FPSAC PROCEEDINGS 2004

ACTES SFCA 2004

Vancouver CANADA

Posters

16th Annual International Conference on
FORMAL POWER SERIES AND ALGEBRAIC COMBINATORICS
June 28 - July 2, 2004
University of British Columbia
(Vancouver B.C., Canada)

16ième Conférence Internationale sur
SÉRIES FORMELLES ET COMBINATOIRE ALGÉBRIQUE
28 juin - 2 juillet 2004
University of British Columbia
(Vancouver C.B., Canada)

Document prepared by M. Mishna with the aid of F. Chyzak's Latex class files.

Programme Committee – Comité de programme

Jonathan Borwein *(Simon Fraser, Dalhousie)*, chair/président
Mike Atkinson *(Otago)*
Nantel Bergeron *(CRC, York)*
Luisa Carini *(Messina)*
William Chen *(Nankai)*
Sylvie Corteel *(CNRS U.V.S.Q.)*
Ira Gessel *(Brandeis)*
Ian Goulden *(Waterloo)*
Michael Joswig *(T.U. Berlin)*
Martin Klazar *(Prague)*
Christian Krattenthaler *(Lyon)*
Gilbert Labelle *(UQAM)*
Soichi Okada *(Nagoya)*
Margaret Readdy *(Kentucky)*
Frank Sottile *(Massachusetts)*
Einar Steingrimsson *(Göteborg)*
Julian West *(Malaspina, Victoria)*, co-chair/co-président
Stephanie van Willigenburg *(UBC)*, co-chair/co-président

Organizing Committee – Comité d'organisation

Julian West *(Malaspina, Victoria)*, chair/président
Nantel Bergeron *(CRC at York)*
Luis Goddyn *(Simon Fraser)*
Marni Mishna *(Bordeaux)*
Tom Roby *(Cal State Hayward)*
Frank Ruskey *(Victoria)*
Stephanie van Willigenburg *(UBC)*

Sponsors – Commanditaires

Pacific Institute for the Mathematical Sciences

UVic

UBC

TABLE OF CONTENTS

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Combinatorial Aspects of Abstract Young Representations (Extended Abstract)

Ron M. Adin, Francesco Brenti, and Yuval Roichman

Abstract. *The goal of this paper is to give a new unified axiomatic approach to the representation theory of Coxeter groups and their Hecke algebras. Building upon fundamental works by Young and Kazhdan-Lusztig, followed by Vershik and Ram, we propose a direct combinatorial construction, avoiding a priori use of external concepts (such as Young tableaux). This is carried out by a natural assumption on the representation matrices. For simply laced Coxeter groups, this assumption yields explicit simple matrices, generalizing the Young forms. For the symmetric groups the resulting representations are completely classified and include the irreducible ones. Analysis involves generalized descent classes and convexity (à la Tits) within the Hasse diagram of the weak Bruhat poset.*

Résumé. *L'objectif de cet article est de donner une nouvelle approche axiomatique unifiée de la théorie des représentation des groupes de Coxeter et de leurs algèbres de Hecke. En utilisant les travaux de Young, Kazhdan-Lusztig ainsi que de Vershik et Ram, nous proposons une construction combinatoire directe qui évite l'introduction de concepts extérieurs (par exemple les tableaux de Young). Cette construction est faite à partir d'une hypothèse naturelle sur les matrices de représentation. Pour les groupes de Coxeter simplement lacé, cette hypothèse donne des matrices simples explicites, généralisant la forme de Young. Pour les groupes symmétriques les représentations associèes sont complètement classifiées, en particulier celles qui sont irréductibles. Ce travail utilise les classes de descente généralisées et la convexité (à la Tits) dans le diagramme de Hasse de l'ordre de Bruhat faible.*

1. Introduction

The goal of the construction of abstract Young representations, presented in [**ABR1**], is to give a new unified axiomatic approach to the representation theory of Coxeter groups and their Hecke algebras.

We want our construction to

(a) apply in a general context;
(b) be simple, direct and **combinatorial**; and
(c) avoid a priori use of concepts external to the group or algebra itself (such as standard Young tableaux).

1991 *Mathematics Subject Classification.* Primary 05E10, 20F55; Secondary 52A01.

Key words and phrases. Coxeter group, Hecke algebra, cell, representation, convexity.

Research of all authors was supported in part by the Israel Science Foundation, founded by the Israel Academy of Sciences and Humanities; by the EC's IHRP Programme, within the Research Training Network "Algebraic Combinatorics in Europe", grant HPRN-CT-2001-00272; and by internal research grants from Bar-Ilan University.

Goals (a) and (c) were stated and pursued by Kazhdan-Lusztig [**KL**] and Vershik [**V**]. Goal (b) was posed in [**BR**].

Our guiding lines are two fundamental methods to construct representations: Young theory and Kazhdan-Lusztig theory.

In *Young Theory* (as explained by James [**J**]) the construction starts with Young tableaux, which are sophisticated ad-hoc combinatorial objects. Modules (in particular, irreducible ones) are generated by symmetrizers of Young tableaux. Representing matrices are obtained as a side benefit.This theory is effective for classical Weyl groups. For a detailed description see [**J**] and [**JK**].

Kazhdan-Lusztig Theory [**KL**] is a very general approach to the construction of Hecke algebra representations. A distinguished *basis*, indexed by group elements, is compatible with the decomposition of the Hecke algebra. The Coxeter group acts on linear spaces with bases indexed by special subsets of the group, called *cells*. The basic tools in this construction are Kazhdan-Lusztig polynomials. Resulting representation matrices are given in terms of coefficients of these polynomials [**Hu2**, §7.14]. Unfortunately, these coefficients (and thus entries of the representing matrices) are very difficult to compute. For an axiomatic approach to this construction via cellular algebras see [**GL**].

The idea of "reversing Young theory", namely, constructing representations using explicit representation matrices for the Coxeter generators, is apparently due to Vershik [**V**], and was further developed in works of Vershik-Okounkov [**V**] [**OV**], Pushkarev [**P**] and Ram [**Ra1**] (see also [**BR**]). In these papers the external objects (Young tableaux, or abstractions thereof) are applied as an important initial ingredient.

Our approach is different. The idea is, again, to reverse Young theory — but along "Kazhdan-Lusztig language". As in Kazhdan-Lusztig theory, we start with a (formal) *basis* indexed by group elements; decomposition is compatible with special subsets of the group, called *cells*. The action is assumed to satisfy a natural condition, as follows.

Let (W, S) be a Coxeter system, and let $\mathcal{K}$ be a subset of W. Let F be a suitable field of characteristic zero (e.g., the field $\boldsymbol{C}(q)$ in the case of the Iwahori-Hecke algebra of type A), and let ρ be a representation of (the Iwahori-Hecke algebra of) W on the vector space $V_{\mathcal{K}} := span_F\{C_w \mid w \in \mathcal{K}\}$, with basis vectors indexed by elements of $\mathcal{K}$. Motivated by goals (a)–(c) above, we want to study the sets $\mathcal{K}$ and representations ρ which satisfy the following axiom:

(A) *For any generator $s \in S$ and any element $w \in \mathcal{K}$ there exist scalars $a_s(w), b_s(w) \in F$ such that*

$$\rho_s(C_w) = a_s(w)C_w + b_s(w)C_{ws}.$$

If $w \in \mathcal{K}$ but $ws \notin \mathcal{K}$ we assume $b_s(w) = 0$.

A pair $(\rho, \mathcal{K})$ satisfying Axiom (A) is called an *abstract Young (AY) pair;* ρ is an *AY representation,* and $\mathcal{K}$ is an *AY cell.* If $\mathcal{K} \neq \emptyset$ and has no proper subset $\emptyset \subset \mathcal{K}' \subset \mathcal{K}$ such that $V_{\mathcal{K}'}$ is ρ-invariant, then $(\rho, \mathcal{K})$ is called a *minimal AY pair.* (This is much weaker than assuming ρ to be irreducible.)

Surprisingly, Axiom (A) leads to very concrete matrices, whose entries are essentially inverse linear. Analysis of the construction involves a convexity theorem of Tits [**T**] and the generalized descent classes introduced by Björner and Wachs [**BW1**].

This extended abstract is based on the paper [**ABR1**]. Main definitions and results of that paper are surveyed in Sections 2 and 3. A new result on boundary conditions, not yet available in preprint form, is proved in Section 4. A combinatorial characterization of minimal AY cells and representations for the symmetric group is given in Section 5. For proofs and more details see [**ABR1**].

Note Added: Having completed the current version of [**ABR1**], we have been informed of the important recent paper [**Ra2**]. Although it differs from our work in context, initial assumptions, motivation and language, there are points of contact and similarity in some of the results. In particular, the linear functional $\langle f, \cdot \rangle$ which appears in the coefficients of a minimal AY pair (see Theorem 3.4 below) is a basic ingredient in [**Ra2**].

2. Abstract Young Cells

Recall the definition of AY cells and representations from the previous section.

Problem 2.1. (Kazhdan [**K**]) *Given a subset $\mathcal{K} \subseteq W$, how many nonisomorphic abstract Young representations may be defined on $V_{\mathcal{K}}$?*

In particular,

Problem 2.2. *Which subsets of W are (minimal) AY cells?*

Observation 2.3. Every nonempty AY cell is a left translate of an AY cell containing the identity element of W.

Let T be the set of all reflections in W, and let $A \subseteq T$ be any subset. The (left) A-*descent set* of an element $w \in W$ is defined by

$$Des_A(w) := \{t \in A \mid \ell(tw) < \ell(w)\}.$$

For $D \subseteq A \subseteq T$, the corresponding *generalized descent class* is

$$W_A^D := \{w \in W \mid Des_A(w) = D\}.$$

These sets were studied by Tits [**T**, Ch. 2] and Björner-Wachs [**BW1, BW2**].

The *right Cayley graph* $X(W,S)$ has W as the set of vertices, and has a directed edge $w \to ws$ whenever $w \in W$ and $s \in S$. A subset $\mathcal{K}$ of W is *convex* in $X(W,S)$ if every shortest path between any two elements of $\mathcal{K}$ has all its vertices in $\mathcal{K}$.

Using [**T**, Theorem 2.19] we prove

Theorem 2.1. *Every minimal AY cell is a generalized descent class; in particular, it is convex in the right Cayley graph $X(W,S)$ (or, equivalently, under right weak Bruhat order).*

3. Abstract Young Representations

In [**ABR1**] it is shown that, under mild conditions (see Theorem 3.1 below), Axiom (A) is equivalent to the following more specific version.

(B) *There exist scalars $\dot{a}_t, \dot{b}_t, \ddot{a}_t, \ddot{b}_t \in F$ $(\forall t \in T)$ such that, for all $s \in S$ and $w \in \mathcal{K}$:*

$$\rho_s(C_w) = \begin{cases} \dot{a}_{wsw^{-1}} C_w + \dot{b}_{wsw^{-1}} C_{ws}, & \text{if } \ell(w) < \ell(ws); \\ \ddot{a}_{wsw^{-1}} C_w + \ddot{b}_{wsw^{-1}} C_{ws}, & \text{if } \ell(w) > \ell(ws). \end{cases}$$

If $w \in \mathcal{K}$ and $ws \notin \mathcal{K}$ we assume that $\dot{b}_{wsw^{-1}} = 0$ (if $\ell(w) < \ell(ws)$) or $\ddot{b}_{wsw^{-1}} = 0$ (if $\ell(w) > \ell(ws)$).

Theorem 3.1. *Let $(\rho, \mathcal{K})$ be a minimal AY pair for the Iwahori-Hecke algebra of (W,S). If $a_s(w) = a_{s'}(w') \Longrightarrow b_s(w) = b_{s'}(w')$ $(\forall s, s' \in S, w, w' \in \mathcal{K})$, then Axioms (A) and (B) are equivalent.*

This theorem shows that the coefficients $a_s(w)$ and $b_s(w)$ in Axiom (A) depend only on the reflection $wsw^{-1} \in T$ and on the relation between w and ws under right weak Bruhat order.

The assumption regarding the coefficients $b_s(w)$ in Theorem 3.1 is merely a normalization condition. Thus, in order to determine an AY representation, it suffices to determine the coefficients $\dot{a}_t$ and $\ddot{a}_t$ (actually, $\dot{a}_t$ will suffice) for all reflections t and to choose a normalization for the $\dot{b}_t$ and $\ddot{b}_t$. One such normalization is defined as follows (assuming, for simplicity, that $\mathcal{K}$ contains the identity element of W). Let

$$\begin{aligned} T_{\mathcal{K}} &:= \{wsw^{-1} \mid s \in S,\, w, ws \in \mathcal{K}\}, \\ T_{\partial\mathcal{K}} &:= \{wsw^{-1} \mid s \in S,\, w \in \mathcal{K},\, ws \notin \mathcal{K}\}. \end{aligned}$$

Fact 3.1.

$$T_{\mathcal{K}} \cap T_{\partial\mathcal{K}} = \emptyset.$$

The *row stochastic* normalization satisfies

$$\dot{a}_t + \ddot{a}_t = 1 - q, \quad \dot{b}_t = 1 - \dot{a}_t, \quad \ddot{b}_t = 1 - \ddot{a}_t \qquad (\forall t \in T_{\mathcal{K}});$$

$$\dot{a}_t \in \{1, -q\}, \quad \dot{b}_t = 0 \qquad (\forall t \in T_{\partial\mathcal{K}}).$$

Problem 3.2. (Kazhdan [**K**]) *Do the coefficients $a_s(w)$ determine all the character values?*

An (affirmative) answer to this problem, independent of the choice of normalization, will be given in [**ABR3**].

It turns out that for simply laced Coxeter groups the coefficients $\dot{a}_t$ are given by a linear functional (see Theorems 3.3 and 3.4 below).

Let V be the root space of W, and let $\langle\ ,\ \rangle$ be an arbitrary positive definite bilinear form on V. For a reflection $t \in T$, let $\alpha_t \in V$ be the corresponding positive root.

Definition 3.2. Let $\mathcal{K}$ be a convex subset of W containing the identity element. A vector f in the root space V is $\mathcal{K}$-generic if:

(i) For all $t \in T_{\mathcal{K}}$,

$$\langle f, \alpha_t \rangle \notin \{0, 1, -1\}.$$

(ii) For all $t \in T_{\partial\mathcal{K}}$,

$$\langle f, \alpha_t \rangle \in \{1, -1\}.$$

(iii) If $w \in \mathcal{K}$, $s, t \in S$, $(st)^3 = 1$ and $ws, wt \notin \mathcal{K}$ then

$$\langle f, \alpha_{wsw^{-1}} \rangle = \langle f, \alpha_{wtw^{-1}} \rangle.$$

By Observation 2.3, every abstract Young representation is isomorphic to one on an AY cell containing the identity element. Therefore, in the following theorems, we assume that $\mathcal{K}$ contains the identity element.

Theorem 3.3. *Let W be an irreducible simply laced Coxeter group, and let $\mathcal{K}$ be a convex subset of W containing the identity element. Let $\langle\ ,\ \rangle$ be an arbitrary positive definite bilinear form on the root space V. If $f \in V$ is $\mathcal{K}$-generic then*

$$\dot{a}_{wsw^{-1}} := \frac{1}{\langle f, \alpha_{wsw^{-1}} \rangle} \qquad (\forall w \in \mathcal{K}, s \in S),$$

together with $\ddot{a}_{wsw^{-1}}$, $\dot{b}_{wsw^{-1}}$ and $\ddot{b}_{wsw^{-1}}$ satisfying appropriate normalization conditions, define a representation ρ such that $(\rho, \mathcal{K})$ is a minimal AY pair.

For $n \in \boldsymbol{Z}$ let

$$[n]_q := \frac{1 - q^n}{1 - q} \in \boldsymbol{Z}[q, q^{-1}].$$

Replacing $\langle f, \alpha_t \rangle$ by its q-analogue $[\langle f, \alpha_t \rangle]_q$ gives representations of the Iwahori-Hecke algebra $\mathcal{H}_q(W)$. See [**ABR1**, Theorem 8.5].

The following theorem is complementary.

Theorem 3.4. *Let W be an irreducible simply laced Coxeter group and let $\mathcal{K}$ be a subset of W containing the identity element. Assume that $\dot{a}_{wsw^{-1}} \neq 0$ $(\forall w \in \mathcal{K}, s \in S\ )$. If $(\rho, \mathcal{K})$ is a minimal AY pair satisfying Axiom (B) then there exists a $\mathcal{K}$-generic $f \in V$ such that*

$$\dot{a}_{wsw^{-1}} = \frac{1}{\langle f, \alpha_{wsw^{-1}} \rangle} \qquad (\forall\ w \in \mathcal{K}, s \in S).$$

For an Iwahori-Hecke algebra analogue see [**ABR1**, Theorem 8.6].

4. Boundary Conditions

In this section it is shown that the action of the group W on the boundary of a cell determines the representation up to isomorphism. As this result is not yet available in preprint form, it is given with a proof.

Definition 4.1. Let W be a finite Coxeter group, and let V be its root space. A *basic (affine) hyperplane* in V has the form

$$H_{t,\varepsilon} := \{f \in V \mid \langle f, \alpha_t \rangle = \varepsilon\},$$

where $t \in T$ and $\varepsilon = \pm 1$.

A *basic (proper) flat* in V is an intersection (other than $\emptyset$ or V) of basic hyperplanes.

For a basic proper flat L, let

$$A = A_L := \{t \in T \mid L \subseteq H_{t,\varepsilon} \text{ for some } \varepsilon = \pm 1\}.$$

Then $\{W_A^D \mid D \subseteq A\}$ (see Section 2 for the definition of W_A^D) is a partition of W into convex subsets, called the *L-partition* of W.

Let f be a $\mathcal{K}$-generic vector in V. Denote by ρ^f the representation determined by f on $\mathcal{K}$ (say with the row stochastic normalization).

Theorem 4.2. *Let W be a finite simply laced Coxeter group. Let L be a basic proper flat in V, and fix some nonempty convex set $\mathcal{K}$ in the L-partition of W. Then, for any two $\mathcal{K}$-generic vectors $f, g \in L$, the representations ρ^f and ρ^g on $\mathcal{K}$ are isomorphic.*

Proof. Choose $f_0 \in L$, and let $\{f_1, \ldots, f_k\}$ be a basis for the linear subspace $L - f_0$ of V. Each $f \in L$ has a unique expression as

$$f = f_0 + r_1 f_1 + \cdots + r_k f_k,$$

where $r_1, \ldots, r_k \in \boldsymbol{R}$. For any $t \in T_{\mathcal{K}} \cup T_{\partial\mathcal{K}}$, $\langle f, \alpha_t \rangle$ is a linear combination of $1, r_1, \ldots, r_k$, and is nonzero if f is $\mathcal{K}$-generic. Thus, for any $\mathcal{K}$-generic $f \in L$ and any $s \in S$, each entry of the matrix $\rho^f(s)$ is a rational function of $r_1, \ldots, r_k$; and the same holds for each entry of $\rho^f(w)$ ($\forall w \in W$) and for the character values $Tr(\rho^f(w))$. Note that the coefficients of these rational functions (unlike the actual values of $r_1, \ldots, r_k$) do not depend on the choice of $\mathcal{K}$-generic $f \in L$, even though the set of all such f may be disconnected (see example below). By discreteness of character values and continuity in a small neighborhood of a $\mathcal{K}$-generic $f \in L$, each character value is constant in each such neighborhood, and is thus represented by a constant rational function. It is therefore the same for all the $\mathcal{K}$-generic vectors in L, as claimed. □

Example 4.1. Take $W = S_3 = \langle s_1, s_2 \rangle$ (type A_2) and the basic flat $L = \{f \in V \mid \langle f, \alpha_{s_1 s_2 s_1} \rangle = -1\}$. Then $A = \{s_1 s_2 s_1\}$, and we may choose $\mathcal{K} = \{1_W, s_1, s_2\}$. In that case, $T_{\mathcal{K}} = \{s_1, s_2\}$ and $T_{\partial\mathcal{K}} = \{s_1 s_2 s_1\} = A$. L is an affine line in $V \cong \boldsymbol{R}^2$, and the $\mathcal{K}$-generic points in L form five disjoint open intervals (three of them bounded). For any $\mathcal{K}$-generic vector $f \in L$, ρ^f is the 3-dimensional representation isomorphic to the direct sum of the sign representation and the unique irreducible 2-dimensional representation of S_3.

5. The Symmetric Group

5.1. Minimal AY Cells.

The following theorem characterizes the minimal AY cells in the symmetric group S_n.

Theorem 5.1. *Let $\mathcal{K}$ be a nonempty subset of the symmetric group S_n, and let $\sigma \in \mathcal{K}$. Then $\mathcal{K}$ is a minimal AY cell if and only if there exists a standard Young tableau Q such that*

$$\sigma^{-1}\mathcal{K} = \{\pi \in S_n \mid Q^\pi \textit{ is standard}\},$$

where Q^π is the tableau obtained from Q by replacing each entry i by $\pi(i)$.

Proof. First observe that any basic proper flat of the symmetric group contains a vector with integer coordinates. Combining this observation with Theorem 4.2 and Observation 2.3 allows one to reduce the discussion to minimal AY cells, containing the identity element, which are determined by integer valued linear functionals. Theorem 5.2 below completes the proof.

□

For a vector $v = (v_1, \ldots, v_n) \in F^n$ denote

$$\Delta\, v := (v_2 - v_1, \ldots, v_n - v_{n-1}) \in F^{n-1}.$$

For a (skew) standard Young tableau T denote $c(k) := j - i$, where k is the entry in row i and column j of T. Call $cont(T) := (c(1), \ldots, c(n))$ the *content vector* of T, and call $\Delta\, cont(T)$ the *derived content vector* (or *axial distance vector*) of T.

Let $w \in W$, and let f be an arbitrary vector in the root space V of W. Let

$$A_f := \{t \in T \mid \langle f, \alpha_t \rangle \in \{1, -1\}\},$$

and denote by $\mathcal{K}^f(w)$ the generalized descent class containing w, taken with respect to $A = A_f$. If f is $\mathcal{K}^f(w)$-generic then the corresponding AY representation of W, with any appropriate normalization, will be denoted $\rho^f(w)$.

Theorem 5.2. *Let $f \in V$ have integer coordinates. Then: the cell $\mathcal{K}^f(1_W)$ is a minimal AY cell for $W = S_n$ if and only if there exists a skew standard Young tableau T of size n such that*

$$f = \Delta\, cont(T).$$

Note that the cell $\mathcal{K}^f(1_W)$ is the generalized descent class $W^{\emptyset}_{A_f}$ (see Section 2). The proof of Theorem 5.2 relies on the following lemmas. The proofs of these lemmas are purely combinatorial, see [**ABR1**]. Here α_{ij} is the positive root corresponding to the reflection (transposition) $(i, j) \in S_n$ $(1 \le i < j \le n)$.

Lemma 5.1. *Under the assumptions of Theorem 5.2, if $i < j$ and $\langle f, \alpha_{ij} \rangle \in \{0, 1, -1\}$ then $w^{-1}(i) < w^{-1}(j)$ for all $w \in \mathcal{K}^f(1_W)$.*

Lemma 5.2. *Let $f \in V$ be an arbitrary vector. The cell $\mathcal{K} := \mathcal{K}^f(1_W)$ is a minimal AY cell for $W = S_n$ if and only if, for all $1 \le i < j \le n$:*

$$\langle f, \alpha_{ij} \rangle = 0 \implies \exists\ r_1, r_2 \in [i+1, j-1] \text{ s.t. } \langle f, \alpha_{ir_1} \rangle = -\langle f, \alpha_{ir_2} \rangle = 1. \tag{5.1}$$

Lemma 5.3. *The vector $c = (c_1, \ldots, c_n) \in \boldsymbol{Z}^n$ is a content vector for some skew standard Young tableaux if and only if for all $1 \le i < j \le n$*

$$c_i = c_j \Longrightarrow \exists\, r_1, r_2 \in [i+1, j-1] \text{ s.t. } c_{r_1} = c_i + 1 \text{ and } c_{r_2} = c_i - 1. \tag{5.2}$$

5.2. Minimal AY Representations of S_n. A direct combinatorial bijection between elements of minimal AY cells and standard Young tableaux follows from Theorem 5.2. This is used to prove the following result.

Theorem 5.3. *The minimal AY representations of the symmetric group S_n are exactly the skew representations, i.e., the representations determined by Young symmetrizers of skew shape. In particular, every irreducible representation of S_n may be realized as a minimal abstract Young representation.*

References

[ABR1] R. M. Adin, F. Brenti and Y. Roichman, *A unified construction of Coxeter group representations (I)*, preprint, 2003, <http://arXiv.org/abs/math.RT/0309364>.

[ABR2] R. M. Adin, F. Brenti and Y. Roichman, *A unified construction of Coxeter Group representations (II)*, in preparation.

[ABR3] R. M. Adin, F. Brenti and Y. Roichman, *A unified construction of Coxeter Group representations (III)*, in preparation.

[ABR4] R. M. Adin, F. Brenti and Y. Roichman, *Descent representations and multivariate statistics*, Trans. Amer. Math. Soc., to appear, <http://arXiv.org/abs/math.CO/0112073>.

[BR] H. Barcelo and A. Ram, *Combinatorial representation theory*, in: New Perspectives in Geometric Combinatorics, MSRI Publications, Vol. 38, 1999, 23–90.

[BB] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Springer Verlag, to appear.

[BW1] A. Björner and M. Wachs, *Generalized quotients in Coxeter groups*, Trans. Amer. Math. Soc. **308** (1988), 1–37.

[BW2] A. Björner and M. Wachs, *Permutation statistics and linear extensions of posets*, J. Combin. Theory (Ser. A) **58** (1991), 85–114.

[Bou] N. Bourbaki, *Elements of Mathematics, Lie Groups and Lie Algebras*, Springer-Verlag, Berlin, 2002.

[GL] J. J. Graham and G. I. Lehrer, *Cellular algebras*, Invent. Math. **123** (1996), 1–34.

[Ho] P. N. Hoefsmit, *Representations of Hecke algebras of finite groups with BN-pairs of classical type*, Ph.D. Thesis, Univ. of British Columbia, 1974.

[Hu1] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, New-York, 1972.

[Hu2] J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Adv. Math. 29, Cambridge Univ. Press, Cambridge, 1990.

[J] G. D. James, *The Representation Theory of the Symmetric Groups*, Lecture Notes in Math. 682, Springer-Verlag, 1978.

[JK] G. D. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Encyclopedia of Math. and its Appl. 16, Addison-Wesley, 1981.

[K] D. Kazhdan, personal communication.

[KL] D. Kazhdan and G. Lusztig, *Representations of Coxeter groups and Hecke algebras*, Invent. Math. **53** (1979), 165–184.

[OV] A. Okounkov and A. M. Vershik, *A new approach to representation theory of symmetric groups*, Selecta Math. (N. S.) **2** (1996), 581–605.

[P] I. A. Pushkarev, *On the theory of representations of the wreath products of finite groups and symmetric groups*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI) **240** (1997), Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 2, 229–244, 294–295; translation in J. Math. Sci. (New York) **96** (1999), 3590–3599.

[Ra1] A. Ram, *Seminormal representations of Weyl groups and Iwahori-Hecke algebras*, Proc. London Math. Soc. **75** (1997), 99–133.

[Ra2] A. Ram, *Affine Hecke algebras and generalized standard Young tableaux*, J. Algebra **230** (2003), 367–415.

[Sa] B. E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms and Symmetric Functions*, Wadsworth and Brooks/Cole, Monterey, CA, 1991.

[So] L. Solomon, *A decomposition of the group algebra of a finite Coxeter group*, J. Algebra **9** (1968), 220–239.

[St] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge Studies in Adv. Math. 62, Cambridge Univ. Press, Cambridge, 1999.

[T] J. Tits, *Buildings of Spherical Type and Finite BN-Pairs*, Lecture Notes in Math. 386, Springer-Verlag, Berlin and New-York, 1974.

[V] A. Vershik, *Local algebras and a new version of Young's orthogonal form*, in: Topics in Algebra, Banach Center Publications, Vol. 26 Part 2, PWN-Polish Scientific Publications, Warsaw 1990, 467–473.

Department of Mathematics and Statistics, Bar-Ilan University, Ramat-Gan 52900, Israel
E-mail address: radin@math.biu.ac.il

Dipartimento di Matematica, Universitá di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Roma, Italy
E-mail address: brenti@mat.uniroma2.it

Department of Mathematics and Statistics, Bar-Ilan University, Ramat-Gan 52900, Israel
E-mail address: yuvalr@math.biu.ac.il

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Equi-distribution over Descent Classes of the Hyperoctahedral Group (Extended Abstract)

Ron M. Adin, Francesco Brenti, and Yuval Roichman

Abstract. *A classical result of MacMahon shows that the length function and the major index are equi-distributed over the symmetric group. Foata and Schützenberger gave a remarkable refinement and proved that these parameters are equi-distributed over inverse descent classes, implying bivariate equi-distribution identities. Type B analogues and further refinements and consequences are given in this paper.*

Résumé. *Un résultat classique de MacMahon montre l'équi-distribution de l'indice majeur et de la fonction de longueur sur le groupe symétrique. Foata et Schützenberger en ont donné une raffinement remarquable et montré l'équidistribution sur les classes de descentes inverses, impliquant ainsi des équidistribu-tions bivariées. Les analogues pour le type B ainsi que d'autres raffinements et conséquences sont donnés dans cet article.*

1. Introduction

Many combinatorial identities on groups are motivated by the fundamental works of MacMahon [**M**]. Let S_n be the symmetric group acting on $1, \ldots, n$. We are interested in a refined enumeration of permutations according to (non-negative, integer valued) combinatorial parameters. Two parameters that have the same generating function are said to be *equi-distributed*. MacMahon [**M**] has shown, about a hundred years ago, that the inversion number and the major index statistics are equi-distributed on S_n (Theorem 2.2 below). In the last three decades MacMahon's theorem has received far-reaching refinements and generalizations. Bivariate distributions were first studied by Carlitz [**C**]. Foata [**F**] gave a bijective proof of MacMahon's theorem; then Foata and Schützenberger [**FS**] applied this bijection to refine MacMahon's identity, proving that the inversion number and the major index are equi-distributed over subsets of S_n with prescribed descent set of the inverse permutation (Theorem 2.3 below). Garsia and Gessel [**GG**] extended the analysis to multivariate distributions. In particular, they gave an independent proof of the Foata-Schützenberger theorem, relying on an explicit and simple generating function (see Theorem 2.4 below). Further refinements and analogues of the Foata-Schützenberger theorem were found recently, involving left-to-right minima and maxima [**RR, FH2**] and pattern-avoiding permutations [**RR, AR2**]. For a representation theoretic application of Theorem 2.3 see [**Roi**].

1991 *Mathematics Subject Classification.* Primary 05A15, 05A19; Secondary 20F55.

Key words and phrases. permutation statistics, type B, major index, length, descent class.

The research of all authors was partially supported by the EC's IHRP Programme, within the Research Training Network "Algebraic Combinatorics in Europe", grant HPRN-CT-2001-00272.

Since the length and descent parameters may be defined via the Coxeter structure of the symmetric group, it is very natural to look for analogues of the above theorems in other Coxeter groups. This is a challenging open problem. In this paper we focus on the hyperoctahedral group B_n, also known as the classical Weyl group of type B.

Despite the fact that an increasing number of enumerative results of this nature have been generalized to the hyperoctahedral group B_n (see, e.g., [**Br, FH1, Re3, Re4, Sta1**]) and that several "major index" statistics have been introduced and studied for B_n (see, e.g., [**CF1, CF2, CF3, Re1, Re2, Ste, FK**]), no generalization of MacMahon's result to B_n has been found until the recent paper [**AR1**]. There a new statistic, the *flag major index*, defined in terms of Coxeter elements, was introduced and shown to be equi-distributed with length, which is the natural analogue of inversion number from a Coxeter group theoretic point of view. A search was then initiated for a corresponding "descent statistic" that would allow the generalization to B_n of the Carlitz identity for descent number and major index [**C**], a problem first posed by Foata. In [**ABR1**] we introduced and studied two families of statistics on the hyperoctahedral group B_n, and showed that they give two generalizations of the Carlitz identity. Another solution of Foata's problem, also involving the flag major index, was presented most recently by Chow and Gessel [**CG**]. Combinatorial and algebraic properties of the flag major index were further investigated in [**AR1, HLR, AGR**]. In particular, it was shown to play an important role in the study of polynomial algebras, see [**AR1, ABR2, Ba**].

A natural goal now is to find a type B analogue of the Foata-Schützenberger theorem (Theorem 2.3); namely, to prove the equi-distribution of the flag major index and the length function on inverse descent classes of B_n. This will be carried out by finding a type B analogue of the Garsia-Gessel theorem (Theorem 2.4), which expresses the refined enumeration of the classical major index on shuffle permutations in terms of q-binomial coefficients.

The last digit parameter is involved in several closely related identities on S_n, see e.g. [**AR2, AGR, RR**]. Theorems 4.3 and 4.4 below present a refinement involving the last digit. This refinement implies a MacMahon type theorem for the classical Weyl group of type D, which is the same as the one recently proved in [**BC**]. See Subsection 5.1.

2. Background and Notation

2.1. Notation. Let $\mathbf{P} := \{1,2,3,\ldots\}$, $\mathbf{N} := \mathbf{P} \cup \{0\}$, and $\mathbf{Z}$ the ring of integers. For $n \in \mathbf{P}$ let $[n] := \{1,2,\ldots,n\}$, and also $[0] := \emptyset$. Given $m,n \in \mathbf{Z}$, $m \le n$, let $[m,n] := \{m, m+1, \ldots, n\}$. For $n \in \mathbf{P}$ denote $[\pm n] := [-n,n] \setminus \{0\}$. For $S \subset \mathbf{N}$ write $S = \{a_1,\ldots,a_r\}_<$ to mean that $S = \{a_1,\ldots,a_r\}$ and $a_1 < \ldots < a_r$. The cardinality of a set A will be denoted by $|A|$.

For $n,k \in \mathbf{N}$ denote

$$\begin{aligned} [n]_q &:= \frac{1-q^n}{1-q}; \\ [n]_q! &:= \prod_{i=1}^{n} [i]_q \qquad (n \ge 1), \qquad [0]_q! := 1; \\ \begin{bmatrix} n \\ k \end{bmatrix}_q &:= \frac{[n]_q!}{[k]_q![n-k]_q!}. \end{aligned}$$

Given a sequence $\sigma = (a_1,\ldots,a_n) \in \mathbf{Z}^n$ we say that a pair $(i,j) \in [n] \times [n]$ is an *inversion* of σ if $i < j$ and $a_i > a_j$. We say that $i \in [n-1]$ is a *descent* of σ if $a_i > a_{i+1}$. We denote by $inv(\sigma)$ (respectively, $des(\sigma)$) the number of inversions (respectively, descents) of σ. We also let

$$maj(\sigma) := \sum_{\{i:\ a_i > a_{i+1}\}} i$$

and call it the *major index* of σ.

Let $M = \{m_1, \dots, m_t\}_< \subseteq [n-1]$. Denote $m_0 := 0$ and $m_{t+1} := n$. A sequence $\sigma = (a_1, \dots, a_n)$ is an *M-shuffle* if it satisfies: if $m_i < a < b \le m_{i+1}$ for some $0 \le i \le t$, then $\sigma = (\dots, a, \dots, b, \dots)$ (i.e. a appears to the left of b in σ).

2.2. Binomial Identities. In this subsection we recall some binomial identities which will be used in the proof of Theorem 3.3.

Lemma 2.1. *For every subset $M = \{m_1, \dots, m_t\}_< \subseteq [n-1]$*

$$\prod_{j=1}^{n}(1+q^j) \cdot \begin{bmatrix} n \\ m_1 - m_0, m_2 - m_1, \dots, m_{t+1} - m_t \end{bmatrix}_q = \tag{2.1}$$

$$\sum_{\{(r_0,\dots,r_t)\,|\,m_i \le r_i \le m_{i+1}\ (\forall i)\}} \begin{bmatrix} n \\ r_0 - m_0, m_1 - r_0, \dots, r_t - m_t, m_{t+1} - r_t \end{bmatrix}_{q^2} \cdot q^{\sum_i (r_i - m_i)},$$

where $m_0 := 0$ and $m_{t+1} := n$.

For $t = 0$ (i.e., $M = \emptyset$), identity (2.1) is equivalent to a well-known classical result of Euler, comparing partitions into distinct parts with partitions into odd parts [**An**, Corollary 1.2]. The proof of the lemma is obtained by induction on t, see [**ABR3**, Lemma 3.1].

The following "q-binomial theorem" is well-known.

Theorem 2.1.

$$\prod_{i=1}^{n}(1+q^i x) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q q^{\binom{k+1}{2}} x^k.$$

2.3. The Symmetric Group. Let S_n be the symmetric group on $[n]$. Recall that S_n is a Coxeter group with respect to the Coxeter generators $S := \{s_i \mid 1 \le i \le n-1\}$, where s_i may be interpreted as the adjacent transposition $(i, i+1)$. The classical combinatorial statistics of $\pi \in S_n$, defined by viewing π as a sequence $(\pi(1), \dots, \pi(n))$, may also be defined via the Coxeter generators.

For $\pi \in S_n$ let $\ell(\pi)$ be the standard *length* of π with respect to the set of generators S. It is well-known that $\ell(\pi) = inv(\pi)$.

Given a permutation π in the symmetric group S_n, the *descent set* of π is

$$Des(\pi) := \{1 \le i < n \,|\, \ell(\pi) > \ell(\pi s_i)\} = \{1 \le i < n \,|\, \pi(i) > \pi(i+1)\}.$$

The *descent number* of $\pi \in S_n$ is $des(\pi) := |Des(\pi)|$.

The *major index*, $maj(\pi)$ is the sum (possibly zero)

$$maj(\pi) := \sum_{i \in Des(\pi)} i.$$

The *inverse descent class* in S_n corresponding to $M \subseteq [n-1]$ is the set $\{\pi \in S_n \,|\, Des(\pi^{-1}) = M\}$. Note the following relation between inverse descent classes and shuffles.

Fact 2.2. For every $M \subseteq [n-1]$,

$$\{\pi \in S_n \mid Des(\pi^{-1}) \subseteq M\} = \{\pi \in S_n \mid (\pi(1), \dots, \pi(n)) \text{ is an } M\text{-shuffle}\}.$$

MacMahon's classical theorem asserts that the length function and the major index are equi-distributed on S_n.

Theorem 2.2. (MacMahon's Theorem)

$$\sum_{\pi \in S_n} q^{\ell(\pi)} = \sum_{\pi \in S_n} q^{maj(\pi)} = [n]_q!.$$

Foata [**F**] gave a bijective proof of this theorem. Foata and Schützenberger [**FS**] applied this bijection to prove the following refinement.

Theorem 2.3. (The Foata-Schützenberger Theorem [**FS**, Theorem 1])
For every subset $B \subseteq [n-1]$,

$$\sum_{\{\pi \in S_n \ | \ Des(\pi^{-1})=B\}} q^{\ell(\pi)} = \sum_{\{\pi \in S_n \ | \ Des(\pi^{-1})=B\}} q^{maj(\pi)}.$$

This theorem implies

Corollary 2.3.

$$\sum_{\pi \in S_n} q^{\ell(\pi)} t^{des(\pi^{-1})} = \sum_{\pi \in S_n} q^{maj(\pi)} t^{des(\pi^{-1})}$$

and

$$\sum_{\pi \in S_n} q^{\ell(\pi)} t^{maj(\pi^{-1})} = \sum_{\pi \in S_n} q^{maj(\pi)} t^{maj(\pi^{-1})}.$$

An alternative proof of Theorem 2.3 may be obtained using the following classical fact [**Sta2**, Prop. 1.3.17].

Fact 2.4. Let $M = \{m_1, \ldots, m_t\}_< \subseteq [n-1]$. Denote $m_0 := 0$ and $m_{t+1} := n$. Then

$$\sum_{\{\pi \in S_n | \ Des(\pi^{-1}) \subseteq M\}} q^{inv(\pi)} = \begin{bmatrix} n \\ m_1 - m_0, m_2 - m_1, \ldots, m_{t+1} - m_t \end{bmatrix}_q.$$

Garsia and Gessel proved that a similar identity holds for the major index.

Theorem 2.4. [**GG**, Theorem 3.1] *Let* $M = \{m_1, \ldots, m_t\}_< \subseteq [n-1]$. *Denote* $m_0 = 0$ *and* $m_{t+1} = n$. *Then*

$$\sum_{\{\pi \in S_n | \ Des(\pi^{-1}) \subseteq M\}} q^{maj(\pi)} = \begin{bmatrix} n \\ m_1 - m_0, m_2 - m_1, \ldots, m_{t+1} - m_t \end{bmatrix}_q.$$

Combining this theorem with Fact 2.4 implies Theorem 2.3.

2.4. The Hyperoctahedral Group. We denote by B_n the group of all bijections σ of the set $[\pm n] := [-n, n] \setminus \{0\}$ onto itself such that

$$\sigma(-a) = -\sigma(a) \qquad (\forall a \in [\pm n]),$$

with composition as the group operation. This group is usually known as the group of "signed permutations" on $[n]$, or as the *hyperoctahedral group* of rank n. We identify S_n as a subgroup of B_n, and B_n as a subgroup of S_{2n}, in the natural ways.

If $\sigma \in B_n$ then write $\sigma = [a_1, \ldots, a_n]$ to mean that $\sigma(i) = a_i$ for $1 \le i \le n$, and let

$$inv(\sigma) := inv(a_1, \ldots, a_n),$$
$$Des_A(\sigma) := Des(a_1, \ldots, a_n),$$
$$des_A(\sigma) := des(a_1, \ldots, a_n),$$
$$maj_A(\sigma) := maj(a_1, \ldots, a_n),$$
$$Neg(\sigma) := \{i \in [n] : \ a_i < 0\},$$

and

$$neg(\sigma) := |Neg(\sigma)|.$$

It is well-known (see, e.g., [**BB**, Proposition 8.1.3]) that B_n is a Coxeter group with respect to the generating set $\{s_0, s_1, s_2, \ldots, s_{n-1}\}$, where

$$s_0 := [-1, 2, \ldots n]$$

and

$$s_i := [1, 2, \ldots, i-1, i+1, i, i+2, \ldots n] \qquad (1 \le i < n).$$

This gives rise to two other natural statistics on B_n (similarly definable for any Coxeter group), namely

$$\ell_B(\sigma) := min\{r \in \mathbf{N} : \ \sigma = s_{i_1} \ldots s_{i_r} \quad \text{for some } \ i_1, \ldots, i_r \in [0, n-1]\}$$

(known as the *length* of σ) and

$$des_B(\sigma) := |Des_B(\sigma)|,$$

where the *B-descent set* $Des_B(\sigma)$ is defined as

$$Des_B(\sigma) := \{i \in [0, n-1] \mid \ell_B(\sigma s_i) < \ell_B(\sigma)\}.$$

Remark 2.5. Note that for every $\sigma \in B_n$

$$Des_A(\sigma) = Des_B(\sigma) \setminus \{0\}.$$

There are well-known direct combinatorial ways to compute the statistics for $\sigma \in B_n$ (see, e.g., [**BB**, Propositions 8.1.1 and 8.1.2] or [**Br**, Proposition 3.1 and Corollary 3.2]), namely

$$\ell_B(\sigma) = inv(\sigma) - \sum_{i \in Neg(\sigma)} \sigma(i)$$

and

$$des_B(\sigma) = |\{i \in [0, n-1] : \ \sigma(i) > \sigma(i+1)\}|,$$

where $\sigma(0) := 0$. For example, if $\sigma = [-3, 1, -6, 2, -4, -5] \in B_6$ then $inv(\sigma) = 9$, $des_A(\sigma) = 3$, $maj_A(\sigma) = 11$, $neg(\sigma) = 4$, $\ell_B(\sigma) = 27$, and $des_B(\sigma) = 4$.

We shall also use the following formula, first observed by Incitti [**I**]:

$$\ell_B(\sigma) = \frac{inv(\overline{\sigma}) + neg(\sigma)}{2} \qquad (\forall \sigma \in B_n), \tag{2.2}$$

where $\overline{\sigma}$ denotes the sequence $(\sigma(-n), \ldots, \sigma(-1), \sigma(1), \ldots, \sigma(n))$. For example, if we take $\sigma = [-3, 5, -7, 1, 2, -4, 6]$ then $inv(\overline{\sigma}) = 35$ and $\ell_B(\sigma) = \frac{35+3}{2} = 19$.

3. Main Results

The *flag major index* of a signed permutation $\sigma \in B_n$ is defined by

$$fmaj(\sigma) := 2 \cdot maj_A(\sigma) + neg(\sigma),$$

where $maj_A(\sigma)$ is the major index of the sequence $(\sigma(1), \ldots, \sigma(n))$ with respect to the natural order $-n < \cdots < -1 < 1 < \cdots < n$.

The following is a type B analogue of the Garsia-Gessel theorem (Theorem 2.4).

Theorem 3.1. *For every subset* $M = \{m_1, \ldots, m_t\}_< \subseteq [0, n-1]$

$$\sum_{\{\sigma \in B_n \mid \ Des_B(\sigma^{-1}) \subseteq M\}} q^{fmaj(\sigma)} = \prod_{i=m_1+1}^{n} (1+q^i) \cdot \begin{bmatrix} n \\ m_1 - m_0, \ldots, m_{t+1} - m_t \end{bmatrix}_q,$$

where $m_0 := 0$ *and* $m_{t+1} := n$.

The following is a type B analogue of a classical result (Fact 2.4).

Theorem 3.2. *For every subset* $M = \{m_1, \ldots, m_t\}_< \subseteq [0, n-1]$

$$\sum_{\{\sigma \in B_n \mid \ Des_B(\sigma^{-1}) \subseteq M\}} q^{\ell_B(\sigma)} = \prod_{i=m_1+1}^{n} (1+q^i) \cdot \begin{bmatrix} n \\ m_1 - m_0, \ldots, m_{t+1} - m_t \end{bmatrix}_q,$$

where $m_0 := 0$ *and* $m_{t+1} := n$.

We deduce a Foata-Schützenberger type theorem for B_n.

Theorem 3.3. *For every subset* $M \subseteq [0, n-1]$

$$\sum_{\{\sigma\in B_n \mid \ Des_B(\sigma^{-1})=M\}} q^{\ell_B(\sigma)} = \sum_{\{\sigma\in B_n \mid \ Des_B(\sigma^{-1})=M\}} q^{fmaj(\sigma)}.$$

The following result refines Theorem 3.3.

Theorem 3.4. *For every subset* $M \subseteq [0, n-1]$ *and* $j \in [\pm n]$

$$\sum_{\{\sigma\in B_n \mid \ Des_B(\sigma^{-1})=M, \ \sigma(n)=j\}} q^{\ell_B(\sigma)} = \sum_{\{\sigma\in B_n \mid \ Des_B(\sigma^{-1})=M, \ \sigma(n)=j\}} q^{fmaj(\sigma)}.$$

An analogue of MacMahon's theorem for D_n follows; see Corollary 5.2 below.

4. Proof Outlines

Observation 4.1. Let $M = \{m_1, \ldots, m_t\}_< \subseteq [n-1]$. (Note: $0 \notin M$.) Let $m_0 := 0$ and $m_{t+1} := n$. For $\sigma \in B_n$, if $Des_A(\sigma^{-1}) = M$ then there exist r_i $(0 \le i \le t)$ such that $m_i \le r_i \le m_{i+1}$ and σ is a shuffle of the following increasing sequences:

$$(-r_0, -r_0+1, \ldots, -1\,(=\,-(m_0+1)\,)),$$
$$(r_0+1, r_0+2, \ldots, m_1),$$
$$(-r_1, -r_1+1, \ldots, -(m_1+1)),$$
$$(r_1+1, r_1+2, \ldots, m_2),$$
$$\vdots$$
$$(-r_t, -r_t+1, \ldots, -(m_t+1))$$

and

$$(r_t+1, r_t+2, \ldots, n\,(=\,m_{t+1}\,)).$$

For every i, if $r_i - m_i = 0$ $(m_{i+1} - r_i = 0)$ then the sequence $(-r_i, \ldots, -(m_i+1))$ (respectively, $(r_i + 1, \ldots, m_{i+1})$) is understood to be empty. Also, with the above notations: $0 \in Des_B(\sigma^{-1})$ if and only if $r_0 > 0$.

First, we prove the following special cases.

Theorem 4.1. *For every subset* $M = \{m_1, \ldots, m_t\}_< \subseteq [n-1]$

$$\sum_{\{\sigma\in B_n \mid \ Des_A(\sigma^{-1})\subseteq M\}} q^{fmaj(\sigma)} = \sum_{\{\sigma\in B_n \mid \ Des_A(\sigma^{-1})\subseteq M\}} q^{\ell_B(\sigma)} =$$

$$= \prod_{i=1}^{n}(1+q^i) \cdot \begin{bmatrix} n \\ m_1 - m_0, \ldots, m_{t+1} - m_t \end{bmatrix}_q,$$

where $m_0 := 0$ *and* $m_{t+1} := n$.

Theorem 4.2. *For every subset* $M = \{m_1, \ldots, m_t\}_< \subseteq [n-1]$

$$\sum_{\{\sigma\in B_n \mid \ Des_B(\sigma^{-1})\subseteq M\}} q^{fmaj(\sigma)} = \sum_{\{\sigma\in B_n \mid \ Des_B(\sigma^{-1})\subseteq M\}} q^{\ell_B(\sigma)} =$$

$$= \prod_{i=m_1+1}^{n}(1+q^i) \cdot \begin{bmatrix} n \\ m_1 - m_0, \ldots, m_{t+1} - m_t \end{bmatrix}_q,$$

where $m_0 := 0$ *and* $m_{t+1} := n$.

The proofs of Theorems 4.1 and 4.2 rely on Theorem 2.4, binomial identities (mentioned in Subsection 2.2), combinatorial properties of the length and descent functions on B_n (mentioned in Subsection 2.4) and Observation 4.1. For detailed proofs see [**ABR3**].

Proof of Theorems 3.1 and 3.2. Combine Theorems 4.1 and 4.2 with Remark 2.5.

□

Proof of Theorem 3.3. Combine Theorems 3.1 and 3.2, and apply the Principle of Inclusion-Exclusion.

□

Theorem 3.4 is an immediate consequence of the following refinements of Theorems 3.1 and 3.2.

Theorem 4.3. *Let* $n \in \mathbf{P}$, $M = \{m_1, m_2, \ldots, m_t\}_< \subseteq [0, n-1]$ *and* $i \in [\pm n]$. *Then*

$$\sum_{\{\sigma \in B_n:\ Des_B(\sigma^{-1}) \subseteq M,\ \sigma(n)=i\}} q^{fmaj(\sigma)} =$$

$$= \begin{cases} \frac{[m_r - m_{r-1}]_q}{[n]_q} \begin{bmatrix} n \\ m_1 - m_0, \ldots, m_{t+1} - m_t \end{bmatrix}_q \cdot q^{n-m_r} \prod_{j=\tilde{m}_1+1}^{n-1} (1+q^j), & \text{if } i = m_r, \\ & \text{for } r \in [t+1]; \\ \frac{[m_{r+1} - m_r]_q}{[n]_q} \begin{bmatrix} n \\ m_1 - m_0, \ldots, m_{t+1} - m_t \end{bmatrix}_q \cdot q^{n+m_r} \prod_{j=m_1+1}^{n-1} (1+q^j), & \text{if } i = -m_r - 1, \\ & \text{for } r \in [t]; \\ 0, & \text{otherwise.} \end{cases}$$

Here $m_0 := 0$, $m_{t+1} := n$, *and*

$$\tilde{m}_1 := \begin{cases} m_1 - 1, & \text{if } i = m_1; \\ m_1, & \text{otherwise.} \end{cases}$$

Theorem 4.4. *Let* $n \in \mathbf{P}$, $M = \{m_1, m_2, \ldots, m_t\}_< \subseteq [0, n-1]$ *and* $i \in [\pm n]$. *Then* $q^{\ell_B(\sigma)}$ *satisfies exactly the same formula as does* $q^{fmaj(\sigma)}$ *in Theorem 4.3.*

The proofs of Theorems 4.3 and 4.4 use case-by-case analysis.

Proof of Theorem 3.4. Combine Theorems 4.3 and 4.4, and apply the Principle of Inclusion-Exclusion.

□

Problem 4.2. *Find combinatorial (bijective) proofs for Theorems 4.3 and 4.4.*

5. Final Remarks

5.1. Classical Weyl Groups of Type D. Let D_n be the classical Weyl group of type D and rank n. For an element $\sigma \in D_n$, let $\ell_D(\sigma)$ be the length of σ with respect to the Coxeter generators of D_n. It is well-known that we may take

$$D_n = \{\sigma \in B_n \mid neg(\sigma) \equiv 0 \bmod 2\}.$$

Let $\sigma = [\sigma(1), \ldots, \sigma(n)] \in D_n$. Biagioli and Caselli [**BC**] introduced a flag major index for D_n:

$$fmaj_D(\sigma) := fmaj(\sigma(1), \ldots, \sigma(n-1), |\sigma(n)|).$$

By definition,

$$\sum_{\sigma \in D_n} q^{fmaj_D(\sigma)} = \sum_{\{\sigma \in B_n \mid\ \sigma(n) > 0\}} q^{fmaj(\sigma)}. \tag{5.1}$$

Proposition 5.1.

$$\sum_{\{\sigma\in B_n \mid \ \sigma(n)>0\}} q^{\ell_B(\sigma)} = \sum_{\sigma\in D_n} q^{\ell_D(\sigma)}.$$

For a proof see [**ABR3**, Proposition 6.1]. We deduce the following type D analogue (first proved in [**BC**]) of MacMahon's theorem.

Corollary 5.2.

$$\sum_{\sigma\in D_n} q^{fmaj_D(\sigma)} = \sum_{\sigma\in D_n} q^{\ell_D(\sigma)}.$$

Proof. Combine identity (5.1) and Proposition 5.1 with Theorem 3.4.

□

Problem 5.3. *Find an analogue of the Foata-Schützenberger theorem for D_n.*

The obvious candidate for such an analogue is false.

5.2. Two Versions of the Flag Major Index. The flag major index of $\sigma \in B_n$was originally defined as the length of a distinguished canonical expression for σ. In [**AR1**] this length was shown to be equal to $2\cdot maj_A(\sigma)+neg(\sigma)$, where the major index of the sequence $(\sigma(1),\dots,\sigma(n))$ was taken with respect to the order $-1 < \cdots < -n < 1 < \cdots < n$. In [**ABR1**] we considered a different order: $-n < \cdots < -1 < 1 < \cdots < n$ (i.e., we defined $fmaj$ as in Section 3 above).

While both versions give type B analogues of the MacMahon and Carlitz identities, only the second one gives an analogue of the Foata-Schützenberger theorem. On the other hand, the first one has the alternative natural interpretation as length, as mentioned above, and also produces a natural analogue of the signed Mahonian formula of Gessel and Simion, see [**AGR**]. The relation between these two versions and their (possibly different) algebraic roles requires further study.

References

[ABR1] R. M. Adin, F. Brenti and Y. Roichman, *Descent numbers and major indices for the hyperoctahedral group*, Special issue in honor of Dominique Foata's 65th birthday (Philadelphia, PA 2000), Adv. Appl. Math. **27** (2001), 210–224.

[ABR2] R. M. Adin, F. Brenti and Y. Roichman, *Descent representations and multivariate statistics*, Trans. Amer. Math. Soc., to appear, <http://arXiv.org/abs/math.CO/0112073>.

[ABR3] R. M. Adin, F. Brenti and Y. Roichman, *A type B analogue of the Foata-Schützenberger theorem*, preprint, 2003.

[AGR] R. M. Adin, I. M. Gessel and Y. Roichman, *Signed Mahonians*, in preparation.

[AR1] R. M. Adin and Y. Roichman, *The flag major index and group actions on polynomial rings*. Europ. J. Combin. **22** (2001), 431–446.

[AR2] R. M. Adin and Y. Roichman, *Equidistribution and sign-balance on* 321*-avoiding permutations*, Sém. Lothar. Combin., to appear, <http://arXiv.org/abs/math.CO/0304429>.

[An] G. E. Andrews, *The Theory of Partitions*, Cambridge Univ. Press, Cambridge, 1984.

[Ba] E. Bagno, *Combinatorial parameters on classical groups*, Ph.D. Thesis, Bar-Ilan Univ., Ramat-Gan, 2003.

[BC] R. Biagioli and F. Caselli, *Invariant algebras and major indices for classical Weyl groups*, preprint, 2002.

[BB] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, Springer-Verlag, to appear.

[Br] F. Brenti, *q-Eulerian polynomials arising from Coxeter groups*, Europ. J. Combin. **15** (1994), 417–441.

[C] L. Carlitz, *A combinatorial property of q-Eulerian numbers*, Amer. Math. Monthly **82** (1975), 51–54.

[CF1] R. J. Clarke and D. Foata, *Eulerian calculus. I. Univariable statistics*, Europ. J. Combin. **15** (1994), 345–362.

[CF2] R. J. Clarke and D. Foata, *Eulerian calculus. II. An extension of Han's fundamental transformation*, Europ. J. Combin. **16** (1995), 221–252.

[CF3] R. J. Clarke and D. Foata, *Eulerian calculus. III. The ubiquitous Cauchy formula*, Europ. J. Combin. **16** (1995), 329–355.

[CG] C. O. Chow and I. M. Gessel, *On the descent numbers and major indices for the hyperoctahedral group*, preprint, 2003.

[F] D. Foata, *On the Netto inversion number of a sequence*, Proc. Amer. Math. Soc. **19** (1968), 236–240.

[FH1] D. Foata and G.-N. Han, *Calcul basique des permutations signées. I. Longueur et nombre d'inversions*, Adv. Appl. Math. **18** (1997), 489–509.

[FH2] D. Foata and G.-N. Han, *Further properties of the second fundamental transformation on words*, preprint, 2003.

[FK] D. Foata and C. Krattenthaler, *Graphical major indices, II*, Sém. Lothar. Combin. **34** (1995), B34k, 16 pp.

[FS] D. Foata and M. P. Schützenberger, *Major index and inversion number of permutations*. Math. Nachr. **83** (1978), 143–159.

[GG] A. M. Garsia and I. Gessel, *Permutation statistics and partitions.* Adv. Math. **31** (1979), 288–305.

[HLR] J. Haglund, N. Loehr and J. Remmel, *Statistics on wreath products, perfect matchings and signed words*, preprint, 2003.

[H] J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Math., no. 29, Cambridge Univ. Press, Cambridge, 1990.

[I] F. Incitti, *The Bruhat order on the involutions of the hyperoctahedral group*, preprint, 2003.

[M] P. A. MacMahon, *Combinatory Analysis I-II.* Cambridge Univ. Press, London/New-York, 1916. (Reprinted by Chelsea, New-York, 1960.)

[RR] A. Regev and Y. Roichman, *Permutation statistics on the alternating group*, Adv. Appl. Math., to appear, `<http://arXiv.org/abs/math.CO/0302301>`.

[Re1] V. Reiner, *Signed permutation statistics*, Europ. J. Combin. **14** (1993), 553–567.

[Re2] V. Reiner, *Signed permutation statistics and cycle type*, Europ. J. Combin. **14** (1993), 569–579.

[Re3] V. Reiner, *Upper binomial posets and signed permutation statistics*, Europ. J. Combin. **14** (1993), 581–588.

[Re4] V. Reiner, *The distribution of descent and length in a Coxeter group*, Electron. J. Combin. **2** (1995), R25.

[Roi] Y. Roichman, *Schubert polynomials, Kazhdan-Lusztig basis and characters*, Discrete Math. **217** (2000), 353–365.

[Ros] D. P. Roselle, *Coefficients associated with the expansion of certain products*, Proc. Amer. Math. Soc. **45** (1974), 144–150.

[Sta1] R. P. Stanley, *Binomial posets, Möbius inversion, and permutation enumeration*, J. Combin. Theory (Ser. A) **20** (1976), 336–356.

[Sta2] R. P. Stanley, *Enumerative Combinatorics*, vol. 1, Wadsworth and Brooks/Cole, Monterey, CA, 1986.

[Sta3] R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge Univ. Press, New York/Cambridge, 1999.

[Ste] E. Steingrimsson, *Permutation statistics of indexed permutations*, Europ. J. Combin. **15** (1994), 187–205.

Department of Mathematics and Statistics, Bar-Ilan University, Ramat-Gan 52900, Israel
E-mail address: `radin@math.biu.ac.il`

Dipartimento di Matematica, Universitá di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Roma, Italy
E-mail address: `brenti@mat.uniroma2.it`

Department of Mathematics and Statistics, Bar-Ilan University, Ramat-Gan 52900, Israel
E-mail address: `yuvalr@math.biu.ac.il`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

A New Representation of Formal Power Series

Kostyantyn V. Archangelsky

Abstract. *This paper is dedicated to the genesis arising at the boundary between the theory of formal power series (FPS) and combinatorics.*

Similarly to combinatorics where any rational sequence of natural numbers $\{r_k\}_{k\geq 0}$ is representable for all k in the form

(1.1) $$r_{k+n} = \sum_{i=1}^{n} r_{k+n-i} X_{n-i}$$

*where X_j – are, generally speaking, complex numbers (Berstel,Reutenauer, [**BR**]), we prove that any rational FPS r is representable in the form (1) where $r_s = \sum_{|w|=s}(r,w)w$, and X_j are elements of some special skew field. As a trivial consequence of such a representation were obtained: 1)truthfulness of Eilenberg's Equality Theorem [**E**], decidability of the equivalence problem of finite multitape deterministic automata (Rabin,Scott [**RS**]) and decidability of problem of whether two given morphisms are equivalent on regular language, (Culik, Salomaa [**CS**]); 2) more simply formulated and proved the results from monographs on FPS (Salomaa, Soittola [**SS**], Berstel, Reutenauer [**BR**], Kuich, Salomaa [**KS**]); 3) solved partial cases of the problem of existence for an inverse element of Hadamard product and others; 4) provided 3 Conjectures and 10 Open problems.*

The conclusion contains a complete comparative analysis of the attempts to utilize linear recurrence in theory of FPS by other authors.

Résumé.

We use the standard notations from monographs Berstel, Reuteneuer [**BR**] and Cohn [**Coh**]. In particular, it will be assumed that $\Sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_t\}$ is a finite alphabet, $\Sigma^{-1} = \{\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_t^{-1}\}$, ε is empty word and unity in semigroup Σ^* and group G, generated by Σ, $\emptyset$ is empty set and zero in semirings and fields, generated by Σ, $\underline{\varepsilon}, \underline{\sigma}_i, \underline{\sigma}_i^{-1}$ are corresponding characteristic FPS, $\boldsymbol{k}$ is commutative zero-divizor-free semiring embeddable in commutative field $\boldsymbol{K}$ (this includes the semirings $\boldsymbol{N}, \boldsymbol{Z}, \boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{C}$).

According to Salomaa, Soittola [**SS**], every FPS $r \in \boldsymbol{k}^{rat} \ll \Sigma^* \gg$ can be represented as a behaviour of $\boldsymbol{k} - \Sigma^*$-automaton

$$\mathfrak{A} = < \{q_1, q_2, \ldots, q_n\}, A, q_1, F >$$

where $A \in \boldsymbol{k}^{n\times n} < \Sigma >$ - transition matrix, q_1-initial state, $F \in \boldsymbol{k}^{n\times 1} < \{\underline{\varepsilon}, \emptyset\} >$ - final states:

$$r_{\mathfrak{A}} = \sum_{i=0}^{\infty} (A^i F)_1.$$

1991 *Mathematics Subject Classification.* Primary 13F25, 11B34; Secondary 13J05, 11B39.

Key words and phrases. formal power series, representation, recurrence.

We denote $q_i^{(j)} = (A^j F)_i$, then $r_{\mathfrak{A}} = \sum_{i=0}^{\infty} q_1^{(i)}$ and

$$q_i^{(j)} = \sum_{s=1}^{n} A_{is} q_s^{(j-1)} \tag{1.2}$$

Let us consider a system of n equations with $(n+1)$ unknowns $X_0, X_1, \ldots, X_n$:

$$\begin{cases} q_i^{(n)} X_n = \sum_{j=1}^{n} q_i^{(n-j)} X_{n-j}, \ i = \overline{1, n} \end{cases} \tag{1.3}$$

and show that it always has a non-zero solution in Malcev-Neumann skew field $\boldsymbol{K}((G))$ of FPS with well-ordered support.

We solve the system (3) following the usual Gauss algorithm by successive excluding unknowns $X_0, X_1, \ldots, X_n$. On step 0 all coefficients of unknowns are $q_i^{(j)} \in \boldsymbol{k} \ll \Sigma^* \gg$ and of course are elements of $\boldsymbol{K}((G))$. Let us assume that on step i an equation for X_i has the form of

$$q_{in} X_n = q_{i(n-1)} X_{n-1} + \ldots + q_{ii} X_i, \ q_{is} \in \boldsymbol{K} \ll G \gg, \ s = \overline{n, i} \tag{1.4}$$

and we can compute a leading term for every q_{is}.

Suppose $q_{ii} \neq \emptyset$ (otherwise we can exchange i-th column with one of $n-1, ..., i+1$; if all $q_{is} = \emptyset$ for $s = \overline{i, n-1}$ then assume $X_n = \emptyset$ and go to an equation for X_{i+1}). Then for all non-zero q_{is} denote $q_{is} = \alpha_{is} + q'_{is}$ where α_{is} is a leading term in q_{is}. It follows that after multiplying both parts of the equation on α_{ii}^{-1} and solving it for X_i we obtain

$$X_i = \left(-\alpha_{ii}^{-1} q'_{ii}\right)^* \left(\alpha_{ii}^{-1} \alpha_{in} + \alpha_{ii}^{-1} q'_{in}\right) X_n - \cdots - (\ldots) X_{i+1} \tag{1.5}$$

where the bracket content $(\ldots)$ is analogous to the coefficient of X_n and is not provided for the sake of simplicity. Substitute equation (5) into the remaining equations for $X_j, \ j = \overline{1, i-1}$:

$$\begin{aligned} (\alpha_{jn} + q'_{jn} &- (\alpha_{ji} + q'_{ji})(-\alpha_{ii}^{-1} q'_{ii})^* (\alpha_{ii}^{-1} \alpha_{in} + \alpha_{ii}^{-1} q'_{in})) X_n = \\ &= (\cdots) X_{n-1} + \cdots + (\cdots) X_{i+1} \end{aligned} \tag{1.6}$$

Since $supp(\alpha_{ii}^{-1} q'_{ii}) > \varepsilon$ then leading term of the coefficient of X_n should be searched for in $\alpha_{jn} - \alpha_{ji} \alpha_{ii}^{-1} \alpha_{in}$. If the coefficient equals $\emptyset$ then we take next in ascending order elements from $supp(q'_{jn}), supp(q'_{ji})$, $supp(\alpha_{ii}^{-1} q'_{in})$ and so on. This process is constructive (see Lewin [**L**], Cohn [**Coh**]), so the inductive hypothesis holds true – at the beginning of next step of Gauss algorithm all coefficients of unknowns $X_n, X_{n-1}, \ldots X_{i+1}$ will be again from $\boldsymbol{K}((G))$ with known leading terms.

At the last step for the equation $q_{nn} X_n = q_{n(n-1)} X_{n-1}$ we have:

(i) if $q_{n(n-1)} \neq \emptyset$ that is $q_{n(n-1)} = \alpha_{n(n-1)} + q'_{n(n-1)}$ then assume
$X_n = \underline{\varepsilon}, X_{n-1} = \left(-\alpha_{n(n-1)}^{-1} q'_{n(n-1)}\right)^* \alpha_{n(n-1)}^{-1} q_{nn}$;

(ii) if $q_{n(n-1)} = \emptyset$, then assume $X_n = \emptyset, X_{n-1} = \underline{\varepsilon}$. So we proved

Theorem 1.1. *A solution of the system (3) in $\boldsymbol{K}((G))$ exists always in the form:*

$$(\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_p, \underline{\varepsilon}, \emptyset, \ldots, \emptyset), \ 1 \leq p \leq n-1 \tag{1.7}$$

while some $\tilde{X}_i$ also can be $\emptyset$. ■

We prove the main theorem of the paper.

Theorem 1.2. *For all $k \in \mathbf{N}$ holds*

$$q_i^{(n+k)} = \sum_{j=1}^{n} q_i^{(n+k-j)} \tilde{X}_{n-j}, \ i = \overline{1,n} \tag{1.8}$$

Proof. For $k = 0$ the statement is proved - suppose it is true for k. Then
$q_i^{(n+k+1)} \overset{(2)}{=} \sum_{j=1}^{n} A_{ij} q_j^{(n+k)} = \sum_{j=1}^{n} A_{ij} \sum_{l=1}^{n} q_j^{(n+k-l)} \tilde{X}_{n-l} =$
$= \sum_{l=1}^{n} \left(\sum_{j=1}^{n} A_{ij} q_j^{(n+k-l)} \right) \tilde{X}_{n-l} \overset{(2)}{=} \sum_{l=1}^{n} q_i^{(n+k+1-l)} \tilde{X}_{n-l}$.■

Definition 1. We call a representation of FPS q_i in the form (8) *a linear recurrence representation* (further referred shortly as LRR), vector-solution (7) - *a stencil*, $q_i^{(j)}$ - j-th *layer* of FPS q_i. ■

Example 1.4. Following the considerations about the solving of system (3) one can find that FPS $s = (\underline{a}^2(\underline{ab})^*\underline{b}^2(\underline{ab})^*)^*$ has LRR

$s_4 = \underline{a}^2\underline{b}^2, s_3 = s_2 = s_1 = \emptyset, s_0 = \underline{\varepsilon},$

$s_{n+5} = s_{n+4} \cdot \emptyset + s_{n+3}(\underline{b}^{-2}\underline{ab}^3 + \underline{ab}) + s_{n+2} \cdot \emptyset + s_{n+1}(\underline{a}^2\underline{b}^2 - \underline{ab}^{-1}\underline{ab}^3) + s_n \cdot \emptyset$.■

Having analyzed the process of solving system (3) it is not difficult to prove:

Theorem 1.3. *There exists a stencil with coefficients from the set $\{-1, 0, 1\}$ for a characteristic series of an arbitrary given rational languages.* ■

The opposite is intresting:

Open problem 1. *("Fatou extension") If : 1) stensil of FPS r has all the coefficients from the set $\{-1, 0, 1\}$, 2) layers of r have coefficients from the set $\{0, 1\}$, then: r is $\boldsymbol{N}$-rational? And $\boldsymbol{Z}$-rational? And K-algebraic?* (we point out to the relationship of this problem with the counter-example Reutenauer [**R**]). ■

Corollary 1 (Eilenberg's Equality Theorem [**E**]). *Let $\mathfrak{A} =< \{q_1, \ldots, q_n\}, A, q_1, F_1 >$ and $\mathfrak{B} =< \{p_1, \ldots, p_m\}, B, p_1, F_2 >$ be $\boldsymbol{k} - \Sigma^*$-automata. Then $r_{\mathfrak{A}} = r_{\mathfrak{B}}$ iff $(r_{\mathfrak{A}}, w) = (r_{\mathfrak{B}}, w)$ for all $w \in \Sigma^*$ of length at most (n+m-1).*

Proof. Consider a system of equations:

$$\begin{cases} q_i^{n+m} = \sum_{j=1}^{n+m} q_i^{(n+m-j)} X_{n+m-j}, & i = \overline{1,n} \\ p_i^{n+m} = \sum_{j=1}^{n+m} p_i^{(n+m-j)} X_{n+m-j}, & i = \overline{1,m} \quad ■ \end{cases} \tag{9}$$

Definition 2. We call the solution of the system (9) and the system itself *a common stencil for automata* $\mathfrak{A}$ and $\mathfrak{B}$. Common stencil exists for any finite number of $\boldsymbol{k} - \Sigma^*$-automata. ■

Corollary 2. *(Equivalence Problem for Multitape Deterministic Finite Autumata, Rabin,Scott* [**RS**]*) Two automata $\mathfrak{A}_1 =< \{q_1, \ldots, q_n\}, \Sigma_1 \cup \Sigma_2 \cup \ldots \cup \Sigma_k, \delta_1, q_1, F_1 >$ and $\mathfrak{A}_2 =< \{p_1, \ldots, p_m\}, \Sigma_1 \cup \Sigma_2 \cup \ldots \cup \Sigma_k, \delta_2, p_1, F_2 >$ are equivalent iff the sets of their acceptable words of length at most $(n + m - 1)$ are equal.*

Proof. Consider common stencil for automata $\mathfrak{A}$ and $\mathfrak{B}$. As the direct product of fully ordered groups equipped with lexicographic order $\Sigma_1 < \Sigma_2 < \ldots < \Sigma_k$ is still fully ordered group G_k (Passman [**Pa**]), hence this common stencil exists in form of solution for system (9) in $\boldsymbol{Z}((G_k))$. ■

Remark 1. *Harju, Karhumaki* [**HK**] *result is also in checking up of all words of length at most $(n+m-1)$. To check the equivalence of two finite multitape deterministic automata an exponential time, therefore, is required. At the same time there exist polynomial algorithms for the checking of the equivalence of $\boldsymbol{k} - \Sigma^*$- automata ($O(n^4)$ – Tzeng* [**T**]*, $O(n^3)$ – Archangelsky* [**A1**]*). This provides a hint that should exist a polynomial algorithm. Indeed not every initial set of layers should be checked up because not all of them in combination with stencil would generate only 'clean' noncommutative polynomials – the ones without σ_i^{-1}.* ■

Open problem 2. *How many 'clean' tuples of layers there exist for a given stencil?* ■

Remark 2. *Corollary 2 could have been proven more simpler by leaving out the process of finding of stencil and the proof of its existence. According to Hebish, Weinert* [**HW**] *the semiring of FPS on partially commutative monoids over* $\boldsymbol{Z}$ *is zero-divisor-free and additivelly- cancellative and multiplicately-left-cancellative. This means that a solution of the system (9) exists over some partially commutative skew field.* ■

Corollary 3. *(Equivalence Problem for Morphisms on Regular Languages, Culik, Salomaa* [**CS**]*) Let* L *be a regular language, defined by the minimal deterministic automaton* $\mathfrak{A} =< \{q_1, \ldots, q_n\},$ $\Sigma, \delta, q_1, F >$*, and* $h, g : \Sigma^* \to \Delta^*$ *be morphisms. If* $h(w) = g(w)$ *for all* $w \in L$ *of length at most* $(2n - 1)$*, then* $h(w) = g(w)$ *for all* $w \in L$*.*

Proof. One may assume that $\Sigma \cap \Delta = \emptyset$ and letters from Σ and Δ commute. We define the transition function δ_h in 2-tape automaton $\mathfrak{A}_h =< \{q_1, \ldots, q_n\}, \Sigma \cup \Delta, \delta_h, q_1, F >$ as follows $\delta_h(q_i, \sigma_j h(\sigma_j)) = q_k \Leftrightarrow \delta(q_i, \sigma_j) = q_k$. Similarly define δ_g and $\mathfrak{A}_g$. Until common stencil of $\mathfrak{A}_h$ and $\mathfrak{A}_g$ is being built we assume for convenience each $\sigma_j h(\sigma_j)$ and $\sigma_j g(\sigma_j)$ to be one unique letter. Thus the length of common stencil will be $2n$. ■

Remark 3. *Proof of Corollary 3 does not use unlike Karhumaki* [**K**] *an Eihrengeucht's conjecture. Actually we have proven a more stronger result – the decidability of morphism equivalence on regular language with multiplicities of words are taken into consideration.*■

Corollary 4. *Let* $r \in \boldsymbol{k}^{rat} \ll \Sigma^* \gg$ *and* p *be number of first nonzero element in stencil of* r*, i.e.* $\tilde{X}_0 = ... = \tilde{X}_{p-1} = \emptyset, \tilde{X}_p \neq \emptyset$, $0 \leq p \leq n$*. Then*

(i) r *is identecically zero iff* $r_i = \emptyset$ *for all* $i = \overline{0, (n-1)}$*;*

(ii) r *is polynomial iff* $r_i = \emptyset$ *for all* $i = \overline{p, (n-1)}$*;*

(iii) r *is ultimately constant iff* $r_i = c\underline{\Sigma}^i$ *for all* $i = \overline{p, (n-1)}$*;*

(iv) r *is identically constant iff* $r_i = c\underline{\Sigma}^i$ *for all* $i = \overline{0, (n-1)}$*.*

Proof. Trivial combinatorical considerations.■

Remark 4. *Proof of Corollary 4 does not use, unlike Salomaa, Soittola* [**SS**]*, Kuich, Salomaa* [**KS**] *Hadamard product and morphisms.* ■

Let us investigate more scrupulously how of the summands with negative powers of letters in $\sum r_i \tilde{X}_i$ annihilate. In the first approximation it can be done by tracing down step-by-step how only 'clean' non-commutative polynomials are left in the following examples.

