes the enumeration of ordinary partitions via the Gaussian polynomials.

In Section 2, we derive a recurrence for the generating function of *any* family of ratio compositions with first part *bounded.* We use this to derive new counting formulas and their q-analogs. Among these, we discover a family of polynomials with several interesting properties which arise in the enumeration of lecture hall partitions. In addition, we find a functional relationship between the generating functions for the ratio compositions constrained by a sequence $[a_1, a_2, \ldots, a_n]$ and those constrained by its reverse sequence $[a_n, a_{n-1} \ldots, a_1]$. This reveals for the first time the relationship between, e.g., lecture hall partitions and anti-lecture hall compositions and between Hickerson partitions and Cayley compositions.

In Section 3, we derive a different recurrence for the enumeration of ratio compositions with first part bounded. By allowing the bound to approach infinity, we get a recurrence for the generating function of *any* family of ratio compositions with first part *unbounded.* As one consequence, we discover new "easy" proofs of the (truncated) lecture hall and (truncated) anti-lecture hall theorems. In contrast to earlier proofs, where deriving a recurrence was a challenge, here the recurrence is generic and the work is moved entirely to "standard" q-series manipulation in an induction proof.

2 Enumeration of ratio compositions with first part bounded

To build a recurrence, we first consider the case where the constraint sequence **a** in (1) satisfies $[a_1, \ldots, a_n] = [s_n, s_{n-1}, \ldots, s_1]$ for an infinite sequence of positive integers $\{s_i\}$. For $n \geq 0$, let S_n be the set of compositions $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ satisfying

$$\frac{\lambda_1}{s_n} \geq \frac{\lambda_2}{s_{n-1}} \geq \cdots \geq \frac{\lambda_{n-1}}{s_2} \geq \frac{\lambda_n}{s_1} \geq 0. \tag{2}$$

Let $S_n^{(j,i)}$ be the set of $\lambda \in S_n$ with $\lambda_1 \leq js_n + i$ and let

$$S_n^{(j,i)}(q) = \sum_{\lambda \in S_n^{(j,i)}} q^{|\lambda|}.$$

Theorem 1 *For $n \geq 0$, $j \geq 0$, and $0 \leq i \leq s_n$,*

$$S_n^{(j,i)}(q) = \begin{cases} 1 & \textit{if } n = 0 \textit{ or } j = i = 0, \textit{ else} \\ S_n^{(j-1,s_n)}(q) & \textit{if } i = 0, \textit{ else} \\ S_n^{(j,i-1)}(q) + q^{js_n+i} S_{n-1}^{(j,\lfloor is_{n-1}/s_n \rfloor)}(q) & \textit{otherwise.} \end{cases}$$

Proof. The theorem is clearly true for $n = 0$ and for $j = i = 0$. Let (n, j, i) satisfy $n > 0$, $(j, i) > (0, 0)$. If $i = 0$, then $j > 0$ and $js_n + i = js_n = (j-1)s_n + s_n$, so the theorem is true. Assume, then, that $1 \leq i \leq s_n$. By definition, $\lambda \in S_n^{(j,i)}$ if and only if either $\lambda \in S_n^{(j,i-1)}$ or $\lambda \in S_n$ and $\lambda_1 = js_n + i$. But $(js_n + i, \lambda_2, \ldots, \lambda_n) \in S_n$ if and only if $(\lambda_2, \ldots, \lambda_n) \in S_{n-1}$ and $(js_n + i)/\lambda_2 \geq s_n/s_{n-1}$. That is,

$$\lambda_2 \leq \frac{s_{n-1}}{s_n}(js_n + i) = js_{n-1} + i\frac{s_{n-1}}{s_n}.$$

So, since λ_2 is an integer,

$$\lambda_2 \leq js_{n-1} + \lfloor i\frac{s_{n-1}}{s_n} \rfloor.$$

Note, since $1 \leq i \leq s_n$, $\lfloor is_{n-1}/s_n \rfloor \leq s_{n-1}$, so $(\lambda_2, \ldots, \lambda_n) \in S_{n-1}^{(j,\lfloor is_{n-1}/s_n \rfloor)}$. □

Remark 1. For ordinary partitions, P_n, into n nonnegative parts, $\{s_i\} = \{1\}$ and $\lfloor is_{n-1}/s_n \rfloor = i$, so the recurrence of Theorem 1 reduces to the recurrence

$$P_n^{(j,i)}(q) = P_n^{(j,i-1)}(q) + q^{j+i} P_{n-1}^{(j,i)}(q),$$

the familiar recurrence for Gaussian polynomials.

The *lecture hall partitions* [5], L_n, are those compositions, necessarily partitions, satisfying

$$\frac{\lambda_1}{n} \geq \frac{\lambda_2}{n-1} \geq \cdots \geq \frac{\lambda_{n-1}}{2} \geq \frac{\lambda_n}{1} \geq 0. \tag{3}$$

Then $L_n = S_n$ with $\{s_i\} = \{i\}$ in (2). Since $\lfloor i(n-1)/n \rfloor = i - 1$ we get the following from Theorem 1.

Corollary 1 *For $n \geq 0$, $j \geq 0$, and $0 \leq i \leq n$, let $L_n^{(j,i)}(q)$ be the generating function for the lecture hall partitions $\lambda \in L_n$ with $\lambda_1 \leq jn + i$.*

$$L_n^{(j,i)} = \begin{cases} 1 & \text{if } n = 0 \text{ or } j = i = 0, \text{ else} \\ L_n^{(j-1,n)}(q) & \text{if } i = 0, \text{ else} \\ L_n^{(j,i-1)}(q) + q^{jn+i} L_{n-1}^{(j,i-1)}(q) & \text{otherwise.} \end{cases}$$

For fixed $n > 0$, any $t \geq 0$ can be written uniquely in the form $t = jn + i$, where $0 \leq i < n$. So, we get a nice counting formula for lecture hall partitions with largest part at most t.

Theorem 2 *For $n \geq 0$, $j \geq 0$, and $0 \leq i \leq n$, the number of lecture hall partitions in L_n with first part bounded by $jn + i$ is*

$$L_n^{(j,i)} = (j+1)^{n-i}(j+2)^i.$$

Proof. If $n = 0$ then $i = 0$, so $(j+1)^{0-0}(j+2)^0 = 1$. If $i = j = 0$, then $(1)^{n-0}(2)^0 = 1$. Let (n, j, i) satisfy $n > 0$, $(j, i) > (0, 0)$ and assume the theorem is true for $(n', i', j') < (n, i, j)$. If $i = 0$, then $j > 0$ and by Corollary 1, $L_n^{(j,0)}(1) = L_n^{(j-1,n)}(1)$, which, by induction, is $(j)^{n-n}(j+1)^n = (j+1)^n(j+2)^0$. Otherwise, by Corollary 1,

$$\begin{aligned} L_n^{(j,i)}(1) &= L_n^{(j,i-1)}(1) + L_{n-1}^{(j,i-1)}(1) \\ &= (j+1)^{n-i+1}(j+2)^{i-1} + (j+1)^{n-i}(j+2)^{i-1} \\ &= (j+1)^{n-i}(j+2)^i. \end{aligned}$$

□

In [5], it was shown that the generating function for the lecture hall partitions, L_n is:

$$L_n(q) = \frac{1}{(q; q^2)_n}, \tag{4}$$

where $(a; q)_n = (1-a)(1-aq)\cdots(1-aq^{n-1})$. Let D be the set of partitions into distinct parts and let O be the set of partitions into odd parts. The sets D and O have generating functions $D(q) = (-q; q)_\infty$ and $O(q) = (q; q^2)_\infty^{-1}$, respectively. Since $\lim_{n\to\infty} L_n = D$, the Lecture Hall Theorem (4) is a finite version of Euler's Theorem which says that $D(q) = O(q)$. The polynomial $L_n^{(j)}(q) = L_n^{(j,0)}(q)$ can be viewed as a q-analog of $(j+1)^n$ that encapsulates a further finitization of Euler's Theorem in the following sense.

Corollary 2 *The lecture hall polynomials $L_n^{(j)}(q)$ satisfy*

(i) $L_n^{(j)}(1) = (j+1)^n$,

(ii) $\lim_{n\to\infty} L_n^{(j)}(q) = (-q; q)_\infty$, *and*

(iii) $\lim_{j\to\infty} L_n^{(j)}(q) = (q; q^2)_n^{-1}$.

Proof. The first equation follows from Theorem 2. The second and third follow from the observations that $\lim_{n\to\infty} L_n^{(j)} = D$ and $\lim_{j\to\infty} L_n^{(j)} = L_n$. □

The *anti-lecture hall compositions* [9], A_n, are those sequences satisfying

$$\frac{\lambda_1}{1} \geq \frac{\lambda_2}{2} \geq \cdots \geq \frac{\lambda_{n-1}}{n-1} \geq \frac{\lambda_n}{n} \geq 0. \tag{5}$$

It was shown in [9] that A_n has generating function $A_n(q) = (-q, q)_n/(q^2; q)_n$. The constraint sequence for A_n, in the sense of (1), is $[1, 2, \ldots, n]$, the *reverse* of the constraint sequence $[n, n-1, \ldots, 1]$ for L_n. We introduce some notation to describe the relationship between their generating functions. Let $S[a_1, a_2, \ldots, a_k]$ be the set of compositions $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ satisfying

$$\frac{\lambda_1}{a_1} \geq \frac{\lambda_2}{a_2} \geq \cdots \geq \frac{\lambda_{n-1}}{a_{n-1}} \geq \frac{\lambda_n}{a_n} \geq 0, \tag{6}$$

with $S^{(j,i)}[a_1, a_2, \ldots, a_n]$ denoting those with $\lambda_1 \leq j a_1 + i$ and let $S^{(j)}[a_1, a_2, \ldots, a_n] = S^{(j,0)}[a_1, a_2, \ldots, a_n]$.

Theorem 3 *The generating functions for $S^{(j)}[a_1, a_2, \ldots, a_n]$ and $S^{(j)}[a_n, a_{n-1}, \ldots, a_1]$ satisfy:*

$$S^{(j)}[a_1, a_2, \ldots, a_n](q) = q^{j(a_1+a_2+\cdots+a_n)} S^{(j)}[a_n, a_{n-1}, \ldots, a_1](1/q).$$

Proof. The result follows if we show that $\lambda \in S^{(j)}[a_n, a_{n-1}, \ldots, a_1]$ if and only if $\mu \in S^{(j)}[a_1, a_2, \ldots, a_n]$, where μ is defined by $\mu_i = j s_i - \lambda_{n+1-i}$.

So, assume $\lambda \in S^{(j)}[a_n, a_{n-1}, \ldots, a_1]$, that is, $\lambda_1 \leq j a_n$ and $a_i \lambda_{n-i} \geq a_{i+1} \lambda_{n+1-i}$ for $1 \leq i \leq n$. Then for $1 \leq i \leq n$,

$$\lambda_{n+1-i} \leq \frac{a_i}{a_{i+1}} \frac{a_{i+1}}{a_{i+2}} \cdots \frac{n-1}{n} nj = a_i j,$$

so $\mu_i = j a_i - \lambda_{n+1-i} \geq 0$. Also, $\mu_1 = j a_1 - \lambda_n$ satisfies $\mu_1 \leq j a_1$. To show $\mu \in S^{(j)}[a_1, a_2, \ldots, a_n]$, it remains to show $a_{i+1}\mu_i \geq a_i \mu_{i+1}$:

$$a_{i+1}\mu_i = a_{i+1}(j a_i - \lambda_{n+1-i}) = j a_i a_{i+1} - a_{i+1}\lambda_{n+1-i} \geq j a_i a_{i+1} - a_i \lambda_{n-i} = a_i \mu_{i+1}.$$

The converse is similar. □

Remark 2. The proof of Theorem 3 also shows that for $1 \leq i \leq n$, $\lambda \in S^{(j)}[a_n, a_{n-1}, \ldots, a_1]$ if and only if $\mu \in S^{(j+1)}[a_1, a_2, \ldots, a_n]$ and $\mu_n \geq a_n - i$.

Corollary 3 *Lecture hall partitions, $L_n^{(j)}$, with first part bounded by jn and anti-lecture hall compositions, $A_n^{(j)}$ with first part bounded by j have the following relationship:*

$$A_n^{(j)}(q) = q^{jn(n+1)/2} L_n^{(j)}(1/q).$$

Proof. Observe that $A_n^{(j)} = S^{(j)}[1, 2, \ldots, n]$ and $L_n^{(j)} = S^{(j)}[n, n-1, \ldots, 1]$ and apply Theorem 3. □

This gives a counting formula for anti-lecture hall compositions.

Corollary 4 *The number of anti-lecture hall compositions in A_n with first part bounded by j is*

$$A_n^{(j)}(1) = (j+1)^n.$$

Proof. By Theorem 3, $A_n^{(j)}(1) = L_n^{(j)}(1)$. By definition, $L_n^{(j)}(q) = L_n^{(j,0)}(q)$, and by Corollary 2, $L_n^{(j,0)}(1) = (j+1)^n(j+2)^0$. □

As another example of the application of Theorems 1 and 3, we consider Hickerson partitions H_n (for $r = 2$) and Cayley compositions, C_n. H_n is the set of compositions into n nonnegative parts satisfying $\lambda_i \geq 2\lambda_{i+1}$ and C_n is the set of compositions into n nonnegative parts satisfying $\lambda_i \geq \lambda_{i+1}/2$. So, $H_n = S[2^{n-1}, 2^{n-2}, \ldots, 1]$ and $C_n = S[1, 2, 4, \ldots, 2^{n-1}]$. Let $B(n)$ be the number of *binary partitions* of n, i.e., the number of partitions of n into powers of 2. It is easy to check that $B(0) = B(1) = 1$, $B(2n) = B(2n-2) + B(n)$, and $B(2n) = B(2n+1)$.

Theorem 4 *For $0 \leq i < 2^{n-1}$, the number of Hickerson partitions with first part at most i is $H_n^{(0,i)} = B(2i)$; with first part at most $2^{n-1} + i$ is $H_n^{(1,i)} = B(2^n + 2i)$; with first part at most 2^n is $H_n^{(2,0)} = B(2^{n+1}) - 1$.*

Proof. Use the recurrence of Theorem 1 with the observation that since $\{s_i\} = \{2^{i-1}\}$ for Hickerson partitions, $\lfloor is_{i-1}/s_i \rfloor = \lfloor i/2 \rfloor$. The theorem follows by induction using the properties of $B(n)$. □

Cayley's Theorem [8] says that the number of compositions in C_n with n *positive* parts and with first part 1 is equal to the number of partitions of $2^{n-1} - 1$ into parts from the set $\{1, 1', 2, 4, \ldots, 2^{n-2}\}$. If we apply Theorem 3 and Remark 2, we get a generalization and reformulation of Cayley's Theorem.

Theorem 5 *For $0 \leq i < 2^{n-1}$, the number of Cayley compositions into n positive parts with first part 1 and last part at least $2^{n-1} - i$ is $B(2i)$. The number with first part at most 2 and last part at least $2^n - i$ is $B(2^n + 2i)$.*

Setting $i = 2^{n-1} - 1$ in Theorem 5 gives:

Corollary 5 *The number of Cayley compositions into n positive parts with first part 1 is $B(2^n - 2)$; with first part at most 2 is $B(2^{n+1} - 2)$.*

These results can be generalized to $r-ary$ Hickerson partitions and Cayley compositions. For other families of ratio compositions, we can expect Theorem 1 to be most useful for sequences $\{s_i\}$ where $\lfloor is_{n-1}/s_n \rfloor$ has a nice form.

As t gets larger, solving for $H_n^{(t)}$ gets harder. This is in spite of the fact that H_n has the nice generating function $H_n(q) = \prod_{t=1}^{n}(1 - q^{2^t-1})^{-1}$ [13]. However, we will see in the next section that if $S_n^{(j)}(q)$ has a nice generating function when $j = 1$, there is hope that $\lim_{j\to\infty} S_n^{(j)}(q)$ will also.

In the next section we show how to get a recurrence for the generating function of ratio compositions when the first part *unrestricted.*

3 Enumeration of ratio sequences with first part unbounded.

We define two slight variations of the set $S^{(j)}[a_1, a_2, \ldots, a_n]$ below:

$$P^{(j)}[a_1, a_2, \ldots, a_n] = \{\lambda \in S^{(j)}[a_1, a_2, \ldots, a_n] \mid \lambda_n \geq 1\};$$

$$R^{(j)}[a_1, a_2, \ldots, a_n] = \{\lambda \in S^{(j)}[a_1, a_2, \ldots, a_n] \mid \lambda_1 < ja_1\}.$$

In $P^{(j)}$, all parts must be positive, whereas in $R^{(j)}$ parts can be nonnegative, but the bound on the first part becomes strict.

Theorem 6 *For $j \geq 1$,*

$$P^{(j)}[a_1, a_2, \ldots, a_k](q) = \sum_{t=0}^{k} q^{(a_1+a_2+\cdots+a_t)} P^{(1)}[a_{t+1}, \ldots, a_k](q) P^{(j-1)}[a_1, \ldots, a_t](q).$$

Proof. Let $\lambda \in P^{(j)}[a_1, a_2, \ldots, a_k]$ and let t be the maximum index such that $\lambda_t > a_t$. Then $(\lambda_1 - a_1, \lambda_2 - a_2, \ldots, \lambda_t - a_t) \in P^{(j-1)}[a_1, \ldots, a_t]$ and $(\lambda_{t+1}, \ldots, \lambda_k) \in P^{(1)}[a_{t+1}, \ldots, a_k]$. □

Note that $P^{(1)}(q)$ can be computed using Theorem 1 as

$$P^{(1)}[s_n, \ldots, s_1](q) = S_n^{(1,0)}[s_1, \ldots, s_n](q) - S_{n-1}^{(1,0)}[s_2, \ldots, s_n](q).$$

Thus, taking the limit as $j \to \infty$ in Theorem 6 gives a recurrence for counting the sequences λ in $P[a_1, a_2, \ldots, a_k]$ without a restriction on the size of the first part.

Theorem 7 *For $n \geq k$,*

$$P[a_1, a_2, \ldots, a_k](q) = \sum_{t=0}^{k} q^{(a_1+a_2+\cdots+a_t)} P^{(1)}[a_{t+1}, \ldots, a_k](q) P[a_1, a_2, \ldots, a_t](q).$$

Remark 3. Theorem 6 and its proof are valid with all occurrences of P replaced by R. We need only change the statement "let t be the maximum index such that $\lambda_t > a_t$" to "let t be the maximum index such that $\lambda_t \geq a_{n-t+1}$". Similarly, Theorem 7 holds with P replaced by R.

Theorem 7 can be used to find an explicit form for the generating function in families where $P^{(1)}[a_1,\ldots,a_k]$ (or $R^{(1)}[a_1,\ldots,a_k]$) is known.

Remark 4. For ordinary partitions into k positive parts, $a_i = 1$, so $P[a_1,a_2,\ldots,a_k](q) = (q;q)_k^{-1}$ and $P^{(1)}[a_{t+1},\ldots,a_k](q) = q^{k-t}$, and the recurrence of Theorem 7 becomes

$$\frac{1}{(q;q)_k} = \sum_{t=0}^{k} \frac{q^k}{(q;q)_t},$$

which counts partitions into k positive parts by summing over the number, t, of parts greater than 1. □

The partitions $L_{n,k} = P[n, n-1, \ldots, n-k+1]$ are called *truncated lecture hall partitions* with all parts positive. The compositions $A_{n,k} = R[n-k+1, n-k+2, \ldots, n]$ are called *truncated anti-lecture hall compositions* with nonnegative parts. These were introduced in [10] where their generating functions were shown to be

$$L_{n,k}(q) = q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q \frac{(-q^{n-k+1};q)_k}{(q^{2n-k+1};q)_k}, \tag{7}$$

$$A_{n,k}(q) = \begin{bmatrix} n \\ k \end{bmatrix}_q \frac{(-q^{n-k+1};q)_k}{(q^{2(n-k+1)};q)_k}. \tag{8}$$

We now show how Theorem 7 can be used to give simple proofs of the (Truncated) Lecture Hall and (Truncated) Anti-lecture Hall Theorems. We begin by computing $P^{(1)}[n, n-1, \ldots, n-k+1](q)$ for $L_{n,k}$ and $R^{(1)}[n-k+1, n-k+2, \ldots, n](q)$ for $A_{n,k}$

Lemma 1 *For $n \geq k \geq 1$, the generating function for truncated lecture hall partitions into k positive parts with first part less than or equal to n is*

$$P^{(1)}[n, n-1, \ldots, n-k+1](q) = q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

Proof. We show $P^{(1)}[n, n-1, \ldots, n-k+1]$ is the set of partitions into k distinct parts from $\{1, 2, \ldots, n\}$, which has the generating function claimed. If $\lambda \in P^{(1)}[n, n-1, \ldots, n-k+1]$, $1 \geq \lambda_1/n \geq \lambda_2/(n-1) \geq \cdots \geq \lambda_k/(n-k+1) > 0$, so the parts of λ are distinct and bounded by n. Conversely, if $\lambda_i \geq \lambda_{i+1} + 1$ for $1 \leq i \leq n-1$ and $\lambda_1 \leq n$, then $\lambda_{i+1} \leq n - i$, so

$$\lambda_i(n-i) \geq (\lambda_{i+1}+1)(n-i) \geq \lambda_{i+1}(n-i) + n - i \geq \lambda_{i+1}(n-i) + \lambda_{i+1} \geq \lambda_{i+1}(n-i+1),$$

that is, $\lambda_i/(n-i+1) \geq \lambda_{i+1}/(n-i)$, so $\lambda \in P^{(1)}[n, n-1, \ldots, n-k+1]$. □

Lemma 2 *For $n \geq k \geq 1$, the generating function for truncated anti-lecture hall compositions into k nonnegative parts with first part less than $n-k+1$ is*

$$R^{(1)}[n-k+1, n-k+2, \ldots, n](q) = \begin{bmatrix} n \\ k \end{bmatrix}_q.$$

Proof. If $\lambda \in R^{(1)}[n-k+1, n-k+2, \ldots, n]$, then for $1 \le i \le n$, $\lambda_i < n-k+i$ and $\lambda_{i-1} \ge \lambda_i(n-k+i-1)/(n-k+i)$, so

$$\lambda_{i-1} - \lambda_i \ge \lambda_i((n-k+i-1)/(n-k+i) - 1) = -\lambda_i/(n-k+i) > -1.$$

Thus, $\lambda_{i-1} - \lambda_i \ge 0$ and λ is a partition into at most k parts of size at most $n-k$. Conversely, any such partition is in $R^{(1)}[n-k+1, n-k+2, \ldots, n]$. □

Corollary 6

$$|L_{n,k}^{(j)}| = (j)^k \binom{n}{k} = |A_{n,k}^{(j)}|.$$

Proof. By induction. It is obviously true for $j = 0$. Then for $j > 0$, by Theorem 6 and Lemma 2,

$$L_{n,k}^{(j)}(1) = \sum_{t=0}^{k} \binom{n-t}{k-t}\binom{n}{t}(j-1)^t.$$

We use the classical identity $\binom{n-t}{k-t}\binom{n}{t} = \binom{n}{k}\binom{k}{t}$ and the binomial theorem and get the result. By Theorem 3, $L_{n,k}^{(j)}(1) = A_{n,k}^{(j)}(1)$. □

Lemma 3 *The generating function for truncated lecture hall partitions satisfies*

$$L_{n,k}(q) = q^{\binom{k+1}{2}} \sum_{t=0}^{k} q^{(n-k)t} \begin{bmatrix} n-t \\ k-t \end{bmatrix}_q L_{n,t}(q),$$

for $k \ge 1$ *and* $L_{n,0}(q) = 1$.

Proof. Since $L_{n,k} = P[n, n-1, \ldots, n-k+1]$, apply Theorem 7 with Lemma 1 to get

$$L_{n,k}(q) = \sum_{t=0}^{k} q^{n+(n-1)+\cdots+(n-t+1)} q^{\binom{k-t+1}{2}} \begin{bmatrix} n-t \\ k-t \end{bmatrix}_q L_{n,t}(q)$$

and simplify. □

This gives the Truncated Lecture Hall Theorem:

Theorem 8 [10]

$$L_{n,k}(q) = q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q \frac{(-q^{n-k+1};q)_k}{(q^{2n-k+1};q)_k}.$$

Proof. Let

$$\mathcal{L}_{n,k}(q) = \frac{L_{n,k}(q)}{q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q}.$$

We show

$$\mathcal{L}_{n,k}(q) = \frac{(-q^{n-k+1};q)_k}{(q^{2n-k+1};q)_k}. \tag{9}$$

Substituting for $L_{n,k}(q)$ and $L_{n,t}(q)$ in the recurrence of Lemma 3 and simplifying gives $\mathcal{L}_{n,1} = 1$ and for $k > 1$,

$$\mathcal{L}_{n,k}(q) = \sum_{t=0}^{k} q^{(n-k)t+\binom{t+1}{2}} \begin{bmatrix} k \\ t \end{bmatrix}_q \mathcal{L}_{n,t}(q), \tag{10}$$

as

$$\begin{bmatrix} n-t \\ k-t \end{bmatrix}_q \begin{bmatrix} n \\ t \end{bmatrix}_q = \begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} k \\ t \end{bmatrix}_q.$$

We make use of (a transformation of) one of the q-Chu Vandermonde identities [1.5.2] from [12]:

$$\frac{(-aq;q)_k}{(cq;q)_k} = \sum_{t=0}^{k} a^t q^{\binom{t+1}{2}} \begin{bmatrix} k \\ t \end{bmatrix}_q \frac{(-(c/a)q^{-t+1};q)_t}{(cq^{k-t+1};q)_t}.$$

We do the substitution $a = q^{n-k}$ and $c = q^{2n-k}$ and get

$$\frac{(-q^{n-k+1};q)_k}{(q^{2n-k+1};q)_k} = \sum_{t=0}^{k} q^{t(n-k)+\binom{t+1}{2}} \begin{bmatrix} k \\ t \end{bmatrix}_q \frac{(-q^{n-t+1};q)_t}{(q^{2n-t+1};q)_t},$$

which shows that $\mathcal{L}_{n,k}(q)$ as given by (9) is the solution to the recurrence (10).

From this we also get the Lecture Hall Theorem:

Theorem 9 [5]

$$L_n(q) = \frac{(-q;q)_n}{(q^{n+1};q)_n} = \frac{1}{(q;q^2)_n}.$$

Proof. Setting $k = n$ in Theorem 8 gives $L_{n,n}$, which is the set of partitions in L_n with all parts positive. So, $L_{n,n}(q) = q^{\binom{n+1}{2}} L_n(q)$. □

The approach for truncated anti-lecture hall compositions is similar.

Lemma 4 *For $n \geq k$,*

$$A_{n,k}(q) = \sum_{t=0}^{k} q^{(n-k)t+\binom{t+1}{2}} \begin{bmatrix} n \\ k-t \end{bmatrix}_q A_{n-k+t,t}(q)$$

with $A^{(0)}_{n,k}(q) = 1$.

Proof. Since $A_{n,k} = R[n-k+1, n-k+2, \ldots, n]$, apply Theorem 7, using Remark 3, then Lemma 2 and simplify. □

Now we get an easy proof of the Truncated Anti-lecture Hall Theorem.

Theorem 10 [10]

$$A_{n,k}(q) = \begin{bmatrix} n \\ k \end{bmatrix}_q \frac{(-q^{n-k+1};q)_k}{(q^{2(n-k+1)};q)_k}.$$

Proof. Let

$$\mathcal{A}_{n,k}(q) = \frac{A_{n,k}(q)}{\begin{bmatrix} n \\ k \end{bmatrix}_q}.$$

Then using Lemma 4, we get that

$$\mathcal{A}_{n,k}(q) = \sum_{t=0}^{k} q^{(n-k)t+\binom{t+1}{2}} \begin{bmatrix} k \\ t \end{bmatrix}_q \mathcal{A}_{n-k+t,t}(q), \tag{11}$$

as

$$\begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} k \\ k-t \end{bmatrix}_q = \begin{bmatrix} n-k+t \\ t \end{bmatrix}_q \begin{bmatrix} n \\ k-t \end{bmatrix}_q.$$

We make use of another transformation of the q-Chu Vandermonde identity [1.5.2] from [12]:

$$\frac{(-c/a;q)_k}{(c;q)_k} = \sum_{t=0}^{k} (c/a)^t q^{\binom{t}{2}} \begin{bmatrix} k \\ t \end{bmatrix}_q \frac{(-a;q)_t}{(c;q)_t}.$$

Set $c = q^{2(n-k+1)}$ and $a = q^{n-k+1}$ and get

$$\frac{(-q^{n-k+1};q)_k}{(q^{2(n-k+1)};q)_k} = \sum_{t=0}^{k} q^{(n-k)t+\binom{t+1}{2}} \begin{bmatrix} k \\ t \end{bmatrix}_q \frac{(-q^{n-k+1};q)_k}{(q^{2(n-k+1)};q)_k},$$

showing $\mathcal{A}_{n,k}(q) = \frac{(-q^{n-k+1};q)_k}{(q^{2(n-k+1)};q)_k}$ satisfies the recurrence. □

Setting $k = n$ in Theorem 10 gives the Anti-Lecture Hall Theorem.

Theorem 11 [9]

$$A_n(q) = \frac{(-q;q)_n}{(q^2;q)_n}.$$

4 Concluding Remarks

The recurrences of Theorems 1 and 7 provide simple computational tools to investigate any family of ratio compositions. More importantly, they supply the foundation for an inductive proof of any conjectured enumeration result. Our experiments suggest that counting formulas and generating functions will be possible for many other families of ratio compositions.

One particular question of interest is to characterize the subfamily of partitions into odd parts in $\{1, 3, \ldots, 2n-1\}$ that is counted by the polynomial $L_n^{(j)}(q)$.

References

[1] George E. Andrews. *The theory of partitions*. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2.

[2] George E. Andrews. The Rogers-Ramanujan reciprocal and Minc's partition function. *Pacific J. Math.*, 95(2):251–256, 1981.

[3] George E. Andrews. MacMahon's partition analysis. I. The lecture hall partition theorem. In *Mathematical essays in honor of Gian-Carlo Rota (Cambridge, MA, 1996)*, volume 161 of *Progr. Math.*, pages 1–22. Birkhäuser Boston, Boston, MA, 1998.

[4] George E. Andrews, Peter Paule, Axel Riese, and Volker Strehl. MacMahon's partition analysis. V. Bijections, recursions, and magic squares. In *Algebraic combinatorics and applications (Gößweinstein, 1999)*, pages 1–39. Springer, Berlin, 2001.

[5] Mireille Bousquet-Mélou and Kimmo Eriksson. Lecture hall partitions. *Ramanujan J.*, 1(1):101–111, 1997.

[6] Mireille Bousquet-Mélou and Kimmo Eriksson. Lecture hall partitions II. *Ramanujan J.*, 1(2):165–185, 1997.

[7] Mireille Bousquet-Mélou and Kimmo Eriksson. A refinement of the lecture hall theorem. *J. Combin. Theory. Ser. A*, 86(1):63–84, 1999.

[8] A. Cayley. On a problem in the partition of numbers. *Philosophical Magazine*, 13:245–248, 1857. reprinted in: The Collected Mathematical Papers of A. Cayley, Vol. III, Cambridge University Press, Cambridge, 1890, 247-249).

[9] Sylvie Corteel and Carla D. Savage. Anti-lecture hall compositions. *Discrete Math.*, 263(1-3):275–280, 2003.

[10] Sylvie Corteel and Carla D. Savage. Lecture hall theorems, q-series and truncated objects. *J. Combin. Theory Ser. A*, 108(2):217–245, 2004.

[11] Sylvie Corteel and Carla D. Savage. Partitions and compositions defined by inequalities. *Ramanujan J.*, 8(3):357–381, 2004.

[12] George Gasper and Mizan Rahman. *Basic hypergeometric series*, volume 35 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1990. With a foreword by Richard Askey.

[13] D. R. Hickerson. A partition identity of the Euler type. *Amer. Math. Monthly*, 81:627–629, 1974.

[14] H. Minc. A problem in partitions: Enumeration of elements of a given degree in the free commutative entropic cyclic groupoid. *Proc. Edinburgh Math. Soc. (2)*, 11:223–224, 1958/1959.

[15] Ae Ja Yee. On the combinatorics of lecture hall partitions. *Ramanujan J.*, 5(3):247–262, 2001.

[16] Ae Ja Yee. On the refined lecture hall theorem. *Discrete Math.*, 248(1-3):293–298, 2002.

RIBBON TABLEAUX, RIGGED CONFIGURATIONS AND HALL-LITTLEWOOD FUNCTIONS AT ROOTS OF UNITY

FRANÇOIS DESCOUENS

Abstract: Hall-Littlewood functions indexed by rectangular partitions, specialized at primitive roots of unity, can be expressed as plethysms. We propose a combinatorial proof of this formula using Schilling's bijection between ribbon tableaux and rigged configurations [13].

Résumé: La spécialisation aux racines de l'unité des fonctions de Hall-Littlewood indexées par des partitions rectangulaires peut s'exprimer à l'aide de pléthysmes. On propose une preuve combinatoire de cette formule en utilisant la bijection de Schilling entre les tableaux de rubans et les configurations [13].

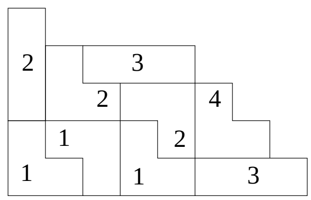
1. Introduction

In [7, 8], Lascoux, Leclerc and Thibon proved a formula for Hall-Littlewood functions, when the parameter is set to a root of unity.
This formula implies a combinatorial interpretation of the plethysms $l_k^{(j)}[h_\lambda]$ and $l_k^{(j)}[e_\lambda]$ where h_λ, e_λ are respectively products of complete and elementary symmetric functions, and $l_k^{(j)}$ the Frobenius characteristics of representations induced by a transitive cyclic subgroup of $\mathfrak{S}_k$.
However, the combinatorial interpretation of the plethysms of Schur functions $l_k^{(j)}[s_\lambda]$ would be far more interesting. This question led the same authors to introduce a new basis $H_\lambda^{(k)}(X;q)$ of symmetric functions, depending on an integer $k \geq 1$ and a parameter q, which interpolate between Schur functions ($k=1$) and Hall-Littlewood functions $Q'_\lambda(X;q)$ (for $k \geq l(\lambda)$). These were conjectured to behave similarly under specialization at root of unity, and to provide a combinatorial expression of the expansion of the plethysm $l_k^{(j)}[s_\lambda]$ in the Schur basis for suitable values of the parameters. This conjecture has been proved only in two cases: the stable case, which reduces to the previous result on Hall-Littlewood functions, and $k=2$ which gives the symmetric and antisymmetric squares $h_2[s_\lambda]$ and $e_2[s_\lambda]$.
The proof given in [1] relies upon the study of diagonal classes of domino tableaux, i.e. sets of domino tableaux having the same diagonals. Carré and Leclerc proved that the spin polynomial of such a class has the form $(1+q)^a q^b$, and from this obtained the specialization $H_{\lambda\cup\lambda}^{(2)}(X;-1)$.
The aim of this note is to provide a similar proof for the stable case, that is, to show that the result on Hall-Littlewood functions at roots of unity follows from an explicit formula for the spin polynomials of certain diagonal classes of ribbon tableaux, which turn out to have a very simple characterization through Schilling's bijection [13].

2. Basic definitions on ribbon tableaux

For a partition $\lambda = (\lambda_1, \dots, \lambda_p)$, we write $l(\lambda)$ its length, $|\lambda|$ its weight and ${}^t\lambda$ its conjugate. With λ is associated a k-core $\lambda_{(k)}$ and a k-quotient $\lambda^{(k)}$. The k-core is the unique partition obtained by removing successively k-ribbons from λ, and the k-quotient is a sequence of k partitions derived from λ (see [3]). A k-ribbon is a connected skew diagram of weight k which does not contain a 2×2 square. The first (north-west) cell of a k-ribbon is called the *head* and the last one (south-east) the *tail*. A k-ribbon tableau of shape λ and weight μ is a tiling of the skew diagram $\lambda/\lambda_{(k)}$ by labelled k-ribbons such that the head of a ribbon labelled i must not be on the right of a ribbon labelled $j > i$ and its tail must not be on the top of a ribbon labelled $j \geq i$. We denote by $Tab_k(\lambda, \mu)$ the set of all k-ribbon tableaux of shape λ and weight μ.

Example: A 3-ribbon tableau of shape (8,7,6,5,1) and weight (3,3,2,1)



In [14], Stanton and White first introduced in the standard case (weight $\mu = (1, \dots, 1)$) a correspondence between k-ribbon tableaux and k-tuples of standard Young tableaux. In the following, we will denote this bijection by sw. This map sends the previous 3-ribbon tableau to the 3-tuple of tableaux:

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

The spin of a k-ribbon R is defined by $sp(R) = \frac{h(R)-1}{2}$ where $h(R)$ is the height of R. The spin of a k-ribbon tableau is the sum of the spins of all its ribbons, and the cospin is the associated co-statistic into $Tab_k(\lambda, \mu)$. We define spin and cospin polynomials as generating polynomials of $Tab_k(\lambda, \mu)$ with spin or cospin statistics:

$$G^{(k)}_{\lambda,\mu}(q) = \sum_{T \in Tab_k(\lambda,\mu)} q^{sp(T)} \quad \text{and} \quad \tilde{G}^{(k)}_{\lambda,\mu}(q) = \sum_{T \in Tab_k(\lambda,\mu)} q^{cosp(T)} .$$

Example: In $Tab_3((8,7,6,5,1),(3,3,2,1))$, these polynomials are:

$$G^{(3)}_{(8,7,6,5,1),(3,3,2,1)}(q) = 3q^2 + 17q^3 + 33q^4 + 31q^5 + 18q^6 + 5q^7$$
$$\tilde{G}^{(3)}_{(8,7,6,5,1),(3,3,2,1)}(q) = 3q^5 + 17q^4 + 33q^3 + 31q^2 + 18q + 5 .$$

By definition, the Hall-Littlewood functions Q'_λ can be written as:

$$Q'_\lambda(X; q) = \prod_{i<j} (1 - qR_{ij})^{-1} s_\lambda(X)$$

where R_{ij} is the raising operator such that $R_{ij} \cdot s_\lambda = s_{R_{ij}\cdot\lambda}$, with

$$R_{ij} \cdot \lambda = (\lambda_1, \dots, \lambda_i + 1, \dots, \lambda_j - 1, \dots, \lambda_p) .$$

In [9], Lascoux, Leclerc and Thibon showed that Hall-Littlewood functions can be expressed in terms of ribbon tableaux by:

$$Q'_\lambda(X;q) \;=\; \sum_{T\in Tab_p(p\lambda)} q^{sp(T)}X^T = \sum_\mu G^{(k)}_{p\lambda,\mu}(q)\; m_\mu \;.$$

The following specialization, with n a positive integer and ζ a primitive k-th root of unity, is proved in [7]:

$$Q'_{n^k}(X;\zeta) = (-1)^{(k-1)n} p_k \circ h_n(X) \;.$$

We shall give a combinatorial proof of this formula using the bijection between ribbon tableaux and rigged configurations given in [13].

3. Diagonal classes and rigged configurations

Let T be a k-ribbon tableau in $Tab_k(\lambda,\mu)$ and $(\Lambda^{(1)},\ldots,\Lambda^{(k)}) = sw(T)$. By writing $\alpha = -\lfloor\frac{\lambda_1}{k}\rfloor$ and $\beta = \lfloor\frac{l(\lambda)}{k}\rfloor$, for all $i\in\{\alpha,\ldots,\beta\}$ we define d_i as the word obtained by concatenation of the i-th diagonals of all the tableaux $\Lambda^{(j)}$ for j in $\{1\ldots k\}$ (we recall that the i-th diagonal of a Young tableau consists of all the cells with coordinates (x,y) such that $y-x=i$). We call *diagonal vector* of T the vector $d_T = (d_\alpha,\ldots,d_\beta)$. Two k-ribbon tableaux T and T' in $Tab_k(\lambda,\mu)$ are said to be *equivalent* if for all i in $\{\alpha,\ldots,\beta\}$ the i-th sorted word in d_T and $d_{T'}$ are the same. A diagonal class in $Tab_k(\lambda,\mu)$ is the set $D^{(k)}_{\lambda,\mu,d}$ of all equivalent ribbon tableaux with diagonal vector d. The set of all diagonal classes is denoted by $\Delta^{(k)}_{\lambda,\mu}$. We also define $G^{(k)}_{\lambda,\mu}(q,d)$ (resp. $\tilde G^{(k)}_{\lambda,\mu}(q,d)$) as the spin (resp. cospin) polynomials of the diagonal class $D^{(k)}_{\lambda,\mu,d}$.

Let $\nu = (\nu^{(1)},\ldots,\nu^{(p)})$ be an increasing p-tuple of partitions and J be such that: for all a in $\{1,\ldots,p-1\}$, $J^{(a)}$ is a $l(\nu^{(a)})$-tuple of partitions $(J^{(a)}_1,\ldots,J^{(a)}_{l(\nu^{(a)})})$ with $l(J^{(a)}_i)\le \nu^{(a)}_i - \nu^{(a)}_{i+1}$ and each part of $J^{(a)}_i$ less than $\nu^{(a+1)}_i - \nu^{(a)}_i$. A *rigged configuration* of shape ν, written (ν,J), is defined by: for all a, top cells of each column of $\nu^{(a)}$ which are in the i-th line are filled with parts of the partition $J^{(a)}_i$. For two partitions μ and δ, we define by $RC(\mu,\delta)$ the set of all the rigged configurations (ν,J) such that $\nu^{(p)} = {}^t\delta$ and $|\nu^{(a)}| = \mu_1+\ldots+\mu_a$ for all a in $\{1..p\}$ (definition as in [13] rather than in [4, 5, 6]).

In the following, $\lambda = (\lambda_1,\ldots,\lambda_p)$ is a partition with its k-core $\lambda_{(k)}$ empty and its k-quotient $\lambda^{(k)}$ equal to a k-tuple of single rows. We also set $m = \max(|\Lambda^{(1)}|,\ldots,|\Lambda^{(k)}|)$. In this special case, Schilling gives in [13] a bijection Ψ between $Tab_k(\lambda,\mu)$ and rigged configurations $RC(\mu,\delta)$, with $\delta_i = |\lambda^{(k)}_i|$. She also defined a co-statistic on these rigged configurations which corresponds to cospin under Ψ. Consequently, by enumeration of

$RC(\mu,\delta)$, she obtains:

$$\tilde{G}^{(k)}_{\lambda,\mu}(q) = \sum_{\{\nu_{\mu,\delta}\}} q^{\Phi(\nu)} \prod_{\substack{1\le a\le n-1\\ 1\le i\le \mu_1}} \begin{bmatrix} \nu_i^{(a+1)}-\nu_{i+1}^{(a)} \\ \nu_i^{(a)}-\nu_{i+1}^{(a)},\ \nu_i^{(a+1)}-\nu_i^{(a)} \end{bmatrix} \qquad (1)$$

where $\{\nu_{\mu,\delta}\}$ represents the set of all shapes appearing in $RC(\mu,\delta)$ and

$$\Phi(\nu) = \sum_{\substack{1\le a\le n-1\\ 1\le i\le \mu_1}} \nu_{i+1}^{(a)}(\nu_i^{(a+1)} - \nu_i^{(a)}) \ . \quad (1')$$

In the following, we will be mainly interested in the shapes of rigged configurations. We shall therefore propose a simpler but similar algorithm for finding only the shape of the rigged configuration $\Psi(T) = (\nu_T, J_T)$ with T in $Tab_k(\lambda,\mu)$. We construct an $m \times l(\mu)$ matrix M^T with the following rule:

$$M^T_{i,j} = \text{number of cells labelled } j \text{ in } d_{-i+1} \ .$$

Then, we construct a matrix N^T where each column $N^T_{\cdot,j}$ is defined by:

$$N^T_{\cdot,j} = \sum_{l\le j} M^T_{\cdot,l} \ .$$

The j-th column is then equal to the j-th partition of ν_T.

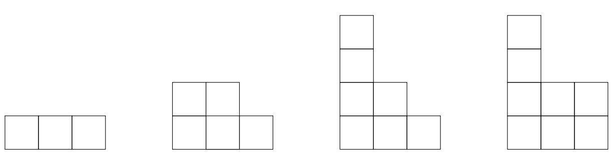
Example: For the 3-ribbon tableau corresponding to the following 3-tuple:

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

we construct the matrices M^T and N^T :

$$M^T = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad N^T = \begin{pmatrix} 3 & 3 & 3 & 3 \\ 0 & 2 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} .$$

The shape of the rigged configuration $\Psi(T)$ is:



as can be read from N^T.

We can remark an additive property of this construction. From a k-tuple of tableaux Λ, we can construct a k-ribbon tableau T_h, which is the Stanton-White inverse image of the k-tuple of tableaux formed by the h-th element of each tableau of Λ.

Lemma 1.

$$N^T = \sum_{h=1}^{m} N^{T_h} \ .$$

Proof: Let $M^T_{i,\cdot}$ be the matrix which has only the i-th line of M^T and zero everywhere else. Thus, by definition of M^{T_i} we have:

$$M^T = \sum_{i=1}^{m} M^T_{i,\cdot} = \sum_{i=1}^{m} M^{T_i} \ .$$

Then, we can write for all j:

$$\sum_{l \leq j} M^T_l = \sum_{i=1}^{m} \sum_{l \leq j} M^{T_i}_l \ .$$

Consequently,

$$N^T_j = \sum_{i=1}^{m} N^{T_i}_j \ . \ \square$$

Proposition 1. *For a given shape λ and weight μ, there is a bijection Γ between $\{\nu_{\lambda,\delta}\}$ and $\Delta^{(k)}_{\lambda,\mu}$ compatible with the statistics. Hence, the explicit expression for the cospin polynomial of a diagonal class is:*

$$\tilde{G}^{(k)}_{\lambda,\mu}(q,d) = q^{\Phi(\nu)} \prod_{\substack{1 \leq a \leq n-1 \\ 1 \leq i \leq \mu_1}} \begin{bmatrix} \nu_i^{(a+1)} - \nu_{i+1}^{(a)} \\ \nu_i^{(a)} - \nu_{i+1}^{(a)}, \ \nu_i^{(a+1)} - \nu_i^{(a)} \end{bmatrix} \qquad (2)$$

where ν is the shape corresponding to the diagonal class.

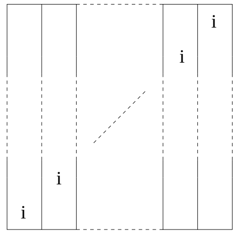
Proof: As the k-quotient consists of single rows, each diagonal class $D^{(k)}_{\lambda,\mu,d}$ is stable under permutation of cells which are in the same positions in each tableau. By construction, this property implies that, for all $l \in \{1\ldots,m\}$, $M^{T_l} = M^{T'_l}$. Then $N^{T_l} = N^{T'_l}$ and $N^T = N^{T'}$, so $\Psi(T)$ and $\Psi(T')$ have the same shape. Consequently, as map Ψ is a bijection, $D^{(k)}_{\lambda,\mu,d}$ is embedded into $\{\nu^T_{\lambda,\mu}\}$. Conversely, let T and T' be two tableaux in $Tab_k(\lambda,\mu)$ which are not in the same diagonal class. Thus, there exists j in $\{1,\ldots,m\}$ such that $\Lambda^T_j \neq \Lambda^{T'}_j$. This implies that $M^T_{\cdot,j} \neq M^{T'}_{\cdot,j}$ and $N^T \neq N^{T'}$ and consequently $\{\nu^T_{\lambda,\mu}\} \neq \{\nu^{T'}_{\lambda,\mu}\}$. Finally, we conclude that

$$\Psi(D^{(k)}_{\lambda,\mu,d}) = \{\nu^T_{\lambda,\mu}\} \ \text{ for all diagonal classes.}$$

The expression of cospin polynomials of diagonal classes in terms of q-supernomial coefficients follows immediately from the properties of Ψ. $\square$

In the following, we consider k-ribbon tableaux of shape $\lambda = (kn)^k$ for some $n \geq 1$. This implies that the image of these tableaux by the Stanton-White map is a k-tuple of semi-standard Young tableaux with the same single row partition of length n as shape.

Corollary 1. *Diagonal classes with only one element correspond to ribbon tableaux which are filled with $k \times k$ blocks of type:*



and the cospin of such a tableau is divisible by k.

Proof: A diagonal class $D^{(k)}_{\lambda,\mu,d}$ has an unique element if and only if there is an unique way to fill $\lambda^{(k)}$ according to the vector d. This implies that, for all i in $\{1,\ldots,k\}$, all letters of d_i are the same. With this property, all the T_i's are the $k\times k$ blocks of the statement. For filling identically each position of d_i, the weight μ must be of the form $\mu = (k\cdot s_1,\ldots,k\cdot s_p)$. Then, we construct the matrices M^T and N^T as :

$$M^T = \begin{pmatrix} k & 0 & \ldots & \ldots & \ldots & 0 \\ \vdots & \vdots & & & & \vdots \\ k & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & k & 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & k & 0 & \ldots & \ldots & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 & k \\ \vdots & & & & \vdots & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 & k \end{pmatrix}$$

$$N^T = \begin{pmatrix} k & k & \ldots & \ldots & \ldots & k \\ \vdots & \vdots & & & & \vdots \\ k & k & \ldots & \ldots & \ldots & k \\ 0 & k & k & \ldots & \ldots & k \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & k & k & \ldots & \ldots & k \\ \vdots & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 & k \\ \vdots & & & & \vdots & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 & k \end{pmatrix}$$

where k occurs s_i times in the i-th column of the matrix M^T. Then the i-th partition in the shape $\nu_{\Psi(T)}$ is the rectangle $k^{s_1+\ldots+s_i}$. This is why each term in the expression (1') is zero or a multiple of k. $\square$

Proposition 2. *For such a shape λ, k-th primitive roots of unity are roots of cospin polynomials for all diagonal classes with strictly more than one element.*

Proof: Let T be a tableau in a diagonal class $D^{(k)}_{\lambda,\mu,d}$ and $\nu = (\nu^{(1)}, \ldots, \nu^{(p)})$ be the shape of $\Psi(T)$, which is the same for all tableaux in this diagonal class. Let $\Lambda = sw(T)$ and h be the last position such that the diagonal vector d_h has at least two different elements. Then the $(h+1)$-th partition in ν_T is a rectangle of width k and height $s \le r$. The last part of $\nu^{(h)} = (\nu^{(h)}_1, \ldots, \nu^{(h)}_l)$ is equal to a with $a < h$ and the following coefficient appears:

$$\begin{bmatrix} \nu_l^{(h+1)} - \nu_{l+1}^{(h)} \\ \nu_l^{(h)} - \nu_{l+1}^{(h)},\ \ \nu_l^{(h+1)} - \nu_l^{(h)} \end{bmatrix} = \begin{bmatrix} k \\ a-0,\ k-a \end{bmatrix} .$$

Consequently, by a known property of the q-binomial $\begin{bmatrix} k \\ a,\ k-a \end{bmatrix}$, all k-th primitive roots of unity annihilate the diagonal class polynomials. $\square$

Theorem 1. *We have the specialization:*

$$Q'_{n^k}(X;\zeta) = (-1)^{(k-1)n} p_k \circ h_n(X) \ .$$

Proof: We use functions H and $\tilde{H}$ as defined in [9]:

$$H^{(k)}_{n^k}(X;q) = \sum_{T \in Tab_k(kn^k)} q^{sp(T)} X^T \ \text{ and } \ \tilde{H}^{(k)}_{n^k}(X;q) = \sum_{T \in Tab_k(kn^k)} q^{cosp(T)} X^T \ .$$

Let ζ be a k-th primitive root of unity. When q is set to ζ^{-1} in the expression of $\tilde{H}$, by Proposition 2 one is left with

$$\tilde{H}^{(k)}_{n^k}(X;\zeta^{-1}) = \sum_T (\zeta^{-1})^{cosp(T)} X^T$$

where T ranges now over k-ribbon tableaux as described in Corollary 1. By definition, these tableaux have maximum spin because they are only constructed with vertical ribbons, so their cospin is zero. Then, if we set $\Lambda_T = sw(T)$ we have

$$\tilde{H}^{(k)}_{n^k}(X;\zeta^{-1}) = \sum_T \underbrace{X^{\Lambda_T^{(1)}} \ldots X^{\Lambda_T^{(1)}}}_{k \text{ times}} = \sum_S \underbrace{X^S \ldots X^S}_{k \text{ times}}$$

where S ranges over all semi-standard Young tableaux with shape a single row of length n. We obtain

$$\tilde{H}^{(k)}_{n^k}(X;\zeta^{-1}) = p_k \circ h_n(X) \ .$$

Using relation between H and $\tilde{H}$ given in [9], we have

$$H^{(k)}_{n^k}(X;\zeta) = \zeta^{\frac{k(k-1)n}{2}} \tilde{H}^{(k)}_{n^k}(X;\zeta^{-1}) = (-1)^{(k-1)n} p_k \circ h_n(X) \ . \square$$

Remark: In the case where $\lambda = (k^{c \cdot k})$ there is a similar bijection between k-ribbon tableaux of shape λ and evaluation μ (see [13]) that allows to prove with the same method the following specialisation:

$$H^{(k)}_{\lambda}(X;\zeta) = (-1)^{(k-1)c} p_k \circ e_c(X) \ .$$

References

[1] C. Carré, B. Leclerc, *Splitting the square of a Schur function into its symmetric and antisymmetric parts*, Journal of Algebraic Combinatorics **4** (1995), 201–231.

[2] J. Désarménien, B. Leclerc, J.-Y. Thibon, *Hall-Littlewood functions and Kotska-Foulkes polynomials in representation theory*, Séminaire Lotharingien de Combinatoire (1992)

[3] G. D. James, A. Kerber, *The representation theory of the symmetric group*, Addison-Wesley, 1981.

[4] S.V. Kerov, A.N. Kirillov, N.Y Reshetikhin, *Combinatorics, the Bethe ansatz and representations of symmetric group*, J. Soviet Math **41** (1988), 916-924.

[5] A.N. Kirillov, N.Y Reshetikhin, *The Bethe ansatz and the combinatorics of Young tableaux*, J. Soviet Math **41** (1988), 925-955.

[6] A.N. Kirillov, N.Y Reshetikhin, *Representations of Yangians and multiplicities of the inclusion of the irreductible components of the tensor product of representations of simple Lie algebras*, J. Soviet Math **52** (1990), 3156-3164.

[7] A. Lascoux, B. Leclerc, J.-Y. Thibon, *Fonctions de Hall-Littlewood et polynomes de Kostka-Foulkes aux racines de l'unité* C.R. Académie des sciences de Paris **316** (1993), 1-6.

[8] A. Lascoux, B. Leclerc, J.-Y. Thibon, *Green polynomials and Hall-Littlewood functions at roots of unity* Europ. J. Combinatorics **15** (1994), 173-180.

[9] 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.

[10] M. Lothaire, *Algebraic Combinatorics on Words*, Vol. 90 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press (2002)

[11] I.G. Macdonald *Symmetric Functions and Hall Polynomials*, second edition, Oxford University Press (1998)

[12] A. Schilling, M. Shimozono, D.E. White, *Branching formula for q-Littlewood-Richardson coefficients*, Advances in Applied Mathematics **30** (2003), 258–272.

[13] A. Schilling, *q-Supernomial Coefficients: From Riggings to Ribbons* (2001) http://arXiv.org/abs/math/0107214.

[14] D. Stanton and D. White, *A Schensted algorithm for rim-hook tableaux*, J. Comb. Theory A **40**, (1985), 211-247.

François Descouens
Institut Gaspard Monge
Université de Marne-la-Vallée
77454 Marne-la-Vallée Cedex 2, France
email: francois.descouens@univ-mlv.fr

OLD AND YOUNG LEAVES ON PLANE AND BINARY TREES
VIEILLES ET JEUNES FEUILLES D'ARBRES PLANAIRES ET BINAIRES

EMERIC DEUTSCH AND SERGI ELIZALDE

Abstract. A leaf of a plane tree is called an old leaf if it is the leftmost child of its parent, and it is called a young leaf otherwise. We enumerate plane trees with a given number of old leaves and young leaves, partly inspired by an idea of Bill Chen. The formula is obtained combinatorially by presenting a bijection between plane trees and 2-Motzkin paths which maps old and young leaves to certain kinds of steps. We derive some implications to the enumeration of restricted permutations with respect to various statistics. Our main bijection is then applied to obtain refinements of two identities of Coker, involving refined Narayana numbers and the Catalan numbers. Finally, we consider the analogous problem for binary trees. We enumerate them with respect to old and young leaves, and describe a bijection to 2-Motzkin paths.

Résumé. On dit qu'une feuille d'un arbre planaire est vieille si c'est l'enfant le plus à gauche de son parent, autrement on dit qu'elle est jeune. Inspirés en partie par une idée de Bill Chen, nous énumérons les arbres planaires avec un nombre donné de vieilles et jeunes feuilles. On obtient la formule de manière combinatoire en présentant une bijection entre ces arbres planaires et des chemins de Motzkin à deux couleurs, qui tient compte des vieilles et jeunes feuilles. Nous dérivons quelques implications reliées à l'énumeration de permutations restreintes par rapport à diverses statistiques. Ensuite, nous utilisons notre bijection pour obtenir des raffinements de deux identités de Coker qui concernent des nombres de Narayana raffinés et les nombres de Catalan. Finalement, nous considérons le problème analogue pour les arbres binaires. Nous les énumérons par rapport aux feuilles vieilles et jeunes, et décrivons une bijection entre ces arbres binaires et les chemins de Motzkin à deux couleurs.

1. Introduction

Plane trees, also referred to as ordered trees, are a basic object frequently used in combinatorics. Many enumerative results about them appear throughout the literature. For example, a well-known interpretation of the Narayana numbers is that they count the number of plane trees with a fixed number of leaves. In this paper we classify the leaves of a plane tree into two different kinds, distinguishing between old leaves and young leaves. This definition, which is introduced in Section 2, naturally gives rise to a refinement of the Narayana numbers.

These refined Narayana numbers also appear in the enumeration of 2-Motzkin paths with respect to the number of up steps and red horizontal steps. Such paths were introduced in [1], and their structure has proved to be useful in the study of lattice paths, noncrossing partitions, plane trees [7], and other combinatorial objects and identities. Our paper gives yet another example of the applicability of 2-Motzkin paths. The key to several of our results is a new bijection between plane trees and 2-Motzkin paths, with very convenient properties. It provides a combinatorial derivation of the expression for the number of plane trees with a given number of old and young leaves.

Another application of our bijection appears in [5], where Chen, Yan and Yang use it to give combinatorial interpretations of two identities involving the Narayana numbers and Catalan numbers, due to Coker [6]. This way they solve the two open problems left in [6]. Here we will show that a more detailed analysis of the bijection and its properties gives refinements of the two identities of Coker, as well as bijective proofs of these refinements.

The paper is structured as follows. In Section 2 we review some definitions and notation about plane trees, Dyck paths, Motzkin paths, and 2-Motzkin paths. We also introduce the concepts of old leaves and young leaves of a plane tree. In Section 3 we give the generating function for plane trees with variables marking the number of old leaves and the number of young leaves, as well as exact formulas for the number of plane trees of a given size when the number of old and young leaves is fixed. In Section 4 we present two bijections from the set of plane trees with n edges to the set of 2-Motzkin paths of length $n-1$. Some interesting properties of these bijections are studied in Section 5. We show that they map old and young leaves of trees into statistics on 2-Motzkin paths that are easier to deal with. In Section 6 we describe some bijections between plane trees and permutations avoiding patterns of length 3, and investigate what old and young leaves are mapped to by these bijections. This allows us to count restricted permutations with respect to certain statistics such as pairs of consecutive deficiencies, double descents, and ascending runs. In Section 7 we apply our bijection to obtain refinements of two combinatorial identities due to Coker [6] and proven combinatorially by Chen, Yan and Yang [5]. Finally, in Section 8 we consider binary trees and give the generating function enumerating them with respect to the number of old and young leaves, which are defined analogously to the case of plane trees. We also present a natural bijection between binary trees and 2-Motzkin paths, and study how our statistics are transformed by the bijection.

2. Preliminaries

2.1. **Plane trees.** A *plane tree* T can be defined recursively as a finite set of vertices such that one distinguished vertex r is called the *root* of T, and the remaining vertices are put into an ordered partition $(T_1, T_2, \ldots, T_m)$ of m disjoint non-empty sets, each of which is a plane tree. We will draw plane trees with the root on the top level, with edges connecting it to the roots of $T_1, T_2, \ldots, T_m$, which will be drawn from left to right on the second level. For each vertex v, the nodes in the next lower level connected to v by an edge are called the *children* or *successors* of v, and v is called the *parent* of its children. Clearly each vertex other than r has exactly one parent. A vertex of T is called a *leaf* if it has no children (by convention, we assume that the *empty* tree, formed by a single node, has no leaves).

We denote by $\mathcal{T}_n$ the set of plane trees with n edges. It is well-known that $|\mathcal{T}_n| = \frac{1}{n+1}\binom{2n}{n}$, the n-th Catalan number, and that the number of trees with n edges and k leaves is the Narayana number $\frac{1}{n}\binom{n}{k}\binom{n}{k-1}$.

We classify the leaves of a plane tree into old and young leaves. We say that a leaf is an *old leaf* if it is the leftmost child of its parent, and that it is a *young leaf* otherwise. For example, the tree in Figure 1 has four young leaves, drawn with black filled circles, and three old leaves, drawn with empty circles. The enumeration of plane trees with respect to the number of old and young leaves is done in Section 3.

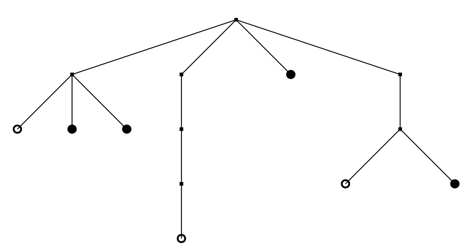


Figure 1. A tree with 3 old leaves and 4 young leaves.

2.2. **Lattice paths.** We review the definitions of Dyck, Motzkin, and 2-Motzkin paths. They are all lattice paths in $\mathbb{Z}^2$ starting at $(0,0)$, ending on the x-axis, and never going below this axis. A *Dyck path* consists of steps $U = (1,1)$ and $D = (1,-1)$. In a *Motzkin path* we allow also horizontal steps $H = (1,0)$, so that the path is a sequence of steps U, D and H. A *2-Motzkin path* consists of up and down steps, and horizontal steps that can be colored either red or blue. We use R to denote a red step, and B for a blue step. In the pictures in this paper, red steps will be drawn with a dashed line to make them clearly distinguishable from blue steps, which will be drawn with a solid line. The length of any of these paths is the total number of steps.

We shall denote by $\mathcal{D}_n$ the set of Dyck paths of length $2n$, by $\mathcal{M}_n$ the set of Motzkin paths of length n, and by $\mathcal{N}_n$ the set of 2-Motzkin paths of length n. The number of paths of each kind is given by $|\mathcal{D}_n| = C_n$, $|\mathcal{M}_n| = M_n$, and $|\mathcal{N}_n| = C_{n+1}$, where $M_n = \sum_{k=0}^{n}\binom{n}{2k}C_k$ is the n-th Motzkin number.

The generating function for Catalan numbers is $C(z) = \sum_{n\geq 0} C_n z^n = \frac{1-\sqrt{1-4z}}{2z}$, and the one for Motzkin numbers is $M(z) = \sum_{n\geq 0} M_n z^n = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2}$.

3. Enumeration of trees with respect to old and young leaves

Here we give an expression for the generating function

$$G(t,s,z) = \sum_T t^{\#\text{old leaves of T}} s^{\#\text{young leaves of T}} z^{\#\text{edges of T}},$$

where the sum is over all plane trees T, and t and s mark the number of old and young leaves respectively.

Theorem 1. *Let $G(t,s,z)$ be defined as above. We have*

$$G(t,s,z) = \frac{1+z-sz-\sqrt{1-2(1+s)z+(1-4t+2s+s^2)z^2}}{2z}.$$

Proof. We will find an equation for $G(t,s,z)$ using a decomposition of plane trees. Let T be any plane tree, and let m be the number of children of the root. If $m=0$, then the tree has no edges, and its contribution to the generating function G is 1. If $m \geq 1$, let $T_1, T_2, \ldots, T_m$ be the sequence of plane trees hanging from left to right from the children of the root. If T_1 has no edges, then it creates an old leaf of T, otherwise all the old (resp. young) leaves of T_1 become old (resp. young) leaves of T. Therefore, the contribution to the generating function of T_1 and the edge connecting it to the root is $z(G(t,s,z)-1+t)$. For $i \geq 2$, old and young leaves of T_i become leaves of T of the same kind as well. However, if T_i is has no edges, then an additional young leaf of T is created. Thus, the contribution to the generating function of each T_i with $i \geq 2$ and the edge connecting it to the root is $z(G(t,s,z)-1+s)$. It follows that for $m \geq 1$, the contribution of the plane trees whose root has degree m is $z^m(G-1+t)(G-1+s)^{m-1}$. Summing over all $m \geq 0$ we obtain

$$G(t,s,z) = 1 + \frac{z(G(t,s,z)-1+t)}{1-z(G(t,s,z)-1+s)}. \tag{1}$$

Isolating G the formula follows. □

Proposition 2. (1) *The number of plane trees with n edges, i old leaves, and j young leaves is*

$$\frac{1}{n}\binom{n}{i}\binom{n-i}{j}\binom{n-i-j}{i-1}.$$

(2) *The number of plane trees with n edges and k old leaves is*

$$\frac{2^{n-2k+1}}{k}\binom{n-1}{2k-2}\binom{2k-2}{k-1}.$$

(3) *The number of plane trees with n edges and k young leaves is*

$$\binom{n-1}{k} M_{n-k-1}.$$

Proof. Applying Lagrange inversion formula to equation (1), we obtain that the coefficient of $t^i s^j z^n$ in $G(t,s,z)$ is $\frac{1}{n}\binom{n}{i}\binom{n-i}{j}\binom{n-i-j}{i-1}$, which is the first expression. For the other two expressions, apply Lagrange inversion to the same equation, after the substitutions $s=1$ and $t=1$ respectively. □

Particular cases of this proposition give rise to the following two statements. The second one appeared already in [8] as a manifestation of the Motzkin numbers.

Corollary 3. (1) *The number of plane trees in $\mathcal{T}_n$ with exactly one old leaf is 2^{n-1}.*
(2) *The number of plane trees in $\mathcal{T}_n$ with no young leaves is M_{n-1}.*

4. A bijection between plane trees and 2-Motzkin paths

In this section we present a bijection Ψ between the set of plane trees with n edges and the set of 2-Motzkin paths of length $n-1$. This bijection has the convenient property that it maps old and young leaves of the tree to certain statistics of the 2-Motzkin path that are very easy to deal with, as shown in the next section. This will allow us to give bijective proofs of Corollary 3 and some parts of Proposition 2.

The bijection consists of three steps. Given a plane tree $T \in \mathcal{T}_n$ (assume $n \geq 1$), we first transform it into a Dyck path using the following well-known bijection, which we denote θ. Starting from the root, traverse the edges of the tree in preorder from right to left. To each edge passed on the way down there corresponds a step U, and to each edge passed on the way up there corresponds a step D. This gives us a Dyck path $\theta(T)$ of length $2n$.

The next step is to replace each peak UD of the path followed by a U step with a red horizontal step R. That is, we traverse the path $\theta(T)$ from left to right replacing each UDU with RU. This gives us a Motzkin path with steps U, D and R, whose length is variable.

Finally, we need to transform this Motzkin path into a 2-Motzkin path $\Psi(T)$ of length $n-1$. The bijection that we will use for this purpose is essentially the same one described by Callan [2] between UDU-free Dyck paths and Motzkin paths, where we "ignore" the steps R of our path and let the new level steps be all B steps. Notice that after the transformation in the previous paragraph, every peak UD in our Motzkin path is followed by a D step, unless it is at the end of the path. This last transformation is done as follows. Place a mark on each U that is followed by a D, on each D that is followed by another D, and on the D at the end of the path. Next, change each unmarked U whose matching D is marked into an B. (The matching D is the step that is encountered directly east of U.) Lastly, delete all the marked steps.

After this procedure we obtain a 2-Motzkin path $\Psi(T)$ with $n-1$ steps. For example, for the tree T in Figure 1, applying the first part of the bijection we get the Dyck path in Figure 2. Replacing each UDU with RU, we get the Motzkin path in Figure 3. In the third part of the bijection, we mark the steps that in Figure 4 are thicker. Changing each unmarked U with a marked matching D to a B, we get $UBR\dot{U}\dot{D}\dot{D}DRUBB\dot{U}\dot{D}\dot{D}\dot{D}DBRR\dot{U}\dot{D}\dot{D}$, where the dots indicate the marked steps. Finally, deleting the marked steps, we obtain the 2-Motzkin path in Figure 5.

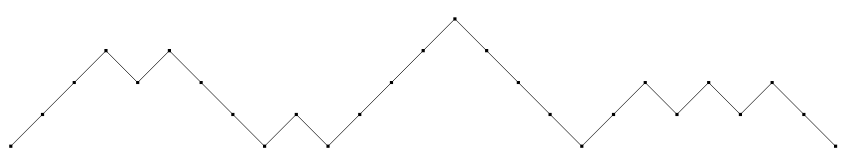


Figure 2. The Dyck path $\theta(T)$ for T in Figure 1.

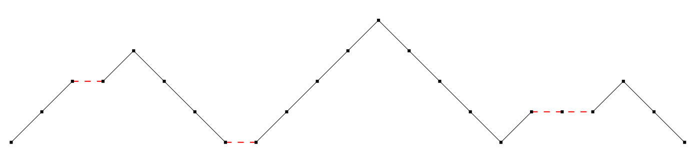


Figure 3. The Motzkin path $UURUDDDRUUUUDDDDURRUDD$.

It is clear that the first two steps of this map are reversible, that is, from the Motzkin path with steps U, D and R it is easy to recover the tree. The fact that the last step is a bijection as well follows from the description of the inverse given in [2]. The only difference here is that we need to disregard the steps R that we have now in the path, since they are not affected by this part of the bijection.

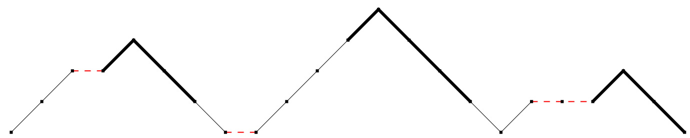


FIGURE 4. The Motzkin path with some steps marked.

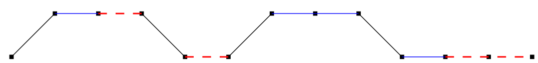


FIGURE 5. The 2-Motzkin path $\Psi(T) = UBRDRUBBDBRR$.

5. CONSEQUENCES OF THE BIJECTION

The main properties of Ψ are given in the following proposition.

Proposition 4. *Let T be a plane tree with $n \geq 1$ edges, and let $P = \Psi(T)$ be the corresponding 2-Motzkin path. We have*
(1) *# of old leaves of $T = 1 +$ # of U steps of P,*
(2) *# of young leaves of $T =$ # of R steps of P.*

Proof. Let us first take a look at how old and young leaves are transformed by the first part θ of the bijection, which consists in reading T in preorder from right to left and building a Dyck path out of it. It is clear that each leaf of T produces a peak in $\theta(T)$. Now, a young leaf of T corresponds to a peak UD followed by a U step, whereas an old leaf of T gives rise to a peak UD not followed by a U.

The second part of the bijection transforms each peak UD followed by a U into a red step R, and these steps remain unchanged by the third part of the bijection. This proves (2). The remaining peaks of the Dyck path are followed either by a D or by nothing, and they are not affected by the second part of the bijection, so these are the only peaks in the Motzkin path. In the final part, we place a mark on each D that is followed by another D or by nothing, and the only D's that are not erased are the unmarked ones. Therefore, the number of U steps (equivalently, the number of D steps) in $\Psi(T)$ equals the number of D's in the Motzkin path that are left unmarked. The D steps in the Motzkin path can be grouped in sequences of consecutive D's, each such sequence immediately following a peak (note that the path has no occurrences of RD, so each D is in one of these sequences). In the sequence of D's following the rightmost peak all the steps are marked. For each remaining peak, among the D steps in the consecutive sequence following it, all but the last one are marked. Thus, only one D step survives for each peak other than the rightmost one. In other words, the number of D steps in $\Psi(T)$ is the number of peaks of the Motzkin path minus one. This implies (1). □

By means of the bijection Ψ and the properties described above, we can now give a combinatorial proof of Corollary 3. To prove the first part, observe that by property (1) of Proposition 4, Ψ induces a bijection between plane trees with exactly one old leaf and 2-Motzkin paths with no U steps. But these paths are just sequences of horizontal steps, each of which can be colored red or blue. Thus, the number of plane trees on n edges with exactly one old leaf is 2^{n-1}.

A direct proof of this nice fact, without using bijections to lattice paths, can be given as follows. Let T be a tree with n edges and exactly one old leaf, call it ℓ. We can find ℓ by following the path that starts at the root and always continues to the leftmost child. Let P be this path. Then ℓ must be at the end of P. Now we claim that the remaining nodes of T are leaves hanging from the nodes of P other than ℓ. Indeed, if a node of P had a child not in P with successors, then following the path that starts at this child and continues always to the leftmost child, we would end at another old leaf, which is a contradiction. Conversely, if only leaves are hanging from P, then no more old leaves appear. Now, the number of trees consisting of a path P with leaves hanging from its nodes is clearly 2^{n-1}. Indeed, one can think of it as a composition of n, say $n = a_1 + a_2 + \cdots$, where a_i is the number of children of the i-th node of P.

More generally, we can use our bijection to give a combinatorial proof of the second part of Proposition 2, namely that the number of plane trees with n edges and k old leaves is $\frac{2^{n-2k+1}}{k}\binom{n-1}{2k-2}\binom{2k-2}{k-1}$. By the first property of Ψ given above, we have to count the number of 2-Motzkin paths of length $n-1$ with $k-1$ U steps. To produce such a path, we can choose in $\binom{n-1}{2k-2}$ ways the positions of the $k-1$ U's and $k-1$ D's in the path. Deciding which of these positions will be filled with a U or with a D is equivalent to choosing a Dyck path with $2k-2$ steps, and this can be done in $\frac{1}{k}\binom{2k-2}{k-1}$ ways. The remaining $n-2k+1$ positions are horizontal steps, which can be colored red or blue in 2^{n-2k+1} ways.

To show the second part of Corollary 3 combinatorially, notice that property (2) of Proposition 4 implies that Ψ maps plane trees with no young leaves into 2-Motzkin paths with no R steps. These are just Motzkin paths with steps U, D and B. Therefore, the number of plane trees on n edges with no young leaves equals the number of Motzkin paths with $n-1$ steps, which is M_{n-1}.

More generally, the same property of Ψ can be used to prove the last part of Proposition 2, namely that the number of plane trees with n edges and k young leaves is $\binom{n-1}{k}M_{n-k-1}$. Indeed, now the problem is equivalent to counting 2-Motzkin paths of length $n-1$ with k R steps. We can choose in $\binom{n-1}{k}$ ways where these R steps go, and then the remaining $n-k-1$ steps can be filled with a Motzkin path with steps U, D and B.

Remark. Another combinatorial proof of part (3) of Proposition 2 can be obtained using the result mentioned in [7] (and proved also in [14]) that $\binom{n-1}{k}M_{n-k-1}$ counts the number of Dyck paths of length $2n$ with k DUD's.

The description of Ψ implicitly contains a bijection between Dyck paths and 2-Motzkin paths. There is a simpler bijection, perhaps the most standard one, that transforms a 2-Motzkin path of length $n-1$ into a Dyck path of length $2n$, by first applying the following rules:

$$U \to UU, \quad D \to DD, \quad R \to UD, \quad B \to DU,$$

and then inserting a U at the beginning and a D at the end of the path. Applying Ψ followed by this bijection, young leaves of the tree are mapped to peaks at even height in the Dyck path. This shows that the statistic 'number of young leaves' in $\mathcal{T}_n$ is equidistributed with the statistic 'number of peaks at even height' in $\mathcal{D}_n$.

6. Some statistics on restricted permutations

Using some known bijections between Dyck paths and permutations avoiding a pattern of length 3, the parameters counting the number of old and young leaves in plane trees correspond to certain statistics on restricted permutations. Given a pattern σ, we denote by $\mathcal{S}_n(\sigma)$ the set of permutations in the symmetric group $\mathcal{S}_n$ avoiding σ. It is well-known [10] that if σ is any pattern of length 3, then $|\mathcal{S}_n(\sigma)| = C_n$, the n-th Catalan number.

We begin with a few definitions. Let π be a permutation. We say that π_i is an *excedance* if $\pi_i > i$, that it is a *weak excedance* if $\pi_i \geq i$, and that it is a *deficiency* if $\pi_i < i$. A *left-to-right minimum* of π is an element π_i such that $\pi_i < \pi_j$ for all $j < i$. We call a *double descent* of π a sequence of three consecutive decreasing elements $\pi_i > \pi_{i+1} > \pi_{i+2}$ (equivalently, two consecutive descents). A *double ascent* is defined analogously. An *ascending run* is a maximal increasing sequence of (at least two) consecutive elements of π, i.e., $\pi_i < \pi_{i+1} < \cdots < \pi_{i+k}$, with $k \geq 1$.

Proposition 5. *There is a bijection $\varphi_1 : \mathcal{T}_n \longrightarrow \mathcal{S}_n(321)$ such that, if $T \in \mathcal{T}_n$ and $\pi := \varphi_1(T) \in \mathcal{S}_n(321)$, then*
(1) *# of young leaves of T = # of pairs of consecutive weak excedances of π,*
(2) *# of old leaves of T = # of weak excedances of π not followed by another weak excedance.*

Proof. We use a bijection ψ between $\mathcal{S}_n(321)$ and $\mathcal{D}_n$ which is very similar to the one given by Krattenthaler [11] from $\mathcal{S}_n(123)$ to $\mathcal{D}_n$. Here is a way to describe it. Let $\pi \in \mathcal{S}_n(321)$, and let $\pi_{i_1}, \pi_{i_2}, \ldots, \pi_{i_k}$

be its weak excedances, from left to right. Define $\psi(\pi)$ to be the path that starts with π_{i_1} up steps, then has, for each j from 2 to k, $i_j - i_{j-1}$ down steps followed by $\pi_{i_j} - \pi_{i_{j-1}}$ up steps, and finally ends with $n+1-i_k$ down steps. It can be checked that this is indeed a bijection between 321-avoiding permutations and Dyck paths.

Our bijection φ_1 is defined as $\varphi_1 = \psi^{-1} \circ \theta$. Recall that θ reads a plane tree in preorder from right to left and creates a Dyck path.

We saw that young leaves of T correspond to occurrences of UDU in the path $\theta(T)$, and that old leaves of T are mapped by θ to either a UDD or a terminal (i.e., at the end of the path) UD. Now, if $\pi \in \mathcal{S}_n(321)$, a UDU is obtained in $\psi(\pi)$ precisely when we have a weak excedance followed by another weak excedance, which causes one of the descending slopes to have length $i_j - i_{j-1} = 1$. Similarly, a UDD corresponds to a weak excedance followed by a deficiency (i.e., an element that is not a weak excedance), and a terminal UD corresponds to the weak excedance $\pi_n = n$. □

For example, if T is the tree in Figure 1, with $\theta(T)$ given in Figure 2, then the corresponding permutation is $\varphi_1(T) = (3,4,1,2,5,9,6,7,8,11,12,13,10) \in \mathcal{S}_{12}(321)$. It has four pairs of consecutive weak excedances, namely $(3,4)$, $(5,9)$, $(11,12)$ and $(12,13)$, and three weak excedances not followed by another weak excedance, namely 4, 9 and 13.

A similar result for 132-avoiding permutations is given next. For $\pi \in \mathcal{S}_n$, let $(n+1)\pi$ (resp. $\pi(n+1)$) be the permutation in $\mathcal{S}_{n+1}$ obtained by inserting $n+1$ at the beginning (resp. at the end) of π.

Proposition 6. *There is a bijection $\varphi_2 : \mathcal{T}_n \longrightarrow \mathcal{S}_n(132)$ such that, if $T \in \mathcal{T}_n$ and $\pi := \varphi_2(T) \in \mathcal{S}_n(132)$, then*

(1) *# of young leaves of T = # of double descents of $(n+1)\pi$,*

(2) *# of old leaves of T = # of ascending runs of $\pi(n+1)$.*

Proof. We use the bijection from $\mathcal{S}_n(132)$ to $\mathcal{D}_n$ denoted by Φ that appears in Krattenthaler [11]. Given $\pi \in \mathcal{S}_n(132)$, let $\pi_{i_1}, \pi_{i_2}, \dots, \pi_{i_k}$ be its left-to-right minima, from left to right. Then $\Phi(\pi)$ is the Dyck path that starts with $n+1-\pi_{i_1}$ up steps, then has, for each j from 2 to k, $i_j - i_{j-1}$ down steps followed by $\pi_{i_{j-1}} - \pi_{i_j}$ up steps, and finally ends with $n+1-i_k$ down steps. It can be checked that this is indeed a bijection between 132-avoiding permutations and Dyck paths. The bijection we are looking for is $\varphi_2 := \Phi^{-1} \circ \theta$.

Each young leaf of T produces an occurrence of UDU in $\theta(T)$. Such an occurrence appears in $\Phi(\pi)$ for each pair of consecutive left-to-right minima. These two elements, together with the entry of $(n+1)\pi$ immediately to their left, form a decreasing sequence of three consecutive elements (a double descent). To see that these are the only double descents of $(n+1)\pi$, notice that from the structure of 132-avoiding permutations it follows that if $\pi_j > \pi_{j+1}$ is a descent of π, then π_{j+1} must be a left-to-right minimum.

The reasoning for old leaves is similar. They correspond in $\theta(T)$ to occurrences of UDD and possibly a UD at the end. Equivalently, to occurrences of UDD in $\theta(T)D$ (i.e., the Dyck path $\theta(T)$ with a D step appended at the end). Each of these occurrences marks the start of a maximal sequence of at least two consecutive D steps in $\theta(T)D$, and each such sequence corresponds to an ascending run of $\pi(n+1)$. □

For example, if T is again the tree in Figure 1, then the corresponding 132-avoiding permutation is $\pi = \varphi_2(T) = (11,10,12,13,9,5,6,7,8,3,2,1,4)$. Note that $(n+1)\pi = (14,\pi)$ has four double descents, namely $(14,11,10)$, $(13,9,5)$, $(8,3,2)$ and $(3,2,1)$, and $(\pi,14)$ has three ascending runs, namely $(10,12,13)$, $(5,6,7,8)$ and $(1,4,14)$.

There is another well-known bijection between plane trees and Dyck paths, which we denote by δ. Given a tree T, traverse it in preorder (from left to right) and build $\delta(T)$ as follows. For each node with r children, draw r up steps followed by one down step; except for the last leaf, for which we do not draw anything. For example, the path corresponding to the tree in Figure 1 is $\delta(T) = UUUUDUUUDDDDUDUDUDDDUDUUDD$.

Define a *drop* of a Dyck path to be a maximal succession of at least two consecutive D steps, and a *triple fall* to be an occurrence of DDD. Then the bijection δ maps each old leaf of T to a drop of $\delta(T)D$, and each young leaf to a triple fall of $\delta(T)D$. In the example from the paragraph above, $\delta(T)D$ has three drops and four triple falls.

Following very similar arguments to the ones in Propositions 5 and 6, but using the bijection δ instead of θ, we obtain the next two results.

Proposition 7. *There is a bijection $\varphi_3 : \mathcal{T}_n \longrightarrow \mathcal{S}_n(321)$ such that, if $T \in \mathcal{T}_n$ and $\pi := \varphi_3(T) \in \mathcal{S}_n(321)$, then*
(1) *# of young leaves of T = # of pairs of consecutive deficiencies of π (+1 if $\pi_n < n$),*
(2) *# of old leaves of T = # of weak excedances of π not followed by another weak excedance.*

Proposition 8. *There is a bijection $\varphi_4 : \mathcal{T}_n \longrightarrow \mathcal{S}_n(132)$ such that, if $T \in \mathcal{T}_n$ and $\pi := \varphi_4(T) \in \mathcal{S}_n(132)$, then*
(1) *# of young leaves of T = # of double ascents of $\pi(n+1)$,*
(2) *# of old leaves of T = # of ascending runs of $\pi(n+1)$.*

7. Refinements of two combinatorial identities

In [6] Coker established the following two identities, involving the Narayana and the Catalan numbers:

$$\sum_{k=1}^{n} \frac{1}{n}\binom{n}{k}\binom{n}{k-1}4^{n-k} = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} C_k \binom{n-1}{2k} 4^k 5^{n-2k-1}, \tag{2}$$

$$\sum_{k=1}^{n} \frac{1}{n}\binom{n}{k}\binom{n}{k-1}x^{2k}(1+x)^{2n-2k} = x^2 \sum_{k=0}^{n-1} C_{k+1}\binom{n-1}{k} x^k (1+x)^k, \tag{3}$$

He stated the open problem of finding a combinatorial interpretation of these identities. In [5], Chen, Yan and Yang proved these identities combinatorially by applying our bijection Ψ to weighted plane trees. In this section, following a suggestion of Chen [3], we use the properties of Ψ given in Proposition 4 to obtain refinements of the identities (2) and (3).

Theorem 9. *For $n \geq 1$, we have*

$$\sum_{i=1}^{n} \sum_{j=0}^{n-2i+1} \frac{1}{n}\binom{n}{i}\binom{n-1}{j}\binom{n-i-j}{i-1}x^{i-1}y^j = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} C_k \binom{n-1}{2k} x^k (1+y)^{n-2k-1}. \tag{4}$$

Proof. We use a very similar reasoning to the one given in [5] to prove equation (2). It will be convenient to use the term *critical leaf* to denote the last old leaf that we encounter when we traverse a plane tree in preorder. Given a plane tree T with n edges, assign weights to the vertices of T as follows: young leaves are given weight y, old leaves other than the critical one are given weight x, and the rest of the vertices (including the critical leaf) are given weight 1. The weight of T is the product of the weights of its vertices. Then, the left hand side of (4) is the sum of the weights of all plane trees with n edges.

By Proposition 4, Ψ is a weight preserving bijection between the set of weighted plane trees on n edges, with weights given as above, and the set of weighted 2-Motzkin paths of length $n-1$ where weights are assigned as follows: U steps are given weight x, R steps are given weight y, and all the remaining steps are given weight 1, defining the weight of a 2-Motzkin path to be the product of weights of its steps. We claim that the right hand side of (4) is the sum of the weights of all 2-Motzkin paths of length $n-1$. Indeed, let $k \leq \lfloor (n-1)/2 \rfloor$ and consider the weighted 2-Motzkin paths with k up steps and k down steps. These up and down steps from a Dyck path of length $2k$, and the positions of these $2k$ steps can be chosen in $\binom{n-1}{2k}$ ways. They contribute x^k to the weight of the path. The remaining $n-2k-1$ steps are either R or B steps. Since R steps have weight y and B steps have weight 1, the total contribution of the horizontal steps in paths with k up steps is $(1+y)^{n-2k-1}$. This justifies the right hand side. $\square$

With the subsitution $y = x$ in equation (4) we recover the result proved in [5], and the particular case $y = x = 4$, together with the symmetry of the Narayana numbers, yields equation (2). A refinement of the second identity (3) is given next.

Theorem 10. *For $n \geq 1$, we have*

$$(5)\quad \sum_{i=1}^{n} \sum_{j=0}^{n-2i+1} \frac{1}{n}\binom{n}{i}\binom{n-1}{j}\binom{n-i-j}{i-1} x^{2(i-1)} y^j z^{n-2i-j+1} = \sum_{k=0}^{n-1} C_{k+1}\binom{n-1}{k} x^k (y+z-2x)^{n-1-k}.$$

Proof. Again we apply the same ideas used in [5] to prove equation (3). Recall the definition of critical leaf from the proof or Theorem 9. Given a plane tree T with n edges, assign weights to the vertices of T in the following way. Old leaves other than the critical one are given weight x, the parents of such leaves are given weight x as well, young leaves are given weight y, the critical leaf and its parent are given weight 1, and the rest of the vertices are given weight z. As before, the weight of T is the product of the weights of its vertices. Notice that two different old leaves cannot have the same parent, so the weight of a tree with i old leaves and j young leaves is $x^{2(i-1)} y^j z^{n-2i-j+1}$. The left hand side of (5) is the sum of the weights of all plane trees with n edges.

By Proposition 4, a tree with i old leaves and j young leaves is mapped by Ψ to a 2-Motzkin path with $i-1$ up steps, $i-1$ down steps, j horizontal R steps, and $n-2i-j+1$ horizontal B steps. To make Ψ a weight preserving bijection between plane trees on n edges with the above weights and 2-Motzkin paths of length $n-1$, we assign weights to the steps of a 2-Motzkin path by giving weight x to U and D steps, weight y to R steps, and weight z to B steps.

Consider now the set of 3-Motzkin paths of length $n-1$, where horizontal steps can be either red, blue or green (call them R, B and G steps respectively). Assign weights to the steps by giving weight $y+z-2x$ to G steps and weight x to all the other steps. This weight assignment in 3-Motzkin paths has the property that the sum of the weights of an R step, a B step and a G step equals the sum of the weights of an R step and a B step in the assignment for 2-Motzkin paths above (namely $y+z$), and also that U and D steps have the same weight x in both assignments. This implies that the sum of weights over all 2-Motzkin paths with the above weight assignment equals the sum of weights over all 3-Motzkin paths with this new assignment. Therefore, all that remains is to show that the right hand side of (5) is the total sum of the weights of 3-Motzkin paths of length $n-1$. But this is clear because if we fix the number of G steps of a 3-Motzkin path to be $n-1-k$, then the positions of these G steps can be chosen in $\binom{n-1}{k}$ ways. The remaining steps, U, D, R and B, form a 2-Motzkin path of length k, and the number of such paths is C_{k+1}. □

To recover identity (3) we only need to substitute $x(1+x)$ for x, x^2 for y, and $(1+x)^2$ for z in equation (5).

8. Old and young leaves in binary trees

In this section we consider binary trees instead of plane trees, and we apply the same idea of classifying its leaves into old and young. A *binary tree T* can be defined recursively as a finite set of vertices such that one distinguished vertex r is called the *root* of T, and the remaining vertices are divided in two (possibly empty) sets T_ℓ and T_r, each of which is a binary tree. The notion of children, parent and leaf are defined as in the case of plane trees. Note that now each vertex v can have at most two children, and that in the case it has one child, we make the distinction of whether it is a *left* or a *right* child of v.

We denote by $\mathcal{B}_n$ the set of binary trees with n edges. It is well-known that $|\mathcal{B}_n| = \frac{1}{n+2}\binom{2n+2}{n+1}$, the $(n+1)$-st Catalan number. As for the case of plane trees, we classify the leaves of a binary tree into old and young leaves. A leaf is old (resp. young) if it is the left (resp. right) child of its parent. For example, the tree in Figure 6 has three young leaves, drawn with black filled circles, and two old leaves, drawn with empty circles. We will now enumerate binary trees with respect to the number of old and young leaves. Let

$$H(x,y,z) = \sum_T x^{\#\text{old leaves of T}} y^{\#\text{young leaves of T}} z^{\#\text{edges of T}},$$

where the sum is over all binary trees T, and x and y mark the number of old and young leaves respectively.

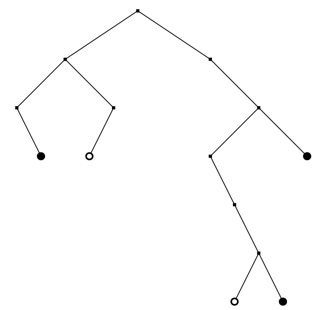


Figure 6. A tree with 2 old leaves and 3 young leaves.

Theorem 11. *Let $H(x,y,z)$ be defined as above. We have*

$$H(x,y,z)=\frac{1-2z+(2-x-y)z^2-\sqrt{1-4z+2(2-x-y)z^2+(x-y)^2z^4}}{2z^2}.$$

Proof. To find an equation for $H(t,s,z)$ we use the following straightforward decomposition. Let T be a binary tree, and consider the left and right subtrees hanging from the root, denoted T_ℓ and T_r respectively. The contribution to the generating function $H(x,y,z)$ of the left subtree and the edge joining it with the root is $1+z(H(x,y,z)-1+x)$. The summand 1 corresponds to the case where the root does not have a left child. In all other cases, z is the weight of the edge joining the root and its left child, and $H(x,y,z)-1+x$ is the contribution of T_ℓ, taking into account that an old leaf of T is created if T_ℓ has no edges. Similarly the contribution of the right part of T is $1+z(H(x,y,z)-1+y)$. This gives the equation

$$H(x,y,z)=[1+z(H(x,y,z)-1+x)][1+z(H(x,y,z)-1+y)], \tag{6}$$

whose solution is the desired expression for H. □

Proposition 12. *The number of binary trees with n vertices (assume $n>1$), k old leaves, and l young leaves is*

$$\frac{1}{n-k-l}\sum_{i=1}^{n-k-l}(-1)^{n-k-l-i}\binom{n-k-l}{i}\binom{i}{k}\binom{i}{l}\binom{2i-k-l}{n-k-l-1}.$$

Proof. Making the substitution $J(u,v,z):=z[H(u/z,v/z,z)-1]$, equation (6) can be expressed in terms of J as

$$J(u,v,z)=z[(1+u+J(u,v,z))(1+v+J(u,v,z)-1]. \tag{7}$$

The coefficient of $x^ky^lz^{n-1}$ in $H(x,y,z)$ equals the coefficient of $u^kv^lz^{n-k-l}$ in $J(u,v,z)$, which can be found applying Lagrange inversion formula to (7). This gives the expression for the number of binary trees with n vertices, k old leaves, and l young leaves. □

Next we describe a bijection Υ from the set of binary trees with n edges to the set of 2-Motzkin paths of length n. Given a binary tree $T\in\mathcal{B}_n$, label its vertices as follows. For each vertex with two children, give the label U to its left child and the label D to its right child. For each vertex with only a left (resp. right) child, give the child the label R (resp. B). This way all vertices except the root get a label. Now traverse the tree in preorder from left to right and write down the labels in the order they are read. Let $\Upsilon(T)$ be the 2-Motzkin path obtained by this procedure. For example, if T is the tree in Figure 6, then $\Upsilon(T)$ is the path in Figure 7.

We now look at how old and young leaves are transformed by the map Υ. For any Motzkin path P and any sequence w of steps U, D, R and B, denote by $w[P]$ the number of occurrences of w in P (in

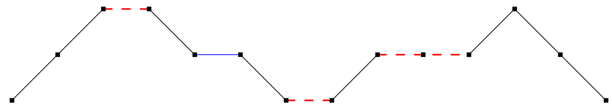


FIGURE 7. The 2-Motzkin path $\Upsilon(T) = UURDBDRURRUDD$.

consecutive positions). For example, $UD[P]$ denotes the number of peaks of P. The following lemma follows easily from the definition of Υ.

Lemma 13. *Let T be a binary tree with $n \geq 1$ edges, let $P = \Upsilon(T)$ be the corresponding 2-Motzkin path, and let $P' = PD$ be the path obtained by appending a down-step at the end of P. We have*
(1) *# of old leaves of $T = RD[P'] + UD[P']$,*
(2) *# of young leaves of $T = BD[P'] + DD[P']$,*
(3) *# of leaves of $T = D[P']$.*

The last part of Lemma 13 together with part (1) of Proposition 4 imply that, for any $n, k \geq 1$ the number of binary trees with n edges and k leaves equals the number of plane trees with $n+1$ edges and k old leaves, which by Proposition 2 is $\frac{2^{n-2k+2}}{k}\binom{n}{2k-2}\binom{2k-2}{k-1}$.

Notice also that red steps in $\Upsilon(T)$ correspond to nodes of T that have only a left child. Thus, it follows from part (2) of Proposition 4 that the number of binary trees with n edges and k nodes having only a left child equals the number of plane trees with $n+1$ edges and k young leaves, which by Proposition 2 is $\binom{n}{k}M_{n-k}$. In particular, Υ induces a bijection between plane trees in which every vertex has at most two successors (see [13, Exercise 6.38i]) and Motzkin paths. This is clear because such plane trees correspond to binary trees in which every vertex has either two children or only a right child. Finally, observe that Υ induces a bijection between binary trees where all vertices have either 0 or 2 children and Dyck paths.

Acknowledgements

We are grateful to William Chen for many of the ideas of the paper, and to Laura Yang for her valuable suggestions.

References

[1] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani, A construction for enumerating k-coloured Motzkin paths, *Lecture Notes in Comput. Sci.*, Vol. 959, Springer, Berlin, 1995.
[2] D. Callan, Bijections from restricted Dyck paths to Motzkin paths, preprint, arxiv:math.CO/0406381.
[3] W.Y.C. Chen, personal communication.
[4] W.Y.C. Chen, A general bijective algorithm for trees, *Proc. Natl. Acad. Sci. USA* 87 (1990), 9635–9639.
[5] W.Y.C. Chen, S.H.F. Yan, L.L.M. Yang, Weighted 2-Motzkin paths, preprint, 2004.
[6] C. Coker, Enumerating a class of lattice paths, *Disc. Math.* 271 (2003), 13–28.
[7] E. Deutsch, L.W. Shapiro, A bijection between ordered trees and 2-Motzkin paths and its many consequences, *Disc. Math.* 256 (2002), 655–670.
[8] R. Donaghey, L.W. Shapiro, Motzkin numbers, *J. Combin. Theory Ser. A*, 23 (1977), 291–301.
[9] S. Elizalde, *Statistics on pattern-avoiding permutations*, Ph.D. Thesis, MIT, 2004.
[10] D. Knuth, *The Art of Computer Programming, Vol. III*, Addison-Wesley, Reading, MA, 1973.
[11] C. Krattenthaler, Permutations with restricted patterns and Dyck paths, *Adv. in Appl. Math.* 27 (2001), 510–530.
[12] R.P. Stanley, *Enumerative Combinatorics, Vol. I*, Cambridge University Press, Cambridge, 1997.
[13] R.P. Stanley, *Enumerative Combinatorics, Vol. II*, Cambridge University Press, Cambridge, 1999.
[14] Y. Sun, The statistic "number of *udu*'s" in Dyck paths, *Disc. Math.* 287 (2004), 177–186.

POLYTECHNIC UNIVERSITY, BROOKLYN, NY 11201.
E-mail address: deutsch@duke.poly.edu

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139; MATHEMATICAL SCIENCES RESEARCH INSTITUTE, BERKELEY, CA 94720.
E-mail address: elizalde@msri.org

RIGIDITY THEORY FOR MATROIDS

MIKE DEVELIN, JEREMY L. MARTIN AND VICTOR REINER

Abstract. Combinatorial rigidity theory seeks to describe the rigidity or flexibility of bar-joint frameworks in $\mathbb{R}^d$ in terms of the structure of the underlying graph G. The goal of this article is to broaden the foundations of combinatorial rigidity theory by replacing G with an arbitrary representable matroid M. The notions of rigidity independence and parallel independence, as well as Laman's and Recski's combinatorial characterizations of 2-dimensional rigidity for graphs, can naturally be extended to this wider setting. As we explain, many of these fundamental concepts really depend only on the matroid associated with G (or its Tutte polynomial), and have little to do with the special nature of graphic matroids or the field $\mathbb{R}$.

Our main result is a "nesting theorem" relating the various kinds of independence. Immediate corollaries include generalizations of Laman's Theorem, as well as the equality of 2-rigidity independence and 2-parallel independence. A key tool in our study is the *space of photos of M*, a natural algebraic variety whose irreducibility is closely related to the notions of rigidity independence and parallel independence. The number of points of this variety, when working over a finite field, turns out to be an interesting Tutte polynomial evaluation.

Resumé. Un des objectifs de la théorie combinatoire de rigidité est de décrire, utilisant la structure du graphe fondamental G, la rigidité ou la flexibilité des cadres des barres et joints dans $\mathbb{R}^d$. Le but de ce travail est d'élargir la théorie combinatoire de rigidité en remplaçant G par un matroïde arbitraire représentable M. Dans ce sens, les idées d'indépendance de rigidité et d'indépendance parallèle, les caractérisations combinatoires de Laman et de Recski de la rigidité 2-dimensionelle pour les graphes, peuvent naturellement être étendues. Comme nous le monterons, beaucoup de ces concepts fondamentaux dépendent seulement du matroïde associé à G (ou à son polynôme de Tutte), et ils sont très peu liés á la nature spéciale des matroïdes graphiques ou du champ $\mathbb{R}$.

Notre principal résultat est un "théorème d'emboîtement" relatif aux divers genres d'indépendance. Quelques conséquences directes de ce théorème sont les généralisations du théorème de Laman et l'équivalence de la propriété d'indépendance 2-rigidité avec celle 2-parallèle. Notre étude est fondamentalement basée sur *l'éspace des photos de M* représentant une variété algébrique naturelle dont l'irréductibilité est étroitement liée aux notions d'indépendance de rigidité et d'indépendance parallèle. Le cardinal de cette variété, en travaillant dans un champ fini, est en fait une évaluation intéressante de polynôme de Tutte.

1. Introduction

1.1. **A brief tour through rigidity theory.** *Combinatorial rigidity theory* is concerned with frameworks built out of bars and joints in $\mathbb{R}^d$, representing the vertices V and edges E of an (undirected, finite) graph G. (For comprehensive treatments of the subject, see, e.g., [5, 18, 19].) The motivating problem is to determine how the combinatorics of G governs the rigidity or flexibility of its frameworks. Typically, one makes a generic choice of coordinates $p = \{p_v : v \in V\} \subset \mathbb{R}^d$ for the vertices of G and considers infinitesimal motions Δp of the vertices. The following two questions are pivotal:

(I.) What is the dimension of the space of infinitesimal motions Δp that preserve all squared edge lengths $Q(p_u - p_v)$, for $\{u, v\} \in E$, where $Q(x) = \sum_{i=1}^{d} x_i^2$?

(II.) What is the dimension of the space of infinitesimal motions Δp that preserve all edge directions $p_u - p_v$, up to scaling?

The answers to these questions are known to be determined by certain linear dependence matroids represented over transcendental extensions of $\mathbb{R}$, as we now explain.

Date: May 1, 2005.

1991 *Mathematics Subject Classification.* 52C25, 14N20.

Key words and phrases. matroid, combinatorial rigidity, parallel redrawing, Laman's Theorem, Tutte polynomial.

First author supported by the American Institute of Mathematics. Second author partially supported by an NSF Postdoctoral Fellowship. Third author partially supported by NSF grant DMS–0245379.

First, the *d-dimensional rigidity matroid* $\mathcal{R}^d(G)$ is represented by the vectors

$$\{(e_u - e_v) \otimes (p_u - p_v) : \ \{u,v\} \in E\} \ \subset \ \mathbb{R}^{|V|} \otimes \mathbb{R}(p)^d, \tag{1}$$

where $\mathbb{R}(p)$ is the extension of $\mathbb{R}$ by a collection of $d|V|$ transcendentals p, thought of as the coordinates of the vertices in a generic framework of G. The $|E| \times d|V|$ *rigidity matrix* $R^d(G)$ has as its rows the $|E|$ vectors in (1). Then the nullspace of $R^d(G)$ is precisely the space of infinitesimal motions of the vertices that preserve all edge distances (because $R^d(G)$ is $\frac{1}{2}$ times the Jacobian in the variables p of the vector of squared edge lengths $Q(p_u - p_v)$; cf. Remark 5.2 below). Since row rank equals column rank, knowing the matroid $\mathcal{R}^d(G)$ represented by the rows of $R^d(G)$ answers question (I).

Second, the *d-dimensional parallel matroid* $\mathcal{P}^d(G)$ is represented by the vectors

$$\{(e_u - e_v) \otimes \eta_{u,v}^{(j)} : \ \{u,v\} \in E, \ j = 1,2,\ldots,d-1\} \ \subset \ \mathbb{R}^{|V|} \otimes \mathbb{R}(p,\eta)^d, \tag{2}$$

where for each edge $\{u,v\} \in E$, the vectors $\eta_{u,v}^{(1)}, \ldots, \eta_{u,v}^{(d-1)}$ are generically chosen normals to $p_u - p_v$ in $\mathbb{R}^d$, and $\mathbb{R}(p,\eta)$ is an extension of $\mathbb{R}$ by $d|V|$ transcendentals p and $(d-1)|E|$ transcendentals η. In analogy to the preceding paragraph, the $|E| \times d|V|$ *parallel matrix* $P^d(G)$ has as its rows the $|E|$ vectors in (2), and its nullspace is the space of infinitesimal motions of the vertices that preserve all edge directions. Consequently, the matroid $\mathcal{P}^d(G)$ represented by the rows of $P^d(G)$ provides the answer to question (II).

For $d = 2$, the rigidity and parallel matroids coincide [18, Corollary 4.1.3]. The matroid $\mathcal{R}^2(G) = \mathcal{P}^2(G)$ has many equivalent combinatorial reformulations, of which the best known is *Laman's condition* [6]: an edge set $A \subset E$ is 2-rigidity-independent if and only if for every subset $A' \subset A$

$$2|V(A')| - 3 \geq |A'|, \quad \text{or equivalently} \quad 2\,(|V(A')| - 1) > |A'| \tag{3}$$

where $V(A')$ denotes the set of vertices incident to at least one edge in A'. We refer to the triple equivalence between the 2-rigidity matroid, the 2-parallel matroid, and the matroid defined by Laman's condition as the *planar trinity*.

For $d > 2$, the parallel matroid has a simple combinatorial characterization that generalizes Laman's condition, while an analogous description for the rigidity matroid is not known.

1.2. **From graphs to matroids.** The purpose of this article is to broaden the scope of rigidity theory by replacing the graph G with a more general object: a matroid M equipped with a representation over a field $\mathbb{F}$. Indeed, the notions of rigidity and parallel independence, as well as Laman's combinatorial characterization, can be naturally generalized to the setting of matroids. In the process, we will see that many of the main results of do not depend on the special properties of graphs (or graphic matroids), nor on the field $\mathbb{R}$, but in fact remain valid for any matroid M and any field $\mathbb{F}$. In the process, we are led naturally to study an algebraic variety, the *space of k-plane-marked d-photos* of M, whose points play the role of "frameworks" of M embedded in $\mathbb{F}^d$.

Whether or not the photo space is irreducible plays a key role in characterizing the matroidal analogues of rigidity independence and parallel independence. In turn, the question of irreducibility can be answered combinatorially. Furthermore, when the field $\mathbb{F}$ is finite, the number of photos of M is given by an evaluation of the Tutte polynomial using q-binomial coefficients. (Theorem 4.1).

In order to summarize our results, we define the main protagonists here. Recall that for a finite set E, an *(abstract) simplicial complex* on E is a collection $\mathcal{I}$ of subsets of E satisfying the following hereditary condition: if $I \in \mathcal{I}$ and $I' \subset I$, then $I' \in \mathcal{I}$. The independent sets of a matroid always form a simplicial complex. From here on we will make free use of standard terminology and notions from matroid theory; background and definitions may be found in standard texts such as [1, 12, 17].

Definition 1.1. Let E be a set of cardinality n, and let M be a (not necessarily representable) matroid on ground set E, with rank function r. Let m be a real number in the open interval $(1,\infty)_{\mathbb{R}}$. Then $A \subset E$ is called *m-Laman independent* if

$$m \cdot r(A') > |A'| \quad \text{for all nonempty subsets } A' \subseteq A. \tag{4}$$

The *m-Laman complex* $\mathcal{L}^m(M)$ is defined to be the abstract simplicial complex of all m-Laman independent subsets of E.

We will prove that if m is a positive integer, then $\mathcal{L}^m(M)$ is the collection of independent sets of a matroid. Moreover, $\mathcal{L}^m(M)$ has several alternate combinatorial descriptions: one of these generalizes Recski's Theorem characterizing rigidity-independent graphs; another is related to Edmonds' classic result on partitioning a matroid into independent subsets [4].

We now consider the case that M is a matroid represented by vectors $v_1, \dots, v_n$ in $\mathbb{F}^r$, where $\mathbb{F}$ is a field. For notational convenience, we identify the ground set E with the numbers $[n] := \{1, 2, \dots, n\}$. When $m > 1$ is a rational number, the Laman complex $\mathcal{L}^m(M)$ is closely related to a certain algebraic variety over $\mathbb{F}$, which we now describe. Denote by $\mathbb{G}r(k,d)$ the Grassmannian of k-planes in $\mathbb{F}^d$, regarded as a projective variety over $\mathbb{F}$ via the usual Plücker embedding.

Definition 1.2. The *space of k-plane-marked d-photos* (or just *(k,d)-photos*) of M is the algebraic set

$$X_{k,d}(M) := \{(\varphi, W) \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d) \times \mathbb{G}r(k, \mathbb{F}^d)^n : \ \varphi(v_i) \in W_i\}. \tag{5}$$

One may think of the map $\varphi \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d)$ as a camera taking a "snapshot" of M on photographic paper that looks like $\mathbb{F}^d$. The k-planes W_i are markings added later to highlight the image vectors $\varphi(v_i)$. Of course, whenever $\varphi(v_i) = 0$ (perhaps the camera φ caught v_i at a bad angle), the k-plane W_i is unconstrained.

The *non-annihilating cellule* of the photo space is defined as the Zariski open subset

$$X^{\varnothing}_{k,d}(M) := \{(\varphi, W) \in X_{k,d}(M) : \ \varphi(v_i) \neq 0 \text{ for } i = 1, 2, \dots, n\}.$$

Its image under the projection map $\pi : \ \mathrm{Hom}(\mathbb{F}^r, \mathbb{F}^d) \times \mathbb{G}r(k,d)^n \to \mathbb{G}r(k,d)^n$ measures the constraints on the W_i when none of the v_i are mapped to zero. Accordingly, we make the following definition.

Definition 1.3. The matroid M is called *(k,d)-slope independent* if $\pi X^{\varnothing}_{k,d}(M)$ is Zariski dense in $\mathbb{G}r(k,d)^n$. The *(k,d)-slope complex* is defined as

$$\mathcal{S}^{k,d}(M) := \{A \subset E : M|_A \text{ is } (k,d) \text{ -slope independent}\}. \tag{6}$$

The third notion of matroidal rigidity generalizes the d-dimensional rigidity matroid $\mathcal{R}^d(G)$ of a graph G. Let φ be a $d \times r$ matrix of algebraically independent transcendentals, regarded as a generic linear transformation $\mathbb{F}^r \to \mathbb{F}^d$. Consider the pseudo-distance quadratic form $Q(x) := \sum_{i=1}^d x_i^2$ on $\mathbb{F}(\varphi)^d$. Provided that the field $\mathbb{F}$ has characteristic $\neq 2$, we wish to define a rigidity matrix $R^d(M)$ whose nullspace consists of the infinitesimal changes of φ that preserve the values $Q(\varphi(v_i))$.

Definition 1.4. The *d-dimensional (generic) rigidity matroid* is the matroid represented by the vectors

$$\{v_i \otimes \varphi(v_i)\}_{i=1}^n \subset \mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}(\varphi)^d. \tag{7}$$

where $\mathbb{F}(\varphi)$ is the purely transcendental field extension of $\mathbb{F}$ by the dr entries of φ. The *d-rigidity complex* $\mathcal{R}^d(M)$ is the complex of independent sets of the d-dimensional rigidity matroid, and the *d-rigidity matrix* $R^d(M)$ is the $n \times dr$ matrix whose rows are given by the vectors (7).

In contrast, if we wish to extend the notion of graph rigidity that keeps track of edge *slopes* instead of edge *lengths* (see Question II above), then we need a matrix $P^d(M)$ whose nullspace consists of the infinitesimal changes $\Delta\varphi$ in the matrix φ which preserve the slopes of all the direction vectors $\varphi(v_i)$.

Definition 1.5. The *d-dimensional hyperplane-marking matroid* is the matroid represented by the vectors

$$\{v_i \otimes \eta_i\}_{i=1}^n \subset \mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}(\varphi, \eta)^d$$

over the field $\mathbb{F}(\varphi, \eta)$, the extension of $\mathbb{F}$ by dr transcendentals φ_{ij} (the entries of the matrix φ) and $(d-1)n$ more transcendentals η_{ij}. The complex $\mathcal{H}^d(M)$ is defined to be the complex of independent sets of this matroid. The *d-dimensional parallel matroid* is defined as

$$\mathcal{P}^d(M) := \mathcal{H}^d((d-1)M),$$

where $(d-1)M$ is the matroid whose ground set consists of $d-1$ parallel copies of each element of E. The *d-parallel matrix* $P^d(M)$ is the $n \times dr$ matrix whose rows represent $\mathcal{P}^d(M)$.

These definitions generalize the ordinary definitions from the rigidity theory of graphs. Strikingly, the geometric constraints on the photo space can be categorized combinatorially: the identity

$$\mathcal{S}^{k,d}(M) = \mathcal{L}^{\frac{d}{d-k}}(M),$$

(Corollary 3.3) provides a geometric interpretation of $\mathcal{L}^m(M)$ for rational m.

The slope complex $\mathcal{S}^{k,d}(M)$ is closely related to the rigidity and parallel matroids. The precise relationship is given by the *Nesting Theorem* (Theorem 5.4):

$$\mathcal{S}^{1,d}(M) \subseteq \mathcal{R}^d(M) \subseteq \mathcal{L}^d(M) = \mathcal{H}^d(M) = \mathcal{S}^{d-1,d}(M)$$

for all integers $d \geq 2$. In particular, when $d = 2$,

$$\mathcal{H}^2(M) = \mathcal{S}^{1,2}(M) = \mathcal{R}^2(M) = \mathcal{L}^2(M). \tag{8}$$

Thus matroid rigidity theory leads to a proof of the planar trinity (the second and third inequalities in (8)).

For $d \geq 3$, the d-rigidity matroid $\mathcal{R}^d(M)$ is the hardest of these objects to understand (as it is for graphic matroids). One fundamental question is whether $\mathcal{R}^d(M)$ depends on the choice of representation of M. It is invariant for $d = 2$ and up to projective equivalence of representations but the problem remains open in general. We also study the behavior of the d-rigidity matroid as $d \to \infty$: it turns out that $R^d(M)$ stabilizes when $d \geq r(M)$.

In this extended abstract, we omit or merely sketch the proofs of many of our results. The complete proofs can be found in the full-length article [3].

2. Laman independence

The main result of this section, Theorem 2.1, states that the generalized Laman's condition (4) always gives a matroid when m is an integer. The proof is completely combinatorial; that is, it is a statement about abstract matroids, not represented matroids. In addition, we describe some useful equivalent characterizations of d-Laman independence: one uses the Tutte polynomial, another is reminiscent of Recski's Theorem, and another is related to Edmonds' theorem on decomposing a matroid into independent sets.

Theorem 2.1. (i) *Let d be a positive integer and let M be any matroid. Then the simplicial complex $\mathcal{L}^d(M)$ is a matroid complex.*

(ii) *Let $m \in (1,\infty)_{\mathbb{R}}$ be a real number which is not an integer. Then there exists a represented matroid M for which $\mathcal{L}^m(M)$ is not a matroid complex.*

We omit the proof, which is technical but not difficult. The difference between the two cases makes itself felt in the following way. If C and C' are distinct minimal d-Laman-dependent sets, then $C \cap C'$ is d-Laman-independent; that is,

$$|C \cap C'| < d \cdot r(C \cap C'), \tag{9a}$$

where r is the rank function of M. If d is an integer, then (9a) implies the logically stronger

$$|C \cap C'| \leq d \cdot r(C \cap C') - 1, \tag{9b}$$

from which it eventually follows that the minimal nonmembers of $\mathcal{L}^d(M)$ satisfy the matroid circuit axioms [1, p. 264, eq. 6.13]. On the other hand, if $d \notin \mathbb{Z}$, then (9b) does not follow from (9a), and one can exploit this to write down a matroid M whose minimal d-Laman-dependent sets fail the circuit axioms.

One of the equivalent phrasings of m-Laman independence involves the *Tutte polynomial* $T_M(x,y)$ of M, a fundamental isomorphism invariant of the matroid M. For background on the Tutte polynomial, see the excellent survey article by Brylawski and Oxley [2]. Given a subset A of the ground set E, denote by $\overline{A}$ the *matroid closure* or *span* of A. If $A = \overline{A}$, then A is called a *flat* of M.

Proposition 2.2. *Let M be a matroid on ground set E with rank function r, and fix $m \in (1,\infty)_{\mathbb{R}}$. Then the following are equivalent:*

(i) *E is m-Laman independent, that is, $\mathcal{L}^m(M) = 2^E$ (the power set of E).*

(ii) *$m \cdot r(\overline{A}) > |\overline{A}|$ for every nonempty subset $A \subset E$. (Equivalently, $m \cdot r(F) > |F|$ for every flat F of M.)*

(iii) *The Tutte polynomial specialization $T_M(q^{m-1}, q)$ is monic of degree $(m-1)r(M)$.*

Sketch of proof. The equivalence of (i) and (ii) is clear from the definition of m-Laman independence since $r(\overline{A}) = r(A)$ and $|\overline{A}| \geq |A|$ for any $A \subset E$. The equivalence of (i) and (iii) arises from expanding $T_M(q^{m-1}, q)$ as a polynomial in q using the Whitney corank-nullity formula [2, eq. 6.13]. $\square$

Note that in (iii) we must allow (non-integral) real number exponents for a "polynomial" in q, but the notions of "degree" and "monic" for such polynomials should still be clear. The connection between the Tutte polynomial and rigidity of graphs was observed by the second author in [8].

Suppose that $m = d$ is a positive integer, so that $\mathcal{L}^d(M)$ is a matroid complex. Here d-Laman independence has two more equivalent formulations, one of which extends a classical result in the rigidity theory of graphs.

Recski's Theorem [13]. *Let $G = (V, E)$ be a graph, and let E' be a spanning set of edges of size $2|V| - 3$. Then E' is a 2-rigidity basis if and only if for any $e \in E'$, we can partition the multiset $E' \cup \{e\}$ (that is, adding an extra copy of e to E') into two disjoint spanning trees of G.*

This notion can be naturally extended to arbitrary matroids and dimensions.

Definition 2.3. Let M be a matroid on E. We say that E is *d-Recski independent* if for any element $e \in E$, the multiset $E \cup \{e\}$ can be partitioned into d disjoint independent sets for M.

We wish to show that this purely matroidal condition is equivalent to d-Laman independence. To prove this, we use a powerful classic result of Edmonds.

Edmonds' Decomposition Theorem [4, Theorem 1]. *Let M be a matroid of rank r on ground set E. Then E has a decomposition into d disjoint independent sets $I_1, \ldots, I_d$ if and only if $d \cdot r(A) \geq |A|$ for every subset $A \subset E$.*

Definition 2.4. Let M be a matroid on E. A *d-Edmonds decomposition* of M is a family of independent sets $I_1, \ldots, I_d$ whose disjoint union is E, with the following property: given subsets $I'_1 \subset I_1, \ldots, I'_d \subset I_d$ with not all I'_i empty, then it is not the case that $\overline{I'_1} = \overline{I'_2} = \cdots = \overline{I'_d}$.

Theorem 2.5. *Let M be a matroid on ground set E, and let d be a positive integer. Then the following are equivalent:*

(1) *E has a d-Edmonds decomposition;*
(2) *E is d-Laman independent;*
(3) *E is d-Recski independent.*

Again, the proof is purely technical.

As we have seen in Theorem 2.1 (ii), when m is not an integer, the Laman complex $\mathcal{L}^m(M)$ need not form the collection of independent sets of a matroid. However, $\mathcal{L}^m(M)$ is related to a more general (and less well-known) object called a *polymatroid* [17, chapter 18], as we now explain. (We will not consider polymatroids in the remainder of the paper.)

Definition 2.6. Fix a ground set $E = [n]$. A function $\rho : 2^E \to \mathbb{R}_{\geq 0}$ is the *ground set rank function of a polymatroid* on E if

- $\rho(A) \leq \rho(B)$ whenever $A \subset B \subset E$ (*monotonicity*);
- $\rho(A \cup B) + \rho(A \cap B) \leq \rho(A) + \rho(B)$ for all $A, B \subset E$ (*submodularity*); and
- $\rho(\varnothing) = 0$ (*normalization*).

The *polymatroid* associated with ρ is the convex polytope

$$P_\rho := \{x \in \mathbb{R}^n_{\geq 0} : \sum_{a \in A} x_a \leq \rho(A) \text{ for all } A \subseteq E\},$$

also called the set of *independent vectors* of the polymatroid.

The connection between Laman independence and polymatroids is as follows.

Proposition 2.7. *For every loopless matroid M on ground set $E = [n]$, and every real number $m \in (1, \infty)_{\mathbb{R}}$, there is a polymatroid rank function ρ on E with the following property: $A \subset E$ is m-Laman independent if and only if its characteristic vector is independent for ρ.*

3. Slope independence and the space of photos

Throughout this section, we work with a matroid M with rank function r, represented over a field $\mathbb{F}$ by nonzero[1] vectors $v_1, \ldots, v_n \in \mathbb{F}^r$. In addition, fix positive integers $k < d$, and let $m = \frac{d}{d-k}$.

Recall (Definition 1.2) that the *space of* (k,d)*-photos* of M is

$$\left\{(\varphi, W) \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d) \times \mathbb{G}r(k, \mathbb{F}^d)^n \ : \ \varphi(v_i) \in W_i \text{ for all } 1 \le i \le n\right\},$$

an algebraic subset of $\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d) \times \mathbb{G}r(k, \mathbb{F}^d)^n$, hence a scheme over $\mathbb{F}$. The symbol $X_{k,d}(M)$ is a slight abuse of notation; as defined, the photo space depends on the representation $\{v_i\}$, and it is not at all clear to what extent it depends only on the structure of M as an abstract matroid. (We will return to this question later.)

For each photo (φ, W), $\ker\varphi$ is a linear subspace of $\mathbb{F}^r$, hence intersects E in some flat F of M. It is useful to classify photos according to what this flat is. Accordingly, for a flat $F \subset E$, we define the corresponding *cellule* as

$$X^F_{k,d}(M) = \{(\varphi, W) \in X_{k,d}(M) \ : \ \ker\varphi \cap E = F\}.$$

Each photo belongs to exactly one cellule; that is, $X_{k,d}(M)$ decomposes as a disjoint union of its cellules.

The cellule $X^{\varnothing}_{k,d}(M)$ corresponding to the empty flat $\varnothing$ is called the *non-annihilating cellule.* It is a Zariski open subset of $X_{k,d}(M)$, defined by the open conditions $\varphi(v_i) \neq 0$ for $i = 1, \ldots, n$. At the other extreme, the cellule $X^E_{k,d}(M)$ corresponding to the improper flat E is called the *degenerate cellule.* It is precisely $\{0\} \times \mathbb{G}r(k, \mathbb{F}^d)^n$, where 0 is the zero map $\mathbb{F}^r \to \mathbb{F}^d$.

The following facts are easy consequences of the preceding discussion.

Proposition 3.1. *Let M and $X_{k,d}(M)$ be as above. Then:*

- (i) *The natural projection map $X^{\varnothing}_{k,d}(M) \to \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d)$ makes $X^{\varnothing}_{k,d}(M)$ into a bundle with fiber $\mathbb{G}r(k-1, \mathbb{F}^{d-1})$ and base the Zariski open subset of $\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d)$ defined by $\varphi(v_i) \neq 0$ for $i = 1, \ldots, n$.*
- (ii) *For each flat F, $X^F_{k,d}(M) \cong X^{\varnothing}_{k,d}(M/F) \times \mathbb{G}r(k, \mathbb{F}^d)^F$. In particular, $X^F_{k,d}(M)$ is irreducible, and*

$$\dim X^F_{k,d}(M) = d(r - r(F)) + (n - |F|)(k-1)(d-k) + |F|k(d-k). \tag{10}$$

Let π denote the projection map

$$\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d) \times \mathbb{G}r(k, \mathbb{F}^d)^n \ \xrightarrow{\pi}\ \mathbb{G}r(k, \mathbb{F}^d)^n, \tag{11}$$

and define M to be (k,d)*-slope independent* if $\pi X^{\varnothing}_{k,d}(M)$ is Zariski dense in $\mathbb{G}r(k,d)^n$.

Theorem 3.2. *The following are equivalent:*

- (i) *M is (k,d)-slope independent, i.e., $\pi X^{\varnothing}_{k,d}(M)$ is dense in $\mathbb{G}r(k,d)^n$.*
- (ii) *M is m-Laman independent, i.e., $m \cdot r(F) > |F|$ for every nonempty flat F of M.*
- (iii) $\dim X^F_{k,d}(M) < \dim X^{\varnothing}_{k,d}(M)$ *for every nonempty flat F of M.*
- (iv) *The photo space $X_{k,d}(M)$ is irreducible.*
- (v) *The photo space $X_{k,d}(M)$ coincides with the Zariski closure of its non-annihilating cellule.*

The result is analogous to Theorem 4.5 of [7], and the proof uses the cellule decomposition in a similar way. In particular, the equivalence of (i) and (ii) immediately gives the following equality between the slope and Laman complexes.

Corollary 3.3. *Let $m \in \mathbb{Q} \cap (1, \infty)_{\mathbb{R}}$. Write m as $\frac{d}{d-k}$, where $0 < k < d$ are integers.*
Then $\mathcal{S}^{k,d}(M) = \mathcal{L}^m(M)$.

Remark 3.4. The condition $d \ge 2$ is implicit in Corollary 3.3. However, there is a sense in which the result is still valid for $d = 1$. When $k = 1$, the result asserts that $\mathcal{S}^{1,d}(M) = \mathcal{L}^{\frac{d}{d-1}}(M)$. Now, if one establishes conventions properly, this equality remains valid as d approaches 1, so that $m = \frac{d}{d-1}$ approaches infinity. That is, $\mathcal{S}^{1,1}(M) = \mathcal{L}^{\infty}(M) = 2^E$. Indeed, the full simplex 2^E is logically equal to $\mathcal{S}^{1,1}(M)$: there is only one possible line through any point in $\mathbb{F}^1$, so the projection map π is dense. Meanwhile, it is easy to see that $\mathcal{L}^{\infty}(M) = 2^E$, where $\mathcal{L}^{\infty}(M) := \lim_{m\to\infty} \mathcal{L}^m(M)$.

[1] That is, we assume that M contains no loops. Our results still hold—with trivial but notationally annoying modifications—when loops are present.

Remark 3.5. For a given matroid M and *irrational* number m, it is not hard to see that there exists a rational number $\tilde{m}$, chosen sufficiently close to m, such that $\mathcal{L}^{\tilde{m}}(M) = \mathcal{L}^m(M)$. Therefore, Corollary 3.3 actually gives a geometric interpretation for *every* instance of Laman independence.

Remark 3.6. Another surprising consequence of Corollary 3.3 is that (k, d)-slope-independence is invariant under simultaneously scaling k and d. That is, $\mathcal{S}^{k,d}(M) = \mathcal{S}^{ak,ad}(M)$ for every integer $a > 0$. Moreover, if d is divisible by k, then $m = \frac{d}{d-k}$ is an integer, and in fact $\mathcal{S}^{k,d}(M) = \mathcal{L}^m(M)$ is a matroid by Theorem 2.1 (i). It is far from clear what the geometry is behind these phenomena.

A natural question is to determine the singularities of the photo space. While we cannot do this in general, we can at least say exactly for which matroids $X_{k,d}(M)$ is smooth. The result and its proof are akin to [8, Proposition 15], and do not depend on the parameters k and d.

Proposition 3.7. *Let M be a loopless matroid equipped with a representation $\{v_1, \ldots, v_n\}$ as above. Then, for all integers $0 < k < d$, the photo space $X_{k,d}(M)$ is smooth if and only if each ground set element is either a loop or a coloop.*

Sketch of proof. If M consists solely of loops and coloops, then its photo space has the structure of an iterated fiber bundle over a point, in which every fiber is smooth (in fact, a copy of a projective space). Otherwise, one can explicitly describe the tangent space to $X_{k,d}(M)$ at a point in the degenerate cellule, and show that its dimension exceeds that of the photo space. □

4. Counting photos

Although it will not be needed in the sequel, we digress to prove an enumerative result, possibly of independent interest: when the field of representation of M is finite, the cardinality of the photo space $X_{k,d}(M)$ is an evaluation of the Tutte polynomial $T(M) = T_M(x, y)$. We refer the reader to [2] for details on the Tutte polynomial; roughly, it is the most general matroid isomorphism invariant satisfying the *deletion-contraction recurrence*

$$T(M) = T(M\backslash v) + T(M/v)$$

for every ground set element v that is neither a loop nor a coloop.

For $n \in \mathbb{N}$, define the q-analogues of n and $n!$ by

$$[n]_q := \frac{1-q^n}{1-q} = 1 + q + q^2 + \cdots + q^{n-1}, \qquad [n]!_q := [n]_q[n-1]_q \cdots [2]_q[1]_q,$$

and define the *q-binomial coefficient*

$$\begin{bmatrix} d \\ k \end{bmatrix}_q := \frac{[d]!_q}{[k]!_q[d-k]!_q}. \tag{12}$$

Theorem 4.1. *Let $\mathbb{F}$ be the finite field with q elements. Let M be a matroid of rank r, represented over $\mathbb{F}$ by vectors $v_1, \ldots, v_n$ spanning $\mathbb{F}^r$, and let $d \geq 2$. Then the number of (k, d)-photos of M is*

$$|X_{k,d}(M)| \;=\; \begin{bmatrix} d-1 \\ k-1 \end{bmatrix}_q^{r(M^\perp)} \left(q^k \begin{bmatrix} d-1 \\ k \end{bmatrix}_q \right)^{r(M)} T_M\left(\frac{\begin{bmatrix} d \\ k \end{bmatrix}_q}{\begin{bmatrix} d-1 \\ k \end{bmatrix}_q}, \frac{\begin{bmatrix} d \\ k \end{bmatrix}_q}{\begin{bmatrix} d-1 \\ k-1 \end{bmatrix}_q} \right).$$

Here $M^\perp$ denotes the *dual* or *orthogonal* matroid to M, defined combinatorially as the matroid on E whose bases are the complements of the bases of M.

The proof uses the commutative diagram

$$\begin{array}{ccc} X_{k,d}(M/e) \times \mathbb{G}r(k, \mathbb{F}^d) & \hookrightarrow & X_{k,d}(M) \\ \Big\downarrow{\scriptstyle\tilde{\pi}} & & \Big\downarrow{\scriptstyle\pi} \\ X_{k,d}(M/e) & \hookrightarrow & X_{k,d}(M-v) \end{array} \tag{13}$$

to give a deletion-contraction recurrence for $|X_{k,d}(M)|$. This recurrence can then be translated into a Tutte polynomial evaluation. The q-binomial coefficient (12) arises as the cardinality of the Grassmannian of k-planes in $\mathbb{F}^n$; see [14, Proposition 1.3.18]. The argument resembles that of [8, Theorem 1], in which the second author used a similar commutative diagram to express the Poincaré series of the picture space of a graph (over $\mathbb{C}$) as an analogous Tutte polynomial evaluation. (In contrast, when $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, the topology of the photo space is much simpler: there is a deformation retraction of $X_{k,d}(M)$ onto its degenerate cellule, which is homeomorphic to $\mathbb{G}r(k, \mathbb{F}^d)^n$.)

Since the Tutte polynomial of M does not depend on the choice of representation, neither does the number of photos. Moreover, there is a curious symmetry between the number of photos of a matroid M and of its dual $M^\perp$. Since $T_{M^\perp}(x,y) = T_M(y,x)$ [2, Prop. 6.2.4] and $\left[\begin{smallmatrix} d \\ k \end{smallmatrix}\right]_q = \left[\begin{smallmatrix} d \\ d-k \end{smallmatrix}\right]_q$, we have

$$q^{d\cdot r(M)}|X_{d-k,d}(M^\perp)| = q^{(d-k)n}|X_{k,d}(M)|. \tag{14}$$

A direct combinatorial explanation of this equality would be of interest.

5. Rigidity and parallel independence

In this section, we examine more closely the special cases $k = 1$ and $k = d-1$ of (k,d)-slope independence for a represented matroid M. It turns out that they are intimately related to the d-dimensional generic rigidity matroid $\mathcal{R}^d(M)$ and the d-dimensional generic hyperplane-marking matroid $\mathcal{H}^d(M)$. As before, let M be a matroid represented by vectors $E = \{v_1, \dots, v_n\}$ spanning $\mathbb{F}^r$, and let $d > 0$ be an integer.

5.1. **Interpreting $\mathcal{R}^d(M)$ and $\mathcal{H}^d(M)$.** Recall (Definition 1.4) that the *d-dimensional rigidity matroid* is represented over $\mathbb{F}(\varphi)$ by the vectors $\{v_i \otimes \varphi(v_i)\}_{i=1}^n \subset \mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}(\varphi)^d$. where $\mathbb{F}(\varphi)$ is the extension of $\mathbb{F}$ by dr transcendentals (the entries of the matrix $\varphi : \mathbb{F}^r \to \mathbb{F}(\varphi)^d$). The complex $\mathcal{R}^d(M)$ is defined to be the complex of independent sets of this matroid. The *d-rigidity matrix* $R^d(M)$ is the $n \times dr$ matrix whose rows represent $\mathcal{R}^d(M)$.

Recall also (Definition 1.5) that the *d-dimensional hyperplane-marking matroid* is represented over $\mathbb{F}(\varphi, n)$ by the vectors $\{v_i \otimes \eta_i\}_{i=1}^n \subset \mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}(\varphi,\eta)^d$. where $\mathbb{F}(\varphi)$ is the extension of $\mathbb{F}$ by $dr + (d-1)n$ transcendentals (the dr entries of the matrix φ, and the $(d-1)n$ coordinates of the normal vectors η_i to $\varphi(v_i)$). The complex $\mathcal{H}^d(M)$ is defined to be the complex of independent sets of this matroid. Denote by $H^d(M)$ the $n \times dr$ matrix whose rows represent $\mathcal{H}^d(M)$.

To interpret $R^d(M)$ and $H^d(M)$, we study their (right) nullspaces. Both matrices have row vectors in $\mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}^d$, so their nullvectors live in the same space. It will be convenient to freely use the identifications

$$\mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}^d \;\cong\; (\mathbb{F}^r)^* \otimes_{\mathbb{F}} \mathbb{F}^d \;\cong\; \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d).$$

The second of these isomorphisms is canonical; the first comes from identifying $\mathbb{F}^r$ and $(\mathbb{F}^r)^*$ by the standard bilinear form $\langle x, y\rangle = \sum_{i=1}^r x_i y_i$ on $\mathbb{F}^r$, whose associated quadratic form is $Q(x) = \langle x, x\rangle = \sum_{i=1}^r x_i^2$. With these identifications, one has

$$\langle v \otimes x, \psi\rangle = \langle x, \psi(v)\rangle.$$

for every $\psi \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d)$, $v \in \mathbb{F}^r$, and $x \in \mathbb{F}^d$. Using this fact, one can prove the following:

Proposition 5.1. *Let M be a matroid represented by E as above, and let $\psi \in \mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}^d \cong \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d)$.*

(i) *The vector ψ lies in $\ker H^d(M)$ if and only if $(\varphi + \psi)(v_i)$ is normal to η_i for every $i = 1, 2, \dots, n$.*

(ii) *Provided that $\mathbb{F}$ does not have characteristic 2, the vector ψ lies in $\ker R^d(M)$ if and only if*

$$Q\big((\varphi + \epsilon\psi)(v_i)\big) \equiv Q\big(\varphi(v_i)\big) \mod \epsilon^2$$

for every $i = 1, 2, \dots, n$.

Remark 5.2. Part (i) of Proposition 5.1 says that the nullspace of $H^d(M)$ is the space of directions in which one can perturb the map φ while keeping every image $\varphi(v_i)$ in the same hyperplane normal to η_i.

In contrast, part (ii) of Proposition 5.1 says that the nullspace of $R^d(M)$ is the space of infinitesimal changes that can be made to φ while keeping $Q(\varphi(v_i))$ constant (up to first order) for every i. (This is a rephrasing of a familiar fact from rigidity theory: the rigidity matrix $R^d(M)$ is just the Jacobian matrix (after scaling by $\frac{1}{2}$) of the map $\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d) \to \mathbb{F}^n$ sending φ to $Q(\varphi(v_i))_{i=1}^n$.)

Denote by $(d-1)M$ the matroid whose ground set consists of $d-1$ copies of each vector in E. The *d-parallel matrix* of M is defined as $H^d((d-1)M)$, and the matroid represented by its rows is the *d-dimensional parallel matroid* $\mathcal{P}^d(M) := \mathcal{H}^d((d-1)M)$. Part (ii) of Proposition 5.1 leads to an interpretation of the geometric meaning carried by the d-parallel matrix:

Corollary 5.3. *Let $\psi \in \mathbb{F}^r \otimes_{\mathbb{F}} \mathbb{F}^d \cong \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^r, \mathbb{F}^d)$. Then $\psi \in \ker P^d(M)$ if and only if $(\varphi+\psi)(v_i)$ is parallel to $\varphi(v_i)$ for all $i = 1, 2, \ldots, n$.*

Proof. Since there are $d-1$ copies of the vector v_i in $(d-1)M$, there will be $(d-1)$ accompanying normal vectors to $\varphi(v_i)$. Because these normals are chosen with generic coordinates, the only vectors normal to all $d-1$ of them are those parallel to $\varphi(v_i)$. Now apply Proposition 5.1. □

5.2. **The Nesting Theorem.** We now give one of our main results, the *Nesting Theorem*, which describes the relationship between the various independence systems associated to an arbitrary representable matroid M. In the special case that M is graphic and the ambient dimension d is 2, the Nesting Theorem gives what we have called the *planar trinity* (Corollary 5.5 below).

Theorem 5.4 (The Nesting Theorem)**.** *Let M be a matroid represented by vectors $E = \{v_1, \ldots, v_n\} \subset \mathbb{F}^r$, and let $d > 1$ be an integer. Then*

$$\mathcal{S}^{1,d}(M) \subseteq \mathcal{R}^d(M) \subseteq \mathcal{L}^d(M) = \mathcal{H}^d(M) \quad (= \mathcal{S}^{d-1,d}(M)).$$

Sketch of proof. To prove that $\mathcal{R}^d(M) \subseteq \mathcal{L}^d(M)$, it suffices to show that whenever $d \cdot r(M) \leq n$, there is an $\mathbb{F}(\varphi)$-linear dependence among the n rows of $R^d(M)$. The construction of $R^d(M)$ implies that these rows lie in a $\mathbb{F}(\varphi)$-vector space of dimension $d \cdot r(M)$. Thus if $d \cdot r(M) < n$, then the desired linear dependence is immediate, while if $d \cdot r(M) = n$, then the form of $R^d(M)$ allows us to exhibit an explicit nullvector. The proof that $\mathcal{H}^d(M) \subseteq \mathcal{L}^d(M)$ is analogous.

To prove that $\mathcal{S}^{1,d}(M) \subseteq \mathcal{R}^d(M)$, we assume that the rows of $R^d(M)$ are dependent and show that M is (k,d)-slope dependent for $k = 1$. Note that $\mathcal{S}^{k,d}(M) = \mathcal{L}^{\frac{d}{d-k}}(M) \subset \mathcal{L}^d(M)$. The equality is Corollary 3.3, and the inclusion follows from the definition of $\mathcal{L}^m(M)$ (because $\frac{d}{d-k} \leq d$). In particular, if M is d-Laman dependent then M is automatically (k,d)-slope dependent; we may therefore assume that M is d-Laman independent. Without loss of generality, $d \cdot r(M) \geq n$, so the dependence of the rows of $R^d(M)$ implies the vanishing of *every* one of its $n \times n$ minors. Moreover, by Theorem 2.5, M admits a d-Edmonds decomposition (see Definition 2.4).

Using the combinatorial properties of an Edmonds decomposition, we construct an $n \times n$ minor ξ of $R^d(M)$ that is a nonzero multihomogeneous polynomial in the coordinates of the vectors $\varphi(v_i)$. If ξ vanishes on the non-annihilating cellule $X^{\varnothing}_{k,d}(M)$ of the photo space, then the projection on $X^{\varnothing}_{k,d}(M) \to \mathbb{G}r(k, \mathbb{F}^d)$ is not Zariski dense, because the homogeneous coordinates of the $\varphi(v_i)$ are in fact the Plücker coordinates on $\mathbb{G}r(k, \mathbb{F}^d)$. This observation, together with Theorem 3.2, implies that $\mathcal{S}^{1,d}(M) \subseteq \mathcal{R}^d(M)$.

Replacing $R^d(M)$ with $H^d(M)$, $k = 1$ with $k = d-1$, and $\varphi(v_i)$ with η_i throughout, the same argument shows that $\mathcal{S}^{d-1,d} \subset \mathcal{H}^d(M)$. This completes the proof, since $\mathcal{S}^{d-1,d}(M) = \mathcal{L}^d(M)$ by Corollary 3.3. □

The case $d = 2$ is very special. Recall that $\mathcal{P}^d(M) = \mathcal{H}^d((d-1)M)$, so $\mathcal{P}^2(M) = \mathcal{H}^2(M)$. Indeed, setting $d = 2$ in the Nesting Theorem gives the following equalities:

Corollary 5.5. $\mathcal{S}^{1,2}(M) = \mathcal{R}^2(M) = \mathcal{L}^2(M) = \mathcal{H}^2(M) = \mathcal{P}^2(M)$.

When $d \geq 3$, the inclusion $\mathcal{R}^d(M) \subset \mathcal{L}^d(M)$ is typically strict. The nullspace of $R^d(M)$ contains the $\binom{d}{2}$-dimensional space of all vectors of the form $\sigma \circ \varphi$, as σ ranges over all skew-symmetric matrices in $\mathbb{F}^{d \times d}$. Consequently, every d-rigidity-independent subset $A \subset E$ must satisfy $|A| \leq d \cdot r(A) - \binom{d}{2}$. On the other hand, there may exist d-Laman independent sets A of cardinality up to $d \cdot r(A) - 1$.

6. Examples: Uniform matroids

Let E be a ground set with n elements. The *uniform matroid* of rank r on E is defined to be the matroid $U_{r,n}$ with independent sets $\{F \subset E : \ |F| \leq r\}$. $U_{r,n}$ may be regarded as the matroid represented by n generically chosen vectors in $\mathbb{F}^r$, where $\mathbb{F}$ is a sufficiently large field.

Predictably, the d-Laman and (k,d)-slope independence complexes on $U_{r,n}$ are also uniform matroids:

$$\mathcal{L}^d(U_{r,n}) = U_{s,n}, \quad \mathcal{S}^{k,d}(U_{r,n}) = U_{t,n} \qquad \text{where} \quad s = \min(\lceil dr-1\rceil, n),\ t = \min(\lceil \tfrac{dr}{d-k} - 1\rceil, n). \tag{15}$$

More striking is that d-Laman independence carries nontrivial geometric information about sets of n generic vectors in r-space: coplanarity for $U_{2,3}$ and the cross-ratio for $U_{2,4}$.

Example 6.1 ($U_{2,3}$)**.** Let $\mathbb{F}$ be any field, and let e_1, e_2 be the standard basis vectors in $\mathbb{F}^2$. The matroid $M = U_{2,3}$ is represented by the vectors $\{e_1, e_1+e_2, e_2\} \subset \mathbb{F}^2$; this representation is unique up to the action of the projective general linear group. By (15),

$$\mathcal{L}^d(U_{2,3}) = \begin{cases} U_{2,3} & \text{if } d \in (1, \frac{3}{2}]_{\mathbb{R}} \\ U_{3,3} & \text{if } d \in (\frac{3}{2}, \infty)_{\mathbb{R}} \end{cases} \qquad \text{and} \qquad \mathcal{S}^{1,d}(U_{2,3}) = \begin{cases} U_{3,3} & \text{if } d = 2 \\ U_{2,3} & \text{if } d \in \{3,4,\ldots\}. \end{cases}$$

We now consider what these equalities mean in terms of slopes. Let $\varphi : \mathbb{F}^2 \to \mathbb{F}^d$ be a linear transformation. If $d = 2$, then the images $\varphi(e_1), \varphi(e_1+e_2), \varphi(e_2)$ can have arbitrary slopes as φ varies. This is why $\mathcal{S}^{1,2}(U_{2,3}) = U_{3,3}$. On the other hand, when $d \geq 3$, those three vectors must be coplanar. This imposes a nontrivial constraint on the homogeneous coordinates for the lines spanned by the three images, and explains why $\mathcal{S}^{1,d}(U_{2,3}) = U_{2,3}$. By direct calculation, the inclusions $\mathcal{R}^d(M) \subseteq \mathcal{L}^d(M)$ given by Theorem 5.4 turn out to be equalities.

Example 6.2 ($U_{2,4}$)**.** Let $\mathbb{F}$ be a field of cardinality > 2, let $\mu \in \mathbb{F} \setminus \{0,1\}$, and let e_1, e_2 be the standard basis vectors in $\mathbb{F}^2$. The four vectors $\{e_1, e_1+e_2, e_2, e_1+\mu e_2\}$ represent $M = U_{2,4}$ over $\mathbb{F}$. Again, this representation is unique up to projective equivalence. By (15),

$$\mathcal{L}^d(U_{2,4}) = \begin{cases} U_{2,4} & \text{if } d \in (1, \frac{3}{2}]_{\mathbb{R}} \\ U_{3,4} & \text{if } d \in (\frac{3}{2}, 2]_{\mathbb{R}} \\ U_{4,4} & \text{if } d \in (2, \infty)_{\mathbb{R}} \end{cases} \qquad \text{and} \qquad \mathcal{S}^{1,d}(U_{2,4}) = \begin{cases} U_{3,4} & \text{if } d = 2 \\ U_{2,4} & \text{if } d \in \{3,4,\ldots\}. \end{cases}$$

Why is this correct from the point of view of slopes? From Example 6.1, we know that when $d \geq 3$, the lines spanned by the images of any three of the four vectors must be coplanar, so there is an algebraic dependence among the homogeneous coordinates for these three lines. For $d = 2$, this does not happen; the slopes of the images of any triple can be made arbitrary. However, applying a linear transformation to the representing vectors does not change their *cross-ratio* (in this case μ), so the fourth image vector is determined by the first three. This is the geometric interpretation of the combinatorial identity $\mathcal{S}^{1,2}(U_{2,4}) = U_{3,4}$.

Direct calculation shows that

$$\mathcal{R}^d(U_{2,4}) = \begin{cases} U_{2,4} & \text{if } d = 1, \\ U_{3,4} & \text{if } d \in \{2,3,\ldots\}. \end{cases}$$

This calculation is independent of the particular coordinates chosen for the representing vectors, even up to projective equivalence (that is, up to the choice of the parameter μ): that is, $\mathcal{R}^d(U_{2,4})$ is a *combinatorial* invariant. On the other hand, unlike the situation for $U_{2,3}$, the inclusions $\mathcal{R}^d(M) \subset \mathcal{L}^d(M)$ given by Theorem 5.4 are strict. (This behavior deviates from the case of graphic matroids; see below.)

7. More about $\mathcal{R}^d(M)$: invariance and stabilization

The examples in the previous section raise some natural questions. Clearly $\mathcal{L}^m(M)$ is a combinatorial invariant of M (that is, it does not depend on the choice of representation), so by Corollary 3.3 the same is true for $\mathcal{S}^{k,d}(M)$ (and in particular $\mathcal{H}^d(M)$ and $\mathcal{P}^d(M)$). But what about $\mathcal{R}^d(M)$? Note that this issue does not arise in classical rigidity theory, where the graphic matroid $M(G)$ is always represented by the vectors $\{e_i - e_j : \{i,j\} \in E(G)\}$, where e_i is the i^{th} standard basis vector in $\mathbb{R}^{|V(G)|}$.

Question 7.1. Is $\mathcal{R}^d(M)$ a combinatorial invariant of M, or does it depend on the choice of representation $\{v_1, \ldots, v_n\}$?

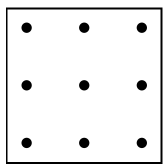
In the special case $d = 2$, the Nesting Theorem implies that $\mathcal{R}^d(M)$ is indeed a combinatorial invariant.

Call two sets of vectors $E = \{v_1, \ldots, v_n\}$, $E' = \{v'_1, \ldots, v'_n\} \subset \mathbb{F}^r$ *projectively equivalent* if there are nonzero scalars $c_1, \ldots, c_n \in \mathbb{F}^\times$ and an invertible linear transformation $g \in GL_r(\mathbb{F})$, such that $v'_i = g(c_i v_i)$

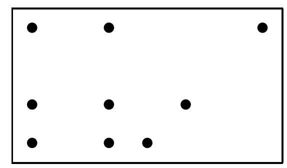
for every i. Then the matroids represented by E and E' are combinatorially identical. It is not hard to prove that their corresponding rigidity matroids are also identical. Unfortunately, this fact provides little insight into Question 7.1, because projective equivalence is a very strong condition.

On the other hand, we have not found a counterexample. We have seen that $\mathcal{R}^d(M)$ is indeed a combinatorial invariant for all d when $M = U_{2,3}$ or $U_{2,4}$. As another example, consider the following two sets of nine coplanar vectors in $\mathbb{R}^3$:

$$E = \{(1,0,0),(1,0,1),(1,0,2),(1,1,0),(1,1,1),(1,1,2),(1,2,0),(1,2,1),(1,2,2)\},$$
$$E' = \{(1,0,0),(1,0,1),(1,0,3),(1,2,0),(1,2,1),(1,2,3),(1,3,0),(1,4,1),(1,6,3)\}.$$



E



E'

Let M, M' be the matroids represented by E, E' respectively. These matroids are combinatorially isomorphic but projectively inequivalent. On the other hand, computations using `Mathematica` have shown that $\mathcal{R}^2(M) = \mathcal{R}^2(M')$ and $\mathcal{R}^3(M) = \mathcal{R}^3(M')$.

It is not hard to show that $\mathcal{R}^d(M) \subset \mathcal{R}^{d+1}(M)$ for all represented matroids M and integers d. Since there are only finitely many simplicial complexes on E, the tower $M = \mathcal{R}^1(M) \subseteq \mathcal{R}^2(M) \subseteq \mathcal{R}^3(M) \subseteq \cdots$ must eventually stabilize. Using standard facts about the transcendence degree of field extensions, we prove that this stabilization occurs no later than the rank $r = r(M)$: that is,

$$\mathcal{R}^d(M) = \mathcal{R}^r(M) \qquad \text{for all } d \geq r.$$

Moreover, if M is the graphic matroid for a graph $G = (V, E)$, equipped with the standard representation $\{e_i - e_j : ij \in E\}$ over an arbitrary field, then $\mathcal{R}^r(M)$ is the Boolean matroid on E.

This observation begs the question of whether $\mathcal{R}^d(M(G))$ depends on the field *before* d reaches the stable range. For an arbitrary representable matroid M, it is not true in general that $\mathcal{R}^\infty(M)$ is Boolean. We have already seen one example for which this fails, namely $U_{2,4}$. Another example is the well-known *Fano matroid* F, represented over the two-element field $\mathbb{F}_2$ by the seven nonzero elements of $\mathbb{F}_2^3$. It is not hard to show that $\mathcal{L}^d(F)$ is Boolean for $d > \frac{7}{3}$. On the other hand, computation with `Mathematica` indicates that $\mathcal{R}^2(F) = U_{5,7}$, but $\mathcal{R}^d(F) = U_{6,7}$ for all integers $d \geq 3$.

8. Open problems

The foregoing results raise many questions that we think are worthy of further study. Some of these have been mentioned earlier in the paper. In this final section, we restate the open problems and add a few more.

Problem 1. Determine the singular locus of the (k,d)-photo space $X_{k,d}(M)$ (perhaps by calculating the dimension of its various tangent spaces, as in Proposition 3.7).

Problem 2. Give a direct combinatorial explanation for the identity (14) (presumably by identifying some natural relationship between photos of M and of $M^\perp$).

Problem 3. Explain the "scaling phenomenon" of Remark 3.6 geometrically.

Problem 4. Determine whether or not the d-rigidity matroid $\mathcal{R}^d(M)$ is a combinatorial invariant of M (Question 7.1). If not, determine which matroids have this property, and to what extent $\mathcal{R}^d(M)$ depends on the field over which M is represented.

Problem 5. Let $M(G)$ be a graphic matroid equipped with the standard representation $\{e_i - e_j : \{i,j\} \in E(G)\}$. Is the rigidity matroid $\mathcal{R}^d(M)$ independent of the ambient field $\mathbb{F}$ for all d?

Problem 6. Generalize other aspects of classical (graph) rigidity theory to non-graphic matroids. One example is Crapo's "$(d+1)\mathbf{T}d$" characterization of the hyperplane-marking matroid of a graph [18, Theorem 8.2.2]. Another is Henneberg's construction of the bases of the 2-rigidity matroid [18, Theorem 2.2.3].

Our last open problem is similar in spirit to the results of [7] and [9], describing the algebraic and combinatorial structure of the equations defining the slope variety of a graph. It is motivated also by the appearance of the cross-ratio in Example 6.2.

Problem 7. Describe explicitly the defining equations (in Plücker coordinates on $\mathbb{G}r(k, \mathbb{F}^d)^n$) for $\overline{\pi X^{\varnothing}_{k,d}(M)}$, where π is the projection map of (11).

Acknowledgments

A substantial portion of this work was completed in July 2004 at the Park City Mathematical Institute, sponsored by the Institute for Advanced Study. The authors thank Neil White for helpful discussions while at PCMI, Walter Whiteley for sharing with them some of his manuscripts on classical rigidity theory, and Ezra Miller for discussions on algebraic geometry.

References

[1] M. Aigner, *Combinatorial Theory.* Springer-Verlag, Berlin, 1997.

[2] T. Brylawski and J. Oxley, *The Tutte polynomial and its applications*, Matroid applications, 123–225, Encyclopedia Math. Appl. **40**, Cambridge Univ. Press, Cambridge, 1992.

[3] M. Develin, J.L. Martin and V. Reiner, *Rigidity theory for matroids*, preprint, `arXiv:math.CO/0503050`.

[4] J. Edmonds, *Minimum partition of a matroid into independent subsets*, J. Res. Nat. Bur. Standards Sect. B **69B** (1965), 69–72.

[5] J. Graver, B. Servatius, and H. Servatius, *Combinatorial rigidity.* Graduate Studies in Mathematics, 2. American Mathematical Society, Providence, RI, 1993.

[6] G. Laman, *On graphs and rigidity of plane skeletal structures*, J. Engrg. Math. **4** (1970), 331–340.

[7] J.L. Martin, *Geometry of graph varieties*, Trans. Amer. Math. Soc. **355** (2003), 4151-4169.

[8] J.L. Martin, *On the topology of graph picture spaces*, Adv. Math. **191**, no. 2 (2005), 312–338.

[9] J.L. Martin, *The slopes determined by n points in the plane*, Duke Math. J., to appear; `arXiV:math.AG/0302106`.

[10] C.St.J.A. Nash-Williams, *Edge-disjoint spanning trees of finite graphs*, J. London Math. Soc. **36** (1961), 445–450.

[11] C.St.J.A. Nash-Williams, *Decompositions of finite graphs into forests*, J. London Math. Soc. **39** (1964), 12.

[12] J.G. Oxley, *Matroid theory.* Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1992.

[13] A. Recski, *A network theory approach to the rigidity of skeletal structures. II. Laman's theorem and topological formulae.* Discrete Appl. Math. **8** (1984), no. 1, 63–68.

[14] R.P. Stanley, *Enumerative combinatorics, Vol. I*, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, Cambridge, 1997.

[15] T.-S. Tay, *A new proof of Laman's theorem*, Graphs Combin. **9** (1993), 365–370.

[16] W.T. Tutte, *On the problem of decomposing a graph into n connected factors.* J. London Math. Soc. **36** (1961), 221–230.

[17] D.J.A. Welsh, *Matroid theory.* London Math. Soc. Monographs **8**, Academic Press [Harcourt Brace Jovanovich, Publishers], London–New York, 1976.

[18] W. Whiteley, *Some matroids from discrete applied geometry*, Contemp. Math. **197** (1996), 171–311.

[19] W. Whiteley, *Rigidity and scene analysis*, Handbook of discrete and computational geometry, 893–916, CRC Press Ser. Discrete Math. Appl., CRC, Boca Raton, FL, 1997.

Mike Develin, American Institute of Mathematics, 360 Portage Ave., Palo Alto, CA 94306-2244, USA
E-mail address: `develin@post.harvard.edu`

Jeremy L. Martin, School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
E-mail address: `martin@math.umn.edu`

Victor Reiner, School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
E-mail address: `reiner@math.umn.edu`

The Weak Order on Pattern-Avoiding Permutations

Brian Drake

Abstract. We consider a partial order on permutations avoiding a set of patterns. The partial order is induced from the weak order of the symmetric group. Some sets of patterns are shown to give well-known posets, including the Tamari lattice, the Boolean lattice, $J(2{\times}n)$, the integer interval lattice, and the lattice of shifted shapes. In the case of a single pattern, we characterize those patterns which give a lattice.

Résumé. Nous considérons un ordre partiel sur des permutations évitant un ensemble des motifs. L'ordre partiel est induit de l'ordre faible du groupe symétrique. Quelques ensembles des motifs sont montres à donner les posets bien connus, y compris le trellis de Tamari, le trellis de Boole, $J(2 \times n)$, et quelques autres trellis. Dans le cas d'un motif simple, nous caractérisons ces motifs qui donnent un trellis.

1. Introduction

As a partially ordered set, the weak order on a finite Coxeter group is a lattice. In some interesting cases this property is retained when passing to an induced subposet. For example, one can obtain the one-skeleta of generalized associahedra of Fomin and Zelevinski [**5**], and the Cambrian lattices of Reading [**9**] as subposets of the weak order. In type A, some of these subposets can be described using pattern avoidance. The Tamari lattice, the one-skeleton of the associahedron, is isomorphic to the weak order on 312 avoiding permutations. The Boolean lattice is isomorphic to the weak order on 312 and 231 avoiding permutations. See [**9**] for these results and some corresponding type B results.

These results lead to two natural questions. For which sets of patterns T is the weak order on permutations avoiding T a lattice? Also, can any other well-known families of lattices be obtained as the weak order on pattern avoiding permutations?

This paper is organized as follows. Section 2 contains some preliminaries about pattern avoidance, lattices, and the weak order. In section 3, we show that $J(2 \times n) \cup \hat{1}$, the integer interval lattice, and the lattice of shifted shapes can be obtained as the weak order on pattern avoiding permutations. We also give two unnamed lattices which may be of independent interest. In section 4 we outline the proof of the following theorem:

Theorem 1.1. *$S_n(\tau)$ is a lattice for all n if and only if*

τ has exactly one descent, which is of magnitude one or two, or
τ has exactly one ascent, which is of magnitude one or two.

In section 5 we give some related results and a conjecture.

2. Preliminaries

A permutation $\pi = \pi(1)\pi(2)\cdots\pi(n) \in S_n$ in the symmetric group on n elements is said to *contain* a pattern τ if there is a subsequence of π in the same relative order as τ. Otherwise, π is said to *avoid* τ. For example, the permutation $\pi = 1423$ contains the pattern 132 twice, as 142 and 143. The permutation 2134 avoids 132. For a set of patterns T, we will use $S_n(T)$ to denote the permutations of length n which avoid all of the patterns in T. For simplicity of notation, we omit the set brackets for a single pattern. There has been much recent interest in enumerative problems in pattern avoidance. For an overview, see [**15**].

A *descent* in a permutation π occurs in position i if $\pi(i) > \pi(i+1)$. The *magnitude* of a descent is $\pi(i+1) - \pi(i)$. Ascents and their magnitudes are defined similarly. An *inversion* in a permutation π is a pair (i, j) such that $i < j$ and $\pi(j) > \pi(i)$.

The weak order on permutations is as follows. For $\pi, \sigma \in S_n$, we say that σ covers π if there is an adjacent transposition $(i, i+1)$ such that $\pi(i, i+1) = \sigma$, and σ has more inversions than π. The weak order is the transitive closure of this relation. Alternatively, we could describe this as $\pi \leq \sigma \iff \pi$ and σ can be written as products of adjacent transpositions, with the product for π appearing as a prefix of the product for σ. We will let $S_n(T)$ denote the set $S_n(T)$ together with its order relation induced from the weak order. See [**2**] for more information on the weak order of Coxeter groups.

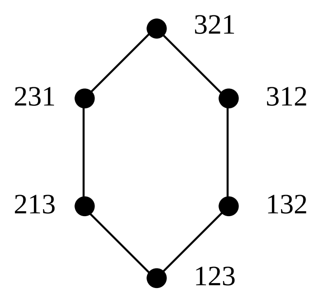


FIGURE 1. The weak order on S_3

A *lattice* L is a partially ordered set with the following property:

For all $x, y \in L$, the set $\{z \in L | z \geq x, z \geq y\}$ has a unique minimal element, called the *join* of x and y and denoted $x \vee y$, and the set $\{z \in L | z \leq x, z \leq y\}$ has a unique maximal element, called the *meet* of x and y and denoted $x \wedge y$.

For background information on lattices, see [**1**] or Chapter 3 of [**11**].

The lattice property of the weak order on S_n was demonstrated in [**7**] and [**16**]. There is a well-known characterization of the weak order which is useful for constructing meets and joins.

LEMMA 1. *For $\sigma \in S_n$, let $I(\sigma) = \{(i, j) |\ i < j, \sigma(j) < \sigma(i)\}$ be the inversion set of σ. Then $\pi \leq \sigma \Leftrightarrow I(\pi) \subseteq I(\sigma)$.*

Let $\pi, \sigma \in S_n$. We will construct the join $\pi \vee \sigma$. First, insert the letter 1. To insert the letter j, insert it immediately to the left of the largest i such that $(i, j) \in I(\sigma) \cup I(\pi)$. If no such i exists, insert j on the right. This gives the unique minimal permutation ω with $I(\pi) \subseteq I(\omega)$ and $I(\sigma) \subseteq I(\omega)$. Since S_n also has a unique minimal element, it is a lattice by the following lemma. This is [**11**], Proposition 3.3.1.

LEMMA 2. *Let P be a finite poset with a unique minimal element. If the join of every pair of elements in P exists, then P is a lattice.*

To consider induced subposets of the weak order, we will use the following easy consequence of Lemma 2.

LEMMA 3. *Let L be a lattice and P an induced subposet. If the following two conditions hold, then P is a lattice:*
1) For all $v \in L \setminus P$, the set $\{v' \in P \mid v' < v\}$ has a unique maximal element, or is the empty set.
2) P has a unique maximal element and a unique minimal element.

We will make use of the following notation. If $\pi \in S_m$ and $\sigma \in S_n$, then $\pi \oplus \sigma$ denotes the permutation in S_{m+n}, where π acts on the first m letters, and σ acts on the last n letters. Similarly, $\pi \ominus \sigma$ denotes the permutation in S_{m+n}, where π acts on the last m letters, and σ acts on the first n letters.

3. Examples of Lattices of Pattern-Avoiding Permutations

To find examples, we tested all sets of patterns including any number of patterns of length three and at most one pattern of length four. Here we give all the resulting lattices, except for the n element chain. In each example, only one representative set of patterns is given.

1. The Tamari Lattice
The Tamari lattice was defined in [**14**] in terms of legal bracketings. See also [**9**]. It can be realized as the weak order on 132 avoiding permutations. This can be seen by using Stanley's representation of the Tamari lattice given in ([**12**], exercise 6.23), and Krattenthaler's bijection [**8**].

2. The Boolean Lattice
The Boolean lattice is isomorphic to the weak order on {132, 213} avoiding permutations. See [**9**].

3. The Lattice of Shifted Shapes
A *shifted shape* is a finite set Q of pairs (i, j), $i < j$, with the following property: If $(i, j) \in Q$, then $(k, j) \in Q \ \forall\ k < i$, and $(i, l) \in Q \ \forall\ l < j$.

Shifted shapes can be thought of as diagrams fitting on top of a staircase (see figure 2). The partial order on shifted shapes with $j \leq n \ \forall\ (i, j) \in Q$ is inclusion of sets. See [**6**] and the references given there.

The lattice of shifted shapes can be realized as $S_n(\{132, 312\})$. A bijection between permutations avoiding 132 and 312 and shifted shapes is given mapping a permutation to its inversion set. The condition σ avoids 132 is equivalent to the first condition for a shifted shape. The condition σ avoids 312 is equivalent to the second.

4. The Integer Interval Lattice
Taking all closed intervals contained in $[1, n]$ with integer endpoints and ordering them by inclusion gives a lattice. This lattice is isomorphic to $S_n(\{231, 312, 2143\})$. Figure 3 shows this lattice for $n = 5$. Permutations avoiding this set of patterns are determined by the pairs $(i, i+1)$ in their inversion sets. Also, for $\sigma \in S_n(\{231, 312, 2143\})$, $(i, i+1), (k, k+1) \in I(\sigma)$ implies $(j, j+1) \in I(\sigma)$ for all $i < j < k$.

5. $J(2 \times n) \cup \hat{1}$
The lattice of order ideals of the poset $2 \times n$ appears often in combinatorics. See [**3**] and [**11**]. This lattice (with an extra maximal element) is isomorphic to $S_n(\{132, 312, 2314\})$. This can

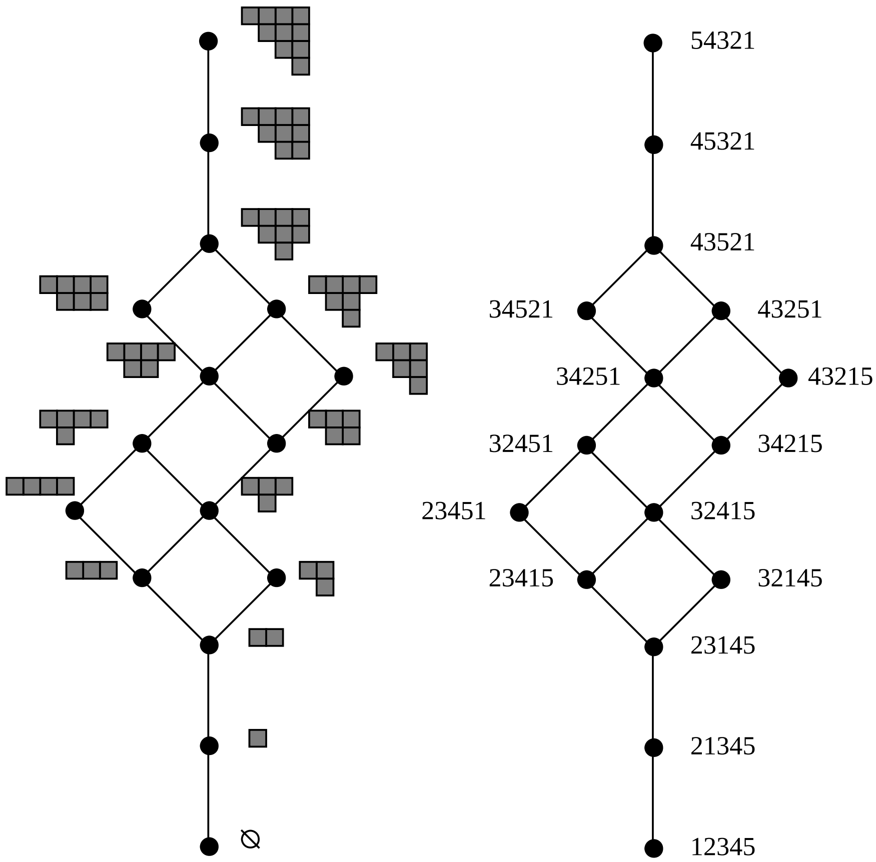


FIGURE 2. $S_n(\{132, 312\})$, the shifted shape lattice

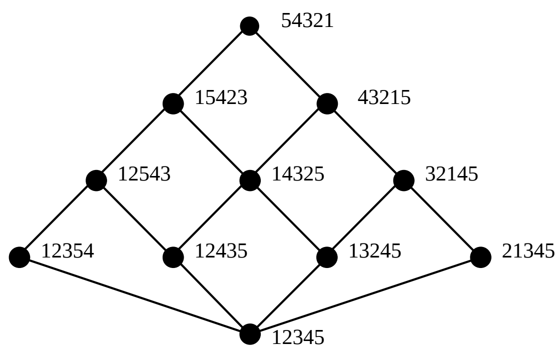


FIGURE 3. $S_n(\{231, 312, 2143\})$, the integer interval lattice

be easily shown by constructing $S_n(\{132, 312, 2314\})$ from $S_{n-1}(\{132, 312, 2314\})$. The same process constructs $J(2 \times n) \cup \hat{1}$ from $J(2 \times (n-1)) \cup \hat{1}$. See figure 4.

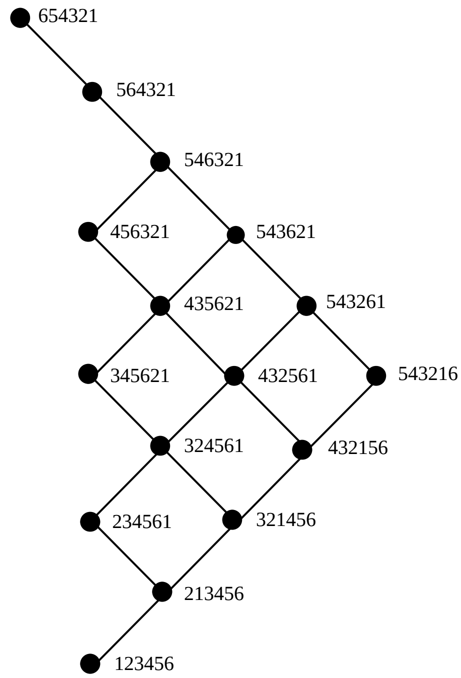


FIGURE 4. $S_n(\{132, 312, 2314\}) \cong J(2 \times n) \cup \hat{1}$

6. A "leaf with ridges" lattice.

The lattice $S_n(\{132, 213, 3421\})$ is graded, and may have some other interesting properties. It has $\binom{n}{2} + 1$ elements. When its Hasse diagram is drawn as in figure 5, it looks somewhat like a leaf with a series of ridges rising out of it.

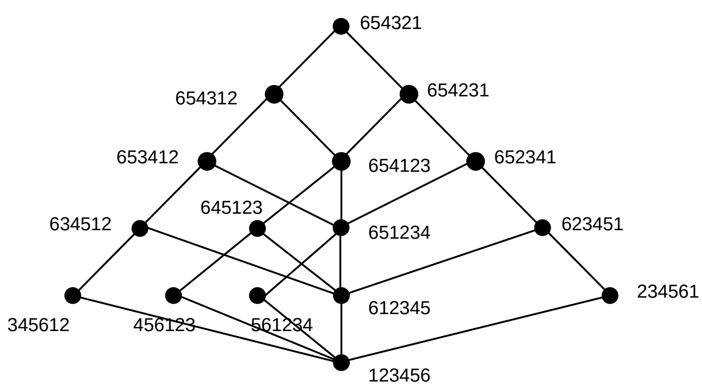


FIGURE 5. $S_n(\{132, 213, 3421\})$

7. A lattice with Fibonacci-many elements.

The lattice $S_n(\{231, 312, 1432\})$ is also graded. It has $F_{n+2} - 1$ elements, where F_n denotes the nth Fibonacci number. Its Hasse diagram is drawn in figure 6 to highlight how the nth lattice in the sequence can be constructed from the $(n-2)$nd and $(n-1)$st. To see this, remove the gray edges.

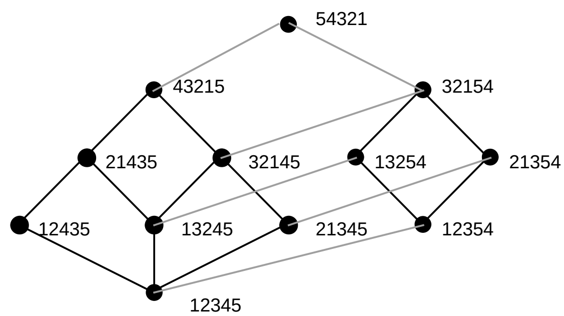


FIGURE 6. $S_n(\{231, 312, 1432\})$

4. Proof Outline of Theorem 1.1

Let τ be a pattern of length k. We show the result in four steps. First, we show that if τ has at least two descents and at least two ascents, then $S_n(\tau)$ is not a lattice for $n \geq k$. To do this, we find the subposet in figure 6, where the edges are covering relations. Then taking the induced subposet in $S_k(\tau)$, shown in figure 7, we find that σ and $\sigma\prime$ do not have a join in $S_k(\tau)$. We can find an isomorphic subposet in S_n for all $n \geq k$.

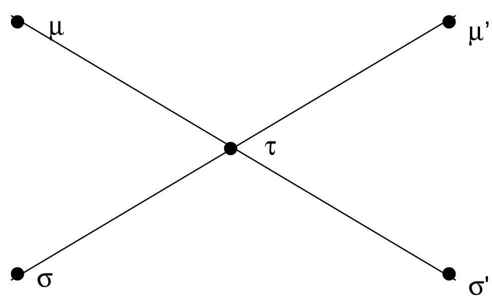


FIGURE 7. Subposet of S_k

Second, we show that if τ has at least two ascents and a descent of magnitude greater than 2, (or vice versa), then $S_n(\tau)$ is not a lattice for $n \geq k+1$. We find similar subposets as in figures 6 and 7, except here the edges are intervals which are chains.

Third, if τ satisfies the conditions of the theorem, we will show that $\pi \in S_n \setminus S_n(\tau)$ implies that the set $\{\pi\prime \leq \pi \mid \pi\prime \in S_n(\tau)\}$ has a maximal element.

Finally we invoke lemma 3 to complete the proof.

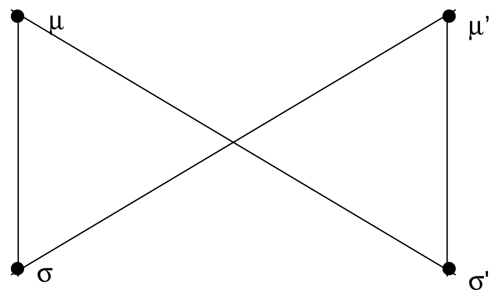


FIGURE 8. Subposet of $S_k(\tau)$

5. Related Results

First let us note that there is an analogous result about meet and join semi-lattices.

THEOREM 5.1. *$S_n(\tau)$ is a meet semi-lattice if and only if τ has at most one descent, which is of magnitude one or two.*

$S_n(\tau)$ is a join semi-lattice if and only if τ has at most one ascent, which is of magnitude one or two.

This implies that the only τ for which $S_n(\tau)$ is a semi-lattice but not a lattice are the strictly increasing and strictly decreasing patterns.

Theorem 1.1 does not generalize immediately to larger sets of patterns. In particular, it is not true that $S_n(\{\tau_1, \tau_2\})$ is a lattice if both $S_n(\tau_1)$ and $S_n(\tau_2)$ are. For example, Stembridge's posets package for Maple [**13**] confirms that $S_5(\{2431, 3124\})$ is not a lattice. Moreover, it is not necessary for both $S_n(\tau_1)$ and $S_n(\tau_2)$ to be lattices in order for $S_n(\{\tau_1, \tau_2\})$ to be a lattice. For example, consider $S_n(\{2134, 2143\})$, which is one case of the following theorem:

THEOREM 5.2. *Let $T = \{21 \oplus \tau \mid \tau \in S_{k-2}\}$. Then $S_n(T)$ is a lattice for all n.*

Proof: Observe that $S_n(T)$ is the set of permutations such that for each descent $\pi(i) > \pi(i+1)$, we have $|\{j \mid j > i, \pi(j) > \pi(i)\}| < k-2$. So if $\pi \in S_n \setminus S_n(T)$, then there is a unique minimal element $\pi\prime$ less than π (in terms of the order on S_n), with $\pi\prime \in S_n(T)$. $\pi\prime$ is obtained by changing all descents which violate the condition above to ascents. Since $12\ldots n \in S_n(T)$ and $n(n-1)\ldots 21 \in S_n(T)$, $S_n(T)$ is a lattice by lemma 3. □

It is probably unfeasable to characterize all sets of patterns T such that $S_n(T)$ is a lattice for all sufficiently large n. However, the following corollary of the proof for Theorem 1.1 might be easier to generalize.

COROLLARY 5.1. *If τ is a pattern of length k, the following are equivalent:*
1) $S_n(\tau)$ is a lattice for all n.
2) $S_{k+1}(\tau)$ is a lattice.

CONJECTURE 1. *There exists an M depending only on the length of the patterns τ_i such that the following are equivalent:*
1) $S_n(\tau_1, \tau_2, \ldots, \tau_r)$ is a lattice for all $n \geq M$.
2) $S_M(\tau_1, \tau_2, \ldots, \tau_r)$ is a lattice.

The Erdös-Szekeres Theorem [**4**] would suggest an M that is roughly the product of the length of the patterns τ_i.

References

[1] G. Birkhoff, *Lattice Theory*, American Mathematical Society Colloquium Publications, Vol 25, 1948.
[2] A. Björner, *Orderings of Coxeter Groups*, in"Contemporary Mathematics" **34** (1984), 175–195.
[3] P. Brändén, *q-Narayana numbers and the flag h-vector of* $J(2 \times n)$, Discrete Mathematics **281** (2004), 67-81.
[4] P. Erdös and G. Szekeres, *A combinatorial theorem in geometry*, Compositio Math **2** (1935), 463-470.
[5] S. Fomin and A. Zelevinsky, *Y-systems and Generalized Associahedra*, Annals of Mathematics **158** (2003), 977-1018.
[6] S. Fomin and D. Stanton, *Rim Hook Lattices*, St. Petersburg Mathematical Journal **9** (1998), 1007-1016.
[7] G.Th. Guilbaud and P. Rosenstiehl, *Analyse algébrique d'un scrutin*, in "Ordres totaux finis", Gauthiers-villars et Mouton, Paris, 1971, pp. 71-100.
[8] C. Krattenthaler, *Permutations with restricted patterns and Dyck paths*, Adv. in Appl. Math. **27** (2001), 510–530.
[9] N. Reading, *Cambrian Lattices*, preprint, 2004.
[10] R. Simion and F. Schmidt, *Restricted Permutations*, Europ. J. Combinatorics **6** (1985), 383–406.
[11] R. Stanley, *Enumerative Combinatorics*, vol. 1. Cambridge University Press, Cambridge, 1999.
[12] R. Stanley, *Enumerative Combinatorics*, vol. 2. Cambridge University Press, Cambridge, 1999.
[13] J. Stembridge *A maple package for posets*, preprint, 2002.
[14] D. Tamari, *The algebra of bracketings and their enumeration*, Nieuw Arch. Wisk. **10** (1962), 131–146.
[15] H. Wilf, *The patterns of permutations*, Discrete Math **257** (2002) no. 2-3, 575–583.
[16] T. Yanagimoto and M. Okamoto, *Partial orderings of permutations and monotonicity of a rank correlation statistic*, Ann. Inst. Statist. Math. **21** (1969), 489–506.

Department of Mathematics, Brandeis University, Waltham, MA, USA, 02454
E-mail address: bdrake@brandeis.edu
URL: http://people.brandeis.edu/~bdrake

Monomial nonnegativity and the Bruhat order

Brian Drake, Sean Gerrish, Mark Skandera

March 8, 2005

Abstract

We show that five nonnegativity properties of polynomials coincide when restricted to polynomials of the form $x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$, where π and σ are permutations in S_n. In particular, we show that each of these properties may be used to characterize the Bruhat order on S_n.

Résumé. Nous démontrons la coïncidence de cinq propriétés des polynômes de la forme $x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$, ou π et σ sont des permutations en S_n. En particulier, nous démontrons que chacune de ces propriétés peut etre employée pour definir l'ordre de Bruhat.

1 Introduction

Let $x = (x_{ij})$ be a generic square matrix and define $\Delta_{I,I'}(x)$ to be the (I, I') minor of x, i.e., the determinant of the submatrix of x corresponding to rows I and columns I'. A real matrix is called *totally nonnegative* (TNN) if each of its minors is nonnegative. (See e.g. [9].) A polynomial $p(x_{11}, \ldots, x_{nn})$ in n^2 variables is called totally nonnegative if it satisfies

$$p(A) \underset{\text{def}}{=} p(a_{1,1}, \ldots, a_{n,n}) \geq 0 \tag{1}$$

for each $n \times n$ totally nonnegative matrix $A = (a_{i,j})$. Some recent interest in total nonnegativity concerns a set of polynomials known in quantum Lie theory as the *dual canonical basis* of $\mathcal{O}(SL(n, \mathbb{C})$ [25]. In particular, Lusztig [17] has proved that these polynomials are TNN.

A polynomial $p(x)$ which is equal to a subtraction-free rational expression in matrix minors must be TNN. (By a result of Whitney [24], we need not be concerned that the denominator vanishes for some TNN matrices.) We shall say that such a polynomial $p(x)$ has the *subtraction-free rational function* (SFR) property. If this subtraction-free rational expression may be chosen so that the denominator is a monomial in matrix minors, we shall say that $p(x)$ has the *subtraction-free Laurent* (SFL) property. One example of a polynomial having the SFL property is

$$x_{1,2}x_{2,1}x_{3,3} - x_{1,2}x_{2,3}x_{3,1} - x_{1,3}x_{2,1}x_{3,2} + x_{1,3}x_{2,2}x_{3,1} = \frac{\Delta_{13,23}(x)\Delta_{23,13}(x) + \Delta_{1,3}(x)\Delta_{3,1}(x)\Delta_{23,23}(x)}{\Delta_{3,3}(x)}.$$

Analogous classes of polynomials may be defined in terms of symmetric functions. (See [21, Ch. 7] for basic definitions concerning symmetric functions.) In particular, any finite submatrix of the infinite matrix $H = (h_{j-i})_{i,j\geq 0}$, where h_k is the kth complete homogeneous symmetric function and $h_k = 0$ for $k < 0$, is called a *Jacobi-Trudi matrix*. We define a polynomial $p(x_{1,1}, \ldots, x_{n,n})$ to be *monomial nonnegative* (MNN) if for each Jacobi-Trudi matrix $A = (a_{i,j})$ the symmetric function $p(A)$ is equal to a nonnegative linear combination of monomial symmetric functions. Defining *Schur nonnegative* (SNN) polynomials analogously, we have that every SNN polynomial is MNN. Some recent interest in SNN polynomials is motivated by problems in algebraic geometry [8, Conj. 2.8, Conj. 5.1], [1].

2 Main result

The five nonnegativity properties defined in Section 1 have been applied most often to *immanants*, polynomials which belong to $\mathrm{span}_{\mathbb{C}}\{x_{1,\sigma(1)} \cdots x_{n,\sigma(n)} \,|\, \sigma \in S_n\}$. (See [11], [12], [13], [20], [19], [22], [23]. The results of [7] may also be stated in these terms.) Curiously, the TNN, MNN, and SNN properties coincide when applied to immanants in the main theorems of the above papers. It is also curious that none of these immanants is known *not* to have the SFL property. It would be interesting to identify immanants which have some of these nonnegativity properties and fail to have others. Nevertheless, our main result shows that the five properties coincide when applied to immanants of the form

$$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}.$$

We shall use the following well-known characterizations of the Bruhat order on S_n. The Bruhat order on S_n is often defined by comparing two permutations $\pi = \pi(1) \cdots \pi(n)$ and $\sigma = \sigma(1) \cdots \sigma(n)$ according to the following criterion: $\pi \leq \sigma$ if σ is obtainable from π by a sequence of transpositions (i, j) where $i < j$ and i appears to the left of j in π. (See e.g. [14, p. 119].) A second well-known criterion compares permutations in terms of their defining matrices. Let $M(\pi)$ be the matrix whose (i, j) entry is 1 if $j = \pi(i)$ and zero otherwise. Defining $[i] = \{1, \ldots, i\}$, and denoting the submatrix of $M(\pi)$ corresponding to rows I and columns J by $M(\pi)_{I,J}$, we have the following.

Theorem 1 *Let π and σ be two permutations in S_n. Then π is less than or equal to σ in the Bruhat order if and only if for all $1 \leq i, j \leq n-1$, the number of ones in $M(\pi)_{[i],[j]}$ is greater than or equal to the number of ones in $M(\sigma)_{[i],[j]}$.*

(See [2], [3], [4], [6], [10, pp. 173-177], [16], [15], [18]. for more characterizations.)

Theorem 2 *Let π and σ be permutations in S_n. The following conditions on π and σ are equivalent.*

1. *$\pi \leq \sigma$ in the Bruhat order.*
2. *$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ is totally nonnegative.*
3. *$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ is Schur nonnegative.*
4. *$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ is monomial nonnegative.*
5. *$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ has the subtraction-free rational function property.*
6. *$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ has the subtraction-free Laurent property.*

Proof: The implications $(3 \Rightarrow 4)$ and $(6 \Rightarrow 5 \Rightarrow 2)$ are immediate. The implication $(2 \Rightarrow 1)$ was estblished in [5, Thm. 2], and the implication $(1 \Rightarrow 6)$ follows trivially from that proof. The implication $(1 \Rightarrow 3)$ was established in [5, Thm. 3]. It will suffice therefore to prove the implication $(4 \Rightarrow 1)$.

Suppose that π is not less than or equal to σ in the Bruhat order. By Theorem 1 we may choose indices $1 \leq k, \ell \leq n-1$ such that $M(\sigma)_{[k],[\ell]}$ contains $q+1$ ones and $M(\pi)_{[k],[\ell]}$ contains q ones. Keeping n fixed, let b be a large nonnegative integer which satisfies

$$\binom{2b}{b} \geq (2b+2n)^{2n^2},$$

(which is possible because $\binom{2b}{b}$ grows exponentially) and consider the Jacobi-Trudi matrix

$$B = \begin{bmatrix} h_{b+k-1} & \cdots & h_{b+k+\ell-2} & h_{2b+k-1} & \cdots & h_{2b+n+k-\ell-2} \\ \vdots & & \vdots & \vdots & & \vdots \\ h_b & \cdots & h_{b+\ell-1} & h_{2b} & \cdots & h_{2b+n-1-\ell} \\ h_{n-k-1} & \cdots & h_{n-k+\ell-2} & h_{b+n-k-1} & \cdots & h_{b+2n-k-\ell-1} \\ \vdots & & \vdots & \vdots & & \vdots \\ h_0 & \cdots & h_{\ell-1} & h_b & \cdots & h_{b+n-\ell-1} \end{bmatrix},$$

defined by the skew shape$(2b+k-\ell-1)^k(b+n-\ell-1)^{n-k}/(b-\ell)^\ell$. Let

$$s = k(n-1+2r) + (n-k)(n-1+r) - \ell r$$

be the number of boxes in this skew shape.

The polynomial $x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ applied to B may be expressed as $h_\lambda - h_\mu$ for some appropriate partitions λ, μ of s, which depend on π, σ, respectively.

We claim that the coefficient of m_{1^s} in the monomial expansion of $h_\lambda - h_\mu$ is negative. Note that the ratio of the coefficients of m_{1^s} in the monomial expansions of h_λ and h_μ is

$$\frac{\binom{s}{\lambda_1,\ldots,\lambda_n}}{\binom{s}{\mu_1,\ldots,\mu_n}} = \frac{\mu_1! \cdots \mu_n!}{\lambda_1! \cdots \lambda_n!}.$$

By the locations of ones in the matrices $M(\pi)$ and $M(\sigma)$, this ratio is less than or equal to

$$\frac{(2b+2n)!^{k-q-1}}{(2b)!^{k-q}} \frac{(b+2n)!^{n-k-\ell+2q+2}}{b!^{n-k-\ell+2q}} \frac{(2n)!^{\ell-q-1}}{0!^{\ell-q}},$$

which in turn is less than or equal to

$$\begin{aligned} \frac{(2b+2n)^{2n(k-q-1)}}{(2b)!}(b+2n)!^2(2b+2n)^{2n(n-k+q-1)} &= \frac{(b+2n)!^2}{(2b)!}(2b+2n)^{2n(n-2)} \\ &\leq \frac{(2b+2n)^{2n(n-1)}}{\binom{2b}{b}}, \end{aligned}$$

which is less than 1 by our choice of b. It follows that the coefficient of m_{1^s} in the monomial expansion of $h_\lambda - h_\mu$ is negative and the polynomial $x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$ is not MNN. $\square$

3 Acknowledgements

The authors are grateful to Arun Ram for helpful conversations.

References

[1] F. Bergeron, R. Biagioli, and M. Rosas. On a conjecture concerning Littlewood-Richardson coefficients, 2003. In progress.

[2] A. Björner. Orderings of Coxeter groups. In *Combinatorics and Algebra* (C. Greene, ed.), vol. 34 of *Contemp. Math.*. American Mathematical Society, Providence, RI, 1984 pp. 175–195.

[3] A. Björner and F. Brenti. An improved tableau criterion for Bruhat order. *Electron. J. Combin.*, **3** (1996). Research paper 22, 5 pp. (electronic).

[4] V. Deodhar. Some characterizations of Bruhat ordering on a Coxeter group and determination of the relative Möbius function. *Inventiones Math.*, **39** (1977) pp. 187–198.

[5] B. Drake, S. Gerrish, and M. Skandera. Two new criteria for comparison in the Bruhat order. *Electron. J. Combin.*, **11** (2004). Note 6, 4 pp. (electronic).

[6] C. Ehresmann. Sur la topologie de certains espaces homogènes. *Ann. Math.*, **35** (1934) pp. 187–198.

[7] S. M. Fallat, M. I. Gekhtman, and C. R. Johnson. Multiplicative principal-minor inequalities for totally nonnegative matrices. *Adv. Appl. Math.*, **30** (2003) pp. 442–470.

[8] S. Fomin, W. Fulton, C. Li, and Y. Poon. Eigenvalues, singular values, and Littlewood-Richardson coefficients, 2003. Preprint math.AG/03013078 on ArXiv.

[9] S. Fomin and A. Zelevinsky. Total positivity: Tests and parametrizations. *Math. Intelligencer*, **22** (2000) pp. 23–33.

[10] W. Fulton. *Young Tableaux; With Applications to Representation Theory and Geometry*, vol. 35 of *London Mathematical Society Student Texts*. Cambridge University Press, New York, 1997.

[11] I. P. Goulden and D. M. Jackson. Immanants of combinatorial matrices. *J. Algebra*, **148** (1992) pp. 305–324.

[12] C. Greene. Proof of a conjecture on immanants of the Jacobi-Trudi matrix. *Linear Algebra Appl.*, **171** (1992) pp. 65–79.

[13] M. Haiman. Hecke algebra characters and immanant conjectures. *J. Amer. Math. Soc.*, **6** (1993) pp. 569–595.

[14] J. E. Humphreys. *Reflection groups and Coxeter groups*. Cambridge University Press, 1990.

[15] A. Lascoux. Potentiel Yin sur le groupe symétrique. *Sém. Lothar. Combin.*, **38** (1996). Art. B38a, 12 pp. (electronic).

[16] A. Lascoux and M. P. Schützenberger. Treillis et bases des groupes de Coxeter. *Electron. J. Combin.*, **3** (1996). Research paper 27, 35 pp. (electronic).

[17] G. Lusztig. Total positivity in reductive groups. In *Lie Theory and Geometry: in Honor of Bertram Kostant*, vol. 123 of *Progress in Mathematics*. Birkhäuser, Boston, 1994 pp. 531–568.

[18] R. A. Proctor. Classical Bruhat orders and lexicographic shellability. *J. Algebra*, **77** (1982) pp. 104–126.

[19] B. Rhoades and M. Skandera. Kazhdan-Lusztig immanants, 2004. In progress.

[20] B. Rhoades and M. Skandera. Temperley-Lieb immanants, 2004. Submitted.

[21] R. Stanley. *Enumerative Combinatorics*, vol. 2. Cambridge University Press, Cambridge, 1999.

[22] R. Stanley and J. R. Stembridge. On immanants of Jacobi-Trudi matrices and permutations with restricted positions. *J. Combin. Theory Ser. A*, **62** (1993) pp. 261–279.

[23] J. Stembridge. Some conjectures for immanants. *Can. J. Math.*, **44** (1992) pp. 1079–1099.

[24] A. Whitney. A reduction theorem for totally positive matrices. *J. Anal. Math.*, **2** (1952) pp. 88–92.

[25] A. Zelevinsky. From Littlewood-Richardson coefficients to cluster algebras in three lectures. In *Symmetric Functions 2001: Surveys of Developments and Perspectives* (S. Fomin, ed.), vol. 74 of *NATO Science Series II: Mathematics, Physics, and Chemistry*. Kluwer, Dordrecht, 2002 pp. 253–273.

ON SOME CLASSES OF PRUDENT WALKS

ENRICA DUCHI*

Abstract. In this paper we consider a class of walks introduced by Pascal Préa, that we call *prudent* walks: they are self-avoiding because they never try to walk towards points that they already visited.

We write functional equations for counting three classes of such prudent walks with respect to their length. The first one is algebraic and we give an object grammar decomposition that explains this fact directly. For the second one we obtain the algebraic generating function but in an indirect way. Whether the third class has an algebraic generating function remains an open problem.

Résumé. Dans cet article nous considérons une classe de chemins introduits par Pascal Préa, que nous appelons les *chemins prudents*, qui sont auto-évitants parce qu'ils ne vont jamais dans la direction d'un point qu'ils ont déjà visité.

Nous écrivons des équations fonctionnelles pour compter trois classes de chemins de ce type en fonction de leur longueur. La première classe est algebrique et nous donnons une explication de ce résultat à l'aide d'une grammaire. Pour la seconde nous obtenons aussi l'algébricité mais de manière indirecte. Le problème de savoir si la troisième classe a une série génératrice algébrique reste ouvert.

1. Prudent walks

The term *walk* is used to denote a sequence of points $s_0, s_1, \dots s_n$ in the plane $\mathbb{Z} \times \mathbb{Z}$. A couple (s_i, s_{i+1}) is said to be a *step* of the walk and the number n of steps is called the *length* of the walk. Given $s_i = (x, y)$ then (s_i, s_{i+1}) is:

- an *east* ($\rightarrow$) step if $s_{i+1} = (x+1, y)$
- a *west* ($\leftarrow$) step if $s_{i+1} = (x-1, y)$
- a *north* ($\uparrow$) step if $s_{i+1} = (x, y+1)$
- a *south* ($\downarrow$) step if $s_{i+1} = (x, y-1)$

From now on all the walks will be on the lattice, that is, they are made of east, west, north, and south steps only. We shall concentrate on some families of self-avoiding walks. A self-avoiding walk is a walk that cannot cross itself, i.e. it never visits two times the same point. Counting self-avoiding walks is a well-known open problem. See [5, 6] for some references. For this reason various subclasses of these walks have been introduced and counted. Here we consider a subclass of self-avoiding walks that we call the class of *prudent walks*. As we learned recently from Mireille Bousquet-Mélou, these walks were introduced by Pascal Préa in [7], where he obtained some recurrences for their enumeration.

Definition 1. *A* prudent walk *is a sequence of east, west, north, and south steps running from* $(0,0)$ *to* (n, m)*, with* $n, m \in \mathbb{Z}$*, defined in the following manner:*

- *The empty walk starting from* $(0,0)$ *and ending in* $(0,0)$ *is a prudent walk.*
- *A prudent walk is obtained from another prudent walk by attaching a new step at its end in such a manner that the extension of this step in the sense of its direction never encounters the walk itself.*

In particular this definition implies that a prudent walk is a self-avoiding walk.

Definition 2. *The* prudent box *of a prudent walk is the smallest rectangle including the walk. We remark that for particular walks this rectangle reduces to a line or also to a point in the case of the empty walk.*

In Figure 1 are given examples of prudent walks with their prudent box and of a non prudent walk.

Proposition 1. *The last point of a prudent walk is always on the border of the prudent box.*

Date: November 21, 2004.
* Supported by a post-doctoral grant from the CNRS.

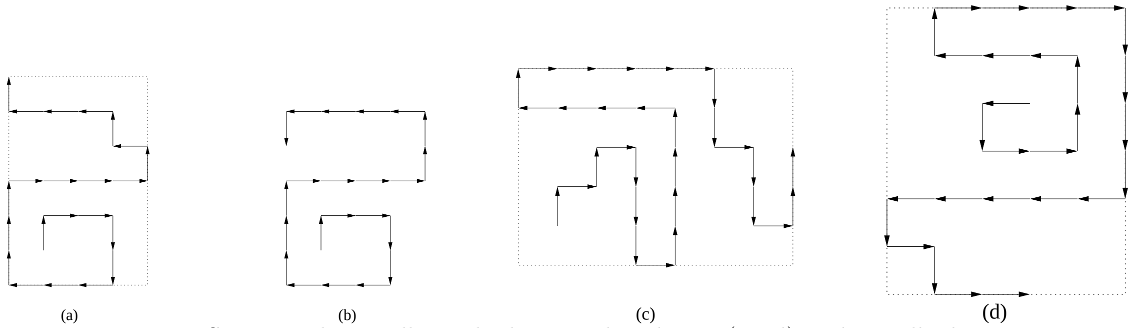


FIGURE 1. Some prudent walks with their prudent boxes (a,c,d) and a walk that is not prudent (b).

Remark that the converse is not true, as illustrated by walk (b) in Figure 1.

In this paper we write functional equations for the generating functions of several classes of prudent walks using the recursive definition of these walks. These equations are equivalent to the recurrences independently obtained by Pascal Préa. For some of these classes we are able to give explicit formulas for the generating functions. First, we deal with prudent walks that can only end on the right side or on the top of their prudent box. For this class, using the kernel method, [1, 2], we compute their generating function which is algebraic. In order to give a direct explanation of this fact we also find an algebraic decomposition [8] (or object grammar [3, 4]) for this class. Then we deal with another subclass of prudent walks, made of walks that can end on the top, right, or bottom side of their prudent box. We show that the generating function of these walks is also algebraic but this result is more complicated to derive. In particular we need several applications of the kernel method and at the moment we do not know a direct algebraic decomposition for this class. The paper is organized as follows: in Section 1 we introduce the first subclass of prudent walks and we give their generating function. In Section 2 we present an object grammar for this class. In Section 3 we introduce the second class of prudent walks and we count them. To conclude, in Section 4 we write a functional equation for the complete class of prudent walks and leave open the problem of knowing whether they have an algebraic generating function.

Given a walk w we define the following parameters (see Figure 2):

- $i(w)$ is the distance between the last point of w and the top of its prudent box;
- $j(w)$ is the distance between the last point of w and the bottom of its prudent box;
- $k(w)$ is the distance between the last point of w and the right side of its prudent box;
- $l(w)$ is the distance between the last point of w and the left side of its prudent box.

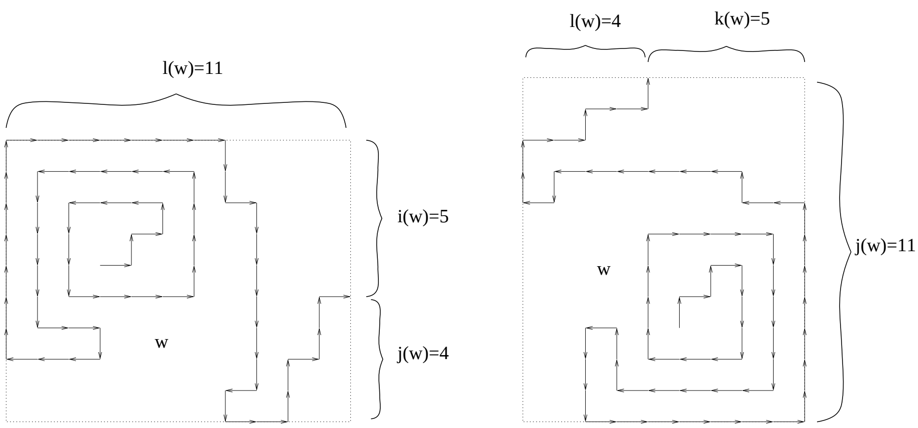


FIGURE 2. The parameters $i(w)$, $j(w)$, $k(w)$, $l(w)$.

2. Prudent walks of the first type

Definition 3. *A prudent walk of the first type is a prudent walk avoiding the following subsequences of steps: a west step followed by a south step* $\genfrac{}{}{0pt}{}{\leftarrow}{\downarrow}$ *and a south step followed by a west step* $\genfrac{}{}{0pt}{}{\downarrow}{\leftarrow}$.

Let us denote by P^1 the class of prudent walks of the first type. See Figure 1(c) for an example.

Remark 1. *The set of prudent walks of the first type is symmetric with respect to the main diagonal. That is, the set is invariant by the following transformation:* $(\leftarrow) \mapsto (\downarrow)$, $(\downarrow) \mapsto (\leftarrow)$, $(\rightarrow) \mapsto (\uparrow)$, $(\uparrow) \mapsto (\rightarrow)$ *for the steps and Bottom* $\mapsto$ *Left, Right* $\mapsto$ *Top, Left* $\mapsto$ *Bottom, Top* $\mapsto$ *Right for the border of the prudent box.*

In order to count the number of walks of P^1 according to their length, the following proposition is useful:

Proposition 2. *A prudent walk w of the first type always ends on the right side or on the top of its prudent box B.*

Counting prudent walks of the first type. We are interested in determining the generating function of the walks of the first type according to their length. However we will need to count them also with respect to $i(w)$ and $k(w)$. We denote by

$$p^1(u, v; t) = \sum_{w \in P^1} u^{k(w)} v^{i(w)} t^{|w|}$$

such a generating function. In order to compute it we distinguish the following subclasses:

- H_n is the subclass of P^1 formed by the prudent walks ending with a $\uparrow$ step, and such that their last step moved the top of the previous prudent box, i.e. the prudent box at the previous step. In particular this means that $i(w) = 0$ for each $w \in H_n$. See Figure 3.
- G_n is the subclass of P^1 formed by the prudent walks ending with a $\uparrow$ step, and such that their prudent box is the same as the one at the previous step. In particular this means that $k(w) = 0$ for each $w \in G_n$. See Figure 3.

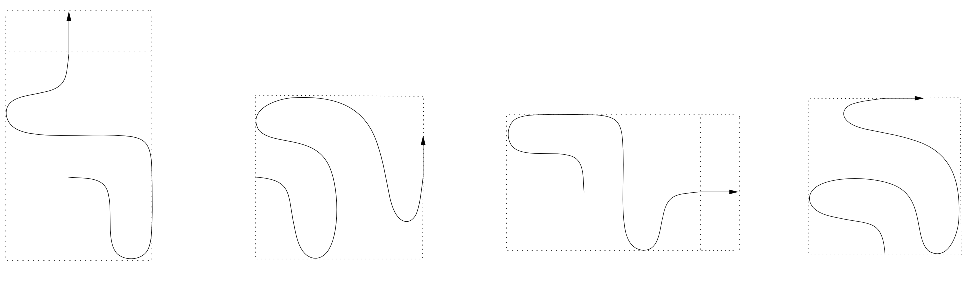


Figure 3. The classes H_n, G_n, H_e and G_e.

- H_e is the subclass of P^1 formed by the prudent walks ending with a $\rightarrow$ step, and such that their last step $\rightarrow$ moved the right border of the previous prudent box. In particular this means that $k(w) = 0$ for each $w \in H_e$. See Figure 3.
- G_e is the subclass of P^1 formed by the prudent walks ending with a $\rightarrow$ step, and such that their prudent box is the same as the one at the previous step. In particular this means that $i(w) = 0$ for each $w \in G_e$. See Figure 3.
- F_o is the subclass of P^1 formed by the prudent walks having as last step a $\leftarrow$ step, therefore ending in the top of B. Then $i(w) = 0$ for each $w \in F_o$. See Figure 4.
- F_s is the subclass of P^1 formed by the prudent walks whose last step is a $\downarrow$ step, therefore ending in the right side of B. Then $k(w) = 0$ for each $w \in F_s$. See Figure 4.

We respectively denote by $h_n(u; t)$, $g_n(v; t)$, $h_e(v; t)$, $g_e(u; t)$, $f_o(u; t)$, $f_s(v; t)$ the generating functions of the previous defined classes, where the variable u marks the parameter $k(w)$, v marks $i(w)$, and t marks the length. Also note that the indices n, s, e, o consistently indicate the direction of the last step of the walks (o stands for *ovest*).

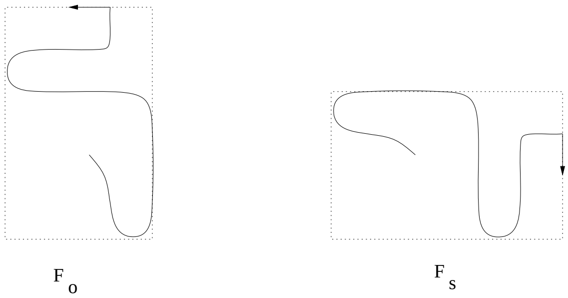


FIGURE 4. The classes F_o and F_s.

Lemma 1. *The set* $S^1 = \{H_n, G_n, H_e, G_e, F_o, F_s\}$ *is a partition of the set* P^1. *Consequently* $p^1(u, v; t) = h_n(u; t) + g_n(v; t) + h_e(v; t) + g_e(u; t) + f_o(u; t) + f_s(v; t)$.

Proof. Let us take a walk w in P^1. If it ends with a $\downarrow$(*resp.* $\leftarrow$) step then it belongs to F_s (*resp.* F_o). If it ends with a $\uparrow$ (*resp.* $\rightarrow$) step then it belongs to H_n or G_n (*resp.* H_e or G_e) if the walk obtained by removing this step ends in the top (*resp.* right side) of its box or not. Hence the set S^1 is a partition of P^1. The equation for $p^1(u, v; t)$ is a direct consequence of this fact. □

Now, by using the recursive definition of prudent walk, we write equations for the previous generating functions. We start from those depending only on the variables u and t. The others are obtained from these ones by symmetry.

- *The equation for* $h_n(u; t)$. The walks of the class H_n end with a $\uparrow$ step which move the top of their previous prudent box. Hence this class is formed by the step $\uparrow$ itself and by the walks obtained by adding a $\uparrow$ step to walks ending on the top of their prudent box. The latter walks belong to the following classes: H_n; the walks belonging to G_n and ending in the top right angle; the walks belonging to H_e and ending in the top right angle; G_e; F_o. The addition of a $\uparrow$ step to these walks increases their length by one. Therefore we have the following equation:

$$h_n(u; t) = t\ (h_n(u; t) + g_n(0; t) + h_e(0; t) + g_e(u; t) + f_o(u; t) + 1) \tag{1}$$

- *The equation for* $g_e(u; t)$. G_e is formed of walks obtained by adding a $\rightarrow$ step to walks ending in the top side of their prudent box, with exclusion of the top right angle. The latter walks belong to the following classes: G_e without the walks ending in the top right angle; H_n without the walks ending in the top right angle. By adding a $\rightarrow$ step, the length of these walks increases by one and their distance to the right side diminishes by one. Therefore we have the following equation:

$$g_e(u; t) = \frac{t}{u}\ (\ g_e(u; t) - g_e(0; t)\ +\ h_n(u; t) - h_n(0; t)\ ) \tag{2}$$

- *The equation for* $f_o(u; t)$. This class is obtained by adding a $\leftarrow$ step to walks ending in the leftmost occupied position of the top of their prudent box (otherwise we would not respect the condition of prudent walk). Then the equation for $f_o(u)$ involves: H_n; F_o; the empty walk. The operation of adding a $\leftarrow$ step increases both the length of these walks and their distance to the right side of their prudent box by one. Therefore we have the following equation:

$$f_o(u; t) = t\, u\ (h_n(u; t) + f_o(u; t) + 1)$$

In order to simplify the system formed of these three equations we replace the expression $g_n(0; t) + h_e(0; t)$ by $I(t)$ in (1) and (2). Then the system becomes:

$$\begin{aligned} h_n(u; t) &= t\ (h_n(u; t) + I(t) + g_e(u; t) + f_o(u; t) + 1) \\ g_e(u; t) &= \tfrac{t}{u}\ (g_e(u; t) + h_n(u; t) - I(t)) \\ f_o(u; t) &= t\, u\ (h_n(u; t) + f_o(u; t) + 1) \end{aligned}$$

By solving this system with respect to $h_n(u;t)$, $g_e(u;t)$, and $f_o(u;t)$ we obtain the following:

$$\begin{aligned} h_n(u;t) &= t\,\frac{(I(t)\,(-2ut^2+u^2t-u+2t)-u+t)}{K(u;t)} \\ g_e(u;t) &= t\,\frac{(I(t)\,(1-2t-ut+ut^2)-t)}{K(u;t)} \\ f_o(u;t) &= -t\,u\,\frac{\left(I(t)(-2t^2+ut)+u-t^2-t\right)}{K(u;t)}, \end{aligned}$$

with $K(u;t) = -u + t + ut + u^2t - ut^2 - ut^3$. By symmetry we also have expressions for $h_e(v;t)$, $g_n(v;t)$, and $f_s(v;t)$, obtained respectively from $h_n(u;t)$, $g_e(u;t)$, and $f_o(u;t)$ by substituting u with v. By using Lemma 1 we obtain the generating function of the class P^1 in terms of $I(t)$:

$$p^1(1,1;t) = \frac{2\,t\,(2+t+t\,I(t))}{1-2t-t^2} \tag{3}$$

We apply the kernel method [1, 2], for instance on the equation for $g_e(u;t)$:

$$g_e(u;t)K(u;t) = t\;(I(t)\;(1-2t-ut+ut^2)-t). \tag{4}$$

The kernel $K(u;t)$ has a unique root $U_0(t)$ which is a formal power series:

$$U_0(t) = \frac{1-t+t^2+t^3-\sqrt{(1-t^4)(1-2t-t^2)}}{2t}$$

Substituting $u = U_0(t)$ in (4) the left hand side is canceled and we obtain $I(t)$:

$$I(t) = \frac{t}{(1-2t-U_0(t)t+U_0(t)t^2)} \tag{5}$$

By using equation (5) and equation (3) we have

$$p^1(1,1;t) = t\frac{(1-2t-t^2)(3+2t-3t^2)+(1-t)\sqrt{(1-t^4)(1-2t-t^2)}}{(1-2t-t^2)(1-2t-2t^2+2t^3)} = 4t+10t^2+26t^3+66t^4+168t^5+O(t^6)$$

3. A grammar for prudent walks of the first type

We counted the number of prudent walks according to their length and we obtained an algebraic generating function. Now we give a direct explanation of the algebraicity of this class by determining an algebraic decomposition (or object grammar) for it. As before we denote classes of paths by capital letters and we use the corresponding small letters for the generating functions with respect to the length. We start by distinguishing four subclasses of prudent walks of the first type:

- The class $W_\uparrow$ of walks of the first type beginning with a $\uparrow$ step.
- The class $W_\downarrow$ of walks of the first type beginning with a $\downarrow$ step.
- The class $W_\rightarrow$ of walks of the first type beginning with a $\rightarrow$ step.
- The class $W_\leftarrow$ of walks of the first type beginning with a $\leftarrow$ step.

Then we have:

$$P^1 \;=\; W_\uparrow \;+\; W_\downarrow \;+\; W_\leftarrow \;+\; W_\rightarrow$$

Let us consider the class $W_\rightarrow$. It consists of the walk reduced to $\rightarrow$ and of all walks obtained by adding a first initial $\rightarrow$ step to a prudent walk of the first type respectively beginning with a $\uparrow$, a $\downarrow$, or a $\rightarrow$ step. Hence we can write the following grammar:

$$W_\rightarrow \;=\; \rightarrow\cdot W_\rightarrow + \;\rightarrow\cdot W_\uparrow + \;\rightarrow\cdot W_\downarrow + \;\rightarrow$$

By symmetry we have a similar decomposition for $W_\uparrow$, and we have the equalities $w_\uparrow(t) = w_\rightarrow(t)$ and $w_\downarrow(t) = w_\leftarrow(t)$. This also implies that it is sufficient to find a decomposition for the class $W_\leftarrow$ or for the class $W_\downarrow$ to determine $p^1(t)$.

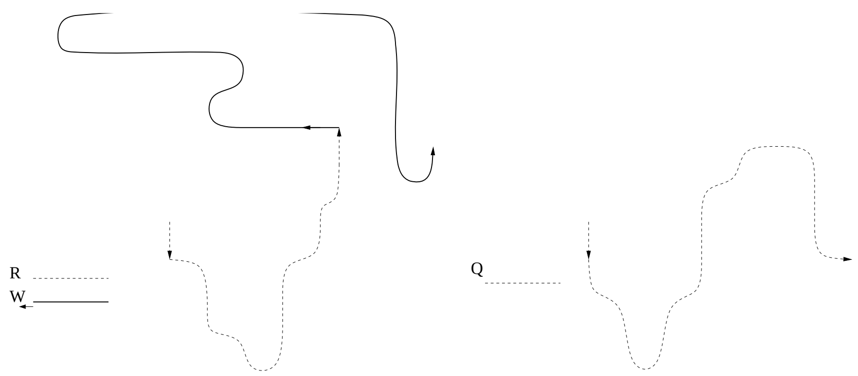


FIGURE 5. The walks belonging to $R \cdot W_{\leftarrow}$ (left) and to Q (right).

3.1. **A decomposition for the class $W_{\downarrow}$.** In order to define this grammar we distinguish the following classes:

i. *The class of walks that contain the subsequence made of a $\uparrow$ step followed by a $\leftarrow$ step.* We divide each walk belonging to this class in two parts after the first $\uparrow$ followed by a $\leftarrow$ (see Figure 5, left hand side):
   - *The first part belongs to the class R:* It is the class of prudent walks of the first type starting with a $\downarrow$ step, not containing any $\leftarrow$ step, ending in the top right angle of their prudent box with a $\uparrow$ step, and such that they moved the top of their prudent box with this last step.
   - *The second part belongs to the class $W_{\leftarrow}$:* It is the class of prudent walks of the first type starting with a $\leftarrow$ step.

ii. *The class of walks of the first type that do not contain any $\leftarrow$ step.* We denote this class of walks by Q (see Figure 5, right hand side).

Then the decomposition for $W_{\downarrow}$ is the following:

$$W_{\downarrow} \;=\; R \cdot W_{\leftarrow} \;+\; Q \tag{6}$$

3.2. **A decomposition for the class R.** We can decompose R in the following manner:

i. The class of walks of R that make a $\downarrow$ step at a later time when they are in the top right angle. In other terms these are the walks of the form $w_0 \cdot \downarrow \cdot w_1$ where w_0 is a walk ending in the top right angle of its box. The decomposition for these walks is the following:

$$\downarrow \cdot C \cdot \uparrow \cdot T \cdot \rightarrow \cdot R$$

where C is essentially the class of Motzkin paths that avoid the subsequences $\uparrow\downarrow$ and $\downarrow\uparrow$,

$$C \;=\; \downarrow \cdot C \cdot \uparrow \;+\; \downarrow \cdot C \cdot \uparrow \cdot \rightarrow \cdot C \;+\; \downarrow \cdot C \cdot \uparrow \cdot \rightarrow \;+\; \rightarrow \cdot C \;+\; \rightarrow \tag{7}$$

and T is the class of staircases,

$$T \;=\; \uparrow \cdot T \;+\; \rightarrow \cdot T \;+\; \epsilon \tag{8}$$

ii. The class of walks of R that after their first step never make anymore a $\downarrow$ step when they are in the top right angle. The decomposition for these walks is the following:

$$\downarrow \cdot C \cdot \uparrow \cdot T \cdot \uparrow$$

In summary the decomposition for R is the following (see Figure 6):

$$R \;=\; \downarrow \cdot C \cdot \uparrow \cdot T \cdot \rightarrow \cdot R \;+\; \downarrow \cdot C \cdot \uparrow \cdot T \cdot \uparrow \tag{9}$$

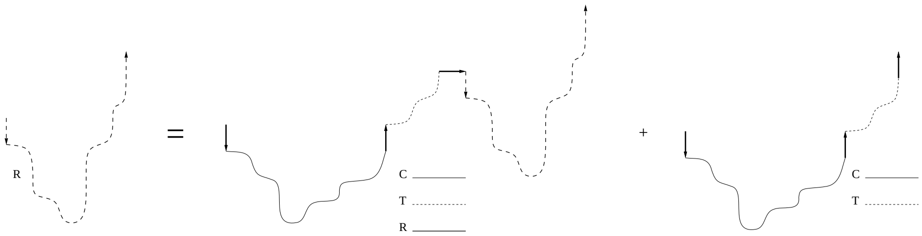


FIGURE 6. The complete decomposition for R.

3.3. **A decomposition for the class Q.** The class Q contains all walks with steps $\uparrow$, $\downarrow$, and $\rightarrow$ that start with a $\downarrow$ step and avoid the subsequences $\uparrow\downarrow$ and $\downarrow\uparrow$. This class satisfies:

$$\begin{aligned}
Q &= \downarrow + \downarrow\cdot Q_{\rightarrow} + \downarrow\cdot Q \\
Q_{\rightarrow} &= \rightarrow + \rightarrow\cdot Q_{\rightarrow} + \rightarrow\cdot Q + \rightarrow\cdot Q_{\uparrow} \\
Q_{\uparrow} &= \uparrow + \uparrow\cdot Q_{\rightarrow} + \uparrow\cdot Q_{\uparrow}
\end{aligned}$$

3.4. **Generating functions.** By using the previous decompositions we obtain the following algebraic equations:

$$\begin{array}{lll lll}
p^1(t) &=& w_{\rightarrow}(t)+w_{\uparrow}(t)+w_{\downarrow}(t)+w_{\leftarrow}(t) & c(t) &=& t+t^2c(t)+t^3c(t)^2+t^2c(t)+tc(t) \\
w_{\rightarrow}(t) &=& t\,w_{\rightarrow}(t)+t\,w_{\uparrow}(t)+t\,w_{\downarrow}(t)+t & T(t) &=& 1+2tT(t) \\
w_{\downarrow}(t) &=& r(t)w_{\leftarrow}(t)\ +\ Q(t) & r(t) &=& t^3c(t)\,T(t)\,r(t)+t^3c(t)\,T(t) \\
w_{\leftarrow}(t) &=& w_{\downarrow}(t) & q(t) &=& t+tq_{\rightarrow}(t)+tq(t) \\
w_{\uparrow}(t) &=& w_{\rightarrow}(t) & q_{\rightarrow}(t) &=& t+tq_{\rightarrow}(t)+2tq(t)
\end{array}$$

From this system we can recover the expressions of the previous section.

4. PRUDENT WALKS OF THE SECOND TYPE

We are now interested in counting a wider class of prudent walks. As we saw in Section 2, prudent walks of the first type can only end on the top or on the right side of their prudent box. We introduce a class of prudent walks which can also end on the bottom of their prudent box.

Definition 4. *A prudent walk of the second type is a prudent walk avoiding the following subsequences of steps when the prudent box is not a line: a west step followed by a south step* $\overset{\leftarrow}{\downarrow}$ *when the walk visits the top of its current prudent box and a west step followed by a north step* $\overset{\uparrow}{\leftarrow}$ *when the walk visits the bottom of its current prudent box.*

In Figure 1(d) there is an example of prudent walk of the second type with its prudent box.

We want to count the number of these walks according to their length. Similarly to the previous section we have the following:

Proposition 3. *A prudent walk w of the second type always ends on the top, right, or bottom side of its prudent box B.*

Counting prudent walks of the second type. Let us call P^2 the class of prudent walks of the second type. We are interested in determining their generating function according to their length and their distances $i(w)$, $j(w)$, and $k(w)$. We denote by $p^2(u,v,w;t)=\sum_{w\in P^2}u^{k(w)}v^{i(w)}w^{j(w)}t^{|w|}$ this generating function. In order to compute it we distinguish the following subclasses:

- H_n is the subclass of P^2 formed by the prudent walks ending with a $\uparrow$ step, and such that their last step $\uparrow$ moved the top of the previous prudent box, i.e. the prudent box at the previous step. In particular this means that $i(w) = 0$ for each $w \in H_n$ (See Figure 7).
- G_n is the subclass of P^2 formed by the prudent walks ending with a $\uparrow$ step, and such that their prudent box is the same as the one at the previous step. In particular this means that $k(w) = 0$ for each $w \in G_n$ (See Figure 7).
- H_s is the subclass of P^2 formed by the prudent walks ending with a $\downarrow$ step, and such that their last step $\downarrow$ moved the bottom of the previous prudent box. In particular this means that $j(w) = 0$ for each $w \in H_s$. (See Figure 7).
- G_s is the subclass of P^2 formed by the prudent walks ending with a $\downarrow$ step, and such that their prudent box is the same as the one at the previous step. In particular this means that $k(w) = 0$ for each $w \in G_s$ (See Figure 7).

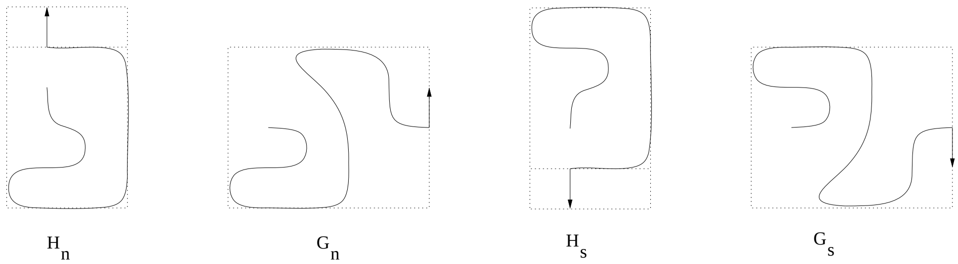


FIGURE 7. The classes H_n, G_n, H_s, and G_s.

- H_e is the subclass of P^2 formed by the prudent walks ending with a $\rightarrow$ step, and such that their last step $\rightarrow$ moved the right side of the previous prudent box. In particular this means that $k(w) = 0$ for each $w \in H_e$ (See Figure 8).
- GT_e is the subclass of P^2 formed by the prudent walks ending with a $\rightarrow$ step on the top of their prudent box, and such that this box is the same as the one at the previous step. In particular this means that $i(w) = 0$ for each $w \in GT_e$ (See Figure 8).
- GB_e is the subclass of P^2 formed by the prudent walks ending with a $\rightarrow$ step on the bottom of their prudent box and such that this box is the same as the one at the previous step. In particular this means that $j(w) = 0$ for each $w \in GB_e$ (See Figure 8).

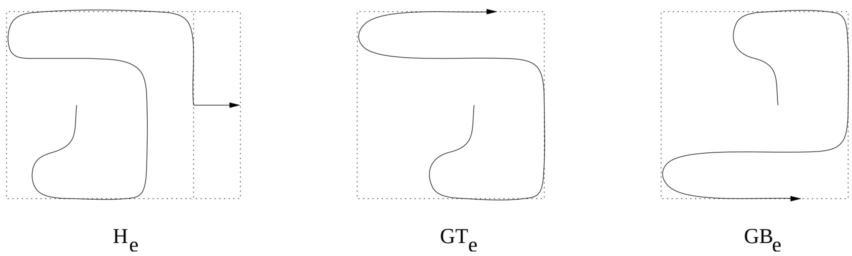


FIGURE 8. The classes H_e, GT_e, and GB_e.

- FT_o is the subclass of P^2 formed by the prudent walks having as last step a $\leftarrow$ step, ending in the top of B and such that B is not reduced to a line. Then $i(w) = 0$ for each $w \in FT_o$ (See Figure 9).
- FB_o is the subclass of P^2 formed by the prudent walks having as last step a $\leftarrow$ step, ending in the bottom of B and such that B is not reduced to a line. Then $j(w) = 0$ for each $w \in FB_o$ (See Figure 9).
- F_o is the subclass of P^2 whose prudent walks are made up just by $\leftarrow$ steps (See Figure 9).

We respectively denote by $h_n(u, w; t)$, $g_n(v, w; t)$, $h_s(u, v; t)$, $g_s(v, w; t)$, $h_e(v, w; t)$, $gt_e(u, w; t)$, $gb_e(u, v; t)$, $ft_o(u, w; t)$, $fb_o(u, v; t)$, $f_o(u; t)$ the generating functions of the previous defined classes with the variables marking as for $p^2(u, v, w; t)$.

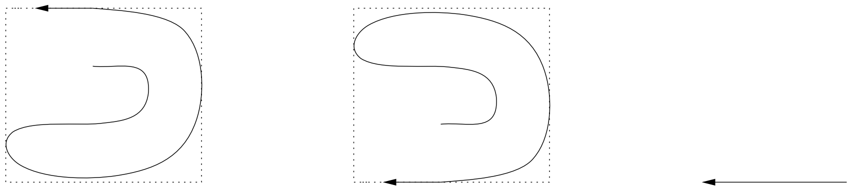


FIGURE 9. The classes FT_o, FB_o, and F_o.

Lemma 2. *The set $S^2 = \{H_n, G_n, H_s, G_s, H_e, GT_e, GB_e, FT_o, FB_o, F\}$ is a partition of the set P^2. Consequently $p^2(u,v,w;t) = h_n(u,w;t)+g_n(v,w;t)+h_s(u,v;t)+g_s(v,w;t)+h_e(v,w;t)+gt_e(u,w;t)+gb_e(u,v;t)+ft_o(u,w;t)+fb_o(u,v;t)+f_o(u;t)$.*

Proof. A prudent walk w in P^2 can finish with a $\downarrow$, $\leftarrow$, $\uparrow$, $\rightarrow$ step. If it ends with a $\leftarrow$ step then we have the following cases: its prudent box is a line, then it belongs to F_o; it ends on the top of its prudent box, then it belongs to belongs to FT_o; it ends on the bottom of its prudent box, then it belongs to FB_o. If it ends with an $\uparrow$ step then it belongs to H_n or G_n (*resp.* H_s or G_s) if the ending point at the previous step was in the top (*resp.* bottom) of the box or not. If it ends with a $\rightarrow$ step then it belongs to H_e if the ending point at the previous step was in the right side of its prudent box. Otherwise it belongs to GT_e(*resp.* GB_e) if its last step is on the top (*resp.* bottom) of its prudent box. Then the set S^2 is a partition of P^2 and the result on $p^2(u,v,w;t)$ is a direct consequence of this fact. ☐

Now we can write equations for the generating functions of elements of S^2 by using the recursive definition of a prudent walk. Let us write the equations for $h_n(u,w;t)$, $g_n(v,w;t)$, $h_e(v,w;t)$, $gt_e(u,w;t)$, $ft_o(u,w;t)$, $f_o(u;t)$. The others are obtained from these ones by symmetry.

- *The equation for $h_n(u,w;t)$.* These walks, ending with a $\uparrow$ step, moved the top of the previous prudent box. Therefore this class is formed by $\uparrow$ itself and by the walks obtained by adding a $\uparrow$ step to walks ending in the top of their prudent box. Such walks belong to the following classes: H_n; the walks belonging to G_n and ending on the top right angle; the walks belonging to H_e and ending on the top right angle; GT_e; FT_o; F_o. Adding a $\uparrow$ step to these walks increases their length by one and their distance from the bottom of the prudent box by one. Then we have the following equation:

$$h_n(u,w;t) = tw\,(h_n(u,w;t) + g_n(0,w;t) + h_e(0,w;t) + gt_e(u,w;t) + ft_o(u,w;t) + f_o(u;t) + 1)$$

- *The equation for $g_n(v,w;t)$.* G_n is formed of walks obtained by adding a $\uparrow$ step to walks ending in the right side of their prudent box, with exclusion of the top right angle. Such walks belong to the following classes: G_n, with the exclusion of walks ending in the top right angle; H_e, with the exclusion of walks ending in the top right angle. With the addition of a $\uparrow$ step the length of the walks increases by one, their distance to the top of the prudent box diminishes by one, and their distance to the bottom of the prudent box increases by one. Then we have the following equation:

$$g_n(v,w;t) = \frac{tw}{v}\,(g_n(v,w;t) - g_n(0,w;t) + h_e(v,w;t) - h_e(0,w;t))$$

- *The equation for $h_e(u,w;t)$.* These walks, ending with a $\rightarrow$ step, moved the right side of the previous prudent box. Therefore this class is formed by $\rightarrow$ itself and by the walks obtained by adding a $\rightarrow$ step to walks ending in the right of their prudent box. Such walks belong to the following classes: H_e; the walks belonging to GT_e and ending on the top right angle; the walks belonging to GB_e and ending on the bottom right angle; the walks belonging to H_n and ending on the top right angle; G_n; the walks belonging to H_s and ending on the bottom right angle; the walks belonging to G_s. Adding a $\rightarrow$ step to these walks increases their length by one. Then we have the following equation:

$$h_e(v,w;t) = t\,(h_e(v,w;t) + gt_e(0,w;t) + gb_e(0,v;t) + h_n(0,w;t) + g_n(v,w;t) + h_s(0,v;t) + g_s(v,w;t) + 1)$$

- *The equation for $gt_e(u,w;t)$.* GT_e is formed of walks obtained by adding a $\rightarrow$ step to walks ending in the top side of their prudent box, with exclusion of the top right angle. Such walks belong to the following classes: GT_e, with the exclusion of walks ending in the top right angle; H_n, with the exclusion of walks ending in the top right angle. With the addition of a $\rightarrow$ step the length of the walks increases by one and their distance to the right of the prudent box diminishes by one. Then we have the following equation:

$$gt_e(u,w;t) = \frac{t}{u}\left(gt_e(v,w;t) - gt_e(0,w;t) + h_n(u,w;t) - h_n(0,w;t)\right)$$

- *The equation for $ft_o(u,w)$* This class is formed by the walks obtained by adding a $\leftarrow$ step to walks ending on the top of their prudent box. Observe that by adding a $\leftarrow$ step to walks G_n ending in the top right angle we do not respect the condition of prudent walk. Then the classes involved in the equation for $ft_o(u)$ are the following: H_n; FT_o. By adding a $\leftarrow$ step both the length of the walks and their distance to the right of the prudent box increase by one. Therefore we have the following equation:
$$ft_o(u,w;t) = t\,u\left(h_n(u,w;t) + ft_o(u,w;t)\right)$$
- *The equation for $f_o(u)$* This is the class $\{\leftarrow\}^+$, therefore:
$$f_o(u;t) = \frac{tu}{1-tu}$$

The other equations are obtained by symmetry with respect to the horizontal axis. In order to simplify the system of these equations, we introduce the following series:

$$L(w;t) = h_n(0,w;t) + gt_e(0,w;t), \text{ or equivalently, } L(v;t) = h_s(0,v;t) + gb_e(0,v;t)$$

$$M(w;t) = h_e(0,w;t) + g_n(0,w;t), \text{ or equivalently, } M(v;t) = h_e(v,0;t) + g_s(v,0;t)$$

Then the system becomes:

$$\begin{array}{lcl}
h_n(u,w;t) & = & tw\left(h_n(u,w;t) + gt_e(u,w;t) + ft_o(u,w;t) + f_o(u;t) + M(w;t) + 1\right) \\
g_n(v,w;t) & = & \frac{tw}{v}\left(h_e(v,w;t) + g_n(v,w;t) - M(w;t)\right) \\
h_e(v,w;t) & = & t\left(h_e(v,w;t) + g_s(v,w;t) + g_n(v,w;t) + L(w;t) + L(v;t) + 1\right) \\
gt_e(u,w;t) & = & \frac{t}{u}\left(h_n(u,w;t) + gt_e(u,w;t) - L(w;t)\right) \\
gb_e(u,v;t) & = & \frac{t}{u}\left(h_s(u,v;t) + gb_e(u,v;t) - L(v;t)\right) \\
h_s(u,v;t) & = & tv\left(h_s(u,v;t) + gb_e(u,v;t) + fb_o(u,v;t) + f_o(u;t) + M(v;t) + 1\right) \\
g_s(v,w;t) & = & \frac{tv}{w}\left(h_e(v,w;t) + g_s(v,w;t) - M(v;t)\right) \\
ft_o(u,w;t) & = & tu(h_n(u,w;t) + ft_o(u,w;t)) \\
fb_o(u,v;t) & = & tu(h_s(u,v;t) + fb_o(u,v;t)) \\
f_o(u;t) & = & \frac{tu}{1-tu}
\end{array}$$

Observe that in this system of linear equations all the series can be expressed in terms of $L(w;t)$, $L(v;t)$, $M(w;t)$, $M(v;t)$. In particular we have:

$$ft_o(u,w;t) = \frac{t^2\,u\,w((1-tu)(1-t)M(w;t) - t(1-ut)L(w;t) - t + u)}{(1-tu)(-tu^2 + t^2u + u + t^3\,w\,u - u\,t\,w - t)}$$

and

$$g_n(v,w;t) = \frac{t\,w\,(t^2v - t\,w + t\,(tv - w)\,(L(w;t) + L(v;t)) + (w - t\,w - t\,v)\,M(w;t) + t^2v\,M(v;t))}{-w\,v + t\,w^2 + w\,t\,v + t\,v^2 - t^2\,v\,w - t^3\,v\,w}$$

In order to compute $p^2(1,1,1;t)$ it is sufficient to determine $L(1;t)$ and $M(1;t)$. We first apply the kernel method to the equation for $ft_o(u,w;t)$. The kernel is $(-tu^2+t^2u+u+t^3\,w\,u-u\,t\,w-t)$. It has two roots and we denote by $U_0(w,t)$ the one which is a formal power series in t. Then

$$U_0(w,t)=\frac{1-t\,w+t^2+t^3\,w-\sqrt{(1-t^2)(1-t-wt-wt^2)(1+t-wt+wt^2)}}{2t}$$

Canceling the kernel by substituting $u=U_0(w;t)$ implies that:

$$(1-tU_0(w,t))(1-t)M(w;t)-t(1-tU_0(w,t))L(w;t)-t+U_0(w,t)=0 \tag{10}$$

Now, by setting respectively $v=1$ and $w=1$ in $g_n(v,w;t)$ we obtain the following system:

$$\begin{aligned}
g_n(1,w;t) &= \frac{t\,w\,(t^2-t\,w+t\,(t-w)\,(L(w;t)+L(1;t))+(w-t\,w-t)\,M(w;t)+t^2\,M(1;t))}{-w+t\,w^2+w\,t+t-t^2\,w-t^3\,w}\\
g_n(v,1;t) &= \frac{t\,(t^2v-t+t\,(tv-1)\,(L(1;t)+L(v;t))+(1-t-t\,v)\,M(1;t)+t^2v\,M(v;t))}{-v+t+t\,v+t\,v^2-t^2\,v-t^3\,v}
\end{aligned}$$

Observe that the two equations have the same denominator up to substituting $w=v$. Then they have a common kernel $K(v;t)=-v+t+t\,v+t\,v^2-t^2\,v-t^3\,v$.

We apply the kernel method to $g_n(1,w;t)$ and $g_n(v,1;t)$, taking the unique formal power series $V_0(t)$ that is solution of $K(v;t)$:

$$V_0(t)=\frac{1-t+t^2+t^3-\sqrt{(1-t^4)(1-2t-t^2)}}{2t} \tag{11}$$

Observe that $U_0(1;t)=V_0(t)$ (remark that it is also equal to the $U_0(t)$ of Section 2). Then, by canceling the kernel in the expressions of $g_n(1,w;t)$ and $g_n(v,1;t)$ we get

$$t^2-tV_0(t)+t\,(t-V_0(t))\,(L(V_0(t);t)+L(1;t))+(V_0(t)-tV_0(t)-t)\,M(V_0(t);t)+t^2\,M(1;t)=0 \tag{12}$$

$$t^2V_0(t)-t+t\,(tV_0(t)-1)\,(L(1;t)+L(V_0(t);t))+(1-t-tV_0(t))\,M(1;t)+t^2V_0(t)M(V_0(t);t)=0 \tag{13}$$

Equations (12), (13), equation (10) with $w=1$, and equation (10) with $w=V_0(t)$ form a system of linear equations that determines $L(1;t)$, $M(1;t)$, $L(V_0(t);t)$, $M(V_0(t);t)$. Finally, by substituting $L(1;t)$ and $M(1;t)$ in the equation for $p^2(1,1,1;t)$ (see Lemma 2) we obtain the following expression for $p^2(1,1,1;t)$:

$$p^2(1,1,1;t)=$$

$$\frac{2\,(-4\,t^4+2\,t^3-2\,t^2-t+2\,t^6-3\,t^5+t^7+1)\,t\,U_0(1)\,U_0(V_0)-2\,(t^4+2\,t^3-4\,t-2\,t^2+1)\,t\,U_0(1)-2\,t^3\,(t^2+t^3-1-3\,t)\,U_0(V_0)+2\,(t+2)\,(-1+t)\,t}{(t^2+2\,t-1)\,(t\,U_0(t)-1)\,(t\,U_0(V_0)-1)\,(t-U_0(V_0)+1-U_0(1))}$$

where $U_0(1)=U_0(1;t)$ and $U_0(V_0)=U_0(V_0(t);t)$. In particular $p^2(1,1,1;t)$ is algebraic of degree 4. The first terms of this series are $4t+12t^2+34t^3+90t^4+236t^5+612t^6+1580t^7+O(t^8)$

5. The class of prudent walks

Now we consider the complete class of prudent walks, i.e. prudent walks without restrictions on the subsequences of steps (see Figure 2). Let

$$p(u,u',v,w;t)=\sum_{w\in P}u^{k(w)}u'^{l(w)}v^{i(w)}w^{j(w)}t^{|w|}$$

be the generating function of P with respect to the distances and the length.

Observe that these prudent walks are symmetric with respect to all directions. Then, in order to count the walks of the class P, we just need to know the generating functions for the following classes:

- H_n is the subclass of P formed by the prudent walks ending with a $\uparrow$ step, and such that their last step moved the top of the previous prudent box. In particular, this means that $i(w) = 0$ for each $w \in H_n$.
- GR_n is the subclass of P formed by the prudent walks ending with a $\uparrow$ step on the right side of their prudent box, and such that their prudent box is the same as the one at the previous step. In particular, this means that $k(w) = 0$ for each $w \in G_n$.

In order to write equations for H_n and GR_n we need to introduce some other classes, as we did for walks of first and second types. However these classes are all symmetric to H_n or GR_n, so we omit their definition. The equations for the associated generating functions $h_n(u, u', w)$ and $gr_n(u', v, w)$ are obtained by using the same arguments as for walks of the first and second types:

$$\begin{aligned} h_n(u,u',w;t) &= t\,w\,(1 + h_n(u,u',w;t) + gr_n(u',0,w;t) + , gl_n(u,0,w;t) \\ &\quad + h_o(u,0,w;t) + gt_o(u,u',w;t) + h_e(u',0,w;t) + gt_e(u,u',w;t)) \\ gr_n(u',v,w;t) &= \tfrac{t\,w}{v}(gr_n(u',v,w;t) - gr_n(u',0,w;t) + h_e(u',v,w;t) - h_e(u',0,w;t)) \end{aligned}$$

Equations for all other classes could be obtained by symmetry from these two ones. Alternatively, by using directly the symmetries to express all generating functions in terms of h_n and gr_n, the two previous equations can be rewritten:

$$\begin{aligned} h_n(u,u',w;t) &= t\,w\,(1 + h_n(u,u',w;t) + gr_n(u',0,w;t) + gr_n(u,0,w;t) \\ &\quad + h_n(0,w,u;t) + gr_n(w,u',u;t) + h_n(w,0,u';t) + gr_n(w,u,u';t)) \\ gr_n(u',v,w;t) &= \tfrac{t\,w}{v}(gr_n(u',v,w;t) - gr_n(u',0,w;t) + h_n(w,v,u';t) - h_n(w,0,u';t)) \end{aligned} \tag{14}$$

This system entirely define the series $h_n(u, u', w; t)$ and $gr_n(u', v, w; t)$. We also have that:

$$p(1,1,1,1;t) = 4\,h_n(1,1,1;t) + 8\,gr_n(1,1,1;t)$$

Then, by using this equation and (14) we can compute the first terms of $p(1, 1, 1, 1; t)$:

$$4t + 12\,t^2 + 36\,t^3 + 100\,t^4 + 276\,t^5 + 748\,t^6 + 2012\,t^7 + 5356\,t^8 + O(t^9)$$

However, we were not able to find a solution for system (14) by similar methods to those in the previous sections.

Acknowledgments. We thank Gilles Schaeffer for suggesting the problem and for stimulating discussions, and Mireille Bousquet-Mélou for communicating [7].

References

[1] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, and D. Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Math.*, 2002.
[2] M. Bousquet-Mélou and M. Petkovsek. Linear recurrences with constant coefficients: the multivariate case. *Discrete Math.*, 2000.
[3] I Dutour and J.M. Fédou. Object grammars and random generation. *Discrete Math. and Theoret. Comput. Sci.*, 2:47–61, 1998.
[4] I Dutour and J.M. Fédou. Object grammars and bijections. *Theoret. Comput. J.*, 290:1915–1929, 2003.
[5] B. Hayes. How to avoid yourself. *Amer. Sci.*, 86, 1998.
[6] N. Madras and G. Slade. *The self-avoiding walk*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1993.
[7] P. Préa. Exterior self-avoiding walks on the square lattice. manuscript.
[8] R. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, 1999.

Enrica Duchi, LIAFA, Paris 7, France
E-mail address: duchi@liafa.jussieu.fr

A COMBINATORIAL APPROACH TO JUMPING PARTICLES: THE PARALLEL TASEP

ENRICA DUCHI* AND GILLES SCHAEFFER†

Abstract. In this paper we continue the combinatorial study of models of particles jumping on a row of cells which we initiated with the standard totally asymmetric exclusion process or TASEP (*Journal of Combinatorial Theory, Series A*, to appear). We consider here the parallel TASEP, in which particles can jump simultaneously. On the one hand, the interest in this process comes from highway traffic modeling: it is the only solvable special case of the Nagel-Schreckenberg automaton, the most popular model in that context. On the other hand, the parallel TASEP is of some theoretical interest because the derivation of its stationary distribution, as appearing in the physics literature, is harder than that of the standard TASEP.

We offer here an elementary derivation that extends the combinatorial approach we developed for the standard TASEP. In particular we show that this stationary distribution can be expressed in terms of refinements of Catalan numbers.

Résumé. L'objet de cet article est de poursuivre l'étude combinatoire d'une famille de modèles de particules sauteuses que nous avons commencé avec le cas du processus d'exclusion totalement asymétrique standard, ou TASEP (*Journal of Combinatorial Theory, Series A*, to appear). Nous traitons ici le TASEP parallèle, dans lequel les particules peuvent sauter simultanément. L'étude de ce processus est motivée par les nombreux travaux de modélisation du trafic automobile qui portent sur l'automate de Nagel-Schreckenberg: le TASEP parallèle est en effet la seule instance de cet automate stochastique dont la distribution stationnaire soit connue. De plus, le TASEP parallèle présente l'intérêt théorique que la détermination de sa distribution stationnaire par des méthodes de physique mathématique est plus délicate que pour le TASEP standard.

Nous utilisons une approche combinatoire qui étend l'approche que nous avions développée pour le TASEP standard. En particulier nous montrons que cette distribution peut-être décrite en termes de raffinements des nombres de Catalan.

1. Jumping particles and the TASEP family

The aim of this article is to continue the combinatorial study of a family of models of particles jumping on a row of cells that are known in the physics and probability literature as *one dimensional totally asymmetric exclusion processes* (TASEPs for short). In order to define TASEPs we first introduce a set of configurations and some rules.

A *TASEP configuration* is a row of n cells, separated by $n+1$ walls (the leftmost and rightmost ones are borders). Each cell is occupied by one particle, and each particle has a type, black or white (see Figure 1).

|●|○|○|●|●|●|○|●|○|○|

Figure 1. A TASEP configuration with $n = 10$ cells, 5 black particles, and 5 white particles.

The transitions of the TASEP are based on a mapping ϑ that modifies a configuration τ near a wall i to produce a configuration $\vartheta(\tau, i)$. Given a pair (τ, i) the following rules define its image $\vartheta(\tau, i)$:

a. Rule $\bullet|\circ \to \circ|\bullet$: If the wall i separates a black particle (on its left) and a white particle (on its right), then two particles swap to give $\vartheta(\tau, i)$.

b. Rule $|\circ \to |\bullet$: If the wall is the left border ($i = 0$) and the leftmost cell contains a white particle, this white particle leaves the row and it is replaced by a black particle.

c. Rule $\bullet| \to \circ|$: If the wall is the right border ($i = n$) and the rightmost cell contains a black particle, this black particle leaves the row and it is replaced by a white particle.

d In the other cases, nothing happens, $\vartheta(\tau, i) = \tau$.

Date: May 4, 2005.

* Supported by a post-doctoral grant of the CNRS.

† Supported in part by EC's IHRP Programm, within the Research Training Network *Algebraic Combinatorics in Europe*, grant HPRN-CT-2001-00272.

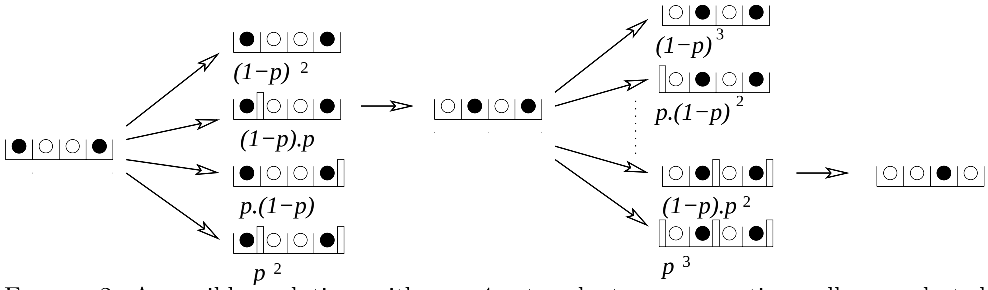


FIGURE 2. A possible evolution, with $n = 4$: at each step, some active walls are selected with the given probabilities, and they trigger a transition.

Given a configuration τ, let $M(\tau)$ be the set of walls on which the previous application ϑ can effectively do something: inner walls with a neighborhood of the form $\bullet|\circ$, or borders with a neighborhood of the form $|\circ$ or $\bullet|$. The definition of ϑ can be extended to any subset $\{i_1, \dots, i_k\}$ of $M(\tau)$ by setting $\vartheta(\tau, i_1, \dots, i_k) = \vartheta(\vartheta(\tau, i_1, \dots, i_{k-1}), i_k)$. Observe that this extended application can be interpreted as performing moves at walls $i_1, \dots, i_k$ in parallel since the basic application acts only locally and $M(\tau)$ never contains two adjacent walls. A pair (τ, A) with $A \subset M(\tau)$ will be referred to as an *active configuration*, and from now on ϑ is considered as a mapping from the set of active configurations into the set of TASEP configurations.

In the previous work [4], we dealt with several variants of *sequential* TASEP. In particular, the standard sequential TASEP with open boundaries is a Markov chain S^{seq} on the set of TASEP configurations with n cells whose dynamic is defined as follows in terms of ϑ:

- Let $\tau = S^{\text{seq}}(t)$ be the current configuration.
- Choose a uniform random wall i in $\{0, \dots, n\}$.
- Set $S^{\text{seq}}(t+1) = \vartheta(\tau, i)$.

The aim of the present article is to extend our approach to a more general model in which particles are allowed to jump simultaneously: the *parallel* TASEP is a Markov chain $S^{/\!/}$ on the set of TASEP configurations whose dynamic is defined as follows in terms of ϑ:

- Let $\tau = S^{/\!/}(t)$ be the current configuration, $M = M(\tau)$, and $m = m(\tau) = |M|$.
- Choose a random subset A of M by independently giving to each wall of M probability p to be taken. In other terms, the probability that $A = \{i_1, \dots, i_k\}$ for some given distinct elements $i_1, \dots, i_k$ of M is $p^k(1-p)^{m-k}$. The walls in A are referred to as the *active* walls of the active configuration (τ, A).
- Then set $S^{/\!/}(t+1) = \vartheta(\tau, A)$.

Figure 2 illustrates the application of these rules, with active walls appearing as $\|$. Observe that the transformation ϑ makes black particles travel from left to right, and makes white particles do the opposite.

The difference between the sequential and the parallel TASEP is thus that in the first process, only one wall can trigger a move at a time, while in the second simultaneous moves can occur. Observe that if p is very small, it is unlikely that more than one particle jump at a time (since $p^2 \ll p$). This implies that in the limit $p \to 0$, the parallel TASEP reduces to a (very slow) sequential TASEP.

We got interested in the TASEP because Derrida *et al.* [2, 3] proved that the stationary distribution of this Markov chain involves Catalan numbers. In [4], we gave a combinatorial explanation of this fact. In the present paper we extend our combinatorial approach to derive the stationary distribution of the parallel TASEP. Another extension, to the partially symmetric sequential TASEP, was recently developed in [7]. Although we concentrate here on TASEPs with open boundary conditions, it is worth indicating that variants with periodic boundaries can be defined and studied similarly (identifying walls 0 and n and concentrating on rule a). For these variants an alternative combinatorial interpretation was proposed in [1].

The stationary distribution of the parallel TASEP was first obtained by Schadschneider *et al.* (see [8] and ref. therein) in the easier case of periodic boundaries and by Evans *et al.* [5] in the case with open boundaries. This last derivation is based on the same *matrix ansatz* approach developed by Derrida *et al.* for the sequential TASEP [3], but requires the introduction of a quartic (instead of quadratic) algebra. This extra complexity reflects in our combinatorial approach, in the sense that, with respect to [4], new ingredients are necessary to construct the covering Markov chain on which we rely. However a nice feature of our approach is that we are able to remain within the realm of Catalan combinatorial structures.

To conclude this introduction, it is worth stressing the fact that the determination of the stationary distribution allows to compute some physical quantities related to the model (densities, flows, phase diagrams,...). Our approach, while providing a new simpler derivation and a combinatorial interpretation of the stationary distribution, leads then to the same computations as far as these next steps are concerned. We thus do not reproduce the corresponding discussions.

2. The combinatorial approach

Our method to study the TASEP consists in the construction of a new *covering* Markov chain $X^{/\!/}$ on a set Ω_n of *complete configurations*, that satisfies two main requirements: on the one hand the stationary distribution of the TASEP chain $S^{/\!/}$ can be simply expressed in terms of that of the covering chain $X^{/\!/}$; on the other hand the stationary distribution of the covering chain $X^{/\!/}$ can be expressed by means of a combinatorial parameter defined on the set Ω_n.

2.1. **The complete Markov chain.** Since the parallel TASEP $S^{/\!/}$ yields back the sequential TASEP S^{seq} for $p \to 0$, we first try to adapt directly the construction of [4] to simultaneous jumps.

Following [4], define a *complete configuration* of Ω_n to be a pair of rows of particles satisfying the following constraints: (i) there is an equal number of black and white particles (the *balance condition)*; (ii) on the left hand side of any vertical wall there are at least as many black particles as white ones (the *positivity condition*). An example of complete configuration is given in Figure 3. The number of elements of Ω_n is the nth Catalan number $\frac{1}{n+1}\binom{2n}{n}$ (see [4], although readers with a background in combinatorics may as well recognize directly bicolored Motzkin paths in disguise).

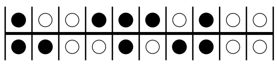


Figure 3. A complete configuration with $n = 14$.

Given a configuration ω of Ω_n, let $\text{top}(\omega)$ denote its first row, which is a TASEP configuration. Still in the steps of [4], we look for a covering chain $X^{/\!/}$ on Ω_n that mimics in the top row the TASEP $S^{/\!/}$. More precisely we would like to define $X^{/\!/}$ exactly as $S^{/\!/}$ with ϑ replaced by a mapping T that has nice combinatorial properties and that extends ϑ in the following sense: given a configuration ω of Ω_n and a set of active walls A, T should produce a new configuration ω' of Ω_n in such a way that if $\text{top}(\omega) = \tau$, then $\text{top}(\omega') = \vartheta(\tau, A)$.

However such a direct extension of the construction of [4] does not seem to be sufficient to account for the more complex dynamic of the parallel TASEP. Instead we need to introduce a further randomization step that is conveniently described in terms of colors.

A *well colored configuration* (ω, A, R, G) consists of a complete configuration ω, a set A of active walls, a set R of red walls and a set G of green walls, such that:

- the set A of active wall is a subset of $M(\omega) = M(\text{top}(\omega))$, the set of walls around which the local configuration in the first row is $\bullet|\circ$, $|\circ$ or $\bullet|$,
- the set R of red walls and the set G of green walls form a partition (R, G) of the subset $C(\omega, A)$ of A consisting of walls around which the local configuration is $\substack{\bullet\\\bullet}|\substack{\circ\\\circ}$.

We now define a parameter Q on complete configurations and its "randomized" version q on the set of well colored configurations. Given (ω, A, R, G) a well colored configuration, we set

$$Q(\omega) \;=\; (1-p)^{h(\omega)-m(\omega)}, \qquad \text{and} \qquad q(\omega, A, R, G) \;=\; p^{|A|}(1-p)^{h(\omega)-|A|} \cdot p^{|R|}(1-p)^{|G|},$$

where $h(\omega)$ denote the number of columns of the form $|\substack{\bullet\\\bullet}|$, $|\substack{\bullet\\\circ}|$ and $|\substack{\circ\\\bullet}|$ in ω, and $m(\omega) = |M(\omega)|$. A configuration ω contains the same number of black and white particles, so that $h(\omega)$ can be rewritten in various ways:

$$\begin{aligned} h(\omega) &= \left|\left\{\text{columns } |\genfrac{}{}{0pt}{}{\bullet}{\circ}| \text{ or } |\genfrac{}{}{0pt}{}{\circ}{\bullet}|\right\}\right| + \frac{1}{2}\cdot\left|\left\{\text{columns } |\genfrac{}{}{0pt}{}{\bullet}{\bullet}| \text{ or } |\genfrac{}{}{0pt}{}{\circ}{\circ}|\right\}\right| \\ &= \frac{n}{2} + \frac{1}{2}\cdot\left|\left\{\text{columns } |\genfrac{}{}{0pt}{}{\bullet}{\circ}| \text{ or } |\genfrac{}{}{0pt}{}{\circ}{\bullet}|\right\}\right| \;=\; n - \left|\left\{\text{columns } |\genfrac{}{}{0pt}{}{\circ}{\circ}|\right\}\right|. \end{aligned}$$

Again, readers with a background in enumerative combinatorics will recognize the statistic $h(\omega)$ as $n/2$ plus half of the number of horizontal steps in the bicolored Motzkin path associated to the configuration ω.

Observe that

$$q(\omega, A, R, G) \quad = \quad Q(\omega)\cdot p^{|A|}(1-p)^{m(\omega)-|A|}\cdot p^{|R|}(1-p)^{|G|}. \tag{1}$$

Given ω, we can thus apply the binomial formula $\sum_{U\subset V} x^{|U|}y^{|V|-|U|} = (x+y)^{|V|}$ to sum over all partitions (R, G) of $C(\omega, A)$, and then again to sum over all subsets $A \subset M(\omega)$:

$$\sum_{A,R,G} q(\omega, A, R, G) \quad = \quad Q(\omega), \tag{2}$$

where the summation is over all triples (A, R, G) such that (ω, A, R, G) is a well colored configuration.

The key of our combinatorial approach is that we can construct an application T that behaves nicely with respect to the parameter q. The construction is given in Section 3.

Theorem 1. *There is a bijection $\bar{T}$ from the set of well colored configurations onto itself such that:*

- *The mapping T, defined as the first component of $\bar{T}$, mimics in the top row the mapping ϑ. More explicitly, if $(\omega', A', R', G') = \bar{T}(\omega, A, R, G)$ is the image of a well colored configuration (ω, A, R, G) by $\bar{T}$, then $\omega' = T(\omega, A, R, G)$ satisfies*

$$\mathrm{top}(\omega') = \vartheta(\mathrm{top}(\omega), A).$$

- *The parameter q is preserved by the bijection $\bar{T}$: for any well colored configuration (ω, A, R, G),*

$$q(\bar{T}(\omega, A, R, G)) = q(\omega, A, R, G).$$

Observe that the first property of $\bar{T}$ completely defines its action on the first row of configurations: in particular, T must move black particles from left to right in the first row. We shall see in Section 3, when we explicitly describe $\bar{T}$ that in the second row it will move black particles in the opposite direction, from right to left.

With Theorem 1 at hand, we are in the position to define our Markov chain $X^{/\!\!/}$ on the set Ω_n:

- Let $\omega = X^{/\!\!/}(t)$ be the current configuration, $M = M(\omega) = M(\mathrm{top}(\omega))$, and $m = m(\omega) = |M|$.
- Choose a random subset A of M by independently giving to each wall of M probability p to be taken. In other terms, the probability that $A = \{i_1, \ldots, i_k\}$ for some given distinct elements $i_1, \ldots, i_k$ of M is $p^k(1-p)^{m-k}$. The walls in the set A are referred to as active walls, and the pair (ω, A) as an active configuration.
- Next, let $C = C(\omega, A)$ be the subset of A consisting of walls around which the local configuration is $\genfrac{}{}{0pt}{}{\bullet}{\bullet}|\genfrac{}{}{0pt}{}{\circ}{\circ}$. Then each wall of C is colored red with probability p or green with probability $1-p$. In other terms we randomly partition C into R (red walls) and G (green walls), and associate to the active configuration (ω, A) a well colored configuration (ω, A, R, G).
- Then set $X^{/\!\!/}(t+1) = T(\omega, A, R, G)$.

See Figure 4 for an illustration.

Let us compare the dynamic of $X^{/\!\!/}$ and $S^{/\!\!/}$. In the chain $X^{/\!\!/}$, a supplementary random coloring step is performed that does not exist in the chain $S^{/\!\!/}$. In particular we allow the action of T to depend on this distinction between colors, and this will be used in the actual construction of the mapping T in Section 3. However the colors only affect the bottom row: the action of T on the top row depends only on A and, as already indicated, mimics ϑ. As a consequence, if one only considers the top row, the coloring step in the definition of $X^{/\!\!/}$ can be ignored, and we obtain the following relation between $X^{/\!\!/}$ and $S^{/\!\!/}$.

Proposition 1. *The chains $S^{/\!\!/}$ and top$(X^{/\!\!/})$ have the same dynamics and the same stationary distributions.*

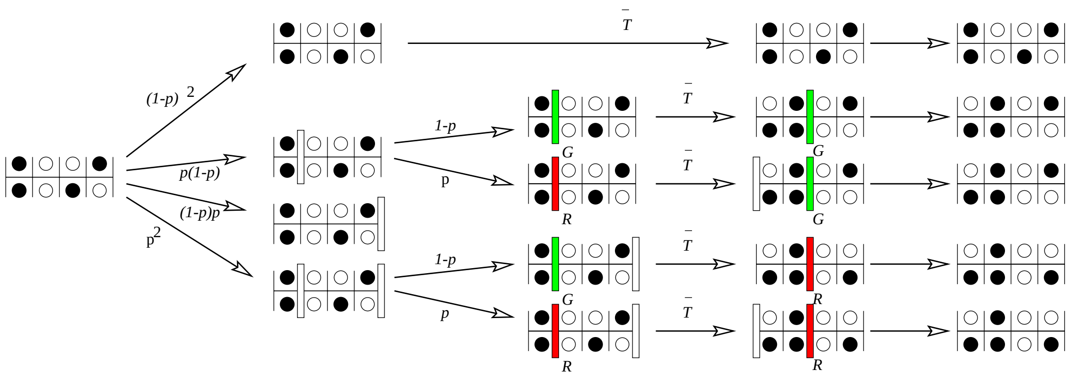


FIGURE 4. The possible transitions for a configuration ω with $M(\omega) = \{1, 4\}$.

2.2. **The stationary distribution.** As we shall see in Section 3, the transitions of the complete chain $X^{/\!/}$ that originate from exactly one wall and such that this wall is not red are exactly the transitions of the chain X studied in [4] in relation with the sequential TASEP. The chain X was proved irreducible there and this implies that the chain $X^{/\!/}$, that has more transitions, is irreducible as well for $0 < p < 1$. Moreover there is a positive probability to stay in any configuration, so that the chain $X^{/\!/}$ is aperiodic. A classical result of the theory of finite Markov chain is that an irreducible aperiodic chain has a unique stationary distribution to which it converges [6]. Our main result is then the following theorem.

Theorem 2. *The stationary distribution of the Markov chain $X^{/\!/}$ is proportional to the parameter*

$$Q(\omega) \;=\; (1-p)^{h(\omega)-m(\omega)},$$

*where $h(\omega)$ is the number of columns of the form $|{}^{\bullet}_{\bullet}|$, $|{}^{\bullet}_{\circ}|$ or $|{}^{\circ}_{\bullet}|$ and $m(\omega)$ the number of walls at which a transition could occur in the first row (*i.e *walls around which the local configuration in the top row is $\bullet|\circ$, $\bullet|$ or $|\circ$). In other terms,*

$$Prob(X^{/\!/}(t)=\omega) \underset{t\to\infty}{\longrightarrow} \frac{1}{Z_n(1-p)}\,(1-p)^{h(\omega)-m(\omega)} \qquad \text{where} \qquad Z_n(x) = \sum_{\omega\in\Omega_n} x^{h(\omega)-m(\omega)}.$$

In particular $Z_n(x)$ is a combinatorial refinement of the Catalan numbers $Z_n(1) = \frac{1}{n+1}\binom{2n}{n}$.

An immediate consequence of this theorem is the following interpretation of the stationary distribution of the Markov chain $S^{/\!/}$ in terms of weighted complete configurations with fixed top row.

Theorem 3. *The stationary distribution of the Markov chain $S^{/\!/}$ is proportional to*

$$\pi(\tau) \;=\; \sum_{top(\omega)=\tau} (1-p)^{h(\omega)-m(\tau)}.$$

Let a balanced substring of τ be a subword $\sigma = \tau_{i_1}\tau_{i_2}\ldots\tau_{i_{2r}}$ of τ that is a balanced parenthesis word with black and white particles respectively viewed as opening and closing parentheses. In this case write $\sigma \vdash \tau$ and $|\sigma| = r$. Then $\pi(\omega)$ can be rewritten as

$$\pi(\tau) = (1-p)^{n-m(\tau)} \sum_{\sigma\vdash\tau} (1-p)^{-|\sigma|}.$$

For instance, for a configuration of the type $\circ\cdots\circ\bullet\cdots\bullet$, we obtain

$$\pi(\underbrace{\circ\cdots\circ}_{k}\underbrace{\bullet\cdots\bullet}_{n-k}) = (1-p)^{n-3}, \qquad \textit{for all } 0<k<n,$$

which is independent of k, whereas for $\bullet\cdots\bullet\circ\cdots\circ$, the probability depends on k as follows:

$$\pi(\underbrace{\bullet\cdots\bullet}_{k}\underbrace{\circ\cdots\circ}_{n-k}) = \sum_{r=0}^{\min(k,n-k)} \binom{k}{r}\binom{n-k}{r}(1-p)^{n-1-r}, \qquad \textit{for all } 0<k<n.$$

Corollary 1. *In the limit $p \to 0$, we recover the stationary distribution of the sequential TASEP: indeed, for $p \to 0$, we get $Q(\omega) = 1$ for all ω, and $Z_n(1) = |\Omega_n| = \frac{1}{n+1}\binom{2n}{n}$:*

$$\Pr(S^{seq} = \tau) = \frac{|\{\omega \mid top(\omega) = \tau\}|}{\frac{1}{n+1}\binom{2n}{n}}.$$

2.3. **Proof of the stationarity.** Our aim is to show that the unnormalized distribution $Q(\omega)$ is stationary. Let us thus assume that $\mathrm{Prob}(X^{/\!/}(t) = \omega) = c_0\, Q(\omega)$ for some constant c_0 and all $\omega \in \Omega_n$, and let us compute $\mathrm{Prob}(X^{/\!/}(t+1) = \omega')$.

By construction, the probability to be in configuration ω' at time $t+1$ is the sum over the probability to be at time t in a configuration ω multiplied by the probability to select subsets A, R, and G such that $T(\omega, A, R, G) = \omega'$. More precisely, the probability to select A from $M(\omega)$ is $p^{|A|}(1-p)^{m(\omega)-|A|}$, and the probability to select R and G is $p^{|R|}(1-p)^{|G|}$. Hence

$$\begin{aligned}\mathrm{Prob}\left(X^{/\!/}(t+1) = \omega'\right) &= \sum_{(w,A,R,G)\in T^{-1}(\omega')} \mathrm{Prob}\left(X^{/\!/}(t) = \omega\right) p^{|A|}(1-p)^{m(\omega)-|A|} p^{|R|}(1-p)^{|G|},\\ &= \sum_{(w,A,R,G)\in T^{-1}(\omega')} c_0\, q(w, A, R, G),\end{aligned}$$

where the second line follows from the hypothesis $\mathrm{Prob}(X^{/\!/}(t) = \omega) = c_0\, Q(\omega)$ and from Formula (1).

In view of Theorem 1, T is the first component of the bijection $\bar{T}$, so that $T^{-1}(\omega')$ is the inverse image by $\bar{T}$ of the set of well colored configurations of the form (ω', A', R', G'). This implies

$$\mathrm{Prob}\left(X^{/\!/}(t+1) = \omega'\right) = \sum_{A',R',G'} c_0\, q(\bar{T}^{-1}(w', A', R', G')) = \sum_{A',R',G'} c_0\, q(w', A', R', G') = c_0\, Q(\omega').$$

where the summations are over all triples (A', R', G') such that (ω', A', R', G') is a well colored configuration, the second equality follows from the invariance of q under the action of $\bar{T}$, as stated in Theorem 1, and the last equality is Formula (2).

3. The bijection $\bar{T}$

In this section we prove Theorem 1 by describing a bijection $\bar{T}$ that transports the parameter q. We first give the definitions of some local operations, and use them to describe an intermediate mapping ψ. Finally we present $\bar{T}$ and check that $\bar{T}$ satisfied the requirements.

3.1. **Local operations.** We shall use two types of local operations: *deletions* map configurations of Ω_n to configurations of Ω_{n-1}, while *insertions* map configurations of Ω_{n-1} to configurations of Ω_n. In the following definitions, the numbering of walls always refers to the configuration of Ω_n:

- A *right deletion* at $i \neq 0$ consists in
  - if $i \neq n$, removing a $|\substack{\circ\\ \bullet}|$-column on the right of i, that is: $\omega_1|\substack{y\\ x}\|_i\substack{\circ\\ \bullet}|\omega_2 \mapsto \omega_1|\substack{y\\ x}|\omega_2$.
  - if $i = n$, removing a $|\substack{\bullet\\ \circ}|$-column at the right border, that is: $\omega_1|\substack{\bullet\\ \circ}\|_n \mapsto \omega_1|$.
- A *left deletion* at $i \neq n$ consists in
  - if $i \neq 0$, removing a ${}^{\bullet}|_{\circ}$-diagonal around i, that is: $\omega_1|\substack{\bullet\\ x}\|_i\substack{y\\ \circ}|\omega_2 \mapsto \omega_1|\substack{y\\ x}|\omega_2$.
  - if $i = 0$, removing a $|\substack{\circ\\ \bullet}|$-column on the left border, that is: $\|_0\substack{\circ\\ \bullet}|\omega_2 \mapsto |\omega_2$.
- A *right insertion* at $j \neq n$ consists in
  - if $j \neq 0$, inserting a $|\substack{\circ\\ \bullet}|$-column on the right of wall j, that is: $\omega_1|\substack{y\\ x}|\omega_2 \mapsto \omega_1|\substack{y\\ x}\|_j\substack{\circ\\ \bullet}|\omega_2$.
  - if $j = 0$, inserting a $|\substack{\circ\\ \bullet}|$-column on the left border, that is: $|\substack{y\\ x}|\omega_2 \mapsto \|_0\substack{\circ\\ \bullet}|\substack{y\\ x}|\omega_2$.
- A *left insertion* at $j \neq 0$ consists in
  - if $j \neq n$, inserting a ${}^{\bullet}|_{\circ}$-diagonal around column j, that is: $\omega_1|\substack{y\\ x}|\omega_2 \mapsto \omega_1|\substack{\bullet\\ x}\|_j\substack{y\\ \circ}|\omega_2$.
  - if $j = n$, inserting a $|\substack{\bullet\\ \circ}|$-column at the right border, that is: $\omega_1|\substack{y\\ x}| \mapsto \omega_1|\substack{y\\ x}|\substack{\bullet\\ \circ}\|_n$.

In the case of the sequential TASEP [4], we used these operations to construct a bijection from the set of complete configurations with exactly one active wall, onto itself. The bijection essentially consisted in applying one deletion at the active wall and one insertion at another nearby wall. Since the main difference between the sequential and the parallel TASEP is the fact that there may be several active walls, one could just try to apply the bijection of [4] to all these walls in parallel. However, this naive approach fails when there are two active walls that are too close, because the transformations applied to nearby walls may interfere.

In order to circumvent this ambiguity a new operation is needed:

- A *block deletion* at a wall i of type $|\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix}\|_{i}\begin{smallmatrix}\circ\\\circ\end{smallmatrix}|$ consists in removing the block around i: $\omega_1|\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix}\|_{i}\begin{smallmatrix}\circ\\\circ\end{smallmatrix}|\omega_2 \mapsto \omega_1|\omega_2$.

Using the corresponding insertion, it would be possible to describe the bijection $\bar{T}$ directly as the simultaneous application of some deletions and insertions near active walls. It will however prove more convenient to give a sequential description in terms of a partial application ψ.

3.2. **The mapping ψ.** Given a configuration (ω, A, R, G), a *pointer* is an element of $A \cup \{\perp, \top\}$. The value $\perp$ and $\top$ are respectively interpreted as positions of the pointer to the left and to the right of the configuration. A pair $(\omega, A, R, G; i)$ is a *right admissible configuration* if (ω, A, R, G) is a well colored configuration and $i \in A \cup \{\top\}$, or if $(\omega, A \setminus \{i+1\}, R, G)$ is a well colored configuration and the local configuration between walls i and $i+2$ is $|_{i}\begin{smallmatrix}\circ\\?\end{smallmatrix}|\begin{smallmatrix}\circ\\\bullet\end{smallmatrix}|$. *Left admissible configurations* are defined similarly with $\top$ replaced by $\perp$.

We are now ready to describe a mapping ψ, that maps right admissible configurations onto left admissible ones. The image $(\omega', A', R', G'; r')$ of a right admissible configuration $(\omega, A, R, G; r)$ by ψ is obtained by applying some local operations near the pointer.

The value $\top$ serves as a initialization case: the image of a pair $(\omega, A, R, G; \top)$ by ψ is $(\omega, A, R, G; \max(A))$ if $A \neq \emptyset$, and $(\omega, A, R, G; \perp)$ otherwise. When the pointer i is in A, the image $(\omega', A', R', G'; i')$ of $(\omega, A, R, G; i)$ depends on the local configuration around i, and, if it exits, on the rightmost active wall $m < i$ on the left of i:

A. Cases $\begin{smallmatrix}\bullet\\\circ\end{smallmatrix}\|_{i}\begin{smallmatrix}\circ\\\circ\end{smallmatrix}$, or $\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix}\|_{i}\begin{smallmatrix}\circ\\\circ\end{smallmatrix}$ with i green, or $\|_{0}\begin{smallmatrix}\circ\\\bullet\end{smallmatrix}$ with $i = 0$: (see Figure 5)

Two operations are performed on ω to produce its image ω'. The first one is a left deletion at i. For the second let $j_2 > i$ be the rightmost wall such that there are only black particles in the top row between walls $i+1$ and j_2. The second operation is a left insertion at j_2. There are two possibilities:

- The wall j_2 was not active ($j_2 \notin A$). The wall j_2 replaces the wall i in the set of active walls: $A' = A \cup \{j_2\} \setminus \{i\}$. Moreover, if there is a black particle on its bottom right in ω, the wall j_2 is colored green: $G' = G \cup \{j_2\} \setminus \{i\}$.
- The wall j_2 was active ($j_2 \in A$). Then i is removed from the set of active walls and the wall j_2 is colored red: $A' = A \setminus \{i\}$, $R' = R \cup \{j_2\}$, $G' = G \setminus \{i\}$.

Other colorings are left unchanged, and in both cases the pointer is set to m if it exists, to $\perp$ otherwise.

B. Cases $\begin{smallmatrix}\bullet\\?\end{smallmatrix}\|_{i}\begin{smallmatrix}\circ\\\bullet\end{smallmatrix}$, or $\begin{smallmatrix}\bullet\\\circ\end{smallmatrix}\|_{n}$ with $i = n$: (see Figure 6)

Two operations are performed on ω to produce its image ω'. The first one is a right deletion at i. To describe the second operation, let $j_1 < i$ be the leftmost wall on the left of i in ω such that there are only white particles in the top row between walls j_1 and $i-1$. There are again several possibilities:

- The wall j_1 was active (so that $m = j_1$). Then the second operation is a right insertion at wall $m+1$, which becomes active: $A' = A \cup \{m+1\} \setminus \{i\}$. (This is the only case in which a wall with local configuration $\circ|\circ$ can become active. As prescribed in the definition of admissible configuration, the pointer will be set on the left of this abnormal active wall.)
- Otherwise, the second operation is a right insertion at the wall j_1. The wall j_1 replaces the wall i in the set of active walls: $A' = A \cup \{j_1\} \setminus \{i\}$.

Other colorings are left unchanged, and in all cases the pointer is set to m if it exists and to $\perp$ otherwise.

C. Case $\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix}\|_{i}\begin{smallmatrix}\circ\\\circ\end{smallmatrix}$ with i red: (see Figure 7)

Three operations are performed on ω to produce its image ω'. The first one consists in removing the block $\begin{smallmatrix}\bullet\\\bullet\end{smallmatrix}\|\begin{smallmatrix}\circ\\\circ\end{smallmatrix}$ around i. The second operation is a left insertion at j_2 as for the cases of type A above, with the same two possibilities for the changes of colors. The third operation is a right insertion at $m+1$ or j_2 as for the cases of type B above. Again the pointer is set to m if m exists and to $\perp$ otherwise.

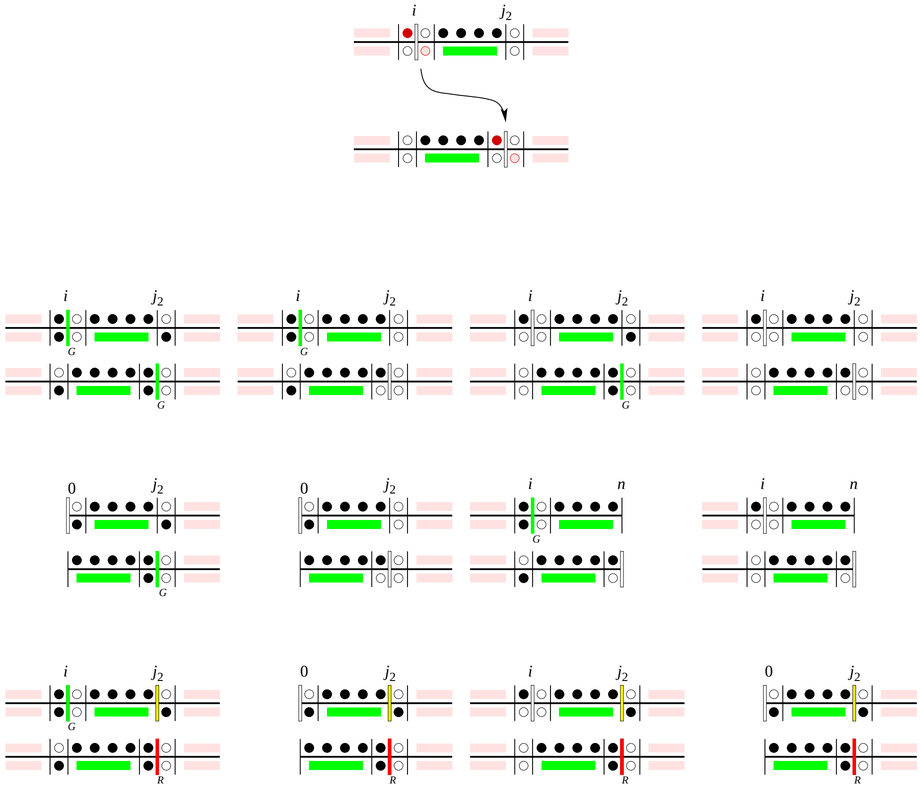


FIGURE 5. The application ψ in cases of type A. The rule is first presented informally. It is then instantiated for all possible local configurations to allow for easy verification of proofs.

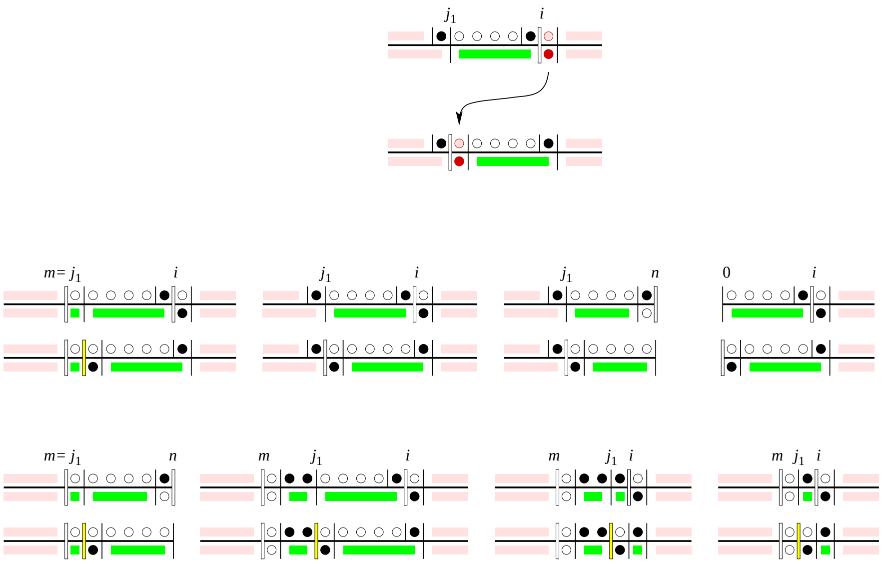


FIGURE 6. The application ψ in cases of type B. The rule is first presented informally. It is then instantiated for all possible local configurations to allow for easy verification of proofs.

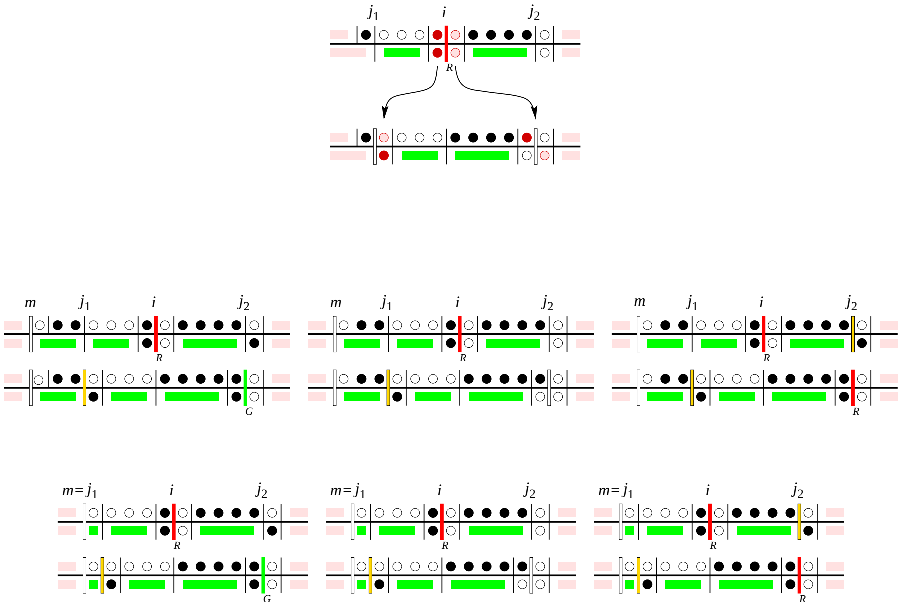


FIGURE 7. The application ψ in cases of type C. The rule is first presented informally. It is then instantiated for a selection of possible local configurations. (All configurations can be retrieved by combining type A to the right and type B to the left.)

Lemma 1. *The mapping ψ is a bijection from the set of right admissible configurations onto the set of left admissible configurations.*

Proof. With the help of Figures 5, 7 and 6, one can easily check the following properties, which together prove the lemma.

First the image of a right admissible configuration by ψ is a left admissible configuration. The fact that the image configuration ω' satisfies the positivity condition follows immediately from the fact that the local operations of Section 3.1 preserve this condition. It should be observed also that if there is an abnormal active wall on the right of the pointer in ω (*i.e.* an active wall with local configuration $\circ|\circ$ in the top row), then the application of ψ brings a black particle to its left, hence turning it back into a normal active wall.

Second a case analysis allows to check that any left admissible configuration has a preimage by ψ. $\square$

3.3. **The bijection $\bar{T}$.** Observe that the mapping ψ always moves the pointer to the left. The mapping $\bar{T}$ is defined by iterating ψ so that the pointer goes from $\top$ to $\bot$.

- Let $\ell := 0$ and $(\omega_0, A_0, R_0, G_0) = (\omega, A, R, G)$, and set the initial pointer to the right of the configuration: $i_0 = \top$.
- Repeat
  - let $(\omega_{\ell+1}, A_{\ell+1}, R_{\ell+1}, G_{\ell+1}; i_{\ell+1}) = \psi(\omega_\ell, A_\ell, R_\ell, G_\ell; i_\ell)$ and $\ell := \ell + 1$.

  until i_ℓ reaches the value $\bot$.
- The image $\bar{T}(\omega, A, R, G)$ is $(\omega_\ell, A_\ell, R_\ell, G_\ell)$.

In view of the properties of the application ψ, the following is immediate.

Proposition 2. *The mapping $\bar{T}$ is a bijection from the set of well colored configurations onto itself. Moreover if $(\omega', A', R', G') = \bar{T}(\omega, A, R, G)$, then* $\mathrm{top}(\omega') = \vartheta(\mathrm{top}(\omega), A)$.

Although it is more convenient to describe $\bar{T}$ in a sequential way, as we did in terms of ψ, it is worth observing again that $\bar{T}$ essentially acts in parallel on all walls, with one pair of particles leaving or arriving at every non-red active wall, and two pairs of particles leaving or arriving at every red active wall.

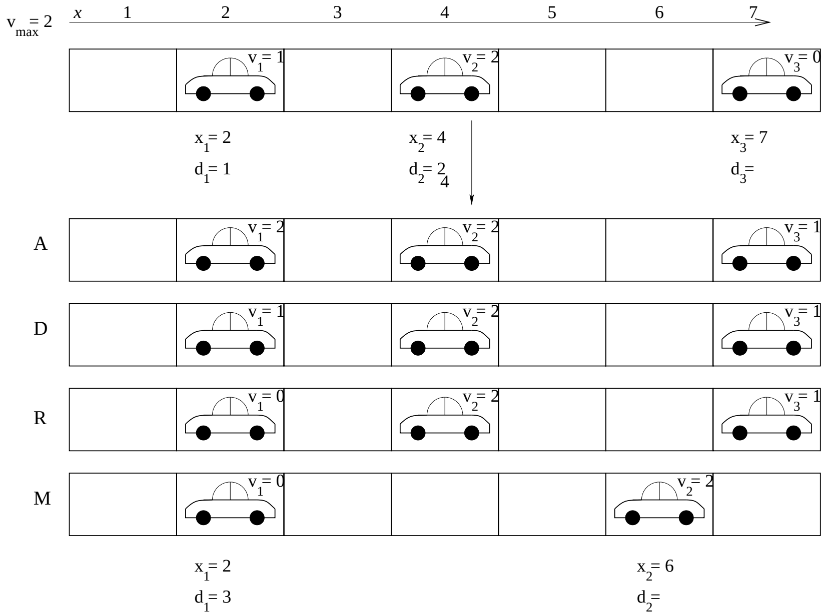


FIGURE 8. An example of evolution for the Nagel-Schreckenberg model

Another remarkable feature of the bijection $\bar{T}$ is that it can be interpreted as moving black particles to the right in the top row and *in the opposite direction in the bottom row.* Indeed, as can be checked on Figure 5, 6, and 7, each time a green region is moved to the right, a black particle is put on its left, so that the global move can be reinterpreted as the displacement of some black particles to the left. The reader is referred to [4] where a similar interpretation in terms of circulating particles is developed for the sequential case.

3.4. **The bijection and the parameter.** In order to conclude the proof of Proposition 1, it remains to check that for any well colored configuration (ω, A, R, G) we have

$$q(\bar{T}(\omega, A, R, G)) = q(\omega, A, R, G).$$

In order to do so, it is sufficient to prove that the parameter q is left unchanged by the application ψ. This can easily be checked on Figures 5, 6 and 7, since all possibilities have been explicitly listed. (The only difficulty is not to forget to count the contribution of all walls of A, even the one that does not belong to $M(\omega)$ when it exists.)

4. Conclusion

The interest in the parallel TASEP originates in the Nagel-Schreckenberg automaton, which is a landmark of highway traffic flow modeling.

The Nagel-Schreckenberg automaton. A configuration of this Markov chain consists of a row of n cells and $n + 1$ walls containing some cars. Each cell can be occupied by a car and cars are numbered from left to right. The jth car is characterized at time t by its position x_j and its velocity $v_j(t) \in \{0, 1, \ldots, v_{\max}\}$, where $v_{\max}$ is the maximal velocity chosen for the system. Moreover, let us denote by $d_j(t)$ the distance, i.e. the number of cells between the jth car and the $(j + 1)$th car.

- At time $t = 0$, the system is in a configuration $NS(0)$ (possibly chosen at random).
- From time t to $t+1$, the system evolves from the configuration $NS(t)$ to the configuration $NS(t+1)$ by applying the following successive transformations to all cars in parallel (see Figure 8):
  - **A Acceleration.** If the jth car is not at the maximal velocity then its velocity increases by one, i.e. $v'_j(t) = \min(v_j(t) + 1, v_{\max})$.
  - **D Safety deceleration.** If the distance $d_j(t)$ to the next car is less than its velocity $v'_j(t)$ then the latter decreases to $d_j(t)$, i.e. $v''_j(t) = \min(d_j(t), v'_j(t))$.

R Random deceleration. The jth car can decelerate by one with probability q if $v''_j(t)$ is not zero:

$$v_j(t+1) = \begin{cases} \max(v''_j(t)-1, 0) & \text{with probability } q, \\ v''_j(t) & \text{otherwise.} \end{cases}$$

M Movement. The jth car moves $v_j(t+1)$ cells to the right, i.e. $x_j(t+1) = x_j(t) + v_j(t+1)$.

The resulting $(x_j(t+1), v_j(t+1))_j$ defines the new configuration $NS(t+1)$, and in particular the new distances $d_j(t+1)$.

Although the Nagel-Schreckenberg automaton is a Markov chain with very simple rules, it appears to be difficult to study from a mathematical point of view. In particular, its stationary distribution is only known for the particular case $v_{\max} = 1$: indeed if $v_{\max} = 1$ then the Nagel-Schreckenberg automaton corresponds to the parallel TASEP, where cars are black particles, and where $q = 1 - p$. Indeed, when $v_{\max} = 1$, the possible velocities for each car are 0 or 1. Therefore, after step **A** all the velocities are equal to 1. At step **D** only cars with $d_j = 0$ (i.e. immediately followed by another car) decrease their velocity to 0. At step **R** a car j such that $v_j = 1$ keeps its non zero velocity with probability equal to $1 - q = p$. At step **M** only cars with $d_j \neq 0$ and $v_j = 1$ move of one cell to the right. This is equivalent to saying that from time t to time $t+1$ cars that have a free cell on their right can move in there with probability p. This corresponds exactly to the parallel TASEP.

A challenging open problem is of course to compute the stationary distribution of the Nagel-Schreckenberg automaton for larger velocities.

References

[1] O. Angel. The stationary measure of a 2-type totally asymmetric exclusion process. arXiv:math.PR/0501005, 2005.

[2] B. Derrida, E. Domany, and D. Mukamel. An exact solution of a one dimensional asymmetric exclusion model with open boundaries. *J. Stat. Phys.*, 69:667–687, 1992.

[3] B. Derrida, M.R. Evans, V. Hakim, and V. Pasquier. Exact solution of a one-dimensional asymmetric exclusion model using a matrix formulation. *J. Phys. A: Math.*, 26:1493–, 1993.

[4] E. Duchi and G. Schaeffer. A combinatorial approach to jumping particles. *Journal of Combinatorial Theory, series A*, 110(1):1–29, 2005.

[5] M. R. Evans, N. Rajewsky, and E. R. Speer. Exact solution of a cellular automaton for traffic. *Journal of Statistical Physics*, 95(1–2):45–96, 1999.

[6] O. Häggström. *Finite Markov Chains and Algorithmic Applications.* Cambridge University Press, 2002.

[7] A. Rechnitzer S. Corteel, R. Brak and J. Essam. A combinatorial derivation of the pasep algebra. *FPSAC'2005, these proceedings.*

[8] A. Schadschneider. Traffic flow: A statistical physics point of view. *Physica A*, 313:153, 2002.

Enrica Duchi, LIAFA, Université Paris 7- Case 7014, 2, place Jussieu, F-75251 Paris Cedex 05
E-mail address: Enrica.Duchi@liafa.jussieu.fr

Gilles Schaeffer, LIX, CNRS, Ecole Polytechnique, F-91128 Palaiseau Cedex
E-mail address: Gilles.Schaeffer@lix.polytechnique.fr

Free quasi-symmetric functions, product actions and quantum field theory of partitions.

Gérard H. E. Duchamp[†], **Jean-Gabriel Luque**[‡], **K A Penson**[⋆], **Christophe Tollu**[†],
([†]) LIPN, UMR CNRS 7030
Institut Galilée - Université Paris-Nord
99, avenue Jean-Baptiste Clément
93430 Villetaneuse, France
([‡]) IGM, Laboratoire d'informatique UMR 8049 IGM-LabInfo
77454 Marne-la-Vallée Cedex 2, France.
⋆ LPTL, Laboratoire de Physique Théorique des Liquides, CNRS UMR 7600
Université Pierre et Marie Curie,
Tour 16, 5ième étage, 4, place Jussieu, F 75252 Paris Cedex 05, France
{ghed@lipn.univ-paris13.fr, luque@univ-mlv.fr,
penson@lptl.jussieu.fr, ct@lipn.univ-paris13.fr}

April 22, 2005

Abstract: We examine two associative products over the ring of symmetric functions related to the intransitive and Cartesian products of permutation groups. As an application, we give an enumeration of some Feynman type diagrams arising in Bender's QFT of partitions. We end by exploring possibilities to construct noncommutative analogues.

Résumé: Nous étudions deux lois produits associatives sur les fonctions symétriques correspondant aux produits intransitif et cartsien des groupes de permutations. Nous donnons comme application l'énumération de certains diagrammes de Feynman apparaissant dans la QFT des partitions de Bender. Enfin, nous donnons quelques pistes possibles pour construire des analogues non-commutatifs.

1 Introduction

In a relatively recent paper, Bender, Brody and Meister introduce a special Field Theory described by

$$G(z) = \left(e^{(\sum_{n\geq 1} L_n \frac{z^n}{n!}\frac{\partial}{\partial x})}\right)\left(e^{(\sum_{m\geq 1} V_m \frac{x^m}{m!})}\right)\Big|_{x=0} \tag{1}$$

in order to prove that any sequence of numbers $\{a_n\}$ can be generated by a suitable set of rules applied to some type of Feynman diagrams [1, 2]. These diagrams actually are bicoloured multigraphs with no isolated vertex.

Expanding one factor of (1), we can observe surprising links between: the normal ordering problem (for bosons), the parametric Stieltjes moment problem and the convolution of kernels, substitution matrices (such as generalised Stirling matrices) and one-parameter groups of analytic substitutions [8, 9, 15].

The aim of this paper is to make explicit miscellaneous connections between noncommutative symmetric functions (here **MQSym**, **FQSym** [6]) and the Feynman diagrams arising in the expansion of formula (1) used in combinatorial physics [15].

The structure of the contribution is the following. In Section 2, we define two associative products in $\mathfrak{S} = \bigsqcup \mathfrak{S}_n$ related to the Intransitive and Cartesian products of permutation groups. These products induce a structure of 2-associative algebra over the symmetric functions. The properties of this algebra are investigated in Section 3. At the end of this section, we give, as an application, an inductive formula for computing generating series of Bender's Feynman diagrams. Noncommutative analogues are proposed in Section 4.

2 Actions of a direct product of permutation groups

2.1 Direct product actions

The actions of the direct product of two permutation groups (in particular, the structure of the cycles) give rise to interesting properties related to the enumeration of unlabelled objets [14]. We open this section with the definition of two actions (namely, Intransitive and Cartesian). For greater detail about these constructions (or for constructions involving the wreath product) the reader can refer to [4].

Consider two pairs (G_1, X_1) and (G_2, X_2), where each G_i is a permutation group acting on X_i. The *intransitive action* of $G_1 \times G_2$ on $X_1 \sqcup X_2$ (here $\sqcup$ means disjoint union) is defined by the rule

$$(\sigma_1, \sigma_2)x = \begin{cases} \sigma_1 x & \text{if } x \in X_1 \\ \sigma_2 x & \text{if } x \in X_2 \end{cases}. \tag{2}$$

This action will be denoted by $(G_1, X_1) \nvrightarrow (G_2, X_2) := (G_1 \times G_2, X_1 \sqcup X_2)$.

The *Cartesian action* of $G_1 \times G_2$ on $X_1 \times X_2$ is defined by

$$(\sigma_1, \sigma_2)(x_1, x_2) = (\sigma_1 x_1, \sigma_2 x_2). \tag{3}$$

This action will be denoted by $(G_1, X_1) \mathbin{\times\!\!\!\!\searrow} (G_2, X_2) := (G_1 \times G_2, X_1 \times X_2)$. Note that neither of the two laws just defined is commutative. A natural question to ask is whether such a structure enjoys some algebraic properties. For example, is the $\mathbin{\times\!\!\!\!\searrow}$ law distributive over $\nvrightarrow$?

In other words, what is the meaning of

$$(G_1, X_1) \mathbin{\times\!\!\!\!\searrow} ((G_2, X_2) \nvrightarrow (G_3, X_3)) = (G_1 \times G_2 \times G_3, X_1 \times (X_2 \sqcup X_3))$$

and

$$((G_1, X_1) \mathbin{\times\!\!\!\!\searrow} (G_2, X_2)) \nvrightarrow ((G_1, X_1) \mathbin{\times\!\!\!\!\searrow} (G_3, X_3)) = (G_1 \times G_2 \times G_1 \times G_3, (X_1 \times X_2) \sqcup (X_1 \times X_3)).$$

The groups $G_1 \times G_2 \times G_1 \times G_3$ and $G_1 \times G_2 \times G_3$ are not isomorphic, so distributivity does not hold, although the set-theoretical Cartesian product is distributive over disjoint union. However an examination of the structure of the cycles (see [4] for the general construction or section 2.2 for a particular case) shows that the cycles are the same. More precisely, a cycle can appear with different multiplicities according to which group is acting, but if we focus on the set of the cycles, the two structures are similar.

Now, let us give a construction which takes such a phenomenon into account.

2.2 Explicit realization

We will denote by $\circ_N$ the natural action of $\mathfrak{S}_n$ on $\{0, \dots, n-1\}$. Let $\mathfrak{S}_n$ and $\mathfrak{S}_m$ be two symmetric groups, we note by $\circ_I$ the intransitive action of $\mathfrak{S}_n \times \mathfrak{S}_m$ on $\{0, \cdots, n+m-1\}$ and by $\circ_C$ the *Cartesian action* of $\mathfrak{S}_n \times \mathfrak{S}_m$ on $\{0, \dots, nm-1\}$. More precisely,

$$(\sigma_1, \sigma_2) \circ_I i = \begin{cases} \sigma_1 \circ_N i & \text{if } 0 \le i \le n-1 \\ \sigma_2 \circ_N (i-n) + n & \text{if } n \le i \le n+m-1 \end{cases}. \tag{4}$$

for $0 \le i \le n+m-1$, and

$$(\sigma_1, \sigma_2) \circ_C (j + nk) = (\sigma_1 \circ_N j) + n(\sigma_2 \circ_N k) \tag{5}$$

for $0 \le j \le n-1$ and $0 \le k \le m-1$.

The *intransitive product* is the map $\nvrightarrow : \mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{n+m}$ defined by

$$\sigma_1 \nvrightarrow \sigma_2 = \sigma_1 \sigma_2[n] \tag{6}$$

where $\sigma_2[n]$ denotes σ_2 composed with the shifted substitution $i \to i+n$ (here permutations are considered as words and $\nvrightarrow$ is nothing else but shifted concatenation).

Example 2.1 Let $\sigma_1 = 1320 \in \mathfrak{S}_4$ and $\sigma_2 = 534120 \in \mathfrak{S}_6$. Here, we denote a permutation of $\mathfrak{S}_n$ by the word whose ith letter is the image of i under the natural action on $\{0, \ldots, n-1\}$). With this notation, we obtain

$$\sigma_1 \nrightarrow \sigma_2 = 1320978564$$

and

$$\sigma_2 \nrightarrow \sigma_1 = 5341207986$$

Clearly, it turns out that $\nrightarrow$ is not commutative.

The following proposition shows that the natural action of $\mathfrak{S}_{n+m}$ coincides with the intransitive action of $\mathfrak{S}_n \times \mathfrak{S}_m$.

Proposition 2.2 $(\sigma_1 \nrightarrow \sigma_2) \circ_N i = (\sigma_1, \sigma_2) \circ_I i$. □

Let us introduce a similar construction for the Cartesian action: we define a map $\times\!\!\!\searrow$: $\mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{nm}$ by

$$\sigma_1 \times\!\!\!\searrow \sigma_2 = \prod_{i,j} c_i \times\!\!\!\searrow c'_j \tag{7}$$

where $\sigma_1 = c_1 \cdots c_k$ (resp. $\sigma_2 = c'_1 \cdots c'_{k'}$) is the decomposition of σ_1 (resp. σ_2) into a product of cycles and

$$c \times\!\!\!\searrow c' = \prod_{s=0}^{l \wedge l'-1} (\phi(s,0), \phi(s+1,1) \cdots, \phi(s+l \vee l'-1, l \vee l'-1)), \tag{8}$$

where $\wedge$ denotes the gcd, $\vee$ denotes the lcm, $c = (i_0, \cdots, i_{l-1})$, $c' = (j_0, \cdots, j_{l'-1})$ are two cycles and $\phi(k,k') = i_{k \bmod l} + nj_{k' \bmod l'}$. Just like the Intransitive action, the Cartesian action coincides with the natural action.

Proposition 2.3 $(\sigma_1 \times\!\!\!\searrow \sigma_2) \circ_N i = (\sigma_1, \sigma_2) \circ_C i$.

Proof — From (7), it suffices to prove the property only when $\sigma_1 = c$ and $\sigma_2 = c'$ are two cycles. But as (8) is equivalent to

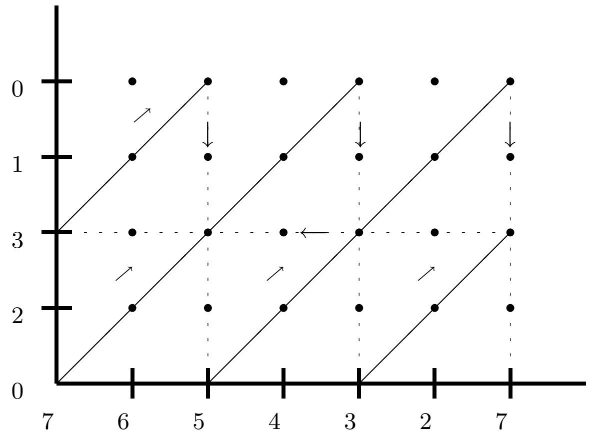
$$\begin{aligned} c \times\!\!\!\searrow c' = & \prod_{s=0}^{l \wedge l'-1} (i_s + nj_0, (c,c') \circ_C (i_s + nj_0), \ldots, (c^{l \vee l'-1}, c'^{l \vee l'-1}) \circ_C (i_s + nj_0)) \\ = & \prod_{s=0}^{l \wedge l'-1} (i_s + nj_0, c \circ_C i_s + nc' \circ_N j_0, \ldots, c^{l \vee l'-1} \circ_N i_s + nc'^{l \vee l'-1} \circ_N j_0)), \end{aligned}$$

which completes the proof. □

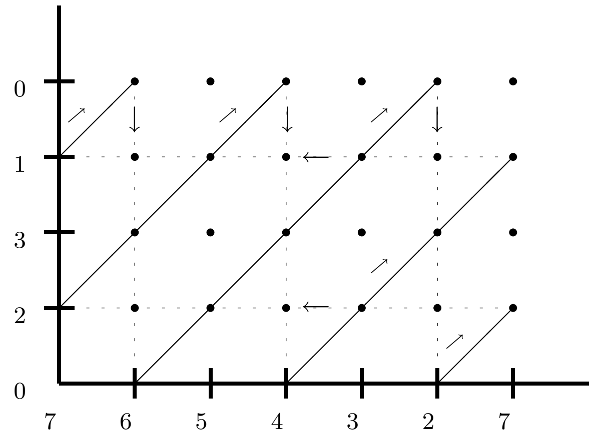
Example 2.4 Consider the two permutations $\sigma_1 = 2031 \in \mathfrak{S}_4$ and $\sigma_2 = 01723456 \in \mathfrak{S}_8$. The permutation σ_1 consists of a unique cycle $c_1 = (0,2,3,1)$ and $\sigma_2 = c'_1 c'_2 c'_3$ is the product of the three cycles $c'_1 = (0)$, $c'_2 = (1)$ and $c'_3 = (7,6,5,4,3,2)$. Hence, the permutation $\sigma_1 \times\!\!\!\searrow \sigma_2$ is the product of four cycles given by

1. $c_1 \times\!\!\!\searrow c'_1 = (0,2,3,1)$
2. $c_1 \times\!\!\!\searrow c'_2 = (4,6,7,5)$
3. $c_1 \times\!\!\!\searrow c'_3 = (28,26,23,17,12,10,31,25,20,18,15,9)(30,27,21,16,14,11,29,24,22,19,13,8)$.

To illustrate proposition 2.3, it suffices to draw the cycles in the Cartesian product $\{0, \ldots, n-1\} \times \{0, \ldots, m-1\}$ whose elements are re- labelled $(i,j) \to i + nj$. For example, the two cycles appearing in $c_1 \times\!\!\!\searrow c'_3$ give the following partition of $\{0,1,2,3\} \times \{2,3,4,5,6,7\}$.



(28, 26, 23, 17, 12, 10, 31, 25, 20, 18, 15, 9)



(30, 27, 21, 16, 14, 11, 29, 24, 22, 19, 13, 8)

On the other hand, the permutation $\sigma_2 \mathbin{\times\!\!\!\!\searrow} \sigma_1$ is the product of the four cycles

1. $c'_1 \mathbin{\times\!\!\!\!\searrow} c_1 = (0, 16, 24, 8)$
2. $c'_2 \mathbin{\times\!\!\!\!\searrow} c_1 = (1, 17, 25, 9)$
3. $c'_3 \mathbin{\times\!\!\!\!\searrow} c_1 = (7, 22, 29, 12, 3, 18, 31, 14, 5, 20, 27, 10)(6, 21, 28, 11, 2, 23, 30, 13, 4, 19, 26, 15)$

Clearly, $\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2 \neq \sigma_2 \mathbin{\times\!\!\!\!\searrow} \sigma_1$: the law $\mathbin{\times\!\!\!\!\searrow}$ is not commutative.

2.3 Algebraic structure

The advantage of the new structures over the ones defined in section 2.1 consists in the omission of the operations over the groups. Hence, algebraic properties come to light quite naturally. First, the two laws are associative.

Proposition 2.5 *Associativity*
Let $\sigma_1 \in \mathfrak{S}_n$, $\sigma_2 \in \mathfrak{S}_m$ and $\sigma_3 \in \mathfrak{S}_p$ be 3 permutations

1. $\sigma_1 \nrightarrow (\sigma_2 \nrightarrow \sigma_3) = (\sigma_1 \nrightarrow \sigma_2) \nrightarrow \sigma_3$
2. $\sigma_1 \mathbin{\times\!\!\!\!\searrow} (\sigma_2 \mathbin{\times\!\!\!\!\searrow} \sigma_3) = (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \mathbin{\times\!\!\!\!\searrow} \sigma_3$

Proof — 1) Set $\eta_1 = \sigma_1 \nrightarrow (\sigma_2 \nrightarrow \sigma_3)$ and $\eta_2 = (\sigma_1 \nrightarrow \sigma_2) \nrightarrow \sigma_3$. One has

$$\eta_1 \circ_N i = \begin{cases} \sigma_1 \circ_N i & \text{if } 0 \leq i \leq n-1 \\ \sigma_2 \circ_N (i-n) + n & \text{if } n \leq i \leq m+n-1 \\ \sigma_3 \circ_N (i-n-m) + n + m & \text{if } n+m \leq i \leq n+m+p-1 \end{cases}$$

for each $0 \leq i \leq n+m-1$, and the same holds for $\eta_2 \circ_N i$. It follows that $\eta_1 = \eta_2$.
2) The strategy is the same. First, we set $\eta_1 = \sigma_1 \mathbin{\times\!\!\!\!\searrow} (\sigma_2 \mathbin{\times\!\!\!\!\searrow} \sigma_3)$ and $\eta_2 = (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \mathbin{\times\!\!\!\!\searrow} \sigma_3$. The action of η_1 can be computed as follows

$$\eta_1 \circ_N (i + ni') = \sigma_1 \circ_N i + n(\sigma_2 \mathbin{\times\!\!\!\!\searrow} \sigma_3) \circ_N i' = \sigma_1 \circ_N i + n\sigma_2 \circ_N j + nm\sigma_3 \circ_N k$$

where $0 \leq i \leq n-1$, $0 \leq i' \leq mp-1$, $0 \leq j \leq m-1$ and $0 \leq k \leq p-1$.
On the other hand, the action of η_2 is

$$\eta_2 \circ_N (k' + nmk) = (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \circ_N k' + nm\sigma_3 \circ_N k = \sigma_1 \circ_N i + n\sigma_2 \circ_N j + nm\sigma_3 \circ_N k$$

where $0 \leq i \leq n-1$, $0 \leq j \leq m-1$, $0 \leq k \leq p-1$ and $0 \leq k' \leq nm-1$. Hence, $\eta_1 \circ_N i = \eta_2 \circ_N i$ for $0 \leq i \leq nmp-1$ and $\eta_1 = \eta_2$. □

From example 2.1 and 2.4, neither $\nrightarrow$ nor $\mathbin{\times\!\!\!\!\searrow}$ is commutative. But, one has the property of left distributivity.

Proposition 2.6 *Semi-distributivity*
Let $\sigma_1 \in \mathfrak{S}_n$, $\sigma_2 \in \mathfrak{S}_m$ and $\sigma_3 \in \mathfrak{S}_p$ be three permutations

$$\sigma_1 \mathbin{\times\!\!\!\!\searrow} (\sigma_2 \nrightarrow \sigma_3) = (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \nrightarrow (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_3)$$

Proof — We use the same method as in the proof of proposition 2.5. First, let us define $\eta_1 = \sigma_1 \mathbin{\times\!\!\!\!\searrow} (\sigma_2 \nrightarrow \sigma_2)$ and $\eta_2 = (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \nrightarrow (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_3)$. The action of η_1 is

$$\eta_1 \circ_N (i+nj) = \eta_1 \circ_N i + n(\sigma_2 \nrightarrow \sigma_3) \circ_N j = \begin{cases} \sigma_1 \circ_N i + n\sigma_2 \circ_N j & \text{if } 0 \leq j \leq m-1 \\ \sigma_1 \circ_N i + n\sigma_3 \circ_N (j-m) + m & \text{if } m \leq j \leq p+m-1 \end{cases} \tag{9}$$

where $0 \leq i \leq n-1$ and $0 \leq j \leq m+p-1$.
On the other hand, one has

$$\eta_2 \circ_N k = \begin{cases} (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \circ_N k & \text{if } 0 \leq k \leq nm-1 \\ (\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_3) \circ_N (k-nm) + nm & \text{if } nm \leq k \leq n(m+p)-1 \end{cases}. \tag{10}$$

If $0 \leq k \leq mn-1$, we set $k = i + nj$ where $0 \leq i \leq n-1$ and $0 \leq j \leq m-1$. Hence,

$$(\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_2) \circ_N k = \sigma_1 \circ_N i + n\sigma_2 \circ_N j. \tag{11}$$

Similarly, if $nm \leq k \leq n(m+p)-1$, we set $(k-nm) = i + nj$ where $0 \leq i \leq n-1$ and $0 \leq j \leq p-1$. Hence,

$$(\sigma_1 \mathbin{\times\!\!\!\!\searrow} \sigma_3) \circ_N (k-nm) + nm = \sigma_1 \circ_N i + n(\sigma_3 \circ_N (j-m) + m). \tag{12}$$

Substituting (11) and (12) in (10), one recovers the right hand side of (9). It follows immediately that $\eta_1 = \eta_2$. □

The two laws can be extended by linearity to the graded vector space $\bigoplus_{n \geq 0} \mathbb{Q}[\mathfrak{S}_n]$ and endow this space with a structure of 2-associative algebra. In the next section, we construct a product $\star$ in Sym (the algebra of symmetric functions) defined on the power sums and appearing when one examines the cycle index polynomial of a Cartesian product. This product is the image of $\mathbin{\times\!\!\!\!\searrow}$ under a particular morphism. We will prove that this last property implies the associativity and the distributivity of $\star$ over $\times$ (the natural product in Sym) and $+$.

3 Cycle index algebra

3.1 Cartesian product in Sym

We first construct a 2-associative morphism $\bigoplus_{n \geq 0} \mathbb{Q}[\mathfrak{S}_n] \mapsto Sym$ (a 2-associative algebra is just a vector space equipped with 2 associative laws [11]).
The arrow maps a permutation $\sigma \in \mathfrak{S}_n$ to a product of power sums. For $j \geq 1$, let $c_j(\sigma)$ be the number of cycles in σ of length j and set

$$\mathfrak{Z}(\sigma) = \prod_{j=0}^{\infty} \psi_j^{c_j(\sigma)} \tag{13}$$

where ψ_i denotes the ith power sum symmetric function. We claim that $\mathfrak{Z}$ is a morphism mapping $\nrightarrow$ to $\times$ (the usual product in Sym) and that $\mathbin{\times\!\!\!\!\searrow}$ is compatible with $\mathfrak{Z}$ to the extent that there exists an associative law on Sym such that $\mathfrak{Z}$ is also a morphism mapping it to $\mathbin{\times\!\!\!\!\searrow}$. This second law is given on the power sums basis by

$$\prod_{1 \leq i \leq \infty} \psi_i^{\alpha_i} \star \prod_{1 \leq j \leq \infty} \psi_j^{\beta_j} = \prod_{1 \leq i,j \leq \infty} \psi_{i \vee j}^{\alpha_i \beta_j (i \wedge j)} \tag{14}$$

(the sequences $(\alpha_i)_{i \geq 1}$, $(\beta_j)_{j \geq 1}$ have finite support). It is straightforward to check that

Proposition 3.1 *i) The mapping* $\mathfrak{Z} : \bigoplus_{n\geq 0}\mathbb{Q}[\mathfrak{S}_n] \mapsto Sym$ *is a morphism of 2-associative algebras sending the two laws* $\dotplus$; $\times\!\!\!\!\searrow$ *respectively to* $\times$; $\star$ *(recall that* $\times$ *denotes the usual product of Sym).*
More precisely, for $\sigma,\ \tau \in \sqcup_{n\geq 0}\mathfrak{S}_n = \mathfrak{S}$ *one has*

$$\mathfrak{Z}(\sigma \dotplus \tau) = \mathfrak{Z}(\sigma)\mathfrak{Z}(\tau)\ ;\ \mathfrak{Z}(\sigma \times\!\!\!\!\searrow \tau) = \mathfrak{Z}(\sigma) \star \mathfrak{Z}(\tau) \tag{15}$$

ii) The law $\star$ *is associative, commutative and distributive over* $\times$.

Proof — i) For the first relation of (15), one just notices that $c_j(\sigma \dotplus \tau) = c_j(\sigma) + c_j(\tau)$. For the second relation, one observes that the Cartesian product of a i-cycle and a j-cycle produces $i \wedge j$ cycles of length $i \vee j$. Thus, one has $c_r(\sigma \times\!\!\!\!\searrow \tau) = \sum_{p\vee q=r}(p\wedge q)c_p(\sigma)c_q(\tau)$, whence (15).
ii) When $\sigma \in \mathfrak{S}_n$ is a cycle of maximum length, one has $\mathfrak{Z}(\sigma) = \psi_n$, hence the image of $\mathfrak{Z}$ contains also all the products of power sums and we get $Im(\mathfrak{Z}) = Sym$. Then, by proposition 3.1(i), $\star$ is distributive on the left over $\times$. Complete distributivity follows from commutativity of $\star$, which straightforwardly follows from the definition. □

The following structural result goes into particulars of the distributivity of $\star$ over $\times$.

Proposition 3.2 *Let* P *be the set of products of power sums (i.e.* $P = \{\prod_{i=1}^{\infty}\psi_i^{\alpha_i}\}_{(\alpha_i)_{i\geq 1}\in\mathbb{N}^{\mathbb{N}^*}}$*). Then* P *is closed by* $\times$ *and* $\star$ *and more precisely* $(P,\times,\star)$ *is isomorphic to a subsemiring of the* $\mathbb{Z}$*-algebra* $\mathbb{Z}[\mathbb{N}^{\mathfrak{p}}]$ *of the monoid* $(\mathbb{N}^{\mathfrak{p}}, sup)$ *(where* $\mathfrak{p}$ *stands for the set of prime numbers).*

Proof — The fact that P is closed by $\times$ and $\star$ follows from the definition and (14). Now P contains the two units (1 and ψ_1), therefore (as a consequence of the properties established for the laws $\times$, $\star$) it is a semiring. For every $p \in \mathfrak{p}$ and $n \in \mathbb{N}^*$, let $\nu_p(n)$ be the exponent of p in the decomposition of n in prime factors ($n = \prod_{p\in\mathfrak{p}} p^{\nu_p(n)}$). Define an arrow $\phi : P \to \mathbb{Z}[(\mathbb{N}^{\mathfrak{p}}]$ by

$$\phi(\prod_{1\leq i\leq\infty}\psi_i^{\alpha_i}) = \sum_{1\leq i\leq\infty} i\alpha_i(p\mapsto\nu_p(i)). \tag{16}$$

As $\phi(m_1m_2) = \phi(m_1) + \phi(m_2)$ by construction (16), it suffices to prove that $\phi(\psi_i \star \psi_j) = \phi(\psi_i) \times_s \phi(\psi_j)$ where $\times_s$ stands for the product in $\mathbb{Z}[(\mathbb{N}^{(\mathfrak{p})}, sup)]$. But

$$\phi(\psi_i \star \psi_j) = \phi(\psi_{i\vee j}^{i\wedge j}) = (i\wedge j)\phi(\psi_{i\vee j}) = (i\wedge j)(i\vee j)(p\mapsto\nu_p(i\vee j)) =$$
$$(i\wedge j)(i\vee j)(p\mapsto sup(\nu_p(i),\nu_p(j))) = ij(p\mapsto sup(\nu_p(i),\nu_p(j))) = \phi(\psi_i)\times_s\phi(\psi_j).$$

The arrow being clearly into the claim is proved. □

3.2 Cycle index polynomial

Let $\mathfrak{S} = \bigsqcup_{n\geq 0}\mathfrak{S}_n$ be the disjoint union of all the symmetric groups and $\mathfrak{S}_{sg} = \bigcup_{n\geq 0}(\mathfrak{S}_n)_{sg}$ be the set of all the subgroups of all symmetric groups (*i.e.* the set of all permutation groups over some interval $[1..n]$). For simplicity, we identify a permutation group $G \in (\mathfrak{S}_n)_{sg}$ with its action $(G, \{0,\ldots,n-1\})$ (see section 2.1). Laws $\dotplus$ and $\times\!\!\!\!\searrow$ can be defined over $\mathfrak{S}_{sg}$ by

$$G_1 \dotplus G_2 := (G_1\times G_2, \{0,\ldots,n+m-1\}) \tag{17}$$

where G_1 acts on $\{0,\ldots,n-1\}$ and G_2 acts on $\{n,\ldots,n+m-1\}$ and

$$G_1 \times\!\!\!\!\searrow G_2 := (G_1\times G_2, \{0,\ldots,nm-1\}) \tag{18}$$

where the action on $\{0,\ldots,nm-1\}$ is given by $(\sigma_1,\sigma_2)k = \phi^{-1}((\sigma_1,\sigma_2)\phi(k))$, the map ϕ being the bijection $\phi : \{0,\ldots,nm-1\} \to \{0,\ldots,n-1\}\times\{0,\ldots,m-1\}$ defined by $\phi(i+nj) = (i,j)$ if $0\leq i\leq n-1$ and $0\leq j\leq m-1$ and $(\sigma_1,\sigma_2)(i,j) = (\sigma_1 i,\sigma_2 j)$. Note that both $\dotplus$ and $\times\!\!\!\!\searrow$ are

associative but $\times\!\!\!\searrow$ is not distributive over $\dotplus$.
Let $Z : \mathfrak{S}_{sg} \to Sym$ be defined by

$$Z(G) = \mathfrak{Z}\left(\frac{1}{|G|}\sum_{\sigma\in G}\sigma\right). \tag{19}$$

Polyà's cycle index polynomial of G is defined to be $Z(G)$.

Example 3.3 1. The cycle index of the symmetric group $\mathfrak{S}_n$ is $Z(\mathfrak{S}_n) = h_n$.

2. The cycle index of the alternating group A_n is $Z(A_n) = h_n + e_n$.

Here h_n (resp. e_n) denotes a complete (resp. elementary) symmetric function. These examples appear as exercices in [12] (ex.9 p 29).

Since $\mathfrak{Z}$ is a morphism of 2-associative algebra, one recovers the classical relations (see [4])

$$Z(G_1 \dotplus G_2) = Z(G_1)Z(G_2) \tag{20}$$

$$Z(G_1 \times\!\!\!\searrow G_2) = Z(G_1) \star Z(G_2) \tag{21}$$

Example 3.4 1. The cycle index polynomial of the Intransitive product of two symmetric groups $\mathfrak{S}_n$ and $\mathfrak{S}_m$ is

$$Z(\mathfrak{S}_n \dotplus \mathfrak{S}_m) = h_n h_m.$$

2. The cycle index polynomial of the Cartesian product of two symmetric groups $\mathfrak{S}_n$ and $\mathfrak{S}_m$ is

$$Z(\mathfrak{S}_n \times\!\!\!\searrow \mathfrak{S}_m) = h_n \star h_m = \sum_{\substack{|\lambda|=n,\\ |\rho|=m}} m_\lambda \star m_\rho = \sum_{\substack{|\lambda|=n,\\ |\rho|=m}} \frac{1}{z_\lambda z_\rho}\prod_{i,j}\psi_{\lambda_i\vee\rho_j}^{\lambda_i\wedge\rho_j},$$

where m_λ denotes a monomial symmetric function and $z_\lambda = \prod i^{n_i} n_i!$ if n_i is the number of parts of λ equal to i.

3.3 Enumeration of a type of Feynman diagrams related to the Quantum Field Theory of partitions

The cycle index polynomials are classic tools used in combination with Polyà's theorem, for the enumeration of unlabelled objects. Let us recall the general process. Consider a permutation group G acting on a finite set $X = \{x_1, \cdots, x_n\}$. Let $L = \{l_0, \dots, l_p, \dots\}$ (possibly infinite) be another set, and $f : X \to L$. The *type* $t(f)$ of f is the vector $(i_0, \dots, i_p, \dots)$ where i_k is the number of elements of X whose image by f is l_k. The *shape* $s(f)$ of f is the partition obtained by sorting in the decreasing order $t(f)$ and erasing the zeroes. For example, a function f having the type $t(f) = (0, 1, 0, 9, 1, 2, 0, \dots, 0, \dots)$ has the shape $s(f) = (9, 2, 1, 1)$. The number $d^s_\lambda(G, L)$ of G-classes on L^X with the shape λ is the coefficient of m_λ in the expansion of $Z(G)$ in the basis of monomial symmetric functions:

$$Z(G) = \sum_\lambda d^s_\lambda(G, L)m_\lambda. \tag{22}$$

Now, let us apply this method to enumerate the Feynman diagrams arising in the expansion of formula (1). These diagrams are bicoloured multigraphs (or bicoloured graphs with edges weighted by positive integers) with no isolated vertex. First, we enumerate all bicoloured multigraphs: Such a computation can be found in [10], so here we only sketch the general case. Of course, the following computations (and more general ones) could be carried out within the framework of the theory od species [3]. Let n and m be the numbers of vertices in each of the two parts. We consider the edges as a function e from $\{0, \dots, n-1\} \times \{0, \dots, m-1\}$ to $\mathbb{N}$. The type (resp. the shape) of a graph is the type (resp. the shape) of its edges, *i.e.* $t(e)$ (resp. $s(e)$).

The number $d_\lambda(n,m)$ of graphs with type λ is equal to the number of orbits with type λ, for the action of $\mathfrak{S}_n \times \mathfrak{S}_m$ on $\mathbb{N}^{\{0,\dots,n-1\}\times\{0,\dots,m-1\}}$. Hence, the generating function of the shape is

$$g(n,m) := \sum_\lambda d^s_\lambda(n,m) m_\lambda = Z(\mathfrak{S}_n) \star Z(\mathfrak{S}_m) \tag{23}$$

Specializing the symmetric functions appearing in (23) to the alphabet $\{y_0,\dots,y_k,\dots,\}$, the coefficient $d^t_I(n,m)$ of $\prod y_k^{i_k}$ in the expansion of $g(n,m)$ is equal to the number of graphs with type $I=(i_0,\dots,i_k,\dots)$,

$$g(n,m) = \sum_{I=(i_0,\dots,i_p,\dots)} d^t_I(n,m) \prod_{k=0}^{\infty} y_k^{i_k}. \tag{24}$$

Note that one can enumerate graphs having edges weighted with integers less than or equal to p by specializing to the finite alphabet $\{y_0,\dots,y_p\}$.
Let us define the generating series of the type of our Feynman diagrams

$$F(n,m) := \sum_{I=(i_0,\dots,i_p,\dots)} f^t_I(n,m) \prod_{k=0}^{\infty} y_k^{i_k}, \tag{25}$$

where $f^t_I(n,m)$ denotes the number of Feynman diagrams of type I. Remark that $F(n,m)$ is a symmetric function over the alphabet $\{y_1,\dots,y_p,\dots\}$ but not over $\{y_0,\dots,y_p,\dots\}$.

Example 3.5 Let us give the first examples of generating series for weight in $\{0,1,2\}$.

1. $F(1,1) = y_1 + y_2$
2. $F(2,1) = F(1,2) = y_1^2 + y_1y_2 + y_2^2$
3. $F(2,2) = y_0^2y_1^2 + y_0^2y_2^2 + y_0^2y_1y_2 + y_0y_1^3 + 3y_0y_1^2y_2 + 3y_0y_1y_2^2 + y_0y_2^3 + y_1^4 + y_1^3y_2 + 3y_1^2y_2^2 + y_1y_2^3 + y_2^4$

One can remark that under this specialization,

$$F(2,2) + F(2,1)y_0^2 + F(1,2)y_0^2 + F(1,1)y_0^3 + y_0^4 = 3m_{22} + m_4 + 3m_{211} + m_{31} = g(2,2).$$

The latter equality could be stated in a more general setting.

Theorem 3.6 *One has the following decomposition of the cycle index polynomial.*

$$Z(\mathfrak{S}_n \times \mathfrak{S}_m) = y_0^{nm} + \sum_{(1,1)\le_{lex}(k,p)\le_{lex}(n,m)} F(k,p) y_0^{nm-kp}. \tag{26}$$

Proof — It suffices to remark that a bicoloured multigraph is either a bicoloured multigraph with no isolated vertex or the union of some isolated vertex and a smaller bicoloured multigraph. □

This yields a nice induction formula for the $F(n,m)$'s.

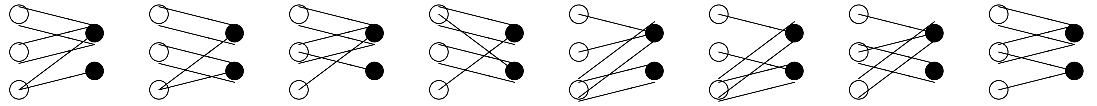
Example 3.7 From theorem 3.6, one has

$$F(3,2) = Z(\mathfrak{S}_3 \times \mathfrak{S}_2) - F(3,1)y_0^3 - F(2,2)y_0^2 - F(2,1)y_0^4 - F(1,2)y_0^4 - F(1,1)y_0^5 - y_0^6.$$

From example 3.5, it suffices to compute $F(3,1) = y_1^3 + y_2^3$ to enumerate Feynman diagrams whose edges are weighted by 0, 1 or 2. After simplification, one obtains

$$\begin{aligned} F(3,2) \;=\; & y_2^6 + y_2^5y_1 + 3y_2^4y_1 + 3y_2^4y_1y_0 + 2y_2^4y_0^2 + 3y_2^3y_1^3 + 6y_2^3y_1^2y_0 + 5y_2^3y_1y_0^2 \\ & + y_2^3y_0^3 + 3y_2^2y_1^4 + 3y_2^2y_1^3y_0 + 8y_2^2y_1^2y_0^2 + 3y_2^2y_1y_0^3 + y_2y_1^5 + 3y_2y_1^4y_0 + 5y_2y_1^3y_0^2 \\ & + 3y_2y_1^2y_0^3 + y_1^6 + y_1^5y_0 + y_1^3y_0^3 + 2y_1^2y_0^4. \end{aligned}$$

For example, there are 8 $(2,2,2)$- Feynman diagrams:



4 Non commutative realizations

4.1 Free quasi-symmetric cycle index algebra

Let $(A, <)$ be an ordered alphabet and $w \in A^*$ a word of length n. One denotes by $Std(w)$, the permutation $\sigma \in \mathfrak{S}_n$ defined by

$$\sigma(i) = (\text{Number of letters} = w[i] \text{ in } w[1..i] + \text{number of letters} < w[i] \text{ in } w) \tag{27}$$

Recall that the algebra **FQSym** is defined by one of its bases, indexed by $\mathfrak{S}$ and defined as follows

$$\mathbf{F}_\sigma = \sum_{Std(w)=\sigma^{-1}} w \in \mathbb{Z}\langle\langle A \rangle\rangle \tag{28}$$

In [6], it is shown that **FQSym** is freely generated by the $\mathbf{F}_\sigma$ where σ runs over the connected permutations (see [5]) (*i.e.* permutations such that $\sigma([1,k]) \neq [1,k]$ for each k). The algebra **FQSym** is spanned by a linear basis, $\{\mathbf{F}^\sigma\}_{\sigma\in\mathfrak{S}}$, whose product implements the Intransitive action $\not\to$:

$$\mathbf{F}^\sigma = \mathbf{F}_{\sigma_1} \cdots \mathbf{F}_{\sigma_n} \tag{29}$$

where $\sigma = \sigma_1 \not\to \cdots \not\to \sigma_n$ is the maximal factorisation of σ in connected permutations. As a consequence of this definition, one has

$$\mathbf{F}^\sigma \mathbf{F}^\tau = \mathbf{F}^{\sigma \not\to \tau}. \tag{30}$$

This naturally induces an isomorphism of algebras

$$\begin{aligned} \underline{\mathfrak{Z}} : \left(\bigoplus_{n\geq 0} \mathbb{Q}[\mathfrak{S}_n], \not\to, + \right) &\to (\mathbf{FQSym}, ., +) \\ \sigma &\mapsto \mathbf{F}^\sigma. \end{aligned} \tag{31}$$

One defines the product $\star$ on **FQSym** by $\mathbf{F}^\sigma \star \mathbf{F}^\tau := \mathbf{F}^{\sigma \times \tau}$. By this way, $\underline{\mathfrak{Z}}$ becomes a morphism of 2-associative algebras. Furthermore, $\star$ is associative, distributive over the sum and semi-distributive over the shifted concatenation.

4.2 Free quasi-symmetric Polyà cycle index polynomial

Let G be a permutation group. The *free quasi-symmetric Polyà cycle index polynomial* of G is its image by $\underline{Z} : \mathfrak{S}_{sg} \to \mathbf{FQSym}$ defined by

$$\underline{Z}(G) := \underline{\mathfrak{Z}} \left(\frac{1}{|G|} \sum_{\sigma\in G} \sigma \right) \mathbf{F}^\sigma. \tag{32}$$

Note 4.1 *There is another basis of* **FQSym** *indexed by permutations, namely* $\{\mathbf{G}^\sigma\}_{\sigma\in\mathfrak{S}}$. *It is obtained by setting* $\mathbf{G}_\sigma = \mathbf{F}_{\sigma^{-1}}$ *and applying the same construction as above (30) to get a basis multiplicative with respect to* $\not\to$, *then*

$$\mathbf{G}^\sigma = \mathbf{G}_{\sigma_1} \cdots \mathbf{G}_{\sigma_n} \tag{33}$$

where $\sigma = \sigma_1 \not\to \cdots \not\to \sigma_n$ *is the maximal factorisation of* σ *into connected permutations. In this case,* σ^{-1} *splits maximally into* $\sigma_1^{-1} \not\to \cdots \not\to \sigma_n^{-1}$, *so one has also* $\mathbf{G}^\sigma = \mathbf{F}^{\sigma^{-1}}$ *and formula (34) can be rewritten*

$$\underline{Z}(G) := \underline{\mathfrak{Z}} \left(\frac{1}{|G|} \sum_{\sigma\in G} \sigma \right) \mathbf{G}^\sigma. \tag{34}$$

The polynomial $\underline{Z}(G)$ has properties similar to that of $Z(G)$, in particular regarding the laws $\not\to$ and $\times$.

Proposition 4.2 *Let $G_1, G_2 \in \mathfrak{S}_{sg}$ be two permutation groups, one has*

1. $\underline{Z}(G_1 \mathbin{\vec{+}} G_2) = \underline{Z}(G_1)\underline{Z}(G_2)$.
2. $\underline{Z}(G_1 \mathbin{\vec{\times}} G_2) = \underline{Z}(G_1) \star \underline{Z}(G_2)$.

Consider the morphism, $z : \mathbf{FQSym} \to Sym$ defined by $z(\mathbf{F}^\sigma) = \mathfrak{Z}(\sigma)$. Note that it is not a morphism of Hopf algebra because $z(\mathbf{F}^{231}) = \psi_3$.
The following diagram is commutative

$$\begin{array}{ccc} \mathfrak{S}_{sg} & \xrightarrow{\underline{Z}} & \mathbf{FQSym} \\ Z\downarrow & {}^{z}\swarrow & \uparrow \underline{\mathfrak{Z}} \\ Sym & \underset{\mathfrak{Z}}{\longleftarrow} & \bigoplus\limits_{n\geq 0}\mathbb{Q}[\mathfrak{S}_n] \end{array} \tag{35}$$

Example 4.3 1. The free quasi-symmetric cycle index of $\mathfrak{S}_n$ is

$$\mathbf{H}_n := \underline{Z}(\mathfrak{S}_n) = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n} \mathbf{F}^\sigma.$$

One can consider it as a free quasi-symmetric analogue of the complete symmetric function h_n: indeed $z(\mathbf{H}_n) = Z(\mathfrak{S}_n) = h_n$.

2. One can define free quasi-symmetric analogues of elementary symmetric functions considering the cycle index polynomial of the alternative groups:

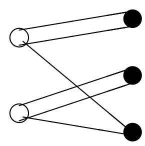
$$\mathbf{E}_n := \underline{Z}(A_n) - \underline{Z}(\mathfrak{S}_n).$$

We get $z(\mathbf{E}_n) = Z(A_n) - Z(\mathfrak{S}_n) = e_n$.

3. The knowledge of analogues of other symmetric functions should be useful to understand the combinatorics of free quasi-symmetric cycle index. In particular, it should be interesting to find free quasi-symmetric functions whose images by z are the monomial symmetric functions.

4.3 Realizations in MQSym

We will call *labelled diagrams* the Feynman diagrams as above but with p white (resp. q black) spots labelled bijectively by $[1..p]$ (resp. by $[1..q]$). When one draws such a diagram, one implicitly assumes that the labelling goes from top to bottom.



Labelled diagram of the matrix $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}$.

Now, to such a $p\times q$ *labelled diagram* we can associate a matrix in $\mathbb{N}^{p\times q}$ and this correspondence is one-to-one. The condition that no vertex be isolated is equivalent to the condition that there be no complete line or column of zeroes, *i.e.* the representative matrix is *packed* [6]. In the same way, the diagrams are in one-to-one correspondence with the classes of packed matrices under the permutations of lines and columns as shown below (the vertical arrows are then one-to-one)

$$\begin{array}{ccc} \text{Packed matrices} & \xrightarrow{Class} & \text{Classes of packed matrices} \\ \big\downarrow & & \big\downarrow \\ \text{Labelled diagrams} & \longrightarrow & \text{Diagrams} \end{array} \tag{36}$$

There is an interesting structure of Hopf algebra (in fact an envelopping algebra) over the diagrams [7] which can be pulled back in a natural way to labelled diagrams.
The correspondence described above allows to construct a new Hopf algebra structure on **MQSym** and a Hopf algebra structure on the space spanned by the classes.

5 Conclusion

Other realizations in Hopf algebras seem feasible. For example, let us consider the Hopf algebras of graphs $GQSym^{110}$ and $GTSym^{110}$ defined in [13]. An interesting mapping from $\bigoplus_{n\geq 0}\mathbb{Q}[\mathfrak{S}_N]$ to $GQSym^{110}$ or $GTSym^{110}$ can be constructed sending each cycle to an equivalent loop.
More precisely, J.-Y. Thibon (personal communication) showed to us how to construct a non-commutative Hopf algebra which is the dual of a quotient of a subalgebra of $GTSym^{110}$. This algebra admits a Hopf morphism to Sym and has two bases indexed by permutations and whose commutative images are respectively proportional to power sums and monomial symmetric functions. This construction gives natural non-commutative analogues of Polyà's cycle index polynomials and will be the subject of a forthcoming paper.

Acknowledgments We are grateful to J.-Y. Thibon for fruitful comments.

References

[1] C.M. Bender, D.C. Brody and B.K. Meister, *Quantum field theory of partitions*, *J.Math. Phys.* **40** (2000), 3239.

[2] C.M. Bender, D.C. Brody, and B.K. Meister, *Combinatorics and field theory*, *Twistor Newsletter* **45** (2000), 36.

[3] F. Bergeron, G. Labelle and P. Leroux, Combinatorial Species and Tree-like Structures, Cambridge University Press (Encyclopedia of Mathematics and its Applications 67), Cambridge, 1997.

[4] P.J. Cameron, D.A. Gewurz, F. Merola, *Product action*, Preprint, `http://www.maths.qmw.ac.uk/~pjc/preprints/product.pdf`

[5] L. Comtet, *Sur les cœfficients de l'inverse de la série formelle* $\sum n!t^n$, C.R. Acad. Sci. Paris, Ser. A **275** (1972), 569-572.

[6] G. Duchamp , F. Hivert , J.-Y. Thibon : *Non commutative functions VI: Free quasi-symmetric functions and related algebras*, International Journal of Algebra and Computation **12** (2000), N 5.

[7] G. H. E. Duchamp, K.A. Penson, A.I. Solomon, A. Horzela and P. Blasiak, *Hopf algebras and Lie group structures inherent in combinatorial field theories*, New symmetries in Mathematical Physics (CIRM - November 22 - 26th 2004). Organizers: R. Kerner (Prof., Univ. Paris 6), J. Kijowski (Prof., Univ. Varsovie)

[8] G. H. E. Duchamp, K.A. Penson, A.I. Solomon, A. Horzela and P. Blasiak, *One-parameter groups and combinatorial physics*, Third International Workshop on Contemporary Problems in Mathematical Physics (COPROMAPH3), Porto-Novo (Benin), November 2003. arXiv : `quant-ph/0401126`

[9] K.A. Penson, P. Blasiak, G. Duchamp, A. Horzela and A.I. Solomon, *Hierarchical Dobiński-type relations via substitution and the moment problem*, J. Phys. A, to appear (2004), arXiv : `quant-ph/0312202`

[10] V. Krishnamurthy, Combinatorics, Theory and Applications, Affilated East-West Press, New Delhi-Madras, 1927.

[11] J.-L. Loday and M. Ronco, *On the structure of cofree Hopf algebras*, preprint arXiv :`math.QA/0405330`

[12] I.G. Macdonald, Symmetric functions and Hall polynomials, second edition, Clarendon Press, Oxford, 1995.

[13] J.-C. Novelli and J.-Y. Thibon, *Algèbres de Hopf de graphes*, C.R. Acad. Sci. Paris Ser. I, to appear Preprint: `http://www-igm.univ-mlv.fr/~jyt/ARTICLES/graphs.ps`.

[14] G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen.*, Acta Math. **68** (1937), 145-254.

[15] A.I. Solomon, P. Blasiak, G. Duchamp, A. Horzela and K.A. Penson, *Combinatorial Physics, Normal Order and Model Feynman Graphs*, Proceedings of the Symposium 'Symmetries in Science XIII', Bregenz, Austria, 2003. Kluwer Publ. (2004) in press.
arXiv : `quant-ph/0310174`

POLYOMINOIDS AND UNIFORM ELECTION

ABDELAAZIZ EL HIBAOUI, NASSER SAHEB-DJAHROMI AND AKKA ZEMMARI*
LABRI, UNIVERSITÉ BORDEAUX I,
351 COURS DE LA LIBÉRATION
33405 TALENCE, FRANCE

ABSTRACT. In this paper, a new structure for polyominoid graph is proposed. This structure is shown to be generated with some rules. A uniform probabilistic election algorithm in polyominoids is developed and studied. Indeed, the election process is considered as a distributed elimination algorithm in a polyominoid, which removes all active vertices one after the other, till there remains one single vertex: the leader.

The elimination algorithm is analyzed as a Markovian random process in continuous time. Our algorithm is totally fair in that all vertices have the same probability of being elected.
Key words: Distributed Algorithms, Election, Fairness.

1. INTRODUCTION

We consider distributed networks of processors [23]. They are presented as a connected graphs where vertices represent processors, and two vertices are connected by an edge if the corresponding processors have a direct communication link. The networks are asynchronous: processors cannot access a global clock and a message sent from a processor to a neighbor arrives within some finite but unpredictable time (asynchronous message passing). Labels are attached to vertices and sometimes to edges. The aim of an election problem is to choose exactly one element in the set of processors. Thus, starting from a configuration where all processors are in the same state, we must obtain a configuration where exactly one processor is in the state "leader" and all other processors are in the state "lost". The leader can be used subsequently to make decisions or to centralize some information. The election problem is well known and many solutions are available [1, 14, 15, 17, 20, 23]. It was first proposed by Le Lann [14].

The networks studied in this paper are anonymous and have a polyominoid topology. A polyominoid combines the tree and polyominos-structure (see Figure 1). Polyominoes have a long history, going back to the beginning of the 20^{th} century, but they were popularized in the present era by Golomb [11, 12] and Gardner [9, 10] in the *Scientific American columns*, "Mathematical Games". They were also studied by mathematicians [2, 3, 5, 6], because they constitute combinatoric objects having interesting properties. They have been the subject of intensive studies by physicists, thanks to their appropriateness for modeling several physical phenomena and are known under the name of animals in statistical physics (see [24] for more details). In computer science, their study has been motivated in different areas such as the VLSI circuit designs (see [16]) and image processing ([4]).

The main motivation behind this study is to introduce a uniform probabilistic distributed election algorithm over polyominoids. This algorithm is *totally fair*, i.e. it gives a same chance of being elected to all vertices of a polyominoid. The algorithm removes vertices of the polyominoid once their random lifetime delay

*e-mail: {hibaoui, saheb, zemmari}@labri.fr.

has been expired (the remaining graph should remain polyominoid). The analysis of the algorithm reveals the surprising fact that, wherever the vertex is placed in the polyominoid, it has the same probability of surviving as the others. The only investigation in this direction, known to the authors, is that of trees [19] .

Our distributed algorithm may be viewed as a randomized extension of a variant of [15], where random delays are introduced.

We consider cellular local computations which allow to modify the state (or label) of a vertex at each step. The new label depends on the previous one and those of its neighbors. The novelty of our approach is the use of random delays for relabeling. These delays are exponential random variables defined independently for active vertices. The parameter of the random variable for a vertex is equal to the attributed value assigned to the vertex. The process of relabeling continues until no more transformation is possible, i.e. a final configuration is reached. In this configuration, there is only one $L-$labeled vertex, considered as elected.

The paper is organized as follows. In Section 2, we introduce the preliminaries and basic notation. Polyominoids are introduced in this section as particular undirected graphs. We provide a set of rules generating the class of all polyominoids. The election algorithm is described in Section 3. The main result in Section 4 is the uniformity of the election on the set of vertices. Due to the space limitation we have to omit simple proofs. For more details, the reader is referred to [13].

2. Preliminaries and Notation

There are many definitions for polyominos and grid-like graphs in the literature, see [21, 18]. Traditionally, a polyomino is the set of cells situated in the interior of an orthogonal polygon drawn on a grid. We define polyominoids as finite graphs whose nodes are points from $\mathbb{Z} \times \mathbb{Z}$, where $\mathbb{Z}$ denotes the set of integers, possibly linked by the neighborhood relationship, defined in the sequel.

Throughout this paper, the vertices are points from $\mathcal{Z} = \mathbb{Z} \times \mathbb{Z}$. We use usual terms such as "up", "down", "right" and "left" on $\mathbb{Z}\times\mathbb{Z}$. The edges are links between pairs of points, i.e. sets of pairs of points of one of the forms $\{(x, y), (x + 1, y)\}$ or $\{(x, y), (x, y + 1)\}$, for some $x \in \mathbb{Z}, y \in \mathbb{Z}$. Two vertices $v = (x, y)$ and $v' = (x', y')$ of $\mathcal{Z}$ are *neighbors* if either $x = x'$ and $|y - y'| = 1$ or else $y = y'$ and $|x - x'| = 1$. We refer to each element of an edge e as its *end*. Let $\mathcal{T}$ be the set of all these edges and set $\mathbf{U}= (\mathcal{Z}, \mathcal{T})$. A *cell* is a subgraph of $\mathbf{U}$, induced by a set $\{(x, y), (x+1, y), (x+1, y+1), (x, y+1)\}$ of four pairwise neighbor vertices. A *path* is a finite alternated sequence $\sigma = v_0, e_1, \ldots, e_k, v_k$ of $k + 1$ vertices and k *different* edges ($k \geq 0$), such that each edge e_i has one end in v_{i-1} and the other one in v_i. We should note that a path may pass several times through a vertex but cannot borrow an edge more than once. The *length* of a path σ as above is k. For the sake of briefness in a path, we may drop edges, identifying σ by the sequence of vertices $v_0, v_1, \ldots, v_k$. If so, any pair of two successive terms v_i and v_{i+1} should constitute a unique set. A *cycle* is a path of length $k \geq 4$ for which the first vertex v_0 and the last one v_k coincide. $\mathbf{U}$ is bipartite i.e. all its cycles are of even length.

Given a cycle γ, one can easily define its inside vertices, see [22]. A vertex (x, y) is said to be *inside* a cycle $\gamma = (x_0, y_0), \ldots, (x_k, y_k)$, with $(x_0, y_0) = (x_k, y_k)$, if $card(\{i \mid y = y_i \text{ and } y \neq y_{i+1} \text{ and } x \leq x_i\})$ is odd (in the addition $i+1$ *modulo* k). According to this definition, the vertices of γ are inside γ.

A *polyominoid* is a partial subgraph $\mathbf{P}= (V, E)$ of $\mathbf{U}$ subject to the following conditions

(i) V is finite,
(ii) $\mathbf{P}$ is connected and

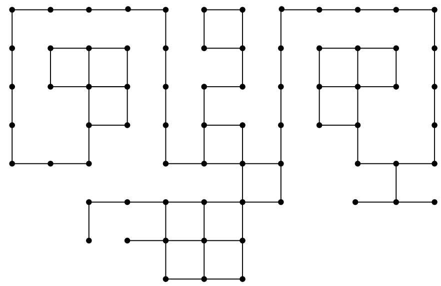


Figure 1. An example of polyominoid

(iii) **P** does not contain any hole, i.e. for all cycle γ in **P**, the vertices inside γ are contained in V and if two neighbor vertices are inside γ, then the linking edge is in E.

It is easy to see that the last property is equivalent to the *tilability* of $\mathbf{P}$=(V, E), i.e. the set of vertices inside γ and their linking edges constitute a subgraph of a grid. The *size* of $\mathbf{P}$= (V, E) is the cardinal of V.

A polyominoid $\mathbf{Q}$= (V', E') is called a *subpolyominoid* of a polyominoid $\mathbf{P}$ = (V, E) if $V' \subseteq V$, $E' \subseteq E$ and $E' = E \cap \{\{u, v\} \mid u \in V', v \in V'\}$.

The class of polyominoids can be defined on **U** by induction in a distributive fashion as follows. The construction is totally distributive in that the application of rewriting rules requires only the knowledge of the neighboring areas in a bull of radius 2. Thus, the set of polyominoids can be generated by a context-free-like grammar. We define the set $\mathcal{P}$ of partial subgraphs of **U** by the following inductive rules:

(a) For any $(x, y) \in \mathcal{Z}$, $\mathbf{P}$=$(\{(x, y)\}, \emptyset)$ is in $\mathcal{P}$.
(b) Let $\mathbf{P}$=$(V, E) \in \mathcal{P}$. Consider two neighbor vertices v and v' such that $v \in V$ and $v' \notin V$. Then, $\mathbf{Q}$=$(V \cup \{v'\}, E \cup \{\{v, v'\}\})$ is in $\mathcal{P}$.
(c) Let $\mathbf{P}$=$(V, E) \in \mathcal{P}$. Suppose V contains 4 neighbor vertices $v_1 = (x, y), v_2 = (x+1, y), v_3 = (x+1, y+1), v_4 = (x, y+1)$, situated on a cell in **U**, such that three edges of the cell on them are in E and the fourth one, say e, is not. Then, $\mathbf{Q}$=$(V, E \cup \{e\})$ is in $\mathcal{P}$.

At this stage, it is not obvious that $\mathcal{P}$ is the class of all polyominoids on **U**. The following proposition shows the equivalence of the two definitions.

Proposition 1. *A partial subgraph* $\mathbf{P}$=(V, G) *of* **U** *is a polyominoid iff it belongs to* $\mathcal{P}$.

3. A Uniform Election Algorithm on Polyominoids

The asynchronous election algorithm, presented in this section, is designed for anonymous networks having a topology of polyominoid. Each vertex only knows the directions of the edges joining it to its neighbors, and knows neither the size of polyominoid nor its own coordinates in the plan. The solution in the general case consists in the computation of a spanning tree, and then election is started for every node. In our study, using the properties of polyominoids, we construct a distributed algorithm which chooses uniformly a vertex as the leader.

3.1. **The Distributed Election.** We now describe the algorithm through a graph relabeling system. Labels (or states) are attached to vertices. Our distributed algorithm is based on a rewriting system, introduced by Litovsky, Métivier and Zielonka [15].

We suppose that initially every vertex has the same label and we look for a noetherian graph rewriting system such that when, after some number of rewriting steps, we get an irreducible labeled graph the there is a special label that is attached to exactly one vertex; this vertex is considered as elected.

In this paper, we use a graph rewriting system enriched by random delays (a rule may be applied if its delay has expired). The graph rewriting system applied here uses forbidden contexts (a rule may be applied if it does not occur in a given forbidden context).

Let $\mathbf{P}= (V,E)$ be a polyominoid and let v a vertex of V. We introduce the set $\mathcal{L}$ of labels $\{N,A,B,L\}$ where N encodes the neutral state, A encodes the active state, B encodes the lost state and lastly L encodes the elected state. Initially, every vertex is of weight $w = 1$ and is $N-$labeled.

Given a polyominoid $\mathbf{P}= (V,E)$, the algorithm works on $\mathbf{P}$ as follows. Any $N-$labeled vertex v decides locally if it is active or not, according to the following rules :

R_0 : If the degree of vertex v is null ($deg(v) = 0$), then the single vertex which constitutes the polyominoid is the elected vertex. This vertex is considered as an *active* vertex.

R_1 : If the degree of vertex v is 1 ($deg(v) = 1$), then the vertex v becomes *active* and it generates its lifetime delay which is an exponentially r.v. (random variable) having its weight as the parameter. Whenever its lifetime has expired, it is removed with its unique incident edge. At this time, the vertex adjacent to the removed one in $\mathbf{P}$, collects the weight of this removed vertex, adding it to its weight.

R_2 : If $deg(v) = 2$, then whenever v is a upper-left most vertex or lower-left most of cell, then it becomes *active* and when its lifetime has expired, it is removed with its incident edges and its right neighbor recuperates its weight.

More precisely, let $\{(x,y),(x+1,y),(x,y+1),(x+1,y+1)\}$ be a cell, if the degree of (x,y) is 2 then (x,y) is *active* and once its lifetime has expired its neighbor $(x+1,y)$ picks up its weight. In the same way and by the horizontal symmetry, if $deg((x,y+1)) = 2$ then $(x,y+1)$ is *active* and its neighbor $(x+1,y+1)$ collects its weight.

R_3 : If $deg(v) = 3$, then if v belongs to two cells and no edge on its left side is found, i.e. v has only one horizontal edge, then v becomes *active* and when its lifetime has expired then its right neighbor recuperates its weight.

So, let $\{(x,y),(x+1,y),(x,y+1),(x+1,y+1)\}$ and $\{(x,y),(x+1,y),(x,y-1),(x-1,y-1)\}$ two cells, if degree (x,y) is 3 then (x,y) becomes *active* and its neighbor $(x+1,y)$ recuperates its weight once its lifetime has expired.

In the sequel, we need the following definition. A vertex which belongs to a maximal cycle in $\mathbf{P}$ is called *peripheral vertex.*

Lemma 1. *Let v be an active vertex of degree 2 or 3 in a polyominoid $\mathbf{P}$. Then v is peripheral.*

Proof. Let $\mathbf{P}= (V,E)$ be a polyominoid and $(x,y) \in V$ an active vertex of degree 2 or 3. By definition, v is situated in a cell, i.e. inside a cycle. Let γ be a maximal cycle having v inside it. If v is on γ, then the proof is complete. Otherwise, there will be a nearest vertex $u = (x',y)$ on γ such that $x' < x$. But $\mathbf{P}$ is a polyominoid and any cycle γ does not contain a hole, i.e, the edges of the segment $[(x',y),(x,y)]$ are in E. This cannot hold, since, v does not admit any edge on its left side.

□

The election algorithm proposed here removes an active vertex once its lifetime has expired. To continue the process, we have to show that the residual graph is still a polyominoid.

Proposition 2. *Let* $\mathbf{P} = (V, E)$ *be a polyominoid of size* ≥ 2 *and let* v *be an active vertex in* $\mathbf{P}$*. The graph* $\mathbf{P}' = (V \setminus \{v\}, E \setminus \{\{v,u\}, u \in V\})$ *is a polyominoid.*

Proof. Let $\mathbf{P}$, v and $\mathbf{P}'$ be as above. To show the proposition, we have to prove that $\mathbf{P}'$ is a connected graph without holes.

- If $deg(v) = 1$, then the suppression of v and its incident edge in $\mathbf{P}$ introduces neither a disconnection nor a hole.
- If $deg(v) = 2$ then let v, v_1, v_2, v_3 be four rectangular vertices of a polyominoid $\mathbf{P}$ such that v is the removable vertex. Consider a vertex $u \in V \setminus \{v, v_1, v_2, v_3\}$. Then, if the vertex u is accessible to a vertex $v_i, 1 \leq i \leq 3$. through a path which passes by v, then when v is removed, u remains accessible to v_i by another path which borrows the vertices $v_{j \neq i}$, $j = 1, 2, 3$. Therefore, by Lemma 1, v is a peripheral vertex and on the other hand its removal generates no hole.
- If $deg(v) = 3$ then the proof is similar to the previous case.

□

3.2. **Standard Spanning Tree.** Let $\mathbf{P}$=(V, E) be a polyominoid. The graph $T = (V, F)$ which traces the weight transmissions in the algorithm is described as follows:

- if $e = \{(x, y), (x + 1, y)\}$ is an edge in E then e belongs to F, i.e. each horizontal edge in E belongs to F:
$$e = \{(x, y), (x + 1, y)\} \in E \Longrightarrow e \in F$$
- if $e = \{(x, y), (x, y + 1)\}$ belongs to E and e is not a left side edge of a cell in $\mathbf{P}$ then e is belong to F, i.e. each vertical edge, who is not a left edge of a cell, belongs to F:
$e = \{(x, y), (x, y + 1)\} \in E \Longrightarrow e \in F$ iff $\{(x, y), (x + 1, y), (x + 1, y + 1), (x, y + 1)\}$ is not a cell in $\mathbf{P}$.

Proposition 3. *The graph* $T = (V, F)$ *described above is a spanning tree of the polyominoid* $\mathbf{P}$*.*

Proof. We can prove this proposition by the inductive construction on $\mathbf{P}$, see [13] for more details. □

Remark. The spanning tree resulting from these rules is unique.

Definition. The spanning tree $T = (V, F)$ is called the *standard spanning tree* of the polyominoid $\mathbf{P}$.

Example 1. Figure 2 gives the standard spanning tree of the polyominoid given in Fig.1.

Proposition 4. *Let* $\mathbf{P} = (V, E)$ *be a polyominoid and* $T = (V, F)$ *be its standard spanning tree. Then, the vertex* $v \in V$ *is active in* $\mathbf{P}$ *iff it is a leaf in* T*.*

Proposition 5. *Let* $\mathbf{P}$ *be a polyominoid of size* ≥ 2*, suppose that* v *is an active vertex in* $\mathbf{P}$ *and let* T *be the standard spanning tree of* $\mathbf{P}$*. Then, let* $\mathbf{P}'$ *denote the residual polyominoid once* v *and its incident edges have been removed and* T' *be the standard spanning tree of* $\mathbf{P}'$*. Then,* T' *can be obtained from* T *by the elimination of the leaf* v *and its incident edge.*

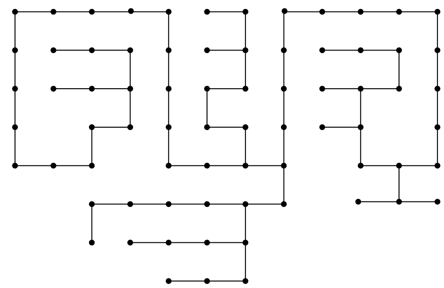


FIGURE 2. Standard spanning tree of the polyominoid given in Fig. 1.

Proof. Let $\mathbf{P}$ be a polyominoid of size ≥ 2, and T its standard spanning tree. Clearly, the residual tree T', after the suppression of a leaf v and its incident edge in T, is a spanning tree of the polyominoid $\mathbf{P}'$ resulting from $\mathbf{P}$ once v and its incident edges are removed. Now, it remains to prove that T' is the standard spanning tree of $\mathbf{P}'$:

- Obviously, the horizontal edges of $\mathbf{P}'$ are in T'.
- The residual vertical edges of $\mathbf{P}'$, which are not situated on the left-hand side of a cell in $\mathbf{P}'$, satisfy the same condition in $\mathbf{P}$ and hence, are in T. Therefore, they are in the standard spanning tree of $\mathbf{P}'$.

□

Putting together the results of this section, we conclude with the following scheme of distributed probabilistic algorithm.

while P is not reduced to a unique vertex
do
- any vertex which active or becomes active (rules $R_0 - R_3$) generates its lifetime according to its weight,
- once the lifetime of an active vertex has expired, it is removed with incident edges and its neighbor in the standard spanning tree collects its weight.

od

4. Analysis of the Algorithm

The election algorithm in a polyominoid is viewed as an election algorithm in its standard spanning tree: as seen in Proposition 4, each active vertex in a polyominoid is a leaf in its standard spanning tree and the weights of the two vertices are equal.

Given $\mathbf{P}$=(V, E) a polyominoid. Initially, all vertices have the same *weight* 1 : $w(v) = 1, \forall v \in V$. According to the rules seen in section 4, when an active vertex vanishes, its successor collects its weight, adding it to its current weight. At the time t when a vertex v becomes active in a residual polyominoid $\mathbf{P}'$, its weight is the number of vanished vertices on its side. The lifetime delay $L(v)$ for v is a r.v. (random variable) having an exponential distribution of parameter $\lambda(v) = w(v)$:

$$Pr\big(L(v) > t\big) = e^{-\lambda(v)t}, \quad \forall t \geq 0.$$

We say that the death of the active vertex v happens according to a Markovian process with the parameter $\lambda(v)$ equal to its weight $w(v)$. This property is equivalent to the one that the death probability of v in the time interval $[t, t+h]$ is $\lambda(v)h+o(h)$, as $h \to 0$ at any time t, and this independent of what is going on elsewhere and of what happened in the past, the assumption which is in agreement with the distributivity of the algorithm. The random process is a variant of pure death processes which are, in turn, special instances of continuous-time-Markov processes (see [7], Chapter XVII).

Theorem 1. *The strategy described below leads to a* **totally** *fair randomized election: in a polyominoid all vertices have the same probability of being elected.*

The proof of this theorem is complicated and is given in the following section, after some preliminary results have been proved.

Uniformity of the Election. The randomized election can mathematically be modeled by a continuous-time Markov process as follows. The initial state of the process is $\mathbf{P}$ (the whole polyominoid). The set of states $\mathcal{E}_{\mathbf{P}}$ is the set of all subpolyominoids $\mathbf{Q}= (U,F)$ of $\mathbf{P}= (V,E)$ satisfying: whenever two diagonal vertices (i.e. of the form (x,y) and $(x+1,y-1)$ or (x,y) and $(x+1,y+1)$) are in $\mathbf{Q}$, then the right vertex $(x+1,y)$, on the cell containing the vertices, is in U, provided that it is in V.

The following proposition shows that $\mathcal{E}_{\mathbf{P}}$ is the set of all subpolyominoids of $\mathbf{P}$ which can be reached from $\mathbf{P}$ by a sequence of active-removal vertices (recall that when an active vertex is removed all incident edges are removed as well).

Proposition 6. *A subpolyominoid* $\mathbf{Q}$ *of a polyominoid* $\mathbf{P}$ *can be reached with a positive probability iff* $\mathbf{Q}$ *is in* $\mathcal{E}_{\mathbf{P}}$.

Proof. Let $\mathbf{Q}$ be a subpolyominoid of $\mathbf{P}$ reachable from $\mathbf{P}$ with a positive probability and prove that $\mathbf{Q} \in \mathcal{E}_{\mathbf{P}}$. According to the process transition definition, $\mathbf{Q}$ must be obtained from $\mathbf{P}$ by $k-$sequence of active-removal vertices ($0 \leq k < n$). For $k = 0$, the proposition obviously holds. Let it be true for k and prove it for $k+1$. So, let $\mathbf{Q}$ be obtained from some polyominoid $\mathbf{R}$ of $\mathbf{P}$ by removing an active vertex v and its incident edges. $\mathbf{R}$ is in $\mathcal{E}_{\mathbf{P}}$ by induction and since, a right most vertex of degree ≥ 2 cannot be active, the resulting subpolyominoid Q will satisfy the condition of being in $\mathcal{E}_{\mathbf{P}}$.

Let now $\mathbf{Q}= (U,F)$ be in $\mathcal{E}_{\mathbf{P}}$. we prove by a decreasing induction over $m = |U|$ that $\mathbf{P}$ can be reached by $n-m$ transitions with a positive probability. For $m = n$, $\mathbf{Q}$ is equal to $\mathbf{P}$ and therefore $\mathbf{Q} \in \mathcal{E}_{\mathbf{P}}$. Suppose that $m < n$, we have to show that there is a vertex $v \in V \setminus U$, such that its addition to U and the addition of all edges with one endpoint v and the other vertex in U, yields a new polyominoid $\mathbf{R}$ belonging to $\mathcal{E}_{\mathbf{P}}$. Since $m < n$ there is a vertex $u \in V \setminus U$. Consider a path from u to a vertex $s \in U$, let v to be the last vertex of the path which does not belong to U. Then, v has a neighbor vertices in U.

- If v has no other neighbor vertex in U, then clearly, v is an active vertex in $\mathbf{R}$ (its degree is 1 in $\mathbf{R}$). Moreover, $\mathbf{R}$ is in $\mathcal{E}_{\mathbf{P}}$.
- Otherwise, v has two or three neighbor vertices in U. In this case, let (x,y) be the coordinates of the vertex v. According to this assumption, $\mathbf{Q}$ is in $\mathcal{E}_{\mathbf{P}}$, there are no neighbor vertices in U, $(x,y+1)$ and $(x-1,y)$ or $(x-1,y)$ and $(x,y-1)$ such that v is a vertex on the right side of a cell in $\mathbf{R}$ containing the vertices. Consequently, v is a vertex on the left side of one or two cells in $\mathbf{R}$. However, v is an active vertex in $\mathbf{R}$. Moreover, $\mathbf{R} \in \mathcal{E}_{\mathbf{P}}$.

□

Let $\mathbf{Q}$ be the state of the system at instant t. According to the distributive random structure of the algorithm, any active vertex v of $\mathbf{Q}$ has a lifetime exponentially distributed with a parameter equal to its weight. This is equivalent to the fact that in the time interval $[t, t+\Delta t]$, v may disappear with all incident edges with probability $w(v)\Delta t + o(h)$, as $h \to 0$, and this independent of what is going on elsewhere and what happened in the past.

One can easily show that the probability of passing from $\mathbf{Q}$ to $\mathbf{R}$ is obtained by the removal of active vertex v and its incident edges is given by:

$$P_{(\mathbf{Q},\mathbf{R})} = \frac{w(v)}{\sum_{u \text{ active in } \mathbf{Q}} w(u)} \tag{1}$$

provided that $\mathbf{Q}$ is not reduced to a vertex. The absorbing states (see [7]) are polyominoids reduced to a vertex (the elected vertex).

A mathematical description of probability of being in state $\mathbf{Q}$ at time t can be given as the solution of a system of differential equations. The following proposition can be proved without any difficulty by a straightforward adaptation of the proof given in [7], Chapter XVII, Section 5.

Proposition 7. *Let $\mathbf{Q}$ be in $\mathcal{E}_{\mathbf{P}}$ and let $P_{\mathbf{Q}}(t)$ denote the probability that the state of the election at time t is $\mathbf{Q}$. We have:*

(i) $\dfrac{dP_{\mathbf{P}}(t)}{dt} = -w(\mathbf{P})P_{\mathbf{P}}(t)$,

(ii) *for all subpolyominoid $\mathbf{Q}{\neq}\mathbf{P}$ of size at least 2 and in $\mathcal{E}_{\mathbf{P}}$,*

$$\frac{dP_{\mathbf{Q}}(t)}{dt} = -w(\mathbf{Q})(t)P_{\mathbf{Q}}(t) + \sum_v w(v)P_{\mathbf{R}}(t),$$

with $\mathbf{R} = \mathbf{Q} \cup (\{v\}, \{\{v,u\}, u \text{ adjacent to } v \text{ in } T\})$, (recall that T is standard spanning tree of $\mathbf{P}$)

where the summation is extended to all vertices v adjacent to $\mathbf{Q}$ in T which do not belong to $\mathbf{Q}$, and

(iii) $$\frac{dP_{(\{v\},\emptyset)}(t)}{dt} = \sum_{u \text{ adjacent to } v \text{ in } \mathbf{P}} w(u)P_{(\{v,u\},\{\{v,u\}\})}(t),$$

with the initial condition $P_{\mathbf{P}}(0) = 1$.

This proposition characterizes *in principle* the distribution probability of states at a given time t. In particular, it should enable us to compute the absorption probabilities [7]. However, no explicit solution is known to the authors.

Propositions 3-5 allow to confirm that any sequence of transitions over $\mathcal{E}_{\mathbf{P}}$ can be simulated, with the same probability, by a sequence of transitions over the set of factor trees in the standard spanning tree of $\mathbf{P}$ (recall that a factor of a tree is a tree obtained by a sequence of leaf removals). Thus, the study of the process is translated into that of the election over a tree, proposed and analyzed in [19]. In this model, initially all vertices have the same weight 1. Each leaf has a lifetime which is an exponentially distributed random variable with a parameter equal to the weight of the leaf. Once the lifetime of the leaf has expired, it is removed with the incident edge and its weight is recuperated by its father. The process continuous on until the tree is reduced to one vertex, which is considered as the elected vertex. Therefore, the probability of being elected for a vertex v in a polyominoid $\mathbf{P}$ is the same as in the standard spanning tree T and this has been shown to be $\frac{1}{n}$, where n is the size of $\mathbf{P}$.

We enumerate here intermediate results and give the outline of the proof. In the sequel, we suppose that T is the spanning tree of polyominoid $\mathbf{P}$ of size n. Leaves of $\mathbf{P}$ are removed following the random process described above until T is reduced to a unique vertex. We have to prove the uniformity of the chance for all vertices of T.

We first introduce a slight modification of the leaf-removal model over T. We translate the model into a variant on directed trees. For a given vertex v, the unique rooted tree at v can be defined. These rooted trees can be used in a natural way to compute the absorption probabilities.

We consider forests of rooted trees. Let F be a forest of rooted trees, we introduce a death process on F as follows. Each leaf v has an exponentially distributed lifetime with a parameter equal to its weight; initially, all vertices of F are of weight 1. At any time interval $[t, t+\Delta t]$, if the lifetime of a leaf has expired, the leaf is removed with its unique incident edge. If the vanishing leaf has a father, then its father picks up its weight, adding it to its weight. The leaf-removal process goes on the reduced forest until the forest totally disappears.

For a given forest F, let $L(F)$ be the vanishing time; it is a positive-real-valued r.v..

The following proposition is surprising. It asserts that $L(F)$ depends only on the size of the forest and not on its structure.

Proposition 8. *Let F be a forest of size n then the distribution function $G_F(t)$ of the r.v. $L(F)$ is given by:*

$$G_F(t) = Pr(L(F) \leq t) = (1-e^{-t})^n, \quad \forall t \geq 0.$$

Proof. By induction on n. If F is reduced to a vertex, then the proposition holds (the lifetime for a single vertex is an exponentially distributed r.v. with parameter 1). Suppose that the proposition holds for forests of size less than n and let us prove it for a forest F of size n ($n \geq 2$).

(i) Suppose that F consists of forests $F_1, F_2, \cdots, F_k$ with $k \geq 2$. Let $n = n_1 + n_2 + \cdots + n_k$, where n_i is the size of F_i, $1 \leq i \leq k$. In this case, $L(F_i)$, $1 \leq i \leq k$, are mutually independent r.v. and hence by the induction hypothesis:

$$\begin{aligned} Pr(L(F) \leq t) &= \prod_{i=1}^k Pr(L(F_i) \leq t) \\ &= \prod_{i=1}^k (1-e^{-t})^{n_i} \\ &= (1-e^{-t})^n. \end{aligned}$$

(ii) Otherwise, suppose that F has size n and consists of a unique root r and rooted trees $A_1, A_2, \cdots, A_k$. Now, let F' be the forest consisting of A_1, $A_2, \cdots, A_k$ (alternatively, let $F' = A_1 \cup \cdots \cup A_k$). F' has size $n-1$ and, by the induction hypothesis, $(1-e^{-t})^{n-1}$ is the distribution function of $L(F')$. But, $L(F)$ is the sum of two independent r.v. $L(F')$ and the lifetime of r. The last one is an exponential r.v. of parameter n (weight of r). Thus, $L(F)$ has the distribution function (see [8], p. 142. Theorem 2) given by:

$$\begin{aligned} G_F(t) &= \int_0^t G_{F'}(t-x)d(1-e^{-nx}) \\ &= \int_0^t G_{F'}(t-x)ne^{-nx}dx, \end{aligned}$$

where $G_{F'}(t-x) = (1-e^{t-x})^{n-1}$.

Hence:

$$\begin{aligned} G_F(t) &= \int_0^t n[1-e^{-(t-x)}]^{n-1}e^{-nx}dx \\ &= \int_0^t n[e^{-x}-e^{-t}]^{n-1}e^{-x}dx \\ &= [-(e^{-x}-e^{-t})^n]_{x=0}^{x=t} \\ &= (1-e^{-t})^n. \end{aligned}$$

The proposition follows. □

Given two forests F_1 and F_2, we say F_1 *beats* F_2, if $L(F_1) \geq L(F_2)$. The next result easily follows from the above lemma.

Corollary 1. *Let F_1 and F_2 be two forests of sizes n_1 and n_2 respectively. The probability that F_1 beats F_2 is $\frac{n_1}{n_1+n_2}$.*

Lemma 2. *Consider a vertex v in T with the adjacent vertices $v_1, \ldots, v_k$. Let F consist of two trees A and B obtained by the suppression of edge $\{v, v_1\}$ rooted at v and v_1 respectively. Then, the probability that v is removed before the whole tree factor on the side of v_1 (i.e. undirected B) in the election process over T is the same as the probability of v_1 beating v in F.*

Proof. The events whose probabilities are to be calculated can be represented as sequences of leaves being removed:

- $\sigma = \langle l_1, \ldots, l_k \rangle$, where $l_i, 1 \leq i \leq k$ are leaves or vertices which become leaves in T (or in F respectively) after the removal of some previous vertices in the sequence,
- $l_k = v$ and
- v_1 does not figure in σ.

On the one hand, it is easy to see that any sequence satisfying the above conditions in T does it in F and vice versa. On the other hand, the probability of such σ according to (1) is:

$$P(\sigma) = \prod_{1 \leq i \leq k} q_i,$$

with

$$q_i = \frac{\lambda(l_i)}{\lambda(T_i)},$$

where T_i is the residual tree (arborescence respectively) just before the l_i removal. In each step of the leaf removal along σ, T and F have the same set of leaves and, hence, the involved quantities on T are the same as the corresponding ones on F. The lemma follows. □

Proposition 9. *Let $q(v)$ denote the probability of being elected in T for a vertex v. We have $q(v) = \frac{1}{n}$.*

Proof. For $n = 1$ or $n = 2$ the proposition is obvious. Otherwise, let $v_1, \ldots, v_k$ be the adjacent vertices to v. Let, on the other hand, $A_1, \ldots, A_k$ be disjoint tree rooted at $v_1, \ldots, v_k$ of sizes $n_1, \ldots, n_k$ respectively. Clearly, v fails iff it vanishes before one of the factors situated on the v_i side for $1 \leq i \leq k$. These last events are pairwise disjoint and therefore, according to the previous lemma, the failure probability of v is the sum of the probabilities of v being beaten by one of its neighbors v_i in the forest consisting of the tree rooted at v and v_i respectively. Hence, according to Corollary 1, we have:

$$1 - q(v) = \sum_{i=1}^{k} \frac{n_i}{n}.$$

Since, $\sum_{i=1}^{k} n_i = n - 1$, the proposition follows. □

Proof of Theorem 1. Straightforward by the similarity of the election process over **P** and over its standard spanning tree T and the previous proposition. □

EFERENCES

[1] D. Angluin. Local and global properties in networks of processors. In *Proceedings of the 1 th Symposium on theory of computing*, pages 82–93, 1980.

[2] M. Bousquet-Mélou. Convex polyominoes and heaps of segments. *J. Phys. A: Math. Gen.*, 25:1925–1934, 1992a.

[3] M. Bousquet-Mélou. A method for the enumeration of various classes of column-convex polygons. *Discrete Mathematics*, 154:1–25, 1996.

[4] A. Bretto. *Introduction to Hypergraph Theory and its Applications to Image Processing. Mongraphy in: Advances in Imaging and Electron Physics*. Academic Press, March 2004.

[5] M. Delest, J.-P. Dubernard, and I. Dutour. Parallelogram polyominoes and corners. *J. Symbolic Comput.*, 20:503 –515, 1995.

[6] M.-P. Delest and G. Viennot. Algebraic languages and polyominoes [sic] enumeration. 34:169–206, 1984.

[7] W. Feller. *An Introduction to Probability Theory and Its Application, Volume I.* John Wiley and Sons, 1950.

[8] W. Feller. *An Introduction to Probability Theory and Its Application, Volume II.* John Wiley and Sons, 1966.

[9] M. Gardner. *Mathematics, magic and mystery.* New York. Dover Publications, 1956.

[10] M. Gardner. *New mathematical diversions.* Washington. Math. Assoc. Amer., 1995.

[11] S.W. Golomb. Checker boards and polyominoes. *American Mathematical Monthly*, 1954.

[12] S.W. Golomb. *Polyominoes: puzzles, patterns, problems, and packings.* Princeton University Press, 1996.

[13] A. El Hibaoui, N. Saheb-Djahromi, and A. Zemmari. Polyominoids and uniform election. Technical Report 1337-04, LaBRI, Université Bordeaux 1, November 2004.

[14] G. Le Lann. Distributed systems – toward a formal approach. In *Proceedings of the I IP Congress 77*, pages 155–160, 1977.

[15] I. Litovsky, Y. Métivier, and W. Zielonka. On the recognition of families of graphs with local computations. *Inform. and Comput.*, 118:110–119, 1995.

[16] A. Del Lungo, A. Frosini, M. Nivat, and L. Vuillon. Discrete tomography: reconstruction under periodicity constraints. In P. Widmayer, F. Triguero, R. Morales, M. Hennessy, S. Eidenbenz, and R. Conejo, editors, *ICALP*, volume 2380 of *Lecture Notes in Computer Science*, pages 38–56. Springer, 2002.

[17] A. Mazurkiewicz. Solvability of the asynchronous ranking problem. In *Inf. Processing Letters*, volume 28, pages 221–224, 1988.

[18] Y. Métivier and N. Saheb. Medians and centres of polyominos. *Inform. Proc. Letters*, 57:175–181., 1996.

[19] Y. Métivier, N. Saheb-Djahromi, and A. Zemmari. A uniform randomized election in trees (extended abstract). In *Proceedings of The 1 th International Colloquium on Structural Information and Communication Complexity (SIROCCO 1)*, pages 259–274. Carleton university press, 2003.

[20] M. Raynal. *Networks and distributed computation : concepts, tools, and algorithms.* MIT Press, Cambridge, 1988.

[21] K. H. Rosen. *Handbook of discrete and combinatorial mathematics.* CRC Press, 2000.

[22] Robert Sedgewick. *Algorithms in C++.* Addison-Wesley Co., 1st edition, 1992.

[23] G. Tel. *Introduction to distributed algorithms.* Cambridge University Press, 2000.

[24] H.N.V. Temperley. Combinatorial problems suggested by the statistical mechanics of domains and of rubber-like molecules. *Phy.*, 103:1–16, 1956.

THE NUMBER OF MONOTONE TRIANGLES WITH PRESCRIBED BOTTOM ROW

ILSE FISCHER

Abstract. We show that the number of monotone triangles with prescribed bottom row $(k_1, \ldots, k_n) \in \mathbb{Z}^n$, $k_1 < k_2 < \ldots < k_n$, is given by a simple product formula which remarkably involves (shift) operators. Monotone triangles with bottom row $(1, 2, \ldots, n)$ are in bijection with $n \times n$ alternating sign matrices.

1. Introduction

An *alternating sign matrix* is a square matrix of 0s, 1s and -1s for which the sum of entries in each row and in each column is 1 and the non-zero entries of each row and of each column alternate in sign. For instance,

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

is an alternating sign matrix. In the early 1980s, Robbins and Rumsey [8] introduced alternating sign matrices in the course of generalizing a determinant evaluation algorithm. Out of curiosity they posed the question for the number of alternating sign matrices of fixed size and, together with Mills, they came up with the appealing conjecture [7] that the number of $n \times n$ alternating sign matrices is

$$\prod_{j=0}^{n-1} \frac{(3j+1)!}{(n+j)!}. \tag{1.1}$$

This turned out to be one of the hardest problems in enumerative combinatorics within the last decades. In 1996, Zeilberger [10] finally succeeded in proving their conjecture. Then, some months later, Kuperberg [5] realized that alternating sign matrices are equivalent to a model in statistical physics for two-dimensional square ice. Using a determinental expression for the partition function of this model discovered earlier by physicists, he was able to provide a shorter proof of the formula. For a nice exposition on this topic see [1].

Alternating sign matrices can be translated into certain triangular arrays of positive integers, called *monotone triangles*. Monotone triangles are probably the right guise of alternating sign matrices for a recursive treatment [1, Section 2.3]. In order to obtain the monotone triangle corresponding to a given alternating sign matrix, replace every entry in the matrix by the sum of entries in the same column above, the entry itself included. In our running example we obtain

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Row by row we record the columns that contain a 1 and obtain the following triangular array.

$$\begin{array}{ccccccccc} & & & & 4 & & & & \\ & & & 2 & & 5 & & & \\ & & 1 & & 3 & & 5 & & \\ & 1 & & 2 & & 4 & & 5 & \\ 1 & & 2 & & 3 & & 4 & & 5 \end{array}$$

This is the monotone triangle corresponding to the alternating sign matrix above. Observe that it is weakly increasing in northeast direction and in southeast direction. Moreover, it is strictly increasing along rows. In general, a monotone triangle with n rows is a triangular array $(a_{i,j})_{1 \le j \le i \le n}$ such that $a_{i,j} \le a_{i-1,j} \le a_{i,j+1}$ and $a_{i,j} < a_{i,j+1}$ for all i, j. It is not too hard to see that monotone triangles with n rows and bottom row $(1, 2, \ldots, n)$, i.e. $a_{n,j} = j$, are in bijection with $n \times n$ alternating sign matrices. Our main theorem provides a formula for the number of monotone triangles with prescribed last row $(k_1, \ldots, k_n) \in \mathbb{Z}^n$.