Example 1.5 (Berstel, Reutenauer [**BR**]). *FPS* $s = \sum_w |w|_a w = \underline{\Sigma}^* \underline{a} \underline{\Sigma}^*$ *has the follows LRR:*

$$s_0 = \emptyset, s_1 = \underline{a}$$

$$s_{n+2} = s_{n+1}(2\underline{a} + \underline{b} + \underline{a}^{-1}\underline{b}\ \underline{a}) + s_n(-\underline{a}^2 - 2\underline{b}\ \underline{a} - \underline{b}\ \underline{a}^{-1}\ \underline{b}\ \underline{a})$$ ■

Example 1.6 (Berstel, Reutenauer [**BR**]). *FPS* $s = \sum_w (|w|_a - |w|_b)w = \underline{\Sigma}^*(\underline{a} - \underline{b})\underline{\Sigma}^*$ has the follows LRR:

$$s_0 = \emptyset, s_1 = \underline{a} - \underline{b},$$

$$s_{n+2} = s_{n+1} \cdot 2(\underline{a}^{-1}\underline{b})^*(\underline{a} - \underline{a}^{-1}\underline{b}^2) + s_n(\underline{a} + \underline{b})(\underline{a} + \underline{b} - 2(\underline{a}^{-1}\underline{b})^*(\underline{a} - \underline{a}^{-1}\underline{b}^2))$$ ■

Example 1.7 (Reutenauer [**R**]). *FPS*

$$s = \sum_w \left(\alpha^{2(|w|_x - |w|_y)} + \alpha^{2(|w|_y - |w|_x)}\right) w = \left(\alpha^2 \underline{x} + \alpha^{-2}\underline{y}\right)^* + \left(\alpha^{-2}\underline{x} + \alpha^2 \underline{y}\right)^*$$

$$\alpha = \frac{1}{2}(\sqrt{5} + 1),$$

has the follows LRR: $s_0 = 2\underline{\varepsilon}, s_1 = 3\underline{x} + 3\underline{y}$,

$$s_{n+2} = s_{n+1} \cdot 3(\underline{x}^{-1}\underline{y})^*(\underline{x} - \underline{x}^{-1}\underline{y}^2) + s_n\left(\alpha^{-2}\underline{x} + \alpha^2\underline{y}\right)\left(\alpha^{-2}\underline{x} + \alpha^2\underline{y} - -3(\underline{x}^{-1}\underline{y})^*(\underline{x} - \underline{x}^{-1}\underline{y}^2)\right) \blacksquare$$

Let us consider arbitrary sequential n-tuple of layers of $r \in \boldsymbol{k}^{rat} << \Sigma^* >>$. One can say that they are n-inert in ring $\boldsymbol{K}((G))$ in several weak sense because $r_{k+n-i} \in \boldsymbol{k} < \Sigma^* >$ (and of course, $r_{k+n-i} \in \boldsymbol{K}((G))$), $\tilde{X}_{n-i} \in \boldsymbol{K}((G))$, but $\sum_{i=1}^{n} r_{k+n-i}\tilde{X}_{n-i} \in \boldsymbol{k} < \Sigma^* >$. And while the inertia theorem is proved (Bergman [**Ber**], Cohn [**Coh**]) also for ring $\boldsymbol{k} < \Sigma^* >$ in ring $\boldsymbol{K} \ll \Sigma^* \gg$, but not in ring $\boldsymbol{K}((G))$, the following analogue seems to be the case.

Conjecture 1. $\boldsymbol{k} < \Sigma^* >$ *is inert in the* $\boldsymbol{K}((G))$. $\blacksquare$

Conjecture 2. *Assuming Conjecture 1 is true - would matrix-trivializer* M *exist such that* $M, M^{-1} \in \boldsymbol{K}^{n\times n} \ll \Sigma^* \gg$*?* $\blacksquare$

Formulae (8) implies the following formulae for computing the coefficients in LRR:

$$(r_{k+n}, w) = \sum_{\substack{(1)\ 1\le i\le n \\ (2) w_{is}\tilde{w}_{is}=w,\ w_{is}\in\Sigma^*,\ \tilde{w}_{is}\in G}} (r_{k+n-i}, w_{is})(\tilde{X}_{n-i}, \tilde{w}_{is}) \tag{10}$$

The second condition of summing means that $w_{is} = \alpha_{is}\beta_{is},\ \beta_{is}^{-1}\gamma_{is} = \tilde{w}_{is},\ \alpha_{is}, \beta_{is}, \gamma_{is} \in \Sigma^*$. Therefore $|\beta_{is}| \le |w_{is}| = k + n - i$ and number of summands in (10) is limited.

Conjecture 3. *Would the length of canceling suffixes and prefixes (like a* β_{is}*) be limited too for each LRR?*$\blacksquare$

Open problem 3 (Archangelsky [**A2**])**.** *For a given* $\tilde{r} \in \boldsymbol{K}((G))$ *determine whether the lengths of all negative subwords of words in supp* $(\tilde{r})$ *are limited (i.e., subwords in alphabet* Σ^{-1} *only).*$\blacksquare$

We apply rule (10) for examining coefficients in the inverse element of Hadamard product.We mean FPS p is *Hadamard inverse* of FPS r iff $r \odot p = \sum_w 1 \cdot w = \underline{\Sigma}^*$. The problem of existence of such an element is still open. All papers on the issue either study FPS on cyclic/commutative semigroups (Cori [**Cor**], Benzaghou [**Ben1, Ben2**],Benzaghou, Bezivin [**BB**], Anselmo, Bertoni [**AB**], Poorten [**Po**]) or simple samples of inversable FPS on $\Sigma^*, |\Sigma| \ge 2$ (Gerardin [**G**]).

Theorem 1.4. *Let* Σ *be alphabet,* $|\Sigma| \ge 2$, $\mathfrak{A}_r, \mathfrak{A}_p$ *be* $\boldsymbol{Q}^+ - \Sigma^*$ *– automata which behaviours are FPS* r, p *and let the coefficients in the common stencil of automata*

$$\begin{cases} \mathfrak{A}_r \\ \mathfrak{A}_p \\ q = \underline{\Sigma} q + \underline{\varepsilon} \end{cases} \tag{11}$$

are in $\mathbf{Q}^+$*. Then* $r \odot p = \underline{\Sigma}^*$ *implies both* r, p *have a finite image.*

Proof.Consider common stencil of automata (11) (t is the sum of states for automata $\mathfrak{A}_r$ and $\mathfrak{A}_p$ plus 1):

$$\begin{cases} r_{n+t} = \Sigma_{i=1}^{t} r_{n+t-i}\tilde{X}_{t-i} \\ p_{n+t} = \Sigma_{i=1}^{t} p_{n+t-i}\tilde{X}_{t-i} \\ \underline{\Sigma}^{n+t} = \Sigma_{i=1}^{t}\underline{\Sigma}^{n+t-i}\tilde{X}_{t-i}, \qquad n \ge 0 \end{cases} \tag{12}$$

According to (12) and (10) a coefficient of the word $w \in supp(r_{n+t})$ in r_{n+t} satisfies the follows:

$$\alpha = \textstyle\sum_{s\in S} \alpha_s x_s \tag{13}$$

where α_s – coefficients of $supp(r_{n+t-i})$ and x_s – coefficients of $supp(\tilde{X}_{t-i})$, and $|S|$ is finite. Respectively,

$$\frac{1}{\alpha} = \textstyle\sum_{s\in S} \frac{1}{\alpha_s} x_s \tag{14}$$

$$1 = \textstyle\sum_{s\in S} 1 \cdot x_s \tag{15}$$

Multiplying (13) and (14) we obtain

$$\alpha \cdot \frac{1}{\alpha} = 1 = \left(\sum_{s\in S} \alpha_s x_s\right)\left(\sum_{s\in S} \frac{1}{\alpha_s} x_s\right) =$$

$$= \sum_{s\in S} x_s^2 + \sum_{i\neq j; i,j\in S} \left(\frac{\alpha_i}{\alpha_j} + \frac{\alpha_j}{\alpha_i}\right) x_i x_j \geq \sum_{s\in S} x_s^2 + 2\sum_{i\neq j; i,j\in S} x_i x_j =$$

$$= \left(\sum_{s\in S} x_s\right)^2 = 1,$$

that is why all $\alpha_i = \alpha_j = \alpha$, i.e. new coefficients do not appear in r_{n+t}. ■

Open problem 4. *Positiveness of all coefficients in all stencils and layers is an essential part of the proof of Theorem 3. In general case this limitation would not exist – therefore one would require to solve (or describe the set of solutions for) the system of Diophantine equations* $\{(13),(14),(15)\}$ $(\alpha_i \in \boldsymbol{N}^+, x_i \in \boldsymbol{Q})$. *For small numbers of unknowns the system above indeed has only trivial solutions. It seems like the class of invertable by Hadamard rational FPS is very narrow.* ■

Method of Theorem 4 may be implememted for the obtaining a necessary condition for the solution of following

Open problem 5 (Restivo,Reutenauer [**RR**]). *Let s be a FPS with integer coefficients and p a prime number; if $\sum_w p^{(s,w)} w$ is rational, then so are s and $\sum_w p^{-(s,w)}$.*■

Corollary 5. *If $s \in \boldsymbol{Q}^{rat} \ll \Sigma^* \gg$, $p \in \boldsymbol{N}$, $s_1 = \sum_w p^{(s,w)}$, $s_2 = \sum_w p^{-(s,w)}$, $s_1, s_2 \in (\boldsymbol{Q}^+)^{rat} \ll \Sigma^* \gg$ and the coefficients of the common stencil of automata*

$$\begin{cases} \mathfrak{A}_{s_1} \\ \mathfrak{A}_{s_2} \\ q = \underline{\Sigma} q + \underline{\varepsilon} \end{cases}$$

are in $\mathbf{R}^+$, *then s, s_1, s_2 have a finite image.*■

Let us try for a given LRR build a FPS, a representation of which the former is:

$$r = \sum_{i=0}^{n-1} r_i + \sum_{i=n}^{\infty} r_i = \sum_{i=0}^{n-1} r_i + \sum_{i=n}^{\infty}\sum_{j=1}^{n} r_{i-j}\tilde{X}_{n-j} =$$

$$= \sum_{i=0}^{n} r_i + \sum_{j=1}^{n}\sum_{i=j-1}^{\infty} r_i \tilde{X}_{j-1} =$$

$$= \sum_{i=0}^{n-1} - \sum_{j=1}^{n-1}\sum_{s=0}^{j-1} r_s \tilde{X}_j + \sum_{j=1}^{n}\sum_{i=0}^{\infty} r_i \tilde{X}_{j-1} =$$

$$= r_0 + \sum_{i=1}^{n-1}(r_i - \sum_{s=0}^{i-1} r_s \tilde{X}_i) + r\sum_{j=1}^{n} \tilde{X}_{j-1} \tag{16}$$

Solve this equation for r:

$$r = (r_0 + \sum_{i=1}^{n-1}(r_i - \sum_{s=0}^{i-1} \tilde{X}_i))(\sum_{j=1}^{n} \tilde{X}_{j-1})^* \tag{17}$$

Unarguably we took too much liberty when applying limit to both parts of identity (16). It still needs to be proved that the obtained expression is indeed the sum of r_i and only them. Because of size limit we would not do that but do illustrate using Example 2 that it is true:

$(\underline{a}+\underline{b})^*\underline{a}(\underline{a}+\underline{b})^* = s \overset{(17)}{=} (s_1 + s_0 - s_0\tilde{X}_1)(\tilde{X}_1 + \tilde{X}_2)^* =$
$= \underline{a}(2\underline{a} + \underline{b} + \underline{a}^{-1}\underline{ba} - \underline{a}^2 - 2\underline{ba} - \underline{ba}^{-1}\underline{ba})^* = ((\underline{\varepsilon} - 2\underline{a} - \underline{b} - \underline{a}^{-1}\ \underline{b}\ \underline{a} +$
$+ \underline{a}^2 + 2\underline{b}\ \underline{a} + \underline{b}\ \underline{a}^{-1}\ \underline{b}\ \underline{a})\underline{a}^{-1})^{-1} = ((\underline{a}^{-1} - \underline{\varepsilon} - \underline{b}\ \underline{a}^{-1})(\underline{\varepsilon} - \underline{a} - \underline{b}))^{-1} =$
$= ((\underline{\varepsilon} - \underline{a} - \underline{b})\underline{a}^{-1}(\underline{\varepsilon} - \underline{a} - \underline{b}))^{-1} = (\underline{a}+\underline{b})^*\underline{a}(\underline{a}+\underline{b})^*$. ■

Brzozowski, Cohen [**BC**] studied a decompositions of rational languages into star languages : $P = R^*S$.One may ask about such decomposition in $\boldsymbol{K}((G))$. Of course, arbitrary regular language R may be trivially decomposed into star FPS in $\boldsymbol{K}((G))$: $\underline{R} = \underline{P}^*(\underline{\varepsilon} - \underline{P})\underline{R}$, where P is arbitrary regular language too.

It is interesting to study a nontrivial case. Consider a common stencil of two arbitrary FPS in the form (17). It implies

Theorem 1.5. *Each two FPS* $r, p \in \boldsymbol{k}^{rat} \ll \Sigma^* \gg$ *have a representation in* $\boldsymbol{K}((G))$ *with nontrivial common star factor :* $r = \tilde{r}_1\tilde{s}^*, p = \tilde{p}_1\tilde{s}^*$. ■

Judging by appearance the regular expression (17) does not represent FPS from $\boldsymbol{k} \ll \Sigma^* \gg$, since it contains inverse elements from Σ^{-1} and $\boldsymbol{K}$. The transition matrix for the corresponding $\boldsymbol{K} - (\Sigma \cup \Sigma^{-1})^*$ - automaton would contain elements from Σ^{-1} and $\boldsymbol{K}$ too – although the behavior of this automaton would be exactly FPS r that is without Σ^{-1} and $\boldsymbol{K}\backslash\boldsymbol{k}$.

Open problem 6 ("Fatou extensions")**.** *Let* $A \in \boldsymbol{Z}^{n\times n} < \Sigma \cup \Sigma^{-1} >$ *but all layers of FPS* $r = \sum_{i=0}^{\infty}(A^i)_{1,n}$ *are in* $\boldsymbol{N} < \Sigma^* >$. *Would* $r \in \boldsymbol{N}^{rat} \ll \Sigma^* \gg$ *be true? And* $\boldsymbol{Z}^{rat} \ll \Sigma^* \gg$*? And* $K^{alg} \ll \Sigma^* \gg$ *?* ■

Open problem 7 (Berstel etc. [**BBCPP**])**.** *Does a function* $n \to r_n$ *preserve a rationality? That is if* $\{a_n\}_{n\geq 0}$ *is a rational sequence of natural numbers, r is rational FPS then would* $\sum_{i=0}^{\infty} r_{a_i}$ *be rational?* ■

Open problem 8. *Based on given LRR of FPS* p, q *build LRR of :* $p^*, p+q, pq$,
$p \odot q, p \sqcup\!\sqcup q$. ■

Open problem 9. *Describe the set of all stencils of given rational FPS.* ■

Open problem 10. *Stencils in their turn are rational FPS. One can be built their LRR and so on. What can be said about the process ?* ■

Conclusion

Many researchers guessed about the existence of a linear dependency between the current value of FPS and a limited number of previous ones, but have failed to express it in a convenient universal form that would allow to obtain trivially results above. Thus for example,

Restivo, Reutenauer [**RR**]: *FPS* $s \in K \ll \Sigma^* \gg$ *is rational iff for any word x there is a common linear recurrence relation over K satisfied by all the sequences* $\{(s, ux^n v)\}_{n\geq 0}, u, v \in \Sigma^*$.

The below listed authors used for stencil the same ring as for represented FPS , what undercut readability and applications:

Salomaa, Soittola [**SS**]: *Assume* $r \in K^{rat} \ll \Sigma^* \gg$ *and N is rank of r. Show that if* $|w_0| = N$ *then there are words* $w_1, \ldots, w_N$ *and elements* $c_1, \ldots, c_N$ *of K such that* $|w_i| < N, i = \overline{1, N}$ *and for all words w:*

$(r, ww_0) = c_1(r, ww_1) + \cdots + c_N(r, ww_N)$

Berstel, Reutenauer [**BR**]: *For any rational series S of rank n there exist a prefix-closed set P of n elements, with an associated prefix set C, and coefficients* $\alpha_{c,p}(c \in C, p \in P)$ *such that, for all words w and all* $c \in C$*:*

$$(S, cw) = \sum_{p\in P} \alpha_{c,p}(S, pw).$$

or limited the domain of definition of the linear relation :

Eilenberg [**E**]: $f = \sum a_n z^n$ *is rational iff the following "recursion formula" holds for all t sufficiently large:*

$$a_{t+m} + c_1 a_{t+m} + \ldots + c_m a_t = 0.$$

On the other hand Cohn [**Coh**] did not lost universality and convenience but to achieve that he had to 'maim' previous layers:

A series $r \in K((X; \alpha))$ *is rational iff there exist integer* m, n_0 *and elements* $c_1, \ldots, c_m \in K$ *such that for all* $n > n_0$*:*

$$r_n = r_{n-1}^{\alpha} c_1 + r_{n-2}^{\alpha^2} c_2 + \ldots + r_{n-m}^{\alpha^m} c_m$$

As for Varricchio [**V**] – he did not go beyond the statement of a linear dependency for initial interval of FPS:

Let $s \in K^{rat} \ll \Sigma^* \gg, \Sigma^{[N]}$ *be the set of words of* Σ *whose length is less then or equal to N,* μ *be matrix interpretation of S. Then one can effectively compute an integer N, depending on S with the property that for any* $u \in \sum^{[N+1]}$ *there exist a set* $T = \{\sigma_v\}_{v\in\sum^{[N]}} \subseteq K$ *such that* $\mu(u) = \sum_{\sigma_v \in T} \mu(v)$.

It is very strange that author failed to discover the attempts to use linear recurrence in FPS on free commutative monoid $c(\Sigma^*), |\Sigma| \geq 2$. According to Kuich, Salomaa [**KS**] $K^{alg} \ll c(\Sigma^*) \gg = K^{rat} \ll c(\Sigma^*) \gg$. Therefore many K - algebraic FPS can be studied with the help of LRR.

As one see the proposed approach contrary to the predecessors is systematic and handy. As a indirect proof of that fact is a large number of correlations between FPS and combinatorics collected by the author and left out the scope of this work.

References

[BR] BerstelJ., Reutenauer C., *Rational Series and Their Languages*, Springer-Verlag(1988).

[E] Eilenberg S., *Automata, Languages, and Machines*, Vol.A, Acad. Press (1974).

[RS] Rabin M.O., Scott D., *Finite Automata and their decision problems*, IBM J.Res.Develop. 3 (1959), 114-125. 7

[CS] Culik K. II, Salomaa A.,*On the decidability of homomorphism equivalence for languages*, J.Comput.Syst.Sci, 17(1978) , 163-175.

[SS] Salomaa A., Soittola M., *Automata-Theoretic Aspects of Formal Power Series*, Springer-Verlag, 1978.

[KS] Kuich W., Salomaa A.*Semirings, Automata, Languages*, 1985.

[Coh] Cohn P.M.,*Free Rings and Their Relations*, 2nd edn. , Acad. Press (1985).

[L] Lewin J.*Fields of fractions for group algebras of free groups*, Trans.AMS , 192 (1974), 339-346.

[R] Reutenauer C., *On a question of S.Eilenberg*, Theor.Comp.Sci., 5(1977), 219.

[Pa] Passman D.S.,*The Algebraic Structure of Group Rings*, Wiley (1977).

[HK] Harju T., Karhumaki J., *The equivalence problem of multitape finite automata*, Theor.Comp.Sci., 78 (1991) 347-355.

[T] Tzeng W.-G.,*A polynomial-time algorithm the equivalence of probabolistic automata*, FOCS (1989).

[A1] Archangelsky K.V., *Efficient algorithm for checking multiplicity equivalence for the finite zet-sigma-automata*, Lect.NotesComp.Sci., 2450 (2002) ,283-289.

[HW] Hebish U., Weinert H.J., *Generalized semigroup semirings which are zero-divisor-free or multiplicatively left-cancellative*, Theor.Comp.Sci., 92 (1992), 269-289.

[K] Karhumaki J.,*Equations over finite sets of words and equivalence problems in automata theory*, Thepr.Comp.Sci., 106(1993), 103-118.

[Ber] Bergman G.M., *Commuting elements in free algebras and related topics in ring theory*, Thesis, Harvard University, 1967.

[A2] Archangelsky K.V., *The length of homogeneous subwords in rational formal power series*, Bull. Eur. Ass.Theor. Comp. Sci., 71(2000) 253.

[Cor] Cori R.,*Formules d'inversion de series formelles*, Ecole de Printemps d'Informatique Theretique "Series formelles en variables non commutatives et applications", Landes, (1977) 55-67.

[Ben1] Benzaghou B., *Sur les series rationnelles a une variable*, Ecole de Printemps d'Informatique Theretique "Series formelles en variables non commutatives et applications", Landes, (1977) 1-3.

[Ben2] Benzaghou B., *Algebres de Hadamard*, Bull. Soc. Math. France, 98 (1970), 209-252.

[BB] Benzaghou B., Bezivin J.-P., *Propertes algebriques de suites differetielement finies*, Bull.Soc.math.France, 120 (1992), 327-346.

[AB] Anselmo M., Bertoni A., *On 2pfa's and the Hadamard quotient of formal power series*, Bull. Belg.Math.Soc. 1 (1994), 165-173.

[Po] Poorten van der A.J., *Solution de la conjecture de Pisot sur le quotient de Hadamard de deux fractions rationnelles*, C.R.Acad.Sc. Paris 306, serie I (1988), 97-102.

[G] Gerardin K., *Quotient de Hadamard de series rationnelles en variables non commutatives*,Ecole de printemps d'Informatique Theretique "Series formelles en variables non commutatives et applications", Landes, (1977) ,119-134.

[RR] Restivo A.,Reutenauer C., *Rational languagrs and the Burnside problem*, Theor.Comp.Sci., 40(1985),13-30.

[BC] Brzozowski J.A.,Cohen R., *On decomposition of regular events*, J.ACM, 16(1969),132-144.

[BBCPP] Berstel J.,Boasson L.,Carton O.,Petazzoni B.,Pin J.-E.,*Operations preserving recognizable languages*, FCT'03.

[V] Varricchio S.,*Rational series with coefficients in commutative ring*, LITP 91.38,Institut Blaise Pascal, 1991.

SR SC "Algorithm-center", P.O.Box 129,Kiev, 04070, Ukraine
E-mail address: akv@zeos.net

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Sign balance for finite groups of Lie type (Extended Abstract)

Eli Bagno and Yona Cherniavsky

Abstract. *A product formula for the parity generating function of the number of $1's$ in invertible matrices over $\mathbb{Z}_2$ is given. The computation is based on algebraic tools such as the Bruhat decomposition. The same technique is used to obtain a parity generating function also for symplectic matrices over $\mathbb{Z}_2$. We present also a generating function for the sum of entries modulo p of matrices over $\mathbb{Z}_p$. This formula is a new appearance of the Mahonian distribution.*

1. Introduction

Let G be a subgroup of $GL_n(\mathbb{Z}_2)$. For every $K \in G$ define $o(K)$ to be the number of 1's in K. A natural problem is to find the number of matrices with a given number of 1's, or in other words, to compute the following generating function:

$$O(G,t) = \sum_{K \in G} t^{o(K)}.$$

It is not hard to see that in the case $G = GL_n(\mathbb{Z}_2)$, $O(G,t)$ has $n!$ as a factor but the complete generating function can be rather hard to compute. A weaker variation of this problem is to evaluate $O(G,-1)$. This is equivalent to determining the difference between the numbers of even and odd matrices, where a matrix is called *even* if it has an even number of 1's and *odd* otherwise. The number $O(G,-1)$ will be called the *parity difference* or the *imbalance* of G. A set S is called *sign-balanced* if $O(S,-1) = 0$.

The notion of sign-balance has recently reappeared in a number of contexts. Simion and Schmidt [**9**] proved that the number of 321-avoiding even permutations is equal to the number of such odd permutations if n is even, and exceeds it by the Catalan number $C_{\frac{1}{2}(n-1)}$ otherwise. Adin and Roichman [**1**] refined this result by taking into account the maximum descent. In a recent paper [**11**], Stanley established the importance of the sign-balance.

In this work we calculate the parity difference for $GL_n(\mathbb{Z}_2)$ as well as for the symplectic groups $Sp_{2n}(\mathbb{Z}_2)$.

We also generalize the problem of sign-balance to matrix groups over prime fields other than $\mathbb{Z}_2$. It turns out that the appropriate parameter for these fields is the sum of non zero entries of the matrix (mod p) rather than just the number of nonzero elements. A generalization of these results to groups over arbitrary finite fields has also been done. It will be published in a future publication.

Another aspect of this work is the occurrence of the Mahonian distribution in our results. Recall that a permutation statistic over S_n is called Mahonian if it has the same distribution as the number of inversions. MacMahon proved that major index has such distribution, explicitly:

$$\sum_{\pi \in S_n} q^{inv(\pi)} = \sum_{\pi \in S_n} q^{maj(\pi)} = [n]_q!$$

1991 *Mathematics Subject Classification.* Primary 54C40, 14E20; Secondary 46E25, 20C20.

where $[n]_q = \frac{1-q^n}{1-q}$.

Our results can also be seen as an example of the *cyclic sieving phenomenon* (See [**8**] for details). They also hint at the existence of permutation statistics theory for finite groups of Lie type. For a pioneering results in this direction see [**5**].

We finally note that another approach to the case of type A was proposed to us by Alex Samorodnitzky [**10**], and will be explained elsewhere.

2. Preliminaries

2.1. The groups of Lie Type A. Let $\mathbb{F}$ be any field and let $G = GL_n(\mathbb{F})$ be the group of invertible $n \times n$ matrices over $\mathbb{F}$. Let H be the subgroup of G consisting of the diagonal matrices. This is a choice of a torus in G. It is easy to show that the normalizer of H, $N(H)$, is the group of monomial matrices (where each row and column contains exactly one non-zero element). The quotient $N(H)/H$ is called the *Weyl group of type A*, and is isomorphic to S_n, the group of permutations on n letters. The *Borel subgroup* $\mathbb{B}^+$ of the group G consists of the upper triangular matrices in G. The *opposite Borel subgroup*, consisting of the lower triangular matrices, is denoted by $\mathbb{B}^-$. We denote by $\mathbb{U}^+$ and $\mathbb{U}^-$ the groups of upper and lower triangular matrices (respectively,) with 1's along the diagonal.

We finish this section with the following:

Proposition 2.1. *(See for example* [**6**, p.20]*) For every finite field $\mathbb{F}$ with q elements the order of $GL_n(\mathbb{F})$ is*

$$q^{\binom{n}{2}}(q-1)^n[n]_q!$$

2.2. Lie Type C. Let J denote the $n \times n$ matrix

$$\begin{pmatrix} 0 & \cdot & \cdot & 1 \\ 0 & \cdot & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & \cdot & 0 \end{pmatrix}$$

and let

$$M = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}.$$

The Lie group of type C, or the symplectic group, is defined over the field $\mathbb{F}$ by:

$$Sp_{2n}(\mathbb{F}) = \{AT \in SL_{2n}(\mathbb{F}) \mid A^T M A = M\}.$$

This is the set of fixed points of the automorphism $\varphi : SL_{2n}(\mathbb{F}) \longrightarrow SL_{2n}(\mathbb{F})$ given by: $\varphi(A) = M^{-1}(A^T)^{-1}M$.

An alternative way to present the symplectic group is the following: We define first a bilinear form on $\mathbb{F}^{2n}$:

Definition 2.1. For every $x = (x_1, ..., x_{2n}), y = (y_1, ..., y_{2n}) \in \mathbb{F}^{2n}$

$$B(x, y) = \sum_{i=1}^{n} x_i \cdot y_{2n+1-i} - \sum_{i=n+1}^{2n} x_i \cdot y_{2n+1-i}.$$

Denoting by $\{x_1, ..., x_{2n}\}$ the set of columns of X it is easy to see that $X \in Sp_{2n}(\mathbb{F})$ if and only if the columns satisfy the following set of equations:

$$B(v_i, v_j) = \begin{cases} (-1)^{i-j} & i+j = 2n+1 \\ 0 & i+j \neq 2n+1 \end{cases}$$

We end this section with the following well known fact:

Proposition 2.2. *(See for example* [**6**, p.35]*)*

For every finite field $\mathbb{F}$ *with* q *elements the order of* $Sp_{2n}(\mathbb{F})$ *is:*

$$q^{n^2}(q-1)^n[2]_q\cdots[2n]_q$$

2.3. The Bruhat Decomposition for type A. The Bruhat decomposition is a way to write every invertible matrix as a product of two triangular matrices and a permutation matrix. We start with the following definitions:

Recall from Section 2.1 the definition of the Borel subgroup $\mathbb{B}^+$ and the unipotent subgroup $\mathbb{U}^-$. For every permutation $\pi \in S_n$ we identify π with the matrix:

$$[\pi]_{i,j} = \begin{cases} 1 & i = \pi(j) \\ 0 & \text{otherwise} \end{cases}$$

Define for every $\pi \in S_n$:

$$\mathbb{U}_\pi = \mathbb{U}^- \cap (\pi \mathbb{U}^- \pi^{-1}).$$

$\mathbb{U}_\pi$ consists of the matrices with 1-s along the diagonal and zeros in place (i,j) whenever $i < j$ or $\pi^{-1}(i) < \pi^{-1}(j)$. This is an affine space of dimension $\binom{n}{2} - \ell(\pi)$ over $\mathbb{F}$. ($\ell(\pi)$ is the length of π with respect to the Coxeter generators).

Now, given $g \in G$, we can column reduce g by multiplying on the right by Borel matrices in order to get an element gb^{-1} satisfying the following condition:

$$(*) \qquad \begin{array}{l}\text{The right most nonzero entry in each row is 1} \\ \text{and it is the first nonzero entry in its column.}\end{array}$$

Those "leading entries" form a permutation matrix corresponding to $\pi \in S_n$.

Now we can use π^{-1} to rearrange the columns of gb^{-1} in order to get $gb^{-1}\pi^{-1} = u \in \mathbb{U}_\pi$, i.e., $g = u\pi b$. This is called the *Bruhat decomposition* of the matrix g. One can prove that this decomposition is unique, and thus we have a partition of G into double cosets indexed by the elements of the Weyl group S_n.

If $\pi \in S_n$ then the double coset indexed by π decomposes into left $\mathbb{B}^+$-cosets in the following way: For every choice of $u \in \mathbb{U}_\pi$, $u\pi$ is a representative of the left coset $u\pi\mathbb{B}^+$. Thus a general representative of the double coset $\mathbb{U}_\pi$ can be taken as matrix of the form $(*)$, with every column filled with free parameters beyond the leading 1.

We summarize the information we gathered about the Bruhat decomposition for type A in the following:

Proposition 2.3. *The group* $GL_n(\mathbb{F})$ *is a disjoint union of double cosets of the form* $\mathbb{U}_\pi \pi \mathbb{B}^+$*, where* π *runs through* S_n*. Every double coset decomposes into cosets of the form* $A\mathbb{B}^+$ *where* A *is a general representative of the form* $(*)$*. The number of free parameters in* A *is equal to* $\binom{n}{2} - \ell(\pi)$.

Here is an example of the coset decomposition for $GL_3(\mathbb{Z}_2)$:

$$\mathbb{U}_1 1 \mathbb{B}^+ = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ \beta & \gamma & 1 \end{pmatrix} \mathbb{B}^+ \mid \alpha, \beta, \gamma \in \mathbb{Z}_2 \right\}$$

$$\mathbb{U}_{s_1} s_1 \mathbb{B}^+ = \left\{ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ \alpha & \beta & 1 \end{pmatrix} \mathbb{B}^+ \mid \alpha, \beta \in \mathbb{Z}_2 \right\}$$

$$\mathbb{U}_{s_2} s_2 \mathbb{B}^+ = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 0 & 1 \\ \beta & 1 & 0 \end{pmatrix} \mathbb{B}^+ \mid \alpha, \beta \in \mathbb{Z}_2 \right\}$$

$$\mathbb{U}_{s_2s_1}s_2s_1\mathbb{B}^+ = \left\{ \begin{pmatrix} 0 & 1 & 0 \\ 0 & \alpha & 1 \\ 1 & 0 & 0 \end{pmatrix} \mathbb{B}^+ \mid \alpha \in \mathbb{Z}_2 \right\}$$

$$\mathbb{U}_{s_1s_2}s_1s_2\mathbb{B}^+ = \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ \alpha & 1 & 0 \end{pmatrix} \mathbb{B}^+ \mid \alpha \in \mathbb{Z}_2 \right\}$$

$$\mathbb{U}_{s_1s_2s_1}s_1s_2s_1\mathbb{B}^+ = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \mathbb{B}^+$$

2.4. Bruhat Decomposition for Type C. In order to be able to present the Bruhat decomposition for type C, we must first define a Borel subgroup for $Sp_{2n}(\mathbb{F})$. We present this subject following [**11**]. Note that although the exposition of [**11**] deals with groups over algebraically closed fields, the results hold also over finite fields. Start with the Borel subgroup $\mathbb{B}^+$, chosen for type A, consisting of the upper triangular matrices.

If $X = \left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right) \in \mathbb{B}^+$, then $\varphi(X) = \begin{pmatrix} J(C^T)^{-1}J & J(C^T)^{-1}B^T(A^T)^{-1}J \\ 0 & J(A^T)^{-1}J \end{pmatrix} \in \mathbb{B}^+$. (The automorphism φ was defined in Section 2.2). Moreover, the automorphism φ keeps the Borel subgroup $\mathbb{B}^+$, as well as the groups of diagonal and monomial matrices (denoted by H and $N(H)$ respectively in Section 2.1) invariant. Thus we can take $\mathbb{B}^+_C = Sp_{2n}(\mathbb{F}) \cap \mathbb{B}^+$ and $\mathbb{B}^-_C = Sp_{2n}(\mathbb{F}) \cap \mathbb{B}^-$ as the Borel subgroup and the opposite Borel subgroup of $Sp_{2n}(\mathbb{F})$ respectively, and similarly for H and $N(H)$.

The Weyl group of type C can be realized as the group of those permutations $\pi \in S_{2n}$ such that $\pi(2n+1-i) = 2n+1-\pi(i)$. This group is isomorphic to the hyperoctahedral group B_n (See definition in Section **??**). The isomorphism can be seen by labelling the basis elements of the space on which $Sp_{2n}(\mathbb{F})$ acts by indices $n, n-1, ..., 1, ..., -n$.

We define also the groups $\mathbb{U}^+_C = \mathbb{U}^+ \cap Sp_{2n}(\mathbb{F})$ and $\mathbb{U}^-_C = \mathbb{U}^- \cap Sp_{2n}(\mathbb{F})$ to be the upper and lower unipotent subgroups respectively. For every $\pi \in B_n$ we define $\mathbb{U}^C_\pi = \mathbb{U}^-_C \cap (\pi \mathbb{U}^-_C \pi^{-1})$. $\mathbb{U}^C_\pi$ is the intersection of $Sp_{2n}(\mathbb{F})$ with the set of matrices with 1's along the diagonal and zeros at entries in location (i,j) whenever $i < j$ or $\pi^{-1}(i) < \pi^{-1}(j)$. This is an affine space of dimension $n^2 - \ell(\pi)$. (Here, $\ell(\pi)$ is the length function of B_n).

Now, we can use the Bruhat decomposition of $GL_{2n}(\mathbb{F})$ to produce the Bruhat decomposition for $Sp_{2n}(\mathbb{F})$. Let $g \in Sp_{2n}(\mathbb{F})$. Consider g as an element of $GL_{2n}(\mathbb{F})$ and write $g = u\pi b$ where $\pi \in S_{2n}$, $u \in \mathbb{U}_\pi$ and $b \in \mathbb{B}^+$. We have:

$$g = \varphi(g) = \varphi(u)\varphi(\pi)\varphi(b),$$

but from the uniqueness of the decomposition in $GL_{2n}(\mathbb{F})$ we have:

$$\varphi(u) = u, \quad \varphi(\pi) = \pi h^{-1}, \quad \varphi(b) = hb$$

where h is diagonal and thus $\pi \in B_n$ and $b \in \mathbb{B}^+_C$. This gives us the Bruhat decomposition. The description of the double cosets and the coset representatives is similar to the one given for type A, with the exception that here we have to intersect with $Sp_{2n}(\mathbb{F})$.

We summarize the information we gathered about the Bruhat decomposition for type C in the following:

Proposition 2.4. *The group $Sp_{2n}(\mathbb{F})$ decomposes into double cosets of the form $\mathbb{U}^C_\pi \pi \mathbb{B}^+_C$, where π runs through B_n. Every double coset decomposes into cosets of the form $A\mathbb{B}^+_C$ where A is a general representative of the form $(*)$. The number of free parameters in A is equal to $n^2 - \ell(\pi)$.*

3. Sign Balance for Type A

3.1. Sign Balance over $\mathbb{Z}_2$. In this section we present the results concerning the imbalance of the groups of type A over the field $\mathbb{Z}_2$. The proofs are written in **??**.

Theorem 3.1.

$$\sum_{K \in GL_n(\mathbb{Z}_2)} (-1)^{o(K)} = -2^{\binom{n}{2}}[n-1]_2!$$

where $[k]_q = \frac{1-q^k}{1-q}$.

The following corollary is immediate:

Corollary 3.1. *The number of even matrices in* $GL_n(\mathbb{Z}_2)$ *is exactly*

$$[n-1]_2! 2^{\binom{n}{2}}(2^{n-1}-1)$$

while the number of odd matrices in $GL_n(\mathbb{Z}_2)$ *is:*

$$[n-1]_2! 2^{\binom{n}{2}+n-1}$$

□

3.2. Sign Balance for Prime Fields. In this section we present the results concerning the imbalance of the groups of type A over the field $\mathbb{Z}_p$. The proofs are written in **??**.

Let p be a prime number and denote by $\mathbb{Z}_p$ the field with p elements. The results of Section 3.1 can be extended to invertible matrices over the field $\mathbb{Z}_p$, provided we substitute a primitive complex p-th root of unity in the generating function of the sum of elements of a matrix mod p. Explicitly, we use the information we gathered in the previous section to get the following:

Theorem 3.2.

$$\sum_{K \in GL_n(\mathbb{Z}_p)} \omega_p^{s(K)} = -(p-1)^{n-1} p^{\binom{n}{2}}[n-1]_p!.$$

where $s(K)$ *is the sum (mod p) of the elements of the matrix* K*, and* ω_p *is a primitive complex p-th root of unity.*

The following corollary is immediate:

Corollary 3.2. *The number of matrices in* $GL_n(\mathbb{Z}_p)$ *whose sum of entries modulo p is* 0 *is exactly*

$$[n-1]_p!(p-1)^{n-1} p^{\binom{n}{2}}(p^{n-1}-1)$$

while for every $1 \leq i \leq p-1$*, the number of matrices in* $GL_n(\mathbb{Z}_p)$ *whose entries add up to* i *modulo* p *is:*

$$[n-1]_p!(p-1)^{n-1} p^{\binom{n}{2}+n-1}.$$

□

4. Sign Balance for Type C

In this section we prove the following result:

Theorem 4.1.

$$\sum_{K \in Sp_{2n}(\mathbb{Z}_2)} (-1)^{o(K)} = -2^{n^2} \cdot [2]_2[4]_2 \cdots [2n-2]_2$$

The following corollary is immediate:

Corollary 4.1. *The number of even matrices in* $Sp_{2n}(\mathbb{Z}_2)$ *is exactly*

$$2^{n^2-1}[2]_2 \cdots [2n-2]_2([2n]_2 - 1)$$

while the number of odd matrices is

$$2^{n^2-1}[2]_2 \cdots [2n-2]_2([2n]_2 + 1). \qquad \square$$

In order to prove the theorem, we take the following direction: Instead of summing over the whole group of matrices, we sum over every coset separately. It turns out that some of the cosets are sign-balanced, while the others have only odd matrices. We start with the following definition:

Definition 4.2. A coset consisting entirely of odd matrices is called an *odd coset.*

The following lemma identifies the sign-balanced cosets.

Lemma 4.2. *Let A be a general representative of the double coset $U_\pi^C \pi \mathbb{B}_C^+$ corresponding to $\pi \in B_n$. Make some substitution in the free parameters of A to get a coset representative, and call it $\tilde{A}$. If $\tilde{A}$ has an odd column which is not the last one, then the coset $[\tilde{A}] = \{\tilde{A}B | B \in \mathbb{B}_C^+\}$ is sign-balanced, i.e.,*

$$\sum_{K \in \tilde{A}\mathbb{B}_C^+} (-1)^{o(K)} = 0.$$

Proof. An element of $\mathbb{B}_C^+$ is an invertible upper triangular matrix which is also symplectic. The condition of being symplectic is expressed by imposing a set of equations on the columns of the matrix. If we take b to be an upper triangular matrix with a set of columns $\{v_1, ..., v_{2n}\}$ then, as was stated in Section 2.2, forcing it to be symplectic is equivalent to imposing the equations (note that we are working over $\mathbb{Z}_2$):

$$B(v_i, v_j) = \begin{cases} 1 & i+j = 2n+1 \\ 0 & i+j \neq 2n+1 \end{cases}$$

As is easy to check, the equations of the form $B(v_i, v_i) = 0$ are trivial over $\mathbb{Z}_2$. The equations of the form $B(v_i, v_{2n+1-i}) = 1$ are also trivial. (Indeed, $B(v_i, v_{2n+1-i}) = \sum_{k=1}^{2n} b_{k,i} \cdot b_{2n+1-k,2n+1-i}$ but since b is upper triangular, over $\mathbb{Z}_2$ we have $b_{ii} \cdot b_{2n+1-i,2n+1-i} = 1$ and the other summands vanish since for $k > i$ one has $b_{ki} = 0$ and for $k > 2n+1-i$ one has $b_{2n+1-k,2n+1-i} = 0$).

Now, the only non trivial equations involving the parameters appearing in the last column are the ones of the form:

$$B(v_i, v_{2n}) = 0, (2 \leq i \leq 2n-1)$$

and each such equation can be written in such a way that the parameters of the last column are free while the parameters of the first row depend on them. Explicitly, we write the equation $B(v_i, v_{2n}) = 0$ as

$$b_{1i} = \sum_{k=2}^{2n} b_{ki} \cdot b_{2n+1-k,2n}.$$

Note that the elements of the last column of the matrix b have no appearance as a part of a linear combination in any place other than the first row. This is justified by the fact that every non trivial equation, involving the first row, which we have not treated yet must be of the form $B(v_i, v_j) = 0$ for $1 \leq i < j \leq 2n-1$. Thanks to the upper triangularity of b, the elements laying in the first row vanish in these equations.

Let us look at the following example:

$$b = \begin{pmatrix} 1 & b_{12} & b_{13} & b_{14} \\ 0 & 1 & b_{23} & b_{24} \\ 0 & 0 & 1 & b_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The only non trivial equations involving the last column are: $B(v_2, v_4) = 0$ and $B(v_3, v_4) = 0$.

These equations can be written as:

$$b_{12} = b_{34}$$
$$b_{13} = b_{24} + b_{23} \cdot b_{34}$$

so after intersecting with $Sp_{2n}(\mathbb{Z}_2)$, the matrix b looks like:

$$b = \begin{pmatrix} 1 & b_{34} & b_{24} + b_{23} \cdot b_{34} & b_{14} \\ 0 & 1 & b_{23} & b_{24} \\ 0 & 0 & 1 & b_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The elements of the last column appear only in the first row and in the equations of the form $B(v_i, v_j) = 0$ the elements located in the first row vanish.

Note that in this case all of the parameters outside the first row are free. This doesn't hold in general. Nevertheless, as we have proven, we can arrange the parameters such that the elements of the last column reappear only in the first row.

Returning now to the proof, we have two cases:

- The first column of $\tilde{A}$ is odd. In this case we can use the element located in place $(1, 2n)$ to construct a bijection between odd and even matrices inside the coset $\tilde{A}\mathbb{B}_C^+$. This is done in the same way described earlier for type A: Divide $\mathbb{B}_C^+$ into two disjoint subsets:

  $$\mathbb{B}_{C0}^+ = \{T = (t_{i,j}) \in \mathbb{B}_C^+ \mid t_{1,2n} = 0\}$$

  $$\mathbb{B}_{C1}^+ = \{T = (t_{i,j}) \in \mathbb{B}_C^+ \mid t_{1,2n} = 1\}.$$

  For every matrix $X \in \tilde{A}\mathbb{B}_C^+$, the k-th column of X is a linear combination of the first k columns of $\tilde{A}$. Now, due to the fact that the parameter appearing in the location $(1, 2n)$ has no other appearance, for every $B \in \mathbb{B}_{C0}^+$ there is some $B' \in \mathbb{B}_{C1}^+$ such that B and B' differ only in the entry numbered $(1, 2n)$.

  Note that $\tilde{A}B$ and $\tilde{A}B'$ are obtained from $\tilde{A}$ by the same sequence of column operations except for the first column which was used in producing AB but was not used in producing $\tilde{A}B'$. Hence $\tilde{A}B$ and $\tilde{A}B'$ have opposite parity. This gives us a bijection between the odd and the even matrices of the coset $A\mathbb{B}_C^+$.
- The first column of $\tilde{A}$ is even. Denoting by j the number of the first odd column of $\tilde{A}$, we use the element located in place $(j, 2n)$ to construct a bijection between the odd and even matrices inside $\tilde{A}\mathbb{B}_C^+$ in the same way as in the previous case. Note that since the element located in place $(j, 2n)$ in the matrices of the Borel subgroup can reappear only in the first row, it affects only the first column of $\tilde{A}$, which is even.

□

We turn now to treat the odd cosets.

Lemma 4.3. *Let $\pi \in B_n$. Let A be a general representative of the double coset $U_\pi \pi \mathbb{B}_C^+$ corresponding to $\pi \in B_n$. Make some substitution in the free parameters of A to get a coset representative, and call it $\tilde{A}$. If all of the first $2n-1$ columns of $\tilde{A}$ are even then all of the matrices belonging to the coset $\tilde{A}\mathbb{B}_C^+$ are odd. The imbalance calculated inside this coset is:*

$$\sum_{K \in \tilde{A}\mathbb{B}_C^+} (-1)^{o(K)} = -|\mathbb{B}_C^+| = -2^{n^2}.$$

Proof. The last column of $\tilde{A}$ is always odd and thus since all other columns of $\tilde{A}$ are even, $\tilde{A}$ itself is an odd matrix and the same holds for $\tilde{A}B$ for every $B \in \mathbb{B}_C^+$. The size of the coset $\tilde{A}\mathbb{B}^+$ is 2^{n^2}, and the result follows. □

Lemma 4.4. *Let $\pi \in B_n$. The double coset $U_\pi^C \pi \mathbb{B}_C^+$ contains odd cosets if and only if $\pi(2n) = 2n$.*

Proof. Let A be a general representative of the double coset $U_\pi^C \pi \mathbb{B}_C^+$. Write $U = A\pi^{-1}$. Then $U \in \mathbb{U}_\pi^C$ is a lower triangular matrix and since $\pi(2n) = 2n$ (which implies also $\pi(1) = 1$), the first column as well as the last row of U contain $2n-1$ parameters. Note that $U^T \in \mathbb{B}_C^+$ and thus by the considerations described in Lemma 4.2, the parameters appearing in the last column of U^T can reappear only in the first row of U^T. We conclude that the parameters of the last row of U can reappear only in the first column of U. Now, for every column numbered $2 \leq k \leq 2n-1$ in U and for every choice of the first elements of the column numbered k, we are free to choose the parameter located in the bottom of this column, $(2n, k)$, in such a way that the column will be even. The parameter located in the place $(2n, 1)$ has no other appearance and thus we can choose all of the first $2n-1$ columns of U to be even. Getting back to the general representative A, since $\pi(2n) = 2n$, we have also $\pi(1) = 1$ and thus A and U differ only in the columns $1 < k < 2n$ so that the proof works also for A.

On the other hand, if $\pi(2n) \neq 2n$ then π contains a column numbered $k < 2n$ which has only one nonzero element, located in place $(2n, k)$. By the construction of the general representative A, there are only zeros above the 1 coming from the permutation and thus this odd column appears also in A. By the previous lemma, the coset $\{\tilde{A}B | B \in \mathbb{B}_C^+\}$ is sign-balanced. □

Now, we have to count the imbalance on the odd cosets. By Lemma 4.4 we are interested only in the double cosets corresponding to the permutations $\pi \in B_{n-1}$. The following lemma shows how to count.

Lemma 4.5. *Let $\pi \in B_n$ such that $\pi(n) = n$. The double coset $\mathbb{U}_\pi^C \pi \mathbb{B}_C^+$ contains exactly $2^{(n-1)^2 - \ell(\pi)}$ odd cosets.*

Proof. Let A be representative of the double coset $\mathbb{U}_\pi^C \pi \mathbb{B}_C^+$. As was shown in the previous lemma, the parity of a each one of the first $2n-1$ columns of A is determined by the free parameter in its bottom. Since there are a total of $n^2 - \ell(\pi)$ free parameters and exactly $2n-1$ 'bottom parameters', the number of substitutions of parameters giving all of the $2n-1$ first columns even is $2^{n^2 - \ell(\pi) - (2n-1)}$. This is also the number of odd cosets in the double coset $\mathbb{U}_\pi^C$. □

□

We turn now to the proof of Theorem 4.1. In order to calculate the imbalance we have to count only odd cosets. By Lemma 4.4, we are interested only in the double cosets corresponding to permutations $\pi \in B_{n-1}$. By Lemma 4.5, every such double coset contains $2^{(n-1)^2 - \ell(\pi)}$ odd cosets. By Lemma 4.3, each odd coset contributes -2^{n^2} to the imbalance, and we have in total:

$$\begin{aligned}
\sum_{K \in Sp_{2n}(\mathbb{Z}_2)} (-1)^{o(K)} &= \sum_{\substack{\pi \in B_n \\ \pi(n)=n}} -2^{n^2} \cdot 2^{(n-1)^2 - \ell(\pi)} \\
&= -2^{n^2} \sum_{\pi \in B_{n-1}} 2^{(n-1)^2 - \ell(\pi)} \\
&= -2^{n^2} \sum_{\pi \in B_{n-1}} 2^{\ell(\pi)} \\
&= -2^{n^2} [n-1]_2! \\
&= -2^{n^2} \cdot [2]_2 [4]_2 \cdots [2n-2]_2
\end{aligned}$$

References

[1] R. M. Adin and Y. Roichman, *Equidistribution and sign-balance on 321 - avoiding permutations*, preprint, 2003, math.CO/0304429.
[2] E. Bagno and Y. Cherniavsky, *Sign balance for finite groups of Lie type*,preprint, 2003.
[3] H. Hiller, *Geometry of Coxeter groups*, Research Notes in Mathematics, no. 54, London, (1982).

[4] J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Mathematics, no. 29, Cambridge Univ. Press, Cambridge, (1990).
[5] J. Kung, *The cycle structure of a linear transformation over a finite field*, Linear Algebra Appl. 36, (1981), 141–155.
[6] O. T. O'Meara, *Symplectic groups*, Mathematical survies, 16, American Mathematical Society, (1978).
[7] A. Ram, T. Halversion, R. Leduc, *Iwahori-Hecke algebras of type A, bitraces and symmetric functions*, Int. Math. Research Notices (1997) No. 9, 401–416.
[8] V. Reiner, D. Stanton, and D. White, *The cyclic sieving phenomen*, preprint, (2003).
[9] R. Steinberg, *Conjugacy Classes in Algebraic Groups*, Lecture Notes in Mathematics, no. 366, Springer-Verlag, (1974).
[10] A. Samorodnitzky, *Personal comunication.*
[11] R. P. Stanley, *Some remarks on sign-balanced and maj-balanced posets.* preprint, (2002), math. CO. /0211113.
[12] L. Solomon, *The orders of the finite Chevalley groups*, J. Algebra, 3, (1966), 376–393.

Department of Mathematics, Hebrew University, Givaat Ram, Jerusalem
E-mail address: bagnoe@math.huji.ac.il

Department of Mathematics and Statistics, Bar-Ilan University, Ramat-Gan, Israel, 52900
E-mail address: cherniy@math.biu.ac.il

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Generating Functions For Kernels of Digraphs (Enumeration & Asymptotics for Nim Games)

Cyril Banderier, Vlady Ravelomanana, and Jean-Marie Le Bars

Abstract. *In this article, we study directed graphs (digraphs) with a coloring constraint due to Von Neumann and related to Nim-type games. This is equivalent to the notion of kernels of digraphs, which appears in numerous fields of research such as game theory, complexity theory, artificial intelligence (default logic, argumentation in multi-agent systems), 0-1 laws in monadic second order logic, combinatorics (perfect graphs)... Kernels of digraphs lead to numerous difficult questions (in the sense of NP-completeness, #P-completeness). However, we show here that it is possible to use a generating function approach to get new informations: we use technique of symbolic and analytic combinatorics (generating functions and their singularities) in order to get exact and asymptotic results, e.g. for the existence of a kernel in a circuit or in a unicircuit digraph. This is a first step toward a generatingfunctionology treatment of kernels, while using, e.g., an approach "à la Wright". Our method could be applied to more general "local coloring constraints" in decomposable combinatorial structures.*

Résumé. *Nous étudions dans cet article les graphes dirigés (digraphes) avec une contrainte de coloriage introduite par Von Neumann et reliée aux jeux de type Nim. Elle équivaut à la notion de noyaux de digraphes, qui apparaît dans de nombreux domaines, tels la théorie des jeux, la théorie de la complexité, l'intelligence artificielle (logique des défauts, argumentation dans les systèmes multi-agents), les lois 0-1 en logique monadique du second ordre, la combinatoire (graphes parfaits)... Les noyaux des digraphes posent de nombreuses questions difficiles (au sens de la NP-complétude ou de la #P-complétude). Cependant, nous montrons ici qu'il est possible de recourir aux séries génératrices afin d'obtenir de nouvelles informations : nous utilisons les techniques de la combinatoire symbolique et analytique (étude des singularités d'une série) afin d'obtenir des résultats exacts ou asymptotiques, par exemple pour l'existence d'un noyau dans un digraphe unicircuit. Il s'agit là de la première étape vers une série génératrilogie des noyaux. Notre méthode peut être appliquée plus généralement à des "contraintes locales" de coloriage dans des structures combinatoires décomposables.*

1. Introduction

Let V and E be the set of vertices and directed edges (also called *arcs*) of a directed graph D without loops or multiarcs (we call such graphs *digraphs* hereafter). A kernel of D is a nonempty subset K of V, such that for any $a, b \in K$, the edge (a, b) does not belong to E, and for any vertex outside the kernel ($a \notin K$), there is a vertex in the kernel ($b \in K$), such that the edge (a, b) belongs to E. In other words,

Key words and phrases. generating functions, analytic combinatorics, kernels of graphs, Nim games, digraphs.
Corresponding author: Cyril.Banderier@lipn.univ-paris13.fr.

K is a nonempty independent and dominating set of vertices in D [**2**]. Not every digraph has a kernel and if a digraph has a kernel, this kernel is not necessarily unique. The notion of kernel allows elegant interpretations in various contexts, since it is related to other well-known concepts from graph theory and complexity theory. In game theory the existence of a kernel corresponds to a winning strategy in two players for famous Nim-type games (cf. [**3, 16, 17, 31**]).

Imagine that two players $\mathcal{A}$ and $\mathcal{B}$ play the following game on D in which they move a token each in turn: $\mathcal{A}$ starts the game by choosing an initial vertex $v_0 \in V$ and then makes a move to a vertex v_1. A move consists in taking the token from the present position v_i and placing it on a child of v_i, *i.e.* a vertex v_{i+1} such that $(v_i, v_{i+1}) \in E$. $\mathcal{B}$ makes a move from v_1 to v_2 and gives the hand to $\mathcal{A}$, which has now to play from v_2, and so on. The first player unable to move loses the game. One of the two players has a winning strategy (as this game is finite in a digraph D without circuit, for circuits one extends the rules by saying that the game is lost for the player who replays a position previously reached). Von Neumann and Morgenstern [**31**] proved that any directed acyclic graph has a unique kernel, which is the set of winning positions for $\mathcal{A}$ ($\mathcal{A}$ always forces $\mathcal{B}$ to play outside the kernel, until $\mathcal{B}$ cannot play anymore). Richardson [**27**] proved later that every digraph without odd circuit has a kernel [**7, 29**]. Berge wrote a chapter on kernels in [**2**]. Furthermore, there is a strong connection between perfect graphs and kernels (see the Berge and Duchet survey [**1**]). Some natural variants of this property are studied in various logic for Intelligence Artificial, some of them are definable in default logic [**8**] and used for argumentation in multi-agents systems, kernels appear there as sets of coherent arguments [**6, 12**].

Fernandez de la Vega [**13**] and Tomescu [**30**] proved independently that dense random digraphs with n vertices and $m = \Theta(n^2)$ edges, have asymptotically almost surely a kernel. In addition, they get the few possible sizes of a kernel and a precise estimation of the numbers of kernels.

Few years ago a new interest for these studies arises by their applications in finite model theory. Indeed variants of kernel are the best properties to provide counterexamples of 0-1 laws in fragments of monadic second-order logic [**21, 22**]. Goranko and Kapron showed in [**19**] that such a variant is expressible in modal logic over almost all finite frames for frame satisfiability; recently Le Bars proved in [**23**] that the 0-1 law fails for this logic.

The existence of a kernel in a digraph has been shown NP-complete, even if one restricts this question to planar graphs with in- and out-degree ≤ 2 and degree ≤ 3 [**9, 11, 15**]. It is somehow related to finding a maximum clique in graphs [**4, 21**], which is known to be difficult for random dense graphs.

In this article, we use some generating function techniques to give some new results on Nim-type games played on directed graphs (or, equivalently, some new informations on kernel of digraphs). More precisely, we deal with a family of planar digraphs with at most one circuit or one cycle and we give enumerative (Theorems 4.1, 4.2, 4.3, 4.4 in Section 4) and asymptotics results (Theorems 5.1, 5.2, 5.3, 5.4 in Section 5) on the size of the kernel, the probability of winning on trees for the first player...

2. Definitions

We give below more precise definitions, readers familiar with digraphs can skip them.

Let $D = (V, E)$ be a digraph. For each $v \in V$, let $v^+ = \{w \in V/(v,w) \in E\}$ and $v^- = \{w \in V/(w,v) \in E\}$, $|v^+|$ is the *out degree* of v and $|v^-|$ is the *in degree* of v.

A vertex with an in degree of 0 is called a *source* (since one can only leave it) and a vertex with an out degree of 0 is called a *sink* (since one cannot leave it). Let $U \subset V$, $U^+ = \cup_{v \in U} v^+$ and $U^- = \cup_{v \in U} v^-$, we denote by $D(U)$ the subgraph induced by the vertices of U.

There is a *path* from vertex v to w means that there exists a sequence $(v_1, \ldots, v_k)$ such that $v_1 = v$, $v_k = w$ and $v_i \in v_{i+1}^+ \cup v_{i+1}^-$, for $i = 1 \ldots k-1$. There is a *directed path* from vertex v to w means that there exists a sequence $(v_1, \ldots, v_k)$ such that $v_1 = v$, $v_k = w$ and $v_i \in v_{i+1}^+$, for $i = 1 \ldots k-1$.

A *cycle* is a path $(v_1, \ldots, v_k)$ such that $v_1 = v_k$. A *circuit* is a directed path $(v_1, \ldots, v_k)$ such that $v_1 = v_k$.

If D contains a directed path from vertex v to w then v is an *ancestor* of w and w is a *descendant* of v. If this directed path is of length 1, then the ancestor v of w is also called a *parent* of w, and v a *child* of w.

D is strongly connected if for each pair of vertices, each one is an ancestor of the other. $D(U)$ is a strongly connected component of D if it is a maximal strongly connected subgraph of D.

U is an *independent set* when $U \cap U^+ = \emptyset$ and a *dominating set* when $v^+ \cap U \neq \emptyset$ for any $v \in V \setminus U$. U is a kernel if it is an independent dominating set.

D is a DAG if it is a directed digraph without circuit (the terminology "directed *acyclic* graph" being popular, we use the acronym *DAG* although it should stands for "directed *acircuit* graph", according to the above definitions of cycles and circuits).

3. How to find the kernel of a digraph

Consider digraphs satisfying the following rules:

- each vertex is colored either in red or in green,
- each green vertex has at least a red child,
- no red vertex has a red child.

It is immediate to see that a digraph satisfying such coloring constraints possesses a kernel, which is exactly the set of its red vertices. It is now easy to see, e.g., that the circuit of length 3 has no kernel, that the circuit of length 4 has 2 kernels, that any DAG has exactly one kernel. For this last point, assume that D is a DAG (directed acircuit graph). Algorithm 1 (below) returns its unique kernel. It begins to color the sinks in red and then goes up toward sources, as it is deterministic and as it colors at least a new vertex at each iteration, this proves that each DAG has a single kernel. Such an algorithm was already considered by Zermelo while studying chessgame.

Algorithm 1 The kernel of a DAG

```
Input: a DAG D = (V, E), Noncolored= V (i.e. no vertex is colored for yet)
Output: the DAG, with all its vertices colored, the red vertices being its kernel
while it remains some non colored vertices (Noncolored ≠ ∅) do
  for all v ∈ Noncolored do
    if v is a sink or if all the children of v are green then
      color v in red
      color all the parents of v in green
      remove the colored vertices from Noncolored
    end if
  end for
end while
```

For sure, it is possible to improve this algorithm by using the poset structure of a DAG, and thus replacing the "for all $v \in$ Noncolored" line by something like "for all $v \in$ Tocolornow" where Tocolornow is a set of candidates much smaller than Noncolored.

More generally, in order to color a digraph which is not a DAG, simply split it in p components which are DAGs. Then, apply the above algorithm on each of these DAGs (excepted the cut points that you arbitrarily fix to be red or green). It finally remains to check the global coherence of these colorings. As one has p cutting points (which can also be seen as p branching points in a backtracking version of this algorithm), this leads to at most 2^p kernels. This also suggests why this problem is NP: for large (dense) graph, one should need to cut at least $p \sim n$ points, which leads to a 2^n complexity (lower bound).

4. Generating functions of well-colored unicircuit digraphs

There exists in the literature some noteworthy results on *digraphs* using generating functions (related e.g. to EGF of acyclic digraphs [**18, 28**], Cayley graphs [**26**], (0,1) matrices [**25**], Erdős–Rényi random digraph model [**24**]), but as fas as we know we give here the first example of application to the *kernel* problem.

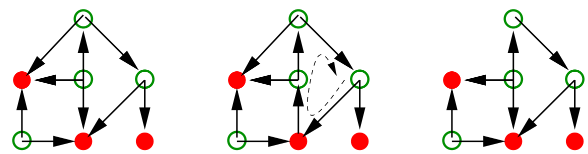


FIGURE 1. The first digraph is a well-colored DAG (it has several cycles, but no circuit). The second digraph is a well-colored digraph (it is not a DAG, as it contains one circuit). The third digraph is a DAG, but is not well colored (the top green vertex misses a red child). [For people who are reading a black & white version of this article, red vertices are fulfilled and green vertices are empty circles.]

The coloring constraints mentioned in Section 3 are "local": they are defined only in function of each vertex and its neighbors. One nice consequence of this "local" viewpoint of kernels is that it opens up a whole range of possibilities for a kind of context-free grammar approach. Indeed if one considers rooted labeled directed trees that are well-colored (*i.e.* which possesses a kernel), one can describe/enumerate them with the help of the five following families of combinatorial structures (all of them being rooted labeled directed trees):

- T: all the trees with the coloring constraint
- $T_r^\uparrow$: well-colored trees with a red root (with an additional out-edge)
- $T_r^\downarrow$: well-colored trees with a red root (with an additional in-edge)
- $T_g^\uparrow$: well-colored trees with a green root (with an additional out-edge)
- $T_g^\downarrow$: well-colored trees with a green root (with an additional in-edge)
- $T_{g_r}^\uparrow$: well-colored trees with a green root (with an additional out-edge which has to be attached to a red vertex)

Those families are related by the following rules:

$$\begin{cases} T = T_g^\uparrow \cup T_r^\uparrow \\ T_g^\uparrow = g^\uparrow \times \mathrm{Set}_{\geq 1}(T_r^\uparrow) \times \mathrm{Set}(T_r^\downarrow \cup T_g^\downarrow \cup T_g^\uparrow) \\ T_g^\downarrow = g^\downarrow \times \mathrm{Set}_{\geq 1}(T_r^\uparrow) \times \mathrm{Set}(T_r^\downarrow \cup T_g^\downarrow \cup T_g^\uparrow) \\ T_r^\uparrow = r^\uparrow \times \mathrm{Set}(T_g^\downarrow \cup T_{g_r}^\uparrow) \\ T_r^\downarrow = r^\downarrow \times \mathrm{Set}(T_g^\downarrow \cup T_{g_r}^\uparrow) \\ T_{g_r}^\uparrow = g^\uparrow \times \mathrm{Set}(T_r^\uparrow \cup T_r^\downarrow \cup T_g^\downarrow \cup T_g^\uparrow) \end{cases}$$

The Set operator reflects the fact that one considers non planar trees, *i.e.* the relative order of the subtrees attached to a given vertex does not matter. The notation $\mathrm{Set}_{\geq 1}$ means one considers non empty set only.

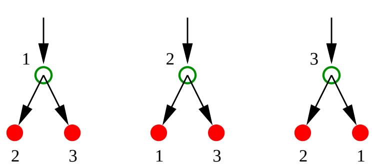


FIGURE 2. A tree $\in T_g^\downarrow$ of size 3 and all its possible labellings. $T_g^\downarrow$ stands for directed trees with a green root with an additional in-edge on this root.

As we are dealing with labeled objects (we refer to Figure 2 for the different labellings of a rooted directed tree), it is more convenient to use exponential generating functions, the above rules are then translated (see e.g. [**20, 14**] for a general presentation of this theory of "graphical enumeration/symbolic combinatorics") into the following set of functional equations (where z marks the vertices):

$$\begin{cases} T(z) = T_g^{\uparrow}(z) + T_r^{\uparrow}(z)\,, \\ T_g^{\uparrow}(z) = T_g^{\downarrow}(z) = z(\exp(T_r^{\uparrow}(z)) - 1)\exp(T_r^{\downarrow}(z) + T_g^{\downarrow}(z) + T_g^{\uparrow}(z))\,, \\ T_r^{\uparrow}(z) = T_r^{\downarrow}(z) = z\exp(T_g^{\downarrow}(z) + T_{g_r}^{\uparrow}(z))\,. \end{cases}$$

Note that $T_{g_r}^{\uparrow} = T$ as one has the trivial bijection "$T_{g_r}^{\uparrow}$ trees with a root without red child" = "$T_r^{\uparrow}$ trees" and "$T_{g_r}^{\uparrow}$ trees with a root with at least a red child" = "$T_g^{\uparrow}$ trees". Define now $T_g(z) := T_g^{\uparrow}(z)$ and $T_r(z) := T_r^{\uparrow}(z)$, the above system simplifies to:

$$\begin{cases} T(z) = T_g(z) + T_r(z) = T_{g_r}^{\uparrow}(z)\,, \\ T_g(z) = z\exp(2T(z)) - z\exp(T(z) + T_g(z))\,, \\ T_r(z) = z\exp(T_g(z) + T(z)) = T(z)\exp(-T_r(z))\,. \end{cases}$$

This system has a unique solution, as the relations can be considered as fixed point equations for power series. Their Taylor expansions are:

$$\begin{aligned} T(z) &= z + 4\frac{z^2}{2!} + 36\frac{z^3}{3!} + 512\frac{z^4}{4!} + 10000\frac{z^5}{5!} + 248832\frac{z^6}{6!} + 7529536\frac{z^7}{7!} + O(z^8)\,, \\ T_g(z) &= \quad 2\frac{z^2}{2!} + 15\frac{z^3}{3!} + 232\frac{z^4}{4!} + 4535\frac{z^5}{5!} + 114276\frac{z^6}{6!} + 3478083\frac{z^7}{7!} + O(z^8)\,, \\ T_r(z) &= z + 2\frac{z^2}{2!} + 21\frac{z^3}{3!} + 280\frac{z^4}{4!} + 5465\frac{z^5}{5!} + 134556\frac{z^6}{6!} + 4051453\frac{z^7}{7!} + O(z^8)\,. \end{aligned}$$

For sure, the i-th coefficients of these series are divisible by i, as we are dealing with rooted object. Here are the 3 generating functions of the corresponding unrooted trees:

$$\begin{aligned} T^{unr.}(z) &= z + 2\frac{z^2}{2!} + 12\frac{z^3}{3!} + 128\frac{z^4}{4!} + 2000\frac{z^5}{5!} + 41472\frac{z^6}{6!} + 1075648\frac{z^7}{7!} + O(z^8)\,, \\ T_g^{unr.}(z) &= \quad \frac{z^2}{2!} + 5\frac{z^3}{3!} + 58\frac{z^4}{4!} + 907\frac{z^5}{5!} + 19046\frac{z^6}{6!} + 496869\frac{z^7}{7!} + O(z^8)\,, \\ T_r^{unr.}(z) &= z + \frac{z^2}{2!} + 7\frac{z^3}{3!} + 70\frac{z^4}{4!} + 1093\frac{z^5}{5!} + 22426\frac{z^6}{6!} + 578779\frac{z^7}{7!} + O(z^8)\,. \end{aligned}$$

Of course, trees are DAG and therefore have a unique kernel. This implies that $T(z)$ is exactly the exponential generating function of directed rooted trees, *i.e.*

$$T(z) = C(2z)/2 \text{ and } T_n = (2n)^{n-1}$$

where $C(z)$ is the Cayley function (see Figure 3 and references [**5, 10**]), defined by

$$C(z) = z\exp(C(z)) = \sum_{n\geq 1} n^{n-1}\frac{z^n}{n!}\,.$$

Solving the set of equations for T, T_g and T_r finally leads to

Theorem 4.1 (Enumeration of well-colored trees)**.**
By ditrees, we mean well-colored rooted labeled directed trees. By well-colored, we mean each green vertex has at least a red child, each red vertex has no red child.
The exponential generating function of ditrees is given by $T(z) = C(2z)/2$,
the EGF of ditrees with a red root is given by

$$T_r(z) = -C(-C(2z)/2),$$

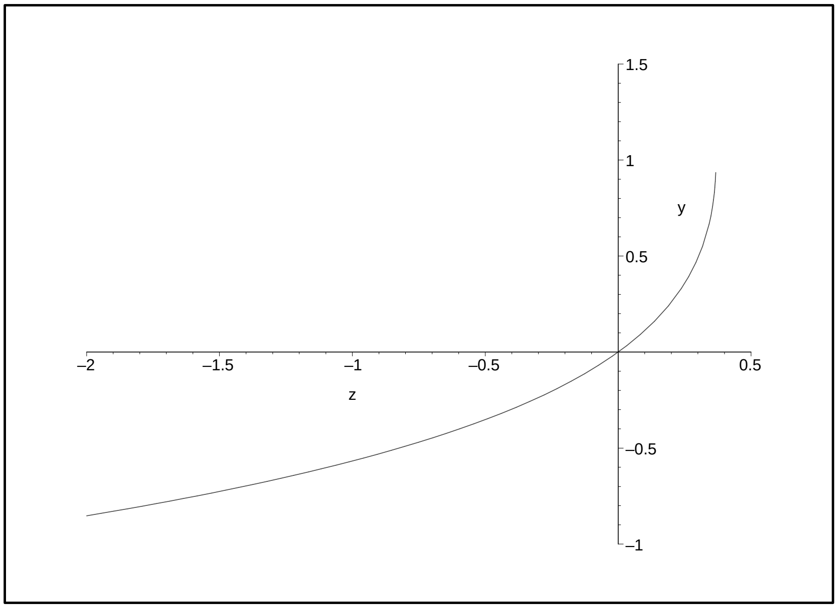


FIGURE 3. The Cayley tree function $C(z)$ goes from $-\infty$ for $z \sim -\infty$ to 1 at $z = \frac{1}{e}$. It satisfies $C(z) = z \exp(C(z))$.

the EGF of ditrees with a green root is given by

$$T_g(z) = C(2z)/2 + C(-C(2z)/2),$$

where $C(z)$ is the Cayley tree function $C(z) = z \exp(C(z))$.
The EGF for the unrooted equivalent objects can be expressed in terms of the rooted ones:

$$T^{unr.} = T - T^2\,, \qquad T_g^{unr.} = T^{unr.} - T_r^{unr.}\,,$$

$$\textit{and} \quad T_r^{unr.} = 2T - 2TT_r + T_r - 2T/T_r + T_r^2/2\,.$$

PROOF. The formulae for T, T_r and T_g can be checked using the definition of $C(z)$ in the fix-point equations in the simplified system above. The fact that the GF for unrooted trees can be expressed in terms of the GF of rooted ones can be proven by integration of the Cayley function, or by a combinatorial splitting argument on trees. □

We can go on and enumerate the different possibilities of circuits for a well-colored digraph. They can be described as

$$\text{Cyc}(g) \quad \cup \quad \text{Cyc}(r \rightarrow \{g \rightarrow\}^+)$$

This reflects the fact that either a circuit is made up of green vertices only, or it contains some red vertices, but they have to be followed by at least a green vertex. NB: Whether one counts or not the cycles of length 1 (*i.e.* a single red or green vertex) will only modify the first term of the generating function. Symbolic combinatorics [**14**] translates the above cycle decompositions in the following function:

$$\ln\left(\frac{1}{1-g}\right) + \ln\left(\frac{1}{1-\frac{rg}{1-g}}\right)$$

where r/g mark the number of red/green vertices. This leads to the following Theorem:

Theorem 4.2 (Enumeration of possible well-colored circuits)**.**
The exponential generating function of possible well-colored circuits is given by

$$L(z) = -\ln(1-z-z^2) = z + 3\frac{z^2}{2!} + 8\frac{z^3}{3!} + 42\frac{z^4}{4!} + 264\frac{z^5}{5!} + 2160\frac{z^6}{6!} + 20880\frac{z^7}{7!} + O(z^8)\,.$$

Its coefficients satisfy $L_n = (n-1)!\,(\phi^n + (1-\phi)^n)$*, where* L_n *are known as the n-th Lucas number (usually defined by the recurrence* $L_{n+2} = L_{n+1} + L_n, L_1 = 1, L_2 = 3$*) and where* $\phi = (1+\sqrt{5})/2$ *is the golden ratio.*

Note that a reverse engineering lecture of this generating function leads to the simpler decomposition $\mathrm{Cyc}(g \cup rg)$, which also explains the recurrence! Now, the following decomposition of possible *cycles* is trivially related to the decomposition of possible *circuits*:

$$\mathrm{Cyc}(r \times \{\,\rightarrow g \cup\ \leftarrow g\}^+ \times \{\,\rightarrow \cup\ \leftarrow\}) \quad \cup \quad \mathrm{Cyc}(g \rightarrow \cup\, g \leftarrow)$$

leads to the EGF $-\ln(1-2z-4z^2)$ whose coefficients are, with no surprise, $2^n L_n$.

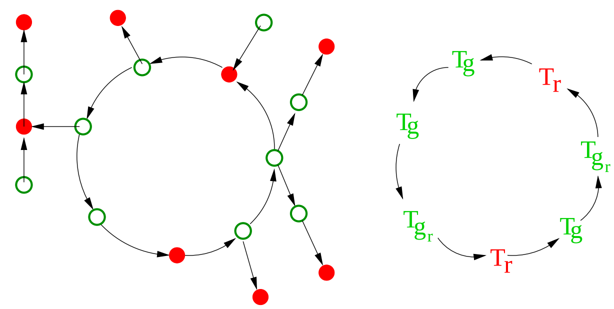


Figure 4. Unicircuit digraphs consist in a circuit with attached trees on it. The left picture above is a unicircuit digraph, to the right, we give its "canonical decomposition" as a circuit of atoms which are trees. Any well-colored unicircuit digraph has such a "canonical decomposition".

Using the decomposition given in Figure 4, one obtains the generating function for unicircuits:
Theorem 4.3 (Enumeration of unicircuit well-colored digraphs)**.**
The EGF of unicircuit well-colored digraphs is

$$\begin{aligned} U(z) &= T^{unr.} - T_g + \ln\left(\frac{1}{1-(T_g + T_{g_r}T_r)}\right) \\ &= -\frac{C(2z)^2}{4} - C(-\frac{C(2z)}{2}) \\ &\quad -\ln\left(1 - \frac{C(2z)}{2} - C(-\frac{C(2z)}{2}) + C\big(-\frac{C(2z)}{2}\big)\frac{C(2z)}{2}\right) \\ &= z + 4\frac{z^2}{2!} + 30\frac{z^3}{3!} + 452\frac{z^4}{4!} + 8840\frac{z^5}{5!} + 224832\frac{z^6}{6!} + 6909784\frac{z^7}{7!} + O(z^8)\,, \end{aligned}$$

where $C(z)$ *is the Cayley tree function* $C(z) = z\exp(C(z))$*.*

Now, consider the larger class of unicycle digraphs (digraphs which have 0 or 1 cycle). Recall that a circuit is a cycle, but a cycle is not necessarily a circuit. In order to get a "canonical decomposition" for unicyle digraphs (similar to the one given in Fig. 4 for unicircuit digraphs), one considers 3 cases:

- Either the graph has no cycle, those graphs are counted by $T^{unr.}$.
- Either it is a cycle with only T_g trees branched on it (*i.e.* no red nodes in the cycle), those graphs are counted by $(\ln\left(\frac{1}{1-2T_g}\right) - 2Tg - 4Tg^2/2)/2 + Tg^2/2$, where $2Tg$ corresponds to $T_g \times \{\rightarrow \cup \leftarrow\}$, one removes cycles of length 1 and 2 from the logarithm (this explains the $-2Tg - 4Tg^2/2$ term)

and one divides the whole formula by 2 because one has to take into account the fact the cycle can be read clockwise or not, and one adds the only legal cycle of length 2.

- Either the graph contains a cycle with some red nodes and then one considers the following possible "bricks":

$$\begin{cases} T_r \leftarrow T_{g_r} \leftarrow \\ T_r \leftarrow T_{g_r} \rightarrow \text{(but not a cycle of length 2, because multiarcs are not allowed)} \\ T_r \rightarrow (T_g\{\rightarrow \cup \leftarrow\})^* \, T_g \leftarrow \qquad \text{(but not a cycle of length 2)} \\ T_r \rightarrow (T_g\{\rightarrow \cup \leftarrow\})^* \, T_{g_r} \rightarrow \\ T_r \leftarrow T_{g_r}\{\rightarrow \cup \leftarrow\} \, (T_g\{\rightarrow \cup \leftarrow\})^* \, T_g \leftarrow \\ T_r \leftarrow T_{g_r}\{\rightarrow \cup \leftarrow\} \, (T_g\{\rightarrow \cup \leftarrow\})^* \, T_{g_r} \rightarrow \end{cases}$$

Theorem 4.4 (Enumeration of unicycle well-colored digraphs)**.**
The EGF of unicycle well-colored digraphs is

$$\begin{aligned} V(z) &= T^{unr.} + \frac{1}{2}\ln\left(\frac{1}{1-2T_g}\right) - Tg - Tg^2/2 - TrTg/2 - T_rT/2 \\ &\quad + \frac{1}{2}\ln\left(\frac{1}{1-\left(2T_rT_{g_r} + \frac{T_rT_g + T_rT_{g_r} + 2T_rT_{g_r}T_g + 2T_rT_{g_r}^2}{1-2T_g}\right)}\right) \\ &= T^{unr.} - T + T_r - T^2/2 - \ln(1+T_r) - \frac{1}{2}\ln(1-2T) \\ &= z + 4\frac{z^2}{2!} + 36\frac{z^3}{3!} + 692\frac{z^4}{4!} + 15920\frac{z^5}{5!} + 458622\frac{z^6}{6!} + 15559264\frac{z^7}{7!} + O(z^8)\,. \end{aligned}$$

where T, T_g, T_r, and $T^{unr.}$ are given in Theorem 4.1.