Theorem 1. *The number of monotone triangles with n rows and bottom row $k_1, k_2, \ldots, k_n$ is given by*

$$\left(\prod_{1 \le p < q \le n} \left(\mathrm{id} + E_{k_p} \Delta_{k_q} \right) \right) \prod_{1 \le i < j \le n} \frac{k_j - k_i}{j - i},$$

where E_x denotes the shift operator, defined by $E_x \, p(x) = p(x+1)$, and $\Delta_x := E_x - \mathrm{id}$ denotes the difference operator.

In order to understand this formula, there are a few things to remark. The product of operators is understood as the composition. Moreover note that the shift operators commute, and consequently, it does not matter in which order the operators in the product $\prod_{1 \le p < q \le n} \left(\mathrm{id} + E_{k_p} \Delta_{k_q} \right)$ are applied. In order to use this formula to compute the number of monotone triangles with bottom row $(k_1, \ldots, k_n)$, one first has to apply the operator $\prod_{1 \le p < q \le n} \left(\mathrm{id} + E_{x_p} \Delta_{x_q} \right)$ to the polynomial $\prod_{1 \le i < j \le n} \frac{x_j - x_i}{j - i}$ and then set $x_i = k_i$. Thus, it is not so clear how to derive (1.1) from this formula.

What is the significance of the formula? In the last decades, the enumeration of plane partitions, alternating sign matrices and related objects subject to a variety of different constraints has attracted a lot of interest. This attraction stems from the fact that now and then these enumerations lead to appealing product formula or hypergeometric series, which are, in spite of their simplicity, pretty hard to prove. At the moment the search for these simple product formulas seems to be a bit exhausted. Therefore, a new challenge is the search for possibilities to give enumeration formulas for the vast majority of enumeration problems for which there exists no closed formula in a traditional sense. The formula in Theorem 1 contributes to this issue.

Also note that the second product in the formula in Theorem 1, i.e. $\prod_{1 \le i < j \le n} \frac{k_j - k_i}{j-i}$, is the number of semistandard tableaux of shape $(k_n - n, k_{n-1} - n, \ldots, k_1 - 1)$ and, equivalently, the number of columnstrict plane partitions of this shape, see [9, p. 375, in (7.105) $q \to 1$]. In fact, these objects are in bijection with monotone triangles with prescribed bottom row $(k_1, k_2, \ldots, k_n)$ that are strictly increasing in southeast direction, see [2, Section 5]. Thus, our formula once more gives an indication of the relation between plane partitions and alternating sign matrices manifested by a number of enumeration formulas which show up in both fields, a phenomenon which is not yet well (i.e. bijectively) understood.

In this extended abstract we sketch the proof of Theorem 1. (See [4] for the full version of this paper.) The method can roughly be described as follows. In the first step, we introduce a recursion, which relates monotone triangles with n rows to monotone triangles with $n-1$ rows. This recursion immediately implies that the enumeration formula is a polynomial in $k_1, k_2, \ldots, k_n$. In the next step we compute the degree of the polynomial. Finally, we deduce enough properties of the polynomial in order to compute it. The polynomial's degree determines how much information is in fact needed. This method is related to the method for proving polynomial enumeration formulas we have introduced in [2] and extended in [3]. In the final section we mention some problems around Theorem 1 we plan to consider next.

2. The recursion

In the following let $\alpha(n; k_1, \ldots, k_n)$, $n \ge 1$, denote the number of monotone triangles with $(k_1, \ldots, k_n)$ as bottom row. If we delete the last row of such a monotone triangle we obtain a monotone triangle with $n-1$ rows and bottom row, say, $(l_1, l_2, \ldots, l_{n-1})$. By the definition of a monotone triangle $k_1 \le l_1 \le k_2 \le l_2 \le \ldots \le k_{n-1} \le l_{n-1} \le k_n$ and $l_i \ne l_{i+1}$. Thus

$$\alpha(n; k_1, \ldots, k_n) = \sum_{\substack{(l_1, \ldots, l_{n-1}) \in \mathbb{Z}^{n-1}, \\ k_1 \le l_1 \le k_2 \le \ldots \le k_{n-1} \le l_{n-1} \le k_n, l_i \ne l_{i+1}}} \alpha(n-1; l_1, \ldots, l_{n-1}). \tag{2.1}$$

We introduce the following abbreviation

$$\sum_{\substack{(l_1, \ldots, l_{n-1}) \in \mathbb{Z}^{n-1}, \\ k_1 \le l_1 \le k_2 \le \cdots \le k_{n-1} \le l_{n-1} \le k_n, l_i \ne l_{i+1}}} =: \sum_{(l_1, \ldots, l_{n-1})}^{(k_1, \ldots, k_n)}$$

for $n \ge 2$. This summation operator is well-defined for all $(k_1, \ldots, k_n) \in \mathbb{Z}^n$ with $k_1 < k_2 < \ldots < k_n$. We extend the definition to arbitrary $(k_1, \ldots, k_n) \in \mathbb{Z}^n$ by induction with respect to n. If $n = 2$ then

$$\sum_{(l_1)}^{(k_1, k_2)} A(l_1) := \sum_{l_1 = k_1}^{k_2} A(l_1),$$

where here and in the following we use the extended definition of the summation over an interval, namely,

$$\sum_{i=a}^{b} f(i) = \begin{cases} f(a) + f(a+1) + \cdots + f(b) & \text{if } a \le b \\ 0 & \text{if } b = a-1 \\ -f(b+1) - f(b+2) - \cdots - f(a-1) & \text{if } b+1 \le a-1 \end{cases} . \tag{2.2}$$

This assures that for any polynomial $p(X)$ over an arbitrary integral domain I containing $\mathbb{Q}$ there exists a unique polynomial $q(X)$ over I such that $\sum_{x=0}^{y} p(x) = q(y)$ for all integers y. We usually write $\sum_{x=0}^{y} p(x)$ for $q(y)$. (We also use the analog extended definition for the product symbol.) If $n > 2$ then

$$\sum_{(l_1,\ldots,l_{n-1})}^{(k_1,\ldots,k_n)} A(l_1,\ldots,l_{n-1}) := \sum_{(l_1,\ldots,l_{n-2})}^{(k_1,\ldots,k_{n-1})} \sum_{l_{n-1}=k_{n-1}+1}^{k_n} A(l_1,\ldots,l_{n-2},l_{n-1}) + \sum_{(l_1,\ldots,l_{n-2})}^{(k_1,\ldots,k_{n-1}-1)} A(l_1,\ldots,l_{n-2},k_{n-1}).$$

We renew the definition of $\alpha(n;k_1,\ldots,k_n)$ after this extension by setting $\alpha(1;k_1) = 1$ and

$$\alpha(n;k_1,\ldots,k_n) = \sum_{(l_1,\ldots,l_{n-1})}^{(k_1,\ldots,k_n)} \alpha(n-1;l_1,\ldots,l_{n-1}).$$

This gives us an extension of our original function $\alpha(n;k_1,\ldots,k_n)$ to arbitrary $(k_1,\ldots,k_n) \in \mathbb{Z}^n$. The recursion implies that $\alpha(n;k_1,\ldots,k_n)$ is a polynomial in $k_1,\ldots,k_n$. We have used this recursion (and a computer) to compute $\alpha(n;k_1,\ldots,k_n)$ for $n = 1,2,3,4$ and obtain the following

$$1, 1-k_1+k_2, \frac{1}{2}(-3k_1+k_1^2+2k_1k_2-k_1^2k_2-2k_2^2+k_1k_2^2+3k_3-4k_1k_3+k_1^2k_3+ 2k_2k_3-k_2^2k_3+k_3^2-k_1k_3^2+k_2k_3^2), \frac{1}{12}(20k_2+11k_1k_2-16k_1^2k_2+3k_1^3k_2+4k_1k_2^2+3k_1^2k_2^2-k_1^3k_2^2+ 4k_2^3-5k_1k_2^3+k_1^2k_2^3-20k_3+16k_1k_3-4k_1^2k_3-27k_2k_3+9k_1^2k_2k_3-2k_1^3k_2k_3-3k_1^2k_2^2k_3+k_1^3k_2^2k_3- 3k_2^3k_3+4k_1k_2^3k_3-k_1^2k_2^3k_3+16k_1k_3^2-12k_1^2k_3^2+2k_1^3k_3^2-9k_1k_2k_3^2+6k_1^2k_2k_3^2-k_1^3k_2k_3^2+9k_2^2k_3^2- 3k_1k_2^2k_3^2-3k_2^3k_3^2+k_1k_2^3k_3^2-4k_3^3+8k_1k_3^3-2k_1^2k_3^3-3k_2k_3^3-2k_1k_2k_3^3+k_1^2k_2k_3^3+3k_2^2k_3^3-k_1k_2^2k_3^3- 27k_1k_4+20k_1^2k_4-3k_1^3k_4+16k_2k_4+24k_1k_2k_4-24k_1^2k_2k_4+4k_1^3k_2k_4-16k_2^2k_4+9k_1k_2^2k_4+ 3k_1^2k_2^2k_4-k_1^3k_2^2k_4+8k_2^3k_4-6k_1k_2^3k_4+k_1^2k_2^3k_4+11k_3k_4-24k_1k_3k_4+15k_1^2k_3k_4-2k_1^3k_3k_4- 9k_2^2k_3k_4+2k_2^3k_3k_4-4k_3^2k_4+9k_1k_3^2k_4-6k_1^2k_3^2k_4+k_1^3k_3^2k_4+3k_2^2k_3^2k_4-k_2^3k_3^2k_4-5k_3^3k_4+ 6k_1k_3^3k_4-k_1^2k_3^3k_4-4k_2k_3^3k_4+k_2^2k_3^3k_4-20k_1k_4^2+9k_1^2k_4^2-k_1^3k_4^2+4k_2k_4^2+15k_1k_2k_4^2-9k_1^2k_2k_4^2+ k_1^3k_2k_4^2-12k_2^2k_4^2+6k_1k_2^2k_4^2+2k_2^3k_4^2-k_1k_2^3k_4^2+16k_3k_4^2-24k_1k_3k_4^2+9k_1^2k_3k_4^2-k_1^3k_3k_4^2+ 9k_2k_3k_4^2-6k_2^2k_3k_4^2+k_2^3k_3k_4^2+3k_3^2k_4^2-3k_1k_3^2k_4^2+3k_2k_3^2k_4^2-k_3^3k_4^2+k_1k_3^3k_4^2-k_2k_3^3k_4^2-3k_1k_4^3+ k_1^2k_4^3+2k_1k_2k_4^3-k_1^2k_2k_4^3-2k_2^2k_4^3+k_1k_2^2k_4^3+3k_3k_4^3-4k_1k_3k_4^3+k_1^2k_3k_4^3+2k_2k_3k_4^3-k_2^2k_3k_4^3+ k_3^2k_4^3-k_1k_3^2k_4^3+k_2k_3^2k_4^3).$$

From this data it is obviously hard to guess a general formula for $\alpha(n;k_1,\ldots,k_n)$. However, it seems plausible that the degree of $\alpha(n;k_1,\ldots,k_n)$ in k_i is $n-1$. In the following two sections we prove that this is indeed true. Note that at first glance the linear growth of the degree is quite surprising: suppose $A(l_1,\ldots,l_{n-1})$ is a polynomial of degree no greater than R in each of l_{i-1} and l_i. Then

$$\deg_{k_i}\left(\sum_{(l_1,\ldots,l_{n-1})}^{(k_1,\ldots,k_n)} A(l_1,\ldots,l_{n-1})\right) =$$

$$\deg_{k_i}\left(\sum_{l_{i-1}=k_{i-1}}^{k_i}\sum_{l_i=k_i}^{k_{i+1}} A(l_1,\ldots,l_{n-1}) - A(l_1,\ldots,l_{i-2},k_i,k_i,l_{i+1},\ldots,l_{n-1})\right) \leq 2R+2$$

and there exist polynomials $A(l_1,\ldots,l_{n-1})$ such that the upper bound $2R+2$ is attained. Consequently, $\alpha(n;k_1,\ldots,k_n)$ must be of a very specific shape.

3. Sketch of the proof of Theorem 1

In this section we sketch the proof of the main theorem by presenting the relevant lemmas without proofs.

Recall that the *shift operator*, denoted by E_x, is defined as $E_x p(x) = p(x+1)$. Clearly E_x is invertible in the algebra of operators of $\mathbb{C}[X]$ and we denote its inverse by E_x^{-1}. Observe that the shift operators with respect to different variables commute, i.e. $E_xE_y = E_yE_x$. The difference operator Δ_x is defined as $\Delta_x = E_x - \mathrm{id}$. However, the difference operator Δ_x is not invertible since it decreases the degree of a polynomial.

If we apply the shift operator or the delta operator to the i-th variable of a function, we sometimes write E_i or Δ_i, respectively, i.e. $\Delta_{k_i} f(k_1,\ldots,k_n) = \Delta_i f(k_1,\ldots,k_n)$. Moreover, $\Delta_2 f(k_3,k_3,k_3)$, for instance, is shorthand for

$$\left(\Delta_{l_2} f(l_1,l_2,l_3)\right)\big|_{l_1=k_3,l_2=k_3,l_3=k_3}.$$

The *swapping operator* $S_{x,y}$ is applicable to functions in (at least) two variables and defined as $S_{x,y} f(x,y) = f(y,x)$. If we apply it to the i-th and j-th variable of a function we sometimes write $S_{i,j}$.

In the following we consider rational functions in shift operators. In order to guarantee that the inverse of the denominator always exists, we need the following lemma.

Lemma 1. *Let $p(x_1,\ldots,x_n)$ be a polynomial in $x_1,x_1^{-1},x_2,x_2^{-1},\ldots,x_n,x_n^{-1}$ over $\mathbb{C}$, and fix an integer i, $1 \le i \le n$. Consider the operator*

$$\mathrm{id} + \Delta_{k_i} p(E_{k_1},E_{k_2},\ldots,E_{k_n}) =: \mathrm{Op}$$

on $\mathbb{C}[k_1,\ldots,k_n]$. Then Op is invertible and the inverse is

$$\mathrm{Op}^{-1} = \sum_{l=0}^{\infty} (-1)^l \Delta_{k_i}^l p(E_{k_1},E_{k_2},\ldots,E_{k_n})^l,$$

where $\Delta_{k_i}^0 p(E_{k_1},E_{k_2},\ldots,E_{k_n})^0 = \mathrm{id}$. Moreover

$$\deg_{k_i} G(k_1,\ldots,k_n) = \deg_{k_i} \mathrm{Op}\, G(k_1,\ldots,k_n) = \deg_{k_i} \mathrm{Op}^{-1} G(k_1,\ldots,k_n).$$

We define two operators applicable to polynomials $G(k_1,\ldots,k_n) \in \mathbb{C}[k_1,\ldots,k_n]$. We set

$$V_{k_i,k_j} = \mathrm{id} + E_{k_i}^{-1} \Delta_{k_i} \Delta_{k_j} = E_{k_i}^{-1}(\mathrm{id} + E_{k_j} \Delta_{k_i})$$

and

$$T_{k_i,k_{i+1}} = (\mathrm{id} + E_{k_{i+1}} E_{k_i}^{-1} S_{k_i,k_{i+1}}) \frac{V_{k_i,k_{i+1}}}{V_{k_i,k_{i+1}} + V_{k_{i+1},k_i}}.$$

By Lemma 1, the inverse $(V_{k_i,k_{i+1}} + V_{k_{i+1},k_i})^{-1}$ is well-defined. The following lemma explains the significance of $T_{k_i,k_{i+1}}$ for the recursion (2.1).

Lemma 2. *Let $A(l_1,l_2)$ be a polynomial in l_1 and l_2 which is of degree at most R in each of l_1 and l_2. Moreover assume that $T_{l_1,l_2} A(l_1,l_2)$ is of degree at most R as a polynomial in l_1 and l_2, i.e. a linear combination of monomials $l_1^m l_2^n$ with $m+n \le R$. Then*

$$\sum_{(l_1,l_2)}^{(k_1,k_2,k_3)} A(l_1,l_2) = \sum_{l_1=k_1}^{k_2} \sum_{l_2=k_2}^{k_3} A(l_1,l_2) - A(k_2,k_2)$$

is of degree at most $R+2$ in k_2. Moreover, if $T_{l_1,l_2} A(l_1,l_2) = 0$ then $\sum_{(l_1,l_2)}^{(k_1,k_2,k_3)} A(l_1,l_2)$ is of degree at most $R+1$ in k_2.

In order to use Lemma 2 to compute the degree of $\sum_{(l_1,l_2)}^{(k_1,k_2,k_3)} A(l_1,l_2)$ in k_2 one has to compute the degree of $T_{l_1,l_2} A(l_1,l_2)$ in l_1 and l_2. However, the operator T_{l_1,l_2} is complicated and thus it is convenient to consider a simplified version of T_{l_1,l_2} for this purpose, which we obtain by multiplying an operator that preserves the degree.

$$T'_{k_i,k_{i+1}} = E_{k_i}(V_{k_i,k_{i+1}} + V_{k_{i+1},k_i})T_{k_i,k_{i+1}} = \\ (\mathrm{id} + S_{k_i,k_{i+1}})E_{k_i}V_{k_i,k_{i+1}} = (\mathrm{id} + S_{k_i,k_{i+1}})(\mathrm{id} + E_{k_{i+1}}\Delta_{k_i})$$

Observe that $\deg_{k_i,k_{i+1}} T_{k_i,k_{i+1}} G(k_1,\dots,k_n) = \deg_{k_i,k_{i+1}} T'_{k_i,k_{i+1}} G(k_1,\dots,k_n)$, since

$$V_{k_i,k_{i+1}} + V_{k_{i+1},k_i} = 2\,\mathrm{id} + (E_{k_i}^{-1} + E_{k_{i+1}}^{-1})\Delta_{k_i}\Delta_{k_{i+1}}$$

and $\Delta_{k_i}\Delta_{k_{i+1}}$ decreases the degree of a polynomial in k_i and k_{i+1}. In particular, $T_{k_i,k_{i+1}} G(k_1,\dots,k_n) = 0$ if and only if $T'_{k_i,k_{i+1}} G(k_1,\dots,k_n) = 0$.

Suppose $A(l_1,\dots,l_n)$ is a function on $\mathbb{Z}^n$. Next we aim to express

$$T'_{k_i,k_{i+1}} \left(\sum_{(l_1,\dots,l_n)}^{(k_1,\dots,k_{n+1})} A(l_1,\dots,l_n) \right) (k_1,k_2,\dots,k_{n+1})$$

in terms of $T'_{l_{i-1},l_i} A(l_1,\dots,l_n)$ and $T'_{l_i,l_{i+1}} A(l_1,\dots,l_n)$. In particular, this shows that if $T'_{l_i,l_{i+1}} A(l_1,\dots,l_n) = 0$ for all $i = 1,2,\dots,n-1$ then

$$T'_{k_i,k_{i+1}} \left(\sum_{(l_1,\dots,l_n)}^{(k_1,\dots,k_{n+1})} A(l_1,\dots,l_n) \right) (k_1,\dots,k_{n+1}) = 0$$

for all $i = 1,2,\dots,n$.

Lemma 3. *Let $f(k_1,k_2,k_3)$ be a function from $\mathbb{Z}^3$ to $\mathbb{C}$ and define*

$$g(k_1,k_2,k_3,k_4) := \sum_{(l_1,l_2,l_3)}^{(k_1,k_2,k_3,k_4)} f(l_1,l_2,l_3).$$

Then

$$\begin{aligned} T'_{2,3}\, g(k_1,k_2,k_3,k_4) &= \\ & -\frac{1}{2}\left(\sum_{l_1=k_2+1}^{k_3} \sum_{l_2=k_2+1}^{k_3} \sum_{l_3=k_2}^{k_4} T'_{1,2}\, f(l_1,l_2,l_3) + \sum_{l_1=k_1}^{k_2+1} \sum_{l_2=k_2}^{k_3-1} \sum_{l_3=k_2}^{k_3-1} T'_{2,3}\, f(l_1,l_2,l_3) \right) \\ & + \frac{1}{2}\left(\sum_{l_1=k_2}^{k_3-1} \sum_{l_2=k_2}^{k_3-1} \Delta_2(\mathrm{id} + E_1) T'_{1,2}\, f(l_1,l_2,k_2) - \sum_{l_2=k_2}^{k_3-1} \sum_{l_3=k_2}^{k_3-1} \Delta_2(\mathrm{id} + E_3) T'_{2,3}\, f(k_2+1,l_2,l_3) \right) \\ & + \frac{1}{2}\Big(T'_{1,2}\, f(k_2,k_2,k_2+1) - T'_{1,2}\, f(k_2,k_2,k_3+1) + T'_{2,3}\, f(k_2,k_2,k_2) - T'_{2,3}\, f(k_3,k_2,k_2) \Big) \\ & \qquad - T'_{1,2}\, f(k_2,k_3,k_2+1) - T'_{2,3}\, f(k_2,k_2,k_3). \end{aligned}$$

Moreover, for a function $h(l_1,l_2)$ on $\mathbb{Z}^2$,

$$T'_{1,2} \left(\sum_{(l_1,l_2)}^{(k_1,k_2,k_3)} h(l_1,l_2) \right) (k_1,k_2,k_3) = -\frac{1}{2} \sum_{l_1=k_1}^{k_2-1} \sum_{l_2=k_1}^{k_2-1} T'_{1,2}\, h(l_1,l_2).$$

This proves the statement preceding the lemma for $n = 2, 3$. It can easily be extended to general n by deriving a merging rule for the recursion (2.1). For this purpose we need another operator. Let $f(x,z)$ be a function on $\mathbb{Z}^2$. Then the operator $I^y_{x,z}$ transforms $f(x,z)$ into a function on $\mathbb{Z}$ by

$$I^y_{x,z} f(x,z) := f(y-1,y) + f(y,y+1) - f(y-1,y+1) = V_{x,z} f(x,z)\big|_{x=y=z}.$$

With this definition we have

$$\sum_{(l_1,\ldots,l_{n-1})}^{(k_1,\ldots,k_n)} A(l_1,\ldots,l_{n-1}) = I^{k_i}_{k'_i,k''_i} \sum_{(l_1,\ldots,l_{i-1})}^{(k_1,\ldots,k_{i-1},k'_i)} \sum_{(l_i,\ldots,l_{n-1})}^{(k''_i,k_{i+1},\ldots,k_n)} A(l_1,\ldots,l_n).$$

If we combine Lemma 3 with this merging rule we obtain formulas for general n. These formulas imply the following corollary.

Corollary 1. *Suppose $A(l_1,\ldots,l_n)$ is a function on $\mathbb{Z}^n$ with $T'_{l_i,l_{i+1}} A(l_1,\ldots,l_n) = 0$ for all i, $1 \le i < n$. Then*

$$T'_{k_i,k_{i+1}} \left(\sum_{(l_1,\ldots,l_n)}^{(k_1,\ldots,k_{n+1})} A(l_1,\ldots,l_n) \right) (k_1,\ldots,k_{n+1}) = 0$$

for all i, $1 \le i \le n$.

We come back to $\alpha(n; k_1,\ldots,k_n)$. By induction with respect to n we conclude that $T'_{k_i,k_{i+1}} \alpha(n; k_1,\ldots,k_n) = 0$ for all i, $1 \le i < n$, if $n \ge 2$. (Note that $\alpha(2; k_1, k_2) = k_2 - k_1 + 1$.) Thus $T_{k_i,k_{i+1}} \alpha(n; k_1,\ldots,k_n) = 0$ for all i. Therefore, by Lemma 2 and by induction with respect to n, the polynomial $\alpha(n; k_1,\ldots,k_n)$ is of degree no greater than $n-1$ in every k_i.

In the following we demonstrate that the property that $T'_{k_i,k_{i+1}} \alpha(n; k_1,\ldots,k_n) = 0$ for all i is not only fundamental for the computation of the polynomial's degree but already determines $\alpha(n; k_1,\ldots,k_n)$ up to a multiplicative constant. Observe that $T'_{k_i,k_{i+1}} A(k_1,\ldots,k_n) = 0$ is equivalent with the fact that $(\mathrm{id} + E_{k_{i+1}} \Delta_{k_i}) A(k_1,\ldots,k_n)$ is antisymmetric in k_i and k_{i+1}. In the following lemma we characterize polynomials $A(k_1,\ldots,k_n)$ with the property that $(\mathrm{id} + E_{k_{i+1}} \Delta_{k_i}) A(k_1,\ldots,k_n)$ is antisymmetric in k_i and k_{i+1} for all i.

Lemma 4. *Let $A(k_1,\ldots,k_n)$ be a polynomial in $(k_1,\ldots,k_n)$. Then*

$$(\mathrm{id} + E_{k_{i+1}} \Delta_{k_i}) A(k_1,\ldots,k_n)$$

is antisymmetric in k_i and k_{i+1} for all i, $1 \le i \le n-1$, if and only if

$$\left(\prod_{1 \le p < q \le n} (\mathrm{id} + E_{k_q} \Delta_{k_p}) \right) A(k_1,\ldots,k_n)$$

is antisymmetric in $k_1,\ldots,k_n$.

Using this lemma we see that

$$\left(\prod_{1 \le p < q \le n} (\mathrm{id} + E_{k_q} \Delta_{k_p}) \right) \alpha(n; k_1,\ldots,k_n) \tag{3.1}$$

is an antisymmetric polynomial in $k_1,\dots,k_n$. The product of shift operators does not increase the polynomial's degree and thus the degree of (3.1) in every k_i is no greater than $n-1$. Every antisymmetric function in $k_1,\dots,k_n$ is a multiple of $\prod_{1\le i<j\le n}(k_j-k_i)$ and since this product is of degree $n-1$ in every k_i, the expression in (3.1) is equal to $C\cdot\prod_{1\le i<j\le n}(k_j-k_i)$, where C is a rational constant. Therefore,

$$\alpha(n;k_1,\dots,k_n)=\left(\prod_{1\le p<q\le n}\frac{1}{\mathrm{id}+E_{k_q}\Delta_{k_p}}\right)C\prod_{1\le i<j\le n}(k_j-k_i).$$

It is not too hard to show that the coefficient of $k_1^0k_2^1\dots k_n^{n-1}$ in $\alpha(n;k_1,\dots,k_n)$ is $C=\prod_{1\le i<j\le n}\frac{1}{j-i}$. Consequently,

$$\alpha(n;k_1,\dots,k_n)=\left(\prod_{1\le p<q\le n}\frac{1}{\mathrm{id}+E_{k_q}\Delta_{k_p}}\right)\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}. \tag{3.2}$$

We need a final lemma in order to derive Theorem 1 from that.

Lemma 5. *Let $P(X_1,\dots,X_n)$ be a polynomial in $(X_1,\dots,X_n)$ over $\mathbb{C}$ which is symmetric in $(X_1,\dots,X_n)$. Then*

$$P(E_{k_1},\dots,E_{k_n})\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}=P(1,\dots,1)\cdot\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}.$$

Observe that $\prod_{1\le p,q\le n}(1+X_q(X_p-1))$ is symmetric in $(X_1,\dots,X_n)$. Thus, by Lemma 5,

$$\prod_{1\le p,q\le n}\left(\mathrm{id}+E_{k_q}\Delta_{k_p}\right)\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}=\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}.$$

Therefore, by (3.2),

$$\begin{aligned}\alpha(n;k_1,\dots,k_n)&=\left(\prod_{1\le p<q\le n}\frac{1}{\mathrm{id}+E_{k_q}\Delta_{k_p}}\right)\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}\\&=\left(\prod_{1\le p<q\le n}\frac{1}{\mathrm{id}+E_{k_q}\Delta_{k_p}}\right)\left(\prod_{1\le p,q\le n}\left(\mathrm{id}+E_{k_q}\Delta_{k_p}\right)\right)\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}\\&=\left(\prod_{1\le p<q\le n}\left(\mathrm{id}+E_{k_p}\Delta_{k_q}\right)\right)\prod_{1\le i<j\le n}\frac{k_j-k_i}{j-i}\end{aligned}$$

and this completes the proof of Theorem 1.

4. Some further projects

In this section we list some further projects around the formula given in Theorem 1 we plan to pursue.

(1) A natural question to ask is whether it is possible to derive the formula for the number of $n \times n$ alternating sign matrices (1.1) from Theorem 1, i.e. to show that

$$\left[\left(\prod_{1 \le p < q \le n} (\mathrm{id} + E_{k_q} \Delta_{k_p}) \right) \prod_{1 \le i < j \le n} \frac{k_j - k_i}{j - i} \right] \Bigg|_{(k_1, k_2, \ldots, k_n) = (1, 2, \ldots, n)} = \prod_{j=0}^{n-1} \frac{(3j+1)!}{(n+j)!}$$

More general, one could try to reprove the refined alternating sign matrix theorem [11], which states that the number of $n \times n$ alternating sign matrices in which the unique 1 in the top row is in the k-th column is given by

$$\frac{(k)_{n-1} (1 + n - k)_{n-1}}{(n-1)!} \prod_{j=1}^{n-1} \frac{(3j-2)!}{(n+j-1)!}. \tag{4.1}$$

An analysis of the correspondence between alternating sign matrices and monotone triangles shows that $\alpha(n-1; 1, 2, \ldots, k-1, k+1, \ldots, n)$ is the number of $n \times n$ alternating sign matrices in which the unique 1 in the bottom row is in the k-th column and this is by symmetry equal to (4.1). This could be a consequence of an even more general theorem: computer experiments suggest that there are other $(k_1, k_2, \ldots, k_n) \in \mathbb{Z}^n$ "near" $(1, 2, \ldots, n)$ for which $\alpha(n; k_1, \ldots, k_n)$ has small prime factors. Small prime factors are an indication for a simple product formula. A similar phenomenon can be observed for some $(k_1, k_2, \ldots, k_n) \in \mathbb{Z}^n$ "near" $(1, 3, \ldots, 2n-1)$. It is not too hard to see that $\alpha(n; 1, 3, \ldots, 2n-1)$ is the number of $(2n+1) \times (2n+1)$ alternating sign matrices, which are symmetric with respect to reflection along the vertical axis. Kuperberg [6] showed that the number of these objects is given by

$$\frac{n!}{(2n)! 2^n} \prod_{j=1}^{n} \frac{(6j-2)!}{(2n+2j-1)!}.$$

(2) Let $\beta(n; k_1, \ldots, k_n)$ denote the number of monotone triangles with prescribed bottom row $(k_1, \ldots, k_n)$ that are strictly increasing in southeast direction. With this notation, Theorem 1 states that

$$\alpha(n; k_1, \ldots, k_n) = \left(\prod_{1 \le p < q \le n} (\mathrm{id} + E_{k_p} \Delta_{k_q}) \right) \beta(n; k_1, \ldots, k_n). \tag{4.2}$$

It would be interesting to find a bijective proof of this formula in the following sense: if we expand the product of operators on the left hand side we obtain a sum of expressions of the form

$$E_{k_1}^{a_1} E_{k_2}^{a_2} \ldots E_{k_n}^{a_n} \Delta_{k_1}^{b_1} \Delta_{k_2}^{b_2} \ldots \Delta_{k_n}^{b_n} \beta(n; k_1, \ldots, k_n)$$

with $a_i, b_i \in \{0, 1, 2, \ldots\}$. We can interpret these expressions as sums and differences of cardinalities of certain subsets of the set of monotone triangles with n rows. For instance,

$$\Delta_{k_q} \beta(n; k_1, \ldots, k_n)$$

is the number of monotone triangles that are strictly increasing in southeast direction and with bottom row $(k_1, \dots, k_q + 1, \dots, k_n)$ such that the $(q-1)$-st part of the $(n-1)$-st row is equal to k_q minus the number of monotone triangles that are strictly increasing in southeast direction and with bottom row $(k_1, \dots, k_n)$ such that the q-th part of the $(n-1)$-st row is equal to k_q. In order to prove (4.2), one has to show that these cardinalities add up to the number of monotone triangles.

(3) This is more a remark than another project: to prove Theorem 1 I have more or less carried out an analysis of the recursion (2.1). I originally started this analysis when considering a somehow reversed situation: let an (r,n) monotone trapezoid be a monotone triangle with the top $n-r$ rows cut off and bottom row $(1,2,\dots,n)$. Let $\gamma(r,n;k_1,\dots,k_{n-r+1})$ denote the number of (r,n) monotone trapzoids with prescribed top row $(k_1,\dots,k_{n-r+1})$. In particular, $\gamma(n,n;k)$ is the number of monotone triangles with n rows, bottom row $(1,2,\dots,n)$ and k as entry in the top row. In the bijection between alternating sign matrices and monotone triangles, the entry in the top row of the monotone triangle corresponds to the column of the unique 1 in the first row of the alternating sign matrix. Thus, $\gamma(n,n;k)$ must be equal to (4.1). On the other hand, we can also use (2.1) to compute $\gamma(r,n;k_1,\dots,k_{n-r+1})$: $\gamma(1,n;k_1,\dots,k_n)=1$ and

$$\gamma(r,n;k_1,\dots,k_{n-r+1}) = \sum_{(l_1,\dots,l_{n-r+2})}^{(1,k_1,\dots,k_{n-r+1},n)} \gamma(r-1,n;l_1,\dots,l_{n-r+2}).$$

With this extended definition, $\gamma(n,n;k)$ is a polynomial in k. In the following we list it for $n=1,2,\dots,6$.

$$\begin{aligned}
\gamma(1,1;k) &= 1\\
\gamma(2,2;k) &= -1+3\,k-k^2\\
\gamma(3,3;k) &= \frac{1}{12}(48-92\,k+103\,k^2-40\,k^3+5\,k^4)\\
\gamma(4,4;k) &= \frac{1}{72}(-2160+5910\,k-5407\,k^2+2940\,k^3\\
&\quad -919\,k^4+150\,k^5-10\,k^6)\\
\gamma(5,5;k) &= \frac{1}{1440}(584640-1644072\,k+1970008\,k^2\\
&\quad -1211172\,k^3+456863\,k^4-111708\,k^5\\
&\quad +17462\,k^6-1608\,k^7+67\,k^8)\\
\gamma(6,6;k) &= \frac{1}{7560}(-73316880+225502200\,k\\
&\quad -284097336\,k^2+204504097\,k^3\\
&\quad -91897169\,k^4+27466950\,k^5\\
&\quad -5651016\,k^6+805518\,k^7\\
&\quad -77646\,k^8+4655\,k^9-133\,k^{10})
\end{aligned}$$

Unfortunately, these polynomials are not equal to (4.1). (For instance, they do not factor over $\mathbb{Z}$.) They only coincide on the combinatorial range $\{1, 2, \ldots, n\}$ of k. However, it might still be possible to compute $\gamma(n, n; k)$ for general n.

Strikingly the degree of $\gamma(n, n; k)$ in k is $2n - 2$ as the degree of (4.1). This linear growth is again unexpected because the application of (2.1) can more than double a polynomial's degree, see Section 2. However, one can use Lemma 2 and an extension of Lemma 3 to show that, more generally, the degree of $\gamma(r, n; k_1, \ldots, k_{n-r+1})$ is $2r - 2$ in every k_i.

(4) Finally we have started to investigate a q-version of the formula in Theorem 1, i.e. a weighted enumeration of monotone triangles with prescribed bottom row $(k_1, \ldots, k_n)$ which reduces to our formula as q tends to 1.

References

[1] D. M. Bressoud, *Proofs and Confirmations, The Story of the Alternating Sign Matrix Conjecture*, Cambridge University Press, Cambridge, 1999.

[2] I. Fischer, A method for proving polynomial enumeration formulas, preprint, `math.CO/0301103`, to appear in *J. Combin. Theory Ser. A*

[3] I. Fischer, Another refinement of the Bender-Knuth (ex-)Conjecture, preprint, `math.CO/0401235`, to appear in *European J. Combin.*

[4] I. Fischer, The number of monotone triangles with prescribed bottom row, preprint, `math.CO/0501102`, to appear in *Adv. Appl. Math.*

[5] G. Kuperberg, Another proof of the alternating sign matrix conjecture, *Internat. Math. Res. Notices* **1996** (1996), 139 – 150.

[6] G. Kuperberg, Symmetry classes of alternating sign matrices under one roof, *Ann. of Math (2)* **156** (2002), no. 3, 835 – 866.

[7] W. H. Mills, D. P. Robbins and H. Rumsey, Alternating sign matrices and descending plane partitions, *J. Combin. Theory Ser. A* **34** (1983), no. 3, 340 – 359.

[8] D. P. Robbins and H. Rumsey, Determinants and alternating sign matrices, *Adv. in Math.* **62** (1986), no. 2, 169 – 184.

[9] R. P. Stanley, Enumerative combinatorics, vol. 2, Cambridge University Press, 1999.

[10] D. Zeilberger, Proof of the alternating sign matrix conjecture, The Foata Festschrift, *Electron. J. Combin.* **3** (1996), no. 2, R13, 84 pp. (electronic)

[11] D. Zeilberger, Proof of the refined alternating sign matrix conjecture, *New York J. Math.* **2** (1996), 59 – 68, electronic.

A Gessel-Viennot-Type Method for Cycle Systems

Christopher R. H. Hanusa

ABSTRACT. We introduce a new determinantal method to count cycle systems in a graph that generalizes Gessel and Viennot's determinantal method on path systems. The presented method gives new insight into the enumeration of domino tilings of Aztec diamonds, Aztec pillows, and other related regions.

1. Introduction

In this article, we present an analogue of the Gessel-Viennot method for counting cycle systems on a type of directed graph we call a hamburger graph. A *hamburger graph* H is made up of two acyclic graphs G_1 and G_2 and a connecting edge set E_3 with the following properties. The graph G_1 has k distinguished vertices $\{v_1, \cdots, v_k\}$ with directed paths from v_i to v_j only if $i < j$. The graph G_2 has k distinguished vertices $\{w_{k+1}, \cdots, w_{2k}\}$ with directed paths from w_i to w_j only if $i > j$. The edge set E_3 connects each vertex v_i to vertex w_{k+i} and vice versa. (See Figure 1 for a visualization.) Hamburger graphs arise naturally in the study of Aztec diamonds, as explained in Section 3.

The Gessel-Viennot method was introduced in [**4, 5**], and has its roots in works by Karlin and McGregor [**7**] and Lindström [**9**]. A nice exposition of the method is given in the article by Aigner [**1**]. The Gessel-Viennot method is a determinantal method to count path systems in an acyclic directed graph G with k sources s_i and k sinks t_j for $1 \le i, j \le k$. A *path system* $\mathcal{P}$ is a collection of k vertex-disjoint paths from s_i to $t_{\sigma(i)}$ for some $\sigma \in S_k$ (where S_k is the symmetric group on k elements).

1991 *Mathematics Subject Classification.* Primary 05B45, 05C30; Secondary 05A15, 05B20, 05C38, 05C50, 05C70, 11A51, 11B83, 15A15, 15A36, 52C20.

Key words and phrases. cycle system, path system, Aztec diamond, Aztec pillow, Schröder numbers.

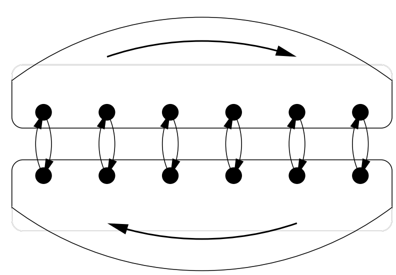


FIGURE 1. A hamburger graph.

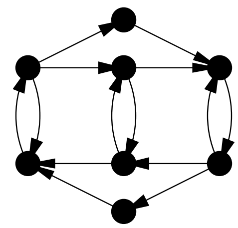


FIGURE 2. A simple hamburger graph H

Call a path system $\mathcal{P}$ *positive* if the sign of this permutation σ satisfies $\operatorname{sgn}(\sigma) = +1$ and *negative* if $\operatorname{sgn}(\sigma) = -1$. Let p^+ be the number of positive path systems and p^- be the number of negative path systems.

Corresponding to this graph G is a $k \times k$ matrix $A = (a_{ij})$ where a_{ij} is the number of paths from s_i to t_j in G. The result of Gessel and Viennot states that $\det A = p^+ - p^-$. For some applications of the Gessel-Viennot method, see [**1, 3, 4, 10**].

This article concerns the following determinantal method for counting cycle systems in a hamburger graph H. A *cycle system* $\mathcal{C}$ is a collection of vertex-disjoint directed cycles in H. Let ℓ be the number of edges in $\mathcal{C}$ that travel from G_2 to G_1 and let m be the number of cycles in $\mathcal{C}$.

Call a cycle system *positive* if $(-1)^{\ell+m} = +1$ and *negative* if $(-1)^{\ell+m} = -1$. Let c^+ be the number of positive cycle systems and c^- be the number of negative cycle systems.

Corresponding to each hamburger graph H is a $2k \times 2k$ block matrix M_H of the form

$$M_H = \begin{bmatrix} A & I_k \\ -I_k & B \end{bmatrix},$$

where the upper triangular matrix $A = (a_{ij})$ represents the number of paths from v_i to v_j in G_1 and the lower triangular matrix $B = (b_{ij})$ represents the number of paths from w_{k+i} to w_{k+j} in G_2. This matrix M_H is referred to as a *hamburger matrix.*

THEOREM 1.1 (The Hamburger Theorem). *If H is a hamburger graph, then* $\det M_H = c^+ - c^-$.

Notice that if the graphs G_1 and G_2 are planar with respect to their embeddings in H, each cycle must use exactly one edge from G_2 to G_1 so that the sign of every cycle system is $+1$. This implies the following corollary.

COROLLARY 1.2. *If H is a hamburger graph such that both G_1 and G_2 are planar,* $\det M_H = c^+$.

The following simple example serves to guide us. Consider the two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, where $V_1 = \{v_1, v_2, v_3\}$, $V_2 = \{w_4, w_5, w_6\}$, $E_1 = \{v_1 \to v_2, v_2 \to v_3, v_1 \to v_3\}$, and $E_2 = \{w_6 \to w_5, w_5 \to w_4, w_6 \to w_4\}$. Our hamburger graph H will be the union of G_1, G_2, and the edge set E_3 consisting of edges $e_i : v_i \to w_{k+i}$ and $e_i' : w_{k+i} \to v_i$. In this example we have that $k = 3$. Figure 2 gives the graphical representation of H.

In terms of this example, the hamburger matrix M_H equals

$$M_H = \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 2 & 1 & 1 \end{bmatrix}.$$

The determinant of M_H is 17, corresponding to the seventeen cycle systems (each with sign $+1$) in Figure 3.

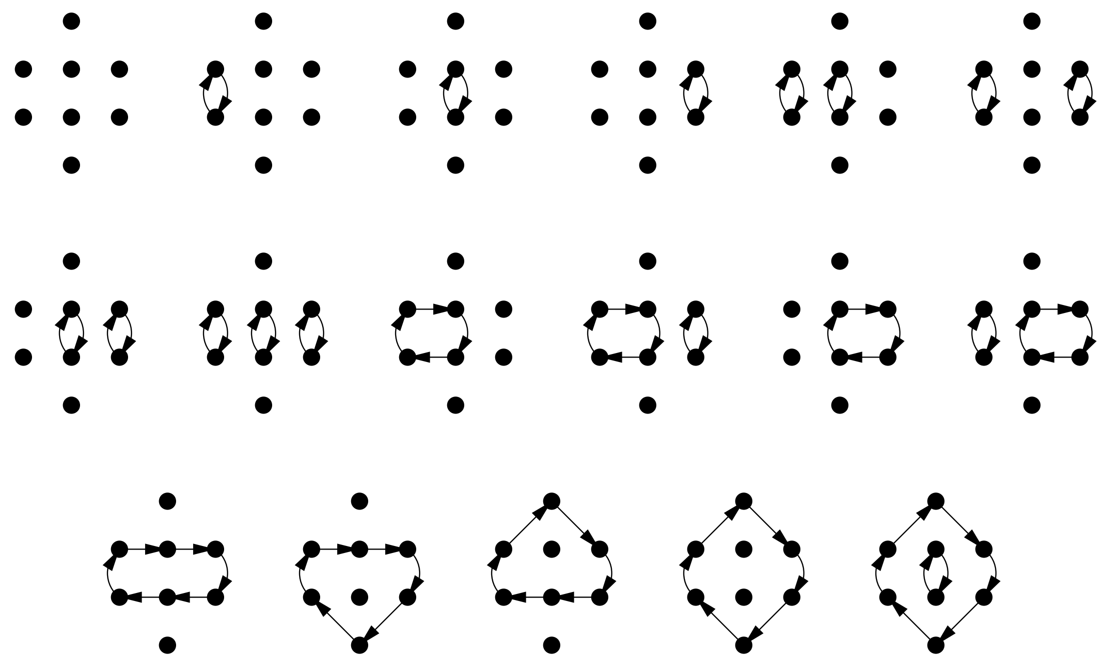


FIGURE 3. The seventeen cycle systems for the hamburger graph in Figure 2

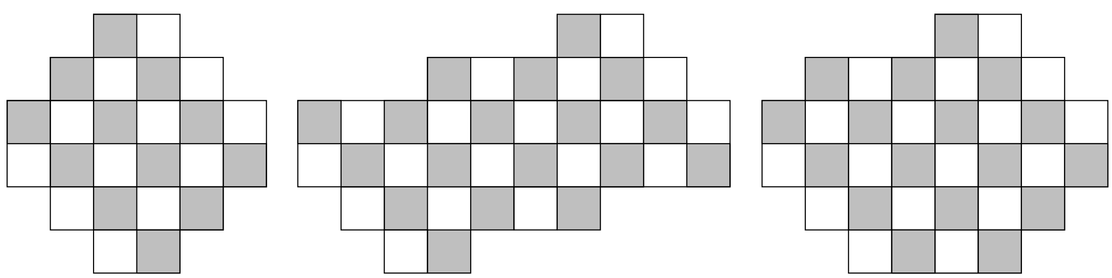


FIGURE 4. Examples of an Aztec diamond, an Aztec pillow, and a generalized Aztec pillow

The graph that inspires the definition of a hamburger graph comes from the work of Brualdi and Kirkland [**2**] in which they give a new proof that the number of tilings of the Aztec diamond is $2^{n(n+1)/2}$.

An *Aztec diamond*, denoted AD_n, is a union of the $2n(n+1)$ unit squares with integral vertices (x, y) such that $|x| + |y| \leq n + 1$. An *Aztec pillow*, as it was initially presented in [**12**], is also a rotationally symmetric region in the plane. On the top left diagonal however, the steps are composed of three squares to the right for every square up. Analogous to an Aztec diamond, we denote the Aztec pillow with $2n$ squares in each of the central rows as AP_n.

We also introduce the idea of a *generalized Aztec pillow*. This will be a horizontally convex and vertically convex region such that each of the steps along each diagonal is composed of an odd number of squares horizontally for every one square vertically. A key fact is that any generalized Aztec pillow can be recovered from a large enough Aztec diamond by the placement of horizontal dominoes. See Figure 4 for examples of an Aztec diamond, an Aztec pillow, and a generalized Aztec pillow.

Brualdi and Kirkland's proof of the Aztec diamond formula starts with the $n(n+1) \times n(n+1)$ Kasteleyn-Percus matrix (see [**8, 11**]) of a particular digraph to enumerate its cycle systems. The

Hamburger Theorem allows us to now enumerate domino tilings by taking the determinant of a $2n \times 2n$ matrix. An analogous reduction in determinant size occurs in all regions to which this theorem applies. In addition, whereas Kasteleyn theory applies only to planar graphs, there is no restriction of planarity for hamburger graphs. For this reason, the Hamburger Theorem introduces a new counting method for cycle systems in some non-planar graphs.

2. Idea of the Proof of the Hamburger Theorem

We now present a sketch of the proof of the Hamburger Theorem. Like the proof of the Gessel-Viennot method, the proof of the Hamburger Theorem hinges on terms canceling in the permutation decomposition of the determinant of M_H. For this proof, we must allow *generic cycle systems*, those unions of directed cycles which may not be vertex-disjoint, since they can appear in the permutation decomposition of the determinant of M_H.

If a cycle system arising from the permutation decomposition of the determinant is M_H is not vertex-disjoint, there are four possibilities. First, the cycle system may contain a cycle that is self-intersecting. Second, the cycle system may have two intersecting cycles, neither of which is a 2-cycle.

LEMMA 2.1. *If a generic cycle system $\mathcal{C}$ contains a cycle that is self-intersecting or contains two intersecting cycles that are not 2-cycles, $\mathcal{C}$ belongs to one well-defined family $\mathcal{F}$ of generic cycle systems of these types that cancel each other in the permutation expansion of the determinant of M_H.*

If either of these first two properties hold, there is some notion of order to determine which holds first. This vertex of first intersection is the basis for a family of cycle systems that all intersect for the first time at this vertex. We have that this family contributes a net zero weight in the determinantal expansion of M_H.

When calculating the number of cycle systems, we notice that the determinantal expansion of M_H has various summands that may represent the same cycle system. Consider the second cycle system in the third row of Figure 3. Since the solitary directed cycle visits vertices v_1, v_2, v_3, w_6, and w_4 in order, this cycle contributes a non-zero weight in the permutation expansion of the determinant corresponding to the permutation cycle (12364) in S_6. Notice that this cycle will also contribute a non-zero weight in the permutation expansion of the determinant corresponding to the permutation cycle (1364) since the cycle follows a path from v_1 to v_3 (by way of v_2), returning to v_1 via w_6 and w_4. Because of this ambiguity, we must introduce the idea of a *minimal permutation cycle* and a *minimal cycle system*, and realize that the determinant of M_H counts minimal cycle systems. The minimal permutation cycle for this second cycle system in the third row of Figure 3 is (1364).

In the case when neither the first nor second properties hold and that the generic cycle system is not vertex-disjoint or not minimal, at least one of two additional properties hold. The third property is that two cycles intersect and one of the cycles is a 2-cycle. The fourth property is that the cycle system may not be minimal.

LEMMA 2.2. *Let $\mathcal{C}$ be a generic cycle system that does not satisfy the conditions of Lemma 2.1. If $\mathcal{C}$ is not minimal or if it contains a cycle that intersects with a 2-cycle, $\mathcal{C}$ belongs to one well-defined family $\mathcal{F}$ of generic cycle systems of these types that cancel each other in the permutation expansion of the determinant of M_H.*

We can determine where the 2-cycle intersections and non-minimalities occur, and corresponding to this set of violations, we create a family of cycle systems each with this set of violations. We have that this family contributes a net zero weight in the determinantal expansion of M_H.

LEMMA 2.3. *The generic cycle system $\mathcal{C}$ is a minimal cycle system with vertex-disjoint cycles if and only if $\mathcal{C}$ does not satisfy the conditions of Lemmas 2.1 and 2.2.*

The cancellation from the above sets of families gives us that only minimal cycle systems contribute to the determinantal expansion of M_H. This contribution is the signed weight of each cycle system, implying that the determinant of M_H exactly equals $c^+ - c^-$. This proves the theorem. •

A weighted version of the Hamburger Theorem also exists. We allow *weights* $\mathrm{wt}(e)$ on edges of the hamburger graph; the simplest weighting which counts of the number of cycle systems assigns $\mathrm{wt}(e) \equiv 1$. We only require that $\mathrm{wt}(e_i)\mathrm{wt}(e_i') = 1$ for all $2 \le i \le k-1$, and do not require this condition for $i = 1$ nor for $i = k$. We now define the $2k \times 2k$ *weighted hamburger matrix* M_H to be the block matrix

$$M_H = \begin{bmatrix} A & D_1 \\ -D_2 & B \end{bmatrix}. \tag{2.1}$$

The upper triangular $k \times k$ matrix $A = (a_{ij})$ represents the sum of the products of the weights of edges over all paths from v_i to v_j in G_1 and the lower triangular $k \times k$ matrix $B = (b_{ij})$ represents the sum of the products of the weights of edges over all paths from w_{k+i} to w_{k+j} in G_2. The diagonal $k \times k$ matrix D_1 has as its entries $d_{ii} = \mathrm{wt}(e_i)$ and the diagonal $k \times k$ matrix D_2 has as its entries $d_{ii} = \mathrm{wt}(e_i')$. Note that when the weights of the edges in E_3 are all 1, these matrices satisfy $D_1 = D_2 = I_k$.

In our hamburger graph H, there are two possible types of cycles. There are k *2-cycles*

$$c : v_i \xrightarrow{e_i} w_{k+i} \xrightarrow{e_i'} v_i$$

and many more *general cycles* in H that alternate between G_1 and G_2. We can think of a cycle of this form as a path P_1 in G_1 connected by an edge $e_{1,1} \in E_3$ to a path Q_1 in G_2, which in turn connects to a path P_2 in G_1 by an edge $e_{1,2}'$, continuing in this fashion until arriving at a final path Q_ℓ in G_2 whose terminal vertex is adjacent to the initial vertex of P_1. We write

$$c : P_1 \xrightarrow{e_{1,1}} Q_1 \xrightarrow{e_{1,2}'} P_2 \xrightarrow{e_{2,1}} \cdots \xrightarrow{e_{\ell,1}} P_\ell \xrightarrow{e_{\ell,2}'} Q_\ell.$$

To each cycle c, we define the weight $\mathrm{wt}(c)$ of c to be the product of all the weights of the edges traversed by c:

$$\mathrm{wt}(c) = \prod_{e \in c} \mathrm{wt}(e).$$

We define a *weighted cycle system* to be a collection $\mathcal{C}$ of m vertex-disjoint cycles. We again define the *sign* of a weighted cycle system to be $\mathrm{sgn}(\mathcal{C}) = (-1)^{\ell+m}$, where ℓ is the total number of edges from G_2 to G_1 in $\mathcal{C}$. We say that a weighted cycle system $\mathcal{C}$ is *positive* if $\mathrm{sgn}(\mathcal{C}) = +1$ and *negative* if $\mathrm{sgn}(\mathcal{C}) = -1$.

For a hamburger graph H, let c^+ be the sum of the weights of positive weighted cycle systems, and c^- be the sum of the weights of negative weighted cycle systems.

THEOREM 2.4 (The weighted Hamburger Theorem). *The determinant of the weighted hamburger matrix M_H equals $c^+ - c^-$.*

3. Applications of the Hamburger Theorem

We discuss first the application of the Hamburger Theorem in the case when the region is an Aztec diamond, mirroring results of Brualdi and Kirkland. Then we discuss the results from the case when the region is an Aztec pillow, and lastly we explain how to implement the Hamburger Theorem when our region is any generalized Aztec pillow.

The Hamburger Theorem applies to the enumeration of domino tilings of Aztec diamonds and generalized Aztec pillows. To illustrate this connection, we count domino tilings of the Aztec diamond by enumerating an equivalent quantity, the number of matchings on the dual graph G of the Aztec diamond. The natural matching N of horizontal neighbors in G as exemplified in Figure 5a on AD_4 is a reference point. Given any other matching M on G, such as in Figure 5b, their symmetric difference is a union of cycles in the graph, such as in Figure 5c. When we orient the edges in N from black vertices to white vertices and in M from white vertices to black vertices, the symmetric

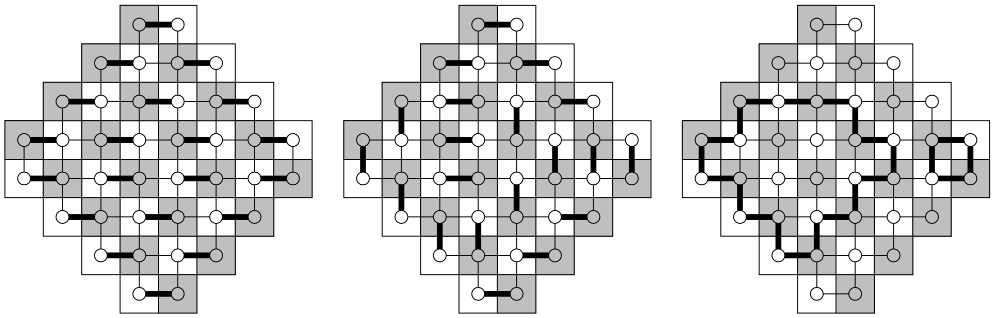


FIGURE 5. The symmetric difference of two matchings gives a cycle system

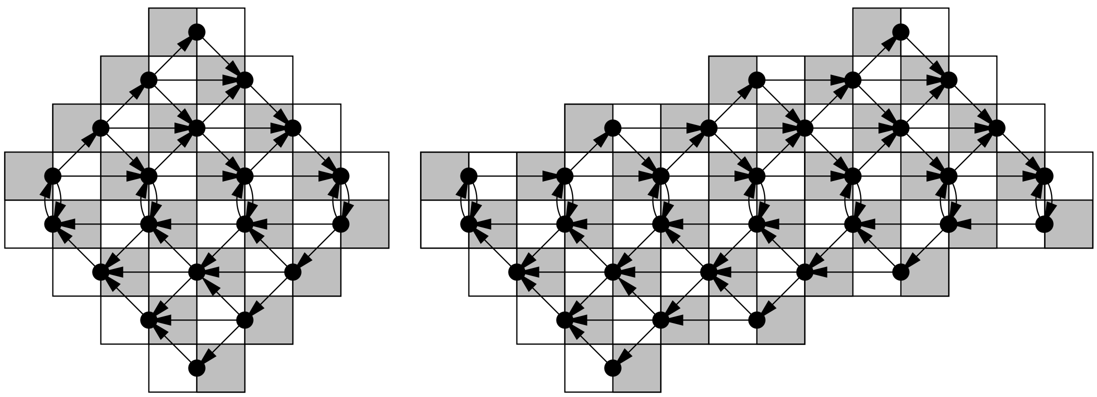


FIGURE 6. The hamburger graph for an Aztec diamond and an Aztec pillow

difference becomes a union of directed cycles. Notice that edges in the upper half of G all go from left to right and the edges in the bottom half of G go from right to left.

Since the edges of N always appear in the cycles, we can contract the edges of N to points and the graph will retain its structure in terms of the cycle systems it produces. This new graph H is of the form in Figure 6a. This argument shows that the number of domino tilings of an Aztec diamond equals the number of cycle systems of this new condensed graph, called the region's *digraph*. In the case of generalized Aztec pillows, the region's digraph is always a hamburger graph.

We wish to concretize this notion of a *digraph* of the Aztec diamond AD_n. Given the natural tiling of an Aztec diamond consisting solely of horizontal dominoes, we place a vertex in the center of every domino. The edges of this digraph are made up of three families of edges. From every vertex in the top half of the diamond, create edges to the east, to the northeast, and to the southeast whenever there is a vertex there. From every vertex in the bottom half of the diamond, form edges to the west, to the southwest, and to the northwest whenever there is a vertex there. Additionally, label the bottom vertices in the top half v_1 through v_n from west to east and the top vertices in the bottom half w_{n+1} through w_{2n}. For all i between 1 and n, create a directed edge from v_i to w_{n+i} and from w_{n+i} to v_i. The result when this construction is applied to AD_4 is a graph of the form in Figure 6a.

THEOREM 3.1. *The digraph of an Aztec diamond is a hamburger graph.*

Since both the upper half of the digraph and the lower half of the digraph are both planar, there are no negative cycle systems. This implies that the determinant of the corresponding hamburger matrix counts exactly the number of cycle systems in the digraph.

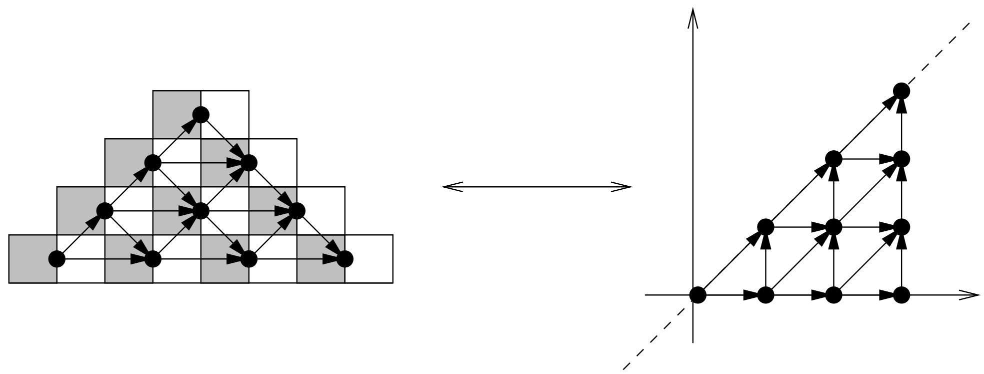


FIGURE 7. The equivalence between paths in D and lattice paths in the first quadrant

To apply Theorem 1.1 to count the number of tilings of AD_n, we need to find the number of paths in the upper half of D from v_i to v_j and the number of paths in the lower half of D from w_{n+j} to w_{n+i}. The key observation is that by the equivalence in Figure 7, we are in effect counting the number of paths from (i, i) to (j, j) using steps of size $(0, 1)$, $(1, 0)$, or $(1, 1)$ that do not pass above the line $y = x$. This is exactly a combinatorial interpretation for the $(j - i)$-th large Schröder number. The large Schröder numbers $s_0, \ldots, s_5$ are $1, 2, 6, 22, 90, 394$, and are referenced as A006318 in the Encyclopedia of Integer Sequences [**13**].

COROLLARY 3.2. *The number of domino tilings of the Aztec diamond AD_n is equal to*

$$\#AD_n = \det \begin{bmatrix} S_n & I_n \\ -I_n & S_n^T \end{bmatrix},$$

where S_n is an upper triangular matrix with the i-th Schröder number on its ith superdiagonal.

For example, when $n = 6$, the matrix S_6 is

$$S_6 = \begin{bmatrix} 1 & 2 & 6 & 22 & 90 & 394 \\ 0 & 1 & 2 & 6 & 22 & 90 \\ 0 & 0 & 1 & 2 & 6 & 22 \\ 0 & 0 & 0 & 1 & 2 & 6 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Brualdi and Kirkland prove a similar determinant formula for the number of tilings of an Aztec diamond in a matrix-theoretical fashion based on the Kasteleyn matrix of the graph H and a Schur complement calculation. The Hamburger Theorem gives a purely combinatorial way to reduce the calculation of the $n(n + 1) \times n(n + 1)$ Kasteleyn determinant to the calculation of a $2n \times 2n$ Hamburger determinant. Following the cues from Brualdi and Kirkland, we can reduce this to an $n \times n$ determinant via a Schur complement calculation.

In the case of the block matrix M_H in Equation 2.1, taking the Schur complement of B in M_H gives that

$$\det M_H = \det B \cdot \det(A + D_1 B^{-1} D_2) = \det(A + D_1 B^{-1} D_2), \tag{3.1}$$

since B is an lower triangular matrix with 1's on the diagonal. In this way, every hamburger determinant can be reduced to a smaller determinant of a Schur complement matrix. In the case of a simple hamburger graph where $D_2 = D_1 = I$, the determinant calculation reduces further to $\det(A + B^{-1})$. Lastly, in the case where the hamburger graph is rotationally symmetric, $B = JAJ$, where J is the exchange matrix. This implies we can write the determinant only in terms of the submatrix A, i.e., $\det(A + JA^{-1}J)$. (Note that $J^{-1} = J$.)

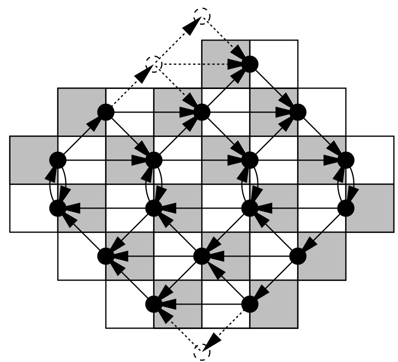


FIGURE 8. The digraph of a generalized Aztec pillow from the digraph of an Aztec diamond

COROLLARY 3.3. *The number of domino tilings of the Aztec diamond AD_n is equal to* $\det(S_n + J_n S_n^{-1} J_n)$*, where J_n is the $n \times n$ exchange matrix.*

In the case of a hamburger graph H, we call this Schur complement $A+B^{-1}$ a *reduced hamburger matrix*. In terms of the Aztec diamond graph example above, we can thus calculate the number of tilings of the Aztec diamond AD_6 as follows. The inverse of S_6 is

$$S_6^{-1} = \begin{bmatrix} 1 & -2 & -2 & -6 & -22 & -90 \\ 0 & 1 & -2 & -2 & -6 & -22 \\ 0 & 0 & 1 & -2 & -2 & -6 \\ 0 & 0 & 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which implies that the determinant of the reduced hamburger matrix

$$M_6 = \begin{bmatrix} 2 & 2 & 6 & 22 & 90 & 394 \\ -2 & 2 & 2 & 6 & 22 & 90 \\ -2 & -2 & 2 & 2 & 6 & 22 \\ -6 & -2 & -2 & 2 & 2 & 6 \\ -22 & -6 & -2 & -2 & 2 & 2 \\ -90 & -22 & -6 & -2 & -2 & 2 \end{bmatrix}$$

gives the number of tilings of AD_6.

Brualdi and Kirkland were the first to find such a determinantal formula for the number of tilings of an Aztec Diamond [**2**]. They were able to calculate the sequence of determinants $\{M_n\}$ using a J-fraction expansion, which only works when matrices are Toeplitz or Hankel.

Since Aztec pillows and generalized Aztec pillows can be created from Aztec diamonds by placement of dominoes, we define the digraph of a generalized Aztec pillow to be the restriction of the digraph of an Aztec diamond to the vertices that are on the interior of the pillow. For a visualization, see the example of Figure 8. Since the generalized Aztec pillow's digraph is a restriction of the Aztec diamond's digraph, we have the following corollary.

COROLLARY 3.4. *The digraph of an Aztec pillow or a generalized Aztec pillow is a hamburger graph.*

Aztec pillows were introduced in part because of an intriguing conjecture for their number of tilings given by Propp [**12**].

CONJECTURE 3.5 (Propp's Conjecture). *The number of tilings of an Aztec pillow AP_n is a larger number squared times a smaller number. We write $\#AP_n = \ell_n^2 s_n$. In addition, depending* *on the parity of n, the smaller number s_n satisfies a simple generating function. For AP_{2m+1}, the generating function is*

$$\sum_{m=0}^{\infty} s_{2m+1}x^m = (5+6x+3x^2-2x^3)/(1-2x-2x^2-2x^3+x^4).$$

while for AP_{2m+2}, the generating function is

$$\sum_{m=0}^{\infty} s_{2m+2}x^m = (5+3x+x^2-x^3)/(1-2x-2x^2-2x^3+x^4),$$

The underlying hope going into the Hamburger Theorem was that it would allow us to prove Propp's Conjecture. We shall see that although we achieve a faster determinantal method to calculate $\#AP_n$, the evaluation of the sequence of said determinants is not able to be calculated explicitly by known methods.

Using the same method as for Aztec diamonds, creating the hamburger graph H for an Aztec pillow gives Figure 6b. Counting the number of paths from v_i to v_j and from w_{k+j} to w_{k+i} in successively larger Aztec pillows gives us the infinite upper-triangular array $S = (s_{i,j})$ of *modified Schröder numbers* defined by the following combinatorial interpretation. Let $s_{i,j}$ be the number of paths from (i,i) to (j,j) using steps of size $(0,1)$, $(1,0)$, and $(1,1)$, not passing above the line $y=x$ nor below the line $y=x/2$. This equivalence is shown in Figure 9. The principal 7×7 minor matrix S_7 of S is

$$S_7 = \begin{bmatrix} 1 & 1 & 2 & 5 & 16 & 57 & 224 \\ 0 & 1 & 2 & 5 & 16 & 57 & 224 \\ 0 & 0 & 1 & 2 & 6 & 21 & 82 \\ 0 & 0 & 0 & 1 & 2 & 6 & 22 \\ 0 & 0 & 0 & 0 & 1 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

THEOREM 3.6. *The number of domino tilings of an Aztec pillow of order n is equal to*

$$\#AP_n = \det \begin{bmatrix} S_n & I_n \\ -I_n & J_nS_nJ_n \end{bmatrix},$$

where S_n is the $n\times n$ principal submatrix of S and J_n is the $n\times n$ exchange matrix.

As in the case of Aztec diamonds, we can calculate the reduced hamburger matrix through a Schur calculation. The inverse of S_6 is

$$S_6^{-1} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & -1 & -2 & -5 \\ 0 & 0 & 1 & -2 & -2 & -5 \\ 0 & 0 & 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the resulting reduced hamburger matrix for AP_6 is

$$M_6 = \begin{bmatrix} 2 & 1 & 2 & 5 & 16 & 57 \\ -2 & 2 & 2 & 5 & 16 & 57 \\ -2 & -2 & 2 & 2 & 6 & 21 \\ -5 & -2 & -2 & 2 & 2 & 6 \\ -5 & -2 & -1 & -2 & 2 & 2 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

This gives us a much faster way to calculate the number of domino tilings of an Aztec pillow than was known previously. We have reduced the calculation of the $O(n^2)\times O(n^2)$ Kasteleyn-Percus determinant to an $n\times n$ reduced hamburger matrix. To be fair, the Kasteleyn-Percus matrix has

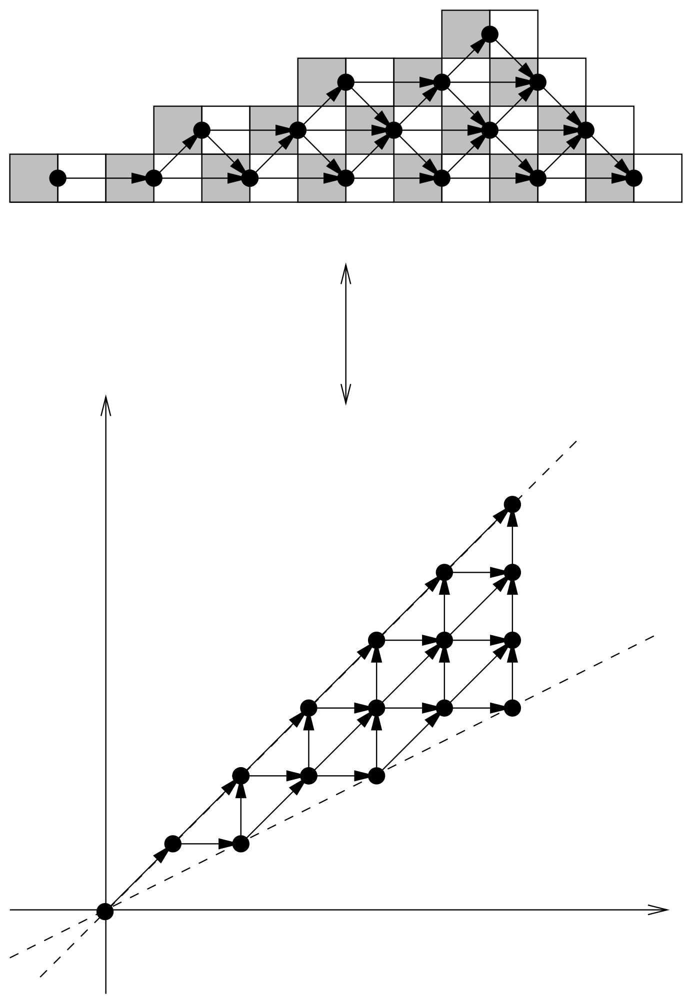


FIGURE 9. The equivalence between paths in D and lattice paths in the first quadrant

-1, 0, and $+1$ entries while the reduced hamburger matrix may have very large entries, which makes running time comparisons difficult theoretically. Experimentally, when calculating the number of domino tilings of AP_{14} using Maple 8.0 on a 447 MHz Pentium III processor, the determinant of the 112×112 Kasteleyn-Percus matrix takes 25.3 seconds while the determinant of the 14×14 reduced hamburger matrix takes less than 0.1 seconds.

Whereas we now have a very understandable determinantal formula for the number of tilings of the region, this does not translate into a proof of Propp's Conjecture because we can not calculate the determinant of the sequence of matrices $\{M_n\}$ explicitly. We can not apply a J-fraction expansion as Brualdi and Kirkland did since the reduced hamburger matrix is not Toeplitz or Hankel.

The Hamburger Theorem allows us to calculate the number of cycle systems in old graphs more quickly and in many new graphs that were inaccessible before, such as non-planar hamburger graphs. Just as Gessel and Viennot's result has found many applications, we hope the Hamburger Theorem to be as useful as well.

References

[1] Martin Aigner. Lattice paths and determinants. In *Computational discrete mathematics*, volume 2122 of *Lecture Notes in Comput. Sci.*, pages 1–12. Springer, Berlin, 2001.

[2] Richard Brualdi and Stephen Kirkland. Aztec diamonds and digraphs, and Hankel determinants of Schröder numbers, available at http://www.math.wisc.edu/~brualdi/aztec2.pdf 2003.

[3] N. Destainville. Entropy and boundary conditions in random rhombus tilings. *J. Phys. A*, 31(29):6123–6139, 1998. arXiv:cond-mat/9804062.

[4] Ira Gessel and Xavier G. Viennot. Binomial determinants, paths, and hook length formulae. *Adv. in Math.*, 58:300–321, 1985.

[5] Ira Gessel and Xavier G. Viennot. Determinants, paths, and plane partitions. 1989. preprint, available at http://www.cs.brandeis.edu/~ira/papers/pp.pdf.

[6] W. Jockusch. Perfect matchings and perfect squares. *J. Combin. Theory Ser. A*, 67(1):100–115, 1994.

[7] Samuel Karlin and James McGregor. Coincidence probabilities. *Pacific J. Math.*, 9:1141–1164, 1959.

[8] P. W. Kasteleyn. The statistics of dimers on a lattice i. the number of dimer arrangements on a quadratic lattice. *Physica*, 27:1209–1225, 1961.

[9] B. Lindström. On the vector representations of induced matroids. *Bull. London Math. Soc.*, 5:85–90, 1973.

[10] M. E. Mays and Jerzy Wojciechowski. A determinant property of Catalan numbers. *Discrete Math.*, 211(1-3):125–133, 2000.

[11] J. Percus. One more technique for the dimer problem. *J. Math. Phys.*, 10:1881–1884, 1969.

[12] James Propp. Enumeration of matchings: problems and progress. In *New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97)*, volume 38 of *Math. Sci. Res. Inst. Publ.*, pages 255–291. Cambridge Univ. Press, Cambridge, 1999. arXiv:math.CO/9904150.

[13] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences, http://www.research.att.com/~njas/sequences/.

Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, USA
E-mail address: `chanusa@math.washington.edu`

CW-spheres encoded by polyspherical coordinates

Gábor Hetyei*
Department of Mathematics and Statistics
UNC Charlotte
Charlotte, NC 28223

April 30, 2005

Abstract

We construct spherical CW-complexes whose face structure may be conveniently described using a system of polyspherical coordinates introduced by Vilenkin, Kuznetsov and Smorodinskii. We prove that these complexes may be constructed by repeated use of CW-suspension, free join, and edge subdivision. We show that all CW-spheres constructed this way have a non-negative cd-index and thus verify Stanley's famous conjecture. Among the particular examples we find a new class of partially ordered sets whose order complexes encode the derivative polynomials for secant of even degree. The geometric constructions presented here generalize CW-complexes whose flag numbers are suitable to encode systems of orthogonal polynomials.

Résumé

Nous construisons des sphères CW dont la structure de faces se décrit d'une manière convenable en utilisant des coordonnées polysphériques de Vilenkin, Kuznetsov, et Smorodinskii. Nous montrons que ces complexes peuvent être construits récursivement en utilisant des suspensions des complexes CW, des joins libres et des sous-division des arêtes. Nous démontrons ques tous nos sphères ont un indexe cd positif, en accord avec la fameuse conjecture de Stanley. Parmis les examples particulières nous retrouvons un nouveau class d'ensembles ordonnés dont le complex des chaînes croissants code les polynômes dérivés pour la fonction secant de degré pair. Nos constructions géométriques généralisent des complexes CW dont les nombres de drapeaux codent des systemes des polynômes orthogonaux.

Introduction

In a recent paper [13] the present author introduced sequences of CW-spheres whose ce-indices may be transformed into sequences of orthogonal polynomials by sending c into x and e into 1.

*On leave from the Rényi Mathematical Institute of the Hungarian Academy of Sciences.
2000 Mathematics Subject Classification: Primary 05E35; Secondary 06A07, 57Q05
Key words and phrases: partially ordered set, Eulerian, flag, polyspherical coordinates, derivative polynomial.

These complexes were shown to be spherically shellable, and the resulting non-negativity of their cd-index implied by Stanley's [21, Theorem 2.2] induces a new proof for the fact that the true interval of orthogonality of any orthogonal polynomial system $\{Q_n(x)\}_{n=0}^{\infty}$, given by a recurrence formula $Q_n(x) = \nu_n \cdot x \cdot Q_{n-1}(x) - (\nu_n - 1) \cdot Q_{n-2}(x)$ where $\nu_i \geq 2$, is a subset of $[-1, 1]$.

The system of spherical coordinates used in that paper is the simplest example of a system of polyspherical coordinates introduced by Vilenkin, Kuznetsov and Smorodinskii [22] to encode the points of a unit sphere. Each such coordinate system may be described by a rooted binary tree. The spherical coordinate system used in [13] corresponds to the situation when the subtree of interior nodes (=the "small tree") is a rooted path.

In Section 2 we describe the faces of our complexes as intersections of certain lunes and hemispheres, and define our polyspherical complexes by explicitly listing their faces. The fact that our constructions yield CW-spheres may be shown by combining all results of Section 3 where we describe our polyspherical complexes recursively.

Our main result is in Section 4: every polyspherical complex we constructed has a non-negative cd-index. Unlike [13], it seems to be extremely hard to find a proof that uses spherical shelling, the main problem being with the spherical shellability of a CW-complex that arises as the free join of two CW-spheres. Fortunately, the dual version of a result of Ehrenborg and Fox [9] (based on the work of Ehrenborg and Readdy [10]) provides an immediate proof of the fact that the non-negativity of the cd-index is preserved by the free join operation. The same question for CW-suspension is trivial. Finally, edge-subdivision does not necessarily preserve the non-negativity of the cd-index (since it involves changing by a multiple of the cd-index of a proper interval in the associated poset). Fortunately, the face posets of our polyspherical complexes belong to a narrower class of Eulerian posets: not only their cd-index but the cd-index of every interval of the form $[x, \widehat{1}]$ turns out to be non-negative. This fact is easily shown by proving that all poset-operations used preserve this stronger positivity property.

Finally, in Section 5 we focus on a special class of polyspherical complexes, that has about the same "degree of freedom" as the ones studied in [13]: we require the underlying small trees to be strongly binary, and we forbid subdividing the intervals of angles associated to interior nodes. Using the dual version of the type B quasisymmetric functions defined by C.-O. Chow [5], we obtain an explicit formula for their flag f-vectors. We show that substituting x into c and 1 into e in the ce-index of a free join of quadrilaterals yields the derivative polynomials for secant of even degree. These polynomials appear in chain enumeration related to some generalization of the Tchebyshev posets introduced in [12] for the second time (the first connection was noted in [14, Section 9]). The appearance of this second, non-isomorphic connection suggests that the Tchebyshev polynomials of the first and second kind and the derivative polynomials for tangent and secant may be more intimately related at a combinatorial level than we ever thought.

1 Preliminaries

1.1 Graded and Eulerian posets, transformations, flag-enumeration

A partially ordered set P is *graded* if it has a unique minimum $\widehat{0}$, a unique maximum $\widehat{1}$, and a rank function ρ. If P has rank $n+1$ and $S \subseteq \{1, \dots, n\}$, $f_S(P)$ is the number of saturated chains in $P_S = \{x \in P \ : \ \rho(x) \in S\} \cup \{\widehat{0}, \widehat{1}\}$. The vector $(f_S(P) \ : \ S \subseteq \{1, \dots, n\})$ is the *flag f-vector* of P. It has several equivalent encodings. The connection between the flag f-vectors of P, Q, and *direct product* $P \times Q$ is most easily expressed by the *flag quasisymmetric function* $F(P)$ [8], given by

$$F(P) = \lim_{m \longrightarrow \infty} \sum_{\widehat{0}=x_0 \leq x_1 \leq \cdots \leq x_m = \widehat{1}} t_1^{\rho(x_1)-\rho(x_0)} t_2^{\rho(x_2)-\rho(x_1)} \cdots t_m^{\rho(x_m)-\rho(x_{m-1})}. \tag{1}$$

By Ehrenborg's result [8, Proposition 4.4], $F(P \times Q) = F(P) \times F(Q)$. A modified version of $P \times Q$ is the *diamond product* $P \diamond Q := (P \setminus \{\widehat{0}\}) \times (Q \setminus \{\widehat{0}\}) \cup \{\widehat{0}\}$. The analogue of [8, Proposition 4.4], recently discovered by Ehrenborg and Readdy [11], is

$$F_B(P \diamond Q) = F_B(P) \cdot F_B(Q), \tag{2}$$

where $F_B(Q)$ is the *type B quasisymmetric function*

$$F_B(P) = \lim_{m \longrightarrow \infty} \sum_{\widehat{0} < x_0 \leq x_1 \leq \cdots \leq x_m = 1} s^{\rho(x_0)-1} t_1^{\rho(x_1)-\rho(x_0)} t_2^{\rho(x_2)-\rho(x_1)} \cdots t_m^{\rho(x_m)-\rho(x_{m-1})}.$$

Another equivalent encoding of the flag f-vector is the *flag h-vector* $(h_S(P) : S \subseteq \{1, \dots, n\})$ (see [21]), given by $h_S(P) = \sum_{T \subseteq S} (-1)^{|S \setminus T|} f_T(P)$.

A graded poset is *Eulerian* if every interval $[x, y]$ of positive rank in it satisfies $\sum_{x \leq z \leq y} (-1)^{\rho(z)} = 0$. All linear relations holding for the flag f-vector of an arbitrary Eulerian poset of rank n were determined by Bayer and Billera [1]. These linear relations were rephrased by J. Fine as follows (see Bayer and Klapper [2]). For any $S \subseteq \{1, \dots, n\}$ define the non-commutative monomial $u_S = u_1 \dots u_n$ by setting

$$u_i = \begin{cases} b & \text{if } i \in S, \\ a & \text{if } i \notin S. \end{cases}$$

Then the polynomial $\Psi_{ab}(P) = \sum_S h_S u_S$ in non-commuting variables a and b, called the *ab-index of* P, is a polynomial $\Phi_{cd}(P)$ of $c = a + b$ and $d = ab + ba$, called the *cd-index* of P. It was noted by Stanley [21] that the existence of the *cd*-index is equivalent to saying that the *ab*-index rewritten as a polynomial of $c = a + b$ and $e = a - b$ involves only even powers of e. Stanley's conjecture [21, Conjecture 2.1] states that the *cd*-index of any Gorenstein* poset has non-negative coefficients. The *cd*-index form is convenient to calculate the flag f-vector of the *join* $P * Q := (P \setminus \{\widehat{1}_P\}) \cup (Q \setminus \{\widehat{0}_Q\})$ of two Eulerian posets. (We place the elements of Q above the elements of P.) By [21, Lemma 1.1] we have $\Phi_{cd}(P * Q) = \Phi_{cd}(P)\Phi_{cd}(Q)$.

1.2 The Tchebyshev transform of a graded poset

We define the *Tchebyshev transform* $T(P)$ of a graded poset P as follows: we adjoin a new minimum element $\widehat{-1} < \widehat{0}$ to P and set $T(P) = \{(x, y) : x < y,\ x, y \in \{\widehat{-1}\} \cup P\} \cup \{\widehat{1}_{T(P)}\}$. Here $\widehat{1}_{T(P)}$ is the maximum element of $T(P)$, and we set $(x_1, y_1) < (x_2, y_2)$ if either $y_1 < x_2$ or $x_1 = x_2$ and $y_1 < y_2$. It was shown in [12] that this operation yields a graded poset and preserves the Eulerian property. It was observed in [14, Section 9] that substituting x into c and 1 into e in the ce index of the Tchebyshev transform of the Boolean algebra of rank n yields the polynomial $(\sqrt{-1})^n \cdot Q_n(x \cdot \sqrt{-1})$ where $Q_n(x)$ is the *n-th derivative polynomial for secant* given by $\frac{d^n}{du^n} \sec(u) = Q_n(\tan u) \sec u$. Further information on these polynomials may be found in Hoffman's papers [15] and [16]. The study of these polynomials goes back to Krichnamachary and Rao [18] and Knuth and Buckholtz [17]. Further results on the Tchebyshev transform are in [11], [12], and [14].

1.3 Free join and suspension of CW-spheres

Given a regular CW-complex Ω, we obtain a graded poset $P_1(\Omega)$ by adjoining a maximum element $\widehat{1}$ to the face poset of Ω. Such posets are called *CW-posets* and were characterized by Björner [3]. $P_1(\Omega)$ is Eulerian if Ω is a CW-sphere. We use two basic operations on CW-complexes: *free join* and *CW-suspension*. The free join $X * Y$ of the topological spaces X and Y is $X \times Y \times [0, 1]/\equiv$, where the only nontrivial equivalence classes of $\equiv$ are $\{(x, y_0, 0) : x \in X\}$ and $\{(x_0, y, 1) : y \in Y\}$. The free join $\Omega * \Omega'$ of two CW-complexes Ω and Ω' may also be given as a CW-complex (see May [19, Chapter 10, Section 2]). This operation satisfies

$$P_1(\Omega * \Omega') = P_1(\Omega) \diamond^* P_1(\Omega') \tag{3}$$

where $P \diamond^* Q = (P^* \diamond Q^*)^*$. The *suspension* of a topological space X is usually defined as $\mathrm{Susp}(X) = X * \mathbb{S}^0$ (see, e.g., Dong [7] or Readdy [20]), because this operation assigns a simplicial complex to a simplicial complex. For CW-complexes there is also a more economical way to create a face structure $\mathrm{Susp}(\Omega)$ on the suspension of the topological space underlying Ω. This is noted by Dold [6, Chapter V, Example 3.10]. Using that construction we obtain

$$P_1(\mathrm{Susp}(\Omega)) \cong B_2 * P_1(\Omega). \tag{4}$$

Here and later B_m is the Boolean algebra of rank m.

1.4 Polyspherical coordinates

Our polyspherical coordinates $(\theta_1, \ldots, \theta_{n-1})$ parameterize the *unit $(n-1)$-sphere* defined by $\sum_{i=1}^{n} x_i^2 = 1$. Recall that a *binary tree* is a planar rooted tree such that each internal node has at most two children. It is *strongly binary* if each internal node has exactly two children. Let us fix strongly binary tree with

n leaves, and label its leaves with the rectangular coordinates $x_1, \dots, x_n$, as shown in Fig. 1. Associate to each internal node an angle θ_i ($i = 1, \dots, n-1$). We call such a labeled tree a *large tree*. For the

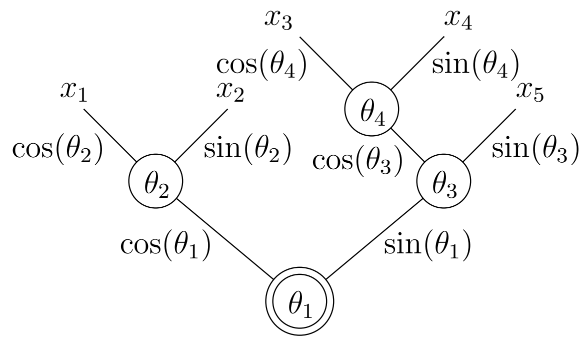


Figure 1: Large tree of polyspherical coordinates for a 4-dimensional sphere

internal node labeled θ_i, we label the edge to its left child with $\cos(\theta_i)$ and the edge to its right child with $\sin(\theta_i)$. Define the value of each x_j as the product of the labels on the edges along the unique path connecting the root with x_j. For example, for the labeled tree in Fig. 1 we set

$$\begin{aligned} x_1 &= \cos(\theta_1) \cdot \cos(\theta_2) \\ x_2 &= \cos(\theta_1) \cdot \sin(\theta_2) \\ x_3 &= \sin(\theta_1) \cdot \cos(\theta_3) \cdot \cos(\theta_4) \\ x_4 &= \sin(\theta_1) \cdot \cos(\theta_3) \cdot \sin(\theta_4) \\ x_5 &= \sin(\theta_1) \cdot \sin(\theta_3) \end{aligned}$$

Using the rule to express the x_i's as products of edge labels, every point of the unit sphere is already determined by the subtree of internal nodes and the labels θ_j associated to them. We call the rooted tree (T, r), consisting of these internal nodes only, a *small tree*. According to [22, (13)], our parameterization provides a single covering of the unit sphere if we set the following restrictions:

(S1) If the node associated to θ_i is a leaf in the small tree, we require $\theta_i \in [0, 2\pi]$.
(S2) If the node of θ_i has only a right child, we require $\theta_i \in [0, \pi]$.
(S3) If the node of θ_i has only a left child, we require $\theta_i \in [-\pi/2, \pi/2]$.
(S4) If the node θ_i has two children, we require $\theta_i \in [0, \pi/2]$.

The representation by polyspherical coordinates may be made unique by factoring by the following equivalence relation.

Definition 1.1 *We say that $(\theta_1, \dots, \theta_{n-1})$ and $(\theta'_1, \dots, \theta'_{n-1})$ are equivalent, if for each j such that $\theta_j \neq \theta'_j$ at least one of the following holds:*

(i) θ_j (and θ'_j) is the label of a leaf in the small tree and $\theta_j, \theta'_j \in \{0, 2\pi\}$,
(ii) θ_j (and θ'_j) is the right descendant of a node whose labels in both vectors satisfy $\theta_i = \theta'_i \in \{0, \pi\}$,
(iii) θ_j (and θ'_j) is the left descendant of a node whose labels in both vectors satisfy $\theta_i = \theta'_i \in \{-\pi/2, \pi/2\}$.

Thus, for example, the labels of the right descendants of a node labeled 0 or π are irrelevant. Each equivalence class may be represented by a *simplified code*, where the irrelevant coordinates are replaced with a dash.

2 The polyspherical complex $C((T,r);m_1,\ldots,m_{n-1})$

Definition 2.1 *Consider a code $(\sigma_1,\ldots,\sigma_{n-1})$ associated to a small tree, such that each σ_j is either a real number, or an interval $[\alpha,\beta]$, or the $*$ sign. We call such a code a* standard lune code *if it satisfies the following conditions:*

(i) Exactly one σ_i is an interval $[\alpha,\beta]$. The node of this σ_i is the root *of the standard lune code.*
(ii) The interval $[\alpha,\beta]$ is a proper *subset of the interval $I(\sigma_i)$ that occurs in the restrictions (S1)–(S4) applied to the node of σ_i. Moreover, $0<\beta-\alpha<\pi$.*
(iii) All descendants of the node of σ_i are labeled with a real number, subject to the restrictions (S1)–(S4).
(iv) All other nodes are labeled with $*$.

We use a standard lune code to denote the set of all polyspherical vectors $(\theta_1,\ldots,\theta_{n-1})$ satisfying $\theta_i\in[\alpha,\beta]$, and $\theta_j=\sigma_j$ whenever $j\neq i$ and $\sigma_j\neq *$. A *standard hemisphere code* is defined analogously, the only difference is setting $\alpha=\beta$. The equivalence relation for polyspherical coordinates may be extended to codes of standard lunes and hemispheres, so we may obtain a simplified code for them, and think of them as subsets of the unit sphere.

Proposition 2.2 *A standard lune code with $d-1$ star signs encodes a d-dimensional closed region, whose boundary is the union of two spheres obtained by replacing the interval $[\alpha,\beta]$ with α or β respectively (thus obtaining standard hemispheres).*

Taking intersections of standard lunes motivates the introduction of canonical regions.

Definition 2.3 *A* canonical region code *is a vector $(\sigma_1,\ldots,\sigma_{n-1})$ such that each σ_i is either a real number, or an interval $[\alpha,\beta]$, or the symbol $*$, or the symbol $-$, subject to the following conditions:*

(i) If σ_i is a real number or an interval, then it is an element or proper subset of $I(\sigma_i)$.
(ii) If $\sigma_i\in\{0,\pi\}$ then every right descendant of the node labeled σ_i is labeled $-$.
(iii) If $\sigma_i\in\{-\pi/2,\pi/2\}$ then every left descendant of the node labeled σ_i is labeled $-$.
(iv) σ_i is $-$ then every descendant of the node labeled σ_i is labeled $-$.
(v) If σ_i is $-$ then the parent of the node labeled σ_i is either labeled with $-$, or it is labeled with an element of $\{0,\pi\}$ and the node of σ_i is the right child, or it is an element of $\{-\pi/2,\pi/2\}$ and the node of σ_i is the left child.

(vi) If $\sigma_i \neq *$ and some descendant of its node is $*$ or an interval, then the node of σ_i has two children in the small tree, and σ_i is a real number belonging to $\{0, \pi/2\}$. (Thus either its left or its right subtree will have all of its nodes labeled with $-$.)

Again we obtain equivalence classes of polyspherical vectors; thus we may think of *canonical regions* as subsets of the unit sphere.

Proposition 2.4 *Every canonical region, except for the entire unit sphere, may be written as an intersection of finitely many standard hemispheres and standard lunes. Conversely, the intersection of any two canonical regions may be written as a union of canonical regions.*

Definition 2.5 *We call a rooted tree (T, r) whose nodes are labeled with positive integers such that the label on each leaf is at least two a* loopless complex code.

Given a loopless complex code, consider the family $C((T, r); m_1, \ldots, m_{n-1})$ consisting of the empty set, and of all canonical regions whose code is subject to the following conditions:

(C1) No leaf is labeled with $*$.

(C2) Each node labeled with a real number has either an ancestor whose label is an interval, or an ancestor labeled with $\sigma_i = *$ for which the corresponding m_i is equal to 1.

(C3) If σ_i is a real number and $I(\sigma_i) = [\gamma, \delta]$, then we have

$$\sigma_i = \gamma + t \cdot \frac{\delta - \gamma}{m_i} \quad \text{for some} \quad t \in \{0, 1, \ldots, m_i\}.$$

(C4) If σ_i is an interval and $I(\sigma_i) = [\gamma, \delta]$, then we have

$$[\alpha, \beta] = \left[\gamma + t \cdot \frac{\delta - \gamma}{m_i}, \gamma + (t+1) \cdot \frac{\delta - \gamma}{m_i}\right] \quad \text{for some} \quad t \in \{0, 1, \ldots, m_i - 1\}.$$

Theorem 2.6 *Given a loopless code associated to a small tree (T, r), $C((T, r); m_1, \ldots, m_{n-1})$ is a CW-complex, homeomorphic to an $(n-1)$-sphere.*

3 A recursive description of the polyspherical complexes

Any binary tree (T, r) having at least two nodes may be reconstructed from knowing the children r_1 and r_2 of the root and the subtrees T_i of descendants of r_i. (Only at least one of the r_i's needs to exist.) In this section we describe the structure of a polyspherical complex in terms of the polyspherical complexes associated to the subtrees of the children of the root in its small tree. The arising operations assign the face poset of a CW-sphere to face posets of CW-spheres. Thus the aggregate of the statements in this section provides proof of Theorem 2.6.

Consider a family of canonical regions $C((T,r);m_1,\ldots,m_{n-1})$. Assume w.l.o.g. that m_1 is associated to the root. We introduce $P_1(C((T,r);m_1,\ldots,m_{n-1}))$ to denote the partially ordered set obtained by taking the elements of $C((T,r);m_1,\ldots,m_{n-1})$, ordered by inclusion, and add a new unique maximum element $\widehat{1}$, which we associate to the canonical region $(*,\ldots,*)$.

Lemma 3.1 *If $m_1=1$ and the root r has only a right child r' then $P_1(C((T,r);1,m_2,\ldots,m_{n-1}))=B_2 * P_1(C((T',r');m_2,\ldots,m_{n-1}))$.*

Proposition 3.2 *Assume that Ω is an $(n-2)$-dimensional CW-sphere, whose faces subdivide the unit sphere $\{(x_2,\ldots,x_n) : x_2^2+\cdots+x_{n-1}^2=1\}$. Then the CW-suspension $\Omega'=\mathrm{Susp}(\Omega)$ may be realized as a CW-complex subdividing the unit sphere $\{(x_1,\ldots,x_n) : x_1^2+\cdots+x_{n-1}^2=1\}$. As a consequence we have $P_1(\Omega')=B_2 * P_1(\Omega)$.*

To handle the case $m_1>1$, we introduce *m-fold edge subdivisions*. Given an edge e in a CW complex, connecting the vertices u and v, we introduce $m-1$ new vertices $u_1,u_2,\ldots,u_{m-1}$, and set $u_0:=u$ and $u_m:=v$. We remove e and introduce m new edges $e_1,\ldots,e_m$ such that e_i connects u_{i-1} and u_i for $i=1,2,\ldots,m$ and a face f contains any of the e_i's in the new CW-complex if and only if it contains e in the original complex. We make the analogous definition for graded posets as well. An m-fold edge subdivision does not change the homeomorphy type of the CW-complex and, for posets, it preserves the Eulerian property. We denote by $E_m(P)$ the poset obtained by applying m-fold edge subdivision to all rank 2elements of the graded poset P.

Theorem 3.3 *Assume that the root r in the small tree associated to $C((T,r);m_1,m_2,\ldots,m_{n-1})$ has only a right child r'. Then $P_1(C((T,r);m_1,m_2,\ldots,m_{n-1}))\cong E_{m_1}(B_2 * P_1(C((T',r');m_2,\ldots,m_{n-1})))$. Here T' is the subtree of descendants of r'.*

The case when the root of the small tree has only a left child is completely analogous. Consider finally the case when the root r has two children: a left child r_1 and a right child r_2, with subtrees of descendants T_1 and T_2. W.l.o.g. we may assume that $m_2,\ldots,m_k$ belong to the nodes in T_1 and $m_{k+1},\ldots,m_{n-1}$ belong to the nodes of T_2. Again we discuss first the case $m_1=1$ separately.

Lemma 3.4 *Under the conditions listed above, we have*

$$P_1(C((T,r);1,m_2,\ldots,m_{n-1}))\cong P_1(C((T_1,r_1);m_1,\ldots,m_k))\diamond^* P_1(C((T_2,r_2);m_{k+1},\ldots,m_{n-1})).$$

Proposition 3.5 *Assume that the $(k-1)$-dimensional CW-sphere Ω_1 subdivides $\{(x_1,\ldots,x_k) : x_1^2+\cdots+x_k^2=1\}$ and the $(n-k-1)$-dimensional CW-sphere Ω_2 subdivides $\{(x_{k+1},\ldots,x_n) : x_{k+1}^2+\cdots+x_n^2=1\}$. Then $\Omega_1 * \Omega_2$ may be realized as a CW-sphere Ω, subdividing $\{(x_1,\ldots,x_n) : x_1^2+\cdots+x_{n-1}^2=1\}$. As a consequence we have $P_1(\Omega)=P_1(\Omega_1)\diamond^* P_1(\Omega_2)$.*

Finally, to state the analogue of Theorem 3.3, we need the following analogue of the operator E_m.

Definition 3.6 *The m-fold edge-subdivided dual diamond product $P \diamond_m^* Q$ is obtained from $P \diamond^* Q$ by applying m-fold edge subdivision to all elements of the form $(p, q) \in P \diamond^* Q$ such that both p and q have rank* 1*.*

Theorem 3.7 *Making the same assumptions as in Lemma 3.4 except for allowing $m_1 > 1$, we have*

$$P_1(C((T, r); m_1, m_2, \ldots, m_{n-1})) \cong P_1(C((T_1, r_1); m_2, \ldots, m_k)) \diamond_{m_1}^* P_1(C((T_2, r_2); m_{k+1}, \ldots, m_{n-1})).$$

When we prove Theorem 2.6 by induction, the basis is the case when the tree T has only one vertex, and $m_1 \geq 2$. The resulting complex $C(\{r\}, r; m_1)$ is a circle, subdivided into m_1 arcs.

4 Non-negativity of the cd-index

Theorem 4.1 *The cd-index of any poset $P_1(C((T, r); m_1, m_2, \ldots, m_{n-1}))$ associated to a loopless complex code is non-negative.*

This may be shown using the recursive description of our polyspherical complexes given in Section 3.

Definition 4.2 *We say that an Eulerian poset P is* upwards cd-positive *if, for every $x \in P \setminus \{\widehat{1}\}$, the interval $[x, \widehat{1}]$ has a non-negative cd-index.*

It is easy to see that $C(\{r\}, r; m_1)$ is upwards cd-positive for all $m_1 > 1$.

Proposition 4.3 *If P is an upwards cd-positive Eulerian poset, then so is $B_2 * P$.*

The following is an easy consequence of [9, Proposition 7.4].

Proposition 4.4 *If the Eulerian posets P and Q are upwards cd-positive, then so is $P \diamond^* Q$.*

Finally, since any m-fold subdivision may be obtained by iterated 2-fold subdivisions, it is sufficient to show the following.

Proposition 4.5 *Assume that P is an upwards cd-positive Eulerian poset of rank $n + 1$, and that $e \in P$ is an element of rank two. Let Q be the Eulerian poset obtained from P by applying* 2*-fold edge-subdivision to e. Then Q is also upwards cd-positive.*

5 Strongly binary small trees

It is hard to convert the *cd*-index to a type B quasisymmetric function; thus finding an explicit general formula for the flag f-vector of an arbitrary polyspherical complex seems difficult. The situation becomes easier when we may restrict ourselves to using only one of these two encodings of the flag f-vector. The special case when the small tree is a path is the subject of [13]. From now on, we assume the "other extreme" that all underlying small trees are strongly binary. To reduce the complexity of the question to a level similar to [13], we we also require that the number m_i associated to any interior node in the small tree has to be 1. Then the shape of the tree becomes irrelevant, because of the following:

Lemma 5.1 *Consider a loopless code $((T,r);m_1,m_2,\ldots,m_{2n-1})$ such that the underlying small tree (T,r), is strongly binary. Assume that $m_{n+1},\ldots,m_{2n-1}$ are associated to the interior nodes and that these numbers are equal to 1. Then we have*

$$P_1(C((T,r);m_1,m_2,\ldots,m_{2n-1}))\cong P_1(C(\{r\},r;m_1))\diamond^* P_1(C(\{r\},r;m_2))\diamond^*\cdots\diamond^* P_1(C(\{r\},r;m_n)).$$

The computation of the flag f-vector of such a poset is possible using *dual type B quasisymmetric functions*

$$F_B^*(P)=\sum_{\widehat{0}\leq x_1\leq\cdots\leq x_m<1} s^{\rho(\widehat{1})-\rho(x_m)-1}\cdot t_1^{\rho(x_1)-\rho(x_0)}t_2^{\rho(x_2)-\rho(x_1)}\cdots t_m^{\rho(x_m)-\rho(x_{m-1})}.$$

Dually to (2) we have the identity $F_B^*(P\diamond^*Q)=F_B^*(P)\cdot F_B^*(Q)$. Direct substitution into the definitions yields

$$F_B^*(P_1(C(\{r\},r;m)))=m\cdot\left(\sum_i t_i+\frac{s}{2}\right)^2-\left(\frac{m}{4}-1\right)\cdot s^2. \tag{5}$$

Corollary 5.2 *Under the assumptions of Lemma 5.1 we have*

$$F_B^*(P_1(C((T,r);m_1,m_2,\ldots,m_{2n-1})))=\prod_{j=1}^{n}\left(m_j\cdot\left(\sum_i t_i+\frac{s}{2}\right)^2-\left(\frac{m_j}{4}-1\right)\cdot s^2\right).$$

Using this Corollary it is possible to calculate the f-vectors of the order complexes, and it will be worthwhile to explore the sequences of polynomials arising, in analogy to [13]. To conclude, consider the special case $m_1=\ldots=m_n=4$. Then all terms $(m_j/4-1)s^2$ vanish from all factors:

Proposition 5.3 *The n-th dual diamond power $\mathcal{Q}_n$ of $P_1(C(\{r\},r;4))$ satisfies*

$$F_B^*(\mathcal{Q}_n)=\left(2\cdot\sum_i t_i+s\right)^{2n}=\sum_{k=0}^{2k}\binom{2n}{k}s^{2n-k}\left(2\cdot\sum_i t_i\right)^k.$$

From this we may deduce

$$f_{\{s_1\ldots,s_k\}}(\mathcal{Q}_n) = \binom{2n}{s_k} \cdot 2^{s_k} \cdot \binom{s_k}{s_1, s_2 - s_1, \ldots, s_k - s_{k-1}} \tag{6}$$

which is also equal to $f_{\{s_1\ldots,s_k\}}(T(B_{2n})$.

Corollary 5.4 *Substituting x into c and 1 into e in the ce-index of $\mathcal{Q}_n$ yields $(-1)^n Q_{2n}(x \cdot \sqrt{-1})$, where $Q_{2n}(x)$ the $2n$-th derivative polynomial for secant.*

Remark 5.5 Among all posets of the form $P_1(C(\{r\}, r; m))$, only $P_1(C(\{r\}, r; 4))$ is the Tchebyshev transform of a poset of rank 2. Thus we may also use a result of Ehrenborg and Readdy [11] stating that for any pair of graded posets (P, Q), $T(P \times Q)$ has the same flag f-vector as $T(P) \diamond^* T(Q)$.

Acknowledgments

I wish to thank Richard Stanley, Margaret Readdy, and Richard Ehrenborg for providing valuable mathematical information and advice.

Final note

To observe the 12 page limit, we omitted all proofs. A preprint with the title "Polyspherical complexes" is available at `http://www.math.uncc.edu/~ghetyei` .

References

[1] M. M. Bayer and L. J. Billera, Generalized Dehn–Sommerville relations for polytopes, spheres and Eulerian partially ordered sets, *Invent. Math.* **79** (1985), 143–157.

[2] M. Bayer and A. Klapper, A new index for polytopes, *Discrete Comput. Geom.* **6** (1991), 33–47.

[3] A. Björner, Posets, regular CW-complexes and Bruhat order *European J. Combin.* **5** (1984), 7–16.

[4] M. Bruggeser and P. Mani, Shellable decompositions of cells and spheres, *Math. Scand.* **29** (1971), 197-205.

[5] C.-O. Chow, Noncommutative symmetric functions of type B, Doctoral dissertation, Massachussetts Institute of Technology, 2001.

[6] A. Dold, “Lectures on Algebraic Topology,” Springer-Verlag New York, 1980.

[7] X. Dong, Topology of bounded-degree graph complexes, *J. Algebra* **262** (2003), 287–312.

[8] R. Ehrenborg, On Posets and Hopf Algebras, *Adv. in Math.* **119** (1996), 1–25.

[9] R. Ehrenborg and H. Fox, Inequalities for cd-indices of joins and products of polytopes, *Combinatorica* **23** (2003), 427–452.

[10] R. Ehrenborg and M. Readdy, Coproducts and the cd-index, J. Algebraic Combin. **8** (1998), 273–299.

[11] R. Ehrenborg and M. Readdy, The Tchebyshev Transforms of the First and Second Kind, manuscript in preparation.

[12] G. Hetyei, Tchebyshev posets, *Discrete & Comput. Geom.* **32** (2004) 493–520.

[13] G. Hetyei, Orthogonal polynomials represented by CW-spheres, *Electron. J. Combin.* **11(2)** (2004), #R4, 28 pp.

[14] G. Hetyei, Matrices of formal power series associate to binomial posets, to appear in *J. Algebraic Combin.*

[15] M. E. Hoffman, Derivative Polynomials for Tangent and Secant, *Amer. Math. Monthly* **102** (1995), 23–30.

[16] M. E. Hoffman, Derivative Polynomials, Euler Polynomials, and Associated Integer Sequences, *Electron. J. Comb.* **6** (1999), #R21.

[17] D. E. Knuth and T. J. Buckholtz, Computation of tangent, Euler and Bernoulli numbers, *Math. Comp.* **21** (1967), 663–688.

[18] C. Krichnamachary and Rao M. Bhimasena, On a table for calculating Eulerian numbers based on a new method, *Proc. London Math. Soc.* (2) **22** (1923), 73–80.

[19] J. P. May, “A concise course in algebraic topology,” The University of Chicago Press, Chicago and London, 1999.

[20] M. Readdy, The pre-WDVV ring of physics and its topology, preprint 2003, `http://www.ms.uky.edu/~readdy/Papers/pre_WDVV.pdf`

[21] R. P. Stanley, Flag f-vectors and the cd-index, *Math. Z.* **216** (1994), 483–499.

[22] N. Ya. Vilenkin, G. I. Kuznetsov, and Ya. A. Smorodinski, Eigenfunctions of the Laplace Operator, Providing Representations of the $U(2)$, $SU(2)$, $SO(3)$, $U(3)$, and $SU(3)$ Groups and the Symbolic Method, *Sov. J. Nucl. Phys.* **2** (1966), 645–652.

Discrete surfaces and infinite smooth words*

Damien Jamet[1] and Geneviève Paquin[2]

[1] LIRMM, Université Montpellier II, 161 rue Ada, 34392 Montpellier Cedex 5 - FRANCE.
jamet@lirmm.fr
[2] LaCIM, Université du Québec à Montréal, Montréal (QC), CANADA
paquin@lacim.uqam.ca

May 1st, 2005

Abstract In the present paper, we study the $(1,1,1)$-discrete surfaces introduced in [Jam04]. In [Jam04], the $(1,1,1)$-discrete surfaces are not assumed to be connected. In this paper, we prove that assuming connectedness is not restrictive, in the sense that, any two-dimensional coding of a $(1,1,1)$-discrete surface is the two-dimensional coding of both connected and simply connected ones. In the second part of this paper, we investigate a particular class of discrete surfaces: those generated by infinite smooth words. We prove that the only smooth words generating such surfaces are $K_{(3,1)}$, $K_{(1,3)}$ and $2K_{(3,1)}$, where $K_{(a,b)}$ is the generalized Kolakoski's word over the two-letter alphabet $\{a,b\}$ with a as first letter.

Résumé Dans cet article, nous étudions les $(1,1,1)$-surfaces discrètes introduites dans [Jam04]. Dans l'article [Jam04], les surfaces ne sont pas supposées discrètes. Nous montrons dans cet article qu'il n'est pas restrictif de faire une telle supposition et que tout codage bi-dimensionel d'une $(1,1,1)$-surface discrète code à la fois une surface connexe et une surface simplement connexe. La seconde partie de cet article est consacrée à l'étude des surfaces discrètes engendrées par des mots lisses. Nous démontrons que les seuls mots lisses engendrant de telles surfaces sont les mots $K_{(3,1)}$, $K_{(1,3)}$ et $2K_{(3,1)}$, où $K_{(a,b)}$ est le mot de Kolakoski généralisé sur l'alphabet $\{a,b\}$ commençant par a.

1 Introduction

A wide literature has been devoted to the study of Sturmian words, that is, the infinite words over a binary alphabet which have $n+1$ factors of length n [Lot02]. These words are also equivalently defined as a discrete approximation of a line with irrational slope. Then, many attempts have been done to generalize this class of infinite words to two-dimensional words. For instance, in [Vui98,BV00,ABS04], it is shown that the orbit of an element $\mu \in [0,1[$ under the action of two rotations codes a discrete plane. Furthermore, the problem of one

* with the support of NSERC (Canada)

or two-dimensional words characterizing discrete lines or planes is investigated in [BV00,Lot02,ABS04,BT04,DGK03]. In [Jam04], the author introduces the $(1,1,1)$-discrete surfaces as a quite natural generalization of the discrete planes of [BV00,ABS04] and shows how to decide whether a given two-dimensional sequence over the three-letter alphabet $\{1,2,3\}$ codes a $(1,1,1)$-discrete surface.

In the present paper, we study the connectedness and the simply-connectedness of the $(1,1,1)$-discrete surfaces and we show that, given a two-dimensional sequence u over the three-letter alphabet $\{1,2,3\}$, then u codes a $(1,1,1)$-discrete surface if and only if u codes a connected one and a simply connected one. Secondly, we study the $(1,1,1)$-discrete surfaces associated with smooth words for the case of two-letter alphabets [BL02,BLL02,BBLP03] and for arbitrary alphabets [BBC04]. These surfaces have local geometric properties and we give an explicit description of the associated smooth words.

This paper is organized as follows. In Section 2, we recall the basic notions concerning $(1,1,1)$-discrete surfaces and the combinatorics on two-dimensional words over a finite alphabet. In Section 3, we prove the first main result of this paper, namely

Theorem. *Let $u \in \{1,2,3\}^{\mathbb{Z}^2}$ be a two-dimensional sequence. The following assertions are equivalent:*

i) the sequence u codes a $(1,1,1)$-discrete surface;
ii) the sequence u codes a connected and simply connected $(1,1,1)$-discrete surface.

We also prove that any connected surface coded by an element of $\{1,2,3\}^{\mathbb{Z}^2}$ is simply connected. Finally, in Section 4, after having recalled basic notions concerning smooth words, we demonstrate the second main result, that is:

Theorem. *Let w be a smooth word over the alphabet $\{1,2,3\}$. The tiling $T(w)$ associated to w is a piece of a discrete surface if and only if $w \in \{K_{(1,3)}, K_{(3,1)}, 2K_{(3,1)}\}$, where $K_{(a,b)}$ is the generalized Kolakoski's word over the two-letter alphabet $\{a,b\}$ with a as first letter.*

2 Basic notions

2.1 Discrete surfaces

In this section we recall the basic notions concerning $(1,1,1)$-discrete surfaces and discrete planes.

Let $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}\}$ denotes the canonical basis of the Euclidean space $\mathbb{R}^3$. An element of $\mathbb{Z}^3$ is called a *voxel*. The *fundamental unit cube* $\mathcal{C}$ is the set defined by:

$$\mathcal{C} = \left\{x_1\overrightarrow{e_1} + x_2\overrightarrow{e_2} + x_3\overrightarrow{e_3} \mid (x_1,x_2,x_3) \in [0,1]^3\right\}.$$

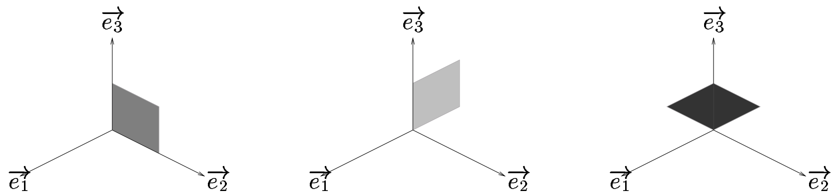
Let $\overrightarrow{x} \in \mathbb{Z}^3$. The set $\overrightarrow{x} + \mathcal{C}$ is called the *unit cube pointed* by $\overrightarrow{x}$.

Let $\mathcal{P}$ be the plane with equation $(\overrightarrow{v}, \overrightarrow{x}) = \mu$ with $\overrightarrow{v} \in \mathbb{R}^3_+$, $\mu \in \mathbb{R}$ and $(\overrightarrow{v}, \overrightarrow{x}) = v_1x_1 + v_2x_2 + v_2x_3$ denoting the usual scalar product of the vectors

$\overrightarrow{v}$ and $\overrightarrow{x}$. Let $\mathcal{C}_{\mathcal{P}}$ be the union of the unit cubes pointed by a voxel $x \in \mathbb{Z}^3$ and intersecting the open half-space $(\overrightarrow{v}, \overrightarrow{x}) < \mu$. We call *discrete plane* associated to $\mathcal{P}$ the set $\mathfrak{P}_{\mathcal{P}} = \overline{\mathcal{C}_{\mathcal{P}}} \setminus \overset{\circ}{\mathcal{C}_{\mathcal{P}}}$, where $\overline{\mathcal{C}_{\mathcal{P}}}$ (resp. $\overset{\circ}{\mathcal{C}_{\mathcal{P}}}$) is the closure (resp. the interior) of the set $\mathcal{C}_{\mathcal{P}}$ in $\mathbb{R}^3$, provided with its usual topology.

Let us now define the three *fundamental faces* (see Fig. 1):

$$\begin{aligned}
E_1 &= \{x_2\overrightarrow{e_2} + x_3\overrightarrow{e_3} \mid (x_2, x_3) \in [0,1[^2\},\\
E_2 &= \{-x_1\overrightarrow{e_1} + x_3\overrightarrow{e_3} \mid (x_1, x_3) \in [0,1[^2\},\\
E_3 &= \{-x_1\overrightarrow{e_1} - x_2\overrightarrow{e_2} \mid (x_1, x_2) \in [0,1[^2\}.
\end{aligned}$$



(a) Face of type 1. (b) Face of type 2 (c) Face of type 3

Fig. 1. The three fundamental faces.

Let $\overrightarrow{x} \in \mathbb{Z}^3$ and $k \in \{1, 2, 3\}$. The set $\overrightarrow{x} + E_k$ is called a *pointed face of type* k. The vector $\overrightarrow{x}$ is called the *distinguished vertex* of $\overrightarrow{x} + E_k$.

Let $\pi : \mathbb{R}^3 \longrightarrow \{\overrightarrow{x} \in \mathbb{R}^3 \mid (\overrightarrow{e_1} + \overrightarrow{e_2} + \overrightarrow{e_3}, \overrightarrow{x}) = 0\}$ be the orthogonal projection map onto the plane $\mathcal{P}_0$ with equation $(\overrightarrow{e_1} + \overrightarrow{e_2} + \overrightarrow{e_3}, \overrightarrow{x}) = 0$.

One has:

Theorem 1. [ABS04] *Let $\mathfrak{P}_{\mathcal{P}}$ be a discrete plane and let $\mathcal{V}_{\mathcal{P}} = \mathfrak{P} \cap \mathbb{Z}^3$ be the set of vertices of $\mathfrak{P}_{\mathcal{P}}$. We suppose that $\mathcal{P}$ admits a normal vector $\overrightarrow{v} \in \mathbb{R}^3_+$.*

1. *The set $\mathfrak{P}_{\mathcal{P}}$ is partitioned by pointed faces.*
2. *The restriction maps $\pi_{|\mathcal{V}_{\mathcal{P}}} : \mathcal{V}_{\mathcal{P}} \longrightarrow \pi(\mathbb{Z}^3)$ and $\pi_{|\mathfrak{P}_{\mathcal{P}}} : \mathfrak{P}_{\mathcal{P}} \longrightarrow \{\overrightarrow{x} \in \mathbb{R}^3 \mid (\overrightarrow{e_1} + \overrightarrow{e_2} + \overrightarrow{e_3}, \overrightarrow{x}) = 0\}$ are bijective.*

Let us now define the $(1, 1, 1)$-discrete surfaces as follows:

Definition 1. [Jam04] *A disjoint union $\mathfrak{S} \subseteq \mathbb{R}^3$ of pointed faces is a* $(1, 1, 1)$-discrete surface *if the map*

$$\begin{aligned}
\pi_{|\mathfrak{S}} : \mathfrak{S} &\longrightarrow \mathcal{P}_0\\
\overrightarrow{x} &\mapsto \pi(\overrightarrow{x})
\end{aligned}$$

is a bijection (see Fig. 2). The set $\mathcal{V}_{\mathfrak{S}} = \mathfrak{S} \cap \mathbb{Z}^3$ is called the set of vertices *of* $\mathfrak{S}$.

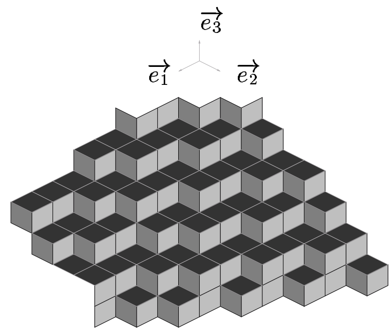


Fig. 2. A piece of a discrete surface

From now on, for clarity issue, we refer to $(1,1,1)$-discrete surface as discrete surface.

Before associating a two-dimensional coding over the three-letter alphabet $\{1,2,3\}$ to any discrete surface $\mathfrak{S}$, let us recall a technical lemma.

Lemma 1. *Let $\mathfrak{S}$ be a discrete surface. The following properties hold:*

i) The map

$$\begin{array}{rccc} \pi_{|\mathcal{V}_S} : & \mathfrak{S} & \longrightarrow & \pi\left(\mathbb{Z}^3\right) \\ & \overrightarrow{x} & \mapsto & \pi(\overrightarrow{x}) \end{array}$$

is a bijection.

ii) Each vertex $\overrightarrow{x}$ of $\mathcal{V}_{\mathcal{P}}$ is the distinguished vertex of one and only one pointed face.

We can now associate a two-dimensional coding over the three letter-alphabet $\{1,2,3\}$ to any discrete surface $\mathfrak{S}$ as follows: let $\Gamma = \pi(\mathbb{Z}^3) = \mathbb{Z}\pi(\overrightarrow{e_1}) \oplus \mathbb{Z}\pi(\overrightarrow{e_2})$. We identify Γ and $\mathbb{Z}^2$ by the lattice isomorphism

$$\begin{array}{rccc} \Phi : & \mathbb{Z}^2 & \longrightarrow & \Gamma \\ & (m,n) & \mapsto & m\pi(\overrightarrow{e_1}) + n\pi(\overrightarrow{e_2}). \end{array} \tag{1}$$

To any discrete surface $\mathfrak{S}$, we associate the two-dimensional coding $u \in \{1,2,3\}^{\mathbb{Z}^2}$ defined by: $\forall (m,n) \in \mathbb{Z}^2$, $\forall k \in \{1,2,3\}$,

$$u_{m,n} = k \text{ if and only if } \pi^{-1}_{|\mathcal{V}_S}(m\pi(\overrightarrow{e_1}) + n\pi(\overrightarrow{e_2})) \text{ is of type } k \text{ in } \mathfrak{S}.$$

In other words, $u_{m,n} = k$ if the pre-image of the points $m\pi(\overrightarrow{e_1}) + n\pi(\overrightarrow{e_2})$ is of type k in $\mathfrak{S}$.

Hence it becomes natural to wonder whether a given two-dimensional sequence $u \in \{1,2,3\}^{\mathbb{Z}^2}$ codes a discrete surface. In order to investigate this problem, we need several notions about combinatorics on two-dimensional words over a finite alphabet.

2.2 Basic notions on two-dimensional sequences over a finite alphabet

In this section, we recall some basic notions of combinatorics on two-dimensional words over a finite alphabet (see for instance [GR97]).

Let Σ be a finite alphabet. Let Ω be a finite subset of $\mathbb{Z}^2$. A function $w : \Omega \longrightarrow \Sigma$ is called a *finite pointed pattern over the alphabet* Σ.

A *shape* $\overline{\Omega}$ of $\mathbb{Z}^2$ is the equivalence class of $\Omega \subseteq \mathbb{Z}^2$ for the following equivalence relation:

$$\Omega \sim \Omega' \iff \exists (v_1, v_2) \in \mathbb{Z}^2,\ \Omega_1 = \Omega_2 + (v_1, v_2).$$

One defines an equivalence relation on the set of finite pointed patterns as follows. Let $w : \Omega \to \Sigma$ and $w' : \Omega' \to \Sigma$ be two finite pointed patterns. Then ω is said to be equivalent to ω' if and only if:

$$\exists (v_1, v_2) \in \mathbb{Z}^2,\ \Omega = \Omega' + (v_1, v_2),\ \forall (m, n) \in \Omega,\ w_{m,n} = w_{m+v_1, n+v_2}. \qquad (2)$$

One can easily convince himself that the previous relation is an equivalence one, where the equivalence classes are called *finite patterns of shape* Ω.

For short, we denote the finite patterns w instead of $\overline{w}$ and we will denote the shapes Ω instead of $\overline{\Omega}$.

Let $u \in \Sigma^{\mathbb{Z}^2}$ be a two-dimensional sequence and let $w : \Omega \longrightarrow \Sigma$ be a finite pointed pattern. We say that w occurs in u if there exists $(m_0, n_0) \in \mathbb{Z}^2$ such that for all $(m, n) \in \Omega$, $w_{m,n} = u_{m_0+m, n_0+n}$. Such a couple (m_0, n_0) is called an *occurrence* of w. We say that a finite pattern $\overline{w}$ occurs in u if one of its pointed represents occurs in u. The set $\mathcal{L}(u)$ (resp. $\mathcal{L}_\Omega(u)$) of finite patterns (resp. of finite patterns of shape Ω) occurring in u is called the *language* (resp. the Ω-*language*) of u. Let Ω be a shape. The Ω-*complexity map of* u is the function $p_\Omega : \Sigma^{\mathbb{Z}^2} \longrightarrow \mathbb{N} \cup \{\infty\}$ defined as follows:

$$\begin{array}{rcl} p_\Omega : \Sigma^{\mathbb{Z}^2} & \longrightarrow & \mathbb{N} \cup \{\infty\} \\ u & \mapsto & |\mathcal{L}_\Omega(u)|, \end{array}$$

where $|\mathcal{L}_\Omega(u)|$ is the cardinality of the set $\mathcal{L}_\Omega(u)$.

2.3 Recognition of discrete surfaces

Let $u \in \{1, 2, 3\}^{\mathbb{Z}^2}$. In this section we investigate the following question: does u code a discrete surface? Let us now introduce the pointed hooks and the hook shape of u.

Definition 2. *The* hook shape *is the equivalence class of the sets* $\{(m, n); (m, n+1); (m+1, n+1)\}$, *for* $(m, n) \in \mathbb{Z}^2$, *for the relation* $\sim$ *defined in Section 2.2. (see Fig. 3)*

The following theorem holds.

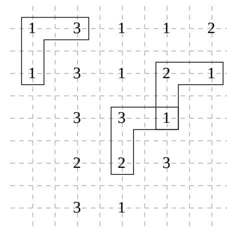
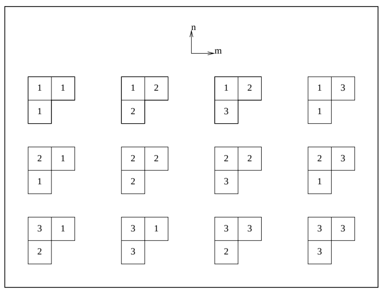


Fig. 3. Examples of hook words occurring in a sequence $u \in \{1,2,3\}^{\mathbb{Z}^2}$.



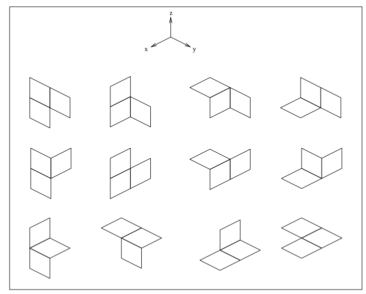


Fig. 4. Left: The permitted hook-words. Right: The 3-dimensional representation of the permitted hook-words.

Theorem 2. [Jam04] *Let $u \in \{1,2,3\}^{\mathbb{Z}^2}$. Then u codes a discrete surface if and only if the hook-language of u is included in the following set of patterns (see Fig. 4).*

A drawback of the previous definition of discrete surface is to be non-intuitive. For instance, let us consider a discrete plane $\mathfrak{P}_{\mathcal{P}}$. By construction, $\mathfrak{P}_{\mathcal{P}}$ is connected and simply connected, that is, it does not contain any hole. Let $\vec{x} \in \mathcal{V}_{\mathcal{P}}$ be a vertex of $\mathfrak{P}_{\mathcal{P}}$ of type k. Then, $\mathfrak{P} \setminus \{\vec{x} + E_k\} \cup \{(\vec{x} + \vec{e_1} + \vec{e_2} + \vec{e_3}) + E_k\}$ is still a discrete surface.

In the following section, we show that assuming the discrete surfaces to be connected is not a restriction. More precisely, we prove that any sequence $u \in \{1,2,3\}^{\mathbb{Z}^2}$ coding a discrete surface also codes a connected and a simply-connected one.

3 The connected discrete surfaces

In this section, we investigate the discrete surfaces introduced in [Jam04] and prove that, for any discrete surface, there exists a connected one with the same two-dimensional coding.

Theorem 3. *Let $u \in \{1,2,3\}^{\mathbb{Z}^2}$ be a two-dimensional sequence. The following assertions are equivalent:*

i) the sequence u codes a discrete surface;
ii) the sequence u codes a connected discrete surface.

Let $x = m\pi(\overrightarrow{e_1}) + n\pi(\overrightarrow{e_2})$, with $(m,n) \in \mathbb{Z}^2$. Using the previously introduced identification of Γ as $\mathbb{Z}^2$ (see (1)), we denote $u_{\overrightarrow{x}}$ instead of $u_{m,n}$.

Let $u \in \{1,2,3\}^{\mathbb{Z}^2}$ be the two-dimensional coding of a discrete surface, that is,

$$\bigcup_{\overrightarrow{x} \in \Gamma} (\overrightarrow{x} + \pi(E_{u_{\overrightarrow{x}}})),$$

where E_i is a fundamental face ($i \in \{1,2,3\}$), is a partition of the plane $\mathcal{P}_0$ (see [Jam04]). We define a partial order relation $\xrightarrow{u}$ over Γ as follows:

$$\forall(\overrightarrow{x}, \overrightarrow{y}) \in \Gamma^2,\ \overrightarrow{x} \xrightarrow{u} \overrightarrow{y} \iff \overrightarrow{y} \in \overline{\overrightarrow{x} + \pi(E_{u_{\overrightarrow{x}}})} \setminus \overrightarrow{x} + \overset{\circ}{\pi}(E_{u_{\overrightarrow{x}}}).$$

Lemma 2. *Let $\overrightarrow{x} \in \Gamma$. Then $\overrightarrow{x} \xrightarrow{u} \overrightarrow{x} - \pi(\overrightarrow{e_1})$ and $\overrightarrow{x} \xrightarrow{u} \overrightarrow{x} - \pi(\overrightarrow{e_1}) - \pi(\overrightarrow{e_2})$ (see Fig. 5).*

Proof. This is immediately deduced from $\pi(\overrightarrow{e_3}) = -(\pi(\overrightarrow{e_1}) + \pi(\overrightarrow{e_2}))$ and from the definitions of E_1, E_2 and E_3 and the projection map π. ■

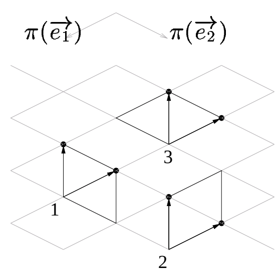


Fig. 5. Lattice representation of the partial order relation $\xrightarrow{u}$.

Finally, Theorem 3 is a direct consequence of the following lemma.

Lemma 3. *Let $\overrightarrow{x} \in \mathbb{Z}^3$ and let $\overrightarrow{l} = \pi(\overrightarrow{x}) - \pi(\overrightarrow{e_1}) - \pi(\overrightarrow{e_2}) \in \Gamma$ (resp. $\overrightarrow{r} = \pi(\overrightarrow{x}) - \pi(\overrightarrow{e_1}) \in \Gamma$) be the left (resp. the right) targets of the arrows whose source is $\pi(\overrightarrow{x})$ in the graph of Fig. 5. Then, the set*

$$\left(\overrightarrow{x} + E_{u_{\pi(\overrightarrow{x})}}\right) + \left(\overrightarrow{y} + E_{u_{\overrightarrow{l}}}\right) + \left(\overrightarrow{z} + E_{u_{\overrightarrow{r}}}\right),$$

with

$$\begin{aligned} \overrightarrow{y} = \overrightarrow{x} + \overrightarrow{e_3} \text{ and } \overrightarrow{z} &= \overrightarrow{x} + \overrightarrow{e_2} + \overrightarrow{e_3} \text{ if } u_{\pi(\overrightarrow{x})} = 1, \\ \overrightarrow{y} = \overrightarrow{x} + \overrightarrow{e_3} \text{ and } \overrightarrow{z} &= \overrightarrow{x} - \overrightarrow{e_1} \text{ if } u_{\pi(\overrightarrow{x})} = 2, \\ \overrightarrow{y} = \overrightarrow{x} - \overrightarrow{e_1} - \overrightarrow{e_2} \text{ and } \overrightarrow{z} &= \overrightarrow{x} - \overrightarrow{e_1} \text{ if } u_{\pi(\overrightarrow{x})} = 3. \end{aligned}$$

is connected and, in each previous case, $\pi(\overrightarrow{y}) = \overrightarrow{l}$ and $\pi(\overrightarrow{z}) = \overrightarrow{r}$.

Proof. In each case, one can verify that $\{\overrightarrow{y}, \overrightarrow{z}\} \subseteq \overline{\left(\overrightarrow{x} + E_{u_{\pi(\overrightarrow{x})}}\right)}$. For instance, let us suppose that $u_{\pi(\overrightarrow{x})} = 1$. Then,

$$\overline{\left(\overrightarrow{x} + E_{u_{\pi(\overrightarrow{x})}}\right)} = \overrightarrow{x} + \left\{x_2\overrightarrow{e_2} + x_3\overrightarrow{e_3} \mid (x_2, x_3) \in [0,1]^2\right\},$$

and $\overrightarrow{y} = \overrightarrow{x} + \overrightarrow{e_3} \in \overline{\left(\overrightarrow{x} + E_{u_{\pi(\overrightarrow{x})}}\right)}$ (see Fig. 6). Idem for $\overrightarrow{z} = \overrightarrow{x} + \overrightarrow{e_2} + \overrightarrow{e_3}$. Finally, $\pi(\overrightarrow{y}) = \pi(\overrightarrow{x} + \overrightarrow{e_3}) = \pi(\overrightarrow{x}) - \pi(\overrightarrow{e_1}) - \pi(\overrightarrow{e_2}) = \pi(\overrightarrow{l})$. ■

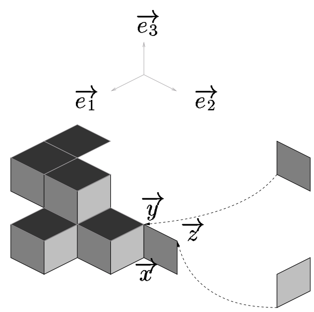


Fig. 6. Computing a connected surface by induction.

We can now prove Theorem 3:

Sketch of the proof. For all $\overrightarrow{x}_0 \in \Gamma$ and $r \in \mathbb{R}_+$, let us denote $B_\Gamma(\overrightarrow{x}_0, r) = \{m\pi(\overrightarrow{e_1}) + n\pi(\overrightarrow{e_2}) \in \Gamma \mid \max\{|m|, |n|\} < r\}$.

Let $\mathfrak{S}_1 = E_{u_{\overrightarrow{0}}}$. Then, the set $\mathfrak{S}_1$ is connected, $\pi(\mathfrak{S}_1 \cap \mathbb{Z}^3) = B_\Gamma(\overrightarrow{0}, 1)$ and for all $\overrightarrow{x} \in \mathbb{Z}^3 \cap \mathfrak{S}_1$, $\overrightarrow{x}$ is of type $u_{\pi(\overrightarrow{x})}$. Let $r \in \mathbb{N}^\star$ and let us suppose that $\mathfrak{S}_r$ is a connected disjoint union of pointed faces such that $\pi(\mathfrak{S}_r) \cap \mathbb{Z}^3 = B_\Gamma(\overrightarrow{0}, r)$ and for all $\overrightarrow{x} \in \mathbb{Z}^3 \cap \mathfrak{S}_r$, $\overrightarrow{x}$ is of type $u_{\pi(\overrightarrow{x})}$. Then, with Lemma 3 and the connectedness of the relation $\xrightarrow{u}$ (see Fig.7), that is the connectedness of the corresponding graph, one can be convinced that it is possible to build, by induction on the sets $\left(B_\Gamma(\overrightarrow{0}, r)\right)_{r\in\mathbb{N}^\star}$, a connected union $\mathfrak{S}_{r+1}$ of pointed faces such that $\pi(\mathfrak{S}_{r+1} \cap \mathbb{Z}^3) = B_\Gamma(\overrightarrow{0}, r+1)$. Indeed, for any element $\overrightarrow{y} \in B_\Gamma(\overrightarrow{0}, r+1) \setminus B_\Gamma(\overrightarrow{0}, r)$, there exists an element $\overrightarrow{x} \in B_\Gamma(\overrightarrow{0}, r)$ with $\overrightarrow{x} \xrightarrow{u} \overrightarrow{y}$. We thus obtain an increasing sequence $(\mathfrak{S}_r)_{r\in\mathbb{N}^\star}$ of connected unions of pointed faces such that, for all $r \in \mathbb{N}^\star$, $\pi_{|\mathfrak{S}_r} : \mathfrak{S}_r \longrightarrow \mathcal{P}$ is 1-1 (remind that we assume u to code a discrete surface). Let

$$\mathfrak{S} = \bigcup_{r\in\mathbb{N}^\star} \mathfrak{S}_r.$$

The set $\mathfrak{S}$ is connected (it is an increasing sequence of connected sets) and $\pi : \mathfrak{S} \longrightarrow \mathcal{P}$ is 1-1. Finally, since $\pi(\mathfrak{S} \cap \mathbb{Z}^3) = \Gamma$ and u codes a discrete surface,

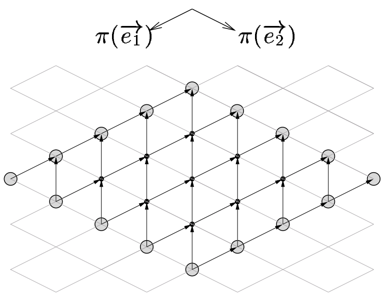


Fig. 7. A part of the graph of the relation $\xrightarrow{u}$.

that is,

$$\bigcup_{\overrightarrow{x} \in \mathfrak{S} \cap \mathbb{Z}^3} \Big(\pi(\overrightarrow{x}) + \pi(E_{u_{\pi(\overrightarrow{x})}}) \Big) = \mathcal{P}_0,$$

we conclude that $\pi : \mathfrak{S} \longrightarrow \mathcal{P}_0$ is onto. ■

An other interesting property of two-dimensional sequences coding discrete surfaces is:

Theorem 4. *Let $\mathfrak{S}$ be a connected discrete surface coded by $u \in \{1,2,3\}^{\mathbb{Z}^2}$. Then $\mathfrak{S}$ is simply-connected, that is, it admits no hole.*

Sketch of the proof. Let $(\mathfrak{S}_r)_{r \in \mathbb{N}^\star}$ be the sequence computed in the proof of Theorem 3. Then the following assertion holds:

$$\forall r \in \mathbb{N}^\star, \, \overline{\mathfrak{S}_r} \subseteq \mathfrak{S}_{r+2}.$$

Then,

$$\mathfrak{S} = \bigcup_{r \in \mathbb{N}^\star} \overline{\mathfrak{S}_r}$$

and $\mathfrak{S}$ is a union of closed sets. On can notice that each set $B(\overrightarrow{x}, r) = \{\overrightarrow{y} \in \mathbb{R}^3 \mid ||\overrightarrow{x} - \overrightarrow{y}||_\infty < r\}$, with $r \in \mathbb{R}_+^\star$, intersects at most a finite number of closed pointed faces. Hence, $\mathfrak{S}$ is closed. Since $\pi : \mathbb{R}^3 \longrightarrow \mathcal{P}_0$ is continuous, it follows that $\pi_{|\mathfrak{S}} : \mathfrak{S} \longrightarrow \mathcal{P}_0$ is continuous. It remains to show that $\pi_{|\mathfrak{S}}^{-1} : \mathcal{P}_0 \longrightarrow \mathfrak{S}$ is continuous. Indeed, $\pi : \mathbb{R}^3 \longrightarrow \mathcal{P}_0$ is a closed map. Finally, we have proved that $\pi_{|\mathfrak{S}} : \mathfrak{S} \longrightarrow \mathcal{P}$ is an homeomorphism. Hence, since Γ is simply-connected,we deduce that so is $\mathfrak{S}$.

4 Discrete surfaces generated by smooth words

In this section, we first recall some notions of combinatorics on words over arbitrary alphabets, as defined in [BBC04]. Then, we study the discrete surfaces generated by a specific class of words, the right infinite smooth words over the alphabet {1,2,3}. We prove that there are only three such discrete surfaces.

Let us consider a finite *alphabet* Σ of *letters*. A *right infinite word* is a sequence $w \in \Sigma^{\mathbb{N}}$. Every word $w \in \Sigma^{\mathbb{N}}$ can be uniquely written as a product of factors as follows:

$$w = {\alpha_1}^{e_1}{\alpha_2}^{e_2}{\alpha_3}^{e_3}\dots$$

with $e_j \in \mathbb{N}^\star$ and $\alpha_i \neq \alpha_{i+1}$. Hence, the *run-length encoding* defined by:

$$\begin{array}{rccl} \Delta : & \Sigma^{\mathbb{N}} & \longrightarrow & \mathbb{N}^{\mathbb{N}} \\ & w = {\alpha_1}^{e_1}{\alpha_2}^{e_2}{\alpha_3}^{e_3}\dots & \mapsto & e_1e_2e_3\dots, \end{array}$$

is well defined on $\Sigma^{\mathbb{N}}$.

Example 1. If $\Sigma = \{1, 2\}$, the operator Δ as two fixpoints, namely

$$\Delta(K_{(1,2)}) = K_{(1,2)}, \qquad \Delta(K_{(2,1)}) = K_{(2,1)},$$

where $K_{(2,1)}$ is the well-known Kolakoski word [Kol66], whose first terms are

$$K_{(2,1)} = 2211212212211211221211212211211212212211212212112112 2\dots$$

and $K_{(1,2)} = 1K_{(2,1)}$.

A right infinite word is said to be smooth if its alphabet is invariant under Δ. More precisely, the set $\mathcal{K}_\Sigma$ of the right infinite smooth words over Σ is:

$$\mathcal{K}_\Sigma = \left\{w \in \Sigma^{\mathbb{N}} \mid \forall k \in \mathbb{N}, \Delta^k(w) \in \Sigma^{\mathbb{N}}\right\}.$$

Given a smooth words w over a finite alphabet Σ, we define the *tiling associated to* w (see [BBLP03]) as the two-dimensional sequence $\left(T(w)_{m,n}\right)_{(m,n)\in\mathbb{N}^2}$ as follows:

$$\forall m \in \mathbb{N},\ (T(w)_{m,\bullet}) = \Delta^m(w).$$

In other words, for any $m \in \mathbb{N}$, the m-th line of $\left(T(w)_{m,n}\right)_{(m,n)\in\mathbb{N}^2}$ is the right infinite word $\Delta^m(w)$.

Let us now state the main result of this section:

Theorem 5. *Let w be a smooth word over the alphabet $\{1,2,3\}$. The tiling $T(w)$ associated to w is a piece of a discrete surface if and only if $w \in \{K_{(1,3)}, K_{(3,1)}, 2K_{(3,1)}\}$.*

Proof. Using the permitted hook-words of Fig. 4 and the smoothness condition, that is,

$$\forall m \in \mathbb{N},\ T(w)_{m+1,\bullet} = \Delta\left(T(w)_{m,\bullet}\right),$$

an exhaustive inspection gives that $T(w)$ must start by one of the patterns of Fig.8.

Clearly, the other 5 patterns are excluded because they do not respect the smoothness condition. We proceed by exhaustive inspection. Let us for instance investigate the first case (see Fig. 9). In the two first extensions of the initial word, the smoothness condition does not hold. In the third extension, the smoothness condition provides a forbidden pattern. In the last extension, we obtain the tiling associated to the word $K_{(1,3)}$.

The other cases can be treated in the same way. For instance, we obtain the tiling $T\left(2K_{(1,3)}\right)$ in the second case, and the tiling $T\left(K_{(3,1)}\right)$ in the third case. None of the other cases leads to a discrete surface. ∎

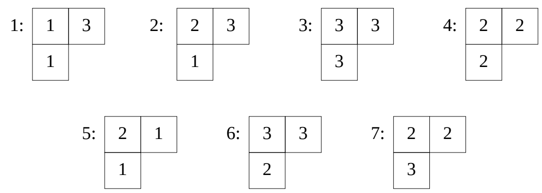


Fig. 8. The possible starting patterns.

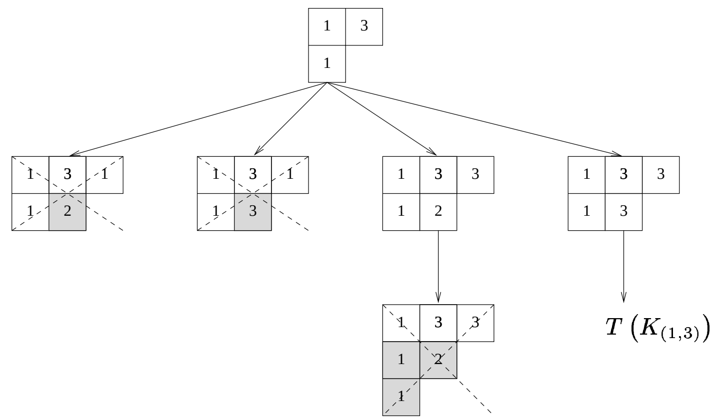


Fig. 9. The different extensions in the first case.

5 Concluding remarks

Since the identification of the three fundamental faces to the letters 1, 2 and 3 is arbitrary, a natural question arises: what is the action of permutation on the coding alphabet? By an exhaustive inspection of the 5 possible permutations, it can be shown that the only smooth tilings describing a discrete surface are generated by a generalized Kolakoski's word. It would be interesting to find a general proof showing which smooth words generate a discrete surface, for an arbitrary permutation on the coding alphabet. The next table gives the results.

Permutation	Smooth words generating discrete surfaces
(123)	$K_{(1,2)}$, $K_{(2,1)}$
(132)	$K_{(2,3)}$, $K_{(3,2)}$
(12)	$K_{(2,3)}$, $K_{(3,2)}$
(13)	$K_{(1,3)}$, $K_{(3,1)}$
(23)	$K_{(1,2)}$, $K_{(2,1)}$, $3K_{(2,1)}$

6 Acknowledgments

The authors wish to thank Valérie Berthé and Srečko Brlek for their useful comments and their support.

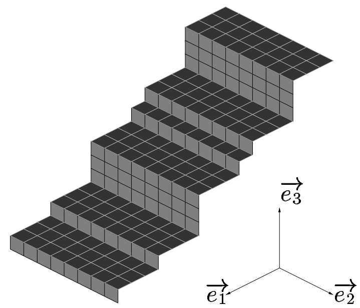


Fig. 10. A connected discrete surface associated to the word $K_{(1,3)}$.

References

[ABS04] P. Arnoux, V. Berthé, and A. Siegel. Two-dimensional iterated morphisms and discrete planes. *Theoret. Comput. Sci.*, 319(1-3):145–176, 2004.

[BBC04] V. Berthé, S. Brlek and P. Choquette. Smooth words over arbitrary alphabets, *Theoret. Comput. Sci.* (2004) submitted.

[BBLP03] A. Bergeron-Brlek, S. Brlek, A. Lacasse, X. Provençal, Patterns in Smooth Tilings, Proc. WORDS'03, 4-th International Conference on Words Turku (Finland) September 10-13, 2003, TUCS General Publication No 27, August (2003) 370-381.

[BLL02] S. Brlek, A. del Lungo, A. Ladouceur. The class of smooth tilings, Journées Montoises d'informatique théorique (Montpellier, France) September 9-11, 2002.

[BL02] S. Brlek, A. Ladouceur. Tiling with a new class of bi-infinite words, Proc. Gian-Carlo Rota Memorial Conference, Barasciano (AQ) Italie, April 25-27, 2002, Università di l'Aquila (2002) 35–46.

[BMP04] S. Brlek, G. Melançon and G. Paquin, Properties of the extremal infinite smooth words, Proc. the Journées Montoises d'informatique théorique (Liège, Belgique) September 8-11, 2004, 291–303

[BT04] V. Berthé and R. Tijdeman. Lattices and multi-dimensional words. *Theoret. Comput. Sci.*, 319(1-3):177–202, 2004.

[BV00] V. Berthé and L. Vuillon. Tilings and rotations on the torus: a two-dimensional generalization of sturmian sequences. *Discrete Math.*, 223(1-3):27–53, 2000.

[DGK03] F. Durand, A. Guerziz, M. Koskas. Words and morphisms with Sturmian erasures. Bulletin of the Belgian Mathematical Society (accepted), 11 pages.

[GR97] D. Giammarresi and A. Restivo. Two-dimensional languages. pages 215–267, 1997.

[Jam04] D. Jamet. On the Language of Discrete Planes and Surfaces. In *Proceedings of the Tenth International Workshop on Combinatorial Image Analysis*, pages 227-241. Springer-Verlag, 2004.

[Kol66] W. Kolakoski, Self Generating Runs, Problem 5304, *American Math. Monthly*, **72** (1965) 674. Solution: *American Math. Monthly* **73** (1966) 681–682.

[Lot02] Lothaire. *Algebraic Combinatorics on Words.* Cambridge University Press, 2002.

[Vui98] L. Vuillon. Combinatoire des motifs d'une suite sturmienne bidimensionelle. *Theoret. Comput. Sci.*, 209(1-2):261–285, 1998.

THE HIVE MODEL AND THE FACTORISATION OF KOSTKA COEFFICIENTS

R.C. KING, C. TOLLU, AND F. TOUMAZET

Version: April 28, 2005.

ABSTRACT. The hive model is used to explore the properties of both Kostka coefficients and stretched Kostka coefficient polynomials. It is shown that both of these may factorise, and that they can then be expressed as products of certain primitive coefficients and polynomials, respectively. It is further shown how to determine a sequence of linear factors $(t+m)$ of the primitive polynomials, where t is the stretching parameter, as well as a bound on their degree in the form of simple formula which is conjectured to be exact.

RÉSUMÉ. Nous utilisons le modèle des ruches pour étudier les propriétés des cœfficients de Kostka et des polynômes associés aux cœfficients de Kostka dilatés. Nous montrons que les uns et les autres peuvent se factoriser : ils s'écrivent comme des produits de cœfficients (resp. polynômes) primitifs. En outre, nous montrons comment établir une suite de facteurs linéaires $(t+m)$ des polynômes primitifs (t est le paramètre de dilatation), et proposons une formule simple donnant une borne supérieure de leur degré.

1. Introduction

There is no doubt that Kostka coefficients, $K_{\lambda\mu}$ are interesting combinatorial objects. They are indexed by pairs of partitions λ and μ, and are non-zero if and only if these partitions have the same weight and λ precedes μ with respect to the dominance partial order on partitions [JK]. They count the number of semistandard Young tableaux of shape determined by λ and of weight determined by μ, see for example [L, M, S2]. They also count Gelfand-Tsetlin patterns, as described for example in [GT, S2]. These patterns are in bijective correspondence with semistandard Young tableaux and also with certain K-hives, introduced comparatively recently [KTT] as a variation on the hives used to calculate Littlewood-Richardson coefficients [KT, KTW, B]. These K-hives are triangular arrays of non-negative integers with borders of length n labelled by the parts of the partitions 0, λ and μ, where $0=(0,0,\dots,0)$. Counting such K-hives gives $K_{\lambda\mu}$.

Multiplying all the parts of the partitions λ and μ by a stretching parameter t, with t a positive integer, gives new partitions $t\lambda$ and $t\mu$. Since the weights of $t\lambda$ and $t\mu$ are simply those of λ and μ multiplied by t, and this scaling preserves dominance partial ordering, it follows that $K_{t\lambda,t\mu}$ is non-zero if and only if $K_{\lambda\mu}$ is non-zero.

By way of a non-trivial example, for $n=9$, $\lambda=(9,5,2,2,2,2,1)$ and $\mu=(5,5,5,3,1,1,1,1,1)$ it is found quite typically that

$$K_{t\lambda,t\mu}=\frac{1}{24}(t+1)^2(t+2)(5t+2)(t^2+3t+6) \tag{1.1}$$

These functions $K_{t\lambda,t\mu}$ are necessarily quasi-polynomials in t since they enumerate integer points of rational polytopes subject to a scaling by t [E, S1]. However, contrary to initial expectations [BK], these rational polytopes may, and indeed sometimes do, possess non-integral vertices [KTT, DeLM]. Despite this, it has been proved [KR, K, BGR] that $P_{\lambda\mu}(t)=K_{t\lambda,t\mu}$ is always a polynomial in t.

A study of such K-polynomials has revealed a number of interesting features that are illustrated in the above example. In particular, it appears that the coefficients in the expansion of $P_{\lambda\mu}(t)$ are all positive rational numbers. This remains a conjecture [KTT].

Secondly, $P_{\lambda\mu}(t)$ often contains a sequence of factors $(t+m)$ for $m=1,2,\dots,M$, with M some non-negative integer. In the above example $M=2$. More generally, both the value of M and the degree, d of the K-polynomial appear to be difficult to predict from a knowledge of n, λ and μ. Although the hive model does give an immediate bound, $d \le (n-1)(n-2)/2$, on the degree, it can be seen in the above $n=9$ example that we have $d=6$, so that the bound is far from being saturated.

It is also found that under certain circumstances there exists a factorisation of the form

$$P_{\lambda\mu}(t) = P_{\sigma\zeta}(t)\ P_{\tau\eta}(t) \tag{1.2}$$

for some σ, τ, ζ and η such that $\lambda = (\sigma,\tau)$ and $\mu = (\zeta,\eta)$, where the notation is intended to signify that the list of parts of λ are simply those of σ followed by those of τ, and similarly for μ. In the above example we have $P_{9522221,555311111}(t) = P_{9522,5553}(t)\ P_{221,11111}(t)$ with

$$P_{9522,5553}(t) = \frac{1}{2}(t+1)(5t+2) \quad\text{and}\quad P_{221,11111}(t) = \frac{1}{12}(t+1)(t+2)(t^2+3t+6). \tag{1.3}$$

In this presentation, our intention is to exploit the hive model to study the properties of stretched Kostka coefficient polynomials. We first derive combinatorially the precise conditions under which such K-polynomials factorise as products of certain primitive K-polynomials. In the case of any primitive K-polynomial we then show how to determine the precise range of values, $1 \le m \le M$, such that the K-polynomial contains a factor $(t+m)$. This is done by giving an interpretation of $P_{\lambda\mu}(t)$ for negative integer values of t. Finally, we obtain a formula for an explicit bound on the degree d of a primitive K-polynomial which we conjecture is always saturated.

Our analysis covers not only the Kostka coefficients $K_{\lambda\mu}$ and the K-polynomials $P_{\lambda\mu}(t)$ in which μ is a partition, but also $K_{\lambda\beta}$ and $P_{\lambda\beta}(t)$ in which β is, more generally, a weight. The combinatorial proof of the factorisation theorem in this case is presented in an Appendix.

2. Kostka coefficients

Let n be a fixed positive integer. For any $\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathbb{Z}^n$ let $|\alpha| = \alpha_1 + \cdots + \alpha_n$ be the weight of α, let $\#\alpha = n$ be the number of components of α and let the symmetric group S_n act naturally on the components of α. Let $\lambda = (\lambda_1,\ldots,\lambda_n)$ be a partition of length $\ell(\lambda) \le n$ and weight $|\lambda|$ with $\lambda_1 \ge \cdots \ge \lambda_n \ge 0$. Then $|\lambda| = \lambda_1 + \cdots + \lambda_n$ with $\lambda_i > 0$ for $i \le \ell(\lambda)$ and $\lambda_i = 0$ for $i > \ell(\lambda)$. It is sometimes convenient to write λ in terms of its distinct parts, that is to set $\lambda = (\kappa_1^{v_1},\ldots,\kappa_m^{v_m})$ with $\kappa_1 > \cdots > \kappa_m \ge 0$ and $v_j > 0$ for $j = 1,\ldots,m$. Then $|\lambda| = v_1\kappa_1 + \cdots + v_m\kappa_m$ with $v_1 + \cdots + v_m = n$, and $\ell(\lambda) = n$ if $\kappa_m > 0$ and $\ell(\lambda) = n - v_m$ if $\kappa_m = 0$.

Definition 2.1. *Let $\mathbf{x} = (x_1,\ldots,x_n)$. Then to each partition λ with $\ell(\lambda) \le n$ there corresponds a Schur function $s_\lambda(\mathbf{x})$ defined by*

$$s_\lambda(\mathbf{x}) = \frac{\left| x_i^{n+\lambda_j-j} \right|_{1\le i,j\le n}}{\left| x_i^{n-j} \right|_{1\le i,j\le n}}. \tag{2.1}$$

Since $s_\lambda(\mathbf{x})$ is a ratio of two alternants, it is a symmetric polynomial in the components $x_1,\ldots,x_n$ of $\mathbf{x}$. Indeed, the Schur functions $\{s_\lambda(\mathbf{x}) \,|\, \ell(\lambda) \le n\}$ constitute a linear basis of the algebra of symmetric polynomials in the set of variables $\{x_1,\ldots,x_n\}$.

Definition 2.2. *The expansion of each Schur function in terms of monomials takes the form*

$$s_\lambda(\mathbf{x}) = \sum_\beta K_{\lambda\beta}\mathbf{x}^\beta, \tag{2.2}$$

where the summation is over all $\beta = (\beta_1,\ldots,\beta_n)$ and $\mathbf{x}^\beta = x_1^{\beta_1}\cdots x_n^{\beta_n}$. The coefficients $K_{\lambda\beta}$ are known as Kostka coefficients.

The fact that $s_\lambda(\mathbf{x})$ is symmetric implies that the coefficients $K_{\lambda\beta}$ are insensitive to the permutations of the components of β, that is they possess:

Property 2.3. *For all $w \in S_n$ we have $K_{\lambda,w(\beta)} = K_{\lambda\beta}$, and there must exist a partition μ such that $w(\beta) = \mu$ for some $w \in S_n$, in which case $K_{\lambda\beta} = K_{\lambda\mu}$.*

To specify those partitions λ and weights β for which $K_{\lambda\beta}$ is non-zero it is convenient to introduce the notion of partial sums of the parts of partitions and weights, and the dominance partial order.

Definition 2.4. *For any partition $\nu = (\nu_1, \nu_2, \ldots, \nu_n)$ and any weight $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, their partial sums are defined by*

$$ps(\nu)_i = \nu_1 + \nu_2 + \cdots + \nu_i \quad \text{and} \quad ps(\alpha)_i = \alpha_1 + \alpha_2 + \cdots + \alpha_i \quad \text{for all } i = 1, 2, \ldots, n. \tag{2.3}$$

More generally, for any subset $I = \{i_1, i_2, \ldots, i_r\}$ of $N = \{1, 2, \ldots, n\}$ of cardinality $\#I = r$ with $1 \le r \le n$ let

$$ps(\nu)_I = \nu_{i_1} + \nu_{i_2} + \cdots + \nu_{i_r} \quad \text{and} \quad ps(\alpha)_I = \alpha_{i_1} + \alpha_{i_2} + \cdots + \alpha_{i_r}. \tag{2.4}$$

Definition 2.5. *Given partitions λ and μ of lengths $\ell(\lambda), \ell(\mu) \le n$, then λ is said to dominate μ, and we write $\lambda \succeq \mu$, or more precisely $\lambda \succeq_n \mu$, if*

$$ps(\lambda)_i \ge ps(\mu)_i \quad \text{for all} \quad i = 1, 2, \ldots, n \text{ and } |\lambda| = |\mu|. \tag{2.5}$$

Moreover, λ is said to strongly dominate μ, and we write $\lambda \succ \mu$, or more precisely $\lambda \succ_n \mu$, if

$$ps(\lambda)_i > ps(\mu)_i \quad \text{for all} \quad i = 1, 2, \ldots, n-1 \quad \text{and} \quad |\lambda| = |\mu|. \tag{2.6}$$

We now have the following important condition for the non-vanishing of Kostka coefficients:

Theorem 2.6. *Let λ and μ be partitions of lengths $\ell(\lambda), \ell(\mu) \le n$, and let $N = \{1, 2, \ldots, n\}$. Then*

$$K_{\lambda\mu} > 0 \iff \lambda \succeq_n \mu \tag{2.7}$$

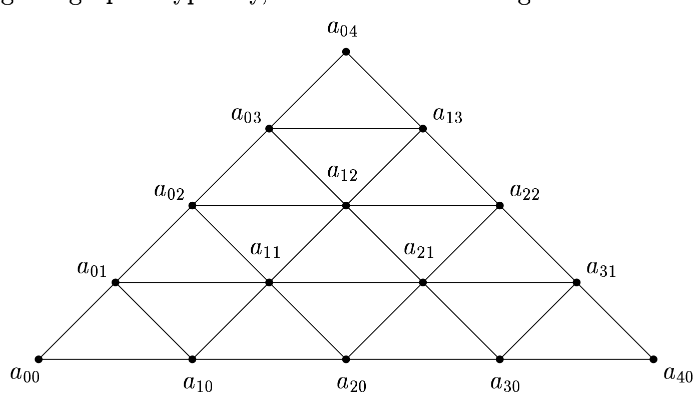
More generally, let λ be a partition of length $\ell(\lambda) \le n$ and let β be a weight with $\#\beta = n$. Then

$$K_{\lambda\beta} > 0 \iff |\lambda| = |\beta| \quad \text{and} \quad ps(\lambda)_i \ge ps(\beta)_I \quad \text{for all} \quad I \subseteq N \quad \text{with} \quad i = \#I > 0. \tag{2.8}$$

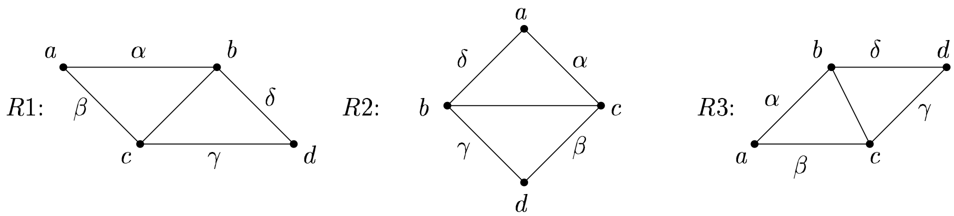
The first set of conditions (2.7) is well-known (see for example [JK]p44), and the second set of conditions (2.8) is a simple corollary following from the fact that if $\mu = w(\beta)$ for some $w \in S_n$ then $ps(\mu)_i = ps(\beta)_I$ for some $I \subseteq N$ with $i = \#I$, and $ps(\mu)_i \ge ps(\beta)_J$ for all $J \subseteq N$ with $i = \#J$.

3. The hive model

An n-hive is an array of numbers a_{ij} with $0 \le i, j \le n$ placed at the vertices of an equilateral triangular graph. Typically, for $n = 4$ their arrangement is as shown below:



Such an n-hive is said to be an integer hive if all of its entries are non-negative integers. Neighbouring entries define three distinct types of rhombus, each with its own constraint condition.



In each case, with the labelling as shown, the hive condition takes the form:

$$b + c \geq a + d \tag{3.1}$$

In what follows we make use of edge labels more often than vertex labels. Each edge in the hive is labelled by means of the difference, $\epsilon = q - p$, between the labels, p and q, on the two vertices connected by this edge, with q always to the right of p. In all the above cases, with this convention, we have $\alpha + \delta = \beta + \gamma$, and the hive conditions take the form:

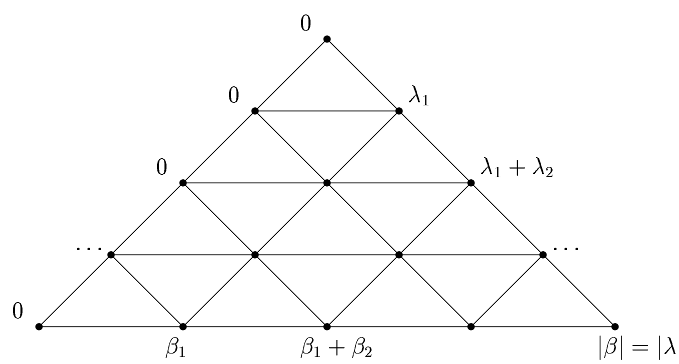
$$\alpha \geq \gamma \quad \text{and} \quad \beta \geq \delta, \tag{3.2}$$

where, of course, either one of the conditions $\alpha \geq \gamma$ or $\beta \geq \delta$ is sufficient to imply the other.

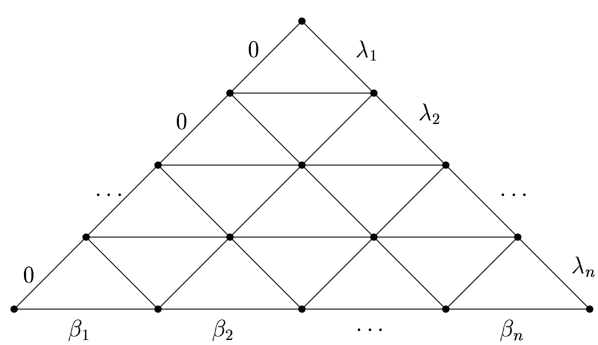
Although, for completeness, we have included hive conditions for all three types of rhombus that appear in a general hive, it is only the type R1 and R2 hive conditions that apply to what we call K-hives. For such K-hives the type R3 hive conditions will, in general, be violated.

Definition 3.1. *A K-hive is an integer hive satisfying the hive conditions (3.1), or equivalently (3.2) for all its constituent rhombi of type R1 and R2 (but not R3), with border labels determined by the zero partition or weight* $0 = (0, 0, \ldots, 0)$ *with* $\#0 = n$*, a partition* λ *with* $\ell(\lambda) \leq n$ *and a weight* β *with* $\#\beta = n$*, satisfying the constraint* $|\lambda| = |\beta|$*, in such a way that* $a_{0i} = 0$ *for* $i = 0, 1, \ldots, n$*,* $a_{j,n-j} = ps(\lambda)_j = \lambda_1 + \lambda_2 + \cdots + \lambda_j$ *for* $j = 1, 2, \ldots, n$*,* $a_{k,0} = ps(\beta)_k = \beta_1 + \beta_2 + \cdots + \beta_k$ *for* $k = 1, 2, \ldots, n$.

Schematically, we have



Alternatively, in terms of edge labels we have:



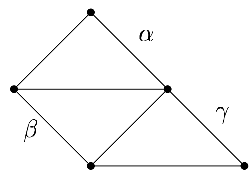
Proposition 3.2. [KTT] *The Kostka coefficient $K_{\lambda\beta}$ is the number of K-hives with border labels determined as above by λ and β.*

For example, if $\lambda = (3,2)$ and $\beta = (2,1,2)$ the corresponding K-hives take the form:

$$\begin{array}{ccccccc} & & & 0 & & & \\ & & 0 & & 3 & & \\ & 0 & & a & & 5 & \\ 0 & & 2 & & 3 & & 5 \end{array} \tag{3.3}$$

where for the sake of clarity all the hive edges have been omitted. Since the only integer values of a satisfying the hive conditions (3.1) for all the constituent rhombi of type R1 and R2 are $a = 2$ and $a = 3$, it follows that $K_{32,212} = 2$.

When expressed in terms of edge labels, the hive conditions (3.2) for all constituent rhombi of types R1 and R2 imply that in every K-hive the edge labels along any line parallel to the right-hand edge of the hive are weakly decreasing from top left to bottom right. This can be seen from the following 5-vertex sub-diagram.



The edge conditions on the rhombi of type R1 and R2 in the above diagram give $\beta \geq \gamma$ and $\alpha \geq \beta$, respectively, so that $\alpha \geq \gamma$, as claimed. This is of course consistent with the edges of the right-hand boundary of each K-hive being specified by a partition λ.

4. Stretched coefficients

4.1. **Polynomial conjectures.** Now we are in a position to define and evaluate stretched Kostka coefficients. The partition obtained from $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ by multiplying all of its parts by the same positive integer, t, is denoted by $t\lambda = (t\lambda_1, t\lambda_2, \ldots, t\lambda_p)$. With this notation, we refer to $K_{t\lambda,t\beta}$ as stretched Kostka coefficients, where t is said to be the stretching parameter. It is not difficult to evaluate these stretched coefficients, particularly through the use of the hive model, for a range of positive integer values of t.

For example, for $\lambda = (3,2)$ and $\mu = (1,1,1,1,1)$ the corresponding stretched Kostka coefficients are given by

$$K_{t\lambda,t\mu} = \frac{1}{2}(t+1)(t^2+2t+2). \tag{4.1}$$

The generating function for these coefficients takes the form:

$$F_{\lambda\mu}(z) = \sum_{t=0}^{\infty} K_{t\lambda,t\mu}\ z^t = \frac{1+z+z^2}{(1-z)^4}. \tag{4.2}$$

On the basis of this and many other examples we were led to the following:

Conjecture 4.1. [KTT] *For all partitions λ and μ such that $K_{\lambda\mu} > 0$ there exists a polynomial $P_{\lambda\mu}(t)$ in t with positive rational coefficients such that $P_{\lambda\mu}(0) = 1$ and $P_{\lambda\mu}(t) = K_{t\lambda,t\mu}$ for all positive integers t.*

Conjecture 4.2. [KTT] *Given that the degree of the polynomial $K_{t\lambda,t\mu}$ is d, the generating function for $K_{t\lambda,t\mu}$ takes the form $F_{\lambda\mu}(z) = G_{\lambda\mu}(z)/(1-z)^{d+1}$, with $G_{\lambda\mu}(z)$ a polynomial of degree $\leq d$ having non-negative integer coefficients.*

Much of these conjectures has now been proved. In particular the fact that $P_{\lambda\mu}(t)$ is polynomial appears to have been proved first by Kirillov and Reshetekin [KR], with a more recent proof provided by Billey, Guillemin and Rassart [BGR]. In fact all that remains to be answered are the questions of the positivity of the coefficients in $P_{\lambda\mu}(t)$ and $G_{\lambda\mu}(z)$.

Before looking in more detail at the nature of the polynomials $P_{\lambda\mu}(t)$ we note:

Saturation Condition 4.3. [KTT]

$$K_{t\lambda,t\mu} > 0 \iff K_{\lambda\mu} > 0. \tag{4.3}$$

Proof This can be seen immediately from Theorem 2.6 by noting that $|t\lambda| = t\,|\lambda|$ and $ps(t\lambda)_r = t\,ps(\lambda)_r$ for all partitions λ and all r. It follows that

$$K_{\lambda\mu} > 0 \iff |\lambda| = |\mu| \text{ and } \lambda \succeq \mu \iff |t\lambda| = |t\mu| \text{ and } t\lambda \succeq t\mu \iff K_{t\lambda,t\mu} > 0, \tag{4.4}$$

as required. □

Turning to the Conjecture 4.1 itself, the following key component has been established:

Theorem 4.4. [K, BGR] *Let λ and μ be partitions of lengths $\ell(\lambda), \ell(\mu) \leq n$ such that $K_{\lambda\mu} > 0$. Then $P_{\lambda\mu}(t) = K_{t\lambda,t\mu}$ is a polynomial of degree at most $(n-1)(n-2)/2$ in t.*

Clearly, Property 2.3 then implies as a corollary of Theorem 4.4 that if $K_{\lambda\beta} > 0$ then $P_{\lambda\beta}(t) = K_{t\lambda,t\beta}$ is also a polynomial in t with the same upper bound on its degree.

5. Factorisation

It was noted in the work of Berenstein and Zelevinsky [BZ] that Kostka coefficients factorise. Here we give a combinatorial proof of this factorisation property based on the use of K-hives.

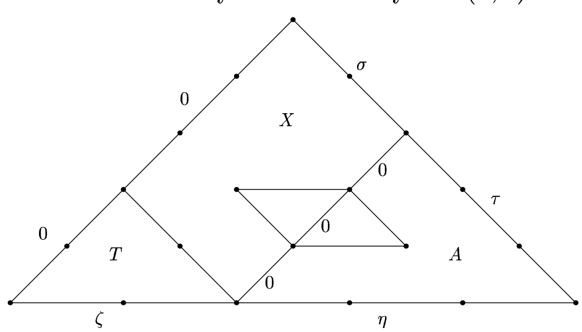
First we deal with the case $K_{\lambda\mu}$ where λ and μ are partitions of the same weight, $|\lambda| = |\mu|$, with $\ell(\lambda), \ell(\mu) \leq n$. Then following Berenstein and Zelevinsky, we say that $K_{\lambda\mu}$ is *primitive* if $\lambda \succ_n \mu$. If $K_{\lambda\mu}$ is not primitive then we have the following factorisation:

Theorem 5.1. *Let λ, μ be partitions such that $|\lambda| = |\mu|$ and $\ell(\lambda), \ell(\mu) \leq n$, with $\lambda \succeq_n \mu$ and $ps(\lambda)_r = ps(\mu)_r$ for some r for which $1 \leq r < n$. Let $\lambda = (\sigma, \tau)$ and $\mu = (\zeta, \eta)$, with $\ell(\sigma) = \ell(\zeta) = r$ and $\ell(\tau) = \ell(\eta) = n - r$, then*

$$K_{\lambda\mu} = K_{\sigma\zeta}\, K_{\tau\eta}. \tag{5.1}$$

Proof First it should be noted that the hypothesis $\lambda \succeq_n \mu$ implies that $K_{\lambda\mu} > 0$. Moreover, we have $ps(\sigma)_i = ps(\lambda)_i \geq ps(\mu)_i = ps(\zeta)_i$ for $i = 1, 2, \ldots, r-1$ and $|\sigma| = ps(\sigma)_r = ps(\lambda)_r = ps(\mu)_r = ps(\zeta)_r = |\zeta|$. Thus $\sigma \succeq_r \zeta$ and hence $K_{\sigma\zeta} > 0$. In addition we have $ps(\tau)_i = ps(\lambda)_{r+i} - ps(\lambda)_r = ps(\lambda)_{r+i} - |\sigma| \geq ps(\mu)_{r+i} - |\zeta| = ps(\mu)_{r+i} - ps(\mu)_i = ps(\eta)_i$ for $i = 1, 2, \ldots, n-r$, and $|\tau| = |\lambda| - |\sigma| = |\mu| - |\zeta| = |\eta|$. Thus $\tau \succeq_{n-r} \eta$ and hence $K_{\tau,\eta} > 0$.

Now consider the K-hive with boundary determined by $\lambda = (\sigma, \tau)$

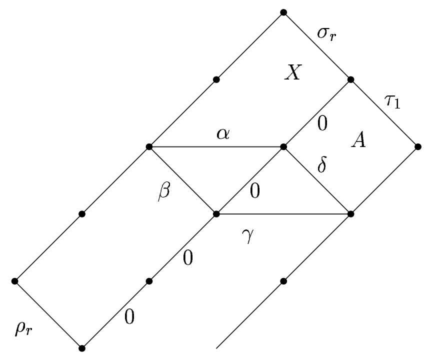


The rules about edge sums are such that on the boundary of the triangular region T we have $|\zeta| = |0| + |\rho| = |\rho|$. Since $|\zeta| = |\sigma|$ we have $|\rho| = |\sigma|$. The same rules applied to the parallelogram X then imply that the sum of the edge lengths on the boundary between X and A must be $|0| + |\sigma| - |\rho| = 0$. Since X can be viewed as a collection of rhombi of type R2, applying the hive condition $\beta \geq \delta$ to each of these rhombi implies that the edge lengths on the lower right boundary of X must be non-negative. Since their sum is zero, they must all be zero, as indicated in the diagram. The $\alpha \geq \gamma$ hive condition for R2 may then be applied to these same rhombi constituting X, yielding the constraints $\rho_i \leq \sigma_i$ for $i = 1, 2, \ldots, r$. Since $|\rho| = |\sigma|$, we must have $\rho_i = \sigma_i$ for $i = 1, 2, \ldots, r$, that is $\rho = \sigma$.

Thus each K-hive H with boundary λ and μ consists of a parallelogram X in which all edge lengths are fixed, together with a K-hive T with boundary σ and ζ, and a second K-hive A with boundary τ and η. The hive conditions for H imply those appropriate to T and A.

In order to prove the required factorisation (5.1) it only remains to show that combining all possible hives T and A with a parallelogram X of the appropriate boundary gives a hive H in which all the hive conditions corresponding to rhombi crossing the boundaries of T with X, and A with X, are automatically satisfied. Since, for K-hives we only use the rhombi of type R1 and R2, it is clear that no such rhombus crosses the boundary between T and X, while only rhombi of type R1 cross the boundary between A and X. One such rhombus has been shown above.

We may look in more detail at such a rhombus by means of the following diagram which shows a strip one edge length wide on either side of the $A-X$ boundary.



With the edge labels as shown, the R2 rhombus condition applied in the region A just below the $A-X$ boundary gives $\delta \leq \tau_1$, while the same rhombus condition applied in the region X just above the $A-X$ boundary gives $\rho_r \leq \beta \leq \sigma_r$. Since $\rho_r = \sigma_r$ and $\tau_1 = \lambda_{r+1} \leq \lambda_r = \sigma_r$, we have $\delta \leq \tau_1 \leq \sigma_r = \beta$. However, this is just what is required to satisfy the R1 condition (3.2).

This completes the proof of the K-hive factorisation theorem. □

Repeated use of this theorem allows any non-primitive Kostka coefficient $K_{\lambda\mu}$ to be expressed as a unique product of primitive Kostka coefficients $K_{\sigma\zeta}$. It is only necessary at each stage to factor out the term meeting the hypotheses of the above Theorem 5.1 with the smallest possible value of r.

Applying the above argument to stretched Kostka coefficients immediately gives the following:

Theorem 5.2. *Let λ, μ be partitions such that $|\lambda| = |\mu|$ and $\ell(\lambda), \ell(\mu) \leq n$, with $\lambda \succeq_n \mu$ and $ps(\lambda)_r = ps(\mu)_r$ for some r for which $1 \leq r < n$. Let $\lambda = (\sigma, \tau)$ and $\mu = (\zeta, \eta)$, with $\ell(\sigma) = \ell(\zeta) = r$ and $\ell(\tau), \ell(\eta) \leq n - r$, then for all positive integers t we have*

$$P_{\lambda\mu}(t) = P_{\sigma\zeta}(t)\ P_{\tau,\eta}(t). \tag{5.2}$$

An example of this type has been provided in the introduction in (1.3).

It is quite instructive from a combinatorial point of view, although not strictly necessary by virtue of Property 2.3, to extend the above analysis to the case of Kostka coefficients $K_{\lambda\beta}$ for which the weight β is not necessarily a partition. In this case $K_{\lambda\beta}$ is said to be primitive if $|\lambda| = |\beta|$ and $ps(\lambda)_r \geq ps(\beta)_I$ for any proper subset I of $N = \{1, 2, \ldots, n\}$ with $r = \#I$.

In the non-primitive case, we have:

Theorem 5.3. *Let λ be a partition with $\ell(\lambda) \leq n$, and let $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ be a weight such that $|\lambda| = |\beta|$ and $ps(\lambda)_i \geq ps(\beta)_I$ for any $I \subset N$, with $i = \#I$. Then $K_{\lambda\beta}$ is not primitive if there exists a proper subset $I = \{i_1, i_2, \ldots, i_r\}$ of N such that $ps(\lambda)_r = ps(\beta)_I$ with $r = \#I$ for some r for which $1 \leq r < n$. Let the complement of I in N be denoted by $\overline{I} = \{j_1, j_2, \ldots, j_{n-r}\}$. In such a case let $\lambda = (\sigma, \tau)$, $\zeta = (\beta_{i_1}, \beta_{i_2}, \ldots, \beta_{i_r})$ and $\eta = (\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_{n-r}})$, so that $\beta = \zeta \cup \eta$, with ζ and η weights, not necessarily partitions, with $\#\zeta = r$ and $\#\eta = n - r$, satisfying the partial sum constraints $|\zeta| = |\sigma|$ and $|\eta| = |\tau|$. Then*

$$K_{\lambda\beta} = K_{\sigma\zeta}\ K_{\tau\eta}. \tag{5.3}$$

A complete combinatorial proof of this theorem, based on the use of hives, has not been included due to space limitations.

Once again scaling everything by t is straightforward. If $K_{\lambda\beta}$ is primitive then so is $K_{t\lambda,t\beta}$ since all the partial sum conditions are scaled by the same factor t. By the same token the factorisation occurs in the stretched non-primitive case just as it does in the unstretched non-primitive case – all boundary edges are simply scaled by t, and as we have shown, it is these boundary edges that completely determine the factorisation. Thus under the hypotheses of Theorem 5.3 we have

$$P_{\lambda\beta}(t) = P_{\sigma\varsigma}(t)\ P_{\tau\eta}(t). \tag{5.4}$$

This is illustrated for $n=6$ and $r=3$ by $\lambda=(9,6,4,4,2,0)$ and $\beta=(2,6,3,7,6,1)$, for which $ps(\lambda)_3 = 19 = ps(\beta)_I$ with $I=\{2,4,5\}$. In this case $\sigma=(9,6,4)$, $\tau=(4,2,0)$, $\zeta=(6,7,6)$ and $\eta=(2,3,1)$. Correspondingly we find

$$P_{96442,263761} = (t+1)(2t+1),\ \ P_{964,676} = (2t+1),\ \ P_{42,231} = (t+1), \tag{5.5}$$

thereby exemplifying the factorisation (5.4).

It is interesting to note that the above factorisation of Kostka coefficients can be extended to the case of q-dependent Kostka-Foulkes polynomials, $K_{\lambda\mu}(q)$. These have a combinatorial definition in terms of the charge statistic on semistandard Young tableaux [M]p242. By way of illustration, in the case of the above example we find

$$K_{t(96442),t(766321)}(q) = q^{3t}\,\frac{q^{t+1}-1}{q-1}\,\frac{q^{2t+1}-1}{q-1} \tag{5.6}$$

$$K_{t(964),t(766)}(q) = q^{2t}\,\frac{q^{2t+1}-1}{q-1} \tag{5.7}$$

$$K_{t(42),t(321)}(q) = q^{t}\,\frac{q^{t+1}-1}{q-1}. \tag{5.8}$$

6. The zeros of stretched Kostka polynomials

As has been noted the stretched Kostka polynomials $P_{\lambda\mu}(t)$ contain factors $(t+m)$ for some sequence of values $m=1,2,\ldots,M$ for some positive integer M. This is no accident. In this section we describe a method of determining M for stretched Kostka polynomials. It will be recalled that the Kostka coefficients are defined in terms of Schur functions by (2.2). It then follows from (2.1) that $P_{\lambda\mu}(t)=K_{t\lambda,t\mu}$ is the coefficient of $x_1^{t\mu_1}x_2^{t\mu_2}\cdots x_n^{t\mu_n}$ in the expansion of

$$s_\lambda(\mathbf{x}) = \frac{\left|x_i^{t\lambda_j+n-j}\right|}{\left|x_i^{n-j}\right|} \tag{6.1}$$

This may be readily extended to the case $t=-m$ with m a positive integer. Then $P_{\lambda\mu}(-m) = K_{-m\lambda,-m\mu}$ is the coefficient of $x_1^{-m\mu_1}x_2^{-m\mu_2}\cdots x_n^{-m\mu_n}$ in the expansion of $s_{-m\lambda}(\mathbf{x})$. However

$$s_{-m\lambda}(\mathbf{x}) = \frac{\left|x_i^{-m\lambda_j+n-j}\right|}{\left|x_i^{n-j}\right|} = \frac{\left|x_i^{-m\lambda_{n-k+1}-n+k}\right|}{\left|x_i^{-n+k}\right|}., \tag{6.2}$$

where first x_i^{n-1} has been extracted as a common factor from the ith row of each determinant for $i=1,2,\ldots,n$ and cancelled from numerator and denominator, and then j replaced by $k=n-j+1$ with an appropriate reversal of order of the columns in both determinants. If we now set $\overline{x}_i = x_i^{-1}$ for $i=1,2,\ldots,n$ and $\overline{\mathbf{x}}=(\overline{x}_1,\overline{x}_2,\ldots,\overline{x}_n)$, this gives

$$s_{-m\lambda}(x_1,x_2,\ldots,x_n) = \frac{\left|\overline{x}_i^{m\lambda_{n-k+1}+n-k}\right|}{\left|\overline{x}_i^{n-k}\right|} = s_{m\lambda_n,\ldots,m\lambda_2,m\lambda_1}(\overline{\mathbf{x}}). \tag{6.3}$$

Since $x_1^{-m\mu_1}x_2^{-m\mu_2}\cdots x_n^{-m\mu_n} = \overline{x}_1^{m\mu_1}\overline{x}_2^{m\mu_2}\cdots\overline{x}_n^{m\mu_n}$ it follows that $P_{\lambda\mu}(-m)=K_{-m\lambda,-m\mu}$ is the coefficient of $\overline{x}_1^{m\mu_1}\overline{x}_2^{m\mu_2}\cdots\overline{x}_n^{m\mu_n}$ in the expansion of the right hand side. It then follows, replacing the dummy variables $\overline{x}_i$ by x_i for $i=1,2,\ldots,n$, that $P_{\lambda\mu}(-m)=K_{-m\lambda,-m\mu}$ is the coefficient of

$x_1^{m\mu_1}x_2^{m\mu_2}\cdots x_n^{m\mu_n}$ in the expansion of $s_{m\tilde{\lambda}}(\mathbf{x})$, where $\tilde{\lambda}=(\lambda_n,\ldots,\lambda_2,\lambda_1)$ is the weight obtained by reversing the order of the parts of λ, including any trailing zeros. Then to exploit the above it is only necessary to standardise

$$s_{m\tilde{\lambda}}(\mathbf{x})=s_{m\lambda_n,\ldots,m\lambda_2,m\lambda_1}(x_1,x_2,\ldots,x_n)=\frac{\left|x_i^{m\lambda_{n-k+1}+n-k}\right|}{\left|x_i^{n-k}\right|}. \tag{6.4}$$

This is carried out in the usual way [L] by reordering the columns of the numerator determinant. However there are two quite different possible outcomes:

$$s_{m\tilde{\lambda}}(\mathbf{x})=\begin{cases}0 & \text{case (i);}\\ \eta_\rho\, s_\rho(\mathbf{x}) \text{ with } \eta_\rho=\pm 1 & \text{case (ii).}\end{cases} \tag{6.5}$$

For example, in the case $n=5$ and $\lambda=(4,2,1,0,0)$ we have $\tilde{\lambda}=(0,0,1,2,4)$. For $m=1$ this gives $s_{m\tilde{\lambda}}(\mathbf{x})=s_{0,0,1,2,4}(\mathbf{x})=0$. Similarly for $m=2$ we have $s_{m\tilde{\lambda}}(\mathbf{x})=s_{0,0,2,4,8}(\mathbf{x})=0$. However, for $m=3$ we obtain $s_{m\tilde{\lambda}}(\mathbf{x})=s_{0,0,3,6,12}(\mathbf{x})=-s_{8,4,3,3,3}(\mathbf{x})$, while for $m=5$ we have $s_{m\tilde{\lambda}}(\mathbf{x})=s_{0,0,4,8,16}(\mathbf{x})=-s_{12,6,4,3,3}(\mathbf{x})$.

More generally, the case (i) result 0 arises if and only if

$$m\,\lambda_{n-k+1}+n-k=m\,\lambda_{n-l+1}+n-l \tag{6.6}$$

for some k and l such that $1\le l<k\le n$. In all other cases the formula given in case (ii) applies for some partition ρ of length at most n and weight $m|\lambda|$, and $\eta_\rho=\pm1$ is a sign factor recording the number of transpositions of columns of the numerator determinant of (6.4) required to standardise the Schur function.

It follows that we can expect there to be a possibility of two types of zero of $P_{\lambda\mu}(t)$ for $t=-m$: type (i) associated with case (i) of (6.5), and type (ii) associated under certain conditions on ρ and $m\,\mu$ with case (ii). Indeed adopting the notation of (6.5) we have the following:

Proposition 6.1. *Let λ and μ be such that $K_{\lambda\mu}$ is primitive. Then $P_{\lambda\mu}(t)=K_{t\lambda,t\mu}$ contains a factor $(t+m)$ if and only if either case (i) applies and*

$$m=(j-i)/(\lambda_i-\lambda_j) \tag{6.7}$$

for some i and j such that $1\le i<j\le n$, or case (ii) applies and

$$ps(\rho)_k<m\,ps(\mu)_k \tag{6.8}$$

for some k such that $1\le k<n$.

Proof We have already noted that $P_{\lambda\mu}(-m)=K_{-m\lambda,-m\mu}$ is the coefficient of $x_1^{m\mu_1}x_2^{m\mu_2}\cdots x_n^{m\mu_n}$ in $s_{m\tilde{\lambda}}(\mathbf{x})=s_{m\lambda_n,\ldots,m\lambda_2,m\lambda_1}(x_1,x_2,\ldots,x_n)$. Thus if case (i) of (6.5) applies then $P_{\lambda\mu}(-m)=0$, while if case (ii) applies then $P_{\lambda\mu}(-m)=\eta_\rho K_{\rho,m\mu}$. The conditions for case (i) to apply are given in (6.6). However, setting $i=n-k+1$ and $j=n-l+1$ leads immediately to the required case (i) conditions (6.7) for $P_{\lambda\mu}(-m)=0$. Turning to case (ii) of (6.5) $P_{\lambda\mu}(-m)=\eta_\rho\,K_{\rho,m\mu}$ will be zero by virtue of Theorem 2.6 if any one of the partial sum conditions $ps(\rho)_k\ge m\,ps(\mu)_k$ for all $k=1,2,\ldots,n$ is violated. Since $|\rho|=m\,|\mu|$ this leaves precisely the required case (ii) conditions (6.8) for $P_{\lambda\mu}(-m)=0$. □

We may extend our earlier example with $n=5$ and $\lambda=(4,2,1,0,0)$ by taking $\mu=(3,1,1,1,1)$ for illustrative purposes. We have already seen that there are type (i) zeros, independent of μ that arise for $m=1$ and $m=2$. These arise from (6.7) for the pairs $(i,j)=(3,4)$ and $(3,5)$, respectively, leading to $(j-i)/(\lambda_i-\lambda_j)=(4-3)/(1-0)=1$ and $(5-3)/(1-0)=2$. For $m=3$ we have a type (ii) zero since $P_{\lambda\mu}(-3)=\eta_\rho K_{\rho,m\mu}=-K_{84333,93333}=0$ by virtue of the fact that $ps(\rho)_1=8<9=ps(m\mu)_1$. On the other hand for $m=4$ we have $P_{\lambda\mu}(-4)=\eta_\rho K_{\rho,m\mu}=-K_{12\,6433,12\,4444}=-3$ which is non-zero, and the same will be true for all $m\ge4$.