Note that in the two theorems above, any given non-colored graph is counted with multiplicity 0, 1 or 2 (if there are 0, 1 or 2 ways to color it). We explained in Section 3 that a multiplicity larger than 2 was not possible for unicycle digraphs. We enumerate in the following proposition those with exactly 2 possible colorations.

Proposition 4.5 (Enumeration of unicycle digraphs with two kernels)**.**
The EGF of unicycle digraphs with 2 kernels is

$$D(z) = -\ln\sqrt{1 + C(-C(2z)/2)^2}\,,$$

where $C(z)$ is the Cayley tree function $C(z) = z\exp(C(z))$.

Remark: From the definition of cycle/circuit, $D(z)$ is also the EGF of *unicircuit* digraphs with 2 kernels.

Proof. Let $\mathcal{D}$ be the set of unicycle digraphs with 2 kernels. First, it is easy to see that $\mathrm{Cyc}(T_r^2) \subset \mathcal{D}$ (with a sligth abuse of notation, as we first only consider the shape, not the coloration of the T_r trees): simply color the nodes in the cycle alternatively in green and red, and switch the colors of a part of attached trees, if needs be.

We now prove the next step $\mathcal{D} \subset \mathrm{Cyc}(T_r^2)$: Take a unicycle graph in $\mathcal{D}$, it means at least one of its vertex can be colored both green and red. Such a vertex v can be taken, without loss of generality, in the circuit (from the above remark, the cycle is in fact a circuit). [If it were not the case, all bi-colorabled vertices would be in the tree components, but then one could split our graph to get DAGs which are known to be uniquely colorable]. But when v is red, it implies it has only T_g trees attached to it, which means than when it gets green, the next node in the circuit has be red (and was previously green!). This implies alternation red/green (and even length for parity reasons) for all the nodes in the circuit.

This leads to a canonical decomposition

$$\mathrm{Cyc}(T_r^2)\,.$$

If one divides by 2 for the (anti)clockwise readings, this leads to the Theorem. □

Most of these results (and also the computations of Section 5 hereafter) were checked with the computer algebra system Maple. A worksheet corresponding to this article is available at
`http://algo.inria.fr/banderier/Paper/kernels.mws`
(or `kernels.html`), it uses the Algolib library, downloadable at
`http://algo.inria.fr/libraries/`.

5. Asymptotics

In this section, we give asymptotic results for $n \to +\infty$.

Theorem 5.1 (Proportion of trees with a green/red root)**.**
Asymptotically $\frac{1-\lambda}{1+\lambda} \approx 47.95\%$ *of the trees have a green root, where the constant* $\lambda \approx 0.351733$ *is defined as the unique real root of* $2\lambda = \exp(-\lambda)$.

A more pleasant way to formulate this Theorem consists in considering Nim-type games (first player who cannot move loses) on directed trees where the tree and the starting position are chosen uniformly at random. The strategies of the two players being optimal, the first player has then a probability of 47.95% (asymptotically) to win the game. (Recall that if the starting position can be chosen by the first player, then he will always win.)

Proof. The key step of this result and the following ones are the following expansions (derived from the expansion of the Cayley function) for T, T_r and T_g:

$$T(z) \sim \frac{5}{6} - \frac{1}{\sqrt{2}}\sqrt{1-2ez} + O(1-2ez)$$

$$T_r(z) \sim \lambda - \frac{\lambda\sqrt{2}}{1+\lambda}\sqrt{1-2ez} + O(1-2ez)$$

$$T_g(z) \sim \frac{1}{2} - \lambda - \frac{1}{\sqrt{2}}\frac{1-\lambda}{1+\lambda}\sqrt{1-2ez} + O(1-2ez)\,,$$

where the constant λ is defined as $\lambda := T_r(\frac{1}{2e}) \approx 0.351733$.

By Pringsheim theorem [**14**], as $T_r(z)$ has nonnegative coefficients, then $T_r(z)$ has a positive singularity. As coefficients of T_r are smaller than coefficients of T, its radius of convergence belongs to $[0, 1/(2e)]$. Now, $-C(2z)/2$ is negative on this interval, and thus $C(-C(2z)/2)$ is analytic there, and $1/(2e)$ is therefore its only possible dominating singularity. The radius of T_g follows from $T = T_r + T_g$. The theorem follows by considering $\frac{[z^n]T_g(z)}{[z^n]T(z)} = \frac{1-\lambda}{1+\lambda} - \frac{\lambda^2(\lambda+4)}{(1+\lambda)^5}\frac{1}{n} + O(\frac{1}{n^2})$. □

Theorem 5.2 (Proportion of red vertices in possible circuits)**.**
Asymptotically $\frac{1}{2} - \frac{1}{2\sqrt{5}} \approx 27.63\%$ *of the vertices of a possible circuits are red.*

Proof. One has to considerer the following bivariate generating function (exponential in z, ordinary in u): $\ln\left(\frac{1}{1-(z+uz^2)}\right)$. The wanted proportion is then given by $\frac{[z^n]\partial_u F(z,1)}{[z^n]F(z,1)}$, where $[z^n]\partial_u F(z,1)$ means the n-th coefficient of "the derivative with respect to u of $F(z,u)$, then evaluated at $u=1$". □

Then, one can wonder if the asymptotic density of well-colored unicircuit graphs is more than 50% or even if it is 100%? The following theorem gives the answer:

Theorem 5.3 (Proportion of well-colored unicircuit digraphs)**.**
The proportion of well-colored graphs amongst unicircuit digraphs is asymptotically:

$$\frac{3\lambda^3+\lambda^2-\lambda-1}{(1+\lambda)^2(\lambda-1))} \approx 92.65\%$$

where λ *is the constant defined in Theorem 5.1.*

Proof. Relies on a singularity analysis of the generating function of Theorem 4.3, with the expansions given in Theorem 5.1. Note that some unicircuit digraphs can have 2 kernels, so one has to perform the following asymptotic expansions:

$$\frac{[z^n]U(z) - D(z)}{[z^n]F(z)} \approx 92.65 - \frac{0.12}{n} + O(\frac{1}{n^2}),$$

where $F(z) = T^{unr}(z) + \ln(\frac{1}{1-T(z)}) - T(z)$ is the EGF of (non-colored) unicircuit digraphs. □

For sure, it one considers now the asymptotic density of well-colored *unicircuit* graphs, the proportion should be larger, as one only adds DAGs (which are all well-colorable). The following theorem gives the noteworthy result that unicircuit graphs are in fact almost surely well-colored:

Theorem 5.4 (Proportion of well-colored unicycle digraphs)**.**
There is asymptotically a proportion of $1 - \frac{2\lambda^3\sqrt{2}}{(1+\lambda)^2(1-\lambda)\sqrt{\pi}}\frac{1}{\sqrt{n}} \approx 1 - \frac{0.05}{\sqrt{n}}$ *of well-colored graphs amongst unicycle digraphs of size* n*, where* λ *is the constant defined in Theorem 5.1.*

Proof. Relies on a singularity analysis of the generating function of Theorem 4.4, with the expansions given in Theorem 5.1. Note that some unicycle digraphs can have 2 kernels, so one has to consider

$$\frac{[z^n]V(z) - D(z)}{[z^n]G(z)},$$

where $G(z) = T^{unr}(z) + \frac{1}{2}\ln(\frac{1}{1-2T(z)}) - T(z) - T(z)^2/2$ is the EGF of (non-colored) unicycle digraphs (one substracts $T^2/2$ because amongst the 4 graphs with a cycle of length 2 created by the $\ln(\frac{1}{1-2T(z)})$ part, 3 are not legal: 1 was already counted because of symmetries, and the other 2 have in fact a multiple arc, whereas it is forbidden for our digraphs). □

Finally, if one considers graphs with at most k cycles, it means one has more cutting points, which relaxes the constraints for well-colarability (=existence of kernel). According to the above results, this implies an asymptotic density of one. This gives as a corolary of our results, that all these families have almost surely a kernel. A kind of "limit case" is dense graphs, for which some results already mentionned by Fernandez de la Vega [**13**] and Tomescu [**30**] give that they have indeed almost surely a kernel.

6. Conclusion

It is quite pleasant that our generating function approach allows to get new results on the kernel problem, giving *e.g.* the proportion of graphs satisfying a given property, and new informations on Nim-type games for some families of graphs.

As a first extension of our work, it is possible to apply classical techniques from analytic combinatorics [**14**] in order to get informations on standard deviation, higher moments, and limit laws for statistics studied in Section 5.

Another extension is to get closed form formulas for bicircuit/bicycles digraphs, (the generating function involves the derivative of the product of two logs and the asymptotics are performed like in our Section 5). It is still possible (for sure with the help of a computer algebra system) to do it for 3 or 4 cycles but the "canonical decompositions" and the computations get cumbersome. In order to go on our analysis far beyond low-cyclic graphs, one needs an equation similar to the one given by E.M. Wright [**32, 33**] for graphs. Let $\mathcal{W}_\ell$ be the family of well-colored digraphs with ℓ edges more than vertices, ($\ell \geq -1$). It is possible to get an equation "à la Wright" for $\mathcal{W}_\ell$ by pointing any edge (except edges linking a green vertex to a red one) in a well-colored digraph. It is however not clear for yet if and how such equations can be simplified in order to get a recurrence as "simple/nice" to the one that Wright got for graphs, thus opening the door to asymptotics and threshold analysis beyond the unicyclic case.

Acknowledgements.

This work was partially realized while the second author (Jean-Marie Le Bars) benefited from a delegation CNRS in the Laboratoire d'Informatique de Paris Nord (LIPN). We would like to thank Guy Chaty and François Lévy for a friendly discussion and their bibliographical references [**7, 29**]. We also thank an anomymous referee for detailed comments on a somehow preliminary version of this work.

References

[1] C. Berge and P. Duchet. Recent problems and results about kernels in directed graphs. In *Applications of discrete mathematics (Clemson, SC, 1986)*, pages 200–204. SIAM, Philadelphia, PA, 1988.

[2] Claude Berge. *Graphs*, volume 6 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1985. Second revised edition of part 1 of the 1973 English version.

[3] Claude Berge and Marcel Paul Schützenberger. Jeux de Nim et solutions. *C. R. Acad. Sci. Paris*, 242:1672–1674, 1956.

[4] Immanuel M. Bomze, Marco Budinich, Panos M. Pardalos, and Marcello Pelillo. The maximum clique problem. In *Handbook of combinatorial optimization, Supplement Vol. A*, pages 1–74. Kluwer Acad. Publ., Dordrecht, 1999.

[5] A. Cayley. A theorem on trees. *Quart. J.*, XXIII:376–378, 1889.

[6] Claudette Cayrol, Sylvie Doutre, and Jérôme Mengin. On decision problems related to the preferred semantics for argumentation frameworks. *Journal of Logic and Computation*, 13(3):377–403, 2003.

[7] Guy Chaty, Y. Dimopoulos, V. Magirou, C.H. Papadimitriou, and Jayme L. Szwarcfiter. Enumerating the kernels of a directed graphs with no odd circuits (erratum). *Information Processing Letters*, To appear, 2003.

[8] Guy Chaty and François Lévy. Default logic and kernels in digraphs. *Université de Paris-Nord, Villetaneuse*, Rapport LIPN 91-9, 1991.

[9] V. Chvátal. On the computational complexity of finding a kernel. *Technical report, Report NO CRM-300, univ. de Montréal, Centre de Recherches Mathématiques*, 1973.

[10] Robert M. Corless, David J. Jeffrey, and Donald E. Knuth. A sequence of series for the Lambert W function. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation (Kihei, HI)*, pages 197–204 (electronic), New York, 1997. ACM.

[11] Nadia Creignou. The class of problems that are linearly equivalent to satisfiability or a uniform method for proving NP-completeness. *Theoretical Computer Science*, 145(1-2):111–145, 1995.

[12] Sylvie Doutre. *Autour de la sémantique préférée des systèmes d'argumentation*. PhD thesis, Université Paul Sabatier, Toulouse, 2002. Thèse de Doctorat.

[13] W. Fernandez de la Vega. Kernels in random graphs. *Discrete Mathematics*, 82(2):213–217, 1990.

[14] Philippe Flajolet and Robert Sedgewick. *Symbolic and Analytic Combinatorics*. To appear (chapters are avalaible as Inria research reports). See `http://algo.inria.fr/flajolet/Publications/books.html`, 2003.

[15] Aviezri S. Fraenkel. Planar kernel and Grundy with $d \leq 3$, $d_{\text{out}} \leq 2$, $d_{\text{in}} \leq 2$ are NP-complete. *Discrete Applied Mathematics*, 3(4):257–262, 1981.

[16] Aviezri S. Fraenkel. Combinatorial game theory foundations applied to digraph kernels. *Electronic Journal of Combinatorics*, 4(2):Research Paper 10, approx. 17 pp. (electronic), 1997. The Wilf Festschrift (Philadelphia, PA, 1996).

[17] Masahiko Fukuyama. A Nim game played on graphs. *Theoretical Computer Science*, 304(1-3):387–419, 2003.

[18] Ira M. Gessel. Counting acyclic digraphs by sources and sinks. *Discrete Mathematics*, 160(1-3):253–258, 1996.

[19] Valentin Goranko and Bruce Kapron. The modal logic of the countable random frame. *Archive for Mathematical Logic*, 42(3):221–243, 2003.

[20] Frank Harary and Edgar M. Palmer. *Graphical enumeration*. Academic Press, New York, 1973.

[21] Jean-Marie Le Bars. Fragments of existential second-order logic without 0-1 laws. In *Thirteenth Annual IEEE Symposium on Logic in Computer Science (Indianapolis, IN, 1998)*, pages 525–536. IEEE Computer Soc., Los Alamitos, CA, 1998.

[22] Jean-Marie Le Bars. Counterexamples of the 0-1 law for fragments of existential second-order logic: an overview. *The Bulletin of Symbolic Logic*, 6(1):67–82, 2000.

[23] Jean-Marie Le Bars. The 0-1 law fails for frame satisfiability of propositional modal logic. In *Proceedings of the 17th IEEE Symposium on Logic in Computer Science*. 2002.

[24] Tomasz Łuczak. Phase transition phenomena in random discrete structures. *Discrete Mathematics*, 136(1-3):225–242, 1994. Trends in discrete mathematics.

[25] B. D. McKay, G. F. Royle, I. M. Wanless, F. E. Oggier, N. J. A. Sloane, and H. Wilf. Acyclic digraphs and eigenvalues of (0,1)-matrices. *Preprint*, 2003.

[26] Marni Mishna. Attribute grammars and automatic complexity analysis. *Advances in Applied Mathematics*, 30(1-2):189–207, 2003. Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).

[27] Moses Richardson. Solutions of irreflexive relations. *Ann. of Math. (2)*, 58:573–590, 1953. (Errata in vol. 60, p. 595, 1954).

[28] Robert W. Robinson. Counting labeled acyclic digraphs. In *New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971)*, pages 239–273. Academic Press, New York, 1973.

[29] Jayme L. Szwarcfiter and Guy Chaty. Enumerating the kernels of a directed graph with no odd circuits. *Information Processing Letters*, 51(3):149–153, 1994.

[30] Ioan Tomescu. Almost all digraphs have a kernel. In *Random graphs '87 (Poznań, 1987)*, pages 325–340. Wiley, Chichester, 1990.

[31] John von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton University Press, Princeton, N.J., third edition, 1980.

[32] E. M. Wright. The number of connected sparsely edged graphs. *J. Graph Theory*, 1(4):317–330, 1977.

[33] E. M. Wright. The number of connected sparsely edged graphs. III. Asymptotic results. *Journal of Graph Theory*, 4(4):393–407, 1980.

Cyril Banderier & Vlady Ravelomanana
CNRS (UMR 7030)/Laboratoire d'Informatique de Paris Nord
Institut Galilée, Université Paris 13
99, avenue Jean-Baptiste Clément,
93430 Villetaneuse (France)
E-mail address: Cyril.Banderier@lipn.univ-paris13.fr
URL: http://www-lipn.univ-paris13.fr/∼banderier
E-mail address: Vlady.Ravelomanana@lipn.univ-paris13.fr
URL: http://www-lipn.univ-paris13.fr/∼ravelomanana

Jean-Marie Le Bars
Laboratoire GREYC
Université de Caen
14032 Caen (France)
E-mail address: Jean-Marie.Lebars@info.unicaen.fr
URL: http://users.info-unicaen.fr/∼lebars

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Negative-descent representations for Weyl groups of type D

Riccardo Biagioli and Fabrizio Caselli

Abstract. *We introduce a monomial basis for the coinvariant algebra of type D, that allows us to define a new family of representations of D_n. We decompose the homogeneous components of the coinvariant algebra into a direct sum of these representations and finally we give the decomposition of them into irreducible components. This algebraic setting is then applied to find new, and generalize various, combinatorial identities.*

Résumé. *On introduit une base monomiale de l'algèbre coinvariante de type D, ce qui nous permet de definir une nouvelle classe de representations de D_n. On decompose les composantes homogènes de l'algèbre coinvariante comme somme directe de ces representations et on decrit leur decomposition en modules irreductibles. Ce contexte algebrique est finalement utilisé pour decouvrir des nouvelles identités combinatoires.*

1. Introduction

Let W be one of the classical Weyl groups A_{n-1}, B_n or D_n and let I_n^W be the ideal of the polynomial ring $\mathbf{P}_n := \mathbf{C}[x_1, \dots, x_n]$ generated by costant-term-free W-invariant polynomials. The quotient $R(W) := \mathbf{P}_n / I_n^W$ is called the coinvariant algebra of W and it's well know that it has dimension $|W|$ as a $\mathbf{C}$-vector space. The problem of finding a basis for the coinvariant algebra has been studied by a number of mathematicians (see, e.g., [**3, 4**], [**5**]). Garsia and Stanton presented a descent basis for a finite dimensional quotient of the Stanley-Reisner ring arising from a finite Weyl group (see [**10**]). For type A, unlike for other types, this quotient is isomorphic to $R(W)$ and in this case the basis elements are monomials of degree equal to the "major index" (maj) of the indexing permutation. On the other hand it is well known that $R(W)$ affords the left regular representation of W (see e.g., [**11**]), i.e. the multiplicity of each irreducible representation is its dimension. Moreover, the action of W preserves the natural grading induced from that of $\mathbf{P}_n$ by total degree, and so it is natural to ask about the multiplicity of each irreducible representation of W in the k-th homogeneous component R_k^W. In the case of the symmetric group S_n, the answer is given by a well known theorem, due independently to Kraskiewicz and Weymann [**13**] and Stanley [**18**], that expresses the multiplicity of the irreducible S_n-representations in $R_k^{S_n}$ in terms of the statistic maj defined on standard Young tableaux (SYT).

For type B these problems have been studied by Adin, Brenti and Roichman in [**1**]. They provide a descent basis of $R(B)$ and an extension of the construction of Solomon's descent representations (see [**17**]) for this type.

In this extended abstract we show how to extend these results to the Weyl groups of type D. We construct an analogue of the descent basis for the coinvariant algebra of type D via a Straightening Lemma. The basis elements are monomials of degree $Dmaj$, that is an analogous statistic of maj for D_n (see [**7**]). This basis leads to the definition of a new family of D_n-modules $R_{D,N}$, which have a basis indexed by the even-signed permutations having D and N as "descent set" and "negative set", respectively. For this reason we call them negative-descent representations. They are analogous but different from Solomon descent representations and Kazhdan-Lusztig representations (see [**12**]). We decompose $R_k^{D_n}$ into a direct sums of these $R_{D,N}$. Finally, we introduce the concept of D-standard Young bitableaux. By extending the definition of $Dmaj$ on them we give an explicit decomposition into irreducible modules of these negative-descent representations, refining a theorem of Stembridge [**20**]. This algebraic setting is then applied to obtain new multivariate combinatorial identities.

2. Notation and preliminaries

In this section we give some definitions, notation and results that will be used in the rest of this work. We let $\mathbf{P} := \{1, 2, 3, \ldots\}$, $\mathbf{N} := \mathbf{P} \cup \{0\}$. For $a \in \mathbf{N}$ we let $[a] := \{1, 2, \ldots, a\}$ (where $[0] := \emptyset$). Given $n, m \in \mathbf{Z}$, $n \leq m$, we let $[n, m] := \{n, n+1, \ldots, m\}$.

2.1. Statistics on Coxeter groups. We always consider the linear order on $\mathbf{Z}$

$$-1 \prec -2 \prec \cdots \prec -n \prec \cdots \prec 0 \prec 1 \prec 2 \prec \cdots \prec n \prec \cdots$$

instead of the usual ordering. Given a finite sequence $\sigma = (\sigma_1, \ldots, \sigma_n) \in \mathbf{Z}^n$ we let

$$Inv(\sigma) := \{(i, j) \ : \ i < j \ , \ \sigma_i \succ \sigma_j\} \ \text{ and } \ inv(\sigma) := |Inv(\sigma)|.$$

The *set of descents* and the *descent number* of σ are respectively

$$Des(\sigma) := \{i \in [n-1] \ : \ \sigma_i \succ \sigma_{i+1}\} \ \text{ and } \ des(\sigma) := |Des(\sigma)|.$$

The number of descents in σ from position i on is denoted by

$$d_i(\sigma) := |\{j \in Des(\sigma) \ : \ j \geq i\}|. \tag{2.1}$$

The *major index* of σ (first defined by MacMahon in [**15**]) is

$$maj(\sigma) := \sum_{i \in Des(\sigma)} i.$$

Note that $d_1(\sigma) = des(\sigma)$ and $\sum_{i=1}^n d_i(\sigma) = maj(\sigma)$. Moreover we let

$$Neg(\sigma) := \{i \in [n] : \sigma_i < 0\} \quad \text{and} \quad neg(\sigma) := |Neg(\sigma)|.$$

The generating function of the joint distribution of des and maj over S_n is given by the following Carlitz's Identity, (see, e.g., [**9**]). Let $n \in \mathbf{P}$. Then

$$\sum_{r \geq 0} [r+1]_q^n t^r = \frac{\sum_{\sigma \in S_n} t^{des(\sigma)} q^{maj(\sigma)}}{\prod_{i=0}^n (1 - tq^i)}$$

in $\mathbf{Z}[q][[t]]$, where $[i]_q := 1 + q + q^2 + \ldots + q^{i-1}$.

Let B_n be the group of all bijections β of the set $[-n, n] \setminus \{0\}$ onto itself such that $\beta(-i) = -\beta(i)$ for all $i \in [-n, n] \setminus \{0\}$, with composition as the group operation. We will usually identify $\beta \in B_n$ with the sequence $(\beta(1), \ldots, \beta(n))$ and we call this the *window* notation of β. Following [**2**] we define the *flag-major index* of $\beta \in B_n$ by $fmaj(\beta) := 2maj(\beta) + neg(\beta)$

It's known that *fmaj* is equidistributed with length on B_n and that it satisfies many other algebraic properties (see, for example, [**1**] and [**2**]).

We denote by D_n the subgroup of B_n consisting of all the signed permutations having an even number of negative entries in their window notation, i.e.

$$D_n := \{\gamma \in B_n \ : \ neg(\gamma) \equiv 0 \,(\mathrm{mod}\ 2)\}.$$

Following [**7**] for $\gamma \in D_n$ we let

$$|\gamma|_n := (\gamma(1), \ldots, \gamma(n-1), |\gamma(n)|) \in B_n,$$

$$D_\gamma := Des(|\gamma|_n) \text{ and } N_\gamma := Neg(|\gamma|_n).$$

Then we define the *D-major index* of $\gamma \in D_n$ by

$$Dmaj(\gamma) := 2 \sum_{i \in D_\gamma} i + |N_\gamma|,$$

and the *D-descent number* of γ by

$$Ddes(\gamma) := 2|D_\gamma| + \eta_1(\gamma)$$

where

$$\eta_1(\gamma) := \begin{cases} 1, & \text{if } \gamma(1) < 0, \\ 0, & \text{otherwise.} \end{cases}$$

For example if $\gamma = [2, -5, 3, 1, -4]$, then $D_\gamma = \{1, 3\}$ and $N_\gamma = \{2\}$ and hence $Dmaj(\gamma) = 9$ and $Ddes(\gamma) = 4$.

The statistic $Dmaj$ is Mahonian (i.e. equidistributed with length) on D_n and the generating function of the pair $(Ddes, Dmaj)$ is given by

$$\sum_{r \geq 0} [r+1]_q^n t^r = \frac{\sum_{\gamma \in D_n} t^{Ddes(\gamma)} q^{Dmaj(\gamma)}}{(1-t)(1-tq^n) \prod_{i=1}^{n-1} (1 - t^2 q^{2i})} \tag{2.2}$$

in $\mathbf{Z}[q][[t]]$, (see [**7**, Theorem 4.3] for a proof).

2.2. Partitions and tableaux. A *partition* λ of a nonnegative integer n is an integer sequence $(\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{\ell(\lambda)}$ and $|\lambda| := \sum_i \lambda_i = n$, denoted also $\lambda \vdash n$. We denote by λ' the conjugate partition of λ. The *dominance order* is a partial order defined on the set of partitions of a fixed nonnegative integer n as follows. Let μ and λ two partitions of n. We define $\mu \trianglelefteq \lambda$ if for all $i \geq 1$

$$\mu_1 + \mu_2 + \cdots + \mu_i \leq \lambda_1 + \lambda_2 + \cdots + \lambda_i.$$

A *standard Young tableau* of shape λ is obtained by inserting the integers $1, 2, \ldots, n$ (where $n = |\lambda|$) as *entries* in the cells of the Young diagram of shape λ in such a way that the entries increase along rows and columns. We denote by $SYT(\lambda)$ the set of all standard Young tableaux of shape λ. For example the tableau T in Figure 1 belongs to $SYT(5, 3, 2, 1)$.

T:=

1	3	5	8	10
2	6	7		
4	11			
9				

FIGURE 1

A *descent* in a 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 $Des(T)$. The *major index* of a tableau T is

$$maj(T) := \sum_{i \in Des(T)} i.$$

In the example in Figure 1 $Des(T) = \{1, 3, 5, 8, 10\}$ and so $maj(T) = 27$.

A *bipartition* of a nonnegative integer n is an ordered pair (λ, μ) of partitions such that $|\lambda| + |\mu| = n$ denoted by $(\lambda, \mu) \vdash n$. The *Young diagram* of shape (λ, μ) is obtained by the union of the Young diagrams of shape λ and μ by positioning the second to the south-west of the first. A *standard Young bitableau* $T = (T_1, T_2)$ of shape $(\lambda, \mu) \vdash n$ is obtained by inserting the integers $1, 2, \ldots, n$ in the corresponding Young diagram increasing along rows and columns.

Definition. Given two partitions λ, μ such that $|\lambda| + |\mu| = n$, we define a *D-standard bitableau* $T = (T_1, T_2)$ of type $\{\lambda, \mu\}$ as a standard Young bitableau of shape (λ, μ) or (μ, λ) such that n is an entry of T_1.

We let $Des(T)$ and $maj(T)$ be as above and we let $Neg(T)$ be the set of entries of T_2. The *D-major index* of a D-standard bitableau is defined by

$$Dmaj(T) := 2 \cdot maj(T) + |Neg(T)|.$$

For example T and S in Figure 2 are two D-standard bitableau of type $\{(3, 1), (2, 2, 1)\}$ and we have $Dmaj(T) = 2 \cdot 15 + 5 = 35$ and $Dmaj(S) = 2 \cdot 13 + 4 = 30$.

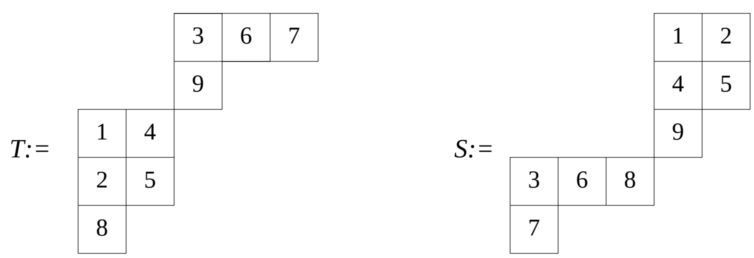


FIGURE 2

We denote by $DSYT\{\lambda, \mu\}$ the set of all D-standard bitableaux of type $\{\lambda, \mu\}$.

2.3. Irreducible representations of classical Weyl groups. Recall that the irreducible representations of the symmetric group S_n are indexed by partitions of n in a classical way (see, for example, [**19**, §7.18]) and denote S^{λ} the irreducible module corresponding to λ

In the case of B_n the irreducible representations are parametrized by ordered pairs of partitions such that the total sum of their parts is equal to n (see, for example, [**14**]), and we denote by $S^{\lambda,\mu}$ the irreducible module corresponding to (λ, μ).

Since D_n is a subgroup of index 2 of the Weyl group B_n, the restrictions of an irreducible representation of B_n to D_n is either irreducible, or splits up into two irreducible components. Let (λ, μ) be a pair of partitions with total size n. If $\lambda \neq \mu$ then the restrictions of the irreducible representations of B_n labeled by (λ, μ) and (μ, λ) are irreducible and equal. If $\lambda = \mu$ then the restriction of the character labeled by (λ, λ) splits into two irreducible components, which we denote by $(\lambda, \lambda)^+$ and $(\lambda, \lambda)^-$. Note that this can only happen if n is even. Hence we may denote all irreducible modules of D_n by $S^{\lambda,\mu,\epsilon}$ where λ and μ are two partitions such that $|\lambda| + |\mu| = n$, $\lambda \preceq \mu$ in some total order $\prec$ on the set of all integer partitions, and ϵ is equal to $\prec$ if $\lambda \neq \mu$ and ϵ is equal to $+$ or $-$ if $\lambda = \mu$.

3. Monomial bases of coinvariant algebras

Let $\mathbf{P}_n := \mathbf{C}[x_1, \ldots, x_n]$ and consider the natural action φ of a classical Weyl group W (with $W = A_{n-1}$, B_n, D_n) on $\mathbf{P}_n$ defined on the generators by

$$\varphi(w) : x_i \mapsto \frac{w(i)}{|w(i)|} x_{|w(i)|},$$

for all $w \in W$ and extended uniquely to an algebra homomorphism. Let I_n^W be the ideal of $\mathbf{P}_n$ generated by the elements in $\mathbf{P}_n^W$ without costant term. The quotient

$$R(W) := \mathbf{P}_n / I_n^W$$

is called the *coinvariant algebra* of W and it is well known that it has dimension $|W|$ as a $\mathbf{C}$-vector space. Moreover, W acts naturally as a group of linear operators on this space and it can be shown that this representation of W is isomorphic to the *regular representation* (see e.g., [**11**, § 3.6]). All these properties naturally lead to the problem of finding a "nice" basis for $R(W)$. A basis for the coinvariant algebra of type A has been found by Garsia and Stanton [**10**]. For $\sigma \in S_n$ they define

$$a_\sigma := \prod_{j \in Des(\sigma)} (x_{\sigma(1)} \cdots x_{\sigma(j)}).$$

It's immediate to see that $a_\sigma := \prod_{i=1}^n x_{\sigma(i)}^{d_i(\sigma)}$ where $d_i(\sigma)$ is defined in (2.1). They show that the set $\{a_\sigma + I_n^{S_n} : \sigma \in S_n\}$ is a basis of $R(S_n)$, called the *descent basis.* Note that the representatives a_σ of this basis are actually monomials with $deg(a_\sigma) = maj(\sigma)$.

Allen ([**4**]) constructed a non-monomial basis for $R(W)$ for all classical Weyl groups and Adin, Brenti and Roichman ([**1**]) defined for any $\beta \in B_n$ a monomial b_β of degree $fmaj(\beta)$ such that the set of the corresponding classes in the coinvariant algebra of type B is a linear basis of this vector space.

The first main goal of this work is to define a family of monomials, indexed by D_n, and to show that the corresponding classes form a basis of the coinvariant algebra of type D. To this end we present a straightening lemma for expanding an arbitrary monomial in $\mathbf{P}_n$ in terms of the descent basis with coefficients in $\mathbf{P}_n^{D_n}$. This algorithm is a generalization of the one presented in [**1**] for type A and B.

For $\gamma \in D_n$ and $i \in [n-1]$, we let

$$\delta_i(\gamma) :=| \{j \in D_\gamma \ : \ j \geq i\} |, \quad \eta_i(\gamma) := \begin{cases} 1, & \text{if } \gamma(i) < 0; \\ 0, & \text{otherwise,} \end{cases}$$

and

$$h_i(\gamma) := 2\delta_i(\gamma) + \eta_i(\gamma).$$

Note that

$$\sum_{i=1}^{n-1} h_i(\gamma) = Dmaj(\gamma) \ \text{ and } \ h_1(\gamma) = Ddes(\gamma). \tag{3.1}$$

Definition. For any $\gamma \in D_n$, we define

$$c_\gamma := \prod_{i=1}^{n-1} x_{|\gamma(i)|}^{h_i(\gamma)}.$$

For example, if $\gamma := (6, -4, -2, 3, -5, -1) \in D_6$, then $(h_1(\gamma), \ldots, h_5(\gamma)) = (6, 5, 3, 2, 1)$ and $c_\gamma = x_6^6 x_4^5 x_2^3 x_3^2 x_5^1$.

The goal of this section is to show how we can prove that the set $\{c_\gamma + I_n^D : \gamma \in D_n\}$ is a linear basis for the coinvariant algebra of type D. We call it the *negative-descent basis.* We denote by

$$f_i(x_1, \ldots, x_n) := \begin{cases} e_i(x_1^2, \ldots, x_n^2), & \text{for } i \in [n-1]; \\ x_1 \cdots x_n, & \text{for } i = n, \end{cases}$$

where e_i is the i-th elementary symmetric function. It is clear that the polynomials f_j are invariant under the action of D_n. Moreover, for any partition $\lambda = (\lambda_1, \ldots, \lambda_t)$ with $\lambda_1 \leq n$, we define $f_\lambda := f_{\lambda_1} \cdots f_{\lambda_t}$. Let's restrict our attention to the quotient $S := \mathbf{P}_n/(f_n)$ and we denote by $\pi : \mathbf{P}_n \to S$ the natural projection. We start by associating to any monomial $M \in S$ an even-signed permutation $\gamma(M)$ and a partition $\mu(M)$. Let M be a monomial such that $\pi(M) \neq 0$, $M = \prod_{i=1}^n x_i^{p_i}$ (note that $p_i = 0$ for some $i \geq 1$). We define $\gamma = \gamma(M) \in D_n$ as the unique even-signed permutation such that, for $i \in [n-1]$,

i) $p_{|\gamma(i)|} \geq p_{|\gamma(i+1)|}$;
ii) $p_{|\gamma(i)|} = p_{|\gamma(i+1)|} \Longrightarrow |\gamma(i)| < |\gamma(i+1)|$;
iii) $p_{|\gamma(i)|} \equiv 0 \,(\mathrm{mod}\ 2) \Longleftrightarrow \gamma(i) > 0$.

Note that the last condition determines also the sign of $\gamma(n)$.

We show how to determine $\gamma(M)$ with an example. For $n = 6$, let $M = x_1^7 x_2 x_3^6 x_5 x_6^4$. Reorder the variables in such a way that the exponents are weakly decreasing without inverting the variables having the same exponent. We obtain $M = x_1^7 x_3^6 x_6^4 x_2^1 x_5^1 x_4^0$. Then $\gamma(M)$ is given by the indices of M reordered in this way and we put a minus sign in the first six entries according to the parity of the corresponding exponent in M. Hence we obtain $\gamma(M) = (-1, 3, 6, -2, -5, -4)$. To define the partition $\mu(M)$ we first need the following observation.

Lemma 3.1. *Let* $M = \prod_{i=1}^n x_i^{p_i}$ *such that* $\pi(M) \neq 0$. *Then the sequence* $(p_{|\gamma(i)|} - h_i(\gamma(M)))$, $i = 1, \ldots, n-1$, *consists of nonnegative even integers and is weakly decreasing.*

We denote by $\mu(M)$ the partition conjugate to $\left(\frac{p_{|\gamma(i)|} - h_i(\gamma)}{2}\right)_{i=1}^{n-1}$, where $\gamma = \gamma(M)$ (note that $\mu(M)_1 < n$). In our running example we have $(h_1(\gamma), \ldots, h_5(\gamma)) = (3, 2, 2, 1, 1)$ and hence $\mu(M) = (3, 2)$.

Now we introduce a technical partial order on the monomials of the same total degree that we will use later on.

Definition. Let M and M' be monomials such that $\pi(M) \neq 0$ and $\pi(M') \neq 0$ with the same total degree and such that the exponents of x_i in M and M' have the same parity 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 $inv(|\gamma(M')|_n) > inv(|\gamma(M)|_n)$.

Lemma 3.2 (Straightening Lemma)**.** *Let M be a monomial in S. Then M admits the following expression*

$$M = f_{\mu(M)} \cdot c_{\gamma(M)} + \sum_{M' < M} n_{M',M} f_{\mu(M')} \cdot c_{\gamma(M')},$$

where $n_{M,M'}$ *are integers.*

For example, let $n = 4$ and $M = x_1^4 x_2 x_4^4$. We have $\gamma(M) = [1, 4, -2, -3]$, $(h_1, h_2, h_3) = (2, 2, 1)$, $c_{\gamma(M)} = x_1^2 x_2 x_4^2$ and $\mu(M) = (2)$. Then, if we set $M_1 = x_1^4 x_2^3 x_4^2$ and $M_2 = x_1^2 x_2^3 x_4^4$, we have that

$$M = c_{\gamma(M)} f_2 - M_1 - M_2$$

in S, with $M_i < M$ for $i = 1, 2$. One can easily verifies that $\gamma(M_1) = [1, -2, 4, -3]$, $\mu(M_1) = \emptyset$, $\gamma(M_2) = [4, -2, 1, -3]$ and $\mu(M_2) = (3)$ and concludes that

$$M = c_{\gamma(M)}\, f_2 - c_{\gamma(M_1)} - c_{\gamma(M_2)}\, f_3.$$

Now the main result of this section is a mere consequence of Lemma 3.2.

Theorem 3.1. *The set*

$$\{c_\gamma + I_n^D \,:\, \gamma \in D_n\}$$

is a basis for $R(D_n)$.

4. Negative-descent representations of D_n

The coinvariant algebra has a natural grading induced from the grading of $\mathbf{P}_n$ by total degree and we denote by R_k its k-th homogeneous component, so that

$$R(W) = \bigoplus_{k \geq 0} R_k.$$

In the case of the symmetric group the major index on standard Young tableaux plays a crucial role in the decomposition of R_k into irreducible representations. The following theorem due independently to Kraskiewicz and Weymann [**13**] and Stanley [**18**, Proposition 4.11] (see also, [**16**, Theorem 8.8]) holds.

Theorem 4.1. *In type A, for $0 \leq k \leq \binom{n}{2}$, the representation R_k is isomorphic to the direct sum $\oplus m_{k,\lambda} S^{\lambda}$, where λ runs through all partitions of n, S^{λ} is the corresponding irreducible S_n-representation, and*

$$m_{k,\lambda} = \mid \{T \in SYT(\lambda) \,:\, maj(T) = k\} \mid .$$

The following is the analogous result for D_n and was proved by Stembridge [**20**] (see also [**4**]). Here we state it in our terminology.

Theorem 4.2. *In type D, for $0 \leq k \leq n^2 - n$, the representation R_k^D is isomorphic to the direct sum $\oplus m_{k,(\lambda,\mu,\epsilon)} S^{\lambda,\mu,\epsilon}$, where $S^{\lambda,\mu,\epsilon}$ is the irreducible representation of D_n labelled as in §2.3, and*

$$m_{k,(\lambda,\mu,\epsilon)} := \mid \{T \in DSYT\{\lambda,\mu\} \,:\, Dmaj(T) = k\} \mid .$$

Now we introduce a new family of D_n-modules $R_{D,N}$. We decompose $R_k^{D_n}$ into a direct sum of these modules and finally we compute the multiplicity of each irreducible representation of D_n in $R_{D,N}$. This result is a refinement of Theorem 4.2.

For any $D \subseteq [n-1]$ we define the partition $\lambda_D := (\lambda_1, \ldots, \lambda_{n-1})$, where $\lambda_i := |D \cap [i, n-1]|$. For $D, N \subseteq [n-1]$, we define the vector

$$\lambda_{D,N} := 2 \cdot \lambda_D + \mathbf{1}_N,$$

where $\mathbf{1}_N \in \{0,1\}^{n-1}$ is the characteristic vector of N. If $\lambda_{D,N}$ is a partition we say that (D,N) is an admissible couple. It is easy to see that (D_γ, N_γ) is admissible for all $\gamma \in D_n$. If (D,N) and (D',N') are two admissible couples then we write $(D,N) \leq (D',N')$ if $\lambda_{D,N} \trianglelefteq \lambda_{D',N'}$. A direct consequence of Lemma 3.2 is that, for all $\gamma, \xi \in D_n$, we have

$$\xi \cdot c_\gamma = \sum_{\{u \in D_n \,:\, (D_u,N_u) \leq (D_\gamma,N_\gamma)\}} n_u c_u + p,$$

where $n_u \in \mathbf{Z}$ and $p \in I_n^D$. It clearly follows that

$$J_{D,N}^{\leq} := \operatorname{span}_{\mathbf{C}}\{c_\gamma + I_n^D \mid \gamma \in D_n,\, (D_\gamma, N_\gamma) \leq (D,N)\}$$

and

$$J_{D,N}^{<} := \operatorname{span}_{\mathbf{C}}\{c_\gamma + I_n^D \mid \gamma \in D_n,\, (D_\gamma, N_\gamma) < (D,N)\}$$

are two submodules of R_k^D, where $k = |\lambda_{D,N}|$, for all admissible couples (D,N). Their quotient is still a D_n-module denoted by

$$R_{D,N} := \frac{J_{D,N}^{\trianglelefteq}}{J_{D,N}^{<}}.$$

If (D,N) is not admissible we let $R_{D,N} := 0$.

Proposition 4.1. *For any D, $N \subseteq [n-1]$, the set*

$$\{\bar{c}_\gamma \,:\, \gamma \in D_n,\ D_\gamma = D \text{ and } N_\gamma = N\},$$

where $\bar{c}_\gamma$ is the image of c_γ in the quotient $R_{D,N}$, is a linear basis of $R_{D,N}$.

By the previous proposition it is natural to call the D_n-module $R_{D,N}$ a *negative-descent representation.* Now we are ready to state the following decomposition of the homogeneous components of the coinvariant algebra.

Theorem 4.3. *For every $0 \leq k \leq n^2 - n$,*

$$R_k^D \cong \bigoplus_{D,N} R_{D,N}$$

as D_n-modules, where the sum is over all $D, N \in [n-1]$ such that $2 \cdot \sum_{i \in D} i + |N| = k$.

Our next goal is to show a simple combinatorial way to compute the multiplicities of the irreducible representations of D_n in $R_{D,N}$.

For any standard Young bitableau $T = (T_1, T_2)$ of shape (λ, μ), following [**1**], we define for $i \in [n]$,

$$h_i(T) := 2 \cdot d_i(T) + \epsilon_i(T), \tag{4.1}$$

where $d_i(T) := |\{j \geq i \ : \ j \in Des(T)\}$, and $\epsilon_i(T) := 1$, if $i \in Neg(T)$ and $\epsilon_i(T) := 0$ otherwise.

The following technical lemma is the key ingredient in the proof of the next theorem.

Lemma 4.2. *Let $T = (T_1, T_2)$ be a Young standard bitableau of total size n such that $n \in T_1$. Then*

$$h_i(T_1, T_2) = h_i(T_2, T_1) + 1$$

for all $i = 1, \ldots, n$.

Theorem 4.4. *For any pair of subset $D, N \subseteq [n-1]$, and a bipartition of n $(\lambda, \mu) \vdash n$, the multiplicity of the irreducible D_n-representation corresponding to $(\lambda, \mu)^\epsilon$ in $R_{D,N}$ is*

$$m_{D,N,(\lambda,\mu)^\epsilon} := |\{T \in DSYT\{\lambda, \mu\} \ : \ Des(T) = D \text{ and } Neg(T) = N\}|.$$

Theorem 4.2 easily follows from this and Theorem 4.3, by observing that $\sum_{i=1}^{n-1} h_i(T) = Dmaj(T)$, for any $T \in DSYT\{\lambda, \mu\}$.

5. Combinatorial Identities

In this last section we compute the Hilbert series of the polynomial ring $\mathbf{P}_n$ with respect to multi-degree rearranged into a weakly decreasing sequence in two different ways and we deduce from this some new combinatorial identities. In particular we obtain one of the main results of [**7**, Corollary 4.4] as a special case of Corollary 5.1.

Following [**6**] we recall the negative statistics on D_n. For $\gamma \in D_n$ we define the *D-negative descent multiset*

$$DDes(\gamma) = Des(\gamma) \biguplus \{Neg(\gamma^{-1})\} \setminus \{n\}. \tag{5.1}$$

and we let

$$ddes(\gamma) := |DDes(\gamma)| \text{ and } dmaj(\gamma) := \sum_{i \in DDes(\gamma)} i.$$

The Hilbert series of $\mathbf{P}_n$ can be computed by considering the even-signed descent basis for the coinvariant algebra of type D and applying the Straightening Lemma. It's easy to see that the map $\mathbf{P}_n \to D_n \times \mathcal{P}(n)$ given by

$$M \mapsto (\gamma(M), \bar{\mu}(M)'), \tag{5.2}$$

is a bijection, where, if $M = f_n^t M'$, with $M' \in S$, then $\bar{\mu}(M) = ((n)^t, \mu(M'))$. For a partition λ we let $m_j(\lambda) := |\{i \in [n] \ : \ \lambda_i = j\}|$, and

$$\binom{n}{\bar{m}(\lambda)} := \binom{n}{m_0(\lambda), m_1(\lambda), \ldots},$$

be the multinomial coefficient.

Theorem 5.1. *Let $n \in \boldsymbol{P}$. Then*

$$\sum_{\ell(\lambda) \leq n} \binom{n}{\bar{m}(\lambda)} \prod_{i=1}^{n} q_i^{\lambda_i} = \frac{\sum_{\gamma \in D_n} \prod_{i=1}^{n-1} q_i^{2\delta_i(\gamma) + \eta_i(\gamma)}}{(1 - q_1 \cdots q_n) \prod_{i=1}^{n-1} (1 - q_1^2 \cdots q_i^2)},$$

in $\boldsymbol{Z}[[q_1, \ldots, q_n]]$.

Now we compute the Hilbert series in a different way using the following observation. Let $T := \{\sigma \in D_n : des(\sigma) = 0\}$. It is well known, and easy to see, that

$$D_n = \biguplus_{u \in S_n} \{\sigma u : \sigma \in T\}, \tag{5.3}$$

where $\biguplus$ denotes disjoint union. Now define $\bar{n}_i(\gamma) := |\{j \geq i \ : \ j \in Neg(|\gamma|_n)\}|$. It follows that

$$ddes(\gamma) = d_1(\gamma) + \bar{n}_1(\gamma). \tag{5.4}$$

Theorem 5.2. *Let* $n \in \boldsymbol{P}$. *Then*

$$\sum_{\ell(\lambda) \leq n} \binom{n}{\bar{m}(\lambda)} \prod_{i=1}^{n} q_i^{\lambda_i} = \frac{\sum_{\gamma \in D_n} \prod_{i=1}^{n-1} q_i^{d_i(\gamma)+\bar{n}_i(\gamma^{-1})}}{\prod_{i=1}^{n-1}(1-q_1^2 \cdots q_i^2)(1-q_1 \cdots q_n)},$$

in $\boldsymbol{Z}[[q_1, \ldots, q_n]]$.

The following beautiful identity easily follows by Theorems 5.1 and 5.2.

Corollary 5.1. *Let* $n \in \boldsymbol{P}$. *Then*

$$\sum_{\gamma \in D_n} \prod_{i=1}^{n-1} q_i^{d_i(\gamma)+\bar{n}_i(\gamma^{-1})} = \sum_{\gamma \in D_n} \prod_{i=1}^{n-1} q_i^{2\delta_i(\gamma)+\eta_i(\gamma)}.$$

□

The two pair of statistics $(ddes, dmaj)$ and $(Ddes, Dmaj)$ have the same distribution on D_n, (see [**7**, Corollary 4.4]) given by (2.2). Now it is clear that this result follows directly by Corollary 5.1 by setting $q_1 = qt$ and $q_i = q$ for $i \geq 2$.

Corollary 5.2. *Let* $n \in \boldsymbol{P}$. *Then*

$$\sum_{\gamma \in D_n} t^{ddes(\gamma)} q^{dmaj(\gamma)} = \sum_{\gamma \in D_n} t^{Ddes(\gamma)} q^{Dmaj(\gamma)}.$$

Acknowledgments. We would like to thank Francesco Brenti for his helpful suggestions and comments and Yuval Roichman for some useful discussions.

References

[1] R. M. Adin, F. Brenti and Y. Roichman, *Descent representations and multivariate statistics*, Trans. Amer. Math. Soc., to appear.

[2] R. M. Adin and Y. Roichman, *The flag major index and group actions on polynomial rings*, European J. Combin., **22** (2001), 431-446.

[3] E. E. Allen, *The diagonal monomials and basis for the diagonally symmetric polynomials*, J. Alg. Combin., **3** (1994), 5-16.

[4] E. E. Allen, *New bases for the decomposition of the graded left regular representation of the reflection groups of type B_n and D_n*, J. Algebra, **173** (1995), 122-143.

[5] H. Barcelo, *Young straightening in a quotient S_n-module*, J. Alg. Combin., **2** (1993), 5-23. 38, 1999.

[6] R. Biagioli, *Major and descent statistics for the even-signed permutation group*, Adv. in Appl. Math., **31** (2003), 163-179.

[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, *Hilbert series of Invariant algebras for classical Weyl groups*, Proc. FPSAC 2003, University of Linkoping, Sweden.

[9] L. Carlitz, *A combinatorial property of q-Eulerian numbers*, Amer. Math. Montly, **82** (1975), 51-54.

[10] A. Garsia and D. Stanton, *Group actions of Stanley-Reisner rings and invariant of permutation groups*, Adv. Math., **51** (1984), 107-201.

[11] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Stud. Adv. Math., no.29, Cambridge Univ. Press, Cambridge, 1990.

[12] D. Kazhdan and G. Lusztig, *Representations of Coxeter groups and Hecke algebras*, Invent. Math., **53** (1979), 165-184.

[13] W. Kraskiewicz and J. Weymann, *Algebra of coinvariants and the action of Coxeter elements*, Bayreuth. Math. Schr., **63** (2001), 265-284.

[14] I. G. Macdonald, *Symmetric functions and Hall polynomials*, second edition, Oxford Math. Monographs, Oxford Univ. Press, Oxford 1995.
[15] P. A. MacMahon, *Combinatory analysis*, vol.1, Cambridge Univ. Press, London, 1915.
[16] C. Reutenauer, *Free Lie algebras,* London Mathematical Soc. Monographs, New series 7, Oxford Univ. Press, 1993.
[17] L. Solomon, *The order of the finite Chevalley groups*, J. Algebra **3** (1966), 376-393.
[18] R. P. Stanley, *Invariants of finite groups and their applications to combinatorics*, Bull. Amer. Math. Soc. (new series), **1** (1979), 475-511.
[19] R. P. Stanley, *Enumerative combinatorics*, vol.2, Cambridge Stud. Adv. Math., no.62, Cambridge Univ. Press, Cambridge, 1999.
[20] J. Stembridge, *On the eigenvalues of representations of reflection groups and wreath products,* Pacific J. Math., **140** (1989), 359-396.

(R. Biagioli) LaCIM, Université du Québec à Montréal, Case Postale 8888, succursale Centre-ville, Montréal, Québec, Canada H3C 3P8
E-mail address: biagioli@math.uqam.ca
URL: http://www.lacim.uqam.ca/~biagioli

(F. Caselli) Institut Girard Desargues, Université Claude Bernard Lyon 1, 21 Av C. Bernard, 69622 Villeurbanne, France
E-mail address: caselli@igd.univ-lyon1.fr
URL: http://igd.univ-lyon1.fr/~caselli/

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Sharper estimates for the number of permutations avoiding a layered or decomposable pattern

Miklós Bóna

Abstract. *We present two methods that for infinitely many patterns q provide better upper bounds for the number $S_n(q)$ of permutations of length n avoiding the pattern q than the recent general result of Marcus and Tardos. While achieving that, we define an apparently new decomposition of permutations .*

Résumé. *Nous montrons deux méthodes qui prouvent des bornes supérieurs pour les nombres $S_n(q)$ dénombrant les permutations de longueur n évitant le motif q. Nos méthodes peuvent être appliquées pour un nombre infini des motifs, et les bornes obtenues sont meilleur que celles découlant du résultat récent de Marcus et Tardos. Nous allons également définir une decomposition des permutations qui semble d'être nouvelle.*

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. The long-standing Stanley-Wilf conjecture claimed that for any given pattern q, there exist an absolute constant c_q so that $S_n(q) < c_q^n$ for all n. See [**3**] or [**6**] for the relevant definitions.

The Stanley-Wilf conjecture was open for more than 20 years. It has recently been proved by a spectacular, yet simple argument [**11**]. That argument actually proved a stronger conjecture, the Füredi-Hajnal conjecture [**8**] , which was shown to imply the Stanley-Wilf conjecture three years ago in [**9**].

Perhaps because the Stanley-Wilf conjecture was proved as a special case of a stronger conjecture, the obtained upper bound seems far away from what is thought to be the truth. Indeed, it is proved in [**11**] (along with another, stronger conjecture from [**1**]), that if q is a pattern of length k, then

$$S_n(q) \leq c_q^n \qquad \text{where} \qquad c_q \leq 15^{2k^4\binom{k^2}{k}}. \tag{1.1}$$

For the rest of this paper, k will denote the length of the pattern q. For instance, if $k = 3$, then the above result shows only that $c_q \leq 15^{13608}$, while in fact it is well-known [**3**] that $c_q = 4$ is sufficient. Therefore, it seems reasonable to think that in the near future significant research will be devoted to the improvement of this upper bound. In fact, R. Arratia [**2**] conjectures that $c_q \leq (k-1)^2$ for any patterns q. There are several patterns, for instance, monotone patterns, for which $(k-1)^2$ is known [**10**] to be the smallest possible value of c_q.

In this paper we present two methods that can prove upper bounds for certain patterns from the upper bounds for certain shorter patterns.

Department of Mathematics, University of Florida, Gainesville FL 32611-8105. Supported by a Young Investigator Grant of the National Security Agency.

For instance, one of our methods will provide upper bounds for all layered patterns, which are patterns consisting of decreasing subsequences that increase among the layers. The other one will work for all decomposable permutations. The arguments will be remarkably simple, compared to previous work on layered patterns. While our upper bounds will still be significantly weaker than the conjectured $(k-1)^{2n}$, they will not be doubly exponential, like the result shown in (1.1).

We mention that it follows from subsequent work of present author ([**4**], to be presented at the subsequent Pattern Avoiding Permutations conference) that for any layered pattern q of length k, we have $L(q) = \lim_{n\to\infty} \sqrt[n]{S_n(q)} \geq (k-1)^2$. In other words, in the asymptotic sense, layered patterns are at least as easy to avoid as the monotone patterns. Present paper complements those results by bounding $L(q)$ from above.

2. The Pattern 1324

We explain our method by demonstrating it on the pattern 1324, but this is only to make our discussion easier to read. The crucial properties of this pattern for our purposes are that it starts with its minimal entry, it ends with its maximal entry, and that if we remove either of these entries, we get a pattern (132 or 213) for which a good exponential upper bound is known.

Our crucial definition is the following.

Definition 2.1. We will say that an n-permutation $p = p_1p_2\cdots p_n$ is *orderly* if $p_1 < p_n$. We will say that p is *dual orderly* if the entry 1 of p precedes the maximal entry n of p.

It is clear that p is orderly if and only if p^{-1} is dual orderly.

The importance of these permutations for us is explained by the following lemma.

Lemma 2.2. *The number of orderly (resp. dual orderly) 1324-avoiding n-permutations is less than $8^n/4(n+1)$.*

Proof. It suffices to prove the statement for orderly permutations as we can take inverses after that to get the other statement.

The crucial idea is this. Each entry p_i of p has at least one of the following two properties.

(a) $p_i \geq p_1$;
(b) $p_i \leq p_n$.

In words, everything is either larger than the first entry, or smaller than the last, possibly both. This would *not* be the case had we not required that p be orderly.

Define $S = \{i|p_i \geq p_1\}$ and $T = \{i|p_i < p_1\}$. Then S and T are disjoint, $S \cup T = [n]$, and crucially, if $i \in T$, then, in particular, $p_i < p_n$. Recall that for any pattern q of length three, we have $S_n(q) = C_n = \binom{2n}{n}/(n+1)$, and that the numbers C_n are the well-known Catalan numbers [**3**]. Let $|S| = s$ and $|T| = t$. Then we have C_{s-1} possibilities for the substring p_S of entries belonging to indices in S, and $C_t = C_{n-s}$ possibilities for the substring p_T of entries belonging to indices in S. Indeed, p_S starts with its smallest entry, and then the rest of it must avoid 213, (otherwise, together with p_1, a 1324-pattern is formed) and p_T must avoid 132 (otherwise, together with p_n, a 1324-pattern is formed). Finally, we have $\binom{n-2}{s-2}$ choices for the set of indices that we denoted by S. Once s is known, we have no liberty in choosing the *entries* p_i, $(i \in S)$ as they must simply be the s largest entries.

Therefore, the total number of possibilities is

$$\sum_{s=2}^{n} \binom{n-2}{s-2} C_{s-1}C_{n-s} < 2^{n-2} \sum_{s=2}^{n} C_{s-1}C_{n-s} < 2^{n-2}C_n < \frac{8^n}{4(n+1)}.$$

□

We have seen that it helps in our efforts to limit the number of 1324-avoiding permutations if a large element is preceded by a small one. To make good use of this observation, look at all non-inversions of a generic permutation $p = p_1p_2\cdots p_n$; that is, pairs (i,j) so that $i<j$ and $p_i<p_j$. Find the non-inversion (i,j) for which

$$\max_{(i,j)}(j-i, p_j-p_i) \tag{2.1}$$

is maximal. If there are several such pairs, take one of them, say the one that is lexicographically first. Call this pair (i,j) the *critical pair* of p.

Recall that an entry of a permutation is called a *left-to-right minimum* if it is smaller than all entries on its left. Similarly, an entry is a *right-to-left maximum* if it is larger than all entries on its right.

The following proposition is obvious, but it will be important in what follows, so we explicitly state it.

Proposition 2.3. *For any permutation $p_1p_2\cdots p_n$, the critical pair (i,j) is always a pair in which p_i is a left-to-right minimum, and p_j is a right-to-left maximum.*

The following definition proved to be useful for treating 1324-avoiding permutations in the past [**7**].

Definition 2.4. We say that two permutations are *in the same class* if they have the same left-to-right minima, and the same right-to-left maxima, and they are in the same positions.

Example 2.5. *The permutations 3612745 and 3416725 are in the same class.*

Proposition 2.6. *The number of nonempty classes of n-permutations is less than 9^n.*

Proof. Each such class contains exactly one 1234-avoiding permutation, namely the one in which all entries that are not left-to-right minima or right-to-left maxima are written in decreasing order. As it is well-known that $S_n(1234) < 9^n$, the statement is proved. □

To achieve our goal, it suffices to find a constant C so that each class contains at most C^n 1324-avoiding n-permutations.

Choose a class A. By Proposition 2.3, we see that the critical pair of any permutation $p\in A$ is the same as it depends only on the left-to-right minima and the right-to-left maxima, and those are the same for all permutations in A.

We will now find an upper bound for the number of 1324-avoiding n-permutations in A.

For symmetry reasons, we can assume that in the critical pair of $p\in A$, we have $j-i\geq p_j-p_i$, in other words, the maximum (2.1) is attained by $j-i$.

We will now reconstruct p from its critical pair. First, all entries that precede p_i must be larger than p_j. Indeed, if there existed $k<i$ so that $p_k<p_j$, then the pair (j,k) would be a "longer" non-inversion than the pair (i,j), contradicting the critical property of (i,j). Similarly, all entries that are on the right of p_j must be smaller than p_i.

This shows that all entries p_t for which $p_i<p_t<p_j$ must be positioned between p_i and p_j, that is, $i<t<j$ must hold for them. However, if $j-i = p_j-p_i+b$, where b is a *positive* integer, then we can select b additional entries that will be located between p_i and p_j. We will call them *excess entries*; that is, an excess entry is an entry p_u that is located between p_i and p_j, but does *not* satisfy $p_i<p_u<p_j$.

The good news is that we do not have too many choices for the excess entries. No excess entry can be smaller than p_i-b. Indeed, if we had $p_u<p_i-b$ for an excess entry, then for the pair (u,j) the value defined by (2.1) would be larger than for the pair (i,j), contradicting the critical property of (i,j). By the analogous argument, no excess entry can be larger than p_j+b. Therefore, the set of b excess entries must be a subset of the at-most-$(2b)$-element set $(\{p_i-b, p_i-b+1,\cdots,p_i-1\}\cup\{p_j+1,p_j+2,\cdots,p_j+b\})\cap[n]$. Therefore, we have at most $\binom{2b}{b}$ choices for the set of excess entries, and consequently, we have $\binom{2b}{b}$ choices for the set of $j-i-1+b$ elements that are located between p_i and p_j. As $p_i<p_j$, the partial permutation $p_ip_{i+1}\cdots p_j$ is orderly, and certainly 1324-avoiding. Therefore, by Lemma 2.2, we have less than $8^{j-i+1}/4(j-i+1)$ choices for it once the set of entries has been chosen.

This proves that altogether, we have less than

$$4^b \cdot \frac{8^{j-i+1}}{4(j-i+1)} < 32^{j-i}$$

possibilities for the string $p_i p_{i+1} \cdots p_j$. We used the fact that $b \leq j-i-1$ as b counts the excess entries between i and j. Note that we have some room to spare here, so we can say that the above upper bound remains valid even if we include the permutations in which the maximum was attained by (p_i, p_j), and not by (i, j).

We can now remove the entries $p_{i+1} \cdots p_{j-1}$ from our permutations. This will split our permutations into two parts, p_L on the left, and p_R on the right. It is possible that one of them is empty. We know exactly what entries belong to p_L and what entries belong to p_R; indeed each entry of p_L is larger than each entry of p_R. Therefore, we do not loose any information if we relabel the entries in each of p_L and p_R so that they both start at 1 (we call this the *standardization of the strings*). This will not change the location and relative value of the left-to-right minima and right-to-left maxima either. The string $p_{i+1} \cdots p_{j-1}$ should not be standardized, however, as that would result in loss of information.

See Figure 1 for the diagram of a generic permutation, its critical pair, and the strings p_L and p_R.

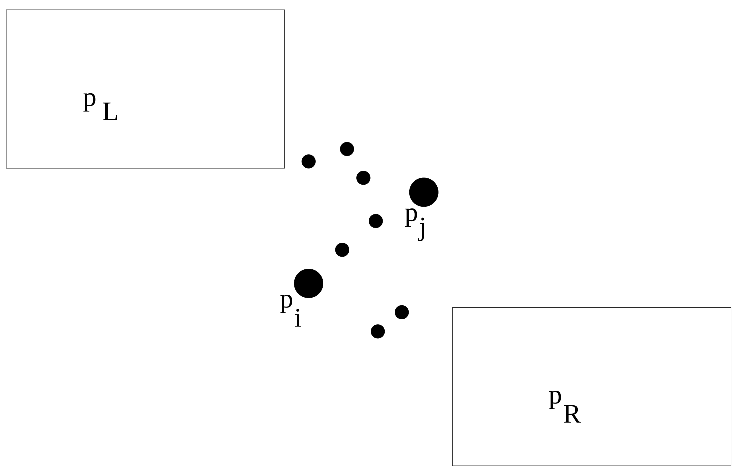


FIGURE 1. A generic permutation and its critical pair.

Then we iterate our procedure. That is, we find the critical pairs of p_L and p_R, denote them by (i_L, j_L) and (i_R, j_R), and prove, just as above, that there are at most $32^{j_L - i_L}$ possibilities for the string between i_L and j_L, and there are at most $32^{j_R - i_R}$ possibilities for the string between i_R and j_R. Then we remove these strings again, cutting our permutations into more parts, and so on, building a binary tree-like structure of strings. Note that the leaves of this tree will be orderly or dual orderly permutations.

Note that this procedure of decomposing of our permutations is injective. Indeed, given the standardized string p_L, the partial permutation $p_i \cdots p_j$, and the standardized string p_R, we can easily recover p.

Iterating this algorithm until all entries of p that are not left-to-right minima or right-to-left maxima are removed, we prove the following.

Lemma 2.7. *The number of 1324-avoiding n-permutations in any given class A is at most 32^n.*

PROOF. The above description of the removal of entries by our method shows that the total number of 1324-avoiding permutations in A is less than

$$32^{\sum_k j_k - i_k}$$

where the summation ranges through all intervals (i_k, j_k) whose endpoints were critical pairs at some point. As these interiors of these intervals are all disjoint, $\sum_k j_k - i_k = n-1$, and our claim is proved. □

Now proving the upper bound for $S_n(1324)$ is a breeze.

Theorem 2.8. *For all positive integers n, we have $S_n(1324) < 288^n$.*

PROOF. As there are less than 9^n classes and less than 32^n n-permutations in each class that avoid 1324, the statement is proved. □

Note that an alternative way of proving our theorem would have been by induction on n. We could have used the induction hypothesis for the class A' that is obtained from A by making p_i and p_j consecutive entries by omitting all positions between them, and setting their values so that each entry on the left of p_i is larger than each entry after p_j.

Finally, we point out that using specific properties of the pattern 1324, we could have decreased the upper bound a little further, but that is not our goal here. Our goal is to find a method that works for many patterns.

3. Layered Patterns

As a generalization, we look at patterns like 14325, 154326, and so on, that is, patterns that start with 1, end with their maximal entry k, and consist of a decreasing sequence all the way between.

Theorem 3.1. *Let $k \geq 4$, and let $q_k = 1\ k-1\ k-2 \cdots 2\ k$. Then for all positive integers n, we have*

$$S_n(q_k) < 72^n (k-2)^{2n} = \left(72(k-2)^2\right)^n.$$

PROOF. We again look at orderly permutations first. If p is orderly and avoids q_k, then define S, p_S and T, p_T just as in the proof of Lemma 2.2. Then p_S starts with its smallest entry, and the rest must avoid $q'_k = k-2 \cdots 2\ 1\ k-1$, whereas p_T must avoid $q''_k = 1\ k-1\ k-2 \cdots 2$. It is known that $S_n(q'_k) = S_n(q''_k) = S_n(12 \cdots (k-1)) < (k-2)^{2n}$, so it follows, just as in Lemma 2.2 that the number of orderly (resp. dual orderly) n-permutations that avoid q_k is less than $(2(k-2)^2)^n$.

The transition from orderly permutations to generic permutations is identical to what we described in the case of $q_4 = 1324$. □

Let us now find an upper bound for all layered patterns. Recall that a permutation is called layered if it is the concatenation of decreasing subsequences $d_1, d_2, \cdots, d_t$ so that each entry of d_i is less than each entry of d_j for all $i < j$. For instance, 321546 is a layered pattern. We will use the following definition and lemma, first used in [**7**].

Definition 3.2. Let q be a pattern, and let y be an entry of q. Then *to replace* y by the pattern w is to add $y-1$ to all entries of w, then to delete y and to succesively insert the entries of w at its position.

Lemma 3.3. *("replacing an element by a pattern") Let q be a pattern and let y be an entry of q so that for any entry x preceding y we have $x < y$ and for any entry z preceded by y we have $y < z$. Suppose that $S_n(q) < K^n$ for some constant K and for all n.*

Let w be a pattern of length m starting with 1 and ending with m so that $S_n(w) < C^n$ holds for all n, for some constant C. Let q' be the pattern obtained by replacing the entry y by the pattern w in q. Then $S_n(q') < (4CK)^n$.

PROOF. Take an n-permutation p which avoids q'. Suppose it contains q. Then consider all copies of q in p and consider their entries y. Color these entries blue, that is, and entry is blue if it can play the role of y in a copy of q. Clearly, these entries must form a permutation which does not contain w. For suppose they do, and denote y_1 and y_m the first and last elements of that purported copy of w. Then the initial segment of the copy of q which contains y_1 followed by the y_2 through y_{k-1} and the ending segment of the copy of q which contains y_k would form a copy of q'.

Therefore, if p avoids q', then it either avoids q, or the substring of its blue entries avoids w. As we have at most 2^{n-1} choices for the set of blue entries, and at most 2^{n-1} choices for their positions, this shows that less than $(4C)^{n-1} \cdot K^n + K^n < (4CK)^n$ permutations of length n can avoid q'. □

Now let $Q = Q(a_1, a_2, \cdots, a_t)$ be the layered pattern whose layers are of length $a_1, a_2, \cdots, a_t$. It is then clear that Q is contained in the pattern $Q' = Q(1, a_1, 1, a_2, 1, \cdots 1, a_t, 1)$. Therefore,

$$S_n(Q) \leq S_n(Q'). \tag{3.1}$$

On the other hand, Q' can be obtained if we take $q_{a_1} = Q(1, a_1, 1)$, then replace the last entry of this pattern by the pattern $q_{a_2} = Q(1, a_2, 1)$, then replace the last entry of the obtained pattern by $q_{a_3} = Q(1, a_3, 1)$, and so on.

Then it follows by iterated applications of Theorem 3.1 and Lemma 3.3 that

$$S_n(Q') \leq 4^{tn} \cdot 72^{tn} \prod_{i=1}^{t} a_i^{2n} = 288^{tn} \prod_{i=1}^{t} a_i^{2n}.$$

So by (3.1), we have

$$S_n(Q) \leq 288^{tn} \prod_{i=1}^{t} a_i^{2n}.$$

For any fixed layered pattern Q, the number of layers t will be fixed, so 288^{tn} is simply exponential. While $\prod_{i=1}^{t} a_i$ can be as large as $3^{k/3}$, which makes c_Q an exponential function of k, it is still not doubly exponential, unlike the general result (1.1).

4. Further Generalizations

We can find a somewhat more general application of our methodology. For a pattern q, let $1q$ denote the pattern obtained from q by adding one to each of the entries and then writing 1 to the front, and let qm denote the pattern that we obtain from q by simply affixing a new maximal element to the end of q. Finally, let $1qm$ denote the pattern $(1q)m = 1(qm)$. So for example, if $q = 2413$, then $1q = 13524$, and $qm = 24135$, while $1qm = 135246$.

Theorem 4.1. *Let q be a pattern so that there exist constants c_1 and c_2 satisfying $S_n(1q) < c_1^n$ and $S_n(qm) < c_2^n$ for all n. Then for all positive integers n, we have*

$$S_n(1qm) < 72^n \cdot (\max(c_1, c_2))^n.$$

Proof. Similar to the proof of Theorem 3.1. The upper bound for orderly permutations is $2^n \cdot (\max(c_1, c_2))^n$, the number of classes is 9^n, and the remaining 4^n comes from the choices for the excess entries. □

This theorem permits a little improvement on the general upper bound (1.1) for all patterns that start with their minimal entry and end in their maximal entry.

Corollary 4.2. *If $r = 1qm$, then*

$$c_r \leq 72^n \cdot 15^{2(k-1)^4 \binom{(k-1)^2}{k-1}}.$$

Proof. Follows from (1.1), applied to the patterns $1q$ and qm, and Theorem 4.1. □

While this last corollary is not a significant improvement as far as principles are concerned, numerically it still decreases c_r by several orders of magnitude.

Another improvement comes from a variation of Lemma 3.3.

Lemma 4.3. *Let q be as in lemma 3.3 and let y be its* last *entry. Replace y by any pattern w which starts with its smallest entry. Then for the pattern q' obtained this way, we have*

$$S_n(q') < (4CK)^n.$$

Proof. This can be proved exactly as Lemma 3.3. The special values and positions of y obviate the omitted restrictions. □

Let us call a permutation v *decomposable* if $v = LR$ so that all entries of L are less than all entries of R, for some nonempty strings L and R. Let $v = LR$ be decomposable, and insert the entry $|L| + 1 = h$ immediately after L, increasing all entries of R by one. Call the obtained permutation pattern q'. Then q' is nothing else but the pattern $q = Lh$ in which we replace the entry h by the pattern $1R$. Therefore, Lemma 4.3 applies, and we have

$$S_n(v) \leq S_n(q') < (4c_q c_R)^n.$$

This leads to significant numerical improvements over the general result, particularly if L and R are also decomposable.

References

[1] N. Alon, E. Friedgut, On the number of permutations avoiding a given pattern, *Journal of Combinatorial Theory Series A,* **89** (2000), 133-140.

[2] R. Arratia, On the Stanley-Wilf Conjecture for the Number of Permutations Avoiding a Given Pattern. *Electronic J. Combin.*, 6 (1): N1, 1999.

[3] M. Bóna, A Walk Through Combinatorics, *World Scientific,* 2002.

[4] M. Bóna, The limit of a Stanley-Wilf sequence is not always an integer, and layered patterns beat monotone patterns. *Preprint, arXiv, math.CO/0403502.*

[5] M. Bóna, The Solution of a Conjecture of Stanley and Wilf for all layered patterns, *Journal of Combinatorial Theory Series A,* **85** (1999) 96-104.

[6] M. Bóna, Exact enumeration of 1342-avoiding permutations; A close link with labeled trees and planar maps, *Journal of Combinatorial Theory, Series A,* **80** (1997), 257-272.

[7] M. Bóna, Exact and Asymptotic Enumeration of Permutations with Subsequence Conditions, *PhD thesis, Massachusetts Institute of Technology*, 1997.

[8] Z. Füredi, P. Hajnal, Davenport-Schinzel theory of matrices. *Discrete Math.* 103 (1992), 233-251.

[9] M. Klazar, The Füredi-Hajnal conjecture implies the Stanley-Wilf conjecture. *Formal Power Series and Algebraic Combinatorics*, Springer, Berlin, 250-255, 2000.

[10] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Advances in Mathematics*, **41** (1981), 115-136.

[11] A. Marcus, G. Tardos, A linear bound on Excluded Permutation matrices, *preprint.*

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

A Generalization of $su(2)$

Brian Curtin

Abstract. *We consider the following generalization of* $su(2)$*. Let* $P(q, x, y, z)$ *denote the associative algebra over any field* K *generated by* A_1*,* A_2*,* A_3 *with relations* $[A_1, A_2]_q = xA_3 + yI + z(A_1 + A_2,$ $[A_2, A_3]_q = xA_1 + yI + z(A_2 + A_3)$*,* $[A_3, A_1]_q = xA_2 + yI + z(A_3 + A_1)$ *for some* q*,* x*,* y*,* $z \in K$*. Assume that* $q \neq 0$ *is either 1 or not a root of unity and that* $x \neq 0$*. We describe the multiplicity-free finite-dimensional representations of this generalized algebra, and we describe an action of the modular group on this algebra.*

Résumé. *Nous considérons la généralisation suivante de* $su(2)$*. Soit* $P(q, x, y, z)$ *l'algèbre associative avec des générateurs* A_1*,* A_2*,* A_3 *et rélations* $[A_1, A_2]_q = xA_3 + yI + z(A_1 + A_2,$ $[A_2, A_3]_q = xA_1 + yI + z(A_2 + A_3)$*,* $[A_3, A_1]_q = xA_2 + yI + z(A_3 + A_1)$ *pour* q*,* x*,* y*,* $z \in K$*. Supposez que* $q \neq 0$ *est 1 ou pas une racine de l'unité, et supposez que Nous décrivons l* $x \neq 0$*.es représentations fini-dimensionnelles sans multiplicité de cette algèbre généralisé, et Nous décrivons une action du groupe modulaire sur cette algèbre.*

1. Introduction

Recall that the special unitary Lie algebra $su(2)$ is the Lie algebra with basis S_1, S_2, S_3 and relations

$$[S_1, S_2] = \mathrm{i}S_3, \qquad [S_2, S_3] = \mathrm{i}S_1, \qquad [S_3, S_1] = \mathrm{i}S_2. \tag{1.1}$$

We generalize $su(2)$ (or rather its enveloping algebra) as follows.

Definition 1.1. Let $\mathbb{K}$ denote any field. Pick q, x, y, $z \in \mathbb{K}$. Let $\mathcal{P} = \mathcal{P}(q, x, y, z)$ be the associative algebra over $\mathbb{K}$ generated by three symbols S_1, S_2, S_3 subject to the relations

$$[S_1, S_2]_q = xS_3 + yI + z(S_1 + S_2), \tag{1.2}$$
$$[S_2, S_3]_q = xS_1 + yI + z(S_2 + S_3), \tag{1.3}$$
$$[S_3, S_1]_q = xS_2 + yI + z(S_3 + S_1), \tag{1.4}$$

where $[x, y]_q = xy - qyx$.

Like the relations of (1.1), the relations (1.2) – (1.4) express (q-)commutators as linear expressions in the three generators (the two in the commutator having the same coefficient) and have a cyclic symmetry.

We describe the multiplicity-free irreducible finite-dimensional representations of $\mathcal{P}(q, x, y, z)$ when $x \neq 0$ and q is some nonzero element of $\mathbb{K}$ which is not a root of unity, other than perhaps 1 itself. We need some notation. Fix a field $\mathbb{K}$ and a vector space V over $\mathbb{K}$ of finite nonnegative dimension. Let $\mathrm{End}(V)$ denote the vector space of all $\mathbb{K}$-linear transformations from V to V. A square matrix over $\mathbb{K}$ is said to be *tridiagonal*

1991 *Mathematics Subject Classification.* Primary 17B37; Secondary 33C45, 33D45, 16G99.
Key words and phrases. special unitary Lie algebra, Leonard pair, Leonard triple, modular Leonard triple.
Project sponsored by the National Security Agency under Grant Number MDA904–03-1-0020.

whenever every nonzero entry appears on the diagonal, the superdiagonal, or the subdiagonal. A tridiagonal matrix is *irreducible* whenever the entries on the sub- and superdiagonals are all nonzero.