In view of our earlier remarks about the sequence of zeros of $P_{\lambda\mu}(t)$ for $t=-m$ with $m=1,2,\ldots,M$, it comes as no surprise that we are able to elaborate on the above and establish that the zeros are indeed consecutive and that for any primitive $K_{\lambda\mu}$ we can always find a finite $M>0$ such that $P_{\lambda\mu}(-m)=0$ for all $m=1,2,\ldots,M$, but thereafter $P_{\lambda\mu}(-m)\ne0$ for all $m>M$.

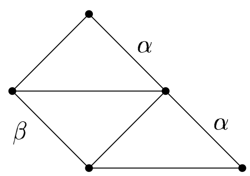
All this is borne out in our example with $n=5$, $\lambda=(4,2,1,0,0)$ and $\mu=(3,1,1,1,1)$ for which

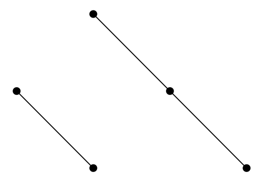
$$P_{\lambda\mu}(t)=K_{t\lambda,t\mu}=\frac{1}{60}(t+1)(t+2)(t+3)(3t^2+7t+10), \tag{6.9}$$

for which $M=3$ and we have consecutive zeros at $t=-m$ with $m=1,2,3$. Of these, as we have seen, the first two are type (i) and the third is type (ii). The first non-zero case occurs with $t=-4$ and we obtain from the explicit formula (6.9) the result $P_{\lambda\mu}(-4)=-3$ that we had found earlier. In fact for $t=-m$ it is clear that (6.9) gives $P_{\lambda\mu}(-m)<0$ for all $m>M=3$.

7. The degrees of stretched Kostka polynomials

As we have seen the calculation of stretched Kostka polynomials may be reduced to that of calculating these polynomials in primitive cases only. Even so the task may be combinatorially formidable. In any given case a knowledge of the degree of the polynomial would be extremely advantageous. Here we establish an upper bound on this degree by means of the following rather innocuous looking result. Let the edge labelling of a particular 5-vertex subset of a K-hive be as shown below in the left hand diagram, with two identical labels α.



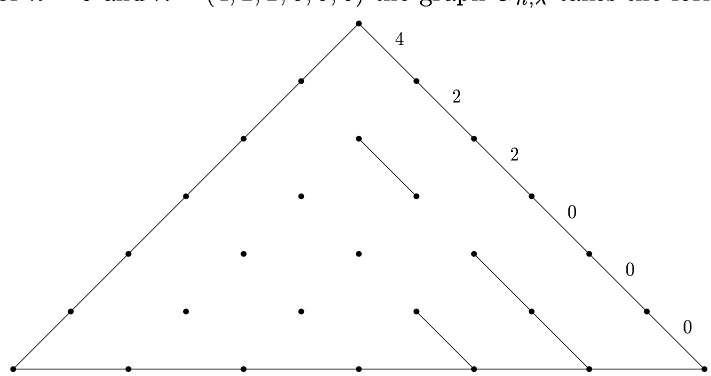


This diagram contains rhombi of type R1 and R2. The hive conditions (3.2) for the former imply $\alpha\le\beta$, while those for the latter give $\beta\le\alpha$. It follows that $\beta=\alpha$. This result is displayed more simply by deleting all the edges except those sharing the same label α, and suppressing the label itself. This gives the right hand diagram where the equality of a pair of edge labels in a linear sequence forces an identical edge label in a neighbouring line.

Applying these notions to the case of K-hives with boundaries of length n and with border labels determined by 0, λ and μ, it follows from the above that any equalities of successive parts of λ propagate as equalities of edge labels within each possible K-hive. To be more precise let all the λ-boundary edges be labelled by the parts of λ. If any sequence of parts of λ share the same value, say α, then we can identify an equilateral sub-hive having the sequence of equally labelled edges as one boundary, with its other boundaries parallel to the 0 and μ-boundaries of the original hive. Within this sub-hive all the vertices along lines parallel to the λ-boundary are to be connected by edges indicating that in any K-hive the differences in values between neighbouring entries along these lines are all α.

This process is to be repeated for all sequences of equal edge labels along the λ-boundary. Finally, all neighbouring vertices on all three boundaries are to be connected by edges. In this way we arrive at a graph $G_{n;\lambda}$ that depends only upon n and λ.

For example, for $n=6$ and $\lambda=(4,2,2,0,0,0)$ the graph $G_{n;\lambda}$ takes the form:



With this notation we have:

Proposition 7.1. *Let λ and μ be partitions such that $K_{\lambda\mu} > 0$. Let $\deg(P_{\lambda\mu}(t))$ be the degree of the corresponding stretched K-polynomial $P_{\lambda\mu}(t) = K_{t\lambda,t\mu}$. Let $d(G_{n;\lambda})$ be the number of connected components of the graph $G_{n;\lambda}$ that are not connected to the boundary. Then*

$$\deg(P_{\lambda\mu}(t)) \leq d(G_{n;\lambda}). \tag{7.1}$$

Moreover, if $\lambda = (w_1^{v_1}, w_2^{v_2}, \ldots, w_t^{v_t})$, with $w_1 > w_2 > \cdots > w_t \geq 0$, $v_s \geq 1$ for all $s = 1, 2 \ldots, t$, with $t \geq 2$, and $v_1 + v_2 + \cdots + v_t = n$, then

$$d(G_{n;\lambda}) = \frac{1}{2}(n-1)(n-2) - \sum_{s=1}^{t} \frac{1}{2} v_s(v_s - 1). \tag{7.2}$$

Proof Any labelling of the vertices of the hive must be such that the difference α in values along each interior edge of the graph $G_{n;\lambda}$ is, by construction of those edges, equal to a corresponding edge value on the parallel λ-boundary, that is $\alpha = \lambda_i = w_s$ for some i and a corresponding s. For each connected component this fixes the values of all vertex labels in terms of that of any one vertex of that component. In the case of a component connected to the boundary, the value at the boundary vertex fixes all the others in that component. The maximum number of degrees of freedom in assigning entries to the corresponding set of K-hives with given boundary is then $d(G_{n;\lambda})$.

The application of the stretching parameter t leaves $G_{n;\lambda}$ unaltered, that is $G_{n;t\lambda} = G_{n;\lambda}$, so that the number of degrees of freedom in assigning entries to the stretched K-hives is still $d(G_{n;\lambda})$. For each connected component of $G_{n;\lambda}$ that is not connected to the boundary we may select any one convenient vertex. The values of its label are not fixed by the hive constraints. They must satisfy a set of linear inequalities that are all scaled by t as both the boundary vertex and edge labels are scaled by the stretching parameter t. It follows that the range of allowed values of the label must be either independent of t or linear in t, thereby giving rise to a corresponding contribution to $P_{\lambda\mu}(t)$ that is at most linear in t. The number of degrees of freedom, $d(G_{n;\lambda})$, that we have identified therefore gives an upper bound on the degree of the corresponding stretched K-polynomial in t.

The number of internal vertices of $G_{n;\lambda}$ is $(n-1)(n-2)/2$. For given n and arbitrary λ this provides a preliminary upper bound on the degree of $P_{\lambda\mu}(t)$. However, for $\lambda = (w_1^{v_1}, w_2^{v_2}, \ldots, w_t^{v_t})$, each s such that $v_s > 1$ specifies a sequence of components of λ having the same value, namely w_s. There is a corresponding set of $v_s(v_s-1)/2$ interior edges within $G_{n;\lambda}$. Whether or not these interior edges reach the boundary, as they do in the two cases $s = 1$ and $s = t$, their introduction reduces the number of connected components not linked to the boundary by $v_s(v_s-1)/2$. The result (7.2) then follows by noting that for different s the sets of vertices linked by the interior edges are disjoint, so that their contributions to the reduction of $d(G_{n;\lambda})$ are independent. □

In the case of our example with $n = 6$ and $\lambda = (4, 2, 2, 0, 0, 0) = (4, 2^2, 0^3)$ we can see from the graph of $G_{n;\lambda}$ that $d(G_{n;\lambda}) = 6$, in agreement with the formula (7.2) that gives $d(G_{n;\lambda}) = 5 \cdot 4/2 - 2 \cdot 1/2 - 3 \cdot 2/2 = 6$. Hence for all μ the degree of the corresponding stretched Kostka coefficient polynomial $P_{\lambda\mu}(t)$ must satisfy

$$\deg(P_{\lambda\mu}(t)) \leq 6. \tag{7.3}$$

By way of example, for $\mu = (3, 1, 1, 1, 1, 1)$ we find

$$P_{\lambda\mu}(t) = \frac{1}{72}(t+1)(t+2)(t+3)(t+4)(t^2+2t+3) \quad \text{so that} \quad F_{\lambda\mu}(z) = \frac{1+3z+6z^2}{(1-z)^7}, \tag{7.4}$$

while for $\mu = (2, 2, 1, 1, 1, 1)$ we have

$$P_{\lambda\mu}(t) = \frac{1}{60}(t+1)(t+2)(t+3)(3t+5)(t^2+2t+2) \quad \text{so that} \quad F_{\lambda\mu}(z) = \frac{1+9z+19z^2+7z^3}{(1-z)^7}. \tag{7.5}$$

On the other hand if $\mu = (3, 2, 1, 1, 1, 0)$ we find

$$P_{\lambda\mu}(t) = \frac{1}{24}(t+1)(t+2)(t+3)(t+4) \quad \text{so that} \quad F_{\lambda\mu}(z) = \frac{1}{(1-z)^5}. \tag{7.6}$$

The third case in which $\deg(P_{\lambda\mu}(t)) = 4 < 6$ is one for which $K_{\lambda\mu}$ is not primitive. In both the other cases $K_{\lambda\mu}$ is primitive, and we have $\deg(P_{\lambda\mu}(t)) = 6$ so that the bound (7.1) is saturated. On the basis of very many calculations of this type we are led to the following:

Conjecture 7.2. *If $K_{\lambda,\mu} > 0$ is primitive and $P_{\lambda\mu}(t) = K_{t\lambda,t\mu}$, then $\deg(P_{\lambda\mu}(t)) = d(G_{n;\lambda})$ for all μ.*

References

[BK] A.D. Berenstein and A.N. Kirillov, *Groups generated by involutions, Gelfand-Tsetlin patterns and combinaatorics of Young tableaux* St Petersburg Math. J. **7** (1966) 77-127.

[BZ] A.D. Berenstein and A.V. Zelevinsky, *When is the multiplicity of a weight equal to* 1*?*, Funct. Anal. Appl. **24** (1990) 259-269.

[BGR] S. Billey, V. Guillemin and E. Rassart, *A vector partition function for the multiplicities of* $sl_k(\mathbb{C})$ J. Algebra **278** (2004) 251-293.

[B] A.S. Buch, *The saturation conjecture (after A. Knutson and T. Tao). With an appendix by W. Fulton* Enseign. Math. **46** (2000) 43-60.

[DeLM] J.A. De Loera and T.B. McAllister, *Vertices of Gelfand-Tsetlin polytopes*, ArXiv:math.CO/0309329, 2003.

[E] E. Ehrhart, *Polynômes arithmétiques et méthode des polyèdres en combinatoire*, Internat. Ser. Numer. Math., Vol. 35., Birkhauser Verlag, Basel, 1977.

[GT] I.M. Gel'fand and M.L. Tsetlin, *Finite-dimensional representations of the group of unimodular matrices*, Dokl. Akad. Nauk. SSSR, **77** (1950) 825-828.

[JK] G. James and A. Kerber, *The representation theory of the symmetric group*, Encyclopedia of mathematics and its applications **16**, Addison-Wesley, Reading, 1981.

[KTT] R.C. King, C. Tollu and F. Toumazet, *Stretched Littlewood-Richardson and Kostka coefficients* CRM Proceedings and Lecture Notes, Vol. 34, Amer. Math. Soc, 2004, pp99-112.

[K] A.N. Kirillov, *Ubiquity of Kostka polynomials*, in Physics and Combinatorics 1999, World Scientific, Singapore, 2001, pp85-200.

[KR] A.N. Kirillov and N.Y. Reshetikhin, *The Bethe ansatz and the combinatorics of Young tableaux* J. Soviet Math. **41** (1988) 925-955.

[KT] A. Knutson and T. Tao, *The honeycomb model of* $GL_n(\mathbb{C})$ *tensor products. I: Proof of the saturation conjecture*, J. Amer. Math. Soc. **12** (1999) 1055-1090.

[KTW] A. Knutson, T. Tao and C. Woodward, *The honeycomb model of* $GL_n(\mathbb{C})$ *tensor products. II: Puzzles determine facets of the Littlewood-Richardson cone*, J. Amer. Math. Soc. **17** (2004) 19-48.

[L] D.E. Littlewood, *The theory of group characters and matrix representations of groups*, 2nd edn., Clarendon Press, Oxford, 1950.

[M] I.G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd edn., Oxford University Press, Oxford, 1995.

[S1] R.P. Stanley, *Enumerative Combinatorics*, Vol. I, Cambridge University Press, Cambridge, 1997.

[S2] R.P. Stanley, *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge, 1999.

School of Mathematics, University of Southampton, Southampton SO17 1BJ, U.K.
E-mail address: R.C.King@maths.soton.ac.uk

Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030, Université Paris 13, 93430 Villetaneuse, France
E-mail address: ct@lipn.univ-paris13.fr

Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030, Université Paris 13, 93430 Villetaneuse, France
E-mail address: ft@lipn.univ-paris13.fr

ON THE ENUMERATION OF PARKING FUNCTIONS BY LEADING NUMBERS

SEN-PENG EU, TUNG-SHAN FU, AND CHUN-JU LAI

ABSTRACT. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a sequence of positive integers. An $\mathbf{x}$-parking function is a sequence $(a_1, \ldots, a_n)$ of positive integers whose non-decreasing rearrangement $b_1 \le \cdots \le b_n$ satisfies $b_i \le x_1 + \cdots + x_i$. In this paper we give a combinatorial approach to the enumeration of $(a, b, \ldots, b)$-parking functions by their leading terms, which covers the special cases $\mathbf{x} = (1, \ldots, 1)$, $(a, 1, \ldots, 1)$, and $(b, \ldots, b)$. The approach relies on bijections between the $\mathbf{x}$-parking functions and labeled rooted forests. To serve this purpose, we present a simple method for establishing the required bijections. Some bijective results between certain sets of $\mathbf{x}$-parking functions of distinct leading terms are also given.

RÉSUMÉ. Soit $\mathbf{x} = (x_1, \ldots, x_n)$ un vecteur d'entiers strictement positifs. Une fonction de $\mathbf{x}$-parking est un vecteur $(a_1, \ldots, a_n)$ d'entiers strictement positifs tel que son réordonnement croissant, noté $b_1 \le \cdots \le b_n$ satisfait $b_i \le x_1 + \cdots + x_i$. Dans cet article, nous proposons une approche combinatoire unifiée pour l'énumération des fonctions de $\mathbf{x}$-parking selon leur terme dominant, dans les cas où $\mathbf{x}$ est égal à $(1, \ldots, 1), (a, 1, \ldots, 1), (b, \ldots, b)$, et $(a, b, \ldots, b)$. Cette énumeration s'appuie sur des bijections entre les fonctions de $\mathbf{x}$-parking et les arbres étiquetés et les forêts étiquetées. À cette fin, nou présentons un mécanisme qui simplifie de façon significative l'établissement des bijections. Nous donnons plusieurs résultats bijectifs entre certains ensembles de fonctions de $\mathbf{x}$-parking ayant des termes dominants distincts.

1. INTRODUCTION

A *parking function* of length n is a sequence $\alpha = (a_1, \ldots, a_n)$ of positive integers such that the non-decreasing rearrangement $b_1 \le \cdots \le b_n$ of α satisfies $b_i \le i$. Parking functions were introduced by Konheim and Weiss in [4] when they dealt with a hashing problem in computer science. They derived that the number of parking functions of length n is $(n+1)^{n-1}$, which coincides with the number of labeled trees on $n+1$ vertices by Cayley's formula. Several bijections between the two sets are known (e.g., see [1, 7, 8]). Parking functions have been found in connection to many other combinatorial structures such as acyclic mappings, polytopes, non-crossing partitions, hyperplane arrangements, etc. Refer to [1, 2, 3, 6, 9, 10] for more information. The notion of parking functions were further generalized in [6]. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a sequence of positive integers. The sequence $\alpha = (a_1, \ldots, a_n)$ is called an $\mathbf{x}$-*parking function* if the non-decreasing rearrangement $b_1 \le \cdots \le b_n$ of α satisfies $b_i \le x_1 + \cdots + x_i$. Thus the ordinary parking function is the case $\mathbf{x} = (1, \ldots, 1)$. The number of $\mathbf{x}$-parking functions for an arbitrary $\mathbf{x}$ was obtained by Kung and Yan [5] in terms of the determinantal formula of Gončarove polynomials. See also [12, 13, 14] for the explicit formulas and properties for some specified cases of $\mathbf{x}$.

Supported in part by National Science Council, Taiwan, ROC (NSC 93-2115-M-390-005 and 93-2115-M-251-001).

Motivated by the work of Foata and Riordan in [1], we give a combinatorial approach to the enumeration of $(a, b, \ldots, b)$-parking functions by their leading terms in this paper, which covers the special cases $\mathbf{x} = (1, \ldots, 1)$, $(a, 1, \ldots, 1)$, and $(b, \ldots, b)$. An $\mathbf{x}$-parking function $(a_1, \ldots, a_n)$ is said to be k-*leading* if $a_1 = k$. Let $p_{n,k}$ denote the number of k-leading ordinary parking functions of length n. Foata and Riordan [1] derived the generating function for $p_{n,k}$ algebraically,

$$\sum_{k=1}^{n} p_{n,k} x^k = \frac{x}{1-x} \left(2(n+1)^{n-2} - \sum_{k=1}^{n} \binom{n-1}{k-1} k^{k-2} (n-k+1)^{n-k-1} x^k \right). \tag{1}$$

One of our main results is that we obtain some bijective results between certain sets of $(a, 1, \ldots, 1)$-parking functions of distinct leading terms, which lead to the explicit formulas for the number of k-leading $(a, 1, \ldots, 1)$-parking functions. In particular, for the case $a = 1$, we deduce (1) combinatorially. These results rely on a bijection φ between $(a, 1, \ldots, 1)$-parking functions and labeled rooted forests, which is a generalization of the second bijection between acyclic mappings and parking functions of Foata and Riordan [1, Section 3]. The key that opens the way is a simple object, called *triplet-labeled rooted forest*, which not only serves as an intermediate stage of the bijection φ but also enables $(a, 1, \ldots, 1)$-parking functions to be manipulated on forests easily. Furthermore, based on the bijection φ, we establish an immediate bijection between $(a, b, \ldots, b)$-parking functions and labeled rooted forests with edge-colorings, which is equivalent to a bijection given by C. Yan in [14], so as to enumerate $(a, b, \ldots, b)$-parking functions by their leading terms. In the end we propose a structure, by using a generalized triplet-labeled rooted forest, for general $\mathbf{x}$-parking functions.

We organize this paper as follows. The notion of triplet-labeled rooted forests and the bijection φ are given in Section 2. How the bijection φ is applied to enumerate $(a, 1, \ldots, 1)$-parking functions (and hence ordinary parking functions) by the leading terms is given in Section 3. Making use of the notion of labeled rooted forests with edge-colorings, we investigate the cases $\mathbf{x} = (b, \ldots, b)$ and $(a, b, \ldots, b)$ in Section 4. Finally, we propose a structure for general $\mathbf{x}$-parking functions in Section 5.

2. Triplet-labeled rooted forests and the bijection φ

For a rooted forest F and two vertices $u, v \in F$, we say that u is a *descendant* of v if v is contained in the path from u to the root of the component that contains u. If also u and v are adjacent, then u is called a *child* of v, and v is called the *parent* of u. Let $T(u)$ denote the subtree of F consisting of u and the descendants of u, and let $F - T(u)$ denote the remaining part of F when the subtree $T(u)$ and the edge uv are removed. For any two integers $m < n$, we use the notation $[m, n] = \{m, m+1, \ldots, n\}$. In particular, we write $[n] = \{1, \ldots, n\}$.

In this section we consider the case $\mathbf{x} = (a, 1, \ldots, 1)$. We call such an $\mathbf{x}$-parking function α an $(a, \overline{1})$-*parking function*. Note that the non-decreasing rearrangement $b_1 \le \cdots \le b_n$ of α satisfies $b_i \le i + a - 1$. Let $\mathcal{P}_n(a, \overline{1})$ denote the set of $(a, \overline{1})$-parking functions of length n. It is known that $|\mathcal{P}_n(a, \overline{1})| = a(a+n)^{n-1}$ (see [6, 14]). Given an $\alpha = (a_1, \ldots, a_n) \in \mathcal{P}_n(a, \overline{1})$, for $1 \le i \le n$, we define

$$\pi_\alpha(i) = \ \mathrm{Card}\{a_j \in \alpha \mid \text{either } a_j < a_i, \text{ or } a_j = a_i \text{ and } j < i\}. \tag{2}$$

Note that $(\pi_\alpha(1), \ldots, \pi_\alpha(n))$ is a permutation of $[n]$. In fact, $\pi_\alpha(i)$ is the position of the term a_i in the non-decreasing rearrangement of α. Moreover, we define $\tau_\alpha(i) = \pi_\alpha(i) + a - 1$, for $1 \le i \le n$. Note that $a_i \le \tau_\alpha(i)$. We associate α with an a-component forest F_α, called *triplet-labeled rooted* *forest.* The vertex set of F_α is the set

$$\{(i, a_i, \tau_\alpha(i)) | \, a_i \in \alpha\} \cup \{(\rho_i, 0, i) | \, 0 \le i \le a-1\}$$

of *triplets* (i.e., here we identify each vertex with a triplet), where $\rho_i \notin [n]$ is just an artificial label for discriminating the additional triplets. Let $(\rho_0, 0, 0), \ldots, (\rho_{a-1}, 0, a-1)$ be the roots of distinct trees of F_α. For any two vertices $v = (x_1, y_1, z_1)$ and $u = (x_2, y_2, z_2)$, u is a child of v whenever $y_2 = z_1 + 1$.

For example, take $a = 2$ and $n = 9$. Consider the $(2, \overline{1})$-parking function $\alpha = (2, 5, 9, 1, 5, 7, 2, 4, 1)$. We have the permutation $(\pi_\alpha(1), \ldots, \pi_\alpha(n)) = (3, 6, 9, 1, 7, 8, 4, 5, 2)$ and the sequence $(\tau_\alpha(1), \ldots, \tau_\alpha(n)) = (4, 7, 10, 2, 8, 9, 5, 6, 3)$. The triplet-labeled rooted forest associated with α is shown on the left of Figure 1.

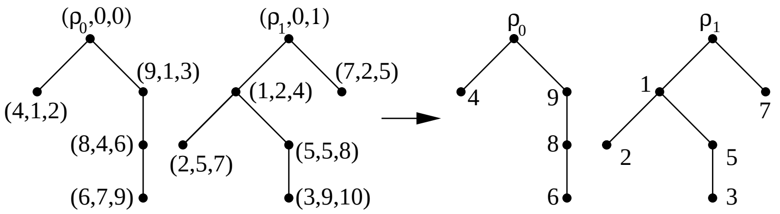


FIGURE 1. the triplet-labeled rooted forest associated with α = (2,5,9,1,5,7,2,4,1) and the corresponding labeled rooted forest $\varphi(\alpha)$ (in the canonical form).

Let $\mathcal{F}_n(a, \overline{1})$ denote the set of a-component rooted forests on the set $\{\rho_0, \ldots, \rho_{a-1}\} \cup [n]$ such that the a specified vertices $\rho_0, \ldots, \rho_{a-1}$ are the roots of distinct trees. We shall establish a bijection φ between $\mathcal{P}_n(a, \overline{1})$ and $\mathcal{F}_n(a, \overline{1})$ with the triplet-labeled rooted forest as an intermediate stage.

The bijection $\varphi : \mathcal{P}_n(a, \overline{1}) \to \mathcal{F}_n(a, \overline{1})$: Define the mapping φ by carrying $\alpha \in \mathcal{P}_n(a, \overline{1})$ into $\varphi(\alpha) \in \mathcal{F}_n(a, \overline{1})$, where $\varphi(\alpha)$ is the same as the triplet-labeled rooted forest F_α associated with α with vertices labeled by the first entries of the triplets of F_α, i.e., $\varphi(\alpha)$ is obtained from F_α simply by erasing the last two entries of all triplets. As illustrated in Figure 1, the forest on the right is the corresponding forest $\varphi(\alpha)$ of the $(2, \overline{1})$-parking function $\alpha = (2, 5, 9, 1, 5, 7, 2, 4, 1)$.

To describe φ^{-1}, for each $F \in \mathcal{F}_n(a, \overline{1})$, first we express F in a form, called *canonical form*, of a plane rooted forest. We write $F = (T_0, \ldots, T_{a-1})$, where T_i denotes the tree of F that is rooted at ρ_i, for $0 \le i \le a-1$. Let $T_0, \ldots, T_{a-1}$ be placed from left to right. If a vertex has more than one child then the labels of these children are increasing from left to right. For example, the forest on the right of Figure 1 is in the canonical form. Then we associate the root ρ_i with the triplet $(\rho_i, 0, i)$, for $0 \le i \le a-1$, and associate each non-root vertex $j \in [n]$ with a triplet (j, p_j, q_j), where p_j and q_j are determined by the following algorithm. Here traversing F by a breadth-first search means that we view F as a rooted tree $\widehat{F}$ by connecting the roots of F to a virtual vertex x and then traverse $\widehat{F}$ by a breadth-first search from x.

Algorithm A.

(i) Traverse F by a breadth-first search and label the third entries q_j of the non-root vertices from a to $a + n - 1$.

(ii) For any two vertices $v = (x_1, y_1, z_1)$ and $u = (x_2, y_2, z_2)$, $y_2 = z_1 + 1$ whenever u is a child of v.

As shown in Figure 1, the forest on the left can be recovered from the one on the right by algorithm A. Note that if u is a child of v then $z_2 > z_1$, and hence $y_2 = z_1 + 1 \le z_2$. Sorting the triplets of non-root vertices by the first entries, the sequence $\varphi^{-1}(F) = (p_1, \dots, p_n)$, which is formed by their second entries, is the required $(a, \overline{1})$-parking function.

Remark: For the special case $a = 1$, φ is a bijection between the set of ordinary parking functions of length n and the set of labeled rooted trees on $[0, n]$ with root $\rho_0 = 0$, which is equivalent to the second bijection between acyclic mappings and parking functions of Foata and Riordan [1, Section 3].

3. Enumerating $(a, \overline{1})$-parking functions by leading terms

Let $\mathcal{P}_{n,k}(a, \overline{1})$ denote the set of k-leading $(a, \overline{1})$-parking functions of length n, and let $p_{n,k}^{(a,\overline{1})} = |\mathcal{P}_{n,k}(a, \overline{1})|$. Let $\mathcal{F}^*_{n,k}(a, \overline{1})$ denote the set of triplet-labeled rooted forests F_α associated with $\alpha \in \mathcal{P}_n(a, \overline{1})$. For each $\alpha = (a_1, \dots, a_n) \in \mathcal{P}_{n,k}(a, \overline{1})$, we observe that $\pi_\alpha(1) + a - 1 \ge a_1 = k$. With the benefit of triplet-labeled rooted forests, we obtain the following bijective result.

Theorem 3.1. *For $a \le k \le a + n - 2$, there is a bijection between the sets $\mathcal{R}$ and $\mathcal{P}_{n,k+1}(a, \overline{1})$, where $\mathcal{R}$ is the set of k-leading $(a, \overline{1})$-parking functions α of length n that satisfy at least one of the two conditions (i) α has more than one term equal to k, (ii) α has at least $k - a + 1$ terms less than k, and $\mathcal{P}_{n,k+1}(a, \overline{1})$ is the set of $(k + 1)$-leading $(a, \overline{1})$-parking functions of length n.*

Proof: Let $\mathcal{F}^*(\mathcal{R}) \subseteq \mathcal{F}^*_{n,k}(a, \overline{1})$ be the set of forests associated with the parking functions in $\mathcal{R}$. It suffices to establish a bijection $\phi : \mathcal{F}^*(\mathcal{R}) \to \mathcal{F}^*_{n,k+1}(a, \overline{1})$. Given an $F_\alpha \in \mathcal{F}^*(\mathcal{R})$, let $u = (1, k, \tau_\alpha(1)) \in F_\alpha$. Note that the condition (ii) can be rephrased as $\tau_\alpha(1) > k$. Since α satisfies at least one of the two conditions (i) and (ii), there are at least $k + 1$ vertices in the subset $F_\alpha - T(u)$. Traverse $F_\alpha - T(u)$ by a breadth-first search and locate the k-th vertex, say v (we mean that the root of the first tree of $F_\alpha - T(u)$ is the 0-th vertex). The mapping ϕ is defined by carrying F_α into $\phi(F_\alpha)$, where $\phi(F_\alpha)$ is obtained from $F_\alpha - T(u)$ with $T(u)$ attached to v so that u is the first child of v. Updating the second and the third entries of all non-root vertices by algorithm A, the triplet of u becomes $(1, k + 1, \tau(1))$, for some $\tau(1) \ge k + 1$. Hence $\phi(F_\alpha) \in \mathcal{F}^*_{n,k+1}(a, \overline{1})$.

To find ϕ^{-1}, given an $F_\beta \in \mathcal{F}^*_{n,k+1}(a, \overline{1})$ for some $\beta \in \mathcal{P}_{n,k+1}(a, \overline{1})$, let $u = (1, k + 1, \tau_\beta(1)) \in F_\beta$ and let v be the parent of u. In F_β we locate the vertex, say w, the third entry of which is equal to $k - 1$. Then $\phi^{-1}(F_\beta)$ is obtained from $F_\beta - T(u)$ with $T(u)$ attached to w so that u is the first child of w. By algorithm A, the updated triplet of u becomes $(1, k, \tau(1))$. We observe that either $\tau(1) = k$ if v is another child of w, or $\tau(1) > k$ otherwise. Hence $\phi^{-1}(F_\beta) \in \mathcal{F}^*(\mathcal{R})$. ☐

For example, take $a = 2$ and $n = 9$. Consider the 2-leading $(2, \overline{1})$-parking function $\alpha = (2, 5, 9, 1, 5, 7, 2, 4, 1)$. On the left of Figure 2 is the forest F_α associated with α. Let $u = (1, 2, 4)$. Note that $v = (4, 1, 2)$ is the second vertex of $F_\alpha - T(u)$ that is visited by a breadth-first search. On the right of Figure 2 is the corresponding forest $\phi(F_\alpha)$, which is obtained from $F_\alpha - T(u)$ with $T(u)$ attached to v and with the second and the third entries of the triplets updated. Sorting the triplets of non-root vertices by the first entries, we retrieve the corresponding 3-leading $(2, \overline{1})$-parking function $(3, 6, 9, 1, 6, 7, 2, 4, 1)$ from their second entries.

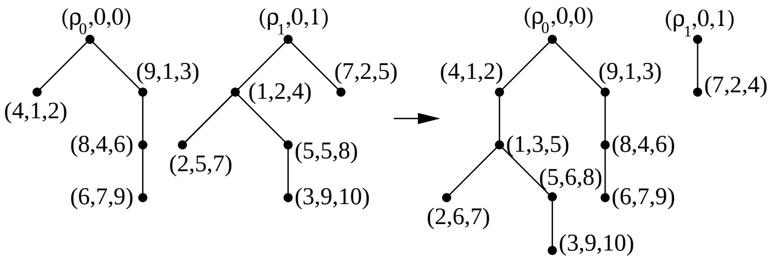


FIGURE 2. the forests associated with the $(2,\overline{1})$-parking functions (2,5,9,1,5,7,2,4,1) and $(3,6,9,1,6,7,2,4,1)$.

The following recurrence relations for $p_{n,k}^{(a,\overline{1})}$ can be derived from Theorem 3.1. Here, unless specified, the labeled rooted forests in $\mathcal{F}_n(a,\overline{1})$ are considered to be in the canonical form. Recall that $|\mathcal{F}_n(a,\overline{1})| = |\mathcal{P}_n(a,\overline{1})| = a(a+n)^{n-1}$.

Theorem 3.2. *For $a \le k \le a+n-2$, we have*

$$p_{n,k}^{(a,\overline{1})} - p_{n,k+1}^{(a,\overline{1})} = \binom{n-1}{k-a} a k^{k-a-1}(n-k+a)^{n-k+a-2}. \tag{3}$$

Proof: By Theorem 3.1, $p_{n,k}^{(a,\overline{1})} - p_{n,k+1}^{(a,\overline{1})}$ is equal to the number of k-leading $(a,\overline{1})$-parking functions α such that the first term of α is the unique term equal to k, and $\tau_\alpha(1) = k$. We shall count the number of forests that are mapped by such parking functions under the mapping φ. Let $F = \varphi(\alpha) \in \mathcal{F}_n(a,\overline{1})$ and let $u = 1 \in F$. We observe that $F - T(u)$ is an a-component labeled forest on k vertices containing the roots $\rho_0, \dots, \rho_{a-1}$, and $T(u)$ is a labeled tree on $n-k+a$ vertices containing u. Since there are $\binom{n-1}{k-a}$ ways to choose $k-a$ numbers from $[2,n]$ for the non-root vertices of $F-T(u)$ and since there are ak^{k-a-1} and $(n-k+a)^{n-k+a-2}$ possibilities for the induced forest $F-T(u)$ and tree $T(u)$, the number of the required forests is $\binom{n-1}{k-a} a k^{k-a-1}(n-k+a)^{n-k+a-2}$. $\square$

In order to evaluate $p_{n,k}^{(a,\overline{1})}$ by Theorem 3.2, we need the following initial conditions (4). Since an $(a,\overline{1})$-parking function of length n with leading term k, $(1 \le k \le a)$ is simply a juxtaposition of k and an $(a+1,\overline{1})$-parking function of length $n-1$, we have

$$p_{n,k}^{(a,\overline{1})} = (a+1)(a+n)^{n-2}, \qquad \text{for } 1 \le k \le a. \tag{4}$$

Now we can derive the explicit formula for $p_{n,k}^{(a,\overline{1})}$ by (3) and (4). In particular, we have $p_{n,a+n-1}^{(a,\overline{1})} = a(a+n-1)^{n-2}$ since an $(a+n-1)$-leading $(a,\overline{1})$-parking function of length n is a juxtaposition of $a+n-1$ and an $(a,\overline{1})$-parking function of length $n-1$. We derive the following enumerator for $(a,\overline{1})$-parking functions by the leading terms.

Theorem 3.3. *If $P^{(a,\overline{1})}(x) = \sum_{k=1}^{a+n-1} p_{n,k}^{(a,\overline{1})} x^k$, then*

$$P^{(a,\overline{1})}(x) = \frac{x}{1-x}\left((a+1)(a+n)^{n-2} - \sum_{k=a}^{a+n-1} \binom{n-1}{k-a} a k^{k-a-1}(n-k+a)^{n-k+a-2} x^k \right).$$

Proof: We have

$$\begin{aligned}\left(\frac{1}{x}-1\right)P^{(a,\bar{1})}(x) &= p_{n,1}^{(a,\bar{1})}-\left(\sum_{k=1}^{a+n-2}\left(p_{n,k}^{(a,\bar{1})}-p_{n,k+1}^{(a,\bar{1})}\right)x^k\right)-p_{n,a+n-1}^{(a,\bar{1})}x^{a+n-1}\\ &= (a+1)(a+n)^{n-2}-\sum_{k=a}^{a+n-1}\binom{n-1}{k-a}ak^{k-a-1}(n-k+a)^{n-k+a-2}x^k,\end{aligned}$$

as required. □

Remark: For the case $a=1$, we deduce that the number $p_{n,k}$ of k-leading ordinary parking functions satisfies the recurrence relations

$$p_{n,k}-p_{n,k+1}=\binom{n-1}{k-1}k^{k-2}(n-k+1)^{n-k-1}, \tag{5}$$

for $1\le k\le n-1$, with the initial condition $p_{n,1}=2(n+1)^{n-2}$. The enumerator (1) for ordinary parking functions by the leading terms is derived anew.

Making use of the bijection φ for the case $a=1$, we derive the following interesting result for ordinary parking functions.

Theorem 3.4. *If n is even, then there is a two-to-one correspondence between the set of 1-leading parking functions of length n and the set of $(\frac{n}{2}+1)$-leading parking functions of length n.*

Proof: Let $\mathcal{A}$ (resp. $\mathcal{B}$) denote the set of labeled trees corresponding to the 1-leading (resp. $(\frac{n}{2}+1)$-leading) parking functions of length n under the mapping φ. We shall establish a two-to-one correspondence ϕ between $\mathcal{A}$ and $\mathcal{B}$. For each $T\in\mathcal{A}$, the two vertices $u=1$ and $v=0$ of T are adjacent. Let $T(v)=T-T(u)$. Since T has $n+1$ vertices and n is even, one of $T(v)$ and $T(u)$ contains more than $\frac{n}{2}$ vertices. If $|T(v)|>\frac{n}{2}$, then we traverse $T(v)$ from v by a breadth-first search and locate the $\frac{n}{2}$-th non-root vertex, say w. The tree $\phi(T)\in\mathcal{B}$ is obtained from $T(v)$ with $T(u)$ attached to w so that u becomes the first child of w. Otherwise $|T(u)|>\frac{n}{2}$. Locate the $\frac{n}{2}$-th non-root vertex of $T(u)$, say w', by a breadth-first search. Then $\phi(T)$ is obtained from $T(u)$ with $T(v)$ attached to w' so that v becomes the first child of w' and with a relabeling $u=0$ and $v=1$.

To find ϕ^{-1}, given a tree $T'\in\mathcal{B}$, let $u=1$, $v=0$, and $T(v)=T'-T(u)$. We retrieve two trees of $\mathcal{A}$ from $T(u)$ and $T(v)$. One is obtained by attaching $T(u)$ to v so that u becomes the first child of v, and the other is obtained by attaching $T(v)$ to u so that v becomes the first child of u and relabeling $u=0$ and $v=1$. □

For example, take $n=6$. On the left of Figure 3 are the labeled trees T corresponding to the 1-leading parking functions $(1,4,1,2,4,1)$ and $(1,5,2,1,5,2)$, respectively. Let $u=1$, $v=0$ and $T(v)=T-T(u)$. For the first tree, we observe that $|T(v)|>3$ and the vertex 2 is the third non-root vertex of $T(v)$ that is visited by a breadth-first search. We attach $T(u)$ to 2 to obtain the required tree on the right. For the second tree, we observe that $|T(u)|>3$ and again the vertex 2 is the third non-root vertex of $T(u)$. Likewise, we attach $T(v)$ to 2 and relabel $u=0$ and $v=1$. Note that the tree on the right of Figure 3 corresponds to the 4-leading parking function $(4,3,1,6,3,1)$.

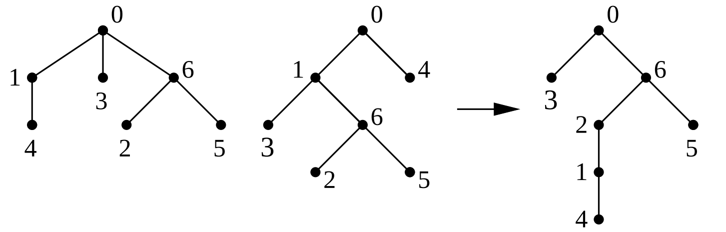


FIGURE 3. an example of the two-to-one correspondence for $n = 6$

4. Edge-colored labeled rooted trees and forests

In this section we shall enumerate $\mathbf{x}$-parking functions by the leading terms for the cases $\mathbf{x} = (b, \ldots, b)$ and $(a, b, \ldots, b)$. First, we define the required structure for the case $\mathbf{x} = (b, \ldots, b)$. Let $\mathcal{T}_n(\bar{b})$ denote the set of labeled trees, called *b-trees*, on the vertex set $[0, n]$, whose edges are colored with the colors $0, 1, \ldots, b-1$. There is no further restriction on the colorings of edges. Unless specified, each $T \in \mathcal{T}_n(\bar{b})$ is rooted at 0 and is in the canonical form regarding the vertex-labeling. Let $\kappa(i)$ denote the color of the edge that connects the vertex i and its parent. It is known that $|\mathcal{T}_n(\bar{b})| = b^n(n+1)^{n-1}$.

For the case $\mathbf{x} = (b, \ldots, b)$, we call such $\mathbf{x}$-parking functions $(\bar{b})$-*parking functions*. Let $\mathcal{P}_n(\bar{b})$ denote the set of $(\bar{b})$-parking functions of length n. Note that the non-decreasing rearrangement $b_1 \le \cdots \le b_n$ of $\alpha \in \mathcal{P}_n(\bar{b})$ satisfies $b_i \le bi$. We shall establish a bijection $\psi_b : \mathcal{P}_n(\bar{b}) \to \mathcal{T}_n(\bar{b})$ based on the bijection φ for the case $a = 1$. Given an $\alpha = (a_1, \ldots, a_n) \in \mathcal{P}_n(\bar{b})$, we associate α with two sequences $\beta = (p_1, \ldots, p_n)$ and $\gamma = (r_1, \ldots, r_n)$, where $p_i = \lceil \frac{a_i}{b} \rceil$ (the least integer that is greater than or equal to $\frac{a_i}{b}$) and $r_i = bp_i - a_i$, for $1 \le i \le n$. It is easy to see that β is an ordinary parking function of length n and $\gamma \in [0, b-1]^n$, and that α is uniquely determined by such a pair (β, γ), i.e., $a_i = bp_i - r_i$, for $1 \le i \le n$. To establish the mapping ψ_b, we first locate the corresponding labeled tree $\varphi(\beta) \in \mathcal{F}_n(1, \bar{1})$ of β, and then define ψ_b by carrying α into $\varphi(\beta)$ with the edge-coloring $\kappa(i) = r_i$, for $1 \le i \le n$.

For example, take $b = 2$. For the $(\bar{2})$-parking function $\alpha = (2, 7, 4, 18, 1, 9, 2, 9, 8)$, we have the associated ordinary parking function $\beta = (1, 4, 2, 9, 1, 5, 1, 5, 4)$ and the sequence $\gamma = (0, 1, 0, 0, 1, 1, 0, 1, 0)$. On the left of Figure 4 is the triplet-labeled rooted tree associated with β. The 2-tree corresponding to α is shown on the right of Figure 4, where the dotted edges and solid edges represent the colors 0 and 1, respectively.

To find ψ_b^{-1}, given a $T \in \mathcal{T}_n(\bar{b})$, we can retrieve the ordinary parking function $(p_1, \ldots, p_n)$ from the vertex-labeling of T, and then derive the require $(\bar{b})$-parking function $\psi_b^{-1}(T) = (a_1, \ldots, a_n) \in \mathcal{P}_n(\bar{b})$ by setting $a_i = bp_i - \kappa(i)$, for $1 \le i \le n$.

Proposition 4.1. *The mapping $\psi_b : \mathcal{P}_n(\bar{b}) \to \mathcal{T}_n(\bar{b})$ mentioned above is a bijection.*

Let $p_{n,m}^{(\bar{b})}$ denote the number of m-leading $(\bar{b})$-parking functions of length n. Recall that we have determined the number $p_{n,k}$ of k-leading ordinary parking functions of length n by (5) and an initial condition. We derive $p_{n,m}^{(\bar{b})}$ as follows.

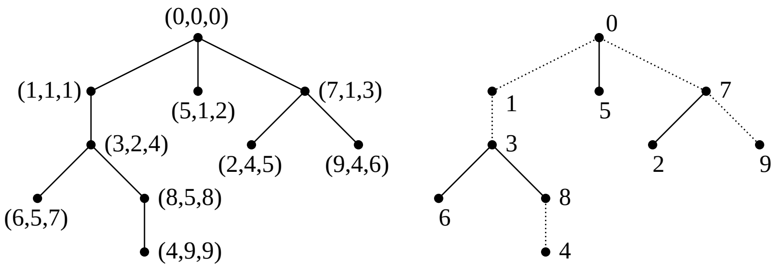


FIGURE 4. the triplet-labeled rooted tree associated with the parking function $(1,4,2,9,1,5,1,5,4)$ and the 2-tree corresponding to $(2,7,4,18,1,9,2,9,8)$

Proposition 4.2. *For $1 \le k \le n$ and $(k-1)b+1 \le m \le kb$, we have $p^{(\bar{b})}_{n,m} = b^{n-1}p_{n,k}$.*

Proof: If $(k-1)b+1 \le m \le kb$, then each m-leading $(\bar{b})$-parking function of length n is associated with a pair (β,γ), where β is a k-leading parking function of length n and $\gamma \in [0, b-1]^n$ with the first term equal to $bk-m$. Since there are $b^{n-1}p_{n,k}$ possibilities for such a pair (β,γ), the assertion follows. □

Let us turn to the case $\mathbf{x} = (a, b, \dots, b)$. In [14] C. Yan introduced the notion of sequences of rooted b-forests in order to generalize a bijection of Foata and Riordan [1]. A *rooted b-forest* is a labeled rooted forest with edges colored with the colors $0, \dots, b-1$ (note that each component may be rooted at any vertex). Consider a sequence $(S_0, \dots, S_t)$ of rooted b-forests on $[n]$ such that (i) each S_i is a rooted b-forest, (ii) S_i and S_j are disjoint if $i \neq j$, and (iii) the union of the vertex sets S_i, $(0 \le i \le t)$ is $[n]$. Let $\widehat{S}_i$ denote the rooted tree obtained by connecting the roots of S_i to a new root vertex ρ_i, where the edges incident to ρ_i are not colored with any color, denoted by -1 for such an edge. Let $\mathcal{F}_n(a,\bar{b})$ denote the set of a-component rooted forests of the form $(\widehat{S_0}, \dots, \widehat{S_{a-1}})$, where $(S_0, \dots, S_{a-1})$ is a sequence of rooted b-forest on $[n]$. We call members of $\mathcal{F}_n(a,\bar{b})$ *extended b-forests*. Let $\kappa(i)$ denote the color of the edge that connects the vertex i and its parent. It is known that $|\mathcal{F}_n(a,\bar{b})| = a(a+nb)^{n-1}$ (see [6, 14]).

For $\mathbf{x} = (a, b, \dots, b)$, we call such an $\mathbf{x}$-parking function α an $(a,\bar{b})$-*parking function*. In this case the non-decreasing rearrangement $b_1 \le \cdots \le b_n$ of α satisfies $b_i \le a + (i-1)b$. Let $\mathcal{P}_n(a,\bar{b})$ denote the set of $(a,\bar{b})$-parking functions of length n. We shall establish a bijection $\varphi_b : \mathcal{P}_n(a,\bar{b}) \to \mathcal{F}_n(a,\bar{b})$ based on the bijection $\varphi : \mathcal{P}_n(a,\bar{1}) \to \mathcal{F}_n(a,\bar{1})$ given in Section 2. Given an $\alpha = (a_1, \dots, a_n) \in \mathcal{P}_n(a,\bar{b})$, we associate α with two sequences $\beta = (p_1, \dots, p_n)$ and $\gamma = (r_1, \dots, r_n)$, where

$$p_i = \begin{cases} a_i & \text{if } a_i \le a, \\ \left\lceil \dfrac{a_i - a}{b} \right\rceil + a & \text{otherwise;} \end{cases} \quad \text{and} \quad r_i = \begin{cases} -1 & \text{if } a_i \le a, \\ b(p_i - a) - a_i + a & \text{otherwise.} \end{cases} \tag{6}$$

It is straightforward to verify that β is an $(a,\bar{1})$-parking function of length n and $\gamma \in [-1, b-1]^n$ with $r_i = -1$ whenever $a_i \le a$. Moreover, α is uniquely determined by such a pair (β,γ), i.e., $a_i = p_i$ if $p_i \le a$, and $a_i = b(p_i - a) - r_i + a$ otherwise. To establish the mapping φ_b, we first locate the corresponding labeled forest $\varphi(\beta) \in \mathcal{F}_n(a, \overline{1})$ of β and then define φ_b by carrying α into $\varphi(\beta)$ with the edge-coloring $\kappa(i) = r_i$, for $1 \le i \le n$.

For example, take $a = 2$ and $b = 2$. For the $\alpha = (2, 7, 15, 1, 8, 12, 2, 5, 1)$, we have the associated pair (β, γ), where $\beta = (2, 5, 9, 1, 5, 7, 2, 4, 1)$ and $\gamma = (-1, 1, 1, -1, 0, 0, -1, 1, -1)$. As shown in Figure 1, we have obtained the labeled forest $\varphi(\beta)$. The required extended b-forest $\varphi_b(\alpha)$ is shown in Figture 5, where the arrowed edges are not colored, and the dotted and solid edges represent the colors 0 and 1, respectively.

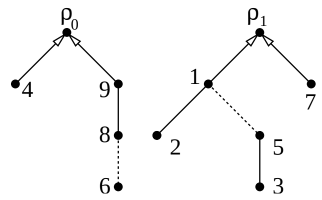


FIGURE 5. the extended 2-forest corresponding to the $(2, \overline{2})$-parking function (2,7,15,1,8,12,2,5,1)

To find φ_b^{-1}, given an $F \in \mathcal{F}_n(a, \overline{b})$, we can retrieve an $(a, \overline{1})$-parking function $(p_1, \dots, p_n)$ from the vertex-labeling of F and then derive the require $(a, \overline{b})$-parking function $\varphi_b^{-1}(F) = (a_1, \dots, a_n)$ by setting $a_i = p_i$ if $p_i \le a$, and $a_i = b(p_i - a) - \kappa(i) + a$ otherwise.

Proposition 4.3. *The mapping $\varphi_b : \mathcal{P}_n(a, \overline{b}) \to \mathcal{F}_n(a, \overline{b})$ mentioned above is a bijection.*

The bijection φ_b is equivalent to Yan's bijection in [14]. Our method not only simplifies the construction but also provides an approach to the enumeration of $(a, b, \dots, b)$-parking functions by the leading terms. Let $\mathcal{P}_{n,m}(a, \overline{b})$ denote the set of m-leading $(a, \overline{b})$-parking functions of length n and let $p_{n,m}^{(a,\overline{b})} = |\mathcal{P}_{n,m}(a, \overline{b})|$.

Lemma 4.4. *For $0 \le k \le n - 2$ and $a + kb + 1 \le i, j \le a + (k + 1)b$, there is a bijection between $\mathcal{P}_{n,i}(a, \overline{b})$ and $\mathcal{P}_{n,j}(a, \overline{b})$, and hence $p_{n,i}^{(a,\overline{b})} = p_{n,j}^{(a,\overline{b})}$.*

Proof: Given an $\alpha = (a_1, \dots, a_n) \in \mathcal{P}_{n,i}(a, \overline{b})$, if the first term ($a_1 = i$) is replaced by j, where $a + kb + 1 \le i, j \le a + (k + 1)b$, then it corresponds to replace the color $\kappa(1) = b(k + 1) - i + a$ of $\varphi_b(\alpha)$ by the color $b(k + 1) - j + a$. Hence there is an immediate bijection between $\mathcal{P}_{n,i}(a, \overline{b})$ and $\mathcal{P}_{n,j}(a, \overline{b})$ by using Proposition 4.3 and an interchange of the color $\kappa(1)$. $\square$

The following result is extended from Theorem 3.1.

Theorem 4.5. *For $0 \le k \le n-2$, there is a bijection between the sets $\mathcal{R}$ and $\mathcal{P}_{n,a+kb+1}(a, \overline{b})$, where $\mathcal{R}$ is the set of $(a + kb)$-leading $(a, \overline{b})$-parking functions α of length n that satisfy at least one of the two conditions (i) α has more than one term belong to the interval $[a + (k - 1)b + 1, a + kb]$ (or α has more than one term equal to a in case $k = 0$), (ii) α has at least $k + 1$ terms less then $a + kb$, and $\mathcal{P}_{n,a+kb+1}(a, \overline{b})$ is the set of $(a + kb + 1)$-leading $(a, \overline{b})$-parking functions of length n.*

Proof: Given an $\alpha \in \mathcal{R}$, let (β, γ) be the pair associated with α determined by (6). We observe that β is an $(a+k)$-leading $(a, \overline{1})$-parking function with at least one of the two properties (i) β has more than one term equal to $a+k$, (ii) $\pi_\beta(1) = \pi_\alpha(1) > k+1$. By Theorem 3.1, there is a bijection ϕ that carries β into an $(a+k+1)$-leading $(a, \overline{1})$-parking function $\phi(\beta)$ of length n. From the pair $(\phi(\beta), \gamma)$, we retrieve an $(a+(k+1)b)$-leading $(a, \overline{b})$-parking function. The assertion follows from Lemma 4.4. $\square$

The following theorem is analogous to Theorems 3.2, and the proof is similar.

Theorem 4.6. *If $0 \le k \le n-2$, then*

$$p^{(a,\overline{b})}_{n,a+kb} - p^{(a,\overline{b})}_{n,a+kb+1} = \binom{n-1}{k} ab^{n-k-1}(a+kb)^{k-1}(n-k)^{n-k-2}. \tag{7}$$

To evaluate $p^{(a,\overline{b})}_{n,m}$ by Lemma 4.4 and Theorem 4.6, we need the following initial conditions (8). Note that an $(a, \overline{b})$-parking function of length n with leading term m, $(1 \le m \le a)$ is a juxtaposition of m and an $(a+b, \overline{b})$-parking function of length $n-1$. Hence

$$p^{(a,\overline{b})}_{n,m} = (a+b)(a+nb)^{n-2}, \qquad \text{for } 1 \le m \le a. \tag{8}$$

By a similar argument of Theorem 3.3, we derive the enumerator for $(a, \overline{b})$-parking functions by the leading terms.

Theorem 4.7. *If $P^{(a,\overline{b})}(x) = \sum_{m=1}^{a+(n-1)b} p^{(a,\overline{b})}_{n,m} x^m$, then*

$$P^{(a,\overline{b})}(x) = \frac{x}{1-x}\left((a+b)(a+nb)^{n-2} - \sum_{k=0}^{n-1}\binom{n-1}{k} ab^{n-k-1}(a+kb)^{k-1}(n-k)^{n-k-2}x^{a+kb}\right).$$

5. A structure for general $\mathbf{x}$-parking functions

In this section we propose a forest structure for general $\mathbf{x}$-parking functions. Given a sequence $\mathbf{x} = (x_1, \dots, x_n)$ of positive integers, let $\mathcal{P}_n(\mathbf{x})$ denote the set of the $\mathbf{x}$-parking functions $\alpha = (a_1, \dots, a_n)$ of length n and let $b_1 \le \cdots \le b_n$ be the non-decreasing rearrangement of α. Likewise, we define the permutation $(\pi_\alpha(1), \dots, \pi_\alpha(n))$ by (2). Let $d_i = x_1 + \cdots + x_i$, for $1 \le i \le n$. We associate α with an x_1-component triplet-labeled rooted forest F_α satisfying the following conditions:

(i) The roots of F_α are the triplets $(0,0,0)$ and $(\rho_{(0,1)}, 0, 1), \dots, (\rho_{(0,x_1-1)}, 0, x_1-1)$.

(ii) For $1 \le j \le n-1$, there are x_{j+1} triplets in F_α associated with b_j. If b_j is the term a_i, for some $i \in [n]$ (i.e., $\pi_\alpha(i) = j$), then F_α contains the triplet (i, a_i, d_j) and the additional $x_{j+1}-1$ triplets $(\rho_{(j,1)}, a_i, d_j+1), \dots, (\rho_{(j,x_{j+1}-1)}, a_i, d_{j+1}-1)$. Finally, there is a triplet (k, a_k, d_n) associated with b_n, where b_n is the term a_k, for some $k \in [n]$ (i.e., $\pi_\alpha(k) = n$).

(iii) For any two vertices $v = (x_1, y_1, z_1)$ and $u = (x_2, y_2, z_2)$, u is a child of v if $y_2 = z_1 + 1$.

Note that the third entries of the triplets are from 0 to $x_1 + \cdots + x_n$. If we erase the second and the third entries of the triplets, then we turn F_α into a rooted forest F with $x_1 + \cdots + x_n + 1$ vertices satisfying the following conditions.

(C.1) The vertex set of F is $[0,n] \cup_{j=0}^{n-1} \{\rho_{(j,1)}, \dots, \rho_{(j,x_{j+1}-1)}\}$.

(C.2) The roots of F are 0 and $\rho_{(0,1)}, \dots, \rho_{(0,x_1-1)}$.

(C.3) The vertices $\rho_{(j,1)}, \ldots, \rho_{(j,x_{j+1}-1)}$ share the same parent with a vertex $u_j \in [n]$, for $1 \le j \le n-1$, such that $u_i \ne u_j$ if $i \ne j$.

(C.4) There is an ordering among $\rho_{(j,i)}$, which is created by a breadth-first search in F, $\rho_{(j,i)} < \rho_{(j',i')}$ if $j < j'$, or $j = j'$ and $i < i'$.

Let $\mathcal{F}_n(\mathbf{x})$ denote the set of rooted forests satisfying conditions (C.1)–(C.4). It is easy to see that there is a bijection $\varphi_{\mathbf{x}}$ between $\mathcal{P}_n(\mathbf{x})$ and $\mathcal{F}_n(\mathbf{x})$, with the triplet-labeled rooted forests as the intermediate stage, which is established in a similar manner to the bijection φ given in Section 2.

Let us consider $\mathbf{x}$-parking functions α for a specified case $\mathbf{x} = (1, \ldots, 1, a, 1, \ldots, 1)$, where a occurs at the k-th entry of $\mathbf{x}$ and $k \ge 2$. We call α a *(k, a)-inflating parking function*. Let $\mathcal{Q}_n(k, a)$ denote the set of (k, a)-inflating parking functions of length n. Given an $\alpha = (a_1, \ldots, a_n) \in \mathcal{Q}_n(k, a)$, we observe that the non-decreasing rearrangement $b_1 \le \cdots \le b_n$ of α satisfies $b_i \le i$ if $i \le k-1$, and $b_i \le i + a - 1$ otherwise. Define $\tau_\alpha(i) = \pi_\alpha(i)$, for $1 \le \pi_\alpha(i) \le k-1$, and $\tau_\alpha(i) = \pi_\alpha(i) + a - 1$, for $k \le \pi_\alpha(i) \le n$.

In addition to the triplets $(0, 0, 0)$ and $\{(i, a_i, \tau_\alpha(i)) | \, 1 \le i \le n\}$, we locate the $(k-1)$-th term in the non-decreasing rearrangement of α, say $b_{k-1} = a_r$, and associate it with another $a - 1$ triplets $(\rho_1, a_r, k), \ldots, (\rho_{a-1}, a_r, k + a - 2)$. The tree T_α associated with α is on the above triplets with the root at $(0, 0, 0)$. For any two vertices $v = (x_1, y_1, z_1)$ and $u = (x_2, y_2, z_z)$, u is a child of v if $y_2 = z_1 + 1$.

For example, take $k = 4$ and $a = 3$. Consider the $(4, 3)$-inflating parking $\alpha = (5, 1, 4, 5, 1, 10, 3, 3, 7)$. We have the permutation $(\pi_\alpha(1), \ldots, \pi_\alpha(n)) = (6, 1, 5, 7, 2, 9, 3, 4, 8)$ and $(\tau_\alpha(1), \ldots, \tau_\alpha(n)) = (8, 1, 7, 9, 2, 11, 3, 6, 10)$. Note that $a_7 = 3$ is the third term in the non-decreasing rearrangement of α. There associate two additional triplets $w_1 = (\rho_1, 3, 4)$ and $w_2 = (\rho_2, 3, 5)$. The tree associated with α is shown on the left of Figure 6.

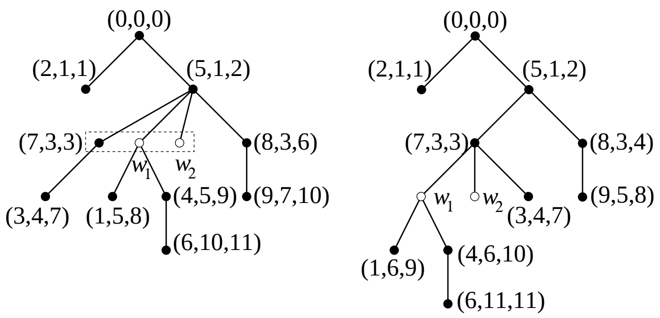


FIGURE 6. the triplet-labeled rooted tree associated with the $(4, 3)$-inflating parking function $\alpha = (5,1,4,5,1,10,3,3,7)$ and the tree associated with $\phi(\alpha)$.

Theorem 5.1. *There is a bijection between the set of (k, a)-inflating parking functions of length n and the set of ordinary parking functions of length $n + a - 1$ with the initial $a - 1$ terms equal to k.*

Proof: We shall establish a bijection ϕ between two sets. Let T_α be the tree associated with α mentioned above. Locate the vertex of T_α with the third entry equal to $k - 1$, say $u = (r, a_r, k - 1)$, and let $w_j = (\rho_j, a_r, k + j - 1) \in T_\alpha$, for $1 \le j \le a - 1$. We form a new triplet-labeled rooted tree from $T_\alpha - (T(w_1) \cup \cdots \cup T(w_{a-1}))$ by attaching $T(w_1), \ldots, T(w_{a-1})$ to u so that $w_1, \ldots, w_{a-1}$ become the first $a-1$ children of u. Then update the second and the third entries of all triplets by algorithm A. Sorting the triplets by the first entries in the order $\rho_1, \ldots, \rho_{a-1}$ and then $1, \ldots, n$, we obtain the corresponding ordinary parking function $\phi(\alpha) = (a_{\rho_1}, \ldots, a_{\rho_{a-1}}, a_1, \ldots, a_n)$ from their second entries, where $a_{\rho_1} = \cdots = a_{\rho_{a-1}} = k$. $\square$

As illustrated in Figure 6, the triplet-labeled rooted tree associated with $\phi(\alpha)$ is shown on the right, where $w_1 = (\rho_1, 4, 5)$ and $w_2 = (\rho_2, 4, 6)$ and $\phi(\alpha) = (4, 4, 6, 1, 4, 6, 1, 11, 3, 3, 5)$. Note that so far we have solved explicitly the number of (k, a)-inflating parking functions for the case $a = 2$. We are interested to know explicit formulas for other cases.

The forest structure reveals that the number of k-leading $\mathbf{x}$-parking functions is a step function in k.

Theorem 5.2. *For any $\mathbf{x} = (x_1, \ldots, x_n)$, let $d_i = x_1 + \cdots + x_i$, $(1 \le i \le n)$ and $d_0 = 0$. If $d_{k-1} + 1 \le p, q \le d_k$, then the number of p-leading $\mathbf{x}$-parking functions of length n is equal to the number of q-leading $\mathbf{x}$-parking functions of length n.*

References

[1] D. Foata, J. Riordan, Mappings of acyclic and parking functions, *Aequationes Math.* **10** (1974), 10–22.
[2] J. Françon, Acyclic and parking functions, *J. Combin. Theory Ser. A* **18** (1975), 27–35.
[3] J. D. Gilbey, L. H. Kalikow, Parking functions, valet functions and priority queues, *Discrete Math.* **197/198** (1999), 351–373.
[4] A. G. Konheim, B. Weiss, An occupancy discipline and applications, *SIAM J. Appl. Math.* **14** (1996), 1266–1274.
[5] J. P. S. Kung, C. H. Yan,887Gončarov polynomials and parking functions, *J. Combin. Theory Ser. A* **102** (2003), 16–37.
[6] J. Pitman, R. Stanley, A polytope related to empirical distributions, plane trees, parking functions, and the associahedron, *Discrete Comput. Geom.* **27(4)** (2002), 603–634.
[7] J. Riordan, Ballots and trees, *J. Combin. Theory* **6** (1969), 408–411.
[8] M. P. Schützenberger, On an enumeration problem. *J. Combin. Theory* **4** (1968), 219–221.
[9] R. P. Stanley, Hyperplane arrangements, interval orders and trees, *Proc. Nat. Acad. Sci.* **93** (1996), 2620–2625.
[10] R. P. Stanley, Parking functions and non-crossing partitions, The Wilf Festschrift, *Electro. J. Combin.* **4** (1997), R20.
[11] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, Cambridge, 1999.
[12] C. H. Yan, Generalized tree inversions and k-parking functions, *J. Combin. Theory Ser. A* **79** (1997), 268–280.
[13] C. H. Yan, On the enumeration of generalized parking functions, *Congressus Numerantium* **147** (2000), 201–209.
[14] C. H. Yan, Generalized parking functions, tree inversions and multicolored graphs, *Adv. in Appl. Math.* **27** (2001), 641–670.

Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 811, Taiwan, ROC
E-mail address: speu@nuk.edu.tw

Mathematics Faculty, National Pingtung Institute of Commerce, Pingtung 900, Taiwan, ROC
E-mail address: tsfu@npic.edu.tw

Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 811, Taiwan, ROC

ON GROWTH OF SOLVABLE LIE SUPERALGEBRAS AND GENERATING FUNCTIONS

S. G. KLEMENTYEV AND V. M. PETROGRADSKY

ABSTRACT. Finitely generated solvable Lie algebras have an intermediate growth between polynomial and exponential. Recently the second author suggested the scale to measure such an intermediate growth of Lie algebras. The growth was specified for solvable Lie algebras $F(\mathbf{A}^q, k)$ with a finite number of generators k, and which are free with respect to a fixed solubility length q. Later, an application of generating functions allowed to obtain more precise asymptotic. These results were obtained in generality of polynilpotent Lie algebras.

Now we consider the case of Lie superalgebras; we announce main results and describe methods. Our goal is to compute the growth for $F(\mathbf{A}^q, m, k)$, the free solvable Lie superalgebra of length q with m even and k odd generators. The proof is based upon a precise formula of the generating function for this algebra obtained earlier. The result is obtained in the generality of free polynilpotent Lie superalgebras.

We also consider almost solvable finitely generated Lie algebras and establish an upper bound on their growth in terms of our scale of subexponential functions.

1. INTRODUCTION

The ground field is denoted by K. Recall that a $\mathbb{Z}_2$-graded algebra $L = L_+ \oplus L_-$ is called a *Lie superalgebra* if it satisfies the following graded identities [27]. Let $\epsilon(L_+, L_+) = \epsilon(L_+, L_-) = \epsilon(L_-, L_+) = 1$, and $\epsilon(L_-, L_-) = -1$. We suppose that

$$[x, y] = -\epsilon(x, y)[y, x], \qquad x, y \in L_\pm \qquad \text{(anticommutativity)};$$
$$[x, [y, z]] = [[x, y], z] - \epsilon(y, z)[[x, z], y], \qquad x, y, z \in L_\pm \qquad \text{(Jacobi identity)}.$$

(In case $\operatorname{char} K = 2$ some more additional assumptions should be imposed [2], also in case $\operatorname{char} K = 3$ one requires $[[y, y], y] = 0$ for all $y \in L_-$, this identity being satisfied in other characteristics). A *variety* of (Lie) (super)algebras is a class of all (Lie) (super)algebras that satisfy some set of (graded or non-graded) identical relations. Concerning varieties of Lie algebras we refer the reader to the monograph [1]. On Lie superalgebras and their varieties see also [2] and [16].

Let L be a Lie (super)algebra. Then the *lower central series* is defined by iteration $L^1 = L$, $L^{i+1} = [L, L^i]$, $i = 1, 2, \dots$. Now L is called *nilpotent* of class s provided that $L^{s+1} = \{0\}$. All Lie algebras nilpotent of class s form the variety denoted by $\mathbf{N}_s$. This notation we also shall use for the variety of nilpotent Lie superalgebras of class s. Recall that L is *polynilpotent* with a tuple $(s_q, \dots, s_2, s_1)$ iff there exists a chain of ideals $0 = L_{q+1} \subset L_q \subset \cdots \subset L_2 \subset L_1 = L$ such that $L_i/L_{i+1} \in \mathbf{N}_{s_i}$, $i = 1, \dots, q$. All polynilpotent Lie (super)algebras with the fixed tuple $(s_q, \dots, s_2, s_1)$ form the variety denoted by $\mathbf{N}_{s_q} \cdots \mathbf{N}_{s_2}\mathbf{N}_{s_1}$. In the case $s_q = \cdots = s_1 = 1$, one obtains as a particular case the variety $\mathbf{A}^q$ of *solvable* Lie (super)algebras of length q. Polynilpotent varieties of groups and Lie algebras as well as their interactions were studied by A.L. Shmelkin [28]. A basis for free polynilpotent Lie algebras was constructed by L.A. Bokut [4], for the case of free solvable Lie algebras, see also the monograph of C. Reutenauer [26].

2000 *Mathematics Subject Classification.* 17B30, 16P90, 17B01, 17B35, 17B65, 05A15, 05A16.
Both authors are partially supported by Grant RFBR-04-01-00739.

Let L be a Lie superalgebra, then by $L = L_+ \oplus L_-$ we denote its decomposition into the even and odd components. Suppose that $\mathbf{M}$ is a variety of Lie superalgebras, then by $F(\mathbf{M}, X)$, $X = X_+ \cup X_-$ we denote its free algebra generated by X. Let $X_+ = \{x_i | i \in I_+\}$, $X_- = \{x_j | j \in I_-\}$. Recall that $F(\mathbf{M}, X)$ is the algebra generated by X and such that for all $H = H_+ \oplus H_- \in \mathbf{M}$ and any $y_i \in H_+$, $i \in I_+$; $y_j \in H_-$, $j \in I_-$, there exists a homomorphism $\phi : F(\mathbf{M}, X) \to H$ with $\phi(x_i) = y_i$, $i \in I_+ \cup I_-$. In case $|X_+| = m$, $|X_-| = k$ we also denote $F(\mathbf{M}, X) = F(\mathbf{M}, m, k)$.

One verifies that each polynilpotent Lie (super)algebra is solvable, i.e. belongs to some $\mathbf{A}^q$ for a sufficiently large q. So, by studying free polynilpotent Lie (super)algebras we study first, some solvable Lie (super)algebras. Second, this setting, as a particular case, includes the free solvable (super)algebras $F(\mathbf{A}^q, X) = F(\underbrace{\mathbf{N}_1 \cdots \mathbf{N}_1}_{q \text{ times}}, X)$.

Let A be a Lie (associative) algebra over a field K, generated by a finite set X. Denote by $A^{(X,n)}$ the subspace spanned by all monomials in X of length not exceeding n. Denote

$$\gamma_A(n) = \gamma_A(X, n) = \dim_K A^{(X,n)}, \qquad n \in \mathbb{N};$$
$$\lambda_A(n) = \gamma_A(n) - \gamma_A(n-1), \qquad n \in \mathbb{N}.$$

where $\dim_K$ stands for the dimension of a vector space over K. If A is an associative algebra with unity then we consider that this unity belongs to $A^{(X,n)}$, $n \geq 0$, and $\gamma_A(0) = \lambda_A(0) = 1$. On functions $f : \mathbb{N} \to \mathbb{R}^+$, where $\mathbb{R}^+ = \{\alpha \in \mathbb{R} \,|\, \alpha > 0\}$, we consider the partial order: $f(n) \overset{\text{a}}{\leq} g(n)$ iff there exists $N > 0$, such that $f(n) \leq g(n)$, $n \geq N$.

Consider two extreme examples of growth. Suppose that A is a free associative algebra (or a free Lie algebra) of finite rank. Then the growth function $\gamma_A(X, n)$ is an exponential. On the other hand, let $A = U(L)$ be the universal enveloping algebra of a finite dimensional Lie algebra L. Then the growth function $\gamma_A(X, n)$ is a polynomial of degree $k = \dim_K L$, and the degree k is extracted by computing of the Gelfand-Kirillov dimension [13].

But there are growths between these two extreme types of growth. The growth less than any exponent is called *subexponential.* If it is also greater than any polynomial growth, then it is called *intermediate.* For study of such growths the following series of dimensions has been suggested [17], [18]. Denote by iteration

$$\ln^{(1)} n = \ln n; \qquad \ln^{(q+1)} n = \ln(\ln^{(q)} n), \qquad q = 1, 2, \ldots.$$

Consider the series of functions $\Phi_\alpha^q(n)$, $q = 1, 2, 3, \ldots$ of the natural argument with the parameter $\alpha \in \mathbb{R}^+$:

$$\Phi_\alpha^1(n) = \alpha; \qquad q = 1;$$
$$\Phi_\alpha^2(n) = n^\alpha; \qquad q = 2;$$
$$\Phi_\alpha^3(n) = \exp(n^{\alpha/(\alpha+1)}); \qquad q = 3;$$
$$\Phi_\alpha^q(n) = \exp\left(\frac{n}{(\ln^{(q-3)} n)^{1/\alpha}}\right), \qquad q = 4, 5, \ldots.$$

Suppose that $f(n)$ is a positive valued function of a natural argument. We define the *(upper) dimension of level* q, $q = 1, 2, 3, \ldots$, and the *lower dimension of level* q by

$$\operatorname{Dim}^q f(n) = \inf\{\alpha \in \mathbb{R}^+ \mid f(n) \overset{\text{a}}{\leq} \Phi_\alpha^q(n)\},$$
$$\underline{\operatorname{Dim}}^q f(n) = \sup\{\alpha \in \mathbb{R}^+ \mid f(n) \overset{\text{a}}{\geq} \Phi_\alpha^q(n)\}.$$

Suppose that A is a finitely generated algebra. We define the *q-dimension* and the *lower q-dimension*, $q = 1, 2, 3, \ldots$, of A by

$$\operatorname{Dim}^q A = \operatorname{Dim}^q \gamma_A(n), \quad \underline{\operatorname{Dim}}^q A = \underline{\operatorname{Dim}}^q \gamma_A(n).$$

Roughly speaking, the condition $\operatorname{Dim}^q A = \underline{\operatorname{Dim}}^q A = \alpha$ means that the growth function $\gamma_A(n)$ behaves like $\Phi_\alpha^q(n)$. These q-dimensions do not depend on a generating set X [18].

Remark that 1-dimension coincides with the dimension of the vector space A over K. Dimensions of level 2 are exactly the upper and lower Gelfand-Kirillov dimensions [7], [13], and [15]. Dimensions of level 3 correspond to the superdimensions of [5] up to normalization (see [18]). If L is a Lie (super)algebra, then by $U(L)$ we denote its universal enveloping algebra. The following two theorems are the crucial facts about this scale of functions and dimensions.

Theorem 1.1 ([17], [18]). *Let L be a finitely generated Lie algebra with* $\mathrm{Dim}^q L = \alpha > 0$, $q = 1, 2, \ldots$ *. Also for $q \geq 3$ suppose that $\underline{\mathrm{Dim}}^q L = \alpha$ and for $q = 2$ suppose that $\underline{\mathrm{Dim}}^2 \lambda_L(n) = \alpha - 1$, and $\alpha \geq 1$. Then*

$$\underline{\mathrm{Dim}}^{q+1} U(L) = \mathrm{Dim}^{q+1} U(L) = \alpha.$$

A. Lichtman proved that the growth of finitely generated solvable Lie algebras is subexponential [14]. The following result specifies the growth of such algebras in terms of our scale of functions.

Theorem 1.2. [18] *Let $L = F(\mathbf{N}_{s_q} \cdots \mathbf{N}_{s_2}\mathbf{N}_{s_1}, k)$ be the free polynilpotent Lie algebra of rank k, $k \geq 2$, $q \geq 2$. Then*

$$\underline{\mathrm{Dim}}^q L = \mathrm{Dim}^q L = s_2 \dim_K F(\mathbf{N}_{s_1}, k).$$

As a particular case, we have the following.

Corollary 1.1. [17] *Let $L = F(\mathbf{A}^q, k)$ be the free solvable Lie algebra of length q and rank k, $k \geq 2$, $q \geq 1$. Then $\underline{\mathrm{Dim}}^q L = \mathrm{Dim}^q L = k$.*

More precise asymptotic for free polynilpotent Lie algebras $L = F(\mathbf{N}_{s_q} \cdots \mathbf{N}_{s_1}, k)$ is found in [19]. As an application, this result gave also an answer to the question of M. I. Kargapolov [12] to describe the lower central series ranks for free polynilpotent finitely generated groups. Earlier exact recursive formulae were found by G.P. Egorychev [6]. Another answer to this problem was given by the second author by describing the asymptotic behaviour of these ranks [19]. In general, the approach [19] heavily relies on the use of generating functions and study of their growth.

Denote by $\zeta(*)$, $\Gamma(*)$, $\mu(*)$ the Riemann zeta-function, the Gamma-function, and the Möbius function, respectively. By $\delta_{i,j}$ denote the Kronecker symbol.

2. Main result: Growth of solvable Lie superalgebras

Now, our goal is to study the growth of solvable Lie *super*algebras. Let us formulate the first result in this direction. Its says that $F(\mathbf{N}_{s_q} \cdots \mathbf{N}_{s_1}, m, k)$ lies, as a rule, also on the level q.

Theorem 2.1. *Let $L = F(\mathbf{N}_{s_q} \cdots \mathbf{N}_{s_2}\mathbf{N}_{s_1}, m, k)$ be the free polynilpotent Lie superalgebra, where $m + k \geq 2$ and $q \geq 2$. Then*

(1) $\underline{\mathrm{Dim}}^q L = \mathrm{Dim}^q L = s_2 \dim F_+(\mathbf{N}_{s_1}, m, k)$.
(2) $\dim F_+(\mathbf{N}_{s_1}, m, k) = 0$ *is equivalent to* $s_1 = 1$ *and* $m = 0$. *In this case we also suppose that $q \geq 3$, then* $\underline{\mathrm{Dim}}^{q-1} L = \mathrm{Dim}^{q-1} L = s_3 \dim F_+(\mathbf{N}_{s_2}\mathbf{A}, 0, k)$.

As a particular case, we obtain.

Corollary 2.1. *Let $L = F(\mathbf{A}^q, m, k)$ be the free solvable Lie superalgebra of length q, where $m + k \geq 2$, $q \geq 2$. Then*

(1) $\underline{\mathrm{Dim}}^q L = \mathrm{Dim}^q L = m$;
(2) *If $m = 0$ and $q \geq 3$ then* $\underline{\mathrm{Dim}}^{q-1} L = \mathrm{Dim}^{q-1} L = 1 + (k-1)2^{k-1}$.

We shall refer to the second case as the *degenerate case*. Now we formulate our main result, that immediately implies Theorem 2.1. Complete proofs of main results will appear in [10].

Theorem 2.2. *Consider the polynilpotent variety* $\mathbf{V} = \mathbf{N}_{s_q} \cdots \mathbf{N}_{s_1}$, $q \geq 2$, *of Lie superalgebras. Suppose that* $L = F(\mathbf{V}, m, k)$ *is its free superalgebra, generated by set* $X = X_+ \cup X_-$, $X_+ = \{x_1, \dots, x_m\}$, $X_- = \{x_{m+1}, \dots, x_{m+k}\}$. *Then*

(1) *If* $s_1 > 1$ *or* $m > 0$ *then*

$$\gamma_L(X,n) = \begin{cases} \dfrac{A+o(1)}{N!} n^N, & q=2; \\ \exp\left((C+o(1))n^{\frac{N}{N+1}}\right), & q=3; \\ \exp\left((B^{1/N}+o(1))\dfrac{n}{(\ln^{(q-3)} n)^{1/N}}\right), & q \geq 4; \end{cases}$$

where the constants are

$$N = s_2 \dim F_+(\mathbf{N}_{s_1}, m, k), \qquad A = \frac{1}{s_2}\left((m+k-1)\frac{2^{\dim F_-(\mathbf{N}_{s_1},m,k)}}{\prod_{j=1}^{s_1} j^{\psi_+(j)}}\right)^{s_2},$$

$$B = s_3 A \zeta(N+1)\left(1 - \frac{1-\delta_{k,0}}{2^{N+1}}\right), \qquad C = \left(1+\frac{1}{N}\right)(BN)^{\frac{1}{1+N}};$$

and $\psi_+(j) = \dim F_{j,+}(\mathbf{N}_{s_1}, m, k)$, $1 \leq j \leq s_1$, *are dimensions of the even parts of the homogeneous components.*

(2) *If* $s_1 = 1$, $m = 0$, *and additionally* $q \geq 3$ *then*

$$\gamma_L(X,n) = \begin{cases} \dfrac{A+o(1)}{N!} n^N, & q=3; \\ \exp\left((C+o(1))n^{\frac{N}{N+1}}\right), & q=4; \\ \exp\left((B^{1/N}+o(1))\dfrac{n}{(\ln^{(q-4)} n)^{1/N}}\right), & q \geq 5; \end{cases}$$

where the constants in this case are

$$N = s_3 \dim F_+(\mathbf{N}_{s_2}\mathbf{A}, 0, k), \qquad A = \frac{1}{s_3}\left((k-1)\frac{2^{\dim F_-(\mathbf{N}_{s_2}\mathbf{A},0,k)}}{\prod_{2|j} j^{\phi(j)}}\right)^{s_3},$$

$$B = s_4 A \zeta(N+1)\left(1 - \frac{1}{2^{N+1}}\right), \qquad C = \left(1+\frac{1}{N}\right)(BN)^{\frac{1}{1+N}};$$

and $\phi(j) = \dim F_j(\mathbf{N}_{s_2}\mathbf{A}, 0, k)$ *are dimensions of the homogeneous components of the finite dimensional algebra* $F(\mathbf{N}_{s_2}\mathbf{A}, 0, k)$.

We draw the readers attention that these asymptotics hold for the growth with respect to the standard generating set X only. One can also easily derive the corollary for the particular case of free solvable Lie superalgebras $L = F(\mathbf{A}^q, m, k)$, $q \geq 2$.

Our approach heavily relies on application of generating functions. Suppose that an algebra A is generated by a finite set X and is homogeneous with respect to the degree in X. So, we have $A = \mathop{\oplus}\limits_{n=0}^{\infty} A_n$ and $\dim A_n = \lambda_A(X,n)$. In this case we define the Hilbert-Poincaré series

$$\mathcal{H}_X(A,t) = \sum_{n=0}^{\infty} \dim A_n t^n.$$

We introduce some more series in the next section. The following result plays an important role in our proof. It is also of independent interest.

Theorem 2.3. *Consider the polynilpotent variety* $\mathbf{V} = \mathbf{N}_{s_q} \cdots \mathbf{N}_{s_1}$, $q \geq 2$, *of Lie superalgebras and* $L = F(\mathbf{V}, m, k)$, $m + k \geq 2$. *Then the Hilbert-Poincaré series with respect to the standard generating set* X *has the following growth, while* $t \to 1 - o$:

(1) *If* $s_1 > 1$ *or* $m > 0$ *then*

$$\mathcal{H}_X(L,t) = \begin{cases} \dfrac{A+o(1)}{(1-t)^N}, & q = 2; \\ \exp^{(q-2)}\left(\dfrac{B+o(1)}{(1-t)^N}\right), & q \geq 3. \end{cases}$$

(2) *If* $s_1 = 1$, $m = 0$, *and additionally* $g \geq 3$ *then*

$$\mathcal{H}_X(L,t) = \begin{cases} \dfrac{A+o(1)}{(1-t)^N}, & q = 3; \\ \exp^{(q-3)}\left(\dfrac{B+o(1)}{(1-t)^N}\right), & q \geq 4. \end{cases}$$

where the constants N, A, *and* B *are the same as in the respective cases of Theorem 2.2.*

3. Generating functions for solvable Lie superalgebras

As it is seen from the previous theorem, the generating functions and their growth play an important role in our arguments. The goal of this section is to formulate a precise formula for generating functions for free solvable (more generally, polynilpotent) Lie superalgebras (Theorem 3.1). The proof of our main result is based on this formula. So, we solve a problem of enumerative combinatorics [8, 29, 30].

Let $X = \{x_i | i \in I\}$, $I = I_+ \cup I_-$ be an at most countable generating set for a superalgebra $A = A_+ \oplus A_-$. We assume that A is multihomogeneous with respect to X. For example, this is the case when A is the relatively free algebra of some variety of superalgebras and $X = X_+ \cup X_-$ is the free generating set. Denote $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. We define the grading $A = \underset{\alpha\in\mathbb{N}_0^I}{\oplus} A_\alpha$ induced by setting $x_i \in A_{\alpha_i}$, where $\alpha_i = (\dots,0,1,0,\dots) \in \mathbb{N}_0^I$ with 1 on the ith place. Also we define a homomorphism $\varepsilon : \mathbb{N}_0^I \to \{\pm 1\}$ by $\varepsilon(\alpha_i) = \pm 1$, $i \in I_\pm$. The sequence $\alpha = \sum_{i\in I} m_i\alpha_i \in \mathbb{N}_0^I$ has only finitely many nonzero entrees m_i, and we denote $\mathbf{t}^\alpha = \prod_{i\in I} t_i^{m_i}$. We consider the formal power series ring $\mathbb{Q}[[\mathbf{t}]] = \mathbb{Q}[[t_i | i \in I]]$. Suppose that $W = \underset{\alpha\in\mathbb{N}_0^I}{\oplus} W_\alpha \subset A$ is a homogeneous subalgebra, then we define its *Hilbert-Poincaré series*

$$\mathcal{H}_X(W,\mathbf{t}) = \mathcal{H}_X(W, t_i | i \in I) = \sum_{\alpha\in\mathbb{N}_0^I} \dim(W_\alpha)\,\mathbf{t}^\alpha \in \mathbb{Q}[[\mathbf{t}]].$$

We use the following operators acting on the formal power series ring $\mathbb{Q}[[\mathbf{t}]]$. These operators were introduced in [22], [23]:

$$\phi^{[-]}(\mathbf{t}) = \phi\Big|_{t_i=\varepsilon(t_i)t_i,\, i\in I};$$

$$\phi^{[m]}(\mathbf{t}) = \phi\Big|_{t_i=\varepsilon(t_i)^{m+1}t_i^m,\, i\in I}, \quad m \in \mathbb{N};$$

$$\mathcal{E}(\phi)(\mathbf{t}) = \exp\left(\sum_{m=1}^{\infty} \frac{1}{m}\phi^{[m]}(\mathbf{t})\right) = \prod_{\alpha\in\mathbb{N}_0^I} (1-\varepsilon(\alpha)\mathbf{t}^\alpha)^{-\varepsilon(\alpha)b_\alpha}, \tag{1}$$

$$\text{where} \quad \phi(\mathbf{t}) = \sum_{\alpha\in\mathbb{N}_0^I} b_\alpha \mathbf{t}^\alpha \in \mathbb{Q}[[\mathbf{t}]]_0;$$

and $\mathbb{Q}[[\mathbf{t}]]_0$ denotes series with zero constant term. For the equality (1) see [22]. For the right hand side of (1) to be defined we also suppose that $b_\alpha \in \mathbb{Z}$, such expressions enumerate universal enveloping algebras [31]. The importance of $\mathcal{E}$ is explained by the following.

Lemma 3.1 ([22])**.** *Let* L *be a Lie superalgebra generated by an at most countable generating set* $X = X_+ \cup X_-$ *and multihomogeneous with respect to it. Then the series for the universal enveloping algebra equals* $\mathcal{H}_X(U(L),\mathbf{t}) = \mathcal{E}(\mathcal{H}_X(L,\mathbf{t}))$.

The proof of our main result is based on the following explicit formula of the generating function for free polynilpotent Lie superalgebras [23]. In a particular case of the free metabelian Lie superalgebra $F(\mathbf{A}^2, m, k)$ this series was found in [3], see also [2].

Theorem 3.1 ([23])**.** *Let $L = F(\mathbf{N}_{s_q} \cdots \mathbf{N}_{s_2}\mathbf{N}_{s_1}, X)$ be the free polynilpotent Lie superalgebra generated by an at most countable generating set $X = \{x_i | i \in I\}$, where $I = I_+ \cup I_-$ and $X_+ = \{x_i | i \in I_+\}$, $X_- = \{x_i | i \in I_-\}$ are the even and odd generators, respectively. We define functions $g_i(\mathbf{t})$, $f_i(\mathbf{t}) \in \mathbb{Q}[[\mathbf{t}]]$, $i = 0, \dots, q$ by $g_0(\mathbf{t}) = 0$, $f_0(\mathbf{t}) = \sum_{i\in I} t_i$, and*

$$g_i(\mathbf{t}) = g_{i-1}(\mathbf{t}) + \sum_{m=1}^{s_i} \frac{1}{m} \sum_{a|m} \mu(a) \left(f_{i-1}^{[a]}(\mathbf{t})\right)^{m/a}, \quad 1 \le i \le q;$$

$$f_i(\mathbf{t}) = 1 + \left(\sum_{i\in I} t_i - 1\right) \cdot \mathcal{E}(g_i(\mathbf{t})), \quad 1 \le i \le q.$$

Then $\mathcal{H}_X(L, \mathbf{t}) = g_q(\mathbf{t})$.

Now assume that the generating set is finite. Let $I = I_+ \cup I_-$, $I_+ = \{1, \dots, m\}$, $I_- = \{m+1, m+2, \dots, m+k\}$, $1 < m+k < \infty$. We have the generating set $X = X_+ \cup X_-$, where $X_+ = \{x_1, \dots, x_m\}$, and $X_- = \{x_{m+1}, \dots, x_{m+k}\}$. Let $W \subset A$ be a multihomogeneous subspace. Then one has components for the gradation by the multidegree W_α, degree W_n, and "superdegree" W_{ij}, where the last space consists of elements of degree i with respect to X_+, and degree j with respect to X_-. So, we obtain the following different Hilbert–Poincaré series

$$\mathcal{H}_X(W, \mathbf{t}) = \mathcal{H}(W, t_1, \dots, t_{m+k}) = \sum_{\alpha \in \mathbb{N}_0^{m+k}} \dim W_\alpha \, t_1^{\alpha_1} \cdots t_{m+k}^{\alpha_{m+k}};$$

$$\mathcal{H}_X(W, x, y) = \sum_{i,j=0}^{\infty} \dim W_{ij} \, x^i y^j;$$

$$\mathcal{H}_X(W, t) = \sum_{n=0}^{\infty} \dim W_n \, t^n = \mathcal{H}(W, x, y)\Big|_{x=y=t}.$$

In [22] we used variables t_+, t_-, now in order to simplify formulas, we use symbols x, y and hope that they are not mixed up with the elements of the generating set X. We can now apply Theorem 3.1 and obtain $\mathcal{H}(L, x, y)$ just by the setting $t_i = x$ for $i \in I_+$, and $t_i = y$ for $i \in I_-$ in the formula $\mathcal{H}(L, \mathbf{t})$. The above operators look in this case as

$$\phi^{[-]}(x, y) = \phi(x, -y);$$

$$\phi^{[m]}(x, y) = \phi(x^m, (-1)^{m+1} y^m);$$

$$\mathcal{E}(\phi(x, y)) = \exp\left(\sum_{m=1}^{\infty} \frac{1}{m} \phi\big(x^m, (-1)^{m+1} y^m\big)\right);$$

$$\mathcal{E}\left(\sum_{i,j \ge 0;\ i+j>0} b_{ij} x^i y^j\right) = \prod_{i,j} \big(1 - (-1)^j x^i y^j\big)^{(-1)^{j+1} b_{ij}}.$$

These and other formulas proved to be useful earlier and allowed to obtain dimension formulas for free Lie superalgebras and study invariants of finite groups acting on free Lie superalgebras [20, 21, 22], see also similar formulas in [9].

Having the explicit formula for $\mathcal{H}(L, x, y)$, we obtain $\mathcal{H}(L, t) = \mathcal{H}(L, x, y)|_{x=t, y=t}$ and we proceed further similar to the case of Lie algebras [19] and using different asymptotic methods [24, 25].

In order to prove main results we study the growth for universal enveloping algebras of Lie superalgebras [10]. Also, some special bases for free Lie superalgebras are constructed [10].

4. Growth of almost solvable Lie algebras

We extend our results on the growth of solvable Lie algebras to *almost solvable* Lie algebras. A Lie algebra is called almost solvable if it has a solvable subalgebra of finite codimension.

Theorem 4.1 ([11])**.** *Let L be a finitely generated Lie algebra such that there exists a subalgebra $H \subseteq L$ of finite codimension and such that H is solvable of length q. Then L is of subexponential growth and*

$$\operatorname{Dim}^{q+1} L \le \dim_K(L/H).$$

By Corollary 1.1 this upper bound is exact. Indeed, we consider $L = F(\mathbf{A}^{q+1}, k)$ and its commutator subalgebra $H = L^2$. Then $\dim_K L/H = k$, H is solvable of step q, and by Corollary 1.1, we have $\operatorname{Dim}^{q+1} L = k$. On the other hand, there are no lower bounds because L can be finite dimensional.

We say that a set $X = \cup_{m=1}^{\infty} X_m$, where $|X_m| < \infty$, $m \in \mathbb{N}$, is *graded.* We extend our definitions for algebras generated by such sets, where we assume that the elements of X_m have *weight* m in computations of all growth functions and series. We define $\mathcal{H}(X,t) = \sum_{n=1}^{\infty} |X_n|\, t^n$. The next result is a particular case of Theorem 3.1.

Theorem 4.2 ([23])**.** *Let $L = F(\mathbf{A}^q, Y)$ be the free solvable Lie algebra of length q, generated by a graded set Y. Denote $g_1(t) = \mathcal{H}(Y,t)$ and $g_i(t) = g_{i-1}(t) + 1 + (\mathcal{H}(Y,t) - 1)\mathcal{E}(g_{i-1}(t))$ for $i = 2,\dots,q$. Then $\mathcal{H}_Y(L,t) = g_q(t)$.*

We show how series are used to prove Theorem 4.1. Let L be generated by $Z = \{z_1,\dots,z_k\}$. By assumption, $\dim_K L/H = N < \infty$ and $H^{(q)} = 0$. Let $F = F(X)$ be the free Lie algebra generated by $X = \{x_1,\dots,x_k\}$. We have an epimorphism $\phi : F \to L$, $\phi(x_i) = z_i$ for $i = 1,\dots,k$. Denote $D = \phi^{-1}(H)$ and $G = \operatorname{Ker}\phi$. Then $\dim_K F/D = N$, $F/G \cong L$ and $D^{(q)} \subseteq G$.

We consider the filtration $F^1 \subseteq F^2 \subseteq \cdots$, given by the degree in X. For any $V \subseteq F$ let $\operatorname{gr} V = \oplus_{n=1}^{\infty} \operatorname{gr}_n V \subseteq F$ denote the associated graded space. The subalgebra $\operatorname{gr} D \subseteq F$ is free by Shirshov-Witt theorem [1]. Let $\bar{Y} = \cup_{n=1}^{\infty} \bar{Y}_n$ be a homogeneous generating set for $\operatorname{gr} D$, where $\bar{Y}_n \subset \operatorname{gr}_n F$. Let $\mathcal{H}_X(\bar{Y},t) = \sum_{n=1}^{\infty} |\bar{Y}_n| t^n$, then we have an analogue of Schreier's formula [23]:

$$\mathcal{H}_X(\bar{Y},t) - 1 = (kt-1)\,\mathcal{E}(\mathcal{H}_X(F/\operatorname{gr} D,t)), \quad \text{where} \tag{2}$$

$\mathcal{H}_X(F/\operatorname{gr} D,t) = \sum_{n=1}^{\infty} c_n t^n$ is a polynomial, because $\sum_n c_n = \dim F/\operatorname{gr} D = \dim F/D = N$. We consider the chain of subspaces $F \supseteq D \supseteq G \supseteq D^{(q)}$. We observe that $\operatorname{gr} D/D^{(q)} \cong F(\mathbf{A}^q, \bar{Y})$. Suppose that $\mathcal{H}(F(\mathbf{A}^q,\bar{Y}),t) = \sum_{n=1}^{\infty} d_n t^n$, denote $\bar{d}_n = d_1 + \cdots + d_n$ for $n \in \mathbb{N}$. We get the following bound on the growth of L: $\gamma_L(n) \le N + \bar{d}_n$, $n \in \mathbb{N}$.

We apply Theorem 4.2 and (2) to algebra $F(\mathbf{A}^q, \bar{Y})$.

$$\begin{aligned} g_1(t) &= \mathcal{H}_X(\bar{Y},t) = 1 + (kt-1)\prod_{n\ge 1}\frac{1}{(1-t^n)^{c_n}} = \frac{A+o(1)}{(1-t)^N}, \quad t \to 1-o;\\ g_i(t) &\le g_{i-1}(t) + g_1(t)\cdot \mathcal{E}(g_{i-1}(t)), \quad 0 \le t < 1; \qquad i = 2,\dots,q. \end{aligned}$$

We apply facts on the growth of functions analytic in the unit circle [19] and obtain the following asymptotics

$$g_p(t) = \exp^{(p-1)}\left(\frac{A\zeta(N+1)+o(1)}{(1-t)^N}\right), \quad t \to 1-o; \quad p = 2,\dots,q.$$

By Theorem 4.2, $\mathcal{H}(F(\mathbf{A}^q,\bar{Y}),t) = g_q(t)$. We use this asympotics along with properties of functions analytic in the unit circle [19] and conclude that $\operatorname{Dim}^{q+1} L \le N$.

References

[1] Yu. A. Bahturin, Identical relations in Lie algebras, VNV Science Press, Utrecht, 1987.

[2] Yu. A. Bahturin, A. A. Mikhalev, V. M. Petrogradsky, and M. V. Zaicev, Infinite dimensional Lie superalgebras, de Gruyter Exp. Math. vol. 7, de Gruyter, Berlin, 1992.

[3] Yu. A. Bakhturin and V. S. Drenski, Identities of free soluble colored Lie superalgebras, *Algebra i Logika* **26** (1987), no 4, 403–418; Engl. transl., *Algebra and Logic* **26** (1987), 229–240.

[4] L. A. Bokut', Basis for free polynilpotent Lie algebras, *Algebra i Logika* **2**, no 4, (1963), 13–19.

[5] W. Borho and H. Kraft, Über die Gelfand-Kirillov-Dimension, *Math. Ann.* **220**, no 1, (1976), 1–24.

[6] G. P. Egorychev, Integral representation and the computation of combinatorial sums, Transl. Math. Monogr. vol. 59, Amer. Math. Soc., Providence, RI, 1984.

[7] I. M. Gelfand and A. A. Kirillov, Sur les corps lies aux algèbres enveloppantes des algèbres de Lie, *Inst. Hautes Études Sci. Publ. Math.* **31**, (1966), 509–523.

[8] I. P. Goulden and D. M. Jackson, Combinatorial enumeration, Wiley, New York, 1983.

[9] Seok-Jin, Kang, Graded Lie superalgebras and the superdimension formula. *J. Algebra* **204** (1998), no 2, 597–655.

[10] S.G. Klementyev and V.M. Petrogradsky, Growth of solvable Lie superalgebras, *Comm. Algebra.*, to appear.

[11] S.G. Klementyev and V.M. Petrogradsky, On growth of almost solvable Lie algebras, *Uspehi Mat. Nauk*, to appear.

[12] Kourovka Notebook, Unsolved problems in group theory, Nauka, Novosibirsk, 1967.

[13] G. R. Krause and T. H. Lenagan, Growth of algebras and Gelfand-Kirillov dimension, Revised edition. Graduate Studies in Mathematics, **22**. AMS, Providence, RI, 2000.

[14] A. I. Lichtman, Growth in enveloping algebras, *Israel J. Math.* **47**, no 4, (1984), 297–304.

[15] J.C. McConnell and J.C. Robson, Noncommutative Noetherian rings, AMS, Providence, RI, 2001.

[16] A.A. Mikhalev and A.A. Zolotykh, Combinatorial aspects of Lie superalgebras, CRC Press, Boca Raton, 1995.

[17] V. M. Petrogradsky, On some types of intermediate growth in Lie algebras, *Uspekhi Mat. Nauk* **48**, No. 5, (1993), 181–182; translation in *Russian Math. Surveys*, **48**, No. 5, (1993), 181–182.

[18] V. M. Petrogradsky, Intermediate growth in Lie algebras and their enveloping algebras, *J. Algebra* **179**, (1996), 459–482.

[19] V. M. Petrogradsky, Growth of finitely generated polynilpotent Lie algebras and groups, generalized partitions, and functions analytic in the unit circle, *Internat. J. Algebra Comput.*, **9** (1999), no 2, 179–212.

[20] V. M. Petrogradsky, On invariants of the action of a finite group on a free Lie algebra, *Sib. Mat. Zh.* **41** (2000), no. 4, 915–925; translation in *Sib. Math. J.* **41** (2000), no. 4, 763–770.

[21] V. M. Petrogradsky, On Witt's formula and invariants for free Lie superalgebras, Formal power series and algebraic combinatorics (Moscow 2000). 543–551, Springer, 2000.

[22] V. M. Petrogradsky, Characters and invariants for free Lie superalgebras, *Algebra i Analiz*, **13** (2001), no 1, 158–181. translation in *St. Petersburg Math. J.*, **13** (2002), no 1, 107–122.

[23] V. M. Petrogradsky, On generating functions for subalgebras of free Lie superalgebras, *Discrete Math.* (Conference proceedings FPSAC'99), **246** (2002), no 1–3, 269–284.

[24] A. G. Postnikov, Tauberian theory and its applications, *Trudy Mat. Inst. Steklov.*, v. 144, (1979); English transl., *Proc. Steklov. Inst. Math.*, no 2, (1980).

[25] S. Ramanujan, Collected papers, Chelsea, New-York, 1962.

[26] C. Reutenauer, Free Lie algebras, Clarendon Press, Oxford, 1993.

[27] M. Scheunert, The theory of Lie superalgebras. An introduction, Lecture Notes in Mathematics. **716**, Springer, 1979.

[28] A. L. Shmel'kin, Free polynilpotent groups, *Izv. Akad. Nauk SSSR Ser. Mat.* **28**, no 1, (1964), 91–122; English transl., *Amer. Math. Soc. Transl. ser.* 2, vol. 55, Amer. Math. Soc., Providence, RI, 1966, 270–304.

[29] R. P. Stanley, Enumerative combinatorics, Brooks/Coole, Monterey, California, 1986.

[30] R. P. Stanley, Enumerative combinatorics. Volume 2. Cambridge University Press. 1999.

[31] V. A. Ufnarovskiy, Combinatorial and asymptotic methods in algebra, Itogi Nauki i Tekhniki, Sovrem. Probl. Mat. Fund. Naprav. vol. 57, Moscow, 1989; Engl. transl., Encyclopaedia Math. Sci., vol. 57, Algebra VI, Springer, Berlin, 1995.

Department of Mechanics and Mathematics, Ulyanovsk State University, Leo Tolstoy 42, Ulyanovsk, 432970 Russia

E-mail address: `klementyev@hotbox.ru`

Department of Mechanics and Mathematics, Ulyanovsk State University, Lev Tolstoy 42, Ulyanovsk, 432970 Russia

E-mail address: `petrogradsky@hotbox.ru, petrogradsky@rambler.ru`

Party Algebra of Type B and Construction of its Irreducible Representations

Masashi KOSUDA

March 1, 2005

Abstract

Suppose that there exist two parties each of which consists of n members. The parties hold meetings splitting into several small groups. Every group consists of even number of members. Some groups may consist of members of just one of the parties. The set of seat-plans of such meetings makes an algebra called *the party algebra of type B*. We show that the party algebra of type B is semisimple by constructing a complete set of irreducible representations.

Supposez que là existent deux parties chacune dont se compose de n membres. Les parties tiennent des réunions coupant en plusieurs petits groupes. Chaque groupe se compose d'un chiffre pair des membres. Quelques groupes peuvent se composer des membres juste d'un des parties. L'ensemble de siège-plans de telles réunions fait une algèbre appelée *l'algèbre de partie du type B*. Nous prouvons que l'algèbre de partie du type B est semisimple en construisant un ensemble complet avec des représentations irréductibles.

1 Introduction

In [3], the author talked about the party algebra $P_{n,\infty} = \mathcal{A}_n$ of type $\tilde{A}$, which was generated by the symmetric group $\mathfrak{S}_n$ together with one special element f. In the talk, he showed that $P_{n,\infty}$ is semisimple and all the irreducible components are indexed by the n-tuple of Young diagrams whose weight some is equal to n. The algebra $P_{n,\infty}$ naturally becomes a subalgebra of the partition algebra $P_{n,1}(Q) = P_n(Q)$ defined by P. Martin in his papers [5, 6]. While the party algebra is isomorphic to the centralizer $\mathrm{End}_{G(1,1,k)}(V^{\otimes n}) = \mathrm{End}_{\mathfrak{S}_k}(V^{\otimes n})$ ($G(1,1,k)$ acts diagonally on $V^{\otimes n}$), the party algebra $P_{n,\infty}$ is isomorphic to the centralizer $\mathrm{End}_{G(r,1,k)}(V^{\otimes n})$ under the condition that $Q = k \geq n$ and $r > n$.

In this talk, we define $P_{n,2}(Q)$ the party algebra of type B slightly changing the definition of the party algebra of type $\tilde{A}$. The standard words of $P_{n,2}(Q)$ will have one to one correspondences with *the type B seat-plans* of size n for the meetings held by two parties under certain conditions (see Section 2). If Q is equal to a positive integer k, then we have a surjective homomorphism from the algebra $P_{n,2}(k)$ onto $\mathrm{End}_{G(2,1,k)}(V^{\otimes n})$. Moreover, if $k \geq n$, then the above homomorphism becomes injective. In particular, in this case we find that the algebra $P_{n,2}(k)$ is semisimple. We show that $P_{n,2}(Q)$ is also semisimple for any generic parameter Q by explicitly constructing a complete set of irreducible representations of the the algebra $P_{n,2}(k)$ and replacing k with Q.

Party algebra $P_{n,r}(Q)$ is also defined in terms of seat-plans and defined from the centralizer algebra of the unitary reflection group $G(r,1,k)$. This generalization is presented in Section 4.

Finally we consider the structure of $\mathrm{End}_{G(2,1,3)}V^{\otimes n}$ in Section 5. This algebra is a surjective image of $P_{n,2}(3)$ and its Bratteli diagram grows periodically in accordance with the growth of n. Our representation is also defined on this diagram. This indicates that $\mathrm{End}_{G(2,1,3)}V^{\otimes \infty}$ may give an example of subfactors.

2 Definition of the party algebra of type B

We consider the following situation. Let $D = \{d_1, d_2, \dots, d_n\}$ and $R = \{r_1, r_2, \dots, r_n\}$ be two sets each of which consists of n distinct elements such that $D \cap R = \emptyset$. We decompose $D \sqcup R$ into subsets $M_1, M_2, \dots, M_n$ (some of M_js might be empty) so that $|M_i| \in \{0, 2, \dots, 2n\}$ for $1 \leq i \leq n$. We call such a partition into subsets a *type B seat-plan* of size n. Let $P(n)$ be a set of partitions of n. If we sort M_is so that they satisfy $|M_1| \geq |M_2| \geq \cdots \geq |M_n|$, then there exists a partition $\lambda \in P(n)$ such that $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) = (|M_1|/2, |M_2|/2, \dots, |M_n|/2)$. Further, if we put $\alpha_i = |\{\lambda_k; \lambda_k = i\}|$, then the number of type B seat-plans of size n is the following:

$$\sum_{\lambda \in P(n)} \left(\frac{(2n)!}{(2\lambda_1)!(2\lambda_2)! \cdots (2\lambda_n)!} \right)^2 \cdot \frac{1}{\alpha_1! \alpha_2! \cdots \alpha_n!}. \tag{1}$$

A type B seat-plan of size n is figured as follows. Consider a rectangle with n marked points on the bottom and the same n on the top as in Figure 1. The

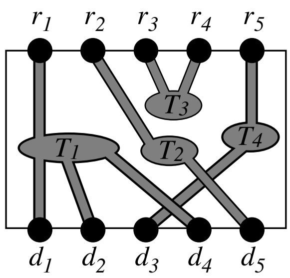


Figure 1: A seat-plan of type B

n marked points on the bottom are labeled by $d_1, d_2, \dots d_n$ from left to right.

Similarly, the n marked points on the top is labeled by $r_1, r_2, \ldots, r_n$ from left to right. If $D \sqcup R$ is divided into non-empty m subsets, then put m shaded circles in the middle of the rectangle so that they have no intersections. Each of the circles corresponds to one of the non-empty M_js. Then we join the $2n$ marked points and the m circles with $2n$ shaded bands so that the marked points labeled by the elements of M_j are connected to the corresponding circle with $|M_j|$ bands.

Now we define the product $w_1 w_2$ between two of rectangles w_1, w_2 (each of which corresponds to a seat-plan) by placing w_1 on w_2, gluing the corresponding boundaries and shrinking half along the vertical axis as in Figure 2. We

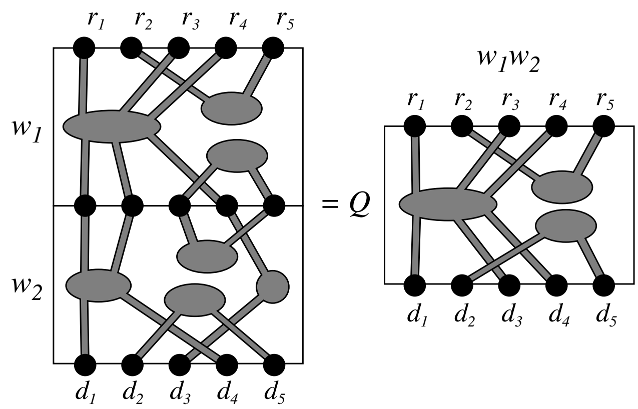


Figure 2: The product of seat-plans

then have a new diagram possibly containing some shaded islands. If there p shaded islands occur in the product, first remove holes in the islands (if they exist) and then multiply the resulting diagram by Q^p removing the p islands. By this definition, a set of linear combinations of seat-plans of size n over $\mathbb{C}$ makes an algebra $P_{n,2}(Q)$. We call it the *party algebra of type B*. We put $P_{0,2}(Q) = P_{1,2}(Q) = \mathbb{C}$.

According to the paper [11], the generators of $P_{n,2}(Q)$ is afforded by the following seat-plans:

$$\begin{aligned} s_i &= \{\{r_1, d_1\}, \{r_2, d_2\}, \ldots, \{r_i, d_{i+1}\}, \{r_{i+1}, d_i\}, \ldots, \{r_n, d_n\}\}, && (2) \\ f &= \{\{r_1, d_1, r_2, d_2\}, \{r_3, d_3\}, \ldots, \{r_{i+1}, d_i\}\}, && (3) \\ e &= \{\{r_1, r_2\}, \{d_1, d_2\}, \{r_3, d_3\}, \ldots, \{r_n, d_n\}\}. \end{aligned}$$

We further have the following proposition.

Proposition 1. *For an integer $n > 1$, the party algebra $P_{n,2}(Q)$ is characterized* *by the following generators and relations:*

$$\begin{array}{rl} \textit{generators}; & s_1, s_2, \ldots, s_{n-1}, f, e, \\ \textit{relations}; & s_i^2 = 1 \quad (1 \le i \le n-1), \\ & s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \quad (1 \le i \le n-2), \\ & s_i s_j = s_j s_i \quad (|i-j| \ge 2), \\ & e^2 = Qe, \quad f^2 = f, \\ & ef = fe = e, \quad es_1 = s_1 e = e, \quad fs_1 = s_1 f = f, \\ & es_i = s_i e, \quad fs_i = s_i f \quad (i \ge 3), \\ & es_2 e = e, \quad fs_2 fs_2 = s_2 fs_2 f, \quad fs_2 es_2 f = fs_2 f, \\ & xs_2 s_1 s_3 s_2 y s_2 s_1 s_3 s_2 = s_2 s_1 s_3 s_2 y s_2 s_1 s_3 s_2 x \quad (x, y \in \{e, f\}). \end{array}$$

Since we have one to one correspondences between the set of *standard words* of the generators above and the set of type B seat-plans, the equation (1) expresses the upper bound of the dimension of the algebra $P_{n,2}(Q)$.

Tanabe also showed the following proposition [11].

Proposition 2. *(Tanabe [11, Theorem 3.1]) Let $G(2,1,k)$ be the group of all the monomial matrices of size n whose non-zero entries are plus or minus one. Let V be a vector space of dimension k with the basis elements $e_1, e_2, \ldots, e_k$ on which $G(2,1,k)$ acts naturally. Let ϕ be the representation of the symmetric group $\mathfrak{S}_n$ on $V^{\otimes n}$ obtained by permuting the tensor product factors, i.e., for $v_1, v_2, \ldots, v_n \in V$ and for $w \in \mathfrak{S}_n = \langle s_1, \ldots, s_{n-1} \rangle$,*

$$\phi(w)(v_1 \otimes v_2 \otimes \cdots \otimes v_n) := v_{w^{-1}(1)} \otimes v_{w^{-1}(2)} \otimes \cdots \otimes v_{w^{-1}(n)}.$$

Define further $\phi(e)$ and $\phi(f)$ as follows:

$$\begin{aligned} \phi(e)(e_{p_1} \otimes e_{p_2} \otimes e_{p_3} \otimes \cdots \otimes e_{p_n}) &:= \delta_{p_1,p_2} \sum_{j=1}^{k} e_j \otimes e_j \otimes e_{p_3} \otimes \cdots \otimes e_{p_n} \\ \phi(f)(e_{p_1} \otimes e_{p_2} \otimes \cdots \otimes e_{p_n}) &:= \delta_{p_1,p_2} e_{p_1} \otimes e_{p_2} \otimes \cdots \otimes e_{p_n} \end{aligned}$$

Then $\mathrm{End}_{G(2,1,k)}(V^{\otimes n})$ is generated by $\phi(\mathfrak{S}_n)$, $\phi(e)$ and $\phi(f)$, and ϕ defines a homomorphism from $P_{n,2}(k)$ to $\mathrm{End}_{G(2,1,k)}(V^{\otimes n})$.

If $k \ge n$, then we can show that the above ϕ is injective. This implies that if Q is a positive integer k such that $k \ge n$ then we find that $P_{n,2}(Q)$ is semisimple.

In the following section, we construct a complete set of irreducible representations of $P_{n,2}(Q)$ for any generic value $Q \in \mathbb{C}$ extending the orthogonal representations of the symmetric group $\mathfrak{S}_n$. In particular, $P_{n,2}(Q)$ becomes semisimple if the parameter $Q \in \mathbb{C}$ is generic.

3 Construction

Fix a positive integer $k \ge n$. We define representations of $P_{n,2}(k)$ which turn out to become a complete set of irreducible representations. The representations are constructed on the Bratteli diagram for the sequence $P_{0,2}(k) \subset P_{1,2}(k) \subset \cdots \subset P_{n,2}(k)$ as in Figure 3. (See for example the papers [1, 7, 12, 13].)

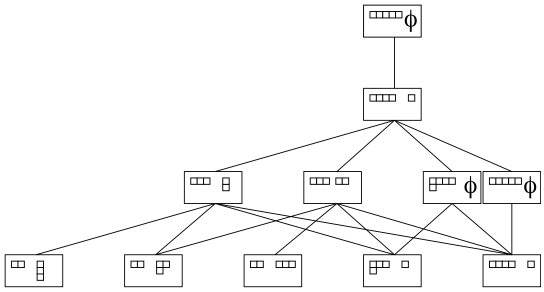


Figure 3: The Bratteli diagram for $P_{3,2}(5)$

Let $\boldsymbol{\beta} = [\alpha, \beta] = [(\alpha_1, \alpha_2, \ldots), (\beta_1, \beta_2, \ldots)]$ be an 2-tuple of Young diagrams. The 1-st [resp. 2-nd] coordinate of the tuple is referred to *the left [resp. right] board*. We consider the following sets:

$$\Lambda_k^B(2i) \quad = \quad \bigcup_{j=0}^{i} \{[\alpha, \beta] \ ; \ |\alpha| = k - 2j, \ \alpha_1 \geq k - i - j, \ |\beta| = 2j\}, \quad (4)$$

$$\Lambda_k^B(2i+1) \quad = \quad \bigcup_{j=0}^{i} \{[\alpha, \beta] \ ; \ |\alpha| = k - 2j - 1, \ \alpha_1 \geq k - i - j - 1, \ |\beta| = 2j + 1\}. \quad (5)$$

Note that $\Lambda_k^B(0) = \{[(k), \emptyset]\}$. Let $\boldsymbol{\beta} \underset{1}{\prec} \tilde{\boldsymbol{\beta}}$ or $\tilde{\boldsymbol{\beta}} \underset{1}{\succ} \boldsymbol{\beta}$ denote that $\tilde{\boldsymbol{\beta}}$ is obtained from $\boldsymbol{\beta}$ by removing one box from the Young diagram on the left board and adding the box to the Young diagram on the right board, or removing one box from the Young diagram on the right board and adding the box to the Young diagram on the left board. We also note that if $\boldsymbol{\beta} \in \Lambda_k^B(m)$, then $\tilde{\boldsymbol{\beta}} \in \Lambda_k^B(m+1)$.

For $\boldsymbol{\beta} \in \Lambda(n)$, a *tableaux* $\mathbb{T}(\boldsymbol{\beta})$ *of shape* $\boldsymbol{\beta}$ is defined by

$$\mathbb{T}(\boldsymbol{\beta}) \quad = \quad \{P = (\boldsymbol{\beta}^{(0)}, \boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(n)}); \boldsymbol{\beta}^{(0)} = [(k), \emptyset] \in \Lambda_k^B(0), \boldsymbol{\beta} \in \Lambda_k^B(n) \ \boldsymbol{\beta}^{(i)} \underset{1}{\prec} \boldsymbol{\beta}^{(i+1)} \text{ for } 0 \leq i \leq n-1\}.$$

Let $V(\boldsymbol{\beta}) = \oplus_{P \in \mathbb{T}(\boldsymbol{\beta})} \mathbb{C} v_P$ be a vector space over $\mathbb{C}$ with the standard basis $\{v_P | P \in \mathbb{T}(\boldsymbol{\beta})\}$.

For a generator s_i of $P_{n,2}(Q)$, we define a linear map on $V(\boldsymbol{\beta})$ giving a matrix $\mathcal{B}_i$ with respect to the basis $\{v_P | P \in \mathbb{T}(\boldsymbol{\beta})\}$. Namely, for a pair of tableaux $P = (\boldsymbol{\beta}^{(0)}, \boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(n)})$ and $Q = (\boldsymbol{\beta}'^{(0)}, \boldsymbol{\beta}'^{(1)}, \ldots, \boldsymbol{\beta}'^{(n)})$ of $\mathbb{T}(\boldsymbol{\beta})$ define $s_i v_P = \sum_{Q \in \mathbb{T}(\boldsymbol{\beta})} (\mathcal{B}_i)_{QP} v_Q$. If there is an $i_0 \in \{1, 2, \ldots, n-1\} \setminus \{i\}$ such that $\boldsymbol{\beta}^{(i_0)} \neq \boldsymbol{\beta}'^{(i_0)}$, then we put $(\mathcal{B}_i)_{QP} = 0$. In the following, we consider the case that $\boldsymbol{\beta}^{(i_0)} = \boldsymbol{\beta}'^{(i_0)}$ for $i_0 \in \{1, 2, \ldots, n-1\} \setminus \{i\}$.

First, we consider the case $\boldsymbol{\beta}^{(i)}$ is obtained from $\boldsymbol{\beta}^{(i-1)}$ by moving a box in the Young diagram on the left [resp. right] board to the Young diagram on the other board and $\boldsymbol{\beta}^{(i+1)}$ is obtained from $\boldsymbol{\beta}^{(i)}$ by moving another box in the Young diagram again on the left [resp. right] board to the Young diagram on the other board. Denote the Young diagram on the left board of $\boldsymbol{\beta}^{(i-1)}$ [resp. $\boldsymbol{\beta}^{(i)}$, $\boldsymbol{\beta}^{(i+1)}$] by $\lambda^{(i-1)}$ [resp. $\lambda^{(i)}$, $\lambda^{(i+1)}$] and denote the Young diagram on the right board of $\boldsymbol{\beta}^{(i-1)}$ [resp. $\boldsymbol{\beta}^{(i)}$, $\boldsymbol{\beta}^{(i+1)}$] by $\mu^{(i-1)}$ [resp. $\mu^{(i)}$, $\mu^{(i+1)}$]. Let $\lambda' \underset{1}{\subset} \lambda$ or $\lambda \underset{1}{\supset} \lambda'$ denote that λ' is obtained from λ by removing one box. Recall that if $\nu \underset{1}{\subset} \mu \underset{1}{\subset} \lambda$, then we can define the *axial distance* $d = d(\nu, \mu, \lambda)$. Namely if μ differs from ν in its r_0-th row and c_0-th column only, and if λ differs from μ in its r_1-th row and c_1-th column only, then $d = d(\nu, \mu, \lambda)$ is defined by

$$d = d(\nu, \mu, \lambda) = (c_1 - r_1) - (c_0 - r_0) = \begin{cases} h_\lambda(r_1, c_0) - 1 & \text{if } r_0 \leq r_1, \\ 1 - h_\lambda(r_0, c_1) & \text{if } r_0 > r_1. \end{cases}$$

Here $h_\lambda(i, j)$ is the *hook-length* at (i, j) in λ and for $\lambda = (\lambda_1, \lambda_2, \ldots)$ the hook-length $h_\lambda(i, j)$ is defined by

$$h_\lambda(i, j) = \lambda_i - j + |\{\lambda_l; \lambda_l \geq j\}| - i + 1.$$

If $\lambda^{(i-1)} \underset{1}{\supset} \lambda^{(i)} \underset{1}{\supset} \lambda^{(i+1)}$, then $\mu^{(i-1)} \underset{1}{\subset} \mu^{(i)} \underset{1}{\subset} \mu^{(i+1)}$. Hence we can define the axial distance $d_1 = d(\lambda^{(i+1)}, \lambda^{(i)}, \lambda^{(i-1)})$ and $d_2 = d(\mu^{(i-1)}, \mu^{(i)}, \mu^{(i+1)})$. If $|d_1| \geq 2$ [resp. $|d_2| \geq 2$], then there is a unique Young diagram $\lambda' \neq \lambda$ [resp. $\mu' \neq \mu$] which satisfies $\lambda^{(i-1)} \underset{1}{\supset} \lambda' \underset{1}{\supset} \lambda^{(i+1)}$ [resp. $\mu^{(i-1)} \underset{1}{\subset} \mu' \underset{1}{\subset} \mu^{(i+1)}$]. Similarly, if $\lambda^{(i-1)} \underset{1}{\subset} \lambda^{(i)} \underset{1}{\subset} \lambda^{(i+1)}$, then $\mu^{(i-1)} \underset{1}{\supset} \mu^{(i)} \underset{1}{\supset} \mu^{(i+1)}$, and we can define the axial distance $d_1 = d(\lambda^{(i-1)}, \lambda^{(i)}, \lambda^{(i)})$ and $d_2 = d(\mu^{(i+1)}, \mu^{(i)}, \mu^{(i-1)})$. If $|d_1| \geq 2$ [resp. $|d_2| \geq 2$], then λ' [resp. $\mu^{(i-1)}$] is defined as before. Let Q_1, Q_2, Q_3 be tableaux of shape $\boldsymbol{\beta}$ which are obtained from P by replacing $\boldsymbol{\beta}^{(i)} = [\lambda^{(i)}, \mu^{(i)}]$ on the j-th and the $(j+1)$-st board of $\boldsymbol{\beta}^{(i)}$ with $[\lambda^{(i)}, \mu']$, $[\lambda', \mu^{(i)}]$, $[\lambda', \mu']$ respectively. For the basis elements given by the above tableaux, we define the linear map by the following matrix:

$$(v_P, v_{Q_1}, v_{Q_2}, v_{Q_3}) \longmapsto (v_P, v_{Q_1}, v_{Q_2}, v_{Q_3}) \mathcal{B}_i,$$

where

$$\mathcal{B}_i = \begin{pmatrix} \frac{1}{d_1 d_2} & \frac{1}{d_1}\sqrt{\frac{d_2^2-1}{d_2^2}} & \sqrt{\frac{d_1^2-1}{d_1^2}}\frac{1}{d_2} & \sqrt{\frac{d_1^2-1}{d_1^2}}\sqrt{\frac{d_2^2-1}{d_2^2}} \\ \frac{1}{d_1}\sqrt{\frac{d_2^2-1}{d_2^2}} & -\frac{1}{d_1 d_2} & \sqrt{\frac{d_1^2-1}{d_1^2}}\sqrt{\frac{d_2^2-1}{d_2^2}} & -\sqrt{\frac{d_1^2-1}{d_1^2}}\frac{1}{d_2} \\ \sqrt{\frac{d_1^2-1}{d_1^2}}\frac{1}{d_2} & \sqrt{\frac{d_1^2-1}{d_1^2}}\sqrt{\frac{d_2^2-1}{d_2^2}} & -\frac{1}{d_1 d_2} & -\frac{1}{d_1}\sqrt{\frac{d_2^2-1}{d_2^2}} \\ \sqrt{\frac{d_1^2-1}{d_1^2}}\sqrt{\frac{d_2^2-1}{d_2^2}} & -\sqrt{\frac{d_1^2-1}{d_1^2}}\frac{1}{d_2} & -\frac{1}{d_1}\sqrt{\frac{d_2^2-1}{d_2^2}} & \frac{1}{d_1 d_2} \end{pmatrix}.$$

Second, we consider the case that the only left boards of $\boldsymbol{\beta}^{(i-1)}$ and $\boldsymbol{\beta}^{(i+1)}$ coincide. Suppose that $\boldsymbol{\beta}^{(i-1)} = [\lambda, \mu]$. Then we can write $\boldsymbol{\beta}^{(i+1)} = [\lambda, \mu']$ $(\mu \neq \mu')$. Let $\{\lambda^+_{(r)} | r = 1, 2, \ldots, b(\lambda)\}$ [resp. $\{\lambda^-_{(r')} | r' = 1, 2, \ldots, b(\lambda)'\}$] be the set of all the Young diagrams which satisfy $\lambda^+_{(r)} \underset{1}{\supset} \lambda$ [resp. $\lambda^-_{(r')} \underset{1}{\subset} \lambda$] and let $P_1, P_2, \ldots, P_{b(\lambda)}$ [resp. $Q_1, Q_2, \ldots, Q_{b(\lambda)'}$] be all the tableaux which are obtained from P by replacing $\boldsymbol{\beta}^{(i)}$ with $[\lambda^+_{(r)}, \mu \cap \mu']$ [resp. $[\lambda^-_{(r')}, \mu \cup \mu']$]. For the basis elements given by the above tableaux, we define the linear map by the following matrix:

$$\begin{aligned}
(\mathcal{B}_i)_{P_r,P_{r'}} &= \sqrt{\frac{h(\lambda)^2}{h(\lambda^+_{(r)})h(\lambda^+_{(r')})}}, \\
(\mathcal{B}_i)_{P_r,Q_{r'}} &= (\mathcal{B}_i)_{Q_{r'},P_r} = \frac{1}{d(\lambda^-_{(r')}, \lambda, \lambda^+_{(r)})}\sqrt{\frac{h(\lambda)^2}{h(\lambda^-_{(r')})h(\lambda^+_{(r)})}}, \\
(\mathcal{B}_i)_{Q_r,Q_{r'}} &= 0.
\end{aligned}$$

Here $h(\nu)$ is the product of all the hook-lengths in ν:

$$h(\nu) = \prod_{(i,j)\in\nu} h_\nu(i,j).$$

If $\boldsymbol{\beta}^{(i-1)} = [\lambda, \mu]$ and $\boldsymbol{\beta}^{(i+1)} = [\lambda', \mu]$, then the matrix $(\mathcal{B}_i)$ is similarly defined by replacing λ with μ in the argument above.

Next, we consider the case $\boldsymbol{\beta}^{(i-1)} = \boldsymbol{\beta}^{(i+1)}$. We put $\boldsymbol{\beta}^{(i-1)} = \boldsymbol{\beta}^{(i+1)} = [\lambda, \mu]$. Let $\{\lambda^+_{(r)}\}$, $\{\lambda^-_{(r')}\}$, $\{\mu^+_{(s)}\}$ and $\{\mu^-_{(s')}\}$ be the sets of Young diagrams previously defined and let $\{Q_{r',s}\}$ and $\{P_{r,s'}\}$ be the sets of tableaux obtained from P by replacing $\boldsymbol{\beta}^{(i)}$ with $[\lambda^-_{(r')}, \mu^+_{(s)}]$ and $[\lambda^+_{(r)}, \mu^-_{(s')}]$ respectively. For the basis elements given by the above tableaux, we define the linear map by the following matrix:

$$(\mathcal{B}_i)_{P,P'} = \begin{cases} \frac{1}{d(\lambda^-_{(r')},\lambda,\lambda^+_{(r)})d(\mu^-_{(s')},\mu,\mu^+_{(s)})}\sqrt{\frac{h(\lambda)^2 h(\mu)^2}{h(\lambda^-_{(r')})h(\lambda^+_{(r)})h(\mu^-_{(s')})h(\mu^+_{(s)})}} & \\ & \text{if } (P,P') = (P_{r,s'}, Q_{r',s}) \text{ or } (Q_{r',s}, P_{r,s'}), \\ \sqrt{\frac{h(\lambda)^2}{h(\lambda^+_{(r)})h(\lambda^+_{(r')})}} & \text{if } (P,P') = (P_{r,s}, P_{r',s}), \\ \sqrt{\frac{h(\mu)^2}{h(\mu^+_{(s)})h(\mu^+_{(s')})}} & \text{if } (P,P') = (Q_{r,s}, Q_{r,s'}), \\ 0 & \text{otherwise.} \end{cases}$$

Finally, we consider the remaining cases. In these cases, we can put $\boldsymbol{\beta}^{(i-1)} = [\lambda, \mu]$ and $\boldsymbol{\beta}^{(i+1)} = [\lambda', \mu']$ ($\lambda \neq \lambda', \mu \neq \mu'$ and $|\lambda| = |\lambda'|, |\mu| = |\mu'|$). Then $\boldsymbol{\beta}^{(i)}$ must be of the form $[\lambda \cup \lambda', \mu \cap \mu']$ or $[\lambda \cap \lambda', \mu \cup \mu']$. If $\boldsymbol{\beta}^{(i)}$ if the former [resp. latter] one, then the tableau P' is obtained from P by replacing $\boldsymbol{\beta}^{(i)}$ with the latter [resp. former] one. For the basis elements given by the above tableaux, we define the linear map by the following matrix:

$$(v_P, v_{P'}) \longmapsto (v_P, v_Q)\mathcal{B}_i = (v_P, v_Q)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now we have completed the preparation, we state the following main result.

Theorem 3. *Fix an integer k such that $k \geq n$. Let $\Lambda_k^B(n)$ be the set defined by (4) and (5). For $\rho_{\boldsymbol{\beta}} \in \Lambda_k^B(n)$ define $\rho_{\boldsymbol{\beta}}$ as follows:*

$$\begin{aligned}
\rho_{\boldsymbol{\beta}}(s_i)v_P &= \sum_{P' \in \mathbb{T}(\boldsymbol{\beta})} (\mathcal{B}_i)_{P'P} v_{P'}, \\
\rho_{\boldsymbol{\beta}}(f)v_P &= \begin{cases} v_P & \text{if } \boldsymbol{\beta}^{(2)} = [(k), \emptyset] \text{ or } [(k-1,1), \emptyset] \\ 0 & \text{otherwise.} \end{cases} \\
\rho_{\boldsymbol{\beta}}(e)v_P &= \begin{cases} kv_P & \text{if } \boldsymbol{\beta}^{(2)} = [(k), \emptyset] \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Then $\{\rho_{\boldsymbol{\beta}}; \boldsymbol{\beta} \in \Lambda_k^B(n)\}$ gives a complete representatives of the irreducible representations of $P_{n,2}(k)$.

In the process of the construction of $\rho_{\boldsymbol{\beta}}$, even if we replace the positive integer k with an indeterminate Q, the matrix elements of $(\mathcal{B}_i)_{P,P'}$ are similarly defined. This means the theorem above is valid for any generic parameter Q. More over if $Q = k$ and $k \geq n$, then by the Schur-Weyl reciprocity, we find that the dimension of $P_{n,2}(k)$ is equal to the square sum of the degree of $\rho_{\boldsymbol{\beta}}$ and also equal to the number of the type B seat-plans, which is presented by the expression (1). Since the degree of $\rho_{\boldsymbol{\beta}}$ does not vary even if we replace the positive integer k with the indeterminate Q, we obtain the following.

Theorem 4. *If $Q \notin \{0, 1, \ldots, n-1\}$, then the party algebra $P_{n,2}(Q)$ is semisimple and $\{\rho_{\boldsymbol{\beta}}; \boldsymbol{\beta} \in \Lambda_k^B(n)\}$ gives a complete representatives of irreducible representations of $P_{n,2}(Q)$.*

4 Party algebra $P_{n,r}(Q)$

In this section we explain how the party algebra $P_{n,r}(Q)$ is introduced from the unitary reflection group $G(r,1,k)$. Although in the paper [11] Tanabe studied the centralizer of the unitary reflection group even for the type $G(r,p,k)$, in the following we consider only the case $p = 1$.

The unitary reflection group $G(r,1,k)$ is the subgroup of $GL(k,\mathbb{C})$ generated by the set of all permutation matrices of size k and $\mathrm{diag}(\zeta, 1, 1, \ldots, 1)$ where ζ is a primitive r-th root of unity. Let V be the vector space of dimension k and suppose that it has the standard basis $\{e_1, \ldots, e_k\}$. The unitary reflection group $G(r,1,k)$ acts on V naturally and it also acts on $V^{\otimes n}$ diagonally. For $X \in \mathrm{End}V^{\otimes n}$, we denote by $X^{f_1,\ldots,f_n}_{m_1,\ldots,m_n}$ the matrix coefficients of X with respect to the basis $\{e_{m_1} \otimes \cdots \otimes e_{m_n} \mid m_1, \ldots, m_n \in [k]\}$. Since we can write $G(r,1,k) = (\mathbb{Z}/r\mathbb{Z}) \wr \mathfrak{S}_k$, in order to check whether X commutes with the action of $G(r,1,k)$ or not we first examine the following action in the tensor space. For $\sigma \in \mathfrak{S}_k$, we have

$$\sigma^{-1}X\sigma(e_{m_1}\otimes\cdots\otimes e_{m_n}) = \sum_{f_1,\dots,f_n\in[k]} X^{\sigma(f_1),\dots,\sigma(f_n)}_{\sigma(m_1),\dots,\sigma(m_n)} e_{f_1}\otimes\cdots\otimes e_{f_n}$$

Hence we have the basis of $\mathrm{End}_{\mathfrak{S}_k}V^{\otimes n}$

$$\left\{ T_\sim \middle| \begin{array}{l} \sim \text{ is an equivalence relation on } \{1,\dots 2n\} \\ \text{whose number of classes is less than or equal to } n \end{array} \right\},$$

where

$$(T_\sim)^{m_{n+1},\dots,m_{2n}}_{m_1,\dots,m_n} := \begin{cases} 1 & \text{if } (m_i = m_j \text{ if and only if } i\sim j), \\ 0 & \text{otherwise.} \end{cases}$$

Here we set $m_{n+i} := f_i$ $(1\le i\le n)$. Note that $\sim$ is zero if the number of classes for $\sim$ is more than k.

In addition to the argument above, considering the action of $\xi\in\mathbb{Z}/r\mathbb{Z}$ we find that the following equivalence relation becomes a basis of the centralizer.

Lemma 5. *(Tanabe [11, Lemma 2.1]) Let* Π_{2n} *be the set of all the partitions of* $[2n]$ *into subsets. For* $B = \{B_1,\dots,B_k\}\in\Pi_{2n}$ *(some of the parts may be empty), let* $\mathrm{bot}(B_i) := B_i\cap[n]$ *and* $\mathrm{top}(B_i) := B_i\cap([2n]\setminus[n])$ *(*$1\le i\le k$*). Let*

$$\begin{aligned} &\Pi_{2n}(r,1,k) \\ &\quad := \ \{B=\{B_1,\dots,B_k\}\in\Pi_{2n}\ ;\ |\mathrm{top}(B_i)|\equiv|\mathrm{bot}(B_i)|(mod\ r)(1\le i\le k)\}. \end{aligned}$$

Then $\{T_{\sim B}\ ;\ B\in\Pi_{2n}(r,1,k)\}$ *is a basis of* $\mathrm{End}_{G(r,1,k)}V^{\otimes n}$.

Similarly to the case $r=2$, if $k\ge n$, then we can understand that $\Pi_{2n}(r,1,k)$ becomes a basis of the party algebra $P_{n,r}(k)$ and geometrically understood. Replacing k with the parameter Q in the geometrical definition of the product, we obtain the party algebra $P_{n,r}(Q)$. We further know by Tanabe's paper [11] the party algebra $P_{n,r}(Q)$ is generated by the seat-plans s_i and f defined by (2) and (3) respectively together with the following e_r:

$$e_r = \{\{r_1,\dots,r_r\},\{d_1,\dots,d_r\},\{r_{r+1},d_{r+1}\},\dots,\{r_n,d_n\}\}.$$

5 $\mathrm{End}_{G(2,1,3)}V^{\otimes n}$

So far, we have assumed that the left coordinate of the top vertex $\boldsymbol{\beta}^{(0)} = [(k),\emptyset]$ has k boxes such that $k\ge n$. It is easy to see that the same diagram will appear even if we begin with $\boldsymbol{\beta}^{(0)} = [(k_1),\emptyset]$ such that $k_1\ge n$ and $k_1\ne k$. On the other hand, in case $k_1<n$, the resulting diagram vary. We mention what happens if we draw a diagram under the condition that $\boldsymbol{\beta}^{(0)} = [(3),\emptyset]$ according to the same recipe. The resulting diagram describes the Bratteli diagram of the centralizer algebra $\mathrm{End}_{G(2,1,3)}V^{\otimes n}$, which is a quotient of the party algebra $P_{n,2}(3)$. This diagram periodically grows in higher levels which

indicates that the centralizer may give an example of subfactors. Hence we can expect that using this algebra the Turaev-Viro-Ocneanu invariants of 3-dimensional manifolds might be calculated in the same way as in the papers [8, 9].

References

[1] Goodman, F. M., de la Harpe, P., and Jones, V. F. R., *Coxeter Graphs and Towers of Algebras.* Springer-Verlag, New York, 1989.

[2] Jones, V. F. R., "The potts model and the symmetric group." *Subfactors (Kyuzeso, 1993)* 259–267, World Sci. Publishing, River Edge, NJ, 1994.

[3] Kosuda, M., "Party algebra and construction of its irreducible representations." *Formal Power Series and Algebraic Combinatorics, FPSAC'01 13th International Conference* Arizona State University, Local Proceedings, 277-283 (2001)

[4] Kosuda, M., "Irreducible representations of the party algebra." *preprint*

[5] Martin, P., "Representations of graph Temperley-Lieb algebras." *Publ. Res. Inst. Math. Sci.* **26**(1990), 485–503.

[6] Martin, P., "Temperley-Lieb algebras for non-planar statistical mechanics – The partition algebra construction." *J. Knot Theory Ramifications* **183**(1996), 319–358.

[7] Murakami, J., "The representations of the q-analogue of Brauer's centralizer algebras and the Kauffman polynomial of links." *Publ. Res. Inst. Math. Sci.* **26**(1990), 935–945.

[8] Sato, N. and Wakui, M, "Computations of Turaev-Viro-Ocneanu invariants of 3-manifolds from subfactors." *J. Knot Theory Ramifications* **12**(2003), 543–574.

[9] Suzuki, K. and Wakui, M, "On the Turaev-Viro-Ocneanu invariant of 3-manifolds derived from the E_6-subfactor." *Kyushu J. Math.* **56**(2002), 59–81.

[10] Shephard, G. C., and Todd, J. A., "Finite unitary reflection groups." *Canad. J. Math.* **6**(1954), 274–304.

[11] Tanabe, K., "On the centralizer algebra of the unitary reflection group $G(m, p, n)$." *Nagoya. Math. J.* **148**(1997), 113–126.

[12] Wenzl, H., "On the structure of Brauer's centralizer algebras." *Ann. of Math.* **128**(1988), 173–193.

[13] Wenzl, H., "Hecke algebras of type A_n and subfactors." *Invent. Math.* **92**(1988), 349–383.

COUNTING UPPER INTERACTIONS IN DYCK PATHS

YVAN LE BORGNE

Abstract. A Dyck word w of size n is a shuffle of n copies of the word $x\overline{x}$. An upper interaction in w is an occurrence of a factor $\overline{x}^k x^k$ where $k \geq 1$. We present different methods to enumerate Dyck words according to the size and the number of upper interactions. The generating function has a rather unusual form: it is the ratio of two q-series where occurs an algebraic term. The first method mainly involves calculation over formal power series. Our next two methods interpret different steps of the previous calculation by manipulations or factorisations of the Dyck words.

Version française. Un mot de Dyck w de taille n est un mélange de n copies du mot $x\overline{x}$. Une interaction supérieure dans w est une occurrence d'un facteur $\overline{x}^k x^k$ où $k \geq 1$. On présente différentes méthodes pour énumérer les mots de Dyck selon leur taille et le nombre d'interactions supérieures. La fonction génératrice, peu usuelle, est un quotient de deux q-séries dans lesquelles apparaît un terme algébrique. La première méthode nécessite principalement des calculs sur des séries formelles. Les deux méthodes suivantes interprètent différentes étapes du calcul précédent par des manipulations ou des factorisations des mots de Dyck.

1. Introduction

A *Dyck word* w is a word over the alphabet $\{x, \overline{x}\}$ that contains as many letters x than $\overline{x}$ and such that any prefix contains at least as many letters x as letters $\overline{x}$. The *size* of w is the number of letters x in w. A *Dyck path* is a walk in the plane, starting from the origin, made up of *rises*, steps $(1, 1)$, and *falls*, steps $(1, -1)$, that remains above the horizontal axis and finishes on it. Figure 1 gives an example of a Dyck path of size 12. The Dyck path related to the Dyck word w is the walk obtained by replacing in w a letter x by a rise, and a letter $\overline{x}$ by a fall. In the rest of the paper we identify the two notions. An *upper interaction*, respectively a *lower*

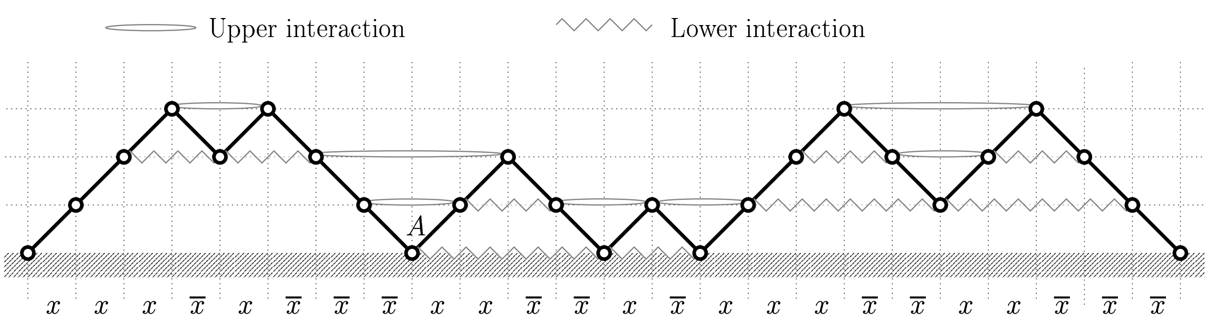


Figure 1. A Dyck path and its upper and lower interactions

Date: 30th January 2005.
Key words and phrases. Dyck paths, Catalytic parameters, Slice functional equations, Linearization of q-algebraic equations, Basic hypergeometric series, Cycle lemma, Heaps models.

interaction, in the Dyck word w is an occurrence of a factor $\overline{x}^k x^k$, respectively $x^k\overline{x}^k$, for any $k \geq 1$. The example of Figure 1 contains 7 upper interactions and 9 lower interactions. These interactions are the translation in terms of Dyck paths of physical quantities defined by physicists [7] in a model of self-interacting partially directed polymers near a surface. Lower interactions are easy to take into account in the enumeration because of the usual decomposition of Dyck paths which splits them at the second vertex on the horizontal axis, the vertex A on the Figure 1. Lower interactions are included in one of the two subwalks. Denise and Simion [6] already used this fact to enumerate Dyck paths according to the size and the number of lower interactions. By contrast, there are upper interactions above the vertex A that belong to the two subwalks. It explains why upper interactions seemingly do not satisfy an algebraic decomposition and are more difficult to take into account. This paper is a summary of a chapter of the author's PhD thesis [8] and presents three methods to find an expression for the generating function

$$A(t,u) = \sum_{\text{non-empty Dyck path } w} t^n u^k \tag{1}$$

where n is the size of w and k the number of upper interactions.

In Section 2, the first method, inspired by a work of Bousquet-Mélou and Rechnitzer [3], consists in building Dyck words by inserting a factor $x^i\overline{x}^i$ after the last letter x. This leads to a functional equation that can be solved through calculations over formal power series involving four main steps: an iteration, the kernel method [2], a division and the use of a relation between roots of a polynomial. The resulting generating function has a rather unusual form: it is a ratio of q-series,with $q = tu$, where occurs an algebraic term σ. Our aim is now to understand better these calculations over formal power series by direct manipulations or factorizations of Dyck paths.

In the next method in Section 3, we consider *Dyck paths with small valleys* that are Dyck words that avoid the factor $\overline{x}\overline{x}xx$. An *ad hoc* valuation of the *valleys*, the factors $\overline{x}x$, allows us to recover the generating function $A(t,u)$. We can recursively split these words to obtain a q-algebraic equation that is solvable after a change of unknown function. The choice of this change is crucial in the solution. A first possibility, inspired by a paper of Janse Van Rensburg [7], leads to a q-linear equation. This equation admits as solution a basic hypergeometric series computable with the algorithm proposed by Abramov, Paule and Petkovšek [1]. We present a second change of unknown function, seemingly a brother of the previous one, that, in our case, requires less calculation to conclude.

In Section 4 we consider more precisely the valuation of each valley: it is the sum of a constant term and a term that depends geometrically on the *height* of the valley. We expand this sum in each valley to consider *two-colored paths with small valleys* where the valuation of a valley is either the constant term for a *white valley*, or the "geometric" term for a *black valley*. Paths with only white valleys give a combinatorial interpretation of the algebraic term σ that comes from the kernel method in Section 2. Paths with only black valleys are in bijection with certain heaps of segments [4] thus their generating function, a ratio of q-series, is an instance of Viennot's heap inversion lemma [10]. The iteration in the calculation of Section 2 seems to correspond to the calculation of trivial heaps of segments. Moreover, we have a combinatorial interpretation of the first change of unknown function in Section 3. Finally we consider two-colored paths with small valleys where any black valley is *isolated* (occurs in a factor $xx\overline{x}x\overline{x}\overline{x}$). A partition of these paths leads to a bijection with other heaps of segments. The inversion lemma gives a ratio of q-series where appears the algebraic term σ. Moreover the relation between roots of a polynomial, used in Section 2, gives rise to a term $(q)_n(qt\sigma^2)_n$ that admits here a combinatorial interpretation. Hence, by considering the three above sets of two-colored paths, we are able to interpret combinatorially the four main steps of the calculation over formal power series. It may be possible to merge these interpretations to obtain a combinatorial interpretation of the generating function of Dyck paths counted according to the size and the number of upper interactions.

2. A catalytic parameter for a "slice" functional equation

In [3], Bousquet-Mélou and Rechnitzer use a factorization of partially directed walks. We adapt their work to the case of Dyck paths. The *length of the last descent* of a Dyck word is the number of letters $\overline{x}$ after the last letter x. We define the generating function of non-empty Dyck paths counted according to the size, the number of upper interactions and the length of the last descent by:

$$B(s) \equiv B(t,u;s) = \sum_{\text{non-empty Dyck path } w} t^n u^k s^j.$$

where n is the size of w, k the number of upper interactions and j the length of the last descent. Our aim is to compute $A(t,u) = B(t,u;1)$ but we need to know the additional parameter to write an equation linking $B(1)$, $B(s)$ and $B(uts)$. Zeilberger [11] calls this kind of parameter a *catalytic parameter*. This decomposition of Dyck paths reminds us of a Temperley-like decompositions, used by certain physicists [9].

Lemma 1. *The generating function $B(s)$ of Dyck paths counted according to the size, the number of upper interactions and the length of the last descent satisfies*

$$\begin{aligned} B(s) \quad = \quad & \frac{ts}{1-ts} + \frac{ut}{1-ut}\frac{ts\,(B(s)-B(uts))}{1-ts} \\ & + \frac{ut}{1-ut}\left(\frac{s\,(B(1)-B(s))}{1-s} - \frac{uts\,(B(1)-B(uts))}{1-uts}\right). \end{aligned} \tag{2}$$

Proof. (sketch) A *peak* is a vertex next to a rise and before a fall. We split the set of Dyck paths into three disjoint subsets illustrated on Figure 2: the paths with at most one peak, the paths where the last peak is strictly higher than the previous one, the paths where the last peak is below the previous one.

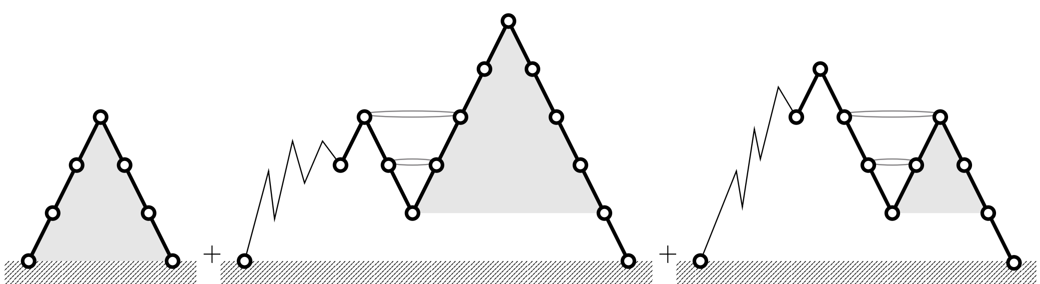


FIGURE 2. The "slice" decomposition of Dyck paths

The generating function of each subset is obtained by adding a factor $x^k\overline{x}^k$: to the empty path for the first subset or to any non-empty Dyck path otherwise. The length of the last descent is sufficient to determine the number of ways one can extend a Dyck path and the number of additional upper interactions in each case. Figure 2 gives extensions leading to each subset. The summation of these extensions over all Dyck paths gives rise to different evaluations of $B(s)$. The generating function of each subset gives one of the three terms in the right-hand side of Equation (2). $\square$

The solution of Equation (2) requires an iteration to remove $B(uts)$ and the kernel method to remove $B(s)$. As in [3], we obtain for $B(1) = A(t,u)$ a ratio of two q-series in which occurs an algebraic term.

Proposition 2. *The generating function of Dyck paths counted according to the size and the number of upper interactions is*

$$A(t,u) = B(1) = -\frac{t\displaystyle\sum_{n\geq 0}\left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n(qt\sigma^2)_n}}{q\displaystyle\sum_{n\geq 0}\left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n(qt\sigma^2)_n}\frac{1-tq^n\sigma}{(1-q^n\sigma)(1-q^{n+1}\sigma)}} \tag{3}$$

where $q = ut$, $(x)_n = \displaystyle\prod_{k=0}^{n-1}(1-q^k x)$ *and* $\sigma = \dfrac{1+t-2q-\sqrt{(1-t)(1-t-4q+4q^2)}}{2t(q-1)}$.

Proof. (sketch) Equation (2) can be rewritten as

$$B(s) = a(s) + b(s)B(1) + c(s)B(qs) + d(s)B(s). \tag{4}$$

We iterate this equation, using its evaluation at $s = q^k s$ to recursively replace the term $B(q^{k+1}s)$ by an expression in terms of $B(q^{k+2}s)$. We show that this process converges, in the sense of formal power series, toward the relation

$$(1-d(s))B(s) = \sum_{n\geq 0}\prod_{k=0}^{n-1}\frac{c(q^k s)}{1-d(q^{k+1}s)}\left(a(q^n s) + b(q^n s)B(1)\right) \tag{5}$$

where the unknown functions $B(q^k s)$, $k \geq 1$, have disappeared. The kernel method consists in choosing $s = \sigma$ such that $1 - d(\sigma) = 0$. Thus σ is one of the roots of the polynomial, deduced from $1 - d(s)$, called the *kernel*:

$$1 + \left(\frac{q-t}{1-q} - 1\right)\sigma + t\sigma^2 = 0. \tag{6}$$

We choose σ to be the unique root of this polynomial that is a formal power series in t and replace s by σ in (5). After the cancellation for $s = \sigma$ of the left side of Equation (5), it remains one division to deduce the expression of $B(1)$. The fact that the product of the two roots of the kernel is $1/t$ enables us to rewrite the products of $c(q^k\sigma)/(1-d(q^{k+1}\sigma))$ as terms where appears $(q)_n(qt\sigma^2)_n$. $\square$

3. A catalytic parameter for a q-algebraic equation

To be able to write a q-algebraic equation whose solution leads to $A(t,u)$ we consider a subset of Dyck paths, the paths with small valleys, and another catalytic parameter. The solution of the q-algebraic equation uses a change of unknown function in the spirit of the works of Brak and Prellberg [5] and Janse Van Rensburg [7].

3.1. Dyck paths with small valleys

A *valley* in a Dyck path is a vertex next to a fall and before a rise. *Dyck paths with small valleys* are Dyck paths that avoid the factor $\overline{x}\overline{x}xx$. We define an *ad hoc* valuation of such paths: there is a weight t on each rise and a weight

$$V_{val}(k) = \frac{q(1-q^{k+1}y)}{t(1-q)} \tag{7}$$

on a valley at height k. The generating function of Dyck paths with small valleys according to these weights is

$$C(y) \equiv C(t,u;y) = \sum_{\text{non-empty Dyck path with small valleys } w} t^n \prod_{k\geq 0} V_{val}(k)^{v_k} \tag{8}$$

where n is the size of w and v_k the number of valleys at height k in w.

Lemma 3. *The generating function $A(t,u)$ of Dyck paths, defined by (1), and the generating function $C(t,u;y)$ of Dyck paths with small valleys satisfy*

$$A(t,u) = C(t,u;1). \tag{9}$$

Proof. We group Dyck paths into sets of paths with the same sequence of heights of the peaks. In each set S there is a single path w_S of minimal size and we use it as the representative of the set. This path is also the single path with small valleys in S. All paths in S are obtained by "digging" the valleys of w_S, that is rewriting recursively factors $\overline{x}x$ of w_S in $\overline{x}\overline{x}xx$ while the path remains above the horizontal axis. In w_S there are as many upper interactions as valleys. Moreover, each rewriting $\overline{x}x \longrightarrow \overline{x}\overline{x}xx$ increases the size and the number of upper interactions by one. Thus the generating function of paths of S according to size and the number of upper interactions corresponds to the weight of w_S where a rise is weighted t and a valley at height k,

$$u + u^2t + \ldots + u^{k+1}t^k = \frac{q(1-q^{k+1})}{t(1-q)} = V_{val}(k)_{|y=1}.$$

We recognize here the valuation of valleys in (8) when $y=1$. The summation over all the sets S, that is over the paths with small valleys, leads to (9). □

The variable y that occurs in the weight of valleys is another example of a catalytic variable since it allows to write a q-algebraic equation for paths with small valleys:

Lemma 4. *The generating function of non-empty paths with small valleys satisfies*

$$C(y) = t + t\left(1 + q\frac{1-qy}{1-q}\right)C(qy) + q\frac{1-qy}{1-q}C(y) + \left(q\frac{1-qy}{1-q}\right)^2 C(qy)C(y). \tag{10}$$

Proof. We split a path with small valleys at the second vertex on the axis, called A in Figure 3. There are five cases due to the avoiding of the factor $\overline{x}\overline{x}xx$ especially around the vertex A.

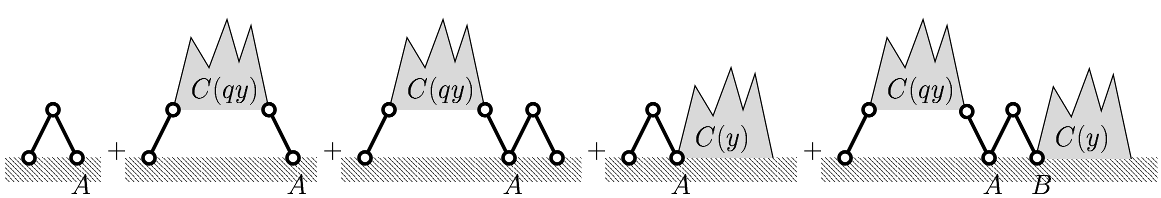


FIGURE 3. Decomposition of paths with small valleys

If w is a path with small valleys of weight $W(t,u,y)$, then $xw\overline{x}$ is a path of weight $tW(t,u,qy)$. This fact explains why we use the catalytic variable y in the weight of paths with small valleys. As an example the valuation of the fifth case is

$$tC(qy)\frac{q(1-qy)}{t(1-q)}t\frac{q(1-qy)}{t(1-q)}C(y)$$

since there are two valleys, A and B. □

3.2. Linearization

The solution of the q-algebraic equation (10) begins with a change of unknown function: we look for solutions of the form

$$C(y) = \frac{J(qy)}{\alpha J(y) + \beta(y) J(qy)} \tag{11}$$

where α is independent of y, $\beta(y)$ a rational function in y and $J(y)$ a formal power series in y such that $J(0) = 1$. A series $H(y) = \sum_{n \geq 0} h_n y^n$ is a *basic hypergeometric series* if there is a rational function $F(X)$ such that $h_{n+1}/h_n = F(q^n)$ for all $n \in \mathbb{N}$.

Proposition 5. *The unique formal power series in t that satisfies the q-algebraic Equation (10) is*

$$C(y) = \frac{(t-q)t\sigma \sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+2}{2}-1}}{(q)_n (qt\sigma^2)_n} y^n}{\sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+1}{2}}(q^n t\sigma - q - q^{2n}t^2\sigma^2)}{(q)_n (qt\sigma^2)_n} y^n} \tag{12}$$

where again $q \equiv ut$ and σ is defined as in Proposition 3.

Proof. We consider Equation (10) where $C(y)$ is replaced by its expression (11). We reduce it to a single rational function R in y, $J(y)$, $J(qy)$ and $J(q^2y)$). The numerator N of R is a linear combination of $J(y)J(qy)$, $J(qy)^2$, $J(q^2y)J(qy)$ and $J(q^2y)J(y)$. We choose $\beta(y) = -\frac{1-qy}{1-q}$ to remove the term $J(q^2y)J(y)$, thus we can factor $J(qy)$ in N. The other factor of N vanishes if and only if the following q-linear equation holds:

$$t\alpha^2 J(y) - \left(1 + \frac{(t-q)(1-qy)}{1-q}\right) \alpha J(qy) + J(q^2y) = 0. \tag{13}$$

The evaluation at $y = 0$ of Equation (13) implies that α is one of the two roots of a polynomial that is the kernel (6) in the proof of Proposition 3. We define $\alpha = 1/(t\sigma)$. We will explain later why we choose this root rather than σ. The change of unknown function defined by this analysis is

$$C(y) = \frac{t\sigma J(qy)}{J(y) - \frac{1-qy}{1-q} t\sigma J(qy)}, \tag{14}$$

and it leads to the q-linear equation (13) where $\alpha = 1/(t\sigma)$. Since (13) is of degree 1 in y and by definition $J(y) = 1 + \sum_{n \geq 1} j_n y^n$, the extraction of the coefficient of y^{n+1} in (13) gives a relation between j_n and j_{n+1} which is

$$j_{n+1} = \frac{(q-t)\sigma q^{n+1}}{(1-q)(1-q^{n+1})(1-q^{n+1}t\sigma^2)} j_n.$$

Thus $J(y)$ is the basic hypergeometric series

$$J(y) = \sum_{n \geq 0} \left(\frac{(q-t)\sigma}{1-q}\right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (qt\sigma^2)_n} y^n.$$

We plug this expression in (14) and we recognize the expression (12). □

Remark 6. We may also compute $C(y)$ using the same method as in Section 2. The additional variable y does not deeply modify the calculations. The change of unknown function (14) was actually conjectured at the sight of this formula. It also explains why we choose $\alpha = 1/(t\sigma)$ instead of σ.

Remark 7. In [7], Janse Van Rensburg suggests a change of unknown function that we generalize in

$$C(y) = \frac{\alpha J(qy) + \beta(y)J(y)}{\gamma(y)J(y)} \tag{15}$$

where α is independent of y and $\beta(y)$ and $\gamma(y)$ polynomials in y. This change also leads to a q-linear equation (L) that admits a basic hypergeometric series as solution. But the solution of (L) requires the algorithm of Abramov, Paule and Petkovšek [1] and much more calculation in our case.

4. Toward a combinatorial interpretation

The weight $V_{val}(k)$ of a small valley is the sum of a constant term $q/(t(1-q))$ and a term $-q^{k+2}y/(t(1-q))$ that depends geometrically on the height k. We expand these sums in the paths with small valleys to define the *two-colored paths (with small valleys)*: they are paths with small valleys where the valleys are either white or black. The weight of a *white valley* is $q/(t(1-q))$, the weight of a *black valley* lying at height k is $-q^{k+2}y/(t(1-q))$ and the weight of a rise remains t. By definition, the generating function of these two-colored paths is also $C(y)$.

We study three subsets of the two-colored paths. The *white paths* are the two-colored paths where all valleys are white. The *black paths* are the two-colored paths where all valleys are black. The *two-colored black-isolated paths* are the two-colored paths where all black valleys belong to a factor $xx\overline{x}x\overline{x}x$. Each of these subsets yields a combinatorial interpretation of some of the four main steps of the calculation that gives Proposition 2.

4.1. The algebraic term

The function $\sigma(t,u)$ in Proposition 2 can be written as $\sum_{n\geq 0} p_n(u)t^n$ where $p_n(u)$ are polynomials whose coefficients are nonnegative integers. Moreover $\sigma(t,1)$ is the generating function of Dyck paths according to the size. These facts suggest the existence of combinatorial interpretations of σ. The first one is the generating function of Dyck paths according to the size, counted by t, and the number of *lower* interactions, counted by u. This result was already obtained by Denise and Simion [6].

Proposition 8. *The generating function $D(t,u)$ of Dyck paths counted according to size and the number of* lower *interactions satisfies*

$$D(t,u) = 1 + \frac{ut}{1-ut}D(t,u) + t\left(D(t,u) - \frac{1}{1-ut}\right)D(t,u) \tag{16}$$

thus $D(t,u) = \sigma$.

We can group Dyck paths into sets with the same sequence of heights of the valleys. The smallest path of each set is a path that avoid $xx\overline{x}\overline{x}$, that is *a path with small peaks.* As for valleys and upper interactions, we can define $D(t,u)$ by a summation over paths with small peaks where a peak is weighted $q/(t(1-q))$ and a rise t.

The white paths correspond to the case $y=0$ for Dyck paths with small valleys. For $y=0$, the q-algebraic equation (10) becomes an algebraic equation whose solution is the generating function of the white paths. This generating function is only "almost" equal to $\sigma(t,u)$. A *double-peak* in a Dyck path is a factor $xx\overline{x}\overline{x}$. To recover exactly σ, we consider the *white paths starting with a double-peak.*

Proposition 9. *There exists a bijection f between paths with small peaks and white paths starting with a double-peak such that the weight of $f(w)$ is exactly the weight of w multiplied by t^2. The generating function of the white paths starting with a double-peak is $t^2\sigma$.*

Proof. (sketch) Let w be a non-empty path with small peaks. Let A be the first vertex of w, B the first vertex of maximal height, C the last vertex and D the vertex preceding B. Using these vertices, we factor w in $w_{AD}xw_{BC}$, where w_{PQ} denotes the subpath of w between the vertices P and Q. We denote by g the morphism on the words over the alphabet $\{x, \overline{x}\}$ defined by $g(x) = \overline{x}$ and $g(\overline{x}) = x$. We define $f(w) = xx\overline{x}\overline{x}g(v_{BC})\overline{x}g(v_{AD})$.

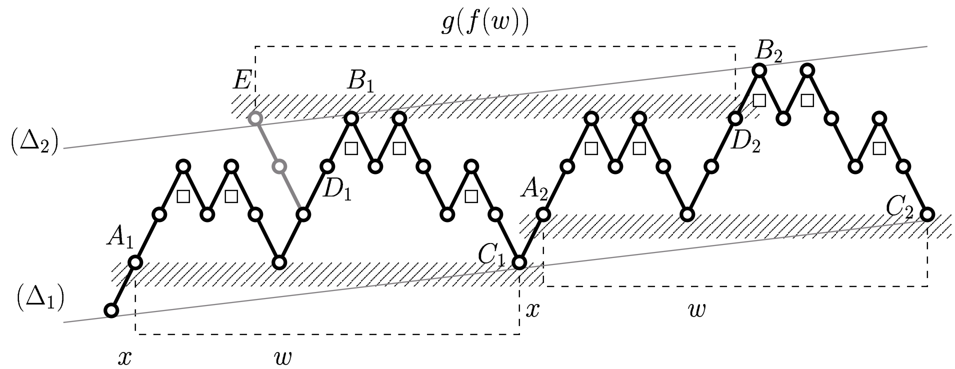


FIGURE 4. A bijection between two interpretations of σ

Figure 4 gives on an example a geometrical definition of this bijection f. It is supposed to convince us that f satisfies all claimed properties. We consider the path $p = xwxw$ and the line (Δ_1) below this path containing exactly three vertices of the path p on it. Then we consider the line (Δ_2) above this path, parallel to (Δ_1), with exactly two vertices B_1 and B_2 of p. Between the vertices E and D_2, we recognize the image of $f(w)$ by g. A path with small peaks w becomes a path with small valleys and the weights $q/(t(1-q))$ remain on the same vertices. $\square$

Remark 10. It is also possible to compute directly the generating function of white paths starting with a double-peak using an algebraic decomposition of these paths.

4.2. THE RATIO OF BASIC HYPERGEOMETRIC SERIES

To enumerate black paths, we use a bijection with heaps of marked segments with an *ad hoc* weight. A *marked segment* $s = ([l(s), r(s)], m(s))$ is defined by an interval $[l(s), r(s)] \subseteq \mathbb{N}$ of $r(s) - l(s) + 1$ *elements* and a subset $m(s) \subseteq [l(s) + 1, r(s) - 1]$ of marked elements. If the marked segment is a singleton $([r(s), r(s)], \emptyset)$ then its weight $V_{seg}(s)$ is $-\frac{q^{r(s)+2}y}{(1-q)}$. Otherwise $l(s) < r(s)$ and the weight of the marked segment $s = ([l(s), r(s)], m(s))$ is

$$V_{seg}(s) = \left(-\frac{q^{l(s)+2}y}{1-q}\right)^2 t^{r(s)-l(s)} \prod_{k\in m(s)} \left(-\frac{q^{k+2}y}{1-q}\right).$$

A *heap of (marked) segments* of size n is a set $\{(s_i, h_i)\}_{i=1\ldots n}$ where s_i is a marked segment and $h_i \in \mathbb{N}$ is its height, such that

- if $h_i = h_j$ then $[l(s_i), r(s_i)] \cap [l(s_j), r(s_j)] = \emptyset$: two segments can not overlap.

- if $h_i > 0$ then there exists $(s_j, h_i - 1)$ such that $[l(s_i), r(s_i)] \cap [l(s_j), r(s_j)] \neq \emptyset$: a segment that is not on the floor lay on a segment just below.

A heap of segments is *a half-pyramid of (marked) segments* if there is at most one segment s_i on the floor, that is $h_i = 0$, and moreover $l(s_i) = 0$. On the right of Figure 5 there is a half-pyramid of marked segments. The generating function of half-pyramids of segments is defined by

$$F(y) = \sum_{\text{half-pyramid } h} \left(\prod_{s \text{ segment of } h} V_{seg}(s) \right).$$

As in the case of white paths, we consider *black paths starting with a double-peak.*

Proposition 11. *There exists a bijection h between black paths starting with a double-peak and half-pyramids of marked segments. This bijection implies that $t^2F(y)$ is the generating function of black paths starting with a double-peak.*

Proof. (sketch)

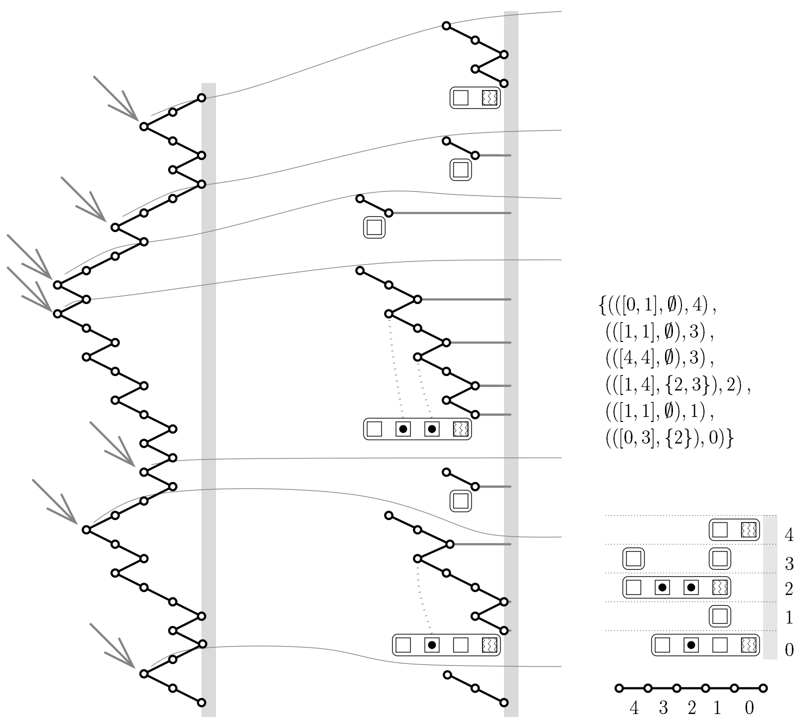


FIGURE 5. Black paths and half-pyramids of marked segments

The bijection is a variant of a previous one between Dyck paths and half-pyramids of segments [4]. We discuss here its specificities. We use the example of Figure 5 to define how a black path w is mapped to a half-pyramid $h(w)$. A *cutting peak* in w, pointed by a grey arrow on Figure 5, is a peak that is not before a factor $\overline{x}xx$. Two consecutive cutting peaks are the endpoints of a *block* b where $l(b)$ is the height of the vertex before the first rise, $r(b) + 1$ is the height of the second cutting peak and $m(b)$ is the set of all heights of valleys in b. The block b is mapped to the marked segment $s = ([l(b), r(b)], m(b) \backslash \{l(b)\})$. Given a sequence of marked segments, we "let them fall on the floor" to obtain an heap of segments $h(w)$. Moreover, since the black path starts with a double-peak, the vertex before the first rise of the first block b is at height 0, leading to a segment $s = ([0, r(s)], m(s))$. By definition of cutting peaks, all the other segments are carried by a previous segment in the sequence. Thus $h(w)$ is a half-pyramid of marked segments. We claim without proving it here that h is a bijection. Since the weight of the marked segment s was defined to be the weight of rises and valleys in the block b, the weight of $h(w)$ is almost the weight of w: only the weight of the two first rises in w is not take into account in $h(w)$. Thus the generating function of black paths starting with a double-peak is $t^2F(y)$. □

A *trivial heap* is a heap $\{(s_i, h_i)\}$ where all segments are on the floor, that is $h_i = 0$ for all i. The *alternating generating function of trivial heaps* is defined by

$$T(y) = \sum_{\text{trivial heap } h} \left(\prod_{s \text{ segment of } h} (-1)V_{seg}(s) \right).$$

Lemma 12. *The alternating generating function of trivial heaps $T(y)$ satisfies the q-linear equation*

$$\begin{aligned} T(y) \quad &= \quad T(qy) + \frac{q^2 y}{1-q} T(qy) - \frac{tqy}{1-q} \left(T(qy) - T(q^2 y) \right) \\ &\quad + t/q^2 \left(T(qy) - T(q^2 y) - \frac{q^3 y}{1-q} T(q^2 y) \right). \end{aligned} \tag{17}$$

and

$$T(y) = \sum_{n \geq 0} \left(\frac{q-t}{1-q} \right)^n \frac{q^{\binom{n+1}{2}}}{(q)_n (t/q)_n} y^n.$$

Proof. (sketch) The weight of a segment $s = ([l(s), r(s)], m(s))$ can be distributed to the elements of $[l(s), r(s)]$ as follows:

$$V_{seg}(s) = \left(-\frac{q^{r(s)+2} y}{1-q} \right) \left(-\frac{q^{l(s)+1} y}{1-q} \right) \prod_{k \in m(s)} \left(-\frac{q^{k+1} y}{1-q} \right) \prod_{k \in [l(s)+1, r(s)-1] \setminus m(s)} \frac{t}{q}.$$

Thus there are three kinds of element in $[l(s), r(s)]$: the *first element* $r(s)$, weighted $-\frac{q^{r(s)+2} y}{1-q}$, the *heavy elements* $k \in m(s) \cup \{l(s)\}$, weighted $-\frac{q^{k+1} y}{1-q}$, and the *light elements* $k \in [l(s)+1, r(s)-1] \setminus m(s)$ weighted t/q. That distribution corresponds to the elevation of the valuation of the first black valley in front of the last rise of the block. A singleton contains only its first element. In the other segments $l(s)$ is an heavy element, $r(s)$ the first element and the elements of $[l(s)+1, r(s)-1]$ are either heavy or light. We split the set of trivial heaps S into four disjoint subsets: the set of trivial heaps S_1 where 0 is not an element of a segment, the set of trivial heaps S_2 where there is a singleton $([0,0], \emptyset)$, the set of trivial heaps S_3 where there is a segment $([0, r(s)], m(s))$ whose 1 is a light element and the set of trivial heaps S_4 where there is a segment $([0, r(s)], m(s))$ whose 1 is either an heavy element or the first element. The alternating generating function of each of these subsets gives a term in the right side of Equation (17). The *translated segment* of a segment s is $t = ([l(s)+1, r(s)+1], \{k+1 | k \in m(s)\})$. The weight $V_{seg}(t)$ is $V_{seg}(s)$ where qy has been substituted to y. The subset S_1 is exactly the set obtained by translating all heaps, thus the alternating generating function of this subset is $T(qy)$. Since we are in an abstract, it is left to the reader to note that the alternating generating functions of the sets are $\frac{q^2 y}{1-q} T(qy)$ for S_2, $t/q^2 \left(T(qy) - (1 + \frac{q^3 y}{1-q}) T(q^2 y) \right)$ for S_3 and $-\frac{tqy}{1-q}(T(qy) - T(q^2 y))$ for S_4.

Equation (17) is of degree 1 in y. Moreover $T(y) = 1$ since only the empty trivial heap is not weighted by a factor y. By a solution similar to that of Equation (13) we compute the basic hypergeometric series $T(y)$. □

The heap inversion Lemma of Viennot [10] states that the generating function of heaps where all segments on the floor belong to a set of segments A is $H_A(y)/H(y)$ where $H(y)$ is the alternating generating function of trivial heaps and $H_A(y)$ the alternating generating function of trivial heaps where no segment belongs to A. In the case of half-pyramids, A is the set of segments $s = ([0, r(s)], m(s))$ and $H_A(y) = T(qy)$. Thus the generating function of half-pyramids is $T(qy)/T(y)$. This fact, Proposition 11 and Lemma 12 lead to the generating function of black paths starting with a double peak. In the next proposition we consider all non-empty black paths.

Proposition 13. *The generating function of non-empty black paths satisfies*

$$G(y) = \frac{(1-q)^2}{tq^4y^2}\left(t^2\frac{T(qy)}{T(y)} - t^2 + \frac{t^2q^2y}{1-q}\right) \tag{18}$$

and

$$G(y) = \frac{t\displaystyle\sum_{n\geq 0}\frac{(q-t)^{n+1}q^{\binom{n+2}{2}}y^n}{(1-q)^n(q)_n(t/q)_{n+2}}}{q^2\displaystyle\sum_{n\geq 0}\frac{(q-t)^n q^{\binom{n+1}{2}}y^n}{(1-q)^n(q)_n(t/q)_n}} \tag{19}$$

Proof. A black path starting with a double-peak is either the path $xx\overline{x}\overline{x}$, or the path $xx\overline{x}\overline{x}x\overline{x}$ or it starts with the factor $xx\overline{x}\overline{x}x\overline{x}$ and followed by any non-empty black path. On the other hand Proposition 11 states that the generating function of black paths starting with a double-peak is $t^2F(y)$. This leads to the equation

$$t^2F(y) = t^2 - t^2\frac{q^2y}{1-q} + t\frac{q^4y^2}{(1-q)^2}G(y). \tag{20}$$

Using the fact that $F(y) = T(qy)/T(y)$, standard calculations lead to (18) and (19). □

Remark 14. Equation (18) is reminiscent of the second change of unknown function (15) used by Janse Van Rensburg to solve a similar question [7]. Moreover the combinatorial interpretation leading to Equation (20) explains why $\gamma(y)$ is the denominator.

4.3. A partial mixing of the two interpretations

Proposition 15. *The generating function of two-colored black-isolated paths starting with a double-peak is*

$$t^2\sigma\frac{\displaystyle\sum_{n\geq 0}\frac{q^{6n}t^n\sigma^n}{(q-1)^{3n}(q)_n(qt\sigma^2)_n}y^n}{\displaystyle\sum_{n\geq 0}\frac{q^{5n}t^n\sigma^n}{(q-1)^{3n}(q)_n(qt\sigma^2)_n}y^n} \tag{21}$$

where $q \equiv ut$ and $\sigma$is defined as in Propositions 2 and 5.

Proof. (sketch) We group two-colored black-isolated paths starting with a double-peak into sets of paths that admit the same sequence of heights of black valleys *and* the same sequence of minimal heights between two black valleys. Let S be one of these sets. There is a single smallest path w in S. In w there is a factor

$x\overline{x}^{2+i}x\overline{x}x^{j+2}\overline{x}$ between consecutive black valleys. Figure 6 gives a example of a path w. The generating function of paths in S is obtained by inserting white paths counted by σ before double rise or double fall in w except for a double rise in a factor $x\overline{x}xx$. The path w is mapped to a half-pyramid of segments where the segment $s = [l(s), r(s)]$ is weighted $\frac{q^5 t\sigma}{(q-1)^3} q^{r(s)} (t\sigma)^{r(s)-l(s)}$. Only a factor $t^2\sigma$ at the start of the path is forgotten in this map. On Figure 6, w is mapped to the half-pyramid defined by $s_1, \ldots, s_4$. We compute the alternating generating function of trivial heaps made up of these segments. The heap inversion lemma leads to the generating function (21) where appears the terms $(q)_n(qt\sigma^2)_n$. $\square$

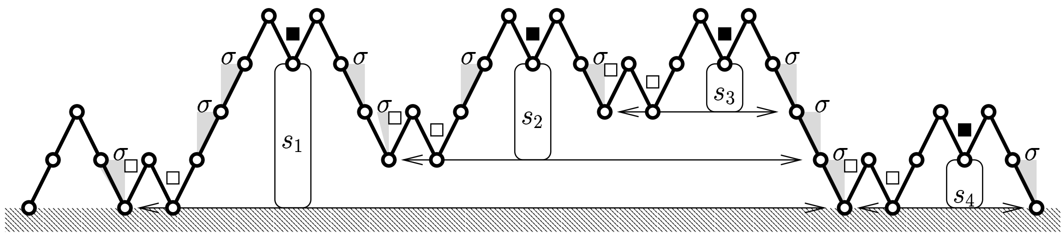


FIGURE 6. A representative of a set of two-colored black-isolated paths

The author thanks Mireille Bousquet-Mélou for fruitful discussions and comments.

References

[1] S. A. Abramov, P. Paule, and M. Petkovšek. q-Hypergeometric solutions of q-difference equations. *Discrete Math.*, 180(1-3):3–22, 1998.

[2] M. Bousquet-Mélou and M. Petkovšek. Walks confined in a quadrant are not always D-finite. *Theoret. Comput. Sci.*, 307(2):257–276, 2003.

[3] M. Bousquet-Mélou and A. Rechnitzer. The site-perimeter of bargraphs. *Adv. in Appl. Math.*, 31(1):86–112, 2003.

[4] M. Bousquet-Mélou and X. G. Viennot. Empilements de segments et q-énumération de polyominos convexes dirigés. *J. Combin. Theory Ser. A*, 60(2):196–224, 1992.

[5] R. Brak and T. Prellberg. Critical exponents from nonlinear functional equations for partially directed cluster models. *J. of Stat. Phys.*, 78(3-4):701–730, 1995.

[6] A. Denise and R. Simion. Two combinatorial statistics on Dyck paths. *Discrete Math.*, 137(1-3):155–176, 1995.

[7] E. J. Janse van Rensburg. Interacting columns: generating functions and scaling exponents. *J. Phys. A*, 33(42):7541–7554, 2000.

[8] Y. Le Borgne. *Variations combinatoires sur des classes d'objets comptées par la suite de Catalan.* PhD thesis, Université Bordeaux 1, 2004.

[9] H. N. V. Temperley. Combinatorial problems suggested by the statistical mechanics of domains and of rubber-like molecules. *Phys. Rev. (2)*, 103:1–16, 1956.

[10] G. X. Viennot. Heaps of pieces. I. Basic definitions and combinatorial lemmas. In *Combinatoire énumérative*, volume 1234 of *Lecture Notes in Math.*, pages 321–350. Springer, Berlin, 1986.

[11] D. Zeilberger. The umbral transfer-matrix method. I. Foundations. *J. Combin. Theory Ser. A*, 91(1-2):451–463, 2000. In memory of Gian-Carlo Rota.

LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France

E-mail address: `borgne@labri.fr`

A COMBINATORIAL MODEL FOR CRYSTALS OF KAC-MOODY ALGEBRAS (EXTENDED ABSTRACT)

CRISTIAN LENART AND ALEXANDER POSTNIKOV

Abstract. We present a simple combinatorial model for the characters of the irreducible representations of Kac-Moody algebras. This model can be viewed as a discrete counterpart to the Littelmann path model. We describe crystal graphs and give a Littlewood-Richardson rule for decomposing tensor products of irreducible representations.

1. Introduction

We have recently given a combinatorial model for the characters of the irreducible representations of a complex semisimple Lie group G, and for the Demazure characters [LP1]. This model was defined in the context of the equivariant K-theory of the generalized flag variety G/B. Our character formulas were derived from a Chevalley-type formula in $K_T(G/B)$. Our model was based on enumerating certain saturated chains in the *Bruhat order* on the corresponding *Weyl group* W. This enumeration is determined by an *alcove path*, which is a sequence of adjacent alcoves for the *affine Weyl group* W_{aff} of the Langland's dual group $G^\vee$. Alcove paths correspond to decompositions of elements in the affine Weyl group. Our Chevalley-type formula was formulated in terms of a certain *R-matrix*, that is, in terms of a collection of operators satisfying the *Yang-Baxter equation*. This setup allowed us to easily explain the independence of our formulas from the choice of an alcove path.

There are other models for Chevalley-type formulas in $K_T(G/B)$ and for the irreducible characters of G. Most notably, there is the *Littelmann path model*. Littelmann [Li1, Li2, Li3] showed that the characters can be described by counting certain continuous paths in $\mathfrak{h}^*_{\mathbb{R}}$. These paths are constructed recursively, by starting with an initial one, and by applying certain *root operators*. By making specific choices for the initial path, one can obtain special cases which have more explicit descriptions. For instance, a straight line initial path leads to the *Lakshmibai-Seshadri paths* (LS paths). These were introduced before Littelmann's work, in the context of *standard monomial theory* [LS]. They have a nonrecursive description as weighted chains in the Bruhat order on the quotient W/W_λ of the corresponding Weyl group W modulo the stabilizer W_λ of the weight λ; therefore, we will use the term *LS chains* when referring to this description. LS paths were used by Pittie and Ram [PR] to derive a K_T-Chevalley formula. Recently, Gaussent and Littelmann [GL], motivated by the study of Mirković-Vilonen cycles, defined another combinatorial model for the irreducible characters of a complex semisimple Lie group. This model is based on *LS galleries*, which are certain sequences of faces of alcoves for the corresponding affine Weyl group. For each LS gallery, there is an associated Littelmann path, and a saturated chain in the Bruhat order on W/W_λ. In [LP1], we explained the way in which our construction, which was developed independently of LS galleries, is related (although not quite equivalent) to the latter in the case of regular weights.

In this paper, we develop the combinatorial model in [LP1] purely in the context of representation theory, and extend it to complex symmetrizable Kac-Moody algebras. Instead of alcove paths (that make sense only in finite types) we now use λ-*chains*, which are chains of roots satisfying a certain *interlacing property*. Note that Littelmann paths and, in particular, LS paths were also defined in this more general context, but LS galleries were not. In fact, we show that LS paths are a certain limiting case of a special case of our model. The latter can be viewed as a discrete counterpart to the Littelmann path model. We define root operators in our model, and study their properties. This allows us to show that our model

Cristian Lenart was supported by NSF grant DMS-0403029 and by SUNY Albany Faculty Research Award 1039703.
Alexander Postnikov was supported by NSF grant DMS-0201494 and by Alfred P. Sloan Foundation research fellowship.

satisfies the axioms of an *admissible system* of Stembridge [Ste]. Thus, we easily derive character formulas, a Littlewood-Richardson rule for decomposing tensor products of irreducible representations, as well as a branching rule. The approach via admissible systems was already applied to LS chains in [Ste, Section 8]. Compared to the proofs in [GL, Li2, Li3], Stembridge's approach has the advantage of making a part of the proof independent of a particular model for Weyl characters, by using a system of axioms for such models.

Our model has several advantages over the Littelmann path model and its specializations mentioned above. First of all, our formulas are equally simple for all weights, regular and nonregular. Note that the (nonrecursive) construction of LS chains and LS galleries usually involves certain choices that add to their computational complexity. Also, it is harder to work with sequences of lower dimensional faces of alcoves (in the case of LS galleries) than with sequences of roots (in our model). We refer to [LP1] for a discussion showing that the computational complexity of our model is significantly smaller than the one of Littelmann paths (constructed recursively via root operators). Our definition of root operators resembles the one for LS paths, which is simpler than the general definition of root operators for Littelmann paths. We think that our model is easier to work with in explicit computations because, being based on certain chains of roots, it has a stronger combinatorial nature than Littelmann paths and, in particular, LS chains. Indeed, even for LS chains, we do need their description as piecewise-linear paths in order to define root operators.

We believe that the properties of our model discussed in this paper represent just a small fraction of a rich combinatorial structure yet to be explored. We will investigate it in a forthcoming paper [LP2].

Acknowledgments. We are grateful to John Stembridge for the explanation of his work, and to V. Lakshmibai and Peter Magyar for helpful conversations.

2. Preliminaries

In this section, we briefly recall the general setup for complex symmetrizable Kac-Moody algebras and their representations. We refer to [Kac, Ku] for more details.

Let V be a finite-dimensional real vector space with a nondegenerate symmetric bilinear form $\langle\cdot\,,\cdot\rangle$, and let $\Phi\subset V$ be a crystallographic root system of rank r with simple roots $\{\alpha_1,\ldots,\alpha_r\}$. By this, we mean that Φ is the set of real roots of some complex symmetrizable Kac-Moody algebra. The finite root systems of this type are the root systems of semisimple Lie algebras.

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.* For each root α, there is a reflection $s_\alpha\,:\,V\to V$ defined by $s_\alpha(\lambda)=\lambda-\langle\lambda,\alpha^\vee\rangle\,\alpha$. More generally, for any integer k, one can consider the affine hyperplane $H_{\alpha,k}:=\{\lambda\in V\mid\langle\lambda,\alpha^\vee\rangle=k\}$, and let $s_{\alpha,k}$ denote the corresponding reflection, that is, $s_{\alpha,k}\,:\,\lambda\mapsto s_\alpha(\lambda)+k\alpha$.

The *Weyl group* W is the subgroup of $GL(V)$ generated by the reflections s_α for $\alpha\in\Phi$. In fact, the Weyl group W is a Coxeter group, which is generated by the *simple reflections* $s_1,\ldots,s_r$ corresponding to the simple roots $s_p:=s_{\alpha_p}$, subject to the *Coxeter relations*: $(s_p)^2=1$ and $(s_ps_q)^{m_{pq}}=1$; here the relations of the second type correspond to the distinct p, q in $\{1,\ldots,r\}$ for which the dihedral subgroup generated by s_p and s_q is finite, in which case m_{pq} is half the order of this subgroup. The Weyl group is finite if and only if Φ is finite.

An expression of a Weyl group element w as a product of generators $w=s_{p_1}\cdots s_{p_l}$ which has minimal length is called a *reduced decomposition* for w; its length $\ell(w)=l$ is called the *length* of w. 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.

Let us note that Φ can be characterized by the following axioms:

(R1) $\{\alpha_1,\ldots,\alpha_r\}$ is a linearly independent set.
(R2) $\langle\alpha_p,\alpha_p\rangle>0$ for all $p=1,\ldots,r$.

(R3) $\langle \alpha_p, \alpha_q^\vee \rangle \in \mathbb{Z}_{\leq 0}$ for all distinct simple roots α_p and α_q.
(R4) $\Phi = \bigcup_{p=1}^{r} W\alpha_p$.

Let $\Phi^+ \subset \Phi$ be the set of positive roots, that is, the set of roots in the nonnegative linear span of the simple roots. 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 $|\alpha| := \mathrm{sgn}(\alpha)\alpha$.

The lattice of (integral) *weights* Λ is given by $\Lambda := \{\lambda \in V \mid \langle \lambda, \alpha^\vee \rangle \in \mathbb{Z} \text{ for any } \alpha \in \Phi\}$. The set Λ^+ of *dominant weights* is given by $\Lambda^+ := \{\lambda \in \Lambda \mid \langle \lambda, \alpha^\vee \rangle \geq 0 \text{ for any } \alpha \in \Phi^+\}$. If we replace the weak inequalities above with strict ones, we obtain the strongly dominant weights. It is known that every W-orbit in V has at most one dominant member. The *fundamental weights* $\omega_1, \ldots, \omega_r$ are defined by $\langle \omega_p, \alpha_q^\vee \rangle = \delta_{pq}$.

We now define a ring R that contains the characters of all integrable highest weight modules for the corresponding Kac-Moody algebra. In the finite case, one may simply take R to be the group ring of Λ, but in general more care is required.

First, we choose a height function $\mathrm{ht} : V \to \mathbb{R}$, that is, a linear map assigning the value 1 to all simple roots. Second, for each $\lambda \in \Lambda$, let e^λ denote a formal exponential subject to the rules $e^\mu \cdot e^\nu = e^{\mu+\nu}$ for all $\mu, \nu \in \Lambda$. We now define the ring R to consist of all formal sums $\sum_{\lambda\in\Lambda} c_\lambda e^\lambda$ with $c_\lambda \in \mathbb{Z}$ satisfying the condition that there are only finitely many weights λ with $\mathrm{ht}(\lambda) > h$ and $c_\lambda \neq 0$, for all $h \in \mathbb{R}$.

For each $\lambda \in \Lambda^+$ with a finite W-stabilizer, we define

$$\Delta(\lambda) := \sum_{w\in W} \mathrm{sgn}(w) e^{w(\lambda)}\,,$$

where $\mathrm{sgn}(w) = (-1)^{\ell(w)}$. It is not hard to check that $\Delta(\lambda)$ is a well-defined member of R. Since the scalar product is nondegenerate, we may select $\rho \in \Lambda^+$ so that $\langle \rho, \alpha_p^\vee \rangle = 1$ for all $p = 1, \ldots, r$. One can verify that $\Delta(\rho)$ is invertible in R. This given, for each $\lambda \in \Lambda^+$ we define

$$\chi(\lambda) := \frac{\Delta(\lambda+\rho)}{\Delta(\rho)} = \frac{\sum_{w\in W} \mathrm{sgn}(w) e^{w(\lambda+\rho)-\rho}}{\sum_{w\in W} \mathrm{sgn}(w) e^{w(\rho)-\rho}} \in R\,.$$

It is easy to show that $w(\rho) - \rho$, and hence $\chi(\lambda)$, do not depend on the choice of ρ. By the Kac-Weyl character formula [Kac], these are the characters of the irreducible highest weight modules for the corresponding Kac-Moody algebra.

3. Crystals

This section closely follows [Ste, Section 2]. We refer to this paper for more details.

Definition 3.1. (cf. [Ste]). A *crystal* is a 4-tuple $(X, \mu, \delta, \{F_1, \ldots, F_r\})$ satisfying Axioms (A1)-(A3) below, where

- X is a set whose elements are called objects;
- μ and δ are maps $X \to \Lambda$;
- F_p are bijections between two subsets of X.

For each $x \in X$, we call $\mu(x)$, $\delta(x)$, and $\varepsilon(x) := \mu(x) - \delta(x)$ the *weight*, *depth*, and *rise* of x.

(A1) $\delta(x) \in -\Lambda^+$, $\varepsilon(x) \in \Lambda^+$.

We define the depth and rise in the direction α_p by $\delta(x,p) := \langle \delta(x), \alpha_p^\vee \rangle$ and $\varepsilon(x,p) := \langle \varepsilon(x), \alpha_p^\vee \rangle$. In fact, we will develop the whole theory in terms of $\delta(x,p)$ and $\varepsilon(x,p)$ rather than $\delta(x)$ and $\varepsilon(x)$.

(A2) F_p is a bijection from $\{x \in X \mid \varepsilon(x,p) > 0\}$ to $\{x \in X \mid \delta(x,p) < 0\}$.

(A3) $\mu(F_p(x)) = \mu(x) - \alpha_p$, $\delta(F_p(x),p) = \delta(x,p) - 1$.

Hence, we also have $\varepsilon(F_p(x), p) = \varepsilon(x, p) - 1$. We let $E_p := F_p^{-1}$ denote the inverse map. The maps E_p and F_p act as raising and lowering operators that provide a partition of the objects into α_p-strings that are closed under the action of E_p and F_p. For example, the *α_p-string through x* is (by definition) $F_p^{\varepsilon}(x), \ldots, F_p(x), x, E_p(x), \ldots, E_p^{-\delta}(x)$, where $\delta = \delta(x, p)$ and $\varepsilon = \varepsilon(x, p)$. Let us now present some additional axioms.

(A4) For all real numbers h, there are only finitely many objects x such that $\mathrm{ht}(\mu(x)) > h$.

Axiom (A4) implies that the generating series $G_X := \sum_{x \in X} e^{\mu(x)}$ is a well-defined member of R.

We define a partial order on X by $x \preceq_p y$ if $x = F_p^k(y)$ for some $k \geq 0$. We call an object of X *maximal* if it is maximal with respect to all partial orders $\preceq_p$, for $p = 1, \ldots, r$.

(A5) X has a unique maximal object.

Stembridge defined *admissible systems* as crystals satisfying Axiom (A4) and an extra axiom, which is related to the existence of a certain map $(x, p) \mapsto t(x, p)$ on pairs (x, p) with $\delta(x, p) < 0$. This map is called *coherent timing pattern*, and is used to construct a certain sign-reversing involution allowing one to cancel the negative terms in the Kac-Weyl character formula. We call an admissible system a *semiperfect crystal* if it satisfies Axiom (A5).

Given $P \subseteq \{1, \ldots, r\}$, let Φ_P denote the root subsystem of Φ with simple roots $\{\alpha_p \mid p \in P\}$. Following [Ste], we let $W_P \subseteq W$, $\Lambda_P \supseteq \Lambda$, and R_P denote the corresponding Weyl group, weight lattice, and character ring. Given $\lambda \in \Lambda_P^+$, we let $\chi(\lambda; P) \in R_P$ denote the Weyl character (relative to Φ_P) corresponding to λ.

Finally, note that one can define on X the structure of a directed colored graph by constructing arrows $x \to y$ colored p for each $F_p(x) = y$.

Definition 3.2. A crystal $(X, \mu, \delta, \{F_1, \ldots, F_r\})$ is called a *perfect crystal* if the associated directed colored graph is isomorphic to the *crystal graph* of an irreducible representation of a quantum group.

4. λ-Chains of Roots

Fix a dominant weight λ. Throughout this paper, we will use the term "sequence" for any map $i \mapsto a_i$ from a totally ordered set I to some other set; we will use the notation $\{a_i\}_{i \in I}$.

Definition 4.1. A *λ-chain (of roots)* is a sequence of positive roots $\{\beta_i\}_{i \in I}$ indexed by the elements of a totally ordered set I, which satisfies the following conditions:

(1) the number of occurrences of any positive root α is $\langle \lambda, \alpha^\vee \rangle$;
(2) for each triple of positive roots (α, β, γ) with $\gamma^\vee = \alpha^\vee + \beta^\vee$, the finite sequence $\{\beta_j\}_{j \in J}$, where $J := \{j \in I \mid \beta_j \in \{\alpha, \beta, \gamma\}\}$ has the induced total order, is a concatenation of pairs (α, γ) and (β, γ) (in any order).

Note that finding a λ-chain amounts to defining a total order on the set

$$I := \{(\alpha, k) \mid \alpha \in \Phi^+,\ 0 \leq k < \langle \lambda, \alpha^\vee \rangle\} \tag{4.1}$$

such that the second condition above holds, where $\beta_i = \alpha$ for any $i = (\alpha, k)$ in I. One particular example of such an order can be constructed as follows. Fix a total order on the set of simple roots $\alpha_1 < \alpha_2 < \ldots < \alpha_r$. For each $i = (\alpha, k)$ in I, let $\alpha^\vee = c_1 \alpha_1^\vee + \ldots + c_r \alpha_r^\vee$, and define the vector

$$v_i := \frac{1}{\langle \lambda, \alpha^\vee \rangle}(k, c_1, \ldots, c_r) \tag{4.2}$$

in $\mathbb{Q}^{r+1}$. The map $i \mapsto v_i$ is easily seen to be injective.

Proposition 4.2. *Consider the total order on the set I in* (4.1) *defined by $i < j$ iff $v_i < v_j$ in the lexicographic order on $\mathbb{Q}^{r+1}$. The sequence $\{\beta_i\}_{i \in I}$ given by $\beta_i = \alpha$ for $i = (\alpha, k)$ is a λ-chain.*

For the rest of our construction (Sections 5–7), let us fix a dominant integral weight λ and fix an arbitrary λ-chain $\{\beta_i\}_{i\in I}$. We will use the notation r_i for the reflection s_{β_i}.

5. Folding Chains of Roots

We start by associating to our fixed λ-chain the closely related object $\Gamma(\emptyset) := (\{(\beta_i, \beta_i)\}_{i\in I}, \rho)$, where ρ is a fixed dominant weight satisfying $\langle \rho, \alpha_p^\vee\rangle = 1$ for all $p = 1, \ldots, r$. Here, as well as throughout this article, we let ∞ be greater than all elements in I. We use operators called *folding operators* to construct from $\Gamma(\emptyset)$ new objects of the form $\Gamma = (\{(\gamma_i, \gamma_i')\}_{i\in I}, \gamma_\infty)$; here (γ_i, γ_i') are pairs of roots with $\gamma_i' = \pm\gamma_i$, any given root appears only finitely many times in Γ, and γ_∞ is in the W-orbit of ρ. More precisely, given Γ as above and i in I, we let $t_i := s_{\gamma_i}$ and we define

$$\phi_i(\Gamma) := (\{(\delta_j, \delta_j')\}_{j\in I}, t_i(\gamma_\infty)), \qquad \text{where} \quad (\delta_j, \delta_j') := \begin{cases} (\gamma_j, \gamma_j') & \text{if } j < i \\ (\gamma_j, t_i(\gamma_j')) & \text{if } j = i \\ (t_i(\gamma_j), t_i(\gamma_j')) & \text{if } j > i\,. \end{cases}$$

Let us now consider the set of all Γ that are obtained from $\Gamma(\emptyset)$ by applying folding operators; we call these objects the *foldings* of $\Gamma(\emptyset)$. Clearly, ϕ_i is an involution on the set of foldings of $\Gamma(\emptyset)$. In order to describe this set, let us note that the folding operators commute. This means that every folding Γ of $\Gamma(\emptyset)$ is determined by the set $J := \{j \mid \gamma_j' = -\gamma_j\}$. More precisely, if $J = \{j_1 < j_2 < \ldots < j_s\}$, then $\Gamma = \phi_{j_1} \ldots \phi_{j_s}(\Gamma(\emptyset))$. We call the elements of J the *folding positions* of Γ, and write $\Gamma = \Gamma(J)$.

Throughout this paper, we will use J and $\Gamma = \Gamma(J)$ interchangeably. For instance, according to the above discussion, we have $\phi_i(\Gamma(J)) = \Gamma(J \triangle \{i\})$, where $\triangle$ denotes the symmetric difference of sets. Hence, it makes sense to define the folding operator ϕ_i on J (compatibly with the action of ϕ_i on $\Gamma(J)$) by $\phi_i : J \mapsto J\triangle\{i\}$.

Remark 5.1. Although a folding Γ of $\Gamma(\emptyset)$ is an infinite sequence if the root system is infinite, we are, in fact, always working with finite objects. Indeed, we are examining Γ by considering only one root at a time.

Given a folding Γ of $\Gamma(\emptyset)$, we associate to each pair of roots (or the corresponding index i in I) an integer l_i, which we call *level*; the sequence $L = L(\Gamma) := \{l_i\}_{i\in I}$ will be called the *level sequence* of Γ. The definition is as follows:

$$l_i := \varepsilon + \sum_{j<i,\, \gamma_j = \gamma_j' = \pm\gamma_i} \operatorname{sgn}(\gamma_j)\,, \qquad \text{where} \quad \varepsilon := \begin{cases} 0 & \text{if } \gamma_i > 0 \\ -1 & \text{otherwise}\,. \end{cases} \tag{5.1}$$

We make the convention that the sum is 0 if it contains no terms. The definition makes sense since the sum is always finite. In particular, we have the level sequence $L_\emptyset = L(\Gamma(\emptyset)) := \{l_i^\emptyset\}_{i\in I}$ of $\Gamma(\emptyset)$. Given a root α, we will use the following notation:

$$I_\alpha = I_\alpha(\Gamma) := \{i \in I \mid \gamma_i = \pm\alpha\}\,, \quad L_\alpha = L_\alpha(\Gamma) := \{l_i \mid i \in I_\alpha\}\,.$$

Remark 5.2. It is often useful to use the following graphical representation. Let $I_\alpha = \{i_1 < i_2 < \ldots < i_n\}$, and let us define the continuous piecewise-linear function $g_\alpha : [0, n] \to \mathbb{R}$ by

$$g_\alpha(0) = -\frac{1}{2}\,, \quad g_\alpha'(x) = \begin{cases} \operatorname{sgn}(\gamma_{i_k}) & \text{if } x \in (k-1, k-\frac{1}{2}) \\ \operatorname{sgn}(\gamma_{i_k}') & \text{if } x \in (k-\frac{1}{2}, k)\,, \end{cases}$$

for $k = 1, \ldots, n$. Then $l_{i_k} = g_\alpha(k - \frac{1}{2})$. For instance, assume that the entries of Γ indexed by the elements of I_α are $(\alpha, -\alpha)$, $(-\alpha, -\alpha)$, (α, α), (α, α), $(\alpha, -\alpha)$, $(-\alpha, -\alpha)$, $(\alpha, -\alpha)$, (α, α), in this order. The graph of g_α is shown on Figure 1.

We will now consider certain affine reflections corresponding to foldings Γ of $\Gamma(\emptyset)$. Let $\widehat{t}_i := s_{|\gamma_i|, l_i}$; recall that the latter is the reflection in the affine hyperplane $H_{|\gamma_i|, l_i}$. In particular, we have the affine reflections $\widehat{r}_i := s_{\beta_i, l_i^\emptyset}$ corresponding to $\Gamma(\emptyset)$.

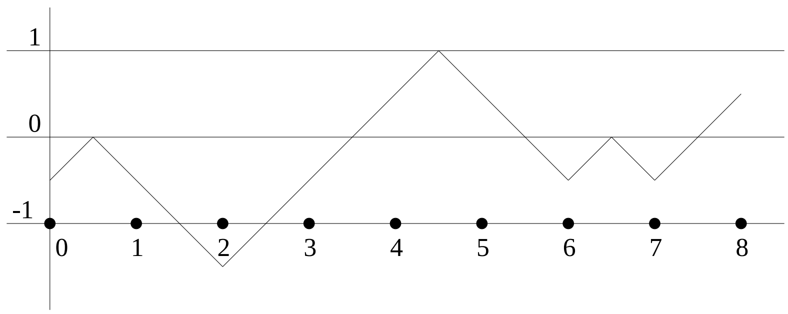


FIGURE 1

Definition 5.3. Given $J = \{j_1 < j_2 < \ldots < j_s\} \subseteq I$ and $\Gamma = \Gamma(J)$, we let

$$\mu = \mu(\Gamma) = \mu(J) := \widehat{r}_{j_1} \ldots \widehat{r}_{j_s}(\lambda)\,,$$

and call μ the *weight* of Γ (respectively J). We also use the notation $w(J) = w(\Gamma) := r_{j_1} \ldots r_{j_s}$ (recall that $r_i := s_{\beta_i}$), and

$$\widehat{I}_\alpha = \widehat{I}_\alpha(\Gamma) := I_\alpha \cup \{\infty\}\,, \quad \widehat{L}_\alpha = \widehat{L}_\alpha(\Gamma) := L_\alpha \cup \{l_\alpha^\infty\}\,, \qquad \text{where} \;\; l_\alpha^\infty := \langle \mu(\Gamma), \alpha^\vee \rangle.$$

The following proposition is our main technical result, which relies heavily on the defining properties of λ-chains.

Proposition 5.4. *Let $\Gamma = \Gamma(J)$ for some $J = \{j_1 < j_2 < \ldots < j_s\} \subseteq I$, and let $j_p < j \leq j_{p+1}$ (the first or the second inequality is dropped if $p = 0$ or $p = s$, respectively). Using the notation above, we have*

$$H_{|\gamma_j|, l_j} = \widehat{r}_{j_1} \ldots \widehat{r}_{j_p}(H_{\beta_j, l_j^\emptyset})\,.$$

Furthermore, if $\Gamma' = \phi_i(\Gamma)$, then $\mu(\Gamma') = \widehat{t}_i(\mu(\Gamma))$.

The next proposition shows that all inner products of $\mu(\Gamma)$ with positive roots can be easily read off from the level sequence $L(\Gamma) = (l_i)_{i \in I}$. Recall that, given $\Gamma = (\{(\gamma_i, \gamma_i')\}_{i \in I}, \gamma_\infty)$, we defined $t_i := s_{\gamma_i}$.

Proposition 5.5. *Given a positive root α, let $m := \max I_\alpha(\Gamma)$, assuming that $I_\alpha(\Gamma) \neq \emptyset$. Then we have*

$$\langle \mu(\Gamma), \alpha^\vee \rangle = \begin{cases} l_m + 1 & \text{if } \gamma_m' > 0 \text{ and } t_{j_1} \ldots t_{j_s}(\alpha) > 0 \\ l_m - 1 & \text{if } \gamma_m' < 0 \text{ and } t_{j_1} \ldots t_{j_s}(\alpha) < 0 \\ l_m & \text{otherwise}\,. \end{cases}$$

On the other hand, if $I_\alpha(\Gamma) = \emptyset$, then we have

$$\langle \mu(\Gamma), \alpha^\vee \rangle = \begin{cases} 0 & \text{if } t_{j_1} \ldots t_{j_s}(\alpha) > 0 \\ -1 & \text{if } t_{j_1} \ldots t_{j_s}(\alpha) < 0\,. \end{cases}$$

Remark 5.6. If $\widehat{I}_\alpha = \{i_1 < i_2 < \ldots < i_n = m < i_{n+1} = \infty\}$, we can extend the definition of the function g_α in Remark 5.2 to the interval $[0, n + \frac{1}{2}]$ in order to express $l_\alpha^\infty := \langle \mu(\Gamma), \alpha^\vee \rangle$, as given by Proposition 5.5. More precisely, letting $g_\alpha'(x) = \mathrm{sgn}(\langle \gamma_\infty, \alpha^\vee \rangle)$ for $x \in (n, n + \frac{1}{2})$, we have $l_\alpha^\infty = g_\alpha(n + \frac{1}{2})$.

We will now define some special foldings $\Gamma(J)$ of $\Gamma(\emptyset)$.

Definition 5.7. An *admissible subset* is a finite subset of I (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:

$$1 \lessdot r_{j_1} \lessdot r_{j_1} r_{j_2} \lessdot \ldots \lessdot r_{j_1} r_{j_2} \ldots r_{j_s}\,.$$

If J is an admissible subset, we will call $\Gamma = \Gamma(J)$ an *admissible folding* (of $\Gamma(\emptyset)$). We denote by $\mathcal{A}$ the collection of all admissible subsets corresponding to our fixed λ-chain.

Admissible foldings have many nice combinatorial properties, such as the ones below.

Proposition 5.8. *Any pair of roots (γ_i, γ_i') in an admissible folding has one of the following forms: (α, α), $(-\alpha, -\alpha)$, or $(\alpha, -\alpha)$, for some positive root α. The first occurence after (α, α) of a pair containing a simple root α or its negative is of the form $(\alpha, \pm\alpha)$ (assuming that such a pair exists). The same is true for the very first occurence of such a pair, if any. If (α, α) is the last occurence of a pair containing a simple root α or its negative, then $\langle \gamma_\infty, \alpha^\vee \rangle > 0$. The same is true if there are no occurences of $(\pm\alpha, \pm\alpha)$.*

6. Root Operators

We will now define *root operators* on the collection $\mathcal{A}$ of admissible subsets corresponding to our fixed λ-chain. Let J be such an admissible subset, let Γ be the associated admissible folding, and $L(\Gamma) = (l_i)_{i \in I}$ its level sequence, denoted as in Section 5.

We will first define a partial operator F_p on admissible subsets J for each p in $\{1, \ldots, r\}$, that is, for each simple root α_p. Let p in $\{1, \ldots, r\}$ be fixed throughout this section. Let $M = M(\Gamma) = M(\Gamma, p) = M(J, p)$ be the maximum of the finite set of integers $\widehat{L}_{\alpha_p}(\Gamma)$. Proposition 5.8 implies that $M \geq 0$. Assume that $M > 0$. Let $m = m_F(\Gamma) = m_F(\Gamma, p)$ be the minimum index i in $I_{\alpha_p}(\Gamma)$ for which we have $l_i = M$. If no such index exists, then $M = \langle \mu(\Gamma), \alpha_p^\vee \rangle$; in this case, we let $m = m_F(\Gamma) = m_F(\Gamma, p) := \infty$. Now let $k = k_F(\Gamma) = k_F(\Gamma, p)$ be the predecessor of m in $\widehat{I}_{\alpha_p}(\Gamma)$. Proposition 5.8 implies that this always exists and we have $l_k = M - 1 \geq 0$.

Let us now define

$$F_p(J) := \phi_k \phi_m(J)\,, \tag{6.1}$$

where ϕ_∞ is the identity map. Note that the folding of $\Gamma(\emptyset)$ associated to $F_p(J)$, which will be denoted by $F_p(\Gamma) = (\{(\delta_i, \delta_i')\}_{i \in I}, \delta_\infty)$, is defined by a similar formula. More precisely, we have

$$(\delta_i, \delta_i') = \begin{cases} (\gamma_i, \gamma_i') & \text{if } i < k \text{ or } i > m \\ (\gamma_i, s_p(\gamma_i')) & \text{if } i = k \\ (s_p(\gamma_i), s_p(\gamma_i')) & \text{if } k < i < m \\ (s_p(\gamma_i), \gamma_i') & \text{if } i = m\,, \end{cases} \qquad \text{and} \qquad \delta_\infty = \begin{cases} \gamma_\infty & \text{if } m \neq \infty \\ s_p(\gamma_\infty) & \text{if } m = \infty\,. \end{cases}$$

We can say that applying the root operator F_p amounts to performing a "folding" in position k, and, if $m \neq \infty$, an "unfolding" in position m.

We now define a partial inverse E_p to F_p. Assume that $M > \langle \mu(\Gamma), \alpha_p^\vee \rangle$. Let $k = k_E(\Gamma) = k_E(\Gamma, p)$ be the maximum index i in $I_{\alpha_p}(\Gamma)$ for which we have $l_i = M$. Proposition 5.8 implies that such indices always exist. Now let $m = m_E(\Gamma) = m_E(\Gamma, p)$ be the successor of k in $\widehat{I}_{\alpha_p}(\Gamma)$. By invoking Proposition 5.8 again, we can see that, if $m = \infty$, then we have $\langle \mu(\Gamma), \alpha_p^\vee \rangle = M - 1$, while, otherwise, we have $l_m = M - 1$. Finally, we define $E_p(J)$ by the same formula as $F_p(J)$, namely (6.1). Hence, the folding of $\Gamma(\emptyset)$ associated to $E_p(J)$ is also defined in the same way as above.

Let us now define

$$\varepsilon(J, p) = \varepsilon(\Gamma, p) := M(J, p)\,, \quad \delta(J, p) = \delta(\Gamma, p) := \langle \mu(J), \alpha_p^\vee \rangle - M(J, p)\,.$$

Proposition 6.1. *If $F_p(J)$ is defined, then it is also an admissible subset. Similarly for $E_p(J)$. Furthermore, the operators F_p and E_p satisfy Axioms* (A2) *and* (A3)*.*

7. Main Results

Recall that $\mathcal{A}$ is the collection of all admissible subsets corresponding to our fixed λ.

Theorem 7.1. *The collection $\mathcal{A}$ of admissible subsets together with the root operators form a semiperfect crystal. Thus we have the following character formula:*

$$\chi(\lambda) = \sum_{J \in \mathcal{A}} e^{\mu(J)}\,.$$

The two corollaries below follow from a general result about admissible systems (Theorem 2.4 in [Ste]).

Corollary 7.2. *(Littlewood-Richardson rule). We have*

$$\chi(\lambda)\cdot\chi(\nu)=\sum\chi(\nu+\mu(J))\,,$$

where the summation is over all J in $\mathcal{A}$ satisfying $\langle\nu+\mu(J),\alpha_p^\vee\rangle\geq M(J,p)$ for all $p=1,\ldots,r$.

Corollary 7.3. *(Branching rule). Given $P\subseteq\{1,\ldots,r\}$, we have the following rule for decomposing $\chi(\lambda)$ as a sum of Weyl characters relative to Φ_P:*

$$\chi(\lambda)=\sum\chi(\mu(J);P)\,,$$

where the summation is over all J in $\mathcal{A}$ satisfying $\langle\mu(J),\alpha_p^\vee\rangle=M(J,p)$ for all $p\in P$.

8. Lakshmibai-Seshadri Chains

In this section, we explain the connection between our model and LS chains. We start with the relevant definitions.

The Bruhat order on the orbit $W\lambda$ of a dominant or antidominant weight is defined by

$$s_\alpha(\mu)<\mu\quad\text{ if }\quad\langle\mu,\alpha^\vee\rangle>0\quad(\mu\in W\lambda,\,\alpha\in\Phi^+)\,.$$

As usual, we write $\nu\lessdot\mu$ to indicate that μ covers ν. Given $\pm\lambda\in\Lambda^+$ and a fixed real number b, one defines the b-*Bruhat order* $<_b$ as the transitive closure of the relations

$$s_\alpha(\mu)<_b\mu\quad\text{ if }\quad s_\alpha(\mu)\lessdot\mu\;\text{ and }\;b\langle\mu,\alpha^\vee\rangle\in\mathbb{Z}\quad(\mu\in W\lambda,\,\alpha\in\Phi^+)\,.$$

Definition 8.1. Given $\pm\lambda\in\Lambda^+$, we say that a pair consisting of a chain $\mu_0<\mu_1<\ldots<\mu_l$ in the W-orbit of λ and an increasing sequence of rational numbers $0<b_1<\ldots<b_l<1$ is a *Lakshmibai-Seshadri chain* (LS chain) if $\mu_0<_{b_1}\mu_1<_{b_2}\ldots<_{b_l}\mu_l$.

Following [Ste], we identify an LS chain (denoted as above) with the map $\gamma\,:\,(0,1]\to W\lambda$ given by $\gamma(t):=\mu_k$ for $b_k<t\leq b_{k+1}$, where $k=0,\ldots,l$ and $b_0:=0$, $b_{l+1}:=1$. To each LS chain γ, we associate the continuous piecewise-linear path $\pi\,:\,[0,1]\to\mathfrak{h}_\mathbb{R}^*$ given by

$$\pi(t):=\int_0^t\gamma(s)\,ds\,.$$

Let us fix λ in Λ^+. Recall the set I in (4.1), and the λ-chain $(\beta_i)_{i\in I}$ given by Proposition 4.2, which depends on a total order on the set of simple roots $\alpha_1<\cdots<\alpha_r$. We will now describe a bijection between the corresponding admissible subsets (cf. Definition 5.7) and the LS chains in the W-orbit of the antidominant weight $-\lambda$.

Given an index $i=(\alpha,k)$, we let $\beta_i:=\alpha$ and $t_i:=k/\langle\lambda,\alpha^\vee\rangle$. Recall the notation $r_i:=s_{\beta_i}$ and $\widehat{r}_i:=s_{\beta_i,l_i^\emptyset}$. Consider an admissible subset $J=\{j_1<j_2<\ldots<j_s\}$ and let

$$\{0=a_0<a_1<\ldots<a_l\}:=\{t_{j_1}\leq t_{j_2}\leq\ldots\leq t_{j_s}\}\cup\{0\}\,.$$

Let $0=n_0\leq n_1<\ldots<n_{l+1}=s$ be such that $t_{j_h}=a_k$ if and only if $n_k<h\leq n_{k+1}$, for $k=0,\ldots,l$. Define Weyl group elements u_h for $h=0,\ldots,s$ and w_k for $k=0,\ldots,l$ by $u_0:=1$, $u_h:=r_{j_1}\ldots r_{j_h}$, and $w_k:=u_{n_{k+1}}$. Let also $\mu_k:=w_k(\lambda)$. For any $k=1,\ldots,l$, we have the following saturated chain in Bruhat order of minimum (left) coset representatives modulo W_λ (the stabilizer of the weight λ):

$$w_{k-1}=u_{n_k}\lessdot u_{n_k+1}\lessdot\ldots\lessdot u_{n_{k+1}}=w_k\,;$$

indeed, none of the reflections $r_{j_1},\ldots,r_{j_s}$ lies in W_λ, since $\langle\lambda,\beta_i^\vee\rangle\neq0$ for all $i\in I$. The above chain gives rise to a saturated increasing chain from $-\mu_{k-1}$ to $-\mu_k$ in the Bruhat order on $-W\lambda$. It is not hard to show that this chain is, in fact, a chain in a_k-Bruhat order. Hence $-\mu_0<_{a_1}-\mu_1<_{a_2}\ldots<_{a_l}-\mu_l$ is an LS chain in the W-orbit of $-\lambda$. We denote it by $\gamma(J)$, and the associated continuous piecewise-linear path by $\pi(J)$.

Theorem 8.2. *The map $J \mapsto \gamma(J)$ is a bijection between the admissible subsets considered above and the LS chains in the W-orbit of the antidominant weight $-\lambda$. Moreover, we have*

$$\pi(J)(1) = -\mu(J)\,, \quad E_p(\pi(J)) = \pi(F_p(J))$$

for all admissible subsets J (here E_p is the root operator on Littelmann paths as defined in [Li1, Li2]*, while F_p in the one defined in Section* 6*).*

Remarks 8.3. (1) The proof of Theorem 8.2 contains the justification of the fact that the minima of the paths associated to LS chains are integers. This justification is based only on the combinatorics in Section 5. Note that the same fact was proved by Littelmann in [Li1] using different methods.

(2) The proof of Theorem 8.2 shows that LS chains can be viewed as a limiting case of a special case of our construction. The special choices of λ-chains that lead to LS chains represent a very small fraction of all possible choices.

Based on the independent results of Kashiwara [Kas], Lakshmibai [La], and Joseph [Jos], we deduce the following corollary.

Corollary 8.4. *Given a complex symmetrizable Kac-Moody algebra $\mathfrak{g}$, consider the colored directed graph defined by the action of root operators (cf. Section* 6*) on the admissible subsets corresponding to the special choice of a λ-chain above. This graph is isomorphic to the crystal graph of the irreducible representation with highest weight λ of the associated quantum group $U_q(\mathfrak{g})$.*

We make the following conjecture, which is the analog of a result due to Littelmann [Li2].

Conjecture 8.5. The colored directed graph defined by the action of root operators on the admissible subsets corresponding to any λ-chain does not depend on the choice of this chain.

This conjecture would imply that any choice of a λ-chain leads to a perfect crystal.

9. The Finite Case

In this section, we discuss the way in which the model in this paper specializes to the one in [LP1] in the case of finite irreducible root systems.

Let Φ be the root system of a simple Lie algebra. Let W_{aff} be the *affine Weyl group* for $\Phi^\vee$, that is, the group generated by the affine reflections $s_{\alpha,k}$ (defined in Section 2). The corresponding affine hyperplanes $H_{\alpha,k}$ divide the real vector space $\mathfrak{h}^*_{\mathbb{R}}$ into open regions, called *alcoves.* The *fundamental alcove* A is given by

$$A \;:= \{\lambda \in \mathfrak{h}^*_{\mathbb{R}} \mid 0 < \langle \lambda, \alpha^\vee \rangle < 1 \text{ for all } \alpha \in \Phi^+\}.$$

We say that two alcoves are adjacent if they are distinct and have a common wall. For a pair of adjacent alcoves, let us write $A \xrightarrow{\alpha} B$ if the common wall of A and B is of the form $H_{\alpha,k}$ and the root $\alpha \in \Phi$ points in the direction from A to B.

Definition 9.1. An *alcove path* is a sequence of alcoves $(A_0, A_1, \ldots, A_l)$ such that A_{j-1} and A_j are adjacent, for $j = 1, \ldots, l$. We say that an alcove path is *reduced* if it has minimal length l among all alcove paths from A_0 to A_l.

Let $A_\lambda = A \; + \lambda$ be the alcove obtained via the affine translation of the fundamental alcove A by a weight λ. The reduced alcove paths from A to A_λ are in bijection with the reduced decompositions of the element v_λ in W_{aff} defined by $v_\lambda(A\;) = A_\lambda$; see [LP1]. Let us fix a dominant weight λ.

Proposition 9.2. *The sequence of roots $\{\beta_i\}_{i \in I}$ with $I = \{1, \ldots, l\}$ is a λ-chain (cf. Definition* 4.1*) if and only if there exists a reduced alcove path $A_0 = A \; \xrightarrow{-\beta_1} \cdots \xrightarrow{-\beta_l} A_l = A_{-\lambda}$.*

Note that, in [LP1], (reduced) λ-chains were defined as chains of roots determined by a reduced alcove path. As we have seen, the mentioned definition is equivalent to the one in this paper.

Definition 9.3. A *gallery* is a sequence $\gamma = (F_0, A_0, F_1, A_1, F_2, \ldots, F_l, A_l, F_{l+1})$ such that $A_0, \ldots, A_l$ are alcoves; F_j is a codimension one common face of the alcoves A_{j-1} and A_j, for $j = 1, \ldots, l$; F_0 is a vertex of the first alcove A_0; and F_{l+1} is a vertex of the last alcove A_l. Furthermore, we require that $F_0 = \{0\}$, $A_0 = A_\circ$, and $F_{l+1} = \{\mu\}$ for some weight $\mu \in \Lambda$, which is called the *weight* of the gallery. The folding operator ϕ_j is the operator which acts on a gallery by leaving its initial segment from A_0 to A_{j-1} intact and by reflecting the remaining tail in the affine hyperplane containing the face F_j. In other words, we define

$$\phi_j(\gamma) := (F_0, A_0, F_1, A_1, \ldots, A_{j-1}, F'_j = F_j, A'_j, F'_{j+1}, A'_{j+1}, \ldots, A'_l, F'_{l+1}),$$

where $F_j \subset H_{\alpha,k}$, $A'_i := s_{\alpha,k}(A_i)$, and $F'_i := s_{\alpha,k}(F_i)$, for $i = j, \ldots, l+1$.

The galleries defined above are special cases of the generalized galleries in [GL].

Let us fix a reduced alcove path $A_0 = A_\circ \xrightarrow{-\beta_1} \cdots \xrightarrow{-\beta_l} A_l = A_{-\lambda}$, which determines the λ-chain $\{\beta_i\}_{i\in I}$ with $I := \{1, \ldots, l\}$. The alcove path also determines an obvious gallery

$$\gamma(\emptyset) = (F_0, A_0, F_1, \ldots, F_l, A_l, F_{l+1})$$

of weight $-\lambda$. We use the same notation as in Sections 4-6. For instance, $r_i := s_{\beta_i}$ and $\widehat{r}_i := s_{\beta_i, l_i^{\emptyset}}$.

Definition 9.4. Given an admissible subset $J = \{j_1 < \cdots < j_s\} \subseteq I$ (cf. Definition 5.7), we define the gallery $\gamma(J)$ as $\phi_{j_1} \cdots \phi_{j_s}(\gamma(\emptyset))$, and call it an *admissible folding* of $\gamma(\emptyset)$.

It is easy to see that the weight of the gallery $\gamma(J)$ is $-\mu(J)$ (cf. Definition 5.3).

Since we assumed that Φ is irreducible, there is a unique *highest coroot* $\theta^\vee \in \Phi^\vee$, i.e., a unique coroot that has maximal height. The *dual Coxeter number* of $\Phi^\vee$ is $h^\vee := \langle \rho, \theta^\vee \rangle + 1$ (in the finite case, the dominant weight ρ considered at the end of Section 2 is unique, and is given by $\frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$). Let Z be the set of the elements of the lattice $\Lambda/h^\vee$ that do not belong to any affine hyperplane $H_{\alpha,k}$. Each alcove A contains precisely one element ζ_A of the set Z (cf. [Kos, LP1]); this will be called the *central point* of A. In particular, $\zeta_{A_\circ} = \rho/h^\vee$. For a pair of adjacent alcoves $A \xrightarrow{\alpha} B$, we have $\zeta_B - \zeta_A = \alpha/h^\vee$.

Let us now associate to the gallery $\gamma(\emptyset)$ a continuous piecewise-linear path. Consider the points $\eta_0 := 0$, $\eta_{2i+1} := \zeta_{A_i}$ for $i = 0, \ldots, l$, $\eta_{2i} := \frac{1}{2}(\eta_{2i-1} + \eta_{2i+1})$ for $i = 1, \ldots, l$, and $\eta_{2l+2} := -\lambda$. Note that η_{2i} lies on F_i for $i = 0, \ldots, l+1$. Let $\pi(\emptyset)$ be the piecewise-linear path obtained by joining $\eta_0, \eta_1, \ldots, \eta_{2l+2}$. Given an admissible subset J, let $\eta'_0 = 0$, $\eta'_1 = \rho/h^\vee$, $\eta'_2, \ldots, \eta'_{2l+2} = -\mu(J)$ be the points on the faces of the gallery $\gamma(J)$ that are obtained (in the obvious way) from $\eta_0, \eta_1, \eta_2, \ldots, \eta_{2l+2}$ in the process of constructing $\gamma(J)$ from $\gamma(\emptyset)$ via folding operators. Clearly, η'_{2i+1} are the central points of the corresponding alcoves in $\gamma(J)$, for $i = 0, \ldots, l$. By joining $\eta'_0, \eta'_1, \ldots, \eta'_{2l+2}$, we obtain a piecewise-linear path that we call $\pi(J)$. Note that $\pi(J)$ can be described using folding operators, once these operators are appropriately defined. The maps $J \mapsto \gamma(J)$ and $J \mapsto \pi(J)$ are one-to-one.

Proposition 9.5. *Let $\Gamma(J) = (\{(\gamma_i, \gamma'_i)\}_{i\in I}, \gamma_\infty)$. Then, for all $i \in I$, we have*

$$\eta'_{2i-1} - \eta'_{2i} = \frac{\gamma_i}{2h^\vee}\,, \qquad \eta'_{2i} - \eta'_{2i+1} = \frac{\gamma'_i}{2h^\vee}\,, \qquad \eta'_{2l+1} - \eta'_{2l+2} = \frac{\gamma_\infty}{h^\vee}\,.$$

It turns out that, in general, the collection of paths $\pi(J)$, for J ranging over admissible subsets, does *not* coincide with the collection of Littelmann paths obtained from $\pi(\emptyset)$ by applying the root operators E_p. Indeed, it is not true in general that $E_p(\pi(J)) = \pi(F_p(J))$, as was the case with the paths corresponding to LS chains (cf. Theorem 8.2). The reason is that the root operators E_p and F_p might act on a Littelmann path $\pi : [0,1] \to \mathfrak{h}^*_{\mathbb{R}}$ by applying the reflection s_p to the direction $\pi'(t)$ of the path for t in more than one subinterval of $[0,1]$; by contrast, the root operators on admissible foldings always apply the reflection s_p to the pairs of roots in an admissible folding corresponding to a single interval of the totally ordered index set I. The situation is the same if we define $\pi(\emptyset)$ by joining the centers of the faces F_i, or the centers of both the alcoves A_i and the faces F_i (in the order they appear in $\gamma(\emptyset)$).

Example 9.6. Suppose that the root system Φ is of type G_2. The positive roots are $\gamma_1 = \alpha_1,\ \gamma_2 = 3\alpha_1 + \alpha_2,\ \gamma_3 = 2\alpha_1 + \alpha_2,\ \gamma_4 = 3\alpha_1 + 2\alpha_2,\ \gamma_5 = \alpha_1 + \alpha_2,\ \gamma_6 = \alpha_2$. The corresponding coroots are $\gamma_1^\vee = \alpha_1^\vee,\ \gamma_2^\vee = \alpha_1^\vee + \alpha_2^\vee,\ \gamma_3^\vee = 2\alpha_1^\vee + 3\alpha_2^\vee,\ \gamma_4^\vee = \alpha_1^\vee + 2\alpha_2^\vee,\ \gamma_5^\vee = \alpha_1^\vee + 3\alpha_2^\vee,\ \gamma_6^\vee = \alpha_2^\vee$.

Suppose that $\lambda = \omega_2$. Proposition 4.2 gives the following ω_2-chain:

$$(\beta_1, \dots, \beta_{10}) = (\gamma_6, \gamma_5, \gamma_4, \gamma_3, \gamma_2, \gamma_5, \gamma_3, \gamma_4, \gamma_5, \gamma_3)\,.$$

Thus, we have $\widehat{r}_1 = s_{\gamma_6,0}$, $\widehat{r}_2 = s_{\gamma_5,0}$, $\widehat{r}_3 = s_{\gamma_4,0}$, $\widehat{r}_4 = s_{\gamma_3,0}$, $\widehat{r}_5 = s_{\gamma_2,0}$, $\widehat{r}_6 = s_{\gamma_5,1}$, $\widehat{r}_7 = s_{\gamma_3,1}$, $\widehat{r}_8 = s_{\gamma_4,1}$, $\widehat{r}_9 = s_{\gamma_5,2}$, $\widehat{r}_{10} = s_{\gamma_3,2}$. There are six saturated chains in the Bruhat order (starting at the identity) on the corresponding Weyl group that can be retrieved as subchains of the ω_2-chain. We indicate each such chain and the corresponding admissible subsets in $\{1, \dots, 10\}$.

(1) 1: $\{\}$;
(2) $1 < s_{\gamma_6}$: $\{1\}$;
(3) $1 < s_{\gamma_6} < s_{\gamma_6}s_{\gamma_5}$: $\{1,2\}$, $\{1,6\}$, $\{1,9\}$;
(4) $1 < s_{\gamma_6} < s_{\gamma_6}s_{\gamma_5} < s_{\gamma_6}s_{\gamma_5}s_{\gamma_4}$: $\{1,2,3\}$, $\{1,2,8\}$, $\{1,6,8\}$;
(5) $1 < s_{\gamma_6} < s_{\gamma_6}s_{\gamma_5} < s_{\gamma_6}s_{\gamma_5}s_{\gamma_4} < s_{\gamma_6}s_{\gamma_5}s_{\gamma_4}s_{\gamma_3}$: $\{1,2,3,4\}$, $\{1,2,3,7\}$, $\{1,2,3,10\}$, $\{1,2,8,10\}$, $\{1,6,8,10\}$;
(6) $1 < s_{\gamma_6} < s_{\gamma_6}s_{\gamma_5} < s_{\gamma_6}s_{\gamma_5}s_{\gamma_4} < s_{\gamma_6}s_{\gamma_5}s_{\gamma_4}s_{\gamma_3} < s_{\gamma_6}s_{\gamma_5}s_{\gamma_4}s_{\gamma_3}s_{\gamma_2}$: $\{1,2,3,4,5\}$.

The weight of each admissible subset is now easy to compute (by applying the corresponding affine reflections above to ω_2, cf. Definition 5.3). This leads to the expression for the character $\chi(\omega_2)$ as the following sum over admissible subsets:

$$\chi(\omega_2) = e^{\omega_2} + e^{\widehat{r}_1(\omega_2)} + e^{\widehat{r}_1\,\widehat{r}_2(\omega_2)} + e^{\widehat{r}_1\,\widehat{r}_6(\omega_2)} + e^{\widehat{r}_1\,\widehat{r}_9(\omega_2)} + \cdots + e^{\widehat{r}_1\,\widehat{r}_6\,\widehat{r}_8\,\widehat{r}_{10}(\omega_2)} + e^{\widehat{r}_1\,\widehat{r}_2\,\widehat{r}_3\,\widehat{r}_4\,\widehat{r}_5(\omega_2)}.$$

Figure 2 displays the galleries $\gamma(J)$ corresponding to the admissible subsets J indicated above, the associated paths $\pi(J)$, as well as the action of the root operators F_p on J. For each path, we shade the fundamental alcove, mark the origin by a white dot "$\circ$", and mark the endpoint of a black dot "$\bullet$". Since some linear steps in $\pi(J)$ might coincide, we display slight deformations of these paths, so that no information is lost in their graphical representations. As discussed above, the weights of the irreducible representation V_{ω_2} are obtained by changing the signs of the endpoints of the paths $\pi(J)$ (marked by black dots). The roots in the corresponding admissible foldings $\Gamma(J)$ can also be read off; see Proposition 9.5. At each step, a path $\pi(J)$ either crosses a wall of the affine Coxeter arrangement or bounces off a wall. The associated admissible subset J is the set of indices of bouncing steps in the path.

References

[GL] S. Gaussent and P. Littelmann. LS-galleries, the path model and MV-cycles. `arXiv:math.RT/0307122`.

[Jos] A. Joseph. *Quantum Groups and Their Primitive Ideals*, Springer-Verlag, New York, 1994.

[Kac] V. G. Kac. *Infinite dimensional Lie algebras.* Cambridge University Press, Cambridge, 1990.

[Kas] M. Kashiwara. Crystal bases of modified quantized enveloping algebra. *Duke Math. J.*, 73:383–413, 1994.

[Kos] B. Kostant. Powers of the Euler product and commutative subalgebras of a complex simple Lie algebra. Preprint.

[Ku] S. Kumar. *Kac-Moody groups, their flag varieties and representation theory*, volume 204 of *Progress in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 2002.

[La] V. Lakshmibai. Bases for quantum Demazure modules. In *Representations of groups (Banff, AB, 1994)*, volume 16 of *CMS Conf. Proc.*, pages 199–216. Amer. Math. Soc., Providence, RI, 1995.

[LS] V. Lakshmibai and C. S. Seshadri. Standard monomial theory. In *Proceedings of the Hyderabad Conference on Algebraic Groups (Hyderabad, 1989)*, pages 279–322, Madras, 1991. Manoj Prakashan.

[LP1] C. Lenart and A. Postnikov. Affine Weyl groups in K-theory and representation theory. `arXiv:math.RT/0309207`.

[LP2] C. Lenart and A. Postnikov. Combinatorial aspects of the alcove path model. In preparation.

[Li1] P. Littelmann. A Littlewood-Richardson rule for symmetrizable Kac-Moody algebras. *Invent. Math.*, 116:329–346, 1994.

[Li2] P. Littelmann. Paths and root operators in representation theory. *Ann. of Math. ( )*, 142:499–525, 1995.

[Li3] P. Littelmann. Characters of representations and paths in $\mathfrak{h}^*_{\mathbb{R}}$. *Proc. Sympos. Pure Math.*, 61:29–49, 1997.

[PR] H. Pittie and A. Ram. A Pieri-Chevalley formula in the K-theory of a G/B-bundle. *Electron. Res. Announc. Amer. Math. Soc.*, 5:102–107, 1999.

[Ste] J. R. Stembridge. Combinatorial models for Weyl characters. *Adv. Math.*, 168:96–131, 2002.

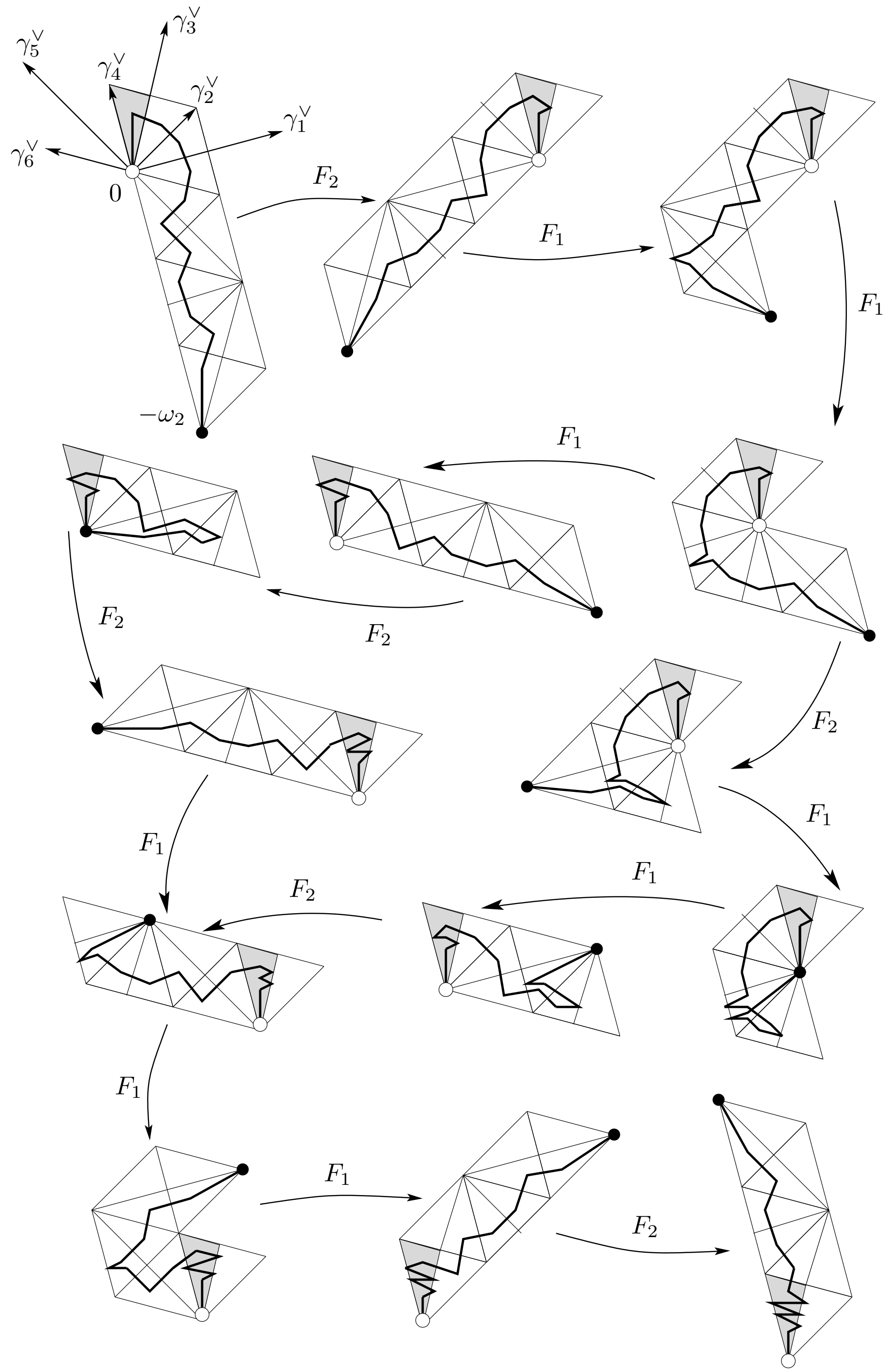


FIGURE 2. The crystal for the fundamental weight ω_2 for type G_2.

DEPARTMENT OF MATHEMATICS AND STATISTICS, STATE UNIVERSITY OF NEW YORK, ALBANY, NY 12222

E-mail address: lenart@csc.albany.edu

DEPARTMENT OF MATHEMATICS, M.I.T., CAMBRIDGE, MA 02139

E-mail address: apost@math.mit.edu

COUNTING OCCURRENCES OF 231 IN AN INVOLUTION

Toufik Mansour

Department of Mathematics
University of Haifa, 31905 Haifa, Israel
toufik@math.haifa.ac.il

Sherry H.F. Yan* and Laura L.M. Yang†
Center for Combinatorics, LPMC
Nankai University, Tianjin 300071, P.R. China
**yanhuifang@sina.com*, †*yanglm@hotmail.com*

Abstract

We study the generating function for the number of involutions on n letters containing exactly $r \geqslant 0$ occurrences of 231. It is shown that finding this function for a given r amounts to a routine check of all involutions of length at most $2r+2$.

Nous étudions la fonction génératrice pour le nombre des involutions sur n lettres en comprenent précisément $r \geqslant 0$ apparitions de 231. Nous démontrons q'il est possible a trouver cette fonction pour un nombre r donné par une vérification de routine des toutes les involutions qui ont leur longueur non plus de $2r+2$.

2000 Mathematics Subject Classification: Primary 05A05, 05A15; Secondary 05C90

1. Introduction

Permutations. Suppose that S_n is the set of permutations of $[n] = \{1, \ldots, n\}$, written in one-line notation. Let $\pi = \pi_1\pi_2 \ldots \pi_n \in S_n$ and $\tau = \tau_1\tau_2 \ldots \tau_k \in S_k$ be two permutations. An *occurrence* of τ in π is a subsequence $\pi_{i_1}\pi_{i_2} \ldots \pi_{i_k}$ such that $1 \leq i_1 < i_2 < \cdots < i_k \leqslant n$ and $\pi_{i_s} < \pi_{i_t}$ if and only if $\tau_s < \tau_t$ for any $1 \leqslant s, t \leqslant k$. In such a context, τ is usually called a *pattern*. We denote the number of occurrences of τ in π by $\tau(\pi)$.

Starting with 1985, much attention has been paid to the counting problem of the number $S_r^\tau(n)$ of permutations of length n which contain the pattern τ exactly $r \geq 0$ times. Most of the authors consider only the case $r = 0$, thus studying permutations *avoiding* a given pattern (see [1, 2, 3, 6, 13, 16, 17, 18, 19, 20]). For the case $r > 0$ there exist only a few papers, usually restricting themselves to the patterns of length three. Using two simple involutions (reverse and complement) on S_n it is immediate that with respect to being equidistributed, the six patterns of S_3 fall into two classes, namely $\{123, 321\}$ and $\{132, 213, 231, 312\}$. In 1996, Noonan [15] has proved that $S_1^{123}(n) = \frac{3}{n}\binom{2n}{n-3}$. A general approach to the counting problem was suggested by Noonan and Zeilberger [16]; they gave another proof of Noonan's result, and conjectured that $S_1^{132}(n) = \binom{2n-3}{n-3}$ and

$$S_2^{123}(n) = \frac{59n^2+117n+100}{2n(2n-1)(n+5)}\binom{2n}{n-4}.$$

The first conjecture was proved by Bóna [5] and the second one was proved by Fulmek [10]. A general conjecture of Noonan and Zeilberger states that the sequence $\{S_r^\tau(n)\}_{n\geqslant 0}$ is P-recursive in n for any r and τ. It was proved by Bóna [4] for $\tau = 132$. However, as stated in [4], a challenging question is to describe $S_r^\tau(n)$, $\tau \in S_3$, explicitly for any given r. In 2002, Mansour and Vainshtein suggested in [14] a new approach to this problem in the case $\tau = 132$, which allows to get an explicit expression for $S_r^{132}(n)$ for any given r. More precisely, they presented an algorithm that computes the generating function $\sum_{n\geqslant 0} S_r^{132}(n)x^n$ for any $r \geqslant 0$. To get the result for a given r, the algorithm performs certain routine checks for each permutation of S_{2r}.

Involutions. An *involution* π is a permutation in S_n such that $\pi = \pi^{-1}$; let $\mathcal{I}_n$ be the set of all the involutions in S_n. We denote $I_{r,n}^\tau$ the number of involutions $\pi \in \mathcal{I}_n$ with $\tau(\pi) = r$, and $I_r^\tau(x)$ the corresponding generating function, that is, $I_r^\tau(x) = \sum_{n\geqslant 0} I_{r,n}^\tau x^n$.

Again, most authors considered the case $r = 0$, namely involutions avoiding a given pattern τ (see [7, 9, 11, 12] and references therein). For the case $r > 0$ there exist only few results. In 2002, Guibert and Mansour [12] gave an explicit expression for $I_{1,n}^{132}$, namely $I_{1,n}^{132} = \binom{n-2}{[(n-3)/2]}$. Egge and Mansour in [8] proved that $I_{1,n}^{231} = (n-1)2^{n-6}$ for $n \geqslant 5$.

In the present paper we suggest a new approach to this problem in the case of $\tau = 231$, which allows to get an explicit expression for $I_{r,n}^{231}$ for any given r. More precisely, we present an algorithm that computes the generating function $I_r^{231}(x) = \sum_{n\geqslant 0} I_{r,n}^{231}x^n$ for any $r \geqslant 0$. To get the result for a given r, the algorithm performs certain routine checks for each element in $\bigcup_{k=1}^{2r+2} I_k$. The algorithm has been implemented in C, and yielded explicit results for $0 \leqslant r \leqslant 7$.

2. Preliminary results

For any involution $\pi \in \mathcal{I}_n$, we can assign a bipartite graph G_π in the following way which is similar to [14].

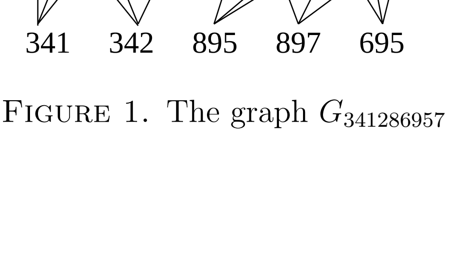


FIGURE 1. The graph $G_{341286957}$

The vertices in one part of G_π, denoted V_1 are the entries of π, and the vertices of the second part, denoted V_3, are the occurrences of 231 in π. Entry $i \in V_1$ is connected by an edge to occurrence $j \in V_3$ if i enters j. For example, let $\pi = 341286957$, then π contains 5 occurrences of 231, and the graph G_π is presented on Figure 1.

Let $\widetilde{G}$ be an arbitrary connected component of G_π, and let $\widetilde{V}$ be its vertex set. We denote $\widetilde{V_1} = \widetilde{V} \bigcap V_1$, $\widetilde{V_3} = \widetilde{V} \bigcap V_3$, $t_1 = |\widetilde{V_1}|$, $t_3 = |\widetilde{V_3}|$. Denote by G_π^n the connected component of G_π containing entry n.

For any $\pi \in \mathcal{I}_n$ with $\pi_j = n$ and $|V_1(G_\pi^n)| > 1$, assume that i_1 is the minimal index such that there exists a subsequence

$$(\pi_{i_1}, \pi_{i_2}, i_1, \pi_{i_3}, i_2, \ldots, i_h, \pi_{i_{h+2}}, i_{h+1}, \ldots, \pi_{i_k}, i_{k-1}, \pi_{i_{k+1}}, i_k, i_{k+1})$$

where $i_1 < i_2 < i_3 < \ldots < i_k < i_{k+1} = j$. we call this subsequence *connected sequence*. For our convenience, we call i_1 the *initial index*. Also, It is obvious that π_{i_1} is the first entry of the subsequence of π contained in G_π^n.

Definition 2.1. *For any $\pi \in \mathcal{I}_n$ and $\pi_j = n$, we define the* 231-tail *by*

$$\chi_\pi = \begin{cases} (n, \pi_{j+1}, \ldots, \pi_{n-1}, j), & \text{if} \quad |V_1(G_\pi^n)| = 1, \\ (\pi_{i_1}, \pi_{i_1+1}, \ldots, \pi_n), & \text{if} \quad |V_1(G_\pi^n)| > 1, \end{cases}$$

where i_1 is the initial index of π.

For example, the 231-tail of the involution 216483957 is 6483957. Denote l_π and c_π the length of π and the number of the occurrences of 231 in π.

In fact, for any $\pi \in \mathcal{I}_n$ with $|V_1(G_\pi^n)| = 1$, the 231-tail χ_π of π can be represented as $\chi_\pi = (n, n-1, \ldots, n-s+1, \lambda)$ where the first entry of λ is not $n-s$. The following lemma holds by the definition of the 231-tail.

Lemma 2.2. *Let $\pi \in \mathcal{I}_n$, the permutation of* 231*-tail of π, χ_π, is an involution, and there exists an involution π' such that $\pi = (\pi', \chi_\pi)$.*

Lemma 2.3. *Let $\pi \in I_n$ with $\chi_\pi = (n, n-1, \ldots, n-s+1, \lambda)$, where λ is nonempty, such that $c_{\chi_\pi} = r$, then $l_{\chi_\pi} \leq 2r+2$. Furthermore, The equality holds if and only if $\chi_\pi = (2r+2, 2r+1, \ldots, r+3, r+1, r+2, r, r-1, \ldots, 1)$.*

Proof. If l_π is maximal, then the first entry of χ_π is l_π. Using Lemma 2.2 we get the last entry of χ_π is 1. By induction we can assume that $\chi_\pi = (n, n-1, \ldots, n-s+1, \mu, s, s-1, \ldots, 1)$ where the first entry of μ is not $n-s$ and μ is nonempty. On the other hand $c_\pi = r$, so $s \leq r$. Hence

$$\chi_\pi = (2r+2, 2r+1, \ldots, r+3, r+1, r+2, r, r-1, \ldots, 1).$$

□

Lemma 2.4. *For any $\pi \in \mathcal{I}_n$ with $|V_1(G_\pi^n)| > 1$, the subsequence of π contained in the connected component G_π^n is just the* 231*-tail χ_π of π.*

Proof. According to the definition of the 231-tail, it is sufficient to prove that the bipartite graph corresponding to χ_π is connected. Assume $\pi_j = n$ and i_1 is the initial index. Suppose that the connected sequence is

$$(\pi_{i_1}, \pi_{i_2}, i_1, \pi_{i_3}, i_2, \ldots, i_h, \pi_{i_{h+2}}, i_{h+1}, \ldots, \pi_{i_k}, i_{k-1}, \pi_{i_{k+1}}, i_k, i_{k+1})$$

where $i_1 < i_2 < i_3 < \ldots < i_k < i_{k+1} = j$. It is obvious that the vertices

$$\pi_{i_1}, \pi_{i_2}, i_1, \pi_{i_3}, i_2, \ldots, i_h, \pi_{i_{h+2}}, i_{h+1}, \ldots, \pi_{i_k}, i_{k-1}, \pi_j, i_k, j$$

are all contained in the connected component G_π^n.

For $i_h < m < i_{h+1}$, if $i_h < \pi_m < \pi_{i_{h+1}}$, then $\pi_m \pi_{i_{h+1}} i_h$ forms a pattern of 231 in π; if $\pi_m > \pi_{i_{h+1}}$, then $\pi_{i_h} \pi_m i_h$ is a subsequence of the pattern 231 in π; if $\pi_m < i_h$, then $m \pi_{i_h} \pi_m$ is a subsequence of the pattern of 231 in π. In these cases, we know that π_m is contained in G_π^n.

For $j < m < n$, if $\pi_m < \pi_{i_k}$, then $\pi_{i_k} n \pi_m$ forms a pattern of 231 in π; if $\pi_m > \pi_{i_k}$, then $\pi_{i_k} \pi_m j$ forms a pattern of 231 in π. In these cases, we know that π_m is contained in G_π^n. □

Studying occurrences of 132 in a permutation which leads to consideration of 231 in a permutation, Mansour and Vainshtein have proved that the relation $t_1 \leq 2t_3 + 1$ in [14]. It is clear that the set of involutions is a subset of permutations. So we have

Lemma 2.5. (see [14, Lemma 2.1]) *For any connected component $\widetilde{G}$ of G_π, one has $t_1 \leq 2t_3 + 1$.*

Remark 2.6. *For any $\pi \in I_n$ with $c_{\chi_\pi} = r$ $(r > 0)$, if $|V_1(G_\pi^n)| = 1$, then $l_{\chi_\pi} \leq 2r + 2$; otherwise $l_{\chi_\pi} \leq 2r + 1$.*

3. Main Theorem and explicit results

Denote by K_t the subset of $\bigcup_{k \leqslant 2t+2} \mathcal{I}_k$ whose elements can be expressed as $(k, k-1, \ldots, k-s+1, \lambda)$ where λ is nonempty, and by H_t be the subset of $\bigcup_{k \leqslant 2t+1} \mathcal{I}_k$ such that the corresponding bipartite graph of each element is connected. It is obvious that $K_t \cap H_t = \emptyset$. Then the main result of this paper can be formulated as follows.

Theorem 3.1. *For any $r \geqslant 1$,*

$$I_r^{231}(x) = \frac{x}{1-x} I_r^{231}(x) + \sum_{\mu \in K_r \cup H_r} x^{l_\mu} I_{r-c_\mu}^{231}(x). \qquad (*)$$

Proof. Denote by $F_r^\mu(x)$ the generating function for the number of involutions $\pi \in \mathcal{I}_n$ such that χ_π is just order-isomorphic to μ. We discuss three cases to find $F_r^\mu(x)$:

- If π is an involution in $\mathcal{I}_n$ with $\chi_\pi = (n, n-1, \ldots, n-s+1)$, then $l_\mu = s$ and $\mu = (s, s-1, \ldots, 1)$, so we have
$$F_r^\mu(x) = x^s I_r^{231}(x).$$
- If π is an involution in $\mathcal{I}_n$ with $\chi_\pi = (n, n-1, \ldots, n-s+1, \lambda)$ where λ is nonempty, then $\mu \in K_r$ by Lemma 2.3, thus we have
$$F_r^\mu(x) = x^{l_\mu} I_{r-c_\mu}^{231}(x).$$
- If π is an involution in $\mathcal{I}_n$ with $\chi_\pi = (\pi_{i_1}, \pi_{i_1+1}, \ldots, \pi_n)$ where i_1 is the initial index of π, then Lemma 2.5 and Lemma 2.4 yield $\mu \in H_r$ and
$$F_r^\mu(x) = x^{l_\mu} I_{r-c_\mu}^{231}(x).$$

Hence, summing over all $\mu \in \{(s, s-1, s-2, \ldots, 2, 1) | s \geqslant 1\} \cup K_r \cup H_r$ we get the desired result. $\square$

Theorem 3.1, Lemma 2.3, and Lemma 2.5 provide a finite algorithm for finding $I_r^{231}(x)$ for any given $r > 0$, since we only have to consider all involutions in I_k, where $k \leqslant 2r+2$, and to perform certain routine operations with all 231-tails found so far.

Remark 3.2. *In fact, according to the Lemma* 2.3*, it is sufficient to check all involutions in I_k, where $k \leq 2r+1$. As a consequence, Formula* $(*)$ *can be reduced as follows:*

$$I_r^{231}(x) = \frac{x}{1-x} I_r^{231}(x) + x^{2r+2} I_0^{231}(x) + \sum_{\mu \in K_r^* \cup H_r} x^{l_\mu} I_{r-c_\mu}^{231}(x),$$

where K_r^* is the set of all involutions of the form $(n, n-1, \ldots, n-s+1, \lambda)$ of in I_k where $k \leq 2t+1$ and λ is nonempty.

Let us start from the case $r = 0$. Observe that $(*)$ remains valid for $r = 0$, provided the left hand side is replaced by $I_0^{231}(x) - 1$; subtracting 1 here accounts for the empty permutation. Note that when $r = 0$, $K_0 \cup H_0$ is empty. Hence we get $I_0^{231}(x) - 1 = \frac{x}{1-x} I_0^{231}(x)$, equivalently

$$I_0^{231}(x) = \frac{1-x}{1-2x}, \qquad (**)$$

which is the result of Simion and Schmidt (see [17, Proposition 6]).

Let now $r = 1$. Observe that $K_1 \cup H_1 = \{4231\}$. Therefore, $(*)$ amounts to

$$I_1^{231}(x) = \frac{x}{1-x} I_1(x) + x^4 I_0^{231}(x),$$

and we get the following result from Formula $(**)$.

Corollary 3.3. (see Egge and Mansour [8, Theorem 4.3]) *The generating function $I_1^{231}(x)$ for the number of involutions containing exactly one occurrence of the pattern* 231 *is given by*

$$I_1^{231}(x) = \frac{x^4(1-x)^2}{(1-2x)^2};$$

equivalently, for $n \geq 5$,

$$I_{1,n}^{231} = (n-1)2^{n-6}.$$

Let $r = 2$. Exhaustive search adds four new elements to the previous list; these are 653421, 52431, 53241, and 3412, therefore we get

Corollary 3.4. *The generating function $I_2^{231}(x)$ is given by*

$$I_2^{231}(x) = \frac{x^4(1-x)^2}{(1-2x)^3}\left(1 - 3x^2 - 2x^3 + x^4 - x^5\right);$$

equivalently, for $n \geqslant 9$,

$$I_{2,n}^{231} = (n^2 + 137n - 234)2^{n-11}.$$

Let $r = 3, 4, 5, 6, 7$; exhaustive search in $\mathcal{I}_8$, $\mathcal{I}_{10}$, $\mathcal{I}_{12}$, $\mathcal{I}_{14}$, and $\mathcal{I}_{16}$ reveals 13, 24, 41, 69, and 103 elements, respectively, and we get

Corollary 3.5. *Let $3 \leqslant r \leqslant 7$, then*

$$I_r^{231}(x) = \frac{(1-x)^2}{(1-2x)^{r+1}} Q_r(x),$$

where

$$\begin{aligned}
Q_3(x) &= x^5(4 - 14x + 8x^2 + 11x^3 - 6x^4 - 2x^5 + 2x^6 + 5x^7 - 2x^8 + x^9);\\
Q_4(x) &= x^6(6 - 32x + 49x^2 + 7x^3 - 73x^4 + 40x^5 + 30x^6 - 37x^7 + 2x^8 + 4x^{10}\\
&\quad -9x^{11} + 3x^{12} - x^{13});\\
Q_5(x) &= x^6(8 - 58x + 146x^2 - 120x^3 - 40x^4 - 24x^5 + 290x^6 - 184x^7 - 197x^8\\
&\quad +228x^9 + 30x^{10} - 132x^{11} + 62x^{12} + 13x^{14} - 16x^{15} + 14x^{16} - 4x^{17} + x^{18});\\
Q_6(x) &= x^6(4 - 31x + 80x^2 - 56x^3 + 4x^4 - 384x^5 + 1097x^6 - 830x^7 - 483x^8\\
&\quad +660x^9 + 685x^{10} - 1091x^{11} - 59x^{12} + 722x^{13} - 195x^{14} - 338x^{15}\\
&\quad +285x^{16} - 92x^{17} + 20x^{18} - 45x^{19} + 35x^{20} - 20x^{21} + 5x^{22} - x^{23});\\
Q_7(x) &= x^7(17 - 199x + 969x^2 - 2502x^3 + 3642x^4 - 3274x^5 + 3324x^6 - 4714x^7\\
&\quad +1874x^8 + 6326x^9 - 8262x^{10} - 231x^{11} + 5474x^{12} - 637x^{13} - 4022x^{14}\\
&\quad +1933x^{15} + 1340x^{16} - 1129x^{17} - 518x^{18} + 982x^{19} - 498x^{20} + 166x^{21}\\
&\quad -92x^{22} + 105x^{23} - 62x^{24} + 27x^{25} - 6x^{26} + x^{27}).
\end{aligned}$$

As an easy consequence of Theorem 3.1 we get the following result.

Corollary 3.6. *For any $r \geq 1$ there exist a polynomial $P_{5r-1}(x)$ of degree $5r-1$ with integer coefficients such that*

$$I_r^{231}(x) = \frac{(1-x)^2}{(1-2x)^{r+1}} P_{5r-1}(x).$$

Proof. Immediately, by the above cases we have the corollary holds for $1 \leq r \leq 7$. Let us assume by induction that the corollary holds for $1, 2, \dots, r-1$; for r the equation $(*)$ give

$$I_r^{231}(x) = \frac{(1-x)^2}{(1-2x)^{r+1}} \sum_{\rho \in K_r \cup H_r} x^{l_\rho} \frac{(1-2x)^r}{1-x} I_{r-c_\rho}^{231}(x).$$

By the induction assumption and $I_0^{231}(x) = \frac{1-x}{1-2x}$ we have that $x^{l_\rho}\frac{(1-2x)^r}{1-x} I_{r-c_\rho}(x)$ is a polynomial with integer coefficients of degree a. So Lemma 2.3 and 2.5 yields

$$a = \max\{b_j | j = 1, \dots, r\},$$

where $b_j = 2j + 2 + r - (r - j + 1) + 1 + 5(r - j) - 1 = 5r - 2j + 1$, which means $a = 5r - 1$, as claimed. □

4. **Further results**

Another direction would be to match the approach of this paper with the previous results on restricted 231-avoiding involutions. Let $\Phi_r(x; k)$ be the generating function for the number of involutions in $\mathcal{I}_n$ containing r occurrences of 231 and avoiding the pattern $12\dots k \in \mathfrak{S}_k$.

We denote by e_λ the length of the longest increasing subsequence of any involution λ. For example, let $\lambda = 3412$, then $e_\lambda = 2$. We denote by $K_t(k) \cup H_t(k)$ the set of all involutions $\lambda \in K_t \cup H_t$ such that $e_\lambda \le k-1$.

Theorem 4.1. *For any $r \geqslant 1$ and $k \geqslant 3$,*

$$\Phi_r(x;k) = \frac{x}{1-x}\Phi_r(x;k-1) + \sum_{\mu \in K_r(k)\cup H_r(k)} x^{l_\mu}\Phi^{231}_{r-c_\mu}(x;k-e_\mu).$$

Besides, $\Phi_r(x;1) = \Phi_r(x;2) = 0$, and $\Phi_0(x;1) = 1$ and $\Phi_0(x;2) = \frac{1}{1-x}$.

Similar to the case of $I_r^{231}(x)$, the statement of the above theorem remains valid for $r = 0$, provided the left hand side is replaced by $\Phi_r(x;k) - 1$. This allows to recover known explicit expressions for $\Phi_r(x;k)$ for $r = 0, 1$, as follows.

Corollary 4.2. (see Egge and Mansour [8]) *For all $k \ge 1$,*

$\Phi_0(x;k) = \sum\limits_{j=0}^{k-1}\left(\frac{x}{1-x}\right)^j$;

$\Phi_1(x;k) = x^4\sum\limits_{j=0}^{k-3}(j+1)\left(\frac{x}{1-x}\right)^j$.

The final direction would be to match the approach of this note with the previous results on restricted 231-avoiding even or odd involutions. We say π an even (resp; odd) involution if the number of inversion in π, namely $21(\pi)$ is even (resp; odd). We denote by $K_r^+ \cup H_r^+$ the set of all the even involutions $\lambda \in K_r \cup H_r$ and denote by $K_r^- \cup H_r^-$ the set of all the odd involutions $\lambda \in K_r \cup H_r$.

Let $I_r^+(x)$ (resp; $I_r^-(x)$) be the generating function for the number of even (resp; odd) involutions in $\mathcal{I}_n$ containing r occurrences of 231. Our new approach allows to get an explicit expression for $I_r^+(x)$ (or $I_r^-(x)$) for any given $r \geqslant 0$.

Theorem 4.3. *For all $r \geqslant 1$,*

$$I_r^+(x) = \frac{x+x^4}{1-x^4}I_r^+(x) + \frac{x^2+x^3}{1-x^4}I_r^-(x) + \sum_{\mu\in K_r^+\cup H_r^+} x^{l_\mu}I^+_{r-c_\mu}(x) + \sum_{\mu\in K_r^-\cup H_r^-} x^{l_\mu}I^-_{r-c_\mu}(x);$$

$$I_r^-(x) = \frac{x+x^4}{1-x^4}I_r^-(x) + \frac{x^2+x^3}{1-x^4}I_r^+(x) + \sum_{\mu\in K_r^+\cup H_r^+} x^{l_\mu}I^-_{r-c_\mu}(x) + \sum_{\mu\in K_r^-\cup H_r^-} x^{l_\mu}I^+_{r-c_\mu}(x).$$

In particular, we have

$$I_0^+(x) - 1 = \frac{x+x^4}{1-x^4}I_0^+(x) + \frac{x^2+x^3}{1-x^4}I_0^-(x)$$

and

$$I_0^-(x) = \frac{x+x^4}{1-x^4}I_0^-(x) + \frac{x^2+x^3}{1-x^4}I_0^+(x).$$

Proof. Here we only prove the result of $I_r^+(x)$ for any $r \geq 1$. By the same method, we can obtain the formula for $I_r^-(x)$. Denote by $F_r^\mu(x)$ the generating function for the number of even involutions in $\pi \in I_n$ such that χ_π is order-isomorphic to μ.

To find $F_r^\mu(x)$, we recall four four cases. If $\mu = (s, s-1, \ldots, 1)$ and μ is even (that is, $21(\mu) = \frac{(s-1)s}{2} = 2k$ for some positive integer k), then in this case we have

$$F_r^\mu(x) = x^s I_r^+(x).$$

If $\mu = (s, s-1, \ldots, 1)$ and μ is odd (that is, $21(\mu) = \frac{(s-1)s}{2} = 2k-1$ for some positive integer k), then in this case we have

$$F_r^\mu(x) = x^s I_r^-(x).$$

If $\mu \in K_r^+ \cup H_r^+$, then

$$F_r^\mu(x) = x^{l_\mu} I_{r-c_\mu}^+(x).$$

If $\mu \in K_r^- \cup H_r^-$, then

$$F_r^\mu(x) = x^{l_\mu} I_{r-c_\mu}^-(x).$$

Hence, if summing over all $\mu \in K_r \cup H_r \cup \{(s, s-1, s-2, \ldots, 2, 1) | s \geq 1\}$ then we get the desired result. When $r = 0$, subtracting 1 here accounts for the empty permutation. $\square$

As an example of the above theorem for $r = 0, 1, 2$, we get

Corollary 4.4.

$$I_r^+(x) = \frac{E_r(x)}{(1-2x)^{r+1}(1-x+2x^2)^{r+1}}, \quad I_r^-(x) = \frac{O_r(x)}{(1-2x)^{r+1}(1-x+2x^2)^{r+1}};$$

where

$E_0(x) = 1 - 2x + 2x^2 - 2x^3$;

$E_1(x) = 2x^6(1 - 2x + 2x^2 - 2x^3)$;

$E_2(x) = x^4(1 - 5x + 11x^2 - 15x^3 + 10x^4 + 5x^5 - 11x^6 - 5x^7 + 47x^8 - 94x^9 + 86x^{10} - 62x^{11} + 16x^{12})$;

$O_0(x) = x^2$;

$O_1(x) = x^4(1 - 4x + 8x^2 - 12x^3 + 13x^4 - 8x^5 + 4x^6)$;

$O_2(x) = x^6(2 - 6x + 6x^2 - 2x^3 - 9x^4 + 4x^5 + 20x^6 - 36x^7 + 53x^8 - 24x^9 + 8x^{10})$.

Again, as an easy consequence of Theorem 4.3 we get the following result.

Corollary 4.5. *For any $r \geqslant 0$, there exists two polynomials $P_{m_r}(x)$ and $P_{n_r}(x)$ of degree m_r and n_r with integer coefficients such that*

$$I_r^+(x) = \frac{P_{m_r}(x)}{(1-2x)^{r+1}(1-x+2x^2)^{r+1}}, \quad I_r^-(x) = \frac{P_{n_r}(x)}{(1-2x)^{r+1}(1-x+2x^2)^{r+1}},$$

where $m_r, n_r \leq \frac{r(r+9)}{2}$.

It can be proved by induction on r as the proof of Corollary 3.6. Here we delete its proof.

As a remark we can derive another results from Theorem 4.3. For example, the generating function for the number of even or odd involution containing exactly r occurrences of the pattern 231 and avoiding $12\dots k$ (or avoiding $k\dots 21$).

References

[1] N. Alon and E. Friedgut, On the number of permutations avoiding a given pattern, *J. Combin. Theory Ser. A* **89** (2000)133–140.

[2] M.D. Atkinson, Restricted permutations, *Discrete Math.* **195** (1999) 27–38.

[3] 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.

[4] M. Bóna, The number of permutations with exactly r 132-subsequences is P-recursive in the size! *Adv. Appl. Math.* **18** (1997) 510–522.

[5] M. Bóna, Permutations with one or two 132-subsequences, *Discrete Math.* **181** (1998) 267–274.

[6] T. Chow and J. West, Forbidden subsequences and Chebyshev polynomials, *Discrete Math.* **204** (1999) 119–128.

[7] E.S. Egge, Restricted 3412-avoiding involutions: Continued fractions, Chebyshev polynomials and enumerations, *Adv. in Appl. Math.*, to appear.

[8] E. Egge and T. Mansour, 231-Avoiding involutions and Fibonacci numbers, *Australasian Journal of Combinatorics*, to appear.

[9] E.S. Egge and T. Mansour, Involutions restricted by 3412, Continued fractions, and Chebyshev polynomials, preprint.

[10] M. Fulmek, Enumeration of permutations containing a presribed number of occurrences of a pattern of length three, *Adv. Appl. Math.* **30** (2003) 607–632.

[11] O. Guibert, Combinatoire des permutations à motifs exclus en liaison avec mots, cartes planaires et tableaux de Young, *PHD-thesis, University Bordeaux 1, France* (1995).

[12] O. Guibert and T. Mansour, Restricted 132-involutions, *Séminaire Lotharingien de Combinatoire* **48** (2002), Article B48a (23pp.).

[13] T. Mansour and A. Vainshtein, Restricted permutations, continued fractions, and Chebyshev polynomials,*Electron. J. Combin.* **7** (2000) #R17.

[14] T. Mansour and A. Vainshtein, Counting occurences of 132 in a permutation, *Adv. Appl. Math.* **28** (2002) 185–195.

[15] J. Noonan, The number of permutations containing exactly one increasing subsequence of length three, *Discrete Math.* **152** (1996) 307–313.

[16] J. Noonan and D. Zeilberger, The enumeration of permutations with a prescribed number of "forbidden" patterns, *Adv. Appl. Math.* **17** (1996) 381–407.

[17] R. Simion, F.W. Schmidt, Restricted Permutations, *Europ. J. of Combin.* **6** (1985) 383–406.

[18] Z. Stankova, Forbidden subsequences, *Discrete Math.* **132** (1994) 291–316.

[19] Z. Stankova, Classification of forbidden subsequences of length 4, *European J. Combin.* **17** (1996) 501–517.

[20] J. West, Generating trees and the Catalan and Schröder numbers, *Discrete Math.* **146** (1995) 247–262.

A Decomposition of the Schur Functions into Non-Symmetric Schur Functions

Sarah Mason

Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104-6395
sarahm2@math.upenn.edu

November 13, 2004

Abstract

The Schur functions, $s_\lambda(x)$, form a basis for the vector space of symmetric functions. Recently Haglund, Haiman and Loehr have derived a combinatorial formula for nonsymmetric Macdonald polynomials, which gives a new decomposition of the Macodnald 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. In this paper, we exhibit a bijection between semi-standard Young tableaux (SSYT) and skyline fillings to give a bijective proof of the formula.

Resumé en Français

Les fonctions de Schur, $s_\lambda(x)$, forment une base de l'espace vectoriel de fonctions symétriques. Les résultats récents de J. Haglund permettent d'introduire un objet nouveau qui est utilisé pour décomposer les fonctions de Schur en fonctions nonsymétriques, $NS_\mu(x)$, numérotées par les compositions au lieu des partitions. Le théorème principal de cet article (qui était conjecturé par J. Haglund) dit que $s_\lambda(x) = \sum_\mu NS_\mu(x)$, sommée sur toutes les transpositions μ de λ. Dans cet article, nous montrons une bijection entre les tableaux de Young semi-standards (SSYT) et les remplissages d'horizon pour démontrer la conjecture.

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 $x_1^{\alpha_1} x_2^{\alpha_2} ...$, and $f(x_{\omega(1)}, x_{\omega(2)}, ...) = f(x_1, x_2, ...)$ 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 $x = (x_1, x_2, ...)$ 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 Tableaux 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 [HHL$^+$]. 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 $\{1^{a_1}, 2^{a_2}, ..., n^{a_n}\}$, 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 [Sta99] 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 [Mac88] introduced them and conjectured that their expansion in terms of Schur polynomials should have positive coefficients. A combinatorial formula for the Macdonald polynomials was conjectured by Haglund and proved by Haglund, Haiman, and Loehr [HHL04].

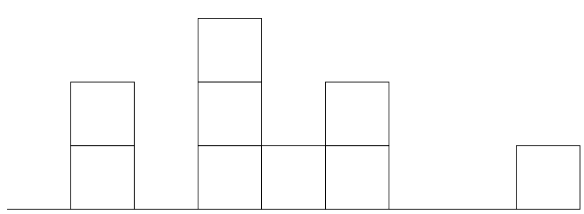
Building on this work, Haglund described [Hag04b] a conjectured combinatorial formula for the non-symmetric Macdonald polynomials. As a consequence of this conjecture he gives a set of objects that decompose the Schur functions into non-symmetric functions indexed by compositions of n instead of partitions of n. They involve statistics generalizing those described in [HHL04]. The weighted sums of these objects are called the *non-symmetric Schur functions*, NS_{λ}. 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. Haglund conjectured that the sum of the non-symmetric Schur functions over all rearrangements of a given partition λ is equal to the ordinary Schur function s_{λ}. In this paper, we prove:

Theorem 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.*

This result gave evidence that Haglund's conjectured formula for the non-symmetric Macdonald polynomials was correct. Haglund, Haiman, and Loehr recently proved this formula [HHL05]. Setting $q = t = 0$ in the formula gives a non-bijective proof of the above theorem.

2 Combinatorial Definition of the non-symmetric Schur Functions

A *composition Young diagram of n* is a figure consisting of n *cells* arranged in n columns. A column may contain anywhere from 0 to n cells, and the number of cells in a column is called the *height* of that column. This means that a composition Young diagram is simply the Young diagram of a composition of n into n parts, allowing zeros. The composition 9=0+2+0+3+1+2+0+0+1 is shown in the example below.



The composition Young diagram for $\lambda = (0, 2, 0, 3, 1, 2, 0, 0, 1)$

A *filling,* σ, of a composition Young 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 ith 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 non-symmetric Schur functions, we need the statistics $\mathrm{maj}(\sigma, \lambda)$ and $\mathrm{inv}(\sigma, \lambda)$. As in [Hag04a], a *descent* of σ is a pair of entries $\sigma(u) > \sigma(v)$, where the cell u is directly above v. In other words, $v = (i, j)$ and $u = (i+1, j)$, where the i^{th} coordinate denotes the height of cell v and the j^{th} coordinate denotes one less than the number of cells to the left of v. Define $\mathrm{Des}(\sigma) = \{u \in \lambda : \sigma(u) > \sigma(v)$ is a descent$\}$.

Three cells $u, v, w \in \lambda$ form a *triple of type A* if they are situated as shown below,

$$\begin{array}{ccc} \boxed{u} & \cdots\cdots & \boxed{w} \\ \boxed{v} & & \end{array}$$

where u and w are in the same row, possibly with cells between them, and the column containing u and v has height greater than or equal to the height of the column containing w.

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 x, y, z be the entries of σ in the cells of a type A triple (u, v, w):

$$\begin{array}{ccc} \boxed{x} & \cdots\cdots & \boxed{z} \\ \boxed{y} & & \end{array}$$

Then the triple (u, v, w) is an *inversion triple of type A* if and only if $I(x, z) + I(z, y) - I(x, y) = 1$.

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 x in the image of σ, the restriction of ξ to $\sigma^{-1}(\{x\})$ is increasing with respect to the reading order. Therefore the triple (u, v, w) is an inversion triple of type A if and only if after standardization, the ordering from smallest to largest of the entries in cells u, v, w induces a counter-clockwise orientation.

Similarly, three cells $u, v, w \in \lambda$ form a *triple of type B* if they are situated as shown below,

$$\begin{array}{ccc} & & \boxed{v} \\ \boxed{u} & \cdots\cdots & \boxed{w} \end{array}$$

Here u and w are in the same row (possibly row 0) and the column containing v and w has greater height than the column containing u.

Let σ be a composition filling and let x, y, z be the entries of σ in the cells of a type B triple (u, v, w):

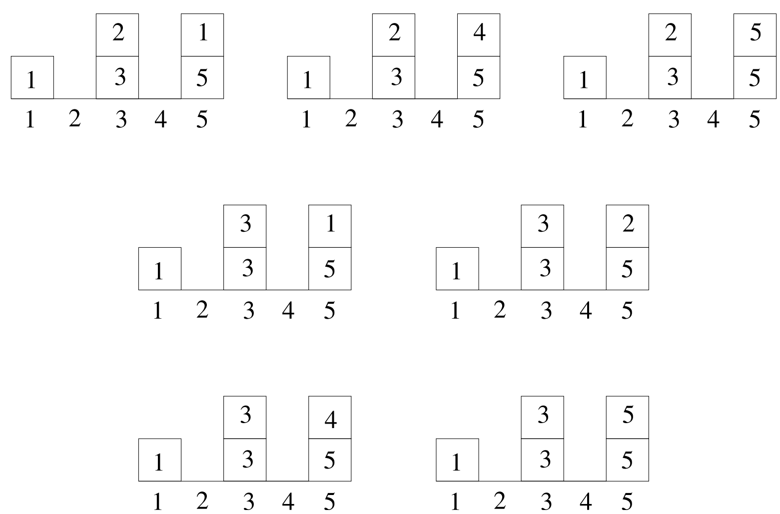
$$\begin{array}{ccc} & & \boxed{y} \\ \boxed{x} & \cdots\cdots & \boxed{z} \end{array}$$

Then the triple (u, v, w) is an *inversion triple of type B* if and only if $I(y, x) + I(x, z) - I(y, z) = 1$. In other words, the triple (u, v, w) is an inversion triple of type B if and only if after standardization, the ordering from smallest to largest of the entries in cells u, v, w induces a clockwise orientation.

Denote by *semi-standard skyline filling* any composition filling K such that $\mathrm{Des}(K) = \emptyset$ and every triple is an inversion triple. These conditions arise by setting $q = t = 0$ in the combinatorial formula for the non-symmetric Macdonald polynomials.

Definition 1 *Let λ be a composition of n into n parts, where some of the parts could be equal to zero. The non-symmetric Schur function $NS_\lambda = NS_\lambda(x)$ in the variables $x = (x_1, x_2, ..., x_n)$ is the formal power series $NS_\lambda(x) = \sum_K x^K$ summed over all semi-standard skyline fillings K of composition λ. Here, $x^K = \prod_{i=1}^n x_{\sigma_i}$ is the weight of σ.*

As an example, take $\lambda = (1, 0, 2, 0, 2)$. The skyline fillings with no descents such that every triple is an inversion triple are as follows:



$\lambda = (1, 0, 2, 0, 2)$

Therefore, $NS_\lambda = x_1^2x_2x_3x_5 + x_1x_2x_3x_4x_5 + x_1x_2x_3x_5^2 + x_1^2x_3^2x_5 + x_1x_2x_3^2x_5 + x_1x_3^2x_4x_5 + x_1x_3^2x_5^2$.

3 Map from SSYTs to Skyline Fillings

The purpose of this paper is to prove that the sum of the non-symmetric Schur functions over all rearrangements of a given partition μ is equal to the ordinary Schur function s_μ. (Here, a *rearrangement* of a partition μ of n is a composition of n into n parts such that when these parts are arranged in decreasing order the partition μ is recovered.) To do this, we must exhibit a bijection between semi-standard young tableaux and skyline fillings which preserves the number of objects with each weight.

Begin with an arbitrary semi-standard young tableau T of shape μ, where $\mu \vdash n$. The cells are labeled by some multiset of positive integers $\{1^{a_1}, 2^{a_2}, ..., n^{a_n}\}$. (Note that some of the a_i might equal 0.) Map the value in each cell to a new value as follows: send α to $n - \alpha + 1$. Call the new filling T'. (Here, a filling of shape μ is a function $\sigma : \mu \to \mathbb{Z}_+$, as defined in [HHL04].) Notice that the column entries are now strictly decreasing and the row entries are weakly decreasing. Let T_i' be the set consisting of the entries of the i^{th} column.

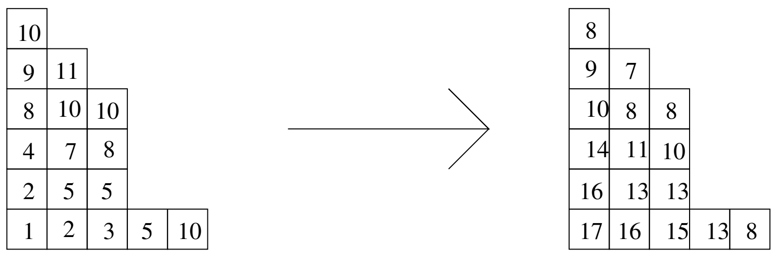
Place the elements of T_i' on top of row $i - 1$ as follows:

Begin with the largest member, α_1, of T_i'. Find the left-most entry of row $i-1$ that is greater than or equal to α_1. We know such an element exists, since in the tableau, the entry to the immediate left of α_1

is greater than or equal to α_1. Place α_1 on top of this element. Next place the second-largest member, α_2 of T_i' in the same way. (Again, an entry greater than α_2 exists because the entry immediately to the left of α_2 in the tableau is greater than or equal to α_2.) Continue in this manner until all the elements of T_i' have been placed. Any remaining cell of row $i-1$ has no cell directly above it.

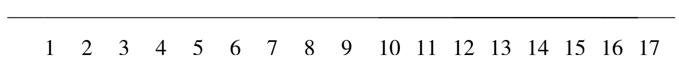
Following this process for each column of T produces a filling of a composition Young diagram, as in the example below:

Example 1 *Begin with a Semi-Standard Young Tableau of shape λ =(5,3,3,3,2,1) (note that $\lambda \vdash$ 17) as pictured below and apply the map described above that sends each of the numbers, α, from 1 through 17, to $17-\alpha+1$.*



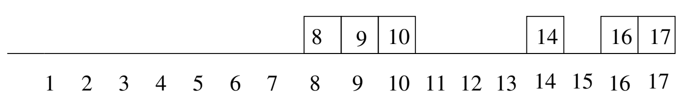
SSYT mapping

Next examine the empty composition filling:



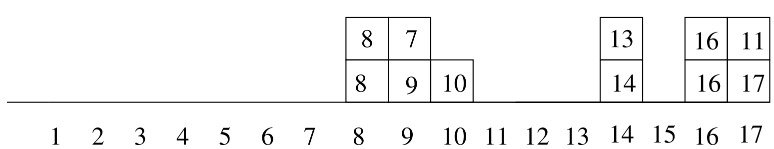
The Empty Composition Filling

We must assign the numbers T_1 ={8,9,10,14,16,17} to the first row of our composition filling according to the map defined above. The figure below shows the placement of these numbers onto the empty composition filling:



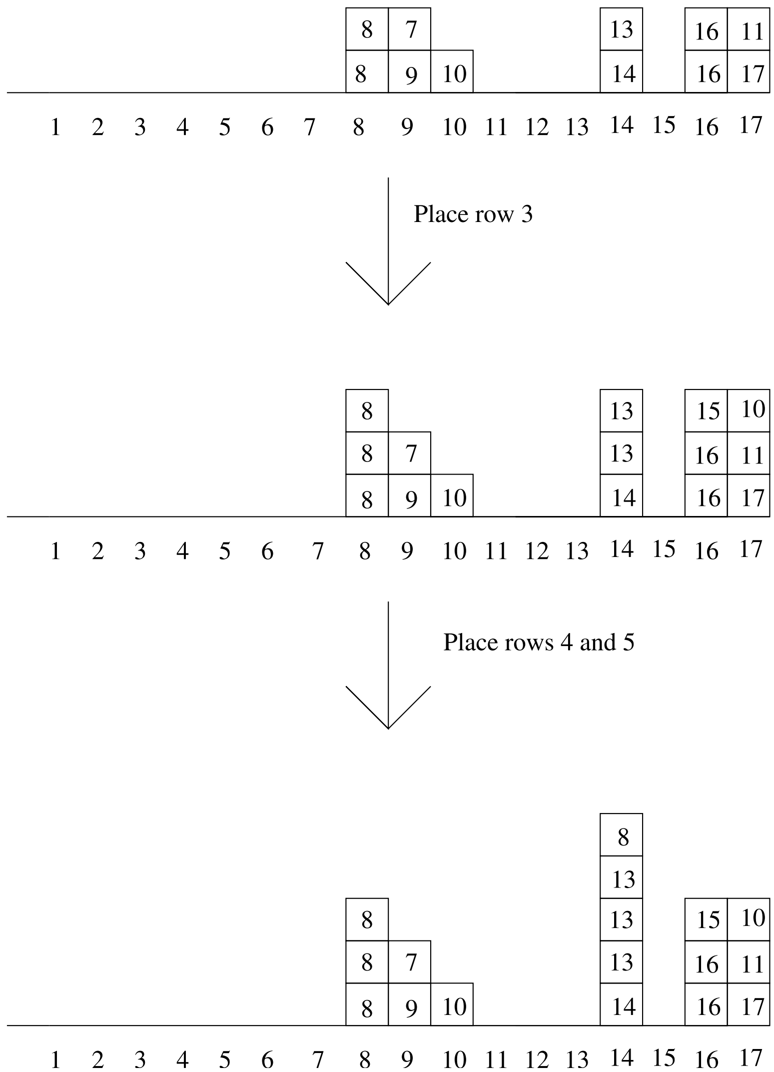
Placement of the first row

The following figure shows the placement of the second row:



Placement of the second row

Next we demonstrate the process of placing each of the additional 3 rows:



Lemma 1 *Once the entries of row $i-1$ have been placed, the arrangement of the elements of T_i' on top of row $i-1$ forms the i^{th} row of a skyline filling, and this placement procedure is the only placement of the elements of T_i' which yields a skyline filling.*

We must show that the following are true:

1. This process yields a skyline filling.

2. This process is the only way to obtain a skyline filling with the given row entries.

Step 1: This process yields a skyline filling.

Proof. By construction, the filling has no descents. Therefore, we must show that all triples are inversion triples. Recall that we can get an non-inversion triple from either of the following two types of triples of cells:

$$\begin{array}{ccc} \boxed{u} & \cdots & \boxed{w} \\ \boxed{v} & & \end{array} \qquad\qquad \begin{array}{ccc} & & \boxed{v} \\ \boxed{u} & \cdots & \boxed{w} \end{array}$$

Type A Type B

In type A, the column containing u and v is weakly taller than the column containing w while in type B, the column containing v and w is taller. After standardization, a type A non-inversion implies a clockwise ordering when the cells are ordered from smallest to largest, and a type B non-nversion implies a counter-clockwise ordering when the cells are ordered from smallest to largest.

First check for type A non-inversion triples. They must look like the cell configuration pictured below, where the column containing u and v has height greater than or equal to the height of the column containing w and t:

$$\begin{array}{ccc} \boxed{u} & \cdots & \boxed{w} \\ \boxed{v} & \cdots & \boxed{t} \end{array}$$

Here, we must have $u \leq v$. Therefore, to get a non-inversion triple, we must have $u \leq w \leq v$. Since the elements of T_i' are all distinct, this implies that $u < w$. But then w would have been placed before u. Since $w \leq v$, w would have been placed on top of v or on top of some entry to the left of v. So this configuration would not happen. Therefore, there are no type A non-inversion triples.

Next check for type B non-inversion triples. This can occur in two ways. Either the left cell in the triple has a cell on top of it (Case 1) or the left cell does not have a cell on top of it (Case 2).

$$\begin{array}{ccc} \boxed{w} & \cdots & \boxed{y} \\ \boxed{x} & \cdots & \boxed{z} \end{array} \qquad\qquad \begin{array}{ccc} & & \boxed{y} \\ \boxed{x} & \cdots & \boxed{z} \end{array}$$

Case 1 Case 2

We know that $y \leq z$. Thus, to get an non-inversion triple, we must have $y \leq x \leq z$. Standardization implies that we may assume $y < x < z$.

In Case 1, since y is less than or equal to x and z, y could be placed on top of either. Since w was placed on top of x, w must have been placed before y. So w must be greater than y.

In order for this triple to be a non-inversion triple of type B, the column containing z and y must be taller when we've completed our composition filling. If the w, x column terminated on the next row, a cell, c, would be added on top of y while nothing was added on top of w. However, since $w > y$, c would not be placed on top of y because w is a cell farther to the left on top of which c could be placed without creating any descents. So the column containing w and x can not terminate on the very next row. However, if it does not terminate, an entry must be placed on top of w and an entry must be placed on top of y. Since $w > y$, the entry on top of w will be greater than the entry on top of y. So we will be dealing with the same situation in row $i + 1$ as we dealt with in row i. Thus the column containing w and x cannot terminate before the column containing y and z. Therefore Case 1 cannot happen.

In Case 2, since y is less than x and z, we could have placed y on top of x instead. Placing y on top of z means y was not placed as far to the left as it should have been. Thus case 2 cannot happen.

Therefore, our process yields a filling with no descents such that all triples are inversion triples. We conclude that our process yields a skyline filling.

□

Step 2: This process is the only way to obtain a skyline filling from the given row entries $\{T_i'\}$.

Proof. Assume there is another way to get a skyline filling from the same row entries $\{T'_i\}$. Denote by K the skyline filling created by the process above. Denote by K' a different skyline filling whose rows contain the same entries (T'_i) as the rows of K but in a different order.

Find the lowest row i of K' whose ordering is not equal to the ordering of row i of K. Consider the largest element of T'_i whose placement in K' does not agree with its placement in K. Call this element u. In K, u was placed in the left-most possible position. Therefore, u must lie in a position further to the right in K'.

Say u lies above the entry v in K and above w in K'. Then this part of the skyline filling looks like the picture below, where x and y might be empty cells:

K K'

Since u is the largest cell of K' to lie in a different place from where it lies in K, y must be less than u. If the column in K' containing y and v were taller than the column containing u and w, then the triple y, u, v would be a non-inversion triple of type A. So the column containing u and w must be taller than the column containing y and v. Then $w < v$, since otherwise u, v, w would be a non-inversion triple of type B.

The 0^{th} row of K' contains the numbers from 1 to n, in increasing order. Therefore, at some row below the row containing v and w, the entry, d, in the column containing v is less than the entry, e, in the column containing w. Find the highest row where this occurs below the row containing v and w. Let f be the entry above d and g be the entry above e. (See Figure 1, below.) Then $g < f < d < e$. So g, d, e is a non-inversion triple of type B.

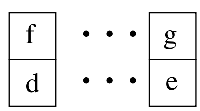


Figure 1

Therefore, regardless of which column is taller, K' contains at least one non-inversion triple. So K' is not a skyline filling. Therefore the skyline filling obtained through the process described at the beginning of this section yields the only possible skyline filling with the given row entries.

□

4 Map from Skyline Fillings to SSYTs

Begin with an arbitrary skyline filling. Select all the entries in the bottom row. Arrange them in a vertical column, sorted into descending order up columns. Then select the entries in the second row, and arrange them in a column immediately to the right of the first column, again in decreasing order. Continue in this manner until there are no more rows left in the composition filling. The shape one gets is clearly a Young diagram, since each column of this figure has height less than or equal to the height of the column to its left.

Lemma 2 *The entries in a column of the Young diagram filling are strictly decreasing as one travels up the column.*

Proof. It is clear by the way we ordered the columns that they are weakly decreasing as one travels up the column. It remains to show that there cannot be two equal entries in a given column. If there were, then in the composition filling there would be two equal entries in a row. (See Figure 2, below).

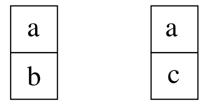


Figure 2

If the column containing b is taller than or equal to the column containing c, then the triple a, a, b would be a type A non-inversion triple. Thus the column containing c has height greater than the column containing b.

If $b \leq c$, then the triple a, b, c (where a is the entry on top of c) is a type B non-inversion triple. So $b > c$. The argument at the end of section 3 demonstrates that this also leads to a non-inversion triple found in lower rows.

□

We just proved that no two entries in the same row of a composition filling can be equal. This implies that all the entries in a column of our young diagram filling must be distinct.

Lemma 3 *Each entry in the Young diagram filling is less than or equal to the entry immediately to its left.*

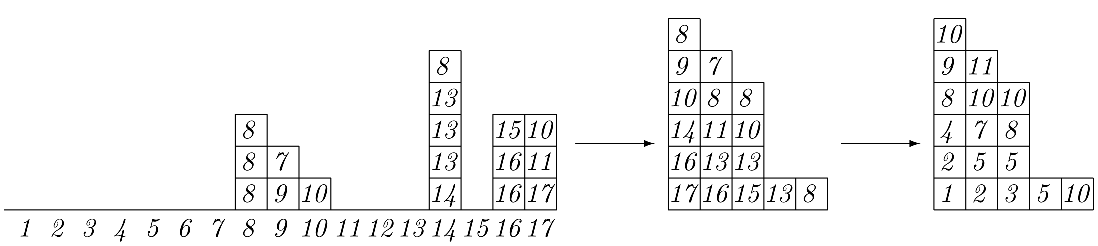
Proof. The entry, j, at height α in the i^{th} column of the Young diagram filling is the α^{th} largest entry in the i^{th} row of the skyline filling. If this value is greater than the value to its left in the Young diagram filling, at most $\alpha - 1$ entries on the $(i-1)^{st}$ row are greater than or equal to j while α entries on the i^{th} row are greater than or equal to j. Then the pigeon-hole principle tells us that at least one entry on the i^{th} row is greater than the entry below it. But then we have a descent and therefore our composition filling is not a skyline filling. Thus, we have a contradiction. So each entry must be less than or equal to the entry immediately to its left.

□

The cells in the Young diagram filling are labelled by the members of the multiset $\{1^{a_1}, 2^{a_2}, ..., k^{a_k}\}$. The total number of cells in the skyline filling, n, is equal to the total number of cells in the Young diagram. Map the value in each cell to a new value by sending α to $n - \alpha + 1$. Before the mapping, the labels were weakly decreasing by column and strictly decreasing by row. Since the map reverses the orders of the labels while preserving the fact that no repeated entries occur within a column, the labels are now weakly increasing by row and strictly increasing by column. Thus, we now have a Semi-Standard Young Tableau.

Say two different composition fillings yield the same SSYT. Then these two composition fillings would have the same set of entries on each row. But we saw in section 3 that once we know the entries on a row, the placement of those entries in a skyline filling is unique. So these two skyline fillings are identical. Thus, our map is injective.

Example 2 *Below we demonstrate the mapping first from a semi-standard skyline filling to a Young diagram filling and then to a semi-standard Young tableau:*



5 The two maps defined above are inverses

Looking back at the two examples, one sees that in this particular case, the two maps are inverses. In fact, this is true in general.

Lemma 4 *The two maps defined in sections 3 and 4 are inverses.*

Proof.
To see this, begin with the map from a SSYT to a skyline filling. This map sends the numbers in a given column to the corresponding row, changing the numbers by mapping α to $n-\alpha+1$, where n is the number which the shape of the SSYT partitions. Then, when we map this skyline filling back to an SSYT, first we take the numbers in each row and place them in the corresponding column in decreasing order. Then we send α to $n-\alpha+1$, which inverts the mapping we did in the first step. So we have the same numbers in each column, arranged in increasing order. Therefore we have the same SSYT that we began with.

Going the other way, we begin with a skyline filling and map each row to a column with the same numbers in decreasing order. We change this shape to a SSYT by mapping α to $n-\alpha+1$. When we map back to a skyline filling, we first send α to $n-\alpha+1$, which inverts the mapping. Next, we enter each column into its corresponding row via the unique map defined in section 3. Since this is the only skyline filling with these particular entries in each row, this is the skyline filling we began with.

□

Thus, the two injective maps are inverses and form a bijection between skyline fillings of rearrangements of μ and SSYT of shape μ. Since the s_λ are symmetric, the number of SSYT of weight $x_1^{a_1}x_2^{a_2}...x_n^{a_n}$ is equal to the number of SSYT of weight $x_{n-1}^{a_1}x_{n-2}^{a_2}...x_1^{a_n}$. Since our map sends each SSYT of weight $x_1^{a_1}x_2^{a_2}...x_n^{a_n}$ to a skyline filling of weight $x_{n-1}^{a_1}x_{n-2}^{a_2}...x_1^{a_n}$, the coefficient of $\prod_{i=1}^n x_i^{\alpha_i}$ in $s_\lambda(\mathrm{x})$ is equal to the coefficient of $\prod_{i=1}^n x_i^{\alpha_i}$ in $\sum_{\lambda'} NS_{\lambda'}(x)$, for all possible multisets $\{\alpha_1,...\alpha_n\}$ with $0 \le \alpha_i \le n$, $\forall i$, and $\sum_{i=1}^n \alpha_i = n$.

This proves that the sum of the non-symmetric Schur functions over all rearrangements of a partition, μ, is equal to the Schur function s_μ.

6 A Basis For the Algebra of degree n Polynomials in n variables

Several other bases for symmetric functions have non-symmetric analogues. For instance, the non-smmyetric 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 non-symmetric 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 non-symmetric monomials, so the non-symmetric monomials form a basis for the algebra of polynomials of degree n in n variables.

Definition 2 *The* **reverse dominance order** *on compositions is defined as follows:*
$\mu \le \gamma \iff \sum_{i=k}^n \mu_i \le \sum_{i=k}^n \gamma_i$ *for* $1 \le i \le n$.

A semi-standard skyline filling is said to have type α if it contains α_i copies of the number i for each i. If γ and α are compositions of n into n parts, let $NK_{\gamma,\alpha}$ denote the number of semi-standard skyline fillings of shape γ and type α. $NK_{\gamma,\alpha}$ is called a *non-symmetric Kostka number.* The ordinary Kostka numbers are obtained as a sum of non-symmetric Kostka numbers: $K_{\lambda,\alpha} = \sum NK_{\gamma,\alpha}$, where the sum is over all rearrangements γ of λ.

Theorem 2 *Suppose that* μ *and* γ *are both compositions of* n *into* n *parts and* $NK_{\mu,\gamma} \ne 0$. *Then* $\mu \ge \gamma$ *in the dominance order. Moreover,* $NK_{\mu,\mu} = 1$.

Proof. Assume that $NK_{\mu,\gamma} \neq 0$. By definition, there exists a semi-standard skyline filling of shape μ and type γ. Assume that a part k appears in one of the first $k-1$ columns. Then this k column would contain a descent, since there is an entry less than k in the column at a lower position than k. Therefore, the parts $k, k+1, ..., n$ all appear in the last $n-k+1$ columns. So $\mu_k + \mu_{k+1} + ... + \mu_n \geq \gamma_k + \gamma_{k+1} + ... + \gamma_n$ for each k, as desired. Moreover, if $\mu = \gamma$, then the i^{th} column must contain only entries with value i, so $NK_{\mu,\mu} = 1$. □

Corollary 1 *The non-symmetric Schur functions form a basis for the algebra of polynomials of degree n in n variables.*

Proof. Theorem 2 is equivalent to the assertion that the transition matrix from the non-symmetric Schur functions to the non-symmetric monomials (with respect to the reverse dominance order) is upper triangular with 1's on the main diagonal. Since this matrix is invertible, the non-symmetric Schur functions of degree n are a basis for polynomials of degree n in n variables. □

References

[Hag04a] J. Haglund. A combinatorial model for the Macdonald polynomials. *Proc. Nat. Acad. Sci. U.S.A.*, 101:16127–16131, 2004.

[Hag04b] J. Haglund. A proof of the q,t-Schroder conjecture. *Internat. Math. Res. Notices*, 11:525–560, 2004.

[HHL$^+$] J. Haglund, M. Haiman, N. Loehr, J. B. Remmel, and A. Ulyanov. A combinatorial formula for the character of the diagonal coinvariants. *Duke J. Math., to appear.* arXiv:math.CO/0310424.

[HHL04] J. Haglund, M. Haiman, and N. Loehr. A combinatorial formula for Macdonald polynomials. *arXiv:math.CO/0310424*, 2004.

[HHL05] J. Haglund, M. Haiman, and N. Loehr. A combinatorial formula for Nonsymmetric Macdonald polynomials. *Jour. Amer. Math. Soc. (to appear)*, 2005.

[Mac88] I. G. Macdonald. A new class of symmetric polynomials. *Actes du* 20^e *Séminaire Lotharingien, Publ. Inst. Rech. Math. Av.*, 372, 1988.

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

Random Strict Partitions and Pfaffian

Sho Matsumoto

Abstract

The shifted Schur measure is a measure on the set of all strict partitions, which is defined by Schur Q-functions. We study distributions of parts of strict partitions. We prove that the correlation function of the measure is given by a Pfaffian in two ways. In the first way, we use commutation relations of operators on an exterior algebra. In the second way, the idea of random point processes is used. As an application, we prove that limit distributions of parts of random strict partitions with respect to specialized shifted Schur measures are given by the Airy ensemble.

2000 Mathematics Subject Classification : 60C05, 05E05,

Key Words : Schur Q-function, strict partition, shifted Schur measure, Pfaffian, Pfaffian point process.

1 Introduction

For a partition (or equivalently, a Young diagram) λ, we denote by f^λ the number of standard Young tableaux of shape λ. The Plancherel measure for the symmetric group $\mathfrak{S}_N$ assigns to each partition λ of N the probability

$$\mathrm{P}_{\mathrm{Plan},N}(\lambda) = \frac{(f^\lambda)^2}{N!}.$$

It is closely related to Ulam's problem for the length $\ell(\pi)$ of the longest increasing subsequence of a random permutation π with respect to the uniform measure $\mathrm{P}_{\mathrm{uniform},N}$ on $\mathfrak{S}_N$, see the survey [AD]. Namely, via the Robinson correspondence (see e.g. [S]), we have

$$\mathrm{P}_{\mathrm{uniform},N}(\{\pi \in \mathfrak{S}_N \ : \ \ell(\pi) = h\}) = \mathrm{P}_{\mathrm{Plan},N}(\{\lambda \in \mathcal{P}_N \ : \ \lambda_1 = h\}),$$

where $\mathcal{P}_N$ is the set of all partitions of N and λ_1 is the largest part of a partition $\lambda = (\lambda_1, \lambda_2, \dots)$. To other parts λ_j, we can also give combinatorial sense related with increasing sequences.

In order to see distributions of λ_j, the correlation function of the poissonized Plancherel measure is calculated in [BOO]. This correlation function is expressed as a determinant. Via the determinantal expression, it is proved in [BOO, Jo2, O1] (see also [BDJ, Jo1, O3, R]) that, as $N \to \infty$, limit distributions of scaled λ_j are described as the Airy ensemble (see §5). In particular, the limit distribution of λ_1 is expressed as the Tracy-Widom distribution F_{GUE} for the Gaussian unitary ensemble.

The Schur measure is the measure on all partitions, which gives each partition λ the probability $s_\lambda(\mathbf{x})s_\lambda(\mathbf{y})$. Here $s_\lambda(\mathbf{x})$ (resp. $s_\lambda(\mathbf{y})$) is the Schur function corresponding to λ in variables $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots)$ (resp. $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots)$). The poissonized Plancherel measure is obtained as a specialization of the Schur measure. In [O2], the correlation function of the Schur measure is calculated and expressed as a determinant as similar as that of the poissonized Plancherel measure is.

In this note, we study random strict partitions. A strict partition is a partition with distinct parts. The shifted Schur measure, introduced in [TW2], is a measure on the set of all strict partitions, which is defined by Schur Q-functions instead of Schur functions (see Definition 1). We prove that, with respect to the shifted Schur measure, the correlation function of random variables $\lambda_1, \lambda_2, \dots$ is expressed by a Pfaffian (see Theorem 1). The Pfaffian of a $2m$ by $2m$ skew-symmetric matrix $B = (b_{ij})_{1 \le i,j \le 2m}$ is defined by

$$\mathrm{pf}(B) = \sum_{\sigma \in \mathfrak{F}_{2m}} \mathrm{sgn}(\sigma) \prod_{j=1}^{m} b_{\sigma(2j-1)\sigma(2j)},$$

where $\mathfrak{F}_{2m}$ is the subset of $\mathfrak{S}_{2m}$ given by

$$\mathfrak{F}_{2m} = \{\sigma = (\sigma(1), \sigma(2), \dots, \sigma(2m)) \in \mathfrak{S}_{2m} \ : \\ \sigma(2j-1) < \sigma(2j) \ (1 \le j \le m), \ \sigma(1) < \sigma(3) < \cdots < \sigma(2m-1)\}.$$

The correlation function is calculated in two ways. First, it is obtained via a representation of a Heisenberg algebra on an exterior algebra. We express the correlation function as a matrix element of an operator on the exterior algebra by using annihilation and creation operators. This idea is used by Okounkov [O2] for the Schur measure. Second, it is a more direct way and we use the idea in [BR]. Since the Schur Q-function has a Pfaffian expression, we can regard the shifted Schur measure as a Pfaffian point process on the set of all non-negative integers. Then our problem is translated into the problem to calculate the inverse of a matrix explicitly. We state the outlines of these proofs in §3 and §4, respectively.

Further, we are interested in limit distributions of λ_j. We see that a limit theorem of the shifted Schur measure is also given by using the Airy ensemble as same as the Schur measure and the Plancherel measure (see Theorem 9). The special case is closely related to the length of the longest ascent pair for a permutation.

Throughout the paper, we denote the set of all positive integers and non-negative integers by $\mathbb{Z}_{>0}$ and $\mathbb{Z}_{\ge 0}$, respectively.

2 Shifted Schur measure

Let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots)$ and $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots)$ be variables. Let $\mathcal{D}$ be the set of all strict partitions;

$$\mathcal{D} = \{\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l) \ : \ l \ge 0, \ \lambda_j \in \mathbb{Z}_{>0} \ (1 \le j \le l), \ \lambda_1 > \lambda_2 > \cdots > \lambda_l > 0\}$$

and $\ell(\lambda)$ be the length of a partition $\lambda \in \mathcal{D}$ (see [Mac]). Set

$$Q_{\mathbf{x}}(z) = \sum_{k=0}^{\infty} Q_{(k)}(\mathbf{x}) z^k = \prod_{j=1}^{\infty} \frac{1 + \mathbf{x}_j z}{1 - \mathbf{x}_j z} = \exp\left(\sum_{n=1,3,5,\dots} \frac{2p_n(\mathbf{x})}{n} z^n \right), \tag{2.1}$$

where $p_n(\mathbf{x}) = \mathbf{x}_1^n + \mathbf{x}_2^n + \cdots$. The Schur Q-function $Q_\lambda(\mathbf{x})$ associated with $\lambda \in \mathcal{D}$ is defined as the coefficient of $z_1^{\lambda_1} \cdots z_n^{\lambda_n}$ in the formal series expansion of

$$Q_{\mathbf{x}}(z_1) \cdots Q_{\mathbf{x}}(z_n) \prod_{1 \le i < j \le n} \frac{z_i - z_j}{z_i + z_j},$$

where $n \geq \ell(\lambda)$ and $\frac{z-w}{z+w} = 1 + 2\sum_{k=1}^{\infty}(-1)^k z^{-k} w^k$. They satisfy the Cauchy-type identity

$$\sum_{\lambda \in \mathcal{D}} 2^{-\ell(\lambda)} Q_\lambda(\mathbf{x}) Q_\lambda(\mathbf{y}) = \prod_{i,j=1}^{\infty} \frac{1+\mathbf{x}_i\mathbf{y}_j}{1-\mathbf{x}_i\mathbf{y}_j}. \tag{2.2}$$

One can define a probability measure on $\mathcal{D}$ via this identity.

Definition 1 ([TW2]). We define the shifted Schur measure by

$$\mathrm{P}_{\mathrm{SS}}(\lambda) = \left(\prod_{i,j=1}^{\infty} \frac{1-\mathbf{x}_i\mathbf{y}_j}{1+\mathbf{x}_i\mathbf{y}_j} \right) 2^{-\ell(\lambda)} Q_\lambda(\mathbf{x}) Q_\lambda(\mathbf{y}) \qquad \text{for } \lambda \in \mathcal{D}.$$

It follows from (2.2) that P_{SS} is a formal probability measure on $\mathcal{D}$; $\sum_{\lambda \in \mathcal{D}} \mathrm{P}_{\mathrm{SS}}(\lambda) = 1$.

Remark 1. The terminology "shifted Schur measure" is used in [TW2] because it is the measure on "shifted" Young diagrams (see §5). However, since the terminology "shifted Schur functions" are already been used in e.g. [OO], the name may confuse.

We are interested in distributions of parts λ_j of random strict partitions $\lambda \in \mathcal{D}$ with respect to the shifted Schur measure P_{SS}. In order to their distributions we study its correlation function. We identify each strict partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l)$ with a finite set $\{\lambda_1, \lambda_2, \dots, \lambda_l\}$ of positive integers. In this sense, we write as $\lambda \supset X$ if the set $\{\lambda_1, \lambda_2, \dots, \lambda_l\}$ contains X for a finite set X of positive integers. Define the correlation function ρ_{SS} by

$$\rho_{\mathrm{SS}}(X) = \mathrm{P}_{\mathrm{SS}}\left(\{\lambda \in \mathcal{D} \ : \ \lambda \supset X\}\right) \qquad \text{for any finite subset } X \subset \mathbb{Z}_{>0}.$$

The following theorem is our main result.

Theorem 1 ([M1, M2]). *For any finite subset $X = \{x_1, x_2, \dots, x_N\} \subset \mathbb{Z}_{>0}$, we have*

$$\rho_{\mathrm{SS}}(X) = \mathrm{pf}(\mathcal{K}(x_i, x_j))_{1 \leq i,j \leq N},$$

where $\mathcal{K}(r,s)$ is a 2 by 2 matrix given by

$$\mathcal{K}(r,s) = \begin{pmatrix} \mathcal{K}_{00}(r,s) & \mathcal{K}_{01}(r,s) \\ \mathcal{K}_{10}(r,s) & \mathcal{K}_{11}(r,s) \end{pmatrix} \qquad \text{for } r,s \in \mathbb{Z}_{>0}.$$

The each entry is given as the coefficient of a formal power series as follows:

$$\mathcal{K}_{00}(r,s) = -\mathcal{K}_{00}(s,r) = \frac{1}{2}[z^r w^s] \frac{Q_{\mathbf{x}}(z) Q_{\mathbf{x}}(w)}{Q_{\mathbf{y}}(z^{-1}) Q_{\mathbf{y}}(w^{-1})} \frac{z-w}{z+w}, \qquad \text{if } r < s,$$

where $\frac{z-w}{z+w} = 1 + 2\sum_{k=1}^{\infty}(-1)^k z^{-k} w^k$ and $[z^r w^s]$ stands for the coefficient of $z^r w^s$, and $\mathcal{K}_{00}(r,r) = 0$.

$$\mathcal{K}_{01}(r,s) = -\mathcal{K}_{10}(s,r) = \frac{1}{2}[z^r w^s] \frac{Q_{\mathbf{x}}(z) Q_{\mathbf{y}}(w)}{Q_{\mathbf{y}}(z^{-1}) Q_{\mathbf{x}}(w^{-1})} \frac{zw+1}{zw-1}, \qquad \text{for any } r \text{ and } s,$$

where $\frac{zw+1}{zw-1} = 1 + 2\sum_{k=1}^{\infty} z^{-k} w^{-k}$.

$$\mathcal{K}_{11}(r,s) = -\mathcal{K}_{11}(s,r) = \frac{1}{2}[z^r w^s] \frac{Q_{\mathbf{y}}(z) Q_{\mathbf{y}}(w)}{Q_{\mathbf{x}}(z^{-1}) Q_{\mathbf{x}}(w^{-1})} \frac{w-z}{w+z}, \qquad \text{if } r < s,$$

where $\frac{w-z}{w+z} = 1 + 2\sum_{k=1}^{\infty}(-1)^k z^k w^{-k}$, and $\mathcal{K}_{11}(r,r) = 0$.

This theorem is a generalization of the Cauchy-type identity (2.2)

$$\sum_{\lambda\in\mathcal{D},\ \lambda\supset X} 2^{-\ell(\lambda)} Q_\lambda(\mathbf{x})Q_\lambda(\mathbf{y}) = \left(\prod_{i,j=1}^{\infty} \frac{1+\mathbf{x}_i\mathbf{y}_j}{1-\mathbf{x}_i\mathbf{y}_j}\right)\cdot \mathrm{pf}(\mathcal{K}(x_i,x_j))_{1\le i,j\le N}, \tag{2.3}$$

where $X=\{x_1,\dots,x_N\}\subset\mathbb{Z}_{>0}$. The formula (2.3) is reduced to (2.2) if $X=\emptyset$.

Example 1. Let $X=\{x\}$. Then Theorem 1 says

$$\sum_{\lambda\in\mathcal{D},\ \lambda\ni x} 2^{-\ell(\lambda)} Q_\lambda(\mathbf{x})Q_\lambda(\mathbf{y}) = \left(\prod_{i,j=1}^{\infty} \frac{1+\mathbf{x}_i\mathbf{y}_j}{1-\mathbf{x}_i\mathbf{y}_j}\right)\cdot \mathcal{K}_{01}(x,x)$$

summed over all strict partitions containing x, where $\mathcal{K}_{01}(r,s)$ is given by

$$\mathcal{K}_{01}(r,s)=\frac{1}{2}\mathcal{J}_r(\mathbf{x},\mathbf{y})\mathcal{J}_s(\mathbf{y},\mathbf{x})+\sum_{n=1}^{\infty}\mathcal{J}_{r+n}(\mathbf{x},\mathbf{y})\mathcal{J}_{s+n}(\mathbf{y},\mathbf{x})$$

with

$$\mathcal{J}_r(\mathbf{x},\mathbf{y})=[z^r]\frac{Q_{\mathbf{x}}(z)}{Q_{\mathbf{y}}(z^{-1})}=\sum_{k=0}^{\infty}(-1)^k Q_{(r+k)}(\mathbf{x})Q_{(k)}(\mathbf{y}).$$

As a corollary of Theorem 1, we can obtain the distribution of the largest part λ_1 of $\lambda\in\mathcal{D}$.

Corollary 2. *For a positive integer h, we have*

$$\mathrm{P}_{\mathrm{SS}}(\lambda_1<h)=\sum_{\lambda_1<h}\mathrm{P}_{\mathrm{SS}}(\lambda)=\mathrm{pf}(J-\mathcal{K})_{\{h,h+1,\dots\}}.$$

Here $\mathrm{pf}(J-\mathcal{K})_{\{h,h+1,\dots,\}}$ *is the Fredholm pfaffian for the kernel $\mathcal{K}$ on $\{h,h+1,\dots\}$ defined by*

$$\mathrm{pf}(J-\mathcal{K})_{\{h,h+1,\dots,\}}=\sum_{n=0}^{\infty}(-1)^n\sum_{h\le x_1<x_2<\cdots<x_n}\mathrm{pf}(\mathcal{K}(x_i,x_j))_{1\le i,j\le n}.$$

3 First Proof of Theorem 1

We state the outline of the first proof of Theorem 1, obtained in [M1].

Let V be a module on $\mathbb{C}[\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{y}_1,\mathbf{y}_2,\dots]$ spanned by $\boldsymbol{e}_k$ $(k=1,2,\dots)$. The exterior algebra $\bigwedge V$ is spanned by vectors

$$\boldsymbol{v}_\lambda=\boldsymbol{e}_{\lambda_1}\wedge\boldsymbol{e}_{\lambda_2}\wedge\cdots\wedge\boldsymbol{e}_{\lambda_\ell},$$

where $\lambda=(\lambda_1,\dots,\lambda_\ell)\in\mathcal{D}$ $(\lambda_1>\cdots>\lambda_\ell\ge 1)$. In particular, we have $\boldsymbol{v}_\emptyset=1$ for the empty partition $\emptyset$. We give $\bigwedge V$ the inner product

$$\langle\boldsymbol{v}_\lambda,\boldsymbol{v}_\mu\rangle=\delta_{\lambda,\mu}2^{-\ell(\lambda)}\qquad\text{for }\lambda,\mu\in\mathcal{D}.$$

Putting $\boldsymbol{e}_k^\vee=2\boldsymbol{e}_k$ and $\boldsymbol{v}_\lambda^\vee=\boldsymbol{e}_{\lambda_1}^\vee\wedge\cdots\wedge\boldsymbol{e}_{\lambda_\ell}^\vee=2^\ell\boldsymbol{v}_\lambda$, the bases $(\boldsymbol{v}_\lambda)_{\lambda\in\mathcal{D}}$ and $(\boldsymbol{v}_\lambda^\vee)_{\lambda\in\mathcal{D}}$ are dual to each other. We define the operator ψ_k on $\bigwedge V$ by

$$\psi_k\boldsymbol{v}_\lambda=\boldsymbol{e}_k\wedge\boldsymbol{v}_\lambda$$

for $k \geq 1$ and let ψ_k^* be the adjoint operator of ψ_k with respect to the inner product defined above. The operator ψ_k^* is then explicitly given by

$$\psi_k^* \boldsymbol{v}_\lambda = \sum_{i=1}^{\ell(\lambda)} \frac{(-1)^{i-1}}{2} \delta_{k,\lambda_i} \boldsymbol{e}_{\lambda_1} \wedge \cdots \wedge \widehat{\boldsymbol{e}_{\lambda_i}} \wedge \cdots \wedge \boldsymbol{e}_{\lambda_\ell},$$

where $\widehat{\boldsymbol{e}_k}$ means that $\boldsymbol{e}_k$ is omitted. Also we define the self-adjoint operator S by $S\boldsymbol{v}_\lambda = (-1)^{\ell(\lambda)}\boldsymbol{v}_\lambda$ for any $\lambda \in \mathcal{D}$. Put

$$\tilde{\psi}_k = \begin{cases} \psi_k, & \text{for} \quad k \geq 1, \\ S/2, & \text{for} \quad k = 0, \\ (-1)^k \psi_{-k}^*, & \text{for} \quad k \leq -1. \end{cases}$$

Then they satisfy the following commutation relation

$$\tilde{\psi}_i \tilde{\psi}_j + \tilde{\psi}_j \tilde{\psi}_i = \frac{(-1)^{|i|}}{2} \delta_{i+j,0} \qquad \text{for } i, j \in \mathbb{Z}$$

and give a projection

$$2\psi_k \psi_k^* \boldsymbol{v}_\lambda = \begin{cases} \boldsymbol{v}_\lambda, & \text{if } k \in \{\lambda_1, \dots, \lambda_l\}, \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

for $k \geq 1$ and $\lambda = (\lambda_1, \dots, \lambda_l) \in \mathcal{D}$. Therefore $(\prod_{k \in X} 2\psi_k \psi_k^*)\boldsymbol{v}_\lambda$ is equal to $\boldsymbol{v}_\lambda$ if $X \subset \lambda$, or to 0 otherwise.

Define $\alpha_n = \sum_{k \in \mathbb{Z}} (-1)^k \tilde{\psi}_{k-n} \tilde{\psi}_{-k}$ for any odd integer $n \in 2\mathbb{Z}+1$. Then they satisfy the Heisenberg relation

$$[\alpha_n, \alpha_m] = \alpha_n \alpha_m - \alpha_m \alpha_n = \frac{n}{2} \delta_{n+m,0}. \tag{3.2}$$

Put

$$\Gamma_\pm(\mathbf{x}) = \exp\left(\sum_{n=1,3,5,\dots} \frac{2p_n(\mathbf{x})}{n} \alpha_{\pm n} \right).$$

It is not hard to obtain the following lemma from commutation relations above.

Lemma 3. *We have*

$$\Gamma_+(\mathbf{x})\boldsymbol{v}_\emptyset = \boldsymbol{v}_\emptyset, \quad (\Gamma_\pm(\mathbf{x}))^* = \Gamma_\mp(\mathbf{x}), \quad \Gamma_+(\mathbf{x})\Gamma_-(\mathbf{y}) = \left(\prod_{i,j=1}^\infty \frac{1+\mathbf{x}_i\mathbf{y}_j}{1-\mathbf{x}_j\mathbf{y}_j} \right) \Gamma_-(\mathbf{y})\Gamma_+(\mathbf{x}).$$

Further, when we put $\psi(z) = \sum_{k \in \mathbb{Z}} z^k \tilde{\psi}_k$, *we have*

$$[\alpha_n, \psi(z)] = z^n \psi(z), \quad \langle 4\psi(z)\psi(w)\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset \rangle = \frac{z-w}{z+w}, \quad \Gamma_\pm(\mathbf{x})\psi(z) = Q_{\mathbf{x}}(z^{\pm 1})\psi(z)\Gamma_\pm(\mathbf{x}).$$

Using the lemma, we can express the Schur Q-function $Q_\lambda(\mathbf{x})$ as the matrix element of $\Gamma_-(\mathbf{x})$ as follows.

Proposition 4. *For each $\lambda \in \mathcal{D}$, we have $\langle \Gamma_-(\mathbf{x})\boldsymbol{v}_\emptyset, \boldsymbol{v}_\lambda^\vee \rangle = Q_\lambda(\mathbf{x})$.*

From (3.1), Lemma 3, and Proposition 4, the correlation function ρ_{SS} of the shifted Schur measure is expressed as

$$\begin{aligned}\rho_{\mathrm{SS}}(X) &= \prod_{i,j=1}^{\infty}\left(\frac{1-\mathbf{x}_i\mathbf{y}_j}{1+\mathbf{x}_i\mathbf{y}_j}\right)\sum_{\lambda\supset X} 2^{-\ell(\lambda)} Q_\lambda(\mathbf{x})Q_\lambda(\mathbf{y}) \\ &= \prod_{i,j=1}^{\infty}\left(\frac{1-\mathbf{x}_i\mathbf{y}_j}{1+\mathbf{x}_i\mathbf{y}_j}\right)\left\langle \Gamma_+(\mathbf{x})\left(\prod_{k\in X} 2\psi_k\psi_k^*\right)\Gamma_-(\mathbf{y})\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset \right\rangle \\ &= \left\langle \left(\prod_{k\in X} 2\Psi_k\Psi_k^*\right)\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset \right\rangle\end{aligned}$$

with $\Psi_k = G\psi_k G^{-1}$, $\Psi_k^* = G\psi_k^* G^{-1}$, and $G = \Gamma_+(\mathbf{x})\Gamma_-(\mathbf{y})^{-1}$. Now we obtain

Proposition 5. *We have $\rho_{\mathrm{SS}}(X) = \mathrm{pf}(\widetilde{\mathcal{K}}(x,y))_{x,y\in X}$ with*

$$\widetilde{\mathcal{K}}(x,y) = \begin{pmatrix} \widetilde{\mathcal{K}}_{00}(x,y) & \widetilde{\mathcal{K}}_{01}(x,y) \\ \widetilde{\mathcal{K}}_{10}(x,y) & \widetilde{\mathcal{K}}_{11}(x,y) \end{pmatrix}.$$

Here we put $\widetilde{\mathcal{K}}_{00}(x,y) = \langle 2\Psi_x\Psi_y\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset\rangle$, $\widetilde{\mathcal{K}}_{01}(x,y) = -\widetilde{\mathcal{K}}_{10}(y,x) = \langle 2\Psi_x\Psi_y^*\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset\rangle$, and $\widetilde{\mathcal{K}}_{11}(x,y) = \langle 2\Psi_x^*\Psi_y^*\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset\rangle$ for $x,y \in \mathbb{Z}_{>0}$.

Finally, by Lemma 3, we can prove $\mathcal{K}(x,y) = \widetilde{\mathcal{K}}(x,y)$ for all $x,y \in \mathbb{Z}_{>0}$. For example, putting $\Psi(z) = G\psi(z)G^{-1}$,

$$\widetilde{\mathcal{K}}_{00}(x,y) = [z^x w^y]\langle 2\Psi(z)\Psi(w)\boldsymbol{v}_\emptyset, \boldsymbol{v}_\emptyset\rangle = [z^x w^y]\frac{1}{2}\frac{Q_{\mathbf{x}}(z)Q_{\mathbf{x}}(w)}{Q_{\mathbf{y}}(z^{-1})Q_{\mathbf{y}}(w^{-1})}\frac{z-w}{z+w} = \mathcal{K}_{00}(x,y).$$

It completes the proof of Theorem 1.

Remark 2. The discussion of this section is very related to the fermion Fock space and vertex operators. Proposition 5 is essentially obtained from the fermion Wick formula, see e.g. [Ji] and [Y].

4 Second Proof of Theorem 1

In this section, we give another proof of Theorem 1, which is obtained in [M2], via a random point process. We recall some fundamental facts of the Pfaffian point process formulated in [BR]. Let $\mathfrak{X}$ be a countable set. Let L be a map

$$L : \mathfrak{X}\times\mathfrak{X} \to \mathfrak{gl}_2(\mathbb{C})\ ;\ (x,y) \mapsto L(x,y) = \begin{pmatrix} L_{00}(x,y) & L_{01}(x,y) \\ L_{10}(x,y) & L_{11}(x,y) \end{pmatrix}$$

such that $L_{ij}(x,y) = -L_{ji}(y,x)$ for any $i,j \in \{0,1\}$ and $x,y \in \mathfrak{X}$. Such L is called a skew-symmetric matrix kernel on $\mathfrak{X}$, see [R, So]. We regard the map L as an operator on the Hilbert space $\ell^2(\mathfrak{X}) \oplus \ell^2(\mathfrak{X})$. Then L is a matrix whose blocks are indexed by elements in $\mathfrak{X}\times\mathfrak{X}$. For any finite subset $X = \{x_1, \cdots, x_n\} \subset \mathfrak{X}$, we denote by $L[X]$ the $2n$ by $2n$ skew-symmetric matrix $(L(x_i, x_j))_{1 \le i,j \le n}$. Let J be the skew-symmetric matrix kernel determined by

$$J(x,y) = \delta_{x,y} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Let $\mathfrak{P}(\mathfrak{X})$ be the set of all finite subsets in $\mathfrak{X}$. We define the Pfaffian point process π_L on $\mathfrak{X}$ determined by L as the probability measure on $\mathfrak{P}(\mathfrak{X})$ given by

$$\pi(X) = \pi_L(X) = \frac{\operatorname{pf}(L[X])}{\operatorname{pf}(J+L)} \qquad \text{for } X \in \mathfrak{P}(\mathfrak{X}).$$

Then its correlation function is expressed as $\rho_L(X) = \sum_{Y \in \mathfrak{P}(\mathfrak{X}),\ Y \supset X} \pi(Y) = \operatorname{pf}(K[X])$, where $K = J + (J+L)^{-1}$.

More generally, let $\mathfrak{Y}$ be a subset in $\mathfrak{X}$ such that $\mathfrak{Y}^c = \mathfrak{X} \setminus \mathfrak{Y}$ is finite. Then we can define the conditional Pfaffian point process on $\mathfrak{Y}$ by

$$\pi_{L,\mathfrak{Y}}(X) = \frac{\operatorname{pf}(L[X \cup \mathfrak{Y}^c])}{\operatorname{pf}(J[\mathfrak{Y}] + L)} \qquad \text{for } X \in \mathfrak{P}(\mathfrak{Y}). \tag{4.1}$$

Here we identify $J[\mathfrak{Y}]$ with the block matrix $\begin{pmatrix} J & 0 \\ 0 & 0 \end{pmatrix}$, where the blocks correspond to the partition $\mathfrak{X} = \mathfrak{Y} \sqcup \mathfrak{Y}^c$. The correlation function is given by $\rho_{L,\mathfrak{Y}}(X) = \sum_{Y \in \mathfrak{P}(\mathfrak{Y}),\ Y \supset X} \pi_{L,\mathfrak{Y}}(Y) = \operatorname{pf}(K[X])$ for $X \in \mathfrak{P}(\mathfrak{Y})$, where

$$K = J[\mathfrak{Y}] + (J[\mathfrak{Y}] + L)^{-1}\Big|_{\mathfrak{Y} \times \mathfrak{Y}}. \tag{4.2}$$

The shifted Schur measure is regarded as a conditional Pfaffian point process as follows. We may assume that the number of variables $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{y}_1, \mathbf{y}_2, \ldots$ is finite. Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$, where n is even. We define the bijective map ϕ from $\mathcal{D}$ to $\mathfrak{P}^{\text{even}}(\mathbb{Z}_{\ge 0}) = \{X \in \mathfrak{P}(\mathbb{Z}_{\ge 0}) \ :\ \#X \text{ is even}\}$ by

$$\phi(\lambda) = \begin{cases} \{\lambda_1, \ldots, \lambda_{\ell(\lambda)}\}, & \text{if } \ell(\lambda) \text{ is even}, \\ \{\lambda_1, \ldots, \lambda_{\ell(\lambda)}, 0\}, & \text{if } \ell(\lambda) \text{ is odd}. \end{cases}$$

The following proposition is proved from the fact that the Schur Q-function is expressed as the quotient of Pfaffians (see [N], also [Mac, III-8]).

Proposition 6. *Define a skew-symmetric matrix kernel L on $\mathfrak{X} = \{1, 2, \ldots, n\} \sqcup \mathfrak{Y}$ by*

$$L = \begin{pmatrix} \mathcal{V} & \mathcal{W}\eta^{-\frac{1}{2}} \\ -\eta^{-\frac{1}{2}}\,{}^t\mathcal{W} & O \end{pmatrix},$$

where $\mathfrak{Y} = \mathbb{Z}_{\ge 0}$, $\mathcal{V} = (\mathcal{V}(i,j))_{1 \le i,j \le n}$ and $\mathcal{W} = (\mathcal{W}(i,r))_{1 \le i \le n,\ r \in \mathbb{Z}_{\ge 0}}$. Their blocks are given by

$$\mathcal{V}(i,j) = \begin{pmatrix} -\frac{\mathbf{x}_i - \mathbf{x}_j}{\mathbf{x}_i + \mathbf{x}_j} & 0 \\ 0 & \frac{\mathbf{y}_i - \mathbf{y}_j}{\mathbf{y}_i + \mathbf{y}_j} \end{pmatrix}, \qquad \mathcal{W}(i,r) = \begin{pmatrix} -\mathbf{x}_i^r & 0 \\ 0 & \mathbf{y}_i^r \end{pmatrix}.$$

Further η is the matrix whose block is given by

$$\eta(r,s) = \delta_{rs} \begin{pmatrix} \eta(r) & 0 \\ 0 & \eta(r) \end{pmatrix} \qquad \text{for } r,s \in \mathbb{Z}_{\geq} 0,$$

where $\eta(r)$ is equal to 1 if $r = 0$, or to $\frac{1}{2}$ if $r \geq 1$. Then the conditional Pfaffian point process on $\mathfrak{Y}$ is agree with the shifted Schur measure on $\mathcal{D}$ via the bijection ϕ. Namely,

$$\pi_{L,\mathfrak{Y}}(\phi(\lambda)) = \frac{\mathrm{pf}(L[\{1,\dots,n\} \sqcup \phi(\lambda)]}{\mathrm{pf}(J[\mathfrak{Y}] + L)} = \mathrm{P_{SS}}(\lambda)$$

for any $\lambda \in \mathcal{D}$.

By (4.2), we have to obtain the explicit expression of the correlation kernel $K = J[\mathfrak{Y}] + (J[\mathfrak{Y}] + L)^{-1}\Big|_{\mathfrak{Y}\times\mathfrak{Y}}$. For that purpose, we employ the following lemma.

Lemma 7. *We have*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} -\mathcal{M}^{-1} & \mathcal{M}^{-1}BD^{-1} \\ D^{-1}C\mathcal{M}^{-1} & D^{-1} - D^{-1}C\mathcal{M}^{-1}BD^{-1} \end{pmatrix},$$

where $\mathcal{M} = BD^{-1}C - A$, if D and $\mathcal{M}$ are invertible.

By this lemma, the kernel $K = J[\mathfrak{Y}] + (J[\mathfrak{Y}] + L)^{-1}\Big|_{\mathfrak{Y}\times\mathfrak{Y}}$ is equal to $J[\mathfrak{Y}]\eta^{-\frac{1}{2}}{}^t\mathcal{W}\mathcal{M}^{-1}\mathcal{W}\eta^{-\frac{1}{2}}J[\mathfrak{Y}]$, with $\mathcal{M} = \mathcal{W}\eta^{-\frac{1}{2}}J[\mathfrak{Y}]\eta^{-\frac{1}{2}}{}^t\mathcal{W} - \mathcal{V}$. We may replace K with $-\eta^{-\frac{1}{2}}{}^t\mathcal{W}\mathcal{M}^{-1}\mathcal{W}\eta^{-\frac{1}{2}}$ because $\mathrm{pf}(-JBJ) = \mathrm{pf}(B)$ for a skew-symmetric matrix B. The explicit expression of entries of the inverse $\mathcal{M}^{-1}$ is obtained by a linear algebraic discussion.

Proposition 8. *Write the skew-symmetric matrix kernel $\mathcal{M}^{-1}$ on $\{1,2,\dots,n\}$ as*

$$\mathcal{M}^{-1}(k,l) = \begin{pmatrix} \mathcal{M}^{-1}_{00}(k,l) & \mathcal{M}^{-1}_{01}(k,l) \\ \mathcal{M}^{-1}_{10}(k,l) & \mathcal{M}^{-1}_{11}(k,l) \end{pmatrix} \qquad \text{for } 1 \leq k,l \leq n.$$

Then we have

$$\begin{aligned}
\mathcal{M}^{-1}_{00}(k,l) &= \prod_{j=1}^{n} \left(\frac{1-\mathbf{x}_k\mathbf{y}_j}{1+\mathbf{x}_k\mathbf{y}_j} \frac{1-\mathbf{x}_l\mathbf{y}_j}{1+\mathbf{x}_l\mathbf{y}_j} \right) \prod_{\substack{1\leq i\leq n,\\ i\neq k}} \left(\frac{\mathbf{x}_k+\mathbf{x}_i}{\mathbf{x}_k-\mathbf{x}_i} \right) \prod_{\substack{1\leq j\leq n,\\ j\neq l}} \left(\frac{\mathbf{x}_l+\mathbf{x}_j}{\mathbf{x}_l-\mathbf{x}_j} \right) \frac{\mathbf{x}_k-\mathbf{x}_l}{\mathbf{x}_k+\mathbf{x}_l}; \\
\mathcal{M}^{-1}_{01}(k,l) &= -\mathcal{M}^{-1}_{10}(l,k) \\
&= \prod_{j=1}^{n} \left(\frac{1-\mathbf{x}_k\mathbf{y}_j}{1+\mathbf{x}_k\mathbf{y}_j} \frac{1-\mathbf{x}_j\mathbf{y}_l}{1+\mathbf{x}_j\mathbf{y}_l} \right) \prod_{\substack{1\leq i\leq n,\\ i\neq k}} \left(\frac{\mathbf{x}_k+\mathbf{x}_i}{\mathbf{x}_k-\mathbf{x}_i} \right) \prod_{\substack{1\leq j\leq n,\\ j\neq l}} \left(\frac{\mathbf{y}_l+\mathbf{y}_j}{\mathbf{y}_l-\mathbf{y}_j} \right) \frac{1+\mathbf{x}_k\mathbf{y}_l}{1-\mathbf{x}_k\mathbf{y}_l}; \\
\mathcal{M}^{-1}_{11}(k,l) &= -\prod_{j=1}^{n} \left(\frac{1-\mathbf{x}_j\mathbf{y}_k}{1+\mathbf{x}_j\mathbf{y}_k} \frac{1-\mathbf{x}_j\mathbf{y}_l}{1+\mathbf{x}_j\mathbf{y}_l} \right) \prod_{\substack{1\leq i\leq n,\\ i\neq k}} \left(\frac{\mathbf{y}_k+\mathbf{y}_i}{\mathbf{y}_k-\mathbf{y}_i} \right) \prod_{\substack{1\leq j\leq n,\\ j\neq l}} \left(\frac{\mathbf{y}_l+\mathbf{y}_j}{\mathbf{y}_l-\mathbf{y}_j} \right) \frac{\mathbf{y}_k-\mathbf{y}_l}{\mathbf{y}_k+\mathbf{y}_l}.
\end{aligned}$$

Finally, we must prove $\mathcal{K}(r,s)=K(r,s)$. Now we assume $\mathbf{x}_1,\dots,\mathbf{x}_n,\mathbf{y}_1,\dots,\mathbf{y}_n$ are $2n$ distinct complex numbers in the unit open disc. Then by changing variables and the residue theorem we obtain

$$\begin{aligned}\mathcal{K}_{00}(r,s)=&\frac{1}{2}\frac{1}{(2\pi\sqrt{-1})^2}\iint_{|z|>|w|>1}\frac{Q_{\mathbf{x}}(z)Q_{\mathbf{x}}(w)}{Q_{\mathbf{y}}(z^{-1})Q_{\mathbf{y}}(w^{-1})}\frac{z-w}{z+w}\frac{dzdw}{z^{r+1}w^{s+1}}\\ =&-2\sum_{k,l=1}^{n}\mathbf{x}_k^r\mathbf{x}_l^s\prod_{j=1}^{n}\left(\frac{1-\mathbf{x}_k\mathbf{y}_j}{1+\mathbf{x}_k\mathbf{y}_j}\frac{1-\mathbf{x}_l\mathbf{y}_j}{1+\mathbf{x}_l\mathbf{y}_j}\right)\prod_{\substack{1\le i\le n,\\ i\neq k}}\left(\frac{\mathbf{x}_k+\mathbf{x}_i}{\mathbf{x}_k-\mathbf{x}_i}\right)\prod_{\substack{1\le j\le n,\\ j\neq l}}\left(\frac{\mathbf{x}_l+\mathbf{x}_j}{\mathbf{x}_l-\mathbf{x}_j}\right)\frac{\mathbf{x}_k-\mathbf{x}_l}{\mathbf{x}_k+\mathbf{x}_l}\\ =&-2\sum_{k,l=1}^{n}\mathbf{x}_k^r\mathcal{M}_{00}^{-1}(k,l)\mathbf{x}_l^s=K_{00}(r,s).\end{aligned}$$

Here the contour in the integral above is $\{z\ :\ |z|=r_1\}\times\{w\ :\ |w|=r_2\}$, where $1+\epsilon>r_1>r_2>1$ and $\epsilon>0$ is very small. Similarly, we can prove $\mathcal{K}_{01}(r,s)=K_{01}(r,s)$ and $\mathcal{K}_{11}(r,s)=K_{11}(r,s)$. Though we have assumed that $\mathbf{x}_i$, $\mathbf{y}_j$ belong to the unit open disc, it is in fact unnecessary, see e.g. [BR]. It completes the proof of Theorem 1 again.

5 Limit Distribution

We study limit distributions of λ_j on special conditions.

We consider the random point process on $\mathbb{R}$ (see the Appendix in [BOO]) whose correlation functions $\rho_{\text{Airy}}(X)=\mathrm{P}_{\text{Airy}}(\{Y\subset\mathbb{R}\ :\ \#Y<\infty,\ X\subset Y\})$ are given by $\rho_{\text{Airy}}(X)=\det(\mathcal{K}_{\text{Airy}}(x_i,x_j))_{1\le i,j\le k}$ for any finite subset $X=\{x_1,\dots,x_k\}\subset\mathbb{R}$. Here $\mathcal{K}_{\text{Airy}}$ is the Airy kernel defined by

$$\mathcal{K}_{\text{Airy}}(x,y)=\int_0^\infty \mathrm{Ai}(x+z)\mathrm{Ai}(y+z)\mathrm{d}z,$$

where $\mathrm{Ai}(x)$ is the Airy function

$$\mathrm{Ai}(x)=\frac{1}{2\pi\sqrt{-1}}\int_{\infty e^{-\pi\sqrt{-1}/3}}^{\infty e^{\pi\sqrt{-1}/3}}\exp\left(\frac{z^3}{3}-xz\right)\mathrm{d}z.$$

Let $\zeta^{\text{Ai}}=(\zeta_1^{\text{Ai}}>\zeta_2^{\text{Ai}}>\cdots)\in\mathbb{R}^\infty$ be its random configuration. The random variables ζ_j^{Ai} are called the Airy ensemble. It is known that the Airy ensemble describes the behavior of the scaled eigenvalues of a large hermitian matrix from the Gaussian unitary ensemble, see [TW1].

We consider the specializations of the shifted Schur measure satisfying the following analytic assumptions.

(0) Let $\theta>0$. We specialize power-sum symmetric functions as $p_k(\mathbf{x})=p_k(\mathbf{y})=p_k^\theta$, where $k=1,3,5,\dots$ and $p_k^\theta\in\mathbb{R}$ satisfies $\lim_{\theta\to+\infty}p_k^\theta/\theta=d_k\ge 0$.

(1) There exists an $\epsilon=\epsilon(\theta)>0$ such that the power series $g^\theta(z):=2\sum_{k\ge1:\text{odd}}\frac{p_k^\theta}{k}z^k$ is holomorphic on $|z|<1+\epsilon$.

(2) Put $g(z):=2\sum_{k\ge1:\text{odd}}\frac{d_k}{k}z^k$. Then the series $g(1)=2\sum_{k\ge1:\text{odd}}\frac{d_k}{k}$ converges. Further $g(z)$ can be extended as a holomorphic function around $z=1$.

Then we have the following theorem.

Theorem 9 ([M1]). *Let* $\mathrm{P}_{\mathrm{SS}}^{\theta}$ *be the shifted Schur measure obtained by the specialization such that satisfies assumptions (0), (1) and (2). Put* $b_1 = 2g'(1)$ *and* $b_2 = g'''(1) + 3g''(1) + g'(1)$*. Then, as* $\theta \to \infty$*, random variables*

$$\frac{\lambda_j - b_1\theta}{(b_2\theta)^{1/3}}, \qquad j = 1, 2, \dots,$$

converge to the Airy ensemble, in joint distribution.

This limit theorem is obtained in [TW2] for only λ_1 and a specialization $p_k^{\theta} = \theta\alpha^k$ with $0 < \alpha < 1$. We now give the simplest example of this theorem. Specialize as $p_k(\mathbf{x}) = p_k(\mathbf{y}) = \delta_{1,k}\sqrt{\frac{\xi}{2}}$ with $\xi > 0$. Then the shifted Schur measure provides (see [M1])

$$\mathrm{P}_{\mathrm{PSP}}^{\xi}(\lambda) = e^{-\xi}\xi^{|\lambda|}2^{|\lambda|-\ell(\lambda)}\left(\frac{g^{\lambda}}{|\lambda|!}\right)^2 \qquad \text{for } \lambda \in \mathcal{D}, \tag{5.1}$$

where $|\lambda|$ stands for the weight $|\lambda| = \sum_{j\geq 1}\lambda_j$, and g^{λ} is the number of the standard shifted tableaux of shifted shape $\mathrm{Sh}(\lambda)$. Here $\mathrm{Sh}(\lambda)$ is the shifted Young diagram associated with a strict partition λ, which is obtained by replacing the i-th row to the right by $i-1$ boxes for $i \geq 1$ from the Young diagram corresponding to λ. A standard shifted tableau T of shifted shape $\mathrm{Sh}(\lambda)$ is an assignment of $1, 2, \dots, |\lambda|$ to each box in $\mathrm{Sh}(\lambda)$ such that entries in T are increasing across rows and down columns. For example,

$$\begin{array}{|c|c|c|c|} \hline 1 & 2 & 4 & 6 \\ \hline \multicolumn{1}{c|}{} & 3 & 5 & 8 \\ \cline{2-4} \multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & 7 & \multicolumn{1}{|c}{} \\ \cline{3-3} \end{array}$$

is a standard shifted tableau of shape $\lambda = (4, 3, 1)$. Then we have the

Corollary 10. *With respect to the probability measure* $\mathrm{P}_{\mathrm{PSP}}^{\xi}$ *defined in* (5.1)*, random variables*

$$\frac{\lambda_j - 2\sqrt{2\xi}}{(2\xi)^{\frac{1}{6}}}, \qquad j = 1, 2, \dots,$$

converge to the Airy ensemble as $\xi \to \infty$*. In particular, the limit distribution of* λ_1 *is given by*

$$\lim_{\xi\to\infty}\mathrm{P}_{\mathrm{PSP}}^{\xi}\left(\frac{\lambda_1 - 2\sqrt{2\xi}}{(2\xi)^{\frac{1}{6}}} < s\right) = \mathrm{P}_{\mathrm{Airy}}(\zeta_1 < s) =: F_{\mathrm{GUE}}(s) \qquad \textit{for } s \in \mathbb{R}.$$

The distribution function $F_{\mathrm{GUE}}(s)$ is called the Tracy-Widom distribution, see e.g. [BOO, Jo1, Jo2]. By Corollary 10, we have $\lambda_1 \sim 2\sqrt{2\xi}$ as $\xi \to \infty$. Now the largest part λ_1 describes the length of the longest ascent pair (see [HH]) for random permutations with respect to the uniform measure of symmetric groups. Therefore we give a solution of an analogue of Ulam's problem.

References

[AD] 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.

[BDJ] J. Baik, P. A. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence in a random permutation, J. Amer. Math. Soc. **12** (1999), 1119–1178.

[BOO] A. Borodin, A. Okounkov, and G. Olshanski, Asymptotics of Plancherel measures for symmetric groups, J. Amer. Math. Soc. **13** (2000), 481–515.

[BR] A. Borodin and E. Rains, Eynard-Mehta theorem, Schur process, and their pfaffian analogs, math-ph/0409059.

[HH] P. N. Hoffman and J. F. Humphreys, Projective Representations of the Symmetric Groups: Q-Functions and Shifted Tableaux, Clarendon Press, Oxford, 1992.

[Ji] N. Jing, Vertex operators, symmetric functions, and the spin group Γ_n, J. Algebra **138** (1991), 340–398.

[Jo1] K. Johansson, Shape fluctuations and random matrices, Commun. Math. Phys. **209** (2000), 437–476.

[Jo2] —–, Discrete orthogonal polynomial ensembles and the Plancherel measure, Ann. of Math. (2) **153** (2001), 259–296.

[Mac] I. G. Macdonald, Symmetric Functions and Hall Polynomials, 2nd Edition, Oxford, 1995.

[M1] S. Matsumoto, Correlation functions of the shifted Schur measure, J. Math. Soc. Japan, To appear, math.CO/0312373.

[M2] —–, α-Pfaffian, pfaffian point process and shifted Schur measure, Linear Alg. Appl., To appear, math.CO/0411277.

[N] J. J. C. Nimmo, Hall-Littlewood symmetric functions and the BKP equation, J. Phys. (A) **23** (1990), 751–760.

[O1] A. Okounkov, Random matrices and random permutations, Internat. Math. Res. Notices (2000), 1043–1095.

[O2] —–, Infinite wedge and random partitions, Selecta Math. (N.S.) **7** (2001), 57–81.

[O3] —–, The uses of random partitions, math-ph/0309015.

[OO] A. Okounkov and G. Olshanski, Shifted Schur functions, St. Petersburg Math. J. **9** (1998), 239–300.

[R] E. M. Rains, Correlation functions for symmetrized increasing subsequences, math.CO/0006097.

[S] B. E. Sagan, The symmetric group: Representations, combinatorial algorithms, and symmetric functions, Wadsworth & Brooks / Cole, Pacific Grove, Calif., 1991.

[So] A. Soshnikov, Janossy densities. II. pfaffian ensemble, J. Stat. Phys. **113** (2003), 611–622.

[TW1] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, Comm. Math. Phys. **159** (1994), 151–174.

[TW2] —–, A limit theorem for shifted Schur measures, Duke Math. J. **123** (2004), 171–208.

[Y] Y. You, DKP and MDKP hierarchy of soliton equations, Phys. D **50** (1991), 429–462.

Sho Matsumoto
Graduate School of Mathematics, Kyushu University.
Hakozaki Fukuoka 812-8581, Japan.
e-mail : shom@math.kyushu-u.ac.jp

ZONAL POLYNOMIALS FOR WREATH PRODUCTS

HIROSHI MIZUKAWA

ABSTRACT. The pair of groups, symmetric group S_{2n} and hyperoctohedral group H_n, is a Gelfand pair. The image of zonal spherical functions of this pair under the characteristic map are a family of symmetric functions called zonal polynomials. In the meaning of wreath products, a generalization of this Gelfand pair is considered in this abstract. Its zonal spherical functions are mapped to products of symmetric functions by characteristic map.

RESUME. La paire des groupes, du groupe symetrique S_{2n} et du groupe hyperoctohedral H_n, est une paire de Gelfand. L'image des fonctions spheriques zonales de cette paire sous characteristic map sont une famille des fonctions symetriques appelees les polynômes zonaux. Dans la signi cation de produits en couronne, une generalisation de cette paire de Gelfand est consideree dans cet abstrait. Ses fonctions spheriques zonales sont tracees aux produits des fonctions symetriques par characteristic map.

Key Words: zonal polynomials, Schur functions, Jack symmetric polynomials, Gelfand pairs of finite groups, zonal spherical functions

1. Introduction

The characteristic map can explain the relation of characters of symmetric groups and symmetric functions. We denote by $R(S_n)$ a complex vector space spanned by the irreducible characters of S_n. An element f of $R(S_n)$ can be identified an element $f = \sum_{x \in S_n} f(x)x$ of the group ring $\mathbb{C}S_n$. $R(S_n)$ has a scalar product defined by

$$\langle f, g\rangle = \frac{1}{|S_n|} \sum_{x \in S_n} f(x)\overline{g(x)}.$$

We put

$$R = \bigoplus_{n \geq 0} R(S_n)$$

and define a scalar product on R as

$$\langle f, g\rangle = \sum_{n \geq 0} n! \langle f_n, g_n\rangle \text{ for } f_n, g_n \in R(S_n).$$

R has a ring structure defined as follows. For $u \in R(S_n)$ and $v \in R(S_m)$, we define the multiplication of R by

$$fg = \mathrm{ind}_{S_n \times S_m}^{S_{n+m}} u \times v.$$

Classi cation number :33C45,05E35,05E05.

Let Λ be a ring of symmetric function. We define a $\mathbb{C}$-linear mapping

$$\mathrm{ch} : R \mapsto \Lambda$$

by

$$\mathrm{ch}(\sum_{x \in S_n} f(x)x) = \sum_{x \in S_n} f(x) p_{\ (x)},$$

where $f(x) \in \mathbb{C}$ and $\sigma(x)$ is a cycle type of x. This mapping is called the *characteristic map*. The characteristic map gives isometric isomorphism of R onto Λ. Let χ_ρ be a irreducible character evaluated at a conjugacy class ρ. We obtain Schur functions as a image of a irreducible character of S_n:

$$\mathrm{ch}(\chi\ \ ) = S\ \ ,\ \lambda \vdash n.$$

In Macdonald's book(Chapter I, Appendix B) the theory above is extended to the character theory of the wreath products of any finite group with a symmetric group. We can also define the characteristic map as isometry isomorphism of the character ring of a wreath product onto the ring of symmetric functions. In this case we obtain c-times product of Schur functions as a image of irreducible characters. Here c is a number of irreducible characters of G.

We know a similar theory for the case of a Gelfand pair (S_{2n}, H_n)[6, VII7-2]. We consider zonal spherical functions of this pair. Although precise definition of zonal spherical functions appear in later (see Section 2). Here H_n is a subgroup of S_{2n} defined to be the centralizer of an element $(1,2)(3,4)\cdots(2n-1,2n)$. Littlewood's formula [5] says that

$$1_{H_n}^{S_{2n}} \sim \bigoplus_{\ \vdash n} \chi^2\ \ .$$

In fact, zonal spherical functions are unique H_n-invariant element of each irreducible component of $1_{H_n}^{S_{2n}}$ and constant on each double coset. It is known that double cosets of this pair are classified by the partition of n [6, VII-2(2.1)]. Let ω_ρ be a zonal spherical function in V_i evaluated on a double coset indexed by ρ. We define *zonal polynomials* (cf. [2, 12, 13]) by

$$Z\ \ = |H_n| \sum_{\rho \vdash n} z_{2\rho}^{-1} \omega_\rho p_\rho,\ \lambda \vdash n.$$

Zonal polynomials are a special family of Jack symmetric function $J^\alpha(x)$ [6, 11] with parameter $\alpha = 2$. In terms of Jack symmetric functions, zonal polynomials are formulated as follows: We consider a inner product on $\Lambda \otimes \mathbb{Q}(\alpha)$ defined by

$$\langle p_\rho, p\ \ \rangle_\alpha = \delta_{\rho,\ } z_\rho \alpha^{\ell(\ )}.$$

Zonal polynomials are the unique homogenous basis of $\Lambda \otimes \mathbb{Q}(2)$ satisfying:

(1) $\langle Z\ , Z\ \rangle_2 = h(2\lambda)\delta_{\ ,\ }$, where $h(\lambda)$ is the hook length product of λ
(2) We write $Z\ = \sum\ v_{\ ,\ } m\ $, where m is a monomial symmetric function. Then $v_{\ ,\ } = 0$ unless μ is less than λ as the dominance order(cf. [6] pp.7 Chapter 1-1).
(3) If $\lambda \vdash v$, then $v_{\ ,1^n} = n!$

The characteristic map make us understood equivalency of two definitions above. We define a graded ring of Hecke algebra

$$\mathcal{H} = \bigoplus_{n \geq 0} e_{H_n} \mathbb{C} S_n e_{H_n},$$

where $e_{H_n} = \frac{1}{|H_n|} \sum_{h \in H_n} h$. The multiplication of $\mathcal{H}$ is defined by

$$uv = e_{n+m}(u \times v)e_{n+m}, \ u \in \mathcal{H}_n, v \in \mathcal{H}_m \text{ and } u \times v \in \mathcal{H}_n \times \mathcal{H}_m.$$

We can define the isometry isomorphism ch of $\mathcal{H}$ onto Λ and obtain

$$|H_n|\mathrm{ch}(\omega\ \ ) = Z\ \ .$$

In this abstract our purpose is to generalize third case in the meaning of wreath products. Then we expect to obtain products of zonal polynomials as images of zonal spherical functions under proper isomorphism like the second case. We will consider $G \wr S_{2n}$ instead of S_{2n}. But what kind of subgroup should be chosen, we argue for this problem in Section 3. In Section 4, we classify double coset of the pair defined at Section 3 by using G-colored graphs. We recall the representation theory of wreath products in Section 5. Section 6 is devoted to the irreducible decomposition of the permutation representations. In Section 7-9 we see that our expectation is true. In fact, we obtain products of zonal polynomials and Schur functions as images of our zonal spherical function under a *'characteristic map'* (cf. Theorem 9.3).

2. Gelfand Pair of Finite Groups and Its Zonal Spherical Functions

We recall the theory of Gelfand pair of finite groups. Through this section we denote G by a finite group and H by its subgroup. Put

$$e_H = \frac{1}{|H|} \sum_{h \in H} h.$$

Let $\mathbb{C}G$ be the group ring of G and $e_H \mathbb{C} G e_H$ a Hecke algebra. We regard

$$f = \sum_{x \in G} f(x)x \in \mathbb{C}G, \ f(x) \in \mathbb{C}$$

as a function $x \mapsto f(x)$ on G. Under this identity, the multiplication on $\mathbb{C}G$ is

$$(f * g)(x) = \sum_{yz = x} f(y)g(z).$$

We assume the induced representation $\mathbb{C}Ge$ is multiplicity free, i.e. (G, H) is a Gelfand pair. Under our assumption $\mathbb{C}Ge$ is direct sum of non-isomorphic irreducible G-module;

$$\mathbb{C}Ge = \bigoplus_{i=1}^{s} V_i,$$

where V_i's are irreducible representations. Let χ_i be a character of V_i. We define

$$e_i = \frac{\dim V_i}{|G|} \sum_{x \in G} \overline{\chi_i(x)}$$

to be a primitive idempotent affording to V_i-isotypic component of $\mathbb{C}Ge$. Then we have next proposition [1].

Proposition 2.1. *In the notation introduced above*

$$V_i = \mathbb{C}Ge_ie_H.$$

The Scalar product on $\mathbb{C}G$ is

$$\langle f, g\rangle_G = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}.$$

We can easy to see that this scalar product is G-invariant Hermitian scalar product. The Frobenius reciprocity gives us that $\frac{\dim V_i}{|G|}e_ie_H$ is a unique H-invariant element of V_i which equals to 1 at unit element of G. We call the functions,

$$\omega_i(x) = \frac{\dim V_i}{|G|}e_ie_H = \langle e_ie_H, xe_ie_H\rangle_G/\langle e_ie_H, e_ie_H\rangle_G,\ (1\le i\le s,\ x\in G),$$

zonal spherical functions of a Gelfand pair (G,H). Zonal spherical functions are constant on each double coset $H\backslash G/H$ and have a orthogonality relation

$$\langle \omega_i, \omega_j\rangle_G = \delta_{ij}\frac{1}{\dim V_i}.$$

Proposition 2.2. [6, VIIpp. 389(1.3)] *Let F be a non-zero H-invariant element of $W \cong V_i$ and $\langle,\rangle$ be a G-invariant Hermitian scalar product on W. Then the zonal spherical function is written as*

$$\omega_i(x) = \langle F, xF\rangle/\langle F, F\rangle.$$

Some zonal spherical functions of Gelfand pair of wreath products are calculated in [7, 8, 9].

3. A pair (SG_{2n}, HG_n)

Through this section, S_{2n} is a permutation group on $[2n] = \{1, 2, \cdots, 2n\}$ and its subgroup H_n is the centralizer of an element $(1,2)(3,4)\cdots(2n-1,2n) \in S_{2n}$. We remark that H_n can be considered the permutation groups on $\{\{2i-1, 2i\}; 1\le i\le n\}$ and

$$H_n \cong W(B_n),$$

where $W(B_n)$ is the Weyl group of type B. Let G be a finite group. We denote by G_* the set of conjugacy class of G. We consider a wreath product

$$SG_{2n} = G \wr S_{2n}.$$

Let ΔG be a diagonal subgroup of $G\times G$ defined by

$$\Delta G = \{(g,g) \mid g\in G\}.$$

We restrict the action of S_{2n} on G^{2n} to H_n and define a subgroup of SG_{2n} by

$$HG_n = (\Delta G)^n \rtimes H_n.$$

From the definition of H_n it is clear that HG_n is well defined.

4. Description of Double Cosets

Through a combinatorial argument, we can describe a complete representatives of each double coset of (SG_{2n}, HG_n) [10]. In this section, without proofs, we introduce a method of identification of each double coset.

For an element $x = (g_1, g_2, \cdots, g_{2n}; \sigma)$ of SG_{2n}, the *G-colored graph* $\Gamma_G(x) = \{V_G(x), E_G(x)\}$ is a graph with vertices

$$V_G(x) = \{g_1, g_2, \cdots, g_{2n}\}$$

and edges

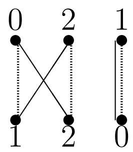
$$E_G(x) = \left\{\{g_{2i-1}, g_{2i}\}, \left\{g_{\ (2j-1)}, g_{\ (2j)}\right\}; 1 \leq i, j \leq n\right\}.$$

Here we call the edge $\{g_{2i-1}, g_{2i}\}$ *"broken"* and $\left\{g_{\ (2i-1)}, g_{\ (2i)}\right\}$ *"staright"*.

Example 4.1. $G = \mathbb{Z}/3\mathbb{Z} = \{0, 1, 2\}$. We consider SG_6 and take an element

$$x = (0, 1, 2, 2, 1, 0; (123)(56)).$$

Then the graph of x is

$\Gamma_G(x) =$  .

This graph gives a two-sided HG_n-invariant: Fix an element $x = (g_1, g_2, \cdots, g_{2n}; \sigma)$ of SG_{2n}. Let L be a cycle of $\Gamma_G(x)$. We assume that L has vertices $\left\{g_{i_j}; 1 \leq j \leq 2k\right\}$. Let

$$\left\{\left\{g_{i_{2j}\ 1}, g_{i_{2j}}\right\}; (1 \leq j \leq k)\right\}$$

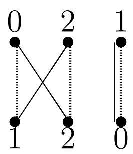
be broken edges of L and

$$\left\{\left\{g_{i_{2j}}, g_{i_{2j+1}}\right\}, \{g_{i_{2k}}, g_{i_1}\}; (1 \leq j \leq k-1)\right\}$$

be staright edges of L. We define a *circuit product* of L by

$$p(L) = \prod_{j=1}^{k} g_{i_{2j}\ 1}^{-1} g_{i_{2j}}.$$

Example 4.2. In the case of example 4.1, from a graph

$\Gamma_G(x) =$  ,

we compute circuit products

$$-0 + 1 - 2 + 2 = 1, -1 + 0 = 2.$$

If L has $2k$ edges then we call $p(L)$ a *circuit product of length k*.

Definition 4.3. Put

$$G_{**} = \left\{ R = C \cup C^{-1} ; C \in G_* \right\},$$

where $C^{-1} = \{g^{-1}; g \in C\}$. We call a conjugacy class *real*(resp. *complex*) when $C = C^{-1}$(resp.$C \neq C^{-1}$). Put

$$m_k(R) = \sharp\{L; L \text{ is a 2k-cycle of } \Gamma(x) \text{ and } p(L) \in G_{**}. \}$$

We define a tuple of partitions

$$\underline{\rho}(x) = (\rho(R); R \in G_{**}),$$

where $\rho(R) = (1^{m_1(R)}, 2^{m_2(R)}, \cdots, n^{m_n(R)})$. This tuple of partitions $\underline{\rho}(x)$ is called *circuit type of* x.

Example 4.4. If $G = \mathbb{Z}/3\mathbb{Z}$ we have $G_{**} = \{\{0\}, \{1,2\}\}$. In the case of example 4.1, we obtain

$$\underline{\rho}(x) = ((\emptyset), (2,1)).$$

Definition 4.3 gives us next theorem.

Theorem 4.5. (1) $x \in HG_n y HG_n \Leftrightarrow \underline{\rho}(x) = \underline{\rho}(y)$.

(2) $\underline{\rho}(x) = \underline{\rho}(x^{-1})$.

Furthermore we can see the cardinality of each double coset.

Proposition 4.6. *Let x be an element such that whose circuit type is $\underline{\rho}(x) = (\rho(R); R \in G_{**})$ where $\rho = (1^{m_1(R)}, 2^{m_2(R)}, \cdots, n^{m_n(R)})$ and $\zeta_C = \frac{|G|}{|C|}$ for $C \in G_*$. Then we have*

$$\begin{aligned}
{\mathcal{Z}_{\underline{\rho}(x)}}^{-1} = |HG_n x HG_n| &= \frac{|H_n|^2 |G|^{2n}}{\prod_{R \in G_{**}} z_{2\rho(R)}} \times \frac{\prod_{R \in G_{**}} |R|^{\ell(\rho(R))}}{|G|^{\ell(\rho)}} \\
&= |H_n|^2 |G|^{2n} \prod_{\substack{R = C \in G_{**} \\ C = C^{-1}}} \frac{1}{z_{2\rho(R)} \zeta_C^{\ell(\rho(R))}} \times \prod_{\substack{R = C \cup C^{-1} \in G_{**} \\ C \neq C^{-1}}} \frac{1}{z_{\rho(R)} \zeta_C^{\ell(\rho(R))}}.
\end{aligned}$$

This result is important to determine the weight of inner product on the ring of symmetric functions see Section 7 .

5. Representation Theory of Wreath Products

In this section we recall the representation theory of wreath products (cf. [4]). Let G be a finite group. We write

$$SG_n = G \wr S_n.$$

Let G^* be a set of irreducible characters of G and c its cardinality. We introduce a construction method of the irreducible representations.

Let

$$\mathcal{C}_n = \left\{ \underline{n} = (n_\chi; \chi \in G^*); \sum_{\chi \in G^*} n_\chi = n, \ n_i \geq 0 \right\}$$

be a set of c-composition of n. We take an element $\underline{n} \in \mathcal{C}_n$ and define a set of c-tuple of partitions;

$$\mathcal{P}(\underline{n}) = \{(\lambda^\chi | \chi \in G^*); \lambda^\chi \vdash n_\chi\}$$

We define a subgroup of SG_n by

$$SG(\underline{n}) = \prod_{\chi \in G^*} SG_{n_\chi}.$$

Taking $\underline{n} \in \mathcal{C}_n$ and $\underline{\lambda} = (\lambda(\chi) | \chi \in G^*) \in \mathcal{P}(\underline{n})$, we define two representations $\mathcal{R}(\underline{n})$ and $\mathcal{S}(\underline{\lambda})$ of $SG(\underline{n})$ as follows:

$$\mathcal{R}(\underline{n}) \cong \bigotimes_{\chi \in G^*} V_\chi^{\otimes n_\chi},$$

$$\mathcal{S}(\underline{\lambda}) \cong \bigotimes_{\chi \in G^*} S^{\ \ (\chi)},$$

where S is the Specht module indexed by a partition λ. The action of $SG(\underline{n})$ is defined by

$$(g_1, \cdots, g_n; \sigma) v_1 \otimes \cdots \otimes v_n = g_1 v_{\ \ {}^{1}(1)} \otimes \cdots \otimes g_n v_{\ \ {}^{1}(n)} \text{ on } \mathcal{R}(\underline{n}),$$

and

$$(g_1, \cdots, g_n; \sigma) v = \sigma v \text{ on } \mathcal{S}(\underline{\lambda}).$$

We consider an irreducible representation of $S(\underline{n})$

$$S(\underline{\lambda}) = \mathcal{R}(\underline{n}) \otimes \mathcal{S}(\underline{\lambda})$$

We write

$$\mathfrak{S}(\underline{\lambda}) = S(\underline{\lambda}) \uparrow_{SG(\underline{n})}^{SG_n}.$$

Theorem 5.1. [4] *The complete system of irreducible representations of SG_n are given by*

$$\{\mathfrak{S}(\underline{\lambda}); \underline{n} \in \mathcal{C}_n,\ \underline{\lambda} \in \mathcal{P}(\underline{n})\}.$$

As can be seen from the theorem above, there is a one-to-one correspondence between SG_n^* and the set of c-tuple of partitions of n.

6. Gelfand pair (SG_{2n}, HG_n)

From the second claim of Theorem 4.5, we see that $x \in SG_{2n}$ and x^{-1} are in same double coset. Therefore we have the following proposition [6, VII(1.2)].

Proposition 6.1. *(SG_{2n}, HG_n) is a Gelfand pair.*

We consider irreducible decomposition of the permutation representation

$$\mathrm{ind}_{HG_n}^{SG_{2n}} 1 = 1_{HG_n}^{SG_{2n}}.$$

Let G^* be a set of irreducible characters of G. We call a character $\chi \in G^*$ *real* (resp. *complex*) if $\chi = \overline{\chi}$(resp. $\chi \neq \overline{\chi}$). Let G_R^* be a set of real characters and G_C^* a set of complex characters. We define a relation in G_C^* as

$$\chi \sim \chi' \Leftrightarrow \overline{\chi} = \chi'$$

and

$$G^{**} = G_R^* \cup G_C^*/_{\sim}.$$

Taking proper representatives, we consider $G^{**}/\sim$ to be a subset of G^*. Next propositions are elementary in this section.

Proposition 6.2. (1) (S_{2n}, H_n) *is a Gelfand pair.*

(2) $(G \times G, \Delta G)$ *is a Gelfand pair.*

(3) *Especially,* $(S_n \times S_n, \Delta S_n)$ *is a Gelfand pair.*

Proposition 6.3. (1) $1_{H_n}^{S_{2n}} = \bigoplus_{\ \vdash n} S^{2\ }$.

(2) $1_{\Delta G}^{G\times G} = \bigoplus_{\chi \in G_R^*} \chi \otimes \chi \oplus \bigoplus_{\chi \in G_C^*} \chi \otimes \overline{\chi}$.

(3) *Especially,* $1_{\Delta S_n}^{S_n \times S_n} = \bigoplus_{\ \vdash n} S^{\ } \otimes S^{\ }$.

We use the notation appeared in Section 5. We put

$$\mathcal{C}_{2n} \supset \mathcal{C}_{2n}^{**} = \{(n_\chi; n_\chi \equiv 0 \pmod 2\ \chi \in G_R^*, n_\chi = n_{\overline{\chi}}\ \chi \in G_C^*)\}.$$

Example 6.4. The case of $G = \mathbb{Z}/2\mathbb{Z}$:

$$\mathcal{C}_{2n}^{**} = \{(2n-2k, 2k); 0 \leq k \leq n\}.$$

The case of $G = \mathbb{Z}/3\mathbb{Z}$:

$$\mathcal{C}_{2n}^{**} = \{(2n-2k, k, k); 0 \leq k \leq n\}.$$

For $\underline{n} \in \mathcal{C}_{2n}^{**}$ we define

$$\mathcal{P}(\underline{n}) \supset \mathcal{P}^{**}(\underline{n}) = \left\{(\lambda^\chi; \lambda^\chi = 2^\exists \mu^\chi,\ \chi \in G_R^* \text{ and } \ \lambda^\chi = \lambda^{\overline{\chi}},\ \chi \in G_C^*)\right\},$$

where 2λ means $(2\lambda_1, 2\lambda_2, \cdots)$ for $\lambda = (\lambda_1, \lambda_2, \cdots)$.

Example 6.5. The case of $G = \mathbb{Z}/2\mathbb{Z}$:

$$\mathcal{P}^{**}(\underline{n}) = \{(2\lambda, 2\mu); |\lambda| + |\mu| = n\}.$$

The case of $G = \mathbb{Z}/3\mathbb{Z}$:

$$\mathcal{P}^{**}(\underline{n}) = \{(2\lambda, \mu, \mu); |\lambda| + |\mu| = n\}.$$

We consider a representation $\chi(\underline{n}) \otimes S(\underline{\lambda})$ for $\underline{n} \in \mathcal{C}_{2n}^{**}$ and $\underline{\lambda} \in \mathcal{P}^{**}(\underline{n})$. Propositions 6.2 and 6.3 give us the following fact.

Proposition 6.6. $\chi(\underline{n}) \otimes S(\underline{\lambda})$ *has* $\prod_{\chi \in G_R^*} HG_{n_\chi} \times \prod_{\chi \in G_C^*/\sim} \Delta SG_{n_\chi}$*-invariant element. Here we think the following embedding*

$$HG_{n_\chi} \subset SG_{2n_\chi},\ \Delta SG_{n_\chi} \subset SG_{n_\chi} \times SC_{n_{\overline{\chi}}}$$

and

$$\prod_{\chi \in G_R^*} HG_{n_\chi} \times \prod_{\chi \in G_C^*/\sim} \Delta SG_{n_\chi} \subset SG(\underline{\lambda}).$$

A construction method of the irreducible representation of wreath product, see Section 5, gives us the reverse of Proposition 6.6. Next proposition is a corollary of a lemma due to Brauer(cf. [3, Chapter6 (6.32)]).

Proposition 6.7.

$$|G^{**}| = |G_{**}|.$$

Therefore we have the following theorem from Proposition 6.6 and 6.7

Theorem 6.8.

$$1_{HG_n}^{SG_{2n}} = \bigoplus_{\underline{n}\in\mathcal{C}_{2n}^{**}} \bigoplus_{\underline{\ }\in\mathcal{P}^{**}(\underline{n})} \mathfrak{S}(\underline{\lambda})$$

The end of this section we see an example.

Example 6.9. The case of $G = \mathbb{Z}/2\mathbb{Z}$:

$$1_{HG_n}^{SG_{2n}} = \bigoplus_{|\ |+|\ |=n} \mathfrak{S}(2\lambda, 2\mu).$$

The case of $G = \mathbb{Z}/3\mathbb{Z}$:

$$1_{HG_n}^{SG_{2n}} = \bigoplus_{|\ |+|\ |=n} \mathfrak{S}(2\lambda, \mu, \mu).$$

7. The Ring $\tilde{\Lambda}(G)$

In this section we define a suitable ring of symmetric function for considering our zonal spherical functions. Let $p_r(R)(r \geq 1)$ be the power sum symmetric function with variables $x(R) = (x(R)_1, x(R)_2, \cdots)$ for $R \in G_{**}$ and $\tilde{\Lambda}(G)$ a ring generated by $p_r(R)(r \geq 1, R \in G_{**})$. Let $\underline{\rho} = (\rho(R); R \in G_{**})$ be a $|G_{**}|$-tuple of partitions. Put

$$P_{\underline{\rho}}(G_{**}) = \prod_{R\in G_{**}} p_{\rho(R)}(R)$$

for $\underline{\rho}$. We change variables $p_r(R)$'s to

$$p_r(\chi) = \sum_{\substack{R=C\cup C^{\ 1}\in G_{**} \\ C=C^{\ 1}}} \frac{\chi(C)}{\zeta_C} p_r(R) + \sum_{\substack{R=C\cup C^{\ 1}\in G_{**} \\ C\neq C^{\ 1}}} \frac{\chi(C)+\overline{\chi(C)}}{\zeta_C} p_r(R),$$

where $\chi \in G^{**}$ and $\chi(C)$ is a value of χ at conjugacy class C. We also put

$$P_{\underline{\ }}(G^{**}) = \prod_{\chi\in G^{**}} p_{\ \chi}(\chi)$$

for a tuple of partition $\underline{\lambda} = (\lambda^\chi; \chi \in G^{**})$. We define an inner product on $\tilde{\Lambda}(G)$ by

$$\langle P_{\underline{\rho}}(G_{**}), P_{\underline{\ }}(G_{**})\rangle_{\tilde{\Lambda}(G)} = \delta_{\underline{\rho}\underline{\ }} \mathcal{Z}_{\underline{\rho}}.$$

$\mathcal{Z}_{\underline{\rho}}$ is given in Proposition 4.6. Here we write a polynomial with variables $p_r(\chi)$ like as $S\ \ (\chi)$.

8. The Ring $\mathcal{H}(G)$

We define a Hecke algebra by

$$\mathcal{H}(SG_{2n}, HG_n) = e_{HG_n}\mathbb{C}SG_{2n}e_{HG_n}.$$

It is true that zonal spherical functions of (SG_{2n}, HG_n) are orthonormal basis of $\mathcal{H}(SG_{2n}, HG_n)$. We define a graded vector space

$$\mathcal{H}(G) = \bigoplus_{n\geq 0}\mathcal{H}(SG_{2n}, HG_n).$$

The multiplication of $\mathcal{H}(G)$ is defined by

$$uv = e_{HG_{n+m}}(u \times v)e_{HG_{n+m}},$$

where we think that $u \times v$ is a function on a diagonal subalgebra $\mathbb{C}S_{2n} \times \mathbb{C}S_{2m}$ of $\mathbb{C}S_{2n+2m}$. Since $\mathcal{H}(G)$ has a structure of a graded algebra. We see that zonal spherical functions are basis of $\mathcal{H}(G)$.

9. Main result

By using group theoretical method as in the book [1] we can obtain zonal spherical functions as a product of some primitive idempotents. To describe details of this fact is a little complicated. So we omit to explain how to get our zonal spherical functions here. We only show the final final form of them.

Our zonal spherical functions can be described as follows. Let $S\ (\chi)$ be an irreducible representation of SG_n isomorphic to $\chi(\underline{n}) \otimes S$ for $\underline{n} = (n_\eta; \eta \in G^*$ and $n_\eta = \delta_{\chi\eta}n)$. We define $e\ (\chi), (\ \lambda \vdash n$ and $\chi \in G^*)$ to be a primitive idempotent of $\mathbb{C}SG_n$ which afford to $S\ (\chi)$-isotypic component in $\mathbb{C}SG_n$. We avoid to write concrete equations in the below so we use the notation "$\propto$". Then we have

Proposition 9.1. *We put* $\underline{n} \in \mathcal{C}^{**}_{2n}$ *and* $\underline{\lambda} = (\lambda^\chi; \lambda^\chi = 2^{\exists}\mu^\chi,\ \chi \in G^*_R$ and $\lambda^\chi = \lambda^{\overline{\chi}}) \in \mathcal{P}^{**}(\underline{n})$. *Then we have zonal spherical functions in a irreducible component* $\chi(\underline{n}) \otimes S(\lambda) \uparrow^{SG_{2n}}_{HG_n}$ *of* $1^{SG_{2n}}_{HG_n}$ *as*

$$\omega^{-} \propto e_{HG_n}\left(\prod_{\chi\in G^*_R} e_{2\ \chi}(\chi) \times \prod_{\chi\in G^*_C/\sim} e_{\ \chi}(\chi) \times e_{\ \chi}(\overline{\chi})\right)e_{HG_n}.$$

We define a characteristic map

$$ch_{\mathcal{H}} : \mathcal{H}(G) \mapsto \tilde{\Lambda}(G)$$

by

$$ch_{\mathcal{H}}(x) = P_{\underline{\rho}(x)}(G^{**}),\ (x \in SG_{2n}).$$

This characteristic map gives an isometric isomorphism of $\mathcal{H}(G)$ onto $\tilde{\Lambda}(G)$. We have the following proposition.