Definition 1.2. Let A_1, A_2, A_3 denote an ordered triple of elements taken from $\mathrm{End}(V)$. We call this triple a *Leonard triple on* V whenever for each $A \in \{A_1, A_2, A_3\}$ there exists a basis of V with respect to which the matrix representing A is diagonal and the matrices representing the other two operators in the triple are irreducible tridiagonal.

By an *antiautomorphism* of $\mathrm{End}(V)$, we mean a $\mathbb{K}$-linear bijection $\tau : \mathrm{End}(V) \to \mathrm{End}(V)$ such that $\tau(XY) = \tau(Y)\tau(X)$ for all $X, Y \in \mathrm{End}(V)$.

Definition 1.3. Let A_1, A_2, A_3 denote a Leonard triple on V. Then this Leonard triple is said to be *modular* whenever for each $A \in \{A_1, A_2, A_3\}$ there exists an antiautomorphism of $\mathrm{End}(V)$ which fixes A and swaps the other two operators in the triple.

Our main result on the representations of $\mathcal{P}(q, x, y, z)$ is the following.

Theorem 1.4. *With reference to Definition 1.1, assume $x \neq 0$. Also assume that $q \neq 0$ is either 1 or not a root of unity. Let V denote an irreducible finite-dimensional module for $\mathcal{P}(q, x, y, z)$. Let $a_1 = S_1|_V$, $a_2 = S_2|_V$, $a_3 = S_3|_V$. Assume that a_1, a_2, a_3 are multiplicity-free. Then a_1, a_2, a_3 is a modular Leonard triple on V.*

The modular Leonard triples are completely characterized–we recall this characterization in Section 3. We conclude by showing that the modular group $\mathrm{PSL}_2(\mathbb{Z})$ acts on $\mathcal{P}(q, x, y, z)$ when $x \neq 0$.

2. Multiplicity-free representations of $\mathcal{P}$

We show that the representations of $\mathcal{P}(q, x, y, z)$ of interest are closely related to Leonard pairs. We begin by recalling the notion of a Leonard pair.

Definition 2.1. Let A_1, A_2 denote an ordered pair of elements taken from $\mathrm{End}(V)$. We call this pair a *Leonard pair on* V whenever for each $A \in \{A_1, A_2\}$ there exists an ordered basis of V with respect to which the matrix representing A is diagonal and the matrix representing the other member of the pair is irreducible tridiagonal.

We need the following criterion.

Theorem 2.2 (Vidunas and Terwilliger [**VT**])**.** *Let V denote a vector space over $\mathbb{K}$ of finite positive dimension. Let A, A_2 denote an ordered pair of elements of* $\mathrm{End}(V)$ *linear operators in* $\mathrm{End}(V)$. *Assume that*

(1) *A_1 and A_2 are multiplicity-free;*
(2) *V is irreducible as an (A_1, A_2)-module;*
(3) *there exist β, γ, γ^*, ρ, ρ^*, ω, η, $\eta^* \in \mathbb{K}$ such that*

$$A_1^2A_2 - \beta A_1A_2A_1 + A_2A_1^2 - \gamma(A_1A_2 + A_2A_1) - \rho A_2 = \gamma^* A_1^2 + \omega A_1 + \eta I, \tag{2.1}$$

$$A_2{}^2A_1 - \beta A_2A_1A_2 + A_1A_2{}^2 - \gamma^*(A_2A_1 + A_1A_2) - \rho^* A_1 = \gamma A_2{}^2 + \omega A_2 + \eta^* I; \tag{2.2}$$

(4) *no q satisfying $q + q^{-1} = \beta$ is a root of unity.*

Then A_1, A_2 is a Leonard pair on V.

Theorem 2.3. *With reference to Definition 1.1, assume $x \neq 0$. Then any two of S_1, S_2, S_3 satisfy (2.1) and (2.2) with*

$$\begin{aligned}
\beta &= q + 1/q, \\
\gamma = \gamma^* &= z(q-1)/q, \\
\rho = \rho^* &= (z^2 - x^2)/q, \\
\omega = \omega^* &= (y(q-1) + z(z-x))/q, \\
\eta = \eta^* &= y(z-x)/q.
\end{aligned}$$

Proof. Each of S_1, S_2, S_3 appears linearly with coefficient x in one of equations one of (1.2)–(1.4). Solve for, say, S_3 in (1.2), and eliminate it in (1.3) and (1.4). □

Lemma 2.4. *With the notation and assumptions of Theorem 1.4, any two of a_1, a_2, and a_3 form a Leonard pair.*

Proof. Observe that V is irreducible as, say, an (a_1, a_2)-module since V is irreducible as a $\mathcal{P}(q, x, y, z)$ and a_3 is expressed using a_1 and a_2. The result follows from Theorems 2.2 and 2.3. □

It turns out that the representations of $\mathcal{P}(q, x, y, z)$ of interest correspond to a special extension of a Leonard pair.

Proof of Theorem 1.4. (sketch) By Lemma 2.4, any two of a_1, a_2, a_3 form a Leonard pair. Thus by Definition 2.1 there is a basis of V with respect to which the matrix representing, say, a_1 is irreducible tridiagonal and the matrix representing a_2 is diagonal. Substituting these forms into (1.2) gives that the matrix representing a_3 is also irreducible tridiagonal. Thus a_1, a_2, a_3 is a Leonard triple. It turns out that all Leonard pairs in Lemma 2.4 are isomorphic. (This follows from the fact that they all satisfy the same Askey-Wilson relations and some facts about canonical forms of a Leonard pair [**T4**]). Composing the antiautomorphism of $\mathrm{End}(V)$ which fixes a_1 and a_2 and the automorphism which swaps a_1 and a_2 gives an antiautomorphism which swaps a_1 and a_2. Applying this map to (1.2) gives that it fixes a_3. □

We conclude this section with some comments on Leonard pairs. Leonard pairs were introduced by P. Terwilliger [**T1, T3**] as an algebraic abstraction of work of D. Leonard concerning the sequences of orthogonal polynomials with discrete support for which there is a dual sequence of orthogonal polynomials. [**Len1, Len2**] (cf. [**BI**]). Leonard characterized these orthogonal polynomials in terms of hypergeometric series. This result is analogous to Askey and Wilson's characterization of similar orthogonal polynomials with continuous support [**AW1, AW2**] (cf. [**KS**]). The reference [**T5**] describes a bijective correspondence between the isomorphism classes of Leonard pairs and the appropriate orthogonal polynomials. In particular, results concerning Leonard pairs can be viewed as results concerning such orthogonal polynomials. This connection is further developed in [**T6**]. Relations (2.1) and (2.2) are called the *Askey-Wilson relations*. They were introduced by Zhedanov et. al. [**GLZ, Z**] in connection with the quadratic Askey-Wilson algebra.

3. The modular Leonard triples

We now recall a characterization of the modular Leonard triples [**C**]. We do so by first describing three examples of modular Leonard triples in Lemmas 3.1, 3.2, and 3.3, and then describing how, up to isomorphism, they are the only examples. We use the following conventions throughout. Given any square matrix X of order n with entries in $\mathbb{K}$, we view X as a linear operator on $\mathbb{K}^n$, acting by $v \mapsto Xv$. Let d denote a nonnegative integer. Write

$$\begin{aligned} A_1 &= \mathrm{tridiag}\begin{pmatrix} b_0 & b_1 & \cdots & b_{d-1} & * \\ a_0 & a_1 & \cdots & a_{d-1} & a_d \\ * & c_1 & \cdots & c_{d-1} & c_d \end{pmatrix}, \\ A_2 &= \mathrm{diag}(\theta_0, \theta_1, \ldots, \theta_d), \\ A_3 &= \mathrm{tridiag}\begin{pmatrix} b_0\nu_1 & b_1\nu_2 & \cdots & b_{d-1}\nu_d & * \\ a_0 & a_1 & \cdots & a_{d-1} & a_d \\ * & c_1/\nu_1 & \cdots & c_{d-1}/\nu_{d-1} & c_d/\nu_d \end{pmatrix}. \end{aligned}$$

Lemma 3.1. *([**C**]) Set*

$$\begin{aligned}
\nu_i &= \nu q^{i-1} \qquad (1 \le i \le d),\\
\theta_i &= \theta_0 + h(1-q^i)(1-\nu^2 q^{i-1})q^{-i} \qquad (0 \le i \le d),\\
b_0 &= -\frac{h(1-q^d)(1+\nu^3 q^{d-1})}{q^d(1-\nu)},\\
b_i &= -\frac{h(1-q^{d-i})(1-\nu^2 q^{i-1})(1+\nu^3 q^{d+i-1})}{q^{d-i}(1-\nu q^i)(1-\nu^2 q^{2i-1})} \qquad (1 \le i \le d-1),\\
c_i &= \frac{h\nu(1-q^i)(1+\nu q^{d-i})(1-\nu^2 q^{d+i-1})}{q^{d-i+1}(1-\nu q^{i-1})(1-\nu^2 q^{2i-1})} \qquad (1 \le i \le d-1),\\
c_d &= \frac{h\nu(1-q^d)(1+\nu)}{q^(1-\nu q^{d-1})},\\
a_i &= \theta_0 - b_i - c_i \qquad (0 \le i \le d)\ (c_0 = 0,\, b_d = 0)
\end{aligned}$$

for some scalars θ_0, h, ν, q in $\mathbb{K}$ such that $h\nu q \ne 0$, $q^i \ne 1$ $(1 \le i \le d)$, $\nu^3 q^{2d-1-i} \ne -1$ $(1 \le i \le d)$, and $\nu^2 q^i \ne 1$ $(0 \le i \le 2d-2)$. Then A_1, A_2, A_3 is a modular Leonard triple on $\mathbb{K}^{d+1}$.

Lemma 3.2. *([**C**]) Assume* $\operatorname{char}\mathbb{K}$ *is 0 or an odd prime greater than d. Set*

$$\begin{aligned}
\nu_i &= -1 \qquad (1 \le i \le d),\\
\theta_i &= \theta_0 + hi(i+1+s) \qquad (0 \le i \le d),\\
b_0 &= \frac{-hd(3s+2d+4)}{4},\\
b_i &= \frac{h(i+1+s)(d-i)(2i+3s+2d+4)}{4(2i+1+s)} \qquad (1 \le i \le d-1),\\
c_i &= \frac{hi(i+s+d+1)(2i-s-2d-2)}{4(2i+1+s)} \qquad (1 \le i \le d-1),\\
c_d &= \frac{-hd(s+2)}{4},\\
a_i &= \theta_0 - b_i - c_i \qquad (0 \le i \le d)\ (c_0 = 0,\, b_d = 0)
\end{aligned}$$

for some scalars θ_0, h, s in $\mathbb{K}$ such that $h \ne 0$, $s \ne -i$ $(2 \le i \le 2d)$, and $3s \ne -2i$ $(d+2 \le i \le 2d+1)$. Then A_1, A_2, A_3 is a modular Leonard triple on $\mathbb{K}^{d+1}$.

Lemma 3.3. *([**C**]) Assume* $\operatorname{char}\mathbb{K} = 0$ *or* $\operatorname{char}\mathbb{K} > d$. *Set*

$$\begin{aligned}
\nu_i &= \nu \qquad (1 \le i \le d),\\
\theta_i &= \theta_0 + hi \qquad (0 \le i \le d),\\
b_i &= -\frac{h(d-i)(1-\nu+\nu^2)}{(1-\nu)^2} \qquad (0 \le i \le d-1),\\
c_i &= \frac{hi\nu}{(1-\nu)^2} \qquad (1 \le i \le d),\\
a_i &= \theta_0 - b_i - c_i \qquad (0 \le i \le d)\ (c_0 = 0,\, b_d = 0)
\end{aligned}$$

for some scalars θ_0, h, ν in $\mathbb{K}$ such that $h\nu \ne 0$, $\nu \ne 1$, and $1 - \nu + \nu^2 \ne 0$. Then A_1, A_2, A_3 is a modular Leonard triple on $\mathbb{K}^{d+1}$.

Definition 3.4. Let V denote a vector space over $\mathbb{K}$ of finite positive dimension. Let A_1, A_2, A_3 denote a modular Leonard triple on V. We say that the triple A_1, A_2, A_3 is of *type I*, *type II*, or *type III*, respectively, whenever there exists a basis of V with respect to which the matrices representing A_1, A_2, A_3 are as in Lemma 3.1, Lemma 3.2, or Lemma 3.3, respectively.

Theorem 3.5 ([**C**]). *Let V denote a vector space over $\mathbb{K}$ of finite positive dimension. Let A_1, A_2, A_3 denote a modular Leonard triple on V. Then A_1, A_2, A_3 is of type I, type II, or type III.*

Theorem 3.6. *Let A_1, A_2, A_3 denote a modular Leonard triple on V. Then there are scalars q, x, y, z in $\mathbb{K}$ with $x \neq 0$ such that (1.2)–(1.4) hold.*

Proof. Direct verification using the above classification of modular Leonard triples. □

4. A modular group action

We describe an action of the modular group $\mathrm{PSL}_2(\mathbb{Z})$ on $\mathcal{P}(q,x,y,z)$. This modular group action was first observed for the modular Leonard triples, hence their name. We begin with describing some antiautomorphisms for $\mathcal{P}(q,x,y,z)$.

Lemma 4.1. *With reference to Definition 1.1, assume $x \neq 0$. Then for any $T \in \{S_1, S_2, S_3\}$, there exists an antiautmorphism of $\mathcal{P}(q,x,y,z)$ which fixes T and swaps the other two generators.*

Proof. Let $\mu : \mathcal{P} \to \mathcal{P}$ denote a linear map which reverses the order of multiplication and swaps S_1 and S_2. Then μ fixes the q-commutator in (1.2). On the right-hand side of (1.2) the linear terms involving I and $S_1 + S_2$ are fixed, so S_3 is fixed by such a map. Observe that μ is indeed an antiautomorphism of $\mathcal{P}$. □

Lemma 4.2. *With reference to Definition 1.1, assume $x \neq 0$. Then for any $T \in \{S_1, S_2, S_3\}$, there exists an antiautmorphism of $\mathcal{P}(q,x,y,z)$ which fixes the elements of $\{S_1, S_2, S_3\} \backslash T$.*

Proof. Let $\alpha : \mathcal{P} \to \mathcal{P}$ denote a linear map which reverses the order of multiplication and swaps S_1 and S_2. Applying α to (1.2) gives an expression for $\alpha(S_3)$. Essentially the same computation as was performed in Theorem 2.3 shows that α is indeed an antiautomorphism of $\mathcal{P}$. □

Recall that $\mathrm{PSL}_2(\mathbb{Z})$ has presentation $\langle s,\, t \,|\, s^2 = 1,\, t^3 = 1 \rangle$.

Lemma 4.3. *With reference to Definition 1.1, assume $x \neq 0$.*

(1) *Let σ denote the composition of the antiautomorphisms of $\mathcal{P}$ which respectively fix and swap S_1 and S_2. Then $\sigma^2 = I$.*
(2) *Let τ denote the composition of the antiautomorphisms of $\mathcal{P}$ which respectively swap S_1 and S_2 and swap S_2 and S_3. Then $\tau^3 = I$.*

In particular, $\mathrm{PSL}_2(\mathbb{Z})$ *acts on $\mathcal{P}$ as a group of automorphisms.*

Proof. It is easy to verify from their constructions that τ sends S_1 to S_3, S_2 to S_1, and S_3 to S_2, and that σ swaps S_1 and S_2. The result follows. □

References

[A] T.M. Apostol, *The Modular Group and Modular Functions.* Ch.Ê2 in Modular Functions and Dirichlet Series in Number Theory, 2nd ed, pp.Ê17 and pp. 26-46, Springer-Verlag, New York, 1997.

[AW1] R. Askey and J. Wilson, *A set of orthogonal polynomials that generalize the Racah coefficients or $6-j$ symbols*, SIAM J. Math. Anal. **10** (1979), pp. 1008–1016.

[AW2] R. Askey and J. Wilson, *Some basic hypergeometric orthogonal polynomials that generalize the Jacobi polynomials*, Mem. Amer. Math. Soc. **54** (1985).

[BI] E. Bannai and T. Ito, *Algebraic Combinatorics I*, Benjamin/Cummings, Menlo Park, 1984.

[BB] J.M. Borwein and P.B. Borwein, *Pi & the AGM: A Study in Analytic Number Theory and Computational Complexity*, p.Ê113, Wiley, New York, 1987.

[C] B. Curtin, *Modular Leonard triples*, preprint.

[GLZ] Ya.A. Granovskii, A.S. Lutzennko and A.S. Zhedanov, *Mutual integrability, quadratic algebras, and dynamical symmetry*, Ann. Phys **217** (1992), pp. 1–20.

[KS] Koekoek and Suwartouw, *The Askey scheme of hypergeometric orthogonal polynomials and its q-analog*, Reports of the Faculty of Technical Mathematics and Informatics, vol. 98-17, Delft, Netherlands, 1998.

[Lan] R.H. Landau *Quantum Mechanics II: A Second Course in Quantum Theory*, 2nd ed. Wiley, New York, 1996.

[Len1] D. Leonard, *Orthogonal polynomials, duality, and association schemes*, SIAM J. Math. Anal. **13** (1982), pp. 656–663.

[Len2] D. Leonard, *Parameters of association schemes that are both P- and Q-polynomial*, J. Combin. Theory Ser. A **36** (1984), pp. 355–363.

[T1] P. Terwilliger, *The subconstituent algebra of an association scheme*, J. Alg. Combin. Part I: **1** (1992), pp. 363–388; Part II: **2** (1993), pp. 73–103; Part III: **2** (1993), pp. 177–210.

[T2] P. Terwilliger, *Leonard pairs from 24 points of view*, Rocky Mountain Journal of Mathematics, **32** (2002), pp. 827–888.

[T3] P. Terwilliger, *Two linear transformations each tridiagonal with respect to an eigenbasis of the other*, Linear Algebra Appl. **330** (2001), pp. 149–203.

[T4] P. Terwilliger, *Two linear transformations each tridiagonal with respect to an eigenbasis of the other: The TD-D and LB-UB canonical forms*, preprint.

[T5] P. Terwilliger, *Two linear transformations each tridiagonal with respect to an eigenbasis of the other: Comments on the parameter array*, preprint.

[T6] P. Terwilliger, *Leonard pairs and the q-Racah polynomials*, preprint.

[VT] R. Vidunas and P. Terwilliger *Leonard pairs and Askey-Wilson relations*, preprint.

[Z] A.S. Zhedanov, *Hidden symmetry of Askey-Wilson polynomials*, Teoret. Mat. Fiz., **89** (1991), pp. 190–204.

Department of Mathematics, University of South Florida, Tampa, FL USA, 33620
E-mail address: bcurtin@math.usf.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Two New Criteria for Comparison in the Bruhat Order

Brian Drake, Sean Gerrish, and Mark Skandera

Abstract. *We give two new criteria by which pairs of permutations may be compared in defining the Bruhat order (of type A). One criterion utilizes totally nonnegative polynomials and the other utilizes Schur functions.*

RÉSUMÉ. *Nous donons deux critères nouveaux avec lesquels on peut comparer couples de permutations en definant l'order de Bruhat (de type A). Un critère utilise les polynômes totallement nonnegatifs et l'autre utilise les fonctions symétriques de Schur.*

1. Main

The Bruhat order on S_n is often defined by comparing 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. [**7**, 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 1.1. *Let π and σ be 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 [**1**], [**2**], [**3**], [**6**, pp. 173-177], [**8**] for more criteria.) Using Theorem 1.1 and our defining criterion we will state and prove the validity of two more criteria.

Our first new criterion defines the Bruhat order in terms of totally nonnegative polynomials. A matrix A is called *totally nonnegative* (TNN) if the determinant of each square submatrix of A is nonnegative. (See e.g. [**5**].) A polynomial in n^2 variables $f(x_{1,1}, \ldots, x_{n,n})$ is called *totally nonnegative* (TNN) if for each $n \times n$ TNN matrix $A = (a_{i,j})$ the number $f(a_{1,1}, \ldots, a_{n,n})$ is nonnegative. Some recent interest in TNN polynomials is motivated by problems in the study of canonical bases. (See [**10**].)

Theorem 1.2. *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 the polynomial*

$$x_{1,\pi(1)} \cdots x_{n,\pi(n)} - x_{1,\sigma(1)} \cdots x_{n,\sigma(n)} \tag{1.1}$$

is totally nonnegative.

PROOF. ($\Rightarrow$) If $\pi = \sigma$ then (1.1) is obviously TNN. Suppose that π is less than σ in the Bruhat order. If π differs from σ by a single transposition (i, j) with $i < j$, then we have $\pi(i) = \sigma(j) < \pi(j) = \sigma(i)$, and

1991 *Mathematics Subject Classification.* 15A15, 05E05.
Key words and phrases. total nonnegativity, Schur nonnegativity, Bruhat order.

the polynomial (1.1) is equal to

$$\frac{x_{1,\pi(1)}\cdots x_{n,\pi(n)}}{x_{i,\pi(i)}x_{j,\pi(j)}}(x_{i,\pi(i)}x_{j,\pi(j)} - x_{i,\pi(j)}x_{j,\pi(i)}) \tag{1.2}$$

which is clearly TNN. If π differs from σ by a sequence of transpositions, then the polynomial (1.1) is equal to a sum of polynomials of the form (1.2) and again is TNN.

($\Leftarrow$) Suppose that π is not less than or equal to σ in the Bruhat order. By Theorem 1.1 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. Now define the matrix $A = (a_{i,j})$ by

$$a_{i,j} = \begin{cases} 2 & \text{if } i \leq k \text{ and } j \leq \ell, \\ 1 & \text{otherwise.} \end{cases}$$

It is easy to see that A is TNN, since all square submatrices of A have determinant equal to $0, 1$, or 2. Applying the polynomial (1.1) to A we have

$$a_{1,\pi(1)}\cdots a_{n,\pi(n)} - a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} = -2^q,$$

and the polynomial (1.1) is not TNN. □

Our second new criterion defines the Bruhat order in terms of Schur functions. (See [**9**, Ch. 7] for definitions.) 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*. Let us define a polynomial in n^2 variables $f(x_{1,1}, \ldots, x_{n,n})$ to be *Schur nonnegative* (SNN) if for each $n \times n$ Jacobi-Trudi matrix $A = (a_{i,j})$ the symmetric function $f(a_{1,1}, \ldots, a_{n,n})$ is equal to a nonnegative linear combination of Schur functions. Some recent interest in SNN polynomials is motivated by problems in algebraic geometry [**4**, Conj. 2.8, Conj. 5.1].

Theorem 1.3. *Let π and σ be permutations in S_n. Then π is less than or equal to σ in the Bruhat order if and only if the polynomial*

$$x_{1,\pi(1)}\cdots x_{n,\pi(n)} - x_{1,\sigma(1)}\cdots x_{n,\sigma(n)} \tag{1.3}$$

is Schur nonnegative.

Proof. ($\Rightarrow$) If $\pi = \sigma$ then (1.3) is obviously SNN. Let A be an $n \times n$ Jacobi-Trudi matrix and suppose that π is less than σ in the Bruhat order. If π differs from σ by a single transposition (i, j), then for some partition ν and some k, ℓ, m ($\ell, m > 0$), the evaluation of the polynomial (1.3) at A is equal to

$$h_\nu(h_{k+\ell}h_{k+m} - h_{k+\ell+m}h_k), \tag{1.4}$$

and (1.3) is clearly SNN. If π differs from σ by a sequence of transpositions, then the evaluation of (1.3) at A is equal to a sum of polynomials of the form (1.4) and again (1.3) is SNN.

($\Leftarrow$) Suppose that π is not less than or equal to σ in the Bruhat order. By Theorem 1.1 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. Now define the nonnegative number $r = (k-q)(n+k-\ell-2)$ and consider the Jacobi-Trudi matrix B defined by the skew shape $(n-1+2r)^k(n-1+r)^{n-k}/r^\ell$,

$$B = \begin{bmatrix} h_{n-1+r} & \cdots & h_{n+\ell-2+r} & h_{n+\ell-1+2r} & \cdots & h_{2n-2+2r} \\ \vdots & & \vdots & \vdots & & \vdots \\ h_{n-k+r} & \cdots & h_{n-k+\ell-1+r} & h_{n-k+\ell+2r} & \cdots & h_{2n-k-1+2r} \\ h_{n-k-1} & \cdots & h_{n-k+\ell-2} & h_{n-k+\ell-1+r} & \cdots & h_{2n-k-2+r} \\ \vdots & & \vdots & \vdots & & \vdots \\ h_0 & \cdots & h_{\ell-1} & h_{\ell+r} & \cdots & h_{n-1+r} \end{bmatrix}.$$

The polynomial (1.3) applied to B may be expressed as $h_\lambda - h_\mu$ for some appropriate partitions λ, μ depending on π, σ, respectively. We claim that λ is incomparable to or greater than μ in the dominance order. Since $M(\pi)_{[k],[\ell+1,n]}$ contains $k-q$ ones we have that

$$\lambda_1 + \cdots + \lambda_{k-q} \geq (k-q)(n-k+\ell+2r). \tag{1.5}$$

Similarly, we have

$$\mu_1 + \cdots + \mu_{k-q} \leq (k-q-1)(2n-2+2r) + \max\{n+\ell-2+r, 2n-k-2+r\}. \tag{1.6}$$

Subtracting (1.6) from (1.5), we obtain

$$(\lambda_1 + \cdots + \lambda_{k-q}) - (\mu_1 + \cdots + \mu_{k-q}) \geq n - \max\{\ell, n-k\} > 0,$$

as desired.

Recall that the Schur expansion of h_μ is

$$h_\mu = s_\mu + \sum_{\nu > \mu} K_{\nu,\mu} s_\nu,$$

where the comparison of partitions $\nu > \mu$ is in the dominance order and the nonnegative *Kostka numbers* $K_{\nu,\mu}$ count semistandard Young tableaux of shape ν and content μ. (See e.g. [**9**, Prop. 7.10.5, Cor. 7.12.4].) It follows that the coefficient of s_μ in the Schur expansion of $h_\lambda - h_\mu$ is -1 and the polynomial (1.3) is not SNN. □

The authors are grateful to Sergey Fomin, Zachary Pavlov, Alex Postnikov, Christophe Reutenauer, Brendon Rhoades, Richard Stanley, John Stembridge, and referees for helpful conversations.

References

[1] 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.

[2] A. Björner and F. Brenti. An improved tableau criterion for Bruhat order. *Electron. J. Combin.*, **3** (1996). Research paper 22, 5 pp. (electronic).

[3] 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.

[4] S. Fomin, W. Fulton, C. Li, and Y. Poon. Eigenvalues, singular values, and Littlewood-Richardson coefficients, 2003. Preprint math.AG/03013078 on ArXiv.

[5] S. Fomin and A. Zelevinsky. Total positivity: Tests and parametrizations. *Math. Intelligencer*, **22** (2000) pp. 23–33.

[6] 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.

[7] J. E. Humphreys. *Reflection groups and Coxeter groups*. Cambridge University Press, 1990.

[8] A. Lascoux and M. P. Schützenberger. Treillis et bases des groupes de Coxeter. *Electron. J. Combin.*, **3** (1996). Research paper 27, 35 pp. (electronic).

[9] R. Stanley. *Enumerative Combinatorics*, vol. 2. Cambridge University Press, Cambridge, 1999.

[10] 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.

Dept. of Mathematics, Brandeis University, MS 050, P.O. Box 9110, Waltham, MA 02454
E-mail address: `bdrake@math.brandeis.edu`

Dept. of Mathematics, University of Michigan, 2074 East Hall, Ann Arbor, MI 48109-1109
E-mail address: `sgerrish@umich.edu`

Dept. of Mathematics, Dartmouth College, 6188 Bradley Hall, Hanover, NH 03755-3551
URL: `http://www.math.dartmouth/~skan`
E-mail address: `mark.skandera@dartmouth.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Restricted Motzkin permutations, Motzkin paths, continued fractions, and Chebyshev polynomials

Sergi Elizalde and Toufik Mansour

Abstract. *We say that a permutation π is a* Motzkin permutation *if it avoids* 132 *and there do not exist $a < b$ such that $\pi_a < \pi_b < \pi_{b+1}$. We study the distribution of several statistics on Motzkin permutations, including the length of the longest increasing and decreasing subsequences and the number of rises and descents. We also enumerate Motzkin permutations with additional restrictions and study the distribution of occurrences of fairly general patterns in this class of permutations.*

Résumé. *On dit qu'une permutation π est une* permutation de Motzkin *si elle évite le motif* 132 *et s'il n'existe pas $a < b$ tels que $\pi_a < \pi_b < \pi_{b+1}$. Nous étudions la distribution de plusieurs statistiques sur permutations de Motzkin, entre autres la longueur des sous-suites croissantes et décroissantes les plus longues et le nombre de montées et descentes. Nous énumérons aussi des permutations de Motzkin avec des contraintes supplémentaires et nous étudions la distribution du nombre d'occurrences de motifs assez généraux dans cette classe de permutations.*

1. Introduction

1.1. Background. Let $\alpha \in S_n$ and $\tau \in S_k$ be two permutations. We say that α *contains* τ if there exists a subsequence $1 \le i_1 < i_2 < \cdots < i_k \le n$ such that $(\alpha_{i_1}, \ldots, \alpha_{i_k})$ is order-isomorphic to τ; in such a context τ is usually called a *pattern*. We say that α *avoids* τ, or is τ*-avoiding*, if such a subsequence does not exist. The set of all τ-avoiding permutations in S_n is denoted $S_n(\tau)$. For an arbitrary finite collection of patterns T, we say that α avoids T if α avoids any $\tau \in T$; the corresponding subset of S_n is denoted $S_n(T)$.

While the case of permutations avoiding a single pattern has attracted much attention, the case of multiple pattern avoidance remains less investigated. In particular, it is natural, as the next step, to consider permutations avoiding pairs of patterns τ_1, τ_2. This problem was solved completely for $\tau_1, \tau_2 \in S_3$ (see [**24**]) and for $\tau_1 \in S_3$ and $\tau_2 \in S_4$ (see [**25**]). Several recent papers [**5, 15, 11, 16, 17, 18**] deal with the case $\tau_1 \in S_3$, $\tau_2 \in S_k$ for various pairs τ_1, τ_2. Another natural question is to study permutations avoiding τ_1 and containing τ_2 exactly r times. Such a problem for certain $\tau_1, \tau_2 \in S_3$ and $r = 1$ was investigated in [**20**], and for certain $\tau_1 \in S_3$, $\tau_2 \in S_k$ in [**22, 15, 11**]. The tools involved in these papers include Catalan numbers, Chebyshev polynomials, and continued fractions.

In [**1**] Babson and Steingrímsson introduced *generalized patterns* that allow the requirement that two adjacent letters in a pattern must be adjacent in the permutation. In this context, we write a classical pattern with dashes between any two adjacent letters of the pattern (for example, 1423 as 1-4-2-3). If we omit the dash between two letters, we mean that for it to be an occurrence in a permutation π, the corresponding letters of π have to be adjacent. For example, in an occurrence of the pattern 12-3-4 in a permutation π, the letters in π that correspond to 1 and 2 are adjacent. For instance, the permutation $\pi = 3542617$ has only one occurrence of the pattern 12-3-4, namely the subsequence 3567, whereas π has two occurrences of the pattern 1-2-3-4, namely the subsequences 3567 and 3467. Claesson [**3**] presented a complete solution for the number of permutations avoiding any single 3-letter generalized pattern with exactly one adjacent pair of letters. Elizalde and Noy [**8**] studied some cases of avoidance of patterns where all letters have to occur in consecutive positions. Claesson and Mansour [**4**] (see also [**12, 13, 14**]) presented a complete solution for the number of permutations avoiding any pair of 3-letter generalized patterns with exactly one adjacent pair of letters. Besides, Kitaev [**9**] investigated simultaneous avoidance of two or more 3-letter generalized patterns without internal dashes.

A remark about notation: throughout the paper, a pattern represented with no dashes will always denote a classical pattern (i.e., with no requirement about elements being consecutive). All the generalized patterns that we will consider will have at least one dash.

1.2. Basic tools. *Catalan numbers* are defined by $C_n = \frac{1}{n+1}\binom{2n}{n}$ for all $n \geq 0$. The generating function for the Catalan numbers is given by $C(x) = \frac{1-\sqrt{1-4x}}{2x}$.

Chebyshev polynomials of the second kind (in what follows just Chebyshev polynomials) are defined by $U_r(\cos\theta) = \frac{\sin(r+1)\theta}{\sin\theta}$ for $r \geq 0$. Clearly, $U_r(t)$ is a polynomial of degree r in t with integer coefficients, and the following recurrence holds:

$$U_0(t) = 1,\ U_1(t) = 2t, \text{ and } U_r(t) = 2tU_{r-1}(t) - U_{r-2}(t) \text{ for all } r \geq 2. \tag{1.1}$$

The same recurrence is used to define $U_r(t)$ for $r < 0$ (for example, $U_{-1}(t) = 0$ and $U_{-2}(t) = -1$). Chebyshev polynomials were invented for the needs of approximation theory, but are also widely used in various other branches of mathematics, including algebra, combinatorics, and number theory (see [**21**]). Apparently, the relation between restricted permutations and Chebyshev polynomials was discovered for the first time by Chow and West in [**5**], and later by Mansour and Vainshtein [**15, 16, 17, 18**], Krattenthaler [**11**].

Recall that a *Dyck path* of length $2n$ is a lattice path in $\mathbb{Z}^2$ between $(0,0)$ and $(2n,0)$ consisting of up-steps $(1,1)$ and down-steps $(1,-1)$ which never goes below the x-axis. Denote by $\overrightarrow{\mathbb{G}}_n$ the set of Dyck paths of length $2n$, and by $\overrightarrow{\mathbb{G}} = \bigcup_{n\geq 0}\overrightarrow{\mathbb{G}}_n$ the class of all Dyck paths. If $D \in \overrightarrow{\mathbb{G}}_n$, we will write $|D| = n$. Recall that a *Motzkin path* of length n is a lattice path in $\mathbb{Z}^2$ between $(0,0)$ and $(n,0)$ consisting of up-steps $(1,1)$, down-steps $(1,-1)$ and horizontal steps $(1,0)$ which never goes below the x-axis. Denote by $\mathcal{M}_n$ the set of *Motzkin paths* with n steps, and let $\mathcal{M} = \bigcup_{n\geq 0}\mathcal{M}_n$. We will write $|M| = n$ if $M \in \mathcal{M}_n$. Sometimes it will be convenient to encode each up-step by a letter u, each down-step by d, and each horizontal step by h. Denote by $M_n = |\mathcal{M}_n|$ the n-th *Motzkin number*. The generating function for these numbers is $M(x) = \frac{1-x-\sqrt{1-2x-3x^2}}{2x^2}$.

Define a *Motzkin permutation* π to be a 132-avoiding permutation in which there do not exist indices $a < b$ such that $\pi_a < \pi_b < \pi_{b+1}$. In such a context, π_a, π_b, π_{b+1} is called an occurrence of the pattern 1-23 (for instance, see [**3**]). For example, there are exactly 4 Motzkin permutations of length 3, namely, 213, 231, 312, and 321. The set of all Motzkin permutations in S_n we denote by $\mathfrak{M}_n$. The main reason for the term "Motzkin permutation" is that $|\mathfrak{M}_n| = M_n$, as we will see in Section 2.

It follows from the definition that the set $\mathfrak{M}_n$ is the same as the set of 132-avoiding permutations $\pi \in S_n$ where there is no a such that $\pi_a < \pi_{a+1} < \pi_{a+2}$. Indeed, assume that $\pi \in S_n(132)$ has an occurrence of 1-23, say $\pi_a < \pi_b < \pi_{b+1}$ with $a < b$. Now, if $\pi_{b-1} > \pi_b$, then π would have an occurrence of 132, namely $\pi_a\pi_{b-1}\pi_{b+1}$. Therefore, $\pi_{b-1} < \pi_b < \pi_{b+1}$, so π has three consecutive increasing elements.

For any subset $A \subseteq S_n$ and any pattern α, define $A(\alpha) := A \cap S_n(\alpha)$. For example, $\mathfrak{M}_n(\alpha)$ denotes the set of Motzkin permutations of length n that avoid α.

1.3. Organization of the paper. In Section 2 we exhibit a bijection between the set of Motzkin permutations and the set of Motzkin paths. Then we use it to obtain generating functions of Motzkin permutations with respect to the length of the longest decreasing and increasing subsequences together with the number of rises. The section ends with another application of the bijection, to the enumeration of fixed points in permutations avoiding simultaneously 231 and 32-1.

In Section 3 we consider additional restrictions on Motzkin permutations. Using a block decomposition, we enumerate Motzkin permutations avoiding the pattern $12\dots k$, and we find the distribution of occurrences of this pattern in Motzkin permutations. Then we obtain generating functions for Motzkin permutations avoiding patterns of more general shape. We conclude the section considering two classes of generalized patterns (as described above), and we study its distribution in Motzkin permutations.

2. The bijection $\Theta : \mathfrak{M}_n \longrightarrow \mathcal{M}_n$

In this section we establish a bijection Θ between Motzkin permutations and Motzkin paths. This bijection allows us to give the distribution of some interesting statistics on the set of Motzkin permutations.

2.1. The bijection Θ. We can give a bijection Θ between $\mathfrak{M}_n$ and $\mathcal{M}_n$. For that we use first the following bijection φ from $S_n(132)$ to $\overrightarrow{\mathbb{G}}_n$, which is essentially due to Krattenthaler [**11**]. Consider $\pi \in S_n(132)$ given as an $n \times n$ array with crosses in the squares (i, π_i). Take the path with *up* and *right* steps that goes from the lower-left corner to the upper-right corner, leaving all the crosses to the right, and staying always as close to the diagonal connecting these two corners as possible. Then $\varphi(\pi)$ is the Dyck path obtained from this path by reading an up-step every time the path goes up and a down-step every time it goes right. Figure 1 shows an example when $\pi = 67435281$.

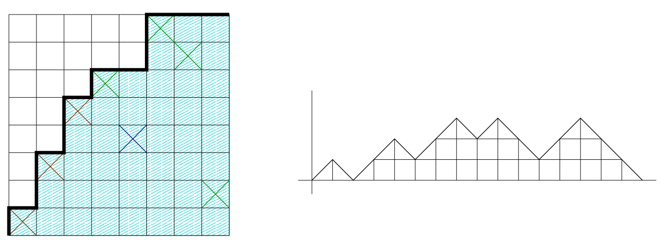


FIGURE 1. The bijection φ.

There is an easy way to recover π from $\varphi(\pi)$. Assume we are given the path from the lower-left corner to the upper-right corner or the array. Row by row, put a cross in the leftmost square to the right of this path such that there is exactly one cross in each column. This gives us π back.

One can see that $\pi \in S_n(132)$ avoids 1-23 if and only if the Dyck path $\varphi(\pi)$ does not contain three consecutive up-steps (a *triple rise*). Indeed, assume that $\varphi(\pi)$ has three consecutive up-steps. Then, the path from the lower-left corner to the upper-right corner of the array has three consecutive vertical steps. The crosses in the corresponding three rows give three consecutive increasing elements in π (this follows from the definition of the inverse of φ), and hence an occurrence of 1-23.

Reciprocally, assume now that π has an occurrence of 1-23. The path from the lower-left to the upper-right corner of the array of π must have two consecutive vertical steps in the rows of the crosses corresponding to '2' and '3'. But if $\varphi(\pi)$ has no triple rise, the next step of this path must be horizontal, and the cross corresponding to '2' must be right below it. But then all the crosses above this cross are to the right of it, which contradicts the fact that this was an occurrence of 1-23.

Denote by $\mathcal{E}_n$ the set of Dyck paths of length $2n$ with no triple rise. We have given a bijection between $\mathfrak{M}_n$ and $\mathcal{E}_n$. The second step is to exhibit a bijection between $\mathcal{E}_n$ and $\mathcal{M}_n$, so that Θ will be defined as the composition of the two bijections. Given $D \in \mathcal{E}_n$, divide it in n blocks, splitting after each down-step. Since

D has no triple rises, each block is of one of these three forms: uud, ud, d. From left to right, transform the blocks according to the rule

$$uud \to u, \qquad ud \to h, \qquad d \to d. \tag{2.1}$$

We obtain a Motzkin path of length n. This step is clearly a bijection.

Up to reflection of the Motkin path over a vertical line, Θ is essentially the same bijection that was given by Claesson [**3**] between $\mathfrak{M}_n$ and $\mathcal{M}_n$, using a recursive definition.

2.2. Statistics in $\mathfrak{M}_n$. Here we show applications of the bijection Θ to give generating functions for several statistics in Motzkin permutations. For a permutation π, denote by $\mathrm{lis}(\pi)$ and $\mathrm{lds}(\pi)$ respectively the length of the longest increasing subsequence and the length of the longest decreasing subsequence of π. The following lemma follows from the definitions of the bijections and from the properties of φ (see [**11**]).

Lemma 2.1. *Let $\pi \in \mathfrak{M}_n$, let $D = \varphi(\pi) \in \overrightarrow{\mathbb{G}}_n$, and let $M = \Theta(\pi) \in \mathcal{M}_n$. We have*

1. $\mathrm{lds}(\pi) = \#\{$*peaks of* $D\} = \#\{$*steps* u *in* $M\} + \#\{$*steps* h *in* $M\}$,
2. $\mathrm{lis}(\pi) =$ *height of* D = *height of* $M + 1$,
3. $\#\{$*rises of* $\pi\} = \#\{$*double rises of* $D\} = \#\{$*steps* u *in* $M\}$.

Theorem 2.2. *The generating function for Motzkin permutations with respect to the length of the longest decreasing subsequence and to the number of rises is*

$$A(v,y,x) := \sum\nolimits_{n\geq 0} \sum\nolimits_{\pi\in\mathfrak{M}_n} v^{\mathrm{lds}(\pi)} y^{\#\{\text{rises of } \pi\}} x^n = \frac{1 - vx - \sqrt{1 - 2vx + (v^2 - 4vy)x^2}}{2vyx^2}.$$

Moreover,

$$A(v,y,x) = \sum\nolimits_{n\geq 0} \sum\nolimits_{m\geq 0} \frac{1}{n+1}\binom{2n}{n}\binom{m+2n}{2n} x^{m+2n} v^{m+n} y^n.$$

Proof. By Lemma 2.1, we can express A as

$$A(v,y,x) = \sum\nolimits_{M\in\mathcal{M}} v^{\#\{\text{steps } u \text{ in } M\} + \#\{\text{steps } h \text{ in } M\}} y^{\#\{\text{steps } u \text{ in } M\}} x^{|M|}.$$

Using the standard decomposition of Motzkin paths, we obtain the following equation for the generating function A.

$$A(v,y,x) = 1 + vxA(v,y,x) + vyx^2A^2(v,y,x). \tag{2.2}$$

Indeed, any nonempty $M \in \mathcal{M}$ can be written uniquely in one of the following two forms: (1) $M = hM_1$ and (2) $M = uM_1dM_2$, where M_1, M_2, M_3 are arbitrary Motzkin paths. In the first case, the number of horizontal steps of hM_1 is one more than in M_1, the number of up steps is the same, and $|hM_1| = |M_1| + 1$, so we get the term $vxA(v,y,x)$. Similarly, the second case gives the term $vyx^2A^2(v,y,x)$. Solving equation (2.2) we get the desired expression. □

Theorem 2.3. *For $k > 0$, let $B_k(v,y,x) := \sum_{n\geq 0}\sum_{\pi\in\mathfrak{M}_n(12\ldots(k+1))} v^{\mathrm{lds}(\pi)} y^{\#\{\text{rises of } \pi\}} x^n$ be the generating function for Motzkin permutations avoiding $12\ldots(k+1)$ with respect to the length of the longest decreasing subsequence and to the number of rises. Then we have the recurrence*

$$B_k(v,y,x) = \frac{1}{1 - vx - vyx^2 B_{k-1}(v,y,x)},$$

with $B_1(v,y,x)=\frac{1}{1-vx}$. *Thus,* B_k *can be expressed as*

$$B_k(v,y,x)=\cfrac{1}{1-vx-\cfrac{vyx^2}{\cfrac{\ddots}{1-vx-\cfrac{vyx^2}{1-vx}}}},$$

where the fraction has k *levels, or in terms of Chebyshev polynomials of the second kind, as*

$$B_k(v,y,x)=\frac{U_{k-1}\left(\frac{1-vx}{2x\sqrt{vy}}\right)}{x\sqrt{vy}U_k\left(\frac{1-vx}{2x\sqrt{vy}}\right)}.$$

Proof. The condition that π avoids $12\dots(k+1)$ is equivalent to the condition $\mathrm{lis}(\pi)\le k$. By Lemma 2.1, permutations in $\mathfrak{M}_n$ satisfying this condition are mapped by Θ to Motzkin paths of height strictly less than k. Thus, we can express B_k as

$$B_k(v,y,x)=\sum_{M\in\mathcal{M}\text{ of height}<k} v^{\#\{\text{steps } u \text{ in } M\}+\#\{\text{steps } h \text{ in } M\}}y^{\#\{\text{steps } u \text{ in } M\}}x^{|M|}.$$

For $k>1$, we use again the standard decomposition of Motzkin paths. In the first of the above cases, the height of hM_1 is the same as the height of M_1. However, in the second case, in order for the height of uM_2dM_3 to be less than k, the height of M_2 has to be less than $k-1$. So we obtain the equation

$$B_k(v,y,x)=1+vxB_k(v,y,x)+vyx^2B_{k-1}(v,y,x)B_k(v,y,x).$$

For $k=1$, the path can have only horizontal steps, so we get $B_1(v,y,x)=\frac{1}{1-vx}$. Now, using the above recurrence and Equation 1.1 we get the desired result. □

2.3. Fixed points in the reversal of Motzkin permutations. Here we show another application of Θ. A slight modification of it will allow us to enumerate fixed points in another class of pattern-avoiding permutations closely related to Motzkin permutations. For any $\pi=\pi_1\pi_2\dots\pi_n\in S_n$, denote its reversal by $\pi^R=\pi_n\dots\pi_2\pi_1$. Let $\mathfrak{M}_n^R:=\{\pi\in S_n:\pi^R\in\mathfrak{M}_n\}$. In terms of pattern avoidance, $\mathfrak{M}_n^R$ is the set of permutations that avoid 231 and 32-1 simultaneously, that is, the set of 231-avoiding permutations $\pi\in S_n$ where there do not exist $a<b$ such that $\pi_{a-1}>\pi_a>\pi_b$. Recall that i is called a *fixed point* of π if $\pi_i=i$.

Theorem 2.4. *The generating function* $\sum_{n\ge0}\sum_{\pi\in\mathfrak{M}_n^R}w^{\mathrm{fp}(\pi)}x^n$ *for permutations avoiding simultaneously* 231 *and* 32-1 *with respect to to the number of fixed points is*

$$\cfrac{1}{1-wx-\cfrac{x^2}{1-x-M_0(w-1)x^2-\cfrac{x^2}{1-x-M_1(w-1)x^3-\cfrac{x^2}{1-x-M_2(w-1)x^4-\cfrac{x^2}{\ddots}}}}},\tag{2.3}$$

where after the second level, the coefficient of $(w-1)x^{n+2}$ *is the Motzkin number* M_n.

Proof. We have the following composition of bijections:

$$\begin{array}{ccccccc}\mathfrak{M}_n^R & \longleftrightarrow & \mathfrak{M}_n & \longleftrightarrow & \mathcal{E}_n & \longleftrightarrow & \mathcal{M}_n\\ \pi & \mapsto & \pi^R & \mapsto & \varphi(\pi^R) & \mapsto & \Theta(\pi^R)\end{array}$$

The idea of the proof is to look at how the fixed points of π are transformed by each of these bijections.

For this we use the definition of tunnel of a Dyck path given in [**6**], and generalize it to Motzkin paths. A *tunnel* of $M \in \mathcal{M}$ (resp. $D \in \vec{\mathbb{G}}$) is a horizontal segment between two lattice points of the path that intersects M (resp. D) only in these two points, and stays always below the path. Tunnels are in obvious one-to-one correspondence with decompositions of the path as $M = XuYdZ$ (resp. $D = XuYdZ$), where $Y \in \mathcal{M}$ (resp. $Y \in \vec{\mathbb{G}}$). In the decomposition, the tunnel is the segment that goes from the beginning of the u to the end of the d. Clearly such a decomposition can be given for each up-step u, so the number of tunnels of a path equals its number of up-steps. The *length* of a tunnel is just its length as a segment, and the *height* is its y-coordinate.

Fixed points of π are mapped by the reversal operation to elements j such that $\pi^R_j = n+1-j$, which in the array of π^R correspond to crosses on the diagonal between the bottom-left and top-right corners. Each cross in this array naturally corresponds to a tunnel of the Dyck path $\varphi(\pi^R)$, namely the one determined by the vertical step in the same row as the cross and the horizontal step in the same column as the cross. It is not hard to see (and is also shown in [**7**]) that crosses on the diagonal between the bottom-left and top-right corners correspond in the Dyck path to tunnels T satisfying the condition $\text{height}(T) + 1 = \frac{1}{2}\text{length}(T)$.

The next step is to see how these tunnels are transformed by the bijection from $\mathcal{E}_n$ to $\mathcal{M}_n$. Tunnels of height 0 and length 2 in the Dyck path $D := \varphi(\pi^R)$ are just hills ud landing on the x-axis. By the rule (2.2) they are mapped to horizontal steps at height 0 in the Motzkin path $M := \Theta(\pi^R)$. Assume now that $k \geq 1$. A tunnel T of height k and length $2(k+1)$ in D corresponds to a decomposition $D = XuYdZ$ where X ends at height k and $Y \in \vec{\mathbb{G}}_{2k}$. Note that Y has to begin with an up-step (since it is a nonempty Dyck path) followed by a down-step, otherwise D would have a triple rise. Thus, we can write $D = XuudY'dZ$ where $Y' \in \vec{\mathbb{G}}_{2(k-1)}$. When we apply to D the bijection given by rule (2.2), X is mapped to an initial segment $\widetilde{X}$ of a Motzkin path ending at height k, uud is mapped to u, Y' is mapped to a Motzkin path $\widetilde{Y'} \in \mathcal{M}_{k-1}$ of length $k-1$, the d following Y' is mapped to d (since it is preceded by another d), and Z is mapped to a final segment $\widetilde{Z}$ of a Motzkin path going from height k to the x-axis. Thus, we have that $M = \widetilde{X}u\widetilde{Y'}d\widetilde{Z}$. It follows that tunnels T of D satisfying $\text{height}(T) + 1 = \frac{1}{2}\text{length}(T)$ are transformed by the bijection into tunnels $\widetilde{T}$ of M satisfying $\text{height}(\widetilde{T}) + 1 = \text{length}(\widetilde{T})$. We will call *good* tunnels the tunnels of M satisfying this last condition. It remains to show that the generating function for Motzkin paths where w marks the number of good tunnels plus the number of horizontal steps at height 0, and x marks the length of the path, is given by (2.3).

To do this we imitate the technique used in [**7**] to enumerate fixed points in 231-avoiding permutations. We will separate good tunnels according to their height. It is important to notice that if a good tunnel of M corresponds to a decomposition $M = XuYdZ$, then M has no good tunnels inside the part given by Y. In other words, the orthogonal projections on the x-axis of all the good tunnels of a given Motzkin path are disjoint. Clearly, they are also disjoint from horizontal steps at height 0. This observation allows us to give a continued fraction expression for our generating function.

For every $k \geq 1$, let $\text{gt}_k(M)$ be the number of tunnels of M of height k and length $k+1$. Let $\text{hor}(M)$ be the number of horizontal steps at height 0. We have seen that for $\pi \in \mathfrak{M}^R_n$, $\text{fp}(\pi) = \text{hor}(\Theta(\pi^R)) + \sum_{k\geq 1} \text{gt}_k(\Theta(\pi^R))$. We will show now that for every $k \geq 1$, the generating function for Motzkin paths where w marks the statistic $\text{hor}(M) + \text{gt}_1(M) + \cdots + \text{gt}_{k-1}(M)$ is given by the continued fraction (2.3) truncated at level k, with the $(k+1)$-st level replaced with $M(x)$.

A Motzkin path M can be written uniquely as a sequence of horizontal steps h and elevated Motzkin paths $uM'd$, where $M' \in \mathcal{M}$. In terms of the generating function $M(x) = \sum_{M\in\mathcal{M}} x^{|M|}$, this translates into the equation $M(x) = \frac{1}{1-x-x^2M(x)}$. The generating function where w marks horizontal steps at height 0 is just

$$\sum\nolimits_{M\in\mathcal{M}} w^{\mathrm{hor}(M)}x^{|M|} = \frac{1}{1-wx-x^2M(x)}.$$

If we want w to mark also good tunnels at height 1, each M' from the elevated paths above has to be decomposed as a sequence of horizontal steps and elevated Motzkin paths $uM''d$. In this decomposition, a tunnel of height 1 and length 2 is produced by each empty M'', so we have

$$\sum\nolimits_{M\in\mathcal{M}} w^{\mathrm{hor}(M)+\mathrm{gt}_1(M)}x^{|M|} = \cfrac{1}{1-wx-\cfrac{x^2}{1-x-x^2[w-1+M(x)]}}. \tag{2.4}$$

Indeed, the $M_0(=1)$ possible empty paths M'' have to be accounted as w, not as 1.

Let us now enumerate simultaneously horizontal steps at height 0 and good tunnels at heights 1 and 2. We can rewrite (2.4) as

$$\cfrac{1}{1-wx-\cfrac{x^2}{1-x-x^2\left[w-1+\cfrac{1}{1-x-x^2M(x)}\right]}}.$$

Combinatorially, this corresponds to expressing each M'' as a sequence of horizontal steps and elevated paths $uM'''d$, where $M'''\in\mathcal{M}$. Notice that since $uM'''d$ starts at height 2, a tunnel of height 2 and length 3 is created whenever $M'''\in\mathcal{M}_1$. Thus, if we want w to mark also these tunnels, such an M''' has to be accounted as wx, not x. The corresponding generating function is

$$\sum_{M\in\mathcal{M}} w^{\mathrm{hor}(M)+\mathrm{gt}_1(M)+\mathrm{gt}_2(M)}x^{|M|} = \cfrac{1}{1-wx-\cfrac{x^2}{1-x-x^2\left[w-1+\cfrac{1}{1-x-x^2[(w-1)x+M(x)]}\right]}}.$$

Now it is clear how iterating this process indefinitely we obtain the continued fraction (2.3). From the generating function where w marks $\mathrm{hor}(M)+\mathrm{gt}_1(M)+\cdots+\mathrm{gt}_k(M)$, we can obtain the one where w marks $\mathrm{hor}(M)+\mathrm{gt}_1(M)+\cdots+\mathrm{gt}_{k+1}(M)$ by replacing the $M(x)$ at the lowest level with

$$\frac{1}{1-x-x^2[M_k(w-1)x^k+M(x)]},$$

to account for tunnels of height k and length $k+1$, which in the decomposition correspond to elevated Motzkin paths at height k. □

3. Restricted Motzkin permutations

In this section we consider those Motzkin permutations in $\mathfrak{M}_n$ that avoid an another pattern τ. More generally, we enumerate Motzkin permutations according to the number of occurrences of τ. Subsection 3.2 deals with the increasing pattern $\tau = 12\dots k$. In Subsection 3.3 we show that if τ has a certain form, we can express the generating function for τ-avoiding Motzkin permutations in terms of the the corresponding generating functions for some subpatterns of τ. Finally, Subsection 3.4 studies the case of the generalized patterns 12-3-…-k and 21-3-…-k.

We begin by setting some notation. Let $M_\tau(n)$ be the number of Motzkin permutations in $\mathfrak{M}_n(\tau)$, and let $N_\tau(x) = \sum_{n\geq 0} M_\tau(n)x^n$ be the corresponding generating function. Let $\pi\in\mathfrak{M}_n$. Using the block decomposition approach (see [**18**]), we have two possible block decompositions of π. These decompositions are described in Lemma 3.1, which is the basis for all the results in this section.

Lemma 3.1. *Let $\pi \in \mathfrak{M}_n$. Then one of the following holds:*
(i) $\pi = (n, \beta)$ *where* $\beta \in \mathfrak{M}_{n-1}$,
(ii) *there exists* t, $2 \le t \le n$, *such that* $\pi = (\alpha, n-t+1, n, \beta)$, *where*

$$(\alpha_1 - (n-t+1), \dots, \alpha_{t-2} - (n-t+1)) \in \mathfrak{M}_{t-2} \text{ and } \beta \in \mathfrak{M}_{n-t}.$$

Proof. Given $\pi \in \mathfrak{M}_n$, take j so that $\pi_j = n$. Then $\pi = (\pi', n, \pi'')$, and the condition that π avoids 132 is equivalent to π' being a permutation of the numbers $n-j+1, n-j+2, \dots, n-1$, π'' being a permutation of the numbers $1, 2, \dots, n-j$, and both π' and π'' being 132-avoiding. On the other hand, it is easy to see that if π' is nonempty, then π avoids 1-23 if and only if the minimal entry of π' is adjacent to n, and both π' and π'' avoid 1-23. Therefore, π avoids 132 and 1-23 if and only if either (i) or (ii) hold. □

3.1. The pattern $\tau = \emptyset$. Here we show the simplest application of Lemma 3.1, to enumerate Motzkin permutations of a given length. This also follows from the bijection to Motzkin paths in Section 2.

Proposition 3.2. *The number of Motzkin permutations of length n is given by M_n, the n-th Motzkin number.*

Proof. As a consequence of Lemma 3.1, there are two possible block decompositions of an arbitrary Motzkin permutation $\pi \in \mathfrak{M}_n$. Let us write an equation for $N_\emptyset(x)$. The first (resp. second) of the block decompositions above contributes as $xN_\emptyset(x)$ (resp. $x^2N_\emptyset^2(x)$). Therefore $N_\emptyset(x) = 1 + xN_\emptyset(x) + x^2N_\emptyset^2(x)$, where 1 is the contribution of the empty Motzkin permutation. Hence, $N_\emptyset(x)$ is the generating function for the Motzkin numbers M_n, as claimed. □

3.2. The pattern $\tau = 12\dots k$. For the first values of k, we have from the definitions that $N_1(x) = 1$ and $N_{12}(x) = \frac{1}{1-x}$. Here we consider the case $\tau = 12\dots k$ for arbitrary k. From Theorem 2.3 we get the following expression for N_τ, for which we also give a direct derivation using the block decomposition of Motzkin permutations.

Theorem 3.3. *For all $k \ge 2$, $N_{12\dots k}(x) = \frac{U_{k-2}\left(\frac{1-x}{2x}\right)}{xU_{k-1}\left(\frac{1-x}{2x}\right)}$.*

Proof. By Lemma 3.1, we have two possibilities for the block decomposition of an arbitrary Motzkin permutation $\pi \in \mathfrak{M}_n$. Let us write an equation for $N_{12\dots k}(x)$. The contribution of the first (resp. second) block decomposition is $xN_{12\dots k}(x)$ (resp. $x^2N_{12\dots(k-1)}(x)N_{12\dots k}(x)$). Therefore,

$$N_{12\dots k}(x) = 1 + xN_{12\dots k}(x) + x^2N_{12\dots k}(x)N_{12\dots(k-1)}(x),$$

where 1 comes from the empty Motzkin permutation. Now, using induction on k and the recursion (1.1) we get the desired result. □

This theorem can be generalized as follows. Let $N(x_1, x_2, \dots)$ be the generating function

$$\sum\nolimits_{n\ge0} \sum\nolimits_{\pi\in\mathfrak{M}_n} \prod\nolimits_{j\ge1} x_j^{12\dots j(\pi)},$$

where $12\dots j(\pi)$ is the number of occurrences of the pattern $12\dots j$ in π.

Theorem 3.4. *The generating function $\sum_{n\ge0}\sum_{\pi\in\mathfrak{M}_n}\prod_{j\ge1} x_j^{12\dots j(\pi)}$ is given by the following continued fraction:*

$$\cfrac{1}{1 - x_1 - \cfrac{x_1^2x_2}{1 - x_1x_2 - \cfrac{x_1^2x_2^3x_3}{1 - x_1x_2^2x_3 - \cfrac{x_1^2x_2^5x_3^4x_4}{\ddots}}}},$$

in which the n-th numerator is $\prod_{i=1}^{n+1} x_i^{\binom{n}{i-1}+\binom{n-1}{i-1}}$ and the monomial in the n-th denominator is $\prod_{i=1}^{n} x_i^{\binom{n-1}{i-1}}$.

Proof. By Lemma 3.1, we have two possibilities for the block decomposition of an arbitrary Motzkin permutation $\pi \in \mathfrak{M}_n$. Let us write an equation for $N(x_1, x_2, \ldots)$. The contribution of the first decomposition is $x_1 N(x_1, x_2, \ldots)$, and the second decomposition gives $x_1^2 x_2 N(x_1 x_2, x_2 x_3, \ldots) N(x_1, x_2, \ldots)$. Therefore,

$$N(x_1, x_2, \ldots) = 1 + x_1 N(x_1, x_2, \ldots) + x_1^2 x_2 N(x_1 x_2, x_2 x_3, \ldots) N(x_1, x_2, \ldots),$$

where 1 is the contribution of the empty Motzkin permutation. The theorem follows now by induction. □

3.2.1. *Counting occurrences of the pattern* $12\ldots k$ *in a Motzkin permutation.* Using Theorem 3.4 we can enumerate occurrences of the pattern $12\ldots k$ in Motzkin permutations.

Theorem 3.5. *Fix* $k \geq 2$*. The generating function for the number of Motzkin permutations which contain* $12\ldots k$ *exactly* r*,* $r = 1, 2, \ldots, k$*, times is given by*

$$\frac{\left(U_{k-2}\left(\frac{1-x}{2x}\right) - xU_{k-3}\left(\frac{1-x}{2x}\right)\right)^{r-1}}{U_{k-1}^{r+1}\left(\frac{1-x}{2x}\right)}.$$

Proof. Let $x_1 = x$, $x_k = y$, and $x_j = 1$ for all $j \neq 1, k$. Let $G_k(x, y)$ be the function obtained from $N(x_1, x_2, \ldots)$ after this substitution. Theorem 3.4 gives

$$G_k(x,y) = \cfrac{1}{1 - x - \cfrac{x^2}{\ddots - \cfrac{\ddots}{1 - x - \cfrac{x^2 y}{1 - xy - \cfrac{x^2 y^{k+1}}{\ddots}}}}}.$$

So, $G_k(x, y)$ can be expressed as follows. For all $k \geq 2$,

$$G_k(x,y) = \frac{1}{1 - x - x^2 G_{k-1}(x,y)},$$

and there exists a continued fraction $H(x, y)$ such that $G_1(x,y) = \frac{y}{1-xy-y^{k+1}H(x,y)}$. Now, using induction on k together with (1.1) we get that there exists a formal power series $J(x, y)$ such that

$$G_k(x,y) = \frac{U_{k-2}\left(\frac{1-x}{2x}\right) - \left(U_{k-3}\left(\frac{1-x}{2x}\right) - xU_{k-4}\left(\frac{1-x}{2x}\right)\right) y}{xU_{k-1}\left(\frac{1-x}{2x}\right) - x\left(U_{k-2}\left(\frac{1-x}{2x}\right) - xU_{k-3}\left(\frac{1-x}{2x}\right)\right) y} + y^{k+1} J(x,y).$$

The series expansion of $G_k(x, y)$ about the point $y = 0$ gives

$$G_k(x,y) = \left[U_{k-2}\left(\tfrac{1-x}{2x}\right) - \left(U_{k-3}\left(\tfrac{1-x}{2x}\right) - xU_{k-4}\left(\tfrac{1-x}{2x}\right)\right) y\right] \cdot \sum_{r\geq 0} \frac{\left(U_{k-2}\left(\frac{1-x}{2x}\right) - xU_{k-3}\left(\frac{1-x}{2x}\right)\right)^r}{xU_{k-1}^{r+1}\left(\frac{1-x}{2x}\right)} y^r + y^{k+1} J(x,y).$$

Hence, by using the identities $U_k^2(t) - U_{k-1}(t)U_{k+1}(t) = 1$ and $U_k(t)U_{k-1}(t) - U_{k-2}(t)U_{k+1}(t) = 2t$ we get the desired result. □

3.2.2. *More statistics on Motzkin permutations.* We can use the above theorem to find the generating function for the number of Motzkin permutations with respect to various statistics.

For another application of Theorem 3.4, recall that i is a *free rise* of π if there exists j such that $\pi_i < \pi_j$. We denote the number of free rises of π by $fr(\pi)$. Using Theorem 3.4 for $x_1 = x$, $x_2 = q$, and $x_j = 1$ for $j \geq 3$, we get the following result.

Corollary 3.6. *The generating function* $\sum_{n\geq 0}\sum_{\pi\in\mathfrak{M}_n} x^n q^{fr(\pi)}$ *is given by the following continued fraction:*

$$\cfrac{1}{1-x-\cfrac{x^2q}{1-xq-\cfrac{x^2q^3}{1-xq^2-\cfrac{x^2q^5}{\ddots}}}},$$

in which the n-th numerator is x^2q^{2n-1} *and the monomial in the n-th denominator is* xq^{n-1}.

For our next application, recall that π_j is a *right-to-left maximum* of a permutation π if $\pi_i < \pi_j$ for all $i > j$. We denote the number of right-to-left maxima of π by $rlm(\pi)$.

Corollary 3.7. *The generating function* $\sum_{n\geq 0}\sum_{\pi\in\mathfrak{M}_n} x^n q^{rlm(\pi)}$ *is given by the following continued fraction:*

$$\cfrac{1}{1-xq-\cfrac{x^2q}{1-x-\cfrac{x^2}{1-x-\cfrac{x^2}{\ddots}}}}.$$

Moreover, $\sum_{n\geq 0}\sum_{\pi\in\mathfrak{M}_n} x^n q^{lrm(\pi)} = \sum_{m\geq 0} x^m(1+xM(x))^m q^m$.

Proof. Using Theorem 3.4 for $x_1 = xq$, and $x_{2j} = x_{2j+1}^{-1} = q^{-1}$ for $j \geq 1$, together with [**2**, Proposition 5] we get the first equation as claimed. The second equation follows from the fact that the continued fraction

$$\cfrac{1}{1-x-\cfrac{x^2}{1-x-\cfrac{x^2}{\ddots}}}$$

is given by the generating function for the Motzkin numbers, namely $M(x)$. □

3.3. General restriction. Let us find the generating function for those Motzkin permutations which avoid τ in terms of the generating function for Motzkin permutations avoiding ρ, where ρ is a permutation obtained by removing some entries from τ.

Theorem 3.8. *Let* $k \geq 4$, $\tau = (\rho', 1, k) \in \mathfrak{M}_k$, *and let* $\rho \in \mathfrak{M}_{k-2}$ *be the permutation obtained by decreasing each entry of* ρ' *by 1. Then*

$$N_\tau(x) = \frac{1}{1-x-x^2N_\rho(x)}.$$

Proof. By Lemma 3.1, we have two possibilities for block decomposition of a nonempty Motzkin permutation in $\mathfrak{M}_n$. Let us write an equation for $N_\tau(x)$. The contribution of the first decomposition is $xN_\tau(x)$, and from the second decomposition we get $x^2N_\rho(x)N_\tau(x)$. Hence $N_\tau(x) = 1 + xN_\tau(x) + x^2N_\rho(x)N_\tau(x)$, where 1 corresponds to the empty Motzkin permutation. Solving the above equation we get the desired result. □

For example, using Theorem 3.8 for $\tau = 23\dots(k-1)1k$ ($\rho = 12\dots(k-2)$) we have

$$N_{23\dots(k-1)1k}(x) = \frac{1}{1-x-x^2N_{12\dots(k-2)}(x)}.$$

Hence, by Theorem 3.3 together with (1.1) we get

$$N_{23\dots(k-1)1k}(x) = \frac{U_{k-3}\left(\frac{1-x}{2x}\right)}{xU_{k-2}\left(\frac{1-x}{2x}\right)}.$$

Corollary 3.9. *For all $k \geq 1$,*

$$N_{k(k+1)(k-1)(k+2)(k-2)(k+3)\dots1(2k)}(x) = \frac{U_{k-1}\left(\frac{1-x}{2x}\right)}{xU_k\left(\frac{1-x}{2x}\right)},$$

and

$$N_{(k+1)k(k+2)(k-1)(k+3)\dots1(2k+1)}(x) = \frac{U_k\left(\frac{1-x}{2x}\right)+U_{k-1}\left(\frac{1-x}{2x}\right)}{x\left(U_{k+1}\left(\frac{1-x}{2x}\right)+U_k\left(\frac{1-x}{2x}\right)\right)}.$$

Proof. Theorem 3.8 for $\tau = k(k+1)(k-1)(k+2)(k-2)(k+3)\dots1(2k)$ gives

$$N_\tau(x) = \frac{1}{1-x-x^2N_{(k-1)k(k-2)(k+1)(k-3)(k+2)\dots1(2k-2)}(x)}.$$

Now we argue by induction on k, using (1.1) and the fact that $N_{12}(x) = \frac{1}{1-x}$. Similarly, we get the explicit formula for $N_{(k+1)k(k+2)(k-1)(k+3)\dots1(2k+1)}(x)$. □

Theorem 3.3 and Corollary 3.9 suggest that there should exist a bijection between the sets $\mathfrak{M}_n(12\dots(k+1))$ and $\mathfrak{M}_n(k(k+1)(k-1)(k+2)(k-2)(k+3)\dots1(2k))$. Finding it remains an interesting open question.

Theorem 3.10. *Let $\tau = (\rho', t, k, \theta', 1, t-1) \in \mathfrak{M}_k$ such that $\rho'_a > t > \theta'_b$ for all a, b. Let ρ and θ be the permutations obtained by decreasing each entry of ρ' by t and decreasing each entry of θ' by 1, respectively. Then*

$$N_\tau(x) = \frac{1-x^2N_\rho(x)\widetilde{N}_\theta(x)}{1-x-x^2(N_\rho(x)+\widetilde{N}_\theta(x))},$$

where $\widetilde{N}_\theta(x) = \frac{1}{1-x-x^2N_\theta(x)}$.

Proof. By Lemma 3.1, we have two possibilities for block decomposition of a nonempty Motzkin permutation $\pi \in \mathfrak{M}_n$. Let us write an equation for $N_\tau(x)$. The contribution of the first decomposition is $xN_\tau(x)$. The second decomposition contributes $x^2N_\rho(x)N_\tau(x)$ if α avoids ρ, and $x^2(N_\tau(x)-N_\rho(x))\widetilde{N}_\theta(x)$ if α contains ρ. This last case follows from Theorem 3.8, since if α contains ρ, β has to avoid $(\theta, 1, t-1)$. Hence,

$$N_\tau(x) = 1 + xN_\tau(x) + x^2N_\rho(x)N_\tau(x) + x^2(N_\tau(x)-N_\rho(x))\widetilde{N}_\theta(x),$$

where 1 is the contribution of the empty Motzkin permutation. Solving the above equation we get the desired result. □

For example, for $\tau = 546213$ ($\tau = \rho46\theta13$), Theorem 3.10 gives $N_\tau(x) = \frac{1-2x}{(1-x)(1-2x-x^2)}$.

The last two theorems can be generalized as follows.

Theorem 3.11. *Let $\tau = (\tau^1, t_1{+}1, t_0, \tau^2, t_2{+}1, t_1, \dots, \tau^m, t_m{+}1, t_{m-1})$ where $t_{j-1} > \tau^j_a > t_j$ for all a and j. We define $\sigma^j = (\tau^1, t_1{+}1, t_0, \dots, \tau^j)$ for $j = 2, \dots, m$, $\sigma^0 = \emptyset$, and $\theta^j = (\tau^j, t_j{+}1, t_{j-1}, \dots, \tau^m, t_m{+}1, t_{m-1})$ for $j = 1, 2, \dots, m$. Then*

$$N_\tau(x) = 1 + xN_\tau(x) + x^2\sum_{j=1}^{m}(N_{\sigma^j}(x) - N_{\sigma^{j-1}})N_{\theta^j}(x).$$

(By convention, if ρ is a permutation of $\{i{+}1, i{+}2, \dots, i{+}l\}$, then N_ρ is defined as $N_{\rho'}$, where ρ' is obtained from ρ decreasing each entry by i.)

PROOF. By Lemma 3.1, we have two possibilities for block decomposition of a nonempty Motzkin permutation $\pi \in \mathfrak{M}_n$. Let us write an equation for $N_\tau(x)$. The contribution of the first decomposition is $xN_\tau(x)$. The second decomposition contributes $x^2(N_{\sigma^j}(x) - N_{\sigma^{j-1}}(x))N_{\theta^j}(x)$ if α avoids σ^j and contains σ^{j-1} (which happens exactly for one value of j), because in this case β must avoid θ^j. Therefore, adding all the possibilities of contributions with the contribution 1 for the empty Motzkin permutation we get the desired result. □

For example, this theorem can be used together with Theorem 3.3 to give the following result.

Corollary 3.12. (i) *For all* $k \geq 3$, $N_{(k-1)k12\ldots(k-2)}(x) = \frac{U_{k-3}\left(\frac{1-x}{2x}\right)}{xU_{k-2}\left(\frac{1-x}{2x}\right)}$;

(ii) *For all* $k \geq 4$, $N_{(k-1)(k-2)k12\ldots(k-3)}(x) = \frac{U_{k-4}\left(\frac{1-x}{2x}\right) - xU_{k-5}\left(\frac{1-x}{2x}\right)}{x\left(U_{k-3}\left(\frac{1-x}{2x}\right) - xU_{k-4}\left(\frac{1-x}{2x}\right)\right)}$;

(iii) *For all* $1 \leq t \leq k-3$, $N_{(t+2)(t+3)\ldots(k-1)(t+1)k12\ldots t}(x) = \frac{U_{k-4}\left(\frac{1-x}{2x}\right)}{xU_{k-3}\left(\frac{1-x}{2x}\right)}$.

3.4. Generalized patterns. In this section we consider the case of generalized patterns (see Subsection 1.1), and we study some statistics on Motzkin permutations.

3.4.1. *Counting occurrences of the generalized patterns* 12-3-...-k *and* 21-3-...-k. We denote by $F(t, X, Y) = F(t, x_2, x_3, \ldots, y_2, y_3, \ldots)$ the generating function $\sum_{n\geq 0}\sum_{\pi\in\mathfrak{M}_n} t^n \prod_{j\geq 2} x_j^{12\text{-}3\text{-}\ldots\text{-}j(\pi)} y_j^{21\text{-}3\text{-}\ldots\text{-}j(\pi)}$, where 12-3-...-$j(\pi)$ and 21-3-...-$j(\pi)$ are the number of occurrences of the pattern 12-3-...-j and 21-3-...-j in π, respectively.

Theorem 3.13. *We have*

$$F(t, X, Y) = 1 - \cfrac{t}{ty_2 - \cfrac{1}{1 + tx_2(1 - y_2y_3) + tx_2y_2y_3F(t, X', Y')}},$$

where $X' = (x_2x_3, x_3x_4, \ldots)$ *and* $Y' = (y_2y_3, y_3y_4, \ldots)$. *In other words, the generating function* $F(t, x_2, x_3, \ldots, y_2, y_3, \ldots)$ *is given by the continued fraction*

$$1 - \cfrac{t}{ty_2 - \cfrac{1}{1 + tx_2 - \cfrac{t^2x_2y_2y_3}{ty_2y_3 - \cfrac{1}{1 + tx_2x_3 - \cfrac{t^2x_2x_3y_2y_3^2y_4}{ty_2y_3^2y_4 - \cfrac{1}{1 + tx_2x_3^2x_4 - \cfrac{t^2x_2x_3^2x_4y_2y_3^3y_4^3y_5}{\ddots}}}}}}}.$$

PROOF. As usual, we consider the two possible block decompositions of a nonempty Motzkin permutation $\pi \in \mathfrak{M}_n$ (see Lemma 3.1). Let us write an equation for $F(t, X, Y)$. The contribution of the first decomposition is $t + ty_2(F(t, X, Y) - 1)$. The contribution of the second decomposition gives t^2x_2, $t^2x_2y_2(F(t, X, Y) - 1)$, $t^2x_2y_2y_3(F(t, X', Y') - 1)$, and $t^2x_2y_2^2y_3(F(t, X, Y) - 1)(F(t, X', Y') - 1)$ for the four possibilities $\alpha = \beta = \emptyset$, $\alpha = \emptyset \neq \beta$, $\beta = \emptyset \neq \alpha$, and $\beta, \alpha \neq \emptyset$, respectively. Hence,

$$\begin{aligned} F(t, X, Y) = 1 + t + ty_2(F(t, X, Y) - 1) + t^2x_2 + t^2x_2y_2y_3(F(t, X'Y') - 1) \\ + t^2x_2y_2(F(t, X, Y) - 1) + t^2x_2y_2^2y_3(F(t, X, Y) - 1)(F(t, X', Y') - 1), \end{aligned}$$

where 1 is as usual the contribution of the empty Motzkin permutation. Simplifying the above equation we get

$$F(t, X, Y) = 1 - \cfrac{t}{ty_2 - \cfrac{1}{1 + tx_2(1 - y_2y_3) + tx_2y_2y_3F(t, X', Y')}}.$$

The second part of the theorem now follows by induction. □

As a corollary of Theorem 3.13 we recover the distribution of the number of rises and number of descents on the set of Motzkin permutations, which also follows easily from Theorem 2.2.

Corollary 3.14. *We have*

$$\sum\nolimits_{n\geq 0}\sum\nolimits_{\pi\in\mathfrak{M}_n} t^n p^{\#\{\text{rises in }\pi\}} q^{\#\{\text{descents in }\pi\}} = \frac{1-qt-2pq(1-q)t^2-\sqrt{(1-qt)^2-4pqt^2}}{2pq^2t^2}.$$

As an application of Theorem 3.13 let us consider the case of Motzkin permutations which contain either 12-3-…-k or 21-3-…-k exactly r times. Using the same arguments as in the proof of Theorem 9, we can apply Theorem 3.13 to obtain the following result.

Theorem 3.15. *Fix $k \geq 2$. Let $N_\tau(x;r)$ be the generating function for the number of Motzkin permutations which contain τ exactly r times. Then*

$$N_{12\text{-}3\text{-}\dots\text{-}k}(x;0) = \frac{U_{k-1}\left(\frac{1-x}{2x}\right)}{xU_k\left(\frac{1-x}{2x}\right)}, \qquad N_{21\text{-}3\text{-}\dots\text{-}k}(x;0) = \frac{U_{k-3}\left(\frac{1-x}{2x}\right)-xU_{k-4}\left(\frac{1-x}{2x}\right)}{x\left(U_{k-2}\left(\frac{1-x}{2x}\right)-xU_{k-3}\left(\frac{1-x}{2x}\right)\right)},$$

and for all $r = 1, 2, \dots, k-1$,

$$N_{12\text{-}3\text{-}\dots\text{-}k}(x;r) = \frac{x^{r-1}U_{k-2}^{r-1}\left(\frac{1-x}{2x}\right)}{(1-x)^r U_{k-1}^{r+1}\left(\frac{1-x}{2x}\right)}, \qquad N_{21\text{-}3\text{-}\dots\text{-}k}(x;r) = \frac{x^r(1+x)^r U_{k-2}^{r-1}\left(\frac{1-x}{2x}\right)}{\left(U_{k-2}\left(\frac{1-x}{2x}\right)-xU_{k-3}\left(\frac{1-x}{2x}\right)\right)^{r+1}}.$$

Acknowledgements. We would like to thank Marc Noy for helpful comments and suggestions. The first author was partially supported by a MAE fellowship.

References

[1] E. Babson and E. Steingrímsson, Generalized permutation patterns and a classification of the Mahonian statistics, *Séminaire Lotharingien de Combinatoire* **44** Article B44b (2000).
[2] P. Brändén, A. Claesson, and E. Steingrímsson, Catalan continued fractions and increasing subsequences in permutations, *Discr. Math.* **258** (2002) 275–287.
[3] A. Claesson, Generalised pattern avoidance, *Europ. J. Combin.* **22** (2001) 961–973.
[4] A. Claesson and T. Mansour, Permutations avoiding a pair of generalized patterns of length three with exactly one dash, *Ars Combinatorica*, to appear, math.CO/0107044.
[5] T. Chow and J. West, Forbidden subsequences and Chebyshev polynomials, *Discr. Math.* **204** (1999) 119–128.
[6] S. Elizalde, Fixed points and excedances in restricted permutations, proceedings of FPSAC'03, math.CO/0212221.
[7] S. Elizalde, Multiple pattern-avoidance with respect to fixed points and excedances, preprint.
[8] S. Elizalde, M. Noy, Consecutive subwords in permutations, *Adv. Appl. Math.* **30** (2003) 110–125.
[9] S. Kitaev, Multi-Avoidance of generalised patterns, *Discr. Math.* **260** (2003) 89–100.
[10] D.E. Knuth, The Art of Computer Programming, 2nd ed. Addison Wesley, Reading, MA (1973).
[11] C. Krattenthaler, Permutations with restricted patterns and Dyck paths, *Adv. Appl. Math.* **27** (2001) 510–530.
[12] T. Mansour, Continued fractions and generalized patterns, *Europ. J. Combin.* **23:3** (2002) 329–344.
[13] T. Mansour, Continued fractions, statistics, and generalized patterns, *Ars Combinatorica*, to appear, math.CO/0110040.
[14] T. Mansour, Restricted 1-3-2 permutations and generalized patterns, *Annals of Combin.* **6** (2002) 65–76.
[15] T. Mansour and A. Vainshtein, Restricted permutations, continued fractions, and Chebyshev polynomials *Electron. J. Combin.* **7** (2000) #R17.
[16] T. Mansour and A. Vainshtein, Restricted 132-avoiding permutations, *Adv. Appl. Math.* **126** (2001) 258–269.
[17] T. Mansour and A. Vainshtein, Layered restrictions and Chebychev polynomials, *Annals of Combin.* **5** (2001) 451–458.
[18] T. Mansour and A. Vainshtein, Restricted permutations and Chebyshev polynomials, *Séminaire Lotharingien de Combinatoire* **47** Article B47c (2002).
[19] T. Mansour and A. Vainshtein, Counting occurrences of 132 in a permutation, *Adv. Appl. Math.* **28:2** (2002) 185–195.
[20] A. Robertson, Permutations containing and avoiding 123 and 132 patterns, *Disc. Math. and Theo. Comp. Sci.* **3** (1999) 151–154.
[21] Th. Rivlin, Chebyshev polynomials. From approximation theory to algebra and number theory, John Wiley, New York (1990).
[22] A. Robertson, H. Wilf, and D. Zeilberger, Permutation patterns and continuous fractions, *Elec. J. Combin.* **6** (1999) #R38.
[23] N.J.A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, New York (1995).

[24] R. Simion, F.W. Schmidt, Restricted Permutations, *Europ. J. Combin.* **6** (1985) 383–406.
[25] J. West, Generating trees and forbidden subsequences, *Discr. Math.* **157** (1996) 363–372.

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge MA 02139.
E-mail address: sergi@math.mit.edu

Department of Mathematics, Haifa University, 31905 Haifa, Israel.
E-mail address: toufik@math.haifa.ac.il

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Dénombrement des classes de symétries des polyominos hexagonaux convexes

Dominique Gouyou-Beauchamps and Pierre Leroux

Abstract. *In this paper, we enumerate the symmetry classes of convex polyominoes on the honeycomb lattice (*hexagonal *polyominoes). Here* convexity *is to be understood as convexity along the three main column directions. We deduce the generating series of* free *(i.e. up to reflection and rotation) and of asymmetric convex hexagonal polyominos, according to area and half-perimeter. See* [**4**] *for a longer version in English.*

Résumé. *Dans ce travail, nous dénombrons les classes de symétrie des polyominos hexagonaux convexes. Ici, la convexité est par rapport aux trois directions principales des colonnes. Nous en deduisons les séries génératrices des polyominos hexagonaux convexes* libres, *c'est-à-dire à réflexions et rotations près, ou encore de ceux qui sont asymétriques, selon l'aire et le demi-périmètre.*

1. Introduction

Un *polyomino hexagonal*, est un ensemble fini connexe de cellules de base d'un réseau hexagonal du plan. Sauf mention contraire, les polyominos considérés ici seront toujours hexagonaux. L'*aire* d'un polyomino est le nombre de cellules qui le composent; son *périmètre* est le nombre de segments qui composent sa frontière. On dit qu'un polyomino est *convexe selon une direction* donnée si l'intersection du polyomino avec toute droite parallèle à cette direction et passant par le centre d'une cellule est connexe. Les directions sont caractérisées par l'angle α $(0 \leq \alpha \leq \pi)$ qu'elles font avec l'axe horizontal positif, calculé dans le sens anti-horaire.

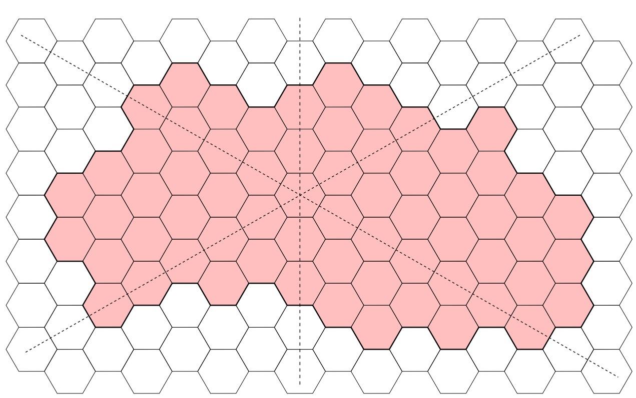


FIGURE 1. Un polyomino convexe et les directions de convexité

1991 *Mathematics Subject Classification.* Primary 05A15 05A30; Secondary 82B20.
Key words and phrases. polyomino enumeration, convex polyominoes, honeycomb lattice, symmetry classes.
With the partial support of CNRS (France), NSERC (Canada), and FCAR (Québec).

Plusieurs concepts de convexité ont été introduits dans la littérature pour les polyominos hexagonaux, selon les directions de convexité demandées. Suivant la nomenclature de Denise, Dür, et Hassani (voir [**2**]), signalons les polyominos EG-convexes, où $\alpha = 0$ et $\pi/2$, étudiés par Guttmann et Enting [**5**] et par Lin et Chang [**9**], les polyominos C^1-convexes, où $\alpha = \pi/2$, énumérés selon plusieurs paramètres par Lin et Wu [**10**] et par Feretić et Svrtan [**3**], les polyominos fortement convexes (F- ou F^3-convexes), où $\alpha = 0$, $\pi/3$ et $2\pi/3$, introduits par Hassani [**6**] et étudiés dans [**2**], et finalement les C- ou C^3-convexes, où $\alpha = \pi/6$, $\pi/2$ et $5\pi/6$, introduits et énumérés suivant le périmètre dans [**6**] et [**2**]. En particulier, Hassani donne explicitement la série algébrique qui énumère les polyominos C-convexes suivant le demi-périmètre.

Ce sont ces derniers polyominos qui nous intéressent ici et que nous appelons tout simplement *convexes*. Ce concept généralise bien la convexité des polyominos dans un réseau carré car les directions de convexité sont celles des colonnes principales du réseau hexagonal. La figure 1 représente un polyomino convexe d'aire 64 et de périmètre 70.

Ces polyominos sont traditionnellement pris à translation près. Il est cependant naturel de les considérer également à rotation et réflexion près, comme des objets qui vivent dans l'espace. Suivant Vöge, Guttmann et Jensen [**15**], nous appelons ces classes d'équivalences *polyominos libres*. En chimie organique, ces objets représentent des molécules d'hydrocarbures benzénoïdes. Voir [**15**] où ces molécules (sans la propriété de convexité) sont énumérées par génération exhaustive.

Notre objectif est donc de dénombrer les polyominos convexes *libres*, selon l'aire et le demi-périmètre. Pour cela, nous les considérons comme les orbites du groupe diédral $\mathcal{D}_6$, des isométries de l'hexagone, agissant sur les polyominos convexes, et nous faisons appel à la Formule de Cauchy-Frobenius (alias Lemme de Burnside). Nous sommes donc amenés à dénombrer les classes de symétries de polyominos convexes, c'est-à-dire les polyominos laissés fixes par chacun des éléments du groupe $\mathcal{D}_6$, suivant la démarche entreprise dans Leroux, Rassart et Robitaille [**8**] pour le réseau carré.

Pour toute classe de polyominos (hexagonaux) convexes $\mathcal{F}$, nous notons $\mathcal{F}(x, q, u, v, t)$ sa série génératrice, où la variable x compte le nombre de colonnes, q compte l'aire, u compte la taille de la première (à gauche) colonne, v, la taille de la dernière colonne, et t le demi-périmètre. Par exemple, le polyomino de la figure 1 est de poids $x^{14}q^{64}u^2v^3t^{35}$. Il est possible que les variables n'apparaissent pas toutes à la fois. Les séries génératrices seront données par des formules explicites ou implicites qui se prètent bien à l'utilisation du calcul formel.

2. Préliminaires

2.1. Classes particulières de polyominos convexes. Quelques classes familières de polyominos convexes du réseau carré se retrouvent naturellement sur le réseau hexagonal et s'avèrent utiles par la suite. C'est le cas notamment des polyominos *partages* et *parallélogrammes*. Par contre, pour les polyominos tas, une variante distincte apparaît.

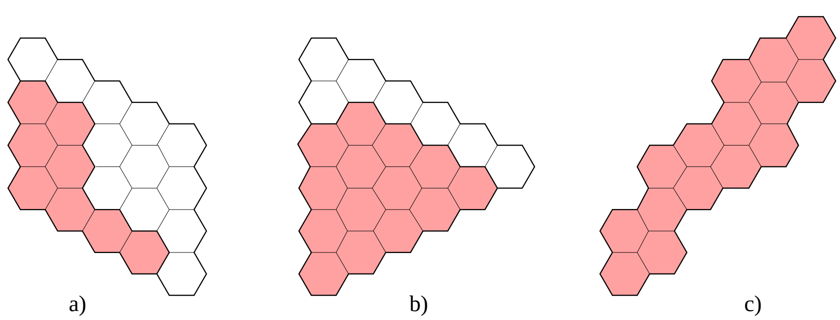


FIGURE 2. polyominos partages et parallélogrammes

2.1.1. *Les polyominos partages.* La figure 2a représente le partage $(4, 2, 2)$ inscrit dans un *rectangle* de taille 5×4 dans le réseau hexagonal. De même, la figure 2b représente le partage en parts distinctes et bornées par 6, $(5, 4, 3, 2)$. Notons par $D_m(u, q)$ le polynôme générateur des partages en parts distinctes bornées par m, où la variable u compte le nombre de parts. On a

$$D_m(u, q) = (1 + uq)(1 + uq^2) \cdots (1 + uq^m) \text{ et } D_0(u, q) = 1. \tag{2.1}$$

2.1.2. *Les polyominos parallélogrammes.* La figure 2c représente un polyomino parallélogramme (*staircase* en anglais) du réseau carré (voir par exemple [**1**] ou [**7**]) reporté sur le réseau hexagonal. On remarque que le demi-périmètre est alors égal à $2p - 1$ où p est le demi-périmètre sur le réseau carré. On sait que ces polyominos sont dénombrés selon le demi-périmètre par les nombres de Catalan et selon l'aire par la suite M1175 de [**13**] (A006958 de [**12**]) dont la série génératrice est un quotient de deux q-fonctions de Bessel.

On note Pa l'ensemble des polyominos parallélogrammes sur le réseau hexagonal et $\mathrm{Pa}(x, q, u, v, t)$ leur série génératrice. En analysant ce qui se passe lorsqu'on ajoute une colonne par la droite, la méthode de M. Bousquet-Mélou donne, pour $\mathrm{Pa}(v) = \mathrm{Pa}(x, q, u, v, t)$ (comparer avec [**1**], Lemma 3.1),

$$\mathrm{Pa}(v) = \frac{xquvt^3}{1 - quvt^2} + \frac{xqvt^2}{(1 - qvt^2)(1 - qv)} \left(\mathrm{Pa}(1) - \mathrm{Pa}(vq)\right) \tag{2.2}$$

et

$$\mathrm{Pa}(v) = \frac{J_1(1) + J_1(v)J_0(1) - J_1(1)J_0(v)}{J_0(1)}, \tag{2.3}$$

où

$$J_1(v) = \sum_{n \geq 0} (-1)^n \frac{x^{n+1} v^{n+1} u t^{2n+3} q^{\binom{n+2}{2}}}{(qvt^2; q)_n (qv; q)_n (1 - q^{n+1} uvt^2)}$$

et

$$J_0(v) = \sum_{n \geq 0} (-1)^n \frac{x^n v^n t^{2n} q^{\binom{n+1}{2}}}{(qvt^2; q)_n (qv; q)_n}$$

La série $\mathrm{Pa}_{i,j}(x, q, t)$ est définie comme le coefficient de $u^i v^j$ dans $\mathrm{Pa}(x, q, u, v, t)$:

$$\mathrm{Pa}(x, q, u, v, t) = \sum_{i \geq 1} \sum_{j \geq 1} \mathrm{Pa}_{i,j}(x, q, t) u^i v^j. \tag{2.4}$$

2.1.3. *Les polyominos tas.* Pour les polyominos *tas*, il existe une variante pour le réseau hexagonal, représentée à la figure 3. Ce sont des empilements pyramidaux d'hexagones, vus de coté. La première classe (figure 3a), notée T, a été considérée dans la littérature sous le nom de *pyramidal stacking of circles*, voir [**11**]. Leur série selon l'aire est référencée sous les numéro M0687 dans [**13**] et A001524 dans [**12**].

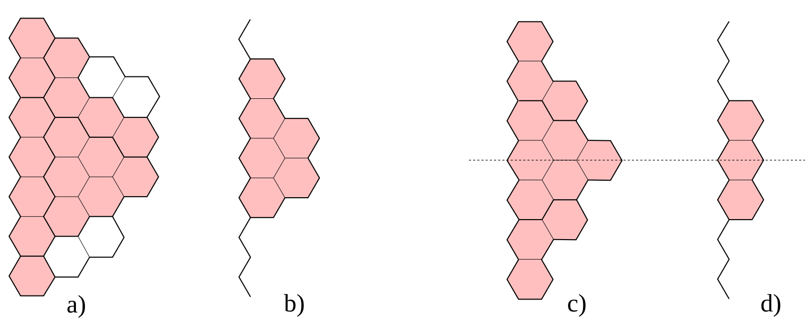


FIGURE 3. Tas et tas symétrique

Soit $T(x,u,q)$, la série génératrice des polyominos tas selon le nombre de colonnes (la *largeur*), la taille de la première colonne (la *hauteur*), et l'aire, et soit $T_n(x,q) = [u^n]T(x,u,q)$, la série génératrice des tas dont la première colonne est de taille n. Noter que le demi-périmètre est égal à deux fois la hauteur plus la largeur de sorte que la série $T(xt, ut^2, q)$ tiendra compte également de ce paramètre.

On a

$$T(x,u,q) = \sum_{m\geq 1} \frac{x^m q^{\binom{m+1}{2}} u^m}{\left((uq;q)_{m-1}\right)^2 (1-uq^m)} \tag{2.5}$$

et

$$T_n(x,q) = \sum_{m=1}^{n} x^m q^{n+\binom{m}{2}} \sum_{j=0}^{n-m} \begin{bmatrix} m+j-1 \\ m-1 \end{bmatrix}_q \begin{bmatrix} n-j-2 \\ m-2 \end{bmatrix}_q . \tag{2.6}$$

Les polynômes $T_n(x,q)$ peuvent aussi être calculés rapidement par récurrence en utilisant la classe $\mathrm{T0}_n$ des polyominos tas dont la première colonne est de taille n en incluant des cellules vides aux deux extrémités. Voir la figure 3b. En effet, on a

$$T_n(x,q) = xq^n \mathrm{T0}_{n-1}(x,q). \tag{2.7}$$

avec $\mathrm{T0}_0(x,q) = 1$, $\mathrm{T0}_1(x,q) = 1+xq$, et, en raisonnant sur l'existence de cellules vides aux deux extrémités,

$$\mathrm{T0}_n(x,q) = (xq^n + 2)\mathrm{T0}_{n-1} - \mathrm{T0}_{n-2}, \tag{2.8}$$

2.1.4. *Les tas symétriques.* Les tas symétriques, par rapport à l'axe horizontal (voir les figures 3c et 3d), constituent les familles TS et TS0. Utilisant les mêmes notations que pour les tas, on a

$$\mathrm{TS}(x,q) = \sum_{m\geq 1} \frac{x^m u^m q^{m(m+1)/2}(1+uq^m)}{(1-u^2q^2)(1-u^2q^4)\cdots(1-u^2q^{2m})}. \tag{2.9}$$

De plus,

$$\mathrm{TS}_n(x,q) = xq^n \mathrm{TS0}_{n-1}(x,q). \tag{2.10}$$

avec $\mathrm{TS0}_0(x,q) = \mathrm{TS0}_{-1}(x,q) = 1$, et

$$\mathrm{TS0}_n(x,q) = xq^n \mathrm{TS0}_{n-1}(x,q) + \mathrm{TS0}_{n-2}(x,q), \tag{2.11}$$

2.2. Le groupe diédral $\mathcal{D}_6$. Le groupe diédral $\mathcal{D}_6$ est défini algébriquement par

$$\mathcal{D}_6 = \langle \rho, \tau \mid \rho^6 = 1,\ \tau^2 = 1,\ \tau\rho\tau = \rho^{-1} \rangle.$$

Ici $\mathcal{D}_6$ est réalisé comme le groupe des isométries de l'hexagone, avec $\rho = \mathrm{r} =$ la rotation de $\pi/3$ radian (dans le sens horaire) et $\tau = \mathrm{ds}_3$, la réflexion selon l'axe horizontal. On a

$$\mathcal{D}_6 = \{\mathrm{id}, \mathrm{r}, \mathrm{r}^2, \mathrm{r}^3, \mathrm{r}^4, \mathrm{r}^5, \mathrm{da}_1, \mathrm{da}_2, \mathrm{da}_3, \mathrm{ds}_1, \mathrm{ds}_2, \mathrm{ds}_3\},$$

où $\mathrm{ds}_2 = \tau\rho^2$, $\mathrm{ds}_1 = \tau\rho^4$, les réflexions selon les axes sommet-sommet, et $\mathrm{da}_3 = \tau\rho$, $\mathrm{da}_2 = \tau\rho^3$, et $\mathrm{da}_1 = \tau\rho^5$, les réflexions selon les axes arêtes-arêtes. Voir la figure 4.

Le groupe diédral $\mathcal{D}_6$ agit sur les polyominos (hexagonaux) de façon naturelle, par rotation ou réflexion. Pour toute classe de polyominos $\mathcal{F}$, munie d'un poids monomial w correspondant à certains paramètres, notons $|\mathcal{F}|_w$ le poids total (i.e. la série génératrice) de cette classe. Si $\mathcal{F}$ est invariante sous l'action de $\mathcal{D}_6$, l'ensemble des orbites de cette action est notée $\mathcal{F}/\mathcal{D}_6$. Le lemme de Burnside permet de dénombrer ces orbites en termes des ensembles $\mathrm{Fix}(g)$ de points fixes de chacun des éléments g de $\mathcal{D}_6$, les *classes de symétries* de $\mathcal{F}$. Notons $fix(g) = |\mathrm{Fix}(g)|_w$. On a évidemment $fix(\mathrm{r}) = fix(\mathrm{r}^5)$, $fix(\mathrm{r}^2) = fix(\mathrm{r}^4)$ et, pour

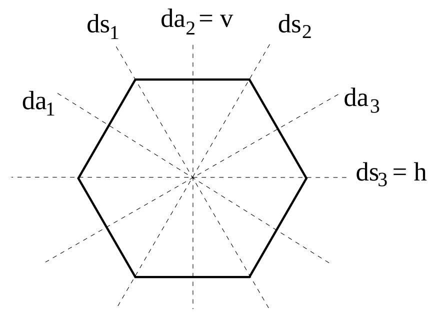


FIGURE 4. Les réflexions de $\mathcal{D}_6$

des raisons de symétries, $fix(\mathrm{da}_1) = fix(\mathrm{da}_2) = fix(\mathrm{da}_3)$ et $fix(\mathrm{ds}_1) = fix(\mathrm{ds}_2) = fix(\mathrm{ds}_3)$. Par la suite, on choisira $\mathrm{v} = \mathrm{da}_2$, l'axe vertical, et $\mathrm{h} = \mathrm{ds}_3$, l'axe horizontal. On a alors

$$|\mathcal{F}/\mathcal{D}_6|_w = \frac{1}{12}\sum_{g\in\mathcal{D}_6} fix(g)$$

$$(2.12) \qquad = \frac{1}{12}\left(|\mathcal{F}|_w + 2fix(\mathrm{r}) + 2fix(\mathrm{r}^2) + fix(\mathrm{r}^3) + 3fix(\mathrm{v}) + 3fix(\mathrm{h})\right).$$

2.3. Phases de croissance des polyominos convexes. Tout polyomino convexe peut être décomposé en blocs selon les phases de croissances, de gauche à droite, de ses profils supérieur et inférieur. La figure 5 donne un exemple de cette décomposition. Le profil supérieur y est représenté par le chemin de A à B le long de la frontière supérieure, et le profil inférieur, par le chemin de C à D. Sur le profil supérieur, on parle d'une *croissance faible* si le niveau monte d'un demi-hexagone seulement par rapport à la colonne précédente, et d'une croissance *forte* si le niveau monte de plus d'un demi-hexagone. On définit de manière analogue la *décroissance faible* ou *forte*. Sur le profil inférieur, une croissance correspond à une descente et une décroissance, à une montée.

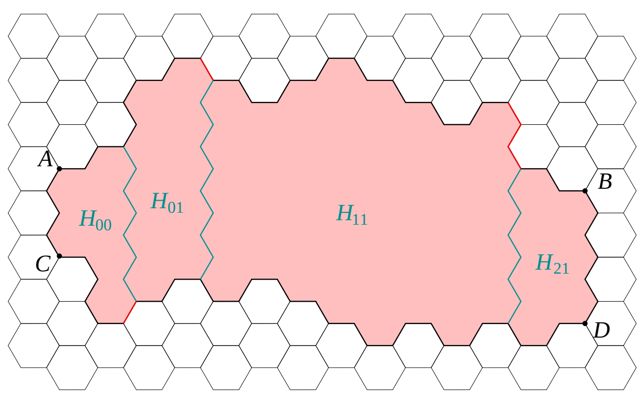


FIGURE 5. Phases d'un polyomino convexe

On décrit l'état dans lequel se trouve une colonne par un couple (i, j), $i, j = 0, 1, 2$, la première composante correspondant au profil supérieur et la deuxième, au profil inférieur. L'état 0 correspond à une croissance faible ou forte, au début du polyomino, l'état 1 à une croissance ou une décroissance faible, dans une phase d'oscillation, et l'état 2, à une décroissance forte ou faible, dans la dernière partie du polyomino. Pour passer de l'état 0 à l'état 1, il doit y avoir une première décroissance, et pour passer de l'état 1 à l'état

2, il doit y avoir une décroissance forte. Enfin, on ne peut passer de l'état 1 à l'état 0 ni de l'état 2 à l'état 1 ou 0. Finalement, un bloc H_{ij} est caractérisé par une suite maximale de colonnes consécutives qui sont dans l'état (i, j).

On peut donc voir un polyomino convexe comme un assemblage de blocs et leur dénombrement passe par celui des H_{ij}. Nous donnons ici les diverses séries génératrices de la forme $H_{ij}(x, q, u, v, t)$.

2.4. Les familles H_{00} et H_{22}. Les polyominos des classes H_{00} et H_{22} sont faciles à énumérer car ce sont en fait des polyominos tas. Ici, une seule des deux variables u et v est utilisée à la fois. On a

$$H_{22}(x, q, u, t) = T(xt, ut^2, q) \text{ et } H_{00}(x, q, v, t) = T(xt, vt^2, q), \tag{2.13}$$

où $T(x, u, q)$ est donnée par la formule (2.5).

2.5. Les familles H_{01}, H_{10}, H_{12}, et H_{21}. Les classes de polyominos H_{01}, H_{10}, H_{12}, et H_{21} sont en bijection entre elles par les réflexions horizontales et verticales et sont donc équivalentes à dénombrer. La figure 6 illustre un polyomino de H_{10}. On trouve facilement que

$$H_{10}(x, q, u, v, t) = \frac{xquvt^3}{1 - quvt^2} + \frac{xt^2(1 + qvt)}{1 - qvt^2} H_{10}(x, q, u, vq, t) \tag{2.14}$$

$$= \sum_{m \geq 1} \frac{x^m q^m uvt^{2m+1}(-qvt; q)_{m-1}}{(qvt^2; q)_{m-1}(1 - q^m uvt^2)}. \tag{2.15}$$

La formule (2.15) se voit directement sur la figure 6. On peut aussi y voir que

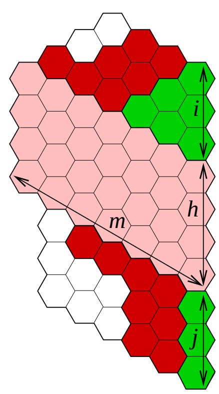


FIGURE 6. Polyomino de H_{10}

$$H_{10}(x, q, u, v, t) = \sum_{h \geq 1} u^h v^h \sum_{m \geq 1} x^m q^{mh} \sum_{i=0}^{m-1} v^i q^{\binom{i+1}{2}} \begin{bmatrix} m-1 \\ i \end{bmatrix}_q \sum_{j \geq 0} v^j q^j t^{2m+2h+i+2j-1} \begin{bmatrix} m-2+j \\ j \end{bmatrix}_q . \tag{2.16}$$

Noter que

$$H_{01}(x, q, u, v, t) = H_{10}(x, q, u, v, t)$$

et que

$$H_{12}(x, q, u, v, t) = H_{21}(x, q, u, v, t) = H_{10}(x, q, v, u, t)$$

.

2.6. Les familles H_{02} et H_{20}. Ces deux classes sont en fait équivalentes aux polyominos parallélogrammes:

$$H_{02}(x, q, u, v, t) = \text{Pa}(x, q, u, v, t) = H_{02}(x, q, u, v, t). \tag{2.17}$$

2.7. La famille H_{11}. La classe H_{11} contient les polyominos convexes dont les profils supérieur et inférieur oscillent tous les deux. Lorsque l'on examine la rangée connexe d'hexagones dans l'axe da_3 (voir Figure 4) contenant la cellule inférieure de la première colonne, on voit apparaître deux sous-classes de H_{11}. La première classe, notée H_{11a}, est celle où cette rangée et celles qui sont à sa droite (jusqu'à la dernière colonne) forment un polyomino parallélogramme (incliné de $\pi/3$); voir la figure 7a. La deuxième classe, notée H_{11b}, est celle où cette rangée est la base d'un rectangle de hauteur au moins 2; voir la figure 7b. Dans les deux cas, on retrouve au-dessus et au-desssous de ces objets (parallélogramme ou rectangle) des partages en parts distinctes qui sont justifiés à gauche et à droite, respectivement.

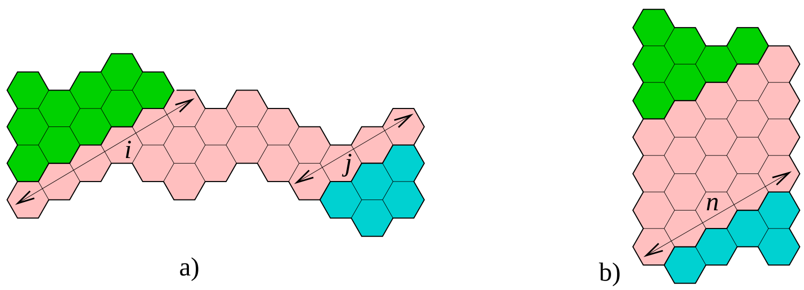


FIGURE 7. Polyominos de H_{11}

Rappelons que dans la série $\mathrm{Pa}_{i,j}(x,q,t)$, définie par (2.4), la variable x marque le nombre de colonnes du parallélogramme(non incliné). Notons plutot le lien entre sa largeur ℓ, lorsqu'incliné, et son demi-périmètre p: $p = 2\ell + 1$. On a donc

$$H_{11}(x,q,u,v,t) = H_{11a}(x,q,u,v,t) + H_{11b}(x,q,u,v,t), \tag{2.18}$$

avec

$$H_{11a}(x,q,u,v,t) = \sum_{i\geq 1\, j\geq 1} x^{-\frac{1}{2}} uv\, \mathrm{Pa}_{i,j}(1,q,tx^{\frac{1}{2}}) D_{i-1}(ut,q) D_{j-1}(vt,q) \tag{2.19}$$

et

$$H_{11b}(x,q,u,v,t) = \sum_{n\geq 1} \frac{x^n q^{2n} u^2 v^2 t^{2n+3} D_{n-1}(ut,q) D_{n-1}(vt,q)}{1-q^n uvt^2}. \tag{2.20}$$

3. Les polyominos convexes

Notons C, la classe de tous les polyominos convexes et C_{ij}, la sous-classe des polyominos dont la dernière colonne est dans l'état (i,j), $i,j = 0,1,2$. Ceci détermine une partition de C. Pour dénombrer C, il faut donc dénombrer chacune des classes C_{ij}. Nous donnons les séries génératrices $C_{ij}(x,q,v,t)$, suivant essentiellement la méthode de Hassani [**6**], utilisant la décomposition d'un polyomino convexe selon les phases de croissance, de gauche à droite.

Nous utilisons la notation $C_{ij} \otimes H_{i'j'}$ pour désigner l'ensemble des polyominos convexes obtenus en recollant de toutes les façons légales possibles un polyomino de C_{ij} avec un de $H_{i'j'}$. On introduit les séries $C_{ij,n}(x,q,t)$ et $H_{ij,n}(x,q,v,t)$ par les extractions de coefficients

$$C_{ij,n}(x,q,t) = [v^n] C_{ij}(x,q,v,t) \text{ et } H_{ij,n}(x,q,v,t) = [u^n] H_{ij}(x,q,u,v,t). \tag{3.1}$$

Par exemple, on a $C_{00} = H_{00}$, $C_{10} = C_{00} \otimes H_{10}$ et

$$\begin{aligned} C_{10}(x,q,v,t) &= \sum_{n\geq 1}\left(\sum_{k=1}^{n}\frac{1}{t^{2k-1}}C_{00,k}(x,q,t)\right)H_{10,n}(x,q,v,t) \\ &= \sum_{n\geq 1}\left(\sum_{k=1}^{n} tT_k(xt,q)\right)H_{10,n}(x,q,v,t) \qquad (3.2) \\ &= C_{01}(x,q,v,t). \end{aligned}$$

De même, $C_{11} = (C_{00} + C_{10} + C_{01}) \otimes H_{11} = C_{00} \otimes H_{11} + C_{10} \otimes H_{11} + C_{01} \otimes H_{11}$ et

$$C_{00} \otimes H_{11}\,(x,q,v,t) = \sum_{n\geq 2}\frac{1}{t^{2n-2}}C_{00,n}(x,q,t)H_{11,n-1}(x,q,v,t), \qquad (3.3)$$

$$C_{10}\otimes H_{11}\,(x,q,v,t) = \sum_{n\geq 1}\frac{1}{t^{2n-1}}C_{10,n}(x,q,t)H_{11,n}(x,q,v,t) + \sum_{n\geq 2}\frac{1}{t^{2n-2}}C_{10,n}(x,q,t)H_{11,n-1}(x,q,v,t). \qquad (3.4)$$

On a aussi $C_{02} = (C_{00} + C_{01}) \otimes H_{02}$,

$$C_{12} = (C_{00} + C_{01} + C_{10} + C_{11} + C_{02}) \otimes H_{12},$$

$$C_{22} = (C_{00} + C_{01} + C_{10} + C_{11} + C_{02} + C_{02}) \otimes H_{22}.$$

Finalement

$$C(x,q,v,t) = (C_{00} + 2C_{10} + C_{11} + 2C_{02} + 2C_{12} + C_{22})(x,q,v,t). \qquad (3.5)$$

4. Les classes de symétrie réflexives

4.1. Symétrie verticale. Considérons un polyomino convexe v-symétrique P. On constate que l'axe de symétrie passe par une colonne centrale. Notons K la région fondamentale gauche de P, incluant la colonne centrale. Voir la figure 8. On a

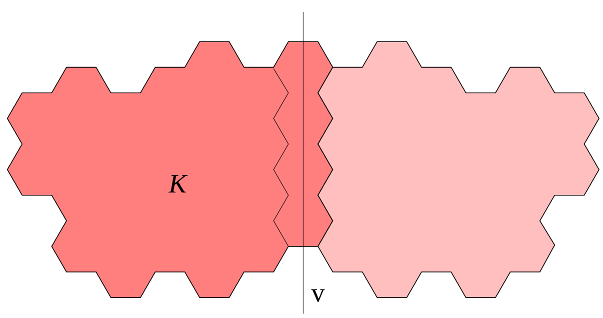


FIGURE 8. Polyomino convexe v-symétrique

$$\begin{aligned} K(x,q,v,t) &= C_{00}(x,q,v,t) + 2C_{10}(x,q,v,t) + C_{1,1}(x,q,v,t) \qquad (4.1) \\ &= \sum_{m\geq 1} K_m(x,q,t)v^m \end{aligned}$$

et

$$|Fix(\mathrm{v})|_{q,t} = \sum_{m\geq 1}\frac{1}{q^m t^{2m+1}}K_m(1,q^2,t^2). \qquad (4.2)$$

4.2. Symétrie horizontale. La classe S des polyominos convexes h-symétriques se partage en trois classes: S_a et S_b, suivant qu'on peut trouver ou non dans la partie oscillante un polyomino *pointe de flèche* (figures 9a et 9b) et la classe S_c, s'il n'existe pas de partie oscillante.

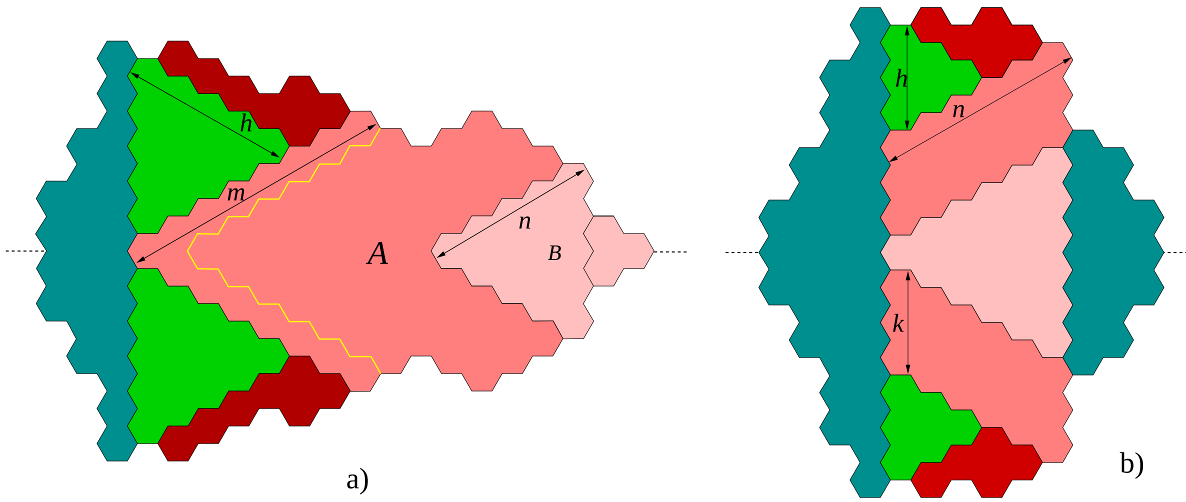


FIGURE 9. Polyominos convexes h-symétriques

Pour construire un polyomino de la classe *pointes de flèches*, notée A, on démarre avec un triangle de coté n auquel un tas symétrique est possiblement attaché pour former la phase H_{22}; notons B, cette classe de polyominos de départ. On construit A à partir de B en attachant successivement des bandes en forme de V par la gauche, comme illustré à la figure 9a. On a donc

$$B(s,x,q,t) = sxqt^3 + \sum_{n\geq 2} s^n x^n q^{n(n+1)/2} t^{3n} \mathrm{TS0}_{n-3}(xt,q), \tag{4.3}$$

où la variable s marque la taille de la partie supérieure gauche du dernier V, et la série génératrice $A(s) = A(s,x,q,t)$ est caractérisée par l'équation fonctionelle suivante, qui se résout par la méthode habituelle:

$$A(s) = B(s) + s^2x^2q^3t^4 \frac{A(1) - A(sq^2)}{1 - sq^2} \tag{4.4}$$

Finalement, posant $A(s,x,q,t) = \sum_{m\geq 1} A_m(x,t,q)s^m$ et tenant compte des décorations supplémentaires et des deux cas de parité de la première colonne de la partie oscillante, on a

$$\begin{aligned} S_a(x,t,q) &= \sum_{h\geq 0} q^{h(h+1)} t^{2h+2} \mathrm{TS}_{2h+2}(xt,q) \sum_{m\geq h+1} {\textstyle\begin{bmatrix} m-1 \\ h \end{bmatrix}_{q^2}} A_m(x,t,q) \\ &\quad + \sum_{h\geq 1} q^{h(h+1)} t^{2h+3} \mathrm{TS}_{2h+1}(xt,q) \sum_{m\geq h} {\textstyle\begin{bmatrix} m \\ h \end{bmatrix}_{q^2}} A_m(x,t,q). \end{aligned} \tag{4.5}$$

Les calculs de $S_b(x,t,q)$ et de $S_c(x,t,q)$ sont plus simples et on trouve

$$\begin{aligned} S_b(x,t,q) = &\sum_{n\geq 1} x^n q^{\binom{n+1}{2}} t^{3n} \sum_{k\geq 1} q^{2kn} t^{4k} \mathrm{TS0}_{n+2k-3}(xt,q) \sum_{h=0}^{n-1} t^{2h+2} q^{h(h+1)} {\textstyle\begin{bmatrix} n-1 \\ h \end{bmatrix}_{q^2}} \mathrm{TS}_{2k+2h+2}(x,t,q) \\ &+ \sum_{n\geq 0} x^n q^{\binom{n+1}{2}} t^{3n} \sum_{k\geq 1} q^{2k(n+1)} t^{4k+1} \mathrm{TS0}_{n+2k-3}(xt,q) \sum_{h=0}^{n} t^{2h+2} q^{h(h+1)} {\textstyle\begin{bmatrix} n \\ h \end{bmatrix}_{q^2}} \mathrm{TS}_{2k+2h+1}(x,t,q) \end{aligned} \tag{4.6}$$

et

$$S_c(x,t,q) = \sum_{h\geq 1} t^{2h}\mathrm{TS}_h(xt,q)\mathrm{TS0}_{h-3}(xt,q). \tag{4.7}$$

Ainsi

$$|Fix(\mathrm{h})|_{q,t} = S_a(1,t,q) + S_b(1,t,q) + S_c(1,t,q). \tag{4.8}$$

5. Symétries de rotation

5.1. Symétrie par rapport à la rotation r de $\pi/3$ radian. Les polyominos symétriques par rapport à la rotation r de $\pi/3$ radian sont essentiellement des polyominos formés de grands hexagones décorés par des tas T0. On trouve

$$|Fix(\mathrm{r})|_{q,t} = \sum_{h\geq 1} t^{3(2h-1)}q^{3h(h-1)+1}\mathrm{T0}_{h-1}(t^6,q^6). \tag{5.1}$$

5.2. Symétrie par rapport à la rotation r^2, de $2\pi/3$ radian. Ce cas est plus complexe. Il faut d'abord distinguer le cas où le centre de rotation est au milieu d'un hexagone de celui où il est en un sommet. Ceci détermine deux sous-classes, notées $\mathcal{P}$ et $\mathcal{Q}$. Dans le premier cas, il y a trois sous-cas selon que $h_1 > h_2$, $h_2 > h_1$ ou $h_1 = h_2$, correspondant aux trois sous-classes $\mathcal{P}_1$, $\mathcal{P}_2$ et $\mathcal{P}_3$. La figure 10 illustre le premier cas $\mathcal{P}_1$. Cette figure ne représente qu'un tiers du polyomino r^2-symétrique, sa région fondamentale. On voit apparaître une nouvelle classe de polyominos hexagonaux convexes, les dirigés (vers le haut), a base diagonale, notée $\mathcal{D}$, qu'il faut d'abord dénombrer avant de déterminer $\mathcal{P}_1(q,t)$. Nous manquons d'espace ici pour donner les détails de ces calculs. Notons que pour des raisons de symétrie, $\mathcal{P}_2(q,t) = \mathcal{P}_1(q,t)$. Les calculs sont semblables pour la classe $\mathcal{Q}$, et

$$|Fix(\mathrm{r}^2)|_{q,t} = 2\mathcal{P}_1(q,t) + \mathcal{P}_3(q,t) + 2\mathcal{Q}_1(q,t) + \mathcal{Q}_3(q,t). \tag{5.2}$$

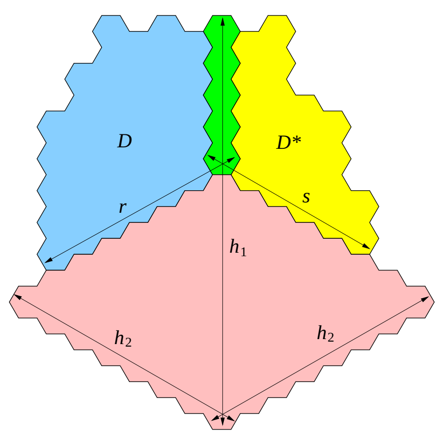


FIGURE 10. Région fondamentale d'un polyomino convexe r^2-symétrique dans $\mathcal{P}_1$

5.3. Symétrie par rapport à la rotation r^3, de π radian. Si le centre de rotation est au milieu d'une arête, il y a trois cas similaires correspondant aux trois types d'arêtes. Prenons le cas de l'arête horizontale et notons $\mathcal{A}$ la classe correspondante. On a

$$\mathcal{A}(x,q,t) = \sum_{k\geq 1} \frac{1}{xq^{2k}t^{4k+1}} (C_{00,2k} + 2C_{01,2k} + C_{11,2k} + 2C_{02,2k})(x^2,q^2,t^2) \tag{5.3}$$

Si le centre de rotation est au milieu d'un hexagone, on note $\mathcal{H}$ la classe correspondante et on a

$$\mathcal{H}(x,q,t) = \sum_{k\geq 0} \frac{1}{xq^{2k+1}t^{4k+3}} (C_{00,2k+1} + 2C_{01,2k+1} + C_{11,2k+1} + 2C_{02,2k+1})(x^2,q^2,t^2). \tag{5.4}$$

Finalement,

$$|Fix(\mathrm{r}^3)|_{q,t} = 3\mathcal{A}(1,q,t) + \mathcal{H}(1,q,t). \tag{5.5}$$

6. Conclusion

Il est maintenant possible d'utiliser la formule de Burnside (2.12), avec $\mathcal{F} = C$, pour dénombrer les polyominos convexes libres, c'est-à-dire à rotation et réflexion près, selon l'aire et le demi-périmètre. Quelques résultats numériques sont présentés dans les tableaux 1 et 2, selon l'aire seule (jusqu'à l'aire 20) ou le demi-périmètre seul (jusqu'au demi-périmètre 16). Voir sous la colonne *Orbites*. Ces résultats ont été vérifiés expérimentalement par une énumération exhaustive par ordinateur.

Il est également possible de dénombrer les polyominos convexes asymétriques ou ayant exactement les symétries d'un sous-groupe donné H de $\mathcal{D}_6$, à l'aide de l'inversion de Möbius dans le treillis des sous-groupes de $\mathcal{D}_6$. Ce treillis et sa fonction de Möbius sont bien décrits dans la thèse de Stockmeyer [**14**] pour tout groupe diédral $\mathcal{D}_n$. Nous suivons ici cette nomenclature et quelques résultats se trouvent dans les tableaux 1 et 2. On voit bien sur ces tableaux que presque tous les polyominos convexes sont asymétriques.

On trouvera plus de détails dans la forme longue de ce résumé substantiel sur le site Web des archives mathématiques `http://arxiv.org` arXiv:math.CO/0403168.

References

[1] Mireille Bousquet-Mélou. A method for the enumeration of various classes of column-convex polygons. *Discrete Mathematics*, 154:1–25, 1996.

[2] A. Denise, C. Dürr, and F. Ibn-Majdoub-Hassani. Enumération et génération aléatoire de polyominos convexes en réseau hexagonal. In *Proceedings of the 9th FPSAC Conference*, pages 222–234, Universität Wien, 1997.

[3] S. Feretić and D. Svrtan. On the number of column-convex polyominoes with given perimeter and number of columns. In *Proceedings of the 5th FPSAC Conference*, pages 201–214, University of Florence, 1993.

[4] D. Gouyou-Beauchamps and P. Leroux. Enumeration of symmetry classes of convex polyominoes on the honeycomb lattice. *arXiv:math.CO/0403168*, pages 1–21, 2004.

[5] A. J. Guttmann and I. G. Enting. The number of convex polygons on the square and honeycomb lattices. *J. Phys. A: Math. Gen.*, 21:L467–L474, 1988.

[6] F. Ibn-Majdoub-Hassani. Combinatoire des polyominos et des tableaux décalés oscillants. PhD Thesis, Université de Paris Sud, Orsay, November 1996.

[7] P. Leroux and É. Rassart. Enumeration of symmetry classes of parallelogram polyominoes. *Ann. Sci. Math. Québec*, 25(1):71–90, 2001.

[8] P. Leroux, É. Rassart, and A. Robitaille. Enumeration of symmetry classes of convex polyominoes in the square lattice. *Advances in Applied Mathematics*, 21:343–380, 1998.

[9] K. Y. Lin and S. J. Chang. Rigorous results for the number of convex polygons on the square and honeycomb lattices. *J. Phys. A: Math. Gen.*, 21:2635–2642, 1988.

[10] K. Y. Lin and F. Y. Wu. Unidirectional convex polygons on the honeycomb lattice. *J. Phys. A: Math. Gen.*, 23:5003–5010, 1990.

[11] V. Privman and N. M. Švrakić. *Directed Models of Polymers, Interfaces, and Clusters: Scaling and Finite-Size Properties.* Lecture Notes in Physics, Springer-Verlag, Volume 338, 1989.

TABLE 1. Classes de symétries des polyominos hexagonaux convexes selon l'aire

Aire	id	h	v	r	r^2	r^3	Orbites	$\mathcal{D}_6$	F_{31}	H_{31}	D_{21}	Asym
1	1	1	1	1	1	1	1	1	1	1	1	0
2	3	1	1	0	0	3	1	0	0	0	1	0
3	11	3	3	0	2	3	3	0	2	0	1	0
4	38	2	4	0	2	12	6	0	0	2	2	24
5	120	6	10	0	0	12	15	0	0	0	2	72
6	348	6	12	0	6	42	38	0	2	0	2	264
7	939	9	27	1	3	37	91	1	1	3	3	816
8	2412	12	30	0	0	126	222	0	0	0	4	2184
9	5973	17	63	0	12	99	528	0	0	0	3	5640
10	14394	20	66	0	6	336	1250	0	2	4	4	13836
11	34056	30	142	0	0	252	2902	0	0	0	6	33324
12	79602	38	140	0	18	840	6751	0	2	0	4	78240
13	184588	46	310	1	13	616	15525	1	1	5	8	182952
14	426036	62	286	0	0	2028	35759	0	0	0	8	423012
15	980961	69	665	0	30	1461	82057	0	2	0	7	977316
16	2256420	100	580	0	18	4788	188607	0	0	6	8	2249640
17	5189577	115	1441	0	0	3435	433140	0	0	0	11	5181540
18	11939804	154	1184	0	50	11142	996255	0	2	0	12	11924676
19	27485271	175	3145	1	27	8005	2291941	1	1	7	13	27467376
20	63308532	238	2458	0	0	25800	5278535	0	0	0	16	63274740

TABLE 2. Classes de symétries des polyominos hexagonaux convexes selon le demi-périmètre

$\frac{1}{2}$pér.	id	h	v	r	r^2	r^3	Orbites	$\mathcal{D}_6$	F_{31}	H_{31}	D_{21}	Asym
3	1	1	1	1	1	1	1	1	1	1	1	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	3	1	1	0	0	3	1	0	0	0	1	0
6	2	2	0	0	2	0	1	0	2	0	0	0
7	12	2	4	0	0	6	3	0	0	0	2	0
8	18	2	0	0	0	0	2	0	0	0	0	12
9	59	5	9	1	5	19	11	1	3	3	3	24
10	120	8	0	0	0	0	12	0	0	0	0	96
11	318	10	24	0	0	48	39	0	0	0	6	204
12	714	14	0	0	12	0	65	0	4	0	0	672
13	1743	25	59	0	0	129	177	0	0	0	7	1368
14	4008	36	0	0	0	0	343	0	0	0	0	3900
15	9433	53	143	2	28	323	867	2	6	8	15	8616
16	21672	76	0	0	0	0	1825	0	0	0	0	21444

[12] N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences.* Published electronically at `http://www.research.att.com/~njas/sequences/`, 2003.
[13] N. J. A. Sloane and S. Plouffe. *The Encyclopedia of Integer Sequences.* Academic Press, Inc., 1995.
[14] P. K. Stockmeyer. Enumeration of graphs with prescribed automorphism group. Ph.D. Thesis, University of Michigan, Ann Arbor, 1971.
[15] M. Vöge, A. J. Guttmann, and I. Jensen. On the number of benzenoid hydrocarbons. *J. Chem. Inf. Comput. Sci.*, 42:456–466, 2002.

LRI, Université de Paris-sud et CNRS, Orsay, France, F91405
E-mail address: dgb@lri.fr
URL: http://www.lri.fr/∼dgb

LaCIM et Département de Mathématiques, Université du Québec à Montréal, Canada, H3C3P8
E-mail address: leroux.pierre@uqam.ca
URL: http://www.lacim.uqam.ca

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

A characterization of the simply-laced FC-finite Coxeter groups

HAGIWARA Manabu, Masao ISHIKAWA, and Hiroyuki TAGAWA

Abstract. *We call an element of a Coxeter group fully covering if its length is equal to the number of the elements covered by it. For the Coxeter groups of type A, an element is fully covering if and only if it is 321-avoiding. In this sense it can be regarded as an extended notion of 321-avoiding. It also can be seen from the definition that a fully covering element is always fully commutative. Also, we call a Coxeter group bi-full when an element of the group is fully commutative if and only if it is fully covering. We show that the bi-full Coxeter groups are of type A, D, E. Note that we do not restrict the type E to E_6, E_7, and E_8. In other words, Coxeter groups of type $E_9, E_{10}, \dots$ are also bi-full. According to a result of Fan, a Coxeter group is a simply-laced FC-finite Coxeter group if and only if it is a bi-full Coxeter group.*

1. Introduction

It is needless to say that the notion of Coxeter groups appears in various mathematical fields and have widely interested people, but they, themselves, are still very interesting objects for study. It is also well known that the Coxeter groups of type A, D, E_6, E_7, and E_8, i.e. simply laced Weyl groups, share a lot of interesting properties and attract many researchers. Usually, when we say the groups of type E_n, we often tend to restrict ourselves to $n = 6, 7$ and 8 cases. However, we sometimes find that the general Coxeter groups of type E_n, which are not restricted to $n = 6, 7, 8$, also share some very interesting properties. For example, we can mention FC-finite Coxeter groups. An element of a Coxeter group is said to be *fully commutative* if any reduced expression for it can be obtained from any other by transposing adjacent commuting generators. A *FC-finite* Coxeter group is, by definition, a Coxeter group which has a *finite* number of fully commutative elements. C. K. Fan proved that the simply-laced FC-finite irreducible Coxeter groups are only of type A, D, and E, and vice versa ([**3**, Proposition 2.]). Here, of course, the Coxeter groups of type E means of type E_n, which are not restricted to $n = 6, 7, 8$.

In this paper, we call an element of a Coxeter group *fully covering* if its length equals the number of elements covered by it. This notion was already appeared in [**4**, Theorem 1]. Further we say a Coxeter group is *bi-full* when each element of the group is fully commutative if and only if it is fully covering. The purpose of our paper is to characterize the bi-full Coxeter groups. Although it is a consequence of the result by Fan, that the Coxeter groups of type A, D, E_6, E_7, and E_8 are bi-full (see [**4**, Theorem 1]) and the Coxeter group of type $\tilde{A}_2$ is not bi-full (see [**4**, Conclusion]), his results do not give a complete characterization of all the bi-full Coxeter groups. In fact, one of our main goals in this paper is to prove that the irreducible bi-full

1991 *Mathematics Subject Classification.* Primary 06A07; Secondary 20F55.

Key words and phrases. Coxeter groups, Bruhat ordering, fully commutative elements, fully covering elements.

The third author was partially supported by Grant-in-Aid for Scientific Research (C) No.15540028, Japan Society for the Promotion of Science.

Coxeter groups are only of type A, D, E and vice versa. Accordingly it immediately implies that a Coxeter group is a simply-laced FC-finite if and only if it is bi-full (see Theorem 2.9).

Now we recall some notation from the symmetric groups. An element σ of the symmetric group of degree n is called a *321-avoiding* if there is no triple $1 \leq i < j < k \leq n$ such that $\sigma(i) > \sigma(j) > \sigma(k)$. Our original motivation is to regard the notion of "fully covering" as an extension of that of 321-avoiding (see [**1**]) to any Coxeter groups. In fact, it is a consequence of some well known facts that, if we restrict our attention to the Coxeter groups of type A, then a permutation is fully covering if and only if it is 321-avoiding. Actually, this observation was the starting point of our research. We should note that there is another interesting extension of the notion of 321-avoiding. In [**5**], Green extended the notion to the affine permutation groups from another point of view, whereas our extension, i.e. fully covering, and his definition of 321-avoiding in the affine permutation groups are not equivalent. Indeed, he defined the notion of 321-avoiding permutations for any affine permutation groups and showed that an element is 321-avoiding if and only if it is fully commutative. In [**6**, Thm. 5.1] Hagiwara proved that a 321-avoiding permutation in an affine permutation group is a minuscule element of the group, and vice versa. Meanwhile, it is not hard to see that, for the affine permutation groups, fully covering implies fully commutative, but the reverse is not true.

We conclude this section by making a remark on the Kazhdan-Lusztig theory. Let W be any Coxeter group and let x, w be elements of W. Let $p_1(x, w)$ denote the coefficient of degree 1 in the Kazhdan-Lusztig polynomial $P_{x,w}$ for the interval $[x, w]$ in the Bruhat ordering of W. M. Dyer showed that $p_1(e, w) = c^-(w) - |\mathrm{supp}(w)|$ and $p_1(e, w) \geq 0$ (see [**2**]). Thus, if W is of type A, D, E and w is a fully commutative element of W, then we can rewrite this result as $p_1(e, w) = \ell(w) - |\mathrm{supp}(w)|$.

This paper is organized as follows: In §2, we recall and provide some basic terminology. In §3, we collect some important properties of a fully commutative element. In §4, we show that Coxeter groups of type A, D, and E are bi-full. Moreover, we show that a Coxeter group which is neither of type A, D nor E cannot be bi-full.

2. Preliminaries and Notations

Now we start with notation and preliminaries again from scratch as this paper become more comprehensive even though it might be slightly repetitious. Throughout this paper, we assume that (W, S) always denote a *Coxeter system* with finite *generator set* S and *Coxeter matrix* $M = [m(s, t)]_{s,t \in S}$. Thus $m(s, t)$ is the order of st in W (possibly $m(s, t) = \infty$). When $m(s, t) = 2$, we say s and t commute. *The Coxeter graph* Γ of (W, S) is, by definition, the simple graph with vertex set S and edges between two non-commuting generators. We may regard (Γ, M) as a weighted graph by interpreting the entries of M as a weight function on the edges of Γ, and call it *the Coxeter diagram* of (W, S). We illustrate a Coxeter diagram by labeling an edge (s, t) of the Coxeter graph Γ with $m(s, t)$ when $m(s, t) \geq 4$.

We denote the set of integers by $\mathbb{Z}$ and denote the set of positive integers by $\mathbb{Z}_{>0}$. For a positive integer n, we put $[n] := \{1, 2, \ldots, n\}$. For a set A, we denote its cardinality by $|A|$ or $\sharp A$.

Notation 2.1. Let w be an element of W and let e be the identity of W. A *length function* ℓ is a mapping from W to $\mathbb{Z}$ defined by $\ell(e)$ equals 0 and $\ell(w)$ equals the smallest m such that there exist elements $s_1, s_2, \ldots, s_m$ of S satisfying $w = s_1 s_2 \ldots s_m$ for $w \neq e$. We call $l(w)$ the *length* of w. Let $x_1, x_2, \ldots, x_m$ be elements of W. If we have $w = x_1 x_2 \ldots x_m$ and $\ell(x_1 x_2 \ldots x_m) = \ell(x_1) + \ell(x_2) + \cdots + \ell(x_m)$, then we call $(x_1, x_2, \ldots, x_m)$ an *extended reduced word* for w and $w = x_1 x_2 \ldots x_m$ an *extended reduced expression* for w. Note that we do not assume that $x_1, x_2, \ldots, x_m$ belong to S. In particular, we call the word $(x_1, x_2, \ldots, x_m)$ *a reduced word for* w and $w = x_1 x_2 \ldots x_m$ *a reduced expression for* w if $x_i \in S$ for $i = 1, 2, \ldots, m$ and $\ell(w) = m$.

Definition 2.2. Let (W, S) and $M = [m(s, t)]_{s,t \in S}$ be as above.

(i) If $\{m(s,t)|s,t \in S\} \subseteq \{1,2,3\}$, then we say (W,S) (resp. W) is a *simply-laced* Coxeter system (resp. a simply-laced Coxeter group).
(ii) If there exist elements $s_1, s_2, \dots, s_m$ of S ($m \geq 3$) such that $m(s_m, s_1) \geq 3$, $m(s_i, s_{i+1}) \geq 3$ for all $i \in [m-1]$, then we say (W,S) (and W) is *cyclic*. If not, then we say it is *acyclic*.
(iii) If the Coxeter graph Γ of (W,S) is connected, then we say (W,S) (and W) is *irreducible*.

Definition 2.3. Let (W,S) be a Coxeter system whose Coxeter diagram is given by Figure 1 (resp. Figure 2). Then we call (W,S) a Coxeter system of type E_{r+4} for $r \geq 2$ (resp. type D_{r+3} for $r \geq 1$).

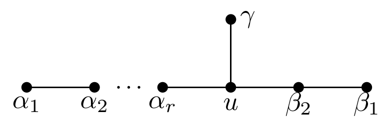


FIGURE 1. Coxeter diagram of type E_{r+4}

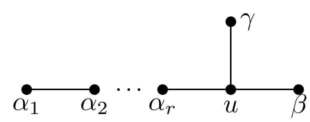


FIGURE 2. Coxeter diagram of type D_{r+3}

For integers $m \geq 0$ and $s,t \in S$, set $\langle s,t\rangle_m$ to be the word $\underbrace{(s,t,s,t,s,\dots)}_{m}$ of length m. We introduce an equivalence relation $\approx$ between the words of S generated by the braid relations $\langle s,t\rangle_{m(s,t)} \approx \langle t,s\rangle_{m(s,t)}$ for all $s,t \in S$ such that $m(s,t) < \infty$. It is an important fact that any reduced word for w can be obtained from any other by the braid relations, i.e. the set of reduced words for w consists of one equivalence class with respect to $\approx$. Following [**9**], we also introduce a weaker equivalence relation $\sim$ on the set of the words of S generated by the relations $(s,t) \sim (t,s)$ for all $s,t \in S$ such that $m(s,t) = 2$. We say that w is *fully commutative* if the set of reduced words for w consists of just one equivalence class with respect to $\sim$, i.e. any reduced word for w can be obtained from any other by transposing adjacent commuting pairs. For a Coxeter group W, we put

$$W^{FC} := \{w \in W | w \text{ is fully commutative}\}.$$

If the cardinality of W^{FC} is finite, then we say (W,S) is (resp. W) a *FC-finite* Coxeter system (resp. a FC-finite Coxeter group).

From now on, we denote a Coxeter group of type X by $W(X)$.

Theorem 2.4 (C. K. Fan)**.** *If W is an irreducible simply-laced FC-finite Coxeter group, then W should be one of $W(A_n)$, $W(D_{n+3})$ and $W(E_{n+5})$ for some $n \geq 1$ (see [**3**]).*

We recall the definition of the Bruhat ordering. Let $T := \{wsw^{-1}|s \in S, w \in W\}$ be the set of reflections in W. Write $y \to z$ if $z = yt$ for some $t \in T$ with $\ell(y) < \ell(z)$. Then define $y < z$ if there is a sequence $y = w_0 \to w_1 \to \cdots \to w_m = z$. It is clear that the resulting relation $y \leq z$ is a partial ordering of W, and we call it the *Bruhat ordering*. We say z *covers* y (or equivalently y is covered by z), denote by $y \lessdot z$, if $y < z$ and $\ell(y) = \ell(z) - 1$.

The following is a well known characterization of the Bruhat ordering which is called the *subword property*. Give a reduced expression $w = s_1 s_2 \cdots s_m$ for $w \in W$, let us call the products (not necessarily reduced, and possibly empty) of the form $s_{i_1} s_{i_2} \dots s_{i_q}$ ($1 \leq i_1 < i_2 < \cdots < i_q \leq m$) *the subexpressions of*

$s_1 s_2 \cdots s_m$. Let $w = s_1 \ldots s_m$ be a fixed, but arbitrary reduced expression for $w \in W$. Then $x \leq w$ if and only if x can be obtained as a subexpression of this reduced expression.

The ordering handled in this paper is always assumed to be the Bruhat ordering.

Notation 2.5. For $w \in W$, we put

$$\begin{aligned} \mathrm{supp}(w) : &= \{s \in S | s \leq w\}, \\ C^-(w) : &= \{x \in W | x \lessdot w\}, \\ c^-(w) : &= |C^-(w)|. \end{aligned}$$

Definition 2.6. For $w \in W$, we call w *fully covering* if $\ell(w) = c^-(w)$.

By the above subword property, the reader easily see that $w \in W$ is fully covering if and only if, given any reduced expression $w = s_1 \cdots s_m$, deleting any one generator from this expression always reduce its length by 1.

If $w \in W$ is not fully commutative, then there exists a reduced expression $w = s_1 \ldots s_m$ including a braid relation $\langle s, t \rangle_{m(s,t)}$ with $m(s,t) \geq 3$. Thus, by discarding one of s or t from this braid relation, we obtain an element $w' < w$ of the form $w' = s_1 \ldots \widehat{s_i} \ldots s_m$ which is not reduced. This immediately shows w is not fully covering, and implies the following proposition.

Proposition 2.7. *A fully covering element w of W is fully commutative.*

The reverse is not always true, and we will give some examples later. If the reverse is true, i.e. a fully commutative element $w \in W$ is always fully covering, we say W (resp. (W, S)) is a *bi-full* Coxeter group (resp. a *bi-full* Coxeter system).

Remark 2.8. Let $(W_1, S_1), (W_2, S_2)$ be bi-full Coxeter systems (resp. FC-finite Coxeter systems). If we have $S_1 \cap S_2 = \emptyset$ and $s_1 s_2 = s_2 s_1$ for any $(s_1, s_2) \in S_1 \times S_2$ then $(W_1 W_2, S_1 \cup S_2)$ is also a bi-full Coxeter system (resp. a FC-finite Coxeter system).

The main result of this paper is the following theorem.

Theorem 2.9. *W is a simply-laced FC-finite Coxeter group if and only if W is a bi-full Coxeter group.*

By Remark 2.8, we can easily reduce Theorem 2.9 to the irreducible cases. By Theorem 2.4, we already know that an irreducible simply-laced FC-finite Coxeter group must be one of type A, D or E. Thus it is enough to show the following theorem to complete the proof of Theorem 2.9.

Theorem 2.10. *Let W be an irreducible Coxeter group. Then, W is bi-full if and only if it is either of type A, D or E.*

By Proposition 2.7, if the following two claims hold then we can obtain Theorem 2.10.

Claim 1. Any fully commutative element of the Coxeter group of type E is fully covering (Theorem 4.3).

Claim 2. If W is neither of type A, D nor E, then there is an element in W which is fully commutative, but not fully covering (Theorem 4.9).

We often use the following notation and facts which the reader may be already familiar with (see [**8**]). For any subset $J \subset S$, let $W_J = \langle J \rangle$ denote the subgroup of W generated by all $s \in J$, which is usually called *the parabolic subgroup of W generated by J*. Put $W^J = \{x \in W | \ell(xy) = \ell(x) + \ell(y) \text{ for all } y \in W_J\}$ and ${}^J W = \{x \in W | \ell(yx) = \ell(y) + \ell(x) \text{ for all } y \in W_J\}$, then the following fact shows W^J (resp. ${}^J W$) is the set of left (resp. right) coset representatives of W with respect to W_J.

Fact 2.11. (i) For any $w \in W$, there is a unique pair $(x, y) \in W^J \times W_J$ such that $w = xy$.
(ii) For any $w \in W$, there is a unique pair $(y, z) \in W_J \times {}^J W$ such that $w = yz$.

3. Properties of fully commutative elements

In this section, we collect some basic and important properties of fully commutative elements which will be concerned with the rest of the paper. Throughout this section we assume that W always denotes any Coxeter group if there is no special mention.

By the definition of the fully commutativity, we have the following.

Lemma 3.1.

(i) *Let w be an element of W. Let $s_1 s_2 \dots s_m$ and $s'_1 s'_2 \dots s'_m$ be reduced expressions for w. If w is fully commutative then we have*

$$\{s_1, s_2, \dots, s_m\} = \{s'_1, s'_2, \dots, s'_m\} \text{ as multisets.}$$

(ii) *Assume $m(s,t)$ is odd or 2 for any $s, t \in S$. For any $w \in W$, w is fully commutative if and only if we have $\{s_1, s_2, \dots, s_m\} = \{s'_1, s'_2, \dots, s'_m\}$ as multisets for any reduced expressions $w = s_1 s_2 \dots s_m = s'_1 s'_2 \dots s'_m$.*

(iii) *An element is fully commutative if it has a unique reduced expression.*

(iv) *Let xyz be an extended reduced expression for w. If w is fully commutative then y is also fully commutative.*

(v) *Let W be a simply-laced Coxeter group and let w be an element of W. Then w is not fully commutative if and only if there is a reduced expression $s_1 s_2 \dots s_m$ for w such that $s_i = s_{i+2}$ for some $1 \le i \le m-2$.*

The following lemma is a key lemma of this paper.

Lemma 3.2. *Let w be a fully commutative element and let $s_1 s_2 \dots s_r$ be a reduced expression for w ($r \ge 2$). If we have $w = s s_1 s_2 \dots s_{r-1}$ for some $s \in S$ then we have the followings:*

(i) $s = s_r$,

(ii) $s s_j = s_j s$ *for any* $j \in [r-1]$,

(iii) $s \not\le s_1 s_2 \dots s_{r-1}$.

The following corollary is useful to find an element which is fully commutative and is not fully covering.

Corollary 3.3. *Let w be an element of W and let $s_1, s_2, \dots, s_m$ be elements of S such that $w = s_1 s_2 \dots s_m$. Note that we do not assume that $s_1 s_2 \dots s_m$ is a reduced expression for w. We define a condition (FC) as follows:*

(FC) *If there exists a pair (i,j) of integers such that $i < j$ and $s_i = s_j$, then there exists a pair (a,b) of integers such that $i < a < b < j$, $s_a s_i \ne s_i s_a$ and $s_b s_i \ne s_i s_b$.*

Then we have the followings.

(i) *If $s_1 s_2 \dots s_m$ satisfies the condition (FC) then $s_1 s_2 \dots s_m$ is a reduced expression for w and w is fully commutative.*

(ii) *If W is a simply-laced Coxeter group, $s_1 s_2 \dots s_m$ is a reduced expression for w and w is fully commutative, then $s_1 s_2 \dots s_m$ satisfies the condition (FC).*

By Corollary 3.3, we have the following.

Corollary 3.4. *Let W be a simply-laced Coxeter group and let w be an element of W such that $\ell(w^2) = 2\ell(w)$ and w^2 is fully commutative. Then for any $k \in \mathbb{Z}_{>0}$ we have $\ell(w^k) = k\ell(w)$ and w^k is fully commutative. In particular, W is not a FC-finite Coxeter group.*

The following lemma holds for any Coxeter system (W,S).

Lemma 3.5. *Let (W,S) be a Coxeter system and let x be an element of W. Let s_1, s_2 be elements of S such that $s_1 s_2 x$ is an extended reduced expression and that $s_2 s_1 s_2$ is reduced (i.e. $m(s_1, s_2) \ge 3$). If we have $s_1 \notin supp(x)$ then $s_2 s_1 s_2 x$ is an extended reduced expression.*

The following lemma holds for any simply-laced Coxeter system.

Lemma 3.6. *Let (W,S) be a simply-laced Coxeter system, and let w be a fully commutative element of W. If $s_1 s_2 \dots s_m$ is a reduced expression for w, then $s_1 \widehat{s_2} s_3 \dots s_m$ is reduced.*

4. Main results

There is a method to derive the following proposition from a well-known fact on 321-avoiding permutations of the symmetric groups. However, here we give a sketch of our proof without the notion of 321-avoiding.

Proposition 4.1. *Let W be a Weyl group of type A_n. Then a fully commutative element w of W is fully covering.*

Let $(s_1, s_2, \ldots, s_m) \in S^*$ be any word from S (i.e. an element of the free monoid generated by S), and let α be an element of S. Then we use the notation:

$$g_\alpha((s_1, s_2, \ldots, s_m)) := \sharp\{i \in [m] \mid s_i = \alpha\}.$$

By Lemma 3.1(i), when w is fully commutative, we can define

$$g_\alpha(w) := g_\alpha((s_1, s_2, \ldots, s_m))$$

without ambiguity where $(s_1, s_2, \ldots, s_m)$ is a reduced word for w.

Lemma 4.2. *Let $w = s_1 s_2 \ldots s_m$ be a reduced expression for $w \in W$. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ be a subset of supp(w) satisfying the following conditions (1),(2), and (3).*

(1) *$\alpha_i s = s\alpha_i$ for any $i \in [r]$ and for any $s \in$ supp$(w) - \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$.*

(2) *$\langle \alpha_1, \alpha_2, \ldots, \alpha_r \rangle$ is a Weyl group of type A_r, whose Coxeter graph is given by Figure 3. (i.e. $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is a connected component of the Coxeter graph of W and α_1 is one of its endpoints.)*

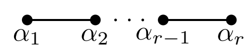


FIGURE 3. Coxeter diagram of type A_r

(3) *$g_{\alpha_1}((s_1, s_2, \ldots, s_m)) \geq 2$.*

Then w is not fully commutative.

Here we don't have enough space to give a detailed proof of this lemma, which the reader can find in our original paper [**7**]. The proof of Proposition 4.1 reduce to this lemma. This fact was the starting point of our main result. In fact the idea of the proof of the following theorem resides in a similar method for type E, while the proof needs more complicated computations. Thus it is worth describing the proof in the case of type A.

Theorem 4.3. *Let W be a Coxeter group of type E and let w be an element of W. If w is fully commutative then w is fully covering.*

By a similar argument as above, the proof of Theorem 4.3 reduce to the following two lemmas, which is a fundamental idea of our proof. Thus the proofs of the following lemmas are main goal of our paper.

Lemma 4.4. *Let (W, S) be a Coxeter system of type D_{r+3}, whose Coxeter graph is given by Figure 2 ($r \geq 1$). (i.e. α_1, β and γ are the endpoints designated in the figure.) Put $J := S - \{\alpha_1\}$. Let $w \in {}^J W$ be a fully commutative element and let $s_1 s_2 \ldots s_m$ be a reduced expression for w. If supp(w) includes the endpoints α_1, β, γ, then the followings hold.*

(i) *$r + 3 \leq m$, $s_1 s_2 \ldots s_{r+3} = \alpha_1 \alpha_2 \ldots \alpha_r u \beta \gamma$.*

(ii) *For any $s \in J$, sw is not fully commutative.*

(iii) *$m \leq 2r + 4$.*

(iv) *If $m \geq r + 4$ then we have $s_{r+4} s_{r+5} \ldots s_m = u \alpha_r \alpha_{r-1} \ldots \alpha_{2r+5-m}$ where $\alpha_{r+1} = u$.*

Lemma 4.5. *Let (W, S) be a Coxeter system of type E_{r+4} ($r \geq 1$) whose Coxeter graph is designated in Figure 1 (i.e. α_1, β_1 and γ are the endpoints in the figure). Put $J := S - \{\alpha_1\}$. Let $w \in {}^J W$ be a fully commutative element and let $s_1 s_2 \ldots s_m$ be a reduced expression for w. Then the followings hold.*

(i) *If supp(w) includes all the end points* α_1*,* β_1 *and* γ*, then sw is not fully commutative for all* $s \in J$*.*
(ii) *Assume that we have* $\alpha_1, \beta_2, \gamma \in supp(w)$*,* $\beta_1 \notin supp(w)$ *and* $s \in J$*. If sw is fully commutative then we have* $s = \beta_1$*.*
(iii) *Assume that we have* $g_{\alpha_1}(w) \geq 2$ *and we have* $s \in J$ *such that sw is fully commutative. Then we have* $w = \alpha_1\alpha_2 \ldots \alpha_r u\gamma\beta_2 u\alpha_r \ldots \alpha_2\alpha_1$ *and* $s = \beta_1$*.*
(iv) *Assume that we have* $g_{\alpha_1}(w) \geq 3$ *and we have* $w \in {}^JW \cap W^J$*. Then there exists an element v of* $W_{S-\{\alpha_1,\alpha_2\}}$ *such that*

$$(\alpha_1\alpha_2 \ldots \alpha_r u\gamma\beta_2 u\alpha_r\alpha_{r-1} \ldots \alpha_2)\alpha_1\beta_1 v\beta_1(\alpha_2 \ldots \alpha_r u\gamma\beta_2 u\alpha_r\alpha_{r-1} \ldots \alpha_1)$$

is an extended reduced expression for w and that $\beta_1 v\beta_1 \in {}^{S-\{\beta_1\}}W \cap W^{S-\{\beta_1\}}$*.*

Remark 4.6. Let w be an element of a Coxeter group. In [**4**], w is said to be *short-braid avoiding* if and only if any reduced expression $s_1s_2 \ldots s_m$ for w satisfies $s_i \neq s_{i+2}$ for all $i \in [m-2]$. It is easy to see that a fully covering element is short-braid avoiding, and that a short-braid avoiding element is fully commutative. By the same method as the one adopted in the proof of [**4**, Theorem 1] and Theorem 4.3, we can easily obtain the following which includes Fan's result [**4**, Theorem 1]. Let (W, S) be a Coxeter system and let (W_0, S_0) be a Coxeter system defined by $S_0 := S$ as a set and $m(s,t) := 3$ if $m(s,t) \geq 3$ in W for $s, t \in S_0$. If W_0 is a Coxeter group of type A, D or E then for $w \in W$, w is a short-braid avoiding element if and only if w is a fully covering element.

Although it is already shown by Fan that a Coxeter group of type E is FC-finite, we can give an explicit upper bound for the maximum length of fully commutative elements.

Proposition 4.7. *For* $n \geq 3$*, we have*

$$\max\{\ell(w) | w \in W(E_n)^{FC}\} \leq 2^{n-1} - 1,$$

where we put $W(E_3) := \langle \beta_1, \beta_2, \gamma \rangle$*. In particular, we have* $|W(E_n)^{FC}| < \infty$*.*

Remark 4.8. In [**10**], H. Tagawa showed that we have $\max\{c^-(x) | x \in W(A_n)\} = \lfloor (n+1)^2/4 \rfloor$, where $\lfloor a \rfloor$ is the largest integer equal or less than a. By the formula, it is easy to show that we have $\max\{\ell(x) | x \in W(A_n)^{FC}\} = \lfloor (n+1)^2/4 \rfloor$. Note that it does not hold on case of type D. In fact, we have $\max\{c^-(x) | x \in W(D_4)\} = 8 > 6 = \max\{\ell(x) | x \in W(D_4)^{FC}\}$.