Proposition 9.2. *Let* λ *be a partition of* n.

$$ch_{\mathcal{H}}(e_2\ (\chi)e_{HG_n}) \propto Z\ (\chi),\ \chi \in G^*_R,$$
$$ch_{\mathcal{H}}(e\ (\chi) \times e\ (\overline{\chi})e_{\Delta SG_n}) \propto h(\lambda)S\ (\chi),\ \chi \in G^*_C/\sim.$$

Bh combining two propositions above, we obtain the main theorem of this abstract.

Theorem 9.3. *We put* $\underline{n} \in \mathcal{C}_{2n}^{**}$ *and* $\underline{\lambda} = (\lambda^\chi; \lambda^\chi = 2\mu^\chi,\ \chi \in G_R^*$ and $\lambda^\chi = \lambda^{\overline{\chi}}) \in \mathcal{P}^{**}(\underline{n})$. *Let* $\omega-$ *be a zonal spherical function in a irreducible component* $\mathfrak{S}(\lambda)$ *of* $1_{HG_n}^{SG_{2n}}$. *Then we have*

$$\begin{aligned} ch_{\mathcal{H}}(\omega-) &\propto \prod_{\chi \in G_R^*} Z_{\ \chi}(\chi) \times \prod_{\chi \in G_C^*/\sim} h(\lambda^\chi) S_{\ \chi}(\chi) \\ &= \prod_{\chi \in G_R^*} J^{(2)}_{\ \chi}(\chi) \times \prod_{\chi \in G_C^*/\sim} J^{(1)}_{\ \chi}(\chi). \end{aligned}$$

Example 9.4. In the case of $G = \mathbb{Z}/2\mathbb{Z} = \{-1, 1\}$: Put $\chi_i(j) = j^i (i = 0, 1, j \in G)$ We obtain

$$\{Z\ (\chi_0) Z\ (\chi_1)\}$$

as images of zonal spherical functions.
In the case of $G = \mathbb{Z}/3\mathbb{Z} = \{1, \xi, \xi^2\}$: Put $\chi_i(j) = j^i (i = 0, 1, 2, j \in G)$ We obtain

$$\{Z\ (\chi_0) S\ (\chi_1)\}$$

as images of zonal spherical functions.

References

[1] C. W. Curtis and I. Reiner, *Methods of Representations Theory, Vol. 1* A Wiley-Interscience, 1981.

[2] N. Bergeron and A. M. Garsia, Zonal polynomials and domino tableaux, Discrete Mathematics 99, 1992 pp. 3-15.

[3] I. M. Issacs, *Character Theory of Finite Groups,* Dover Publications Inc. 1994.

[4] G. James and A. Kerber, *The Representation Theory of the Symmetric Group,* Encyclopedia of Mathematics and its Applications, 16, 1981.

[5] D. E. Littlewood, *The Theory of Group Characters, 2nd. ed. ,* Oxford, 1950.

[6] I. G. Macdonald, *Symmetric Functions and Hall Polynomials, 2nd. ed. ,* Oxford, Clarendon Press, 1995.

[7] H. Mizukawa, Zonal spherical functions on the complex re ection groups and $(m+1, n+1)$-hypergeometric functions Advances in Math. 184, 2004 pp. 1-17.

[8] Hiroki Akazawa and Hiroshi Mizukawa, Orthogonal polynomials arising from the wreath products of dihedral group J. Combin. Theory Ser. A 104 2003 pp. 371-380.

[9] Hiroshi Mizukawa and Hajime Tanaka , $(n+1, m+1)$-hypergeometric functions associated to character algebras, Proc. AMS. 132, 2004, no. 9, pp. 2613-2618.

[10] H. Mizukawa, Zonal polynomials for wreath products, preprint 2004.

[11] R. Stanley, Some combinatorial properties of Jack symmetric functions, Advances in Math. vol. 77,1989, pp.76-115

[12] J. R. Stembridge, On Schur's Q-functions and the primitive idempotents of a commutative Hecke algebra", J. Alg. Comb. 1 1992, pp.71-95.

[13] A. Takemura *Zonal Polynomials,* Inst. of Math. Statistics Lecture Notes,-Monograph Series vol.4 .Hayward, California,1984

Department of Mathematics, National Defense Academy in Japan, Yokosuka 239-8686, Japan.

E-mail address: mzh@nda.ac.jp

The Bruhat ordering on the Coxeter group of type $\widetilde{C}_n$

Masayuki Nanba

Graduate School of Mathematics, Nagoya University.
e-mail : m00012h@math.nagoya-u.ac.jp

Abstract

This paper deals with two topics about the Bruhat ordering on the Coxeter groups.

The first topic is a general theory concerning with Coxeter graph automorphisms.A Coxeter graph automorphism σ of a Coxeter system (W, S) extends to a group automorphism on W, and gives rise to a subgroup W_σ of W consisting of all elements fixed by σ. Then we prove that the Bruhat ordering on W_σ is the restriction of the Bruhat ordering on W. We study a relation between the Bruhat ordering and a Coxeter graph automorphism.

The second topic is an application of the above theory. By applying this theory to a Coxeter group $W(\widetilde{A}_{2n-1})$ of type $\widetilde{A}_{2n-1}$ and a combinatorial description of the Bruhat ordering on it due to Björner–Brenti and Lascoux, we give a combinatorial description of the Bruhat ordering on $W(\widetilde{C}_n)$ of type $\widetilde{C}_n$.

1 Introduction

Let (W, S) be a Coxeter system, and ℓ be the length function on W. Let $T = \{wsw^{-1} \mid w \in W,\ s \in S\}$ be the set of reflections. The Bruhat ordering on W is a partial ordering on W defined as follows:

For $w, w' \in W$, we define $w \leq w'$ if there exist elements $t_1, \cdots, t_r \in T$ such that

(1) $w' = t_r \cdots t_1 w$,

(2) $\ell(t_i \cdots t_1 w) \geq \ell(t_{i-1} \cdots t_1 w) \quad (1 \leq \forall i \leq r)$.

A Coxeter graph automorphism of (W, S) is a bijection from S to itself such that

$$m(\sigma(s), \sigma(s')) = m(s, s') \qquad (\forall s, s' \in S),$$

where $m(s, s')$ is the order of the product ss' for $s, s' \in S$. Such σ extends to an automorphism on W, which is called *a Coxeter graph automorphism of* (W, S).

In this paper, we are interested in the fixed-point subgroup

$$W_\sigma = \{w \in W \mid \sigma(w) = w\}$$

of W. R. Steinberg proved that W_σ is a Coxeter group. More precisely,

Theorem 1.1 (R. Steinberg [12])
Let (W, S) be a Coxeter system, and let σ be a Coxeter graph automorphism of (W, S). If we put

$$S_\sigma = \{w_X \mid X \textit{ is a } \langle\sigma\rangle\textit{-orbit in } S \textit{ with } W_X = \langle X \rangle \textit{ finite}\},$$

where w_X is the longest element in W_X, then (W_σ, S_σ) is a Coxeter system. □

From Theorem 1.1, W_σ has its own Bruhat ordering with respect to S_σ. The author clarifies the relation between the Bruhat ordering on W_σ and that on W. The following is the one of our main results:

Theorem 1.2 (M. Nanba [10])
In the setting of Theorem 1.1, the Bruhat ordering on (W_σ, S_σ) is the restriction of the Bruhat ordering on (W, S) to W_σ. That is, if $\leq$ is the Bruhat ordering on W with respect to S, and $\leq_\sigma$ is the Bruhat ordering on W_σ with respect to S_σ, then

$$w \leq w' \iff w \leq_\sigma w'$$

for all $w, w' \in W_\sigma$.

The outline of the proof of this theorem is explained in Section 2.2.

As application of this theorem, we give a combinatorial description of the Bruhat ordering on the Coxeter group $W(\widetilde{C}_n)$ of type $\widetilde{C}_n$, which is embedded into the Coxeter group $W(\widetilde{A}_{2n-1})$ of type $\widetilde{A}_{2n-1}$.

Let $W(\widetilde{A}_{N-1})$ be the group of affine permutations, i.e., an element π in $W(\widetilde{A}_{N-1})$ is a bijection from $\mathbb{Z}$ to $\mathbb{Z}$ such that

(1) $\pi(x) + N = \pi(x + N)$ for all $x \in \mathbb{Z}$,

(2) $\pi(1) + \pi(2) + \cdots + \pi(N) = \dfrac{1}{2}N(N+1)$.

This group $W(\widetilde{A}_{N-1})$ is known to be *a Coxeter group of type* $\widetilde{A}_{N-1}$. ([8].) Note that an affine permutation π is determined by its values $\pi(1), \pi(2), \cdots, \pi(N)$. Moreover, it is obvious from (2) that an affine permutations π has the following property:

For $i, j \in \mathbb{Z}$,

$$\pi(i) \equiv \pi(j) \pmod N \iff i \equiv j \pmod N. \tag{1.1}$$

In order to describe the Bruhat ordering on $W(\widetilde{A}_{N-1})$, we need investigate the Bruhat ordering on the following subset:

$$W(\widetilde{A}_{N-1})^{J_0} = \{\pi \in W(\widetilde{A}_{N-1}) \mid \pi(1) < \pi(2) < \cdots < \pi(N)\} \subset W(\widetilde{A}_{N-1}).$$

This subset $W(\widetilde{A}_{N-1})^{J_0}$ is the set of distinguished coset representatives with respect to a maximal parabolic subgroup. Moreover, the elements in $W(\widetilde{A}_{N-1})^{J_0}$ are in one-to-one correspondence with N-cores. In particular, this encoding enables us to describe combinatorially the Bruhat ordering on $W(\widetilde{A}_{N-1})^{J_0}$. (See Theorem 4.5.)

Let $N = 2n$ be an even integer and put

$$W(\widetilde{C}_n) = \{w \in W(\widetilde{A}_{2n-1}) \mid w(i) + w(2n+1-i) = 2n+1 \ (\forall i \in \mathbb{Z})\}.$$

This subgroup $W(\widetilde{C}_n)$ of $W(\widetilde{A}_{2n-1})$ is the fixed-point subgroup under a certain Coxeter graph automorphism, and it it is the Coxeter group of type $\widetilde{C}_n$. (See Theorem 3.3.) Hence we can apply the above Theorem 1.2 to obtain a combinatorial description of the Bruhat ordering on it.

If we put

$$W(\widetilde{C}_n)^{I_0} = \{w \in W(\widetilde{C}_n) \mid w(1) < w(2) < \cdots < w(2n)\} = W(\widetilde{C}_n) \cap W(\widetilde{A}_{2n-1})^{J_0},$$

then $W(\widetilde{C}_n)^{I_0}$ is the set of a distinguished coset representatives of $W(\widetilde{C}_n)$ with respect to a maximal parabolic subgroup and has also the encoding with $2n$-core. Another main result is the following theorem. (See Theorem 5.3.)

Theorem 1.3 *The elements of $W(\widetilde{C}_n)^{I_0}$ are in one-to-one correspondence with symmetric $2n$-cores. Moreover, if w and $v \in W(\widetilde{C}_n)$ correspond to symmetric $2n$-cores λ and μ respectively, then $w \leq v$ if and only if $\lambda \subseteq \mu$.* $\square$

This paper is organized as follows.

Section 2: In this section, we review some facts about Coxeter groups and the outline of the proof of Theorem 1.2.

Section 3: We consider two Coxeter groups: $W(\widetilde{A}_{N-1})$ of type $\widetilde{A}_{N-1}$ and $W(\widetilde{C}_n)$ of type $\widetilde{C}_n$.

Section 4: We review Lascoux's encoding of the Bruhat ordering on $W(\widetilde{A}_{N-1})^{J_0}$.

Section 5: In Section 5, we give the combinatorial description of the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ by using Lascoux's description of that on $W(\widetilde{A}_{2n-1})$ and Theorem 1.2.

Section 6: By Theorem 1.3 and a certain Coxeter graph automorphism ω, we find the combinatorial description of the Bruhat ordering on the whole group $W(\widetilde{C}_n)$.

2 General theory

In this section, we review some key facts on Coxeter groups and give an outline of the proof of Theorem 1.1 and 1.2. Throughout this section, let (W, S) be a Coxeter system and σ a Coxeter graph automorphism, otherwise stated.

2.1 Key facts

Since a Coxeter graph automorphism σ is a bijection an S, we have

$$\ell(\sigma(w)) = \ell(w) \qquad (\text{for all } w \in W), \tag{2.1}$$

$$\sigma(T) = T. \tag{2.2}$$

Hence it follows from the definition of Bruhat ordering that σ is order-preserving automorphism of W.

Proposition 2.1 *Let (W, S) be a Coxeter system, and σ be a Coxeter graph automorphism of (W, S). Let $\leq$ be the Bruhat ordering on W, then for $w, w' \in W$,*

$$w \leq w' \iff \sigma(w) \leq \sigma(w').$$

$\square$

For a subset $J \subseteq S$, let $W_J = \langle J \rangle$ be the parabolic subgroup and W^J be the set of *distinguished (left) coset representatives*. For each $w \in W$, we can find a unique pair of elements $w^J \in W^J$ and $w_J \in W_J$ such that $w = w^J w_J$ and $\ell(w) = \ell(w^J) + \ell(w_J)$. (See J. E. Humphreys [5]) In this decomposition, $w^J \in W^J$ is called the *W^J-part* of w. The equation (2.1) can also leads the following proposition:

Proposition 2.2 *For a Coxeter system (W, S), let σ be a Coxeter graph automorphism of (W, S). For all $J \subseteq S$ and $w \in W$, $\sigma(w^J) = (\sigma(w))^{\sigma(J)}$.* $\square$

We define the Bruhat ordering on W^J by restricting the Bruhat ordering on W to W^J. The following theorem can be deduced immediately from [4]:

Theorem 2.3 (See V. V. Deodhar [4])
Let (W, S) be a Coxeter system with the identity element e and J be a subset of S. Let $\leq$ be a relation on W^J. Then the following are equivalent:

(1) *The relation $\leq$ is the Bruhat ordering on W^J.*

(2) *The relation $\leq$ satisfies (a) and (b) as follows:*

   (a) *$w \leq e$ if and only if $w = e$.*

   (b) *The following is satisfied :*

Property $Z^J(s, w_1, w_2)$: *For* $w_1, w_2 \in W^J$ *and* $s \in S$ *such that* $\ell(sw_2) \leq \ell(w_2)$ *and* $\ell((sw_1)^J) \leq \ell(w_1)$*, one has* $w_1 \leq w_2 \Leftrightarrow (sw_1)^J \leq w_2 \Leftrightarrow (sw_1)^J \leq (sw_2)^J$.

□

In particular, if $J = \emptyset$, then the Property $Z^J(s, w_1, w_2)$ is the same as the Property $Z(s, w_1, w_2)$ which appeared in [4].

2.2 Outline of the proof of Theorem 1.1 and 1.2

R. Steinberg proved Theorem 1.1 by using a root system. Here we give another proof to Theorem 1.1 based on the Exchange Condition:

Theorem 2.4 (See N. Bourbaki [3].) *Let* W *be a group, and* S *be a subset of* W *such that* S *generates* W *and* $m(s, s) = 1$ *for all* $s \in S$*. Then the pair* (W, S) *is a Coxeter system if and only if* (W, S) *satisfies the Exchange Condition (EC) :*

(EC) Let $w \in W$ *and* $w = s_1 \cdots s_q$ *be an arbitrary expression. If* $\ell(sw) \leq \ell(w)$ *for* $s \in S$*, then there exists* j *with* $1 \leq j \leq q$ *such that* $sw = s_1 \cdots s_{j-1}s_{j+1} \cdots s_q$. □

The following lemma is obtained from the Exchange Condition and (2.1):

Lemma 2.5 *For a Coxeter system* (W, S) *with a Coxeter group automorphism* σ*, we have*

(1) *Let* $X \subset S$ *be a* $\langle\sigma\rangle$*-orbit. Then, for all* $w \in W_\sigma$,

$$\exists s \in X \ s.t. \ \ell(sw) \leq \ell(w) \iff \ell(s'w) \leq \ell(w) \ (\forall s' \in X),$$
$$\exists s \in X \ s.t. \ \ell(sw) \geq \ell(w) \iff \ell(s'w) \geq \ell(w) \ (\forall s' \in X)$$

(2) *For each* $w \in W_\sigma$*, there exist* $w_{X_1}, w_{X_2}, \cdots, w_{X_r} \in S_\sigma$ *such that* $w = w_{X_1}w_{X_2} \cdots w_{X_r}$ *and* $\ell(w) = \ell(w_{X_1}) + \ell(w_{X_2}) + \cdots + \ell(w_{X_r})$. □

The second part of this lemma shows that S_σ generates W_σ. Therefore we can define the length function ℓ_σ on W_σ with respect to S_σ. We have the following proposition from the Exchange Condition:

Proposition 2.6 *Suppose that* $w \in W_\sigma$ *has an expression* $w = w_{X_1}w_{X_2} \cdots w_{X_r}$ *as a product of* $w_{X_1}, w_{X_2}, \cdots, w_{X_r} \in S_\sigma$*. Let* X *be a subset* $\langle\sigma\rangle$*-orbit with* W_X *finite. If* $\ell(sw) \leq \ell(w)$ *for an element* $s \in X$*, there exists an integer* k *with* $1 \leq k \leq r$ *such that*

$$w_X w = w_{X_1} \cdots w_{X_{k-1}} w_{X_{k+1}} \cdots w_{X_r}.$$

. □

Corollary 2.7 *For* $w_X \in S_\sigma$ *and* $w \in W_\sigma$,

(1) $w = w_{X_1}w_{X_2} \cdots w_{X_r}$ *is reduced if and only if* $\ell(w) = \ell(w_{X_1}) + \ell_\sigma(w_{X_2}) + \cdots + \ell_\sigma(w_{X_r})$.

(2) $\ell_\sigma(w_X w) \leq \ell_\sigma(w) \iff \exists s \in X \ s.t. \ \ell(sw) \leq \ell(w)$*. Then* $\ell(w_X w) = \ell(w) - \ell(w_X)$.

(3) $\ell_\sigma(w_X w) \geq \ell_\sigma(w) \iff \exists s \in X \ s.t. \ \ell(sw) \geq \ell(w)$*. Then* $\ell(w_X w) = \ell(w) + \ell(w_X)$. □

Now we are in position to give proofs of Theorem 1.1 and 1.2.

Proof of Theorem 1.1. It follows from Proposition 2.6 and Corollary 2.7 that (W_σ, S_σ) satisfies the Exchange Condition. □

Proof of Theorem 1.2. It is enough to show that the restriction to W_σ of the Bruhat order $\leq$ on W satisfies the condition (a) and (b) in Theorem 2.3 (2) for $J = \emptyset$.

It is obvious that the restriction $\leq$ satisfies Theorem 2.3 (2) (a).

To prove the condition (b), take elements $w_X \in S_\sigma$ and $w_1, w_2 \in W_\sigma$ satisfying $\ell_\sigma(w_X w_1) \leq \ell_\sigma(w_1)$ and $\ell_\sigma(w_X w_2) \leq \ell_\sigma(w_2)$. Let $w_X = s_k \cdots s_1$ be a reduced expression. Then $s_1, s_2, \cdots, s_k \in X$. Since

$\ell_\sigma(w_X w_1) \leq \ell_\sigma(w_1)$, we see that $\ell(w_X w_1) = \ell(w_1) - \ell(w_X)$ by Corollary 2.7 (2). Thus $\ell(s_i \cdots s_1 w_1) \leq \ell(s_{i-1} \cdots s_1 w_1)$ for $1 \leq i \leq k$. Similarly, $\ell(s_i \cdots s_1 w_2) \leq \ell(s_{i-1} \cdots s_1 w_2)$ for $1 \leq i \leq k$. In particular, since X is a $\langle \sigma \rangle$-orbit, $\ell(s_1 w_2) \leq \ell(w_2)$ and Lemma 2.5 imply that $\ell(s_i w_2) \leq \ell(w_2)$ for $1 \leq i \leq k$. Now, using induction, we show that, for $1 \leq i \leq k$, the following condition $(**_i)$ holds:

$$w_1 \leq w_2 \quad \Leftrightarrow \quad s_i \cdots s_1 w_1 \leq w_2 \quad \Leftrightarrow \quad s_i \cdots s_1 w_1 \leq s_i \cdots s_1 w_2 \qquad (**_i)$$

If $i = 1$, then $\ell(s_1 w_1) \leq \ell(w_1)$, $\ell(s_1 w_2) \leq \ell(w_2)$ and Property $Z(s_1, w_1, w_2)$ imply that the condition $(**_1)$ holds.

Suppose that $i > 1$ and the condition $(**_{i-1})$ holds. Since $\ell(s_i s_{i-1} \cdots s_1 w_1) \leq \ell(s_{i-1} \cdots s_1 w_1)$ and $\ell(s_i w_2) \leq \ell(w_2)$, Property $Z(s_i, s_{i-1} \cdots s_1 w_1, w_2)$ implies that $s_{i-1} \cdots s_1 w_1 \leq w_2$ is equivalent to $s_i \cdots s_1 w_1 \leq w_2$. Thus $w_1 \leq w_2$ is equivalent to $s_i \cdots s_1 w_1 \leq w_2$. And since $\ell(s_i \cdots s_1 w_1) \leq \ell(s_{i-1} \cdots s_1 w_1)$ and $\ell(s_i \cdots s_1 w_2) \leq \ell(s_{i-1} \cdots s_1 w_2)$, Property $Z(s_i, s_{i-1} \cdots s_1 w_1, s_{i-1} \cdots s_1 w_2)$ implies that $s_{i-1} \cdots s_1 w_1 \leq s_{i-1} \cdots s_1 w_2$ is equivalent to $s_i \cdots s_1 w_1 \leq s_i \cdots s_1 w_2$. Thus $w_1 \leq w_2$ is equivalent to $s_i \cdots s_1 w_1 \leq s_i \cdots s_1 w_2$. Thus the condition $(**_i)$ holds.

It is follows from the condition $(**_k)$ that the restriction satisfies the condition (b) in Theorem 2.3 (2) for $J = \emptyset$.

□

3 A Coxeter group of type $\widetilde{C}_n$

3.1 Affine permutation of Type A_{N-1}

Let $W(\widetilde{A}_{N-1})$ be the group of affine permutations introduced in Section 1. Recall that an affine permutation π is determined by its values $\pi(1), \pi(2), \cdots, \pi(n)$. Therefore we denote $\pi \in W(\widetilde{A}_{N-1})$ by writing $\pi = [\pi(1), \cdots \pi(N)]$, and this notation of affine permutations is called the *window*.

For each integer i with $0 \leq i \leq N-1$, let s_i be the affine permutation defined by

$$s_i(j) = \begin{cases} j+1 & \text{if } j \equiv i \pmod N, \\ j-1 & \text{if } j \equiv i+1 \pmod N, \\ j & \text{otherwise.} \end{cases}$$

Put $S(\widetilde{A}_{N-1}) = \{s_0, s_1, \cdots, s_{N-1}\}$. Then G. Lusztig proved, in [8], that $(W(\widetilde{A}_{N-1}), S(\widetilde{A}_{N-1}))$ is a Coxeter system of type $\widetilde{A}_{N-1}$ with the Coxeter graph in Figure 1.

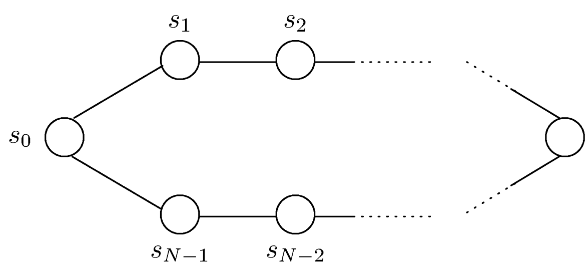


Figure 1: The Coxeter graph of type $\widetilde{A}_{N-1}$.

In the case of $(W(\widetilde{A}_{N-1}), S(\widetilde{A}_{N-1}))$, the length function ℓ is given by [2]:

$$\ell(\pi) = \sum_{1 \leq i < j \leq N} \left| \left\lfloor \frac{\pi(j) - \pi(i)}{N} \right\rfloor \right|,$$

where, for a rational number x, $\lfloor x \rfloor$ is the largest integer not exceeding x. Therefore *the descent set* $D(\pi)$ of π is given by the following:

$$D(\pi) \overset{\text{def}}{=} \{s \in S(\widetilde{A}_{N-1}) \mid \ell(\pi s) \leq \ell(\pi)\} = \{s_i \in S(\widetilde{A}_{N-1}) \mid \pi(i) > \pi(i+1)\} \qquad (3.1)$$

Let $J_0 = S(\widetilde{A}_{N-1}) \setminus \{s_0\}$. By (3.1), $W(\widetilde{A}_{N-1})^{J_0}$ is a subset of $W(\widetilde{A}_{N-1})$ that consists of the elements π satisfying

$$\pi(1) < \pi(2) < \cdots < \pi(N). \tag{3.2}$$

For each $w \in W(\widetilde{A}_{N-1})_{J_0}$, we have

$$\{w(1), w(2), \cdots, w(N)\} = \{1, 2, \cdots, N\}.$$

Thus, given $\pi = [\pi(1), \cdots, \pi(N)] \in W(\widetilde{A}_{N-1})$, the window of πw can be written by permuting $\pi(1), \pi(2), \cdots, \pi(N)$. From (3.2), this implies that the window of π^{J_0} is given as follows:

Lemma 3.1 *For $\pi = [\pi(1), \cdots, \pi(N)] \in W(\widetilde{A}_{N-1})$, the window of π^{J_0} is given by sorting $\pi(1), \cdots, \pi(N)$ in increasing order.* □

3.2 $W(\widetilde{C}_n)$ as a subgroup of $W(\widetilde{A}_{2n-1})$

Let $N = 2n$ and we take a Coxeter group of $W(\widetilde{A}_{2n-1})$.

Let σ be a Coxeter graph automorphism of $(W(\widetilde{A}_{2n-1}), S(\widetilde{A}_{2n-1}))$ given by

$$\sigma(s_i) = \begin{cases} s_0 & i = 0, \\ s_{2n-i} & 1 \leq i \leq 2n-1. \end{cases} \tag{3.3}$$

This action of σ on $S(\widetilde{A}_{2n-1})$ is described as Figure 2.

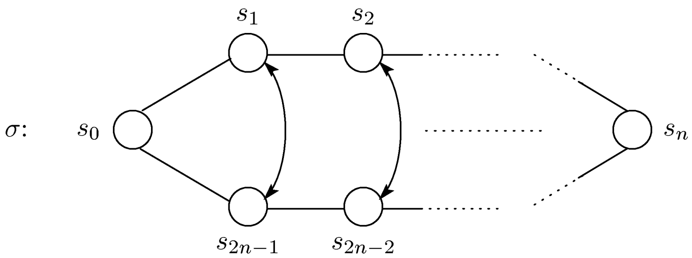


Figure 2: The action of σ on $S(\widetilde{A}_{2n-1})$.

Then σ extends to an automorphism on $W(\widetilde{A}_{2n-1})$ as follows:

Proposition 3.2 *For each $\pi \in W(\widetilde{A}_{2n-1})$, this Coxeter graph automorphism σ defined as* (3.3) *maps $\pi = [\pi(1), \pi(2), \cdots, \pi(2n+1)]$ to the following element:*

$$\sigma(\pi) = [2n+1-\pi(2n), 2n+1-\pi(2n-1), \cdots, 2n+1-\pi(1)]. \tag{3.4}$$

Proof. We can prove this proposition by using induction on $r = \ell(\pi)$. □

For this automorphism σ, Proposition 3.2 implies that

$$\begin{aligned} W(\widetilde{A}_{2n-1})_\sigma &= \{\pi \in W(\widetilde{A}_{2n-1}) \mid \sigma(\pi) = \pi\} \\ &= \{\pi \in W(\widetilde{A}_{2n-1}) | \ \pi(i) + \pi(2n+1-i) = 2n+1 \ (\forall i \in \mathbb{Z})\} \\ &= W(\widetilde{C}_n). \end{aligned}$$

We set $S(\widetilde{C}_n) = \{w_{X_0}, w_{X_1}, w_{X_2}, \cdots, w_{X_{n-1}}, w_{X_n}\}$, where

$$w_{X_i} = \begin{cases} s_0 & i = 0, \\ s_i s_{2n-i} & 1 \leq i \leq n-1, \\ s_n & i = n. \end{cases}$$

Then Theorem 1.1 and Theorem 1.2 imply that we obtain the following result:

Theorem 3.3 *$(W(\widetilde{C}_n), S(\widetilde{C}_n))$ is a Coxeter system of type $\widetilde{C}_n$. And the Bruhat ordering on $W(\widetilde{C}_n)$ is the restriction of the Bruhat ordering on $W(\widetilde{A}_{2n-1})$ to $W(\widetilde{C}_n)$.*

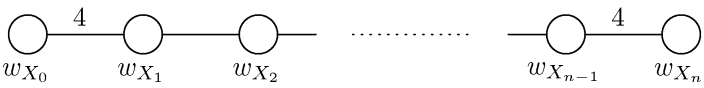


Figure 3: The Coxeter graph of type $\widetilde{C}_n$.

Proof. It remains to show that type of $(W(\widetilde{C}_n), S(\widetilde{C}_n))$ is $\widetilde{C}_n$ and it is easy. □

Let $I_0 = S(\widetilde{C}_n) \setminus \{w_{X_0}\}$. Then Lemma 2.5 leads to

$$\begin{aligned} w \in W(\widetilde{C}_n)^{I_0} &\iff \ell_\sigma(w w_{X_i}) \geq \ell_\sigma(w) \quad \text{for all } i = 1, 2, \cdots, n. \\ &\iff \ell(w s_j) \geq \ell(w) \text{ for all } j = 1, 2, \cdots, 2n+1. \\ &\iff w \in W(\widetilde{A}_{2n-1})^{J_0} \end{aligned}$$

for $w \in W(\widetilde{C}_n)$. Thus we have the following lemma:

Lemma 3.4

$$W(\widetilde{C}_n)^{I_0} = W(\widetilde{C}_n) \cap W(\widetilde{A}_{2n-1})^{J_0}.$$

That is,

$$W(\widetilde{C}_n)^{I_0} = \{w \in W(\widetilde{C}_n) \mid w(1) < w(2) < \cdots < w(2n)\}.$$

□

4 The Bruhat ordering on $W(\widetilde{A}_{N-1})^{J_0}$

By Theorem 3.3 and Lemma 3.4, the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ is closely related to the Bruhat ordering on $W(\widetilde{A}_{2n-1})^{J_0}$ in a certain sense. A. Lascoux, in [7], described the Bruhat ordering on $W(\widetilde{A}_{N-1})^{J_0}$ by N-core encoding. In this section, we review Lascoux's encoding of $W(\widetilde{A}_{N-1})^{J_0}$ and give a description of the Bruhat ordering on it.

By Lemma 3.1, $W(\widetilde{A}_{N-1})^{J_0}$ consists of elements $\pi \in W(\widetilde{A}_{N-1})$ satisfying

$$\pi(1) < \pi(2) < \cdots < \pi(N).$$

Let λ be a Young diagram. To each box (i, j) in λ, we associate an integer p (called color) defined by

$$p \equiv j - i \pmod{N} \qquad 0 \leq p \leq N-1.$$

We introduce the notation of addable or removable i-corner of a Young diagram as follows.

(a) A box in λ is called a *removable corner* if we get another Young diagram by deleting it from λ. If a removable corner of λ has the color i, we call it a *removable i-corner* of λ.

(b) A box in λ is called an *addable corner* if we get another Young diagram by adding it to λ. If an addable corner of λ has the color i, we call it an *addable i-corner.*

Example 4.1 *If $N = 5$ and $\lambda = (4, 3, 1, 1, 1)$, then the coloring for λ and removable/addable corners are given by* Figure 4.

If a diagram λ has no hook whose length is a multiple of N, then λ is called an *N-core*. For example, $\lambda = (4, 3, 1, 1, 1)$ is a 5-core. From the definition, λ is an N-core if and only if λ satisfies the condition $(C_{N\text{-core}})$.

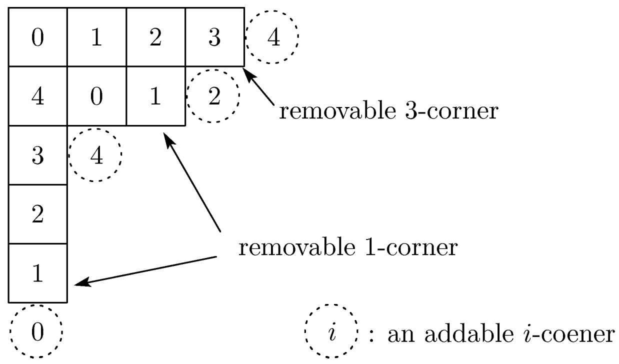


Figure 4: The coloring for $\lambda = (4,3,1,1,1)$ $(N = 5)$, removable corners and addable corners.

<u>$(C_{N\text{-core}})$</u> For each integer i with $0 \leq i \leq N-1$, λ can not have both removable i-corners and addable i-corners.

That is,

(1) If λ has a removable i-corner, then λ can not have any addable i-corners.

(2) If λ has an addable i-corner, then λ can not have any removable i-corners.

Let $\mathcal{C}[N]$ be the set of all N-cores. We define the action of $W(\widetilde{A}_{N-1})$ to $\mathcal{C}[N]$ as follows:

First, for an N-core λ and $s_i \in S(\widetilde{A}_{N-1})$, we define a new diagram $s_i \cdot \lambda$ as follows:

- If λ has removable i-corners, then $s_i \cdot \lambda$ is the diagram obtained by removing all of them from λ.
- If λ has addable i-corners, then $s_i \cdot \lambda$ is the diagram obtained by adding boxes to the corners of λ.
- If λ doesn't have either removable or addable i-corners, then $s_i \cdot \lambda = \lambda$.

Then it is obvious from the condition $C_{N\text{-core}}$ that $s_i \cdot \lambda$ is an N-core if λ is an N-core. Moreover, we can easily deduce that

$$s_i \cdot (s_i \cdot \lambda) = \lambda \qquad (0 \leq \forall i \leq N-1), \tag{4.1}$$

$$s_i \cdot (s_j \cdot \lambda) = s_j \cdot (s_i \cdot \lambda), \qquad \text{if } |i-j| \geq 2, \tag{4.2}$$

and

$$s_i \cdot (s_{i+1} \cdot (s_i \cdot \lambda)) = s_{i+1} \cdot (s_i \cdot (s_{i+1} \cdot \lambda)) \qquad (0 \leq \forall i \leq N-1), \tag{4.3}$$

where we set $s_N = s_0$.

For general $\pi \in W(\widetilde{A}_{N-1})$, we take an expression $\pi = s_{i_1} s_{i_2} \cdots s_{i_r}$ with respect to $S(\widetilde{A}_{N-1})$, and define

$$\pi \cdot \lambda = s_{i_1} \cdot (s_{i_2} \cdot (\cdots (s_{i_r} \cdot \lambda) \cdots)).$$

This action of $W(\widetilde{A}_{N-1})$ on $\mathcal{C}[N]$ is well-defined from (4.1), (4.2) and (4.3).

Example 4.2 *For $N = 5$ and 5-core $\lambda = (4, 3, 1, 1, 1)$,*

$$
\begin{array}{llll}
0 & 1 & 2 & 3 \\
4 & 0 & 1 & \\
3 & & & \\
2 & & & \\
1 & & &
\end{array}
\xrightarrow{s_4}
\begin{array}{lllll}
0 & 1 & 2 & 3 & 4 \\
4 & 0 & 1 & & \\
3 & 4 & & & \\
2 & & & & \\
1 & & & &
\end{array}
\xrightarrow{s_3}
\begin{array}{lllll}
0 & 1 & 2 & 3 & 4 \\
4 & 0 & 1 & & \\
3 & 4 & & & \\
2 & 3 & & & \\
1 & & & &
\end{array}
\xrightarrow{s_1}
\begin{array}{lllll}
0 & 1 & 2 & 3 & 4 \\
4 & 0 & & & \\
3 & 4 & & & \\
2 & 3 & & &
\end{array}
$$

Therefore, for $\pi = s_1 s_3 s_4 = [2, 1, 4, 5, 3]$, we have $\pi \cdot \lambda = (5, 2, 2, 2)$.

This action induces the map $\mathcal{C}_0 : W(\widetilde{A}_{N-1}) \to \mathcal{C}[N]$ as follows:

$$\mathcal{C}_0(\pi) = \pi \cdot \emptyset \in \mathcal{C}[N]$$

where the diagram $\emptyset$ is the empty diagram, i.e., the partition of 0. By the definition of $\mathcal{C}_0$, it is obvious that the map $\mathcal{C}_0$ is compatible with the action of $W(\widetilde{A}_{N-1})$. That is, for $\tau, \pi \in W(\widetilde{A}_{N-1})$,

$$\tau \cdot \mathcal{C}_0(\pi) = \mathcal{C}_0(\tau\pi). \tag{4.4}$$

Then the equation (4.4) leads to the following properties:

Proposition 4.3 (1) *This map $\mathcal{C}_0$ is a surjection from $W(\widetilde{A}_{N-1})$ to $\mathcal{C}[N]$.*

(2) *For $\pi \in W(\widetilde{A}_{N-1})$, π fixes $\emptyset$ if and only if $\pi \in W(\widetilde{A}_{N-1})_{J_0}$. Moreover, for $\pi \in W(\widetilde{A}_{N-1})$, we have*

$$\mathcal{C}_0(\pi) = \pi^{J_0} \cdot \emptyset = \mathcal{C}_0(\pi^{J_0}).$$

Proof. The statement (1) is clear. So it is sufficient to show (2). As for the proof of (2), we need some arguments, see G. James–A. Kerber [6] or M. Nanba [11]. □

Corollary 4.4 *The restriction of $\mathcal{C}_0$ to $W(\widetilde{A}_{N-1})^{J_0}$ is a bijection from $W(\widetilde{A}_{N-1})^{J_0}$ to $\mathcal{C}[N]$.* □

For this bijection, A. Lascoux deduced the following result:

Theorem 4.5 (A. Lascoux [7])
For $\pi, \tau \in W(\widetilde{A}_{N-1})^{J_0}$,

$$\pi \leq \tau \iff \mathcal{C}_0(\pi) \subseteq \mathcal{C}_0(\tau).$$

□

A. Lascoux proved this theorem by using Theorem 2.3 in Section 2.1 for $J = J_0$.

5 The Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$

By Theorem 3.3 and Lemma 3.4, $W(\widetilde{C}_n)^{I_0}$ is the intersection of a Coxeter group $W(\widetilde{C}_n)$ $(= W(\widetilde{A}_{2n-1})_\sigma)$ and $W(\widetilde{A}_{2n-1})^{J_0}$. Then Theorem 1.2 implies that the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ is the restriction of the Bruhat ordering on $W(\widetilde{A}_{2n-1})^{J_0}$ to $W(\widetilde{C}_n)^{I_0}$. Therefore the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ can be described by using the map $\mathcal{C}_0$ on $W(\widetilde{A}_{2n-1})^{J_0}$. In this section, we give a combinatorial description of the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$.

We study a relation between the automorphism σ and the map $\mathcal{C}_0$. For a Young diagram λ, let λ^c be the conjugate diagram of λ.

Proposition 5.1 *For $s_i \in S(\widetilde{A}_{2n-1})$ and an $2n$-core λ, then we have $(s_i \cdot \lambda)^c = s_{2n-i} \cdot \lambda^c$. That is, $(s_i \cdot \lambda)^c = \sigma(s_i) \cdot \lambda^c$.*

Proof. Since the action of s_i to λ has three ways as the case may be, we will show this proposition for each case. It is obvious that the color of the box (q, p) in λ^c is $2n - i$ if the box (p, q) is colored i in λ.

Case 1. If $s_i \cdot \lambda = \lambda$, then λ can't have either removable or addable i-corner. This implies that λ^c can't have either removable or addable $(2n - i)$-corner. Therefore we have $s_{2n-i} \cdot \lambda^c = \lambda^c$ and

$$(s_i \cdot \lambda)^c = \lambda^c = s_{2n-i} \cdot \lambda^c.$$

Case 2. If $\lambda \supsetneq s_i \cdot \lambda$, then λ has some removable i-corners and the action s_i to λ is removing them from λ. Therefore λ^c has some removable $(2n-i)$-corners and the action of s_{2n-i} to λ^c is removing them from λ^c. In particular, the component of each removable $(2n - i)$-corner in λ^c is conjugate to the one of the removable i-corners in λ. This forces $(s_i \cdot \lambda)^c$ to be $s_{2n-i} \cdot \lambda^c$.

Case 3. If $\lambda \subsetneq s_i \cdot \lambda$, then the action of s_i to λ is adding boxes to all addable i-corners in λ. By the same argument as Case 2, we can deduce that $(s_i \cdot \lambda)^c = s_{2n-i} \cdot \lambda^c$.

□

Corollary 5.2 *For $\pi \in W(\widetilde{A}_{N-1})$,*

$$\mathcal{C}_0(\sigma(\pi)) = \mathcal{C}_0(\pi)^c.$$

□

Then we have a combinatorial description of the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ from Lemma 3.4, Theorem 4.5 and Corollary 5.2:

Theorem 5.3 *Let $\mathrm{Sym}\mathcal{C}[2n] = \{\lambda \in \mathcal{C}[2n] \ : \ \lambda^c = \lambda\}$ and $\leq_\sigma$ be the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ with respect to $S(\widetilde{C}_n)$. Then we have*

(1) *The map $\mathcal{C}_0 : W(\widetilde{C}_n)^{I_0} \longrightarrow \mathrm{Sym}\mathcal{C}[2n]$ is a bijection.*

(2) *For $w, v \in W(\widetilde{C}_n)^{I_0}$,*

$$w \leq_\sigma v \iff \mathcal{C}_0(w) \subseteq \mathcal{C}_0(v).$$

□

6 The Bruhat ordering on $W(\widetilde{C}_n)$

In this section, we use the combinatorial description of the Bruhat ordering on $W(\widetilde{C}_n)^{I_0}$ in the previous section, to find a combinatorial description of the Bruhat ordering on the whole group $W(\widetilde{C}_n)$ ($= W(\widetilde{A}_{2n-1})_\sigma$). In order to attain this purpose, the following theorem is very useful:

Theorem 6.1 (V. Deodhar [4])
Let $\mathcal{H}$ be a family of subsets of S such that $\emptyset \notin \mathcal{H}$ and $\bigcap_{J \in \mathcal{H}} J = \emptyset$. Then $w, v \in W$ are $w \leq v$ in the Bruhat order on W if and only if $w^J \leq v^J$ in the Bruhat order on W^J for all $J \in \mathcal{H}$. □

Now, we consider a Coxeter system $(W(\widetilde{A}_{N-1}), S(\widetilde{A}_{N-1}))$ of type $\widetilde{A}_{N-1}$.

Let $J_k = S(\widetilde{A}_{N-1}) \setminus \{s_k\}$, for each integer k with $1 \leq k \leq N - 1$, then by the same argument in case of J_0, $W(\widetilde{A}_{N-1})^{J_k}$ consists of all affine permutations which satisfy

$$\pi(k+1) < \pi(k+2) < \cdots < \pi(N) < \pi(1+N) < \pi(2+N) < \cdots < \pi(k+N).$$

Therefore, for $\tau \in W(\widetilde{A}_{N-1})$, the window of τ^{J_k} satisfies the following condition:

(i) $\tau^{J_k}(1) < \tau^{J_k}(2) < \cdots < \tau^{J_k}(k)$ and $\tau^{J_k}(k+1) < \tau^{J_k}(k+2) < \cdots < \tau^{J_k}(N)$,

(ii) $\{\tau^{J_k}(1), \tau^{J_k}(2), \cdots, \tau^{J_k}(k)\} = \{\tau(1), \tau(2), \cdots, \tau(k)\}$ and $\{\tau^{J_k}(k+1), \tau^{J_k}(k+2), \cdots, \tau^{J_k}(N)\} = \{\tau(k+1), \tau(k+2), \cdots, \tau(N)\}$.

Now, we consider a Coxeter graph automorphism ω of $S(\widetilde{A}_{N-1})$ as follows:

$$\omega(s_i) = \begin{cases} s_{N-1} & \text{if } i = 0, \\ s_{i-1} & \text{otherwise.} \end{cases}$$

Then we can prove the following Proposition by using the induction with respect to the length:

Proposition 6.2 *This automorphism ω extends to a group automorphism on $W(\widetilde{A}_{N-1})$ as follows: For $\pi \in W(\widetilde{A}_{N-1})$,*

$$\omega(\pi) = [\pi(2) - 1, \pi(3) - 1, \cdots, \pi(N) - 1, \pi(1) - 1 + N].$$

$\square$

Moreover, it is obvious that $\omega(J_k) = J_{k-1}$ for $1 \leq k \leq N-1$. Then Proposition 2.1 and 2.2 imply the following proposition:

Proposition 6.3 *Let k be an integer with $0 \leq k \leq N-1$, then the map $\omega^k : W(\widetilde{A}_{N-1})^{J_k} \to W(\widetilde{A}_{N-1})^{J_0}$ is an order preserving bijection.* $\square$

For each integer k with $0 \leq k \leq N-1$, we define the bijection map $\mathcal{C}_k : W(\widetilde{A}_{N-1})^{J_k} \to \mathcal{C}[N]$ by

$$\mathcal{C}_k = \mathcal{C}_0 \circ \omega^k.$$

Then Theorem 6.1 and Proposition 6.3 lead to the following theorem:

Theorem 6.4 *Let $\leq$ be the Bruhat ordering on $W(\widetilde{A}_{N-1})$ with respect to $S(\widetilde{A}_{N-1})$. For $\pi, \tau \in W(\widetilde{A}_{N-1})$,*

$$\pi \leq \tau \iff \mathcal{C}_k(\pi^{J_k}) \subseteq \mathcal{C}_k(\tau^{J_k})$$

for $0 \leq k \leq N-1$. $\square$

Let $N = 2n$, we consider the Coxeter graph automorphism σ defined by (3.3). Since $\sigma(J_k) = J_{2n-k}$, we have

$$\omega^k \circ \sigma = \sigma \circ \omega^{2n-k}$$

for $0 \leq k \leq 2n-1$. Therefore if $\pi \in W(\widetilde{C}_n)$, then we have

$$\mathcal{C}_{2n-k}(\pi^{J_{2n-k}}) = \mathcal{C}_k(\pi^{J_k})^c \tag{6.1}$$

for $0 \leq k \leq 2n-1$. In particular, we should note that the equation (6.1) is sufficient for $\pi \in W(\widetilde{C}_n)$.

Finally, we can give a combinatorial description of the Bruhat ordering on $W(\widetilde{C}_n)$, i.e., a Coxeter group of type $\widetilde{C}_n$:

Theorem 6.5 *Let $\leq_\sigma$ be the Bruhat ordering on $W(\widetilde{C}_n)$ with respect to $S(\widetilde{C}_n)$. For $w, v \in W(\widetilde{C}_n) \subset W(\widetilde{A}_{2n-1})$,*

$$w \leq_\sigma v \iff \mathcal{C}_k(w^{J_k}) \subseteq \mathcal{C}_k(v^{J_k})$$

for all integers k with $0 \leq k \leq n$. $\square$

References

[1] S. Ariki, "Representations of Quantum Algebras and Combinatorics of Young Tableaux", University Lecture Series, Vol. 26, A.M.S., (2002).

[2] A.Björner and F.Brenti, "Affine permutations of type A", Electronic Journal of Combinatorics 3 no.2 (1996), ♯R18.

[3] N. Bourbaki, “Groupes et algèbres de Lie, Chapitres 4, 5 et 6 ”, Hermann, Paris, (1968).

[4] V. V. Deodhar, “Some characterization of Bruhat ordering on a Coxeter group and determination of the relative Möbius function”, Invent. math. 39 (1977), 187–198.

[5] J. E. Humphreys, “Reflection Groups and Coxeter Groups”, Cambridge University Press, (1990).

[6] G. James and A. Kerber, “The Representation Theory of the Symmetric Group”, Encyclopedia of Mathmatics, Addisson–Wesleey, Reading, MA, (1991).

[7] A. Lascoux, “Ordering the affine symmetric group”, Algebraic Combinatorics and Applications, Springer, (2001), 219–231.

[8] G. Lusztig, “Some examples of square integrable representations of semisimple p-adic group”, Trans. Amer. Math. Soc. 277 (1983), 623–653.

[9] I. G. Macdonald, “Symmetric Functions and Hall Polynomials, Second Edition” Claredon Press. (1995).

[10] M. Nanba, “Bruhat order on the fixed-point subgroup by a Coxeter graph automorphism”, Journal of Algebra 285, (2005), 470–480.

[11] M. Nanba, “The Bruhat ordring on the Coxeter group of type $\widetilde{C}_n$”, Preprint.

[12] R.Steinberg, “Endmorphisms of Linear Algebraic Groups”, Memoirs of American Mathmatical Society No.80, (1968).

Raising and Lowering Maps and Modules for the Quantum Affine Algebra $U_q(\widehat{\mathfrak{sl}}_2)$

Darren Neubauer

Abstract

Let $\mathbb{K}$ denote an algebraically closed field and let q denote a nonzero scalar in $\mathbb{K}$ that is not a root of unity. Let V denote a vector space over $\mathbb{K}$ with finite positive dimension. Let $V_0, V_1, \ldots, V_d$ denote a sequence of nonzero subspaces whose direct sum is V. Suppose $R : V \to V$ and $L : V \to V$ are linear transformations such that

(i) $RV_i \subseteq V_{i+1} \qquad (0 \le i \le d-1), \qquad RV_d = 0,$

(ii) $LV_i \subseteq V_{i-1} \qquad (1 \le i \le d), \qquad LV_0 = 0,$

(iii) for $0 \le i \le d/2$ the restriction $R^{d-2i}|_{V_i} : V_i \to V_{d-i}$ is a bijection,

(iv) for $0 \le i \le d/2$ the restriction $L^{d-2i}|_{V_{d-i}} : V_{d-i} \to V_i$ is a bijection,

(v) $R^3L - [3]R^2LR + [3]RLR^2 - LR^3 = 0,$

(vi) $L^3R - [3]L^2RL + [3]LRL^2 - RL^3 = 0,$

where $[3] = (q^3 - q^{-3})/(q - q^{-1})$. Let $K : V \to V$ be the linear transformation such that, for $0 \le i \le d$, V_i is an eigenspace for K with eigenvalue q^{2i-d}. We show that there exists a unique $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V such that each of $R - e_1^-$, $L - e_0^-$, $K - K_0$, and $K^{-1} - K_1$ vanish on V, where e_1^-, e_0^-, K_0, K_1 are Chevalley generators for $U_q(\widehat{\mathfrak{sl}}_2)$. We determine which $U_q(\widehat{\mathfrak{sl}}_2)$-modules arise from our construction.

1 The quantum affine algebra $U_q(\widehat{\mathfrak{sl}}_2)$

Throughout this paper $\mathbb{K}$ will denote an algebraically closed field. We fix a nonzero scalar $q \in \mathbb{K}$ that is not a root of unity. We will use the following notation.

$$[n] = \frac{q^n - q^{-n}}{q - q^{-1}}, \qquad n = 0, 1, \ldots$$

We now recall the definition of $U_q(\widehat{\mathfrak{sl}}_2)$.

Keywords: Quantum group, quantum affine algebra, affine Lie algebra $\widehat{\mathfrak{sl}}_2$, raising and lowering maps, tridiagonal pair.

2000 Mathematics Subject Classification. Primary: 17B37. Secondary: 16W35, 20G42, 81R50

Definition 1.1 [2, p. 262] The quantum affine algebra $U_q(\widehat{\mathfrak{sl}}_2)$ is the unital associative $\mathbb{K}$-alegbra with generators $e_i^{\pm}$, $K_i^{\pm 1}$, $i \in \{0, 1\}$ which satisfy the following relations:

$$
\begin{aligned}
K_i K_i^{-1} &= K_i^{-1} K_i = 1, & (1)\\
K_0 K_1 &= K_1 K_0, & (2)\\
K_i e_i^{\pm} K_i^{-1} &= q^{\pm 2} e_i^{\pm}, & (3)\\
K_i e_j^{\pm} K_i^{-1} &= q^{\mp 2} e_j^{\pm}, \qquad i \neq j, & (4)\\
e_i^{+} e_i^{-} - e_i^{-} e_i^{+} &= \frac{K_i - K_i^{-1}}{q - q^{-1}}, & (5)\\
e_0^{\pm} e_1^{\mp} &= e_1^{\mp} e_0^{\pm}, & (6)
\end{aligned}
$$

$$
(e_i^{\pm})^3 e_j^{\pm} - [3](e_i^{\pm})^2 e_j^{\pm} e_i^{\pm} + [3] e_i^{\pm} e_j^{\pm} (e_i^{\pm})^2 - e_j^{\pm} (e_i^{\pm})^3 = 0 \qquad i \neq j. \tag{7}
$$

We call $e_i^{\pm}$, $K_i^{\pm 1}$, $i \in \{0, 1\}$ the *Chevalley generators* for $U_q(\widehat{\mathfrak{sl}}_2)$ and refer to (7) as the *q-Serre relations*.

2 The Main Theorem

In this section we state our main result. We begin with two definitions.

Definition 2.1 Let V be a vector space over $\mathbb{K}$ with finite positive dimension. By a *decomposition* of V we mean a sequence $V_0, V_1, \ldots, V_d$ consisting of nonzero subspaces of V such that $V = \sum_{i=0}^{d} V_i$ (direct sum). For notational convenience we let $V_{-1} = 0, V_{d+1} = 0$.

Definition 2.2 Let V be a vector space over $\mathbb{K}$ with finite positive dimension. Let $V_0, V_1, \ldots, V_d$ be a decomposition of V. Let $K : V \to V$ denote the linear transformation such that, for $0 \leq i \leq d$, V_i is an eigenspace for K with eigenvalue q^{2i-d}. We refer to K as *the linear transformation corresponding to the decomposition* $V_0, V_1, \ldots, V_d$.

Note 2.3 With reference to Definition 2.2, we note that K is invertible. Moreover, for $0 \leq i \leq d$, V_i is the eigenspace for K^{-1} with eigenvalue q^{d-2i}. We observe that K^{-1} is the linear transformation corresponding to the decomposition $V_d, V_{d-1}, \ldots, V_0$.

We will be concerned with the following situation.

Assumption 2.4 Let V be a vector space over $\mathbb{K}$ with finite positive dimension. Let $V_0, V_1, \ldots V_d$ be a decomposition of V. Let K denote the linear transformation corresponding to $V_0, V_1, \ldots V_d$ as in Definition 2.2. Let $R : V \to V$ and $L : V \to V$ be linear transformations such that

(i) $RV_i \subseteq V_{i+1} \qquad (0 \leq i \leq d)$,

(ii) $LV_i \subseteq V_{i-1} \qquad (0 \le i \le d)$,

(iii) for $0 \le i \le d/2$ the restriction $R^{d-2i}|_{V_i} : V_i \to V_{d-i}$ is a bijection,

(iv) for $0 \le i \le d/2$ the restriction $L^{d-2i}|_{V_{d-i}} : V_{d-i} \to V_i$ is a bijection,

(v) $R^3L - [3]R^2LR + [3]RLR^2 - LR^3 = 0$,

(vi) $L^3R - [3]L^2RL + [3]LRL^2 - RL^3 = 0$.

We now state our main result.

Theorem 2.5 *Adopt Assumption 2.4. Then there exists a unique $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V such that $(R - e_1^-)V = 0$, $(L - e_0^-)V = 0$, $(K - K_0)V = 0$, $(K^{-1} - K_1)V = 0$, where e_1^-, e_0^-, K_0, K_1 are Chevelley generators for $U_q(\widehat{\mathfrak{sl}}_2)$.*

Theorem 2.5 is related to a result of G. Benkart and P. Terwilliger [1]. In [1] the authors adopt Assumption 2.4(i),(ii),(v),(vi). They replace Assumption 2.4(iii),(iv) with the assumption that V is irreducible as a (K, R, L)-module. From this assumption they obtain a $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V as in Theorem 2.5. The $U_q(\widehat{\mathfrak{sl}}_2)$-module structure that they obtain is irreducible while the $U_q(\widehat{\mathfrak{sl}}_2)$-module structure given by Theorem 2.5 is not necessarily irreducible. As far as we know Theorem 2.5 does not imply the result in [1] nor does the result in [1] imply Theorem 2.5. Both this paper and [1] use an adaptation of a construction which T. Ito and P. Terwilliger used to get $U_q(\widehat{\mathfrak{sl}}_2)$-modules from a certain type of tridiagonal pair [4]. In fact, the motivation for our work on $U_q(\widehat{\mathfrak{sl}}_2)$-modules came from the study of tridiagonal pairs [3].

The plan for the paper is as follows. In section 3 we present an overview of the argument used to prove Theorem 2.5. In sections 4 through 9 we summarize the proof of Theorem 2.5. In sections 10 and 11 we determine which $U_q(\widehat{\mathfrak{sl}}_2)$-modules arise from the construction in Theorem 2.5. Since the rest of the paper is meant to provide a summary, many of the proofs are omitted.

3 An outline of the proof of Theorem 2.5

We begin by adopting Assumption 2.4. To start the construction of the $U_q(\widehat{\mathfrak{sl}}_2)$-action on V we require that the linear transformations $R - e_1^-$, $L - e_0^-$, $K^{\pm 1} - K_0^{\pm 1}$, $K^{\pm 1} - K_1^{\mp 1}$ vanish on V. This gives the actions of the elements $e_1^-, e_0^-, K_0^{\pm 1}, K_1^{\pm 1}$ on V. We define the actions of e_0^+, e_1^+ on V as follows. First we prove that $K + R$ and $K^{-1} + L$ are diagonalizable on V. Then we show that the set of distinct eigenvalues of both $K + R$ and $K^{-1} + L$ on V is $\{ q^{2i-d} \,|\, 0 \le i \le d \}$. For $0 \le i \le d$, we let W_i (resp. W_i^*) denote the eigenspace of $K + R$ (resp. $K^{-1} + L$) on V associated with the eigenvalue q^{2i-d}. Then $W_0, \ldots, W_d$

(resp. $W_0^*, \ldots, W_d^*$) is a decomposition of V. We show that the decomposition $V_0, V_1, \ldots, V_d$ satisfies

$$\begin{aligned} V_0 + \cdots + V_i &= W_{d-i}^* + \cdots + W_d^* && (0 \le i \le d) \\ V_i + \cdots + V_d &= W_i + \cdots + W_d && (0 \le i \le d). \end{aligned}$$

Then for $0 \le i \le d$ we define subspaces $Z_i = (W_0 + \cdots + W_{d-i}) \cap (W_{d-i}^* + \cdots + W_d^*)$. We argue that $Z_0, Z_1, \ldots, Z_d$ is a decomposition of V and that

$$\begin{aligned} Z_0 + \cdots + Z_i &= W_{d-i}^* + \cdots + W_d^* && (0 \le i \le d) \\ Z_i + \cdots + Z_d &= W_0 + \cdots + W_{d-i} && (0 \le i \le d). \end{aligned}$$

Next for $0 \le i \le d$ we define subspaces $Z_i^* = (W_{d-i} + \cdots + W_d) \cap (W_0^* + \cdots + W_{d-i}^*)$. We argue that $Z_0^*, Z_1^*, \ldots, Z_d^*$ is a decomposition of V and that

$$\begin{aligned} Z_0^* + \cdots + Z_i^* &= W_{d-i} + \cdots + W_d && (0 \le i \le d) \\ Z_i^* + \cdots + Z_d^* &= W_0^* + \cdots + W_{d-i}^* && (0 \le i \le d). \end{aligned}$$

We then define the linear transformation $B : V \to V$ (resp. $B^* : V \to V$) such that for $0 \le i \le d$, Z_i (resp. Z_i^*) is an eigenspace for B (resp. B^*) with eigenvalue q^{2i-d}. We let e_1^+ act on V as $I - K^{-1}B$ times $q^{-1}(q - q^{-1})^{-2}$. We let e_0^+ act on V as $I - KB^*$ times $q^{-1}(q - q^{-1})^{-2}$. Finally, we display some relations that are satisfied by $B, B^*, L, R, K^{\pm 1}$. Using these relations, we argue that the above actions of $e_0^\pm$, $e_1^\pm$, $K_0^{\pm 1}$, $K_1^{\pm 1}$ satisfy the defining relations for $U_q(\widehat{\mathfrak{sl}}_2)$. In this way, we obtain the required action of $U_q(\widehat{\mathfrak{sl}}_2)$ on V.

4 The linear transformations A and A^*

In this section we define and discuss two linear transformations that will be useful.

Definition 4.1 In reference to Assumption 2.4 let $A : V \to V$ and $A^* : V \to V$ denote the following linear transformations.

$$A = K + R, \qquad A^* = K^{-1} + L. \tag{8}$$

Lemma 4.2 *With reference to Definition 4.1 and Assumption 2.4 the following (i),(ii) hold.*

(i) $(A - q^{2i-d}I)V_i \subseteq V_{i+1}, \qquad 0 \le i \le d,$

(ii) $(A^* - q^{d-2i}I)V_i \subseteq V_{i-1}, \qquad 0 \le i \le d.$

Lemma 4.3 *With reference to Definition 4.1 and Assumption 2.4, the following (i),(ii) hold.*

(i) A is diagonalizable with eigenvalues $q^{-d}, q^{2-d}, \ldots, q^d$. Moreover, for $0 \le i \le d$, the dimension of the eigenspace for A associated with q^{2i-d} is equal to the dimension of V_i.

(ii) A^* *is diagonalizable with eigenvalues* $q^{-d}, q^{2-d}, \ldots, q^d$*. Moreover, for* $0 \le i \le d$*, the dimension of the eigenspace for* A^* *associated with* q^{2i-d} *is equal to the dimension of* V_{d-i}*.*

Proof: (i) We start by displaying the eigenvalues for A. Notice that the scalars q^{2i-d} $(0 \le i \le d)$ are distinct since q is not a root of unity. Using Lemma 4.2(i) we see that, with respect to an appropriate basis for V, A can be represented as a lower triangular matrix that has diagonal entries $q^{-d}, q^{2-d}, \ldots, q^d$, with q^{2i-d} appearing $\dim(V_i)$ times for $0 \le i \le d$. Hence for $0 \le i \le d$, q^{2i-d} is a root of the characteristic polynomial of A with multiplicity $\dim(V_i)$. It remains to show that A is diagonalizable. To do this we show that the minimal polynomial of A has distinct roots. Recall that $V_0, V_1, \ldots, V_d$ is a decomposition of V. Using Lemma 4.2(i) we see that $\prod_{i=0}^{d}(A - q^{2i-d}I)V = 0$. By this and since q^{2i-d} $(0 \le i \le d)$ are distinct we see that the minimal polynomial of A has distinct roots. We conclude that A is diagonalizable and the result follows.
(ii) Similar to (i). □

Definition 4.4 With reference to Definition 4.1 and Lemma 4.3, for $0 \le i \le d$, we let W_i (resp. W_i^*) be the eigenspace for A (resp. A^*) with eigenvalue q^{2i-d}. Using Lemma 4.3 we observe that $W_0, W_1, \ldots, W_d$ (resp. $W_0^*, W_1^*, \ldots, W_d^*$) is a decomposition of V.

Lemma 4.5 *With reference to Assumption 2.4 and Definition 4.4, the following (i)–(iii) hold.*

(i) $V_0 + \cdots + V_i = W_{d-i}^* + \cdots + W_d^*, \qquad 0 \le i \le d,$

(ii) $V_i + \cdots + V_d = W_i + \cdots + W_d, \qquad 0 \le i \le d,$

(iii) $V_i = (W_i + \cdots + W_d) \cap (W_{d-i}^* + \cdots + W_d^*), \qquad 0 \le i \le d.$

Proof: (i) Let i be given. Define $T = V_0 + \cdots + V_i$ and $S = W_{d-i}^* + \cdots + W_d^*$. We show that $T = S$. First we show that $S \subseteq T$. Let $X = \prod_{h=0}^{d-i-1}(A^* - q^{2h-d}I)$. Recall that $W_0^*, W_1^*, \ldots, W_d^*$ is a decomposition of V and so we have $XV = S$. By Lemma 4.2(ii) we have $XV_j \subseteq T$ for $0 \le j \le d$. By this and since $V_0, V_1, \ldots V_d$ is a decomposition of V we find that $XV \subseteq T$. By these comments $S \subseteq T$. By Lemma 4.3(ii) and Definition 4.4 we have $\dim(W_{d-i}^*)$=$\dim(V_i)$. Thus $\dim(S)$=$\dim(T)$. We conclude $T = S$ and the result follows.
(ii) Similar to (i).
(iii) Immediate from (i),(ii) and the fact that $V_0, V_1, \ldots, V_d$ is a decomposition of V. □

5 The Subspaces Z_i, Z_i^*

Definition 5.1 With reference to Definition 4.4, for $0 \le i \le d$, we let Z_i denote the following subspace of V.

$$Z_i = (W_0 + \cdots + W_{d-i}) \cap (W_{d-i}^* + \cdots + W_d^*).$$

Using Assumption 2.4(i),(iii) and Lemma 4.5 the following theorem can be proven.

Theorem 5.2 *With reference to Definition 5.1, the following (i)–(iii) hold.*

(i) $Z_0, Z_1, \cdots, Z_d$ *is a decomposition of* V,

(ii) $Z_0 + \cdots + Z_i = W^*_{d-i} + \cdots + W^*_d, \qquad 0 \leq i \leq d,$

(iii) $Z_i + \cdots + Z_d = W_0 + \cdots + W_{d-i}, \qquad 0 \leq i \leq d.$

Definition 5.3 With reference to Definition 4.4, for $0 \leq i \leq d$, we let Z^*_i denote the following subspace of V.

$$Z^*_i = (W_{d-i} + \cdots + W_d) \cap (W^*_0 + \cdots + W^*_{d-i}).$$

Theorem 5.4 *With reference to Definition 5.3, the following (i)–(iii) hold:*

(i) $Z^*_0, Z^*_1, \cdots, Z^*_d$ *is a decomposition of* V,

(ii) $Z^*_0 + \cdots + Z^*_i = W_{d-i} + \cdots + W_d, \qquad 0 \leq i \leq d,$

(iii) $Z^*_i + \cdots + Z^*_d = W^*_0 + \cdots + W^*_{d-i}, \qquad 0 \leq i \leq d.$

6 The linear transformations B and B^*

Definition 6.1 With reference to Definition 5.1 and Definition 5.3, we define the following linear transformations.
(i) Let $B : V \to V$ be the unique linear transformation such that for $0 \leq i \leq d$, Z_i is an eigenspace for B with eigenvalue q^{2i-d}.
(ii) Let $B^* : V \to V$ be the unique linear transformation such that for $0 \leq i \leq d$, Z^*_i is an eigenspace for B^* with eigenvalue q^{2i-d}.

7 Some relations involving A,A^*,B,B^*,$K^{\pm 1}$

Lemma 7.1 *In reference to Definition 4.1 and Definition 6.1, the following hold.*

$$\frac{qAB - q^{-1}BA}{q - q^{-1}} = I, \tag{9}$$

$$\frac{qA^*B^* - q^{-1}B^*A^*}{q - q^{-1}} = I, \tag{10}$$

$$\frac{qBA^* - q^{-1}A^*B}{q - q^{-1}} = I, \tag{11}$$

$$\frac{qB^*A - q^{-1}AB^*}{q - q^{-1}} = I. \tag{12}$$

Lemma 7.2 *With reference to Assumption 2.4 and Definition 6.1, the following hold.*

$$\frac{qBK^{-1} - q^{-1}K^{-1}B}{q - q^{-1}} = I, \tag{13}$$

$$\frac{qB^*K - q^{-1}KB^*}{q - q^{-1}} = I. \tag{14}$$

Lemma 7.3 *With reference to Defintion 6.1, the following (i),(ii) hold.*

(i) $B^3B^* - [3]B^2B^*B + [3]BB^*B^2 - B^*B^3 = 0$,

(ii) $B^{*3}B - [3]B^{*2}BB^* + [3]B^*BB^{*2} - BB^{*3} = 0$.

8 The proof of Theorem 2.5 (existence)

This section is devoted to proving the existence part of Theorem 2.5.

Definition 8.1 With reference to Assumption 2.4 and Definition 6.1, let $r : V \to V$, $l : V \to V$ be the following linear transformations.

$$r = \frac{I - KB^*}{q(q - q^{-1})^2}, \qquad l = \frac{I - K^{-1}B}{q(q - q^{-1})^2}.$$

Lemma 8.2 *With reference to Definition 8.1, the following (i),(ii) hold.*

(i) $B = K - q(q - q^{-1})^2Kl$,

(ii) $B^* = K^{-1} - q(q - q^{-1})^2K^{-1}r$.

Proof: Immediate from Definition 8.1. □

Theorem 8.3 *With reference to Assumption 2.4 and Definition 8.1, the following (i)–(ix) hold.*

(i) $KK^{-1} = K^{-1}K = I$,

(ii) $KR = q^2RK, \qquad KL = q^{-2}LK$,

(iii) $Kr = q^2rK, \quad Kl = q^{-2}lK$,

(iv) $rR = Rr, \qquad lL = Ll$,

(v) $lR - Rl = \frac{K^{-1}-K}{q-q^{-1}}$, $rL - Lr = \frac{K-K^{-1}}{q-q^{-1}}$,

(vi) $R^3L - [3]R^2LR + [3]RLR^2 - LR^3 = 0$,

(vii) $L^3R-[3]L^2RL+[3]LRL^2-RL^3=0$,

(viii) $r^3l-[3]r^2lr+[3]rlr^2-lr^3=0$,

(ix) $l^3r-[3]l^2rl+[3]lrl^2-rl^3=0$.

Proof: (i) Immediate from Note 2.3.
(ii) Since $V_0, V_1, \ldots, V_d$ is a decompositon of V to prove the first equation it suffices to show that $KR-q^2RK$ vanishes of V_i for $0 \le i \le d$. Let i be given, and let $v \in V_i$. Recall that v is an eigenvector for K with eigenvalue q^{2i-d}. By Assumption 2.4(i), Rv is an eigenvector for K with eigenvalue q^{2i+2-d}. From these comments we see that $(KR-q^2RK)v=0$. The second equation follows in a similar fashion.
(iii) Evaluate the equations in Lemma 7.2 using Lemma 8.2.
(iv),(v) Evaluate (9)–(12) of Lemma 7.1 using Definition 4.1, Lemma 8.2, and Theorem 8.3(ii),(iii).
(vi),(vii) These relations hold by Assumption 2.4(v),(vi).
(viii), (ix) Substiute the expressions in Lemma 8.2 into Lemma 7.3(i),(ii), and simply using Theorem 8.3(iii). □

Theorem 8.4 *With reference to Assumption 2.4 and Definition 8.1, V supports a $U_q(\widehat{\mathfrak{sl}}_2)$-module structure for which the Chevalley generators act as follows.*

generator	e_1^-	e_0^-	e_0^+	e_1^+	K_0	K_1	K_0^{-1}	K_1^{-1}
action on V	R	L	r	l	K	K^{-1}	K^{-1}	K

Proof: To see that the above action on V determines a $U_q(\widehat{\mathfrak{sl}}_2)$-module compare the equations in Theorem 8.3 with the defining relations for $U_q(\widehat{\mathfrak{sl}}_2)$ in Definition 1.1. □

Proof of Theorem 2.5 (existence): The existence part of Theorem 2.5 is immediate from Theorem 8.4. □

9 The proof of Theorem 2.5 (uniqueness)

This section is devoted to proving the uniqueness part of Theorem 2.5.

In proving uniqueness we will make use of the quantum algebra $U_q(sl_2)$ and its representations. We now recall the definition of $U_q(sl_2)$.

Definition 9.1 [5, p. 9] The quantum algebra $U_q(sl_2)$ is the unital associative $\mathbb{K}$-algebra generated by k, k^{-1}, e, f subject to the following relations:

$$\begin{aligned} kk^{-1} = k^{-1}k &= 1, \\ ke &= q^2 ek, \\ kf &= q^{-2} fk, \\ ef - fe &= \frac{k-k^{-1}}{q-q^{-1}}. \end{aligned}$$

We now recall the irreducible finte-dimensional modules for $U_q(sl_2)$.

Lemma 9.2 *[5, p. 20] With reference to Defintion 9.1, there exist a family*

$$V_{\epsilon,d} \qquad \epsilon \in \{-1,1\}, \qquad d = 0,1,2,\ldots$$

of irreducible finite-dimensional $U_q(sl_2)$-modules with the following properties. The module $V_{\epsilon,d}$ has a basis $u_0, u_1, \ldots, u_d$ satisfying:

$$ku_i = \epsilon q^{d-2i} u_i, \qquad 0 \le i \le d, \tag{15}$$

$$fu_i = [i+1]u_{i+1}, \qquad 0 \le i \le d-1, \qquad fu_d = 0, \tag{16}$$

$$eu_i = \epsilon[d-i+1]u_{i-1}, \qquad 1 \le i \le d, \qquad eu_0 = 0. \tag{17}$$

Moreover, every irreducible finite-dimensional $U_q(sl_2)$-module is isomorphic to exactly one of the modules $V_{\epsilon,d}$.

We now show how $U_q(sl_2)$-modules and $U_q(\widehat{\mathfrak{sl}}_2)$-modules are related.

Lemma 9.3 *Let V be a finite-dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-module. For $i \in \{0,1\}$, V supports a $U_q(sl_2)$-module structure such that each of $K_i - k$, $e_i^+ - e$, $e_i^- - f$ vanish on V, where k, e, f are the generators from Definition 9.1.*

Proof: Immediate from Definition 1.1 and Definition 9.1. □

Lemma 9.4 *Let k, e, f be the generators for $U_q(sl_2)$ as in Definition 9.1. Let V be a finite-dimensional $U_q(sl_2)$-module. Assume the action of k on V is diagonalizable. Suppose $e' : V \to V$ is a linear transfomation such that*

$$ke' = q^2 e'k, \tag{18}$$

$$e'f - fe' = \frac{k - k^{-1}}{q - q^{-1}}, \tag{19}$$

hold on V. Then $(e - e')V = 0$.

Proof of Theorem 2.5 (uniqueness): By the existence part of Theorem 2.5 we know that there exists a $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V under the action of the Chevalley generators $e_i^{\pm}, K_i^{\pm 1}, i \in \{0,1\}$ such that each of $R - e_1^-$, $L - e_0^-$, $K - K_0$, and $K^{-1} - K_1$ vanish on V. Now suppose there exists another $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V under the action of the Chevalley generators $(e_i^{\pm})', (K_i^{\pm 1})', i \in \{0,1\}$ such that each of $R - (e_1^-)'$, $L - (e_0^-)'$, $K - K_0'$, and $K^{-1} - K_1'$ vanish on V. To prove uniqueness it suffices to show that for $i \in \{0,1\}$, $e_i^{\pm} - (e_i^{\pm})'$ and $K_i^{\pm 1} - (K_i^{\pm 1})'$ vanish on V. Since $R - e_1^-$ and $R - (e_1^-)'$ vanish on V then $(e_1^- - (e_1^-)')V = 0$. Similarly, we have that

$$e_0^- - (e_0^-)', \quad K_0^{\pm 1} - (K_0^{\pm 1})', \quad K_1^{\pm 1} - (K_1^{\pm 1})', \tag{20}$$

vanish on V. We now show that $(e_0^+ - (e_0^+)')V = 0$. By Lemma 9.3 we can view V as a $U_q(sl_2)$-module under the action of K_0, e_0^-, e_0^+. Using Definition 1.1 and (20) we see that $K_0(e_0^+)' = q^2(e_0^+)'K_0$ and $(e_0^+)'e_0^- - e_0^-(e_0^+)' = \frac{K_0 - K_0^{-1}}{q-q^{-1}}$. Therefore, by Lemma 9.4, we have $(e_0^+ - (e_0^+)')V = 0$. The proof that $(e_1^+ - (e_1^+)')V = 0$ is similar. □

10 Which $U_q(\widehat{\mathfrak{sl}}_2)$-modules arise from Theorem 2.5?

Theorem 2.5 gives a way to constuct finite dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-modules. Not all finite dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-modules arise from this construction; in this section we determine which ones do.

Definition 10.1 Let V denote a nonzero finite dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-module. Let d denote a nonnegative integer. We say V is *basic of diameter* d whenever there exists a decomposition $V_0, V_1, \ldots, V_d$ of V and linear transformations $R : V \to V$ and $L : V \to V$ satisfying Assumption 2.4(i)–(vi) such that the given $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V agrees with the $U_q(\widehat{\mathfrak{sl}}_2)$-module structure on V given by Theorem 2.5.

Our goal for the remainder of this section is to determine which $U_q(\widehat{\mathfrak{sl}}_2)$-modules are basic.

Lemma 10.2 *Let V be a finite dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-module. If the characteristic of $\mathbb{K}$ is not equal to 2 then the actions of K_0 and K_1 on V are diagonalizable.*

Theorem 10.3 *Let d be a nonnegative integer and let V be a finite dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-module. With reference to Definition 10.1, the following are equivalent.*

(i) V is basic of diameter d.

(ii) $(K_0K_1 - I)V = 0$, the action of K_0 on V is diagonalizable, and the set of distinct eigenvalues for K_0 on V is $\{q^{2i-d}, 0 \le i \le d\}$.

11 The relationship between general $U_q(\widehat{\mathfrak{sl}}_2)$-modules and basic $U_q(\widehat{\mathfrak{sl}}_2)$-modules

Throughout this section V will denote a finite dimensional $U_q(\widehat{\mathfrak{sl}}_2)$-module (not necessarily irreducible) on which the actions of K_0 and K_1 are diagonalizable (see Lemma 10.2).

In this section we will show, roughly speaking, that V is made up of basic $U_q(\widehat{\mathfrak{sl}}_2)$-modules. We will use the following definition.

Definition 11.1 Let $\epsilon_0, \epsilon_1 \in \{1, -1\}$. Define $V_{even}^{(\epsilon_0,\epsilon_1)}$ (resp. $V_{odd}^{(\epsilon_0,\epsilon_1)}$) to be the subspace of V spanned by all the vectors $v \in V$ such that $K_0 v = \epsilon_0 q^i v$, $K_1 v = \epsilon_1 q^{-i} v$, $i \in \mathbb{Z}$, i even (resp. i odd).

Lemma 11.2 *With reference to Definition 11.1 the following holds.*

$$V = \sum_{(\epsilon_0,\epsilon_1)} \sum_{\sigma} V_{\sigma}^{(\epsilon_0,\epsilon_1)} \qquad (\textit{direct sum of } U_q(\widehat{\mathfrak{sl}}_2) - \textit{modules}) \tag{21}$$

where the first sum is over all ordered pairs (ϵ_0, ϵ_1) with $\epsilon_0, \epsilon_1 \in \{1, -1\}$ and the second sum is over all $\sigma \in \{even, odd\}$.

Lemma 11.3 *With reference to Definition 10.1, Definition 11.1, and Lemma 11.2 the following are equivalent.*

(i) $V = V_{even}^{(1,1)}$.

(ii) V *is basic of even diameter.*

(iii) The spaces $V_{even}^{(-1,1)}$, $V_{even}^{(1,-1)}$, $V_{even}^{(-1,-1)}$, $V_{odd}^{(1,1)}$, $V_{odd}^{(-1,1)}$, $V_{odd}^{(1,-1)}$, $V_{odd}^{(-1,-1)}$ *are all zero.*

Lemma 11.4 *With reference to Definition 10.1, Definition 11.1, and Lemma 11.2 the following are equivalent.*

(i) $V = V_{odd}^{(1,1)}$.

(ii) V *is basic of odd diameter.*

(iii) The spaces $V_{odd}^{(-1,1)}$, $V_{odd}^{(1,-1)}$, $V_{odd}^{(-1,-1)}$, $V_{even}^{(1,1)}$, $V_{even}^{(-1,1)}$, $V_{even}^{(1,-1)}$, $V_{even}^{(-1,-1)}$ *are all zero.*

Refering to (21) even though the six terms $V_{even}^{(-1,1)}$, $V_{even}^{(1,-1)}$, $V_{even}^{(-1,-1)}$, $V_{odd}^{(-1,1)}$, $V_{odd}^{(1,-1)}$, $V_{odd}^{(-1,-1)}$ are not basic modules they can easily be modified to become basic modules. We now state a lemma that makes this precise.

Lemma 11.5 *[2, Prop. 3.2] For any choice of scalars $\epsilon_0, \epsilon_1 \in \{1, -1\}$ there exists a $\mathbb{K}$-algebra automorphism of $U_q(\widehat{\mathfrak{sl}}_2)$ such that*

$$K_i \to \epsilon_i K_i, \qquad e_i^+ \to e_i^+, \qquad e_i^- \to \epsilon_i e_i^-.$$

for $i \in \{1, -1\}$.

Remark 11.6 With reference to Definition 10.1 and Definition 11.1 we can alter each of the modules $V_{even}^{(-1,1)}$, $V_{even}^{(1,-1)}$, $V_{even}^{(-1,-1)}$ to a basic $U_q(\widehat{\mathfrak{sl}}_2)$-module of even diameter by applying an automorphism as in Lemma 11.5. Furthermore, we can alter each of the modules $V_{odd}^{(-1,1)}$, $V_{odd}^{(1,-1)}$, $V_{odd}^{(-1,-1)}$ to a basic $U_q(\widehat{\mathfrak{sl}}_2)$-module of odd diameter by applying an automorphism as in Lemma 11.5.

References

[1] G. Benkart and P. Terwilliger. Irreducible modules for the quantum affine algebra $U_q(\widehat{\mathfrak{sl}}_2)$ and its Borel subalgebra, Journal of Algebra 282 (2004) 172–194.

[2] V. Chari and A. Pressley, Quantum affine algebras, *Commun. Math. Phys.* **142** (1991), 261–283.

[3] T. Ito, K. Tanabe, and P. Terwilliger. Some algebra related to P- and Q-polynomial association schemes. In *Codes and association schemes (Piscataway NJ, 1999)*, 167–192, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **56**, Amer. Math. Soc., Providence RI, 2001.

[4] T. Ito and P. Terwilliger. Tridiagonal pairs and the quantum affine algebra $U_q(\widehat{\mathfrak{sl}}_2)$. *Ramanujan J.* Submitted.

[5] J.C. Jantzen, *Lectures on Quantum Groups.* American Mathematical Society, Providence RI, 1996.

Darren Neubauer (email: `neubauer@math.wisc.edu`)

Department of Mathematics
University of Wisconsin
480 Lincoln Drive
Madison, WI 53706-1388 USA

On the isomorphism problem, indecomposability and the automorphism groups of Coxeter groups

Koji NUIDA

Abstract

Let (W, S) and (W', S') be Coxeter systems, $\{W_\lambda\}_{\lambda \in \Lambda}$ and $\{W'_{\lambda'}\}_{\lambda' \in \Lambda'}$ the sets of the irreducible components of W relative to S and of W' relative to S' respectively, and let $f : W \to W'$ be an isomorphism of abstract groups. Their Coxeter graphs may not be isomorphic. We show that $f(\prod_{\lambda \in \Lambda, |W_\lambda| < \infty} W_\lambda) = \prod_{\lambda' \in \Lambda', |W'_{\lambda'}| < \infty} W'_{\lambda'}$, and that there is a unique bijection $\varphi : \{\lambda \in \Lambda \mid |W_\lambda| = \infty\} \to \{\lambda' \in \Lambda' \mid |W'_{\lambda'}| = \infty\}$ such that $f(W_\lambda) \equiv W'_{\varphi(\lambda)} \bmod Z(W')$ for every $\lambda \in \Lambda$ with $|W_\lambda| = \infty$, where $Z(W')$ is the center of W'. We also determine which two finite Coxeter groups are isomorphic. Our result reduces the problem of deciding whether two Coxeter groups are isomorphic to the case of infinite irreducible Coxeter groups. As a corollary we determine which irreducible Coxeter group is directly indecomposable as an abstract group. In particular, any infinite irreducible Coxeter group is directly indecomposable.

Soient (W, S) et (W', S') deux systèmes de Coxeter, $\{W_\lambda\}_{\lambda \in \Lambda}$ et $\{W'_{\lambda'}\}_{\lambda' \in \Lambda'}$ les ensembles des composantes irréductibles de W relative à S et de W' relative à S' respectivement, et soit $f : W \to W'$ un isomorphisme de groupes abstraits. Leur graphes de Coxeter peuvent être non isomorphes. Nous montrons que $f(\prod_{\lambda \in \Lambda, |W_\lambda| < \infty} W_\lambda) = \prod_{\lambda' \in \Lambda', |W'_{\lambda'}| < \infty} W'_{\lambda'}$, et qu'il y a une bijection unique $\varphi : \{\lambda \in \Lambda \mid |W_\lambda| = \infty\} \to \{\lambda' \in \Lambda' \mid |W'_{\lambda'}| = \infty\}$ telle que $f(W_\lambda) \equiv W'_{\varphi(\lambda)} \bmod Z(W')$ pour chaque $\lambda \in \Lambda$ avec $|W_\lambda| = \infty$, où $Z(W')$ est le centre du W'. De plus, nous déterminons quel deux groupes de Coxeter finis sont isomorphes. Nôtre résultat ramène le problème de juger si deux groupes de Coxeter sont isomorphes àu cas des groupes de Coxeter infinis irréductibles. Par conséquence, nous déterminous quel groupe de Coxeter irréductible est directement indécomposable comme un groupe abstrait. En particulier, un groupe de Coxeter infini irréductible est directement indécomposable.

Introduction

A pair (W, S) with a group W and its generating set S is called a *Coxeter system* if W has a presentation of the following form:

$$W = \langle S \mid (st)^{m(s,t)} = 1 \text{ (for } s, t \in S \text{ such that } m(s,t) < \infty) \rangle$$

where $m : S \times S \to \{1, 2, 3, \dots\} \sqcup \{\infty\}$ is a symmetric map such that $m(s,t) = 1$ if and only if $s = t$. (Note that we need not assume the finiteness of the set S.) We refer this S as a *Coxeter generating set* of W, and a group W is called a *Coxeter group* if it has a Coxeter generating set. One of the most famous examples of Coxeter groups is a (n-th) symmetric group $\mathcal{S}_n$, where the set of $n-1$ adjacent transpositions forms a Coxeter generating set. Moreover, many of the other important groups, such as elementary abelian 2-groups, dihedral groups, signed-permutation groups, Weyl groups and finite (real) reflection groups, all belong the class of Coxeter groups.

The map m above is usually given in the form of a *Coxeter graph*. This is an unoriented simple graph on the vertex set S, and two vertices s, t are joined by an edge labelled $m(s,t)$ if and only if $3 \leq m(s,t) \leq \infty$ (by convention, the labels '3' are usually omitted). For example, the Coxeter graph corresponding to $\mathcal{S}_n$ (with Coxeter generating set as above) is a simple path with $n-1$ vertices, where all edges are unlabelled. It is nontrivial and crucial that the value $m(s,t)$ for $s, t \in S$ coincides with the order of the element st in W (in particular, every generator $s \in S$ has order 2). This fact implies that the Coxeter graph is indeed determined uniquely by a Coxeter system (W, S); in other words, we have the 1-1 correspondence between Coxeter systems and Coxeter graphs (up to isomorphism).

Now some group-theoretic questions arise naturally. The first one is:

Problem 1. *Given a Coxeter group W, is the corresponding Coxeter graph, with respect to a Coxeter generating set S, determined uniquely by the group W and independent on the choice of S?*

In other words, in which case do two Coxeter graphs define isomorphic Coxeter groups? Or more primitively, when are two Coxeter groups isomorphic? This problem is called the *isomorphism problem* of Coxeter groups.

The second problem relates to the notion of irreducible decompositions of Coxeter groups. Given a Coxeter system (W, S) with Coxeter graph Γ, a subgroup of W of the form $W_I = \langle I \rangle$, where I is the vertex set of a connected component of Γ, is called an *irreducible component* of W (with respect to S). It is well known that W is the (restricted) direct product of all the irreducible components (in the viewpoint, W is called *irreducible* (with respect to S) if W has no proper irreducible component). Now the second problem is:

Problem 2. *Given a decomposition $W = \prod_\lambda W_\lambda$ of a Coxeter group W into irreducible components, is this a finest decomposition of W as an abstract group? In other words, is each irreducible component of W directly indecomposable as an abstract group?*

It has been well known that these two problems have counterexamples and are never easy or trivial. Here we give two classical examples.

Example 3. *We consider the group $W(B_n)$ of signed-permutations on n letters (or the hyperoctahedral group), the finite irreducible Coxeter group of type B_n. It contains the even signed-permutation group $W(D_n)$ (the finite irreducible Coxeter group of type D_n) and the center $Z(W(B_n))$ (which has order 2 and so is $\simeq \mathcal{S}_2 = W(A_1)$) as normal subgroups. Now it is easily checked that if $n \geq 3$ is odd, then $W(B_n)$ is a direct product of these two subgroups. This means that $W(B_n)$ is directly decomposable, and $W(B_n)$ and $W(D_n) \times W(A_1)$ are isomorphic Coxeter groups defined by non-isomorphic Coxeter graphs. This is a counterexample of Problems 1 and 2.*

Example 4. *We consider the dihedral group $\mathcal{D}_m$ of order $2m$, the finite irreducible Coxeter group $W(I_2(m))$ of type $I_2(m)$. Recall that $\mathcal{D}_m$ is the symmetry group of a regular m-gon Δ. If $m = 2k$ is even, then we can obtain an inscribed regular k-gon Δ' in Δ by joining every other vertex of Δ. In this case, the symmetry group $\mathcal{D}_k$ of Δ' is embedded (as a normal subgroup) into $\mathcal{D}_m$, while $\mathcal{D}_m$ has the center of order 2 generated by the half-rotation (rotation of degree π). Now if k is odd, then $\mathcal{D}_m$ is a direct product of these two subgroups, so that $W(I_2(m)) \simeq W(I_2(k)) \times W(A_1)$. This is the second counterexample.*

Moreover, in a paper [8], Bernhard Mühlherr gives an interesting example of two isomorphic non-finite irreducible Coxeter groups on four generators, which are defined by non-isomorphic Coxeter graphs (this is probably the first counterexample of Problem 1 for non-finite irreducible case).

The aim of this short report is to announce some recent results of the author (and also some other related results) on these topics, which is presented in the poster-session of the FPSAC'05. I would like to express my deep gratitude to the organizers of this conference for giving me the opportunity of the presentation, to the referees for their precise reading and precious comments for my report, and to Prof. Itaru Terada (my supervisor) and Prof. Kazuhiko Koike for their several advice and encouragement (especially for suggestion of application for this conference).

Main results

Recall the well-known classification of finite irreducible Coxeter groups (see [7], Chap. 2, etc.). Given a decomposition $W = \prod_\lambda W_\lambda$ of a Coxeter group W into irreducible components W_λ (with respect to a Coxeter generating set S), we define the *finite part* W_{fin} of W as the product of all finite irreducible components W_λ (note that W_{fin} may not be a finite group, in the case that W has infinitely many irreducible components). For Problem 1, the author proved the followings:

Theorem 5 ([10], Theorem 3.4). *Given two Coxeter systems (W, S), (W', S'), we have $W \simeq W'$ (as abstract groups) if and only if the following two conditions are satisfied:*

1. *$W_{\mathrm{fin}} \simeq W'_{\mathrm{fin}}$,*
2. *there is a bijection between the set of non-finite irreducible components of W and the set of those of W', such that the corresponding irreducible components are isomorphic (as abstract groups) to each other.*

Theorem 6 ([10], Theorem 3.4). *Given two Coxeter systems (W, S), (W', S'), let a_n, $b_n, \dots, h_4$, i_m denote the cardinality of the set of all irreducible components of W (with respect to S) of type A_n, $B_n, \dots, H_4$, $I_2(m)$, respectively. Define a'_n, $b'_n, \dots, i'_m$ similarly from (W', S'). Then we have $W_{\mathrm{fin}} \simeq W'_{\mathrm{fin}}$ if and only if all of the following equalities hold:*

$$a_1 + \sum_{n \geq 1} b_{2n+1} + e_7 + h_3 + \sum_{m \geq 1} i_{4m+2} = a'_1 + \sum_{n \geq 1} b'_{2n+1} + e'_7 + h'_3 + \sum_{m \geq 1} i'_{4m+2},$$

$$b_3 + a_3 = b'_3 + a'_3,\ b_{2n+1} + d_{2n+1} = b'_{2n+1} + d'_{2n+1} \text{ for } n \geq 2,$$

$$a_2 + i_6 = a'_2 + i'_6,\ i_{2m+1} + i_{4m+2} = i'_{2m+1} + i'_{4m+2} \textit{ for } m \geq 2,$$
$$a_n = a'_n \textit{ for } n \geq 4,\ b_{2n} = b'_{2n} \textit{ for } n \geq 1,\ d_{2n} = d'_{2n} \textit{ for } n \geq 2,$$
$$e_n = e'_n \textit{ for } n = 6, 7, 8,\ f_4 = f'_4,\ h_3 = h'_3,\ h_4 = h'_4,$$
$$i_{2m+1} = i'_{2m+1} \textit{ for } m \geq 2,\ i_{4m+2} = i'_{4m+2} \textit{ for } m \geq 1.$$

These theorems imply that now the isomorphism problem of Coxeter groups are reduced completely to the case of non-finite irreducible Coxeter groups. Moreover, as a byproduct, Theorem 5 shows that the set W_{fin} (not only its group structure) is uniquely determined by W and independent on the choice of S.

For Problem 2, the author also proved the following:

Theorem 7 ([10], Theorem 3.3). *All nontrivial direct product decompositions of irreducible Coxeter groups are one of the followings:*

1. $W(B_n) \simeq W(D_n) \times W(A_1)$ *for* $n \geq 3$ *odd (where we put* $D_3 = A_3$*),*
2. $W(I_2(2k)) \simeq W(I_2(k)) \times W(A_1)$ *for* $k \geq 3$ *odd (where* $I_2(3) = A_2$*),*
3. $W(E_7) = W(E_7)^+ \times W(A_1)$*, where* W^+ *denotes the subgroup of a Coxeter group* W *of elements of even length,*
4. $W(H_3) = W(H_3)^+ \times W(A_1)$.

In particular, all non-finite irreducible Coxeter groups are directly indecomposable as abstract groups, and the center of a directly decomposable irreducible Coxeter group is always a nontrivial direct factor .

Remark 8. *In view of this theorem, the equalities in Theorem 6 mean that, when we decompose (owing to Theorem 7) each of* W_{fin} *and* W'_{fin} *into directly indecomposable factors, there is a 1-1 correspondence between the factors of* W_{fin} *and those of* W'_{fin} *such that the corresponding factors are isomorphic.*

Note that the factors $W(E_7)^+$ and $W(H_3)^+$ in Theorem 7 are not Coxeter groups (namely the simple groups $S_6(2)$ and A_5, respectively). Thus Examples 3 and 4 are the only nontrivial direct product decompositions of irreducible Coxeter groups into other Coxeter groups.

We give further results related to Problem 1. The proofs of Theorems 5 and 6 given in [10] in fact describe the structure of arbitrary isomorphisms between two isomorphic Coxeter groups. Thus, by taking the Coxeter groups as the same group W, we can obtain a description of the automorphism group $\operatorname{Aut} W$ of W. The following results are deduced in this way.

We prepare some notations. For a group G and a group homomorphism $f \in \operatorname{Hom}(G, Z(G))$ from G to its center $Z(G)$, we define an endomorphism $f^\flat \in \operatorname{End} G$ of G by
$$f^\flat(w) = wf(w)^{-1} \text{ for } w \in G.$$

Lemma 9 ([10], Lemma 2.2). *The map* $f \mapsto f^\flat$ *is injective. Moreover,* $f^\flat$ *is invertible (i.e.* $f^\flat \in \operatorname{Aut} G$*) if and only if the restriction* $f^\flat|_{Z(G)}$ *of* $f^\flat$ *to* $Z(G)$ *is an automorphism of* $Z(G)$.

Given a Coxeter group W (with a Coxeter generating set S), we have a decomposition $W = W_{\mathrm{fin}} \times \prod_{\lambda \in \Lambda} W_\lambda$, where W_λ runs over all non-finite irreducible components of W (with respect to S). Put $W_{\mathrm{inf}} = \prod_{\lambda \in \Lambda} W_\lambda$. Then we give a (unique) partition $\Lambda = \bigsqcup_{\xi \in \Xi} \Lambda_\xi$ of the index set Λ such that $W_\lambda \simeq W_\mu$ (as abstract groups) if and only if λ and μ belong the same part Λ_ξ. Now let H_1 be the set of all $f^\flat$ such that $f \in \mathrm{Hom}(W_{\mathrm{inf}}, Z(W))$. Let $H_2 = \mathrm{Aut}\, W_{\mathrm{fin}}$, H_3 the (complete) direct product of all $\mathrm{Aut}\, W_\lambda$ ($\lambda \in \Lambda$) and H_4 the (complete) direct product of all symmetric groups $\mathrm{Sym}(\Lambda_\xi)$ of the sets Λ_ξ ($\xi \in \Xi$). Note that all of H_2, H_3 and H_4 are naturally embedded into $\mathrm{Aut}\, W$. Now we have:

Theorem 10 ([10], Theorem 3.10). *The group* $\mathrm{Aut}\, W$ *decomposes as*

$$\mathrm{Aut}\, W = H_1 \rtimes (H_2 \times H_3) \rtimes H_4.$$

Moreover, the action of H_4 fixes H_2 pointwise and leaves H_3 invariant.

Note that the group homomorphisms from a Coxeter group W to a group of order 2 (and so the homomorphisms from W to the center $Z(W)$ which is an elementary abelian 2-group) are characterized as follows, by using an odd-Coxeter graph. Here an *odd-Coxeter graph* is the graph obtained from a Coxeter graph by removing all but the edges with label odd. Then the following fact is easy to check (by definition of Coxeter groups).

Lemma 11. *Let W be a Coxeter group with odd-Coxeter graph Γ^{odd} (with respect to a Coxeter generating set S). Then for any map f from W to a group of order 2, f is a group homomorphism if and only if $f(s) = f(t)$ whenever $s, t \in S$ are in the same connected component of Γ^{odd}.*

Next, we consider the automorphism group of the finite part W_{fin}, which is the factor H_2 in Theorem 10. Owing to Theorem 7, we obtain a direct product decomposition $W_{\mathrm{fin}} = G_0 \times \prod_{i \in I} G_i$ of W_{fin} such that G_0 is an elementary abelian 2-group (namely the product of all factors $\simeq W(A_1)$) and each G_i ($i \in I$) is either a finite irreducible Coxeter group of type other than A_1, B_n (n odd), $I_2(2k)$ ($k \geq 3$ odd), E_7 and H_3, or a simple group $W(E_7)^+$ or $W(H_3)^+$ (and so each G_i is directly indecomposable as an abstract group). We give a partition $I = \bigsqcup_{\upsilon \in \Upsilon} I_\upsilon$ of the index set I similarly. Let H'_1 be the set of all $f^\flat \in \mathrm{Aut}\, W_{\mathrm{fin}}$ such that $f \in \mathrm{Hom}(W_{\mathrm{fin}}, Z(W_{\mathrm{fin}}))$. Let H'_2 be the (complete) direct product of all $\mathrm{Aut}\, G_i$ ($i \in I$) and H'_3 the (complete) direct product of all $\mathrm{Sym}(I_\upsilon)$ ($\upsilon \in \Upsilon$). Moreover, let H'_4 be the set of all $f^\flat$ such that $f \in \mathrm{Hom}(W_{\mathrm{fin}}, Z(W_{\mathrm{fin}}))$, $f(G_0) = 1$ and $f(G_i) \subset Z(G_i)$ for any $i \in I$. Then we have:

Theorem 12 ([10], Theorem 3.10). *We have*

$$\mathrm{Aut}\, W_{\mathrm{fin}} = (H'_1 H'_2) \rtimes H'_3 \text{ and } H'_1 \cap H'_2 = H'_4.$$

Moreover, H'_1 is normal in $\mathrm{Aut}\, W_{\mathrm{fin}}$ and the action of H'_3 leaves H'_2 invariant.

Note that the structure of $\mathrm{Aut}\, W_{\mathrm{fin}}$ is still complicated because of the existence of the intersection H'_4 of the factors H'_1 and H'_2. However, we can compute the order of $\mathrm{Aut}\, W$ for finite Coxeter groups W; see later sections.

Moreover, we consider a decomposition $W = \prod_{\lambda' \in \Lambda'} W_{\lambda'}$ of a Coxeter group W into (possibly finite) irreducible components $W_{\lambda'}$. We give a similar partition $\Lambda' = \sqcup_{\xi' \in \Xi'} \Lambda'_{\xi'}$ of the index set Λ'. Then the (complete) direct product H of all

$\operatorname{Aut} W_{\lambda'}$ ($\lambda' \in \Lambda'$) and all $\operatorname{Sym}(\Lambda'_{\xi'})$ ($\xi' \in \Xi'$) is also embedded naturally into $\operatorname{Aut} W$. (This H can be regarded as the set of automorphisms of W which are easily seen by the decomposition of W.) Then the author proved the following:

Theorem 13 ([10], Theorem 3.10). *The subgroup H of $\operatorname{Aut} W$ has finite index in $\operatorname{Aut} W$ if and only if either $Z(W) = 1$ or the odd-Coxeter graph of W consists of only finitely many connected components.*

Note that Theorems 5, 6, 7 and 13 are also proved independently by Luis Paris in [12], only for finitely generated Coxeter groups (in fact, in his proofs the finiteness of the rank of the Coxeter group is essential).

Outline of the proof

Theorems 5 and 6 are proved by using Theorem 7. To prove Theorem 7, we first show the following property of the centralizers of certain subgroups in Coxeter groups.

Proposition 14 ([10], Theorem 3.1). *Let H be a normal subgroup of a Coxeter group W which is generated by involutions. Then the structure of the centralizer $Z_W(H)$ of H in W is described completely. In particular, if $Z(W) \subsetneq Z_W(H) \subsetneq W$, then there is a proper subgroup of W containing both H and $Z_W(H)$.*

For the proof, it follows from some group theory and properties of Coxeter groups that $Z_W(H)$ is the intersection of core subgroups $\operatorname{Core}_W(N_W(W_I))$ (where $\operatorname{Core}_W(G)$ is defined as the unique largest normal subgroup of W contained in G) of the normalizers $N_W(W_I)$ of certain parabolic subgroups W_I. Then the proof of Proposition 14 is reduced to the computation of the groups $\operatorname{Core}_W(N_W(W_I))$. This is done by a certain graph-theoretical argument about the Coxeter graph owing to some properties of the normalizers $N_W(W_I)$ examined by Brigitte Brink and Robert B. Howlett in a paper [3].

Once Proposition 14 is proved, a part of Theorem 7, namely the direct indecomposability of non-finite irreducible Coxeter groups, is deduced immediately. In fact, if a non-finite irreducible Coxeter group W admits a decomposition $W = G_1 \times G_2$ into subgroups, then G_1 is generated by involutions (since it is a quotient of W) and $W = G_1 Z_W(G_1)$. Since now $Z(W) = 1$, we have (by Proposition 14) either $Z_W(G_1) = 1$ or $Z_W(G_1) = W$ (and so $G_1 = 1$). This implies the desired direct indecomposability of W. The remaining part of the theorem follows from a case-by-case argument based on the classification of finite irreducible Coxeter groups.

By a similar argument based on Proposition 14, if G_1 and G_2 are groups generated by involutions and $f : G_1 \times G_2 \to W$ is a surjective homomorphism from $G_1 \times G_2$ to a Coxeter group W, then either $f(G_1) \subset Z(W)$ or $f(G_2) \subset Z(W)$. Owing to this observation, Theorems 5, 6, 10 and 12 are deduced by similar arguments to the proof of the Remak-Krull-Schmidt Theorem about direct product decompositions of groups (see [15], Section 4.6–4.7). Theorem 13 is deduced from Theorems 10 and 12 together with Lemma 11.

$\mathrm{Aut}\, W$ of finite Coxeter groups W

As we predicted in a preceding section, we compute the order of the automorphism group $\mathrm{Aut}\, W$ (or its 'growth' $|\mathrm{Aut}\, W|/|W|$) of an arbitrary finite Coxeter group W by using Theorem 12. First, since H_1' is normal in $\mathrm{Aut}\, W$, the product $H_1'H_2'$ consists of elements of the form w_1w_2 ($w_1 \in H_1'$, $w_2 \in H_2'$) and so we have

$$\frac{|\mathrm{Aut}\, W|}{|W|} = \frac{|H_1'| \cdot |H_2'| \cdot |H_3'|}{|H_4'| \cdot |W|}.$$

Let $a_n, \dots, i_m$ be as in Theorem 6 (now all but finitely many of those are 0), and let $W = G_0 \times \prod_{i \in I} G_i$ be the decomposition of W introduced before the statement of Theorem 12. Note that G_0 is the direct product of the factors $\simeq W(A_1)$, where the number N of factors is (by Theorem 7)

$$N = a_1 + \sum_{n \geq 1} b_{2n+1} + e_7 + h_3 + \sum_{m \geq 1} i_{4m+2}.$$

Similarly, the numbers of factors G_i ($i \in I$) isomorphic to $W(A_2)$, $W(A_3)$, $W(D_{2n+1})$ ($n \geq 2$), $W(I_2(2m+1))$ ($m \geq 2$), $W(E_7)^+$, $W(H_3)^+$ are $a_2 + i_6$, $a_3 + b_3$, $b_{2n+1} + d_{2n+1}$, $i_{2m+1} + i_{4m+2}$, e_7, h_3, respectively.

For the factor H_1' of $\mathrm{Aut}\, W$, we use the following easy group-theoretic lemma:

Lemma 15. *Let G be a directly indecomposable non-abelian group such that $|Z(G)| \leq 2$. Then we have $f(Z(G)) = 1$ for any homomorphism f from G to a group of order 2.*

Owing to this lemma, we have $f(Z(G_i)) = 1$ for any $f \in \mathrm{Hom}(W, Z(W))$ (since $Z(W)$ is an elementary abelian 2-group and each G_i is either simple or a directly indecomposable non-abelian Coxeter group). Thus we have

$$f^\flat(w) = w \text{ for } w \in Z(G_i), \;\; f^\flat(w) = wf(w) \text{ for } w \in G_0.$$

By regarding elementary abelian 2-groups as vector spaces over a finite field $\mathbb{F}_2$, this implies (owing to Lemma 9) that the order of H_1' is equal to the number of invertible matrices with coefficients in $\mathbb{F}_2$ of the form

$$\begin{pmatrix} I_N + X & O \\ Y & I_M \end{pmatrix}$$

where N is as above,

$$M = |\{i \in I \mid Z(G_i) \neq 1\}| = \sum_{n \geq 1} b_{2n} + \sum_{n \geq 2} d_{2n} + e_8 + f_4 + h_4 + \sum_{m \geq 2} i_{4m}$$

(see Theorem 7) and X, Y are matrices of appropriate size with coefficients in $\mathbb{F}_2$. Thus we have

$$|H_1'| = |\mathrm{GL}_N(\mathbb{F}_2)| \cdot 2^{NM} = 2^{\binom{N}{2}} \prod_{j=1}^{N} (2^j - 1) \cdot 2^{NM}.$$

For the factor H_2', note that for any finite group G, we have

$$\frac{|\mathrm{Aut}\, G|}{|G|} = \frac{|\mathrm{Aut}\, G|}{|\mathrm{Inn}\, G| \cdot |Z(G)|} = \frac{|\mathrm{Out}\, G|}{|Z(G)|},$$

where $\operatorname{Inn} G$, $\operatorname{Out} G = \operatorname{Aut} G / \operatorname{Inn} G$ denote the groups of inner, outer automorphisms of G, respectively. Then by Table I in [1] of outer automorphism groups of finite irreducible Coxeter groups, $|\operatorname{Aut} G_i| / |G_i|$ (for $i \in I$) is

$$\begin{cases} 2 & \text{if } G_i \simeq W(A_5), W(B_{2n})\ (n \geq 2), W(D_{2n})\ (n \geq 3), W(H_4) \text{ or } W(H_3)^+ \\ 4 & \text{if } G_i \simeq W(F_4) \\ 6 & \text{if } G_i \simeq W(D_4) \\ \frac{\varphi(m)}{2} & \text{if } G_i \simeq W(I_2(m)) \\ 1 & \text{otherwise} \end{cases}$$

(where φ is the Euler function). Note that $\varphi(4m+2) = \varphi(2m+1)$ since $2m+1$ is odd, and $\varphi(6)/2 = 1$. Thus by the above observation on the numbers of factors G_i of each type, we have

$$\begin{aligned} \frac{|H_2'|}{|W|} &= |G_0|^{-1} \cdot \prod_{i \in I} \frac{|Aut\, G_i|}{|G_i|} \\ &= \frac{1}{2^N} \cdot 2^{a_5 + \sum_{n \geq 2} b_{2n} + \sum_{n \geq 3} d_{2n} + h_3 + h_4} \cdot 4^{f_4} \cdot 6^{d_4} \\ &\quad \cdot \prod_{m \geq 2} \left(\left(\frac{\varphi(2m+1)}{2} \right)^{i_{2m+1} + i_{4m+2}} \left(\frac{\varphi(4m)}{2} \right)^{i_{4m}} \right) \\ &= 2^{-N + a_5 + \sum_{n \geq 2} (b_{2n} + d_{2n}) + 2f_4 + h_3 + h_4 - \sum_{m \geq 5} i_m} \cdot 3^{d_4} \cdot \prod_{m \geq 5} \varphi(m)^{i_m}. \end{aligned}$$

For the order of H_3', it follows immediately from definition that

$$\begin{aligned} |H_3'| = \big(a_2 + i_6\big)!\, (a_3 + b_3)! \prod_{n \geq 4} a_n! \prod_{n \geq 1} b_{2n}! \prod_{n \geq 2} d_{2n}! \prod_{n \geq 2} (b_{2n+1} + d_{2n+1})! \\ \cdot e_6!\, e_7!\, e_8!\, f_4!\, h_3!\, h_4! \prod_{m \geq 2} \big(i_{2m+1} + i_{4m+2}\big)! \prod_{m \geq 2} i_{4m}!. \end{aligned}$$

Moreover, by definition, $|H_4'|$ is equal to the product of the size of all $\operatorname{Hom}(G_i, Z(G_i))$, $i \in I$. Since $|Z(G_i)| \leq 2$, it follows from Lemma 11 that the size of $\operatorname{Hom}(G_i, Z(G_i))$ is

$$\begin{cases} 2 & \text{if } G_i \simeq W(D_{2n})\ (n \geq 2), W(E_8) \text{ or } W(H_4) \\ 4 & \text{if } G_i \simeq W(B_{2n})\ (n \geq 1), W(F_4) \text{ or } W(I_2(4m))\ (m \geq 2) \\ 1 & \text{otherwise.} \end{cases}$$

Thus we have

$$|H_4'| = 2^{2 \sum_{n \geq 1} b_{2n} + \sum_{n \geq 2} d_{2n} + e_8 + 2f_4 + h_4 + 2 \sum_{m \geq 2} i_{4m}}.$$

Now we compress the data $a_n, b_n, \dots, i_m$ into a symbol $\mathbf{k}$, and write $W = W_{\mathbf{k}}$, $N = N_{\mathbf{k}}$ and $M = M_{\mathbf{k}}$. Then in the following generating function

$$F(\mathbf{X}) = \sum_{\mathbf{k}} \frac{|\operatorname{Aut} W_{\mathbf{k}}|}{|W_{\mathbf{k}}|} {X_{A_1}}^{a_1} \prod_{n \geq 2} \frac{{X_{A_n}}^{a_n}}{a_n!} \cdots \prod_{m \geq 5} \frac{{X_{I_2(m)}}^{i_m}}{i_m!}$$

(where the sum runs over all $\mathbf{k}$ such that all but finitely many contents are 0), the coefficient of $\mathbf{X}^{\mathbf{k}} = \prod_{n\geq 1} {X_{A_n}}^{a_n} \cdots \prod_{m\geq 5} {X_{I_2(m)}}^{i_m}$ is (by the above arguments)

$$2^{\binom{N_{\mathbf{k}}}{2}+N_{\mathbf{k}}M_{\mathbf{k}}-a_1+a_5-b_2-\sum_{n\geq 2} b_n-e_7-e_8-\sum_{m\geq 2}(i_{4m-3}+2i_{4m-2}+i_{4m-1}+3i_{4m})}$$
$$\cdot 3^{d_4}\prod_{j=1}^{N_{\mathbf{k}}}(2^j-1)\prod_{m\geq 5}\varphi(m)^{i_m}\cdot\binom{a_2+i_6}{a_2}\binom{a_3+b_3}{a_3}$$
$$\cdot\prod_{n\geq 2}\binom{b_{2n+1}+d_{2n+1}}{b_{2n+1}}\prod_{m\geq 2}\binom{i_{2m+1}+i_{4m+2}}{i_{2m+1}}.$$

As a simple example, we consider the case that every irreducible component of W is of type A; in other words, W is a Young subgroup of a symmetric group. The corresponding generating function F_Y is obtained from F by substituting $X_{B_n} = \cdots = X_{I_2(m)} = 0$. Put $X_j = X_{A_j}$ $(j \geq 1)$. Then we have

$$\begin{aligned} F_Y(X_1, X_2, \dots) &= \sum_{a_1,a_2,\dots} 2^{\binom{a_1}{2}-a_1+a_5}\prod_{j=1}^{a_1}(2^j-1){X_1}^{a_1}{X_2}^{a_2}\cdots \\ &= G\Big(\frac{X_1}{2}\Big)\frac{1}{1-2X_5}\sum_{\lambda} m_\lambda(X_2, X_3, X_4, X_6, X_7, \dots) \end{aligned}$$

where $G(x) = \sum_n |\mathrm{GL}_n(\mathbb{F}_2)|x^n$ denotes the generating function of the order of $\mathrm{GL}_n(\mathbb{F}_2)$, m_λ is the monomial symmetric function and the sum runs over all partitions λ.

Further remarks

As we have seen above, we now have a complete solution to Problem 2 (Theorem 7). On the other hand, Problem 1, as well as computation of the automorphism group, is basically reduced to the case of non-finite irreducible Coxeter groups (see Theorems 5, 6, 10 and 12). Nevertheless, it is still true that this problem is difficult to solve generally.

The study of the isomorphism problem of general Coxeter groups has been developing rapidly only in this decade, especially in the last five years. Here we summarize some recent results on this topic.

First we prepare some terminology. For a Coxeter system (W, S), let $R_S(W)$ be the set of *reflections* in W with respect to S; namely,

$$R_S(W) = \{w \in W \mid w \text{ is conjugate to some } s \in S\}.$$

W is said to be *rigid* if all Coxeter generators S' of W yield isomorphic Coxeter graphs. W is said to be *strongly rigid* if all Coxeter generators S' of W are conjugate in W with each other (note that this name is valid, since strong rigidity is actually stronger than rigidity). W (precisely, (W, S)) is said to be (*strongly*) *reflection rigid* if the conclusion in definition of (strong) rigidity holds for all Coxeter generators S' contained in $R_S(W)$. Moreover, W is said to be *reflection independent* if the set $R_S(W)$ of reflections is independent on the choice of the Coxeter generating set S.

The followings are examples of the known results on these properties:

Theorem 16. *Let (W,S) be a Coxeter system.*

1. *([2], Theorem 3.10) If W is finite, then (W,S) is reflection rigid.*
2. *([2], Theorem 3.9) If (W,S) is 'even' (that is, $m(s,t)$ is not odd for any distinct $s,t \in S$), then (W,S) is reflection rigid.*
3. *([13]) If $|S| < \infty$ and (W,S) is 'right-angled' (that is, $m(s,t) \in \{2,\infty\}$ for any distinct $s,t \in S$), then W is rigid.*
4. *([4]) If $|S| < \infty$ and W is capable of acting effectively, properly and cocompactly on some contractible manifold, then W is strongly rigid. In particular, affine Coxeter groups are strongly rigid.*
5. *([9]) If $|S| < \infty$ and W is non-finite, irreducible and '2-spherical' (that is, $m(s,t) < \infty$ for any $s,t \in S$), then W is strongly rigid.*

On the other hand, for the counterexamples for these properties, we summarize the properties of finite irreducible Coxeter groups in Table 1.

Table 1: List of properties of finite irreducible Coxeter groups

type		strongly rigid	rigid	strongly reflection rigid	reflection independent
A_n	$(n \neq 5)$	◯	◯	◯	◯
A_5		×	◯	◯	×
B_2		◯	◯	◯	◯
B_n	$(n \geq 3$ odd$)$	×	×	◯	×
B_n	$(n \geq 4$ even$)$	×	◯	◯	×
D_n	$(n \geq 5$ odd$)$	◯	◯	◯	◯
D_n	$(n \geq 4$ even$)$	×	◯	◯	×
E_6, E_7		◯	◯	◯	◯
E_8		×	◯	◯	×
F_4		×	◯	◯	×
H_3		×	◯	×	◯
H_4		×	◯	×	×
$I_2(m)$	$(m \not\equiv 2 \pmod 4)$	×	◯	×	◯
$I_2(6)$		×	×	◯	×
$I_2(4k+2)$	$(k \geq 2)$	×	×	×	×

In this list, the reflection rigidity is omitted since it always holds.
The description of $\operatorname{Aut} W$ for finite irreducible W given in [1] is used.

We introduce an important operation, *diagram twisting*, on Coxeter graphs and a related conjecture for the isomorphism problem. Let (W,S) be a Coxeter system with Coxeter graph Γ. A diagram twisting on Γ, with respect to two subsets $I,J \subset S$ satisfying certain conditions, is an operation which changes, for each edge e of Γ between I and J, the terminal vertex $s \in J$ of e to the new terminal vertex $w_0(J)sw_0(J)$ (where $w_0(J)$ denotes the longest element of W_J; its existence is due to the definition of diagram twistings). For the precise definition, see the paper [2] in which diagram twistings are first introduced. It is shown in [2] that, if we obtain another Coxeter graph Γ' from Γ by a diagram twisting, then there is canonically another Coxeter generating set S' of W which corresponds to Γ'. In this case, we have $R_S(W) = R_{S'}(W)$ but Γ and Γ' may be non-isomorphic. This means that, by using diagram twistings, we can obtain many examples of non-rigid Coxeter groups (note that Mühlherr's example given in [8] which we mentioned above is one of such examples).

For the converse direction, the following conjecture is presented in [2]:

Conjecture 17 ([2], Conjecture 8.1). *Let W be a Coxeter group and S, S' two finite Coxeter generating sets of W. Let Γ, Γ' be the Coxeter graphs of W with respect to S, S', respectively. If $R_S(W) = R_{S'}(W)$, then Γ' is obtained from Γ by a finite number of consecutive diagram twistings. In other words, finitely-generated Coxeter groups are reflection rigid up to diagram twistings.*

Moreover, it is hoped by many researchers (including the author) that there are (not too many) classes of special isomorphisms, described explicitly, between Coxeter groups (such as diagram twistings) such that, an arbitrary isomorphism between two Coxeter groups is made up from those special ones. If this is fortunately true, then we can reduce the study of relations between some combinatorial properties of two isomorphic Coxeter groups to the study of those special isomorphisms.

Finally, we state some more recent results of the author. Note that, most of the results on the isomorphism problem of Coxeter groups which have been known now are only limited to the case of finitely generated Coxeter groups. One of the reason is that now a main strategy for this problem is to analize the maximal finite subgroups of given Coxeter groups, but in the non-finitely generated cases, it happens very often that the Coxeter group has no maximal finite subgroups. In contrast with those cases, the results of the author are applicable to non-finitely generated cases as well as finitely generated cases.

We prepare some notations. For a generator $x \in S$ of a Coxeter system (W, S), let $W^{\perp x}$ be the subgroup of W generated by all reflections $t \in R_S(W)$, $t \neq x$, which commutes with x. By a general result of Vinay V. Deodhar [5] or of Matthew Dyer [6], $W^{\perp x}$ forms a Coxeter group with a canonical Coxeter generating set. Let ${W^{\perp x}}_{\mathrm{fin}}$ denote the finite part of this Coxeter group $W^{\perp x}$.

Theorem 18 ([11]). *Let (W, S), (W', S') be two Coxeter system (where S or S' may be infinite), $f : W \xrightarrow{\sim} W'$ an isomorphism and $x \in S$.*

1. *The structure of ${W^{\perp x}}_{\mathrm{fin}}$ is completely determined.*
2. *(See [14]) There are an inner automorphism g of W' and a finite subset $I' \subset S'$ of (-1)-type (i.e. every irreducible component of $W'_{I'}$ has nontrivial center) such that $g \circ f(x)$ is the longest element $w_0(I')$ of $W'_{I'}$.*
3. *Suppose that there is a finite subset $I' \subset S'$ of (-1)-type such that $f(x) = w_0(I')$. Then $f^{-1}(W'_{I'}) \subset \langle x \rangle \times {W^{\perp x}}_{\mathrm{fin}}$.*
4. *Suppose that ${W^{\perp x}}_{\mathrm{fin}}$ is either trivial or generated by a single reflection $t \in R_S(W)$ conjugate to x. Then $f(x) \in R_{S'}(W')$.*

As an application of this theorem, the author obtained recently the following results on reflection independence of some (possibly non-finitely generated) Coxeter groups.

Theorem 19 ([11]). *Let (W, S) be a Coxeter system (where S may be infinite).*

1. *If W is non-finite, irreducible and 2-spherical, then W is reflection independent.*
2. *If W is non-finite and 'odd-connected' (that is, the odd-Coxeter graph Γ^{odd} is connected), then W is reflection independent.*

References

[1] E. Bannai, *Automorphisms of irreducible Weyl groups*, J. Fac. Sci. Univ. Tokyo Sect. I 16 (1969) 273–286.

[2] N. Brady, J. P. McCammond, B. Mühlherr, W. D. Neumann, *Rigidity of Coxeter groups and Artin groups*, Geom. Dedicata 94 (2002) 91–109.

[3] B. Brink, R. B. Howlett, *Normalizers of parabolic subgroups in Coxeter groups*, Invent. Math. 136 (1999) 323–351.

[4] R. Charney, M. Davis, *When is a Coxeter system determined by its Coxeter group?*, J. London Math. Soc. (2) 61 (2000) 441–461.

[5] V. V. Deodhar, *A note on subgroups generated by reflections in Coxeter groups*, Arch. Math. 53 (1989) 543–546.

[6] M. Dyer, *Reflection subgroups of Coxeter systems*, J. Algebra 135 (1990) 57–73.

[7] J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge U.P., 1990.

[8] B. Mühlherr, *On isomorphisms between Coxeter groups*, Des. Codes Cryptogr. 21 (2000) 189–189.

[9] B. Mühlherr, *On the isomorphism problem for Coxeter groups*, The Coxeter Legacy: Reflections and Projections, Toronto, May 12–16, 2004.

[10] K. Nuida, *On the direct indecomposability of infinite irreducible Coxeter groups and the Isomorphism Problem of Coxeter groups*, arXiv:math.GR/0501276 (2005).

[11] K. Nuida, *Centralizers of reflections and reflection independence of Coxeter groups*, preprint, 2005.

[12] L. Paris, *Irreducible Coxeter groups*, arXiv:math.GR/0412214 (2004).

[13] D. G. Radcliffe, *Rigidity of graph products of groups*, Algebr. Geom. Topol. 3 (2003) 1079–1088.

[14] R. W. Richardson, *Conjugacy classes of involutions in Coxeter groups*, Bull. Austral. Math. Soc. 26 (1982) 1–15.

[15] W. R. Scott, *Group Theory*, Dover, 1987.

Koji Nuida
Graduate School of Mathematical Sciences, University of Tokyo, Japan

Supported by JSPS Research Fellowship (No. 16-10825)

E-mail: *nuida@ms.u-tokyo.ac.jp*

SOME RELATIONAL STRUCTURES WITH POLYNOMIAL GROWTH AND THEIR ASSOCIATED ALGEBRAS

MAURICE POUZET AND NICOLAS M. THIÉRY

This paper is dedicated to Adriano Garsia at the occasion of his 75th birthday

ABSTRACT. The *profile* of a relational structure R is the function φ_R which counts for every integer n the number, possibly infinite, $\varphi_R(n)$ of substructures of R induced on the n-element subsets, isomorphic substructures being identified. Several graded algebras can be associated with R in such a way that the profile of R is simply the Hilbert function. An example of such graded algebra is the *age algebra* $\mathbb{K}.\mathcal{A}(R)$, introduced by P. J. Cameron. In this paper, we give a closer look at this association, particularly when the relational structure R decomposes into finitely many monomorphic components. In this case, several well-studied graded commutative algebras (e.g. the invariant ring of a finite permutation group, the ring of quasi-symmetric polynomials) are isomorphic to some $\mathbb{K}.\mathcal{A}(R)$. Also, φ_R is a quasi-polynomial, this supporting the conjecture that, with mild assumptions on R, φ_R is a quasi-polynomial when it is bounded by some polynomial.

Keywords: Relational structure, profile, graded algebra, Hilbert function, Hilbert series, polynomial growth, invariant ring, permutation group.

1. Presentation

A *relational structure* is a realization of a language whose non-logical symbols are predicates, that is a pair $R := (E, (\rho_i)_{i\in I})$ made of a set E and of a family of n_i-ary relations ρ_i on E. The family $\mu := (n_i)_{i\in I}$ is the *signature* of R. The *profile* of R is the function φ_R which counts for every integer n the number $\varphi_R(n)$ of substructures of R induced on the n-element subsets, isomorphic substructures being identified. Clearly, this function only depends upon the set $\mathcal{A}(R)$ of finite substructures of R considered up to an isomorphism, a set introduced by R. Fraïssé under the name of *age* of R (see [Fra00]). If I is finite $\varphi_R(n)$ is necessarily finite. As we will see, in order to capture examples coming from algebra and group theory, we cannot preclude I to be infinite. Since the profile is finite in these examples, we will always make the assumption that $\varphi_R(n)$ is finite, no matter how large I is.

A basic result about the behavior of the profile is this:

Theorem 1.1. *If R is a relational structure on an infinite set, then φ_R is non-decreasing.*

This result was obtained in 1971 by the first author (see Exercise 8 p. 113 [Fra71]). A proof based on linear algebra is given in [Pou76].

Provided that the relational structures satisfies some mild conditions, there are jumps in the behavior of the profile:

Date: 10th May 2005.

Theorem 1.2. [Pou78] *Let $R := (E, (\rho_i)_{i\in I})$ be a relational structure. The growth of φ_R is either polynomial or as fast as every polynomial provided that either the signature $\mu := (n_i)_{i\in I}$ is bounded or the kernel $K(R)$ of R is finite.*

Note that a map $\varphi : \mathbb{N} \to \mathbb{N}$ has *polynomial growth*, of *degree* k, if $an^k \le \varphi(n) \le bn^k$ for some $a, b > 0$ and n large enough. The *kernel* of R is the set $K(R)$ of $x \in E$ such that $\mathcal{A}(R_{|E\setminus\{x\}}) \ne \mathcal{A}(R)$. Relations with empty kernel are the *inexhaustible* relations of R. Fraïssé (see [Fra00]). We call *almost inexhaustible* those with finite kernel. The hypothesis about the kernel is not ad hoc. As it turns out, if the growth of the profile of a relational structure with a bounded signature is bounded by a polynomial then its kernel is finite. Some hypotheses on R are needed, indeed *for every increasing and unbounded map $\varphi : \mathbb{N} \to \mathbb{N}$, there is a relational structure R such that φ_R is unbounded and eventually bounded above by φ*(cf. [Pou81]).

The consideration of examples suggested that in order to study the profile of a relational structure R, the right object to consider is rather the generating series

$$\mathcal{H}_{\varphi_R} := \sum_{n=0}^{\infty} \varphi_R(n) Z^n$$

This innocuous change leads immediately to several:

Questions 1.3.

(1) For which relational structures is the series $\mathcal{H}_{\varphi_R}$ a rational fraction? A rational fraction of the form

$$\frac{P(Z)}{(1-Z)(1-Z^{n_2})\cdots(1-Z^{n_k})},$$

with $k \ge 1$, $1 = n_1 \le n_2 \le \cdots \le n_k$, $P(0) = 1$, and $P \in \mathbb{Z}[Z]$?

(2) Is this the case of relational structures with bounded signature or finite kernel for which the profile is bounded by some polynomial? When can P be taken with nonnegative coefficients?

(3) For which relational structures is this series convergent? Is this the case of relational structures R whose age $\mathcal{A}(R)$ is well-quasi-ordered by embeddability?

Remark 1.4. When $\mathcal{H}_{\varphi_R}$ is a rational fraction of the form above then, for n large enough, $\varphi_R(n)$ is a *quasi-polynomial* of degree k', with $k' \le k-1$, that is a polynomial $a_{k'}(n)n^{k'} + \cdots + a_0(n)$ whose coefficients $a_{k'}(n), \ldots, a_0(n)$ are periodic functions. Since the profile is non-decreasing, it follows that $a_{k'}(n)$ is eventually constant. Hence the profile has polynomial growth: $\varphi_R(n) \sim an^{k'}$ for some non-negative real a.

With the contribution of P. J. Cameron, this also gives links to some quite venerable fields of mathematics. Indeed, P. J. Cameron [Cam97] associates to the age of R, $\mathcal{A}(R)$, its *age algebra*, a graded commutative algebra $\mathbb{K}.\mathcal{A}(R)$ over a field $\mathbb{K}$ of characteristic zero, and shows that the dimension of the homogeneous component of degree n of $\mathbb{K}.\mathcal{A}(R)$ is $\varphi_R(n)$, hence the generating series above is simply the Hilbert series of $\mathbb{K}.\mathcal{A}(R)$. Other graded algebras than the age algebra enjoy this property. In any case, the association between relational structures and graded algebras via the profile seems to be an interesting topic.

The purpose of this paper is to document this association. We do that for a special case of relational structures that we introduce here for the first time: *those admitting a finite monomorphic decomposition.* Despite the apparent simplicity of these relational structures, the corresponding age algebras include familiar objects like invariant rings of finite permutation groups. The profile of these relational structures is a quasi-polynomial (cf. Theorem 2.16). This supports the conjecture that the profile of a relational structures with bounded signature or finite kernel is a quasi-polynomial whenever the profile is bounded by some polynomial (for more on the profile, see [Pou02]).

We particularly study the special case of relational structures associated with a permutation groupoid G on a finite set X; as it turns out, their age algebra is a subring of $\mathbb{K}[X]$, the *invariant ring* associated to G (cf. Theorem 4.4). This setting provides a close generalization of invariant rings of permutation groups which includes other famous algebras like quasi-symmetric polynomials and their generalizations. We analyze in details which properties of invariant rings of permutation groups carry over — or not — to permutation groupoids (cf. Propositions 4.18 and 4.9, and Theorems 4.12 and 4.17). To this end, we use in particular techniques from [GS84].

A strong impulse to this research came from the paper by Garsia and Wallach [GW03], which proves that, like invariant rings of permutation groups, the rings of quasi-symmetric polynomials are Cohen-Macaulay. Indeed, a central long term goal is to characterize those permutation groupoids whose invariant ring is Cohen-Macaulay (cf. Problem 4.14), this algebraic property implying that the profile can be written as a quasi-polynomial whose numerator has non-negative coefficients. As a first step in this direction, we analyze several examples, showing in particular that not all invariant rings of permutation groupoids are Cohen-Macaulay.

2. Relational structures admitting a finite monomorphic decomposition

A *monomorphic decomposition* of a relational structure R is a partition $\mathcal{P}$ of E into blocks such that for every integer n, the induced structures on two n-elements subsets A and A' of E are isomorphic whenever the intersections $A \cap B$ and $A' \cap B$ over each block B of $\mathcal{P}$ have the same size.

2.1. **Some examples of relational structures admitting a finite monomorphic decomposition.** .

Example 2.1. Let $R := (\mathbb{Q}, \leq, u_1, \ldots, u_{k+1})$ where $\mathbb{Q}$ is the chain of rational numbers, $u_1, \ldots, u_{k+1}$ are $k+1$ unary relations which divide Q into $k+1$ intervals. Then $\varphi_R(n) = \binom{n+k}{k}$ and $\mathcal{H}_{\varphi_R} = \frac{1}{(1-Z)^{k+1}}$.

Example 2.2. A graph $G := (V, \mathcal{E})$ being considered as a binary irreflexive and symmetric relation, its profile φ_G is the function which counts, for each integer n, the number $\varphi_G(n)$ of induced subgraphs on n elements subsets of $V(G)$, isomorphic subgraphs counting for one.

Trivially φ_G is constant, equal to 1, if and only if G is a clique or an infinite independent set. A bit less trivial is the fact that φ_G is bounded if and only if G is *almost constant* in the sense of R. Fraïssé [Fra00], that is there is a finite subset F_G of vertices such that two pairs of vertices having the same intersection on F_G are both edges or both non-edges.

Example 2.3. Let G be the direct sum $K_\omega \oplus K_\omega$ of two infinite cliques; then $\varphi_G(n) = \lfloor \frac{n}{2} \rfloor + 1$.

Example 2.4. Let G be the direct sum of $k+1$ many infinite cliques; then $\varphi_G(n) = p_{k+1}(n) \simeq \frac{n^k}{(k+1)!k!}$.

Examples 2.5. Let G be the direct sum $K_{(1,\omega)} \oplus \overline{K}_\omega$ of an infinite wheel and an infinite independent set, or the direct sum $K_\omega \oplus \overline{K}_\omega$ of an infinite clique and an infinite set; then $\varphi_G(n) = n$. Hence $\mathcal{H}_{\varphi_G} = 1 + \frac{Z}{(1-Z)^2}$, that we may write $\frac{1-Z-Z^2}{(1-Z)^2}$, as well as $\frac{1+Z^3}{(1-Z)(1-Z^2)}$.

The first example above has three monomorphic components, one being finite, whereas the second one has two component, both infinite; still, the generating series coincide, with one representation as a rational fraction with a numerator with some negative coefficient, and another with all coefficients non-negative.

A more involved example is the following:

Example 2.6. Let $R := (E, (\rho, U_2, U_3))$, where $E := \mathbb{N}\times\{0,1,2,3\}$, $\rho := \{((n,i),(m,j)) : i = 0, j \in \{1,2\}$ or $i = 1, j = 3\}$; $U_i := \mathbb{N} \times \{i\}$ for $i \in \{0,1,2,3\}$. Then R has four monomorphic components, namely $\mathbb{N}\times\{0\}, \mathbb{N}\times\{1\}, \mathbb{N}\times\{2\}, \mathbb{N}\times\{3\}$. Let S be the induced structure on four elements of the form $(x_i, i), i \in \{0,1,2,3\}$. A crucial property is that S has only two non-trivial local isomorphisms, namely the map sending $(x_0, 0)$ onto $(x_1, 1)$ and its inverse. From this follows that the induced substructures on two n-element subsets E are isomorphic if either they have the same number of elements on each $\mathbb{N}\times\{i\}$ or one subset is included into $\mathbb{N}\times\{0\}$, the other into $\mathbb{N}\times\{1\}$. Hence, the generating series $\mathcal{H}_{\varphi_R}$ is $\frac{1}{(1-Z)^4} - \frac{Z}{1-Z} = \frac{1-Z+3Z^2-3Z^3+Z^4}{(1-Z)^4}$. We may write it $\mathcal{H}_{\varphi_R} = \frac{Q_1}{(1-Z)(1-Z^4)(1-Z^5)(1-Z^5)}$ where $Q_1 := 1+2Z+6Z^2+10Z^3+14Z^4+17Z^5+18Z^6+14Z^7+10Z^8+6Z^9+Z^{10}$, as well as $\mathcal{H}_{\varphi_R} = \frac{Q}{(1-Z)(1-Z^5)^3}$ where $Q_2 := 1+2Z+6Z^2+10Z^3+15Z^4+18Z^5+22Z^6+18Z^7+15Z^8+10Z^9+6Z^{10}+Z^{12}+Z^{16}$.

Here is an example of relational structure R such that $\mathcal{H}_{\varphi_R}$ is a rational fraction, for which there is no way of choosing the numerator with non-negative coefficients.

Example 2.7. Let $R := (E, (\rho_0, \rho_1, \rho_2))$ be defined as follows. First, $E := \mathbb{N} \times \{0,1,2,3\} \setminus \mathbb{N}^* \times \{0,1\}$, that is E is the union of two one-element sets, $E_0 := \{(0,0)\}, E_1 := \{(0,1)\}$, and of two infinite sets, $E_2 := \mathbb{N}\times\{2\}, E_3 := \mathbb{N}\times\{3\}$. Next $\rho_i := E_i \times (E_2 \cup E_3)$ for $i = 0,1$ and $\rho_2 := E_0 \times E_1 \times (E_1 \cup E_2)$. Then R has four monomorphic components, namely $E_0, E_1, E_2 E_3$. The crucial property is that the induced structures on two n-element subsets A, B of E are isomorphic if either A and B includes $E_0 \cup E_1$ and their respective traces on E_2, E_3 have the same size, or A and B include exactly one of the sets E_0 E_1 and not the other, or exclude both. From this $\varphi_R(0) = 1$ and $\varphi_R(n) = n+2$ for $n \geq 1$, hence $\mathcal{H}_{\varphi_R} = \frac{1+Z-Z^2}{(1-Z)^2}$. If $\mathcal{H}_{\varphi_R} = \frac{P}{(1-Z)(1-Z^k)}$ with $k \geq 2$, then

$$P = 1 + 2Z + \sum_{j=2}^{k-2} Z^j - Z^{k+1}$$

hence has a negative coefficient .

Another example with the same property, with two monomorphic component which are both infinite.

Example 2.8. Let $R := (E, \mathcal{H})$, where $E := \mathbb{N} \times \{0,1\}$, $\mathcal{H} := [\mathbb{N} \times \{0\}]^3 \cup [\mathbb{N} \times \{1\}]^3$. Then R has two monomorphic components, namely $\mathbb{N} \times \{0\}$ and $\mathbb{N} \times \{1\}$. Each type of n-element restriction has a representative made of a $m+k$ element subset of $\mathbb{N} \times \{0\}$ and of a m-element subset of $\mathbb{N} \times \{1\}$ such that $n = 2m+k$; these representatives are non-isomorphic, except if $n = 2$ (in the later case, all 2 -element restrictions are isomorphic, hence we may eliminate the representative corresponding to $m = 1, k = 0$). With this observation, a straightforward computation shows that $\varphi_R(0) = \varphi_R(1) = \varphi_R(2) = 1$ and $\varphi_R(n) = \lfloor \frac{n}{2} \rfloor + 1$ for $n \geq 3$. Hence the generating series $H_{\varphi_R} = \frac{1}{(1-x)(1-x^2)} - x^2 = \frac{1-x^2+x^3+x^4-x^5}{(1-x)(1-x^2)}$.

But, then H_{φ_R} cannot be written as a quotient of the form $\frac{P}{(1-x)(1-x^k)}$ where P is a polynomial with non-negative integer coefficients. Indeed, suppose, by contradiction, that H_{φ_R} is of this form. We may suppose k even (otherwise, multiply P and $(1-x)(1-x^k)$ by $(1+x^k)$. Set $k' := \frac{k}{2}$. Multiplying $1 - x^2 + x^3 + x^4 - x^5$ and $(1-x)(1-x^2)$ by $1 + x^2 + \cdots + x^{2(k'-1)}$, we get $P = (1 - x^2 + x^3 + x^4 - x^5)(1 + x^2 + \cdots + x^{2(k'-1)})$. Hence, the term of largest degree has a negative coefficient, a contradiction.

See also Example 4.13 for a last example with three infinite monomorphic components and additional algebraic structure which still has this property.

2.1.1. *Examples coming from group actions.* The *orbital profile* of a permutation group G acting on a set E is the function θ_G which counts for each integer n the number, possibly infinite, of orbits of the n-element subsets of E.

As it is easy to see, orbital profiles are special cases of profiles. Indeed, for every G there is a relational structure such that $\mathrm{Aut}R = \overline{G}$ (the topological closure of G in the symmetric group $\mathfrak{S}(E)$, equipped with the topology induced by the product topology on E^E, E being equipped with the discrete topology). Groups for which the orbital profile takes only finite values are said *oligomorphic*, cf. P. J. Cameron book [Cam90]. These groups are quite common. Indeed, *if G is a group acting on a denumerable set E and R is a relational structure such that* $\mathrm{Aut}R = \overline{G}$ *Then G is oligomorphic if and only if the complete theory of M is $\aleph_0$-categorical (Ryll-Nardzewski, 1959).*

Even in the special case of groups, questions we ask in section 1 have not been solved yet. For an example, let G be a group acting on a denumerable set E; if the orbital profile of G is bounded above by some polynomial, is the generating series of this profile a rational fraction? a rational fraction of the form given in Question 1? with a numerator with non-negative coefficients?

Several examples come from relational structures which decompose into finitely many monomorphic components.

Example 2.9. Let G be is the identity group on a m element set E. Set $R := (E, U_1, \ldots, U_m)$ where $U_1, \ldots, U_m$ are m-unary relations defining the m elements of E; then $\theta_G(n) = \varphi_R(n) = \binom{m}{n}$.

Example 2.10. Let $G := \mathrm{Aut}\mathbb{Q}$, where $\mathbb{Q}$ is the chain of rational numbers. Then $\theta_G(n) = \varphi_{\mathbb{Q}}(n) = 1$ for all n.

Example 2.11. Let G' be the wreath product $G' := G \wr \mathfrak{S}_{\mathbb{N}}$ of a permutation group G acting on $\{1, \ldots, k\}$ and of $\mathfrak{S}_{\mathbb{N}}$, the symmetric group on $\mathbb{N}$. Looking at G' as a permutation group acting on $E' := \{1, \ldots, k\} \times \mathbb{N}$, then $G' = \mathrm{Aut}R'$ for some relational structure R' on E'; moreover, for all n, $\theta_{G'}(n) = \varphi_{R'}(n)$. Among the possible R' take $R \wr \mathbb{N} := (E', \equiv, (\overline{\rho}_i)_{i \in I})$ where $\equiv$ is $\{((i,n),(j,m)) \in E'^2 : i = j\}$, $\overline{\rho}_i := \{((x_1, m_1), \dots, (x_{n_i}, m_{n_i})) : (x_1, \dots, x_{n_i}) \in \rho_i, (m_1, \dots, m_{n_i}) \in \mathbb{N}^{n_i}\}$, and $R := (\{1, \dots, k\}, (\rho_i)_{i \in I})$ is a relational structure having signature $\mu := (n_i)_{i \in I}$ such that $\mathrm{Aut}R = G$. The relational structure $R \wr \mathbb{N}$ decomposes into k monomorphic components, namely the equivalence classes of $\equiv$.

As it turns out, $\mathcal{H}_{\varphi_{R \wr \mathbb{N}}}$ is the Hilbert series $\sum_{n=0}^{\infty} \dim \mathbb{K}[X]_n^G Z^n$ of the invariant ring $\mathbb{K}[X]^G$ of G (that is the subring of the polynomials in the indeterminates $X := (x_1, \dots, x_k)$ which are invariant under the action of G) (Cameron [Cam90]). As it is well known, this Hilbert series is a rational fraction of the form indicated in Question 1.3, where the coefficients of $P(Z)$ are non-negative.

Problem 2.12. Find an example of a permutation group G' acting on a set E with no finite orbit, such that the orbital profile of G' has polynomial growth, but the generating series is not the Hilbert series of the invariant ring $\mathbb{K}[X]^G$ of a permutation group G acting on a finite set X.

2.1.2. *Quasi-symmetric polynomials and the like.*

Example 2.13. Let $X_k := (x_1, \dots, x_k)$ be k indeterminates and $n_1, \dots, n_l$ be a sequence of positive integers, $l \leq k$. The polynomial

$$\sum_{1 \leq i_1 < \dots < i_l \leq k} x_{i_1}^{n_1} \dots x_{i_l}^{n_l}$$

is a *quasi-symmetric monomial* of degree $n := n_1 + \dots + n_l$. The vector space spanned by the quasi-symmetric monomials forms the space $\mathrm{QSym}(X_k)$ of *quasi-symmetric polynomials* as introduced by I. Gessel. As in the example above, the Hilbert series of $\mathrm{QSym}(X_k)$ is defined as

$$\mathcal{H}_{\mathrm{QSym}(X_k)} := \sum_{n=0}^{\infty} \dim \mathrm{QSym}(X_k)_n Z^n \ .$$

As shown by F. Bergeron and C. Reutenauer (cf. [GW03]), this is a rational fraction of the form $\frac{P_k}{(1-Z)*(1-Z^2)*\dots(1-Z^k)}$ where the coefficients P_k are non negative. Let R be the poset product of a k-element chain by a denumerable antichain. More formally, $R := (E, \rho)$ where $E := \{1, \dots, k\} \times \mathbb{N}$ and $\rho := \{((i,n),(j,m)) \in E$ such that $i \leq j\}$. Each isomorphic type of an n-element restriction may be identified to a quasi-symmetric polynomial, hence the generating series associated to the profile of R is the Hilbert series defined above.

Example 2.14. A relational structure $R := (E, (\rho_i)_{i \in I})$ is *categorical for its age* if every R' having the same age as R is isomorphic to R. It was proved in [HM88] that for relational structure with finite signature (I finite) this happens just in case E is countable and can be divided into finitely many blocks such that every permutation of E which preserves each block is an automorphism of R.

2.2. **Results and problems about relational structures admitting a finite monomorphic decomposition.** The following result motivates the introduction of the notion under review.

Theorem 2.15. *The profile of a relational structure R is bounded by some integer if and only if R has a monomorphic decomposition into finitely many blocks, at most one being infinite.*

Relational structures satisfying the second condition of the above sentence are the so-called *almost-monomorphic relational structures* of R. Fraïssé. Theorem 2.15 above was proved in [FP71] for finite signature and in [Pou81]) for arbitrary signature by means of Ramsey theorem and compactness theorem of first order logic. From Theorem 1.1 and Theorem 2.15, it follows that a relational structure R has a monomorphic decomposition into finitely many blocks, at most one being infinite if and only if

$$\mathcal{H}_{\varphi_R} = \frac{1 + b_1 Z + \cdots + b_l Z^l}{1 - Z} \ ,$$

where $b_1, \ldots, b_l$ are non negative integers.

It is trivial that, if an infinite relational structure R has a monomorphic decomposition into finitely many blocks, whereof k are infinite, then the profile is bounded by some polynomial, whose degree itself is bounded by $k - 1$.

Theorem 2.16. *Let R be an infinite relational structure R with a monomorphic decomposition into finitely many blocks $(E_i, i \in X)$, k of which being infinite. Then, the generating series $\mathcal{H}_{\varphi_R}$ is a rational fraction of the form:*

$$\frac{P(Z)}{(1-Z)(1-Z^2)\cdots(1-Z^k)} \ .$$

In particular, remark 1.4 applies.

To each subset A of size d of E, we associate the monomial

$$x^{d(A)} := \prod_{i \in X} x_i^{d_i(A)} \ ,$$

where $d_i(A) = |A \cap E_i|$ for all i in X. Obviously, A is isomorphic to B whenever $x^{d(A)} = x^{d(B)}$. The shape of a monomial $x^d = \prod x_i^{d_i}$ is the partition obtained by sorting decreasingly $(d_i, i \in X)$. We define a total order on monomials by comparing their shape w.r.t. the degree reverse lexicographic order, and breaking ties by the usual lexicographic order on monomials w.r.t. some arbitrary fixed order on X. To each orbit of sets, we associate the unique maximal monomial $\mathrm{lm}(A)$, where A ranges through the orbit; we call this monomial *leading monomial*. To prove the theorem, we essentially endow the set of leading monomials with an ideal structure in some appropriate polynomial ring. This is reminiscent of the chain-product technique as defined in Subsection 4.1.1. The key property of leading monomials is this:

Lemma 2.17. *Let m be a leading monomial, and $S \subset X$ a layer of m. Then, either $d_i = |E_i|$ for some i in S, or mx_S is again a leading monomial.*

The proof of this result relies on Proposition 2.19 below for which we introduce the following definition. Let R be a relational structure on E; a subset B of E is a *monomorphic part* of R if for every integer n and every pair A, A' of n-element subsets of E the induced structures on A and A' are isomorphic whenever $A \setminus B = A' \setminus B$. The following lemma, given without proof, rassembles the main properties of monomorphic parts.

Lemma 2.18. *(i) The emptyset and the one element subsets of E are monomorphic parts of R;*
(ii) If B is a monomorphic part of R then every subset of B too;
(iii) Let B and B' be two monomorphic parts of R; if B and B' intersect, then $B \cup B'$ is a monomorphic part of R;
(iv) Let $\mathcal{B}$ be a family of monomorphic parts of R; if $\mathcal{B}$ is up-directed (that is the union of two members of $\mathcal{B}$ is contained into a third one), then their union $B := \bigcup \mathcal{B}$ is a monomorphic part of R.

Let $x \in E$, let $R(x)$ be the set-union of all the monomorphic parts of R containing x. By (i) of Lemma 2.18 this set contains x and by (iii) and (iv) this is a monomorphic part, thus the largest monomorphic part of R containing x.

Proposition 2.19. *The largest monomorphic parts form a monomorphic decomposition of R off which every monomorphic decomposition of R is a refinement.*

Proof of Lemma 2.17. Let $e := |X|$, $\overline{d} := (d_{i_1}, \dots, d_{i_e})$ be the shape of m sorted decreasingly and $s := |S|$. Suppose that $d_i < |E_i|$ for every i in S. Let A, B , B' be subsets of E such that $x^{d(A)} = m$, $x^{d(B)} = mx_S$, and $m' := x^{d(B')}$ is the leading monomial in the orbit of B and let R_A, R_B, $R_{B'}$ be the corresponding induced structures.

Clearly, the shape of mx_S is $\overline{d_1} := (d_{i_1} + 1, \dots, d_{i_s} + 1, d_{i_{s+1}}, \dots, d_{i_e})$. Let $\overline{d'} := (d'_{i'_1}, \dots, d'_{i'_s}, d'_{i'_{s+1}}, \dots, d'_{i'_e})$ be the shape of B'. Our first goal is to prove that these two shapes are the same.

Claim 1 *$d_{i_p} = d'_{i'_p}$ for all $p > s$.*

Proof of Claim 1 Suppose this does not hold. Let p be the largest such that $d_{i_p} \neq d'_{i'_p}$. Since, by definition, we have $\overline{d'} \geq \overline{d_1}$ it follows that $d'_{i'_p} < d_{i_p}$; thus, $\overline{d'} > \overline{d}$. However, since $R_{B'}$ contains a copy of R_A we have $\overline{d'} \leq \overline{d}$, a contradiction. □

Set $U := \bigcup\{E_i \cap B : i \notin S\}$, $S' := \{i'_1, \dots, i'_s\}$, $U' := \bigcup\{E_i \cap B' : i \notin S'\}$. Let φ be an isomorphism from R_B onto $R_{B'}$.

Claim 2 *U is the set of $x \in B$ such that the induced structure $R_{B\setminus\{x\}}$ on $B \setminus \{x\}$ contains no copy of R_A. Moreover φ transforms U into U'.*

Proof of Claim 2 From the definition of U, $R_{B\setminus\{x\}}$ contains a copy of R_A for every element $x \in B \setminus U$. Conversely, let $x \in U$ and let $\overline{d''}$ be the shape of $B \setminus \{x\}$. Clearly, for the largest p such that $d''_{i''_p} \neq d_{i_p}$ we have $p > s$. Hence $\overline{d}'' > \overline{d}$, thus $R_{B\setminus\{x\}}$ cannot contains a copy of R_A. This proves the first part of Claim 2.

Since, from Claim 1, $d_{i_p} = d'_{i'_p}$ for all $p \geq s$, the same argument show that if $x' \in U'$ then $R_{B'\setminus\{x'\}}$ cannot contains a copy of R_A. Since from Claim 1, U and U' have the same size we get that U' is the set of $x' \in B'$ such that $R_{B'\setminus\{x'\}}$, contains no copy of R_A. The second part of Claim 2 follows immediately.

Claim 3 *Let $i \notin S$ and $j \in S$ then every monomorphic part containing $E_i \cap B$ is disjoint from $E_j \cap B$.*

Proof of Claim 3 According to Claim 2 a monomorphic part containing $E_i \cap B$ must be disjoint from U.

Claim 4 For each $i \in S$, $E_i \cap B$ is a largest monomorphic part of R_B.

Proof of Claim 4 Suppose not. Then this largest monomorphic part, say C, contains some other $E_j \cap B$. From Claim 3, $j \in S$. It follows that all induced substructures on $C \setminus \{x, y\}$, where $\{x, y\}$ is a pair of distinct elements of C, are isomorphic. Suppose $d_i \geq d_j$. Since the shape of A is maximal then for $x, y \in E_j$ the induced structure does not contain a copy of R_A. But if $x \in E_i$ and $y \in E_j$ then trivially the induced structure contains a copy of R_A. A contradiction.

Claim 5 φ *transforms* $(E_i \cap B, i \in S)$ *into* $(E_i \cap B', i \in S')$

Proof of Claim 5. The $E_i \cap B$'s for $i \in S$ are the largest monomorphic parts of $R_{B\setminus A}$. Via φ there are transformed into the s largest monomorphic parts of $R_{B'\setminus U'}$. Since $(E_i \cap B', i \in S')$ is a decomposition of $R_{B'\setminus U'}$ into s monomorphic parts, this decomposition coincides with this decomposition into largest parts.

From Claim 1 and Claim 5, we have $\overline{d'} = \overline{d_1}$. Suppose that $m' > mx_S$. Let T be a transversal of the $E_i \cap B$'s for $i \in S$. Then, from Claim 5, $T' := \varphi(T)$ is a transversal of the $E_{i'} \cap B'$'s for $i' \in S'$. Let m_T, resp. $m_{T'}$, be the monomial associated with $B \setminus T$, resp. $B' \setminus T'$. We have $m'_{T'} > m_T$. Since $m_T = m$ and $B' \setminus T'$ is in the orbit of B, we get a contradiction.

Proof of theorem 2.16. Fix a chain $C = (\emptyset \subsetneq S_1 \subsetneq \cdots \subsetneq S_r \subset X)$ of non empty subsets of X. Let lm_C be the set of leading monomials with chain support C. The plan is essentially to realize lm_C as the linear basis of some ideal of a polynomial ring, so that the generating series of lm_C is realized as an Hilbert series.

Consider the polynomial ring $\mathbb{K}[S_1, \ldots, S_l]$, with its natural embedding in $\mathbb{K}[X]$ by $S_j \mapsto \prod_{i \in S_j} x_i$. Let I be the subspace spanned by the monomials $m = S_1^{r_1} \ldots S_l^{r_l}$ such that $d_i(m) > |E_i|$ for some i; it is obviously a monomial ideal. When all monomorphic components are infinite, I is the trivial ideal $\{0\}$. Consider the subspace $\mathbb{K}.\mathrm{lm}_C$ of $\mathbb{K}[S_1, \ldots, S_l]$ spanned by the monomials in lm_C. Lemma 2.17 exactly states that $J = \mathbb{K}.\mathrm{lm}_C \oplus I$ is in fact a monomial ideal of $\mathbb{K}[S_1, \ldots, S_l]$. Both I and J have finite free resolution as modules over $\mathbb{K}[S_1, \ldots, S_l]$, so that their Hilbert series are rational fractions of the form:

$$\frac{P}{(1 - Z^{|S_1|}) \cdots (1 - Z^{|S_l|})} \ .$$

Hence, the same hold for $\mathcal{H}_{\mathbb{K}.\mathrm{lm}_C} = \mathcal{H}_J - \mathcal{H}_I$. Furthermore, whenever S_j contains i with $|E_i| < \infty$, the denominator $(1 - Z^{|S_l|})$ can be canceled out in $\mathcal{H}_{\mathbb{K}.\mathrm{lm}_C}$. The remaining denominator divides $(1 - Z) \cdots (1 - Z^k)$.

By summing up those Hilbert series $\mathcal{H}_{\mathbb{K}.\mathrm{lm}_C}$ over all chains C of subsets of X, we get the generating series of all the leading monomials, that is the profile of R. Hence, this profile is a rational fraction of the form:

$$\mathcal{H}_{\mathbb{K}.\mathrm{lm}_C} = \frac{P}{(1 - Z) \cdots (1 - Z^k)} \ .$$

$\square$

Remark 2.20. As Examples 2.8 and 4.13 illustrates, it is not true that if all blocks of a monomorphic decomposition of R are infinite, then the numerator P in the above fraction can be choosen with non-negative coefficients.

We do not know for which relational structures having a finite monomorphic decomposition the numerator P can be choosen with non-negative coefficients. A possible approach is to look for some sensible Cohen-Macaulay graded algebra whose Hilbert series is $\mathcal{H}_{\varphi_R}$ (by proposition 4 of [BM04] such a Cohen-Macaulay algebra always exists as soon as P has non-negative coefficients). This is one of our motivations for the upcoming study of the *age algebras*.

3. The age algebra of a relational structure

3.1. **The set-algebra.** Let E be a set and let $[E]^{<\omega}$ be the set of finite subsets of E (including the empty set). Let $\mathbb{K}$ be a field, and $\mathbb{K}^{[E]^{<\omega}}$ be the set of maps $f : [E]^{<\omega} \to \mathbb{K}$. Endowed with the usual addition and scalar multiplication of maps, $\mathbb{K}^{[E]^{<\omega}}$ is a $\mathbb{K}$-vector space. Let $f, g \in \mathbb{K}^{[E]^{<\omega}}$; according to Cameron, we set:

$$fg(P) := \sum_{M\in[P]^{<\omega}} f(P)g(P \setminus M) \ ,$$

for all $P \in [E]^{<\omega}$ [Cam97]. With this operation added, $\mathbb{K}^{[E]^{<\omega}}$ becomes a ring. This ring is commutative and has a unit: denoted by 1, this is the map taking the value 1 on the empty set and the value 0 everywhere else.
Let $\equiv$ be an equivalence relation on $[E]^{<\omega}$. A map $f : [E]^{<\omega} \to \mathbb{K}$ is $\equiv$-*invariant* or, briefly, *invariant* if f is constant on each equivalence class. Invariant maps form a subspace of the vector space $\mathbb{K}^{[E]^{<\omega}}$. We give a condition below which insures that they form a subalgebra too.

Lemma 3.1. *Let $\equiv$ be an equivalence relation on $[E]^{<\omega}$ and $D, D' \in [E]^{<\omega}$. Then the following properties are equivalent:*

(i) There exists some bijective map $f : D \hookrightarrow\!\!\!\!\!\rightarrow D'$ such that $D \setminus \{x\} \equiv D' \setminus \{f(x)\}$ for every $x \in D$;

(ii) 1) $|D| = |D'| = d$ for some d;
2) $|\{X \in [D]^{d-1} : X \equiv B\}| = |\{X \in [D']^{d-1} : X \equiv B\}|$ for every $B \subseteq E$.

An equivalence relation on $[E]^{<\omega}$ is *hereditary* if every pair D, D' of equivalent elements satisfies one of the two equivalent conditions of Lemma 3.1.

Remark 3.2. Hereditary equivalences are introduced in [PR86] with Condition (ii) 2) of Lemma 3.1 replaced by the condition:

$$|\{X \subseteq D : X \equiv B\}| = |\{X \subseteq D' : X \equiv B\}| \ \text{ for every } B \subseteq E \ .$$

It follows from the next Lemma that this condition is not stronger.

Let $\equiv$ be an equivalence relation on $[E]^{<\omega}$. We denote by $[E]^{<\omega}_{/\equiv}$ the set of equivalence classes. Let $a, b, c \in [E]^{<\omega}_{/\equiv}$ and $D \in [E]^{<\omega}$. Set

$$\chi_{a,b,c}(D) := |\{(A, B) \in a \times b : A \cup B = C, C \subseteq D, C \in c\}| \ .$$

If all subsets of E belonging to some equivalence class a have the same size, we denote by $|a|$ this common size.

Lemma 3.3. *If $\equiv$ is an hereditary equivalence relation on $[E]^{<\omega}$ then*

$$\chi_{a,b,c}(D) := \chi_{a,b,c}(D') \text{ whenever } D \equiv D'. \tag{1}$$

Proposition 3.4. *Let $\equiv$ be an hereditary equivalence relation on $[E]^{<\omega}$. Then the product of two invariant maps is invariant.*

3.2. **The age algebra.** Let R be a relational structure with domain E. Set $F \equiv F'$ for $F, F' \in [E]^{<\omega}$ if the restrictions $R\restriction_F$ and $R\restriction_{F'}$ are isomorphic. The resulting equivalence on $[E]^{<\omega}$ is hereditary, hence the set of invariant maps $f : [E]^{<\omega} \to \mathbb{K}$ form a subalgebra of $\mathbb{K}^{[E]^{<\omega}}$. Let $\mathbb{K}.\mathcal{A}(R)$ be the subset made of the invariant maps which are everywhere zero except on a finite number of equivalence classes. Then $\mathbb{K}.\mathcal{A}(R)$ forms an algebra, the *age algebra* of Cameron.

4. Invariant rings of permutation groupoids

Let $\mathbb{K}$ be a field of characteristic 0. In this section, we study in more details a specific class of age algebras which can be realized as graded subrings of polynomial rings $\mathbb{K}[X]$ that we call *invariant rings of permutation groupoids.* This class extends the class of invariant rings of permutation groups (Example 2.11), and contains other interesting examples like the rings of quasi-symmetric polynomials (Example 2.13). Our long-term motivations are twofold. On one hand, relate, in this simpler yet rich setting, the properties of the profile to algebraic properties of the invariant ring. In particular, find conditions under which the invariant ring is Cohen-Macaulay. On the other hand, generalize the theory, algorithms, and techniques of invariant rings of permutation groups to a larger class of subrings of $\mathbb{K}[X]$. In particular, find new properties of the ring of quasi-symmetric polynomials, one specific goal being to find a simpler proof that this ring is Cohen-Macaulay.

4.1. **Permutation groupoids.** Let X be a finite set. A *local bijection* of X is a bijective function $f : \mathrm{dom} f \hookrightarrow\!\!\!\!\rightarrow \mathrm{im} f$ whose domain $\mathrm{dom} f$ and image $\mathrm{im} f$ are subsets of X. The *rank* of f is the size of its domain, so that f is a permutation of X if it is of maximal rank $|X|$. The inverse f^{-1} of a local bijection f, its restriction $f_{\restriction X'} : X' \hookrightarrow\!\!\!\!\rightarrow f(X')$ to a subset X' of $\mathrm{dom} f$, and the composition $f \circ g$ of two local bijections f and g such that $\mathrm{im} g = \mathrm{dom} f$ are defined in the natural way. A set G of local bijections of X is called a *permutation groupoid* if it contains the identity and is stable by restriction, inverse, and composition. It can be seen as a category: the objects are the subsets of X and the morphisms $G(A, B)$ from A to B are the local bijections $f : A \hookrightarrow\!\!\!\!\rightarrow B$ in G. Those morphisms are by definition isomorphisms, and G satisfies the usual *groupoid* axioms.

The *underlying permutation group* is the subset $G(X, X)$ of all permutations in G; those are exactly the invertible elements w.r.t. the composition product.

Examples 4.1. The set $\downarrow\!\mathfrak{S}(X)$ of all local bijections of X is a permutation groupoid.

The closure $\downarrow\!G$ of a permutation group G by restriction is a permutation groupoid. In the following, we say that $\downarrow\!G$ *comes from the permutation group* G.

Let $X := \{1, \dots, n\}$. The set G of strictly increasing local bijections of X forms a permutation groupoid. Obviously G does not come from a permutation group since its underlying permutation group is reduced to the identity.

Let R be a relational structure on X. The local isomorphisms of R form a permutation groupoid. Its underlying permutation group is the automorphism group of R. Typically, the previous example is obtained by taking as relational structure R the chain $1 < 2 < \cdots < n$. Also, $\downarrow\!\mathfrak{S}(X)$ is obtained by taking the trivial relational structure on X. In fact, *any permutation groupoid* G *can be obtained from a suitable relational structure* R_G *on* X (recall that X is finite!).

4.1.1. *The invariant ring of a permutation groupoid.* Let G be a permutation groupoid acting on a finite set X, and $\mathbb{K}[X]$ be the polynomial ring whose variables x_i are indexed by the elements i of X. Given a local function f of G, and a monomial $x^d := \prod_{i\in X} x_i^{d_i}$ whose support support(x^d) is contained in dom$f$, we set

$$f.x^d := \prod_{i\in X, d_i>0} x_{f(i)}^{d_i} ,$$

generalizing the usual action of a permutation on a monomial. This partial action of G on monomials does not extend to a global action of G. Still, notions like *G-isomorphic monomials* and *G-orbits* are well defined. The *orbit sum* $o(x^d)$ of a monomial x^d is the sum of all the monomials in its orbit.

Our object of study is the *invariant ring* $\mathbb{K}[X]^G$ of G, which is defined as the linear subspace of $\mathbb{K}[X]$ spanned by the orbitsums of all monomials.

Examples 4.2. Let G be a permutation group. Then, $\mathbb{K}[X]^{\downarrow G}$ is the usual invariant ring of G.

Let G be the permutation groupoid of the strictly increasing local bijections of $\{1,\dots,n\}$. Then $\mathbb{K}[X]^G$ is the ring $\mathrm{QSym}(X_n)$ of quasi-symmetric polynomials on the ordered alphabet $X_n := (x_1,\dots,x_n)$.

Taking the same groupoid G as in the previous example, and letting it act naturally on respectively pairs, couples, k-subsets, or k-tuples of elements of $\{1,\dots,n\}$, yields respectively the (un)oriented (hyper)graph quasi-symmetric polynomials of [NTT04].

Remarks 4.3. The orbitsums form a linear basis of $\mathbb{K}[X]^G$.

It is not obvious from the definition that $\mathbb{K}[X]^G$ is indeed a graded algebra. In the following we prove this by making G into a monoid and the action of G on polynomials into a multiplicative linear representation of G. An other way is to encode G by some relational structure R_G on X and, as in Example 2.11, to define a relational structure $R_G \wr \mathbb{N}$ on $E := X\times\mathbb{N}$ with monomorphic components $E_i := \{i\}\times\mathbb{N}$ for $i\in X$. Let $\phi : \mathbb{K}[X] \hookrightarrow \mathbb{K}^{[X\times\mathbb{N}]^{<\omega}}$ defined by setting $\phi(x^d) := d!\chi_{O_{\mathfrak{G}}(x^d)}$, where $d! := \prod_{i\in X} d_i!$, and $\chi_{O_{\mathfrak{G}}(x^d)}$ is the characteristic function of $O_{\mathfrak{G}}(x^d) := \{A \subset X\times\mathbb{N}, |A\cap E_i| = d_i, \forall i\in X\}$. Once $\mathbb{K}^{[X\times\mathbb{N}]^{<\omega}}$ is equipped with its set-algebra stucture, ϕ is a morphism of algebras. Applying Proposition 3.4, we get:

Theorem 4.4. *The invariant ring $\mathbb{K}[X]^G$ is isomorphic via ϕ to the age algebra $\mathbb{K}.\mathcal{A}(R_G\wr\mathbb{N})$. In particular, the generating series of the orbits, the generating series of the profile of $R_G\wr\mathbb{N}$, and the Hilbert series of $\mathbb{K}[X]^G$ coincide.*

4.1.2. *Restrictions of permutation groupoids.* The *restriction* $G_{\restriction X'}$ of a permutation groupoid G to a subset X' is the set of all local functions f in G such that dom$f \subset X'$ and im$f \subset X'$, which is again a permutation groupoid. Furthermore, the orbits of monomials in $\mathbb{K}[X']$ are unchanged by this restriction. In particular, the invariant ring of $G_{\restriction X'}$ is simply the quotient of the invariant ring of G obtained by killing all the variables x_i with $i \notin X'$. This simple fact is one of the points of considering permutation groupoids instead of just permutation groups (for which the restriction to a subset is not clearly defined). This may indeed give opportunities for induction techniques on the size of the underlying set.

Proposition 4.5. *Any permutation groupoid comes from the restriction of a permutation group of some superset. However, this superset may need to be infinite.*

Examples 4.6. (a) The permutation groupoid on $\{1,2,3\}$ generated by the rank 1 local bijection $1 \mapsto 2$ is the restriction of the permutation group on $\{1,2,3,4\}$ generated by the permutation $(1,2)(3,4)$.

(b) The local automorphism permutation groupoid of the chain $a < b$ is the restriction of the cyclic group C_3 on $\{a,b,c\}$.

(c) Consider a relational structure R such that there exists three elements a,b,c and a binary relation $<$ which restricts on $\{a,b,c\}$ to the chain $a < b < c$. Typically, R is a chain of length at least 3 (giving $\mathrm{QSym}(X)$ as invariants) or a poset of height at least 3. Then, there exists no relational structure $\overline{R}$ on a finite superset where all local isomorphisms extend to global isomorphisms.

4.1.3. *The monoid of a permutation groupoid.* The goal is now to turn G into a monoid, and to make the partial actions of G into a linear representations of this monoid. The composition of two local functions f and g can be extended when $\mathrm{im} f \neq \mathrm{im} g$ by setting it to the local function with the largest domain on which $f(g(x))$ is well-defined:

$$f \circ g : \begin{cases} g^{-1}(\mathrm{im} g \cap \mathrm{dom} f) & \hookrightarrow\!\!\!\!\rightarrow f(\mathrm{im} g \cap \mathrm{dom} g) \\ x \mapsto f(g(x)) \end{cases} .$$

With this composition product, G turns into a monoid whose unit is the identity of X. Now we can extend the partially defined action of local bijections on monomials into a linear action on polynomials by setting:

$$f.x^d := \begin{cases} \displaystyle\prod_{i \in X, d_i > 0} x_{f(i)}^{d_i} & \text{if } \mathrm{support}(x^d) \subset \mathrm{dom} f, \\ 0 & \text{otherwise.} \end{cases}$$

We leave it as exercise to check that this defines a linear representation of the monoid G, which is *multiplicative*: for any f in G and P and Q in $\mathbb{K}[X]$,

$$f.(PQ) = (f.P)(f.Q) .$$

Corollary 4.7. *The invariant ring $\mathbb{K}[X]^G$ is, as its name suggests, indeed a ring.*

Proof. Consider a product of two orbitsums $o(m_1)o(m_2)$, and take two isomorphic monomials m and $f.m$, $f \in G$. Whenever m occurs as a product $m = m_1'm_2'$, $m_1' \in G.m_1$, $m_2' \in G.m_2$, the monomial $f.m$ occurs simultaneously as the product $f.m = f.(m_1'm_2') = f.m_1' f.m_2'$, and reciprocally. Hence m and m' occur with the same coefficient in $o(m_1)o(m_2)$. □

Note that the monoid algebra of G is isomorphic to its groupoid algebra $\mathbb{K}.G$ which is semi-simple. This linear representation of G extends into a linear representation of $\mathbb{K}.G$.

4.1.4. *Groupoid and monoid algebra of a permutation groupoid.* Let G be a permutation groupoid, and $\mathbb{K}$ a field (of characteristic zero; typically $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). By definition, its *groupoid algebra* $\mathbb{K}.G$ is the $\mathbb{K}$-vector space whose basis $\{\mathrm{gr} f, f \in G\}$ is indexed by the elements f of G, and whose product is given by:

$$\mathrm{gr} f \ \mathrm{gr} g = \begin{cases} f \circ g & \text{if } \mathrm{im} g = \mathrm{dom} f, \\ 0 & \text{otherwise.} \end{cases}$$

We call $\{\mathrm{gr} f, f \in G\}$ the *graded basis* of $\mathbb{K}.G$. Similarly the *monoid algebra* of G is defined as the $\mathbb{K}$-vector with basis $\{f, f \in G\}$ equipped with the extended composition product $\circ$.

Remarks 4.8. As the notation suggests, the groupoid and the monoid algebra of G are isomorphic.

The groupoid algebra $\mathbb{K}.G$ is semi-simple, and decomposes as a direct sum of non-unitary algebras:

$$\mathbb{K}.G = \sum_{k=0}^{n} \mathbb{K}.G_k \ ,$$

where $G_k := \{f \in G, \mathrm{rank} f = k\}$.

Proof. The isomorphism from the monoid algebra to the groupoid algebra is given by:

$$f \mapsto \sum_{A \subset \mathrm{dom} f} \mathrm{gr} f_{\restriction A} \ .$$

The inverse isomorphism is obtained by Möbius inversion:

$$\mathrm{gr} f \mapsto \sum_{A \subset \mathrm{dom} f} (-1)^{|\mathrm{dom} f| - |A|} f_{\restriction A} \ .$$

Checking the compatibility with the product rule is straightforward.

The semi-simplicity is a general property of groupoid algebras, which is easily checked using Dickson's Lemma. □

The linear representations of the monoid G on polynomials extend directly to linear representations of its algebra $\mathbb{K}.G$. In particular, this defines the actions of the graded basis. Its characteristic is that $\mathrm{gr} f$ kills all monomials whose support is not exactly $\mathrm{dom} f$, whereas f kills only those monomials whose support is not contained in $\mathrm{dom} f$.

Important note: the action of $\mathrm{gr} f$ is *not* multiplicative on polynomials! Take for example $f := \mathrm{id}_{\{1,2\}}$, $P := x_1$ and $Q := x_2$. This is in fact the main reason for considering the monoid algebra and not just only the groupoid algebra.

4.2. **Invariants of permutation groupoids.** In this section, we review which properties of invariants of permutation groups extend to permutation groupoids.

4.2.1. *The Reynolds operator.* The first essential feature of invariant rings is the so-called *Reynolds operator*, which is a projector on the invariant ring. The following proposition states that this operator still exists for invariants of permutation groupoids, albeit missing the important property of being a $\mathbb{K}[X]^G$-module morphism. In particular, although $\mathbb{K}[X]^G$ still contains the ring of symmetric polynomials $\mathrm{Sym}(X)$, R is not anymore a $\mathrm{Sym}(X)$-module morphism.

Proposition 4.9. *There exists an idempotent R in the groupoid algebra $\mathbb{K}.G$ which projects $\mathbb{K}[X]$ on the invariant ring $\mathbb{K}[X]^G$:*

$$R := \sum_{A \subset X} \frac{1}{|g \in G, \mathrm{dom} g = A|} \sum_{g \in G, \mathrm{dom} g = A} g r g \ .$$

Furthermore, the four following conditions are equivalent: R is a $\mathrm{Sym}(X)$*-module morphism, R is a $\mathbb{K}[X]^G$-module morphism,* $\ker R$ *is a* Sym*-module, and G comes from a permutation group.*

4.2.2. *The chain product.* We now define another product $\star$ on the invariant ring $\mathbb{K}[X]^G$, called the *chain product*, which preserves a finer grading. In fact, $(\mathbb{K}[X]^G, \star)$ is a simple realization of the Stanley-Reisner ring of a suitable poset. Such rings have been studied intensively, in particular by Garsia and Stanton [GS84] to construct $\mathrm{Sym}(X)$-module generators for the invariant rings of certain permutation groups, and prove the degree bound for permutation groups $\beta(G) \leq \binom{|X|}{2}$ (recall that the *degree bound* $\beta(A)$ of a finitely generated graded algebra A is the smallest integer such that A is generated by its elements of degree at most $\beta(A)$). This tool is characteristic free: all statements below actually hold over any ground ring.

Given a subset S of X, set $x_S := \prod_{i \in S} x_i$. By square-free decomposition, any monomial x^d can be identified uniquely with a *multichain* $S_1 \subset \cdots \subset S_k$ of nested subsets of X, so that:

$$x^d = x_{S_1} \dots x_{S_k} \ .$$

We call each S_k a *layer* of x. The *fine degree* of the monomial x^d is the integer vector $(r_1, \dots, r_n)$ where each r_i counts the (possibly null) number of repetitions of the layer of size i in x^d. The fine degree defines a filtration on $\mathbb{K}[X]$. The *chain product* $\star$ of two monomials $x^d = x_{S_1} \dots x_{S_k}$ and $x^{d'} = x_{S'_1} \dots x_{S'_k}$ is defined by:

$$x^d \star x^{d'} := \begin{cases} x^d x^{d'} & \text{if } \{S_1, \dots, S_k, S'_1, \dots, S'_k\} \text{ is again a multichain of subsets,} \\ 0 & \text{otherwise.} \end{cases}$$

For example, $x_1 \star x_1 = x_1^2$, $x_1 \star x_2 = 0$, $x_1 x_3^2 \star x_1 x_2 x_3^2 = x_1^2 x_2 x_3^4$, and $x_1 x_3^2 \star x_1 x_2 = 0$.

The chain product endows $(\mathbb{K}[X], \star)$ with a second algebra structure (in fact $(\mathbb{K}[X], \star)$ is isomorphic to the quotient $\mathbb{K}[x_S, S \subset X]/\{x_S x_{S'} = 0, S \not\subset S' \text{ and } S' \not\subset S\}$). It is also finely graded (fine degrees being added term-by-term). In fact, $(\mathbb{K}[X], \star)$ is exactly the associated graded algebra of $\mathbb{K}[X]$ w.r.t. the fine degree filtration. Beware that $(\mathbb{K}[X], \star)$ is not an integral domain.

The elementary symmetric functions

$$e_d := \sum_{S \subset X, |S| = d} x_S$$

are still algebraically independent and generate $(\mathrm{Sym}(X)_n, \star)$. Note that this does not hold for, say, the symmetric powersums. The following simple fact turns out to be an essential key:

Remark 4.10. Consider the chain product of a monomial $x_{S_1} \dots x_{S_k}$ by the elementary symmetric function e_d. It is the sum of all monomials $x_{S_1} \dots x_S \dots x_{S_k}$, where S is of size k, and fits in the chain $S_1 \subset \cdots \subset S \subset \cdots \subset S_k$. In particular, if $x_{S_1} \dots x_{S_k}$ readily contains a layer S of size k, then $x_{S_1} \dots x_{S_k} \star e_k$ is the unique monomial obtained by replicating this layer.

More generally, $(\mathbb{K}[X]^G, \star)$ is a subring of $(\mathbb{K}[X], \star)$. In particular, $(\mathbb{K}[X]^G, \star)$ is a $\mathrm{Sym}(X)$-module. Furthermore, we may transfer the following algebraic properties from $(\mathbb{K}[X], \star)$ to $\mathbb{K}[X]^G$, as in the case of permutation groups [GS84].

Proposition 4.11. (a) *A family F of finely homogeneous invariants of positive degree which generates $(\mathbb{K}[X]^G, \star)$, also generates $\mathbb{K}[X]^G$;*
(b) *$\beta(\mathbb{K}[X]^G, \star) \geq \beta(\mathbb{K}[X]^G)$;*
(c) *A family F of finely homogeneous invariants which generates $(\mathbb{K}[X]^G, \star)$ as a $\mathrm{Sym}(X)$-module also generates $\mathbb{K}[X]^G$ as a $\mathrm{Sym}(X)$-module;*
(d) *If $(\mathbb{K}[X]^G, \star)$ is a free $\mathrm{Sym}(X)$-module, then so is $\mathbb{K}[X]^G$.*

Proof. This is a standard fact about filtrations and associated graded connected algebras. The key of the proof is that, if p and q are finely homogeneous, the maximal finely homogeneous component of pq is exactly $p \star q$. (a) and (c) follow by induction over the fine grading. Then, (b) follows straightaway from (a), and (d) from (c) by a simple Hilbert series argument. □

The converse of (a) and (b) do not hold. In fact, with most permutation groups, the degree bound $\beta(\mathbb{K}[X]^G, \star)$ is much larger than $\beta(\mathbb{K}[X]^G)$. We conjecture that the converse of (c) and (d) hold. However (d) does not hold anymore in a slightly larger setting which includes the r-quasi-symmetric polynomials of F. Hivert [Hiv04], a counter example being $\mathrm{QSym}^2(X_3)$ (there is an obstruction in the fine Hilbert series).

Theorem 4.12. *The invariant ring $\mathbb{K}[X]^G$ is a finitely generated algebra and $\mathrm{Sym}(X)$-module, in degree at most $\frac{|X|(|X|+1)}{2}$. This degree bound is tight.*

Note that, as usual, when G does not act transitively on the variables, the degree bound can be greatly improved by considering the elementary symmetric polynomials on each transitive component instead.

Proof. The set of orbit sums $o(x_{S_1} \dots x_{S_k})$, where $S_1 \subsetneq \cdots \subsetneq S_k$ is a chain, generate $(\mathbb{K}[X]^G, \star)$ as a $(\mathrm{Sym}, \star)$-module. This transfers back to $\mathbb{K}[X]^G$ and Sym.

Note that we may need to consider chains with $S_k = X$; hence the degree bound of $\frac{|X|(|X|+1)}{2}$ instead of $\binom{|X|}{2}$ for permutation groups. For an example where the bound is achieved, consider the group G made of the identity together with all the local bijections of $X = \{1, \dots, n\}$ whose domain is of size at most $|X| - 1$; then, $\mathbb{K}[X]^G$ is freely generated as a Sym-module by 1 and the "staircase" monomials $x_1^{d_1}, \dots, x_n^{d_n}$ with $1 \leq d_i \leq i$. □

4.2.3. *The Cohen-Macaulay property.* Invariant rings of permutation groups are always Cohen-Macaulay, and in fact free $\mathrm{Sym}(X)$-modules. This follows easily from the fact that the Reynolds operator is a $\mathrm{Sym}(X)$-module morphism. A recent and more involved result is that, for all n, $\mathrm{QSym}(X_n)$ is also a free $\mathrm{Sym}(X_n)$-module [GW03].

As the following example will show, this property does not hold for all permutation groupoids G. Still, $\mathbb{K}[X]^G$ and $(\mathbb{K}[X]^G, \star)$ being finitely generated over $\mathrm{Sym}(X)$, they are Cohen-Macaulay if and only if they are free $\mathrm{Sym}(X)$-modules.

Example 4.13. Let G be the permutation groupoid on $\{1, 2, 3\}$ of example 4.6 (a), generated by the local bijection $1 \mapsto 2$. Then, G is the restriction of a finite permutation group whose invariant ring is Cohen-Macaulay. However, $\mathbb{K}[X]^G$ itself is not Cohen-Macaulay. Computing the Hilbert series shows right away that this module is not free:

$$\mathcal{H}_{\mathbb{K}[X]^G} = \frac{1}{(1-Z)^3} - \frac{Z}{1-Z} = \frac{1+Z+2Z^2+2Z^3+Z^4-Z^6}{(1-Z)(1-Z^2)(1-Z^3)} \ .$$

To be more explicit, the transitive components of G being $\{1,2\}$ and $\{3\}$, we may replace $\mathrm{Sym}(X)$ by $R = \mathrm{Sym}(x_1, x_2) \otimes \mathrm{Sym}(x_3)$, and view $\mathbb{K}[X]^G$ as a finitely generated R-module. Then, as suggests the Hilbert series,

$$\mathcal{H}_{\mathbb{K}[X]^G} = \frac{1+Z^2+Z^3-Z^4}{(1-Z)^2(1-Z^2)} \ .$$

$\mathbb{K}[X]^G$ is minimally generated as an R-module by $(1, x_1x_3, x_1^2x_2)$, subject to the single relation $x_3.(x_1^2x_2) = (x_1x_2).(x_1x_3)$.

Finally, there is no way of choosing the numerator of the Hilbert series with non-negative coefficients. Indeed, $\mathcal{H}_{\mathbb{K}[X]^G} = \frac{1-Z+2Z^2-Z^3}{(1-Z)^3}$, and the coefficient of highest degree in the product of the numerator by $\frac{(1-Z^{n_1})(1-Z^{n_2})(1-Z^{n_3})}{(1-Z)^3}$ is always -1.

Problem 4.14. Characterize the permutation groupoids G whose invariant rings $\mathbb{K}[X]^G$ (or $(\mathbb{K}[X]^G, \star)$) are Cohen-Macaulay.

The following theorem is a straightforward extension of a theorem of [GS84].

Theorem 4.15. *$(\mathbb{K}[X]^G, \star)$ is a free $\mathrm{Sym}(X)$-module if and only if the incidence matrix between generators and maximal chains is invertible. In particular, for a set F of finely homogeneous invariants whose fine degrees are given by the Hilbert series of $\mathbb{K}[X]^G$, the three following conditions are equivalent: F spans $\mathbb{K}[X]^G$ as a $\mathrm{Sym}(X)$-module, F is a free $\mathrm{Sym}(X)$-family, and F is a $\mathrm{Sym}(X)$-basis of $\mathbb{K}[X]^G$.*

This readily gives us a necessary condition on the number of generators.

Corollary 4.16. *If $(K[X]^G, \star)$ is a free $\mathrm{Sym}(X)$-module, then it is of rank $\frac{|X|!}{|G(X,X)|}$, where $G(X,X)$ is the underlying permutation group of G.*

4.2.4. *SAGBI bases.* SAGBI bases (Subalgebra Analog of a Gröbner Bases for Ideals) were introduced in [KM89, RS90] to develop an elimination theory in subalgebras of polynomial rings. Unlike Gröbner bases, not all subalgebras have a finite SAGBI basis, and it remains a long open problem to characterize those subalgebras which have a one. The following theorem states that, as in the case of permutation groups, invariant rings of permutation groupoids seldom have finite SAGBI bases. The proof follows the short proof given by the second author in [TT04] for permutation groups, with some adaptations. For example $\mathrm{QSym}(X_n)$, represented as a subring of $\mathbb{K}[X]$, has no finite SAGBI basis whenever $n > 1$. In particular, $\mathrm{QSym}(X_2)$ becomes the smallest example of finitely generated algebra which has no finite SAGBI basis (the standard example being the invariant ring of the alternating group A_3). Still, SAGBI bases and SAGBI-Gröbner bases provide a useful device in the computational study of invariant rings of permutation groups [Thi01], and most likely play the same role with permutation groupoids.

Theorem 4.17. *Let G be a permutation groupoid, and $<$ be any admissible term order on $\mathbb{K}[X]$. Then, the invariant ring $\mathbb{K}[X]^G$ has a finite SAGBI basis w.r.t. $<$ if, and only if, G comes from a permutation group generated by reflections (that is transpositions).*

The following proof is a close variant on the short proof given by the second author in [TT04] in the special case of permutation groups. For the sake of readability and completeness, we include it in full here. The key fact is that a submonoid M of $\mathbb{N}^n$ is finitely generated if, and only if, the convex cone $C := \mathbb{R}_+M$ it spans in $\mathbb{R}_+^n$ is finitely generated (that is C is a *polyhedral cone*). For details, see for example [BG05, Corollary 2.8]. In particular C must be the intersection of finitely many half spaces, and thus closed for the euclidean topology.

Proof. The if-part is easy, a finite SAGBI basis being given by the elementary symmetric polynomials in the variables in each G-transitive components.

Without loss of generality, we may assume $X = \{1, \dots, n\}$ with $x_1 > \dots > x_n$. Let M be the monoid of initial monomials in $\mathbb{K}[X]^G$, seen as a submonoid of $\mathbb{N}^n$, and $C := \mathbb{R}_+M$ be the convex cone it spans in $\mathbb{R}_+^n$.

At this stage, we cannot give an explicit description of C, but we can construct a convex cone C' which approximates it closely enough for our purposes. By the standard characterization of admissible term orders on $\mathbb{K}[X]$, there exists a family of n linear forms $l = (l_1, \dots, l_n)$ such that $x^d > x^{d'}$ if and only if $l(d) >_{\text{lex}} l(d')$, where we denote by $l(d)$ the n-uple $l_1(d_1, \dots, d_n), \dots, l_n(d_1, \dots, d_n)$. Given two vectors v and v' in $\mathbb{R}_+^n$, we write $v > v'$ if $l(v) >_{\text{lex}} l(v')$. The partial action of G on monomials extends naturally to a partial action on $\mathbb{R}_+^n$: whenever the support of $v = (v_1, \dots, v_n)$ in contained in the domain of a local bijection $f \in G$, $f.v$ is the vector obtained by permuting the non zero entries of v according to f. Let C' be the subset of all vectors v of $\mathbb{R}_+^n$ such that $v > f.v$ for all $f.v$ in the G-orbit of v. In fact, C' is a convex cone with non empty interior (it contains the n linearly independent vectors $(1, 0, \dots, 0), (1, 1, 0, \dots, 0), \dots, (1, \dots, 1)$). By construction, M consists of the points of C' with integer coordinates. It follows that $C \subset C'\overline{C}$, where $\overline{C}$ is the topological closure of C.

Assume now that M is finitely generated. Then, C is a closed convex cone, and C and C' simply coincide.

Assume further that G is not generated by transpositions. Then, there exists $a < b$ such that the transposition (a, b) is not in G, while a is in the G-orbit of b. Choose such a pair $a < b$ with b minimal. We claim that there is no transposition (a', b) in G with $a' < b$. Otherwise, a and a' are in the same G-orbit, and by minimality of b, $(a, a') \in G$; thus, $(a, b) = (a, a')(a', b)(a, a') \in G$. Pick $g \in G$ such that $g.b = a$, and for $t \geq 0$, define the vector in $\mathbb{R}_+^n$:

$$u_t := (nt,\ (n-1)t,\ \dots,\ (n-b+2)t,\ n-b+1,\ (n-b)t,\ \dots,\ t,\ 1)\ .$$

Note that $u_1 = (n, \dots, 1)$ is in C, whereas $u_0 = (0, \dots, 0, n-b+1, 0, \dots 0)$ is not in C because $g.u_0 > u_0$.