Moreover, we can show the following.

Theorem 4.9. *Let* W *be an irreducible Coxeter group which is neither of type* A*,* D *nor* E*. Then* W *is not a bi-full Coxeter group. In other words, there is an element of* W *which is fully commutative and which is not fully covering. In particular, if* W *is a simply-laced Coxeter group then we have* $|W^{FC}| = \infty$*.*

This theorem is easily obtained by the following proposition.

Proposition 4.10. *Let* (W_1, S_1) *(resp.* (W_2, S_2)*,* (W_3, S_3)*,* (W_4, S_4)*,* (W_5, S_5)*) be a Coxeter system of type* $\tilde{A}_n$ $(n \geq 2)$ *(resp.* $\tilde{D}_{r+3}$ $(r \geq 1)$*,* $\tilde{E}_6$*,* $\tilde{E}_7$*,* $I_2(m)$ $(m \geq 4)$*). Then for each* $1 \leq i \leq 5$ *there exists an element* w_i *of* W_i *such that* w_i *is fully commutative and* w_i *is not fully covering. Furthermore we have* $|W_i^{FC}| = \infty$ *for any* $1 \leq i \leq 4$*.*

References

[1] S. Billey, W. Jockusch and R. Stanley, *Some combinatorial properties of Schubert polynomials*, J. Algebraic Combinatorics, 2 (1993), 345-374.
[2] M. Dyer, *Hecke algebras and reflections in Coxeter groups*, Ph.D. Thesis, Univ. of Sydney (1987).
[3] C. K. Fan, *A Hecke Algebra Quotient and Properties of Commutative Elements of a Weyl Group*, Ph. D. thesis, MIT (1995).
[4] C. K. Fan, *Schubert Varieties And Short Braidedness*, Transformation Groups, Vol 3, No.1 (1998), pp. 51-56.
[5] R. M. Green, *On 321-avoiding permutations in affine Weyl groups*, J. Algebraic Combinatorics, 15 (2002), 241-252.
[6] Hagiwara M., *Minuscule Heaps Over Dynkin diagrams of type* $\tilde{A}$, preprint.
[7] Hagiwara M., M. Ishikawa, and H. Tagawa, *A characterization of the simply-laced FC-finite Coxeter groups*, preprint.
[8] J. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Univ. Press (1990).
[9] J. R. Stembridge, *On the Fully Commutative Elements of Coxeter Groups*, J. Algebraic Combinatorics, 5 (1996), 353-385.

[10] H. Tagawa, *On the maximum value of the first coefficients of Kazhdan-Lusztig polynomials for symmetric groups*, J. Math. Sci. Univ. Tokyo, 1 (1994), 461-469.

Institute of Industrial Science, University of Tokyo, Komaba, Meguro-ku, Tokyo, Japan, 153-8505
E-mail address: manau@imailab.iis.u-tokyo.ac.jp

Department of Mathematics, Tottori University, Koyama, Tottori, Japan, 680-8551
E-mail address: ishikawa@fed.tottori-u.ac.jp

Department of Mathematics, Wakayama University, Sakaedani, Wakayama, Japan, 640-8510
E-mail address: tagawa@math.edu.wakayama-u.ac.jp

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Cyclic Resultants

Christopher J. Hillar

Abstract. *Let k be a field of characteristic zero and let $f \in k[x]$. The m-th cyclic resultant of f is $r_m = Res(f, x^m - 1)$. We characterize polynomials having the same set of nonzero cyclic resultants. Generically, for a polynomial f of degree d, there are exactly 2^{d-1} distinct degree d polynomials with the same set of cyclic resultants as f. However, in the generic monic case, degree d polynomials are uniquely determined by their cyclic resultants. Moreover, two reciprocal ("palindromic") polynomials giving rise to the same set of nonzero r_m are equal. The reciprocal case was stated many years ago (for $k = \mathbb{R}$) and has many applications stemming from such disparate fields as dynamics, number theory, and Lagrangian mechanics. In the process, we also prove a unique factorization result in semigroup algebras involving products of binomials.*

1. Introduction

Let k be a field of characteristic zero and let $f(x) = a_0x^d + a_1x^{d-1} + \cdots + a_d \in k[x]$. The m-th *cyclic resultant* of f is $r_m(f) = \text{Res}(f, x^m - 1)$. We are primarily interested here in the fibers of the map $r : k[x] \to k^{\mathbb{N}}$ given by $f \mapsto (r_m)_{m=0}^{\infty}$. In particular, what are the conditions for two polynomials to give rise to the same set of cyclic resultants? For technical reasons, we will only consider polynomials f that do not have a root of unity as a zero. With this restriction, a polynomial will map to a set of all nonzero cyclic resultants.

One motivation for the study of cyclic resultants comes from the theory of dynamical systems. Sequences of the form r_m arise as the cardinalities of sets of periodic points for toral endomorphisms. Let f be monic of degree d with integral coefficients and let $X = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ denote the d-dimensional additive torus. Then, the companion matrix A_f of f acts on X by multiplication mod 1; that is, it defines a map $T : X \to X$ given by

$$T(\mathbf{x}) = A_f\mathbf{x} \mod 1.$$

Let $\text{Per}_m(T) = \{\mathbf{x} \in \mathbb{T}^d : T^m(\mathbf{x}) = \mathbf{x}\}$ be the set of points fixed under the map T^m. Under the ergodicity condition that no zero of f is a root of unity, it follows (see [**3**]) that $|\text{Per}_m(T)| = |\det(A_f^m - I)|$, in which I is the d-by-d identity matrix, and both of these quantities are given by $|r_m(f)|$. As a consequence of our results, we characterize when the sequence $|\text{Per}_m(T)|$ determines the spectrum of the linear map $A : \mathbb{R}^d \to \mathbb{R}^d$ that lifts T.

In connection with number theory, such sequences were also studied by Pierce and Lehmer [**3**] in the hope of using them to produce large primes. As a simple example, the polynomial $f(x) = x - 2$ gives the Mersenne sequence $M_m = 2^m - 1$. Indeed, we have $M_m = |\det(A_f^m - I)|$, and these numbers are precisely

Key words and phrases. cyclic resultant, binomial factorization, group rings.
This work is supported under a National Science Foundation Graduate Research Fellowship.

the cardinalities of the sets $\mathrm{Per}_m(T)$ for the map $T(x) = 2x \mod 1$. Further motivation comes from knot theory [**9**] and Lagrangian mechanics [**6, 7**].

The principal result in the direction of our main characterization theorem was discovered by Fried [**4**] although certain implications of Fried's result were known to Stark [**2**]. One of our motivations for this work was to present a complete and satisfactory proof of this result. Fried's argument in [**4**], while elegant, is difficult to read and not as complete as one would like. Given a polynomial f of degree d, the *reversal* of f is the polynomial $x^d f(1/x)$. Additionally, f is called *reciprocal* if $a_i = a_{d-i}$ for $0 \leq i \leq d$ (sometimes such a polynomial is called *palindromic*). Alternatively, f is reciprocal if it is equal to its own reversal. Fried's result may be stated as follows.

Theorem 1.1 (Fried)**.** *Let $p(x) = a_0x^d + \cdots + a_{d-1}x + a_d \in \mathbb{R}[x]$ be a real reciprocal polynomial of even degree d with $a_0 > 0$, and let r_m be the m-th cyclic resultants of p. Then, $|r_m|$ uniquely determine this polynomial of degree d as long as the r_m are never 0.*

2. Statement of Results

As far as we know, the general (non-reciprocal) case has not received much attention. We begin by stating our main characterization theorem for cyclic resultants.

Theorem 2.1. *Let k be a field of characteristic zero, and let f and g be polynomials in $\overline{k}[x]$. Then, f and g generate the same sequence of nonzero cyclic resultants if and only if there exist $u, v \in \overline{k}[x]$ with $deg(u)$ even, $u(0) \neq 0$, and nonnegative integers $l_1 \equiv l_2 \pmod 2$ such that*

$$f(x) = x^{l_1}v(x)u(x^{-1})x^{deg(u)}$$
$$g(x) = x^{l_2}v(x)u(x).$$

Although the theorem statement appears somewhat technical, we present a natural interpretation of the result. Suppose that $g(x) = x^{l_2}v(x)u(x)$ is a factorization of a polynomial g with nonzero cyclic resultants. Then, another polynomial f giving rise to this same sequence of resultants is obtained from v by multiplication with the reversal of u and a factor x^{l_1} in which $l_1 \in \mathbb{N}$ has the same parity as l_2. In other words, $f(x) = x^{l_1}v(x)u(x^{-1})x^{\deg(u)}$, and all such f must arise in this manner.

Example 2.2. One can check that the polynomials

$$f(x) = x^3 - 10\,x^2 + 31\,x - 30$$
$$g(x) = 15\,x^5 - 38\,x^4 + 17\,x^3 - 2\,x^2$$

both generate the same cyclic resultants. This follows from the factorizations

$$f(x) = (x-2)\left(15x^2 - 8x + 1\right)$$
$$g(x) = x^2(x-2)\left(x^2 - 8x + 15\right).$$

The following is a direct corollary of our main theorem to the generic case.

Corollary 2.3. *Let k be a field of characteristic zero and let g be a generic polynomial in $k[x]$ of degree d. Then, there are exactly 2^{d-1} distinct degree d polynomials with the same set of cyclic resultants as g.*

Proof. If g is generic, then g will not have a root of unity as a zero nor will $g(0) = 0$. Theorem 2.1, therefore, implies that any other degree d polynomial $f \in \overline{k}[x]$ giving rise to the same set of cyclic resultants is determined by choosing an even cardinality subset of the roots of g. Such polynomials will be distinct since g is generic. Since there are 2^d subsets of the roots of g and half of them have even cardinality, the theorem follows. □

Example 2.4. Let $g(x) = (x-2)(x-3)(x-5) = x^3 - 10\,x^2 + 31\,x - 30$. Then, there are $2^{3-1} - 1 = 3$ other degree 3 polynomials with the same set of cyclic resultants as g. They are:

$$15\,x^3 - 38\,x^2 + 17\,x - 2$$

$$10\,x^3 - 37\,x^2 + 22\,x - 3$$
$$6\,x^3 - 35\,x^2 + 26\,x - 5.$$

If one is interested in the case of generic monic polynomials, then Theorem 2.1 also implies the following uniqueness result.

Corollary 2.5. *Let k be a field of characteristic zero and let g be a generic monic polynomial in $k[x]$ of degree d. Then, there is only one monic, degree d polynomial with the same set of cyclic resultants as g.*

Proof. Again, since g is generic, it will not have a root of unity as a zero nor will $g(0) = 0$. Theorem 2.1 forces a constraint on the roots of g for there to be a different polynomial f with the same set of cyclic resultants as g. Namely, a subset of the roots of f has product 1, a non-generic situation. □

As to be expected, there are analogs of Theorem 2.1 and Corollary 2.5 to the real case involving absolute values.

Theorem 2.6. *Let f and g be polynomials in $\mathbb{R}[x]$. If f and g generate the same sequence of nonzero cyclic resultant absolute values, then there exist $u, v \in \mathbb{C}[x]$ with $u(0) \neq 0$ and nonnegative integers l_1, l_2 such that*

$$f(x) = \pm x^{l_1} v(x) u(x^{-1}) x^{deg(u)}$$
$$g(x) = x^{l_2} v(x) u(x).$$

Corollary 2.7. *Let g be a generic monic polynomial in $\mathbb{R}[x]$ of degree d. Then, g is the only monic, degree d polynomial in $\mathbb{R}[x]$ with the same set of cyclic resultant absolute values as g.*

The generic real case without the monic assumption is somewhat more subtle than that of Corollary 2.3. The difficulty is that we are restricted to polynomials in $\mathbb{R}[x]$. However, there is the following

Corollary 2.8. *Let g be a generic polynomial in $\mathbb{R}[x]$ of degree d. Then there are exactly $2^{\lceil d/2 \rceil + 1}$ distinct degree d polynomials in $\mathbb{R}[x]$ with the same set of cyclic resultant absolute values as g.*

Proof. If d is even, then genericity implies that all of the roots of g will be nonreal. In particular, it follows from Theorem 2.6 (and genericity) that any other degree d polynomial $f \in \mathbb{R}[x]$ giving rise to the same set of cyclic resultant absolute values is determined by choosing a subset of the $d/2$ pairs of conjugate roots of g and a sign. This gives us a count of $2^{d/2+1}$ distinct real polynomials. When d is odd, g will have exactly one real root, and a similar counting argument gives us $2^{\lceil d/2 \rceil + 1}$ for the number of distinct real polynomials in this case. This proves the corollary. □

A surprising consequence of this result is that the number of polynomials with equal sets of cyclic resultant absolute values is significantly smaller than the number predicted in Corollary 2.3.

Example 2.9. Let $g(x) = (x-2)(x+i+2)(x-i+2) = x^3+2\,x^2-3\,x-10$. Then, there are $2^{\lceil 3/2 \rceil + 1} - 1 = 7$ other degree 3 real polynomials with the same set of cyclic resultant absolute values as g. They are:

$$-x^3 - 2\,x^2 + 3\,x + 10$$
$$\pm(-2\,x^3 - 7\,x^2 - 6\,x + 5)$$
$$\pm(5\,x^3 - 6\,x^2 - 7\,x - 2)$$
$$\pm(-10\,x^3 - 3\,x^2 + 2\,x + 1).$$

It is important to realize that while

$$\begin{aligned} f(x) =& (1-2x)(1+(i+2)x)(x-i+2) \\ =& (-4-2\,i)\,x^3 - (10-i)\,x^2 + (2+2\,i)\,x + 2 - i \end{aligned}$$

has the same set of actual cyclic resultants (by Theorem 2.1), it does not appear in the count above since it is not in $\mathbb{R}[x]$.

As an illustration of the usefulness of Theorem 2.1, we prove a uniqueness result involving cyclic resultants of reciprocal polynomials. Fried's result also follows in the same way using Theorem 2.6 in place of Theorem 2.1.

Corollary 2.10. *Let f and g be reciprocal polynomials with equal sets of nonzero cyclic resultants. Then, $f = g$.*

PROOF. Let f and g be reciprocal polynomials having the same set of nonzero cyclic resultants. Applying Theorem 2.1, it follows that $d = \deg(f) = \deg(g)$ and that

$$f(x) = v(x)u(x^{-1})x^{\deg(u)}$$
$$g(x) = v(x)u(x)$$

($l_1 = l_2 = 0$ since $f(0), g(0) \neq 0$). But then,

$$\frac{u(x^{-1})}{u(x)}x^{\deg(u)} = \frac{f(x)}{g(x)} = \frac{x^d f(x^{-1})}{x^d g(x^{-1})} = \frac{u(x)}{u(x^{-1})}x^{-\deg(u)}.$$

In particular, $u(x) = \pm u(x^{-1})x^{\deg(u)}$. If $u(x) = u(x^{-1})x^{\deg(u)}$, then $f = g$ as desired. In the other case, it follows that $f = -g$. But then $\mathrm{Res}(f,x-1) = \mathrm{Res}(g,x-1) = -\mathrm{Res}(f,x-1)$ is a contradiction to f having all nonzero cyclic resultants. This completes the proof. □

We now switch to the seemingly unrelated topic of binomial factorizations in semigroup algebras. The relationship to cyclic resultants will become clear later. Let A be a finitely generated abelian group and let $a_1, \ldots, a_n$ be distinguished generators of A. Let Q be the semigroup generated by $a_1, \ldots, a_n$. If k is a field, the *semigroup algebra* $k[Q]$ is the k-algebra with vector space basis $\{\mathbf{s}^a : a \in Q\}$ and multiplication defined by $\mathbf{s}^a \cdot \mathbf{s}^b = \mathbf{s}^{a+b}$. Let L denote the kernel of the homomorphism $\mathbb{Z}^n$ onto A. The *lattice ideal* associated with L is the following ideal in $S = k[x_1, \ldots, x_n]$:

$$I_L = \langle x^u - x^v \ : \ u, v \in \mathbb{N}^n \text{ with } u - v \in L\rangle.$$

It is a well-known fact that $k[Q] \cong S/I_L$ (e.g. see [**8**]). We are primarily concerned here with certain kinds of factorizations in $k[Q]$.

Question 2.11. When is a product of binomials in $k[Q]$ equal to another product of binomials?

The answer to this question is turns out to be fundamental for the study of cyclic resultants. Our main result in this direction is a certain kind of unique factorization of binomials in $k[Q]$.

Theorem 2.12. *Let k be a field of characteristic zero and let $\alpha \in k$. Suppose that*

$$\boldsymbol{s}^a \prod_{i=1}^{e} (\boldsymbol{s}^{u_i} - \boldsymbol{s}^{v_i}) = \alpha \boldsymbol{s}^b \prod_{i=1}^{f} (\boldsymbol{s}^{x_i} - \boldsymbol{s}^{y_i})$$

are two factorizations of binomials in the ring $k[Q]$. Furthermore, suppose that for each i, $u_i - v_i$ ($x_i - y_i$) has infinite order as an element of A. Then, $\alpha = \pm 1$, $e = f$, and up to permutation, for each i, there are elements $c_i, d_i \in Q$ such that $\boldsymbol{s}^{c_i}(\boldsymbol{s}^{u_i} - \boldsymbol{s}^{v_i}) = \pm \boldsymbol{s}^{d_i}(\boldsymbol{s}^{x_i} - \boldsymbol{s}^{y_i})$.

Of course, when each side has a factor of zero, the theorem fails. There are other obstructions, however, that make necessary the supplemental hypotheses concerning order. For example, take $k = \mathbb{Q}$, and let $A = \mathbb{Z}/2\mathbb{Z}$. Then, $k[Q] = k[A] \cong \mathbb{Q}[s]/\langle s^2 - 1\rangle$, and we have that

$$(1 - s)(1 - s) = 2(1 - s).$$

This theorem also fails when the characteristic is not 0.

Example 2.13. $L = \{0\}$, $I_L = \langle 0 \rangle$, $A = \mathbb{Z}$, $Q = \mathbb{N}$, $k = \mathbb{Z}/3\mathbb{Z}$,

$$(1 - t^3) = (1 - t)(1 - t)(1 - t).$$

One might wonder what happens when the binomials are not of the form $\mathbf{s}^u - \mathbf{s}^v$. The following example exhibits some of the difficulty in formulating a general statement.

Example 2.14. $L = \{(0, b) \in \mathbb{Z}^2 : b \text{ is even}\}$, $I_L = \langle s^2 - 1 \rangle \subseteq k[s, t]$, $A = \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$, $Q = \mathbb{N} \oplus \mathbb{Z}/2\mathbb{Z}$, $k = \mathbb{Q}(i)$. Then,

$$(1 - t^4) = (1 - st)(1 + st)(1 - ist)(1 + ist) = (1 - st^2)(1 + st^2)$$

are three different binomial factorizations of the same semigroup algebra element.

Example 2.15. $L = \{0\}$, $I_L = \langle 0 \rangle$, $A = \mathbb{Z}$, $Q = \mathbb{N}$, $k = \mathbb{C}$. If

$$\prod_{i=1}^{r} (1 - t^{m_i}) = \prod_{i=1}^{s} (1 - t^{n_i})$$

for positive integers m_i,n_i, then $r = s$ and up to permutation, $m_i = n_i$ for all i.

We now are in a position to outline our strategy for characterizing those polynomials f and g having the same set of nonzero cyclic resultants (this strategy is similar to the one employed in [**4**]). Given a polynomial f and its sequence of r_m, we construct the generating function $E_f(z) = \exp\left(-\sum_{m \geq 1} r_m \frac{z^m}{m}\right)$. This series turns out to be rational with coefficients depending explicitly on the roots of f. Since f and g are assumed to have the same set of r_m, it follows that their corresponding rational functions E_f and E_g are equal. Let G be the (multiplicative) group of units in the algebraic closure of k. Then, the divisors of these two rational functions are group ring elements in $\mathbb{Z}[G]$ and their equality forces a certain binomial group ring factorization that is analyzed explicitly. The results above follow from this final analysis.

3. Binomial Factorizations in Semigroup Algebras

To prove our factorization result, we will pass to the full group algebra $k[A]$. As above, we represent elements $\tau \in k[A]$ as $\tau = \sum_{i=1}^{m} \alpha_i \mathbf{s}^{g_i}$, in which $\alpha_i \in k$ and $g_i \in A$. The following lemma is quite well-known.

Lemma 3.1. *If $\alpha \in k^*$ and $g \in A$ has infinite order, then $1 - \alpha \boldsymbol{s}^g \in k[A]$ is not a 0-divisor.*

Proof. Let $\alpha \in k^*, g \in A$ and $\tau = \sum_{i=1}^{m} \alpha_i \mathbf{s}^{g_i} \neq 0$ be such that

$$\tau = \alpha \mathbf{s}^g \tau = \alpha \mathbf{s}^{2g} \tau = \alpha \mathbf{s}^{3g} \tau = \cdots .$$

Suppose that $\alpha_1 \neq 0$. Then, the elements $\mathbf{s}^{g_1}, \mathbf{s}^{g_1+g}, \mathbf{s}^{g_1+2g}, \ldots$ appear in τ with nonzero coefficient, and since g has infinite order, these elements are all distinct. It follows, therefore, that τ cannot be a finite sum, and this contradiction finishes the proof. □

Since the proof of the main theorem involves multiple steps, we record several facts that will be useful later. The first result is a verification of the factorization theorem for a generalization of the situation in Example 2.15.

Lemma 3.2. *Let k be a field of characteristic zero and let C be an abelian group. Let $k[C]$ be the group algebra with k-vector space basis given by $\{\boldsymbol{s}^c : c \in C\}$ and set $R = k[C][t, t^{-1}]$. Suppose that $c_1, \ldots, c_e, d_1, \ldots, d_f, b \in C$, $m_1, \ldots, m_e, n_1, \ldots, n_f$ are nonzero integers, $q \in \mathbb{Z}$, and $z \in k$ are such that*

$$\prod_{i=1}^{e} (1 - \boldsymbol{s}^{c_i} t^{m_i}) = z \boldsymbol{s}^b t^q \prod_{i=1}^{f} (1 - \boldsymbol{s}^{d_i} t^{n_i})$$

holds in R. Then, $e = f$ and after a permutation, for each i, either $\boldsymbol{s}^{c_i} t^{m_i} = \boldsymbol{s}^{d_i} t^{n_i}$ or $\boldsymbol{s}^{c_i} t^{m_i} = \boldsymbol{s}^{-d_i} t^{-n_i}$.

Proof. Let sgn : $\mathbb{Z} \setminus \{0\} \to \{-1, 1\}$ denote the standard sign map $\text{sgn}(n) = n/|n|$ and set $\gamma = z\mathbf{s}^b t^q$. Rewrite the left-hand side of the given equality as:

$$\prod_{i=1}^{e} (1 - \mathbf{s}^{c_i} t^{m_i}) = \prod_{\text{sgn}(m_i)=-1} -\mathbf{s}^{c_i} t^{m_i} \prod_{i=1}^{e} \left(1 - \mathbf{s}^{\text{sgn}(m_i)c_i} t^{|m_i|}\right).$$

Similarly for the right-hand side, we have:

$$\prod_{i=1}^{f} (1 - \mathbf{s}^{d_i} t^{n_i}) = \prod_{\text{sgn}(n_i)=-1} -\mathbf{s}^{d_i} t^{n_i} \prod_{i=1}^{f} \left(1 - \mathbf{s}^{\text{sgn}(n_i)d_i} t^{|n_i|}\right).$$

Next, set

$$\eta = \gamma \prod_{\text{sgn}(m_i)=-1} -\mathbf{s}^{-c_i} t^{-m_i} \prod_{\text{sgn}(n_i)=-1} -\mathbf{s}^{d_i} t^{n_i}$$

so that our original equation may be written as

$$\prod_{i=1}^{e} \left(1 - \mathbf{s}^{\text{sgn}(m_i)c_i} t^{|m_i|}\right) = \eta \prod_{i=1}^{f} \left(1 - \mathbf{s}^{\text{sgn}(n_i)d_i} t^{|n_i|}\right).$$

Comparing the lowest degree term (with respect to t) on both sides, it follows that $\eta = 1$. It is enough, therefore, to prove the claim in the case when

$$\prod_{i=1}^{e} (1 - \mathbf{s}^{c_i} t^{m_i}) = \prod_{i=1}^{f} (1 - \mathbf{s}^{d_i} t^{n_i}) \tag{3.1}$$

and the m_i, n_i are positive. Without loss of generality, suppose the lowest degree nonconstant term on both sides of (3.1) is t^{m_1} with coefficient $-\mathbf{s}^{c_1} - \cdots - \mathbf{s}^{c_u}$ on the left and $-\mathbf{s}^{d_1} - \cdots - \mathbf{s}^{d_v}$ on the right. Here, u (v) corresponds to the number of m_i (n_i) with $m_i = m_1$ ($n_i = m_1$).

Since the set of distinct monomials $\{\mathbf{s}^c : c \in C\}$ is a k-vector space basis for the ring $k[C]$, equality of the t^{m_1} coefficients above implies that $u = v$ and that up to permutation, $\mathbf{s}^{c_j} = \mathbf{s}^{d_j}$ for $j = 1, \ldots, u$ (recall that the characteristic of k is zero). Using Lemma 3.1 and induction completes the proof. □

Lemma 3.3. *Let $P = (p_{ij})$ be a d-by-n integer matrix such that every row has at least one nonzero integer. Then, there exists $\mathbf{v} \in \mathbb{Z}^n$ such that the vector $P\mathbf{v}$ does not contain a zero entry.*

Proof. Let P be a d-by-n integer matrix as in the hypothesis of the lemma, and for $h \in \mathbb{Z}$, let $\mathbf{v}_h = (1, h, h^2, \ldots, h^{n-1})^T$. Assume, by way of contradiction, that $P\mathbf{v}$ contains a zero entry for all $\mathbf{v} \in \mathbb{Z}^n$. Then, in particular, this is true for all $\mathbf{v}_h$ as above. By the (infinite) pigeon-hole principle, there exists an infinite set of $h \in \mathbb{Z}$ such that (without loss of generality) the first entry of $P\mathbf{v}_h$ is zero. But then,

$$f(h) := \sum_{i=1}^{n} p_{1i} h^{i-1} = 0$$

for infinitely many values of h. It follows, therefore, that $f(h)$ is the zero polynomial, contradicting our hypothesis and completing the proof. □

Lemma 3.3 will be useful in verifying the following fact.

Lemma 3.4. *Let A be a finitely generated abelian group and $a_1, \ldots, a_d$ elements in A of infinite order. Then, there exists a homomorphism $\phi : A \to \mathbb{Z}$ such that $\phi(a_i) \neq 0$ for all i.*

Proof. Write $A = B \oplus C$, in which C is a finite group and B is free of rank n. If $n = 0$, then there are no elements of infinite order; therefore, we may assume that the rank of B is positive. Since $a_1, \ldots, a_d$ have infinite order, their images in the natural projection $\pi : A \to B$ are nonzero. It follows that we may assume that A is free and a_i are nonzero elements of A.

Let $e_1, \ldots, e_n$ be a basis for A, and write

$$a_t = p_{t1}e_1 + \cdots + p_{tn}e_n$$

for (unique) integers $p_{ij} \in \mathbb{Z}$. To determine a homomorphism $\phi : A \to \mathbb{Z}$ as in the lemma, we must find integers $\phi(e_1), \ldots, \phi(e_n)$ such that

$$\begin{aligned} &0 \neq p_{11}\phi(e_1) + \cdots + p_{1n}\phi(e_n) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &0 \neq p_{d1}\phi(e_1) + \cdots + p_{dn}\phi(e_n). \end{aligned} \tag{3.2}$$

This, of course, is precisely the consequence of Lemma 3.3 applied to the matrix $P = (p_{ij})$, finishing the proof. □

Recall that a *trivial unit* in the group ring $k[A]$ is an element of the form $\alpha\mathbf{s}^a$ in which $\alpha \in k^*$ and $a \in A$. The main content of Theorem 2.12 is contained in the following result. The technique of embedding $k[A]$ into a Laurent polynomial ring is also used by Fried in [**4**].

Lemma 3.5. *Let A be an abelian group and let k be a field of characteristic 0. Two factorizations in $k[A]$,*

$$\prod_{i=1}^{e} (1 - \mathbf{s}^{g_i}) = \eta \prod_{i=1}^{f} \left(1 - \mathbf{s}^{h_i}\right),$$

in which η is a trivial unit and $g_i, h_i \in A$ all have infinite order are equal if and only if $e = f$ and there is some nonnegative integer p such that, up to permutation,

(1) $g_i = h_i$ *for* $i = 1, \ldots, p$
(2) $g_i = -h_i$ *for* $i = p+1, \ldots, e$
(3) $\eta = (-1)^{e-p}\,\mathbf{s}^{g_{p+1}+\cdots+g_e}$.

Proof. The if-direction of the claim is a straightforward calculation. Therefore, suppose that one has two factorizations as in the lemma. It is clear we may assume that A is finitely generated. By Lemma 3.4, there exists a homomorphism $\phi : A \to \mathbb{Z}$ such that $\phi(g_i), \phi(h_i) \neq 0$ for all i. The ring $k[A]$ may be embedded into the Laurent ring, $R = k[A][t, t^{-1}]$, by way of

$$\psi\left(\sum_{i=1}^{m} \alpha_i \mathbf{s}^{a_i}\right) = \sum_{i=1}^{m} \alpha_i \mathbf{s}^{a_i} t^{\phi(a_i)}.$$

Write $\eta = \alpha\mathbf{s}^b$. Then, applying this homomorphism to the original factorization, we have

$$\prod_{i=1}^{e} \left(1 - \mathbf{s}^{g_i} t^{\phi(g_i)}\right) = \alpha \mathbf{s}^b t^{\phi(b)} \prod_{i=1}^{f} \left(1 - \mathbf{s}^{h_i} t^{\phi(h_i)}\right).$$

Lemma 3.2 now applies to give us that $e = f$ and there is an integer p such that up to permutation,

(1) $g_i = h_i$ for $i = 1, \ldots, p$
(2) $g_i = -h_i$ for $i = p+1, \ldots, e$.

We are therefore left with verifying statement (3) of the lemma. Using Lemma 3.1, we may cancel equal terms in our original factorization, leaving us with the following equation:

$$\begin{aligned}\prod_{i=p+1}^{e}(1-\mathbf{s}^{g_i}) &= \eta \prod_{i=p+1}^{e}(1-\mathbf{s}^{-g_i}) \\ &= \eta(-1)^{e-p}\prod_{i=p+1}^{e}\mathbf{s}^{-g_i}\prod_{i=p+1}^{e}(1-\mathbf{s}^{g_i}).\end{aligned}$$

Finally, one more application of Lemma 3.1 gives us that $\eta = (-1)^{e-p}\mathbf{s}^{g_{p+1}+\cdots+g_e}$ as desired. This finishes the proof. □

We may now prove Theorem 2.12.

Proof of Theorem 2.12. Let

$$\mathbf{s}^a\prod_{i=1}^{e}(\mathbf{s}^{u_i}-\mathbf{s}^{v_i}) = \alpha\mathbf{s}^b\prod_{i=1}^{f}(\mathbf{s}^{x_i}-\mathbf{s}^{y_i})$$

be two factorizations in the ring $k[Q]$. View this expression in $k[A]$ and factor each element of the form $(\mathbf{s}^u - \mathbf{s}^v)$ as $\mathbf{s}^u(1-\mathbf{s}^{v-u})$. By assumption, each such $v-u$ has infinite order. Now, apply Lemma 3.5, giving us that $\alpha = \pm 1$, $e = f$, and that after a permutation, for each i either $\mathbf{s}^{v_i-u_i} = \mathbf{s}^{y_i-x_i}$ or $\mathbf{s}^{v_i-u_i} = \mathbf{s}^{x_i-y_i}$. It easily follows from this that for each i, there are elements $c_i, d_i \in Q$ such that $\mathbf{s}^{c_i}(\mathbf{s}^{u_i}-\mathbf{s}^{v_i}) = \pm\mathbf{s}^{d_i}(\mathbf{s}^{x_i}-\mathbf{s}^{y_i})$. This completes the proof of the theorem. □

4. Cyclic Resultants and Rational Functions

We begin with some preliminaries concerning cyclic resultants. Let $f(x) = a_0x^d + a_1x^{d-1} + \cdots + a_d$ be a degree d polynomial over k, and let the companion matrix for f be given by:

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_d/a_0 \\ 1 & 0 & \cdots & 0 & -a_{d-1}/a_0 \\ 0 & 1 & \cdots & 0 & -a_{d-2}/a_0 \\ 0 & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1/a_0 \end{bmatrix}.$$

Also, let I denote the d-by-d identity matrix. Then, we may write [**1**, p. 77]

$$r_m = a_0^m \det(A^m - I). \tag{4.1}$$

Extending to a splitting field of f, this equation can also be expressed as,

$$r_m = a_0^m \prod_{i=1}^{d}(\alpha_i^m - 1), \tag{4.2}$$

in which $\alpha_1, \ldots, \alpha_d$ are the roots of $f(x)$.

Let $e_i(y_1, \ldots, y_d)$ be the i-th elementary symmetric function in the variables $y_1, \ldots, y_d$ (we set $e_0 = 1$). Then, we know that $a_i = (-1)^i a_0 e_i(\alpha_1, \ldots, \alpha_d)$ and that

$$r_m = a_0^m \sum_{i=0}^{d}(-1)^i e_{d-i}(\alpha_1^m, \ldots, \alpha_d^m). \tag{4.3}$$

We first record an auxiliary result.

Lemma 4.1. *Let* $F_k(z) = \prod_{1 \le i_1 < \cdots < i_k \le d} (1 - a_0 \alpha_{i_1} \cdots \alpha_{i_k} z)$ *with* $F_0(z) = 1 - a_0 z$. *Then,*

$$\sum_{m=1}^{\infty} a_0^m e_k \left(\alpha_1^m, \ldots, \alpha_n^m\right) z^m = -z \cdot \frac{F_k'}{F_k},$$

in which F_k' *denotes* $\frac{dF_k}{dz}$.

Proof. For $k = 0$, the equation is easily verified. When $k > 0$, the calculation is still fairly straightforward:

$$\begin{aligned}
\sum_{m=1}^{\infty} a_0^m e_k \left(\alpha_1^m, \ldots, \alpha_d^m\right) z^m &= \sum_{m=1}^{\infty} \sum_{i_1 < \cdots < i_k} a_0^m \alpha_{i_1}^m \cdots \alpha_{i_k}^m \cdot z^m \\
&= \sum_{i_1 < \cdots < i_k} \sum_{m=1}^{\infty} a_0^m \alpha_{i_1}^m \cdots \alpha_{i_k}^m \cdot z^m \\
&= \sum_{i_1 < \cdots < i_k} \frac{a_0 \alpha_{i_1} \cdots \alpha_{i_k} z}{1 - a_0 \alpha_{i_1} \cdots \alpha_{i_k} z} \\
&= \frac{-z \cdot \frac{d}{dz} \left[\prod_{i_1 < \cdots < i_k} (1 - a_0 \alpha_{i_1} \cdots \alpha_{i_k} z) \right]}{\prod_{i_1 < \cdots < i_k} (1 - a_0 \alpha_{i_1} \cdots \alpha_{i_k} z)} \\
&= -z \cdot \frac{F_k'}{F_k}.
\end{aligned}$$

□

We may now state and prove the rationality result mentioned in the introduction.

Lemma 4.2. $R_f(z) = \sum_{m=1}^{\infty} r_m z^m$ *is a rational function in* z.

Proof. We simply compute that

$$\begin{aligned}
\sum_{m=1}^{\infty} r_m z^m &= \sum_{m=1}^{\infty} \sum_{i=0}^{d} (-1)^i a_0^m e_{d-i} \left(\alpha_1^m, \ldots, \alpha_d^m\right) \cdot z^m \\
&= \sum_{i=0}^{d} (-1)^i \sum_{m=1}^{\infty} a_0^m e_{d-i} \left(\alpha_1^m, \ldots, \alpha_d^m\right) \cdot z^m \\
&= -z \cdot \sum_{i=0}^{d} (-1)^i \cdot \frac{F_{d-i}'}{F_{d-i}}.
\end{aligned}$$

□

Let us remark at this point that Lemma 4.2 implies the following curious determinantal identity.

Corollary 4.3. *Let* d *be a positive integer and set* $n = 2^d + 1$. *Then,*

$$A = \left(\prod_{l=1}^{d} \left(\alpha_l^{n+i-j} - 1 \right) \right)_{i,j=1}^{n}$$

has determinant 0.

PROOF. Let $r_m = \prod_{l=1}^{d} (\alpha_l^m - 1)$ for $m \in \{1, 2, \ldots\}$. From above, $\sum_{m=1}^{\infty} r_m z^m$ is a rational function of z with numerator and denominator each having degree at most 2^d. This implies a linear recurrence for the r_m of length at most 2^d, and therefore it follows that $\det(A) = 0$. □

Manipulating the expression for $R_f(z)$ occurring in Lemma 4.2, we also have the following fact.

Corollary 4.4. *If d is even, let $G_d = \frac{F_d F_{d-2} \cdots F_0}{F_{d-1} F_{d-3} \cdots F_1}$ and if d is odd, let $G_d = \frac{F_d F_{d-2} \cdots F_1}{F_{d-1} F_{d-3} \cdots F_0}$. Then,*

$$\sum_{m=1}^{\infty} r_m z^m = -z \frac{G_d'}{G_d}.$$

In particular, it follows that

$$\exp\left(-\sum_{m=1}^{\infty} r_m \frac{z^m}{m}\right) = G_d. \tag{4.4}$$

Example 4.5. Let $f(x) = x^2 - 5x + 6 = (x-2)(x-3)$. Then, $r_m = (2^m - 1)(3^m - 1)$ and $F_0(z) = 1 - z$, $F_1(z) = (1-2z)(1-3z)$, $F_2(z) = 1 - 6z$. Thus,

$$R_f(z) = -z\left(\frac{F_2'}{F_2} - \frac{F_1'}{F_1} + \frac{F_0'}{F_0}\right) = \frac{6z}{1-6z} - \frac{2z}{1-2z} - \frac{3z}{1-3z} + \frac{z}{1-z}$$

and

$$\exp\left(-\sum_{m=1}^{\infty} r_m \frac{z^m}{m}\right) = \frac{(1-6z)(1-z)}{(1-2z)(1-3z)}.$$

Following [**4**], we discuss how to deal with absolute values in the $k = \mathbb{R}$ case. Let $f \in \mathbb{R}[x]$ have degree d such that the r_m as defined above are all nonzero. We examine the sign of r_m using equation (4.2). First notice that a complex conjugate pair of roots of f does not affect the sign of r_m. A real root α of f contributes a sign factor of $+1$ if $\alpha > 1$, -1 if $-1 < \alpha < 1$, and $(-1)^m$ if $\alpha < -1$. Let E be the number of zeroes of f in $(-1, 1)$ and let D be the number of zeroes in $(-\infty, -1)$. Also, set $\epsilon = (-1)^E$ and $\delta = (-1)^D$. Then, it follows that

$$\frac{r_m}{|r_m|} = \epsilon \cdot \delta^m.$$

In particular,

$$|r_m| = \epsilon(\delta a_0)^m \prod_{i=1}^{d} (\alpha_i^m - 1). \tag{4.5}$$

In other words, the sequence of $|r_m|$ is obtained by multiplying each cyclic resultant of the polynomial $\tilde{f} := \delta f = \delta a_0 x^d + \delta a_1 x^{d-1} + \cdots + \delta a_d$ by ϵ. Denoting by $\widetilde{G}_d$ the rational function determined by $\tilde{f}$ as in (4.4), it follows that

$$\exp\left(-\sum_{m=1}^{\infty} |r_m| \frac{z^m}{m}\right) = \left(\widetilde{G}_d\right)^{\epsilon}. \tag{4.6}$$

5. Proofs of the Main Theorems

Let G be the multiplicative group generated by the nonzero roots $\alpha_1, \ldots, \alpha_d$ of f. Vector space basis elements of the group ring $k[G]$ will be represented by $[\alpha]$, $\alpha \in G$. The divisor (in $k[G]$) of the rational function G_d defined by Corollary 4.4 is

$$(-1)^{d+1}\left(\sum_{k \text{ odd}} \sum_{i_1 < \cdots < i_k} \left[(a_0 \alpha_{i_1} \cdots \alpha_{i_k})^{-1}\right] - \sum_{k \text{ even}} \sum_{i_1 < \cdots < i_k} \left[(a_0 \alpha_{i_1} \cdots \alpha_{i_k})^{-1}\right]\right) \tag{5.1}$$

$$= \left[a_0^{-1}\right] \prod_{i=1}^{d} \left(\left[\alpha_i^{-1}\right] - [1]\right).$$

Let us remark that for ease of presentation above, when $k = 0$, we have assigned

$$\sum_{i_1 < \cdots < i_k} \left[(a_0 \alpha_{i_1} \cdots \alpha_{i_k})^{-1}\right] = [a_0^{-1}],$$

which corresponds to the factor of $F_0(z) = 1 - a_0 z$ in G_d. With this computation in hand, we now prove our main theorems.

Proof of Theorem 2.1. Examining the statement of the theorem, we may assume that k is algebraically closed. Let f and g be polynomials in $k[x]$ such that the multiplicity of 0 as a root of f (g) is l_1 (l_2). Then, $f(x) = x^{l_1}(a_0 x^{d_1} + \cdots + a_{d_1})$ and $g(x) = x^{l_2}(b_0 x^{d_2} + \cdots + b_{d_2})$ in which a_0 and b_0 are not 0. Let $\alpha_1, \ldots, \alpha_{d_1}$ and $\beta_1, \ldots, \beta_{d_2}$ be the nonzero roots of f and g, respectively, and let G be the multiplicative group generated by these elements. Since $f(x)$ and $g(x)$ both generate the same sequence of cyclic resultants, it follows that the divisor (in the group ring $k[G]$) of their corresponding rational functions (see (4.4)) are equal. By above, such divisors factor, giving us that

$$(-1)^{d_1}[a_0^{-1}] \prod_{i=1}^{d_1} \left([1] - [\alpha_i^{-1}]\right) = (-1)^{d_2}[b_0^{-1}] \prod_{i=1}^{d_2} \left([1] - [\beta_i^{-1}]\right).$$

Since we have assumed that f and g generate a set of nonzero cyclic resultants, neither of them can have a root of unity as a zero. Therefore, Lemma 3.5 applies to give us that $d := d_1 = d_2$ and that up to a permutation, there is a nonnegative integer p such that

(1) $\alpha_i = \beta_i$ for $i = 1, \ldots, p$
(2) $\alpha_i = \beta_i^{-1}$ for $i = p+1, \ldots, d$
(3) $(-1)^{d-p} = 1$, $a_0 b_0^{-1} = \beta_{p+1} \cdots \beta_d$.

Set $u(x) = (x - \beta_{p+1}) \cdots (x - \beta_d)$, which has even degree, and let $v(x) = b_0(x - \beta_1) \cdots (x - \beta_p)$ (note that if $p = 0$, then $v(x) = b_0$) so that $g(x) = x^{l_2} v(x) u(x)$. Now,

$$u(x^{-1}) x^{\deg(u)} = (-1)^{d-p} \beta_{p+1} \cdots \beta_d (x - \beta_{p+1}^{-1}) \cdots (x - \beta_d^{-1}),$$

and thus

$$\begin{aligned} f(x) &= x^{l_1} a_0 b_0^{-1} v(x)(x - \beta_{p+1}^{-1}) \cdots (x - \beta_d^{-1}) \\ &= x^{l_1} v(x) u(x^{-1}) x^{\deg(u)}. \end{aligned}$$

It remains only to argue that $l_1 \equiv l_2 \pmod 2$. However, from formula (4.2) with $m = 1$, it is easily seen that $(-1)^{l_1} = (-1)^{l_2}$. The converse is also straightforward from (4.2), and this completes the proof of the theorem. □

The proof of Theorem 2.6 is similar, employing equation (4.6) in place of (4.4).

Proof of Theorem 2.6. Since multiplication of a real polynomial by a power of x does not change the absolute value of a cyclic resultant, we may assume $f, g \in \mathbb{R}[x]$ have distinct roots. The result now follows from (4.6) and the argument used to prove the if-direction of Theorem 2.1. □

6. Algorithms Related to Cyclic Resultants

In the proof of Theorem 2.1, the multiplicative group generated by the roots of f played an important role; which leads us to the following natural question. Given a polynomial $f \in \mathbb{Z}[x]$ of degree d, can one devise an algorithm to determine the structure of the group G generated by the roots of f? Of course, G will be a direct sum of a free abelian group and a finite cyclic group, so one possible output would consist of two numbers: the rank of the free part and the order of the cyclic component. Another description would be to give generators for the lattice L, where L is the kernel of the homomorphism sending the generators of $\mathbb{Z}^d$ to the roots of f.

It turns out that an algorithm does indeed exist, however, it is exponential in d. The result is due to Ge [**5**], although our question is a special case of a more general problem he studied. Given a finite list of nonzero elements of an algebraic number field K, Ge has an algorithm that determines a generating set for the group of all multiplicative relations between those elements (and therefore the structure of the subgroup they generate). It would be nice to know if there is a better (polynomial) time procedure to solve our special case, however, we do not know of any work in this direction.

7. Acknowledgement

We would like to thank Bernd Sturmfels for bringing this problem to our attention and for the idea of reformulating Lemma 3.5 into the statement of Theorem 2.12.

References

[1] D. Cox, J. Little, D. O'Shea, *Using Algebraic Geometry*, Springer, New York, 1998.
[2] J.J. Duistermaat and V. Guillemin, *The spectrum of positive elliptic operators and periodic bicharacteristics*, Inv. Math. 25 (1975) 39-79.
[3] G. Everest and T. Ward. Heights of Polynomials and Entropy in Algebraic Dynamics. Springer-Verlag London Ltd., London, 1999.
[4] D. Fried, *Cyclic resultants of reciprocal polynomials*, in Holomorphic Dynamics (Mexico 1986), Lecture Notes in Math. 1345, Springer Verlag, 1988, 124-128.
[5] Guoqiang Ge, *Algorithms related to multiplicative representations of algebraic numbers*, PhD thesis, Math Dept, U. C. Berkeley, 1993.
[6] V. Guillemin, *Wave trace invariants*, Duke Math. J. 83 (1996), 287-352.
[7] A. Iantchenko, J. Sjöstrand, and M. Zworski, *Birkhoff normal forms in semi-classical inverse problems*, preprint.
[8] E. Miller and B. Sturmfels, *Combinatorial Commutative Algebra*, Springer, 2004.
[9] W. H. Stevens, *Recursion formulas for some abelian knot invariants*, Journal of Knot Theory and Its Ramifications, Vol. 9, No. 3 (2000) 413-422.

Department of Mathematics, University of California, Berkeley, CA 94720.
E-mail address: chillar@math.berkeley.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Rectangular Schur Functions and Fermions

Takeshi Ikeda, Hiroshi Mizukawa, Tatsuhiro Nakjima, and Hiro-Fumi Yamada

Abstract. *We give an expression of the Schur function* $S_{\square(m,n)}$*, indexed by the rectangular partition* $\square(m,n) = (n^m)$ *as a sum of products of certain Schur functions and Schur's Q-functions.*
Résumé. *Nous donnons une expression de la fonction de Schur* $S_{\square(m,n)}$*, indexée par la partition rectangulaire* $\square(m,n) = (n^m)$*, comme une somme de produits de fonctions de Schur et de Q-fonctions de Schur.*

1. Introduction

Let λ be a strict partition. Draw the Young diagram of λ and fill each cell with 0 or 1 in such a way that, in each row the sequence (0110) repeats from the left as long as possible. For a positive integer ℓ, set $A_\ell = (4\ell-3, 4\ell-7, \cdots, 5, 1)$. Let $\mathcal{F}_1^n(A_\ell)$ be the set of strict Young diagrams which are obtained by appending n $\boxed{1}$'s to A_ℓ. Our formula reads

$$\sum_{\mu \in \mathcal{F}_1^n(A_\ell)} \delta(\mu) Q_{\mu^b[0]}(t) S_{\mu^b[1]}(t') = S_{\square(2\ell-n,n)}(t),$$

where $(\mu^b[0], \mu^b[1])$ is the 4-bar quotient of μ, $\delta(\mu)$ is a sign, $S_\nu(t)$ and $Q_\nu(t)$ are the Schur function and the Q-function, respectively, corresponding to the partition ν, expressed as polynomials of the power sum symmetric functions (or the so-called Sato variables) $t = (t_1, t_2, t_3, \cdots)$ and $t' = (t_2, t_4, t_6, \cdots)$.

We understand this formula from the viewpoint of the basic representation $L(\Lambda_0)$ of the affine Lie algebra of type $A_1^{(1)}$, or more suitably, type $D_2^{(2)}$. It is known that the weight vectors of the basic representation of $D_2^{(2)}$ are, in the principal picture, best described by means of the Q-functions ([**6**]). In particular the maximal weight vectors are the Q-functions $Q_\lambda(t)$ with $\lambda = A_\ell = (4\ell-3, 4\ell-7, \cdots, 5, 1)$ or $\lambda = B_\ell = (4\ell-1, 4\ell-5, \cdots, 7, 3, 0) \quad (\ell = 0, 1, 2, \cdots)$. To give the intertwining operator between the principal and homogeneous realizations we make use of 4-bar quotients of the strict partitions, which arise naturally from the parting the neutral fermions into the neutral and charged fermions. Through this intertwining operator, one sees that, in the homogeneous realization, $f_i^n v \quad (i = 0, 1, n = 0, 1, 2, \cdots)$ is expressed as a rectangular Schur function for any maximal weight vector v. For the identification we will employ some fermion calculus.

2. Combinatorics of strict partitions

Let P_n denote the set of all partitions of n, SP_n the set of all strict partitions of n, and OP_n the set of those partitions of n whose parts are odd numbers. For $\lambda \in P_n$, $\ell(\lambda)$ denotes the number of non-zero parts

1991 *Mathematics Subject Classification.* Primary 17B67, 05E05; Secondary 37K30.
Key words and phrases. Schur function, affine Lie algebra, fermion.

of λ. We need the following "4-bar abacus":

0	①	3
②		
4	5	7
⑥		
8	⑨	⑪
10		
12	13	15

For a strict partition λ we put a set of beads on the assigned positions. The above figure is the 4-bar abacus representing the strict partition $\lambda = (11, 9, 6, 2, 1)$. From the 4-bar abacus of the given strict partition λ, we read off a triplet of partitions $(\lambda^{bc}, \lambda^b[0], \lambda^b[1])$ as follows: The strict partition $\lambda^b[0]$ is obtained by reading the halves of even positions of the beads. In the above example we see that $\lambda^b[0] = (3, 1)$. For the right two runners, horizontal levels are numbered as $0, 1, 2, \cdots$ from the top. We mark (by writing 1) the levels of the beads on the central runner and unmark (by writing 0) the vacancies. In the above example we get $(1, 0, 1, \underline{0}) = (1, 0, 1, 0, 0, \cdots)$. Also we unmark the levels of the beads on the rightmost runner and mark the vacancies. In the example we get $(1, 1, 0, \underline{1}) = (1, 1, 0, 1, 1, \cdots)$. Arrange the two obtained infinite $(0, 1)$-sequences:

$$\underline{1}011|101\underline{0}$$

On the right of the bar "|", the sequence of the central runner comes, and on the left the reversed sequence of the rightmost runner comes. Counting the 0's to the left of each 1, we get the partition $\lambda^b[1]$. The above $(0, 1)$-sequence shows that $\lambda^b[1] = (2, 1, 1, 1)$. Finally, if $\ell \in \mathbb{Z}$ is the number of beads on the central runner minus that on the rightmost runner, we set $\lambda^{bc} = A_\ell = (4\ell - 3, 4\ell - 7, \cdots, 5, 1)$ for $\ell \geq 0$ ($A_0 = \emptyset$), and $\lambda^{bc} = B_{|\ell|} = (4|\ell| - 1, 4|\ell| - 5, \cdots, 7, 3, 0)$ for $\ell < 0$. In the above example we see that $\lambda^{bc} = A_1 = (1)$. Note that $|\lambda^{bc}| + 2|\lambda^b[0]| + 4|\lambda^b[1]| = |\lambda|$. The above procedure is invertible and the correspondence between SP_n and the set $\{(\lambda^{bc}, \lambda^b[0], \lambda^b[1]); |\lambda^{bc}| + 2|\lambda^b[0]| + 4|\lambda^b[1]| = n\}$ is shown to be one-to-one. The strict partition λ^{bc} is called the "4-bar core" of λ and the pair $(\lambda^b[0], \lambda^b[1])$ is called the "4-bar quotient" of λ (cf. [**7**]).

For a strict partition λ we draw the Young diagram and fill each cell with 0 or 1 in such a way that, in each row the sequence (0110) repeats from the left as long as possible. Let $\mathcal{F}_i^n(\lambda)$ $(i = 0, 1)$ denote the set of the strict partitions obtained by appending n $\boxed{i}$'s to the Young diagram λ. It is easy to see that the cardinality of $\mathcal{F}_1^n(A_\ell)$ is the coefficient of x^n of $(1 + x + x^2)^\ell$, and that of $\mathcal{F}_0^n(B_\ell)$ is the sum of coefficients of x^n and x^{n-1} of the same polynomial.

Each strict partition μ in $\mathcal{F}_1^n(A_\ell)$ or $\mathcal{F}_0^n(B_\ell)$ has its own sign $\delta'(\mu) = (-1)^g$, where g is, in the 4-bar abacus of μ, the number of beads on the central runner at the positions bigger than that of each bead on the leftmost runner. For example, for $\mu = (9, 7, 2) \in \mathcal{F}_1^3(A_3)$, whose 4-bar abacus looks

0	1	3
②		
4	5	⑦
6		
8	⑨	11
10		
12	13	15

the sign is $\delta'(\mu) = -1$, since $g = 1$.

3. The formula

Let χ_ρ^λ be the irreducible character of the symmetric group S_n, indexed by $\lambda \in P_n$ and evaluated at the conjugacy class $\rho \in P_n$. And let ζ_ρ^λ be the irreducible negative character of the double cover $\tilde{S_n}$ of the symmetric group, indexed by $\lambda \in SP_n$ and evaluated at the conjugacy class $\rho \in OP_n$. In our context the Schur function indexed by $\lambda \in P_n$ is defined, as a polynomial of $u = (u_1, u_2, u_3, \cdots)$, by

$$S_\lambda(u) = \sum_{\rho \in P_n} \chi^\lambda_\rho \frac{u_1^{m_1} u_2^{m_2} \cdots}{m_1! m_2! \cdots},$$

where the summation runs over the partitions $\rho = (1^{m_1} 2^{m_2} \cdots) \in P_n$ (cf. [**5**]). Schur's Q-function indexed by $\lambda \in SP_n$ appears, as a polynomial of $t = (t_1, t_3, t_5, \cdots)$, in the form

$$Q_\lambda(t) = \sum_{\rho \in OP_n} 2^{\frac{\ell(\lambda) - \ell(\rho) + \epsilon}{2}} \zeta^\lambda_\rho \frac{t_1^{m_1} t_3^{m_3} \cdots}{m_1! m_3! \cdots},$$

where the summation runs over the partitions $\rho = (1^{m_1} 3^{m_3} \cdots) \in OP_n$, and $\epsilon = 0$ or 1 according to that $n - \ell(\lambda)$ is even or odd (cf. [**2**]). It is sometimes convenient to normalize Q-functions as

$$P_\lambda(t) = 2^{-\ell(\lambda)} Q_\lambda(t).$$

These functions are called Schur's P-functions. We can now state our formula.

Theorem 3.1. *For non-negative integers ℓ and n, we have*

$$\sum_{\mu \in \mathcal{F}_1^n(A_\ell)} \delta'(\mu) Q_{\mu^b[0]}(t) S_{\mu^b[1]}(t') = S_{\Box(2\ell - n, n)}(t),$$

$$\sum_{\mu \in \mathcal{F}_0^n(B_\ell)} \delta'(\mu) Q_{\mu^b[0]}(t) S_{\mu^b[1]}(t') = S_{\Box(n, 2\ell + 1 - n)}(t),$$

where $t = (t_1, t_2, t_3, \cdots)$ and $S_\nu(t') = S_\nu(u)|_{u_j \mapsto t_{2j}}$.

4. Basic representation

In this section we connect our formula with the basic representation of the affine Lie algebra of type $A_1^{(1)}$. It turns out to be that our formula describes certain weight vectors in the homogeneous realization of the basic representation $L(\Lambda_0)$.

Associated with the Cartan matrix

$$(a_{ij})_{i,j=0,1} = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

the Lie algebra $\mathfrak{g}$ of type $A_1^{(1)}$ is generated by e_i, f_i, h_i and d subject to the relations

$$[h_i, h_j] = 0, \qquad [h_i, e_j] = a_{ij} e_j, \qquad [h_i, f_j] = -a_{ij} f_j,$$
$$[e_i, f_j] = \delta_{i,j} h_i, \qquad (\mathrm{ad} e_i)^{1 - a_{ij}} e_j = (\mathrm{ad} f_i)^{1 - a_{ij}} f_j = 0 \quad (i \neq j),$$

and

$$[d, h_i] = 0, \qquad [d, e_j] = \delta_{j,0} e_j, \qquad [d, f_j] = -\delta_{j,0} f_j.$$

The Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is spanned by h_0, h_1 and d. Choose the basis $\{\alpha_0, \alpha_1, \Lambda_0\}$ for the dual space $\mathfrak{h}^*$ of $\mathfrak{h}$ by the pairing

$$< h_i, \alpha_j > = a_{ij}, \qquad < h_i, \Lambda_0 > = \delta_{i,0},$$
$$< d, \alpha_j > = \delta_{0,j}, \qquad < d, \Lambda_0 > = 0.$$

The fundamental imaginary root is $\delta = \alpha_0 + \alpha_1$.

The basic representation $L(\Lambda_0)$ of $\mathfrak{g}$ is by definition the irreducible highest weight $\mathfrak{g}$-module with highest weight Λ_0 (cf. [**4**]). The weight system of $L(\Lambda_0)$ is well-known:

$$P(\Lambda_0) = \{\Lambda_0 - p\delta + q\alpha_1; p, q \in \mathbb{Z}, p \geq q^2\}.$$

A weight Λ on the parabola $\Lambda_0 - q^2\delta + q\alpha_1 \quad (q \in \mathbb{Z})$ is said to be maximal in the sense that $\Lambda + \delta$ is no longer a weight.

We discuss a twisted version of the principal realization of $L(\Lambda_0)$, or more suitably, the basic representation of the affine Lie algebra of type $D_2^{(2)}$, which is isomorphic to $A_1^{(1)}$.

The basic representation in principal grading is realized on the space

$$V^{(2)} = \mathbb{C}[t_j; j \geq 1, \text{odd}]$$

of the P-functions ([**6**]). In fact, the P-functions form a weight basis of $L(\Lambda_0) = V^{(2)}$. Given a strict partition λ, fill the Young diagram with 0 or 1 as in Section 2. If the number of $\boxed{i}$'s is m_i $(i = 0, 1)$, then the weight of the corresponding P-function $P_\lambda(t)$ equals $\Lambda_0 - m_0\alpha_0 - m_1\alpha_1$. In particular the weight of $P_{A_\ell}(t)$ (resp. $P_{B_\ell}(t)$) equals $\Lambda_0 - \ell^2\delta + \ell\alpha_1$ (resp. $\Lambda_0 - \ell^2\delta - \ell\alpha_1$), which is maximal for any $\ell \geq 0$. The action of $f_i \in \mathfrak{g} \quad (i = 0, 1)$ to the P-function $P_\lambda(t)$ is easily described:

$$f_i P_\lambda = \sum_{\mu \in \mathcal{F}_i^1(\lambda)} P_\mu.$$

Strict partitions in $\mathcal{F}_i^n(\lambda)$ occur in the expression of $f_i^n P_\lambda$.

Another realization of the basic representation is known, one in the homogeneous grading. The representation space turns out $\mathcal{B} = V \otimes \mathbb{C}[q, q^{-1}]$, where $V = \mathbb{C}[t_j; j \geq 1]$. Define the linear isomorphism Ψ by

$$\Psi : V^{(2)} \longrightarrow \mathcal{B}$$
$$P_\lambda(t) \mapsto 2^{\frac{\ell(\lambda^b[0])}{2}} \delta(\lambda) P_{\lambda^b[0]}(t) S_{\lambda^b[1]}(t') \otimes q^{m(\lambda)}$$

where $m(\lambda)$ is determined by drawing the 4-bar abacus of λ:

$$\begin{aligned} m(\lambda) = &\ (\text{number of beads on the central runner of } \lambda) \\ &- (\text{number of beads of the rightmost runner of } \lambda) \end{aligned}$$

and $\delta(\lambda)$ is certain sign which is naturally determined by arranging the fermion operators corresponding to λ. Here we only remark that $\delta(\lambda)$ coincides with $\delta'(\lambda)$ for λ in $\mathcal{F}_1^n(A_\ell)$.

The representation of $\mathfrak{g}$ on $\mathcal{B}$, which is induced by Ψ, is the basic representation in the homogeneous grading. In fact, if we define the degree in $\mathcal{B}$ by

$$\deg(f(t) \otimes q^m) = \deg f(t) + m^2,$$

then $\deg\Psi(P_\lambda)$ is equal to the number of $\boxed{0}$'s in λ.

Now our first formula can be translated into

$$\Psi(\frac{1}{n!} f_1^n P_{A_\ell}) = 2^{-\frac{n}{2}} S_{\square(2\ell-n,n)} \otimes q^{\ell-n}.$$

As for the second formula we need to extend Ψ to a superspace $V^{(2)} \oplus V^{(2)}\theta$.

5. Fermionic Fock space

In this section we look at the formula from the fermionic point of view. Although we will not give a proof of the formula in this extended abstract, we emphasize that the discussion of this section is essential for our proof.

We first recall how to realize the basic representation of $\mathfrak{g}$ in terms of free fermions. Let $\mathbb{B}$ be the $\mathbb{C}$-algebra generated by $\beta_n \quad (n \in \mathbb{Z})$ subject to the relations

$$\beta_n\beta_m + \beta_m\beta_n = (-1)^n \delta_{n+m,0} \quad (n, m \in \mathbb{Z}).$$

Note that $\beta_0^2 = 1/2$. These generators are often called the neutral free fermions. Define the degree on $\mathbb{B}$ by $\deg\beta_n = 1$. If we let $\mathbb{B}_0$ (resp. $\mathbb{B}_1$) be the subspace consisting of the elements of even (resp. odd) degree, then $\mathbb{B} = \mathbb{B}_0 \oplus \mathbb{B}_1$ has a structure of a superalgebra. Let $\mathcal{F} = \mathcal{F}_0 \oplus \mathcal{F}_1 = \mathbb{B}_0|0\rangle \oplus \mathbb{B}_1|0\rangle$ be the fermionic Fock space, where the vacuum $|0\rangle$ is defined by $\beta_n|0\rangle = 0 \quad (n < 0)$. The vacuum expectation value $\langle 0|a|0\rangle \quad (a \in \mathbb{B})$ is uniquely determined by putting $\langle 0|1|0\rangle = 1, \langle 0|\beta_0|0\rangle = 0$. The normal ordering for the quadratic elements is defined by $:\beta_n\beta_m: = \beta_n\beta_m - \langle 0|\beta_n\beta_m|0\rangle$.

Set

$$e_0 = \sqrt{2}\sum_{n\in\mathbb{Z}} \beta_{4n}\beta_{-4n-1}, \quad e_1 = \sqrt{2}\sum_{n\in\mathbb{Z}} \beta_{4n+2}\beta_{-4n-3}$$

$$f_0 = -\sqrt{2}\sum_{n\in\mathbb{Z}} \beta_{4n}\beta_{-4n+1}, \quad f_1 = -\sqrt{2}\sum_{n\in\mathbb{Z}} \beta_{4n+2}\beta_{-4n-1}$$

$$h_1 = 2\sum_{n\in\mathbb{Z}} :\beta_{4n+3}\beta_{-4n-3}:$$

and $h_0 = 1 - h_1$. These elements generate a Lie algebra inside $\mathbb{B}_0$, which is known to be isomorphic to the affine Lie algebra $\mathfrak{g}$ of type $A_1^{(1)}$. The representation of $\mathfrak{g}$ on $\mathcal{F}$ via the action of $\mathbb{B}_0$ turns out to be the direct sum $V^{(2)} \oplus V^{(2)}\theta$ of the basic representation, where θ is a symbol satisfying $\theta^2 = 1$. One often identifies θ with $\sqrt{2}\beta_0$. The isomorphism Φ_P from $\mathcal{F}$ to $V^{(2)} \oplus V^{(2)}\theta$ is given by

$$\Phi_P : a|0\rangle \mapsto \langle 0|e^{H_B(t)}a|0\rangle + \langle 0|\sqrt{2}\beta_0 e^{H_B(t)}a|0\rangle\theta \quad (a \in \mathbb{B}),$$

where $H_B(t) = \frac{1}{2}\sum_{j\geq 1,\mathrm{odd}}\sum_{n\in\mathbb{Z}}(-1)^{n+1}t_j\beta_n\beta_{-n-j}$. This type of isomorphism is often called the boson-fermion correspondence (cf [**1**] or [**6**]). A standard fermion calculus shows that, putting $\beta_\lambda|0\rangle = \beta_{\lambda_1}\cdots\beta_{\lambda_\ell}|0\rangle$, $\Phi_P(\beta_\lambda|0\rangle) = \sqrt{2}^{-\ell}Q_\lambda(t)\theta^\epsilon$ for a strict partition $\lambda = (\lambda_1, \cdots, \lambda_\ell) \quad (\lambda_1 > \cdots > \lambda_\ell > 0)$, where $\epsilon = 0$ or 1 according to that ℓ is even or odd. In order to give a boson-fermion correspondence in the homogeneous grading, we make parting of the neutral free fermions into three groups: $\{\psi_n; n \in \mathbb{Z}\}, \{\psi_n^*; n \in \mathbb{Z}\}$ and $\{\phi_n; n \in \mathbb{Z}\}$, where

$$\psi_n = i\beta_{4n+1}, \psi_n^* = i\beta_{-4n-1}, \phi_{2n} = \beta_{4n}, \phi_{2n+1} = i\beta_{4n+2} \quad (i = \sqrt{-1}).$$

A product of β's is rewritten as a word of ψ, ψ^* and ϕ's.

For an integer m and for $\epsilon = 0, 1$, we set

$$\langle m, \epsilon| = \begin{cases} \langle 0|\psi_0^*\cdots\psi_{m-1}^*\theta^\epsilon & (m \geq 1) \\ \langle 0|\theta^\epsilon & (m = 0) \\ \langle 0|\psi_{-1}\cdots\psi_m\theta^\epsilon & (m \leq -1). \end{cases}$$

The homogeneous boson-fermion correspondence

$$\Phi_H : \mathcal{F} \longrightarrow (V \oplus V\theta) \otimes \mathbb{C}[q, q^{-1}] \cong \mathcal{B} \oplus \mathcal{B}\theta$$

is given by

$$\Phi_H(a|0\rangle) = \sum_{m\in\mathbb{Z},\epsilon=0,1} \langle m, \epsilon|e^{H_A(t)+H_B(t)}a|0\rangle\theta^\epsilon \otimes q^m \quad (a \in \mathbb{B}),$$

where $H_A(t) = \frac{1}{2}\sum_{j,n\in\mathbb{Z}} t_{2j} :\psi_n\psi_{n+j}^*:$. It is easily seen that

$$\Phi_H(\mathcal{F}_0) = V \otimes \mathbb{C}[q^2, q^{-2}] \oplus V\theta \otimes \mathbb{C}[q^2, q^{-2}]q,$$

$$\Phi_H(\mathcal{F}_1) = V\theta \otimes \mathbb{C}[q^2, q^{-2}] \oplus V \otimes \mathbb{C}[q^2, q^{-2}]q.$$

Both are isomorphic to $V \otimes \mathbb{C}[q, q^{-1}] = \mathcal{B}$.

Here we give an example illustrating how to associate a polynomial with an element of the fermionic Fock space $\mathcal{F}$. Take $A_3 = (9,5,1)$ and $\mathcal{F}_1^1(A_3) = \{(10,5,1),(9,6,1),(9,5,2)\}$. For $\mu = (9,5,2)$, we consider $\beta_\mu = \beta_9\beta_5\beta_2|0\rangle = \phi_1\psi_2\psi_1|0\rangle \in \mathcal{F}$ and see that

$$\begin{aligned}\Phi_H(\beta_9\beta_5\beta_2|0\rangle) &= \Phi_H(\phi_1\psi_2\psi_1|0\rangle)\\ &= \langle 2,1|e^{H_A(t)+H_B(t)}\phi_1\psi_2\psi_1|0\rangle\theta \otimes q^2\\ &= \langle 0|\psi_0^*\psi_1^*\sqrt{2}\phi_0 e^{H_A(t)+H_B(t)}\phi_1\psi_2\psi_1|0\rangle\theta \otimes q^2\\ &= \sqrt{2}\langle 0|e^{H_B(t)}\phi_1\phi_0|0\rangle\theta\langle 0|e^{H_A(t)}\psi_0^*\psi_1^*\psi_2\psi_1|0\rangle \otimes q^2\\ &= \sqrt{2}^{-1}Q_{(1)}(t)S_{(1,1)}(t')\theta \otimes q^2.\end{aligned}$$

Likewise one computes

$$\Phi_H(\beta_{10}\beta_5\beta_1|0\rangle) = \sqrt{2}^{-1}Q_{(5)}(t)\theta \otimes q^2,$$
$$\Phi_H(\beta_9\beta_6\beta_1|0\rangle) = -\sqrt{2}^{-1}Q_{(3)}(t)S_{(1)}(t')\theta \otimes q^2.$$

A combinatorial calculation shows that, for a strict partition λ,

$$\Phi_H(\beta_\lambda|0\rangle) = cQ_{\lambda^b[0]}(t)S_{\lambda^b[1]}(t')\theta^\epsilon \otimes q^{m(\lambda)},$$

where $\epsilon = 0$ or 1, and $c = \pm\sqrt{2}^g$ with $g \in \mathbb{Z}$. The sign ($\pm$) comes from the arranging the fermions to the normal form, which we shall discuss in the next section. The action of the Lie algebra $\mathfrak{g}$ on $\mathcal{F}$ (and on $\mathcal{B}\oplus\mathcal{B}\theta$ via Φ_H) is best described in terms of the vertex operator. To prove the formula we employ a calculation of the vertex operators (cf. [**3**]), which we shall omit here.

6. Determining the sign

We see that the sign which appears in our formula can be easily determined by looking at the 4-bar abacuses. We explain this fact through an example. Take the partition

$$\lambda = (35,31,25,23,18,15,11,6,1) \in \mathcal{F}_1^{12}(A_9)$$

and write the corresponding state

$$\beta_\lambda|0\rangle = \psi_{-9}^*\psi_{-8}^*\psi_6\psi_{-6}^*\phi_9\psi_{-4}^*\psi_{-3}^*\phi_3\psi_0|0\rangle.$$

We rewrite this state into the normal form as follows.

Step 1. Equate the number of ψ's and ψ^*'s by "shifting the vacuum".
Step 2. Move ϕ's to the left of ψ's and ψ^*'s.
Step 3. Make pairs $\psi^*\psi$:

$$\begin{aligned}&\psi_{-9}^*\psi_{-8}^*\psi_6\psi_{-6}^*\phi_9\psi_{-4}^*\psi_{-3}^*\phi_3\psi_0|0\rangle\\ &\underset{\text{step1}}{=} (-1)^4\psi_{-9}^*\psi_{-8}^*\psi_6\psi_{-6}^*\phi_9\psi_{-4}^*\phi_3\psi_0\psi_{-1}\psi_{-2}|-3\rangle\\ &\underset{\text{step2}}{=} (-1)^{4+9}\phi_9\phi_3\psi_{-9}^*\psi_{-8}^*\psi_6\psi_{-6}^*\psi_{-4}^*\psi_0\psi_{-1}\psi_{-2}|-3\rangle\\ &\underset{\text{step3}}{=} (-1)^{4+9+a'}\phi_9\phi_3(\psi_{-9}^*\psi_6)(\psi_{-8}^*\psi_0)(\psi_{-6}^*\psi_{-1})(\psi_{-4}^*\psi_{-2})|-3\rangle,\end{aligned}$$

where the shifted vacuum is, by definition,

$$|m\rangle = \begin{cases}\psi_{m-1}\cdots\psi_0|0\rangle & (m \geq 1)\\ |0\rangle & (m=0)\\ \psi_m^*\cdots\psi_{-1}^*|0\rangle & (m \leq -1).\end{cases}$$

We observe that step 1 does not change the sign. Therefore we only have to consider the sign change which comes from step 2 and step 3. We express the state by the following (modified) 4-bar abacus and attach a number to each bead:

$$
\begin{array}{ccc}
\psi & \phi & \psi^* \\
 & & \\
 & & \\
 & & \\
\textcircled{0} & 1 & -1 \\
1 & \textcircled{3} & -2 \\
2 & 5 & \textcircled{-3} \\
3 & 7 & \textcircled{-4} \\
4 & \textcircled{9} & -5 \\
5 & 11 & \textcircled{-6} \\
\textcircled{6} & 15 & -7 \\
7 & 17 & \textcircled{-8} \\
8 & 19 & \textcircled{-9}
\end{array}
\quad \underset{\text{step1}}{\longrightarrow} \quad
\begin{array}{ccc}
\psi & \phi & \psi^* \\
-3 & & \\
\textcircled{-2}_{10} & & \\
\textcircled{-1}_{8} & & \\
\textcircled{0}_{6} & 1 & \\
1 & \textcircled{3}_{2} & \\
2 & 5 & \\
3 & 7 & \textcircled{-4}_{9} \\
4 & \textcircled{9}_{1} & -5 \\
5 & 11 & \textcircled{-6}_{7} \\
\textcircled{6}_{4} & 15 & -7 \\
7 & 17 & \textcircled{-8}_{5} \\
8 & 19 & \textcircled{-9}_{3}
\end{array}
\quad \underset{\text{word}}{\longrightarrow} \quad
\begin{array}{c}
w = 3\ 5\ 4\ 7\ 1\ 9\ 2\ 6\ 8\ 10, \\
\text{number of inversions of } w = 13.
\end{array}
$$

The numbering of the beads is given in the following way.

1. Number ϕ's from bottom to top.
2. Number ψ^*'s and ψ's according to the *layers*.

We read the numbers by rows from bottom to top and count the inversions involved in the obtained word.

One can see that this inversion number gives the number $a(= a' + 9)$ of the interchanges of fermions.

Next we consider the sign which comes from the boson-fermion correspondence, i.e.,

$$
\Phi_H(\phi_9\phi_3(\psi^*_{-9}\psi_6)(\psi^*_{-8}\psi_0)(\psi^*_{-6}\psi_{-1})(\psi^*_{-4}\psi_{-2})|-3\rangle) \\
= (-1)^{1+3+5+6}\frac{1}{2}Q_{(9,3)}S_{(10,5^3,3,2)}.
$$

We read $1, 3, 5, 6$ by renumbering beads on the rightmost runner.

$$
\begin{array}{c}
\textcircled{-4}_{1} \\
-5_{2} \\
\textcircled{-6}_{3} \\
-7_{4} \\
\textcircled{-8}_{5} \\
\textcircled{-9}_{6}
\end{array}
\to (1, 3, 5, 6) \to b = 1 + 3 + 5 + 6.
$$

Finally we get the desired sign $\delta(\lambda) = (-1)^{a+b}$. A more careful case-by-case check shows that $\delta(\lambda) = \delta'(\lambda)$(see Section 2) for $\lambda \in \mathcal{F}_1^n(A_\ell) \cup \mathcal{F}_0^n(B_\ell)$.

References

[1] E. Date, M. Kashiwara, M. Jimbo and T. Miwa, Transformation groups for soliton equations, in Nonlinear Integrable Systems—Classical Theory and Quantum Theory, ed. by M. Jimbo and T. Miwa, 39-119, World Sci. Publishing, Singapore, 1983.

[2] P. N. Hoffman and J. F. Humphreys, *Projective Representations of the Symmetric Groups,* Oxford, 1992.

[3] T. Ikeda and H.-F. Yamada, Polynomial τ-functions of the NLS-Toda hierarchy and the Virasoro singular vectors, Lett. Math. Phys. 60 (2002), 147-156.

[4] V. G. Kac, *Infinite Dimensional Lie Algebras, 3rd. ed.* , Cambridge, 1990.

[5] I. G. Macdonald, *Symmetric Functions and Hall Polynomials, 2nd. ed.* , Oxford, 1995.

[6] T. Nakajima and H.-F. Yamada, Schur's Q-functions and twisted affine Lie algebras, Adv. Stud. in Pure Math. 28 (2000), 241-259.

[7] J. B. Olsson, Combinatorics and Representations of Finite Groups, Essen, 1993.

Takeshi IKeda, Department of Applied Mathematics, Okayama University of Science, Okayama 700-0005, Japan
E-mail address: ike@xmath.ous.ac.jp

Hiroshi Mizukawa, Department of Mathematics, National Defense Academy in Japan, Yokosuka 239-8686, Japan
E-mail address: mzh@math.okayama-u.ac.jp

Tatsuhiro Nakajima, Faculty of Economics, Meikai University, Urayasu 279-8550, Japan
E-mail address: nakajima@phys.metro-u.ac.jp

Hiro-Fumi Yamada, Department of Mathematics, Okayama University, Okayama 700-8530, Japan
E-mail address: yamada@math.okayama-u.ac.jp

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Bruhat Order on the Involutions of Classical Weyl Groups

Federico Incitti

Abstract. *It is known that a Coxeter group W, partially ordered by the Bruhat order, is a graded poset, with rank function given by the length, and that it is EL-shellable, hence Cohen-Macaulay, and Eulerian. In this work we consider the subposet of W induced by the set of involutions of W, denoted by $Invol(W)$. Our main result is that, if W is a classical Weyl group, then the poset $Invol(W)$ is graded, with rank function given by the average between the length and the absolute length, and that it is EL-shellable, hence Cohen-Macaulay, and Eulerian. In particular we obtain, as new results, a combinatorial description of the covering relation in the Bruhat order of the hyperoctahedral group and the even-signed permutation group, and a combinatorial description of the absolute length of the involutions in classical Weyl groups.*

Résumé. *Il est bien connu qu'un groupe de Coxeter W, munis de l'ordre de Bruhat, est un poset gradué, avec fonction rang donnée par la longueur, et qu'il est EL-shellable, donc de Cohen-Macaulay, et Eulerien. Dans cet article on considère le sous-poset induit par l'ensemble des involutions de W, noté $Invol(W)$. Nous montrons que, si W est un groupe de Weyl classique, alors le poset $Invol(W)$ est gradué, avec fonction rang égale à la moyenne entre la longueur et la longueur absolue, et qu'il est EL-shellable, donc de Cohen-Macaulay, et Eulerien. Nous obtenons en particulier deux résultats nouveaux: une description combinatoire de la relation de couverture dans l'ordre de Bruhat de B_n et D_n, et une description combinatoire de la longueur absolue des involutions dans les groupes de Weyl classiques.*

1. Introduction

It is known that a Coxeter group W, partially ordered by the Bruhat order, is a graded poset, with rank function given by the length, and that it is EL-shellable, hence Cohen-Macaulay, and Eulerian. The aim of this work is to investigate whether a particular subposet of W, namely that induced by the set of involutions of W, which we denote by $Invol(W)$, is endowed with similar properties.