Take t such that $0 < t \leq 1$. Then, the vector u_t has no zero coefficients, and in particular its G-orbit coincides with its orbit w.r.t. the underlying permutation group $G(X, X)$. Furthermore, the entries of u_t are all distinct, except when $t = \frac{n-b}{n-a'}$ for some $a' < b$, in which case the a'-th and b-th entries are equal. Since $(a', b) \notin G$, the orbit of u_t is of size $|G(X, X)|$, and there exists a unique permutation $f_t \in G(X, X)$ such that $f_t.u_t$ is in C.

Let $t_0 = \inf\{t \geq 0, u_t \in C\}$. If $u_{t_0} \notin C$, then u_{t_0} is in the closure of C, but not in C, a contradiction. Otherwise, $u_{t_0} \in C$, and $t_0 > 0$ because $u_0 \notin C$. For any permutation f, $\{f.u_t, t \geq 0\}$ is a half-line; so, C being convex and closed, $I_f := \{t, f.u_t \in C\}$ is a closed interval $[x_f, y_f]$. For example, $I_{\mathrm{id}} = [t_0, 1] \subsetneq [0, 1]$. Since the interval $[0, 1]$ is the union of all the I_f, there exists $f \neq \mathrm{id}$ such that $t_0 \in I_f$. This contradicts the uniqueness of f_{t_0}. □

4.3. **Stability by derivation.** We denote by ∂_i the derivative w.r.t. the variable x_i, and consider the derivation $D := \sum_{i \in X} \partial_i$ on $\mathbb{K}[X]$.

Proposition 4.18. *Let G be a permutation groupoid. Then $\mathbb{K}[X]^G$ is stable by the derivation D if and only if G comes from a permutation group. On the other hand, $\mathbb{K}[X]^G$ is always stable w.r.t. the action of the* rational Steenrod operators $S_k := \sum_i x_i^{k+1} \partial_i$ *for $k \geq 0$ (see* [HT04] *for details on the rational Steenrod operators).*

Proof. The if part is trivial, since D commutes with the action of the symmetric group $\mathfrak{S}_X$ on $\mathbb{K}[X]$. Similarly, the rational Steenrod operators always stabilize $\mathbb{K}[X]^G$ because they commute with the action of any local bijection on $\mathbb{K}[X]^G$.

Assume now that $\mathbb{K}[X]^G$ is stable by derivation. Let $f : A \mapsto B$ be a local bijection such that $A \subsetneq X$, and take i in $X \backslash A$. We just need to prove that f extends to a local bijection g in G with domain $A \cup \{i\}$. Applying induction, any local bijection in G will then extend to a permutation, as desired.

Take a monomial m whose support is A and whose exponents are all distinct and at least 2, and consider the derivation $p = D(o(mx_i))$ of the orbitsum of the monomial mx_i in $\mathbb{K}[X]^G$. The monomial m occurs in p; hence, by invariance of p, $f(m)$ also occurs in p, as the derivative of some monomial $g(mx_i)$ in the orbit of mx_i. By the choice of the exponents of m, f and g must coincide on A, while at the same time i belongs to the domain of g. □

Example 4.19. $\mathrm{QSym}(X_2)$ has no graded derivation of degree -1.

References

[BG05] Winfried Bruns and Joseph Gubeladze. Polytopes, rings, and k-theory. February 2005.

[BM04] Rikard Bogvad and Thomas Meyer. On algorithmically checking whether a hilbert series comes from a complete intersection. Research Reports in Mathematics 5, Department of Mathematics, Stockholm University, may 2004.

[Cam90] Peter J. Cameron. *Oligomorphic permutation groups*, volume 152 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1990.

[Cam97] Peter J. Cameron. The algebra of an age. In *Model theory of groups and automorphism groups (Blaubeuren, 1995)*, volume 244 of *London Math. Soc. Lecture Note Ser.*, pages 126–133. Cambridge Univ. Press, Cambridge, 1997.

[FP71] Roland Fraïssé and Maurice Pouzet. Interprétabilité d'une relation pour une chaîne. *C. R. Acad. Sci. Paris Sér. A-B*, 272:A1624–A1627, 1971.

[Fra71] Roland Fraïssé. *Cours de logique mathématique. Tome 1: Relation et formule logique.* Gauthier-Villars Éditeur, Paris, 1971. Deuxième édition revue et modifiée, Collection de "Logique Mathématique", Série A, No. 23.

[Fra00] Roland Fraïssé. *Theory of relations*, volume 145 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, revised edition, 2000. With an appendix by Norbert Sauer.

[GS84] A. M. Garsia and D. Stanton. Group actions of Stanley - Reisner rings and invariants of permutation groups. *Adv. in Math.*, 51(2):107–201, 1984.

[GW03] A. M. Garsia and N. Wallach. Qsym over Sym is free. *J. Combin. Theory Ser. A*, 104(2):217–263, 2003.

[Hiv04] Florent Hivert. Local actions of the symmetric group and generalisations of quasi-symmetric functions. Preprint, October 2004.

[HM88] I. M. Hodkinson and H. D. Macpherson. Relational structures determined by their finite induced substructures. *J. Symbolic Logic*, 53(1):222–230, 1988.

[HT04] F. Hivert and N. M. Thiéry. Deformation of symmetric functions and the rational Steenrod algebra. In *Invariant theory in all characteristics*, volume 35 of *CRM Proc. Lecture Notes*, pages 91–125. Amer. Math. Soc., Providence, RI, 2004.

[KM89] Deepak Kapur and Klaus Madlener. A completion procedure for computing a canonical basis for a k-subalgebra. In *Computers and mathematics (Cambridge, MA, 1989)*, pages 1–11. Springer, New York, 1989.

[NTT04] Jean-Christophe Novelli, Jean-Yves Thibon, and Nicolas M. Thiéry. Algèbres de Hopf de graphes. *C. R. Math. Acad. Sci. Paris*, 339(9):607–610, 2004.

[Pou76] Maurice Pouzet. Application d'une propriété combinatoire des parties d'un ensemble aux groupes et aux relations. *Math. Z.*, 150(2):117–134, 1976.

[Pou78] Maurice Pouzet. *Sur la théorie des relations*. PhD thesis, Thèse d'état, Université Claude-Bernard, Lyon 1, 1978.

[Pou81] Maurice Pouzet. Application de la notion de relation presque-enchaînable au dénombrement des restrictions finies d'une relation. *Z. Math. Logik Grundlag. Math.*, 27(4):289–332, 1981.

[Pou02] Maurice Pouzet. The profile of relations. Preprint, lecture at Simon Bolivar University, August 2002.

[PR86] M. Pouzet and I. G. Rosenberg. Sperner properties for groups and relations. *European J. Combin.*, 7(4):349–370, 1986.

[RS90] Lorenzo Robbiano and Moss Sweedler. Subalgebra bases. In *Commutative algebra (Salvador, 1988)*, pages 61–87. Springer, Berlin, 1990.

[Thi01] Nicolas M. Thiéry. Computing minimal generating sets of invariant rings of permutation groups with SAGBI-Gröbner basis. In *Discrete models: combinatorics, computation, and geometry (Paris, 2001)*, Discrete Math. Theor. Comput. Sci. Proc., AA, pages 315–328 (electronic). Maison Inform. Math. Discrèt. (MIMD), Paris, 2001.

[TT04] N. M. Thiéry and S. Thomassé. Convex cones and SAGBI bases of permutation invariants. In *Invariant theory in all characteristics*, volume 35 of *CRM Proc. Lecture Notes*, pages 259–263. Amer. Math. Soc., Providence, RI, 2004.

UFR de Mathématiques, Université Claude-Bernard, 43, Bd. du 11 Novembre 1918, 69622 Villeurbanne, France, Fax 33 4 37 28 74 80

E-mail address: pouzet@univ-lyon1.fr, nthiery@users.sf.net

MAGIC SQUARES, ROOK POLYNOMIALS AND PERMUTATIONS

FANJA RAKOTONDRAJAO

Abstract. We study in this paper the vector space of magic squares and their relation with some restricted permutations.

Résumé. Nous étudions dans cet article l'espace vectoriel des carrés magiques et leur relation avec des permutations spéciales.

1. Introduction

The *oldest magic square* $\begin{pmatrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{pmatrix}$ first appeared in ancient Chinese literature under the name **Lo Shu** two thousands years BC. The reader is likely to have encountered such objects, which following Ehrhart [2] are referred as *historical magic squares.* These are square matrices of order n whose entries are nonnegative integers $\{1, \cdots, n^2\}$ and whose rows and columns and the main diagonals sum up to the same number, which is called *the magic sum.* MacMahon [7] and Stanley [10] defined magic squares in modern combinatorics as square matrices of order n whose entries are nonnegative integers and whose rows and columns sum up to the same number, which is called *the line sum.* In this paper we will study the magic squares following the next definition.

Definition 1.1. A magic square is a square matrix of order n, whose entries are nonnegative integers and the sum of each row, each column and both the main diagonals adds up to the same number, which is called *the magic total.*

Example 1.2. $\begin{pmatrix} 3 & 6 & 0 \\ 0 & 3 & 6 \\ 6 & 0 & 3 \end{pmatrix}$ is a magic square of order 3 and whose magic total is equal to 9.

MacMahon[7] has already enumerated the number of all magic squares of order 3 in 1915, and it was in 2002 that Ahmed et al.[1] could find the number of magic squares of order 4 for a given magic total. The number of magic squares of order $n \geq 5$ with magic total $s \geq 2$ is a challenge ! We will introduce notions on magic permutations which are generators of all magic squares. In 1879, Hertzsprung [5] defined the number of magic permutations as well as the number of permutations without fixed points and without reflected points, well before the development of rook theory ([3],[4],[6],[8],[11]) as a method for enumeration of permutations with restricted positions and it was Riordan [8](1958) and Simpson [9](1995) who recalled these recurrence relations. We will use weighted rook polynomials to generalize the results on generalized restricted permutations and we will give an unexpected relation which relates derangements and restricted permutations. We will denote by MS_n the set (or vector space) of magic squares of order n.

2. Magic permutations

We will recall the following definitions :

Definition 2.1. A permutation σ of order n is a bijection over n objects.

We will denote by $[n]$ the set $\{1, \cdots, n\}$, and by $\mathfrak{S}_n$ the set of all permutations over $[n]$.

Definition 2.2. We say that an integer i is a fixed point for the permutation σ if $\sigma(i) = i$.

Definition 2.3. We say that an integer i is a reflected point for the permutation σ if $\sigma(i) = n - i + 1$.

Key words and phrases. Magic squares, magic total, magic permutations, rook polynomial.

The author was supported by the ISP during her stay at the Mittag-Leffler Institute, Sweden on March 2005.

We will denote by $Fix(\sigma)$ the set of the fixed points of the permutation σ, and by $Rfl(\sigma)$ the set of its reflected points.

Definition 2.4. We say that an integer i is a pivot point if i is a fixed reflected point.

Remark 2.5. If n is even, all permutation σ of length n does not have a pivot point.

Remark 2.6. The only pivot point of a permutation of length n is the integer $\dfrac{n+1}{2}$ if n is odd.

Example 2.7. For the permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 6 & 5 & 4 & 2 & 7 & 3 \end{pmatrix}$, we have

$$Fix(\sigma) = \{1,\, 4\} \text{ and } Rfl(\sigma) = \{2,\, 3,\, 4\}\,.$$

We will write a permutation σ of length n as a square matrix of order n such that the i-th column is presented by the vector column $e_{\sigma(i)}$ where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad e_{n-1} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix} \text{ and } e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Example 2.8. If we consider the permutation in Example 2.7, we have :

$$\sigma = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Remark 2.9. The reflected points and the fixed points of a permutation σ are shown in the matrix representation of the permutation σ as the occurrence of the integer 1 on the main diagonals.

Definition 2.10. A magic permutation is a permutation σ whose matrix representation is a magic square of magic total 1.

Proposition 2.11. *A permutation σ is magic if σ has one fixed point and one reflected point.*

Example 2.12. The following permutations σ_1 and σ_2 of length 9 are magic :

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 2 & 5 & 6 & 8 & 1 & 4 & 9 & 7 \end{pmatrix}$$

and

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 9 & 7 & 5 & 1 & 8 & 3 & 6 \end{pmatrix}$$

Proposition 2.13. *There does not exist a magic permutation of length n for $n = 2,\ 3$.*

Proposition 2.14. (1) *All magic squares of order 2 have the form $\left(\begin{smallmatrix} n & n \\ n & n \end{smallmatrix}\right)$, for $n \in \mathbb{N}$.*

(2) *For all magic squares $\left(\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & k \end{smallmatrix}\right)$ of magic total s, we have :*

$$\left(\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & k \end{smallmatrix}\right) = (g-f)\left(\begin{smallmatrix} 0 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{smallmatrix}\right) + \frac{h}{2}\left(\begin{smallmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{smallmatrix}\right) + \frac{f}{2}\left(\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{smallmatrix}\right),$$

with $2g - f + h = 2/3s$.

Proof. (1) It is easy to verify that if the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is magic square, then we have the system of equations $a+b = c+d = a+c = b+d = a+d = b+c$ which involves $a = b = c = d$.

(2) It is left to the reader to prove that these three matrices are independant i.e. if

$$\begin{pmatrix} 0&0&0\\0&0&0\\0&0&0 \end{pmatrix} = \alpha_1 \begin{pmatrix} 0&2&1\\2&1&0\\1&0&2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1&0&2\\2&1&0\\0&2&1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1&2&0\\0&1&2\\2&0&1 \end{pmatrix},$$

then $\alpha_1 = \alpha_2 = \alpha_3 = 0$. We leave as exercise to prove that for a given magic square $\begin{pmatrix} a&b&c\\d&e&f\\g&h&k \end{pmatrix}$, if we have

$$\begin{pmatrix} a&b&c\\d&e&f\\g&h&k \end{pmatrix} = \alpha_1 \begin{pmatrix} 0&2&1\\2&1&0\\1&0&2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1&0&2\\2&1&0\\0&2&1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1&2&0\\0&1&2\\2&0&1 \end{pmatrix},$$

then $\alpha_1 = g - f$, $\alpha_2 = \dfrac{h}{2}$, $\alpha_3 = \dfrac{f}{2}$.

□

Corollary 2.15. *$DimMS_2 = 1$ and $DimMS_3 = 3$.*

MacMahon [7] has established the following theorem :

Theorem 2.16. *The number $M_3(s)$ of magic squares of order 3 of magic total s is defined by :*

$$M_3(s) = \begin{cases} \frac{2}{9}s^2 + \frac{2}{3}s + 1 \textit{ if } 3 \textit{ divides } s, \\ 0 \textit{ otherwise.} \end{cases}$$

and their generating function has the closed form :

$$\sum_s M_3(s)t^s = \frac{(1+t^3)^2}{(1-t^3)^2}.$$

Proposition 2.17. *All magic squares of order 4 of magic total s can be written as linear combination of the following seven magic permutations as below :*

$$\begin{pmatrix} a&b&c&d\\e&f&g&h\\j&k&l&m\\n&p&q&r \end{pmatrix} = (a-m)\begin{pmatrix} 1&0&0&0\\0&0&0&1\\0&1&0&0\\0&0&1&0 \end{pmatrix} + m\begin{pmatrix} 1&0&0&0\\0&0&1&0\\0&0&0&1\\0&1&0&0 \end{pmatrix} + (d-f)\begin{pmatrix} 0&0&0&1\\1&0&0&0\\0&0&1&0\\0&1&0&0 \end{pmatrix} +$$
$$f\begin{pmatrix} 0&0&0&1\\0&1&0&0\\1&0&0&0\\0&0&1&0 \end{pmatrix} + (b-n)\begin{pmatrix} 0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&0&1 \end{pmatrix} + c\begin{pmatrix} 0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1 \end{pmatrix} + n\begin{pmatrix} 0&1&0&0\\0&0&0&1\\0&0&1&0\\1&0&0&0 \end{pmatrix},$$

with $a+b+c+d = s$, $0 \le a+n,\ b+f,\ d+m,\ d+n \le s$.

Proof. Similarily to the previous proof, it is left to the reader to prove that those seven magic permutations are independant and for a given magic square $\begin{pmatrix} a&b&c&d\\e&f&g&h\\j&k&l&m\\n&p&q&r \end{pmatrix}$, if we have

$$\begin{pmatrix} a&b&c&d\\e&f&g&h\\j&k&l&m\\n&p&q&r \end{pmatrix} = x_1\begin{pmatrix} 1&0&0&0\\0&0&0&1\\0&1&0&0\\0&0&1&0 \end{pmatrix} + x_2\begin{pmatrix} 1&0&0&0\\0&0&1&0\\0&0&0&1\\0&1&0&0 \end{pmatrix} + x_3\begin{pmatrix} 0&0&0&1\\1&0&0&0\\0&0&1&0\\0&1&0&0 \end{pmatrix} +$$
$$x_4\begin{pmatrix} 0&0&0&1\\0&1&0&0\\1&0&0&0\\0&0&1&0 \end{pmatrix} + x_5\begin{pmatrix} 0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&0&1 \end{pmatrix} + x_6\begin{pmatrix} 0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1 \end{pmatrix} + x_7\begin{pmatrix} 0&1&0&0\\0&0&0&1\\0&0&1&0\\1&0&0&0 \end{pmatrix}.,$$

then $x_1 = a - m$, $x_2 = m$, $x_3 = d - f$, $x_4 = f$, $x_5 = b - n$, $x_6 = c$, $x_7 = n$. □

Corollary 2.18. *$DimMS_4 = 7$.*

Ahmed et al. [1] have established the following theorem :

Theorem 2.19. *The generating function of the numbers $M_4(s)$ of magic squares of order 4 of magic total s is defined by :*

$$\sum_s M_4(s)t^s = \frac{t^8 + 4t^7 + 18t^6 + 36t^5 + 50t^4 + 36t^3 + 18t^2 + 4t + 1}{(1-t)^4(1-t^2)^4}.$$

Proposition 2.20. *For all integers $n \ge 4$, we have $DimMS_n = (n-1)^2 - 2$.*

Proof. Since the dimension of the Birkhoff polytope is equal to $(n-1)^2$, this is also the dimension of the vector space of magic squares with Stanley's definition. While adding two equalities for the diagonals, we have $(n-1)^2 - 2 \leq DimMS_n \leq (n-1)^2$. Now, let us consider a magic square $A = (a_{ij})_{1\leq i,j\leq n}$ of order $n \geq 4$ and of magic total s. We will give the minimal entries that are necessary to create the matrix A. If the first $n-2$ entries of the first $n-1$ columns are given, we deduce the first $n-2$ entries of the last column of the matrix A by the equations

$$\sum_{j=1}^{n} a_{ij} = s, \qquad \text{for all} \quad i = 1..n-2.$$

If we give the entry $a_{n-1\,1}$, we deduce $a_{n\,1}$, $a_{n-1\,2}$ and $a_{n\,2}$. When we give the entries $a_{n-1\,j}$ for $3 \leq j \leq n-3$, we deduce the entries $a_{n\,j}$ for $3 \leq j \leq n-3$, and when we give the entry $a_{n-1\,n-1}$, we deduce the remaining entries $a_{n\,n-1}$, $a_{n\,n}$, $a_{n-1\,n}$, , and $a_{n-1\,n-2}, a_{n\,n-2}$ of the matrix A. It is easy to verify that these given $(n-1)^2 - 2$ entries suffice to create the magic square A and the integer $(n-1)^2 - 2$ is also the dimension of the vector space MS_n. This gives the result. □

Proposition 2.21. *If a permutation σ is magic, then :*

(1) *σ^{-1} is magic,*
(2) *the reflected permutation σ' of the permutation σ, defined by $\sigma'(i) = n - \sigma(i) + 1$, is magic.*

Proof. Notice that a fixed point of the permutation σ remains a fixed point for σ^{-1} and becomes a reflected point for the reflected permutation σ' and vice-versa. Notice also that if the integer i is a reflected point for the permutation σ, then the integer i is a fixed point for the reflected permutation σ' and the integer $n-i+1$ is a reflected point for σ^{-1} and vice-versa. □

If we denote by a_n and x_n the number of magic permutations and the number of permutations without fixed points and without reflected points of length n respectively, we can find in the following table the first values of these numbers :

n	0	1	2	3	4	5	6	7	8
a_n	0	1	0	0	8	20	96	656	5568
x_n	1	0	0	0	4	16	80	672	4752

Hertzsprung [5] established the following theorem :

Theorem 2.22. *The numbers a_n and x_n satisfy the following reccurrences :*

$$\begin{array}{lllllll} a_{2n} & = & n(x_{2n} & - & (2n-3)x_{2n-1}) & & \\ a_{2n+1} & = & (2n+1)x_{2n} & + & 3nx_{2n-1} & - & 2n(n-1)x_{2n-2}, \\ x_n & = & (n-1)x_{n-1} & + & 2(n-d)x_{n-e}, & & \\ & & \multicolumn{3}{l}{\textit{where } (d,e) = (2,4) \textit{ if } n \textit{ is even,}} & \multicolumn{2}{l}{(1,2) \textit{ if } n \textit{ is odd.}} \end{array}$$

We will generalize the rook polynomial notions to enumerate some restricted permutations.

3. Rook polynomial

We will study in this section the number of permutations $\sigma \in \mathfrak{S}_n$ where for each i, certain values of $\sigma(i)$ are disallowed (namely, $\sigma(i) \neq i$ and $\sigma(i) \neq n-i+1$). We have a board $\mathfrak{B} \subset [n] \times [n]$. Each square s on $\mathfrak{B}$ has a weight ω_s. We define the rook numbers (actually polynomials) of $\mathfrak{B}$ by

$$r_k = \sum_{|A|=k} \prod_{s\in A} \omega_s$$

where the sum is over all subsets $A \subset \mathfrak{B}$ of cardinality k with no two squares on the same row or column. We define the generalized hit numbers h_i :

$$h_k = \sum_{\pi} \omega(\pi)$$

where the sum is over all permutations π of $[n]$ with k hits (values of i such that $(i, \pi(i)) \in \mathfrak{B}$) and the weight $\omega(\pi)$ of π is the product $\prod_{i=1}^{n} \omega_{(i,\pi(i))}$ where $\omega_{(i,\pi(i))}$ is the weight of $(i, \pi(i))$ if $(i, \pi(i)) \in \mathfrak{B}$ and is 1 otherwise. The generalized hit polynomial is

$$H = \sum_k h_k.$$

We can find a relation between H and the rook numbers r_i just as in the usual case. Claim that

$$\sum_k r_k (n-k)! = H^+ \qquad (\star)$$

where H^+ is the result of replacing each weight ω_s for $s \in \mathfrak{B}$ with $\omega_s + 1$. To see this, note that $r_k(n-k)!$ counts pairs (A, π) where A is a rook placement in $\mathfrak{B}$ of size k, and π extends A to a permutation of $[n]$. If we fix π and sum over all possible A, we are summing over all of the subsets of the hits of A and this gives H^+. If we replace each weight ω_s in $(\star)$ by $\omega_s - 1$ we get

$$H = \sum_k r_k^- (n-k)!$$

where r_k^- is the result of replacing each ω_s with $\omega_s - 1$.

Example 3.1. Let $n = 2m$ and let $\mathfrak{B}$ the following board with weights as indicated. (This is the case $n = 6$)

α					β
	α			β	
		α	β		
		β	α		
	β			α	
β					α

By permuting the rows and columns we get

α	β				
β	α				
		α	β		
		β	α		
				α	β
				β	α

We see that

$$\sum_k r_k X^k = \left[1 + (2\alpha + 2\beta)X + (\alpha^2 + \beta^2)X^2\right]^m,$$

so

$$\sum_k r_k^- X^k = \left[1 + (2\alpha + 2\beta - 4)X + ((\alpha-1)^2 + (\beta-1)^2)X^2\right]^m,$$

and therefore $H = \sum_k r_k^-(n-k)!$ with r_k^- as above. For example, to count permutations with no reflected points and no fixed points, we set $\alpha = \beta = 0$ to get

$$\sum_k r_k^- X^k = (1 - 4X + 2X^2)^m$$

and

$$x_{2m} = \sum_k r_k^- (2m-k)!.$$

To count permutations with no reflected points and one fixed point we set $\beta = 0$ and look at the coefficient of α. So we want the coefficient of α in

$$\left[1 + (2\alpha - 4)X + ((\alpha-1)^2 + 1)X^2\right]^m$$

which is easily computed to be

$$2mX(1-X)(1-4X+2X^2)^{m-1},$$

and to count permutations with two fixed points without reflected points, we look at the coefficient of α^2 which is easily computed to be

$$2m(m-1)X^2(1-X)^2(1-4X+2X^2)^{m-2}+mX^2(1-4X+2X^2)^{m-1}.$$

To count permutations with one reflected point and one fixed point we look at the coefficient of $\alpha\beta$, which is

$$4m(m-1)X^2(1-X)^2(1-4X+2X^2)^{m-2}$$

and

$$a_{2m}=\sum_k r_k^-(2m-k)!.$$

and so on. For n odd we take the following board where we have a separate weight for the middle square

α				β
	α		β	
		γ		
	β		α	
β				α

By permuting the rows and columns we get

α	β			
β	α			
		α	β	
		β	α	
				γ

We see that

$$\sum_k r_k X^k=(1+\gamma X)\left[1+(2\alpha+2\beta)X+(\alpha^2+\beta^2)X^2\right]^m,$$

so

$$\sum_k r_k^- X^k=(1+(\gamma-1)X)\left[1+(2\alpha+2\beta-4)X+((\alpha-1)^2+(\beta-1)^2)X^2\right]^m,$$

and therefore $H=\sum_k r_k^-(n-k)!$ with r_k^- as above and $n=2m+1$. For example, to count permutations with no reflected points and no fixed points, we set $\alpha=\beta=\gamma=0$ to get

$$\sum_k r_k^- X^k=(1-X)(1-4X+2X^2)^m$$

and

$$x_{2m+1}=\sum_k r_k^-(2m+1-k)!.$$

To count permutations with no reflected points and one fixed point we set $\beta=\gamma=0$ and look at the coefficient of α. So we want the coefficient of α in

$$(1-X)\left[1+(2\alpha-4)X+((\alpha-1)^2+1)X^2\right]^m$$

which is easily computed to be

$$2mX(1-X)(1-4X+2X^2)^{m-1},$$

and to count permutations with no reflected points and two fixed points, we look at the coefficient of α^2 which is easily computed to be

$$2m(m-1)X^2(1-X)^3(1-4X+2X^2)^{m-2}+mX^2(1-X)(1-4X+2X^2)^{m-1}.$$

To count permutations with a pivot point we set $\alpha=\beta=0$ and look at the coefficient of γ, which is

$$X(1-4X+2X^2)^m.$$

We deduce that the number x_{2m} enumerates also permutations of order $2m+1$ having a pivot point. To count permutations with one fixed point and one reflected point without pivot points, we set $\gamma = 0$ and look at the coefficient of $\alpha\beta$ in

$$(1-X)\left[1+(2\alpha+2\beta-4)X+((\alpha-1)^2+(\beta-1)^2)X^2\right]^m$$

to get

$$4m(m-1)X^2(1-X)^3(1-4X+2X^2)^{m-2}$$

and

$$a_{2m+1} = \sum_k r_k^-(2m+1-k)! + x_{2m}.$$

To count derangements, that is permutations without fixed points, we set $\alpha = \gamma = 0$ and $\beta = 1$ to get

$$(1-X)^n$$

and so on.

Definition 3.2. We say that a subset F of the set $[n]$ is :

(1) semi-reflected if there exists at least one element $i \in F$ such that $n-i+1 \in F$.
(2) self-reflected if $i \in F$ and $n-i+1 \in F$, for all elements i in the subset F.

The proof of the following lemmas is a simple exercise of combinatorics.

Lemma 3.3. *For disjoint subsets F and R of the set $[2n]$ such that $\#F \cup R = 2k$ and $F \cup R$ is self-reflected, the number of pair (F,R) is equal to $2^k\binom{n}{k}$.*

Lemma 3.4. *For disjoint subsets F and R of the set $[2n]$ or $[2n+1]$ such that $\#F \cup R = n$ and $F \cup R$ is not semi-reflected, the number of pair (F,R) is equal to 2^{2n}.*

Theorem 3.5. *The number of permutations of length $2n$ having set of fixed points and reflected points of cardinality $2k$, and which is a self-reflected set, is equal to $\binom{n}{k}2^k x_{2(n-k)}$.*

Proof. We consider the following board with weights as indicated. We illustrate it with the case for $2n = 6$.

α_1					β_1
	α_2			β_2	
		α_3	β_3		
		β_4	α_4		
	β_5			α_5	
β_6					α_6

By permuting the rows and columns we get

α_1	β_1				
β_6	α_6				
		α_2	β_2		
		β_5	α_5		
				α_3	β_3
				β_4	α_4

We see that

$$\sum_k r_k X^k = \prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1})X+(\alpha_i\alpha_{2n-i+1}+\beta_i\beta_{2n-i+1})X^2],$$

so

$$\sum_k r_k^- X^k = \prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1}-4)X+((\alpha_i-1)(\alpha_{2n-i+1}-1)+(\beta_i-1)(\beta_{2n-i+1}-1))X^2],$$

and therefore $H = \sum_k r_k^-(n-k)!$ with r_k^- as above. To count permutations of length $2n$ having set of fixed points and reflected points of cardinality $2k$, and which is a self-reflected set, we look first at the coefficient of $\prod_{s=1}^{p} \mu_{i_s} \prod_{s=p+1}^{k} \nu_{i_s}$ where $\mu_{i_s} = \alpha_{i_s}\alpha_{2n-i_s+1}$ and $\nu_{i_s} = \beta_{i_s}\beta_{2n-i_s+1}$ which is easily computed to be

$$2^k X^{2k}(1-4X+2X^2)^{(n-k)},$$

and we will consider all products of the form $\prod_{s=1}^{p} \mu_{i_s} \prod_{s=p+1}^{k} \nu_{i_s}$ whose number is equal to $\binom{n}{k}$, and this gives the result. □

Theorem 3.6. *The number of permutations of length $2n+1$ having set of fixed points and reflected points of cardinality $2k+1$, and which is a self-reflected set, is equal to $\binom{n}{k}2^k x_{2(n-k)}$.*

Proof. Notice that if the cardinality of a self-reflected subset of the set $[2n+1]$ is odd, this subset contains the integer $n+1$. We consider the following board with weights as indicated. We illustrate it with the case for $2n+1 = 5$.

α_1				β_1
	α_2		β_2	
		γ		
	β_4		α_4	
β_5				α_5

By permuting the rows and columns we get

α_1	β_1			
β_5	α_5			
		α_2	β_2	
		β_4	α_4	
				γ

We see that

$$\sum_k r_k X^k = (1+\gamma X)\prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1})X+(\alpha_i\alpha_{2n-i+1}+\beta_i\beta_{2n-i+1})X^2],$$

so

$$\begin{aligned}\textstyle\sum_k r_k^- X^k = \quad & (1+(\gamma-1)X)\textstyle\prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1}-4)X \\ & +((\alpha_i-1)(\alpha_{2n-i+1}-1)+(\beta_i-1)(\beta_{2n-i+1}-1))X^2],\end{aligned}$$

and therefore $H = \sum_k r_k^-(n-k)!$ with r_k^- as above. To count permutations of length $2n+1$ having set of fixed points and reflected points of cardinality $2k$, and which is self-reflected, we set $\gamma = 0$ and we look at the coefficient of $\prod_{s=1}^{p} \mu_{i_s} \prod_{s=p+1}^{k} \nu_{i_s}$ where $\mu_{i_s} = \alpha_{i_s}\alpha_{2n-i_s+1}$ and $\nu_{i_s} = \beta_{i_s}\beta_{2n-i_s+1}$ which is easily computed to be

$$2^k X^{2k}(1-X)(1-4X+2X^2)^{(n-k)},$$

and we will consider all the product of the form $\prod_{s=1}^{p} \mu_{i_s} \prod_{s=p+1}^{k} \nu_{i_s}$ whose number is equal to $\binom{n}{k}$, and this gives the result. □

4. Derangements

We will conclude this paper with an unexpected relation which relates derangements and restricted permutations.

Theorem 4.1. *The number of permutations of length $2n$ having set of fixed points and reflected points of cardinality n, and which is not a semi-reflected set, is equal to $2^{2n}d_n$.*

Proof. We consider again a board as in the proof of Theorem 3.6. To count permutations of length $2n$ having set of fixed points and reflected points of cardinality n, and which is not a semi-reflected set, we look first at the coefficient of $\prod_{s=1}^{p}\alpha_{i_s}\prod_{s=p+1}^{n}\beta_{i_s}$ such that if $\ell\neq m$, then $i_\ell\neq 2n-i_m+1$ in

$$\prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1}-4)X+((\alpha_i-1)(\alpha_{2n-i+1}-1)+(\beta_i-1)(\beta_{2n-i+1}-1))X^2]$$

which is easily computed to be

$$X^n(1-X)^n.$$

If we consider the coefficient of all such products whose number is equal to 2^{2n} by Lemma 3.3, we obtain the result. □

Theorem 4.2. *The number of permutations of length $2n+1$ having set of fixed points and reflected points of cardinality n, and which is not a semi-reflected set, is equal to $2^{2n}d_{n+1}$.*

Proof. We consider again a board as in the proof of Theorem 3.6. To count permutations of length $2n+1$ having set of fixed points and reflected points of cardinality n, and which is not a semi-reflected set, we set $\gamma=0$ and we look first at the coefficient of $\prod_{s=1}^{p}\alpha_{i_s}\prod_{s=p+1}^{n}\beta_{i_s}$ such that if $\ell\neq m$, then $i_\ell\neq 2n-i_m+1$ in

$$(1-X)\prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1}-4)X+((\alpha_i-1)(\alpha_{2n-i+1}-1)+(\beta_i-1)(\beta_{2n-i+1}-1))X^2]$$

which is easily computed to be

$$X^n(1-X)^{n+1}.$$

If we consider the coefficient of all such products whose number is equal to 2^{2n} by Lemma 3.4, we deduce the result. □

Theorem 4.3. *The number of permutations of length $2n+1$ having set of fixed points and reflected points of cardinality $n+1$ containing $n+1$, and which is not a semi-reflected set if the element $n+1$ is deleted, is equal to $2^{2n}d_n$.*

Proof. We consider again a board as in the proof of Theorem 3.6. To count permutations of length $2n+1$ having set of fixed points and reflected points of cardinality $n+1$, and which is not a semi-reflected set if the integer $n+1$ is deleted, we look first at the coefficient of $\gamma\prod_{s=1}^{p}\alpha_{i_s}\prod_{s=p+1}^{n}\beta_{i_s}$ such that if $\ell\neq m$, then $i_\ell\neq 2n-i_m+1$ in

$$\prod_{i=1}^{n}[1+(\alpha_i+\alpha_{2n-i+1}+\beta_i+\beta_{2n-i+1}-4)X+((\alpha_i-1)(\alpha_{2n-i+1}-1)+(\beta_i-1)(\beta_{2n-i+1}-1))X^2]$$

which is easily computed to be

$$X^{n+1}(1-X)^n.$$

If we add the coefficients of all such products whose number is equal to 2^{2n} by Lemma 3.4, we deduce the result. □

Acknowledgements. The author is very grateful to Ira Gessel for his kind advices during the author's stay at the Institut Mittag-Leffler, Sweden. Thanks also to the two referees for their comments and corrections to the original version of this paper.

References

[1] M. Ahmed, J. De Loera, R. Hemmecke, *Polyhedral Cones of Magic Cubes and Squares*, Algorithm combinat. 25(2003), 25-41 arXiv:math.CO/0201108, 2002

[2] E. Ehrhart, *Sur les carrés magiques*, C. R. Acad.Sci. Paris, **227 A**, 1973, pp. 575-577.

[3] D. C. Fielder, *A Generator of Rook Polynomials.* Mathematica J. **9**, 2004, pp. 371-375, .

[4] I. Gessel, *Generalised rook polynomials and orthogonal polynomials*, in q-Series and Partitions, ed. D. Stanton, The IMA Volumes in Mathematics and Its Applications, Vol.**18**, Springer-Verlag, New York, 1989, pp. 159-176.

[5] S. Hertzsprung, *L∅sning og Udvidelse af Opgave 402*, Tidsskrift for Math. **4**, 1879, pp. 134-140.

[6] C. L. Liu, *Introduction to Combinatorial Mathematics.* New York: McGraw-Hill, 1968.

[7] P. A. MacMahon, *Combinatory Analysis*, Cambridge University Press, 1915.
[8] J. Riordan, *An Introduction to Combinatorial Analysis*, John Wiley & Sons, New York, 1958.
[9] T. Simpson, *Permutations with unique fixed and reflected points.* Ars Combin. **39**, 97-108, 1995.
[10] R. P. Stanley, *Enumerative Combinatorics,* Vol **1**, Cambridge, 1997.
[11] A. Tucker, *Applied Combinatorics*, John Wiley & Sons, New York, 2002.

Département de Mathématiques et Informatique, Université d'Antananarivo, 101, Antananarivo, Madagascar
E-mail address: frakoton@univ-antananarivo.mg

A Generalization of the Cayley-Hamilton Theorem

Brendon Rhoades
University of Michigan
Ann Arbor, MI, USA
brhoades@umich.edu

April 28, 2005

Abstract

We present a generalization of the Cayley-Hamilton Theorem using a collection of immanants which naturally generalize the determinant.

For $f : S_n \to \mathbb{C}$, and $x = (x_{ij})_{1 \le i,j \le n}$, define the *f-immanant*, $Imm_f(x)$, by

$Imm_f(x) = \sum_{\sigma \in S_n} f(\sigma) x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$.

Given $\xi \in \mathbb{C}$ and a positive integer n, define the to be the n^{th} *Temperly-Lieb algebra* $T_n(\xi)$ to be the multiplicative, associative $\mathbb{C}$-algebra with unity 1 generated by $t_1, t_2, \ldots, t_{(n-1)}$ subject to the relations

$$\begin{cases} t_i t_i = \xi t_i, i \in [n-1] \\ t_i t_j t_i = t_i, |i-j| = 1 \\ t_i t_j = t_j t_i, |i-j| > 1. \end{cases}$$

It is well known that the multiplicative monoid generated by $t_1, t_2, \ldots t_{(n-1)}$ is a basis for $T_n(\xi)$. We refer this basis as *the standard basis* and to its elements as *the basis elements of* $T_n(\xi)$. Given basis elements $\tau_1 = t_{i_1} \ldots t_{i_k}$ of $T_r(\xi)$ and $\tau_2 = t_{j_1} \ldots t_{j_l}$ of $T_s(\xi)$, define $\tau_1 \oplus \tau_2$ to be the basis element of $T_{r+s}(\xi)$ given by $\tau_1 \oplus \tau_2 = t_{i_1} \ldots t_{i_k} t_{j_1+r} \ldots t_{j_l+r}$

Let for $i \in [n-1]$, let s_i denote the element of the symmetric group S_n written $(i, i+1)$ in cycle notation. Define a map $\theta : S_n \to T_n(2)$ by mapping s_i into $(t_i - 1)$ for every $i \in [n-1]$. It is easy to check that this induces a well defined homomorphism from S_n into the multiplicative monoid of $T_n(2)$. (see, for example, [1]) For every basis element τ of $T_n(2)$, define a map $f_\tau : S_n \to \mathbb{C}$ by sending σ to the coefficient of τ in the expansion of $\theta(\sigma)$ in the standard basis. The immanants $Imm_{f_\tau}(x)$ induced by the functions f_τ are called the *Temperly-Lieb immanants*.

In [3] and [4] Rhoades and Skandera show that the Temperly-Lieb immanants are totally, monomial, and Schur nonnegative and may be used to study positivity properties of linear combinations of products of matrix minors. In [2], Lam, Postnikov, and Pylavyskyy use these results to resolve several Schur positivity conjectures. In [3], Rhoades and Skandera also give some generalizations of results from linear algebra using these immanants. In this spirit, we give here a generalization of the Cayley-Hamilton theorem.

Let V be a n-dimensional vector space and let $T \in End(V)$. For any ordered basis γ of V and any basis element τ of $T_n(2)$, define the (τ, γ)-*polynomial* to be the polynomial $g_{(\tau,\gamma)}(X) \in \mathbb{C}[X]$ given by

$$g_{\tau,\gamma}(X) = Imm_{f_\tau}(I_n X - [T]_\gamma),$$

where I_n is the $n \times n$ identity matrix. Let β be a rational canonical basis for V with invariant factor degrees $d_1 \leq d_2 \leq \ldots \leq d_k$. For $j \in [k]$, call an ordered basis γ of V (β, j)-*respecting* if the matrix $[T]_\gamma$ is of the form $diag(A, B, C)$, where A is a $d_1 + \ldots + d_{(j-1)}$ matrix and $B = [T|span(\beta_j)]_{\beta_j}$, where β_j is the subset of β corresponding to the j^{th} invariant factor.

We are now ready to state our result.

Theorem. *Let V be a finite dimensional $\mathbb{C}$-vector space and let $T \in End(V)$. Let β be a rational canonical basis for V and let the invariant factor degrees of T be $d_1 \leq d_2 \leq \ldots \leq d_k$. Let $j \in [k]$ and let γ be a $(\beta$, j$)$-respecting ordered basis of V. Set $s = d_1 + \cdots + d_{(j-1)}$ and $r = d_{(j+1)} + \cdots + d_k$. If τ_1 and τ_2 are basis elements of $T_s(2)$ and $T_r(2)$, respectively, then*

$$rank(g_{\tau_1 \oplus 1 \oplus \tau_2, \gamma}(T)) \leq r.$$

Proof. Define a $\mathbb{C}[X]$-module structure on V by linearly extending the action

$$X \cdot v = T(v) \text{ for all } v \in V.$$

This makes V into a module over a Principal Ideal Domain. Since V is finite dimensional, we have the following isomorphism of $\mathbb{C}[X]$-modules:

$$V \cong \bigoplus_{i=1}^{k} \mathbb{C}[X]/(p_i(X)),$$

where $p_1|p_2|\ldots|p_k$. Recall that the polynomials $p_1(X), \ldots, p_k(X)$ are the invariant factors of T. We now compute the $(\tau_1 \oplus\ 1 \oplus \tau_2, \gamma)$-polynomial of T.

Since γ is $(\beta$, j$)$-respecting, the matrix $[T]_\gamma$ has the form $diag(A, B, C)$, were A is a square matrix of size s and B is the restriction of $[T]_\beta$ to its j^{th} diagonal block. By Proposition 3.15 of Rhoades and Skandera [3], we have that

$$g_{\tau_1 \oplus\ 1 \oplus \tau_2, \gamma}(X) =$$
$$Imm_{f_{\tau_1 \oplus 1 \oplus \tau_2}}(IX - [T]_\gamma) =$$
$$Imm_{f_{\tau_1}}(IX - A) Imm_{f_1}(IX - B) Imm_{f_{\tau_2}}(IX - C).$$

It is easy to check that for $\sigma \in S_n$, $f_1(\sigma) = (-1)^{inv(\sigma)}$. That is, $Imm_{f_1}(x) = det(x)$. So, the center factor in the above product is equal to $p_j(X)$, which implies that $g_{\tau_1 \oplus\ 1 \oplus \tau_2, \gamma}(X)$ lies in the ideal $(p_j(X))$. However, since we have the chain of divisibilities, $p_1|p_2|\ldots|p_k$, $g_{\tau_1 \oplus\ 1 \oplus \tau_2, \gamma}(X)$ also lies in $(p_i(X))$ for every $i < k$. This implies that $g_{\tau_1 \oplus\ 1 \oplus \tau_2, \gamma}(T)$ kills every vector in the subspaces of V corresponding to $\mathbb{C}[X]/(p_i(X))$ for every $i \leq k$. The desired inequality follows.

Aknowledgements

The author is grateful to Mark Skandera for many helpful conversations.

References

1. J. Humphreys. *Reflection Groups and Coxeter Groups*. Cambridge University Press, 1990.

2. Lam, Postnikov, Pylavyskyy. *Schur Positivity Conjectures* : 2 1/2 *Are No More*!. Preprint available on ArXiV: `http://arxiv.org/pdf/math.CO/0502446`.

3. B. Rhoades, M. Skandera. *Temperly-Lieb Immanants*. To appear, *Ann. Comb*. Preprint available at `http://www.math.dartmouth.edu/~skan/papers.htm`.

4. B. Rhoades, M. Skandera. *Kazhdan-Lusztig Immanants*. Preprint available at `http://www.math.dartmouth.edu/~skan/papers.htm`.

On the degree-adjacency matrix of a graph

J. A. Rodríguez and J. M. Sigarreta

Abstract. The aim of this paper is to study some parameters of simple graphs related with the degree of the vertices. So, our main tool is the $n \times n$ matrix $\mathcal{A}$ whose (i,j)-entry is

$$a_{ij} = \begin{cases} \frac{1}{\sqrt{\delta_i \delta_j}} & \text{if} \quad v_i \sim v_j; \\ 0 & \text{otherwise,} \end{cases}$$

where δ_i denotes the degree of the vertex v_i. We study the Randić index and some interesting particular cases of conditional excess, conditional Wiener index, and conditional diameter. In particular, using the matrix $\mathcal{A}$ or its eigenvalues, we obtain tight bounds on the studied parameters.

1. Introduction

In order to deduce properties of graphs from results and methods of algebra, firstly we need to translate properties of graphs into algebraic properties. In this sense, a natural way is to consider algebraic structures or algebraic objects as, for instance, groups or matrices. In particular, the use of matrices allows us to use methods of linear algebra to derive properties of graphs. There are various matrices that are naturally associated with graphs, such as the adjacency matrix, the Laplacian matrix, and the incidence matrix [**1, 3, 8**]. One of the main aims of algebraic graph theory is to determine how, or whether, properties of graphs are reflected in the algebraic properties of such matrices [**8**]. The aim of this paper is to study the Randić index and some interesting particular cases of conditional excess, conditional Wiener index, and conditional diameter. All these parameters are related with the degree of the vertices of the graph. So, our main tool will be a suitable adjacency matrix that we call *degree-adjacency matrix*.

The plan of the paper is the following: in Section 2 we emphasize some of the main properties of the degree-adjacency matrix. The remaining sections are devoted to study the relationship between the degree-adjacency matrix (or its eigenvalues) and several parameters of graphs. More precisely, in Section 3 we obtain bounds on the Randić index, in Section 4 we obtain bounds on a particular case of conditional

---

2000 *Mathematics Subject Classification.* Primary 05C50, 05C12, 15A18.

Key words and phrases. Adjacency matrix, Conditional diameter, Excess, Randić index, Wiener index, Graph eigenvalues, Alternating polynomials.

The final version of this paper has been published in *MATCH Communications in Mathematical and in Computer Chemistry* **54** (2) 2005, with the title "On the Randić index and conditional parameters of a graph".

excess, Section 5 is devoted to bound the degree diameter and, finally, in section 6 we obtain bounds on a particular case of conditional Wiener index.

We begin by stating some notation. In this paper all graphs $\Gamma = (V, E)$ will be finite, undirected and simple. We will assume that $|V| = n$ and $|E| = m$. The distance between vertices $u, v \in V(\Gamma)$ will be denoted by $\partial(u, v)$. The degree of a vertex $v_i \in V(\Gamma)$ will be denoted by $\delta(v_i)$ (or by δ_i for short), the minimum degree of Γ will be denoted by δ and the maximum by Δ.

2. Degree-adjacency matrix

We define the *degree-adjacency matrix* of a graph Γ of order n as the $n \times n$ matrix $\mathcal{A}$ whose (i, j)-entry is

$$a_{ij} = \begin{cases} \frac{1}{\sqrt{\delta_i \delta_j}} & \text{if} \quad v_i \sim v_j; \\ 0 & \text{otherwise.} \end{cases}$$

The matrix $\mathcal{A}$ can be regarded as the adjacency matrix of a weighted graph in which the edge-weight $R(v_i v_j)$ of the edge $v_i v_j$ is equal to $R(v_i v_j) = \frac{1}{\sqrt{\delta_i \delta_j}}$, thus justifying the terminology used. The weight $R(v_i v_j)$ will be called the *Randić weight* of the edge $v_i v_j \in E$. We will say that a graph is *weight-regular* if each of its edges has the same Randić weight. Particular cases of weight-regular graphs are the class of regular graphs and the class of semi-regular bipartite graphs.

If we consider the vector $\nu = (\sqrt{\delta_1}, \sqrt{\delta_2}, ..., \sqrt{\delta_n})$, then we have $\mathcal{A}\nu = \nu$. Thus, $\lambda = 1$ is an eigenvalue of $\mathcal{A}$ and ν is an eigenvector associated to λ. Hence, as $\mathcal{A}$ is non-negative and irreducible in the case of connected graphs, by the Perron-Frobenius theorem, $\lambda = 1$ is a simple eigenvalue and $\lambda = 1 \geq |\lambda_j|$ for every eigenvalue λ_j of $\mathcal{A}$. Therefore, we have

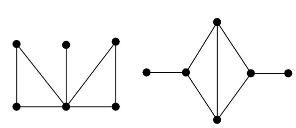
$$\|\mathcal{A}x\| \leq \|x\|, \quad \forall x \in \mathbb{R}^n. \tag{1}$$

Notice that the above inequality holds also in the case of non-connected graphs.

Hereafter the eigenvalues of $\mathcal{A}$ will be called *degree-adjacency eigenvalues of* Γ.

It is well-known that there are non-isomorphic graphs that have the same standard adjacency eigenvalues with the same multiplicities (the so called cospectral graphs). For instance, two connected graphs, both having the characteristic polynomial $P(x) = x^6 - 7x^4 - 4x^3 + 7x^2 + 4x - 1$, are shown in Figure 1. Therefore,

FIGURE 1. Two cospectral graphs but not cospectral with regard to $\mathcal{A}$



we can try to study cospectral graphs by using an alternative matrix, for instance, the degree-adjacency matrix $\mathcal{A}$. If we consider the matrix $\mathcal{A}$, the eigenvalues of both graphs are different: the left hand side graph has degree-adjacency eigenvalues 1, $\pm\frac{1}{2}$ and $-\frac{1}{4}\left(1 \pm \sqrt{2.6}\right)$ (where the eigenvalue $-\frac{1}{2}$ has multiplicity 2), on the other hand, the right hand side graph has degree-adjacency eigenvalues 1, $\frac{-1\pm\sqrt{2}}{3}$, $\pm\frac{\sqrt{3}}{3}$ and $-\frac{1}{3}$. Even so, the degree-adjacency eigenvalues do not determine the graph. That is, there are non-isomorphic graphs (and non-cospectral) that are cospectral with regard to the degree-adjacency matrix. For instance, the degree-adjacency eigenvalues of the cycle graph C_4 and the semi-regular bipartite graph $K_{1,3}$ are the same: $1, 0, 0, -1$. However, the standard eigenvalues are $2, 0, 0, -2$, in the case of C_4, and $\sqrt{3}, 0, 0, -\sqrt{3}$ in the case of $K_{1,3}$.

It is easy to see that there are some classes of graphs in which the standard eigenvalues, $\vartheta_1 \geq \vartheta_2 \geq \cdots \geq \vartheta_n$, and the degree-adjacency eigenvalues, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, are directly related. For instance, in the case of weight-regular graphs, of weight w^{-1}, the adjacency matrix, $\mathbf{A}$, and the degree-adjacency matrix are related by $\mathcal{A}(\Gamma) = \frac{1}{w}\mathbf{A}(\Gamma)$. Thus, the eigenvalues of both matrices are related by

$$\lambda_l = \frac{\vartheta_l}{w}, \quad l \in \{1, ..., n\}. \tag{2}$$

As in the case of the adjacency matrix, there are some classes of graphs in which we can deduce a formula to compute the characteristic polynomial, Ψ, of the degree-adjacency matrix. For instance, from the degree-adjacency matrix of the path graph, $\Gamma = P_n$, we deduce that

$$\Psi(\mathcal{A}(P_n), \lambda) = \lambda\Phi_n(\lambda) + \frac{1}{2}\Phi_{n-1}(\lambda), \quad n \geq 3,$$

where

$$\Phi_n(\lambda) = -\lambda\Phi_{n-1}(\lambda) - \frac{1}{4}\Phi_{n-2}(\lambda), \quad \Phi_2(\lambda) = -\lambda \quad \text{and} \quad \Phi_3(\lambda) = \lambda^2 - \frac{1}{2}.$$

Hereafter, in the general case of an arbitrary graph, we will consider that the characteristic polynomial, $\Psi(\mathcal{A}, \lambda) = det(\lambda\mathbf{I} - \mathcal{A})$, is of the form

$$\Psi(\mathcal{A}, \lambda) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n.$$

We can compute the first coefficients of Ψ by using a well-known result of theory of matrices: all the coefficients can be expressed in terms of the principal minors of $\mathcal{A}$.

Proposition 1. *Let Γ be a graph. The coefficients of the characteristic polynomial of $\mathcal{A} = \mathcal{A}(\Gamma)$ satisfy:*

$$c_1 = 0; \tag{3}$$

$$-c_2 = \frac{1}{2}\sum_{l=1}^{n}\lambda_l^2 = \sum_{i\sim j}\frac{1}{\delta_i\delta_j}; \tag{4}$$

$$c_3 = \sum_{\langle i,j,k\rangle \simeq K_3}\frac{-2}{\delta_i\delta_j\delta_k}, \tag{5}$$

where $\langle i, j, k\rangle \simeq K_3$ runs over all subgraphs of Γ induced by $\{v_i, v_j, v_k\}$ and isomorphic to K_3.

Proof. For each $r = 1, 2, ..., n$, the number $(-1)^r c_r$ is the sum of those principal minors of $\mathcal{A}$ which have order r. Thus, we derive the result as follows. Since

$\mathcal{A}$ has diagonal entries all zero, $c_1 = 0$. A principal non-null minor of order 2 must be of the form

$$\begin{vmatrix} 0 & \frac{1}{\sqrt{\delta_i \delta_j}} \\ \frac{1}{\sqrt{\delta_i \delta_j}} & 0 \end{vmatrix} = \frac{-1}{\delta_i \delta_j}.$$

There is one such minor for each edge of Γ. Moreover, since the trace of a square matrix is also equal to the sum of its eigenvalues, we have

$$\sum_{l=1}^{n} \lambda_l^2 = tr\left(\mathcal{A}^2\right) = 2 \sum_{i \sim j} \frac{1}{\delta_i \delta_j}.$$

Thus, (4) follows. On the other hand, the only non-null principal minor of order 3 is

$$\begin{vmatrix} 0 & \frac{1}{\sqrt{\delta_i \delta_j}} & \frac{1}{\sqrt{\delta_i \delta_k}} \\ \frac{1}{\sqrt{\delta_j \delta_i}} & 0 & \frac{1}{\sqrt{\delta_j \delta_k}} \\ \frac{1}{\sqrt{\delta_k \delta_i}} & \frac{1}{\sqrt{\delta_k \delta_j}} & 0 \end{vmatrix} = \frac{2}{\delta_i \delta_j \delta_k}.$$

There is one such minor for each triangle of Γ. Hence, (5) follows. □

Notice that the coefficient c_2 is immediately bounded from (4):

$$\frac{m}{\Delta^2} \leq -c_2 \leq \frac{m}{\delta^2}. \tag{6}$$

Corollary 2. *A graph is regular if, and only if, its order is* $2|c_2|\Delta$.

In Section 3 we will show the relationship between c_2 and the generalized Randić index.

We remark that the spectrum of $\mathcal{A}$ can be computed directly from the adjacency matrix $\mathbf{A}$ and the degree sequence. That is,

$$det(\mathcal{A} - \lambda \mathbf{I}) \cdot \prod_{j=1}^{n} \delta_j = det(\mathbf{A} - \lambda \mathbf{D}), \tag{7}$$

where $\mathbf{D} = diag(\delta_1, \delta_2, ..., \delta_n)$ is the diagonal matrix whose diagonal entries are the degrees of the vertices of Γ.

There are other properties of the degree-adjacency matrix that have been obtained previously (see, for instance [**3**]), in the following theorem we cite some of them.

Theorem 3. [**3**].

- *The number of connected components of* Γ *is equal to the multiplicity of the eigenvalue* 1 *of* $\mathcal{A}$.
- *Let* Γ *be a graph without isolated vertices.* Γ *is bipartite if and only if* $\Psi(\Gamma, \lambda) = \Psi(\Gamma, -\lambda)$.
- *Let* Γ *be a connected graph.* Γ *is bipartite if and only if* -1 *is an eigenvalue of* $\mathcal{A}$.

We identify the degree-adjacency matrix $\mathcal{A}$ with an endomorphism of the "vertex-space" of Γ, $l^2(V(\Gamma))$ which, for any given indexing of the vertices, is isomorphic to $\mathbb{R}^n$. Thus, for any vertex $v_i \in V(\Gamma)$, e_i will denote the corresponding unit vector of the canonical base of $\mathbb{R}^n$.

If for two vertices $v_i, v_j \in V(\Gamma)$ we have $\partial(v_i, v_j) > k$ then $(\mathcal{A}^k(\Gamma))_{ij} = 0$. Thus, for a real polynomial P of degree k, we have

$$\partial(v_i, v_j) > k \Rightarrow P(\mathcal{A}(\Gamma))_{ij} = 0. \tag{8}$$

Through this fact we will study some metric parameters of graphs.

3. Randić index

The *Randić index*, $R(\Gamma)$, of a graph Γ was introduced by the chemist Milan Randić in 1975 [**10**] as

$$R(\Gamma) = \sum_{v_i \sim v_j} \frac{1}{\sqrt{\delta_i \delta_j}}.$$

This topological index, sometimes called *connectivity index,* has been successfully related to physical and chemical properties of organic molecules and became one of the most popular molecular descriptors.

The Randić index has the following trivial bounds:

$$\frac{m}{\Delta} \leq R(\Gamma) \leq \frac{m}{\delta}. \tag{9}$$

Equality holds if, and only if, Γ is regular. Moreover, there are non-trivial bounds as the following [**2**]:

$$\sqrt{n-1} \leq R(\Gamma) \leq \frac{n}{2}. \tag{10}$$

Equality on the right-hand side holds if, and only if, Γ is a graph whose all components are regular of (not necessarily equal) degrees greater than zero. Equality on the left-hand side holds if, and only if, Γ is a star [**2**].

We emphasize that the degree-adjacency matrix allows us to obtain a short proof of the right hand side of (10): by the Cauchy-Schwarz inequality and (1) we have

$$2R(\Gamma) = \langle \mathcal{A}\mathbf{j}, \mathbf{j} \rangle \leq \|\mathcal{A}\mathbf{j}\| \|\mathbf{j}\| \leq \|\mathbf{j}\|^2 = n,$$

where $\mathbf{j} = (1, 1, ..., 1) \in \mathbb{R}^n$.

The *zeroth-order* Randić index is defined as

$$R_0(\Gamma) = \sum_{v \in V(\Gamma)} \frac{1}{\sqrt{\delta(v)}}.$$

Trivially, $R_0(\Gamma)$ is bounded by

$$\frac{n}{\sqrt{\Delta}} \leq R_0(\Gamma) \leq \frac{n}{\sqrt{\delta}}. \tag{11}$$

The equality holds if, and only if, Γ is regular of degree greater than zero.

The Randić index has been generalized [**9**] as

$$R_\alpha(\Gamma) = \sum_{v_i \sim v_j} (\delta_i \delta_j)^\alpha, \quad \alpha \neq 0.$$

Obviously, the standard Randić index is obtained when $\alpha = -\frac{1}{2}$.

In the chemical literature the quantity

$$R_1(\Gamma) = \sum_{v_i \sim v_j} \delta_i \delta_j$$

is called *the second Zagreb index* [**4**]. The second Zagreb index was bounded in [**2**] by

$$R_1(\Gamma) \leq m\left(\frac{\sqrt{8m+1}-1}{2}\right)^2. \tag{12}$$

Moreover, by (4) we have

$$R_{-1}(\Gamma) = |c_2|. \tag{13}$$

The *higher-order Randić index* or *higher-order connectivity index* is also of interest in molecular graph theory. For $t \geq 1$, the higher-order Randić index is defined as

$$R^{(t)}(\Gamma) = \sum_{v_{i_1}-v_{i_2}-\cdots-v_{i_{t+1}}} \frac{1}{\sqrt{\delta_{i_1}\delta_{i_2}\cdots\delta_{i_{t+1}}}},$$

where $v_{i_1} - v_{i_2} - \cdots - v_{i_{t+1}}$ runs over all paths of length t in Γ.

Now we are going to obtain tight bounds on $R_\alpha(\Gamma)$. Moreover, we are going to obtain tight bounds on $R_\alpha(\Gamma)$, $\alpha \neq -1, 0$, and $R^{(2)}(\Gamma)$ in terms of c_2 (the coefficient of λ^{n-2} in the characteristic polynomial,$\Psi(\mathcal{A}, \lambda)$, of the degree-adjacency matrix of Γ).

THEOREM 4. *Let Γ be a simple graph of order n and size m.*

(a) *The zeroth-order Randić index is bounded by*

$$\frac{n^3}{2m} \leq \mathrm{R}_0^2(\Gamma).$$

The equality holds if, and only if, Γ is regular.

(b) *Let $\alpha_1, \alpha_2 \in \mathbb{R} \setminus \{0\}$ such that $\alpha_1 < \alpha_2$. Then*

$$\alpha_1\alpha_2 > 0 \Rightarrow \mathrm{R}_{\alpha_1}^{\alpha_2}(\Gamma)m^{\alpha_1} \leq \mathrm{R}_{\alpha_2}^{\alpha_1}(\Gamma)m^{\alpha_2}. \tag{14}$$

$$\alpha_1\alpha_2 < 0 \Rightarrow \mathrm{R}_{\alpha_2}^{\alpha_1}(\Gamma)m^{\alpha_2} \leq \mathrm{R}_{\alpha_1}^{\alpha_2}(\Gamma)m^{\alpha_1}. \tag{15}$$

The equalities hold if, and only if, Γ is weight-regular.

(c) *Let $\vartheta_1 \geq \vartheta_2 \geq \cdots \geq \vartheta_n$ be the standard eigenvalues of Γ and let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the degree-adjacency eigenvalues of Γ, then*

$$\mathrm{R}(\Gamma) \leq \frac{1}{2}\sum_{i=1}^{n} |\lambda_i\vartheta_i|.$$

PROOF.

(a) Application of the Jensen's inequality to the convex function $f(x) = x^{-2}$ leads to the result. That is

$$\frac{n^2}{R_0^2(\Gamma)} = f\left(\frac{R_0(\Gamma)}{n}\right) \leq \frac{1}{n}\sum_{v_i \in V(\Gamma)} \delta_i = \frac{2m}{n}.$$

(b) Let $g(x) = x^{\frac{\alpha_2}{\alpha_1}}$, where $x > 0$. If ($\alpha_1 < 0$ and $\alpha_2 > 0$) or ($0 < \alpha_1 < \alpha_2$), application of the Jensen's inequality to the convex function g leads to

$$\left(\frac{R_{\alpha_1}(\Gamma)}{m}\right)^{\frac{\alpha_2}{\alpha_1}} = g\left(\frac{R_{\alpha_1}(\Gamma)}{m}\right) \leq \frac{R_{\alpha_2}(\Gamma)}{m}. \tag{16}$$

Thus, by (16), if $\alpha_1 < 0$ and $\alpha_2 > 0$ we obtain

$$R_{\alpha_2}^{\alpha_1}(\Gamma)m^{\alpha_2} \leq R_{\alpha_1}^{\alpha_2}(\Gamma)m^{\alpha_1} \tag{17}$$

and, if $0 < \alpha_1 < \alpha_2$, we obtain

$$R_{\alpha_1}^{\alpha_2}(\Gamma) m^{\alpha_1} \leq R_{\alpha_2}^{\alpha_1}(\Gamma) m^{\alpha_2}. \tag{18}$$

Analogously, if $\alpha_1 < \alpha_2 < 0$, application of the Jensen's inequality to the concave function g leads to (18). Hence, the result follows.

(c) The result is obtained by $2R(\Gamma) = Tr(\mathcal{A}\mathbf{A}) \leq \sum_{i=1}^{n} |\lambda_i \vartheta_i|$.

□

Notice that, in the case of weight-regular graphs, the bound (c) is attained. Moreover, as a particular case of (b), by (13), we deduce de following result.

Corollary 5. *Let Γ be a simple graph of size m. Then*

$$\alpha \in \mathbb{R} \setminus [-1, 0] \Rightarrow \frac{m^{\alpha+1}}{|c_2|^{\alpha}} \leq \mathrm{R}_{\alpha}(\Gamma) \tag{19}$$

$$\alpha \in (-1, 0) \Rightarrow \mathrm{R}_{\alpha}(\Gamma) \leq \frac{m^{\alpha+1}}{|c_2|^{\alpha}} \tag{20}$$

The equalities hold if, and only if, Γ is weight-regular.

As a particular case of above corollary we obtain

$$R(\Gamma) \leq \sqrt{m|c_2|}. \tag{21}$$

Theorem 6. *Let Γ be a simple and connected graph of order n and size m. Let ϕ denotes de graph invariant defined as $\phi = \left(\sum_{i=1}^{n} \sqrt{\delta_i}\right)^2 / 2m$, and let $\delta_* = \min\limits_{\delta_j > 1}\{\delta_j\}$. Then*

$$\left(\frac{(2R(\Gamma) - \phi)^2}{2(n - \phi)} + \frac{\phi}{2} + c_2\right)\sqrt{\delta_*} \leq R^{(2)}(\Gamma) \leq \sqrt{\Delta}\left(\frac{n}{2} + c_2\right),$$

where the lower bound holds only in the case of a non-regular graph.

Proof. For the vector $\mathbf{j} = (1, 1, ..., 1) \in \mathbb{R}^n$ we consider the following decomposition

$$\mathbf{j} = \frac{\langle \mathbf{j}, \nu \rangle}{\|\nu\|^2}\nu + z = \frac{\sum_{i=1}^{n} \sqrt{\delta_i}}{\sum_{i=1}^{n} \delta_i}\nu + z, \tag{22}$$

where $z \in \nu^{\perp}$. Then we have

$$\begin{aligned}
2R(\Gamma) &= \langle \mathcal{A}\mathbf{j}, \mathbf{j} \rangle \\
&= \left\langle \frac{\sum_{i=1}^{n} \sqrt{\delta_i}}{\sum_{i=1}^{n} \delta_i}\nu, \frac{\sum_{i=1}^{n} \sqrt{\delta_i}}{\sum_{i=1}^{n} \delta_i}\nu \right\rangle + \langle \mathcal{A}z, z \rangle \\
&= \frac{\left(\sum_{i=1}^{n} \sqrt{\delta_i}\right)^2}{\sum_{i=1}^{n} \delta_i} + \langle \mathcal{A}z, z \rangle \\
&= \phi + \langle \mathcal{A}z, z \rangle .
\end{aligned}$$

Thus, $2R(\Gamma) - \phi = \langle \mathcal{A}z, z \rangle$ and by the Cauchy-Schwarz inequality we obtain $|2R(\Gamma) - \phi| \leq \|\mathcal{A}z\|\|z\|$ and from $\|z\| = \sqrt{n - \phi}$ and $\|\mathcal{A}z\| = \sqrt{\|\mathcal{A}\mathbf{j}\|^2 - \phi}$ we obtain

$$|2R(\Gamma) - \phi| \leq \sqrt{(\|\mathcal{A}\mathbf{j}\|^2 - \phi)\,(n - \phi)}. \tag{23}$$

Moreover,

$$\|\mathcal{A}\mathbf{j}\|^2 = 2\sum_{v_i\sim v_j}\frac{1}{\delta_i\delta_j} + 2\sum_{v_i-v_j-v_k}\frac{1}{\sqrt{\delta_j\delta_i\delta_j\delta_k}} \tag{24}$$

$$\leq 2\sum_{v_i\sim v_j}\frac{1}{\delta_i\delta_j} + \frac{2}{\sqrt{\delta_*}}\sum_{v_i-v_j-v_k}\frac{1}{\sqrt{\delta_i\delta_j\delta_k}}. \tag{25}$$

Hence, by (4) we obtain

$$\|\mathcal{A}\mathbf{j}\|^2 \leq 2\left(\frac{R^{(2)}(\Gamma)}{\sqrt{\delta_*}} - c_2\right).$$

Thus, if Γ is non-regular, by the above inequality and (23) we conclude the proof of the left hand side inequality. On the other hand, by (1) we have $\|\mathcal{A}\mathbf{j}\|^2 \leq \|\mathbf{j}\|^2$=n, then, by (24) and (4) we have

$$-2c_2 + \frac{2R^{(2)}(\Gamma)}{\sqrt{\Delta}} \leq -2c_2 + 2\sum_{v_i-v_j-v_k}\frac{1}{\sqrt{\delta_j\delta_i\delta_j\delta_k}} \leq n.$$

Hence, the result follows. □

The above bounds are attained, for instance, in the case of the star graphs. Moreover, the upper bound is attained also in the case of regular graphs. The reader is referred to [**14**] for a complementary study on the Randić index.

4. Conditional excess

Let $D(\Gamma)$ denotes the diameter of Γ. We define, for any $k = 0, 1, \dots, D(\Gamma)$, the *$k$-excess* of a vertex $u \in V(\Gamma)$, denoted by $\mathbf{e}_k(u)$, as the number of vertices which are at distance greater than k from u. That is,

$$\mathbf{e}_k(u) = |\{v \in V : \partial(u,v) > k\}|.$$

Then, trivially, $\mathbf{e}_0(u) = n-1$, $\mathbf{e}_{D(\Gamma)}(u) = \mathbf{e}_{\varepsilon(u)}(u) = 0$ and $\mathbf{e}_k(u) = 0$ if and only if $\varepsilon(u) \leq k$, where $\varepsilon(u)$ denotes the eccentricity of u. The name "*excess*" is borrowed from Biggs [**1**], in which he gives a lower bound, in terms of the adjacency eigenvalues of a graph, for the excess $\mathbf{e}_r(u)$ of any vertex u in a δ-regular graph with odd girth $g = 2r+1$. The excess of a vertex was studied by Fiol and Garriga [**7**] using the adjacency eigenvalues of a graph, and by Yebra and the first author of this paper in [**12**] using the Laplacian eigenvalues.

The *k-excess* of Γ, denoted by $\mathbf{e}_k$, is defined as

$$\mathbf{e}_k = \max_{v_i \in V(\Gamma)}\{\mathbf{e}_k(v_i)\}.$$

This parameter was studied by Yebra and the first author of this paper in [**11**] using the Laplacian spectrum and the k-alternating polynomials.

We define the *conditional excess* of a vertex $v \in V(\Gamma)$ as follows:

$$\mathbf{e}_k^{\wp}(u) := |\{v \in \wp : \partial(u,v) > k\}|,$$

where $\wp$ is a property of some vertices of Γ and $v \in \wp$ means that the vertex v satisfies the property $\wp$. In this section we study the following particular case of conditional excess:

$$\mathbf{e}_k^{\beta}(u) := |\{v \in V(\Gamma) : \partial(u,v) > k \quad \text{and} \quad \delta(v) \geq \beta\}|.$$

To begin with, firstly we will recall the main properties of the k-alternating polynomials.

The k-alternating polynomials, defined and studied in [**6**] by Fiol, Garriga and Yebra, can be defined as follows: let $\mathcal{M} = \{\mu_1 > \cdots > \mu_b\}$ be a mesh of real numbers. For any $k = 0, 1, ..., b-1$ let P_k denote the *k-alternating polynomial* associated to $\mathcal{M}$. That is, the polynomial of $\mathbb{R}_k[x]$ with norm $\|P_k\|_\infty = \max_{1\le i\le b}\{|P_k(\mu_i)|\}$, such that

$$P_k(\mu) = \sup\{P(\mu) : P \in \mathbb{R}_k[x], \quad \|P\|_\infty \le 1\}$$

where μ is any real number greater than μ_1. We collect here some of its main properties, referring the reader to [**6**] for a more detailed study.

- For any $k = 0, 1, ..., b-1$ there is a unique P_k which, moreover, is independent of the value of $\mu(> \mu_1)$;
- P_k has degree k;
- $P_0(\mu) = 1 < P_1(\mu) < \cdots < P_{b-1}(\mu)$;
- P_k takes $k+1$ alternating values ± 1 at the mesh points;
- If $z \in \nu^\perp$ then $\|P_k(\mathcal{A}(\Gamma))z\| \le \|P_k\|_\infty \|z\|$ where $\nu = (\sqrt{\delta_1}, \sqrt{\delta_2}, ..., \sqrt{\delta_n})$, see [**13**];
- There are explicit formulae for $P_0(=1)$, P_1, P_2, and P_{b-1}, while the other polynomials can be computed by solving a linear programming problem (for instance by the simplex method).

THEOREM 7. *Let $\Gamma = (V, E)$ be a simple and connected graph of size m. Let $u \in V$ and let P_k be the k-alternating polynomial associated to the mesh of the degree-adjacency eigenvalues of* Γ. *Then,*

$$\mathbf{e}_k^\beta(u) \le \left\lfloor \frac{2m(2m - \delta(u))}{\beta\left[\delta(u)P_k^2(1) + 2m - \delta(u)\right]} \right\rfloor .$$

PROOF. Let $S = \{v_{i_1}, v_{i_2}, ..., v_{i_s}\} \subset V$ such that $\delta_{i_l} \ge \beta$ $(l = 1, ..., s)$, and $\partial(S, u) > k$. Let $\sigma = \sum_{l=1}^s e_{i_l}$, where e_{i_l} denotes the canonical vector associated to the vertex v_{i_l}, and let e be the canonical vector associated to the vertex u. From $\partial(S, u) > k \Rightarrow \langle P_k(\mathcal{A})\sigma, e\rangle = 0$, using the following decompositions

$$\sigma = \frac{\langle \sigma, \nu\rangle}{\|\nu\|^2}\nu + w_s = \frac{\sum_{l=1}^s \sqrt{\delta_{i_l}}}{2m}\nu + w_s, \tag{26}$$

$$e = \frac{\langle e, \nu\rangle}{\|\nu\|^2}\nu + w_u = \frac{\sqrt{\delta(u)}}{2m}\nu + w_u, \tag{27}$$

where $\nu = (\sqrt{\delta_1}, \sqrt{\delta_2}, ..., \sqrt{\delta_n})$ and $w_s, w_u \in \nu^\perp$, we obtain

$$P_k(1)\frac{\sqrt{\delta(u)}\sum_{l=1}^s \sqrt{\delta_{i_l}}}{2m} = -\langle P_k(\mathcal{A})w_s, w_u\rangle.$$

Hence, by the Cauchy-Schwarz inequality we have

$$P_k(1)\frac{\sqrt{\delta(u)}\sum_{l=1}^s \sqrt{\delta_{i_l}}}{2m} \le \|P_k(\mathcal{A})w_s\|\|w_u\|.$$

Thus,

$$P_k(1)\frac{\sqrt{\delta(u)}\sum_{l=1}^s \sqrt{\delta_{i_l}}}{2m} \le \|w_s\|\|w_u\|. \tag{28}$$

Moreover, the decompositions (26) and (27) lead to

$$s = \|\sigma\|^2 = \frac{\left(\sum_{l=1}^{s}\sqrt{\delta_{i_l}}\right)^2}{2m} + \|w_s\|^2 \Rightarrow \|w_s\| = \sqrt{s - \frac{\left(\sum_{l=1}^{s}\sqrt{\delta_{i_l}}\right)^2}{2m}}$$

and

$$1 = \|e\|^2 = \frac{\delta(u)}{2m} + \|w_u\|^2 \Rightarrow \|w_u\| = \sqrt{1 - \frac{\delta(u)}{2m}}.$$

So, by (28), we obtain

$$P_k(1)\sqrt{\delta(u)}\sum_{l=1}^{s}\sqrt{\delta_{i_l}} \leq \sqrt{(2m-\delta(u))\left(2ms - \left(\sum_{l=1}^{s}\sqrt{\delta_{i_l}}\right)^2\right)}. \tag{29}$$

Therefore,

$$P_k(1)s\sqrt{\delta(u)\beta} \leq \sqrt{(2m-\delta(u))\left(2ms - s^2\beta\right)}. \tag{30}$$

Solving (30) for s, and considering that it is an integer, we obtain the result. □

The above bound is tight for different values of k and β, as we can see in the following example. Let Γ be the graph of 5 vertices obtained by joining one vertex of the cycle C_4 to the vertex of the trivial graph K_1. The degree-adjacency eigenvalues of Γ are ± 1, $\pm\frac{\sqrt{6}}{6}$ and 0, from which we obtain $P_1(1) = 1.84...$ and $P_2(1) = 5.899....$ Hence, the values of the excess $\mathbf{e}_k^{\beta}(v)$ are attained whenever: $\delta(v) = 1$, $k = 0, 1, 2$ and $\beta = 2, 3$; $\delta(v) = 2$, $k = 1$ and $\beta = 3$; $\delta(v) = 2$, $k = 2$ and $\beta = 1, 2, 3$; $\delta(v) = 3$, $k = 2$ and $\beta = 2, 3$.

As we can see in Section 6, the above result becomes an important tool in the study of the conditional Wiener index.

An analogous upper bound on the standard excess is obtained by replacing, in above theorem, β by the minimum degree δ. Moreover, in the case of regular graphs, the above theorem becomes the following result.

Corollary 8. *Let Γ be a simple and connected graph of order n and let P_k be the k-alternating polynomial associated to the mesh of the degree-adjacency eigenvalues of Γ. Then,*

$$\mathbf{e}_k \leq \left\lfloor \frac{n(n-1)}{P_k^2(1) + n - 1} \right\rfloor.$$

The above result is analogous to the previous one obtained by the first author of this paper and Yebra in [**11**], for non-necessarily regular graphs, by using the Laplacian eigenvalues.

5. Degree diameter

In this section we study the problem of finding how far apart can be two vertices of given degrees in a connected graph. More precisely, the problem is to find

$$D^{(\alpha,\beta)}(\Gamma) := \max_{v_i, v_j \in V}\{\partial(v_i, v_j) : \delta_i \geq \alpha, \delta_j \geq \beta\}$$

We call this parameter (α, β)-*degree diameter*.

As in the case of the standard diameter, the study of this parameter is of interest in the design of interconnection networks when we need to minimize the communication delays between two nodes of given degrees.

In this section we obtain a tight bound on the (α,β)-degree diameter by using the k-alternating polynomials on the mesh of eigenvalues of the degree-adjacency matrix.

THEOREM 9. *Let $\Gamma=(V,E)$ be a simple and connected graph of size m. Let P_k be the k-alternating polynomial associated to the mesh of the degree-adjacency eigenvalues of Γ. Then,*

$$P_k(1) > \sqrt{\left(\frac{2m}{\alpha}-1\right)\left(\frac{2m}{\beta}-1\right)} \Rightarrow D^{(\alpha,\beta)}(\Gamma) \le k. \tag{31}$$

PROOF. Let e_i and e_j be the canonical vectors of $\mathbb{R}^n$ associated to the vertices v_i and v_j. Using the following decomposition

$$e_i = \frac{\langle e_i,\nu\rangle}{\|\nu\|^2}\nu + u = \frac{\sqrt{\delta_i}}{2m}\nu + u, \quad e_j = \frac{\langle e_j,\nu\rangle}{\|\nu\|^2}\nu + w = \frac{\sqrt{\delta_j}}{2m}\nu + w, \tag{32}$$

where $\nu=(\sqrt{\delta_1},\sqrt{\delta_2},...,\sqrt{\delta_n})$ and $u,w\in\nu^{\perp}$, we obtain

$$\begin{aligned}\partial(v_i,v_j)>k &\Rightarrow \left(P_k\left(\mathcal{A}\right)\right)_{ij}=0\\ &\Rightarrow \langle P_k(\mathcal{A})e_i,e_j\rangle = 0\\ &\Rightarrow P_k(1)\frac{\sqrt{\delta_i\delta_j}}{2m} + \langle P_k(\mathcal{A})u,w\rangle = 0\\ &\Rightarrow P_k(1)\frac{\sqrt{\delta_i\delta_j}}{2m} = -\langle P_k(\mathcal{A})u,w\rangle.\end{aligned}$$

Then, by the Cauchy-Schwarz inequality we have

$$\partial(v_i,v_j)>k \Rightarrow P_k(1)\frac{\sqrt{\delta_i\delta_j}}{2m} \le \|P_k(\mathcal{A})u\|\|w\| \tag{33}$$

$$\Rightarrow P_k(1)\frac{\sqrt{\delta_i\delta_j}}{2m} \le \|P_k\|_\infty\|u\|\|w\|. \tag{34}$$

Moreover, the decomposition (32) leads to

$$1=\|e_i\|^2=\frac{\delta_i}{2m}+\|u\|^2 \Rightarrow \|u\| = \sqrt{1-\frac{\delta_i}{2m}}$$

and

$$1=\|e_j\|^2=\frac{\delta_j}{2m}+\|w\|^2 \Rightarrow \|w\| = \sqrt{1-\frac{\delta_j}{2m}}$$

So, by (34), we obtain

$$\partial(v_i,v_j)>k \Rightarrow P_k(1)\sqrt{\delta_i\delta_j} \le \sqrt{(2m-\delta_i)(2m-\delta_j)}. \tag{35}$$

The converse of (35) leads to

$$P_k(1) > \sqrt{\frac{(2m-\delta_i)(2m-\delta_j)}{\delta_i\delta_j}} \Rightarrow \partial(v_i,v_j)\le k. \tag{36}$$

The result follows from (36). $\square$

As we can see in the following example, the above bound is attained for several values of α and β. The graph of Figure 2 has degree-adjacency eigenvalues

$$\left\{1, \tfrac{-3+\sqrt{249}}{24}, \tfrac{1}{4}, 0, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{-3-\sqrt{249}}{24}\right\}$$

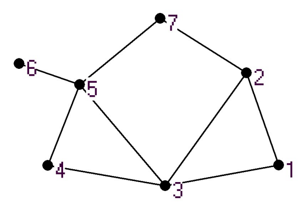
from which we obtain

$$P_1(1) = 1.7,\ P_2(1) = 5,\ P_3(1) = 15.2 \text{ and } P_4(1) = 58.$$

Thus, the following bounds are attained:

$$D^{(1,2)}(\Gamma) \leq 3,\ D^{(3,4)}(\Gamma) \leq 2 \text{ and } D^{(4,4)}(\Gamma) \leq 1.$$

FIGURE 2



As particular cases of above theorem we derive the following results in which the expression (31) is simplified.

COROLLARY 10. *Let $\Gamma = (V, E)$ be a simple and connected graph of order n and size m. Let P_k be the k-alternating polynomial associated to the mesh of the degree-adjacency eigenvalues of Γ. Then,*

$$P_k(1) > \frac{2m}{\beta} - 1 \Rightarrow D^{(\beta,\beta)}(\Gamma) \leq k. \tag{37}$$

The standard diameter is bounded by

$$P_k(1) > \frac{2m}{\delta} - 1 \Rightarrow D(\Gamma) \leq k. \tag{38}$$

If Γ is regular, the standard diameter is bounded by

$$P_k(1) > n - 1 \Rightarrow D(\Gamma) \leq k. \tag{39}$$

As we can see in next section, the bound (37) becomes an important tool in the study of the conditional Wiener index. Moreover, the bound (39) is an analogous result to the previous one given by Fiol, Garriga and Yebra in [**6**] by using the standard adjacency matrix. The reader is referred to [**13**] for a more general study on the conditional diameter.

6. Conditional Wiener index

The *Wiener index* $W(\Gamma)$ of a graph Γ with vertex set $\{v_1, v_2, ..., v_n\}$ defined as the sum of distances between all pairs of vertices of Γ,

$$W(\Gamma) := \frac{1}{2} \sum_{i=1, j=1}^{n} \partial(v_i, v_j),$$

is the first mathematical invariant reflecting the topological structure of a molecular graph.

This topological index has been extensively studied, for instance, a comprehensive survey on the direct calculation, applications and the relation of the Wiener index of trees with other parameters of graphs can be found in [**5**]. Moreover, a list of 120 references of the main works on the Wiener index of graphs can be found in the referred survey.

Alternatively, the Wiener index can be defined as

$$W(\Gamma) = \frac{1}{2} \sum_{v \in V(\Gamma)} S(v),$$

where $S(v)$ denotes the *distance of the vertex* v:

$$S(v) := \sum_{u \in V(\Gamma)} \partial(u, v).$$

We define the *conditional Wiener index*

$$W_{\wp}(\Gamma) := \frac{1}{2} \sum_{v \in \wp} S_{\wp}(v),$$

where $\wp$ is a property and $v \in \wp$ means that the vertex v satisfies the property $\wp$, and

$$S_{\wp}(v) := \sum_{u \in \wp} \partial(u, v)$$

is the *conditional distance* of v. In particular, if $\wp$ requires that $\delta(v) \geq \beta$, the conditional Wiener index will be denoted by $W_\beta(\Gamma)$, moreover, the conditional distance of v will be denoted by $S_\beta(v)$. Clearly, if β is the minimum degree of Γ, then $W_\beta(\Gamma)$ and the standard Wiener index coincides.

Lemma 11. *The conditional Wiener index of a graph* Γ*,* $W_\beta(\Gamma)$*, satisfies*

$$W_\beta(\Gamma) = \frac{1}{2} \sum_{\delta(v) \geq \beta} \sum_{k=0}^{D^{(\beta,\beta)}(\Gamma)-1} \mathbf{e}_k^\beta(v)$$

Proof. For each vertex $v \in V(\Gamma)$ of degree $\delta(v) > \beta$ we have

$$S_\beta(v) = \sum_{k=1}^{D^{(\beta,\beta)}(\Gamma)} k(\mathbf{e}_{k-1}^\beta(v) - \mathbf{e}_k^\beta(v)).$$

Moreover, by a simple calculation we have

$$S_\beta(v) = \sum_{k=0}^{D^{(\beta,\beta)}(\Gamma)-1} \mathbf{e}_k^\beta(v). \tag{40}$$

Hence, by (40) we obtain the result. □

Therefore, it follows from Lemma 11 that bounds on $\mathbf{e}_k^\beta$ lead to bounds on the conditional Wiener index W_β.

Theorem 12. *Let* $\Gamma = (V, E)$ *be a simple and connected graph of size* m*. Let* P_k *be the* k*-alternating polynomial associated to the mesh of the degree-adjacency eigenvalues of* Γ *and let* $x = |\{v \in V(\Gamma) : \quad \delta(v) \geq \beta\}|$*. If* $P_k(1) > \frac{2m}{\beta} - 1$*, then*

$$W_\beta(\Gamma) \leq \frac{x}{2} \sum_{l=0}^{k-1} \left\lfloor \frac{2m(2m-\beta)}{\beta\left(\beta P_l^2(1) + 2m - \beta\right)} \right\rfloor.$$

PROOF. By Lemma 11 and Theorem 7 we have

$$W_\beta(\Gamma) \leq \frac{x}{2} \sum_{k=0}^{D^{(\beta,\beta)(\Gamma)-1}} \left\lfloor \frac{2m(2m-\beta)}{\beta\left(\beta P_k^2(1)+2m-\beta\right)} \right\rfloor . \tag{41}$$

Therefore, by (37) we conclude the proof. □

An analogous upper bound on the standard Wiener index is obtained by replacing, in above theorem, β by δ, and x by n. Moreover, in the case of regular graphs, the above theorem becomes the following result.

COROLLARY 13. *Let Γ be a simple and connected δ-regular graph of order n. Let P_k be the k-alternating polynomial associated to the mesh of the degree-adjacency eigenvalues of Γ. If $P_k(1) > n-1$, then*

$$W(\Gamma) \leq \frac{n}{2} \sum_{l=0}^{k-1} \left\lfloor \frac{n(n-1)}{P_l^2(1)+n-1} \right\rfloor .$$

The reader is referred to [**15**] for a more general study on the Wiener index of hypergraphs.

References

[1] N. Biggs, *Algebraic graph theory*, Cambridge University Press, 1993.

[2] B. Bollobás and P. Erdös, Graphs of extremal weights, *Ars Combinatoria* **50** (1998), 225-233.

[3] D. M. Cvetković, M. Doob and H. Sachs, *Spectra of graphs*, Academic Press Inc., New York, 1979.

[4] K. C. Das and I. Gutman, Some properties of the second Zagreb index, *MATCH Commun. Math. Comput. Chem.* **52** (2004), 103-112.

[5] A. A. Dobrynin, R. Entringer and I. Gutman, Wiener Index of Trees: Theory and Applications, *Acta Applicandae Mathematicae* **66** (2001), 211-249.

[6] M.A. Fiol, E. Garriga and J.L.A. Yebra, On a class of polynomials and its relation with the spectra and diameters of graphs, *J. Combin. Theory Ser. B* **67** (1996), 48-61.

[7] M.A. Fiol and E. Garriga, From local adjacency polynomials to locally pseudo-distance-regular-graphs. *J. Combin. Theory Ser. B.* **71** (1997), 162-183.

[8] C. Godsil and G. Royle, *Algebraic graph theory*, Springer-Verlang New York, Inc. 2001.

[9] I. Gutman and M. Lepović, Choosing the exponent in the definition of the connectivity, *J. Serb. Chem. Soc.* **66** (2001), 605-611.

[10] M. Randić, On the characterization of molecular branching. *J. Amer. Chem. Soc.* **97** (1975), 6609-6615.

[11] J. A. Rodríguez and J.L.A. Yebra, Laplacian eigenvalues and the excess of a graph *Ars Combinatoria.* **64** (2002), 249-258.

[12] J. A. Rodríguez and J. L. A. Yebra, Cotas espectrales del exceso y la excentricidad de un conjunto de vértices de un grafo. *EAMA-97* (1997), 383-390.

[13] J. A. Rodríguez, The (α, β)-degree (s,t)-diameter of graphs: A particular case of conditional diameter. Submitted 2004.

[14] J. A. Rodríguez, A spectral approach to the Randić index. *Linear Algebra and its Applications* **400** (2005) 339-344.

[15] J. A. Rodríguez, On the Wiener index and the eccentric distance sum of hypergraphs, *MATCH. Communications in Mathematical and in Computer Chemistry.* **54** (1) (2005) 213-224.

DEPARTMENT OF MATHEMATICS, CARLOS III OF MADRID UNIVERSITY, AVDA. DE LA UNIVERSIDAD 30, 28911 LEGANÉS (MADRID), SPAIN

E-mail address: juanalberto.rodriguez@uc3m.es

E-mail address: josemaria.sigarreta@uc3m.es

BORREL ORBITS OF X^2=0 IN gl_n

BRIAN ROTHBACH

ABSTRACT. We analyzes the structure of Borel orbits in the subvariety of gl_n defined by $X^2 = 0$. The number of Borel orbits is finite, and is in one to one correspondence with certain partial permutation matrices. Equations are found up to radical for the Zariski closure of each orbit and these equations are shown to be generically reduced. The orbits are given a poset structure, which can also be described in terms of certain words. The Zariski closure of an orbit can be determined from the poset. The dimension of an orbit (as an algebraic variety) is given by a rank function for the poset, which is defined in terms of a statistic of the word of an orbit. An algorithm for calculating the degree of the Zariski closure of a given orbit is discussed.

Section 1. Orbital Varieties

Fix a positive integer n and a an algebraically closed field K (but we make no restrictions on the characteristic). Let $B_n(K)$ (or simply B_n) be the set of upper triangular $n \times n$ matrices. Let X be a nilpotent $n \times n$ matrix, that is a matrix satisfying $X^m = 0$ for some positive integer m. Notice that the eigenvalues of X are all 0, so that X is determined up to conjugacy by its Jordan canonical form. In turn, these canonical forms are parameterized by partitions of n boxes (with column lengths corresponding to sizes of the Jordan blocks).

The set of all nilpotent matrices in gl_n corresponding to a given partition is a natural object of study. Let K^n be the n dimensional K vector space with basis $e_1, \ldots, e_n$. Recall that gl_n acts on K^n by left multiplication. (Concretely, matrices act on column vectors). By a result of Gerstenhaber (see [3]), the Zariski closure of a conjugacy class of nilpotent matrices is defined by power-rank conditions, which are equations coming from conditions of the form $\dim_k(X^m K^n) \leq a_m$ for various nonnegative integers a_m. (However, this set of equations is generally not reduced).

A natural object to consider is the intersection of a conjugacy class of a nilpotent matrix with the set B_n of upper triangular matrices. These varieties arise in the study of Steinberg's triple variety [10] and their irreducible components classify the irreducible components of Springer fibers arising from the resolution of the flag variety. Understanding the irreducible components will also help in quantizing nilpotent conjugacy classes.

As the previous paragraph suggests, the intersection of a conjugacy class with the upper triangular matrices is not usually irreducible as an algebraic variety. The simplest example is the set of all 3×3 matrices corresponding to the partition 21, which has two components. An **orbital variety** is an irreducible component of the intersection of a nilpotent conjugacy class of gl_n and the upper triangular matrices B_n.

Example

For the partition 21, there are two orbital varieties corresponding to the following linear subspaces of B_n.

1. $\begin{pmatrix} 0 & x_{12} & x_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 2. $\begin{pmatrix} 0 & 0 & x_{13} \\ 0 & 0 & x_{23} \\ 0 & 0 & 0 \end{pmatrix}$

By the work of Spaltenstein [7] and Springer [10], the orbital varieties arising from a partition λ are in bijective correspondence with the standard Young tableau of shape λ and content 1^n. Given an orbital variety, one can obtain a tableau in the following manner. Take a generic matrix in the orbital variety. Then for each $1 \leq i \leq n$, the upper left square $i \times i$ submatrix is nilpotent, so one gets a sequence of partitions $\lambda_1, \ldots, \lambda_n = \lambda$, with $\lambda_i \subset \lambda_{i+1}$ for all i. Now one obtains a tableau of the appropriate shape and content by placing i in the unique box of $\lambda_i / \lambda_{i-1}$.

The inverse map from tableau (of shape λ and content 1^n) to orbital varieties can be described in terms of RSK. Given a tableau T, one obtains an element w of the Weyl group by applying RSK to the pair (T, T) (in fact, one gets an involution). Then the corresponding orbital variety is $O_\lambda \cap \overline{n_+ \cap (w \cdot n_+)}$, where O_λ is the corresponding nilpotent orbit and n_+ is the set of strictly upper triangular matrices.

The construction of the previous paragraph allows one to determine nice geometric information about orbital varieties. For example, for a given λ any orbital variety corresponding to λ has dimension $\frac{1}{2} \dim O_\lambda$. However, this construction gives little algebraic information. No algorithm is known for the equations of an orbital variety, and similarly it is difficult to determine if one orbital variety is contained in the closure of a second orbital variety.

One nice properties of orbital varieties is that they are stable under the conjugation action of the Borel group of invertible upper triangular matrices. Our general philosophy is to try to find and understand nice Borel stable subvarieties of the nilpotent cone. The natural inclination is to understand all Borel orbits; however there are infinitely many orbits, and moreover there exist continuous families of orbits. The right idea seems to be to use Borel invariants to define nice sets of Borel stable varieties of the nilpotent cone. However, the goal of our paper is more modest; we only examine the action of the Borel group in the set of matrices X with $X^2 = 0$.

Section 2. Borel orbits in $X^2 = 0$

Remark

Much of the work found in this paper was done independently by Melnikov ([5], [6]), but with the extra limitations that the matrices to be considered are upper triangular. Our work does not have this restriction. Also, Melnikov states her description of the Borel orbit poset in terms of involutions; our use of words to describe the Borel orbit poset make certain results easier to describe.

Let $M_n = gl_n$ be the set of all $n \times n$ matrices. Then M_n forms a variety with coordinate ring $K[x_{ij}]$ (where x_{ij} corresponds to the entry in the ith column and jth row of the generic matrix). Let V_n be the subvariety of M_n consisting of all $n \times n$ matrices with $X^2 = 0$. Notice that V_n can be described set theoretically by the ideal I_n generated by polynomials $p_{ij} = \Sigma_{k=1}^{k=n} x_{ik}x_{kj}$ for each ordered pair (i, j). (These equations do *not* generate V_n scheme theoretically, see section 12).

Let B_n be the $n \times n$ Borel group of invertible upper triangular matrices. B_n acts on V_n by $b(X) = bXb^{-1}$ for all $b \in B$, $X \in V_n$. We wish to study the Borel orbits of V_n. One way such orbits arise is from certain partial permutation matrices.

Definition

A partial permutation matrix is a matrix such that all entries are 0 or 1and such that each row and each column has at most one nonzero entry.

Notice that the concept of a partial permutation matrix is a generalization of a permutation matrix. Recall that a permutation matrix is determined uniquely by a word of length n in the letters $1, \ldots, n$. By generalizing this definition, a partial permutation can be described by a word of length n, but now using the alphabet $0, 1, \ldots, n$. This construction will also allow us to define a useful statistic later on.

Definition
Given a partial permutation matrix P, define the word of P to be $W_P = w_1, \ldots, w_n$, where $w_j = i$ if $Pe_j = e_i$ and $w_j = 0$ if $Pe_j = 0$.

Comment. Notice that the word of a partial permutation may have multiple zeroes, but each nonzero term occurs at most once.

Obviously not all partial permutation matrices give rise to orbits in V_n. We now classify all such matrices in terms of their words.

Definition

A word W is a valid X^2 word if for some partial permutation matrix P with $P^2 = 0$, $W = W_P$. Alternatively, a word W is a valid X^2 word if and only if the following two conditions hold:

1. No nonzero number of W appears more than once.

2. If $w_j = i > 0$, then $w_i = 0$.

The second condition suggest the following important definition that will be used later.

Definition

Let $W = w_1, \dots, w_n$ be a valid X^2 word. Suppose $w_i = 0$. We call w_i a bound zero (or just bound) if for some j, $w_j = i$. A letter w_i is said to be free if it is not a bound letter (in particular all nonzero letters are free).

Example.
The word 0103 is a valid X^2 word assigned to the partial permutation matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Note that it is easy to count the number of valid X^2 words. They are in obvious bijection with the set of directed partial matchings on n labeled points. (Given a valid X^2 word W, construct the graph with edges pointing from i to j if and only if $w_i = j$). Moreover, the exponential generating function for these words is e^{x^2+x}.

In general not all Borel orbits contain a partial permutation matrix. But in the case of $X^2 = 0$, we have the following theorem.

Theorem 1 Let $X \in V_n$. The orbit $B \cdot X$ contains a unique partial permutation matrix P_X, and this partial permutation matrix is given by a valid X^2 word. In particular, there are finitely many orbits, indexed by the valid X^2 words.

The idea behind the proof of this theorem is to construct Borel invariants. Given a matrix X, one can then use these invariants to determine a partial permutation matrix P_X. To finish the proof, one uses the fact that $X^2 = 0$ to inductively show that X and P_X are conjugate.

Section 3. Flags and Borel invariants

In order to construct the Borel invariants, we will have to recall the definition of a complete flag.

Definition. A complete flag of the vector space K^n is a sequence $V_0 \subset V_1 \subset \cdots \subset V_n = K^n$ of vector subspaces of K^n such that V_i has dimension i.

Example. An example of a complete flag is the standard complete flag, where $V_i = K^i$, the i dimensional vector space with basis $e_1, \dots, e_i$. (The vector spaces K^i will also be called the standard i flag.) Notice that a matrix b is an invertible upper triangular matrix if and only if $bK^i = K^i$ for all $1 \leq i \leq n$, that is if and only if it preserves the standard complete flag.

Now we can define a collection of Borel invariants that will allow us to classify all the orbits of V_n.

Definition. For any $0 \leq i, j \leq n$, let $r_{i,j}(X) = \dim(K^i + XK^j) - i$. Equivalently, $r_{i,j}(X)$ is the rank of the lower left $n-i \times j$ submatrix of X.

Notice that the first definition of the $r_{i,j}(X)$'s shows that they are invariant under the action of the Borel group, that is for all $b \in B$, we have $r_{i,j}(X) = r_{i,j}(bXb^{-1})$. One can see this by noting that the action of b just corresponds to a change of basis that fixes the standard flag.

For a partial permutation matrix P, $r_{i,j}(P)$ equals the number of ones in the lower left $n-i \times j$ submatrix of P. Alternatively, $r_{i,j}(P)$ can be calculated from W_P as the number of elements of $w_1, \ldots, w_j$ that are greater than i.

Example.
For the word 0103, one gets the following matrix of $r_{i,j}$'s.

$$\begin{pmatrix} & j=1 & j=2 & j=3 & j=4 \\ i=0 & 0 & 1 & 1 & 2 \\ i=1 & 0 & 0 & 0 & 1 \\ i=2 & 0 & 0 & 0 & 1 \\ i=3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We have seen that the word W_P determines the $r_{i,j}$'s. However, the converse is also true.

Lemma
Given P a partial permutation matrix with word $w_1, \ldots, w_n$, we have $w_j = i > 0$ if and only if $r_{i,j}(P) = r_{i,j-1}(P) = r_{i-1,j}(P) = r_{i-1,j-1}(P) + 1$. For a given j, $w_j = 0$ if and only if there is no i such that the condition in the previous sentence holds.

In particular, the Borel invariants $r_{i,j}(P)$ distinguish between any two partial permutation matrices.

Suppose that we have $X^2 = 0$. If X was conjugate to a partial permutation matrix P, then we could determine the word of P (and thus P itself) by noting that $r_{i,j}(P) = r_{i,j}(X)$ for all i and j. By mimicking this procedure, we can assign a potential partial permutation P_X to X.

The final step of the proof is to show that X and P_X are actually conjugate. Note that we must use the fact that $X^2 = 0$ here; for example the matrices

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

have the same $r_{i,j}$'s, but are not conjugate. In this case the first matrix has $X^3 = 0$ but not $X^2 = 0$.

Now we use the word W_{P_X} of P_X and the fact that $X^2 = 0$ to find an upper triangular change of coordinates where K^n decomposes into a direct sum of an X stable two dimensional vector space and an X stable $n-2$ dimensional vector space. (The two dimensional vector space will be spanned by e_i, e_{w_i}, where w_i is the first nonzero letter of W_{P_X}, while the $n-2$ is spanned by the other vectors). By induction, we see that X is conjugate to P_X.

Section 4. Algebraic considerations and the combinatorial Borel poset

Because of the previous theorem, we can identify a Borel orbit of V_n with either the unique partial permutation of that orbit or with the valid X^2 word of that partial permutation matrix. Notice that each Borel orbit is a locally closed algebraic set, since the condition $r_{i,j}(X) = r_{ij}$ is defined the vanishing and non vanishing of certain minors. One would like to determine the Zariski closure of any Borel orbit, and also certain geometric information such as the dimension of any orbit.

First we attempt to determine the Zariski closure of any orbit. Here we have an obvious candidate. Suppose that an orbit associated to P is defined by the equations $r_{i,j}(X) = r_{i,j}(P)$ and $X^2 = 0$. Notice that the condition $r_{i,j}(X) \leq r_{i,j}(P)$ is an algebraic condition, defined by the vanishing of all r_{ij} of the lower left $n-i \times j$ submatrix. Now we can conjecture that the Zariski closure of the orbit should be the variety $Cl(P)$ defined by $r_{i,j}(X) \leq r_{i,j}(P)$ and $X^2 = 0$. Clearly, this variety contains the Zariski closure of P.

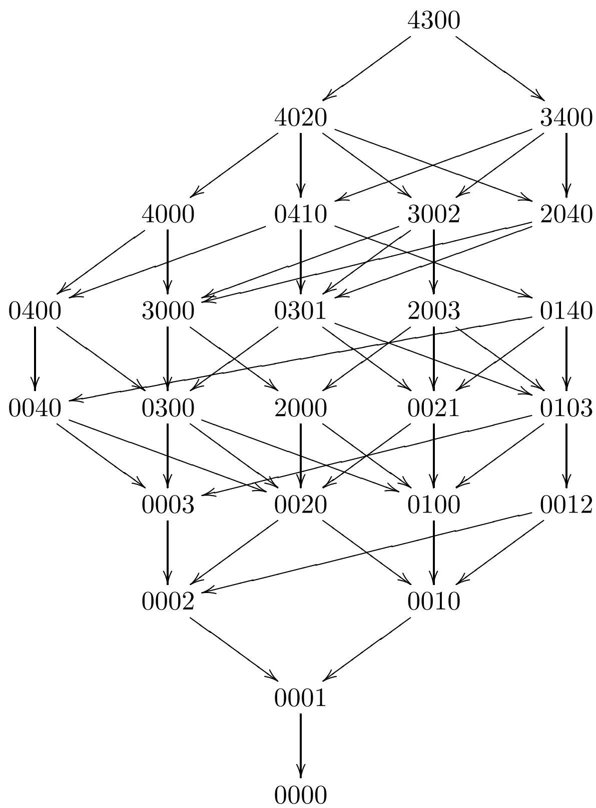
Now, the Zariski closure of a Borel orbit of V_n is a union of Borel orbits of V_n. So we need to determine which partial permutation matrices Q are contained the Zariski closure of P. Our approach will be to consider the set of all Q such that $r_{i,j}(Q) \leq r_{i,j}(P)$ for all i and j, and then to show that all such Q are contained in the Zariski closure of P. This conjecture for the Zariski closure of an orbit inspires us to define the following poset.

Definition. The **combinatorial Borel orbit poset** is a poset on the set of B-orbits in V_n, with the relation that $Q \leq P$ if and only if $Q \subset Cl(P)$. Equivalently, $Q \leq P$ if and only if $r_{i,j}(Q) \leq r_{i,j}(P)$ for all i and j.

The condition that $Q \leq P$ can also be interpreted as a combinatorial condition on the words $W_Q = v_1, \dots, v_n$ and $W_P = w_1, \dots, w_n$. Namely $Q \leq P$ if and only if for each $1 \leq l \leq n$, the elements $v_1, \dots, v_l$ of the initial l subword of W_Q cover the elements $w_1, \dots, w_l$ of the initial l subword of W_P, in the sense that there exists a permutation $\sigma_l \in S_l$, such that $w_i \leq v_{\sigma_l(i)}$ for all $1 \leq i \leq l$.

Example

We give the Hasse diagram for the combinatorial poset of valid X^2 words of length $n = 4$.



Section 5. Standard Moves and finding the Zariski closure

As part of the proof that $Cl(P)$ is the Zariski closure of P, we show that the combinatorial Borel poset is the transitive closure of a finite set of standard moves. Recall that $w_i = 0$ is bound if $w_j = i$ for some j, and w_i is free otherwise.

Definition. A standard move in the combinatorial Borel poset replaces a valid X^2 word W by a smaller valid X^2 word V in one of the four following ways.

1. Let w_i be any nonzero letter of W. Then a standard move of type 1 has $V = w_1, \ldots, w_{i-1}, w_i^*, w_{i+1}, \ldots, w_n$ where w_i^* is any value strictly less than w_i such that V is a valid X^2 word. In other words, V is obtained from W by decreasing the value of one particular nonzero letter. Note that for any nonzero letter w_i of W, there exists a word V obtained by from W by decreasing the letter w_i, since we can always replace w_i by 0 to get a valid X^2 word.

2. Let $w_i > w_j$ be any free letters with $i < j$. (Recall that a letter w_k is free unless for some $w_l = k$. In particular, any nonzero letter is free). Then a standard move of type 2 has $V = w_1, \ldots, w_{i-1}, w_j, w_{i+1}, \ldots, w_{j-1}, w_i, w_{j+1}, \ldots, w_n$. In other words, V is obtained from W by switching the free letters w_i and w_j. Notice that w_j may equal 0 if the zero is free, but w_i is never zero.

3. Let $w_i = k$ be any free letter and let $w_j = 0$ be any bound zero with $i < j$. Recall that since $w_j = 0$ is bound, then for some l, $w_l = j$. Then a standard move of type 3 has $V = w_1, \ldots, w_{i-1}, 0, w_{i+1}, \ldots, w_{j-1}, w_i, w_{j+1}, \ldots, w_{l-1}, i, w_{l+1}, \ldots, w_n$. In other words, V is obtained from W by switching the nonzero letter w_i and the bound zero $w_j = 0$ and replacing $w_l = j$ with i. (In general, we do not assume either $i < l$ or $j < l$ for a standard move of type 3).

4. Let $w_i = j$ be any nonzero letter such that $j > i$ (note $w_j = 0$). Then a standard move of type 4 has $V = w_1, \ldots, w_{j-1}, 0, w_{j+1}, \ldots, w_{i-1}, j, w_{i+1}, \ldots, w_n$. In other words, V is obtained from W by replacing w_i with 0 and replacing w_j with i.

We give several examples of standard moves, along with a sequence of permutations that show that the larger word covers the smaller word.

Examples

1a. $00 < 01$. This a standard move of type 1, with $i = 2$ and $w_i^* = 0$. A covering sequence is given by $\sigma_1 = id_1$, $\sigma_2 = id_2$.

1b. $001 < 002$. This is a standard move of type 1, with $i = 3$ and $w_i^* = 1$. A covering sequence is given by $\sigma_1 = id_1$, $\sigma_2 = id_2$, $\sigma_3 = id_3$.

2a. $0012 < 0021$. This is a standard move of type 2, with $i = 3$ and $j = 4$. A covering sequence is given by $\sigma_1 = id_1$, $\sigma_2 = id_2$, $\sigma_3 = id_3$, $\sigma_4 = (34)$.

2b. $001 < 010$. This is a standard move of type 2, with $i = 2$ and $j = 3$. Notice w_j is a free zero. A covering sequence is given by $\sigma_1 = id_1$, $\sigma_2 = id_2$, $\sigma_3 = (1)(23)$.

3. $0012 < 0103$. This is a standard move of type 3, with $i = 2$, $j = 3$ and $l = 4$. A covering sequence is given by $\sigma_1 = id_1$, $\sigma_2 = id_2$, $\sigma_3 = (23)$, $\sigma_4 = (23)$.

4. $01 < 20$. This is a standard move of type 4, with $i = 1$ and $j = 2$. A covering sequence is given by $\sigma_1 = id_1$, $\sigma_2 = (12)$.

One can generalize the covering sequences given above to show that any standard move gives rise to a relation in the combinatorial Borel poset. The following theorem shows that these moves generate the poset.

Theorem

The combinatorial Borel orbit poset relation $\leq$ is the transitive closure of the standard moves.

The proof of this theorem involves some difficult combinatorics, and the construction of a nice algorithm to construct a covering. As a consequence of this theorem, in order to show that $Cl(P)$ is the Zariski closure of P, it suffices to show that for any Q obtained from P by a standard move, Q is in the Zariski closure of P. One can now use geometric methods to finish the proof; namely one constructs an affine line such that the general point lies in P but a special point lies in Q.

Corollary

The Zariski closure of P is $Cl(P)$.

Section 6. A dimension statistic

The Hasse diagram from the $n = 4$ case suggests that the combinatorial Borel poset is ranked. In fact, that is the case. First, we construct the rank statistic. Recall that in a valid X^2 word, a letter $w_i = 0$ is a bound zero if for some j, $w_j = i$. A letter is free if it is not bound.

Definition. Let $W = w_1, \ldots, w_n$ be a valid X^2 word. A free inversion of W is a pair (i, j) with $1 \leq i < j \leq n$ such that w_i and w_j are both free letters and with $w_i > w_j$. We define $FI(W)$ to be the number of free inversions of W and we define $\pi(W) = FI(W) + \Sigma_{i=1}^{n} w_i$.

Theorem

The combinatorial Borel poset is a ranked poset, with rank function $\pi(W)$. Also, $\pi(W)$ is the Krull dimension of the Zariski closure the orbit associated to W.

We have two methods of proof for the second statement. One method just involves analyzing the conditions for P to cover Q, and showing $\pi(P) = \pi(Q) + 1$ in this case. The alternative method is to compute the Borel stabilizer of the partial permutation matrix.

Section 7. Hyperplanes and algebraic considerations

We know that the Zariski closure of an orbit corresponding to a partial permutation matrix P is defined set-theoretically by the conditions $X^2 = 0$ and $r_{i,j}(X) \leq r_{i,j}(P)$ for all i, j. One wants to prove that this set of equations is reduced, or to compute the radical if it is not reduced. In fact, for nonupper triangular orbits, the radical contains additional traces arising from the Levi factor of the appropriate parabolic subgroup.

Similarly, one would like to know the degree of an orbit closure as an algebraic variety. Also, these orbits seem to be Cohen-Macaulay in general.

Our general strategy to attack these questions is to look for Borel invariant hyperplane sections and try to set up an induction. Suppose $i = w_j$ is the largest element of the valid X^2 word W. Then $x_{i,j} = 0$ is such a Borel invariant hyperplane section, corresponding to the condition the $r(i-1, j) = 0$.

Using this method, one can show inductively by the method of principal radical systems that the ideals we have constructed are reduced for certain upper triangular orbits. (For nonupper triangular orbits, one must consider orbits of partial permutation matrices under the action of certain linear subgroups of the Borel group.)

As a corollary, one can show that any unions of matrix Schubert varieties coming from a rank table is reduced.

Once one can show that the hyperplane sections are reduced, it is relatively easy to show that the upper triangular orbits are Cohen-Macaulay.

Finally, the hyperplane sections also give us an inductive formula for computing degree.

Theorem

Let W be a valid X^2 word. Let $i = w_j$ be the largest nonzero letter of W.
a. If $W = 0, 0, \ldots, 0$, $\deg(Cl(O_W)) = 1$.
b. Otherwise,

$$\deg(Cl(O_W)) = \sum_{V \leq W, \pi(V)+1=\pi(W), v_i < w_i} m_V \deg(Cl(O_V)),$$

where $m_V = 1$ if V is obtained from W by a standard move of type 1, 2, or 3, and $m_V = 2$ if V is obtained from W by a standard move of type 4.

References

[1] E. Benlolo. *Etude sur les variétés orbitales dans* $\mathfrak{sl}_n(\mathbb{C})$. PhD thesis, Haifa University, 1990.

[2] J. Eagon and M. Hochster. Cohen-macaulay rings, invariant theory, and the generic perfection of determinantal loci. *American Journal of Mathematics*, 93(4):1020–1058, 1971.

[3] W. Hesselink. Singularities in the nilpotent scheme of a classical group. *Trans. Am. Math. Soc.*, 222:1–32, 1976.

[4] A. Knutson and E. Miller. Gröebner geometry of schubert polynomials. *math.AG/ 11 8*, 2001.

[5] A. Melnikov. B-orbits in solutions to the equation $x^2 = 0$ in triangular matrices. *J. of Algebra*, 223:101–108, 2000.

[6] A. Melnikov. Description of b-orbit closure of order 2 in upper triangular matrices. *math.RT/ 31 9* , 2003.

[7] N. Spaltenstein. The fixed point set of a unipotent transformation on the flag manifold. In *Proc. Konin. Nederl. Akad.*, pages 452–456, 1976.

[8] N. Spaltenstein. *Classes unipotents de sous-groupes de Borel*, volume 964 of *LN in Math.* Springer-Verlag, 1982.

[9] R. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, Cambridge, 1999.

[10] R. Steinberg. On the desingularization of the unipotent variety. *Invent. Math.*, 36:59–113, 1976.

[11] R. Steinberg. An occurance of robinson-schensted correspondence. *J. of Algebra*, 113:523–528, 1988.

[12] P. Thijsse. Upper triangular similarity of upper triangular matrices. *Linear Algebra Appl.*, 260:119–149, 1997.

MULTIPLICITY FREE EXPANSIONS OF SCHUR P-FUNCTIONS

KRISTIN M. SHAW AND STEPHANIE VAN WILLIGENBURG

Abstract. After deriving inequalities on coefficients arising in the expansion of a Schur P-function in terms of Schur functions we give criteria for when such expansions are multiplicity free. From here we study the multiplicity of an irreducible spin character of the twisted symmetric group in the product of a basic spin character with an irreducible character of the symmetric group, and determine when it is multiplicity free.

Résumé. Par l'obtention d'inégalités sur les coefficients du développement des P fonctions de Schur en termes de fonctions de Schur ordinaires, nous donnons une caractérisation des cas où ces développements n'ont pas de multiplicités. D'ici nous étudions des multiplicités des caractères projectifs du groupe symétrique en les développements des produits des caractères linéaire avec des caractères projectifs.

1. Introduction

In [4], Stembridge determined when the product of two Schur functions is multiplicity free, which yielded when the outer product of characters of the symmetric groups did not have multiplicities. Meanwhile, in [1] Bessenrodt determined when the product of two Schur P-functions is multiplicity free. This led to an analogous classification with respect to projective outer products of spin characters of double covers of the symmetric groups. In this article we interpolate between these two results to determine when a Schur P-function expanded in terms of Schur functions is multiplicity free. As an application we give criteria for when the multiplicity of an irreducible spin character of the twisted symmetric groups in the product of a basic spin character with an irreducible character of the symmetric groups is 0 or 1.

2000 *Mathematics Subject Classification.* Primary 05E05, 05A17; Secondary 05A19, 05E10.

Key words and phrases. multiplicity free, Schur functions, Schur P-functions, spin characters, staircase partitions.

Both authors were supported in part by the National Sciences and Engineering Research Council of Canada.

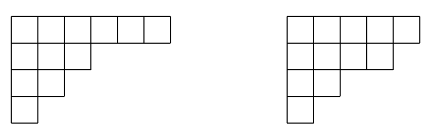
The remainder of this paper is structured as follows. We review the necessary definitions in the rest of this section. Then in Section 2 we derive some equalities and inequalities concerning certain coefficients. In Section 3 we give criteria for a Schur P-function to have a multiplicity free Schur function expansion before applying this to character theory in Section 4.

1.1. **Partitions.** A *partition* $\lambda = \lambda_1\lambda_2\ldots\lambda_k$ of n is a list of integers $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0$ whose sum is n, denoted $\lambda \vdash n$. We say k is the *length* of λ denoted by $l(\lambda)$, n is the *size* of λ and call the λ_i *parts*. Also denote the set of all partitions of n by $P(n)$. Contained in $P(n)$ is the subset of partitions $D(n)$ consisting of all the partitions whose parts are distinct i.e. $\lambda_1 > \lambda_2 > \ldots > \lambda_k > 0$. We call such partitions *strict*. A strict partition that will be of particular interest to us will be the *staircase* (of length k): $k(k-1)(k-2)\ldots 321$. Two other partitions that will be of interest to us are $\lambda+1^r$ and $\lambda\cup r$ for any given partition λ, and positive integer r. The partition $\lambda+1^r$ of formed by adding 1 to the parts $\lambda_1, \lambda_2, \ldots, \lambda_r$ and $\lambda \cup r$ is formed by sorting the multiset of the union of parts of λ and r. With these concepts in mind we are able to define a final partition that will be of interest to us, known as a near staircase. A partition is a *near staircase* if it is of the form $\lambda+1^r$, $1 \leq r \leq k$ or $\lambda \cup r$, $r \geq k+1$ and λ is the staircase of length k.

Example 1.1. 6321 and 5421 are both near staircases of 12.

1.2. **Diagrams and tableaux.** For any partition $\lambda \vdash n$ the associated *(Ferrers) diagram*, also denoted by λ, is an array of left justified boxes with λ_i boxes in the i-th row, for $1 \leq i \leq l(\lambda)$. Observe that in terms of diagrams a near staircase is more easily visualised as a diagram of a strict partition such that the deletion of exactly one row or column yields the diagram of a staircase.

Example 1.2. The near staircases 6321 and 5421.



Given a diagram λ then the *conjugate* diagram of λ, λ', is formed by transposing the rows and columns of λ. The resulting partition λ' is also known as the conjugate of λ. The *shifted* diagram of λ, $S(\lambda)$ is formed by shifting the i-th row $(i-1)$ boxes to the right. If we are given two diagrams λ and μ such that if μ has a box in the (i,j)-th position then λ has a box in the (i,j)-th position then the *skew* diagram λ/μ is formed by the array of boxes

$$\{c|c \in \lambda, c \notin \mu\}.$$

Now that we have introduced the necessary diagrams we are now in a position to fill the boxes and form tableaux.

Consider the alphabet

$$1' < 1 < 2' < 2 < 3' < 3 < \dots$$

For convenience we call the integers $\{1, 2, 3, \dots\}$ *unmarked* and the integers $\{1', 2', 3', \dots\}$ *marked*. Any filling of the boxes of a diagram λ with letters from the above alphabet is called a *tableau* of shape λ. If we fill the boxes of a skew or shifted diagram we similarly obtain a skew or shifted tableau. Given any type of tableau, T, we define the *reading word* $w(T)$ to be the entries of T read from right to left and top to bottom, and define the *augmented reverse reading word* $\hat{w}(T)$ to be $w(T)$ read backwards with each entry increased by one according to the total order on our alphabet e.g. if $T = \begin{array}{cc} 1' & 1 \\ 1 & 2' \\ 2 & \end{array}$ then $w(T) = 11'2'12$ and $\hat{w}(T) = 3'2'212'$. When there is no ambiguity concerning the tableau under discussion we refer to the reading word and augmented reverse reading word as w and $\hat{w}$ respectively. We also define the content of T, $c(T)$, to be the sequence of integers $c_1c_2\dots$ where

$$c_i = |i| + |i'|$$

and $|i|$ is the number of is in $w(T)$ and $|i'|$ is the number of i's in $w(T)$. For our previous example $c(T) = 32$. Given a word we say it is *lattice* if as we read it if the number of is we have read is equal to the number of $(i+1)$s we have read then the next symbol we read is neither an $(i+1)$ nor an $(i+1)'$ e.g. $11'2'23$ is lattice, however, $11'22'3$ is not.

Let T be a (skew or shifted) tableau, then we say T is *amenable* if it satisfies the following [2, p259]:

(1) The entries in each row of T weakly increase.
(2) The entries in each column of T weakly increase.
(3) Each row contains at most one i' for $i \geq 1$.
(4) Each column contains at most one i for each $i \geq 1$.
(5) The word $w\hat{w}$ is lattice.
(6) In w the rightmost occurrence of i is to the right of the rightmost occurrence of i' for all i.

Example 1.3. The first tableau is amenable whilst the second is not as it violates the lattice condition.

$$\begin{array}{ll} 1' & 1 \\ 1 & 2' \\ 2 & \end{array} \qquad \begin{array}{ll} 1' & 1 \\ 1 & 2 \\ 2 & \end{array}$$

Before we define Schur P-functions we make two observations about amenable tableaux.

Lemma 1.1. *Let T be an amenable tableau then if i or i' appear in row j then $j \geq i$.*

Proof. We proceed by induction on the number of rows of T. If T has one row then the result is clear. Assume the result holds up to row $(k-1)$. Consider row k. If it has an entry in it greater than k, l or l', then it must lie in the rightmost box since the rows of T weakly increase. However, this ensures that $w\hat{w}$ is not lattice as when we first read l or l' in w we will have read no $(l-1)$ or l. $\square$

Lemma 1.2. *Let T be an amenable tableau with $c(T) = k(k-1)(k-2)\dots(k-j)$ then in $w(T)$*

$$|i'| \leq |(i+1)'| \qquad 1 \leq i \leq j.$$

Proof. To prove this we consider the lattice condition on $w\hat{w}$. Assume $|i'| > |(i+1)'|$ then as we read w there will be a rightmost occurrence when $|i| = |(i+1)|$. However, because of $c(T)$ before we read another i in $w\hat{w}$ we must read $(i+1)$ or $(i+1)'$. $\square$

1.3. **Schur P-functions.** Given commuting variables $x_1, x_2, x_3, \dots$ let the r-th elementary symmetric function, e_r, be defined by

$$e_r = \sum_{i_1 < i_2 < \dots < i_r} x_{i_1} x_{i_2} \dots x_{i_r}.$$

Moreover, for any partition $\lambda = \lambda_1 \lambda_2 \dots \lambda_k$ let

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

then the algebra of symmetric functions, Λ, is the algebra over $\mathbb{C}$ spanned by all e_λ where $\lambda \vdash n, n > 0$ and $e_0 = 1$. Another well-known basis for Λ is the basis of *Schur functions*, s_λ, defined by

$$s_\lambda = \det(e_{\lambda'_i - i + j})_{1 \leq i,j \leq k}$$

where $e_r = 0$ if $r < 0$. It is these that can be used to define the subalgebra of Schur P-functions, Γ. More precisely, let λ be a strict partition of $n > 0$, then Γ is spanned by $P_0 = 1$ and all

$$P_\lambda = \sum_{\mu \vdash n} g_{\lambda\mu} s_\mu$$

where $g_{\lambda\mu}$ is the number of amenable tableaux T of shape μ and content λ.

Amenable tableaux can also be used to describe the multiplication rule for the P_λ as follows. Let λ, μ, ν be strict partitions then

$$P_\mu P_\nu = \sum f^\lambda_{\mu\nu} P_\lambda$$

where $f^\lambda_{\mu\nu}$ is the number of amenable shifted skew tableaux of shape λ/μ and content ν. Further details on Schur and Schur P-functions can be found in [2].

1.4. **Acknowledgements.** The authors would like to thank the referees for their valuable comments.

2. Relations on Stembridge coefficients

The combinatorial descriptions of the $g_{\lambda\mu}$ and $f^\lambda_{\mu\nu}$ from the previous section were discussed by Stembridge [3] who also implicitly observed the following useful relationship between them.

Lemma 2.1. *If $\lambda \in D(n)$, $\mu \in P(n)$, and ν is a staircase of length $l(\mu)$ then*

$$f^{\mu+\nu}_{\lambda\nu} = g_{\lambda\mu}.$$

Proof. By definition $f^{\mu+\nu}_{\lambda\nu}$ is the number of amenable shifted skew tableaux of shape $\mu+\nu/\lambda$ and content ν. However, since $P_\lambda P_\nu = P_\nu P_\lambda$ we know $f^{\mu+\nu}_{\lambda\nu}$ is also the number of amenable shifted skew tableaux of shape $\mu+\nu/\nu$ and content λ. In addition, the shifted skew diagram $\mu+\nu/\nu$ is simply the diagram μ. Thus $f^{\mu+\nu}_{\lambda\nu}$ is the number of amenable tableaux of shape μ and content λ, and this is precisely $g_{\lambda\mu}$. □

There are also equalities between the $g_{\lambda\mu}$.

Lemma 2.2. *If $\lambda \in D(n)$ and $\mu \in P(n)$ then*

$$g_{\lambda\mu} = g_{\lambda\mu'}.$$

Proof. Let $\omega : \Lambda \to \Lambda$ be the involution on symmetric functions such that $\omega(s_\lambda) = s_{\lambda'}$. In Exercise 3 [2, p259] it was proved that $\omega(P_\lambda) = P_\lambda$. Consequently,

$$\omega(P_\lambda) = \sum g_{\lambda\mu}\omega(s_\mu) = \sum g_{\lambda\mu} s_{\mu'}$$

and the result follows. □

Finally, we present two inequalities that will be useful in the following sections and relate amenable tableaux of different shape and content.

Lemma 2.3. *Given $\lambda \in D(n)$ and $\mu \in P(n)$, if $r \leq l(\lambda)$ and $s \geq \lambda_1$ then*

$$g_{\lambda\mu} \leq g_{(\lambda+1^r)(\mu+1^r)}$$

and

$$g_{\lambda\mu} \leq g_{(\lambda\cup s)(\mu\cup s)}.$$

Proof. Consider an amenable tableau T of shape μ and content λ. To prove the first inequality, for $1 \leq i \leq r$ append a box containing i to row i on the right side of T. By Lemma 1.1 it is straightforward to verify that this is an amenable tableau of shape $\mu + 1^r$ and content $\lambda + 1^r$. For the second inequality, replace each entry i with $(i+1)$ and each entry i' with $(i+1)'$ for $1 \leq i \leq l(\lambda)$ to form T'. Then append a row of s boxes each containing 1 to the top of T'. Again, it is straightforward to check this is an amenable tableau of shape $\mu \cup s$ and content $\lambda \cup s$. $\square$

Example 2.1. Consider the amenable tableau $\begin{matrix} 1' & 1 & 1 \\ 1' & 2 & 2 \\ 1 & & \end{matrix}$. Then the following two tableaux illustrate the operations utilised in proving the first and second inequalities, respectively.

$$\begin{matrix} 1' & 1 & 1 & \mathbf{1} \\ 1' & 2 & 2 & \mathbf{2} \\ 1 & \mathbf{3} & & \end{matrix} \qquad \begin{matrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 2' & 2 & 2 & & & \\ 2' & 3 & 3 & & & \\ 2 & & & & & \end{matrix}$$

3. Multiplicity free Schur expansions

Despite the number of conditions amenable tableaux must satisfy, it transpires that most Schur P-functions do not have multiplicity free expansions in terms of Schur functions.

Example 3.1. Neither P_{541} nor P_{654} are multiplicity free. We see this in the first case by observing there at least two amenable tableaux of shape 4321 and content 541:

$$\begin{matrix} 1' & 1 & 1 & 1 \\ 1 & 2' & 2 & \\ 2 & 2 & & \\ 3 & & & \end{matrix} \qquad \begin{matrix} 1' & 1 & 1 & 1 \\ 1 & 2' & 2 & \\ 2' & 3 & & \\ 2 & & & \end{matrix}.$$

In the second case we note that there are at least two amenable tableaux of shape 54321 and content 654:

$$\begin{matrix} 1' & 1 & 1 & 1 & 1 \\ 1 & 2' & 2 & 2 & \\ 2' & 3' & 3 & & \\ 2 & 3' & & & \\ 3 & & & & \end{matrix} \qquad \begin{matrix} 1' & 1 & 1 & 1 & 1 \\ 1 & 2' & 2 & 2 & \\ 2 & 2 & 3' & & \\ 3' & 3 & & & \\ 3 & & & & \end{matrix} \quad .$$

However, in some cases it is easy to deduce a certain Schur P-function expansion is multiplicity free as we can give a precise description of it in terms of Schur functions.

Proposition 3.1.

$$P_n = \sum_{1\le k\le n} s_{k1^{n-k}}$$

and

$$P_{(n-1)1} = \sum_{1\le k\le n} s_{k1^{n-k}} + \sum_{2\le k\le n-2} s_{k21^{n-k-2}}.$$

Proof. For the first result observe that $g_{n(k1^{n-k})} = 1$ since the tableau filled with one $1'$ and $(k-1)$ 1s in the first row and one 1 and $(n-k)$ $1'$s in the first column is the only amenable tableau of shape $k1^{n-k}$ and content n. Then observe that $g_{n\mu} = 0$ for any other μ since we have no way to fill a 2×2 rectangle with only 1 or $1'$ and create an amenable tableau.

For the second result note that since the content of any tableau we create is $(n-1)1$ we will be filling our diagram $(n-1)$ 1 or $1'$s and one 2. As in the previous case if our resulting tableau is to be amenable the 1s must appear in the first row and column marked or unmarked as necessary. The unmarked 2 can now only appear in one of two places, either at the end of the second row, or the end of the first column. □

A third multiplicity free expansion is obtained from the determination of when a Schur P-function is equal to a Schur function.

Theorem 3.2.

$$P_\lambda = s_\lambda \textit{ if and only if } \lambda \textit{ is a staircase.}$$

Proof. The reverse implication is proved in Exercise 3(b) [2, p 259]. For the forward implication assume that $\lambda = \lambda_1 \dots \lambda_k \in D(n)$ is not a staircase. It follows there must exist at least one $1 \le i \le k-1$ for which $\lambda_i \ge \lambda_{i+1} + 2$. We are going to show that in this situation there are at least two amenable tableaux with content λ.

Consider the tableau, T, of shape and content λ where the j-th row is filled with unmarked js. Clearly T is amenable. Now consider the first row i for which $\lambda_i \geq \lambda_{i+1} + 2$. Delete the rightmost box from this row and append it to the first column of T to form a tableau T' of shape $\lambda_1\lambda_2 \ldots \lambda_i - 1 \ldots \lambda_k 1$ and content λ. Now alter the entries in the first column of T' as follows. In row i change i to i' and in rows $i+1 \leq j \leq k$ change j to $j-1$. Finally in row $k+1$ change i to k. Now T' is an amenable tableau, and we are done. $\square$

As we will see, staircases play an important role in the determination of multiplicity free expansions of Schur P-fuctions.

Theorem 3.3. *For $\lambda \in D(n)$ the Schur function expansion of P_λ is multiplicity free if and only if λ is one of the following*

(1) *staircase*
(2) *near staircase*
(3) $k(k-1)(k-2)\ldots 43$
(4) $k(k-1)$.

Proof. If $\lambda \in D(n)$ is not one of the partitions listed in Theorem 3.3 then it must satisfy one of the following:

(1) There exists an $1 < i < l(\lambda)$ such that $\lambda_i \geq \lambda_{i+1} + 3$.
(2) For $1 \leq i < l(\lambda)$, $\lambda_i = \lambda_{i+1} + 1$ and $\lambda_{l(\lambda)} \geq 4$.
(3) There exists $i < j$ such that $\lambda_i = \lambda_{i+1} + 2$ and $\lambda_j = \lambda_{j+1} + 2$.

If λ satisfies the first criterion then by observing that P_{541} is not multiplicity free and Lemma 2.3 it follows that P_λ is not multiplicity free. Similarly if λ satisfies the second criterion then by observing P_{654} has multiplicity and Lemma 2.3, P_λ again has multiplicity. Lastly, if λ satisfies the third criterion then consider the partition $\nu = k(k-2)(k-3)\ldots 431 \in D(n)$ and the partition $\mu = (k-1)(k-2)(k-3)\ldots 432 \in D(n)$ for $k \geq 6$. We now show that $g_{\nu\mu} > 1$. Take a diagram μ and fill the first row with one $1'$ and $(k-2)$ 1s. For $i > 1$ fill the i-th row with one $(i-1)$ and the rest is. This is clearly an amenable tableau T. If we now change the $(k-3)$ to a $(k-3)'$ in the second column of the penultimate row of T we obtain another amenable tableau of shape μ and content ν. Thus $g_{\nu\mu} > 1$ and P_ν is not multiplicity free. This combined with Lemma 2.3 yields that P_λ is not multiplicity free if λ satisfies the third criterion.

Finally it remains to show that if $\lambda \in D(n)$ is one of the partitions listed in Theorem 3.3 then $g_{\lambda\mu} \leq 1$ for all μ. If λ is a staircase the result follows from Theorem 3.2. If λ is a near staircase of the form $mk(k-1)\ldots 321$ then by Lemma 1.2 it follows that any amenable tableau of content λ must contain no i' for $i > 1$ and thus there can exist at most one amenable tableau of shape μ and content λ for any given μ. Consequently $g_{\lambda\mu} \leq 1$ for all μ. Similarly if λ is the other type of near staircase or of the form $k(k-1)\ldots 43$ then by Lemma 2.1 we can calculate $g_{\lambda\mu}$ by enumerating all amenable shifted skew tableaux, T, of shape $\mu + \nu/\lambda$ and *staircase* content. By Lemma 1.2 it follows that no entry in T can be marked. From this we can deduce that there can exist at most one amenable shifted skew tableau and so $g_{\lambda\mu} \leq 1$. Finally a proof similar to that of Proposition 3.1 yields that $P_{k(k-1)}$ is multiplicity free. $\square$

Remark 3.2. The reverse direction of the above theorem can also be proved via Lemma 2.1 and [1, Theorem 2.2].

4. Multiplicity free spin character expansions

The twisted symmetric group $\tilde{S}_n$ is presented by

$$\langle z, t_1, t_2, \ldots, t_{n-1} |\ z^2 = 1, t_i^2 = (t_i t_{i+1})^3 = (t_i t_j)^2 = z\ |i-j| \geq 2\rangle.$$

Moreover, the ordinary representations of $\tilde{S}_n$ are equivalent to the projective representations of the symmetric group S_n and in [3] Stembridge determined the product of a basic spin character of $\tilde{S}_n$ with an irreducible character of S_n, whose description we include here for completeness. If $\lambda = \lambda_1 \ldots \lambda_k \in D(n)$ then define

$$\varepsilon_\lambda = \begin{cases} 1 & \text{if } n-k \text{ is even} \\ \sqrt{2} & \text{if } n-k \text{ is odd} \end{cases}.$$

Let ϕ^λ be an irreducible spin character of $\tilde{S}_n$, χ^μ for $\mu \in P(n)$ be an irreducible character of S_n and $\langle \cdot, \cdot \rangle$ be defined on Λ by $\langle s_\mu, s_\nu \rangle = \delta_{\mu\nu}$ then we have

Theorem 4.1. [3, Theorem 9.3] *If $\lambda \in D(n)$, $\mu \in P(n)$ then*

$$\langle \phi^n \chi^\mu, \phi^\lambda \rangle = \frac{1}{\varepsilon_\lambda \varepsilon_n} 2^{(l(\lambda)-1)/2} g_{\lambda\mu} \tag{4.1}$$

unless $\lambda = n$, n even, and $\mu = k1^{n-k}$ in which case the multiplicity is 0 *or* 1.

Using this formula we can deduce

Theorem 4.2. *If $\lambda \in D(n), \mu \in P(n)$ then the coefficient of ϕ^λ in $\phi^n\chi^\mu$ is multiplicity free for all μ if and only if λ is one of the following*

(1) n
(2) $(n-1)1$
(3) $k(k-1)$
(4) $(2k+1)21$

(5) 543
(6) 431.

Proof. Considering Equation 4.1 we first show no λ exists such that $g_{\lambda\mu} \geq 2$ but $\langle \phi^n \chi^\mu, \phi^\lambda \rangle$ is multiplicity free. If such a λ did exist then

$$2^{(l(\lambda)-1)/2} < \varepsilon_\lambda \varepsilon_n$$

where $\varepsilon_\lambda \varepsilon_n = 1, \sqrt{2}, 2$ depending on λ and its size. However, if $\varepsilon_\lambda \varepsilon_n = 1$ then $l(\lambda) = 0$ and if $\varepsilon_\lambda \varepsilon_n = \sqrt{2}$ then $l(\lambda) = 1$ so $\lambda = n$ but then $\varepsilon_\lambda \varepsilon_n \neq \sqrt{2}$ so we must have that

$$2^{(l(\lambda)-1)/2} < 2$$

and hence $l(\lambda) < 3$. Since $\varepsilon_\lambda, \varepsilon_n = \sqrt{2}$ it follows n is even and we must in fact have $\lambda = n$. By Proposition 3.1 we know in this case $g_{\lambda\mu} \leq 1$ for all μ and we have our desired contradiction. Consequently if $\langle \phi^n \chi^\mu, \phi^\lambda \rangle$ is multiplicity free then $g_{\lambda\mu} \leq 1$. Additionally we must have $2^{(l(\lambda)-1)/2} = \varepsilon_\lambda \varepsilon_n$ and so it remains for us to check three cases.

(1) $\varepsilon_\lambda \varepsilon_n = 1$: We have $l(\lambda) = 1$ and so by Theorem 3.3 $\lambda = n$, n odd.
(2) $\varepsilon_\lambda \varepsilon_n = \sqrt{2}$: We have $l(\lambda) = 2$ and so by Theorem 3.3 $\lambda = (n-1)1$, or $\lambda = k(k-1)$.
(3) $\varepsilon_\lambda \varepsilon_n = 2$: We have $l(\lambda) = 3$ and since $\varepsilon_n = \sqrt{2}$ it follows that n is even. Hence by Theorem 3.3 $\lambda = (2k+1)21$, $\lambda = 543$, or $\lambda = 431$.

□

References

[1] C. Bessenrodt, On multiplicity-free products of Schur P-functions, *Ann. Combin.* **6**, (2002), 119–124.
[2] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Second Edition, Oxford University Press, Oxford, 1995.
[3] J. Stembridge, Shifted tableaux and the projective representations of symmetric groups, *Adv. in Math.* **74**, (1989), 87–134.
[4] J. Stembridge, Multiplicity-free products of Schur functions, *Ann. Combin.* **5**, (2001), 113–121.

E-mail address: krishaw@math.ubc.ca

E-mail address: steph@math.ubc.ca

Department of Mathematics, Univ. of British Columbia, Vancouver, BC V6T 1Z2

The Discrete Fundamental Group of the Order Complex of B_n

Shelly Smith

Abstract

A-theory is a recently developed area of algebraic combinatorics that takes concepts from algebraic topology and transfers them to a combinatorial setting. It contains discrete analogues to continuity, homotopy, and fundamental group, defined on graphs and simplicial complexes. We provide a construction for a graph arising from the order complex of the direct product of two graded lattices. With this construction, we show that the rank of the abelianization of the discrete fundamental group of the order complex of the Boolean lattice, B_n, is $2^{n-3}(n^2 - 5n + 8) - 1$. This result recovers a formula from Björner and Welker's work on the computational complexity of the k-equal problem, a computer science application.

1 Introduction

An early appearance of a discrete homotopy theory can be found in the work of Atkin [1, 2] in the early 1970s. A physicist modeling social networks using simplicial complexes, Atkin developed Q-analysis, a discrete topological theory used to measure the combinatorial connectivity of a complex and identify combinatorial "holes" in the complexes. In 1972, Maurer [8] developed a similar concept of discrete deformation of paths in graphs while working on his dissertation, developing a characterization of basis graphs of matroids. In 1983, Malle [7] also defined a notion of equivalence of graph maps, as well as discrete fundamental group. These authors were apparently unaware of each other's work, but in fact the concepts they created are all equivalent. More recently, Laubenbacher and Kramer [6] became aware of Atkin's work while conducting research in social and communications networks. With Barcelo and Weaver [3], they pursued Atkin's ideas in Q-analysis and extended them to include graphs and discrete analogues to higher homotopy groups. They also named their work A-theory in his honor.

Let $\Gamma = (V, E)$ and $\Gamma' = (V', E')$ be *simple* graphs, with no loops or parallel edges. A *graph map* $f : \Gamma \to \Gamma'$ is a set map $V \to V'$ that preserves adjacency, that is, if $vw \in E$, then either $f(v)$ is adjacent to $f(w)$ in Γ', denoted by $f(v) \sim_{\Gamma'} f(w)$, or $f(v) = f(w)$. Let $v \in V$ and $v' \in V'$ be distinguished vertices. A *based graph map* is a graph map $f : (\Gamma, v) \to (\Gamma', v')$ such that $f(v) = v'$. The *box product* $\Gamma \square \Gamma'$ of two graphs, Γ and Γ', is the graph with vertex set $V \times V'$ and an edge between (v, v') and (w, w') if either

1. $v = w$ and $v' \sim_{\Gamma'} w'$, or
2. $v' = w'$ and $v \sim_{\Gamma} w$.

Let I_m be the path on $m + 1$ vertices, with vertices labeled from 0 to m. The boundary, ∂I_m, is the set of vertices 0 and m. This path has the same role as that of the unit interval in classical homotopy theory.

Definition 1.1. [3]
Let $\Gamma = (V, E)$ and $\Gamma' = (V', E')$ be simple graphs with distinguished vertices $v_0, v_1 \in V$ and $v_0', v_1' \in V'$. Let f and g be based graph maps $\Gamma \to \Gamma'$ such that $f(v_0) = g(v_0) = v_0'$ and $f(v_1) = g(v_1) = v_1'$. We say that f and g are G-homotopic relative to v_0' and v_1', denoted by $f \simeq_G g \quad rel(v_0', v_1')$ if there is an integer n and a graph map $F : \Gamma \square I_n \to \Gamma'$ that discretely deforms f into g, specifically

1. $F(v, 0) = f(v) \qquad \forall\, v \in V$
2. $F(v, n) = g(v) \qquad \forall\, v \in V$
3. $F(v_0, j) = v_0' \qquad 0 \leq j \leq n$
4. $F(v_1, j) = v_1' \qquad 0 \leq j \leq n.$

If $v_0' = v_1'$, then we write $f \simeq_G g \quad rel(v_0')$, or simply $f \simeq_G g$ of the base vertex is clear.

While G-homotopy is defined for graph maps in general, for the remainder of this discussion, we will limit our investigation to graph maps defined on the discrete interval I_m. If a based graph map $f : (I_m, \partial I_m) \to (\Gamma, v)$ sends ∂I_m to the base vertex v in Γ, then the image of f is a *string loop* in Γ, or simply a *loop*, based at v. Furthermore, we can "stretch" a graph map $f : (I_m, \partial I_m) \to (\Gamma, v)$ to define it on a larger discrete interval by sending vertices with labels $> m$ to v. We can view these graph maps as being defined on $\mathbb{Z}$, with only finitely many images not equal to v, so we may drop the subscript m.

$$f : (I_m, \partial I_m) \to (\Gamma, v) \simeq_G \tilde{f} : (I_p, \partial I_p) \to (\Gamma, v), \quad p > m$$

Multiplication of graph maps f and g is equivalent to the concatenation of the loops corresponding to the maps. Furthermore, G-homotopy is an equivalence relation on the set of based graph maps from the discrete interval I to Γ. Barcelo, Kramer, Laubenbacher, and Weaver [3] showed that these equivalence classes, with multiplication, form a group, denoted by $A_1^G(\Gamma, v)$, and referred to simply as the A_1-group of Γ. As in classical topology, if Γ is connected, the discrete fundamental group $A_1^G(\Gamma, v)$ is independent of the choice of base vertex.

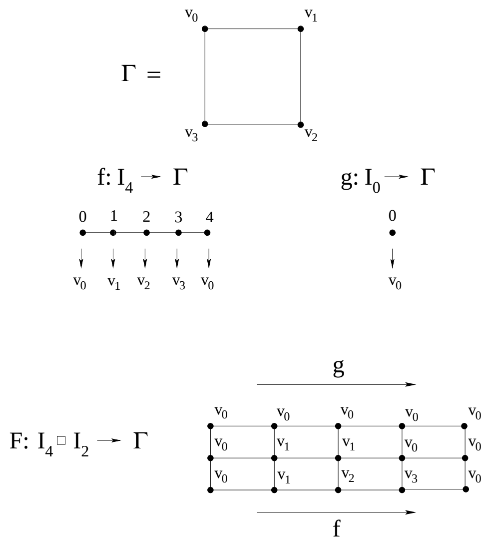


Figure 1: A G-homotopy from f to g.

Figure 1 shows an example of two G-homotopic graph maps; the image of f is the 4-cycle Γ, and the image of g is the single vertex v_0. The vertices of the graph (grid) $I_4 \Box I_2$ are labeled with the image of a G-homotopy from f to g, where g has been stretched so that it is also defined on I_4. The G-homotopy F is itself a graph map which must preserve adjacency, thus each edge in the grid corresponds to an edge or a single vertex in Γ.

Furthermore, is is straightforward to show that any based graph map from the discrete interval to the 4-cycle is G-homotopic to the constant map g, so the A_1 group of the 4-cycle, and similarly the 3-cycle, is trivial. Barcelo et al. [3] also showed that $A_1^G(\Gamma, v) \simeq \pi_1(\Gamma, v)/N$, where $\pi_1(\Gamma, v)$ is the classical fundamental group of Γ when considered as a 1-dimensional simplicial complex and N is the normal subgroup generated by 3-and 4-cycles. Thus, computing the A_1 group of a graph is equivalent to attaching 2-cells to the 3- and 4-cycles of the graph and computing the classical fundamental group of the resulting topological space.

2 Constructing the Graph for the Boolean Lattice

There is an equivalent definition of discrete homotopy for simplicial complexes that we will use in order to compute the discrete fundamental group of the order complex of B_n, the Boolean lattice. This definition includes a graded version of the discrete fundamental group, related to the dimension of the intersection of simplices in a simplicial complex. This complete definition can be found in [3], however, here we will only be concerned with the highest of these groups. In general, to compute the discrete fundamental group of a simplicial complex Δ, we first construct a graph $\Gamma(\Delta)$ and then we compute $A_1^G(\Gamma(\Delta))$.

The simplicial complex we will consider here is the order complex of B_n. The i-faces of $\Delta(B_n)$ correspond to the i-chains of B_n. When we construct the graph, $\Gamma((\Delta(B_n))$, or simply $\Gamma(B_n)$, associated with the highest of the discrete fundamental groups, the vertices of the graph correspond to the maximal faces of $\Delta(B_n)$. These maximal faces are the maximal chains in B_n after the $\hat{0}$ and $\hat{1}$ are removed, or equivalently, permutations in S_n. Two vertices in $\Gamma(B_n)$ are adjacent if the two chains in B_n differ in precisely one element. In this case, the associated permutations in S_n differ by multiplication on the right by a simple transposition $(i, i+1)$, for some $1 \leq i \leq n-1$.

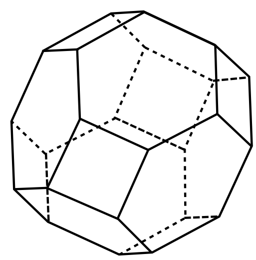


Figure 2: The 1-skeleton of the permutahedron P_3.

We note that $\Gamma(B_n)$ is the 1-skeleton of the permutahedron P_{n-1} [11], and we see the graph for $n = 4$ in Figure 2. We can see that if we attach 2-cells to the 4-cycles in $\Gamma(B_4)$, we are left with 6-cycles. Thus, in order to compute the rank $A_1^G(\Gamma(B_n))^{ab}$, the abelianization of the A_1-group, we are looking for a way to define and count equivalence classes of based graph maps defined on the discrete interval whose images are 6-cycles in $\Gamma(B_n)$. Unfortunately, it is not easy to see a relationship just by looking at the permutahedron.

The breakthrough that allows us to understand the G-homotopy relation on $\Gamma(B_n)$ is the simple observation that B_n is isomorphic to the direct product $B_{n-1} \times B_1$. We will use this observation to define a method for constructing $\Gamma(B_n)$ that will make it easier for us to compute the rank of $A_1^G(\Gamma(B_n))^{ab}$. The graph $\Gamma(B_n)$ is not isomorphic to $\Gamma(B_{n-1})\Box\Gamma(B_1)$ because a maximal chain in B_n corresponds to a shuffle of the edges of a maximal chain in B_{n-1} with the single edge from B_1. These edges can be shuffled in more than one way, so the product of the graphs for the smaller lattices will not have enough vertices.

To solve this problem, we introduce the *shuffle graph*. The vertices of the shuffle graph $\Gamma_{shuffle}^{n-1,1}$ correspond to shuffles of a maximal chain in B_{n-1} with the single edge from B_1. Two vertices are adjacent if the shuffles differ by a switch of two consecutive edges, one from B_{n-1} and the other from B_1. We note that $\Gamma(B_1)$ is a single vertex and $\Gamma_{shuffle}^{n-1,1}$ is a path on n vertices. We use this shuffle graph in the construction of another graph, $\widetilde{\Gamma}(B_n)$:

$$\mathbf{\widetilde{\Gamma}(B_n) = \Gamma(B_{n-1})\Box\Gamma(B_1)\Box\Gamma^{n-1,1}_{shuffle}}$$

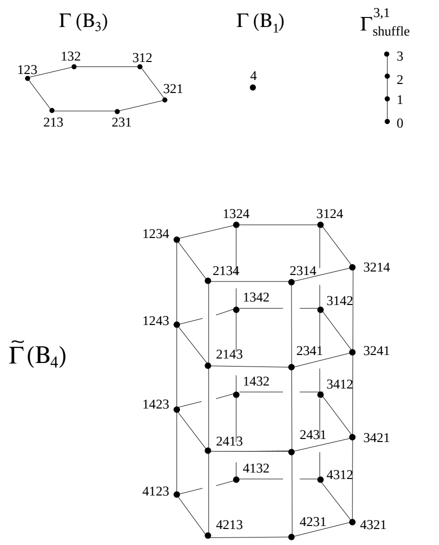


Figure 3: The intermediate graph $\widetilde{\Gamma}(B_4)$.

The vertices in $\widetilde{\Gamma}(B_n)$, the box product of three graphs, are ordered triples. The first coordinate is a permutation in S_{n-1} corresponding to a maximal chain in B_{n-1}. The second coordinate is the element n, corresponding to the single edge in B_1. The third coordinate is an integer i, $0 \leq i \leq n-1$, and it uniquely defines the shuffle of the two chains by indicating how many edges of the chain from B_{n-1} are below the edge from B_1 in the resulting chain. Two vertices in $\widetilde{\Gamma}(B_n)$ are adjacent if either the chains in B_{n-1} are the same and the shuffles are adjacent in $\Gamma^{n-1,1}_{shuffle}$, or the chains differ in one element and the shuffles are the same.

The graph $\widetilde{\Gamma}(B_n)$ now has the right number of vertices, but there are too many edges because some edges are incident to pairs of chains in B_n that differ in more than one element. This problem, however, is easily solved by removing a well-defined set of edges from $\widetilde{\Gamma}(B_n)$ to obtain the desired graph $\Gamma(B_n)$ [9]; we can use $\Gamma(B_4)$ to illustrate the following properties of $\Gamma(B_n)$.

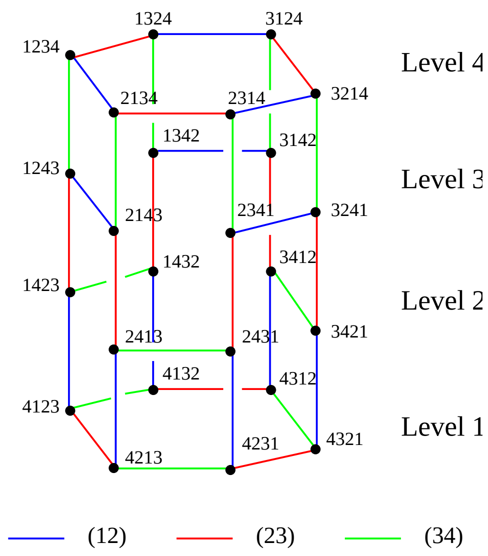


Figure 4: The final graph $\Gamma(B_4)$.

1. **Vertices.** We label the vertices with permutations in S_n, written in single line notation.

2. **Edges.** Each edge corresponds to a simple transposition. The graph is bipartite and $(n-1)$-regular, with each vertex incident to precisely one edge for each of the $n-1$ simple transpositions in S_n.

3. **Bipartite.** The graph is bipartite, with vertices partitioned into even and odd permutations, thus all cycles in the graph are of even length.

4. **Cycles.** The set of transpositions labeling the edges of a cycle correspond to a representation of the identity in S_n. Each 4-cycle in the graph correspond to a pair of disjoint transpositions such as (12) and (34). Each 6-cycle corresponds to a pair of consecutive simple transpositions, such as (12) and (23). All other cycles of length ≥ 8 can be expressed as the concatenation of 4- and 6-cycles. Therefore, we can limit our investigation to 6-cycles which are not the concatenation of 4-cycles. We want to count the G-homotopy equivalence classes of 6-cycles in B_n, which form a basis for $A_1^G(\Gamma(B_n))^{ab}$.

5. **Levels.** The may view the graph as having n levels; each level was initially a copy of $\Gamma(B_{n-1})$ before we removed edges from $\widetilde{\Gamma}(B_n)$. All vertices in a single level resulted from the use of the same shuffle, so all permutations in level i have the element n as entry i when the permutations are written in single line notation. We can classify each edge in the graph as *horizontal* if it is incident to two vertices within the same level of $\Gamma(B_n)$, or *vertical* if it is incident to vertices in consecutive levels of the graph. We note that all vertical edges between two consecutive levels correspond to the same transposition. We can similarly define horizontal and vertical 6-cycles. All vertices in a horizontal cycle are in the same level. A vertical 6-cycle contains two vertices in each of three consecutive levels. We identify each vertical 6-cycle with the middle of the three levels. For example, 1243-2143-2413-4213-4123-1423 is a vertical 6-cycle at level 2 in $\Gamma(B_4)$, and its edges correspond to the transpositions (12) and (23).

3 Equivalence Classes

In this section, we describe how to distinguish between different G-homotopy equivalence classes so that we may count them. The proof of both Lemma 3.1 and Theorem 3.2 rely heavily on checking many possible cases of the labelings of G-homotopy grids, the precise details of which we will not go into here, but we give a brief outline of each proof to capture the flavor of the argument. The complete details of the proofs are contained in [9].

Let C_1 and C_2 be two distinct 6-cycles in $\Gamma(B_n)$, and suppose that the edges of C_1 are associated with the transpositions $(i-1, i)$ and $(i, i+1)$ for some i, $2 \leq i \leq n-1$. If $C_1 \simeq_G C_2$, then, as in our example in Figure 1, we must be able to construct a G-homotopy grid so that the image of the first row of the grid is C_1 and the image of the last row is C_2. Recall that a G-homotopy is itself a graph map and must preserve adjacency. When we consider the various changes that we can make from row to row in the grid that will preserve adjacency, we find that they will also preserve the parity of the number edges in each row that are associated with $(i-1,i)$ and $(i,i+1)$. In particular, the last row must also contain an odd number of edges associated with each of $(i-1,i)$ and $(i,i+1)$, and consequently the edges of C_2 are also associated with this same pair of simple transpositions. This leads us to our initial description of equivalence classes of 6-cycles.

Lemma 3.1. *Let C_1 and C_2 be two distinct 6-cycles in $\Gamma(B_n)$. If $C_1 \simeq_G C_2$, then they are associated with the same pair of transpositions, $(i-1,i)$ and $(i,i+1)$, for some i, $2 \leq i \leq n-1$.*

Association with the same pair of transpositions is a necessary condition for G-homotopy of 6-cycles, but it turns out not to be sufficient. In order to guarantee that two 6-cycles, C_1 and C_2, are G-homotopic, they must also differ by a sequence of simple transpositions, $\tau_1\tau_2...\tau_k$, where each τ_j is disjoint from $(i-1,i)$ and $(i,i+1)$. That is, if we multiply each of the six permutations in C_1 by the same sequence $\tau_1\tau_2\cdots\tau_k$, the result is precisely the six permutations in C_2. To indicate this relationship, we write $C_2 = C_1\tau_1\tau_2\cdots\tau_k$.

Theorem 3.2. *Let C_1 and C_2 be two distinct 6-cycles in $\Gamma(B_n)$. Then $C_1 \simeq_G C_2$ iff there exists an integer $k \geq 1$ such that $C_2 = C_1\tau_1\ldots\tau_k$ where C_1 and C_2 are both associated with $(i-1,i)$ and $(i,i+1)$ for some i, $2 \leq i \leq n-1$, and the τ_j are simple transpositions in S_n that are disjoint from $(i-1,i)$ and $(i,i+1)$.*

Proof sketch. The first part of the proof is constructive: assuming $C_2 = C_1\tau_1\ldots\tau_k$, we construct a G-homotopy from C_1 to C_2 whose image is a sequence of 6-cycles connected by 4-cycles. Figure 5 is the image of a such a G-homotopy from C_1 to $C_2 = C_1\tau_1\tau_2\tau_3$.

In the second part of the proof we assume $C_1 \simeq_G C_2$, which means there is a path P such that $C_1PC_2^{-1}P^{-1} \simeq_G \sigma$, where σ is a permutation in C_1. The edges of P correspond to simple transpositions, and the product of these transpositions is a permutation in S_n. Therefore we must be able to construct another valid G-homotopy grid, this time with the first row corresponding to $C_1PC_2^{-1}P^{-1}$ and the last row corresponding to the single vertex σ. By again performing a check of all possible cases, we can show that

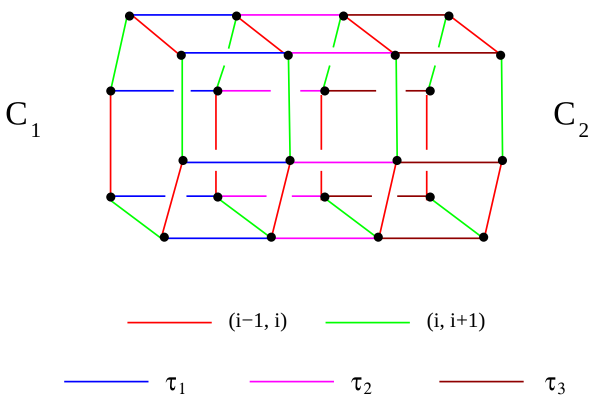


Figure 5: A G-homotopy from C_1 to C_2.

this permutation can be written using only transpositions that are disjoint from the pair $(i-1, i)$ and $(i, i+1)$ associated to C_1 and C_2.

□

Theorem 3.2 stems from the definition of a G-homotopy from C_1 to C_2, and the limitations on the types of changes we are able to make from row to row of the G-homotopy grid. We can combine this theorem with our new understanding of the structure of $\Gamma(B_n)$ to make the following observations about equivalence classes:

1. **Horizontal and vertical 6-cycles are in different equivalence classes.** All permutations in a single horizontal 6-cycle have the element n in the same position when they are written in single-line notation because they result from the use of the same shuffle. In a vertical 6-cycle, the element n will be in different positions in the permutations, depending on which of three consecutive levels each permutation is in. Consequently a horizontal 6-cycle cannot be G-homotopic to a vertical 6-cycles, so we may count horizontal and vertical equivalence classes separately.

2. **Counting Horizontal Equivalence Classes.** We can count the horizontal equivalence classes in level n, which remains a copy of $\Gamma(B_{n-1})$ even after we removed edges in the construction of $\Gamma(B_n)$. We can show that a horizontal 6-cycle in another level of $\Gamma(B_n)$ is G-homotopic to the concatenation of a 6-cycle in level n with vertical 4- and 6-cycles,

and will therefore be contained in the product of a horizontal equivalence class counted at level n with vertical equivalence classes described below.

3. **Vertical 6-cycles at different levels of Γ_{B_n} are in different equivalence classes.** This is a direct consequence of Lemma 3.1 and the observation we made in Section 2 that a vertical 6-cycle at level i, $2 \leq i \leq n-1$ is associated with transpositions $(i-1, i)$ and $(i, i+1)$.

4. **There are $\binom{n-1}{i}\binom{i}{2}$ equivalence classes at level i of Γ_{B_n}, $2 \leq i \leq n-1$.** We can count the number of vertical equivalence classes at level i by using the order of the subgroup of S_n generated by transpositions disjoint from $(i-1, i)$ and $(i, i+1)$ to determine the number of 6-cycles in each equivalence class.

Using standard techniques [10] to count the equivalence classes described above gives us a total of $2^{n-3}(n^2 - 5n + 8) - 1$ classes. Each cycle of length ≥ 8 is contained in the product of one or more of these classes, thus the collection of equivalence classes forms a basis for $A_1^G(\Gamma(B_n))^{ab}$.

4 Related Questions

In the beginning of Section 2, we noted that computing the A_1 group of $\Gamma(B_n)$ is equivalent to attaching 2-cells to the 4-cycles of the graph and computing the classical fundamental group of the resulting 2-dimensional topological space. Eric Babson noted in 2001 that attaching 2-cells to the 4-cycles in $\Gamma(B_n)$ results in a topological space that is homotopy equivalent to the complement (in $\mathbb{R}^n$) of the 3-equal hyperplane arrangement. The k-equal hyperplane arrangements have been extensively studied, and it turns out that these two problems are in fact related (for more details see [4]). Our computation of the rank of $A_1^G(\Gamma(B_n))^{ab}$ recovers a formula of Björner and Welker [5] for the dimensions of the homology groups of the 3-equal arrangements.

The definition of the shuffle graph $\Gamma^{n-1,1}_{shuffle}$ can be generalized to a graph $\Gamma^{k,l}_{shuffle}$ and used to construct the graph associated with the $\Delta(L_1 \times L_2)$, the order complex of the direct product of finite ranked lattices of rank k and l, respectively. The arguments in the computation of the rank of $A_1^G(\Gamma(B_n))^{ab}$ depended heavily on the structure of S_n in determining what changes are permissible from row to row in a valid G-homotopy grid. Nevertheless, using the construction described in Section 2 to build $\Gamma(\Delta(L))$ from smaller graphs may prove useful in obtaining results for lattices other than B_n.

References

[1] R. Atkin, *An Algebra for Patterns on a Complex, I*, Int. J. Man-Machine Studies, **6** (1974), 285-307.

[2] R. Atkin, *An Algebra for Patterns on a Complex, II*, Int. J. Man-Machine Studies, (1976), 448-483.

[3] H. Barcelo, X. Kramer, R. Laubenbacher, and C. Weaver, *Foundations of a Connectivity Theory for Simplicial Complexes*, Adv. Appl. Math., **26** (2001), 97-128.

[4] H. Barcelo and R. Laubenbacher, *Perspectives on A-Homotopy and its Applications*, to appear in Journal of Discrete Mathematics, 2004.

[5] A. Björner and V. Welker, *The Homology of "k-Equal" Manifolds and Related Partition Lattices*, Adv. Math., **110** (1995), 277-313.

[6] X. Kramer and R. Laubenbacher, *Combinatorial Homotopy of Simplicial Complexes and Complex Information Networks*, Applications of Computational Algebraic Geometry (D. Cox and B. Sturmfels, eds.), Proc. Sympos. in Appl. Math., American Mathematical Society, vol. 53, Providence, RI, 1998.

[7] G. Malle, *A Homotopy Theory for Graphs*, Glasnik Matematicki, **18** (1983), 3-25.

[8] S. Maurer, *Matroid Basis Graphs. I*, J. Combin. Theory (B), **14** (1973), 216-240.

[9] S. Smith, *A Discrete Homotopy Theory for Graphs, with Applications to Order Complexes of Lattices*, Ph.D. thesis, Arizona State University, http://faculty.gvsu.edu/smithshe, 2004.

[10] H. Wilf, *Generatingfunctionality*, Academic Press, Inc., San Diego, CA, 1994.

[11] G. Zeigler, *Lectures on Polytopes*, Springer-Verlag, New York, 1995.

On the Littlewood-Richardson rule applying

I.Sviridova
Department of Mathematics, Ulyanovsk State University,
Ulyanovsk,432970, Russia
sviridovaiyu@sv.ulsu.ru

November 20, 2004.

Introduction

The subject we will consider for the most part concerns the combinatorics of Young diagrams and Young tableaus. As a combinatorial object Young diagrams have a wide application in various fields of mathematics, in particular in the representation theory of symmetric groups ([10]) which is also actively used. One of the field of applying of the symmetric group representation theory is the PI-theory (see [4], [9], [12], [13]).

We will consider associative algebras over a field of zero characteristic. Let $F\langle X\rangle$ be the free associative algebra over a field F of zero characteristic with a countable set of generators $X = \{x_1, x_2, \dots\}$. Let $f(x_1, \dots, x_n) \in F\langle X\rangle$ be any associative noncommutative polynomial on variables $x_1, \dots, x_n$. They say an associative algebra A over a field F satisfies the polynomial identity $f(x_1, \dots, x_n) \equiv 0$ if $f(a_1, \dots, a_n) = 0$ holds in A for any $a_i \in A$. The algebra satisfying some nontrivial polynomial identity is called a PI-algebra. For example, a commutative associative algebra is a PI-algebra because it satisfies the identity $xy - yx \equiv 0$, also a nilpotent algebra is PI, it satisfies the identity $x_1 \cdots x_n \equiv 0$ for some natural n. It is well known ([4], [9], [12], [13]) all polynomial identities of an associative PI-algebra A form a T-ideal of the algebra $F\langle X\rangle$ (i.e., an ideal invariant under all endomorphisms of $F\langle X\rangle$). We will denote by $\Gamma = T[A]$ the T-ideal of polynomial identities of A and by $Var(A)$ the variety of all associative algebras over the field F satisfying all polynomial identities of the algebra A.

*The work is partially supported by grants RFFI 04-01-00739

Let us consider the multilinear part $P_n(A) = P_n/(P_n \bigcap T[A])$ of the relatively free algebra $F\langle X\rangle/T[A]$ which is left FS_n- module ([4], [13]). Here $P_n = \langle x_{\sigma(1)} \cdots x_{\sigma(n)} | \sigma \in S_n\rangle$. The S_n-character of $P_n(A)$ $\chi_n(A) = \sum_{\lambda \vdash n} m_\lambda \chi_\lambda$ is called the n-th cocharacter of A (respectively of T[A] or of Var(A)). We will consider the multiplicity series for the algebra A

$$f_A(t_1, t_2, \dots) = \sum_{\lambda=(\lambda_1,\dots,\lambda_k)} m_\lambda t_1^{\lambda_1} \cdots t_k^{\lambda_k}.$$

It is well known ([1]) in the case of a finitely generated algebra the height of Young diagrams in the cocharacter decomposition formula is restricted. It means in this case the number of variables t_i of the multiplicity series f_A is finite. We will consider only this case to make the presentation of the methods simpler.

Let A_1 and A_2 be any finitely generated PI-algebras over a field F of zero characteristic, $\Gamma_1 = T[A_1]$ and $\Gamma_2 = T[A_2]$ be correspondingly their ideals of polynomial identities. Let us consider the multiplicity series for these algebras

$$f^{(1)}(t_1, \dots, t_k) = \sum_\lambda m_\lambda^{(1)} t_1^{\lambda_1} \cdots t_k^{\lambda_k} \text{ - the multiplicity series}$$

for the algebra A_1 and the T-ideal Γ_1,

$$f^{(2)}(t_1, \dots, t_r) = \sum_\lambda m_\lambda^{(2)} t_1^{\lambda_1} \cdots t_r^{\lambda_r} \text{ - the multiplicity series}$$

for the algebra A_2 and the T-ideal Γ_2.

Let us denote

$$\eta(\Gamma_1, \Gamma_2) = \sum_{n\geq 0} \sum_{i=0}^{n} \chi_i(\Gamma_1) \hat{\otimes} \chi_{n-i}(\Gamma_2) = \sum_{i\geq 0, j\geq 0} \chi_i(\Gamma_1) \hat{\otimes} \chi_j(\Gamma_2).$$

Here $\chi_i(\Gamma_1)\hat{\otimes}\chi_j(\Gamma_2) = (\chi_i(\Gamma_1)\otimes\chi_j(\Gamma_2)) \uparrow^{S_{i+j}}$ is the outer product of characters and can be computed by the Littlewood-Richardson rule ([10, 11]). We will present an algorithm counting the multiplicity series $f_{\eta(\Gamma_1,\Gamma_2)}(t_1, \dots, t_{k+r})$ for the character $\eta(\Gamma_1, \Gamma_2)$ if the multiplicity series $f^{(1)}$ and $f^{(2)}$ for the T-ideals Γ_1 and Γ_2 are given. The similar formulas for the case of two-variable multiplicity series were introduced and applied in ([5, 6]).

1 The basic transformations.

Let us consider five basic transformations of multiplicity series used in the algorithm.

Let $f(t_1,\dots,t_k)$ be any function depending on k variables $t_1,\dots,t_k$ (notice, it can also depend on another variables). To be short we will use sometimes for a set of variables $t_1,\dots,t_k$ a notation $(t)_k$.

1. $D^{(1)}_{(t)_k\to(z)_k}(f)=\hat f(t_1,\dots,t_k,z_1,\dots,z_k)=f(t_1,\dots,t_k)|_{t_i=t_i\cdot z_i,\ i=\overline{1,k}}$;

2. $D^{(2)}_{(t)_k\to(z)_k}(f)=\tilde f(z_1,\dots,z_k)=f(t_1,\dots,t_k)|_{t_i=z_i/z_{i-1},z_0=1,\ i=\overline{1,k}}$;

3. $D^{(3)}_{(t)_k\to(z)_k}(f)=\bar f(z_1,\dots,z_k)=f(t_1,\dots,t_k)|_{t_i=z_i\cdots z_k,\ i=\overline{1,k}}$;

4. $D^{(4)}_{(t)_k\to(z)_{k+1}}(f)=\breve f(z_1,\dots,z_{k+1})=$
$$\frac{\left.\sum\limits_{\substack{1\le i_1<\dots<i_m\le k\\0\le m\le k}}(-1)^m(z_{i_1+1}\cdots z_{i_m+1})f((t)_k)\right|_{t_j=\begin{cases}1, j\notin\{i_1,\dots,i_m\}\\ z_{j+1}, j\in\{i_1,\dots,i_m\}\end{cases}}}{(1-z_2)\cdots(1-z_{k+1})},$$

5. $D^{(5)}_{(s)_k,(z)_k}(f)=\frac{1}{(2\pi)^k}\int\limits_0^{2\pi}\cdots\int\limits_0^{2\pi}f((s)_k;(z)_k)\Big|_{\substack{s_j=2e^{i\varphi_j},\\ z_j=\frac12 e^{-i\varphi_j}}}d\varphi_1\dots d\varphi_k.$

Here and later for transformations the superscript enumerates the transformation and the subscript determines the set of variables which are modified.We will omit the subscripts if the sets of variables are not essential for the understanding of a matter.

Examples.
Let us consider $f(t_1,t_2,s_1,s_2,s_3)=t_1t_2^2+t_2s_3-\frac{s_1t_2+2t_1}{1-t_1s_1s_2}$,
$g(t_1,t_2,t_3,s_1,s_2,s_3,p_1)=p_1t_2t_3s_2s_3-2p_1t_2t_3s_1$. Then

1. $f_1(t_1,t_2,s_1,s_2,s_3,p_1,p_2)=D^{(1)}_{(t)_2\to(p)_2}(f)=f|_{\substack{t_1=t_1\cdot p_1,\\ t_2=t_2\cdot p_2}}=t_1p_1(t_2p_2)^2+t_2p_2s_3-\frac{s_1t_2p_2+2t_1p_1}{1-t_1p_1s_1s_2}$;

2. $f_2(t_1,t_2,t_3)=D^{(2)}_{(s)_3\to(t)_3}(f)=f(t_1,t_2,s_1,s_2)|_{\substack{s_1=t_1,s_2=t_2/t_1,\\ s_3=t_3/t_2}}=t_1t_2^2+t_2(t_3/t_2)-\frac{t_1t_2+2t_1}{1-t_1^2(t_2/t_1)}=t_1t_2^2+t_3-\frac{t_1t_2+2t_1}{1-t_1t_2}$;

3. $f_3(t_1,t_2,z_1,z_2,z_3) = D^{(3)}_{(s)_3\to(z)_3}(f) = f(t_1,t_2,s_1,s_2,s_3)\big|_{\substack{s_1=z_1z_2z_3,\\ s_2=z_2z_3,s_3=z_3}} = t_1t_2^2 + t_2z_3 - \frac{z_1z_2z_3t_2+2t_1}{1-t_1z_1z_2^2z_3^2}$;

4. $g_1((t)_3,(z)_4,p_1) = D^{(4)}_{(s)_3\to(z)_4}(g) = \frac{1}{(1-z_2)\cdot(1-z_3)\cdot(1-z_4)}\big(g|_{s_1=s_2=s_3=1} - z_2\cdot g\big|_{\substack{s_1=z_2,\\ s_2=s_3=1}} - z_3\cdot g\big|_{\substack{s_2=z_3,\\ s_1=s_3=1}} - z_4\cdot g\big|_{\substack{s_3=z_4,\\ s_1=s_2=1}} + z_2z_3\cdot g|_{\substack{s_1=z_2,s_2=z_3,\\ s_3=1}} + z_2z_4\cdot g|_{\substack{s_1=z_2,s_3=z_4,\\ s_2=1}} + z_3z_4\cdot g|_{\substack{s_2=z_3,s_3=z_4,\\ s_1=1}} - z_2z_3z_4\cdot g|_{\substack{s_1=z_2,s_2=z_3,\\ s_3=z_4}}\big) = \frac{p_1t_2t_3}{(1-z_2)\cdot(1-z_3)\cdot(1-z_4)}\big(-1-z_2+2z_2^2-z_3^2+2z_3-z_4^2+2z_4+z_2z_3^2-2z_2^2z_3+z_2z_4^2-2z_2^2z_4+z_3^2z_4^2-2z_3z_4-z_2z_3^2z_4^2+2z_2^2z_3z_4\big) = p_1t_2t_3(-1-2z_2+z_3+z_4+z_3z_4)$;

5. $g_2(p_1) = D^{(5)}_{(t)_3,(s)_3}(g) = \frac{1}{(2\pi)^3}\int\limits_0^{2\pi}\int\limits_0^{2\pi}\int\limits_0^{2\pi} g\Bigg|_{\substack{t_j=2e^{i\varphi_j},\\ s_j=\frac{1}{2}e^{-i\varphi_j}}} d\varphi_1 d\varphi_2 d\varphi_3 =$

$\frac{1}{(2\pi)^3}\int\limits_0^{2\pi}\int\limits_0^{2\pi}\int\limits_0^{2\pi}\big(4p_1e^{i\varphi_2}e^{i\varphi_3}\cdot\frac{1}{4}e^{-i\varphi_2}e^{-i\varphi_3} - 2p_1\cdot 4e^{i\varphi_2}e^{i\varphi_3}\cdot\frac{1}{2}e^{-i\varphi_1}\big)d\varphi_1 d\varphi_2 d\varphi_3 = \frac{1}{(2\pi)^3}\int\limits_0^{2\pi}\int\limits_0^{2\pi}\int\limits_0^{2\pi} p_1 d\varphi_1 d\varphi_2 d\varphi_3 - 4p_1\big(\frac{1}{2\pi}\int\limits_0^{2\pi} e^{-i\varphi_1}d\varphi_1\big)\cdot\big(\frac{1}{2\pi}\int\limits_0^{2\pi} e^{i\varphi_2}d\varphi_2\big)\times$

$\times(\frac{1}{2\pi}\int\limits_0^{2\pi} e^{i\varphi_3}d\varphi_3) = p_1.$

The transformations $D^{(1)}$, $D^{(2)}$, $D^{(3)}$ are the simple substitutions of variables, and $D^{(1)}$ adds the new set of variables, while $D^{(2)}$, $D^{(3)}$ exchange the old variables by the new ones. The transformation $D^{(4)}$ also changes variables and their number increase by 1. The transformation $D^{(5)}$ acts on two sets of variables $s_1,\dots,s_k$ and $z_1,\dots,z_k$.

2 The derived transformations.

We also will use some derived transformations.

1. $S^{(1)}_{(t)_k\to(z)_k} = D^{(2)}_{(s)_k\to(z)_k}\circ D^{(1)}_{(t)_k\to(s)_k}$,

2. $h_z(f) = \frac{1}{1-z_1}f$,

3. $S^{(2)}_{(s)_k\to(z,t)_{k+1}} = D^{(1)}_{(z)_{k+1}\to(t)_{k+1}}\circ h_z\circ D^{(4)}_{(s)_k\to(z)_{k+1}}$,

4. $S^{(3)}_{(t)_k\to(z)_k} = D^{(3)}_{(s)_k\to(z)_k}\circ D^{(1)}_{(t)_k\to(s)_k}$.

Here "$\circ$" denotes the usual composition of maps.

Examples.

Let us take $f(s_1, s_2, s_3) = 2s_1 s_2^2 + s_2 s_3$, then

1. $f_1(s_1, s_2, s_3, z_1, z_2, z_3) = S^{(1)}_{(s)_3 \to (z)_3}(f) = f(s_1, s_2, s_3)|_{\substack{s_i = s_i(z_i/z_{i-1}) \\ i=\overline{1,3},\ z_0=1}} =$ $2s_1 z_1 (s_2(z_2/z_1))^2 + s_2(z_2/z_1) s_3(z_3/z_2) = 2s_1 s_2^2 \frac{z_2^2}{z_1} + s_2 s_3 \frac{z_3}{z_1}$;

2. $f_2((z)_4, (t)_4) = S^{(2)}_{(s)_3 \to (z,t)_4}(f(s_1, s_2, s_3)) =$
$$\frac{\sum\limits_{\substack{1 \le i_1 < \cdots < i_m \le 3 \\ 0 \le m \le 3}} (-1)^m \left(\prod\limits_{r=1}^{m} z_{i_r+1} t_{i_r+1}\right) f((s)_3) \Bigg|_{s_j = \begin{cases} 1, j \notin \{i_1, \ldots, i_m\} \\ z_{j+1} t_{j+1}, j \in \{i_1, \ldots, i_m\} \end{cases}}}{(1 - z_1 t_1)(1 - z_2 t_2) \cdots (1 - z_4 t_4)} =$$
$\frac{1}{1 - z_1 t_1}\big(z_4 t_4 (1 + z_3 t_3) + 2z_2 t_2 (1 + z_3 t_3 + z_3^2 t_3^2) + 3 + 3z_3 t_3 + 2z_3^2 t_3^2\big)$;

3. $f_3(s_1, s_2, s_3, z_1, z_2, z_3) = S^{(3)}_{(s)_3 \to (z)_3}(f((s)_3)) = f(s_1, s_2, s_3)|_{\substack{s_i = s_i z_i \cdots z_3 \\ i=\overline{1,3}}} =$ $2(s_1 z_1 z_2 z_3)(s_2 z_2 z_3)^2 + (s_2 z_2 z_3)(s_3 z_3) = 2s_1 s_2^2 z_1 z_2^3 z_3^3 + s_2 s_3 z_2 z_3^2$.

3 The algorithm.

Before counting the multiplicity series $f_{\eta(\Gamma_1, \Gamma_2)}$ we need to modify the first generating function $f^{(1)}(t_1, \ldots, t_k)$.

<u>The 1-st stage.</u>

On the entrance we have a function $F_{11} = f^{(1)}(t_1, \ldots, t_k)$.

1. $F_{12} = S^{(1)}_{(t)_k \to (s)_k}(F_{11})$,

2. $F_{21} = S^{(2)}_{(s)_k \to (\alpha, t)_{k+1}}(F_{12})$.

<u>The j-th stage $(2 \le j \le r)$.</u>

On the entrance we have a function $F_{j1}((t)_{k+j-1}; (\varepsilon)_{j-2}; (\alpha)_{k+j-1})$ (F_{21} does not depend on ε).

1. $F_{j2} = S^{(1)}_{(t)_{k+j-1} \to (s)_{k+j-1}}(F_{j1})$,

2. $F_{j3} = S^{(3)}_{(\alpha)_{k+j-1} \to (y)_{k+j-1}}(F_{j2})$,

3. $F_{j4} = D^{(4)}_{(s)_{k+j-1} \to (p)_{k+j}}(F_{j3})$,

4. $F_{j5} = S^{(3)}_{(p)_{k+j}\to(s)_{k+j}}(F_{j4})$,

5. $F_{j6} = D^{(4)}_{(y)_{k+j-1}\to(z)_{k+j}}(F_{j5})$,

6. $F_{j7} = D^{(5)}_{(s)_{k+j},(z)_{k+j}}(F_{j6})$,

7. $F_{j8} = D^{(1)}_{(p)_{k+j}\to(t)_{k+j}}(F_{j7})$,

8. $F_{j9} = F_{j8}((t)_{k+j};(\varepsilon)_{j-2};\alpha_1,\dots,\alpha_{k+j-1};(p)_{k+j})|_{\alpha_1=\dots=\alpha_{k+j-1}=\varepsilon_{j-1}}$,

9. $F_{j+11}((t)_{k+j};(\varepsilon)_{j-1};(\alpha)_{k+j}) = F_{j9}((t)_{k+j};(\varepsilon)_{j-1};p_1,\dots,p_{k+j})|_{p_i=\alpha_i}$.

Now we can go to the next $(j+1)$-th stage.

When we have finished the last r-th stage and obtain the function $F_{r+11}((t)_{k+r};(\varepsilon)_{r-1},(\alpha)_{k+r})$ we can find the multiplicity series $f_{\eta(\Gamma_1,\Gamma_2)}$.

$$F^*((t)_{k+r};(\varepsilon)_r) = F_{r+11}((t)_{k+r};(\varepsilon)_{r-1};\alpha_1,\dots,\alpha_{k+r})|_{\alpha_1=\dots=\alpha_{k+r}=\varepsilon_r},$$

$$f_{\eta(\Gamma_1,\Gamma_2)}(t_1,\dots,t_{k+r}) = D^{(5)}_{(\varepsilon)_r,(s)_r}\big(F^*((t)_{k+r},(\varepsilon)_r)\cdot f^{(2)}((s)_r)\big). \qquad (1)$$

4 On the rationality of some multiplicity series.

The next statement is obvious.

Lemma 1 *The algebra of rational functions is closed under the basic transformations $D^{(1)}$, $D^{(2)}$, $D^{(3)}$, $D^{(4)}$.*

We will call such transformations by rational transformations.

Corollary 2 *The compositions $S^{(1)}$, $S^{(2)}$, $S^{(3)}$ of rational transformations are also rational.*

Definition 3 *We will call a rational function $f = \frac{P}{Q}$, (P, Q are polynomials) specific on variables $(s)_m$ and $(z)_m$ if the denominator Q has a form $Q = \prod_{j=1}^{d}(1-\omega_j)$, where all ω_j are words on variables and for any $i = 1,\dots,m$ and for all $j = 1,\dots,d$ $\deg_{s_i}\omega_j + \deg_{z_i}\omega_j \le 1$.*

Lemma 4 *The image of the transformation $D^{(5)}_{(s)_m,(z)_m}$ of a rational function specific on variables $(s)_m,(z)_m$ is also a rational function.*

Theorem 5 *If f_{A_1}, f_{A_2} are rational functions specific on all variables then $f_{\eta(\Gamma_1,\Gamma_2)}$ is also a rational function, specific on all variables.*

Evidently it is enough to be sure that all functions on the entrance of the transformation D^5 (on the 6-th step of any stage and in (1)) remain specific on all corresponding variables.

Theorem 6 *If the multiplicity series f_{Γ_1}, f_{Γ_2} for T-ideals Γ_1 and Γ_2 are rational specific functions then the multiplicity series f_Γ for the product $\Gamma = \Gamma_1 \cdot \Gamma_2$ of the T-ideals is a rational specific function.*

Proof The theorem follows from the previous theorem and the Berele and Regev formula for the cocharacter of the product of T-ideals [3]. We can realize this formula for the multiplicity series as follows

$$f_\Gamma((t)_{k+r}) = f_{\Gamma_1} + f_{\Gamma_1} + B(f_{\eta(\Gamma_1,\Gamma_2)}) - f_{\eta(\Gamma_1,\Gamma_2)}, \qquad \text{where}$$
$$\check{f}((t)_m,(s)_m) = S^{(1)}_{(t)_k\to(s)_k}(f((t)_m)), \qquad \text{and}$$
$$B(f((t)_m)) = (\sum_{i=1}^{m+1} t_i) f((t)_m) - \sum_{i=1}^{m} t_{i+1}\cdot \check{f}((t)_m,(s)_m)\Bigg|_{s_j=\begin{cases}1, j\neq i,\\ 0, j=i.\end{cases}}$$

The transformation B here is evidently rational. $\Box$

Theorem 7 *Any minimal variety of associative algebras over a field of zero characteristic of the matrix type not greater than 2 and generated by a finitely generated algebra has a rational multiplicity series.*

Proof By [8] the ideal of polynomial identities Γ of such variety is the product of T-ideals $\Gamma = \prod_{j=1}^{m} \Gamma_j$, where any $\Gamma_j = T[M_2(F)]$ is the ideal of identities of full matrix algebra of the 2-nd order over the base field, or $\Gamma_j = \{[x,y]\}^T$ is the commutator ideal. Then in the first case the multiplicity series for Γ_j can be obtained using the description of multiplicities for 2×2 matrices given by V.Drensky [7]

$$f_{M_2(F)} = \frac{1}{(1-t_1)^2(1-t_1t_2)^2(1-t_1t_2t_3)^2(1-t_1t_2t_3t_4)} - \frac{1}{(1-t_1)^2(1-t_1t_2)} - \frac{t_1t_2t_3 + t_1t_2t_3t_4 - 1}{(1-t_1)}.$$

In the second case the multiplicity series is trivial $f_{\{[x,y]\}^T} = \frac{1}{1-t_1}$. It is obvious in the both cases the multiplicity series are rational specific on all variables functions. Then by the previous theorem the multiplicity series for a product Γ of these T-ideals is also rational. $\square$

Taking into account this result, the results of V.Drensky and G.K.Genov [5, 6] and the rationality of a Hilbert series of any relatively free algebra [2] the question whether any associative PI-algebra over a field of zero characteristic has a rational multiplicity series becomes quite natural.

Notice at the end the presenting algorithm also can be applied for counting of exact formulas for cocharacters of some PI-algebras using some mathematical software.

References

[1] S.A.Amitsur, A.Regev, P.I. algebras and their cocharacters, J. Algebra, 78(1982), 248-254.

[2] A.Ya.Belov On the rationality of Hilberts series of relatively free algebras, Uspechi Mat. Nauk, 52(1997), N.2, 153-154 (russian).

[3] A.Berele, A.Regev, Codimensions of products and of intersections of verbally prime T-ideals, Israel J. Math. 103 (1998), 17-28.

[4] V.Drensky, Free algebras and PI-algebras, Springer-Verlag Singapore, Singapore, 2000.

[5] V.Drensky, G.K.Genov, Multiplicities of Shur functions in invariants of two 3×3 matrices, J. Algebra, 264(2003), 496-519.

[6] V.Drensky, G.K.Genov, Multiplicities of Shur functions with applications to invariant theory and PI-algebras, Comptes rendus de l'Académie bulgare des Sciences, 57(2004), No 3, 5-10.

[7] V.Drensky Polynomial identities for 2×2 matrices, Acta Applicandae Mathematicae, 21(1990), 137-161.

[8] A.Giambruno, M.Zaicev, Codimension growth and minimal superalgebras, Transactions of the AMS 355 (2003), N. 12, 5091-5117.