The problem arises from a geometric question. It is known that the symmetric group, partially ordered by the Bruhat order, encodes the cell decomposition of Schubert varieties. Richardson and Springer ([**RS1**], [**RS2**]) introduced a vast generalization of this partial order, in relation to the cell decomposition of certain symmetric varieties. In a particular case they obtained the poset $Invol(S_n)$.

In this work the problem is completely solved for an important class of Coxeter groups, namely that of classical Weyl groups. Our main result is that, if W is a classical Weyl group, then the poset $Invol(W)$ is graded, with rank function given by the average between the length and the absolute length, and that it is EL-shellable, hence Cohen-Macaulay, and Eulerian.

1991 *Mathematics Subject Classification.* Primary 05E15; Secondary 06A07.
Key words and phrases. Bruhat order, involutions, classical Weyl groups, EL-shellability.

The proofs (see [**Inc1**], [**Inc2**], [**Inc3**] for details) are combinatorial and use the descriptions of classical Weyl groups in terms of permutation groups: the symmetric group for type $\mathbf{A}_n$, the hyperoctahedral group for type $\mathbf{B}_n$ and the even-signed permutation group for type $\mathbf{D}_n$.

In particular we obtain, as new results, a combinatorial description of the covering relation in the Bruhat order of the hyperoctahedral group and the even-signed permutation group, and a combinatorial description of the absolute length of the involutions in classical Weyl groups.

Finally it is conjectured that the result proved for classical Weyl groups actually holds for every Coxeter group.

2. Notation and preliminaries

We let $\mathbf{N} = \{1, 2, 3, \ldots\}$ and $\mathbf{Z}$ be the set of integers. For $n, m \in \mathbf{Z}$, with $n \leq m$, we let $[n, m] = \{n, n+1, \ldots, m\}$. For $n \in \mathbf{N}$, we let $[n] = [1, n]$ and $[\pm n] = [-n, n]\backslash\{0\}$.

2.1. Posets. We follow [**Sta1**, Chapter 3] for poset notation and terminology. In particular we denote by $\lhd$ the *covering relation*: $x \lhd y$ means that $x < y$ and there is no z such that $x < z < y$. A poset is *bounded* if it has a minimum and a maximum, denoted by $\hat{0}$ and $\hat{1}$ respectively. If $x, y \in P$, with $x \leq y$, we let $[x, y] = \{z \in P : x \leq z \leq y\}$, and we call it an *interval* of P. If $x, y \in P$, with $x < y$, a *chain* from x to y of *length* k is a $(k+1)$-tuple $(x_0, x_1, ..., x_k)$ such that $x = x_0 < x_1 < ... < x_k = y$. A chain $x_0 < x_1 < ... < x_k$ is said to be *saturated* if all the relations in it are covering relations ($x_0 \lhd x_1 \lhd ... \lhd x_k$).

A poset is said to be *graded* of *rank* n if it is finite, bounded and if all maximal chains of P have the same length n. If P is a graded poset of rank n, then there is a unique *rank function* $\rho : P \rightarrow [0, n]$ such that $\rho(\hat{0}) = 0$, $\rho(\hat{1}) = n$ and $\rho(y) = \rho(x) + 1$ whenever y covers x in P. Conversely, if P is finite and bounded, and if such a function exists, then P is graded of rank n.

Let P be a graded poset and let Q be a totally ordered set. An *EL-labelling* of P is a function $\lambda : \{(x, y) \in P^2 : x \lhd y\} \rightarrow Q$ such that for every $x, y \in P$, with $x < y$, two properties hold:

1. there is exactly one saturated chain from x to y with non decreasing labels:
$$x = x_0 \underset{\lambda_1}{\lhd} x_1 \underset{\lambda_2}{\lhd} \ldots \underset{\lambda_k}{\lhd} x_k = y,$$
with $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k$;
2. this chain has the lexicographically minimal labelling: if
$$x = y_0 \underset{\mu_1}{\lhd} y_1 \underset{\mu_2}{\lhd} \ldots \underset{\mu_k}{\lhd} y_k = y$$
is a saturated chain from x to y different from the previous one, then
$$(\lambda_1, \lambda_2, \ldots, \lambda_k) < (\mu_1, \mu_2, \ldots, \mu_k).$$

A graded poset P is said to be *EL-shellable* if it has an EL-labelling.

Connections between EL-shellable posets and shellable complexes, Cohen-Macaulay complexes and Cohen-Macaulay rings can be found, for example, in [**Bac**], [**BGS**], [**Bjö**], [**Gar**], [**Hoc**], [**Rei**] and [**Sta2**]. Here we only recall the following important result, due to Björner.

Theorem 2.1. *Let P be a graded poset. If P is EL-shellable then P is shellable and hence Cohen-Macaulay.*

A graded poset P with rank function ρ is said to be *Eulerian* if
$$|\{z \in [x, y] : \rho(z) \text{ is even}\}| = \{z \in [x, y] : \rho(z)| \text{ is odd}\}|,$$
for every $x, y \in P$ such that $x < y$.

In an EL-shellable poset there is a necessary and sufficient condition for the poset to be Eulerian. We state it in the following form (see [**Bjö**, Theorem 2.7] and [**Sta3**, Theorem 1.2] for proofs of more general results).

Theorem 2.2. *Let P be a graded EL-shellable poset and let λ be an EL-labelling of P. Then P is Eulerian if and only if for every $x, y \in P$, with $x < y$, there is exactly one saturated chain from x to y with decreasing labels.*

2.2. Coxeter groups. About Coxeter groups we recall some basic definitions. Let W be a Coxeter group, with set of generators S. The *length* of an element $w \in W$, denoted by $l(w)$, is the minimal k such that w can be written as a product of k generators.

A *reflection* in a Coxeter group W is a conjugate of some element in S. The set of all reflections is usually denoted by T:

$$T = \{wsw^{-1} : s \in S,\ w \in W\}.$$

The *absolute length* of an element $w \in W$, denoted by $al(w)$, is the minimal k such that w can be written as a product of k reflections.

2.3. Bruhat order. Let W be a Coxeter group with set of generators S. Let $u, v \in W$. Then $u \to v$ if and only if $v = ut$, with $t \in T$, and $l(u) < l(v)$. The *Bruhat order* of W is the partial order relation so defined: given $u, v \in W$, then $u \leq v$ if and only if there is a chain

$$u = u_0 \to u_1 \to u_2 \to \ldots \to u_k = v.$$

The Bruhat order of Coxeter groups has been studied extensively (see, e.g., [**BW**], [**Deo**], [**Ede**], [**Ful**], [**Pro**], [**Rea**], [**Ver**]). In particular it is known that it gives to W the structure of a graded poset, whose rank function is the length. It has been also proved that this poset is EL-shellable, hence Cohen-Macaulay (see [**Ede**], [**Pro**], [**BW**]), and Eulerian (see [**Ver**]).* The aim of this work is to investigate whether the induced subposet $Invol(W)$ is endowed with similar properties. The problem is solved for classical Weyl groups, to which next subsection is dedicated.

2.4. Classical Weyl groups. The finite irreducible Coxeter groups have been completely classified (see, e.g., [**BB**], [**Hum**]). Among them we find the classical Weyl groups, which have nice combinatorial descriptions in terms of permutation groups: the symmetric group S_n is a representative for type $\mathbf{A}_{n-1}$, the hyperoctahedral group B_n for type $\mathbf{B}_n$ and the even-signed permutation group D_n for type $\mathbf{D}_n$.

2.4.1. *The symmetric group.* We denote by S_n the *symmetric group*, defined by

$$S_n = \{\sigma : [n] \to [n] : \sigma \text{ is a bijection}\}$$

and we call its elements *permutations*. To denote a permutation $\sigma \in S_n$ we often use the *one-line notation*: we write $\sigma = \sigma_1\sigma_2 \ldots \sigma_n$, to mean that $\sigma(i) = \sigma_i$ for every $i \in [n]$. We also write σ in *disjoint cycle form*, omitting to write the 1-cycles of σ: for example, if $\sigma = 364152$, then we also write $\sigma = (1, 3, 4)(2, 6)$. Given $\sigma, \tau \in S_n$, we let $\sigma\tau = \sigma \circ \tau$ (composition of functions) so that, for example, $(1, 2)(2, 3) = (1, 2, 3)$. Given $\sigma \in S_n$, the *diagram* of σ is a square of $n \times n$ cells, with the cell (i, j) (that is, the cell in column i and row j, with the convention that the first column is the leftmost one and the first row is the lowest one) filled with a dot if and only if $\sigma(i) = j$. For example, in Figure 1 the diagram of $\sigma = 35124 \in S_5$ is represented.

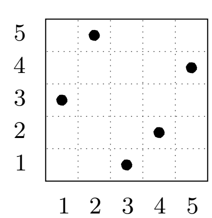


FIGURE 1. Diagram of $\sigma = 35124 \in S_5$.

The *diagonal* of the diagram is the set of cells $\{(i,i) : i \in [n]\}$.

As a set of generators for S_n, we take $S = \{s_1, s_2, \ldots, s_{n-1}\}$, where $s_i = (i, i+1)$ for every $i \in [n-1]$. It is known that the symmetric group S_n, with this set of generators, is a Coxeter group of type $\mathbf{A}_{n-1}$ (see, e.g., [**BB**]).

The length of a permutation $\sigma \in S_n$ is given by

$$l(\sigma) = inv(\sigma),$$

where

$$inv(\sigma) = |\{(i,j) \in [n]^2 : i < j,\ \sigma(i) > \sigma(j)\}|$$

is the number of *inversions* of σ.

In the symmetric group the reflections are the transpositions:

$$T = \{(i,j) \in [n]^2 : i < j\}.$$

In order to give a characterization of the covering relation in the Bruhat order of the symmetric group, we introduce the following definition.

Definition 2.1. Let $\sigma \in S_n$. A *rise* of σ is a pair $(i,j) \in [n]^2$ such that

1. $i < j$,
2. $\sigma(i) < \sigma(j)$.

A rise (i,j) is said to be *free* if there is no $k \in [n]$ such that

1. $i < k < j$,
2. $\sigma(i) < \sigma(k) < \sigma(j)$.

For example, the rises of $\sigma = 35124 \in S_5$ are $(1,2)$, $(1,5)$, $(3,4)$, $(3,5)$ and $(4,5)$. They are all free except $(3,5)$. The following is a well-known result.

Proposition 2.2. Let $\sigma, \tau \in S_n$, with $\sigma < \tau$. Then $\sigma \lhd \tau$ in S_n if and only if

$$\tau = \sigma(i,j),$$

where (i,j) is a free rise of σ.

2.4.2. *The hyperoctahedral group.* We denote by $S_{\pm n}$ the symmetric group on the set $[\pm n]$:

$$S_{\pm n} = \{\sigma : [\pm n] \to [\pm n] : \sigma \text{ is a bijection}\}$$

(which is clearly isomorphic to S_{2n}), and by B_n the *hyperoctahedral group*, defined by

$$B_n = \{\sigma \in S_{\pm n} : \sigma(-i) = -\sigma(i) \text{ for every } i \in [n]\}$$

and we call its elements *signed permutations*. To denote a signed permutation $\sigma \in B_n$ we use the *window notation*: we write $\sigma = [\sigma_1, \sigma_2, \ldots, \sigma_n]$, to mean that $\sigma(i) = \sigma_i$ for every $i \in [n]$ (the images of the negative entries are then uniquely determined). We also denote σ by the sequence $|\sigma_1|\,|\sigma_2|\,\ldots\,|\sigma_n|$, with the negative entries underlined. For example, $\underline{3}\,\underline{2}\,1$ denotes the signed permutation $[-3,-2,1]$. We also write σ in disjoint cycle form. Signed permutations are particular permutations of the set $[\pm n]$, so they inherit the notion of diagram. Note that the diagram of a signed permutation is symmetric with respect to the center. In Figure 2, the diagram of $\sigma = \underline{3}\,\underline{2}\,1 \in B_3$ is represented.

The (*main*) *diagonal* of the diagram is the set of cells $\{(i,i) : i \in [\pm n]\}$, and the *antidiagonal* is the set of cells $\{(i,-i) : i \in [\pm n]\}$.

As a set of generators for B_n, we take $S = \{s_0, s_1, \ldots, s_{n-1}\}$, where $s_0 = (1,-1)$ and $s_i = (i, i+1)(-i, -i-1)$ for every $i \in [n-1]$. It is known that the hyperoctahedral group B_n, with this set of generators, is a Coxeter group of type $\mathbf{B}_n$ (see, e.g., [**BB**]).

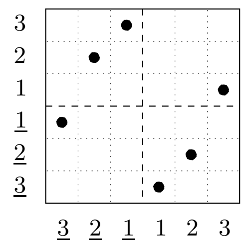


FIGURE 2. Diagram of $\sigma = \underline{3}\,\underline{2}\,1 \in B_3$.

There are various known formulas for computing the length in B_n (see, e.g., [**BB**]). In [**Inc2**] we introduced a new one: the length of $\sigma \in B_n$ is given by

$$l_B(\sigma) = \frac{inv(\sigma) + neg(\sigma)}{2}, \tag{2.1}$$

where

$$inv(\sigma) = |\{(i,j) \in [\pm n]^2 : i < j,\ \sigma(i) > \sigma(j)\}|$$

(the length of σ in the symmetric group $S_{\pm n}$), and

$$neg(\sigma) = |\{i \in [n] : \sigma(i) < 0\}|.$$

For example, for $\sigma = \underline{3}\,\underline{2}\,1 \in B_3$, we have $inv(\sigma) = 8$, $neg(\sigma) = 2$, so $l_B(\sigma) = 5$.

Finally, it is known (see, e.g., [**BB**]) that the set of reflections of B_n is

$$T = \{(i, -i) : i \in [n]\} \cup \{(i,j)(-i,-j) : 1 \leq i < |j| \leq n\}.$$

2.4.3. *The even-signed permutation group.* We denote by D_n the *even-signed permutation group*, defined by

$$D_n = \{\sigma \in B_n : neg(\sigma) \text{ is even}\}.$$

Notation and terminology are ineherited from the hyperoctahedral group. For example the signed permutation $\sigma = \underline{3}\,\underline{2}\,1$, whose diagram is represented in Figure 2, is also in D_3.

As a set of generators for D_n, we take $S = \{s_0, s_1, \ldots, s_{n-1}\}$, where $s_0 = (1,-2)(-1,2)$ and $s_i = (i, i+1)(-i, -i-1)$ for every $i \in [n-1]$. It is known that the even-signed permutation group D_n, with this set of generators, is a Coxeter group of type $\mathbf{D}_n$ (see, e.g., [**BB**]).

About the length function in D_n, it is known (see, e.g., [**BB**]) that

$$l_D(\sigma) = l_B(\sigma) - neg(\sigma).$$

Thus, by (2.1), the length of $\sigma \in D_n$ is given by

$$l_D(\sigma) = \frac{inv(\sigma) - neg(\sigma)}{2}.$$

For example, for $\sigma = \underline{3}\,\underline{2}\,1 \in D_3$, we have $l_D(\sigma) = 3$.

Finally, it is known (see, e.g., [**BB**]) that the set of reflections of D_n is

$$T = \{(i,j)(-i,-j) : 1 \leq i < |j| \leq n\}.$$

3. The main problem

It is known that a Coxeter group W, partially ordered by the Bruhat order, is a graded poset, with rank function given by the length, and that it is also EL-shellable, hence Cohen-Macaulay, and Eulerian.* The aim of this work is to investigate whether a particular subposet of W, namely that induced by the set of involutions of W, is endowed with similar properties.

3.1. Motivation. The problem arises from a geometric question. It is known that the symmetric group, partially ordered by the Bruhat order, encodes the cell decomposition of Schubert varieties (see [**Ful**]). In 1990 Richardson and Springer (see [**RS1**] and [**RS2**]) considered a vast generalization of this partial order, in relation to the cell decomposition of certain symmetric varieties. In a particular case they obtained the subposet of S_n induced by the involutions.

3.2. An example. In Figure 3 the example of the poset S_4 with the induced subposet $Invol(S_4)$ is illustrated. Even in this simple case it is not obvious why the poset $Invol(S_4)$ is graded and who the rank function is. Note that, for example, the involutions 2143 and 4231 have distance 3 in the Hasse diagram of S_4, while they are in covering relation in $Invol(S_4)$.

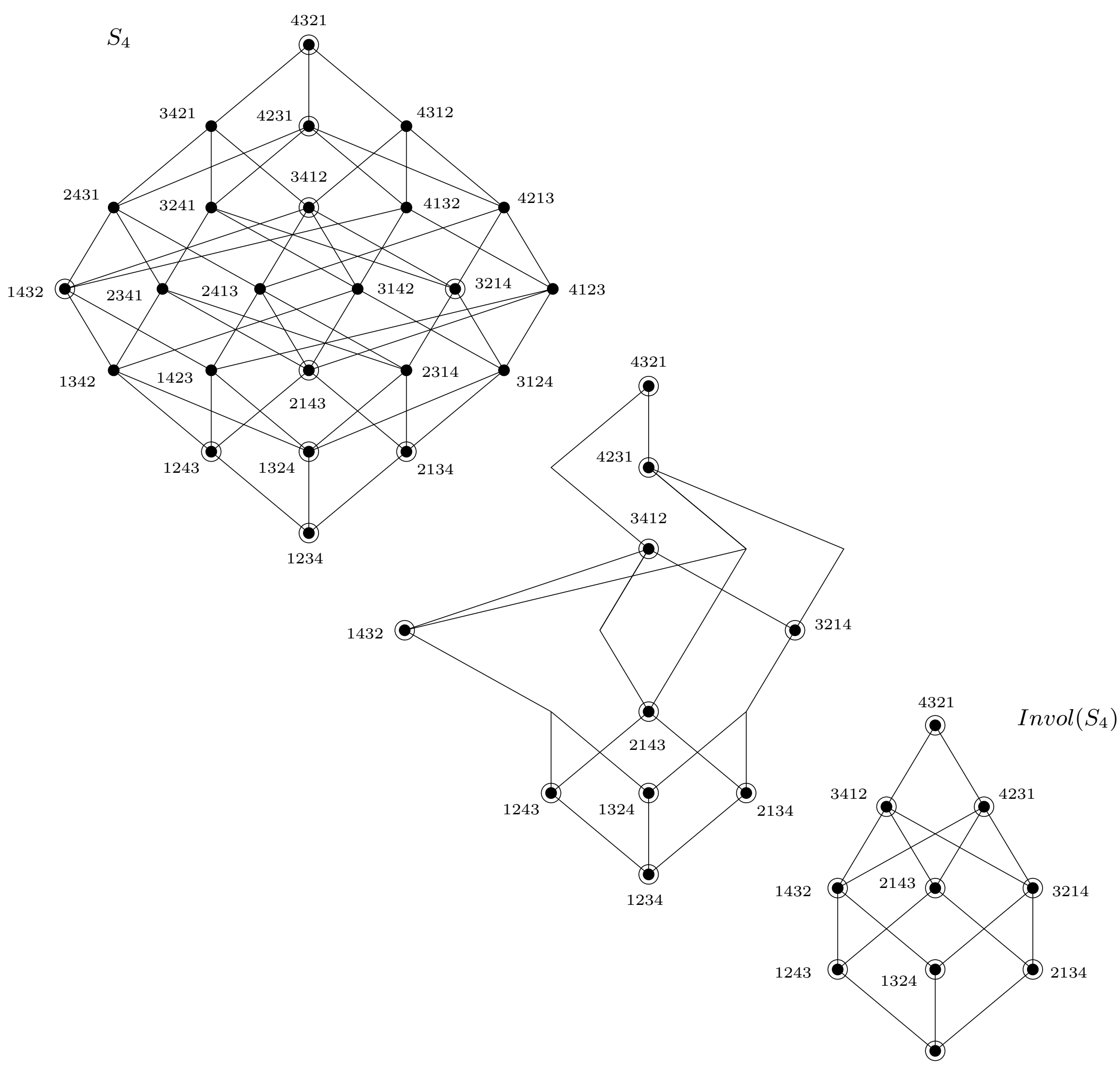


FIGURE 3. From S_4 to $Invol(S_4)$.

3.3. The main result. The following is the main result of this work.

> **Theorem 3.1.** *Let W be a classical Weyl group. The poset $Invol(W)$ is*
> 1. *graded, with rank function given by*
> $$\rho(w) = \frac{l(w) + al(w)}{2},$$
> *for every $w \in Invol(W)$;*
> 2. *EL-shellable, hence Cohen-Macaulay;*
> 3. *Eulerian.*

We will give a sketch of the proof in Section 5.

4. Preliminary results

In this section we discuss some new results, which play a crucial role in the proof of the main result of this work. Precisely, we describe the covering relation in the groups B_n and D_n, and we give a combinatorial description of the absolute length of the involutions in classical Weyl groups.

4.1. Covering relation in the Bruhat order of B_n and D_n.

Definition 4.1. Let $\sigma \in B_n$. A rise (i, j) of σ is *central* if

$$(0, 0) \in [i, j] \times [\sigma(i), \sigma(j)].$$

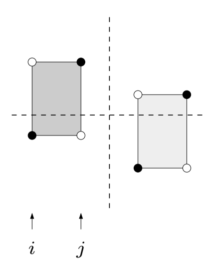
A central rise (i, j) of σ is *symmetric* if $j = -i$.

The characterization of the covering relation in B_n is then the following.

Theorem 4.1. *Let $\sigma, \tau \in B_n$. Then $\sigma \lhd \tau$ in B_n if and only if either*

1. *$\tau = \sigma(i, j)(-i, -j)$, where (i, j) is a not central free rise of σ, or*
2. *$\tau = \sigma(i, -i)$, where $(i, -i)$ is a central symmetric free rise of σ.*

Theorem 4.1 is illustrated in Figure 4, where black dots and white dots denote respectively σ and τ, inside the gray areas there are no other dots of σ and τ than those indicated, and the diagrams of the two permutations are supposed to be the same anywhere else.

1. 

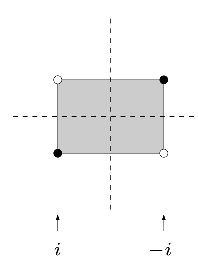
2. 

FIGURE 4. Covering relation in B_n.

For the even-signed permutation group we introduce the following definition.

Definition 4.2. Let $\sigma \in D_n$. A central rise (i, j) is *semifree* if

$$\{k \in [i, j] : \sigma(k) \in [\sigma(i), \sigma(j)]\} = \{i, -j, j\}.$$

An example of central semifree rise is illustrated in Figure 5 (3).

Theorem 4.2. *Let $\sigma, \tau \in D_n$. Then $\sigma \lhd \tau$ in D_n if and only if*

$$\tau = \sigma(i, j)(-i, -j),$$

where (i, j) is either

1. *a not central free rise of σ, or*
2. *a central not symmetric free rise of σ, or*
3. *a central semifree rise of σ.*

Theorem 4.2 is illustrated in Figure 5, with the same notation as in Figure 4.

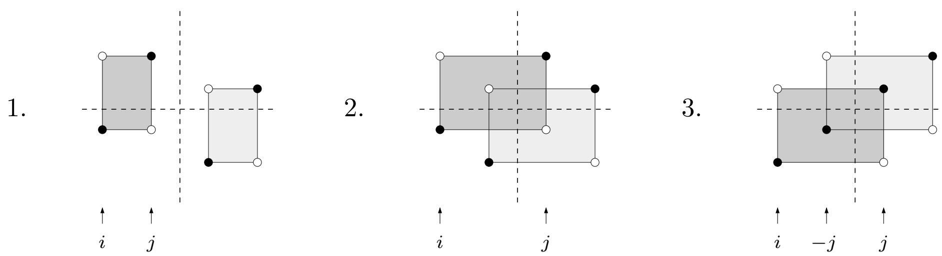


FIGURE 5. Covering relation in D_n.

4.2. Absolute length of involutions in classical Weyl groups. In classical Weyl groups there is a nice combinatorial description for the absolute length of the involutions. In the symmetric group it is simply given by the number of excedances. Note that an involution of S_n has the diagram symmetric with respect to the diagonal.

Proposition 4.3. Let $\sigma \in Invol(S_n)$. Then

$$al(\sigma) = exc(\sigma),$$

where

$$exc(\sigma) = |\{i \in [n] : \sigma(i) > i\}|$$

is the number of *excedances* of σ.

For example, for $\sigma = 32154 \in Invol(5)$, we have $al(\sigma) = exc(\sigma) = 2$. In fact

$$\sigma = \underbrace{(1, 3)}_{t_1} \cdot \underbrace{(4, 5)}_{t_2}$$

is a minimal decomposition of σ as a product of reflections of S_5.

We now define a new statistic on a signed permutation σ. Note that an involution of B_n has the diagram symmetric with respect to both the diagonals.

Definition 4.4. Let $\sigma \in B_n$. The number of *deficiencies-not-antideficiencies* of σ is

$$dna(\sigma) = |\{i \in [n] : -i \leq \sigma(i) < i\}|.$$

For example, consider $\sigma = 4\,\underline{7}\,\underline{3}\,1\,5\,\underline{6}\,\underline{2} \in B_7$, whose diagram is shown in Figure 6. Looking at the picture, $dna(\sigma)$ is the number of dots which lie in the gray area. In this case

$$dna(\sigma) = 4.$$

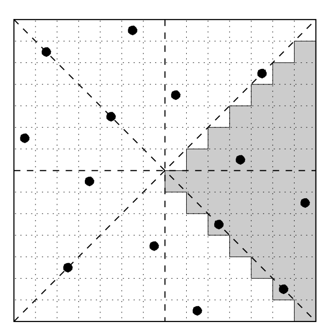


FIGURE 6. The *dna* statistic.

A surprising fact is that in the hyperoctahedral group and in the even-signed permutation group, the combinatorial description for the absolute length of an involution is exactly the same: in both cases it is given by the *dna* statistic. But the reasons are different.

Proposition 4.5. Let $\sigma \in Invol(B_n)$. Then

$$al_B(\sigma) = dna(\sigma).$$

For example, for the involution of Figure 6, we have $al_B(\sigma) = dna(\sigma) = 4$. In fact

$$\sigma = \underbrace{(1,4)(-1,-4)}_{t_1} \cdot \underbrace{(7,-2)(-7,2)}_{t_2} \cdot \underbrace{(3,-3)}_{t_3} \cdot \underbrace{(6,-6)}_{t_4} \tag{4.1}$$

is a minimal decomposition of σ as a product of reflections of B_7.

Proposition 4.6. Let $\sigma \in Invol(D_n)$. Then

$$al_D(\sigma) = dna(\sigma).$$

For example, for the involution of Figure 6, which is also in $Invol(D_7)$, we have $al_D(\sigma) = dna(\sigma) = 4$. Note that the decomposition in (4.1) does not work in D_7, since $(3,-3)$ and $(6,-6)$ are not elements of D_7. But in general an involution σ of D_n necessarily has an even number of antifixed points (that is, indices $i > 0$ such that $\sigma(i) = -i$), so we can consider them in pairs. In the example, σ has the two antifixed points 3 and 6 and

$$\sigma = \underbrace{(1,4)(-1,-4)}_{t_1} \cdot \underbrace{(7,-2)(-7,2)}_{t_2} \cdot \underbrace{(3,6)(-3,-6)}_{t_3} \cdot \underbrace{(3,-6)(-3,6)}_{t_4}$$

is a minimal decomposition of σ as a product of reflections of D_7.

5. Sketch of proofs

5.1. Gradedness. To prove that the posets $Invol(S_n)$, $Invol(B_n)$ and $Invol(D_n)$ are graded with rank function ρ we follow two steps:

1. we first give a characterization of the covering relation in the poset (this is done starting from the description of the covering relation in S_n, B_n and D_n);
2. then we prove that in every covering relation the variation of ρ is 1 (this is done using the combinatorial description of the absolute length of the involutions).

The following are the characterizations of the covering relations in the posets.

Theorem 5.1. *Let $\sigma, \tau \in Invol(S_n)$. Then $\sigma \lhd \tau$ in $Invol(S_n)$ if and only if there exists a rectangle $R = [i,j] \times [\sigma(i), \tau(i)]$ such that σ and τ have the same diagram except for the dots in R, and in its symmetric with respect to the diagonal, for which the situation, depending on the position of R with respect to the diagonal, is described in Figure 7: black dots and white dots denote respectively σ and τ, and the rectangle R (darker gray rectangle) contains no other dots of σ and τ than those indicated.*

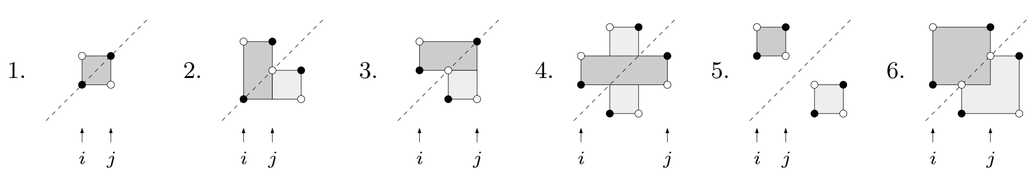


FIGURE 7. Covering relation in $Invol(S_n)$.

Looking at the diagram of a signed permutation, with *orbit* of an object (which can be a dot, a cell or a rectangle of cells), we mean the set made of that object and its symmetric with respect to the main diagonal, to the antidiagonal and to the center.

Theorem 5.2. *Let $\sigma, \tau \in Invol(B_n)$. Then $\sigma \lhd \tau$ in $Invol(B_n)$ if and only if there exists a rectangle $R = [i, j] \times [\sigma(i), \tau(i)]$ such that σ and τ have the same diagram except for the dots in R, and in the rectangles of its orbit, for which the situation, depending on the position of R with respect to the antidiagonal and to the main diagonal, is described in Figure 8, with the same notation as in Figure 7.*

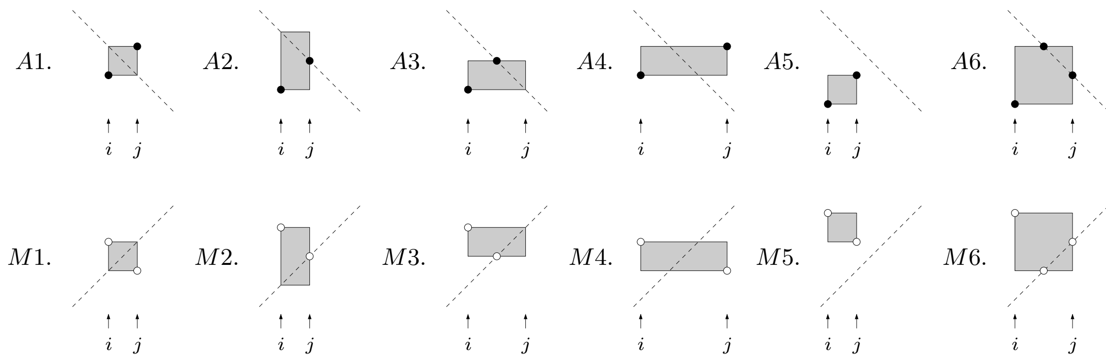
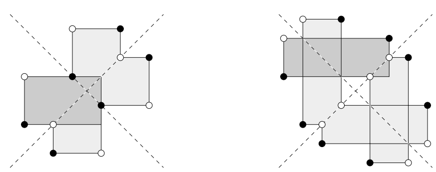


FIGURE 8. Covering relation in $Invol(B_n)$.

The *case* of (σ, τ) is (Ah, Mk), with $h, k \in [6]$, where Ah and Mk refer to the cases of Figure 8. Note that for geometrical reasons not all the 36 pairs are possible cases. In Figure 9 two examples are shown.



$(A6, M3)$ $(A4, M6)$

FIGURE 9. Two examples of covering relation in $Invol(B_n)$.

Theorem 5.3. *Let $\sigma, \tau \in Invol(D_n)$. Then $\sigma \lhd \tau$ in $Invol(D_n)$ if and only if there exists a rectangle $R = [i, j] \times [\sigma(i), \tau(i)]$, either not central or central not symmetric, such that the same conditions as in Theorem 5.2 are satisfied, with the exceptions, if R is central not symmetric, that:*

1. *in cases $(A6, M1)$ and $(A6, M3)$, picture $A6$ is replaced by picture $A6'$, and in cases $(A1, M6)$ and $(A3, M6)$, picture $M6$ is replaced by picture $M6'$, as shown in Figure 10;*

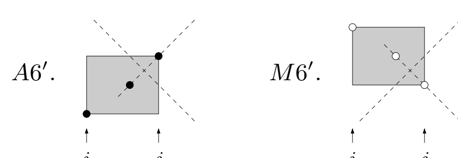
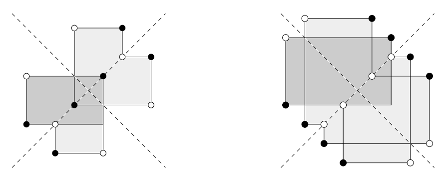


FIGURE 10. Covering relation in $Invol(D_n)$: new cases.

2. *in the remaining cases,* $(A3, M4)$, $(A4, M3)$, $(A4, M4)$, $(A4, M6)$, $(A6, M4)$, *the presence in* R *of one more dot either of* σ *or of* τ*, which is in the orbit of one of those indicated in the pictures, is allowed.*

In Figure 11 two examples are shown.



$(A6', M3)$ $(A4, M6)$

FIGURE 11. Two examples of covering relation in $Invol(D_n)$.

In the following the gradedness of the posets is stated.

Theorem 5.4. *The poset* $Invol(S_n)$ *is graded, with rank function given by*

$$\rho(\sigma) = \frac{inv(\sigma) + exc(\sigma)}{2},$$

for every $\sigma \in Invol(S_n)$*. In particular* $Invol(S_n)$ *has rank*

$$\rho(Invol(S_n)) = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Theorem 5.5. *The poset* $Invol(B_n)$ *is graded, with rank function given by*

$$\rho(\sigma) = \frac{inv(\sigma) + neg(\sigma) + 2dna(\sigma)}{4},$$

for every $\sigma \in Invol(B_n)$*. In particular* $Invol(B_n)$ *has rank*

$$\rho(Invol(B_n)) = \frac{n^2 + n}{2}.$$

Theorem 5.6. *The poset* $Invol(D_n)$ *is graded, with rank function given by*

$$\rho(\sigma) = \frac{inv(\sigma) - neg(\sigma) + 2dna(\sigma)}{4},$$

for every $\sigma \in Invol(D_n)$*. In particular* $Invol(D_n)$ *has rank*

$$\rho(Invol(D_n)) = \left\lfloor \frac{n^2}{2} \right\rfloor.$$

5.2. EL-shellability and Eulerianity. Let P be one of $Invol(S_n)$, $Invol(B_n)$ or $Invol(D_n)$.

The characterization of the covering relation gives rise in a natural way to the definition of a "standard labelling" of P. In fact, for every $\sigma, \tau \in P$, with $\sigma \lhd \tau$, we call *main rectangle* of the pair (σ, τ) the rectangle $R = [i,j] \times [\sigma(i), \tau(i)]$, mentioned in each of the Theorems 5.1, 5.2 and 5.3. Note that this rectangle necessarily is unique. Then we can give the following definition.

Definition 5.7. The *standard labelling* of P is the function

$$\lambda : \{(\sigma, \tau) \in P^2 : \sigma \lhd \tau\} \to \{(i,j) \in I^2 : i < j\}$$

(where $I = [n]$ if $P = Invol(S_n)$, and $I = [\pm n]$ otherwise) so defined: for every $\sigma, \tau \in P$, with $\sigma \lhd \tau$, if $R = [i,j] \times [\sigma(i), \tau(i)]$ is the main rectangle of (σ, τ), then we set

$$\lambda(\sigma, \tau) = (i, j).$$

To prove that the poset P is EL-shellable, we show that the standard labelling actually is an EL-labelling. This is proved first describing the lexicographically minimal saturated chains, and then showing that those are the unique with the property of having non decreasing labels.

Theorem 5.8. *The poset P is EL-shellable, hence Cohen-Macaulay.*

To prove that the poset P is Eulerian, we show that the standard labelling satisfies the condition of Theorem 2.2, that is, for every $\sigma, \tau \in P$, with $\sigma < \tau$, there is a unique saturated chain from σ to τ with decreasing labels. This is proved starting from the EL-shellability and considering the lexicographically minimal descending chains.

Theorem 5.9. *The poset P is Eulerian.*

6. Conjecture

It is natural to conjecture that our main result actually holds for every Coxeter group.

Conjecture 7.1. *Let W be a Coxeter group. The poset $Invol(W)$ is*

1. *graded, with rank function given by*

$$\rho(w) = \frac{l(w) + al(w)}{2},$$

for every $w \in Invol(W)$;

2. *EL-shellable, hence Cohen-Macaulay;*
3. *Eulerian.**

After a preliminary investigation on the affine Weyl groups (which also have nice combinatorial descriptions), we feel that our techniques may be applied also to this class of Coxeter groups. There is another class of Coxeter groups, which are not Weyl groups, for which the result is valid: the class of dihedral groups.

* In the infinite cases, we mean that every interval $[\hat{0}, x]$ of the poset has the mentioned properties.

References

[Bac] K. Baclawski, Cohen-Macaulay ordered sets, *J. Algebra* **63** (1980), 226-258

[BB] A. Björner, F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, Springer-Verlag, New York/Berlin, to appear

[BGS] A. Björner, A.M. Garsia, R.P. Stanley, *An introduction to Cohen-Macaulay partially ordered sets*, Ordered Sets (Banff., Alta., 1981), pp. 583-615

[Bjö] A. Björner, Shellable, Cohen-Macaulay partially ordered sets, *Trans. Amer. Math. Soc.* **260** (1980), 159-183

[BW] A. Björner, M. Wachs, Bruhat order of Coxeter groups and shellability, *Advances in Math.* **43** (1982), 87-100

[Deo] V.V. Deodhar, Some characterizations of Bruhat ordering on a Coxeter group and determination of the relative Möbius function, *Invent. Math.* **39** (1977), no. 2, 187-198

[Ede] P.H. Edelman, The Bruhat order of the symmetric group is lexicographically shellable, *Proc. Amer. Math. Soc.* **82** (1981), 355-358

[Ful] W. Fulton, *Young tableaux. With applications to representation theory and geometry*, London Mathematical Society Student Texts, 35. Cambridge University Press, Cambridge, 1997

[Gar] A.M. Garsia, Combinatorial methods in the thoery of Cohen-Macaulay rings, *Advances in Math.* **38** (1980), no. 3, 229-266

[Hoc] M. Hochster, *Cohen-Macaulay rings, combinatorics, and simplicial complexes*, Ring Theory II (Proc. Second Conf. Univ. Oklahoma, 1975), pp. 171-223

[Hum] J.E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Univ. Press, Cambridge, 1990

[Inc1] F. Incitti, The Bruhat order on the involutions of the symmetric group, *J. Algebraic Combin.*, to appear

[Inc2] F. Incitti, The Bruhat order on the involutions of the hyperoctahedral group, *European J. Combin.* **24** (2003), 825-848

[Inc3] F. Incitti, Bruhat order on the involutions of classical Weyl groups, Ph.D. Thesis, Dipartimento di Matematica, Università di Roma "La Sapienza", 2003

[Pro] R.A. Proctor, Classical Bruhat orders and lexicographic shellability, *J. Algebra* **77** (1982), 104-126

[Rea] N. Reading, Order dimension, strong Bruhat order and lattice properties for posets, *Order* **19** (2002), 73-100

[Rei] G.A. Reisner, Cohen-Macaulay quotients of polynomial rings, *Advances in Math.* **21** (1976), 30-49

[RS1] R.W. Richardson, T.A. Springer, The Bruhat order on symmetric varieties, *Geom. Dedicata* **35** (1990), no. 1-3, 389-436

[RS2] R.W. Richardson, T.A. Springer, Complements to: The Bruhat order on symmetric varieties, *Geom. Dedicata* **49** (1994), no. 2, 231-238

[Sta1] R.P. Stanley, *Enumerative Combinatorics*, Vol. 1, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1986

[Sta2] R.P. Stanley, *Cohen-Macaulay complexes*, Higher Combinatorics (Proc. NATO Advanced Study Inst., Berlin, 1976), pp. 51-62

[Sta3] R.P. Stanley, Supersolvable lattices, *Algebra Universalis* **2** (1972), 197-217

[Ver] D.N. Verma, Möbius inversion for the Bruhat ordering on a Weyl group, *Ann. Sci. École Norm. Sup.* **4** (1971), 393-398

Dipartimento di Matematica, Università "La Sapienza", 00185 Roma (Italy)
E-mail address: incitti@mat.uniroma1.it

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Alternating Sign Matrices With One -1 Under Vertical Reflection

Pierre Lalonde

Abstract. *We define a bijection that transforms an alternating sign matrix A with one -1 into a pair (N, E) where N is a (so called)* neutral *alternating sign matrix (with one -1) and E is an integer. The bijection preserves the classical parameters of Mills, Robbins and Rumsey as well as three new parameters (including E). It translates vertical reflection of A into vertical reflection of N. A hidden symmetry allows the interchange of E with one of the remaining two new parameters. A second bijection transforms (N, E) into a configuration of lattice paths called "mixed configuration".*

RÉSUMÉ. *On définit une bijection qui transforme une matrice à signes alternants A ayant un seul -1 en une paire (N, E) constituée d'une matrice à signes alternants dite* neutre *N (elle aussi à un seul -1) et d'un paramètre entier E. La bijection préserve les paramètres classiques de Mills, Robbins et Rumsey ainsi que trois nouveaux paramètres (dont E). Elle transforme la réflexion verticale de A en la réflexion verticale de N. Une symétrie cachée permet l'échange de E avec un des deux autres nouveaux paramètres. Une seconde bijection transforme (N, E) en une configuration de chemins dite "configuration mixte".*

1. Introduction

Recall that a square matrix $A = (a_{ij})_{1 \le i,j \le n}$ is an order n *alternating sign matrix* if $a_{ij} \in \{1, 0, -1\}$ and if, in each row and each column, the non-zero entries alternate in sign, beginning and ending with a 1. Thus, the entries of each row and of each column add up to 1.

The entries in the first row of an alternating sign matrix are all 0 except for one, which must be a 1. It will be called the *first* 1.

In their paper [**MRR**], Mills, Robbins and Rumsey defined the following parameters on order n alternating sign matrices $A = (a_{ij})$:

- $r(A)$ is the number of entries to the left of the first 1. We have $0 \le r(A) \le n - 1$.
- $s(A)$ is the number of entries that are equal to -1.
- $i(A) = \sum_{k>i,\ell<j} a_{ij} a_{k\ell} = \sum_{i,j} a_{ij} \left(\sum_{k>i,\ell<j} a_{k\ell} \right)$ is the *number of inversions* of A. If A is a permutation matrix, $i(A)$ reduces to the usual number of inversions.

We will use the following notation: $\mathcal{A}_n$ denotes the set of order n alternating sign matrices and $\mathcal{A}_{n,s}$ the set of order n alternating sign matrices A with $s(A) = s$.

1991 *Mathematics Subject Classification.* Primary 05A15.
Key words and phrases. alternating sign matrices, lattice paths.
Supported by a grant from NSERC..

One of the Mills, Robbins and Rumsey conjectures asserts that $|\mathcal{A}_n|$ is also the number of order n descending plane partitions. In this form, the conjecture was solved by Zeilberger (see [**Ze1**], [**Ze2**]) with subsequent simplifications by Kuperberg (see [**Ku**]). Bressoud (see [**Br**]) gives an historical and mathematical account of the whole subject.

Stronger forms of the conjectures involve the parameters (defined above), which should translate into known combinatorially significant parameters on descending plane partitions. In that direction, only special cases of the conjectures are solved. This is well known, of course, for $\mathcal{A}_{n,0}$ (permutation matrices). The conjectures are also true for $\mathcal{A}_{n,1}$ (see [**La1**]). This was done by encoding descending plane partitions into configurations of non-intersecting paths (so called TB-configurations), which allows enumeration by a determinant. After application of an algebraic transformation, the determinant is reinterpreted as enumerating another kind of lattice paths (mixed configurations), the set of which follows the same recurrences that describe $\mathcal{A}_{n,1}$. Mixed configurations are seen to generalize inversion tables of permutations.

In the present paper, we will give a bijective version of the last step, transforming $A \in \mathcal{A}_{n,1}$ into a pair (N, E), where $N \in \mathcal{A}_{n,1}$ is "neutral" (to be defined in the next section) and E is an integer (that can thus be seen as a measure of the "difference" between A and N). A second bijection will transform the pair (N, E) into a mixed configuration Ω. The bijections translate the already defined parameters (as well as three new ones) in a way that is coherent with the Mills, Robbins and Rumsey conjectures. Morover, the bijective link $A \leftrightarrow \Omega$ generalizes the encoding of permutations into inversion tables.

Let $A = (a_{ij})_{1 \le i,j \le n} \in \mathcal{A}_n$. We write $\overline{A} = (a_{i,n+1-j})_{1 \le i,j \le n}$ to denote the matrix obtained from A by vertical reflection. The classical parameters r, i and s applied to A and to $\overline{A}$ are easily related (see [**MRR**]):

- $r(A) + r(\overline{A}) = n - 1$,
- $i(A) + i(\overline{A}) = \binom{n}{2} + s(A)$,
- $s(\overline{A}) = s(A)$.

Vertical reflection can be included in the conjectures. It is then believed to correspond to an operation that can be interpreted as a kind of "complementation" operation on descending plane partitions. In [**La2**], it is shown that this operation takes a simple form in terms of Gessel-Viennot path duality (see [**GV**]) on TB-configurations. (Krattenthaler (see [**Kr**]) has an even simpler interpretation in terms of rhombus tilings.) Our bijections behave similarly: if $A \in \mathcal{A}_{n,1}$ is sent to (N, E) and then to the mixed configuration Ω, then $\overline{A}$ is sent to $(\overline{N}, -E)$, which is sent to $\overline{\Omega}$, the Gessel-Viennot dual of Ω.

This is of course a first step toward an eventual general bijection between unrestricted alternating sign matrices and mixed configurations. The results suggest that the three new parameters will play an important role in the general bijection. The natural guess is that each -1 of an alternating sign matrix will be associated to three similar parameters collectively encoded into a second neutral alternating sign matrix. Moreover, the bijections introduced here are, in some sense, the simplest possible and thus should appear in some form in the general bijection.

A few words on the organization of the paper: In section 2, we will define the sign (positive, neutral, negative) of an alternating sign matrix with one -1, define the new parameters and describe the various parts of the matrix that intervene in the bijections. The first bijection (to neutral matrices) is introduced in section 3. In section 4, the equivalence of two of the new parameters is shown. Section 5 describe the second bijection (to mixed configurations). Finally, in section 6 we show the translation of vertical reflection (on matrices) into path duality (on mixed configurations).

FIGURE 1. Schematic view of a positive matrix (left) and a neutral one (right), with some of the related regions as defined in this section. Only the significant non-zero entries are depicted. The 0 region contains only 0's.

2. Three new parameters

In what follows, we will introduce the three new parameters defined for a matrix $A \in \mathcal{A}_{n,1}$. These parameters are related to various sub-matrices of A, which we describe below (see also figure 1).

- The *opening column* of A is the column of its (unique) -1. The highest 1 in this column is the *opening* 1 and the corresponding row, the *opening row*. The *closing row* is the row of the -1. The opening column divides A into a *left side* and a *right side* (both excluding the opening column).
- The closing row is the only row that contains two 1, one in each side. These 1 will be referred to as the *left* 1 and the *right* 1.
- If any, the rows between the opening and the closing rows are the *enclosed rows*. If there are no enclosed rows, A is said to be *neutral*; otherwise A is *charged*. In the latter case, define the *charged side* to be the side (left or right) where we find the 1 of the lowest enclosed row, the other side being the *neutral side*. If the charged side is the right side (respectively: left side), we say that A is *positive* (respectively: *negative*).

In fact, we can define more generally $A = (a_{ij}) \in \mathcal{A}_n$ to be *neutral* if $a_{ij} = 1$ when $a_{i+1\,j} = -1$.

Let $\mathcal{A}^+_{n,1}$ (respectively: $\mathcal{A}^0_{n,1}$, $\mathcal{A}^-_{n,1}$) be the set of positive (respectively: neutral, negative) matrices $A \in \mathcal{A}_{n,1}$. These sets are mutually disjoint and form a partition of $\mathcal{A}_{n,1}$. Moreover, $\mathcal{A}^+_{n,1}$ and $\mathcal{A}^-_{n,1}$ are mirror-images of one another: $A \in \mathcal{A}^+_{n,1}$ iff $\overline{A} \in \mathcal{A}^-_{n,1}$.

We further define the following for $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$:

- The intersections of the enclosed rows with the right (respectively: left) side define the *charged* (respectively: *neutral*) *cell*. The *extended* neutral cell includes the intersection of the opening and of the closing rows with the left side. If $A \in \mathcal{A}^0_{n,1}$, the charged and the neutral cells are empty.
- The highest 1 in the left side below the opening row is the *leading* 1. Its column is the *leading column*. The sub-matrix between the leading and the opening column and below the opening row is the *leading cell*. The sum of the entries of the leading cell is denoted $\ell(A)$.
- Finally, the right 1 (in the closing row) is also called the *closing* 1. Its column is the *closing column*. The sub-matrix of A between the closing and the opening column and below the closing row is the

FIGURE 2. In each of the above matrices, the leading, charged and closing cells are emphasized. Matrix A_0 is positive, with $E(A_0) = 3$, $B(A_0) = -1$ and $J(A_0) = 7$. Matrix N_0 is neutral, with $E(N_0) = 0$, $B(N_0) = 2$ and $J(N_0) = 7$. The classical parameters are: $r(A_0) = r(N_0) = 6$ and $i(A_0) = i(N_0) = 30$.

closing cell. The *extended* closing cell includes the parts of opening and of the closing columns that are below the closing row. The sum of the entries of the closing cell is denoted $c(A)$.

Remark 2.1. It should be observed that $\ell(\overline{A}) = c(A)$ and $c(\overline{A}) = \ell(A)$ when $A \in \mathcal{A}^0_{n,1}$.

We can now define the new parameters (see figure 2):

- If $A \in \mathcal{A}^+_{n,1}$, its *electric charge*, $E(A)$, is the sum of the entries of the charged cell of A. In that case, $E(A) > 0$. Define $E(A) = 0$ if $A \in \mathcal{A}^0_{n,1}$ and $E(A) = -E(\overline{A})$ if $A \in \mathcal{A}^-_{n,1}$. Thus A is positive, neutral or negative according to the sign of $E(A)$.
- If $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$, define its *magnetic charge* by $B(A) = c(A) - \ell(A)$. If $A \in \mathcal{A}^0_{n,1}$, we clearly have $B(\overline{A}) = -B(A)$. Extend this property to define $B(A)$ for $A \in \mathcal{A}^-_{n,1}$.
- If $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$, define $J(A) = c(A) + \ell(A) + |E(A)| + 1$. Notice that $J(A) = J(\overline{A})$ if $A \in \mathcal{A}^0_{n,1}$. Extend this property to define $J(A)$ for $A \in \mathcal{A}^-_{n,1}$.

Clearly, with respect to vertical reflection, E and B are anti-invariants, while J is invariant. Algebraically:

$$E(M) + E(\overline{M}) = 0, \qquad B(M) + B(\overline{M}) = 0 \qquad \text{and} \qquad J(\overline{M}) = J(M)\,.$$

3. Neutralizing alternating sign matrices

Our first task will be to learn how to "neutralize" a given matrix $A \in \mathcal{A}^+_{n,1}$. This requires many steps based on the *horizontal/vertical displacement* procedure, which applies to some of the entries of a $(0,1)$-matrix.

Horizontal displacement (H): Let $P = (p_{ij})_{1\le i\le m, 1\le j\le n}$ be a $(0,1)$-matrix. Suppose that the non-zero columns occupy positions $j_1 < j_2 < \cdots < j_k$ with $j_1 = 1$ and $j_k < n$.

Its horizontal displacement, $H(P)$, is the matrix obtained from P by displacing column j_i to column j_{i+1} (for $1 \le i \le k$), where $j_{k+1} = n$. Column j_1 is replaced by a column of 0's. Clearly, $H(P)$ is a $(0,1)$-matrix of the same dimension as P, with non-zero columns in positions $j_2 < \cdots j_k < j_{k+1} = n$. The procedure is obviously injective.

FIGURE 3. The rectangles in these matrices enclose the extended neutral cell, the charged cell and the extended closing cell. The first matrix result from A_0 (figure 2) after applying the first three steps of δ. The last matrix is $P_0 = \delta(A_0)$.

We define similarly the *vertical displacement* $V(P)$ for $(0,1)$-matrices P such that the first row is 0 and the last, non-zero. (The rows are displaced from bottom to top.)

We will apply the horizontal/vertical displacement to some of the cells of a given matrix $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$ (or to some modifications of A). This will give the *discharging procedure* which essentially transforms A into a permutation matrix P of the same dimension. The opening column, closing cell,... of any transformation of A refer to sub-matrices of the transformed matrix that occupies the same position as in A.

Definition 3.1. (Partial discharging procedure δ) Let $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$. We obtain $\delta(A)$ from A by:

(1) Erasing the -1 and the closing 1 of A.
(2) Applying H to the extended closing cell of A.
(3) Applying V to the extended neutral cell. (See figure 3 (left).)
(4) Lower the 1's in the extended neutral and in the charged cells by one row (erasing or writing 0's when necessary). (See figure 3 (right).)

The resulting matrix is $\delta(A)$.

Remark 3.2. Carefully keeping track of the number of 1's in each row and each column after each step, we see that the resulting matrix $\delta(A)$ is always a permutation matrix. It is clear that $r(\delta(A)) = r(A)$ since the procedure never affects (permanently) the first row. A fine analysis will show that $\ell(\delta(A)) = \ell(A)$ and that $i(\delta(A)) = i(A) - 1 - (c(A) + E(A))$.

Definition 3.3. (Complete discharging procedure Δ) Let $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$. We define

$$\Delta(A) = (k, \delta(A), c(A), E(A))\,,$$

where k is the position of the opening row of A ($1 \le k \le n$).

Example 3.4. For instance, referring to figure 2, we have $\Delta(A_0) = (3, P_0, 1, 3)$ and $\Delta(N_0) = (3, P_0, 4, 0)$ where P_0 is the last matrix of figure 3.

Notice that we can recover A from its image $\Delta(A) = (k, \delta(A), c(A), E(A))$. In fact, we can apply δ backward to the permutation $P = \delta(A)$ provided that:

(1) we can identify the opening column and the opening 1. (The latter being unaffected by δ.) This is determined by k, the position of the opening row.

(2) we determine the closing row. This is given by $E(A)$: the closing row is the highest row below the opening row such that the elements between (and including) these rows in the right side sum up to $E(A)$.
(3) we determine the closing column. But it is the leftmost column to the right of the opening column such that the elements between (and including) these columns and below (strictly) the closing row sum up to $c(A) + 1$.

It is always possible to do so since, by construction, $\delta(A)$ contains at least $E(A)+c(A)+1$ non-zero elements below the opening row and to the right of the opening column.

This shows that δ is injective. More generally, let $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$ with $\Delta(A) = (k, \delta(A), c(A), E(A))$ then, for any $c, E \geq 0$ such that $c + E \leq c(A) + E(A)$, there is a unique $B \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$ such that $\Delta(B) = (k, \delta(A), c, E)$ (we will write $B = \Delta^{-1}(k, \delta(A), c, E)$). In particular, we can take $N = \Delta^{-1}(k, \delta(A), c(A) + E(A), 0)$, which will be neutral by construction. In that case, N is quite easy to find from $\delta(A)$ by applying δ backward: steps 4 and 3 cancel each other.

Definition 3.5. (Neutralizing procedure Λ) Let $A \in \mathcal{A}_{n,1}$.
(1) If $A \in \mathcal{A}^+_{n,1} \cup \mathcal{A}^0_{n,1}$, let $(k, P, c, E) = \Delta(A)$, define $\Lambda(A) = (\Delta^{-1}(k, P, c + E, 0), E)$.
(2) If $A \in \mathcal{A}^-_{n,1}$, notice that $\overline{A} \in \mathcal{A}^+_{n,1}$. Writing $(\overline{N}, -E) = \Lambda(\overline{A})$, define $\Lambda(A) = (N, E)$.

Example 3.6. Referring to figures 2 (and 3), we have: $\Lambda(A_0) = (N_0, 3)$.

Theorem 3.7. *The neutralizing procedure is a bijection*

$$\Lambda : \mathcal{A}_{n,1} \longrightarrow \mathcal{N}_{n,1} := \{(N, E) \mid N \in \mathcal{A}^0_{n,1}, E \in \mathbb{Z}, -\ell(N) \leq E \leq c(N)\}.$$

Moreover, let $A \in \mathcal{A}_{n,1}$ and $\Lambda(A) = (N, E)$, then:
(1) $A \in \mathcal{A}^0_{n,1}$ *iff* $N = A$.
(2) $\Lambda(\overline{A}) = (\overline{N}, -E)$.
(3) *The matrices A and N are the same, from the first row to the opening row (included).*
(4) *The following relations hold:*
   (a) $r(N) = r(A)$,
   (b) $i(N) = i(A)$,
   (c) $E = E(A)$,
   (d) $B(N) = B(A) + E(A)$,
   (e) $J(N) = J(A)$.

Indeed, if A is positive (or neutral), we have $r(N) = r(A)$, $\ell(N) = \ell(A)$ and $i(N) = i(A)$, by remark 3.2. By construction, $c(N) = c(A) + E(A)$; thus $B(N) = B(A) + E(A)$ and $J(N) = J(A)$. In particular, $0 \leq E = E(A) \leq c(N)$. If A is negative, use remark 2.1 and the formulae of section 1.

4. Exchanging the electric charge and the magnetic charge

Using the neutralizing procedure, we define an involution on $\mathcal{A}_{n,1}$ that exchanges E and B. Thus the two charges play the same role and are completely interchangeable.

Lemma 4.1. *Let $A \in \mathcal{A}_{n,1}$ and $(N, E) = \Lambda(A)$. Then $-\ell(N) \leq B(A) \leq c(N)$.*

Theorem 4.2. *Let $\Xi = \Lambda^{-1} \circ \xi \circ \Lambda$ where ξ is defined by $\xi(N, E) = (N, c(N) - \ell(N) - E)$. Then ξ is an involution on $\mathcal{N}_{n,1}$ and Ξ an involution on $\mathcal{A}_{n,1}$. Moreover, if $A \in \mathcal{A}_{n,1}$ and $\Xi(A) = A'$, we have:*
(1) $\Xi(\overline{A}) = \overline{A'}$.
(2) *The matrices A and A' are the same, from the first row to the opening row (included).*

(3) *The involution* Ξ *exchanges the charges; namely:* $E(A') = B(A)$ *and* $B(A') = E(A)$.
(4) *All other defined parameters* (r, i *and* J) *take the same value on* A *as on* A'.

Of course, this leads to another bijection, $\xi \circ \Lambda : \mathcal{A}_{n,1} \longrightarrow \mathcal{N}_{n,1}$, which focuses on the parameter B instead of E. In fact $\xi \circ \Lambda(A) = (N, B(A))$.

5. Encoding elements of $\mathcal{N}_{n,1}$ into mixed configurations

It is well known that a permutation matrix $P = (p_{ij}) \in \mathcal{A}_{n,0}$ can be bijectively encoded by a sequence $(a_i)_{i=1}^n$ of non-negative integers called its inversion table. In fact, a_i is the sum of the entries of P that are below row $n+1-i$ and to the left of the (unique) 1 in that row. With this convention, we have $0 \le a_i < i$ for $1 \le i \le n$. The classical parameters are easily recovered: $r(P) = a_n$ and $i(P) = a_1 + \cdots + a_n$. Moreover, if $(\overline{a}_i)_{i=1}^n$ is the inversion table of $\overline{P}$, then $\overline{a}_i = i - 1 - a_i$ (for $1 \le i \le n$). We define a generalization of inversion table that applies to $\mathcal{A}_{n,1}$.

Definition 5.1. Let $(N, E) \in \mathcal{N}_{n,1}$. Let $n+1-k$ be the position of the opening row of N (thus the position of the closing row is $n+2-k$). For $1 \le i \le n$, define a_i as the sum of the entries of N that are below row $n+1-i$ and to the left of the unique 1 (or the leftmost 1 if $i = k-1$) in row $n+1-i$. Let $b = c(N)$ and $\beta = E + \ell(N)$. The sequence of integers $(k; a_1, \ldots, a_n; b, \beta)$ is called the *generalized inversion table* of (N, E).

Remark 5.2. Clearly, $\ell(N) = a_k - 1 - a_{k-1}$, an observation that we will often use later.

Example 5.3. For instance, the generalized inversion table of $(N_0, 3)$ (from figure 2) is:

$$(10;\ 0,0,2,2,0,0,1,5,0,3,6,6;\ 4,5)\,.$$

Inversion tables are encoded as sequences of non-intersecting lattice paths called *mixed configuration* (introduced in [**La1**]).

We consider lattice-paths on the strict half-grid $\mathcal{G}_n = \{(k,\ell) \,|\, 0 \le k < \ell \le n\}$. (For more symmetry in the figures, the grid will be slightly shifted so that its boundary forms a reversed equilateral triangle.) *Mixed paths* on $\mathcal{G}_n$ are composed of two consecutive parts: Left and Right, where:

- the Left part is composed of South steps (S) and East steps (E).
- the Right part is composed of (another kind of) East steps (F) and North-East steps (N).

Notice that each path contains a vertex that belongs both to the Left part and to the Right part. Such vertices are called *junctions*.

An order n *mixed configuration* is a sequence of mixed paths $\Omega = (\omega_1, \ldots, \omega_n)$ on $\mathcal{G}_n$ such that:

- There is a permutation σ such that ω_i starts from $(0, i)$ and ends at $(\sigma(i)-1, \sigma(i))$.
- The sub-configuration obtained by deleting the Right part of paths is non-intersecting (no common vertex).
- The sub-configuration obtained by deleting the Left part of paths is non-intersecting.

Let $\mathcal{M}_{n,s}$ be the set of order n mixed configurations with s N-steps.

Given a generalized inversion table $(k; a_1, \ldots, a_n; b, \beta)$, we define the corresponding mixed configuration $\Omega = (\omega_1, \ldots, \omega_n) \in \mathcal{M}_{n,1}$. In that case, Ω has two consecutive special paths ω_{k-1} (which contains the N-step) and ω_k (which contains a S-step. The paths, written as sequences of steps, are:

- $\omega_i = E^{a_i} F^{i-1-a_i}$, for i such that $1 \le i \le n$ and $i \ne k-1, k$. This path joins $(0,i)$ to $(i-1,i)$.
- $\omega_{k-1} = E^{a_{k-1}} F^{\beta} N F^{k-2-a_{k-1}-\beta}$. This path joins $(0,k-1)$ to $(k-1,k)$.

FIGURE 5. The combinatorial interpretations of the parameters E and J on mixed configurations.

- $\omega_k = E^{a_k} S E^b F^{k-2-a_k-b}$. This path joins $(0,k)$ to $(k-2,k-1)$.

(see figure 4 (right)). It is easy to check that this defines a mixed configuration.

Theorem 5.4. *The encoding of the generalized inversion table of an element $(N,E) \in \mathcal{N}_{n,1}$ into a mixed configuration $\Omega \in \mathcal{M}_{n,1}$ defines a bijection $\Phi : \mathcal{N}_{n,1} \longrightarrow \mathcal{M}_{n,1}$.*

Moreover, let $A \in \mathcal{A}_{n,1}$, $(N,E) = \Lambda(A)$, $(k; a_1, \ldots, a_n; b, \beta)$ its generalized inversion table and $\Omega = (\omega_1, \ldots, \omega_n) = \Phi(N,E)$. Then

1. *$r(A) = r(N)$ is the number of E-steps of Ω that are at level n (all occurring in path ω_n).*
2. *$i(A) = i(N)$ is the total number of E-steps and of N-steps of Ω.*
3. *$E(A) = E = a_{k-1} + \beta + 1 - a_k$ is the signed distance from the beginning of the S-step to the end of the N-step of Ω (see figure 5).*
4. *$B(A) = B(N) - E = b - \beta$.*
5. *$J(A) = J(N) = a_k - a_{k-1} + b$ is the (non-signed) distance between the junctions of path ω_{k-1} and of path ω_k.*

Corollary 5.5. *Let $A \in \mathcal{A}_{n,1}$, $(k; a_1, \ldots, a_n; b, \beta)$ the generalized inversion table of $\Lambda(A)$ and $\Omega = (\omega_1, \ldots, \omega_n) = \Phi(\Lambda(A))$. Let $A' = \Xi(A)$ then $(k; a_1, \ldots, a_n; b, \beta')$ is the generalized inversion table of $\Lambda(A')$ (where $\beta' = b - \beta + a_k - a_{k-1} - 1$).*

Thus $\Omega' = (\omega'_1, \ldots, \omega'_n) = \Phi(\Lambda(A'))$ is obtained from Ω by replacing β by β' (this only changes the Right part of ω_{k-1}).

FIGURE 7. Duality on $\mathcal{M}_{n,1}$ (before the final reflection) and its relation with vertical reflection on matrices.

6. Duality and Mixed Configurations

First, we examine path duality for mixed configurations $\Omega = (\omega_1, \ldots, \omega_n) \in \mathcal{M}_{n,0}$. We saw how to extract the inversion table $(a_i)_{i=1}^n$, (which then corresponds to a unique permutation matrix P). The dual $\overline{\Omega}$ of Ω is obtained by "complementing" the Left and the Right parts (separately) of each path, leading to the sequence $(\overline{a}_i)_{i=1}^n = (i-1-a_i)_{i=1}^n$ which is the inversion table of $\overline{P}$. Graphically, we obtain $\overline{\Omega}$ from Ω by starting from the right edge of the grid $\mathcal{G}_n$, putting (reversed) E-steps until we reach a junction. We then continue by putting (reversed) F-steps until we touch the left edge of the grid. We get a reversed mixed configuration; a vertical reflection gives the (ordinary) mixed configuration $\overline{\Omega}$ (see figure 6).

For mixed configurations $\Omega = (\omega_1, \ldots, \omega_n) \in \mathcal{M}_{n,1}$ (or more generally $\mathcal{M}_{n,s}$), the graphical procedure is similar, with the added rules:

- replace every S-step by a reversed S-step with the same starting vertex.
- replace every N-step by a reversed N-step with the same ending vertex.

Figure 7 (left) shows how this is done. Observe that duality (before the final vertical reflection) preserves the positions of the starting vertex of the S-step, of the ending vertex of the N-steps and of the junctions.

Theorem 6.1. *Let* $A \in \mathcal{A}_{n,1}$ *and* $\Omega = (\omega_1, \ldots, \omega_n) = \Phi(\Lambda(A))$. *Then* $\overline{\Omega} = \Phi(\Lambda(\overline{A}))$.

References

[Br] Bressoud D. M., *Proofs and Confirmations* MAA, Spectrum series, Cambridge Univ. Press, New York, (1999).

[GV] Gessel I., Viennot X., *Binomial determinants, paths and hook length formulae*, Adv. in Math., **58**, (1985), 300–321.

[Kr] Krattenthaler, C., *Descending plane partitions and rhombus tilings of a hexagon with triangular hole* Preprint, arXiv: math.CO/0310188v1.

[Ku] Kuperberg, G., *Another proof of the alternating sign matrix conjecture* International Mathematics Research Notes, (1996), 139–150.

[La1] Lalonde, P., *q-Enumeration of alternating sign matrices with exactly one 1*, Disc. Math., **256**, #3, 2002, 759–773.

[La2] Lalonde, P., *Lattice paths and the antiautomorphism of descending plane partitions*, Disc. Math., **271**, 2003, 311–319.

[MRR] Mills W.H., Robbins D.P., Rumsey H., *Alternating sign matrices and descending plane partitions*, J. Comb. Th.(A), **34**, (1983), 340–359.

[Vi] Viennot X.G., *Une théorie combinatoire des polynômes orthogonaux généraux*, Publication du LaCIM, H1, Montréal (1983).

[Ze1] Zeilberger, D., *Proof of the alternating sign matrix conjecture*, Electronic Journal of Combinatorics, **3**, (1996), R13.

[Ze2] Zeilberger, D., *Proof of the refined alternating sign matrix conjecture*, New York J. Math., **2**, (1996), 59–68.

LaCIM, UQÀM, C.P. 8888, Succ. "A", Montréal, Qc, Canada, H3C 3P8
E-mail address: `lalonde@math.uqam.ca`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Strict Partitions and Discrete Dynamical Systems

LE Minh Ha and PHAN Thi Ha Duong

Abstract. *We prove that the set of partitions with distinct parts of a given positive integer under dominance ordering can be considered as a configuration space of a discrete dynamical model with two transition rules and with initial configuration being the singleton partition. This allows us to characterize its lattice structure, fixed point, longest chains as well as their length, using Chip Firing Game theory. Finally, two extensions and their applications are discussed.*

Résumé. *Nous montrons que l'ensemble des partitions avec differents parts d'un entier donné n muni de l'ordre de dominance peut être considérećomm l'espace de configurations d'un système dynamique discret avec deux règles de transitions et avec la configuration initialle étant la partition (n). Cela nous permet de caractériser sa structure de treillis, son point fixe, les chaînes les plus longues ainsi que leurs longueur, en utilisant la theorie de Chip Firing Game. Enfin, deux extensions et leurs applications sont données.*

1. Introduction

A partition of a positive integer n is a sequence of non-increasing positive integers $a = (a_1, \dots a_m)$ such that $a_1 + \cdots + a_m = n$. The set of all such partitions of n is denoted by $\mathcal{P}(n)$. $\mathcal{P}(n)$ is equipped with a partial order called *dominance order* as follows : $a \geq b$ if its partial sums is greater than that of b, *i.e.* $\sum_{i=1}^{j} a_i \geq \sum_{i=1}^{j} b_i$. This order has been showed to have many applications to problems in combinatorics as well as group representation theory, among other fields. The structure of this poset was studied by Brylawski [**Bry73**] who showed in particular that it is a lattice. Since then, other properties such as maximal chains, fixed point have also been characterized in [**Bry73, GK86, GK93**]. In [**LP01**], Phan and Latapy constructed its infinite extension and obtained a construction algorithm.

In this paper, we study the structure of an interesting class $\mathcal{SP}(n)$ of partitions of n called strict partitions, or partitions with distinct parts, from the point of view of discrete dynamical systems. For any strict partition a of n, one can apply on a the following transition rules so that the resulting partition is also strict :

- Vertical transition (V-transition):

$$(a_1, \dots, a_i, a_{i+1}, \dots, a_n) \rightarrow (a_1, \dots, a_i - 1, a_{i+1} + 1, \dots, a_n),$$

  if $a_i - a_{i+1} \geq 3$.

1991 *Mathematics Subject Classification.* Primary 06A07, 05C20, 05A17, 37E25; Secondary 05C85, 68Q45, 82C20.
Key words and phrases. Strict partition, Chip Firing Game, poset, lattice, discrete dynamical systems.

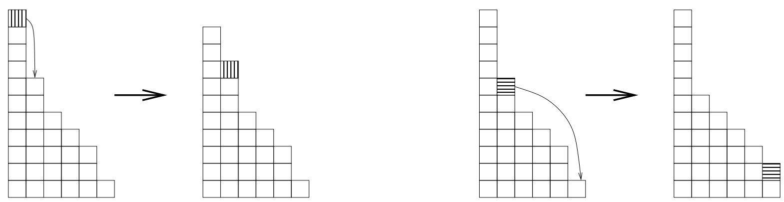


FIGURE 1. Vertical transition and horizontal transition

- Horizontal transition (H-transition):

$$(a_1, \ldots, p+l+1, p+l-1, p+l-2, \ldots, p+2, p+1, p-1, \ldots, a_n) \rightarrow$$
$$(a_1, \ldots, p+l, p+l-1, p+l-2, \ldots, p+2, p+1, p, \ldots, a_n).$$

These rules then define a partial order on $\mathcal{SP}(n)$ by declaring that $b \leq_S a$ if b can be obtained from a via a sequence of transitions. In particular, we will show that all strict partitions can be obtained in this way if the initial configuration is the singleton partition (n). Moreover, the poset $\mathcal{SP}(n)$ which corresponds to the above order and initial configuration (n) turns out to be the same as the poset $\mathcal{SP}(n)$ with dominance order, so that the two orders can now be identified. We then show that $\mathcal{SP}(n)$ is also a lattice, but it is not a sublattice of $\mathcal{P}(n)$. Furthermore, unlike $\mathcal{P}(n)$, $\mathcal{SP}(n)$ is not self-dual. Using the fact that our dynamical model can be viewed as a "composition" of two Chip Firing Games in the sense of [**BL92**] (see also [**LP01**], [**GMP02**]), we are able to characterize explicitly the fixed point, longest chains as well as their length in $\mathcal{SP}(n)$. Finally, we present two generalizations : We obtain similar results for the set of k-strict partitions, *i.e.* partitions where two parts differ by at least $k > 0$. We also obtain an infinite extension of $\mathcal{SP}(n)$ and an algorithm to construct $\mathcal{SP}(n+1)$ from $\mathcal{SP}(n)$ in linear time.

2. Lattice structure of $\mathcal{SP}(n)$

Theorem 2.1. *The set $\mathcal{SP}(n)$ is exactly the set of all strict partitions reachables from (n) by applying two transitions rule V and H. .*

PROOF. Let $a = (a_1, \ldots, a_m)$ be a strict partition. It suffices to show that if a is different from (n) itself, then there exist another strict partition a' such that one can recover a by applying a transition on a'.

First of all, observe that if there is a subsequence $(a_i, a_{i+1}, \ldots, a_j)$ of consecutive numbers in a, where $i = 1$, or else $a_{i-1} - a_i \geq 2$, similarly $j = m$ or else $a_j - a_{j+1} \geq 2$. Then we can choose

$$a' = (a_1, \ldots, a_{i-1}, a_i + 1, a_{i+1}, \ldots, a_{j-1}, a_j - 1, a_{j+1}, \ldots, a_m),$$

so that a' is again strict. Furthermore, one recovers a from a' by applying a H-transition.

On the other hand, if no such subsequence exists, then $a_1 - a_2 \geq 2$ and either $m = 2$ or $a_2 - a_3 \geq 2$. In this case, we can simply choose

$$a' = (a_1 + 1, a_2 - 1, a_3, \ldots, a_m).$$

It is easy to check that a' is a strict partition and that a V-transition applied on a' at the first position gives back a. The theorem is proved. □

Proposition 2.2. *$\mathcal{SP}(n)$ is a subposet of $\mathcal{P}(n)$.*

PROOF. It is sufficient to show that if $a, b \in \mathcal{SP}(n)$ and $a > b$ then $a >_S b$, *i.e.* there exists a sequence of transitions from a to b. For this purpose, it suffices to prove that one can apply a transition on a to obtain a new strict partition a' such that one still has $a' \geq b$.

Since $a > b$, we have $\sum_{i=1}^{j} a_i \geq \sum_{i=1}^{j} b_i$ for all $1 \leq j \leq n$. Let j be the smallest index where $a_j > b_j$. Then let ℓ be the smallest index such that $\ell > j$ and $\sum_{i=1}^{l} a_i = \sum_{i=1}^{l} b_i$. Such a number ℓ exists because $\ell = n$ satisfies both conditions above. It is clear that $a_l < b_l$ because of the choice of ℓ.

We claim that we can apply a transition on a at some positions between j and ℓ, so that the newly constructed partition a' are identical with a outside this range . If this is possible, then we are done, because it is easy to verify, using the definition of j and ℓ,that $a' > b$ in $\mathcal{P}(n)$.

To construct a', observe that if there is an index $j \leq i \leq \ell$ such that $a_i - a_{i+1} \geq 3$, then a V-transition can be applied at position i and we are done.

Suppose now that $a_i - a_{i+1} \leq 2$ for all $j \leq i < \ell$. Since b is a strict partition and $b_i \geq 1$ for all i, we have $b_\ell - b_j \geq \ell - j$. But $a_\ell > b_\ell$ and a_j, b_j, hence $a_\ell - a_j \geq \ell - j + 2$. It follows that there exists at least two indices $j \leq r < s < \ell$ such that $a_r - a_{r+1} = a_s - a_{s+1} = 2$. Furthermore, by choosing a different pair of indices if necessary, we can even assume that $a_i - a_{i+1} = 1$ for all $r < i < s$. But in this case, the subsequence $(a_r, \ldots, a_s)$ is of exactly the form where one can apply a H-transition. The proof is finished. □

Because of the above result, we can now write $b \leq a$ instead of $b \leq_S a$ for any two strict partition a and b.

Theorem 2.3. *$\mathcal{SP}(n)$ is a lattice. Moreover, the meet operation in $\mathcal{SP}(n)$ is the same as that in $\mathcal{P}(n)$, i.e. $a \wedge_S b = a \wedge b$ for any two strict partitions a and b.*

Proof. Since $\mathcal{SP}(n)$ contains a maximal element, it is enough to prove that any pair of element in $\mathcal{SP}(n)$ has a greatest lower bound. Of course, their greatest lower bound $c = a \wedge b$ in $\mathcal{P}(n)$ does exist, but is it true that c is again a strict partition? We will show that this is the case for any pair of strict partitions a and b.