[9] N.Jacobson, PI-algebras: An introductions, Lecture Notes in Math., 441, Springer-Verlag, Berlin-New York,1975.

[10] G.James, A.Kerber, The representation theory of the symmetric group, Encyclopedia of Mathematics and its Applications, V. 16, Addison-Wesley, London,1981.

[11] I. Macdonald, Symmetric functions and Hall polynomials, Oxford Univ. Press (Clarendon), Oxford, 1979.

[12] C.Procesi, Rings with polynomial identities, Marcel Dekker, New York,1973.

[13] L.H.Rowen, Polinomial identities in ring theory, Academic Press, 1980.

A Non-Messing-Up Phenomenon for Posets

Bridget Eileen Tenner

Abstract. We classify finite posets with a particular sorting property, generalizing a result for rectangular arrays. Each poset is covered by two sets of disjoint saturated chains such that, for any original labeling, after sorting the labels along both sets of chains, the labels of the chains in the first set remain sorted. This gives a linear extension of the poset. We also characterize posets with more restrictive sorting properties.

Résumé. Nous classifions les ensembles partiellement ordonnés ayant une certain propriété de triage, généralisant ainsi un résultat connu pour les tables rectangulaires. Chaque ensemble partiellement ordonné est couvert par deux ensembles de chaînes saturées disjointes de telle sorte que, pour tout étiquetage, trier le long des chaînes du premier ensemble, puis celles du second, produit un étiquetage où les étiquettes sont toujours bien ordonnées par rapport au premier ensemble de chaînes. Nous obtenons de cette façon une extension linéaire de l'ensemble partiellement ordonné. Nous caractérisons aussi les ensembles partiellement ordonnés possédant des propriétés de triage plus contraignantes.

1. Introduction

The so-called Non-Messing-Up Theorem is a well known sorting result for rectangular arrays. In [**5**], Donald E. Knuth attributes the result to Hermann Boerner, who mentions it in a footnote in Chapter V, §5 of [**1**]. Later, David Gale and Richard M. Karp include the phenomenon in [**2**] and in [**3**], where they prove more general results about order preservation in sorting procedures. The first use of the term "non-messing-up" seems to be due to Gale and Karp, as suggested in [**4**]. One statement of the result is as follows.

Theorem 1. *Let $A = (a_{ij})$ be an m-by-n array of real numbers. Put each row of A into non-decreasing order. That is, for each $1 \leq i \leq m$, place the values $\{a_{i1}, \ldots, a_{in}\}$ in non-decreasing order (henceforth denoted* row-sort*). This yields the array $A' = (a'_{ij})$. Column-sort A'. Each row in the resulting array is in non-decreasing order.*

Applying the theorem to the transpose of the array A, the sorting can also be done first in the columns, then in the rows, and the columns remain sorted.

Example.

$$\begin{array}{cccc} 4 & 9 & 7 & 8 \\ 12 & 5 & 1 & 10 \\ 2 & 6 & 11 & 3 \end{array} \xrightarrow{\text{row-sort}} \begin{array}{cccc} 4 & 7 & 8 & 9 \\ 1 & 5 & 10 & 12 \\ 2 & 3 & 6 & 11 \end{array} \xrightarrow{\text{column-sort}} \begin{array}{cccc} 1 & 3 & 6 & 9 \\ 2 & 5 & 8 & 11 \\ 4 & 7 & 10 & 12 \end{array}$$

Answering a question posed by Richard P. Stanley, the author's thesis advisor, this paper defines a notion of non-messing-up for posets and Theorem 7 generalizes Theorem 1 by characterizing all posets with this property.

Key words and phrases. Non-messing-up, partially ordered set, sorting, linear extension.

Throughout this paper, we will use standard terminology from the theory of partially ordered sets. A good reference for these terms and other information about posets is Chapter 3 of [**6**].

The rectangular array in Theorem 1 can be viewed as the poset $\boldsymbol{m} \times \boldsymbol{n}$ (where $\boldsymbol{j}$ denotes a j-element chain). The *rows* and *columns* are two different sets of disjoint saturated chains, each covering this poset. Sorting a chain orders the chain's labels so that the minimum element in the chain has the minimum label. Thus, sorting the labels in this manner gives a linear extension of $\boldsymbol{m} \times \boldsymbol{n}$.

Definition. An *edge* in a poset P is a covering relation $x \lessdot y$. Two elements in P are *adjacent* if there is an edge between them.

Definition. A *chain cover* of a poset P is a set of disjoint saturated chains covering the elements of P.

Definition. A finite poset P has the *non-messing-up* property if there exists an unordered pair of chain covers $\{\mathcal{C}_1, \mathcal{C}_2\}$ such that

(1) For any labeling of the elements of P, $\mathcal{C}_i$-sorting and then $\mathcal{C}_{3-i}$-sorting leaves the labels sorted along the chains of $\mathcal{C}_i$, for $i = 1$ and 2; and
(2) Every edge in P is contained in an element of $\mathcal{C}_1$ or $\mathcal{C}_2$.

The set $\mathcal{N}_2$ consists of all posets with the non-messing-up property, where the subscript indicates that an unordered *pair* of chain covers is required. For a non-messing-up poset P with chain covers as defined, write $P \in \mathcal{N}_2$ *via* $\{\mathcal{C}_1, \mathcal{C}_2\}$.

Let us clarify the difference between this result and Gale and Karp's work in [**2**] and [**3**]. Gale and Karp consider a poset P and a partition F of the elements of P. The elements in each block of F are linearly ordered, not necessarily in relation to comparability in P. Given P and F, the authors determine whether each natural labeling of P, sorted within each block of F, yields a labeling that is still natural. In this paper, we do not require that the original labeling be natural. In fact, it is the labelings that are not natural and that do not become natural after the first sort that determine membership in $\mathcal{N}_2$. Additionally, the partition blocks in $\mathcal{N}_2$ are saturated chains, and every covering relation must be in at least one of these chains. The goal of this paper is to determine, for a given poset, when there *exist* chain covers with the non-messing-up property, not if a given pair of chain covers has the property.

It is important to emphasize that $\{\mathcal{C}_1, \mathcal{C}_2\}$ is an unordered pair and that there is a symmetry between the chain covers. We will refer to elements of $\mathcal{C}_1$ and their edges as *red*, and elements of $\mathcal{C}_2$ and their edges as *blue*. If an edge belongs to both chain covers, it is *doubly colored*. The symmetry between the chain covers may be expressed by a statement about red and blue chains and an indication that a *color reversed* version of the statement is also true.

A central object in the classification of $\mathcal{N}_2$ is the following.

Definition. Let $N \geq 3$ be an integer, and consider the poset $P = \boldsymbol{N} \times \boldsymbol{N} = \{(i,j) : 1 \leq i,j \leq N\}$. For integers k_1 and k_2, $3 \leq k_1 \leq k_2 \leq N$, let P' be

$$P \setminus \Big(\{(i,j) : j \geq i + k_1 \text{ or } i \geq j + k_2\} \cup \{(i,j) : j \geq N + k_1 - k_2 + 1\}\Big).$$

Let the poset $\widehat{P}$ be obtained from P' by identifying $(i, k_1+i-1) \sim (k_2+i-1, i)$ for $i = 1, \ldots, N-k_2+1$.

The poset $\widehat{P}$ is $\boldsymbol{N} \times \boldsymbol{N}$ *on the cylinder*. This definition is independent of the values k_1 and k_2.

The classification in Theorem 7 states that $\mathcal{N}_2$ is the set of disjoint unions of connected posets that each can be "reduced" to a convex subposet of $\boldsymbol{N} \times \boldsymbol{N}$ or of $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder for some N, subject to a technical constraint. Informally speaking, P reduces to Q if P is formed by replacing particular elements of Q with chains of various lengths. Sample Hasse diagrams for elements of $\mathcal{N}_2$ are shown in Figures 5(a), 6(a), 7(a), and 8(a).

In Section 2 of this paper, we address definitions and preliminary results. The definitions describe the objects and operations needed for the classification, and the results will be the fundamental tools

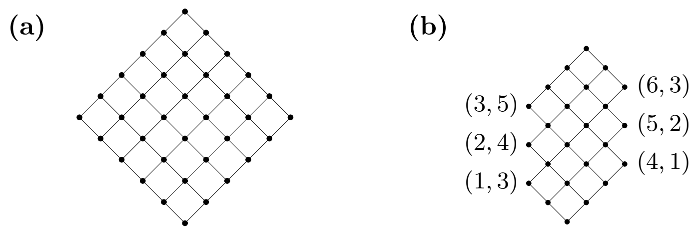


FIGURE 1. (**a**) $P = \mathbf{6} \times \mathbf{6}$. (**b**) P' for $k_1 = 3$ and $k_2 = 4$. To form $\widehat{P}$, identify $(1,3) \sim (4,1)$, $(2,4) \sim (5,2)$ and $(3,5) \sim (6,3)$.

for defining $\mathcal{N}_2$. The main theorem is proved in Section 3 by induction on the size of a connected poset. The final section of the paper discusses further directions for the study of non-messing-up posets, including several open questions.

2. Preliminary results

The definition of a non-messing-up poset requires that every edge be colored. Therefore, as in the case of the product of two chains, $\mathcal{C}_i$-sorting any labeling and then $\mathcal{C}_{3-i}$-sorting yields a linear extension of the poset. The chains of $\mathcal{C}_i$ are disjoint, so each element of a non-messing-up poset is covered by at most two elements, and covers at most two elements.

It is sufficient to consider connected posets, as a poset is in $\mathcal{N}_2$ if and only if each of its connected components is in $\mathcal{N}_2$. Key to determining membership in $\mathcal{N}_2$ is the following fact.

THEOREM 2. *Every convex subposet of an element of $\mathcal{N}_2$ is also in $\mathcal{N}_2$.*

The coloring of a convex subposet Q of $P \in \mathcal{N}_2$ is inherited from the coloring of P in the sense that the chain covers in Q are as in Q when considered as a subposet of P.

LEMMA 3. *If a convex subposet of a non-messing-up poset is a chain, then there is a red chain or a blue chain containing this entire subposet.*

DEFINITION. A *diamond* in a poset is a convex subposet that is the union of distinct (saturated) chains which only intersect at a common minimal element and a common maximal element.

LEMMA 4. *Let Q be a diamond consisting of chains $\boldsymbol{a}$ and $\boldsymbol{b}$ in a non-messing-up poset. Let x be the minimal element in $\boldsymbol{a}$ and $\boldsymbol{b}$, denoted* $\min(\boldsymbol{a})$ *and* $\min(\boldsymbol{b})$*, and let $y = \max(\boldsymbol{a}) = \max(\boldsymbol{b})$, with similar notation. Up to color reversal, one of the following is true (where $\boldsymbol{c} \setminus z$ is taken to mean $\boldsymbol{c} \setminus \{z\}$).*

(1) *There exists a red chain containing $\boldsymbol{a} \setminus y$, a blue chain containing $\boldsymbol{b} \setminus y$, a red chain containing $\boldsymbol{b} \setminus x$ and a blue chain containing $\boldsymbol{a} \setminus x$; or*
(2) *There exists a red chain containing $\boldsymbol{a}$ and a blue chain containing $\boldsymbol{b}$.*

Call the former of these a Type I *diamond and the latter a* Type II *diamond.*

DEFINITION. A diamond with *bottom chain* of length k and *top chain* of length l is a convex subposet that is a diamond with minimum x and maximum y, a chain of k elements covered by x, and a chain of l elements covering y, with no other elements or relations among the elements already mentioned.

The technical condition mentioned in the introduction is due to the following requirement.

LEMMA 5. *Let $Q \subseteq P \in \mathcal{N}_2$ be a diamond consisting of chains $\boldsymbol{a}$ and $\boldsymbol{b}$. Suppose there is a coloring of P for which Q has Type I, with bottom and top chains $\boldsymbol{C}$ and $\boldsymbol{D}$. Then there are chains in that coloring such that, up to color reversal, $(\boldsymbol{C} \cup \boldsymbol{a}) \setminus y$ is red, $(\boldsymbol{C} \cup \boldsymbol{b}) \setminus y$ is blue, $(\boldsymbol{a} \cup \boldsymbol{D}) \setminus x$ is blue, and $(\boldsymbol{b} \cup \boldsymbol{D}) \setminus x$ is red. Also,* $\max\{|\boldsymbol{C}|, |\boldsymbol{D}|\} < \min\{|\boldsymbol{a}| - 2, |\boldsymbol{b}| - 2\}$.

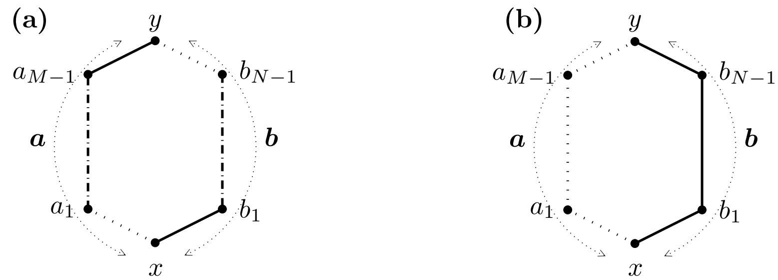


FIGURE 2. (**a**) Type I diamond coloring. (**b**) Type II diamond coloring, where the intervals $\boldsymbol{a} \setminus \{x, y\}$ and $\boldsymbol{b} \setminus \{x, y\}$ may be partially or totally doubly colored.

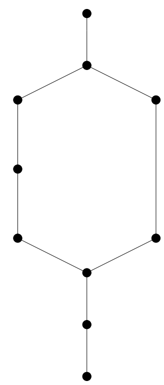


FIGURE 3. A diamond with bottom chain of length 2 and top chain of length 1

If $\max\{|\boldsymbol{C}|, |\boldsymbol{D}|\} = \min\{|\boldsymbol{a}| - 3, |\boldsymbol{b}| - 3\}$, then the described chains have the non-messing-up property, so the bounds in Lemma 5 are sharp.

Recall the definition of $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder. As suggested by the main result, this object is crucial in the study of non-messing-up posets.

THEOREM 6. *The poset $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder is in $\mathcal{N}_2$ for all N. The chain covers for this poset are of the same form as the chain covers in Theorem 1.*

Before discussing the main theorem, it remains to rigorously define the notion of reduction.

DEFINITION. The process of *splitting* the element $x \in Q'$ gives a poset Q where

(1) $x \in Q'$ is replaced by $\{x_1 \lessdot \cdots \lessdot x_{s(x)}\}$ for some positive integer $s(x)$;
(2) All elements and relations in $Q' \setminus x$ are unchanged in Q;
(3) If $y \gtrdot x$ in Q', then $y \gtrdot x_{s(x)}$ in Q; and
(4) If $y \lessdot x$ in Q', then $y \lessdot x_1$ in Q.

If Q is formed by splitting elements of $\widetilde{Q}$, then Q *reduces* to $\widetilde{Q}$, denoted $Q \rightsquigarrow \widetilde{Q}$.

DEFINITION. Let $P \rightsquigarrow \widetilde{P} \in \mathcal{N}_2$. The coloring of $\widetilde{P}$ *induces* the coloring of P if the edge $\widetilde{u} \lessdot \widetilde{v}$ in $\widetilde{P}$ and its image, the edge $u \lessdot v$ in P, are colored in the same way. Edges in the chain into which an element splits get doubly colored.

3. Characterization of $\mathcal{N}_2$

The classification of the set $\mathcal{N}_2$ is done in two steps. The first direction shows that any poset reducing to a convex subposet of $\boldsymbol{N} \times \boldsymbol{N}$ or of $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder, subject to a technical

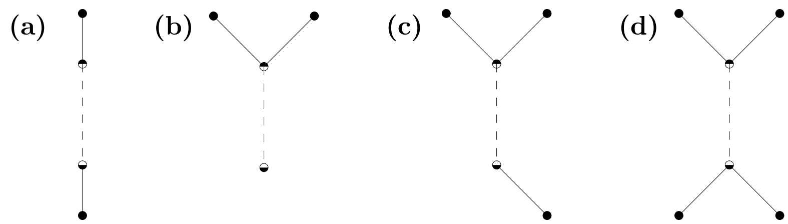


FIGURE 4. How to split a vertex.

constraint imposed by Lemma 5, has the non-messing-up property. The second step shows the reverse inclusion. Both directions are proved by induction on the size of a connected poset.

THEOREM 7. *The collection $\mathcal{N}_2$ is exactly the set of posets each of whose connected components P reduces to $\widetilde{P}$, a convex subposet of $\boldsymbol{N} \times \boldsymbol{N}$ or of $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder for some N, given the following stipulation:*

TECHNICAL CONDITION. *For any diamond $\{w \lessdot x, y \lessdot z\}$ in $\widetilde{P}$ that does not realize a generator of the fundamental group of the cylinder,*

$$\max\{s(w), s(z)\} \leq \min\{s(x), s(y)\}.$$

The required coloring of the connected poset $P \in \mathcal{N}_2$ is induced by the coloring of $\widetilde{P}$, which is inherited from the coloring in Theorem 1 or Theorem 6.

Both directions of the proof consider a subposet P' formed by removing either a maximal or a minimal element from P. Thus P' is convex in P, and it is not hard to see that the suppositions for P must hold for P' as well. Each connected component in P' has fewer than $|P|$ elements, so the theorem holds for P' by the inductive assumption.

One case considered in the proof is when a maximal or minimal element of P is adjacent to two other elements but is not in a diamond, and its removal does not disconnect the poset. Observe that this describes a poset P that can only reduce to a poset on the cylinder, while a maximal proper subposet of P reduces to a convex subposet in the plane.

Examples of posets with the non-messing-up property are depicted in Figures 5(a), 6(a), 7(a), and 8(a). The first two of these reduce to convex subposets of $\boldsymbol{N} \times \boldsymbol{N}$, and the last two reduce to convex subposets of $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder.

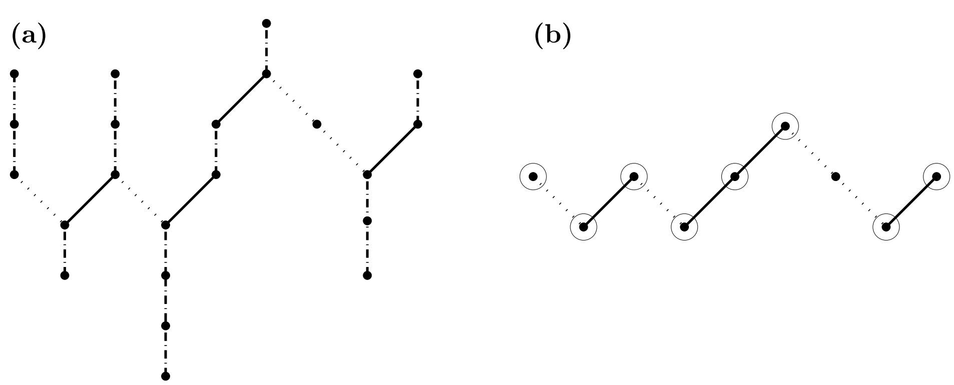


FIGURE 5. (**a**) A poset $P \in \mathcal{N}_2$. (**b**) The reduced poset $\widetilde{P}$, where the elements that split to form P are circled.

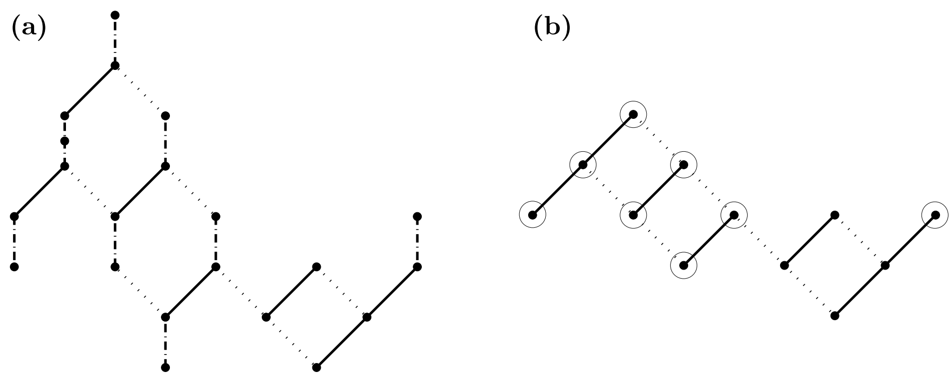


FIGURE 6. (**a**) A poset $P \in \mathcal{N}_2$. (**b**) The reduced poset $\widetilde{P}$, where the elements that split to form P are circled.

Notice that a Type II diamond as described in Lemma 4 occurs only in elements of $\mathcal{N}_2$ that reduce to posets on the cylinder. Moreover, such a diamond must realize a generator of the fundamental group of the cylinder because of the definition of an induced coloring. This explains the technical condition.

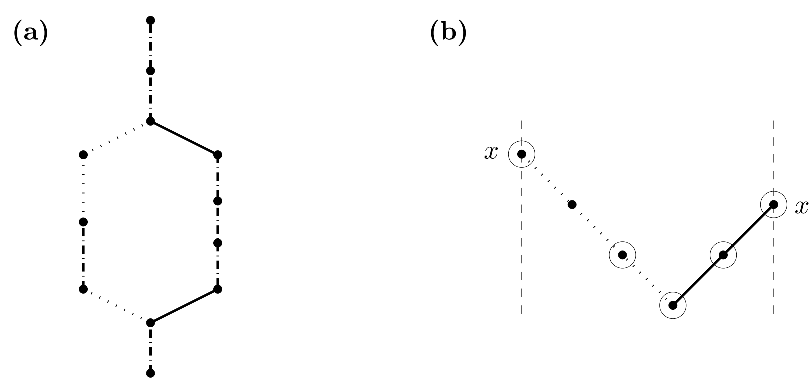


FIGURE 7. (**a**) A poset $P \in \mathcal{N}_2$. (**b**) The reduced poset $\widetilde{P}$ as viewed with identified sides, where the elements that split to form P are circled and the elements that are identified are both labeled x.

The requirement for membership in $\mathcal{N}_2$ is the existence of a pair of chain covers $\{\mathcal{C}_1, \mathcal{C}_2\}$ with particular properties. We might also ask if there are other choices for $\mathcal{C}_i$. A poset of the form depicted in Figure 7(a), that is, a poset consisting of a single diamond and its bottom and top chains, can also be colored so that the diamond has Type I if the bounds of Lemma 5 are satisfied. Otherwise, the only freedom in defining the chain covers arises from the various ways to reduce P due to splits as depicted in Figure 4(a).

4. Further directions

The classification of $\mathcal{N}_2$ prompts further questions relating to the non-messing-up property. In the final section of this paper, we suggest several such questions and provide answers to some.

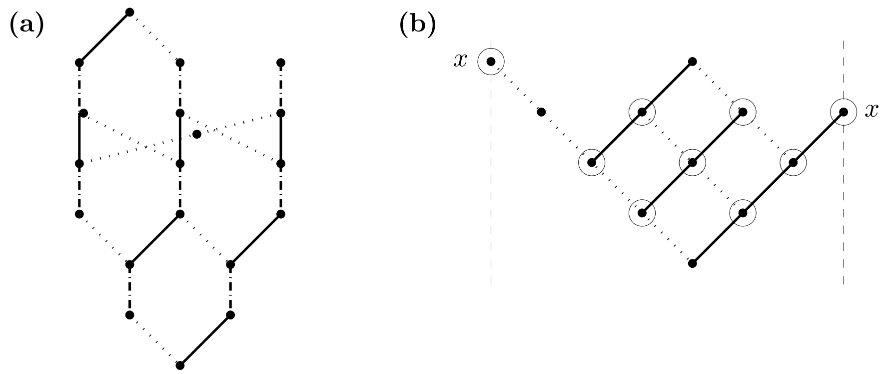


FIGURE 8. (**a**) A poset in $\mathcal{N}_2$. (**b**) The reduced poset $\widetilde{P}$ as viewed with identified sides, where the elements that split to form P are circled and the elements that are identified are both labeled x.

4.1. The set $\mathcal{N}_2{}' \subsetneq \mathcal{N}_2$ with reduced redundancy.
In the classification of $\mathcal{N}_2$, there were instances of a $\mathcal{C}_i$ chain entirely contained in a $\mathcal{C}_{3-i}$ chain. These chain covers have the non-messing-up property, but there is a certain redundancy: this particular $\mathcal{C}_i$ chain adds no information about the relations in the poset since its labels are already ordered after the $\mathcal{C}_{3-i}$-sort.

DEFINITION. The class $\mathcal{N}_2{}'$ consists of all posets $P \in \mathcal{N}_2$ via $\{\mathcal{C}_1, \mathcal{C}_2\}$ such that $\boldsymbol{c}_i \not\subseteq \boldsymbol{c}_{3-i}$ for all $\boldsymbol{c}_1 \in \mathcal{C}_1$ and $\boldsymbol{c}_2 \in \mathcal{C}_2$.

Because the coloring of a non-messing-up poset is induced by its reduced poset, the elements of $\mathcal{N}_2{}'$ can be determined by looking at these reduced posets. Call a chain that shares no covering relation with any diamond a *branch chain* and a maximal such chain a *maximal branch chain*.

THEOREM 8. *The collection $\mathcal{N}_2{}'$ is the set of posets in $\mathcal{N}_2$ where every maximal branch chain in the reduced poset $\widetilde{P}$ consists of exactly two elements, and every element of $\widetilde{P}$ is adjacent to at least two other elements in $\widetilde{P}$.*

4.2. The set $\mathcal{N}_2{}'' \subseteq \mathcal{N}_2$ with reduced redundancy.
In the Non-Messing-Up Theorem as stated in Theorem 1, the rows and columns have minimal redundancy in the sense that for any row $\boldsymbol{r}$ and any column $\boldsymbol{c}$, $\#(\boldsymbol{r} \cap \boldsymbol{c}) = 1$.

DEFINITION. The class $\mathcal{N}_2{}''$ consists of all posets $P \in \mathcal{N}_2$ via $\{\mathcal{C}_1, \mathcal{C}_2\}$ such that $\#(\boldsymbol{c}_1 \cap \boldsymbol{c}_2) \leq 1$ for all $\boldsymbol{c}_i \in \mathcal{C}_i$.

THEOREM 9. *The collection $\mathcal{N}_2{}''$ is the set of posets each of whose connected components is a convex subposet of $\boldsymbol{N} \times \boldsymbol{N}$ or of $\boldsymbol{N} \times \boldsymbol{N}$ on the cylinder.*

4.3. Open questions.
This paper studies finite posets and saturated chains, but interesting results may arise if we relax one or both of these restrictions. Similarly, we could study posets with some variation of the non-messing-up phenomenon. For example, we could consider more than two sets of chains, or expand beyond identities like $S_i S_{3-i} S_i(\mathcal{L}(P)) = S_{3-i} S_i(\mathcal{L}(P))$ for all labelings $\mathcal{L}$ of P and $i \in \{1, 2\}$, where $S_i(\mathcal{L}(P))$ represents $\mathcal{C}_i$-sorting a labeling $\mathcal{L}$ of a poset P.

Additionally, as stated earlier, any labeling of a poset $P \in \mathcal{N}_2$ produces a linear extension of P after performing the two sorts. It would be interesting to understand the distribution of the linear extensions that arise in this way.

References

[1] H. Boerner, Darstellungen von Gruppen, Springer-Verlag, Berlin, 1955, second edition 1967.
[2] D. Gale and R. M. Karp, The Non-Messing-Up Theorem, preprint.
[3] D. Gale and R. M. Karp, A Phenomenon in the Theory of Sorting, J. Comput. System Sci. 6 (1972) 103-115.
[4] D. E. Knuth, private communication, May 2004.
[5] D. E. Knuth, The Art of Computer Programming, vol. 3, Sorting and Searching, Addison-Wesley, Reading, MA, 1973, second edition 1998.
[6] R. P. Stanley, Enumerative Combinatorics, vol. 1, Cambridge Studies in Advanced Mathematics, no. 49, Cambridge University Press, Cambridge, 1997.

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
E-mail address: bridget@math.mit.edu

Braids and tableaux for unipotent Hecke algebras

Nathaniel Thiem

Abstract

This talk describes a family of Hecke algebras that generalize the classical Iwahori-Hecke algebra. While many of the results extend to other groups of Lie type, this talk focuses on the case where the underlying group is the general linear group over a finite field. The main results include (a) an indexing of the basis elements in terms of row and column degree-sum matrices, (b) a set of braid-like relations for multiplying basis elements, and (c) a generalization of the RSK-correspondence that maps sets of monomial matrices to multi-tableaux.

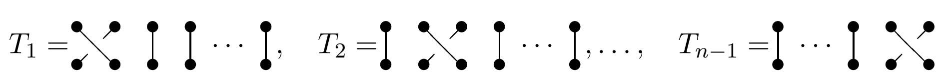
Cet exposé décrit une famille d'algèbres de Hecke qui généralisent l'algèbre classique d'Iwahori-Hecke. Tandis que plusieurs deces résultats se prolongent à d'autres groupes de type de Lie, cet exposé se concentre sur le cas où le groupe sous-jacent est un groupe linéaire général sur un corps fini. Les résultats principaux incluent (a) une indexation des éléments de la base par des colonnes et des rangées de matrices "degree-sum," (b) un ensemble de "braid-like" relations pour multiplier des éléments de la base, et (c) une généralisation de la correspondance RSK qui met en correspondance les ensembles de matrices de monôme avec les multi-tableaux.

1 Introduction

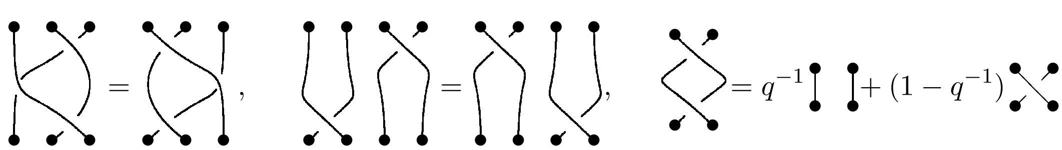
Iwahori [Iw] and Iwahori-Matsumoto [IM] introduced the Iwahori-Hecke algebra as a first step in classifying the irreducible representations of finite Chevalley groups and reductive p-adic Lie groups. Subsequent work (e.g. [Cu] [KL] [LV]) has established Hecke algebras as fundamental tools in the representation theory of Lie groups and Lie algebras, and advances on subfactors and quantum groups by Jones [Jo1], Jimbo [Ji], and Drinfeld [Dr] gave Hecke algebras a central role in knot theory [Jo2], statistical mechanics [Jo3], mathematical physics, and operator algebras. This paper considers a generalization of the classical Iwahori-Hecke algebra obtained by replacing the Borel subgroup B with a maximal unipotent subgroup U.

The Iwahori-Hecke algebra for the general linear group over a finite field ($GL_n(\mathbb{F}_q)$) has a presentation that generalizes the braid-like relations of the symmetric group S_n, given by

Generators.



Relations.



Such a presentation facilitates computations in the Iwahori Hecke algebra and leads to an explicit construction of its representation theory based on the combinatorics of the symmetric group.

This paper examines a family of Hecke algebras that both preserve more of the group structure of $GL_n(\mathbb{F}_q)$ in their representation theory, but also maintain the underlying braid structure of the symmetric group.

Let $G = GL_n(\mathbb{F}_q)$ be the general linear group over the finite field $\mathbb{F}_q$ with q elements. Define subgroups

$$\begin{aligned} T &= \left\{ \begin{array}{c} \text{diagonal} \\ \text{matrices} \end{array} \right\}, \quad N = \left\{ \begin{array}{c} \text{monomial} \\ \text{matrices} \end{array} \right\}, \\ W &= \left\{ \begin{array}{c} \text{permutation} \\ \text{matrices} \end{array} \right\}, \quad \text{and} \quad U = \left\{ \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \right\}, \end{aligned} \tag{1.1}$$

where a monomial matrix is a matrix with exactly one nonzero entry in each row and column.

Fix a composition $\mu = (\mu_1, \mu_2, \ldots, \mu_\ell)$ of n and a nontrivial linear character $\psi : \mathbb{F}_q^+ \to \mathbb{C}^*$ of the additive group of the field. Place a 1 in the last box of every row in the Ferrer's diagram of μ and let all the other boxes contain 0. Let $\mu_{(1)}, \mu_{(2)}, \ldots, \mu_{(n)}$ be the sequence of 0's and 1's obtained by reading left to right, top to bottom. Then the map

$$\begin{array}{cccc} \psi_\mu : & U & \longrightarrow & \mathbb{C}^* \\ & \begin{pmatrix} 1 & t_1 & * & * \\ 0 & 1 & \ddots & * \\ \vdots & \ddots & \ddots & t_{n-1} \\ 0 & \cdots & 0 & 1 \end{pmatrix} & \mapsto & \psi\big(\mu_{(1)} t_1 + \mu_{(2)} t_2 + \cdots + \mu_{(n-1)} t_{n-1}\big) \end{array}$$

is a linear character of U. Let

$$e_\mu = \frac{1}{|U|} \sum_{u \in U} \psi_\mu(u^{-1}) u \qquad \in \mathbb{C}G$$

be the corresponding idempotent. Then the unipotent Hecke algebra $\mathcal{H}(G, U, \psi_\mu)$ is

$$\mathcal{H}_\mu = e_\mu \mathbb{C}G e_\mu \qquad \big(= \mathrm{End}_{\mathbb{C}G}(\mathrm{Ind}_U^G(\psi_\mu))\big),$$

with a natural "double-coset" basis given by

$$\{e_\mu v e_\mu \mid v \in N_\mu\}, \qquad \text{where} \quad N_\mu = \{v \in N \mid e_\mu v e_\mu \neq 0\}.$$

The main results

The main results of this paper are

Section 3 An enumeration of N_μ in terms of matrices with

(1) monic polynomials entries in $\mathbb{F}_q[X]$
(2) row degree-sums and column degree-sums equal to μ.

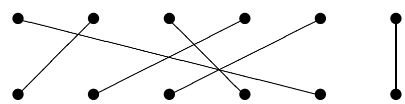
Section 4 A set of braid-like relations for multiplying elements of the double coset basis of $\mathcal{H}_\mu$.

Section 5 An RSK-correspondence between N_μ and column strict multi-tableaux.

This abstract focuses on results rather than proofs; for a more in depth analysis, see [Th1, Th2].

2 Skein model

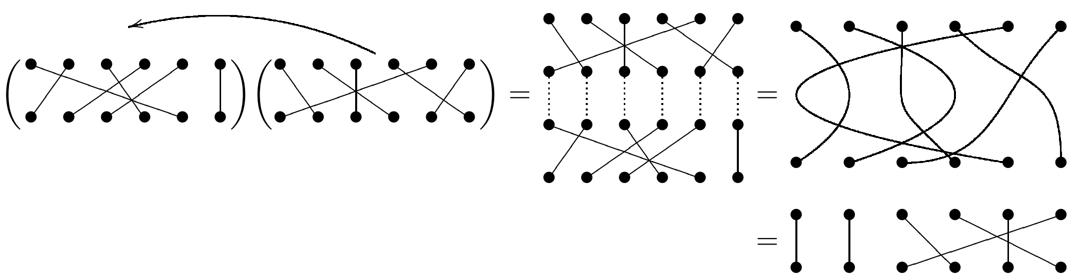
For the results that follow, it will be useful to view elements of $\mathbb{C}G$ as braid-like diagrams instead of matrices. The basic idea is to depict an $n \times n$ permutation matrix w as two rows of n vertices each, with an edge (called a strand) from the ith top vertex to the jth bottom vertex if $w(i) = j$. For example,



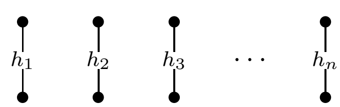
corresponds to

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Matrix multiplication corresponds to concatenation of diagrams, so

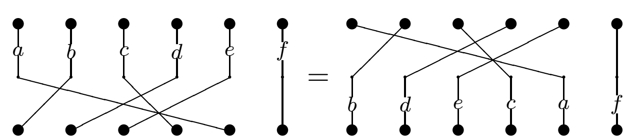


We generalize these diagrams to N by adding "beads" to these diagrams that slide along the strands. Thus, a diagonal matrix corresponds to the identity permutation with a bead on each strand, such as



corresponds to $\quad \mathrm{diag}(h_1, h_2, \ldots, h_n)$.

The advantage of this approach is that it allows a visual shortcut to computing products (such as the permutations above) and commutations in N. For example, by simply pushing the beads of $h \in T$ along the strands of $w \in W$,



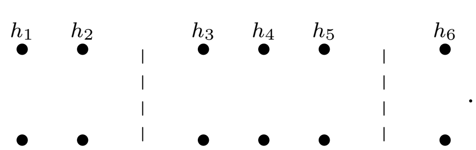
gives

$$s_4 s_3 s_4 s_2 s_3 s_1 \mathrm{diag}(a, b, c, d, e, f) = \mathrm{diag}(b, d, e, c, a, f) s_4 s_3 s_4 s_2 s_3 s_1,$$

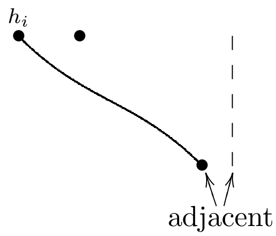
where s_i is the simple transposition $(i, i+1)$.

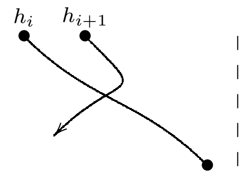
3 A natural basis for $\mathcal{H}_\mu$

We may characterize the elements of N_μ in the following fashion. Suppose $v \in N$. Partition the top vertices by μ; for example, $\mu = (2, 3, 1)$ gives

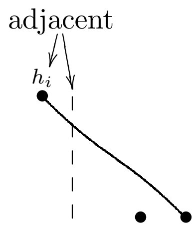


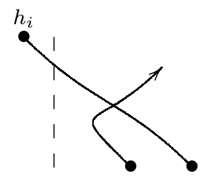
Then $e_\mu v e_\mu \neq 0$ if and only if the diagram for v satisfies

(1) if 

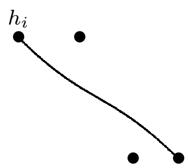
, then 

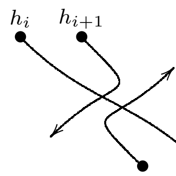
,

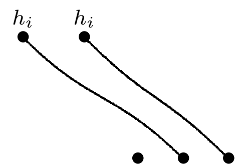
(2) if 

, then 

,

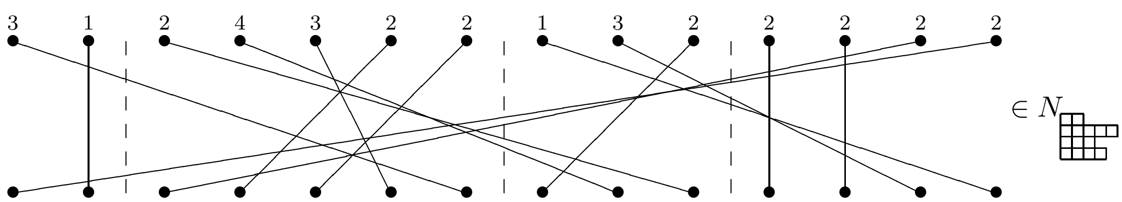
(3) if 

, then either 

or 

.

Example. The element



.

Let $\mu = (\mu_1, \mu_2, \ldots, \mu_\ell)$ be a composition of n, and define

$$M_\mu = \left\{ a = (a_{ij}) \in M_\ell(\mathbb{F}_q[X]) \;\middle|\; \begin{array}{l} a_{ij}(X) \text{ is monic with } a_{ij}(0) \neq 0 \\ \sum_{j=1}^{\ell} deg(a_{ij}) = \mu_i, \sum_{i=1}^{\ell} deg(a_{ij}) = \mu_j \end{array} \right\}$$

$$m_\mu = \left\{ a = (a_{ij}) \in M_\ell(\mathbb{Z}_{\geq 0}) \;\mid\; \sum_{j=1}^{\ell} a_{ij} = \mu_i, \sum_{i=1}^{\ell} a_{ij} = \mu_j \right\},$$

where the degree of a polynomial f is denoted by $deg(f)$.

Theorem 3.1. *Let μ be a composition of n. Then there is a bijection*

$$N_\mu \stackrel{1-1}{\longleftrightarrow} M_\mu.$$

Corollary 1. *Let μ be a composition of n. Then*

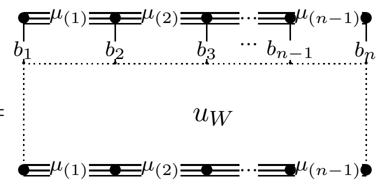
$$\dim(\mathcal{H}_\mu) = \sum_{a \in m_\mu} (q-1)^{\ell(a)} q^{n-\ell(a)},$$

where $\ell(a) = |\{a_{ij} \neq 0 \;\mid\; 1 \leq i, j \leq \ell\}|$.

4 Multiplication relations in $\mathcal{H}_\mu$

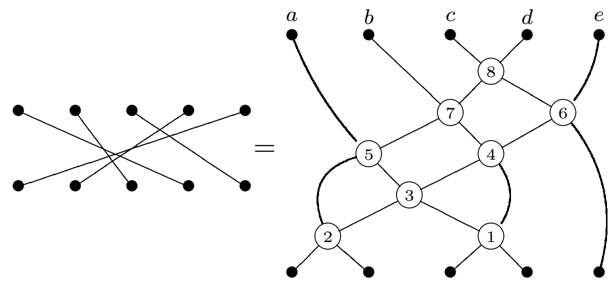
Suppose $u \in N_\mu$ and write $u = u_W u_T$ where $u_W \in W$ and $u_T = \mathrm{diag}(b_1, b_2, \dots, b_n) \in T$. Depict the corresponding unipotent Hecke algebra element as

$e_\mu u e_\mu =$



.

It will be necessary to select a specify a minimal decomposition of u_W in W. Depict this choice, by numbering the crossings from 1 to the length of u_W. For example, if $u_W = s_3 s_1 s_2 s_3 s_1 s_4 s_2 s_3$, then write

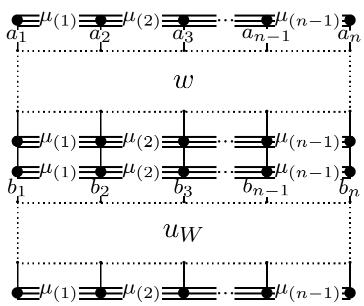


For $1 \le i < j \le \ell$, let

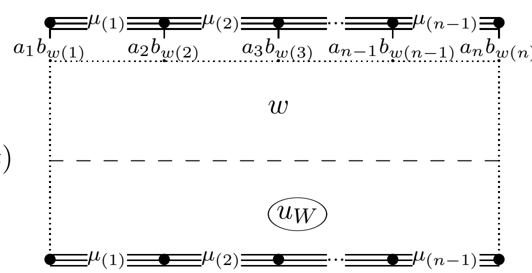
$$\mu_{ij} = \begin{cases} \mu_{(i)}, & \text{if } j = i+1, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose $v \in N_\mu$ such that $v = w\mathrm{diag}(a_1, a_2, \dots, a_n)$ with $w \in W$. Then

Relation 0



$$= \frac{1}{q^r} \sum_{t \in \mathbb{F}_q^r} (\psi \circ f_u)(t)$$



where $u_W = s_{i_1} s_{i_2} \dots, s_{i_r}$ according to some choice of minimal decomposition in W, and $f_u \in \mathbb{F}_q[y_1, y_2, \dots, y_r]$ is given by

$$f_u(y_1, y_2, \dots, y_r) = -\mu_{i_1 j_1} b_{i_1}^{-1} b_{j_1} y_1 - \mu_{i_2 j_2} b_{i_2}^{-1} b_{j_2} y_2 - \dots - \mu_{i_r j_r} b_{i_r}^{-1} b_{j_r} y_r, \tag{4.1}$$

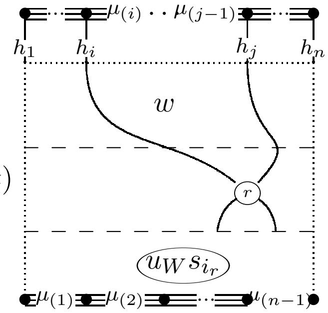
where $(i_k, j_k) = (l, m)$, if the kth crossing in u crosses the strands coming from the lth and mth top vertices in u.

To simplify the concatenated product, apply one of the following two relations to crossing r in u_W.

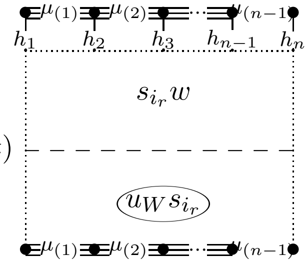
Relation 1

If the strands that cross at ⓡ do not cross in w, then for any $f \in \mathbb{F}_q[y_1^{\pm 1}, \ldots, y^{\pm r}]$,

$$\frac{1}{q^r} \sum_{t \in \mathbb{F}_q^r} (\psi \circ f)(t)$$



$$= \frac{1}{q^r} \sum_{t \in \mathbb{F}_q^r} (\psi \circ f^{(+0)})(t)$$



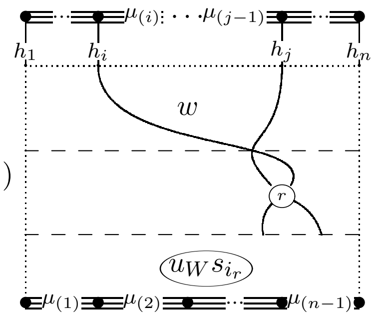
.

where $f^{(+0)} = f + \mu_{ij} h_i^{-1} h_j y_r$. Note that $f^{(+0)} = f$ unless $j = i + 1$.

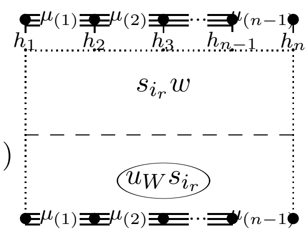
Relation 2

If the strands that cross at ⓡ cross in w, then

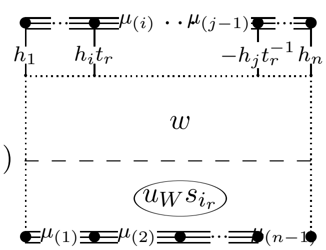
$$\frac{1}{q^r} \sum_{\substack{t' \in \mathbb{F}_q^{r-1} \\ t_r \in \mathbb{F}_q}} (\psi \circ f)(t', t_r)$$



$$= \frac{1}{q^r} \sum_{\substack{t' \in \mathbb{F}_q^{r-1} \\ t_r = 0}} (\psi \circ f)(t', t_r)$$



$$+ \frac{1}{q^r} \sum_{\substack{t' \in \mathbb{F}_q^{r-1} \\ t_r \in \mathbb{F}_q^*}} (\psi \circ f^{(1)})(t', t_r)$$

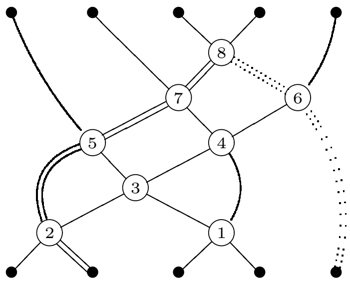


where $f^{(1)} = \varphi_r(f) + \mu_{ij} h_j h_i^{-1} y_r^{-1}$, and $\varphi_r : \mathbb{F}_q[y_1^{\pm 1}, \ldots, y_r^{\pm 1}] \to \mathbb{F}_q[y_1^{\pm 1}, \ldots, y_r^{\pm 1}]$ is computed in Lemma 4.1 below. Note that we could have applied these steps for any f, u, and v, so we can iterate the process with each sum until we have applied either Relation 1 or Relation 2 to to every numbered crossing in u_W.

Relation $2'$: A combinatorial way to compute $\varphi_k(f)$.

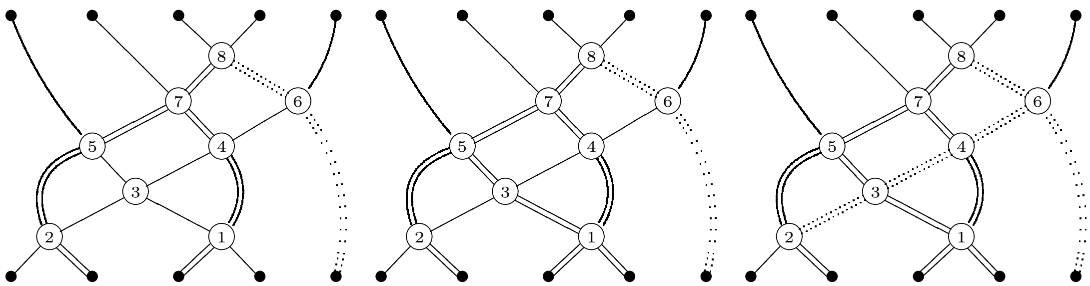
Paint the strands below crossing ⓚ in u_W. Suppose $u_W = s_{i_1} s_{i_2} \cdots s_{i_k} \in W$ is a minimal expression. Each step is illustrated with the example $u_W = s_3 s_1 s_2 s_3 s_1 s_4 s_2 s_3$.

(1) Paint the left [respectively right] strand exiting k below red [blue] all the way to the bottom of the diagram.



where red is $=$, blue is $\cdots$, and k is 8.

(2) For each crossing that the red [blue] strand passes through, paint the right [left] strand (if possible) red [blue] until that strand either reaches the bottom or crosses the blue [red] strand of (1).

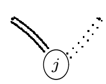


(4.2)

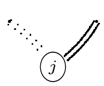
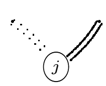
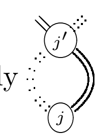
(3) Set

$$u_W^{(k)} = \text{ the diagram } u_W \text{ painted according to (1) and (2).} \tag{4.3}$$

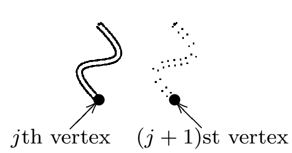
Sinks. The diagram $u_W^{(k)}$ has a *crossed sink at* j if j is a crossing between a red strand and a blue one, or



Note that since u_W is decomposed according to a minimal expression in W, there will be no crossings of the form

 (since  would imply  for some $j' \geq j$.)

The diagram $u_W^{(k)}$ has a *bottom sink at* j if a red strand enters jth bottom vertex *and* a blue strand enters the $(j+1)$st bottom vertex, or



Example (continued) In the running example above $u_W^{(8)}$ has crossed sinks at ②, ③, and ④, and a bottom sink at 4. Note that ① is not a crossed sink since both strands are red.

Paths. A *red [respectively blue] path* p from a sink s (either crossed or bottom) in u_W^{k} is an increasing sequence

$$j_1 < j_2 < \cdots < j_l = k,$$

such that in u_W^{k}

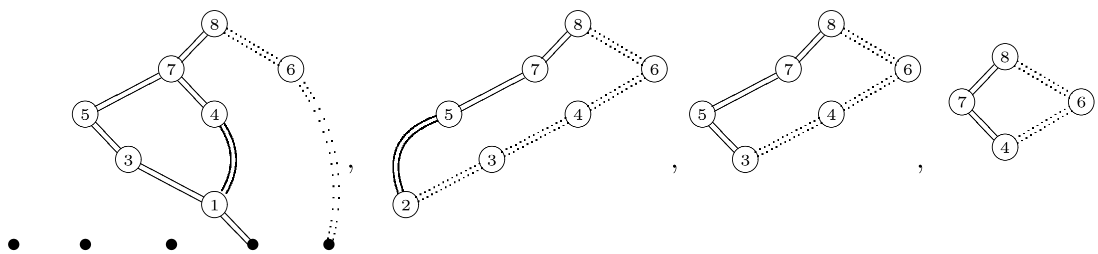
(a) j_m is directly connected (no intervening crossings) to j_{m+1} by a red [blue] strand,

(b) if s is a crossed sink, then $j_1 = s$,

(b′) if s is a bottom sink, then

- in a red path, the sth bottom vertex connects to the crossing j_1 with a red strand.
- in a blue path, the $(s+1)$st bottom vertex connects to the crossing j_1 with a blue strand.

Example (continued). The sinks with their corresponding paths for u_W^{8} are



Let

$$P_{=}(u_W^{k}, s) = \left\{\begin{matrix}\text{red paths from} \\ s \text{ in } u_W^{k}\end{matrix}\right\} \quad \text{and} \quad P_{:}(u_W^{k}, s) = \left\{\begin{matrix}\text{blue paths from} \\ s \text{ in } u_W^{k}\end{matrix}\right\} \tag{4.4}$$

The *weight* of a path p is

$$\mathrm{wt}(p) = \begin{cases} \displaystyle\prod_{\substack{p \text{ switches} \\ \text{strands at } i}} y_i, & \text{if } p \in P_{=}(u_W^{k}, s), \\ \displaystyle\prod_{\substack{p \text{ switches} \\ \text{strands at } i}} (-y_i), & \text{if } p \in P_{:}(u_W^{k}, s). \end{cases} \tag{4.5}$$

Each sink s in u_W^{k} (either crossed j or bottom j) has an associated polynomial $g_s \in \mathbb{F}_q[y_1, y_2, \ldots, y_{k-1}, y_k^{-1}]$ given by

$$g_s = \sum_{p \in P_{=}(u_W^{k}, s)} \; \sum_{p' \in P_{:}(u_W^{k}, s)} \mathrm{wt}(p) y_k^{-1} \mathrm{wt}(p'). \tag{4.6}$$

Example (continued). Consider the weights of the above paths,

Sink	4	4	4	2	2
Path	$1 < 3 < 5 < 7 < 8$	$1 < 4 < 7 < 8$	$6 < 8$	$2 < 5 < 7 < 8$	$2 < 3 < 4 < 6 < 8$
Weight	y_5	$y_1 y_7$	1	1	$-y_6$

Sink	③	③	④	④
Path	$3<5<7<8$	$3<4<6<8$	$4<7<8$	$4<6<8$
Weight	y_5	$-y_6$	y_7	$-y_6$

The corresponding polynomials are

$$g_4 = y_5 y_8^{-1} + y_1 y_7 y_8^{-1}, \quad g_{②} = -y_8^{-1} y_6, \quad g_{③} = -y_5 y_8^{-1} y_6, \quad g_{④} = -y_7 y_8^{-1} y_6. \tag{4.7}$$

Lemma 4.1. *Let u_W and φ_r be as in Relation 2; suppose $u_W^{(r)}$ is painted as above. Then*

$$\varphi_r(f) = f\Big|_{\{y_j \mapsto y_j - g_{(j)} \,\mid\, (j) \text{ a crossed sink}\}} + \sum_{\substack{j \text{ a bottom} \\ \text{sink}}} \mu_{(j)} g_j .$$

For example (see (4.7)),

$$\varphi_8(f) = f\Big|_{\substack{y_4 \mapsto y_4 - g_{④} \\ y_3 \mapsto y_3 - g_{③} \\ y_2 \mapsto y_2 - g_{②}}} + \mu_4 g_4 = f\Big|_{\substack{y_4 \mapsto y_4 + y_7 y_8^{-1} y_6 \\ y_3 \mapsto y_3 + y_5 y_8^{-1} y_6 \\ y_2 \mapsto y_2 + y_8^{-1} y_6}} + \mu_4 (y_5 y_8^{-1} + y_1 y_7 y_8^{-1}).$$

5 A generalized RSK-correspondence

Let

$$\Phi = \left\{ f \in \mathbb{C}[t] \ : \ \begin{matrix} f \text{ is monic, irre-} \\ \text{ducible and } f(0) \neq 0 \end{matrix} \right\}. \tag{5.1}$$

A Φ-*partition* $\lambda = (\lambda^{(f_1)}, \lambda^{(f_2)}, \ldots)$ is a sequence of partitions indexed by Φ. The *size* of λ is

$$|\lambda| = \sum_{f \in \Phi} deg(f) |\lambda^{(f)}|.$$

A *column strict tableau* $P = (P^{(f_1)}, P^{(f_2)}, \ldots)$ *of shape* λ is a column strict filling of λ by positive integers. That is, $P^{(f)}$ is a column strict tableau of shape $\lambda^{(f)}$. Write $\mathrm{sh}(P) = \lambda$. The *weight of* P is the composition $\mathrm{wt}(P) = (\mathrm{wt}(P)_1, \mathrm{wt}(P)_2, \ldots)$ given by

$$\mathrm{wt}(P)_i = \sum_{f \in \Phi} deg(f) \begin{pmatrix} \text{number of} \\ i \text{ in } P^{(f)} \end{pmatrix}.$$

If λ is a Φ-partition and μ is a composition, then let

$$\hat{\mathcal{H}}_\mu^\lambda = \{\text{column strict tableaux } P \ : \ \mathrm{sh}(P) = \lambda, \mathrm{wt}(P) = \mu\} \tag{5.2}$$

and

$$\hat{\mathcal{H}}_\mu = \{\lambda \text{ a } \Phi\text{-partition} \ : \ \hat{\mathcal{H}}_\mu^\lambda \text{ is not empty}\}. \tag{5.3}$$

The following theorem is a consequence of double centralizer theory and [Ze, Theorem 5.5].

Theorem 5.1. *The set $\hat{\mathcal{H}}_\mu$ indexes the irreducible $\mathcal{H}_\mu$-modules $\mathcal{H}_\mu^\lambda$ and*

$$\dim(\mathcal{H}_\mu^\lambda) = |\hat{\mathcal{H}}_\mu^\lambda|.$$

The $(\mathcal{H}_\mu, \mathcal{H}_\mu)$-bimodule decomposition

$$\mathcal{H}_\mu \cong \bigoplus_{\lambda \in \hat{\mathcal{H}}_\mu} \mathcal{H}_\mu^\lambda \otimes \mathcal{H}_\mu^\lambda \quad \text{implies} \quad |N_\mu| = \dim(\mathcal{H}_\mu) = \bigoplus_{\lambda \in \hat{\mathcal{H}}_\mu} \dim(\mathcal{H}_\mu^\lambda)^2 = \sum_{\lambda \in \hat{\mathcal{H}}_\mu} |\hat{\mathcal{H}}_\mu^\lambda|^2.$$

Theorem 5.2, below, gives a combinatorial proof of this identity.

Encode each matrix $a \in M_\mu$ as a Φ-sequence

$$(a^{(f_1)}, a^{(f_2)}, \ldots), \qquad f_i \in \Phi,$$

where $a^{(f)} \in M_{\ell(\mu)}(\mathbb{Z}_{\geq 0})$ is given by

$$a_{ij}^{(f)} = \text{highest power of } f \text{ dividing } a_{ij}.$$

Note that this is an entry by entry "factorization" of a such that

$$a_{ij} = \prod_{f \in \Phi} f^{a_{ij}^{(f)}}.$$

Let the classical RSK correspondence be given by

$$\begin{array}{ccc} M_\ell(\mathbb{Z}_{\geq 0}) & \longrightarrow & \left\{ \begin{array}{l} \text{Pairs } (P,Q) \text{ of column strict} \\ \text{tableaux of the same shape} \end{array} \right\} \\ b & \mapsto & (P(b), Q(b)). \end{array}$$

Theorem 5.2. *For $a \in M_\mu$, let $P(a)$ and $Q(a)$ be the Φ-column strict tableaux given by*

$$P(a) = (P(a^{(f_1)}), P(a^{(f_2)}), \ldots) \qquad \text{and} \qquad Q(a) = (Q(a^{(f_1)}), Q(a^{(f_2)}), \ldots) \qquad \text{for } f_i \in \Phi.$$

Then the map

$$\begin{array}{ccccc} N_\mu & \longrightarrow & M_\mu & \longrightarrow & \left\{ \begin{array}{l} \textit{Pairs } (P,Q) \textit{ of } \Phi\textit{-column} \\ \textit{strict tableaux of the same} \\ \textit{shape and weight } \mu \end{array} \right\} \\ v & \mapsto & a_v & \mapsto & (P(a_v), Q(a_v)), \end{array}$$

is a bijection, where the first map is the inverse of the bijection of Theorem 3.1.

By the construction above, the map is well-defined and since all the steps are invertible, the map is a bijection.

For example, suppose $\mu = (7,5,3,2)$ and $f, g, h \in \Phi$ are such that $deg(f) = 1$, $deg(g) = 2$, and $deg(h) = 3$. Then

$$a_v = \begin{pmatrix} g & f^2h & 1 & 1 \\ h & 1 & g & 1 \\ 1 & 1 & f & f^2 \\ g & 1 & 1 & 1 \end{pmatrix} \in M_{\ydiagram{7,5,3,2}}$$

corresponds to the sequence

$$(a_v^{(f_1)}, a_v^{(f_2)}, \ldots) = \left(\begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}^{(f)}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{(g)}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}^{(h)} \right)$$

and

$$(P(a_v), Q(a_v)) = \left(\ytableaushort{224,34}^{(f)}, \ytableaushort{11,3}^{(g)}, \ytableaushort{12}^{(h)} \right) \left(\ytableaushort{113,33}^{(f)}, \ytableaushort{14,2}^{(g)}, \ytableaushort{12}^{(h)} \right).$$

References

[Cu] Curtis, C. "Representations of Hecke algebras." *Astérisque* **9** (1988): 13-60.

[Dr] Drinfeld, V. "Quantum Groups." *Proceedings of the International Congress of Mathematicians, Berkeley, California, 1-2 1986.* Providence, RI: American Mathematical Society, 1987. 798-820.

[Iw] Iwahori, N. "On the structure of a Hecke ring of a Chevalley group over a finite field." *Journal of the Faculty of Science, University of Tokyo* **10** (1964): 215-236.

[IM] Iwahori, N. and Matsumoto, H. "On some Bruhat decomposition and the structure of Hecke rings of p-adic Chevalley groups." *Institut des Hautes Études Scientifiques, Publications Mathématiques* **25** (1965): 5-48.

[Ji] Jimbo, M. "A q-analogue of $U(\mathfrak{gl}(N+1))$, Hecke algebra, and the Yang-Baxter equation." *Letters in Mathematical Physics* **11** (1986): 247-252.

[Jo1] Jones, V. "Index for subfactors." *Inventiones Mathematicae* **72** (1983): 1-25.

[Jo2] Jones, V. "Hecke algebra representations of Braid groups and link polynomials." *Annals of Mathematics* **126** (1987): 103-111.

[Jo3] Jones, V. "On knot invariants related to some statistical mechanical models." *Pacific Journal of Mathematics* **137** (1989): 311-334.

[KL] Kazhdan, D. and Lusztig G. "Representations of Coxeter groups and Hecke algebras." *Inventiones Mathematicae* **53** (1979): 165-184.

[LV] Lusztig, G. and Vogan, D. "Singularities of closures of K-orbits on flag manifolds." *Inventiones Mathematicae* **21** (1983): 365-379.

[Th1] Thiem, N. "Unipotent Hecke algebras of $GL_n(\mathbb{F}_q)$." *Journal of Algebra* **284** (2005): 559-577.

[Th2] Thiem, N. "A skein-like multiplication algorithm for unipotent Hecke algebras." accepted, *Transactions of the American Mathematical Society.*

[Ze] Zelevinsky, A. *Representations of Finite Classical Groups.* New York: Springer Verlag, 1981.

EQUALITY OF SCHUR AND SKEW SCHUR FUNCTIONS

STEPHANIE VAN WILLIGENBURG

ABSTRACT. We determine the precise conditions under which any skew Schur function is equal to a Schur function over both infinitely and finitely many variables.

RÉSUMÉ. Nous déterminons l'égalité d'une fonction Schur et une fonction skew Schur.

1. INTRODUCTION

Littlewood-Richardson coefficients arise in a variety of contexts. The first of these is that they are the structure constants in the algebra of symmetric functions with respect to the basis of Schur functions. Another instance is as the multiplicities of irreducible representations in the tensor product of representations of the symmetric group. A third occurrence is as intersection numbers in the Schubert Calculus on a Grassmanian. Thus knowing their values has an impact on a number of fields. In this paper we calculate when certain coefficients are 0 or 1 by determining when a skew Schur function is equal to a Schur function. Although multiplicity free products have been studied in [4] our determination will reveal more precisely when certain coefficients are 0 and when they are 1. The related question of when two ribbon Schur functions are equal has been answered recently in [1], which revealed many new equalities of Littlewood-Richardson coefficients.

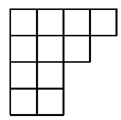
The remainder of this note is structured as follows. In the rest of Section 1 we review the definitions required. This is followed by the two main theorems in which we give straightforward conditions that prescribe when a skew Schur function is equal to a Schur function over both infinitely and finitely many variables.

2000 *Mathematics Subject Classification.* 05E05, 05E10.

Key words and phrases. Schur function, skew Schur function, Littlewood-Richardson coefficients.

The author was supported in part by the National Sciences and Engineering Research Council of Canada.

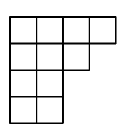
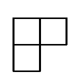
1.1. **Schur and skew Schur functions.** We say that a list of positive integers $\lambda = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ whose sum is n is a *partition* of n, denoted $\lambda \vdash n$. We call the λ_i the *parts* of λ. A partition with at most one part size is called a *rectangle* and a partition with exactly two different part sizes is called a *fat hook*. If $\lambda = \lambda_1\lambda_2\ldots\lambda_k \vdash n$ then we define the *(Ferrers) diagram* D_λ to be the array of left justified boxes with λ_i boxes in the i-th row for $1 \leq i \leq k$. If we transpose D_λ we obtain another diagram $D_{\lambda'}$ known as the *conjugate* of D_λ and we refer to λ and λ' as *conjugate partitions*. Furthermore, for any column c in D_λ we denote by $l(c)$ the number of boxes in c and refer to $l(c)$ as the *length* of c. Where the context is clear we abuse notation and refer to D_λ as λ.

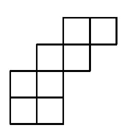
Example 1.1. $D_{4322} =$ 

We define a *(Young) tableau* T of *shape* λ to be a filling of the boxes of D_λ with positive integers. If the filling is such that the integers in each row weakly increase, whilst the integers in each column strictly increase we say that T is a *semi-standard* tableau.

Example 1.2. $\begin{array}{cccc} 1 & 1 & 2 & 3 \\ 2 & 3 & 4 & \\ 4 & 4 & & \\ 5 & 7 & & \end{array}$ is a semi-standard tableau of shape 4322.

If $\mu = \mu_1 \ldots \mu_k \vdash m$, $\lambda = \lambda_1 \ldots \lambda_l \vdash n$ where $k \leq l$ and $m \leq n$ such that $\mu_i \leq \lambda_i$ for $1 \leq i \leq k$ then we define the *skew diagram* $D_{\lambda/\mu}$ to be the array of boxes that appear in D_λ but not in D_μ. For our purposes $D_{\lambda/\mu}$ will always be *connected*, that is to say, for any pair of adjacent rows in $D_{\lambda/\mu}$ there exists at least one column in which they both have a box.

Example 1.3. $D_{4322} =$  $D_{21} =$ 

$D_{4322/21} =$ 

Again when the context is clear we refer to $D_{\lambda/\mu}$ as λ/μ. Similarly we define *skew tableaux* and *semi-standard skew tableaux* by appropriately inserting the adjective skew in the above definitions for tableaux and semi-standard tableaux.

Definition 1.1. *Given a (skew) tableau T and a set of variables $x_1, x_2, \ldots$ we define the monomial x^T to be*

$$x^T := x_1^{t_1} x_2^{t_2} \ldots$$

where t_i is the number of times i appears in T. Let λ, μ be partitions such that λ/μ is a (skew) diagram, then we define the corresponding (skew) Schur function $s_{\lambda/\mu}$ *to be*

$$s_{\lambda/\mu} = \sum x^T$$

where the sum is over all semi-standard (skew) tableaux T of shape λ/μ.

The set of all Schur functions s_λ (i.e. $s_{\lambda/\mu}$ where $\mu = \emptyset$) forms a basis for the algebra of symmetric functions, Λ, which is a subalgebra of $\mathbb{Q}[x_1, x_2, \ldots]$.

Since it can be easily shown that skew Schur functions are symmetric it follows that skew Schur functions can be written as a linear combination of Schur functions. To be more precise we need to recall two more notions: that of the reading word and the content of a (skew) tableau. Firstly, given a (skew) tableau, T, we say its *reading word*, $w(T)$, is the entries of the tableau read from top to bottom and *right* to *left*. Given a reading word we say it is *lattice* if as we read it from left to right the number of i's we have read is at least as large as the number of $i+1$'s we have read e.g. 1213 is lattice, however, 1132 is not as when we have read 113 the number of 3's we have read is greater than the number of 2's. Secondly, the *content* of a (skew) tableau, $c(T)$, is a list $t_1 t_2 t_3 \ldots$ where, as before, t_i is the number of times i appears in T.

We are now ready to express any skew Schur function as a linear combination of Schur functions.

Proposition 1.2. [3, A1.3.3] *Let λ, μ, ν be partitions such that λ/μ is a (skew) diagram then*

$$s_{\lambda/\mu} = \sum_\nu c^\lambda_{\mu\nu} s_\nu$$

where $c^\lambda_{\mu\nu}$ is the number of semi-standard (skew) tableaux T such that

(1) *the shape of T is λ/μ*
(2) *$c(T) = \nu$*
(3) *$w(T)$ is lattice.*

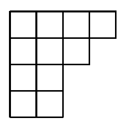
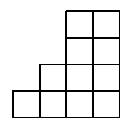
Example 1.4.

$$s_{32/1} = s_{31} + s_{22}.$$

Remark 1.5. The $c^{\lambda}_{\mu\nu}$ are known as Littlewood-Richardson coefficients, and the above method of computing them is known as the Littlewood-Richardson rule. There are many other methods for computing the $c^{\lambda}_{\mu\nu}$ such as Zelevinsky's pictures or Remmel and Whitney's reverse numbering, and the interested reader may wish to consult, say, [2] for further details. However, it is the Littlewood-Richardson rule that will allow us to determine our results most succinctly.

2. Equality of Schur and skew Schur functions

Before we state our main result let us define an involution on diagrams. Given a diagram λ let λ° be the (skew) diagram that is the diagram λ rotated by 180°.

Example 2.1. $4322 =$  $4322^{\circ} =$ 

Theorem 2.1. *For partitions λ, μ, ν*

$$s_{\lambda/\mu} = s_{\nu} \text{ if and only if } \lambda/\mu = \nu \text{ or } \nu^{\circ}.$$

Proof. The reverse implication follows by Exercise 7.56(a) [3], which yields that

$$s_{\nu} = s_{\nu^{\circ}}.$$

For the forward implication we need only show that if λ/μ is not ν or ν° for some diagram ν then λ/μ has more than one filling whose reading word is lattice.

Consider the skew tableau T of shape λ/μ where each column c is filled with the integers $1, \ldots, l(c)$ in increasing order. This filling is clearly lattice. Now since λ/μ is not a diagram ν nor a (skew) diagram ν° where ν is a diagram, consider the first row i where λ/μ fails to be either ν or ν° for some diagram ν (i.e. if λ/μ is truncated at row $i-1$ then we obtain a (rotated) diagram, but this is no longer true if λ/μ is truncated at row i). Moving from right to left note the first entry j in T which does not have $i-1$ entries above it. Form the reading word of T upto this entry and note the smallest integer $i \geq k > j$ for which the number of occurrences of k is strictly less than the number of occurrences of $k-1$. Change j to k and change all entries below it in that column by adding $k-j$ to the existing entry to form a new skew tableau T' of shape λ/μ. Since $w(T')$ is clearly lattice, we are done. □

3. Equality and $GL(n)$ or $SL(n)$ characters

The set of all Schur functions restricted to the variables $x_1, \ldots, x_n$, obtained by setting $x_m = 0$ for $m > n$, forms a basis for the algebra of symmetric polynomials Λ_n. Skew Schur functions in Λ_n can be expressed in terms of the Schur functions by

$$s_{\lambda/\mu}(x_1, \ldots, x_n) = \sum_{\nu} c^{\lambda}_{\mu\nu} s_\nu(x_1, \ldots, x_n)$$

and thus we can ask when $s_{\lambda/\mu}(x_1, \ldots, x_n) = s_\nu(x_1, \ldots, x_n)$. In terms of representation theory this yields when certain multiplicities in the tensor products of irreducible representations of $SL(n, \mathbb{C})$ (or polynomial representations of $GL(n, \mathbb{C})$) will be 0 and when they will be 1.

Clearly if $s_{\lambda/\mu} = s_\nu$ in Λ then the result holds in Λ_n, however the converse may not be true as an $s_{\lambda/\mu}$ comprising of a sum of s_ν only one of which has less than $n+1$ parts could exist. However, the search for such an $s_{\lambda/\mu}$ is greatly reduced as the converse may not be true only when the length of the longest column in λ/μ is equal to n by

Lemma 3.1. *Let λ, μ be partitions such that the length of the longest column in λ/μ is m.*

1. *If $m > n$ then $s_{\lambda/\mu} = 0$ in Λ_n.*
2. *If $m < n$ then $s_{\lambda/\mu} = s_\nu$ in Λ_n if and only if $s_{\lambda/\mu} = s_\nu$ in Λ.*

Proof. The first result is immediate from the definitions. The reverse direction of the second result has already been discussed, thus it only remains to show that if $m < n$ and $s_{\lambda/\mu} \neq s_\nu$ in Λ then $s_{\lambda/\mu} \neq s_\nu$ in Λ_n.

Consider the skew tableau T of shape λ/μ where each column c is filled with the integers $1, \ldots, l(c)$ in increasing order. Since $s_{\lambda/\mu} \neq s_\nu$ in Λ this implies λ/μ is not a diagram ν nor a (skew) diagram ν°, so as in the proof of Theorem 2.1 consider the first row i where λ/μ fails to be ν or ν° for some diagram ν. Moving from right to left note the first entry j in column c of T which does not have $i-1$ entries above it, form the reading word upto this entry, w', and note the smallest integer $i \geq k > j$ for which the number of occurrences of k is strictly less than the number of occurrences of $k-1$. If $l(c) + k - j \leq n$ then change the j to k and change all entries below it in c by adding $k - j$ to the existing entry to form a new skew tableau T' of shape λ/μ. If not then find the largest entry in w', $k' < n$, and fill the $n - k'$ lowest boxes in c with $k'+1, \ldots, n$ to form T'. Since $w(T')$ is clearly lattice, the result follows. $\square$

Example 3.1. If $n = 5$ then the following semi-standard skew tableaux illustrate T' in the situation $l(c) + k - j \leq n$ and $l(c) + k - j > n$ respectively.

$$\begin{array}{cccc} & & 1 & 1 \\ & 2 & 2 & \\ 1 & 3 & & \\ 2 & 4 & & \end{array} \qquad \begin{array}{ccc} & & 1 \\ & & 2 \\ & & 3 \\ & 1 & 4 \\ & 2 & \\ 1 & 5 & \end{array}$$

Thus, from here on we shall assume that the length of the longest column in λ/μ is n. Before we reveal the analogous result to Theorem 2.1 let us define two operations on (skew) diagrams.

Definition 3.2. *Let $\lambda = \lambda_1 \ldots \lambda_k$ and $\mu = \mu_1 \ldots \mu_l$ be partitions such that λ/μ is a (skew) diagram. Let c be a column of longest length in λ/μ and $\overline{\lambda}$ and $\overline{\mu}$ be partitions such that*

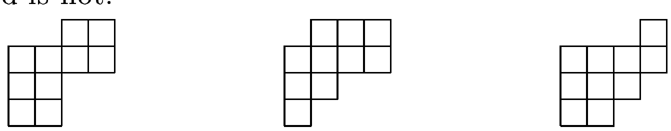
(1) $\overline{\lambda}' = (\lambda'_1 + r)(\lambda'_2 + r)(\lambda'_3 + r) \ldots (\lambda'_c + r)\lambda'_{c+1} \ldots \lambda'_{\lambda_1}$
$\overline{\mu}' = (\mu'_1 + r)(\mu'_2 + r)(\mu'_3 + r) \ldots (\mu'_c + r)\mu'_{c+1} \ldots \mu'_{\mu_1}$ *or*

(2) $\overline{\lambda}' = (\lambda'_1 + r)(\lambda'_2 + r)(\lambda'_3 + r) \ldots (\lambda'_{c-1} + r)\lambda'_c \ldots \lambda'_{\lambda_1}$
$\overline{\mu}' = (\mu'_1 + r)(\mu'_2 + r)(\mu'_3 + r) \ldots (\mu'_{c-1} + r)\mu'_c \ldots \mu'_{\mu_1}$

and λ/μ is a (skew) diagram then we say $\overline{\lambda}/\overline{\mu}$ is a shearing *of λ/μ.*

Remark 3.2. Intuitively we can interpret this definition as creating a diagram $\overline{\lambda}/\overline{\mu}$ from λ/μ by choosing a column of longest length and sliding it and every column to the left of it down r boxes, or sliding it and every column to the right of it up r boxes.

Example 3.3. The first two skew diagrams are shearings of 442 whilst the third is not.



Definition 3.3. *Let $\underline{\lambda}$ and $\underline{\mu}$ be partitions such that*

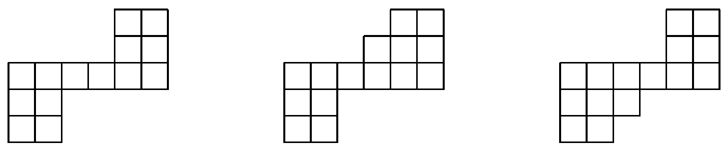
(1) $\underline{\lambda}' = (a+b)^i c^{j+k}$ *and* $\underline{\mu}' = b^{i+k}$ *if $a \neq c$ or*

(2) $\underline{\lambda}' = (a+b)^i(\nu_1 + b) \ldots (\nu_k + b)c^j$ *and* $\underline{\mu}' = b^{i+k}$*, where $\nu = \nu_1 \ldots \nu_k$ is a partition, if $a = c$*

and $\underline{\lambda}/\underline{\mu}$ is a (skew) diagram then we say $\underline{\lambda}/\underline{\mu}$ is a fattening *of $(a^i c^j)'$ and $(\underline{\lambda}/\underline{\mu})^\circ$ is a* fattening *of $((a^i c^j)')^\circ$.*

Remark 3.4. Intuitively we can interpret this definition as creating a diagram say $\underline{\lambda}/\underline{\mu}$ from a fat hook or rectangle in the following way. If we have a fat hook $(a^i c^j)'$ then we shear the rightmost column of length a and all the columns to the left of it down by b boxes. We then insert a rectangle k^{c-b} such that the result is a (skew) diagram. If we have a rectangle $(a^{i+j})'$ then we shear a column and all the columns to the left of it down by b boxes. We then insert a diagram such that the result is a (skew) diagram. A similar interpretation follows for $(\underline{\lambda}/\underline{\mu})^{\circ}$.

Example 3.5. All three skew diagrams are fattenings of 444.



For convenience we extend the notion of fattening to all (skew) diagrams by defining the fattening of λ/μ to be λ/μ if λ/μ is any (skew) diagram other than those referred to in Definition 3.3. In addition, for clarity of exposition, we denote by $\widetilde{\lambda/\mu}$ any (skew) diagram that has been derived from λ/μ via some combination of shearings or fattenings.

Theorem 3.4. *For partitions $\lambda, \mu, \nu, \eta = \eta_1\eta_2\dots$ where η_i is the number of columns of length at least i in λ/μ*

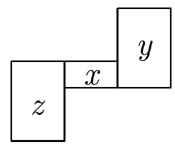
$$s_{\lambda/\mu} = s_\eta \text{ in } \Lambda_n \text{ if and only if } \lambda/\mu = \widetilde{\nu} \text{ or } \widetilde{\nu^{\circ}}.$$

Proof. For the reverse implication it is straightforward to check that if $\lambda/\mu = \widetilde{\nu}$ or $\widetilde{\nu^{\circ}}$ then the only semi-standard skew tableau T of shape λ/μ whose reading word is lattice has each column c filled with the integers $1, \dots, l(c)$ in increasing order.

The forward implication will follow once we show that if $\lambda/\mu \neq \widetilde{\nu}$ or $\widetilde{\nu^{\circ}}$ for some diagram ν then λ/μ has more than one filling utilising the integers $1, \dots, n$ whose reading word is lattice.

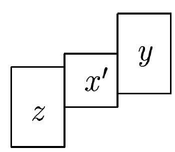
Consider the skew tableau T of such a shape λ/μ whose reading word is lattice where each column c is filled with the integers $1, \dots, l(c)$ in increasing order. If λ/μ has a column c_1 such that $l(c_1) < n$ and c_1 contains at least one box with no box to the right of it, and a column c_2 to the right of c_1 such that $l(c_2) < n$ then if $l(c_1) \leq l(c_2)$ change the entry in the last box of c_1 to $l(c_2) + 1$ otherwise, unless the box at the head of c_1 and the column immediately to the right of it are in the same row, change the entry $l(c_2)$ in c_1 to $l(c_2) + 1$ and increase all entries below it in c_1 by 1 to form a new skew tableau T' of shape λ/μ whose reading word is lattice.

Thus if λ/μ does not satisfy these criteria then λ/μ must be of the form



where x is a rectangle, and by Lemma 3.1 and what we have already proved y (and z) must consist of a non-rectangular (skew) diagram δ or δ°, for some diagram δ, whose column lengths are all less than n with a non-negative number of columns to the right or left of it that consist of n boxes. If c_s is the column of shortest length $l(c_s)$ in y then change the entry $l(c_s)$ in the first column of x to $l(c_s)+1$ and increase all the entries below it in that column by 1. If the first column of x does not contain the entry $l(c_s)$ then change the entry in the last box to $l(c_s)+1$. In both instances we form a new skew tableau T' of shape λ/μ whose reading word is lattice.

Hence the number of columns in x must be zero and λ/μ must be of the form



where the length of every column of x' is n and y (and z) consists of a non-rectangular (skew) diagram δ or δ°, for some diagram δ, whose column lengths are all less than n with a non-negative number of columns to the right (respectively left) that consist of n boxes. For clarity of exposition we identify y and z with the sub (skew) diagram δ or δ° that they contain. Let δ and ε be diagrams and c_l, c_s be the columns of longest or shortest length in δ respectively.

If $y=\delta$ and $z=\varepsilon$ then it follows that x' must contain zero columns otherwise we can change the entries in the first column of ε; every column in ε must be at least as long as the longest column of δ; and the box at the head of the leftmost column of δ and the rightmost column of ε are not in the same row. Change the entry $l(c_l)$ in the first column of ε to $l(c_l)+1$ and increase all the entries below it in that column by 1.

If $y=\delta^\circ$ and $z=\varepsilon^\circ$ then x' must contain zero columns; every column in ε° must be no longer than the shortest column in δ°; and the box at the base of the leftmost column of δ° and the rightmost column of ε°

are not in the same row. Change the last entry in the first column of ε° to $l(c_l)+1$.

If $y=\delta$ and $z=\varepsilon^\circ$ then either x' must contain zero columns or the box at the base of the rightmost column of ε° is in the same row as the adjacent column of length n. In either case we can apply an argument similar to that above to either change the entries of the rightmost column of ε° from $l(c_s)$ downwards by increasing them by 1 or change the entry in the last box to $l(c_s)+1$.

Finally if $y=\delta^\circ$ and $z=\varepsilon$ then the same conditions must be satisfied as for the case $y=\delta$ and $z=\varepsilon$. However, we can change the entries of the rightmost column of ε from $l(c_l)$ downwards by increasing them by 1.

In each situation we have been able to create a new skew tableau T' of shape λ/μ whose reading word is lattice, and having eliminated all possibilities the result follows. $\square$

References

[1] L. Billera, H. Thomas and S. van Willigenburg, Decomposable Compositions, Symmetric Quasisymmetric Functions and Equality of Ribbon Schur Functions, preprint in http://arXiv.org/abs/math/0405434.

[2] W. Fulton, *Young Tableaux*, Cambridge University Press, Cambridge, UK, 1997.

[3] R. Stanley, *Enumerative Combinatorics, Vol. 2*, Cambridge Studies in Advanced Mathematics, Vol. 62, Cambridge University Press, Cambridge, UK, 1999.

[4] J. Stembridge, Multiplicity-free products of Schur functions, *Ann. Combin.* **5**, (2001), 113–121.

Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2

E-mail address: `steph@math.ubc.ca`

THE COMBINATORICS OF TRANSLATION FUNCTORS FOR THE VIRASORO ALGEBRA

EMILIE B. WIESNER

Abstract. In this paper, I consider the representation theory of the Virasoro algebra Vir from a combinatorial perspective. For the Virasoro algebra, the irreducible modules in Category $\mathcal{O}$ are indexed by the weights $\lambda \in \mathfrak{h}^\star = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ of Vir, where $\mathfrak{h}$ is a certain abelian subalgebra of Vir. The weights of Vir can be naturally partitioned in blocks. Using the results of Feigin and Fuchs [1], I give a combinatorial characterization of the blocks for the Virasoro algebra. The module $M(\lambda) \otimes L(\mu)$, where $M(\lambda)$ is the Verma module indexed by the weight $\lambda \in \mathfrak{h}^\star$ and $L(\mu)$ is the irreducible module indexed by the weight $\mu \in \mathfrak{h}^\star$, decomposes into a direct sum of submodules according to blocks. I give a description of the submodules corresponding to each block. The main tools I use to describe blocks and the decomposition of $M(\lambda) \otimes L(\mu)$ are the Shapovalov form and the Shapovalov determinant.

Ce travail est une étude de la théorie combinatoire des représentations de l'algèbre de Lie Virasoro Vir. Les modules simples de la catégorie $\mathcal{O}$ pour Vir sont parametrisés par des poids $\mathfrak{h}^\star = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ de Vir, où $\mathfrak{h}$ est une sous-algèbre commutative spécifique. Les poids sont partitionnés en blocs. En utilisant le travail de Feigin et Fuchs [1], je donne une caractérisation combinatoire pour les blocs de l'algèbre de Virasoro. Si $M(\lambda)$ est un module Verma et $L(\mu)$ le module simple de poids maximal μ, le produit $M(\lambda) \otimes L(\mu)$ a une décomposition en somme directe de sous-modules indexés par les blocs. Je donne une description des sous-modules correspondant aux blocs. Les techniques principales utilisées sont la forme de Shapovalov et le déterminant de Shapovalov.

1. Introduction

The motivation for this paper comes from similar results for semisimple Lie algebras. Let $\mathfrak{g}$ be a finite-dimensional complex semisimple Lie algebra. Shapovalov [7] showed that the Shapovalov determinant associated to a Verma module $M(\lambda)$ of $\mathfrak{g}$ reflects the block structure of Category $\mathcal{O}$ for the Lie algebra $\mathfrak{g}$. Jantzen [3] used the Shapovalov determinant to analyze translation functors for $\mathfrak{g}$, that is, the decomposition of $M(\lambda) \otimes L(\mu)$ as a direct sum of submodules each corresponding to a block. Specifically, he produced determinant formulas for each of these submodules. Using these formulas, Gabber and Joseph [2] were able to contruct a Hecke algebra action on the blocks of $\mathfrak{g}$ using certain translation functors. They suggest this approach may lead to a proof of the Kazhdan-Lusztig Conjecture.

Date: April 13, 2005.

My results are analogous to the results of Jantzen and may allow for contructions for the Virasoro algebra similar to those of Gabber and Joseph. In particular, it may be possible to formulate a Hecke algebra and Kazhdan-Lusztig polynomials for the Virasoro algebra.

2. The Virasoro algebra and a definition of blocks

Lie algebras with triangular decomposition generalize complex finite-dimensional semisimple Lie algebras. The Virasoro algebra is an important example of this class of Lie algebras. The *Virasoro algebra*, Vir, is the complex vector space with basis

$$z \quad \text{and} \quad d_k, k \in \mathbb{Z}$$

and with relations

$$[d_j, d_k] = (j-k)d_{j+k} + \delta_{j,-k}\frac{j^3-j}{12}z, \qquad [z, d_k] = 0.$$

The triangular decomposition of Vir is given by

$$Vir = Vir_- \oplus \mathfrak{h} \oplus Vir_+,$$

where Vir_- is the subalgebra of Vir generated by $d_k, k \in \mathbb{Z}_{<0}$; $\mathfrak{h}$ is the subalgebra generated by z and d_0; and Vir_+ is the subalgebra generated by $d_k, k \in \mathbb{Z}_{>0}$.

Note that $\mathfrak{h}$ acts diagonally on $Vir_\pm$:

$$[d_0, d_k] = -kd_k, \quad [z, d_k] = 0.$$

The *enveloping algebra $U(Vir)$ of Vir* is the associative algebra with generators

$$z \quad \text{and} \quad d_k, k \in \mathbb{Z}$$

and relations

$$\begin{aligned} d_jd_k - d_kd_j &= (j-k)d_{j+k} + \delta_{j,-k}\frac{j^3-j}{12}z \\ zd_k - d_kz &= 0 \end{aligned}$$

The algebra $U(Vir_+)$ is the subalgebra of $U(Vir)$ generated by d_k, $k > 0$, and $U(Vir_-)$ is the subalgebra of $U(Vir)$ generated by d_k, $k < 0$. The algebra $U(Vir_-)$ has a basis

$$d_{-\lambda} = d_{-\lambda_1}d_{-\lambda_2}\cdots d_{-\lambda_k}, \quad \lambda = (\lambda_1 \geq \cdots \geq \lambda_k) \text{ a partition.}$$

For example, if $\lambda = (5, 5, 3, 2)$, then $d_{-\lambda} = d_{-5}d_{-5}d_{-3}d_{-2}$.

Let $\mathfrak{h}^\star = \mathrm{Hom}_{\mathbb{C}}(\mathfrak{h}, \mathbb{C})$, linear maps from $\mathfrak{h}$ to $\mathbb{C}$. There is a bijection

$$\begin{aligned} \mathfrak{h}^\star &\to \mathbb{C}^2 \\ \lambda &\mapsto (\lambda(d_0), \lambda(z)). \end{aligned}$$

For $\lambda, \mu \in \mathfrak{h}^\star$, define $\lambda > \mu$ if $\lambda(d_0) - \mu(d_0) \in \mathbb{Z}_{<0}$ and $\lambda(z) = \mu(z)$. This partial order on $\mathfrak{h}^\star$ is the analogue for Vir of the classical dominance order on partitions.

Category $\mathcal{O}$ consists of those Vir-modules M such that

- M is $\mathfrak{h}$-diagonalizable:

$$M = \bigoplus_{\mu \in \mathfrak{h}^\star} M^\mu, \quad \text{where} \quad M^\mu = \{m \in M | hm = \mu(x)m \text{ for all } x \in \mathfrak{h}\};$$

- with respect to dominance, the $\mu \in \mathfrak{h}^\star$ such that $M^\mu \neq 0$ lie in a lower order ideal with a finite number of maximal elements;
- $\dim M^\mu < \infty$ for all $\mu \in \mathfrak{h}^\star$.

Category $\mathcal{O}$ reflects much of the structure of Vir. Within Category $\mathcal{O}$, Verma modules are of special interest. For $(h, c) \in \mathbb{C}^2$, define the Verma module $M(h, c)$ to be the $U(Vir)$-module generated by v^+ with relations

$$d_0 v^+ = hv^+, \quad zv^+ = cv^+, \quad d_k v^+ = 0 \quad \text{for } k > 0.$$

Then

$$M(h, c) = \bigoplus_{(h', c') \leq (h, c)} M(h, c)^{(h', c')} = \bigoplus_{n \in \mathbb{Z}_{\geq 0}} M(h, c)^{(h+n, c)}.$$

Proposition 1. [6] *The set* $\{\, d_{-\lambda} v^+ |\, \lambda \vdash n\}$ *is a basis for* $M(h, c)^{(h+n,c)}$.

For each $\lambda \in \mathfrak{h}^\star$, $M(\lambda)$ has a unique proper maximal submodule $J(\lambda)$. This implies that $M(\lambda)$ has a unique irreducible quotient

$$L(\lambda) = M(\lambda)/J(\lambda).$$

Theorem 1. [6] *Up to isomorphism, the modules* $L(\lambda)$ *are all of the irreducible modules in Category* $\mathcal{O}$.

In general, modules M in Category $\mathcal{O}$ do not have finite composition series. However, one can make appropriate definitions to say that "$L(\mu)$ appears in M" if $L(\mu)$ is a factor in a "local" composition series for M. Let $\sim$ be the equivalence relation generated by $\lambda \sim \mu$ if $L(\mu)$ appears in $M(\lambda)$. The *blocks* are the equivalence classes $[\lambda]$ of $\mathfrak{h}^\star$ with respect to this equivalence relation.

3. A Description of Blocks

In this section we explicitly describe the blocks for the Virasoro algebra. The main tools we use are the Shapovalov form and the Shapovalov determinant.

The following result provides an alternative way to define blocks.

Proposition 2. [6] *For* $\mu, \lambda \in \mathfrak{h}^\star$, $L(\mu)$ *appears in* $M(\lambda)$ *if and only if* $M(\mu) \subseteq M(\lambda)$.

Verma module embeddings prove to be a useful way to characterize blocks. In particular, the Shapovalov form and the Shapovalov determinant can be used to provide a complete description of Verma module embeddings.

For $(h,c) \in \mathbb{R}^2$ the *Shapovalov form*, $\langle , \rangle : M(h,c) \times M(h,c) \to \mathbb{R}$, is the form defined by the conditions

$$\begin{aligned} \langle v^+, v^+ \rangle &= 1 \\ \langle d_{-\lambda_1} \cdots d_{-\lambda_k} v^+, d_{-\mu_1} \cdots d_{-\mu_j} v^+ \rangle &= \langle v^+, d_{\lambda_k} \cdots d_{\lambda_1} d_{-\mu_1} \cdots d_{-\mu_j} v^+ \rangle \end{aligned}$$

and

$$\begin{aligned} \langle \alpha v + u, w \rangle &= \overline{\alpha} \langle v, w \rangle + \langle u, w \rangle \\ \langle w, \alpha v + u \rangle &= \alpha \langle w, v \rangle + \langle w, u \rangle \end{aligned}$$

for $\alpha \in \mathbb{C}$, $u, v, w \in M(h,c)$. For example,

$$\langle d_{-1} v^+, d_{-1} v^+ \rangle = \langle v^+, d_1 d_{-1} v^+ \rangle = \langle v^+, (d_{-1} d_1 + 2d_0) v^+ \rangle = \langle v^+, 2hv^+ \rangle = 2h.$$

Define the *radical* of $\langle , \rangle$ by

$$Rad\langle , \rangle = \{ v \in M(h,c) | \langle v, w \rangle = 0 \ \text{ for all } w \in M(h,c) \}.$$

Lemma 1. [6]

- $Rad\langle , \rangle = J(h,c)$.
- $M(h,c)^{(h+n,c)} \perp M(h,c)^{(h+m,c)}$ *with respect to* $\langle , \rangle$ *if* $m \neq n$.

This lemma has the following implications: the form provides information about submodules of $M(h,c)$, and we can consider how the form acts weight space by weight space. Also note that this lemma implies the Shapovalov form is well-defined on the quotient $L(h,c)$.

The *Shapovalov determinant* for $M(h,c)^{(h+n,c)}$ is

$$det M(h,c)^{(h+n,c)} = det \left(\langle d_{-\lambda} v^+, d_{-\mu} v^+ \rangle \right)_{\lambda, \mu \vdash n}.$$

The determinant $det M(h,c)^{(h+n,c)}$ gives information about the submodule structure of $M(h,c)$:

$$(J(h,c))^{(h+n,c)} \neq 0 \quad \text{if and only if} \quad det M(h,c)^{(h+n,c)} = 0.$$

Kac [5] and Feigin and Fuchs [1] have found explicit formulas for $det M(h,c)^{(h+n,c)}$.

Theorem 2. [4], [1] *Define*

$$\begin{aligned} \mathcal{C}_{r,s}(h,c) &= \left(48h - \left((13-c)(r^2+s^2) - 24rs - 2 + 2c\right)\right)^2 - (c-1)(c-25)(r^2-s^2)^2 \text{ if } r \neq s, \\ \mathcal{C}_{r,r}(h,c) &= 48h - \left(2r^2(13-c) - 24r^2 - 2 + 2c\right) = 48h - 2(r^2-1)(1-c), \end{aligned}$$

and

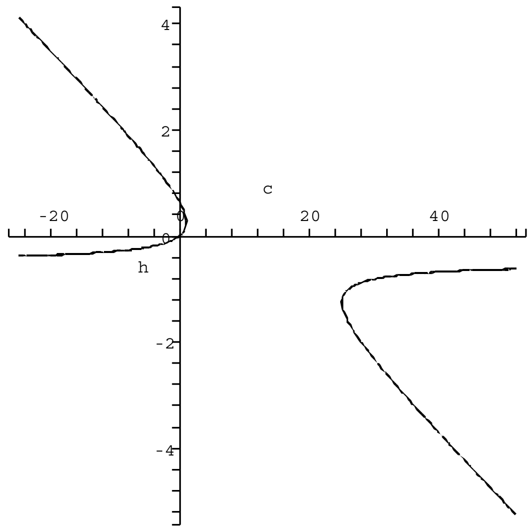
$$K_n = \prod_{\substack{r,s \in \mathbb{Z}_{>0}, \\ 1 \le r \le s \le n}} ((2r)^s s!)^{p(n-rs) - p(n-r(s+1))},$$

where $p(k)$ *is the number of partitions of* k. *Then for* $(h,c) \in \mathbb{R}^2$,

$$det M(h,c)^{(h+n,c)} = K_n \prod_{\substack{r,s \in \mathbb{Z}_{>0}, \\ 1 \le r \le s \le n}} (\mathcal{C}_{r,s}(h,c))^{p(n-rs)}.$$

Feigin and Fuchs [1] have shown that $detM(h,c)^{(h+n,c)}$ actually contains *all* the information about the submodule structure of $M(h,c)$. That is, the formulas for $detM(h,c)^{(h+n,c)}$ can be used to completely describe the submodule structure of $M(h,c)$. The statement and proof of this result (theorem 3) depend on the following analysis of the curves $\mathcal{C}_{r,s}(h,c) = 0$.

For fixed r, s, $r \neq s$, the curves $\mathcal{C}_{r,s}(h,c) = 0$ are hyperbolas. Below is the curve $\mathcal{C}_{1,2}(h,c) = 0$.



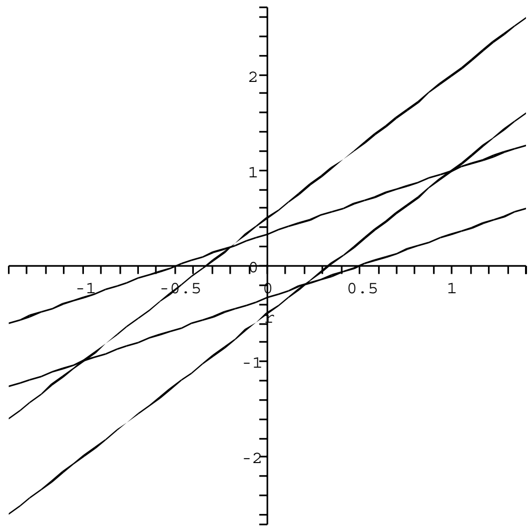
For fixed (h,c), the equation $\mathcal{C}_{r,s}(h,c) = 0$ can be factored into linear terms

$$0 = \mathcal{C}_{r,s}(h,c) = K(pr + qs + m)(pr + qs - m)(ps + qr + m)(ps + qr - m)$$

where $K, p, q, m \in \mathbb{C}$ such that

$$\frac{p}{q} + \frac{q}{p} = \frac{c-13}{6} \quad \text{and} \quad 4pqh + (p+q)^2 = m^2.$$

Thus, for fixed (h,c),the solutions to the equation $\mathcal{C}_{r,s}(h,c)$ form two sets of parallel lines. The figure below illustrates the example $\mathcal{C}_{r,s}(0,0) = 0$.



To find all integer solutions to $\mathcal{C}_{r,s}(h,c) = 0$, we only need to consider one line, say $pr+qs+m = 0$. (If (r,s) is a point on any of the other lines, $(-r,-s)$, (s,r) or $(-s,-r)$ will lie on the line $pr+qs+m = 0$.) We fix one of the lines and call it $\mathcal{L}_{(h,c)}$. The integer points (r,s) on this line encode the embeddings $M(h',c') \subseteq M(h,c) \subseteq M(h'',c'')$.

The line $\mathcal{L}_{(h,c)}$ passes through 0, 1, or infinitely many integer points. (If the line passes through two integer points, it has rational slope and therefore passes through infinitely many integer points.) Also, the line $\mathcal{L}_{(h,c)}$ has nonzero slope. Therefore, if it passes through infinitely many integer points (r,s) with $rs > 0$ it must pass through finitely many points (r,s) with $rs < 0$, and vice versa.

Theorem 3. [1] *Fix a pair $(h,c) \in \mathbb{R}^2$, and let $\mathcal{L}_{(h,c)}$ be one of the lines defined by this pair. Then the Verma module embeddings involving $M(h,c)$, and thus the block containing (h,c), are described by one of the following four cases.*

(1) *Suppose $\mathcal{L}_{(h,c)}$ passes through no integer points. Then the block $[(h,c)]$ is given by*

$$[(h,c)] = \{(h,c)\}.$$

The Verma module embeddings for $M(h,c)$ look like

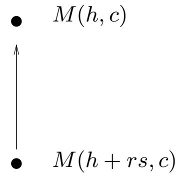
$$\bullet \quad M(h,c)$$

i.e, the Verma module $M(h,c)$ is irreducible and does not embed in any other Verma modules.

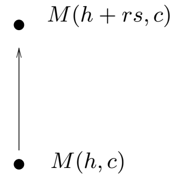
(2) *Suppose $\mathcal{L}_{(h,c)}$ passes through exactly one integer point (r,s). Then the block $[(h,c)]$ is given by*

$$[(h,c)] = \{(h,c), (h+rs,c)\}.$$

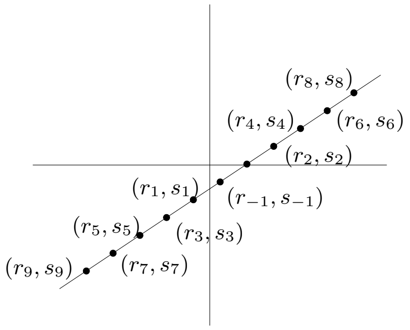
- *If $rs > 0$, the embeddings for $M(h,c)$ look like*



  where the arrow indicates inclusion.
- *If $rs < 0$, the embeddings for $M(h,c)$ look like*



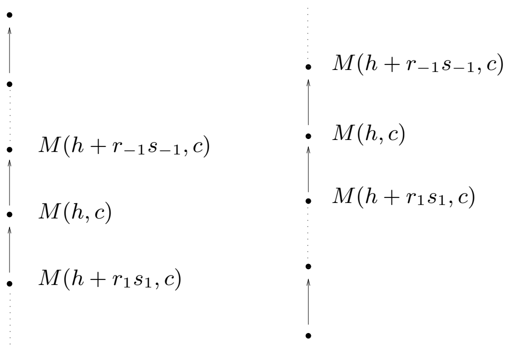
(3) *Suppose $\mathcal{L}_{(h,c)}$ passes through infinitely many integer points and crosses an axis at an integer point. Label these points (r_i,s_i) so that $\ldots < r_{-2}s_{-2} < r_{-1}s_{-1} < 0 < r_1s_1 < r_2s_2 \ldots$. (We exclude points (r,s) where $r=0$ or $s=0$.)*



Then, the block $[(h,c)]$ is given by

$$[(h,c)] = \{(h,c),(h+r_is_i,c)\}.$$

The embeddings between the corresponding Verma modules take one of the following forms:

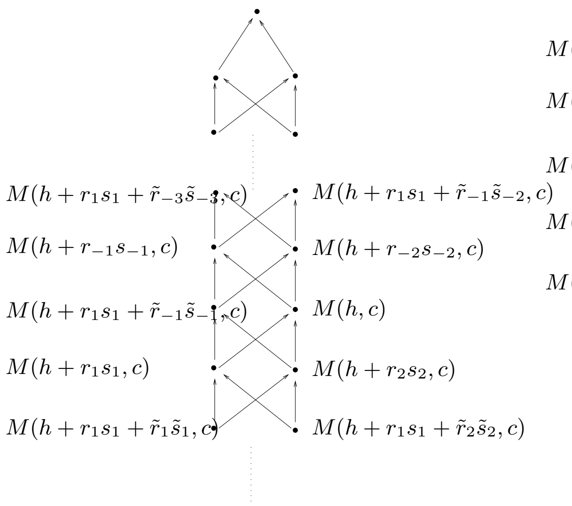


slope $(\mathcal{L}_{(h,c)}) > 0$ *slope* $(\mathcal{L}_{(h,c)}) < 0$

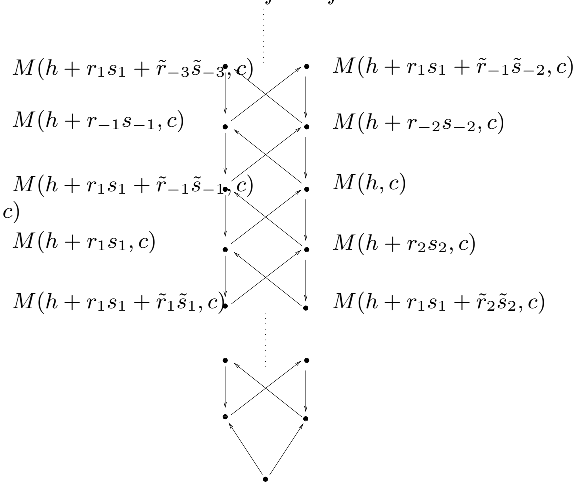
(4) *Suppose* $\mathcal{L}_{(h,c)}$ *passes through infinitely many integer points and does not cross either axis at an integer point. Again label the integer points* (r_i, s_i) *on* $\mathcal{L}_{(h,c)}$ *so that* $\ldots < r_{-2}s_{-2} < r_{-1}s_{-1} < 0 < r_1s_1 < r_2s_2 \ldots$. *Also consider the auxiliary line* $\tilde{\mathcal{L}}_{(h,c)}$ *with the same slope as* $\mathcal{L}_{h,c}$ *passing through the point* $(-r_1, s_1)$. *Label the integer points on this line* $(\tilde{r}_j, \tilde{s}_j)$ *as above. Then,*

$$[(h,c)] = \{(h + r_i s_i, c), (h + r_1 s_1 + \tilde{r}_j \tilde{s}_j, c)\}.$$

The embeddings between the corresponding Verma modules take one of the forms



slope $(\mathcal{L}_{(h,c)}) > 0$



slope $(\mathcal{L}_{(h,c)}) < 0$

We can use the line $\mathcal{L}_{(h,c)}$ to generate lines corresponding to the entire block $[(h,c)]$. If (r,s) is an integer point on the line $\mathcal{L}_{(h,c)}$, let $\tilde{\mathcal{L}}_{(h,c)}$ be the line with the same slope as $\mathcal{L}_{(h,c)}$ and passing through the point $(-r,s)$. Then $\tilde{\mathcal{L}}_{(h,c)}$ corresponds to the weight $(h + rs, c) \in [(h,c)]$. This allows us to identify maximal or minimal elements of a block. The weight (h,c) is a maximal (respectively minimal) element of its block if and only if there are no integer points (r,s) on $\mathcal{L}_{(h,c)}$ such that $rs < 0$ (respectively $rs >$).

In the following proposition, we identify a line $\mathcal{L}$ with the triple (μ, a, b), where μ is the slope of the line and (a,b) is a point on the line. Then $\mathcal{L}$ determines a weight (h,c) by

$$h = \frac{(a\mu - b)^2 - (\mu - 1)^2}{4\mu}, \quad c = 13 - 6\left(\mu + \frac{1}{\mu}\right).$$

Theorem 4. • *Blocks of size two are indexed by triples*

$$\{(\mu, a, b) | \mu \in \mathbb{R} - \mathbb{Q} \text{ with } |\mu| \leq 1 \text{ and } a, b \in \mathbb{Z}_{>0}\}.$$

The weights in a block of size two are determined by triples $\{(\mu, \pm a, b)\}$.

- *Infinite blocks with a maximal element are indexed by triples*

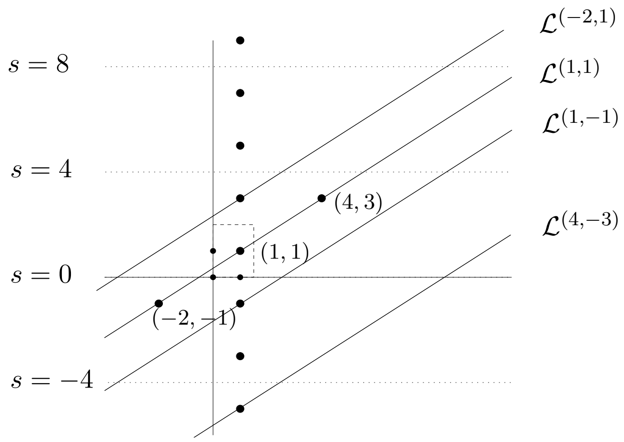
$$\left\{ \left(\frac{p}{q}, a, b\right) \middle| \begin{array}{l} p, q \in \mathbb{Z}_{>0},\ \textit{with } gcd(p,q) = 1, p < q \\ \textit{If } 2 \nmid q,\ \textit{then } 0 \le a < \frac{q}{2},\ 0 \le b < p \\ \textit{If } 2|q,\ \textit{then } 0 \le a < q,\ 0 \le b < \frac{p}{2} \end{array} \right\}.$$

Infinite blocks with a minimal element are indexed by triples

$$\left\{ \left(-\frac{p}{q}, -a, b\right) \middle| \begin{array}{l} p, q \in \mathbb{Z}_{>0},\ \textit{with } gcd(p,q) = 1, p < q \\ \textit{If } 2 \nmid q,\ \textit{then } 0 \le a < \frac{q}{2},\ 0 \le b < p \\ \textit{If } 2|q,\ \textit{then } 0 \le a < q,\ 0 \le b < \frac{p}{2} \end{array} \right\}.$$

The weights in an infinite block are determined by triples $\left\{ \left(\frac{p}{q}, a, \pm b + 2kp\right) | k \in \mathbb{Z} \right\}$ *(for blocks with a maximal element) and* $\left\{ \left(-\frac{p}{q}, -a \pm b + 2kp\right) | k \in \mathbb{Z} \right\}$ *(for blocks with a minimal element).*

Consider the example with $\mu = \frac{2}{3}$. We denote the line with slope $\frac{2}{3}$ and passing through the point (a,b) by $\mathcal{L}^{(a,b)}$.



The set of integer points $\{(a,b) \in \mathbb{Z}^2 | 0 \le a < \frac{3}{2}, 0 \le b < 2\}$ indexes the infinite blocks with $c = 13 - 6(\frac{2}{3} + \frac{3}{2}) = 0$. The line $\mathcal{L}^{(1,1)}$ determines the weight $(0,0)$. From the integer points $(1,1)$, $(-2,-1)$, and $(4,3)$ on the line $\mathcal{L}^{(1,1)}$, we get the lines $\mathcal{L}^{(1,-1)}$, $\mathcal{L}^{(-2,1)}$, and $\mathcal{L}^{(4,-3)}$; these lines determine the weights $(1,0)$, $(2,0)$, and $(12,0)$ respectively. In general, the set of points $\{(1, 4k \pm 1) | k \in \mathbb{Z}\}$ correspond to the block

$$\left\{ \left(\frac{(12k + 2 \pm 3)^2 - 1}{24}, 0\right) | k \in \mathbb{Z} \right\} = \left\{ \left(\frac{j(3j \pm 1)}{2}, 0\right) | j \in \mathbb{Z}_{\ge 0} \right\}.$$

4. Translation Functors

We now consider $M(h,c)\otimes L(h',c')$ and its decomposition by blocks. The module $M(h,c)\otimes L(h',c')$ can be written as a direct sum of submodules

$$M(h,c)\otimes L(h',c')=\bigoplus_{[\mu]\in[\mathfrak{h}^\star]}\left(M(h,c)\otimes L(h',c')\right)^{[\mu]}$$

where $(M(h,c)\otimes L(h',c'))^{[\mu]}$ is such that $L(\nu)$ appears in $(M(h,c)\otimes L(h',c'))^{[\mu]}$ only if ν is in the block $[\mu]$. For a given choice of $(h',c')\in\mathbb{C}^2$ and $[\mu]\in[\mathfrak{h}^\star]$, a *translation functor* is the map sending $M(h,c)$ to $(M(h,c)\otimes L(h',c'))^{[\mu]}$. Our goal in this section is to describe the submodules $(M(h,c)\otimes L(h',c'))^{[\mu]}$ corresponding to each block $[\mu]$. The Shapovalov determinant will again prove to be a useful tool.

We begin by observing that for all $(h,c),(h',c')\in\mathbb{C}^2$, the weights of $M(h,c)\otimes L(h',c')$ are of the form $(h+h'+n,c+c')$, $n\in\mathbb{Z}_{\geq 0}$. Therefore, $(M(h,c)\otimes L(h',c'))^{[\mu]}\neq 0$ only if $[\mu]=[(h+h'+n,c+c')]$ for some $n\in\mathbb{Z}_{\geq 0}$.

Proposition 3. [6] *For $(h,c),(h',c')\in\mathbb{C}^2$ and $[\mu]\in[\mathfrak{h}^\star]$, $(M(h,c)\otimes L(h',c'))^{[\mu]}$ has a filtration*

$$0=M_0\subseteq M_1\subseteq M_2\subseteq\cdots$$

where

- $(M(h,c)\otimes L(h',c'))^{[\mu]}=\bigcup_i M_i$;
- $M_{i+1}/M_i\cong M(h+h'+\tilde{n},c+c')$ *for some* $(h+h'+\tilde{n},c+c')\in[\mu]$;
- *the multiplicity of the factor* $M(h+h'+\tilde{n},c+c')$ *is* $\dim L(h',c')^{(h'+\tilde{n},c')}$.

We will use the Shapovalov form and the Shapovalov determinant to say more about $(M(h,c)\otimes L(h',c'))^{[\mu]}$.

Let $(h,c),(h',c')\in\mathbb{R}^2$. Recall that the Shapovalov form is defined on both $M(h,c)$ and $L(h',c')$. Define a form on $M(h,c)\otimes L(h',c')$ by

$$\langle v\otimes w,\tilde{v}\otimes\tilde{w}\rangle=\langle v,\tilde{v}\rangle\langle w,\tilde{w}\rangle.$$

Observe that we still have

$$\langle d_k(v\otimes w),\tilde{v}\otimes\tilde{w}\rangle=\langle v\otimes w,d_{-k}(\tilde{v}\otimes\tilde{w})\rangle$$

for all $x\in Vir$. For each weight space of $L(h',c')^{(h'+j,c')}$, we fix a basis $\{w_{j,i}|1\leq i\leq \dim L(h',c')^{(h'+j,c')}\}$ and define $det L(h',c')^{(h'+j,c')}=det\left(\langle w_{j,i},w_{j,k}\rangle\right)$. Then

$$\bigcup_{j=0}^{n}\{d_{-\lambda}v^+\otimes w_{j,i}|1\leq i\leq\dim L(h',c')^{(h'+j,c')},\lambda\vdash n-j\}$$

is a basis for $(M(h,c)\otimes L(h',c'))^{(h+h'+n,c+c')}$. Define

$$det\left(M(h,c)\otimes L(h',c')\right)^{(h+h'+n,c+c')}=det\left(\langle d_{-\lambda}v^+\otimes w_{j,i},d_{-\mu}\otimes w_{j',k}\rangle\right)$$

where $0\leq j,j'\leq n$, $1\leq i\leq\dim L(h',c')^{(h'+j,c')}$, $1\leq k\leq\dim L(h',c')^{(h'+j',c')}$, $\lambda\vdash n-j$, and $\mu\vdash n-j'$. The following lemma is a straightforward calculation, given this definition.

Lemma 2. *Let $(h,c),(h',c') \in \mathbb{R}^2$ and $n \in \mathbb{Z}_{\geq 0}$. Then*

$$det\left(M(h,c)\otimes L(h',c')\right)^{(h+h'+n,c+c')}$$

is given by

$$\prod_{0\leq j\leq n} \left(a_j^{(h',c')}(h,c)\ detL(h',c')^{(h'+j,c')}\right)^{p(n-j)} \times \left(detM(h+h'+j,c+c')^{(h+h'+n,c+c')}\right)^{\dim L(h',c')^{(h'+j,c')}},$$

where

$$a_j^{(h',c')}(h,c) = \prod_{\substack{1\leq r\leq s\\ rs\leq j}} \left(\frac{\mathcal{C}_{r,s}(h,c)}{\mathcal{C}_{r,s}(h+h'+j-rs,c+c')}\right)^{\dim L(h',c')^{(h'+j-rs,c')}}.$$

Proposition 4. *Let $(h',c') \in \mathbb{R}^2$ and let $(h,c) \in \mathbb{R}^2$ be generic. For a block $[\mu]$ and $n \in \mathbb{Z}_{\geq 0}$,*

$$det\left(\left(M(h,c)\otimes L(h',c')\right)^{[\mu]}\right)^{(h+h'+n,c+c')}$$

is given by

$$\prod_j \left(a_j^{(h',c')}(h,c) detL(h',c')^{(h'+j,c')}\right)^{p(n-j)} \left(detM(h+h'+j,c+c')^{(h+h'+n,c+c')}\right)^{\dim L(h',c')^{(h'+j,c')}}.$$

where the product is over j such that $0\leq j\leq n$ and $(h+h'+j,c+c') \in [\mu]$.

Proof. With respect to the Shapovalov form

$$\left(M(h,c)\otimes L(h',c')\right)^{[\lambda]} \perp \left(M(h,c)\otimes L(h',c')\right)^{[\mu]} \quad \text{for } [\lambda]\neq[\mu]. \tag{1}$$

Now fix $n \in \mathbb{Z}_{\geq 0}$ and $(h',c') \in \mathbb{R}^2$.

For generic (h,c), $det\left(M(h,c)\otimes L(h',c')\right)^{(h+h'+n,c+c')}$ is nondegenerate. Given (1), the Shapovalov form can be used to construct explicit projection maps

$$Pr_{[\mu]} : M(h,c)\otimes L(h',c') \to \left(M(h,c)\otimes L(h',c')\right)^{[\mu]}.$$

By applying these maps to the basis for $\left(M(h,c)\otimes L(h',c')\right)^{(h+h'+n,c+c')}$ given above, we obtain bases for the weight spaces of each block . We use these bases to define $det\left(\left(M(h,c)\otimes L(h',c')\right)^{[\mu]}\right)^{(h+h'+n,c+c')}$. From (1), we have

$$det\left(M(h,c)\otimes L(h',c')\right)^{(h+h'+n,c+c')} = \prod_{[\mu]\in[\mathfrak{h}^\star]} det\left(\left(M(h,c)\otimes L(h',c')\right)^{[\mu]}\right)^{(h+h'+n,c+c')}.$$

Suppose $(h,c)\in\mathbb{R}^2$ is such that the block $[\mu] = [(h+h'+j,c+c')]$ $(j\leq n)$ has size one. Proposition 3 then implies that

$$\left(M(h,c)\otimes L(h',c')\right)^{[\mu]} \cong M(h+h'+j,c+c')^{\oplus \dim L(h',c')^{(h'+j,c')}}.$$

By inducting on n, we show that in this case $det\left((M(h,c)\otimes L(h',c'))^{[\mu]}\right)^{(h+h'+n,c+c')}$ is given by

$$\begin{aligned} (2) \qquad & \left(a_j^{(h',c')}(h,c) det L(h',c')^{(h'+j,c')}\right)^{p(n-j)} \\ \times \quad & \left(det M(h+h'+j,c+c')^{(h+h'+n,c+c')}\right)^{\dim L(h',c')^{(h'+j,c')}}. \end{aligned}$$

Now, for a given $c \in \mathbb{R}$, there are infinitely many $h \in \mathbb{R}$ so that $[(h+h'+j,c+c')]$ $(j \le n)$ has size one. We use (2) to prove the result for all generic $(h,c) \in \mathbb{R}^2$. $\square$

This determinant formula reflects the Verma filtration of $(M(h,c)\otimes L(h',c'))^{[\mu]}$, with the $a_j^{(h',c')}(h,c)$ acting as error terms. In fact, these error terms contain more information about the structure of $(M(h,c)\otimes L(h',c'))^{[\mu]}$. For example, the following proposition shows that the error terms count the dimension of the radical.

Proposition 5. *Let $(h,c),(h',c') \in \mathbb{R}^2$, and suppose (h,c) does not belong to an infinite block. Define*

$$Rad\langle,\rangle_{[\mu]} = \left\{m \in \left(M(h,c)\otimes L(h',c')\right)^{[\mu]} | \langle m,x\rangle = 0 \text{ for all } x \in M(h,c)\otimes L(h,c)\right\}.$$

Then $\dim\left(Rad\langle,\rangle_{[\mu]}\right)^{(h+h'+n,c+c')}$ *is the number of zeros in*

$$\prod_{\substack{0 \le j \le n \\ (h+h'+j,c+c')\in[\mu]}} \left(a_j^{(h',c')}(h,c)\right)^{p(n-j)}.$$

References

[1] B. L. Feĭgin and D. B. Fuchs. Representations of the Virasoro algebra. 7:465–554, 1990.

[2] O. Gabber and A. Joseph. Towards the Kazhdan-Lusztig conjecture. *Ann. Sci. École Norm. Sup. (4)*, 14(3):261–302, 1981.

[3] J. C. Jantzen. Darstellungen halbeinfacher algebraischer gruppen und zugeordnete kontravariante formen. Thesis, 1973.

[4] V. Kac and A. Raina. *Bombay Lectures on highest weight representations of infinite dimensional Lie algebras*, volume 2 of *Advanced Series in Mathematical Physics*. World Scientific, New Jersey.

[5] V. G. Kac. Highest weight representations of infinite-dimensional Lie algebras. In *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, pages 299–304, Helsinki, 1980. Acad. Sci. Fennica.

[6] R. Moody and A. Pianzola. *Lie Algebras with Triangular Decompositions*. Canadian Mathematical Society series of monographs and advanced texts. John Wiley & Sons, Inc., New York, 1995.

[7] N. N. Shapovalov. On a bilinear form on the universal enveloping algebra of a complex semisimple lie algebra. *Func. Anal. Appl.*, 6:307–312, 1972.

University of Wisconsin-Madison, 480 Lincoln Dr, Madison, WI 53706, USA

E-mail address: wiesner@math.wisc.edu

WHEN IS A SCHUBERT VARIETY GORENSTEIN?

ALEXANDER WOO AND ALEXANDER YONG

ABSTRACT. A variety is Gorenstein if it is Cohen-Macualay and its canonical sheaf is a line bundle. This property implies a variety behaves like a smooth one for various algebro-geometric purposes. We introduce a new notion of pattern avoidance involving Bruhat order and use it to characterize which Schubert varieties are Gorenstein. We also give an explicit description as a line bundle of the canonical sheaf of a Gorenstein Schubert variety.

1. INTRODUCTION

This extended abstract is a shortened version of the paper [31], with details of the proofs omitted. The main goal of the paper is to give an explicit combinatorial characterization of which Schubert varieties in the complete flag variety are Gorenstein.

Let $\mathrm{Flags}(\mathbb{C}^n)$ denote the variety of complete flags $F_\bullet : \langle 0\rangle \subseteq F_1 \subseteq \ldots \subseteq F_n = \mathbb{C}^n$. Fix a basis $e_1, e_2, \ldots, e_n$ of $\mathbb{C}^n$ and let $E_\bullet$ be the *anti*-canonical reference flag $E_\bullet$, that is, the flag where $E_i = \langle e_{n-i+1}, e_{n-i+2}, \ldots, e_n\rangle$. For every permutation w in the symmetric group S_n, there is the **Schubert variety**

$$X_w = \big\{F_\bullet \mid \dim(E_i \cap F_j) \geq \#\{k \geq n-i+1, w(k) \leq j\}\big\}.$$

These conventions have been arranged so that the codimension of X_w is $\ell(w)$, that is, the length of any expression for w as a product of simple reflections $s_i = (i \leftrightarrow i+1)$.

Gorensteinness is a well-known technical condition which implies that a variety behaves like a smooth one for various algebro-geometric purposes. In particular, smooth varieties are Gorenstein and Gorenstein varieties are by definition Cohen-Macaulay. A variety is **Gorenstein** if it is Cohen-Macaulay and its canonical sheaf is a line bundle. (Throughout this paper we freely identify vector bundles and their sheaves of sections for convenience.) Recall that on a *smooth* variety X, the **canonical sheaf**, denoted ω_X is $\wedge^{\dim(X)}\Omega_X$, where Ω_X is the cotangent bundle of X. For a possibly singular but normal variety, it is the pushforward of the canonical sheaf $\omega_{X_{\mathrm{smooth}}}$ of the smooth part X_{smooth} of X under the inclusion map. In fact, every Schubert variety is normal [10, 26] and Cohen-Macaulay [27], therefore, the above remarks actually suffice to give a complete definition of Gorensteinness for our purposes.

Gorensteinness can also be determined locally using a free resolution. Examples of Gorenstein varieties include all (normal) hypersurfaces. The variety of $m \times n$ matrices of rank at most r is Gorenstein iff $m = n$ (or, trivially, $r = 0$ or $r = \min(m, n)$); this follows either from the characterization of Gorenstein Schubert varieties on the Grassmannian, originally due to Svanes [30] and recoverable from our results here, or in characteristic 0 from the construction of a free resolution due originally to Lascoux [19].

Date: May 9, 2005.

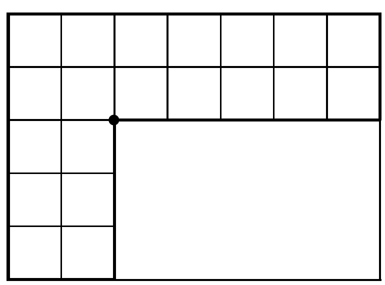
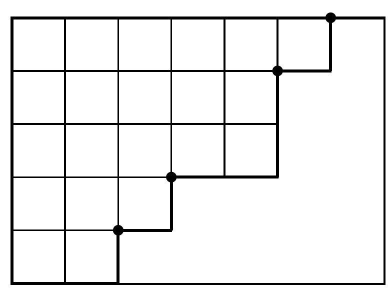
2000 *Mathematics Subject Classification.* 14M15; 14M05, 05E99.

Smoothness and Cohen-Macaulayness of Schubert varieties have been extensively studied in the literature; see, for example, [3, 27] and the references therein. While all Schubert varieties are Cohen-Macaulay, actually very few Schubert varieties are smooth. (See the table near the end of this extended abstract.) Explicitly, X_w is smooth if and only if w is "1324-pattern avoiding" and "2143-pattern avoiding" [20].

Our main result (Theorem 1) gives an explicit combinatorial characterization of Gorensteinness similar to the above smoothness criteria. This answers a question raised by M. Brion and S. Kumar, and that was passed along to us by A. Knutson; see also [28, p. 88]. Our answer uses a generalized notion of pattern avoidance that we will define below.

2. The Grassmannian Case

First let us present the answer for Schubert varieties on the Grassmannian, due originally to T. Svanes in [30]; this case will illustrate one of the conditions given in our main theorem, and can also be derived as a special case of it. On the Grassmannian, Schubert varieties X_λ are indexed by partitions λ sitting inside an $\ell \times (n-\ell)$ rectangle; we use the convention that $|\lambda|$ is the codimension of X_λ in $Gr(\ell, n)$. The smooth Schubert varieties are those indexed by partitions λ whose complement in $\ell \times (n-\ell)$ is a rectangle, see, for example, [3] and the references therein. For example, $\lambda = (7,7,2,2,2)$ indexes a smooth Schubert variety in $\mathrm{Gr}(5,12)$.

$\lambda =$  $\mu =$ 

Alternatively, smooth Schubert varieties are those with exactly one inner corner. View the lower border of partition as a lattice path from the lower left-hand corner to the upper right-hand corner of $\ell \times (n-\ell)$; then an **inner corner** is a lattice point that has a lattice point of the path both directly below and directly to the right of it. The inner corners for the partitions λ and μ above are marked by "dots".

Therefore, the partition $\mu = (6,5,5,3,2)$ above does not index a smooth Schubert variety. However, it does index a Gorenstein Schubert variety; in general, a Grassmannian Schubert variety X_μ is Gorenstein if and only if all of the inner corners of μ sit on the same antidiagonal.

3. Main Definitions

In order to state our main results for a Schubert variety X_w of the flag variety $\mathrm{Flags}(\mathbb{C}^n)$, we will need some preliminary definitions. First we need to associate a partition to each descent of w, and define the associated inner corner distance. Secondly we need to define a new notion of pattern avoidance which we call Bruhat-restricted pattern avoidance.

Let d be a **descent** of w, that is, an index where $w(d) > w(d+1)$. Now write w in one-line notation as $w(1)w(2)\cdots w(n)$, and construct a subword $v_d(w)$ of w by concatenating the right-to-left minima of the segment strictly to the left of $d+1$ with the left-to-right maxima of the segment strictly to the right of d. In particular, $v_d(w)$ will necessarily include $w(d)$ and $w(d+1)$. Let $\widetilde{v}_d(w)$ denote the **flattening** of $v_d(w)$, that is, the unique permutation whose relative position of its entries matches that of $v_d(w)$.

Example 1. Let $w = 314972658 \in S_9$. This permutation has descents at positions 1, 4, 5 and 7. We see that $v_1(w) = 3149$, $v_4(w) = 14978$, $v_5(w) = 147268$, and $v_7(w) = 12658$, so therefore $\widetilde{v}_1(w) = 2134$, $\widetilde{v}_4(w) = 12534$, $\widetilde{v}_5(w) = 135246$, and $\widetilde{v}_7(w) = 12435$.

By construction, $\widetilde{v}_d(w) \in S_m$ is a **Grassmannian permutation**, that is, it has a unique descent at, say, position e. For any Grassmannian permutation $w \in S_m$ with its unique descent at e, let $\lambda(w) \subseteq e \times (m-e)$ denote the associated partition. This is obtained by drawing a lattice path starting from the lower left-hand corner of $e\times(m-e)$ and drawing a unit horizontal line segment at step $i = 1, 2, \ldots, m$ if i appears strictly *after* position e, and a unit vertical line segment otherwise. For example, the Grassmannian permutation $w = 3589\,11 \mid 12467\,10\,12$ corresponds to the partition $\lambda(w) = \mu = (6, 5, 5, 3, 2)$ depicted above. Now, given an inner corner of a partition $\lambda(w)$, let its **inner corner distance** be the sum of the distances from the inner corner to the top and left edges of the rectangle $e \times (m-e)$. Furthermore, suppose that $\lambda(w)$ has all its inner corners on the same antidiagonal; this is equivalent to requiring that the inner corner distance be the same for all inner corners. In this case we call this common inner corner distance $\mathfrak{I}(w)$; if there are no inner corners, we set $\mathfrak{I}(w) = 0$ by convention. For example, in μ above, all the inner corner distances equal 6.

Now we define Bruhat-restricted pattern avoidance. First we recall the classical notion of pattern avoidance and the Bruhat order on S_n. For $v \in S_\ell$ and $w \in S_n$, with $\ell \le n$, an **embedding of** v **into** w is a sequence of indices $i_1 < i_2 < \cdots < i_\ell$ such that, for all $1 \le a < b \le \ell$, $w(i_a) > w(i_b)$ if and only if $v(a) > v(b)$. Then w **pattern avoids** v if there are no embeddings of v into w.

The **Bruhat order** on S_n, which we will denote by $\succ$, is defined as follows. First we say that $w(i \leftrightarrow j)$ **covers** w if $i < j$, $w(i) < w(j)$, and, for each k with $i < k < j$, either $w(k) < w(i)$ or $w(k) > w(j)$; then the Bruhat order is the transitive closure of this covering relation. The Bruhat order is graded by the length of a permutation, and one can check that v can cover w only if $\ell(v) = \ell(w) + 1$.

Now let $\mathcal{T}_v = \{(m_1 \leftrightarrow n_1), \ldots, (m_k \leftrightarrow n_k)\}$ be a set of **Bruhat transpositions** for v, where a Bruhat transposition $(m_j \leftrightarrow n_j)$ is one such that $v \cdot (m_j \leftrightarrow n_j)$ covers v in the Bruhat order. A $\mathcal{T}_v$**-restricted embedding** of v into w is an embedding of v into w such that $w \cdot (i_{m_j} \leftrightarrow i_{n_j})$ covers w (in the Bruhat order) for all $(m_j \leftrightarrow n_j) \in \mathcal{T}_v$. Then w **pattern avoids** v **with Bruhat restrictions** $\mathcal{T}_v$ if there are no $\mathcal{T}_v$-restricted embeddings of v into w. For example, the Bruhat transpositions for 31524 are $(1 \leftrightarrow 3)$, $(1 \leftrightarrow 5)$, $(2 \leftrightarrow 3)$, $(2 \leftrightarrow 4)$, and $(4 \leftrightarrow 5)$. This can be indicated by brackets drawn under the permutation as in Figure 1; the emptiness of the shaded rectangles in the graph of the permutation shows that $(1 \leftrightarrow 5)$ and $(2 \leftrightarrow 3)$ are indeed Bruhat transpositions. Using these "bracket diagrams", being a $\mathcal{T}_v$-restricted embedding means that the brackets associated to the transpositions in $\mathcal{T}_v$ are present in the "bracket diagram" for w. For example, there is no $\{(2 \leftrightarrow 4)\}$-restricted embedding of 2143 into 31524 since such an embedding would require a bracket between the 1 and the 4.

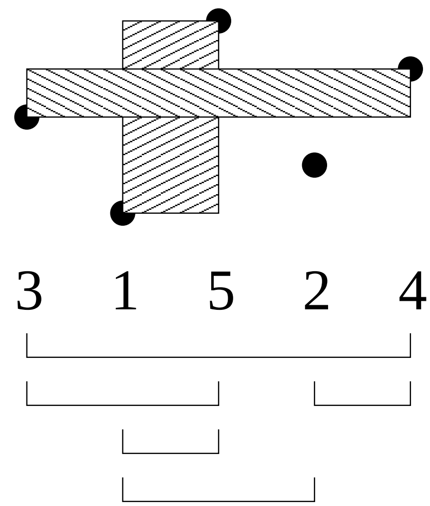


FIGURE 1. Bruhat transpositions for $w = 31524$

4. MAIN THEOREMS

Now we are ready to state our combinatorial characterization of Gorensteinness for Schubert varieties in $\mathrm{Flags}(\mathbb{C}^n)$:

Theorem 1. *Let $w \in S_n$. The Schubert variety X_w is Gorenstein if and only if for each descent d of w, $\lambda(\widetilde{v}_d(w))$ has all of its inner corners on the same antidiagonal and w pattern avoids both 31524 and 24153 with Bruhat restrictions $\{(1 \leftrightarrow 5), (2 \leftrightarrow 3)\}$ and $\{(1 \leftrightarrow 5), (3 \leftrightarrow 4)\}$ respectively.*

It is possible to replace the conditions on the descents with an infinite list of patterns to be avoided with certain Bruhat restrictions; this list contains 2 patterns in S_k for each odd $k \geq 5$.

By combining Theorem 1 with the descriptions of the singularities along the "maximal singular locus" of a Schubert variety X_w given in [9, 22], we obtain the following purely geometric corollary.

Corollary 1. *A Schubert variety X_w is Gorenstein if and only if it is Gorenstein along its maximal singular locus.*

The proof follows from classical results characterizing which of the singularities described in [9, 22] are Gorenstein.

In comparing the smoothness characterization of [20] with Theorem 1, the considerations from our description of the Grassmannian case allow one to check that the 1324-pattern avoidance condition of the former implies the "inner corner condition" of the latter. It is also easy to see that the 2143-pattern avoidance condition of the former implies each of the Bruhat-restricted pattern avoidance conditions of the latter. We mention that Fulton [12] has characterized 2143-pattern avoidance in terms of the essential set of a permutation. A similar characterization can be given for the Bruhat-restricted pattern avoidance conditions of Theorem 1.

Example 2. The permutation $w = \underline{5}13\underline{28}4\underline{6}\underline{7} \in S_8$ has descents at positions $1, 3$ and 5 and we have

$$\widetilde{v}_1(w) = 3124, \widetilde{v}_3(w) = 1324, \text{ and } \widetilde{v}_5(w) = 126345.$$

Hence one checks that w satisfies the inner corner condition with

$$\mathfrak{I}(\widetilde{v}_1(w)) = 2,\ \mathfrak{I}(\widetilde{v}_3(w)) = 1,\ \text{and } \mathfrak{I}(\widetilde{v}_5(w)) = 1.$$

The Schubert variety X_w is Gorenstein, since there are no forbidden 31524 and 24153 patterns with Bruhat restrictions $\{(1 \leftrightarrow 5), (2 \leftrightarrow 3)\}$ or $\{(1 \leftrightarrow 5), (3 \leftrightarrow 4)\}$ respectively. Note that the underlined subword of w is a 31524-pattern, but since $w(1 \leftrightarrow 8)$ does not cover w, it does not prevent X_w from being Gorenstein.

We now describe the canonical sheaf of a Gorenstein Schubert variety as a line bundle. Let $T \cong (\mathbb{C}^*)^{n-1}$ be the subgroup of invertible diagonal matrices of determinant 1 in $\mathrm{SL}_n(\mathbb{C})$; the Borel-Weil construction associates to each integral weight $\alpha \in \mathrm{Hom}(T, \mathbb{C}^*)$ a line bundle $\mathcal{L}_\alpha$. Let $\mathcal{L}_\alpha\big|_{X_w}$ denote the restriction of this line bundle to X_w. We will write weights additively in terms of the $\mathbb{Z}$-basis of fundamental weights Λ_r, defined by

$$\Lambda_r\left(\begin{bmatrix} t_1 & & 0 \\ & \ddots & \\ 0 & & t_n \end{bmatrix}\right) = t_1 \cdots t_r.$$

Theorem 2. *If X_w is Gorenstein, then $\omega_{X_w} \cong \mathcal{L}_\alpha\big|_{X_w}$ where $\alpha = \sum_{r=1}^{n-1} \widetilde{\alpha}_r \Lambda_{n-r}$ and*

$$\widetilde{\alpha}_r = \begin{cases} -2 + \mathfrak{I}(\widetilde{v}_r(w)) & \text{if } r \text{ is a descent} \\ -2 & \text{otherwise.} \end{cases} \tag{1}$$

5. Applications and Problems

Theorem 1 extends with no difficulty to Schubert varieties on partial flag varieties. Theorem 2 also extends, though some further calculations are needed. Our results can also be extended to the matrix Schubert varieties originally defined in [12], and thereby used to recover previously known results on Gorensteiness of ladder determinantal varieties [8, 15].

More geometrically, Theorem 2 calculates the sheaf cohomology of some line bundles on Gorenstein Schubert varieties, and gives small hints towards a final theorem on this open problem. Theorem 2 also allows us to characterize which smooth Schubert varieties are Fano, and gives new examples of higher-dimensional Fano varieties.

Further study of the relations between the geometry of Gorensteinness of Schubert varieties and related combinatorics should have potential. The most natural question is:

Problem 1. *Give analogues of Theorems 1 and 2 for generalized flag varieties corresponding to Lie groups other than $\mathrm{SL}_n(\mathbb{C})$.*

We expect that the methods given in this paper will extend to solve Problem 1. For the case of the odd orthogonal groups $\mathrm{SO}_{2n+1}(\mathbb{C})$, the solution for $\mathrm{SL}_n(\mathbb{C})$ leads to an answer which, however, is not entirely in terms of a good generalization of Bruhat-restricted pattern avoidance. The other classical types are not completely understood as of this writing.

In analogy with the determination of the singular loci of singular Schubert varieties [4, 9, 14, 18, 20, 21, 22], it should also be interesting to determine the "non-Gorenstein locus" of a non-Gorenstein Schubert variety; as in the case of singular loci this will for geometric reasons be a union of Schubert subvarieties X_v of our Schubert variety X_w. Therefore, we ask:

Problem 2. *Give a combinatorial characterization for the minimal v in the Bruhat order for which X_w is non-Gorenstein at X_v.*

Presumably, the eventual answer (for $\mathrm{SL}_n(\mathbb{C})$) will have some interesting relationship with the combinatorial characterization of the maximal singular locus. Indeed, in view of Corollary 1, one can hope that the maximal non-Gorenstein locus of X_w is simply the union of those Schubert cells in the maximal singular locus at which X_w is not Gorenstein.

A geometric explanation was recently given in [2] for the appearance of pattern avoidance in characterizations of smooth Schubert varieties. However, this explanation does not have an obvious modification to take into account Bruhat-restrictions. This leads to the following:

Problem 3. *Give a geometric explanation of Bruhat-restricted pattern avoidance which explains its appearance in Theorem 1.*

Finally, for those interested in combinatorial enumeration:

Problem 4. *Give a combinatorial formula (for example, a generating series) computing the number of Gorenstein Schubert varieties in* $\mathrm{Flags}(\mathbb{C}^n)$.

Using the methods of this paper, we computed the number of Gorenstein Schubert varieties in $\mathrm{Flags}(\mathbb{C}^n)$ for some small values of n (see below). We compare this to the number of smooth Schubert varieties computed using the result of [20] (by the recursive formulas found in [5, 29]).

n	$n! = \#$ Cohen-Macaulay X_w	$\#$ Gorenstein X_w	$\#$ Smooth X_w
1	1	1	1
2	2	2	2
3	6	6	6
4	24	24	22
5	120	116	88
6	720	636	366
7	5040	3807	1552
8	40320	24314	6652
9	362880	163311	28696

6. Acknowledgements

We are very grateful to M. Brion, A. Knutson and S. Kumar for bringing the problem addressed by Theorem 1 to our attention, for outlining the geometric portion of our proof, which had previously been folklore, and for many other suggestions. We also thank A. Bertram, S. Billey, A. Buch, R. Donagi, S. Fomin, M. Haiman, R. MacPherson, E. Miller, R. Stanley, B. Sturmfels, J. Tymoczko, and anonymous referees for helpful remarks. This work was partially completed while the two authors were in residence at the Park City Mathematics Institute program on "Geometric Combinatorics" during July 2004.

REFERENCES

[1] V. Balaji, S. Senthamarai Kannan, K. V. Subrahmanyam, *Cohomology of line bundles on Schubert varieties-I.* Transform. Groups **9** (2004), no. 2, 105–131.
[2] S. Billey and T. Braden, *Lower bounds for Kazhdan-Luzstig polynomials from patterns*, Transform. Groups **8** (2003), 321–332.
[3] S. Billey and V. Lakshmibai, *Singular loci of Schubert varieties*, Progr. Math. **182** (2000), Birkhäuser, Boston.
[4] S. Billey and G. Warrington, *Maximal singular loci of Schubert varieties in* $\mathrm{SL}(n)/B$, Trans. Amer. Math. Soc. **335** (2003), 3915–3945.
[5] M. Bóna, *The permutation classes equinumerous to the smooth class*, Elect. J. Combinatorics, **5** (1998).
[6] R. Bott, *Homogeneous vector bundles*, Annals of Math. **66** (1957), no. 2, 203–248.
[7] M. Brion, *Lectures on the geometry of flag varieties*, arXiv:math.AG/0410240.
[8] A. Conca *Gorenstein ladder determinantal rings*, J. London Math. Soc. (2) **54** (1996), 453–474.
[9] A. Cortez, *Singularités génériques et quasi-résolutions des variétés de Schubert pour le groupe linéaire*, Adv. Math. **178** (2003), 396–445.
[10] C. DeConcini and V. Lakshmibai, *Arithmetic Cohen-Macaulayness and arithmetic normality for Schubert varieties*, Am. J. Math. **103** (1981), 835–850.
[11] M. Demazure, *Désingularisation des variétés de Schubert généralisées*, Ann. Sci. Ècole Norm. Sup. (4) **7** (1974), 53–88.
[12] W. Fulton, *Flags, Schubert polynomials, degeneracy loci, and determinantal formulas*, Duke Math. J. **65** (1992), no. 3, 381–420.
[13] ———, *Intersection theory*, second edition, Springer-Verlag (1998).
[14] V. Gasharov, *Sufficiency of Lakshmibai–Sandhya singularity conditions for Schubert varieties*, Compositio Math. **126** (2001), 47–56
[15] N. Gonciulea and C. Miller, *Mixed ladder determinantal varieties*, J. Algebra **231** (2000), 104–137.
[16] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York-Heidelberg, 1977.
[17] J. C. Jantzen, *Representations of Algebraic groups*, Pure and Appl. Math., Academic Press, 1987.
[18] C. Kassel, A. Lascoux and C. Reutenauer, *The singular locus of a Schubert variety*, J. Algebra **269** (2003), 74–108.
[19] A. Lascoux, *Syzygies des variétés déterminantales*, Adv. Math. **30** (1978), 202–237.
[20] V. Laskshmibai and B. Sandhya, *Criterion for smoothness of Schubert varieties in* $\mathrm{SL}(n)/B$, Proc. Indian Acad. Sci. Math. Sci. **100** (1990), no. 1, 45–52.
[21] L. Manivel, *Le lieu singulier des variétés de Schubert*, Internat. Math. Res. Notices **16** (2001), 849–871.
[22] ———, *Generic singularities of Schubert varieties*, arXiv:math.AG/0105239.
[23] O. Mathieu, *Formules de caractères pour les algèbres de Kac-Moody générales*, Astérisque **159-160** (1988).
[24] D. Monk, *The geometry of flag manifolds*, Proc. London Math. Soc., **9** (1959), 253–286.
[25] C. Musili and C.S. Seshadri, *Schubert varieties and the variety of complexes*, in *Arithmetic and Geometry, Vol. II*, Progr. Math. **36** (1983), Birkhäuser, Boston, 329–359.
[26] S. Ramanan and A. Ramanathan, *Projective normality of flag varieties and Schubert varieties*, Invent. Math. **79** (1985), no. 2, 217–224.
[27] A. Ramanathan, *Schubert varieties are arithmetically Cohen-Macaulay*, Invent. Math., **80** (1985), 283–294.
[28] ———, *Equations defining Schubert varieties and Frobenius splitting of diagonals*, Publ. Math. IHES **65** (1987), 61–90.
[29] Z. Stankova, *Forbidden subsequences*, Discrete Math. **132** (1994), 291–316.
[30] T. Svanes, *Coherent cohomology on Schubert subschemes of flag schemes and applications*, Adv. Math. **14** (1974), 369–453.
[31] A. Woo and A. Yong, *When is a Schubert variety Gorenstein?*, submitted, arXiv:math.AG/0409490.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720-3840, USA

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

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720-3840, USA

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

Vector partition function and representation theory

Charles Cochet

Abstract. We apply some recent developments of Baldoni-Beck-Cochet-Vergne [**BBCV05**] on vector partition function, to Kostant's and Steinberg's formulae, for classical Lie algebras A_r, B_r, C_r, D_r. We therefore get efficient `Maple` programs that compute for these Lie algebras: the multiplicity of a weight in an irreducible finite-dimensional representation; the decomposition coefficients of the tensor product of two irreducible finite-dimensional representations. These programs can also calculate associated Ehrhart quasipolynomials.

Nous appliquons des résultats récents de Baldoni-Beck-Cochet-Vergne [**BBCV05**] sur la fonction de partition vectorielle, aux formules de Kostant et de Steinberg, dans le cas des algèbres de Lie classiques A_r, B_r, C_r, D_r. Ceci donne lieu à des programmes `Maple` efficaces qui calculent pour ces algèbres de Lie : la multiplicité d'un poids dans une représentation irréductible de dimension finie ; les coefficients de décomposition du produit tensoriel de deux représentations irréductibles de dimension finie. Ces programmes permettent également d'évaluer les quasipolynômes d'Ehrhart associés.

1. Introduction

In this note, we are interested in the two following computational problems for classical Lie algebras A_r, B_r, C_r, D_r:

- The multiplicity c_λ^μ of the weight μ in the representation $V(\lambda)$ of highest weight λ.
- Littlewood-Richardson coefficients, that is the multiplicity ${c_\lambda}^\nu{}_\mu$ of the representation $V(\nu)$ in the tensor product of representations of highest weights λ and μ.

Softwares LiE (from van Leeuwen *et al.* [**vL94**]) and GAP [**GAP**]), and `Maple` packages `coxeter/weyl` (from Stembridge [**S95**]), use Freudenthal's and Klimyk's formulae, and work for any semi-simple Lie algebra (not only for classical Lie algebras). Unfortunately, these formulae are really sensitive to the size of coefficients of weights. Moreover, they do not lead to the computation of associated quasipolynomials $(\lambda, \mu) \mapsto c_\lambda^\mu$ and $(\lambda, \mu, \nu) \mapsto {c_\lambda}^\nu{}_\mu$.

Here the approach to these two problems is through *vector partition function*, that is the function computing the number of ways one can decompose a vector as a linear combination with nonnegative integral coefficients of a fixed set of vectors.

2000 *Mathematics Subject Classification.* Primary 22E46, 52B55.

Key words and phrases. weight multiplicity, Littlewood-Richardson coefficients, Kostka numbers, Ehrhart quasipolynomial, convex polytopes, partition function.

For example the number $p(x)$ of ways of counting x euros with coins, that is

$$p(x) = \sharp\{n \in \mathbb{Z}_+^8 \,;\, x = 200n_1 + 100n_2 + 50n_3 + 20n_4 + 10n_5 + 5n_6 + 2n_7 + n_8\},$$

can be seen as the partition of the 1-dimensional vector (x) with respects to the set $\{(200), (100), (50), (20), (10), (5), (2), (1)\}$ of 1-dimensional vectors. In the case of the decomposition with respects to the set of positive roots of a simple Lie algebra, we speak of *Kostant partition function*.

Recall that any d-dimensional rational convex polytope can be written as the set $P(\Phi, a)$ of nonnegative solutions $x = (x_i) \in \mathbb{R}^N$ of an equation $\sum_{i=1}^N x_i\phi_i = a$, for a matrix Φ with columns $\phi_i \in \mathbb{Z}^r$ and $a \in \mathbb{Z}^r$ $(d = N - r)$. It follows that evaluating the vector partition is equivalent to computing the number of integral points in a rational convex polytope.

The vector partition function arises in many areas of mathematics: representation theory, flows in networks, magic squares, statistics, crystal bases of quantum groups. Its complexity is polynomial in the size of input when the dimension of the polytope is fixed, and NP-hard if it can vary [**B94, B97, BP99**].

There are several approaches to the vector partition problem. For example Barvinok's decomposition algorithm [**B94**], recently implemented by the `LattE` team [**DHTY03, L**], works for general sets of vectors. Beck-Pixton [**BP03**] also created an algorithm dedicated to the vector set arising from the Birkhoff polytope, counting the number of semi-magic squares.

In this note, we use recent results of Baldoni-Beck-Cochet-Vergne [**BBCV05**] to obtain a fast algorithm for Kostant partition function *via* inverse Laplace formula. These results involve DeConcini-Procesi's *maximal nested sets* (or in short MNSs [**DCP04**]) and iterated residues of rational functions computed by formal power series development.

We combine resulting procedures with Kostant's and Steinberg's formulae giving c_λ^μ and ${c_\lambda}^\nu{}_\mu$ in terms of vector partition function. We then obtain a `Maple` program computing for classical Lie algebras (A_r, B_r, C_r, D_r), the multiplicity of a weight in an irreducible finite-dimensional representation, as well as decomposition coefficients of the tensor product of two irreducible finite-dimensional representations. To the best of our knowledge, they are also the only ones able to compute associated piecewise-defined quasipolynomials $(\lambda, \mu) \mapsto c_\lambda^\mu$ and $(\lambda, \mu, \nu) \mapsto {c_\lambda}^\nu{}_\mu$.

These programs (available at [**C**]) are specially designed for large parameters of weights. Indeed although only written in `Maple` they can perform examples with weights with 5 digits coordinates, far beyond classical softwares written in `C++`. We also stress that our programs are absolutely clear, easy to use and require no installation of exotic package or program. Retro-compatibility has been checked downto `Maple Vr5`. They are fully commented, so that a curious user can figure out their internal mechanisms.

However, certain other softwares and packages are not limited by the rank of the algebra like our programs. For example computation of non-trivial examples in Lie algebras of rank 10 is possible with the software LiE, whereas our programs are efficient up to rank 5–7. These facts make our programs complementary to traditional softwares.

Remark that Kostant's and Steinberg's formulae have already been implemented once in the case of A_r [**C03**]. This previous program relies on results of Baldoni-Vergne [**BV01**] implemented by Baldoni-DeLoera-Vergne [**BdLV03**], computing Kostant partition function only in the case of A_r. Tools were *special permutations* and again iterated residues of rational fraction.

A new technique for Littlewood-Richardson coefficients has been recently designed by DeLoera-McAllister [**DM05**]. For A_r, they wrote an algorithm using hive polytopes [**KT99**]. For B_r, C_r, D_r, they implemented Berenstein-Zelevinsky polytopes [**BZ01**]. They can also evaluate stretched Littlewood-Richardson coefficients ${c_{t\lambda}}^{t\nu}{}_{t\mu}$. These two methods consist in computing a tensor product coefficient as the number of lattice points in just one specific convex rational polytope. However our programs based on multidimensional residues are faster, and can reach examples not available by their method.

This paper is organized as follows. Section 2 recalls representation theory problems we are interested in and links them with algebraic combinatorics. Section 3 describes more precisely rational convex polytopes and formulae counting their integral points. Section 4 introduces maximal nested sets and formulae that were used in our programs. Finally in Section 5 we perform tests of our programs.

2. Representation theory and convex polytopes

Let us fix the notations once and for all. Let $\mathfrak{g}$ be a semi-simple Lie algebra of rank r. Choose a Cartan subalgebra $\mathfrak{t}$ of $\mathfrak{g}$ and denote by $L \subset \mathfrak{t}^*$ the weight lattice.

Let Δ^+ be a positive roots system. The root lattice is defined as $\mathbb{Z}[\Delta^+]$. Let $C(\Delta^+)$ be the cone spanned by linear combinations with nonnegative coefficients of positive roots. The Weyl group of $\mathfrak{g}$ for $\mathfrak{t}$ is denoted by W.

There exist only four simple Lie algebras A_r, B_r, C_r, D_r of rank r, called *classical Lie algebras* of rank r [**Bou68**], and determined by their positive roots systems:

$$\begin{array}{ll} A_r: & \Delta^+ = \{e_i - e_j \,|\, 1 \le i < j \le r+1\} \subset \mathbb{R}^{r+1}, \\ B_r: & \Delta^+ = \{e_i - e_j \,|\, 1 \le i < j \le r\} \cup \{e_i \,|\, 1 \le i \le r\} \subset \mathbb{R}^r, \\ C_r: & \Delta^+ = \{e_i - e_j \,|\, 1 \le i < j \le r\} \cup \{2e_i \,|\, 1 \le i \le r\} \subset \mathbb{R}^r, \\ D_r: & \Delta^+ = \{e_i - e_j \,|\, 1 \le i < j \le r\} \cup \{e_i + e_j \,|\, 1 \le i < j \le r\} \subset \mathbb{R}^r. \end{array}$$

The character of a representation V of $\mathfrak{g}$ is $\mathrm{ch}(V) = \sum_{\mu \in L} \dim(V_\mu) e^\mu$. Recall that the irreducible finite-dimensional representation of $\mathfrak{g}$ of highest weight λ is denoted by $V(\lambda)$. Hence the weight multiplicity c_λ^μ is defined as $\dim(V(\lambda)_\mu)$ for any weight μ such that $\lambda - \mu$ is in the root lattice. Multiplicities c_λ^μ are called *Kostka numbers* when $\mathfrak{g} = A_r = \mathfrak{sl}_{r+1}(\mathbb{C})$.

On the other hand, multiplicities of representations $V(\nu)$ in the tensor product $V(\lambda) \otimes V(\mu)$ are called *Littlewood-Richardson coefficients* (or *Clebsch-Gordan coefficients*). Here ν is a dominant weight such that $\lambda + \mu - \nu$ is in the root lattice.

Evaluating weight multiplicities and Littlewood-Richardson coefficients is a difficult task. For A_1, computing Kostka numbers is immediate and Clebsch-Gordan's formula gives Littlewood-Richardson coefficients. For A_2, one can still compute some small examples. But for general X_r ($r \ge 3$) or for weights which components are big (say, with two digits), direct computation is usually intractable.

There exist many formulae from representation theory for c_λ^μ and ${c_\lambda}^\nu{}_\mu$. The first one, valid in any complex semi-simple Lie algebra $\mathfrak{g}$, is Weyl's character formula

$$\mathrm{ch}(V(\lambda)) \quad = \quad \frac{A_{\lambda+\rho}}{A_\rho}, \qquad \text{where } A_\mu = \sum_{w \in W} (-1)^{\varepsilon(w)} e^{w(\mu)},$$

where ρ is half the sum of positive roots for $\mathfrak{g}$. Littlewood-Richardson coefficients are obtained from this formula, since the character of $V(\lambda) \otimes V(\mu)$ is

$$\mathrm{ch}(V(\lambda) \otimes V(\mu)) \;=\; \mathrm{ch}(V(\lambda)) \times \mathrm{ch}(V(\mu)) \;=\; \sum_{\nu \in L\,;\,\lambda+\mu-\nu \in \mathbb{Z}[\Delta^+]} c_{\lambda}{}^{\nu}{}_{\mu}\, \mathrm{ch}(V(\nu)).$$

But these two formulae do not lead to efficient computations when the rank of $\mathfrak{g}$ or the size of coefficients of weights grow. Moreover, computing the whole character is untractable: for $\mathfrak{g} = A_3 = \mathfrak{sl}_4(\mathbb{C})$ and $\lambda = (2,1,0,-3)$, the character $\mathrm{ch}(V(\lambda))$ has 9 monomials but the character $\mathrm{ch}(V(10\lambda))$ has 2903 monomials.

Let us describe Kostant's and Steinberg's formulae in the case of any semi-simple Lie algebra $\mathfrak{g}$. Denote by $k_{\mathfrak{g}}(a)$ the number of ways one can write a vector a as a nonnegative linear combination of positive roots. Remark that $k_{\mathfrak{g}}(a) = 0$ unless a is in the root lattice $\mathbb{Z}[\Delta^+]$. This number satisfies the equation

$$\frac{1}{\prod_{\alpha \in \Delta^+}(1 - e^{-\alpha})} = \sum_{a \in \mathbb{Z}[\Delta^+]} k_{\mathfrak{g}}(a)\; e^{-a}.$$

Let λ and μ be respectively a dominant weight and a weight such that $\lambda - \mu \in \mathbb{Z}[\Delta^+]$. A Weyl group element $w \in W$ is *valid* for λ and μ if the root lattice element $w(\lambda+\rho) - (\mu+\rho)$ is in the cone $C(\Delta^+)$. The set of such w's is denoted by $\mathrm{Val}(\lambda,\mu)$. Then Kostant's formula asserts that the weight multiplicity c_{λ}^{μ} equals

$$c_{\lambda}^{\mu} = \sum_{w \in \mathrm{Val}(\lambda,\mu)} (-1)^{\varepsilon(w)}\, k_{\mathfrak{g}}(w(\lambda+\rho) - (\mu+\rho)). \tag{2.1}$$

Similarly let λ, μ, ν, be three dominant weights such that $\lambda + \mu - \nu \in \mathbb{Z}[\Delta^+]$. The couple $(w, w') \in W \times W$ is *valid* for λ, μ, ν, if the root lattice element $w(\lambda+\rho) + w'(\mu+\rho) - (\nu+2\rho)$ is in $C(\Delta^+)$. The set of such couples is denoted by $\mathrm{Val}(\lambda,\mu,\nu)$. Then Steinberg's formula asserts that the Littlewood-Richardson coefficient equals

$$c_{\lambda}{}^{\nu}{}_{\mu} = \sum_{(w,w') \in \mathrm{Val}(\lambda,\mu,\nu)} (-1)^{\varepsilon(w)+\varepsilon(w')} k_{\mathfrak{g}}(w(\lambda+\rho) + w'(\mu+\rho) - (\nu+2\rho)). \tag{2.2}$$

Sets of valid Weyl group elements and valid couples of Weyl group elements turn out to be relatively small, when compared to W and $W \times W$ (which size is exponential in the rank). Remark that Kostant's (resp. Steinberg's) formula also work when $\lambda - \mu$ (resp. $\lambda + \mu - \nu$) is not in the root lattice, since Kostant partition function vanishes on vectors that are not in the root lattice.

From now on, let X_r be a classical Lie algebra of rank r. Here X stands for A, B, C, D. Its positive roots system will be denoted by X_r^+.

Multiplicities c_{λ}^{μ} and $c_{\lambda}{}^{\nu}{}_{\mu}$ behave nicely, in function of the parameters. More precisely, there exists a decomposition of the space $\mathfrak{t}^* \oplus \mathfrak{t}^* \oplus \mathfrak{t}^*$ in union of closed cones C, such that the restriction of $c_{\lambda}{}^{\nu}{}_{\mu}$ to each cone C is given by a quasi-polynomial function. This follows from theorems of Knutson-Tao [**KT99**] (for A_r), Berenstein-Zelevinsky [**BZ01**] (for any semi-simple Lie algebra) giving $c_{\lambda}{}^{\nu}{}_{\mu}$ as the number of points in a rational convex polytope. In the case of A_r, the fact that $c_{\lambda}{}^{\nu}{}_{\mu}$ is given on each cone C by a polynomial function is proven in Rassart [**Ras04**], and the case of A_3 is treated as an illustration. The description of the decomposition of $\mathfrak{t}^* \oplus \mathfrak{t}^*$ in cones C, where the function c_λ^μ is polynomial for A_r, was given for low ranks by Billey-Guillemin-Rassart [**BGR03**]. See also Rassart's website [**R**] for wonderful slides.

The common point to Kostant's and Steinberg's formulae is the function counting the number of decompositions of a root lattice element as a linear combination with nonnegative integral coefficients of positive roots of the Lie algebra. The next section deals with an efficient method to compute it.

3. Counting integral points in rational convex polytopes

3.1. Vector partition function. Let $E \simeq \mathbb{R}^r$ and Φ be an integral matrix with set of columns $\Delta^+ = \{\phi_1, \dots, \phi_N\} \subset E^*$. Choose $a \in \mathbb{Z}^r$. The rational convex polyhedron associated to Φ and a is

$$P(\Phi, a) \quad = \quad \left\{ x \in \mathbb{R}^N \,;\, \sum_{i=1}^N x_i \phi_i = a,\, x_i \geq 0 \right\}.$$

Remark 3.1. Every convex polyhedron can be realized under the form $P(\Phi, a)$, that is as a set satisfying equality constraints on nonnegative variables. Indeed any inequality can be replaced by an equality by adding a new variable. For example polytopes $\{(x, y) \in \mathbb{R}^2 \,;\, x \geq 0, y \geq 0, x + y \leq 1\}$ and $\{(x, y, z) \in \mathbb{R}^3 \,;\, x \geq 0, y \geq 0, z \geq 0, x + y + z = 1\}$ are isomorphic and have the same number of integral points.

We assume that a is in the cone $C(\Phi)$ spanned by nonnegative linear combinations of the vectors ϕ_i, so that $P(\Phi, a)$ in non-empty. We also assume that the kernel of Φ intersects trivially with the positive orthant $\mathbb{R}_+^N$, so that the cone $C(\Phi)$ is acute and $P(\Phi, a)$ is a polytope (*i.e.* bounded). Finally, we assume that Φ has rank r. The *vector partition function* is by definition

$$k(\Phi, a) \quad = \quad \left| P(\Phi, a) \cap \mathbb{Z}_+^N \right|,$$

that is the number of nonnegative integral solutions $(x_1, \dots, x_N)$ of the equation $\sum_{i=1}^N x_i \phi_i = a$. If $\Phi = \Phi(X_r)$ is the matrix which columns are positive roots for a classical Lie algebra X_r, then $a \mapsto k(\Phi(X_r), a)$ is the *Kostant partition function*. For example

$$\Phi(A_2) = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix} \quad \text{and} \quad \Phi(B_2) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}.$$

Note that the matrix $\Phi(A_r)$ has rank r (and not $r + 1$), since sums on lines are zero.

A *basic subset* of Δ^+ is a basis $\sigma = \{\alpha_1, \dots, \alpha_r\}$ of E^* constituted with elements of Δ^+. Let $B(\Delta^+)$ be the collection of all basic subsets of Δ^+. For such a σ, let $C(\sigma)$ be the cone of linear combinations with nonnegative coefficients of α_i's. Denote by $\mathrm{Sing}(\Delta^+)$ the reunion of the facets of cones $C(\sigma)$, $\sigma \in B(\Delta^+)$; this is the set of *singular* vectors. Let $C_{\mathrm{reg}}(\Delta^+) := C(\Delta^+) \setminus \mathrm{Sing}(\Delta^+)$ be the set of *regular* vectors. A *combinatorial chamber* $\mathfrak{c}$ is by definition a connected component of $C_{\mathrm{reg}}(\Delta^+)$. Combinatorial chambers are regions of quasi-polynomiality of the vector partition function $a \mapsto k(\Phi, a)$. Figure 1 represents cones $C(A_3^+)$ and $C(B_3^+)$, and their chamber decompositions.

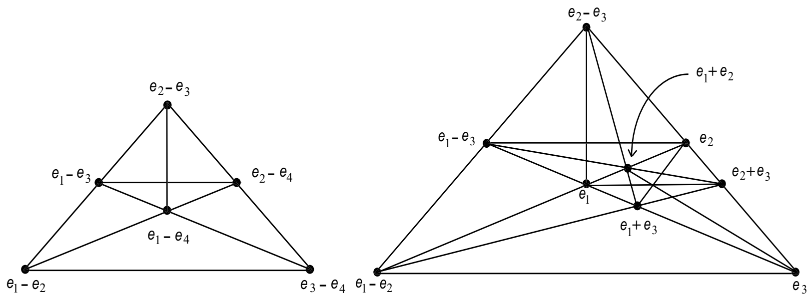


FIGURE 1. The 7 chambers for A_3 and the 23 chambers for B_3

3.2. Brion-Szenes-Vergne formula for classical Lie algebras. Let us describe the formula, computing the number of integral points in rational convex polytopes $P(\Phi(X_r), a)$ associated to a classical algebra X_r, that was implemented in our program.

Let $E = \mathfrak{t}$ and consider the set Δ^+ of positive roots for X_r. Denote by Δ the set $\Delta^+ \cup (-\Delta^+)$ of all roots. Let R_Δ be the vector space of fractions with poles on the hyperplanes defined as kernels of forms $\alpha \in \Delta$. Let S_Δ be the vector space generated by fractions $f_\sigma := \frac{1}{\prod_{\alpha\in\sigma}\alpha}$, $\sigma \in B(\Delta^+)$. Brion-Vergne [**BV97**] proved that R_Δ decomposes as the direct sum $S_\Delta \oplus \partial(R_\Delta)$. We define the *Jeffrey-Kirwan residue* of the chamber $\mathfrak{c}$ as the linear form $\mathrm{JK}_\mathfrak{c}$ on S_Δ:

$$\mathrm{JK}_\mathfrak{c}(f_\sigma) := \begin{cases} \mathrm{vol}(\sigma)^{-1}, & \text{if } \mathfrak{c} \subset C(\sigma), \\ 0, & \text{if } \mathfrak{c} \cap C(\sigma) = \emptyset, \end{cases}$$

where $\mathrm{vol}(\sigma)$ is the volume of the parallelopiped $\sum_{\alpha\in\sigma}[0,1]\alpha$. We extend the JK residue to a linear form on R_Δ by setting it to 0 on $\partial(R_\Delta)$, and to a linear form on the space of formal series $\widehat{R_\Delta}$ by setting it to 0 on homogeneous elements of degree different from $-r$. For example, for the system $\Delta^+ = \{e_1, e_2, e_1+e_2, e_1-e_2\} \subset \mathbb{R}^2$ of positive roots for B_2 and the chamber $\mathfrak{c} = \mathbb{Z}_+ e_1 \oplus \mathbb{Z}_+(e_1+e_2)$ we have

$$\mathrm{JK}_\mathfrak{c}\left(\frac{e^{x-y}}{xy^2}\right) = \mathrm{JK}_\mathfrak{c}\left(\frac{x-y}{xy^2}\right) = \mathrm{JK}_\mathfrak{c}\left(\frac{1}{y^2} - \frac{1}{xy}\right) = -1,$$

since $\mathfrak{c} \subset C(\{e_1, e_2\})$.

Let T be the torus $E/E_\mathbb{Z}$, where $E_\mathbb{Z} \subset E$ is the dual of the root lattice. Given a basic subset σ, we define $T(\sigma)$ as the set of elements $g \in T$ such that $e^{\langle \alpha, 2i\pi G\rangle} = 1$ for all $\alpha \in \sigma$; here G is a representative of $g \in E/E_\mathbb{Z}$. Now let

$$\mathcal{F}(g,a)(u) := \frac{e^{\langle a, 2i\pi G + u\rangle}}{\prod_{\alpha\in\Delta}(1 - e^{-\langle \alpha, 2i\pi G+u\rangle})}.$$

THEOREM 3.2 (Brion-Szenes-Vergne [**BV99, SV04**]). *Let $F \subset T$ be a finite set such that $T(\sigma) \subset F$ for all $\sigma \in B(\Delta^+)$. Fix a combinatorial chamber $\mathfrak{c}$. Then for all $a \in \mathbb{Z}[\Delta^+] \cap \bar{\mathfrak{c}}$, we have:*

$$k(\Phi, a) = \sum_{g\in F} \mathrm{JK}_\mathfrak{c}(\mathcal{F}(g,a)).$$

Now that we linked vector partition function and Jeffrey-Kirwan residue, we describe in Section 4 an efficient way to compute the latter.

4. DeConcini-Procesi's maximal nested sets (MNS) [**DCP04**]

We keep the same notations as in Section 3. A subset $S \subset \Delta^+$ is *complete* if $S = \langle S \rangle \cap \Delta^+$. A complete subset is *reducible* if one can find a decomposition $E = E_1 \oplus E_2$ such that $S = S_1 \cup S_2$ with $S_1 \subset E_1$ and $S_2 \subset E_2$; else S is said *irreducible*. Let $\mathcal{I}$ be the collection of irreducible subsets.

A collection $M = \{I_1, I_2, \ldots, I_s\}$ of irreducible subsets I_j of Δ^+ is *nested*, if: for every subset $\{S_1, \ldots, S_m\}$ of M such that there exist no i, j with $S_i \subset S_j$, the union $S_1 \cup S_2 \cup \cdots \cup S_m$ is complete and the S_i's are its irreducible components. Note that a maximal nested set (MNS in short) has exactly r elements.

Assume Δ^+ irreductible and fix a total order on it. For $M = \{I_1, \ldots, I_s\}$, $I_j \in \Delta^+$, take for every j the maximal element $\beta_j \in I_j$. This defines an application $\phi(M) := \{\beta_1, \ldots, \beta_s\} \subset \Delta^+$. A maximal nested set M is *proper* if $\phi(M)$ is a basis of E^*. Denote by $\mathcal{P}$ the collection of maximal proper nested sets (MPNS in short). We sort $\phi(M)$ and get an ordered list $\theta(M) = [\alpha_1, \ldots, \alpha_r]$. Thus θ is an application from the collection of MPNSs to the collection of ordered basis of E^*. For a given M, let then

$$\begin{aligned} C(M) &:= C(\alpha_1, \ldots, \alpha_r), \\ \mathrm{vol}(M) &:= \mathrm{vol}\left(\oplus_{i=1}^{r} [0,1]\alpha_i\right), \\ \mathrm{IRes}_M &:= \mathrm{Res}_{\alpha_r=0} \cdots \mathrm{Res}_{\alpha_1=0}\,. \end{aligned}$$

Example 4.1. Let e_i be the canonical basis of $\mathbb{R}^r$, with dual basis e^i ($i = 1, \ldots, r$), and define E as the subspace of vectors which sum of coordinates vanish. Consider the set $\Delta^+ = \{e^i - e^j \,|\, 1 \le i < j \le r\}$ of positive roots for A_{r-1}. Irreducible subsets of Δ^+ are indexed by subsets S of $\{1, 2, \ldots, r\}$, the corresponding irreducible subset being $\{e^i - e^j \,|\, i, j \in S, i < j\}$. For instance $S = \{1, 2, 4\}$ parametrizes the set of roots given by $\{e^1 - e^2, e^2 - e^4, e^1 - e^4\}$.

A nested set is represented by a collection $M = \{S_1, S_2, \ldots, S_k\}$ of subsets of $\{1, 2, \ldots, r\}$ such that if S_i, $S_j \in M$ then either $S_i \cap S_j$ is empty, or one of them is contained in another.

For example one can easily compute that for the set of positive roots for A_3 (see Figure 1) there are only 7 MPNS, namely

$$\begin{aligned} &M_1 = \{[1,2],[1,2,3],[1,2,3,4]\}, \quad &M_2 = \{[2,3],[1,2,3],[1,2,3,4]\}, \\ &M_3 = \{[2,3],[2,3,4],[1,2,3,4]\}, \quad &M_4 = \{[3,4],[2,3,4],[1,2,3,4]\}, \\ &M_5 = \{[1,3],[2,4],[1,2,3,4]\}, \quad &M_6 = \{[1,2],[3,4],[1,2,3,4]\}. \end{aligned}$$

Now we can quote the Theorem for the Jeffrey-Kirwan residue computation:

Theorem 4.2 (DeConcini-Procesi). *Let $\mathfrak{c}$ be a combinatorial chamber and fix $f \in R_\Delta$. Take any regular vector $v \in \mathfrak{c}$. Then:*

$$\mathrm{JK}_{\mathfrak{c}}(f) = \sum_{M \in \mathcal{P}\,:\, v \in C(M)} \frac{1}{\mathrm{vol}(M)} \,\mathrm{IRes}_M(f).$$

See [**BBCV05**] for a detailed description of how formulae from Theorems 3.2 and 4.2 were implemented.

5. Our programs

5.1. Description and implementation. Initial data for weight multiplicity and Littlewood-Richardson coefficients are only vectors (respectively two and three). Our programs work with weights represented in the canonical basis of E^*, and not in the fundamental weights basis for X_r. Translation between these two bases is performed *via* straightforward procedures `FromFundaToCanoX(r,v')` and `FromCanoToFundaX(r,v)` (where one replaces `X` by `A`, `B`, `C`, `D`, according to the algebra).

Computation of the weight multiplicity c_λ^μ and of the Littlewood-Richardson coefficient $c_{\lambda}{}^{\nu}{}_{\mu}$ is done by typing in

```
MultiplicityX(lambda,mu);
TensorProductX(lambda,mu,nu);
```

where λ, μ, ν are suitable weights. The syntax for computing quasipolynomials is slightly different. Assume that we want to evaluate $(\lambda', \mu') \mapsto c_{\lambda'}^{\mu'}$ in a neighborhood of a couple (λ, μ), and $(\lambda', \mu', \nu') \mapsto c_{\lambda'}{}^{\nu'}{}_{\mu'}$ in a neighborhood of a triple (λ, μ, ν). Let $\lambda_F = [x_1, \ldots, x_r]$, $\mu_F = [y_1, \ldots, y_r]$, $\nu_F = [z_1, \ldots, z_r]$, be three formal vectors where x_i's, y_i's and z_i's are variables. Then we use the command lines

```
PolynomialMultiplicityX(lambda, lambdaF, mu, muF);
PolynomialTensorProductX(lambda, lambdaF, mu, muF, nu, nuF);
```

So for the polynomial $(\lambda', \mu') \mapsto c_{\lambda'}^{\mu'}$ with $\lambda = (3, 2, 1, -6)$ and $\mu = (2, 2, -2, -2)$ for A_3 we enter

```
PolynomialMultiplicityA(
    [3, 2, 1, -6], [x[1], x[2], x[3], x[4]], [2, 2, -2, -2], [y[1], y[2], y[3], y[4]]);
```

and get instantly

$$\frac{1}{6}(3x_1 - 2y_1 + 1)(3x_1 - 2y_1 + 2)(3x_1 + 6x_2 - 2y_1 - 6y_2 + 3).$$

Remark that quasipolynomials $c_{t\lambda}^{t\mu}$ and $c_{t\lambda}{}^{t\nu}{}_{t\mu}$ are obtained by setting $x_i = t\lambda_i$, $y_i = t\mu_i$, $z_i = t\nu_i$, so that

```
PolynomialMultiplicityA(
    [3, 2, 1, -6], [3t, 2t, t, -6t], [2, 2, -2, -2], [2t, 2t, -2t, -2t]);
```

returns $(t + 1)(t + 2)(t + 3)/6$.

Now some words about implementation. There are two main parts in our programs. The first one is the implementation of Theorems 3.2 and 4.2; it is described in [**BBCV05**]. The second one is the implementation of Kostant's (2.1) and Steinberg's (2.1) formulae using valid Weyl group elements and valid couples of Weyl group elements; it is a generalization for classical Lie algebras of what has been done for A_r in [**C03**].

5.2. Comparative tests. Figure 2 describes efficiency area of the software LiE and of our programs using MNS; any area located to the left of a colored line represents the range where a program can compute examples in a reasonable time. Figures 3–5 present precise comparative tests of the software LiE, of DeLoera-McAllister's script [**DM05**] using LattE [**L**] and of our programs using MNS.

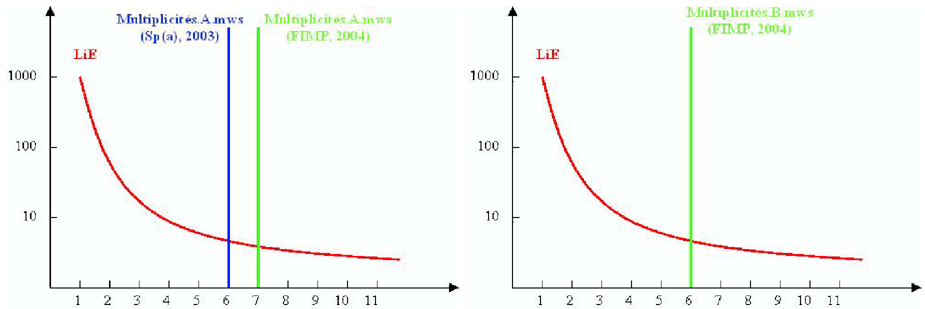


FIGURE 2. To the left, comparison for tensor product coefficients for A_r: with LiE, with Sp($\mathbf{a}$) and with MNS. To the right, comparison for weight multiplicity of a weight for B_r: with LiE and with MNS. Similar Figures for C_r and D_r.

All examples were runned on the same computer, a Pentium IV 1,13GHz with 2Go of RAM memory. Remark that computation times for LattE and LiE are slower than those shown in [**DM05**], due to different computers. However, we performed exactly same examples for comparison purposes.

As in [**DM05**], in Tables 3–4 weights are for $\mathfrak{gl}_{r+1}(\mathbb{C})$ and not $\mathfrak{sl}_{r+1}(\mathbb{C})$ (coordinates do not add to zero). However the sum of coordinates of $\lambda+\mu-\nu$ vanish.

Now some words about quasipolynomials computation. Let us examine the first example for B_3 in [**DM05**], that is the evaluation of the quasipolynomial $c_{t\lambda}{}^{t\nu}{}_{t\mu}$ for weights $\lambda=[0,15,5]$, $\mu=[12,15,3]$ and $\nu=[6,15,6]$ expressed in the basis of fundamental weights. In canonical basis, these data become $\lambda=(35/2,35/2,5/2)$, $\mu=(57/2,33/2,3/2)$, $\nu=(24,18,3)$. The program using the MNS algorithm returns the quasipolynomial

$$\begin{aligned} c_{t\lambda}{}^{t\nu}{}_{t\mu} &= \left(\frac{203}{256}+\frac{53}{256}(-1)^t\right)+\left(\frac{1515}{128}+\frac{197}{128}(-1)^t\right)t \\ &+\left(\frac{35353}{384}+\frac{881}{128}(-1)^t\right)t^2+\left(\frac{13405}{32}\right)t^3 \\ &+\left(\frac{407513}{384}\right)t^4+\left(\frac{68339}{64}\right)t^5 \end{aligned}$$

in 1099,4s. On the other hand, the computation of the full quasipolynomial $c_\lambda{}^\nu{}_\mu$ with formal vectors $[x_1,x_2,x_3]$, $[y_1,y_2,y_3]$, $[z_1,z_2,z_3]$ leads to a 87 pages result, obtained in only 1158,6s. With LattE, on our computer, one obtains the quasipolynomial $c_{t\lambda}{}^{t\nu}{}_{t\mu}$ in only 825,8s.

As announced in the introduction, our program is really efficient for weights with huge coefficients. Note that in the particular case of A_r the MNS algorithm allows us to compute examples one rank further than the Sp($\mathbf{a}$) algorithm.

The translation of the program using MNS in the language of the symbolic calculation software MuPAD is in progress. A version using distributed calculation on a grid of computers is in the air; it will considerably increase the speed of computations.

λ, μ, ν	$c_{\lambda\ \mu}^{\ \nu}$	MNS	LattE	LiE
(9,7,3,0,0), (9,9,3,2,0), (10,9,9,8,6)	2	8,0s	3,0s	<0,1s
(18,11,9,4,2), (20,17,9,4,0), (26,25,19,16,8)	453	2,8s	8,8s	<0,1s
(30,24,17,10,2), (27,23,13,8,2), (47,36,33,29,11)	5231	2,2s	11,4s	0,5s
(38,27,14,4,2), (35,26,16,11,2), (58,49,29,26,13)	16784	1,3s	12,8s	1,5s
(47,44,25,12,10), (40,34,25,15,8), (77,68,55,31,29)	5449	1,3s	8,8s	1,4s
(60,35,19,12,10), (60,54,27,25,3), (96,83,61,42,23)	13637	1,0s	8,4s	9,1s
(64,30,27,17,9), (55,48,32,12,4), (84,75,66,49,24)	49307	2,5s	9,5s	15,9s
(73,58,41,21,4), (77,61,46,27,1), (124,117,71,52,45)	557744	2,1s	12,3s	284,1s

FIGURE 3. For A_r, comparison of running times between the MNS algorithm, LattE and LiE

λ, μ, ν	$c_{\lambda\ \mu}^{\ \nu}$	MNS	LattE
(935,639,283,75,48) (921,683,386,136,21) (1529,1142,743,488,225)	1303088213330	1,7s	12,8s
(6797,5843,4136,2770,707) (6071,5175,4035,1169,135) (10527,9398,8040,5803,3070)	459072901240524338	3,1s	15,1s
(859647,444276,283294,33686,24714) (482907,437967,280801,79229,26997) (1120207,699019,624861,351784,157647)	11711220003870071391294871475	2,0s	11,9s

FIGURE 4. For A_r, comparison of running times for large weights between the MNS algorithm and LattE

	λ, μ, ν	$c_{\lambda\ \mu}^{\ \nu}$	MNS	LattE	LiE
B_3	(46,42,38), (38,36,42), (41,36,44)	354440672	6,4s	22,5s	229,0s
	(46,42,41), (14,58,17), (50,54,38)	88429965	2,7s	15,2s	102,6s
	(15,60,67), (58,70,52), (57,38,63)	626863031	7,8s	17,0s	713,5s
	(5567,2146,6241), (6932,1819,8227), (3538,4733,3648)	215676881876569849679	7,0s	16,3s	–
C_3	(25,42,22), (36,38,50), (31,33,48)	87348857	5,6s	18,1s	52,9s
	(34,56,36), (44,51,49), (37,51,54)	606746767	5,1s	20,4s	516,0s
	(39,64,58), (65,15,72), (70,41,44)	519379044	8,7s	18,3s	1096,9s
	(5046,5267,7266), (7091,3228,9528), (9655,7698,2728)	1578943284716032240384	8,2s	18,3s	–
D_4	(13,20,10,14), (10,20,13,20), (5,11,15,18)	41336415	131,0s	185,8s	224,7s
	(12,22,9,30), (28,14,15,26), (10,24,10,26)	322610723	78,6s	192,7s	1184,8s
	(37,16,31,29), (40,18,35,41), (36,27,19,37)	18538329184	64,3s	258,7s	21978,4s
	(2883,8198,3874,5423), (1901,9609,889,4288), (5284,9031,2959,5527)	1891293256704574356565149344	27,7s	165,2s	–

FIGURE 5. For B_r, C_r, D_r, comparison of running times between LattE, the MNS algorithm and LiE

References

[BdLV03] Baldoni-Silva W., De Loera J.A., and Vergne M., *Counting Integer lows in Networks*, Found. Comput. Math. **4** (2004), 277–314.

[BV01] Baldoni-Silva W. and Vergne M., *Residues ormulae for Volumes and Ehrhart Polynomials of Convex Polytopes*, 81 pages, 2001, available at arXiv:math.CO/0103097.

[BBCV05] Baldoni W., Beck M., Cochet C. and Vergne M., *Volume of Polytopes and Partition unction*, 2005, 55 pages, available at arXiv:math.CO/0504231.

[BP99] Barvinok A. and Pommersheim J., *An Algorithmic Theory of Lattice Points in Polyhedra*, in: *New Perspectives in Algebraic Combinatorics* (Berkeley, CA, 1996-1997), Math. Sci. Res. Inst. Publ. **38** (1999), Cambridge Univ. Press, Cambridge, 91-147.

[B94] Barvinok A. I., *A Polynomial Time Algorithm for Counting Integral Points in Polyhedra when the Dimension is ixed*, Math. Oper. Res. **19** (1994), 769–779.

[B97] Barvinok A., *Lattice Points and Lattice Polytopes*, in *Handbook of discrete and computational geometry*, CRC Press Ser. Discrete Math. Appl., CRC, Boca Raton, FL (1997), 133–152.

[BP03] Beck M. and Pixton D., *The Ehrhart Polynomial of the Birkhoff Polytope*, Discrete Comp. Geom. **30** (2003), 623–637.

[BZ01] Berenstein A. and Zelevinsky A., *Tensor Product Multiplicities, Canonical Bases and Totally Positive Varieties*, Invent. Math. **143** (2001) 77–128.

[BGR03] Billey S., Guillemin V. and Rassart É., *A Vector Partition unction for the Multiplicities of* $\mathfrak{sl}_k(\mathbb{C})$, J. Algebra **278** (2004), 251–293.

[Bou68] N. Bourbaki, *Éléments de mathématique. asc. XXXIV. Groupes et algèbres de Lie. Chapitre IV: Groupes de Coxeter et systèmes de Tits. Chapitre V: Groupes engendrés par des réflexions. Chapitre VI: systèmes de racines.* Hermann, Paris, 1968.

[BV97] M. Brion and M. Vergne, *Residue ormulae, Vector Partition unctions and Lattice Points in Rational Polytopes*, J. Amer. Math. Soc. **10** (1997) 797–833.

[BV99] M. Brion and M. Vergne, *Arrangement of Hyperplanes. I. Rational unctions and Jeffrey-Kirwan Residue*, Ann. Sci. École Norm. Sup. (4) **32** (1999) 715–741.

[C] `http://www.math.jussieu.fr/~cochet`

[C03] Cochet C., *Kostka Numbers and Littlewood-Richardson Coefficients*, Proceedings of the AMS–IMS–SIAM Joint Summer Research Conference “Integer Points in Polyhedra – Geometry, Number Theory, Algebra, Optimization”, July 13–17, 2003, Snowbird, UT, to appear in *Contempory Mathematics*.

[DCP04] De Concini C. and Procesi C., *Nested Sets and Jeffrey-Kirwan Cycles*, 10 pages, available at arXiv:math.AG/0406290.

[DHTY03] DeLoera J., Hemmecke R., Tauzer J. and Yoshida R., *Effective Lattice Point Counting in Rational Convex Polytopes*, J. Symbolic Comput. **38** (2004), no. 4, 1273–1302.

[DM05] DeLoera J. and McAllister T., *On the Computation of Clebsch–Gordan Coefficients and the Dilation Effect*, 2005, 21 pages, available at arXiv:math.RT/0501446.

[GAP] `http://www.math.rwth-aachen.de/~GAP`

[KT99] Knutson A. and Tao T., *The Honeycomb Model of* $GL_n(\mathbb{C})$ *Tensor Products. I. Proof of the saturation conjecture*, J. Amer. Math. Soc. **12** (1999), 1055–1090.

[L] `http://www.math.ucdavis.edu/~totalresidue/`

[R] `http://www.math.ias.edu.de/~rassart/presentations.html`

[Ras04] Rassart É., *A Polynomiality Property for Littlewood-Richardson Coefficients*, J. Combin. Theory Ser. A **107** (2004), 161–179.

[S95] Stembridge, J., *A Maple Package for Symmetric unctions*, J. Symbolic Computation **20** (1995), 755–768.

[SV04] Szenes A. and Vergne M., *Toric Reduction and a Conjecture of Batyrev and Materov*, Invent. Math. **158** (2004), 453–495.

[vL94] van Leeuwen M. A. A., *LiE, a Software Package for Lie Group Computations*, Euromath Bull. **1** (1994), 83–94.

Institut de Mathématiques de Jussieu, Université Paris 7 – Denis Diderot, Case 7012, 75251 Paris cedex 05, France

E-mail address: `cochet@math.jussieu.fr`