By definition, c is a partition defined by the formulae

$$\sum_{i=1}^{m} c_i = min(\sum_{i=1}^{m} a_i, \sum_{i=1}^{m} b_i)$$

for all $1 \leq m$. Suppose that $c_m > 0$. Without loss of generality, assume that $\sum_{i=1}^{m} c_i = \sum_{i=1}^{m} a_i$. Then $c_{m+1} \leq a_{m+1}$ while $a_m \leq c_m$. Thus $c_{m+1} < c_m$ because $a_{m+1} < a_m$. Hence c is also a strict partition. The proof above clearly also implies that the meet operation in $\mathcal{SP}(n)$ is the same as that in $\mathcal{P}(n)$. □

Remark 2.4. $\mathcal{SP}(n)$ is **not** a sublattice of $\mathcal{P}(n)$. In fact, the joint operations in $\mathcal{SP}(n)$ and $\mathcal{P}(n)$ are different. For example, $(8,4,3,1) \vee (7,5,4) = (8,4,4)$ which is not a strict partition. Nevertheless, we still have $a \vee_S b \geq a \vee b$ for any a and b.

Since $\mathcal{SP}(n)$ is a lattice, it has an unique minimal element (or fixed point). We finish this section by giving an explicit formula for this minimal partition. Let p be the unique number such that

$$\frac{1}{2}p(p+1) \leq n < \frac{1}{2}(p+1)(p+2).$$

Then let $q = n - \frac{1}{2}p(p+1)$. One verifies easily that $q < p$. Now let Π be the following partition

$$\Pi = ((p+1), p, \ldots, (p-q+2), (p-q), (p-q-1), \ldots, 2, 1). \tag{2.1}$$

It is evident that Π is a strict partition on which no transition can be applied. Thus we have the following proposition:

Proposition 2.5. *Π is the fixed point of the lattice $\mathcal{SP}(n)$.*

3. Longest chains

In this section, we characterize longest chains in $\mathcal{SP}(n)$ as well as their length. The longest chains in $\mathcal{P}(n)$ were characterized by Greene and Kleitman [**GK86**] where they introduced the notion of VH-chain (*i.e.* a chain of V-transitions followed by a chain of H-transitions) and proved that all VH-chains are longest chains. It turns out that the same is true for strict partitions. Our proof, however, is different. The proof in [**GK86**] makes use of a series of delicate lemmas which basically consider the differences of consecutive parts of partitions. We believe that our proof, which is based on the theory of Chip Firing Game on directed graph (CFG) [**BL92**], is simpler and probably can be adapted in other contexts.

3.1. V(H)-chain. Let us first introduce some definitions. A V(resp. H)-chain is a chain of V(resp. H)-transitions, and a VH-chain is a concatenation of a V-chain and a H-chain. If there is a V-chain from a strict partition a to another b, then we say that b is V-reachable from a. But a partition c is H-reachable from d means that there is an H-transitions from d back to c, or equivalently an inverse H-transition from c to d.

We will also need the two functions *V-weight* $w_V(a)$ and *H-weight* $w_H(a)$ on a strict partition a. From the Ferrers diagram for a, let

$$w_V(a) = \sum (i-1)a_i \tag{3.1}$$

and

$$w_H(a) = \sum (k-1)\tilde{a}_k, \tag{3.2}$$

where $\tilde{a}_k$ is the number of cells (i,j) on the segment $i+j=k+1, i \geq 0, j \geq 0$. It is easy to see that a V-transition increases V-weight by 1, but decrease H-weight by at least 1. On the other hand, an H-transition decreases H-weight by 1, and increases V-weight by at least 1. This simple observation shows that V-chains (or H-chains) between two partitions are longest chains.

3.2. Chip Firing Game. We now give a brief overview of the theory of Chip Firing Game (CFG for short). In particular, we show that the dynamical model consisting of only the V-transition (resp. H-transition) are examples of CFG. For more details account of theory of Chip Firing Game, we refer to [**BLS91, BL92, LP01, GLM$^+$ar**].

A Chip Firing Game is a discrete dynamical system defined on a (directed) graph $G=(V,E)$, where each configuration consists of a partition of n chips on the vertices V, and obeys the following rule, called *firing rule*: a vertex containing at least at many chips as its outgoing degree (*i.e.* the number of outgoing edges) transfers one chip along each of its outgoing edges.

This rule defines a natural partial order on the space of configurations by declaring that a configuration b is smaller than a if b can be obtained from a by iterating the firing rule. A *fixed point* of a CFG is a configuration where no firing is possible. The following is the fundamental result in the theory of CFG,

Theorem 3.1. [**BL92, LP01**] *The set of all configurations reachable from the initial one of a CFG with no closed component is a lattice.*

A closed component of a graph is a strongly connected component without outgoing edge.

One can also characterize the natural order defined above using the notion of shot vector. If $b < a$, then the *shot vector* $k(a,b)$ is the vector in $\mathbb{N}^{|V|}$ whose entry $k_v(a,b)$ is the number of firings at vertex v to obtain b from a. This vector depends only on a and b but not on a chosen sequence of firings. We then have:

Lemma 3.2. [**LP01**] *Let c and d be two configurations reachable from the same initial configuration a in a CFG. Then $c \geq d$ if and only if $k_v(a,c) \leq k_v(a,d)$ for all vertices $v \in V$.*

Here are two important examples of CFG.

Example V: The dynamical model consisting to only the V-transition is a CFG. Indeed, consider the graph $G=(V,E)$ with $n+1$ vertices defined pictorially as follows:

$$\overset{v_0}{\circ} \leftarrow \overset{v_1}{\circ} \leftrightarrows \overset{v_2}{\circ} \ \ldots \leftrightarrows \overset{v_n}{\circ}$$

Thus each vertex of G, beside v_0 and v_n has outgoing degree 2. Now let a be a configuration *i.e.* a strict partition of n, we put $d_i = a_i - a_{i+1} - 1$ chips at vertex v_i for all $i \geq 1$ and no chip at v_0.

The necessary condition to apply a V-transition at position i on a is $a_i - a_{i+1} \geq 3$, or equivalently $d_i \geq 2$ which is the same as the condition to apply the CFG firing rule on v_i. It is easy to see that the space of reachable configurations of this CFG is exactly the set of partitions that are V-reachable from a. In particular, the unique fixed point of this CFG corresponds to the smallest partition which is V-reachable from a. In fact, in any interval $b \leq a$ in $\mathcal{SP}(n)$ there exists a unique smallest strict partition $\lfloor a \rfloor_b$ which is V-reachable from a.

Example H: The inverse H-transition defines a CFG on the same graph as in previous example in a similar way. For each initial configuration b, we put $\tilde{d}_i = \tilde{b}_i - \tilde{b}_{i+1}$ chips at vertex v_i for all $i \geq 1$ and no chip at v_0. Here $\tilde{b}_i$ is defined as in (3.1). One verifies again that the space of configurations of this game is the set of H-reachable from b and the fixed point corresponds to unique greatest partition which is H-reachable from b. Furthermore, there is a unique greatest strict partition $\lceil b \rceil^a$ which is H-reachable from b in any interval $b \leq a$.

The reader of [**GK86**] may find it interesting to compare the definition of H-transition there and ours, as in Example H. While an H-transition in the sense of [**GK86**] is dual to a V-transition in the usual sense (row vs column), our H-transition is defined in terms of the "diagonal" in the Ferrers diagram of the corresponding partition.

3.3. VH-chains are longest chains. First of all, it is not hard to show, as in [**GK86**, Lemma 3] that any longest chain must be a VH-chain. The point is that any sequence of two transitions (H,V) (H follows by V) equals sequence of the form either (V,H) or (V,V,H), proof by direct inspection. Thus for any chain of transitions between two partitions, there is a VH-chain of at least the same length.

It remains to show that any VH-chain is a longest chain. We begin with the following key lemma which explains the relevance of dominance order.

Lemma 3.3. *Let c and d be two partitions which are V-reachable from a. If $d \leq c$, then d is V-reachable from c.*

Proof. We compute the shot vector $k(a,c)$ and $k(a,d)$ in the corresponding CFG. It is easy to see that $k_i(a,c) = k_{i-1}(a,c) + a_i - c_i$ for all $i \geq 1$, which implies that $k_i(a,c) = \sum_{j=1}^{i} a_j - \sum_{j=1}^{i} c_j$. Similarly, $k_i(a,d) = \sum_{j=1}^{i} a_j - \sum_{j=1}^{i} d_j$. On the other hand, $\sum_{j=1}^{i} c_j \geq \sum_{j=1}^{i} d_j$ because $c \geq d$. It follows that $k_i(a,c) \leq k_i(a,d)$ and so d is V-reachable from c by Lemma 3.2. □

Lemma 3.4. *If $a \geq b$, then $\lfloor a \rfloor_b$ is H-reachable from b and $\lceil b \rceil^a$ is V-reachable from a.*

Proof. There is a VH-chain from $\lfloor a \rfloor_b$ to b. Since $\lfloor a \rfloor_b$ is the smallest strict partition which is V-reachable from a in interval $a \leq b$, there can not be no V-transition in this chain and $\lfloor a \rfloor_b$ is H-reachable from b. Similar argument applied for $\lceil b \rceil^a$. □

As an immediate corollary, we see that there is a VH-chain $a \to \lceil b \rceil^a \to b$ from a to b of length $w_V(a, \lceil b \rceil^a) + w_H(\lceil b \rceil^a, b)$.

We can now state the main result of this section:

Theorem 3.5. *All VH-chains from a to b in $\mathcal{SP}(n)$ have the same length and this length is maximal.*

Proof. Suppose that $a \xrightarrow{V} c \xrightarrow{H} b$ is a VH-chain from a to b with length $w_V(c) - w_V(a) + w_H(b) - w_H(c)$. We will show that it has the same length as that of the VH-chain $a \xrightarrow{V} \lceil b \rceil^a \xrightarrow{H} b$. In particular, its length only depends on a and b and is maximal.

It is clear from the definition of $\lfloor a \rfloor_b$ and $\lceil b \rceil^a$ that $\lfloor a \rfloor_b \leq c \leq \lceil b \rceil^a$. Since both $\lceil b \rceil^a$ and c are V-reachable from a and $\lceil b \rceil^a \geq c$, then there is a V-chain from $\lceil b \rceil^a$ to c by Lemma 2. On the other hand, there is also an H-chain from $\lceil b \rceil^a$ to c because $\lceil b \rceil^a$ is the minimum element of the lattice of all

H-reachable strict partitions from b which contains c. The two chains are both of maximal length, hence $w_V(c) - w_V(\lceil b\rceil^a) = w_H(\lceil b\rceil^a) - w_H(c)$. The required result immediately follows from the equalities :

$$\begin{aligned} w_V(c) - w_V(a) &= w_V(\lceil b\rceil^a) - w_V(a) + w_V(c) - w_V(\lceil b\rceil^a) \\ w_H(c) - w_H(b) &= w_H(\lceil b\rceil^a) - w_H(b) - (w_H(\lceil b\rceil^a) - w_H(c)). \end{aligned}$$

□

3.4. The length of a longest chain. Once we know all VH-chains are longest chains, it is sufficient to calculate the length of a well-chosen VH-chains from (n) to Π. The VH-chain that we will use is $(n) \xrightarrow{V} \lfloor (n) \rfloor_\Pi \xrightarrow{H} \Pi$. For the point $P = \lfloor (n) \rfloor_\Pi$, which is the fixed point of the CFG in Example V with initial configuration (n), together with the length of the V-chain from $(n) \to \lfloor (n) \rfloor_\Pi$ was already computed in [**GMP02**]. Our model corresponds to the model named $L(n, 3)$ in that article. To describe P and $w_V((n), P)$, first write n in the form $n = k(k+1) + \ell(k+1) + h$, where $0 \le \ell \le 1, 0 \le h \le k$. The integers k, ℓ, h are all uniquely determined from n. We have

Proposition 3.6.

$$P = (\ell + 2k, \ell + 2(k-1), \ldots, \ell + 2h, \ell + 2(h-1) + 1, \ldots, \ell + 2 + 1, \ell + 1), \tag{3.3}$$

and

$$w_V(P) = \frac{(k-1)k(k+1)}{3} + \ell\frac{k(k+1)}{2} + h\frac{2k-h+1}{2}. \tag{3.4}$$

We can now state the following result:

Proposition 3.7. *Let p, q the unique integers such that $n = \frac{1}{2}p(p+1) + q, 0 \le q \le p$ and let k, ℓ, h the unique integers such that $n = k(k+1) + \ell(k+1) + h, 0 \le \ell \le 1, 0 \le h \le k$. We have the following formula for the length L of longest chains in $\mathcal{SP}(n)$:*

$$L = \frac{k(k+1)(8k-5)}{6} + 2\ell k(k+1) + (2k+l)h - \frac{(p-1)p(p+1)}{3} - qp.$$

Proof. Since $L = w_V(P) + w_H(P) - w_H(\Pi)$, we have from (2.1) and (3.2):

$$w_H(\Pi) = \sum_{i=1}^{p}(i-1)i + qp = \frac{(p-1)p(p+1)}{3} + qp.$$

and from (3.3) and (3.2):

$$\begin{aligned} w_H(P) &= \sum_{i=1}^{k}(i-1)i + \sum_{i=k}^{1}(2k-i)i + l\sum_{i=k}^{2k} i + \sum_{i=0}^{h-1}(k+\ell+i) \\ &= \frac{1}{2}k(k+1)(2k-1) + \frac{3}{2}\ell k(k+1) + \frac{1}{2}(2k+2\ell+h-1)h. \end{aligned}$$

□

4. Infinite extension of $\mathcal{SP}(n)$

It is natural to ask whether one can construct the lattice $\mathcal{SP}(n+1)$ from $\mathcal{SP}(n)$. More generally, what is the precise relationship between the lattices $\mathcal{SP}(n)$ for various n. Our solution to these questions is to assemble them together into a lattice $\mathcal{SP}(\infty)$ called lattice of strict partitions of infinity. Indeed, this lattice is constructed in a similar way as $\mathcal{SP}(n)$ by pretending that n can be as large as needed. More precisely, it is the lattice obtained from the dynamical system with two transitions rules as those for $\mathcal{SP}(n)$, and the initial configuration is infinity. Equivalently, one can also define $\mathcal{SP}(\infty)$ in terms of dominance order : A strict partition of infinity is just a sequence of finitely many strictly decreasing positive integers, except the

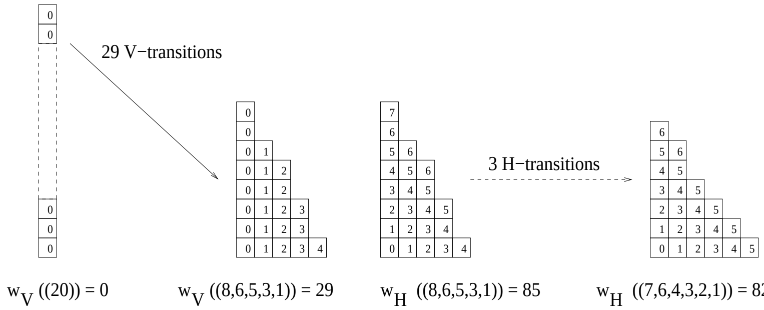


FIGURE 2. A longest chain in $\mathcal{SP}(23)$: $P = (8, 6, 5, 3, 1)$ and $\Pi = (7, 6, 4, 3, 2, 1)$. A longest chain in $\mathcal{SP}(23)$ is a chain containing a V-chain from (23) to P and an H-chain from P to Π, and its length is $w_H(8, 6, 5, 3, 1) + w_V(8, 6, 5, 3, 1) - w_V(7, 6, 4, 3, 2, 1) = 29 - 0 + 85 - 82 = 32$.

first entry : $(\infty, a_2, a_3, \ldots a_k)$. The partial order is defined by declaring that $a \geq_\infty b$ if $\sum_{i \geq j} a_i \leq \sum_{i \geq j} b_i$ for all $j \geq 2$. By convention, we put $a_n = 0$ for $n > k$.

Many results presented in this section are obtained initially in the case normal partitions in [**LP99**]. However, the proofs are not completely similar since we must be careful that our operations are within the set of strict partition. In fact, even though $\mathcal{SP}(n)$ can be embedded in a $\mathcal{P}(n)$, the structure of the infinite lattices or infinity trees are different.

4.1. Notations and definitions. If $a = (a_1, a_2, \ldots, a_k)$ is a strict partition, then the partition obtained from a by adding one grain on its i-th column is denoted by $a^{\downarrow_i}$. Notice that $a^{\downarrow_i}$ is not necessarily a strict partition. If S is a set of strict partitions, then $S^{\downarrow_i}$ denotes the set $\{a^{\downarrow_i} | a \in S\}$. We denote $a \xrightarrow{i} b$ if b is obtained from a by applying a transition at position i and by $Succ(a)$ the set of configurations directly reachable from a.

Write $d_i(a) = a_i - a_{i+1}$ with the convention that $a_{k+1} = 0$. We say that a has a *cliff* at position i if $d_i(a) \geq 3$. If there exists an $\ell > i$ such that $d_j(a) = 1$ for all $i \leq j < \ell$ and $d_\ell(a) = 2$, then we say that a has a *slippery plateau* at i with *length* $(\ell - i)$. Likewise, a has a *non-slippery plateau* at i if $d_j(a) = 1$ for all $i \leq j < \ell$ and it has a cliff at ℓ. The integer $\ell - i$ is called the *length* of the non-slippery plateau at i. The partition a has a *(non)-slippery step* at i if there is a strict partition b such that $b^{\downarrow_i} = a$ and b has a (non)-slippery plateau at i. See Figure 3 for some illustrations. The set of elements of $\mathcal{SP}(n)$ that begin

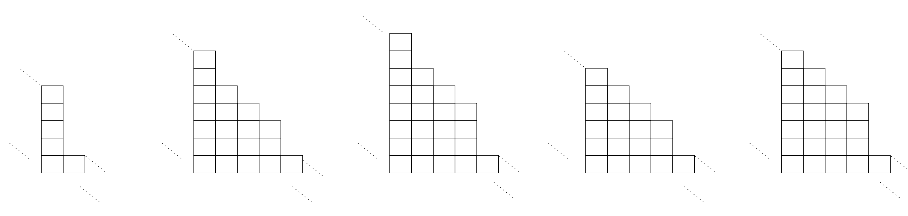


FIGURE 3. From left to right: a cliff, a slippery step, a non-slippery step, a slippery plateau and a non-slippery plateau.

with a cliff, a slippery step, a non-slippery step, a slippery plateau of length l and a non-slippery plateau of length l are denoted by C, SS, nSS, SP_l, nSP_l respectively.

4.2. Constructing $\mathcal{SP}(n+1)$ from $\mathcal{SP}(n)$. Let $a = (a_1, a_2, \ldots, a_k)$ be a strict partition. It is clear that $a^{\downarrow_1}$ is again a strict partition. This define an embedding $\pi : \mathcal{SP}(n) \to \mathcal{SP}(n)^{\downarrow_1} \subset \mathcal{SP}(n+1)$ which can be proved, by using infimum formula of $\mathcal{SP}(n)$ and $\mathcal{SP}(n+1)$, as a lattice map.

Proposition 4.1. *$\mathcal{SP}(n)^{\downarrow_1}$ is a sublattice of $\mathcal{SP}(n+1)$.*

Our next result characterizes the remaining elements of $\mathcal{SP}(n+1)$ that are not in $\mathcal{SP}(n)^{\downarrow_1}$.

Theorem 4.2. *For all $n \geq 1$, we have*

$$\mathcal{SP}(n+1) = \mathcal{SP}(n)^{\downarrow_1} \sqcup SS^{\downarrow_2} \sqcup nSS^{\downarrow_2} \ \sqcup_l SP_l^{\downarrow_{l+1}}.$$

PROOF. It is easy to check that each element in one of the sets $\mathcal{SP}(n)^{\downarrow_1}$, $SS^{\downarrow_2}$, $nSS^{\downarrow_2}$ and $SP_l^{\downarrow_{l+1}}$ is an element of $\mathcal{SP}(n+1)$, and that these sets are disjoint.

Now let us consider an element b of $\mathcal{SP}(n+1)$. If b begins with a cliff or a step then b is in $\mathcal{SP}(n)^{\downarrow_1}$. If b begins with a slippery plateau of length 2 then b is in $SS^{\downarrow_2}$, if b begins with a non-slippery of length 2 then b is in $nSS^{\downarrow_2}$. And if b begins with a plateau of length $l+1, l \geq 2$, then b is in $SP_l^{\downarrow_{l+1}}$. □

Finally, we describe an algorithm to compute the successors of any given element of $\mathcal{SP}(n+1)$, thus giving a complete construction of $\mathcal{SP}(n+1)$ from $\mathcal{SP}(n)$.

Proposition 4.3. *Let x be an element of $\mathcal{SP}(n+1)$.*

(1) *Suppose $x = a^{\downarrow_1} \in \mathcal{SP}(n)^{\downarrow_1}$.*
   - *If a is in C or nSP then $Succ(a^{\downarrow_1}) = Succ(a)^{\downarrow_1}$,*
   - *If a is in SP_l then $Succ(a^{\downarrow_1}) = Succ(a)^{\downarrow_1} \cup \{a^{\downarrow_{l+1}}\}$,*
   - *If a is in SS then let b be such that $a \xrightarrow{1} b$. We have $Succ(a^{\downarrow_1}) = (Succ(a) \setminus \{b\})^{\downarrow_1} \cup \{a^{\downarrow_2}\}$.*

(2) *If $x = a^{\downarrow_2} \in SS^{\downarrow_2}$ where $a \in SS$: Let b be such that $a \xrightarrow{1} b$, then $Succ(a^{\downarrow_2}) = (Succ(a) \setminus \{b\})^{\downarrow_2} \cup \{b^{\downarrow_1}\}$.*

(3) *If $x = a^{\downarrow_2} \in nSS^{\downarrow_2}$ with $a \in nSS$, then $Succ(a^{\downarrow_2}) = Succ(a)^{\downarrow_2}$.*

(4) *Finally, if $x = a^{\downarrow_{l+1}} \in SP_l^{\downarrow_{l+1}}$ for some $a \in SP_l$, then*
   - *If a has a cliff at $l+1$ or a non-slippery step at l, then $Succ(a^{\downarrow_{l+1}}) = Succ(a)^{\downarrow_{l+1}}$,*
   - *If a has a slippery step at l, let b such that $a \xrightarrow{l} b$ in $\mathcal{SP}(n)$, then $Succ(a^{\downarrow_{l+1}}) = (Succ(a) \setminus \{b\})^{\downarrow_{l+1}} \cup \ \{b^{\downarrow_l}\}$.*

PROOF. We will give the proof for the two most difficult cases (1) and (4). Consider $x = a^{\downarrow_1}$ where $a \in C$: notice first that the transitions possible from a on columns other than the first one are still possible from $a^{\downarrow_1}$, and on the other hand the addition of one grain on a cliff does not allow any new transition from the first column, since such a transition was already possible.

In the last case: $x = a^{\downarrow_{l+1}}$ where $a \in SP_l^{\downarrow_{l+1}}$ and a has a slippery step of length l' at l. Then, $a \xrightarrow{l} b$ in $\mathcal{SP}(n)$. The possible transitions from $a^{\downarrow_{l+1}}$ are the same as the possible ones from a, except the transition on the column l. All the elements directly reachable from a except b have a slippery plateau at 1, therefore the elements of $(Succ(a) \setminus \{b\})^{\downarrow_{l+1}} \in Succ(a^{\downarrow_{l+1}})$ The only one missing transition is: $a^{\downarrow_{l+1}} \xrightarrow{l+1} a^{\downarrow_{l+l'+1}}$. But we can verify that $a^{\downarrow_{l+l'+1}} = b^{\downarrow_l}$. □

Proposition 4.3 makes it possible to write an algorithm to construct the lattice $\mathcal{SP}(n+1)$ in linear time (with respect to its size).

4.3. The infinite lattice $\mathcal{SP}(\infty)$. Imagine that (∞) is the initial configuration where the first column contains infinitely many grains and all the other columns contain no grain. Then the transitions V and H defined in the first section can be performed on (∞) just as if it is finite, and we call $\mathcal{SP}(\infty)$ as the set of all the configurations reachable from (∞). A typical element a of $\mathcal{SP}(\infty)$ has the form $(\infty, a_2, a_3, \ldots, a_k)$. As in the previous section, we find that the dominance ordering on $\mathcal{SP}(\infty)$ (when the first component is ignored) is equivalent to the order induced by the dynamical model. The first partitions in $\mathcal{SP}(\infty)$ are given

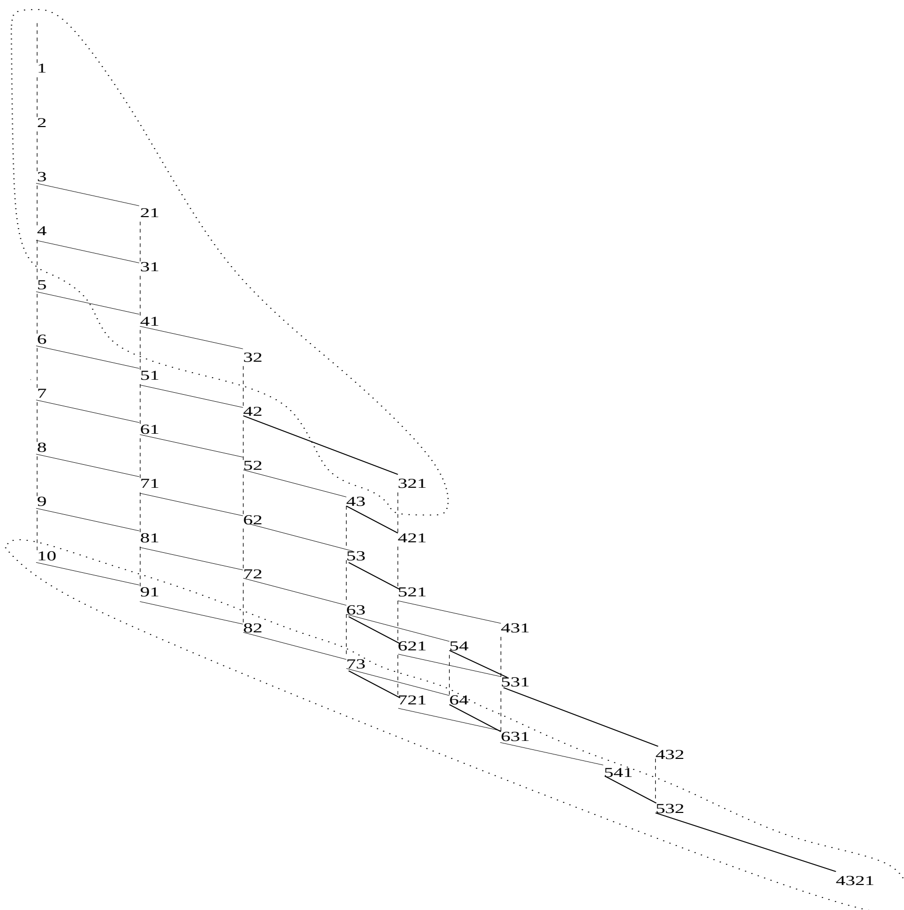


FIGURE 4. The first elements and transitions of $\mathcal{SP}(\infty)$. As shown on this figure for $n = 10$, we will see two ways to find parts of $\mathcal{SP}(\infty)$ isomorphic to $\mathcal{SP}(n)$ for any n.

in Figure 4, along with their covering relations (the first component, equal to ∞, is not represented on this diagram).

We start by showing that $\mathcal{SP}(\infty)$ is a lattice. We also obtain a formula for the minimum in $\mathcal{SP}(\infty)$. Furthermore, for any n, there are two different ways to find sublattices of $\mathcal{SP}(\infty)$ isomorphic to $\mathcal{SP}(n)$. We will also give a way to compute some other special sublattices of $\mathcal{SP}(\infty)$, using its self-similarity.

Theorem 4.4. *The set $\mathcal{SP}(\infty)$ is a lattice. Moreover, for any two elements $a = (\infty, a_2, \ldots, a_k)$ and $b = (\infty, b_2, \ldots, b_\ell)$ of $\mathcal{SP}(\infty)$, then $\inf_{\mathcal{SP}(\infty)}(a, b) = c$ in $\mathcal{SP}(\infty)$, where c is defined by:*

$$c_i = max(\sum_{j \geq i} a_j, \sum_{j \geq i} b_j) - \sum_{j > i} c_j \quad \textit{for all } i \textit{ such that } 2 \leq i \leq max(k, l).$$

PROOF. One just needs to check that c is an element of $\mathcal{SP}(\infty)$, *i.e.* $c_1 = \infty$ and $c_i > c_{i+1}$ for all $i > 1$. ☐

Now for any $n > 1$, there are two canonical embeddings of $\mathcal{SP}(n)$ in $\mathcal{SP}(\infty)$, defined by

$$\begin{array}{cccc} \pi: & \mathcal{SP}(n) & \longrightarrow & \mathcal{SP}(\infty) \\ & a=(a_1,a_2,\ldots,a_k) & \mapsto & \pi(a)=(\infty,a_2,\ldots,a_k) \end{array}$$

$$\begin{array}{cccc} \chi: & \mathcal{SP}(n) & \longrightarrow & \mathcal{SP}(\infty) \\ & a=(a_1,a_2,\ldots,a_k) & \mapsto & \chi(a)=(\infty,a_1,a_2,\ldots,a_k) \end{array}$$

Theorem 4.5. *Both π and χ are embedding of lattices.*

Proof. The first embedding π comes from our result in the Proposition 1. The second is clear by noting that for all $a,b \in \mathcal{SP}(n)$, $a \stackrel{l}{\longrightarrow} b$ if and only if $\chi(a) \stackrel{l+1}{\longrightarrow} \chi(b)$ in $\mathcal{SP}(\infty)$. □

Corollary 4.6. *Let*

$$\mathcal{SP}(\leq n) \;=\; \bigsqcup_{0\leq i\leq n} \mathcal{SP}(i),$$

then $\mathcal{SP}(\leq n)$ is a sublattice of $\mathcal{SP}(\infty)$ (by the embedding χ).

So by using the embedding χ, one can consider $\mathcal{SP}(\infty)$ as the union disjoint of $\mathcal{SP}(n)$ for all n, $\mathcal{SP}(\infty) = \bigsqcup_{n\geq 0} \mathcal{SP}(n)$.

4.4. Self-reference property: the infinite binary tree $T_B(\infty)$. Observe that each element a of $\mathcal{SP}(n+1)$ can be obtained from an element b of $\mathcal{SP}(n)$ by addition of one grain at some position i; that is $a = b^{\downarrow i}$. We will represent this relation by a tree where $a \in \mathcal{SP}(n+1)$ is a child of $b \in \mathcal{SP}(n)$ if and only if $a = b^{\downarrow i}$ for some $i \geq 0$, and we label the edge $b \longrightarrow a$ by i. We denote this tree by $\mathcal{ST}(\infty)$. The root of $\mathcal{ST}(\infty)$ is the empty partition. We will describe two ways to compute all strict partitions of a given positive integer n in $\mathcal{ST}(\infty)$. As an application, we derive an efficient and simple algorithm to compute them. Moreover, this tree has a special property which we called 'self-reference' from which we can deduce a recursive formula for the cardinality of $\mathcal{SP}(n)$ and some special classes of strict partitions.

First of all, it is easy to see from the construction of $\mathcal{SP}(n+1)$ from $\mathcal{SP}(n)$ that the each node $a \in \bigsqcup_{n\geq 0} \mathcal{SP}(n)$ has at least one child, which is $a^{\downarrow 1}$. Furthermore, if a begins with a slippery plateau of length l, then it has another child which is the element $a^{\downarrow l+1}$. It follows that $\mathcal{ST}(\infty)$ is a binary tree. We will call *left child* the first of two children, and *right child* the other (if it exists). We call the level n of the tree the set of elements of depth n. The first levels of $\mathcal{ST}(\infty)$ are shown in Figure 5.

By using the embedding χ and π in Theorem 4.5, we have:

Proposition 4.7. *The level n of $\mathcal{ST}(\infty)$ is exactly the set of the elements of $\mathcal{SP}(n)$. Moreover the set $\mathcal{SP}(n)$ is in a bijection with a subtree of $\mathcal{ST}(\infty)$ having the same root.*

We will now give a recursive description of $ST(\infty)$. This will allow us to obtain a new recursive formula to calculate the cardinality of $\mathcal{SP}(n)$, as well as for some special classes of strict partitions. We first define a certain kind of subtrees of $ST(\infty)$. Afterward, we show how the whole structure of $\mathcal{ST}(\infty)$ can be described in terms of such subtrees.

We call X_k *subtree* any left subtree of an element beginning with a slippery plateau of length k. Moreover, we define X_0 as a simple node.

The next proposition shows that all the X_k subtrees are isomorphic (see Figure 6).

Proposition 4.8. *A X_k subtree, with $k \geq 1$, is composed by a chain of $k+1$ nodes (the rightmost chain) whose edges are labeled $1, 2, \ldots, k$ and whose i-th node having an out going edge labeled with 1 to a X_i subtree for all i between 1 and k.*

This recursive structure and the above propositions allows us to give a compact representation of the tree $\mathcal{ST}(\infty)$ by a chain (see Figure 7).

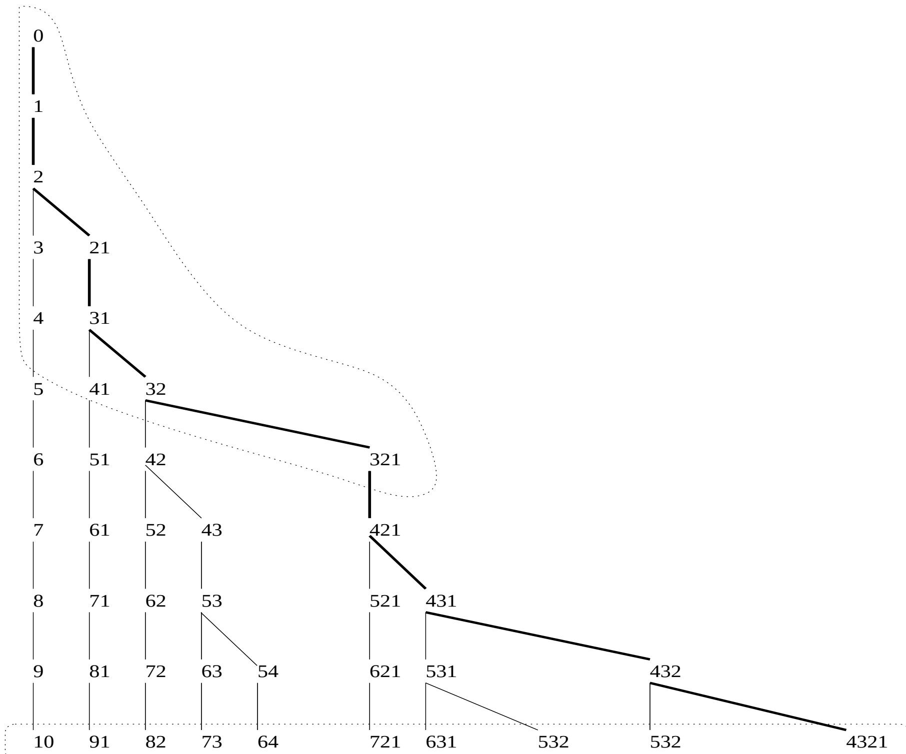


FIGURE 5. The first levels of the tree $\mathcal{ST}(\infty)$ (to clarify the picture, the labels are omitted). As shown on this figure for $n = 10$, we will see two ways to find the elements of $\mathcal{SP}(n)$ in $\mathcal{ST}(\infty)$ for any n.

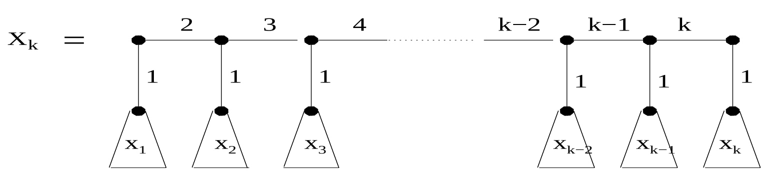


FIGURE 6. Self-referencing structure of X_k subtrees.

Theorem 4.9. *The tree $\mathcal{ST}(\infty)$ can be represented by the infinite chain* $(), 1, 2, 21, 31, 32, 321, \ldots, (n-1, n-2, \ldots, 1), (n, n-2, \ldots, 2, 1), \ldots,$ $(n, n-1, \ldots, 3, 2), (n, n-1, \ldots, 3, 2, 1), \ldots$ *with corresponding edges* $1, 1, 2, 1, 2, 3, \ldots, 1, 2, \ldots, n, \ldots$*; each node before an edge k having an out going edge labeled with 1 to the root of a X_{k-1} subtree.*

Moreover, we can prove a stronger property of each subtree in this chain:

Theorem 4.10. *The subtree (of the form $(k, k-1, \ldots 2, 1) \xrightarrow{1} X_k$) of $\mathcal{ST}(\infty)$ contains exactly the partitions of length k.*

Different to the case of infinite tree of partitions, the distance of this root to the root of $\mathcal{ST}(\infty)$ is equal to $\frac{k(k-1)}{2}$. We can now state our last result:

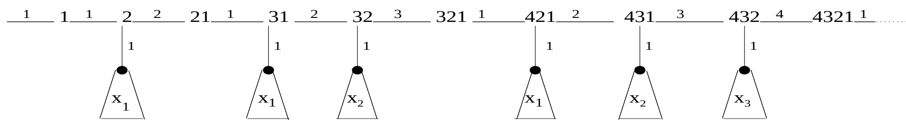


FIGURE 7. Representation of $\mathcal{ST}(\infty)$ as a chain.

Theorem 4.11. *Let $c(l,k)$ denote the number of paths in a X_k tree originating from the root and having length l. We have:*

$$c(l,k) = \begin{cases} 1 & \text{if } l = 0 \text{ or } k = 1 \\ \sum_{i=1}^{inf(l,k)} c(l-i,i) & \text{otherwise} \end{cases}$$

Moreover, $|\mathcal{SP}(n)| = \sum_{0 \le k \le \sqrt{(2n)}+1} c(n - \frac{k(k-1}{2}, k)$ and the number of partitions of n with length exactly k is $c(n - \frac{k(k-1)}{2}, k)$.

□

References

[And76] G.E. Andrews. *The Theory of Partitions.* Addison-Wesley Publishing Company, 1976.

[BL92] A. Bjorner and L. Lovász. *Chip firing games on directed graphes.* Journal of Algebraic Combinatorics, 1:305–328, 1992.

[BLS91] A. Bjorner, L. Lovász, and W. Shor. *Chip-firing games on graphes.* European J. Combin., 12:283–291, 1991.

[Bry73] T. Brylawski. *The lattice of interger partitions.* Discrete Mathematics, 6:201–219, 1973.

[GK86] C. Greene and D.J. Kleiman. *Longest chains in the lattice of integer partitions ordered by majorization.* European J. Combin., 7:1–10, 1986.

[GK93] E. Goles and M.A. Kiwi. *Games on line graphes and sand piles.* Theoret. Comput. Sci., 115:321–349, 1993.

[GLM⁺ar] E. Goles, M. Latapy, C. Magnien, M. Morvan, and H. D. Phan. *Sandpile models and lattices: a comprehensive survey.* Theoret. Comput. Sci., 2001. To appear.

[GMP02] E. Goles, M. Morvan, and H.D. Phan. *The structure of linear chip firing game and related models.* Theoret. Comput. Sci., 270:827–841, 2002.

[Las] A. Lascoux. *Sylvester's bijection between strict and odd partitions.* Preprint, http://phalanstere.univ-mlv.fr/∼al.

[LP99] M. Latapy and H.D. Phan. *The lattice of integer partitions and its infinite extension.* To appear in LNCS special issue, proceedings of ORDAl '99.

[LP01] M. Latapy and H.D. Phan. *The lattice structure of chip firing games.* Physica D, 115:69–82, 2001.

[Syl73] J. J. Sylvester. *Three acts, and interact and an exodion.* Collected Work, IV:1–96, 1973.

Current address: Université de Lille I, UFR de Mathématiques, UMR 8524, 59655 Villeneuve d'Ascq Cédex, France
E-mail address: Minh-Ha.Le@math.univ-lille1.fr

LIAFA UNIVERSITÉ DENIS DIDEROT, PARIS 7 - CASE 7014-2, PLACE JUSSIEU-75256 PARIS CEDEX 05-FRANCE
E-mail address: phan@liafa.jussieu.fr

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Counting Unrooted Loopless Planar Maps

Valery A. Liskovets and Timothy R. Walsh

Abstract. *We present a formula for the number of n-edge unrooted loopless planar maps considered up to orientation-preserving isomorphism. The only sum contained in this formula is over the divisors of n.*

RÉSUMÉ. *Nous présentons une formule pour le nombre de cartes planaires sans boucles avec n arêtes, à isomorphisme près préservant l'orientation. La seule somme contenue dans cette formule est prise parmi les diviseurs de n.*

1. Introduction

At the end of the 1970s the first-named author developed a general method of counting planar maps up to orientation-preserving isomorphism ("unrooted") which is based on using quotient maps [**Li81**] (cf. also [**Li85, Li98**]). It results in a formula which represents the number of unrooted planar n-edge maps of a given class in terms of the numbers of rooted maps of the same class and of their quotient maps with respect to orientation-preserving isomorphism. Based on Burnside's (orbit counting) lemma, this reductive formula contains a sum over the orders of automorphisms of the maps under consideration; as a rule, these are the divisors of n. Generally the formula may contain other summations and need not be very simple since quotient maps may form a fairly complicated class of maps.

Until now, this method was applied successfully to several natural classes of planar maps. Namely, simple formulae have been obtained for counting all maps, homogeneous maps and so called strongly self-dual maps; this last formula contains no sums [**Li81**]. (We add also that two related problems were solved in [**BoLL00, B-MS00**]. Moreover, a formula of this kind has been obtained for the first time in another way in [**Wk72**] for plane trees. It is a particular case of the formula for homogeneous maps, and in [**BnBLL00**] it was generalized to planar m-ary cacti.) Later on, we applied this method to obtain similar formulae for non-separable maps [**LiW83**] (see also [**LiW87**]) and for eulerian and unicursal planar maps [**LiW02**]. All these classes have a remarkable property in common: the number of rooted maps in them is expressed by a simple sum-free formula. This feature immediately implies a simple explicit form of the formula for counting unrooted maps in the simplest cases when the class of quotient maps coincides, or almost coincides, with the initial class. These cases include in particular the types of maps considered in [**Li81**]. However for non-separable, eulerian and unicursal maps the quotient maps are not identical, or even nearly so, to the original maps; so we cannot assume a priori that the corresponding rooted quotient maps are also enumerated by

2000 *Mathematics Subject Classification.* Primary 05C30.

Key words and phrases. Lagrange inversion, rooted planar maps, loopless map, quotient map.

The first author was supported in part by the INTAS under grant INTAS-BELARUS 03-50-5975.

The second author was supported by National Science and Engineering Research Council (NSERC) under grant RGPIN-44703.

simple sum-free formulae which eliminate additional sums and auxiliary terms in the formula for unrooted maps and simplify it significantly. Nevertheless, quite unexpectedly, this property is shared by all the cases considered so far and we have found, for the corresponding unrooted maps, counting formulae that contain only a sum over the divisors of n and a bounded number of additional terms.

The aim of the present article is to investigate one more natural class of maps, loopless maps, which have the same property with respect to rooted enumeration and to establish the existence of a similar simple formula for the number of unrooted maps in it. Again, we do not have any direct explanation for this phenomenon.

Loopless maps have attracted much attention in enumerative combinatorics. Let $L'(n)$ be the number of rooted loopless planar maps with n edges. It was shown in [**WlL75**] and in several more recent publications (see, in particular, [**Wo80, BeW85**]) that

$$L'(n) = \frac{2(4n+1)!}{(n+1)!(3n+2)!} = \frac{2(4n+1)}{(n+1)(3n+1)(3n+2)}\binom{4n}{n}, \quad n \geq 0. \tag{1.1}$$

Let $L^+(n)$ denote the number of unrooted loopless planar maps with n edges counted up to orientation-preserving isomorphism. In this article we prove

Theorem 1.1. *For $n \geq 1$,*

$$\begin{aligned} L^+(n) &= \frac{1}{2n}\Bigg[L'(n) + \sum_{t<n,\, t|n} \phi\left(\frac{n}{t}\right)\frac{(t+1)(3t+1)(3t+2)}{2(4t+1)}\,L'(t) \\ &+ \begin{cases} n^2\, L'\left(\dfrac{n-1}{2}\right) & \text{if } n \text{ is odd} \\ \dfrac{4n(n-1)(2n-1)}{3(3n+2)}\, L'\left(\dfrac{n-2}{2}\right) & \text{if } n \text{ is even} \end{cases}\Bigg], \end{aligned} \tag{1.2}$$

where $\phi(n)$ is the Euler totient function.

Substituting (1.1) into formula (1.2) we can represent it in the following explicit form:

Corollary 1.2.

$$\begin{aligned} L^+(n) &= \frac{1}{2n}\Bigg[\frac{2(4n+1)}{(n+1)(3n+1)(3n+2)}\binom{4n}{n} + \sum_{t<n,\, t|n} \phi\left(\frac{n}{t}\right)\binom{4t}{t} \\ &+ \begin{cases} \dfrac{2n}{n+1}\dbinom{2n}{\frac{n-1}{2}} & \text{if } n \text{ is odd} \\ \dbinom{2n}{\frac{n-2}{2}} & \text{if } n \text{ is even} \end{cases}\Bigg]. \end{aligned} \tag{1.3}$$

The article is organized as follows. Section 2 contains a general description of planar maps, their automorphisms and quotient maps. Section 3 contains a general "reductive" enumerative formula for loopless planar maps and a description of their quotient maps. These quotient maps are enumerated in Section 4. From these results, formula (1.2) is derived in Section 5, which also includes a table of values and some open questions.

2. Maps and quotient maps

A *map* is a 2-cell imbedding of a connected planar graph in a closed orientable surface; if the surface is a sphere, then the map is *planar*. A well-known combinatorial model of maps on an orientable surface represents a map as a pair of permutations (σ, α) acting on a finite set D of *darts* or *edge-ends* such that α is a fixed-point-free involution and the group generated by σ and α is transitive on D. The vertices, edges and faces are, respectively, the cycles of σ, α and $\sigma\alpha$; σ corresponds to counter-clockwise rotation around a vertex from one dart to the next, α corresponds to going from one end of an edge to the other, and $\sigma\alpha$ corresponds to walking clockwise from one edge to the next around the boundary of a face. A map is planar if it satisfies Euler's formula:

$$\#(vertices) + \#(faces) - \#(edges) = 2. \tag{2.1}$$

In what follows, a map is assumed to be planar. An *automorphism* of a combinatorial map is a permutation of D that commutes with σ and α; it corresponds to an orientation-preserving homeomorphism of a (topological) map. Topological and combinatorial models of maps are known to be equivalent (see [**JoS78**]); we will need them both.

A map is *rooted* by distinguishing a dart as the *root*. It was shown in [**Tu63**] (and follows easily from the combinatorial model) that only the trivial automorphism of a planar map fixes the root. Consequently, rooted maps can be counted without considering their symmetries. By *counting unrooted maps* we mean counting isomorphism classes of maps (with respect to orientation-preserving isomorphism).

The method developed in [**Li81, Li85**] (and slightly simplified and modified by form in [**Li98**]) makes it possible to count unrooted maps of classes more complex than plane trees. It relies on constructing and counting quotient maps and uses significantly the familiar property that for any non-trivial orientation-preserving automorphism ρ of a map Γ, the map can be drawn on the sphere so that ρ represents a (geometrical) *rotation* of the sphere about a well-defined axis which intersects the map in two *elements* (vertices, edges or faces) called *axial*, which, for the sake of brevity, we call the *poles* (see loc. cit. for the necessary references). Geometrically, the points of intersection of the axis with edges and faces are their midpoints. On the other hand, as follows from the combinatorial model (the transitivity property), any automorphism of a map is *regular* - all the dart-cycles are of the same length. There is a bijection between the maps fixed by an automorphism and the isomorphic submaps into which the automorphism divides the maps, and this fact provides a way for counting unrooted maps using Burnside's lemma.

Given a map Γ and a non-trivial (orientation-preserving) automorphism ρ of it which is presented geometrically as a rotation and determined by the pair of poles, the order $p \geq 2$ (the period of rotation) and the angle of rotation $2\pi k/p$ (where $k,\ 1 \leq k < p$, is prime to p), the *quotient map* Δ of Γ with respect to ρ is constructed by cutting the sphere into p identical sectors whose common edge is the axis of rotation, choosing one of those sectors, expanding it into a sphere and closing it. In fact, Δ depends only on the cyclic group generated by ρ. If a pole of Γ is an edge, then it turns into a "half-edge" in Δ, that is into an edge which contains a single dart; so an additional vertex of valency 1 (endpoint), called a *singular vertex*, is created. A singular vertex contains no darts and it is identified with the corresponding pole. If Δ contains one or two singular vertices, then $p = 2$. If Γ is rooted, then among the p sectors we choose the one that contains the root, so that Δ is also rooted.

We define a *q-map* to be a planar map with 0, 1 or 2 vertices of valency 1 distinguished as singular vertices and two elements distinguished as axial (poles) which are either vertices or faces and must include all the singular vertices. Given a q-map Δ and an integer $p \geq 2$, the map Γ and the pair of poles such that Δ is the quotient map with respect to an automorphism of order p about an axis intersecting that pair of poles can be retrieved by a process called *lifting*: a semicircular cut whose diameter intersects the two poles is made in the sphere containing Δ, the sphere is then shrunk into a sector of dihedral angle $2\pi/p$, any singular vertex (if any for $p = 2$) is deleted leaving its incident edge with a single dart, and p copies of this sector are pasted together to make a sphere containing Γ. If Δ is rooted, then the root of one of these copies is chosen to be the root of Γ.

3. The quotient map of a loopless map

A map is called *loopless* if its graph does not contain loops. Below we give a construction for a quotient map of a loopless map. Let $L'_0(n)$, $L'_1(n)$ and $L'_2(n)$ be the number of rooted n-edge q-maps with 0, 1 or 2 singular vertices, respectively, whose liftings are rooted loopless maps. The following formula is a direct consequence of the general enumerative scheme of [**Li81, Li85**] described in the form presented in [**Li98**, Sect. 8.7]:

Proposition 3.1.

$$2nL^+(n) = L'(n) + \sum_{t<n,\, t|n} \phi\left(\frac{n}{t}\right) L'_0(t) + \begin{cases} L'_1\big((n+1)/2\big) & \textit{if } n \textit{ is odd} \\ L'_2\big((n/2)+1\big) & \textit{if } n \textit{ is even.} \end{cases} \tag{3.1}$$

□

Each term in the sum in (3.1) is contributed by the automorphisms of order $p = n/t$ and the factor $\phi(n/t)$ is the number of such automorphisms. Below we prove (1.2) by finding expressions for $L'_i(n)$, $i = 0, 1, 2$, which are sums of one or two terms, and substituting them into (3.1).

To evaluate $L'_i(n)$, $i = 0, 1, 2$, we must consider two cases: either the quotient map has no loops or it has at least one loop. The former case is easily tractable by adding 0, 1 or 2 singular vertices to a rooted loopless map; the latter requires a characterization of a rooted q-map that has at least one loop but is lifted into a rooted map with no loops.

Lemma 3.2. *A loop ℓ in a q-map Δ is destroyed by lifting if and only if ℓ separates the poles of Δ.*

PROOF. Suppose that the loop ℓ separates the poles. Then ℓ can be drawn as a circle that separates the poles. The cut made as the first step in lifting Δ (see Section 2) will intersect ℓ and it can be arranged not to intersect the vertex v incident to ℓ. The sector will contain v with one dart of ℓ on either side of it. When p such sectors are pasted together, each one will have a copy of v, and adjacent copies of v will be joined by a *link* (non-loop edge) consisting of one dart of ℓ from one sector and the other dart of ℓ from the adjacent sector. The loop ℓ will thus be replaced by p links.

Suppose that ℓ does not separate the poles. Then the cut can be arranged not to intersect ℓ. The sector will have the loop ℓ and the lifted map will have p loops. □

Lemma 3.3. *Lifting a q-map Δ creates a loop if and only if one pole of Δ is a singular vertex v and the other pole is the vertex adjacent to v.*

PROOF. Suppose that one pole is a singular vertex v and the other pole is the vertex adjacent to v. The sector will contain v and a single dart d. Since v is a pole, the lifted map will have a single copy of v with two copies of d joined together into a single edge, a loop incident to v.

Suppose that Δ has no singular vertex. Then each link of Δ is lifted into p links of Γ. Now suppose that Δ has a singular vertex, which must of necessity be a pole. If the adjacent vertex v is not the other pole, then it will be lifted into two vertices joined together by the link whose darts are the two copies of the edge joining v to the singular vertex of Δ. In either case, no loop will be created. □

Theorem 3.4. *A map Γ lifted from a q-map Δ is loopless if and only if Δ satisfies the following two conditions:*

(1) *if Δ has loops, each of them separates the poles, and*
(2) *if one pole is a singular vertex s, then the other pole is not the vertex adjacent to s.*

PROOF. An easy consequence of Lemmas 3.2 and 3.3. □

Definition 3.5. An *ℓ-map* is a q-map which has at least one loop but whose liftings have no loops.

We present a construction of an ℓ-map Δ, an analogue of the s-map of [**LiW83**] that consists of a chain of blocks.

Suppose that Δ contains $k-1$ loops, $k > 1$. Arbitrarily call one pole the outer pole and the other one the inner pole. By Theorem 3.4, condition (1), the $k-1$ loops are all nested one inside the other in linear order $\ell_1, \ell_2, \ldots, \ell_{k-1}$, with the outer pole strictly outside the outermost loop ℓ_1, and therefore not the vertex incident to ℓ_1 and the inner pole strictly inside the innermost loop ℓ_{k-1}, and therefore not the vertex incident to ℓ_{k-1}. The outer pole belongs to the submap M_1 of Δ that is outside of ℓ_1 (if the outside of ℓ_1 is empty, then M_1 is just the vertex-map). The inner pole belongs to the submap M_k that is inside of ℓ_{k-1}. For $i = 2, 3, ..., k-1$ we denote by M_i the submap that is inside of ℓ_{i-1} but outside of ℓ_i. All the M_i are loopless. We call the M_i the *components* of Δ; M_1 and M_k are the *extremal* components and the other components are the *internal* ones.

An example of a loopless map and its quotient map, which is an ℓ-map, is depicted in Fig. 1.

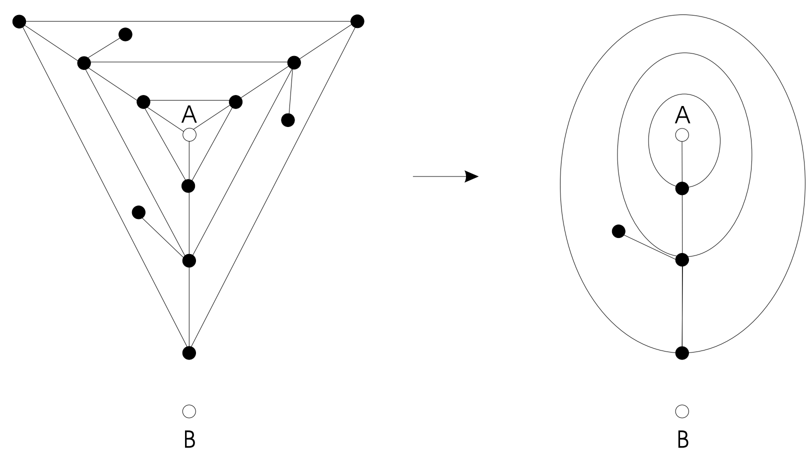


FIGURE 1. A loopless map (left) and its quotient map (right) with respect to rotations of order 3 around the axis AB (where B is the midpoint of the outer face)

We construct Δ from the outside in. We begin with a loopless map M_1. We insert an empty loop ℓ_1 into M_1, at the one vertex of M_1 if M_1 is a vertex-map or between two consecutive darts of a vertex of M_1 otherwise. After this insertion, one of the darts d of ℓ_1 will still have the property that $\sigma(d) = \alpha(d)$ (a counter-clockwise rotation about the vertex incident to ℓ_1 starting at d traverses the empty inside of ℓ_1 and then encounters the other dart of the loop) - we call d the *right* dart of ℓ_1 and $\alpha(d)$ its *left* dart. Then a rooted loopless map M_2 is inserted into ℓ_1 so that the root of M_2 (if M_2 is not a vertex-map) becomes $\sigma(d)$. If $k > 2$, then another empty loop ℓ_2 is inserted into M_2 and another rooted loopless map M_3 inserted into ℓ_2 and so on until the innermost loop ℓ_{k-1} has been inserted into M_{k-1} and the innermost rooted loopless map M_k has been inserted into ℓ_{k-1}. If Δ is not to have any singular vertices, then the outer pole is chosen to be some vertex or face of M_1 but not the vertex incident with ℓ_1 and the inner pole is chosen to be some vertex or face of M_k but not the vertex incident with ℓ_{k-1}. The modification of this construction to account for singular vertices is discussed in Sections 4.3 and 4.4.

4. Enumeration of quotient maps of rooted loopless maps

4.1. Enumeration of rooted ℓ-maps by the sizes of the extremal components. We now proceed to enumerate rooted ℓ-maps with n edges and no singular vertices such that M_1 has a edges, M_k has b edges, and for $2 \le i \le k-1$, M_i has n_i edges, so that $n_2 + \cdots + n_{k-1} = n - (a + b) - (k - 1)$. For the moment we distinguish the poles as outer and inner (to distinguish between M_1 and M_k) but in this subsection we do not include the number of choices of poles in the enumeration formulae.

Suppose for the moment that the root of Δ belongs to M_1 (if M_1 is a vertex-map, then the root is the left dart of ℓ_1). If M_1 is not a vertex-map, then there are $2a$ places to insert ℓ_1; otherwise there is one place. For $2 \le i \le k$ there is one place to insert M_i into ℓ_{i-1}. For $2 \le i \le k-1$ there are $2n_i + 1$ places to insert

ℓ_i into M_i: for any dart d of M_i, ℓ_i can be inserted between d and $\sigma(d)$, or else ℓ_i can be inserted between the root of M_i and the right dart of ℓ_{i-1}. The number of ℓ-maps whose root belongs to M_1 is thus

$$L'(a)L'(b)\prod_{i=2}^{k-1}(2n_i+1)L'(n_i)\cdot\begin{cases}2a & \text{if } a>0\\ 1 & \text{if } a=0.\end{cases} \tag{4.1}$$

Before continuing with the enumeration we formally state a folkloric lemma and provide two proofs, special cases of which appear in numerous places in the literature.

Lemma 4.1 (the Little Labeling Lemma)**.** *Suppose that there are two sorts of labels for a combinatorial object, each with the property that only the trivial automorphism preserves the labels. If the object can be labeled in x ways with labels of the first sort and y ways with labels of the second sort, then the numbers x' and y' of equivalence classes of labelings of the two sorts, where two labelings are equivalent if the object has an automorphism taking one set of labels into the other, are in the same proportion $x : y$. This proportion extends by summation to any set of objects with the same proportion $x : y$ of ways of labeling them with labels of the two sorts.*

PROOF. *Proof 1:* Let A be the number of automorphisms of the object. Then $x' = x/A$ and $y' = y/A$.

Proof 2: We count the number of inequivalent ways to apply labels of both sorts at once. Since either sort of labeling destroys all non-trivial automorphisms, once labels of one sort have been applied, all the ways of applying labels of the other sort are inequivalent. There are thus $x'y$ inequivalent ways to apply labels of the first sort followed by labels of the second sort and $y'x$ ways to apply labels of the second sort followed by labels of the first sort; so $x'y = y'x$. □

Resuming the enumeration, suppose now that the root can be any dart of the map. Then the factor $2a$ or 1 of (4.1) is replaced by $2n$ (here we are using Lemma 4.1, where one sort of labeling is a rooting in M_1 and the other sort is a rooting anywhere in the map) provided that the poles are actually distinguished as outer and inner. If $a \neq b$, we can call the outer pole the one that belongs to the component with more edges by insisting that $a > b$. If $a = b$, then the distinction between the poles remains arbitrary; removing it is equivalent to dividing the number of rooted ℓ-maps by 2 (here we are using Lemma 4.1, where one sort of labeling includes distinguishing the poles as well as rooting the map and the other sort does not).

Let $C'(a, b; n)$ be the number of rooted n-edge ℓ-maps whose extremal components have a and b edges (we no longer distinguish the poles as inner and outer). Applying to (4.1) the discussion of the previous paragraph and then summing first over all sequences of $n_2, \ldots, n_{k-1}$ which add to $n - (a + b) - (k - 1)$ and then over k from 2 to $n - 2$ we obtain

$$C'(a,b;n) = nL'(a)L'(b)\sum_{k=2}^{n-2}\sum_{\substack{n_2+\cdots+n_{k-1}\\=n-(a+b)-(k-1)}}\prod_{i=2}^{k-1}(2n_i+1)L'(n_i)\cdot\begin{cases}2 & \text{if } a\neq b\\ 1 & \text{if } a=b.\end{cases} \tag{4.2}$$

Let

$$g(x) = \sum_{n=0}^{\infty} L'(n)x^n. \tag{4.3}$$

It was shown in [**WlL75**] that

$$g(x) = 1 + z - z^2 - z^3, \tag{4.4}$$

where

$$z = x(1 + z)^4. \tag{4.5}$$

We will use these formulae repeatedly. By differentiating (4.3) we find that

$$\sum_{n=0}^{\infty}(2n+1)L'(n)x^n = 2xg'(x) + g(x), \tag{4.6}$$

We evaluate $g'(x)$ by differentiating (4.4) with respect to z and then dividing by dx/dz as evaluated from (4.5) and then we multiply by x, again from (4.5), and simplify to obtain

$$xg'(x) = z(1+z)^2. \tag{4.7}$$

Substituting from (4.7) and (4.4) and simplifying, we obtain

$$2xg'(x) + g(x) = (1+z)^3. \tag{4.8}$$

We denote by $[x^n]\,f$ the coefficient of x^n in the power series f. Substituting from (4.8) and (4.6) we find that the inner sum in (4.2) is $[x^{n-(a+b)-(k-1)}]\,(1+z)^{3(k-2)} = [x^{n-(a+b)-1}]\,x^{k-2}(1+z)^{3(k-2)}$, so that the outer sum is

$$[x^{n-(a+b)-1}]\,(1-x(1+z)^3)^{-1}. \tag{4.9}$$

Substituting from (4.5) for x into (4.9), simplifying and substituting into (4.2), we find that

$$C'(a,b;n) = nL'(a)L'(b)\cdot[x^{n-(a+b)-1}]\,(1+z)\cdot\begin{cases} 2 & \text{if } a \neq b \\ 1 & \text{if } a = b. \end{cases} \tag{4.10}$$

We could use Lagrange inversion [**La81**] to evaluate $C'(a,b;n)$ explicitly but we do not need that formula in what follows.

In the rest of Section 4 the choice of poles will be included in the enumeration formulae.

4.2. No singular vertices. A pole of an ℓ-map can be any vertex or face of an extremal component except the vertex the component shares with a loop. If the component has m edges, then by Euler's formula (2.1) there are a total of $m+1$ vertices and faces, not counting the forbidden vertex; so the number of rooted ℓ-maps with n edges and no singular vertices is

$$\sum_{\substack{a\geq b\geq 0\\ a+b\leq n}} (a+1)(b+1)C'(a,b;n), \tag{4.11}$$

which, by (4.10), is equal to

$$n\sum_{\substack{a\geq b\geq 0\\ a+b\leq n}} (a+1)L'(a)(b+1)L'(b)\cdot[x^{n-(a+b)-1}]\,(1+z). \tag{4.12}$$

In a manner similar to the derivation of (4.6) and (4.8) we find that

$$\sum_{a=0}^{\infty}(a+1)L'(a)x^a = xg'(x) + g(x) = (1+z)^2. \tag{4.13}$$

Substituting from (4.13) into (4.12) and simplifying, we obtain

$$n\cdot[x^{n-1}]\,(1+z)^5. \tag{4.14}$$

The derivative of $(1+z)^5$ is $5(1+z)^4$; so by Lagrange inversion (4.14) is equal to

$$5\frac{n}{n-1}\cdot[z^{n-2}]\,(1+z)^{4n} = \frac{5n}{n-1}\binom{4n}{n-2} = \frac{5n(4n)!}{(n-1)!(3n+2)!}. \tag{4.15}$$

Comparing the right side of (4.15) with (1.1), we express the number of rooted n-edge ℓ-maps as a multiple of $L'(n)$:

$$\frac{5nL'(n)}{4n+1}\binom{n+1}{2}. \tag{4.16}$$

The number of rooted loopless n-edge q-maps is

$$\binom{n+2}{2} L'(n) \tag{4.17}$$

because by Euler's formula (2.1) there are a total of $n+2$ faces and vertices and any pair can be chosen to be the poles.

Adding (4.16) and (4.17) we find that

$$L_0'(n) = \frac{(n+1)(3n+1)(3n+2)}{2(4n+1)} L'(n) = \binom{4n}{n} \tag{4.18}$$

(the last equality is obtained by using (1.1)).

4.3. One singular vertex. We construct a rooted ℓ-map with n edges and one singular vertex by taking a rooted ℓ-map with $n-1$ edges and no singular vertices, inserting a singular vertex and its incident edge into M_1 (which has a edges) making the singular vertex the outer pole and choosing one of the $b+1$ possible inner poles in M_k which has b edges. There are $2a+1$ slots into which to insert the dart opposite the singular vertex: as $\sigma(d)$, where d can be either any dart of M_1 or else the right dart of ℓ_1. This augmented ℓ-map has $2n-1$ darts that can be the root, as opposed to $2n-2$ for the original ℓ-map. To get the number of maps we substitute $n-1$ for n in (4.10), multiply by $(2n-1)/(2n-2)$ to account for the extra possible root (by Lemma 4.1), by $2a+1$ to account for the insertions and by $b+1$ to account for the inner pole, and then sum over a and b. We obtain

$$\frac{2n-1}{2} \sum_{\substack{a,b\geq 0\\ a+b\leq n-1}} (2a+1)L'(a)(b+1)L'(b) \cdot [x^{n-(a+b)-2}]\,(1+z) \tag{4.19}$$

$$= \frac{2n-1}{2} \cdot [x^{n-2}]\,(2xg'(x)+g(x))\,(xg'(x)+g(x))\,(1+z). \tag{4.20}$$

Substituting from (4.8) and (4.13) into (4.20) and simplifying we obtain

$$\frac{2n-1}{2} \cdot [x^{n-2}]\,(1+z)^6. \tag{4.21}$$

The derivative of $(1+z)^6$ is $6(1+z)^5$; so by Lagrange inversion, (4.21) is equal to

$$\frac{3(2n-1)}{n-2} \cdot [z^{n-3}]\,(1+z)^{4n-3} = \frac{3(2n-1)}{n-2}\binom{4n-3}{n-3}. \tag{4.22}$$

Comparing the right side of (4.22) with (1.1), we find that the number of rooted ℓ-maps with n edges and one singular vertex is

$$(n-1)(2n-1)L'(n-1). \tag{4.23}$$

We construct a rooted loopless q-map with n edges and one singular vertex by taking a rooted loopless map with $n-1$ edges, inserting a vertex of valency 1 and its incident edge into one of the $2(n-1)$ possible slots, making the singular vertex one pole, choosing another pole and letting the set of possible roots include the dart opposite the singular vertex. The number $L'(n-1)$ gets multiplied by $2n-2$ for the insertions, by $(2n-1)/(2n-2)$ for the extra possible root (by Lemma 4.1), and by n for the choice of the second pole: by Euler's formula (2.1) there are a total of $n+1$ vertices and faces aside from the first pole, but by Theorem 3.4, condition (2), the vertex adjacent to it is ineligible to be a pole. The number of rooted loopless q-maps with n edges and one singular vertex is thus

$$n(2n-1)L'(n-1). \tag{4.24}$$

By adding (4.23) and (4.24), we find that

$$L_1'(n) = (2n-1)^2 L'(n-1). \tag{4.25}$$

4.4. Two singular vertices. We construct a rooted ℓ-map with two singular vertices by taking a rooted ℓ-map with $n-2$ edges and no singular vertices, inserting a singular vertex and its incident edge into M_1 (which has a edges) and another one into M_k (which has b edges), making each of these singular vertices a pole and allowing the set of possible roots to include the darts opposite both singular vertices. There are $2a+1$ possible insertions into M_1 and $2b+1$ possible insertions into M_k. To get the number of maps, we substitute $n-2$ for n in (4.10), multiply by $(2a+1)(2b+1)$ to account for the insertions, multiply by $(2n-2)/(2n-4)$ to account for the two extra possible roots (by Lemma 4.1) and sum over a and b. We obtain

$$(n-1) \sum_{\substack{a,b \geq 0 \\ a+b \leq n-2}} (2a+1)L'(a)(2b+1)L'(b) \cdot [x^{n-(a+b)-3}]\,(1+z) \tag{4.26}$$

$$= (n-1) \cdot [x^{n-3}]\,\left(2xg'(x)+g(x)\right)^2 (1+z). \tag{4.27}$$

Substituting from (4.8) into (4.27) and simplifying we obtain

$$(n-1) \cdot [x^{n-3}]\,(1+z)^7. \tag{4.28}$$

The derivative of $(1+z)^7$ is $7(1+z)^6$; so, by Lagrange inversion, (4.28) is equal to

$$\frac{7(n-1)}{n-3} \cdot [z^{n-4}]\,(1+z)^{4n-6} = \frac{7(n-1)}{n-3}\binom{4n-6}{n-4}. \tag{4.29}$$

We construct a rooted loopless q-map with n edges and two singular vertices by taking a rooted loopless map with $n-2$ edges and inserting two singular vertices and their incident edges into $2n-4$ possible slots, making both the singular vertices poles, and allowing the set of possible roots to include the darts opposite the two singular vertices. The number $L'(n-2)$ gets multiplied by $(2n-2)/(2n-4)$ to account for the two extra possible roots (by Lemma 4.1); to account for the insertions it gets multiplied by $(2n-4)(2n-3)/2$ instead of $(2n-4)(2n-5)/2$ because both opposite darts can be inserted into the same slot. The number of rooted loopless q-maps with n edges and two singular vertices is thus

$$(n-1)(2n-3)L'(n-2). \tag{4.30}$$

Adding (4.29) and (4.30) and comparing with (1.1) we get two expressions for $L'_2(n)$:

$$L'_2(n) = \frac{4(n-1)(2n-3)(4n-5)}{3(3n-2)}\,L'(n-2) = \binom{4n-4}{n-2}, \tag{4.31}$$

and we keep them both because the one that is not a multiple of $L'(n-2)$ is simpler.

5. The result. Discussion

Substituting from (4.18), (4.25) and (4.31) into (3.1) we obtain (1.2), thus proving Theorem 1.1. □

Table 1 contains the values of $L'(n)$ and $L^+(n)$ for $0 \leq n \leq 20$. These latter values were verified for up to 7 edges by comparison with the number of unrooted loopless maps generated by computer [**Wl83**].

We note here that there is another way to derive formula (1.2): we express an ℓ-map as a chain of blocks, at least one of which is a loop, whose extremal components contain the poles as internal elements, with a rooted loopless map inserted between each pair of darts d, $\sigma(d)$.

An interesting open problem would be a proof of formulae (1.2) or (1.3) (and the analogous formulae for unrooted non-separable, eulerian and unicursal maps) that involves natural bijections instead of Lagrange inversion, thus possibly explaining the absence of a rational factor to be multiplied by $\binom{4t}{t}$ for $t < n$ in (1.3), which is a special case of a general phenomenon discussed in more detail in [**Li04**]. Another open problem is counting unrooted loopless maps (as well as eulerian and unicursal maps) by number of edges and vertices. This problem is probably easier to solve than the previous one because it has already been solved quite effectively for all maps and non-separable maps by the second-named author [**Wl03**]. In general,

TABLE 1. The number of rooted and unrooted loopless planar maps

#(edges)	#(rooted maps)	#(unrooted maps)
0	1	1
1	1	1
2	3	2
3	13	5
4	68	14
5	399	49
6	2530	240
7	16965	1259
8	118668	7570
9	857956	47996
10	6369883	319518
11	48336171	2199295
12	373537388	15571610
13	2931682810	112773478
14	23317105140	832809504
15	187606350645	6253763323
16	1524813969276	47650870538
17	12504654858828	376784975116
18	103367824774012	2871331929096
19	860593023907540	22647192990256
20	7211115497448720	180277915464664

there is no necessity to restrict oneself to classes of maps for which rooted maps are enumerated by sum-free formulae. For instance, it would be interesting to count unrooted n-edge planar maps without either loops or isthmuses; for counting such rooted maps see [**WlL75**].

References

[BeW85] E. A. Bender and N. C. Wormald, *The number of loopless planar maps*, Discr. Math. **54**:2 (1985), 235–237.

[BnBLL00] M. Bóna, M. Bousquet, M., G. Labelle and P. Leroux, *Enumeration of m-ary cacti*, Adv. Appl. Math. **24**:1 (2000), 22–56.

[BoLL00] M. Bousquet, G. Labelle and P. Leroux, *Enumeration of planar two-face maps*, Discr. Math. **222** (2000), 1–25.

[B-MS00] M. Bousquet-Mélou and G. Schaeffer, *Enumeration of planar constellations*, Adv. Appl. Math. **24**:4 (2000), 337–368.

[JoS78] G. A. Jones and D. Singerman, *Theory of maps on orientable surfaces*, Proc. London Math. Soc. (3) **37**:2 (1978), 273–307.

[La81] G. Labelle, *Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange*, Advances in Math. **42** (1981), 217–247.

[Li81] V. A. Liskovets, *A census of non-isomorphic planar maps*, Colloq. Math. Soc. J. Bolyai **25** (1981), Proc. Intern. Conf. *"Algebr. Methods in Graph Th."*, Szeged, Hungary, 1978, 479–494.

[Li85] V. A. Liskovets, *Enumeration of nonisomorphic planar maps*, Selecta Math. Soviet. **4**:4 (1985), 303–323.

[Li98] V. Liskovets, *Reductive enumeration under mutually orthogonal group actions*, Acta Appl. Math. **52** (1998), 91–120.

[Li04] V. A. Liskovets, *Enumerative formulae for unrooted planar maps: a pattern*, submitted.

[LiW83] V. A. Liskovets and T. R. S. Walsh, *The enumeration of non-isomorphic 2-connected planar maps*, Canad. J. Math. **35**:3 (1983), 417–435.

[LiW87] V. A. Liskovets and T. R. S. Walsh, *Ten steps to counting planar graphs*, Congr. Numer. **60** (1987), 269–277.

[LiW02] V. A. Liskovets and T. R. S. Walsh, *Enumeration of eulerian and unicursal planar maps*, Discr. Math., to appear. (See also preliminary versions in: Proc. 14th Intern. Conf. *"Formal Power Series and Algebr. Combin. (FPSAC'02)"*, Univ. of Melbourne, Australia, 2002, 425–433, and Rapport technique No 2002-03, UQAM, Montréal, Canada (2002), 12 p.)

[Tu63] W. T. Tutte, *A census of planar maps*, Canad. J. Math. **15**:2 (1963), 249–271.
[Wk72] W. Walkup, *The number of planar trees*, Mathematika **19**:2 (1972), 200–204.
[Wl83] T. R. S. Walsh, *Generating nonisomorphic maps without storing them*, SIAM J. Alg. Discr. Meth. **4**:2 (1983), 161–178.
[Wl03] T. R. Walsh, *Efficient enumeration of sensed maps*, submitted.
[WlL75] T. R. S. Walsh and A. B. Lehman, *Counting rooted maps by genus. III: Non-separable maps*, J. Combin. Th., Ser. B **18**:3 (1975), 222–259.
[Wo80] N. C. Wormald, *A correspondence for rooted planar maps*, Ars Combin. **9** (1980), 11–28.

Institute of Mathematics, National Academy of Sciences, Minsk, BELARUS, 220072
E-mail address: `liskov@im.bas-net.by`

Département D'Informatique, Université du Québec à Montréal, Montréal (Québec), CANADA, H3C 3P8
E-mail address: `walsh.timothy@uqam.ca`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Descents, Major Indices, and Inversions in Permutation Groups

Anthony Mendes and Jeffrey Remmel

Abstract. *We give a new proof of a multivariate generating function involving the descent, major index, and inversion statistic due to Gessel. We then show how one can easily modify this proof to give new generating functions involving these three statistics over Young's hyperoctahedral group, the Weyl group of type D, and multiples of permutations. All of our proofs are combinatorial in nature and exploit fundamental relationships between the elementary and homogeneous symmetric functions.*

1. Introduction

Let $\sigma = \sigma_1 \cdots \sigma_n$ be an element of the symmetric group S_n written in one line notation. The descent, major index, and inversion statistics are defined by

$$des(\sigma) = \sum_{i=1}^{n-1} \chi(\sigma_{i+1} < \sigma_i),\ maj(\sigma) = \sum_{i=1}^{n-1} i\,\chi(\sigma_{i+1} < \sigma_i), \quad \text{and} \quad inv(\sigma) = \sum_{j<i} \chi(\sigma_i < \sigma_j),$$

where for any statement A, $\chi(A)$ is 1 if A is true and 0 if A is false. These definitions also hold for any finite sequence. The past century has witnessed a beautiful theory develop from the study of these (and other) permutation statistics. To this day, new generalizations and variations of these statistics are investigated. In this work, we will create multivariate generating functions involving the three statistics defined above. They will follow from the combinatorial manipulation of objects arising from fundamental relationships between bases of symmetric functions.

Standard notation from hypergeometric function theory will be used. For $n \geq 1$, $\lambda \vdash n$, and an indeterminate q, let

$$[n]_q = q^0 + \cdots + q^{n-1},\ [n]_q! = [n]_q \cdots [1]_q, \quad \text{and} \quad \begin{bmatrix} n \\ \lambda \end{bmatrix}_q = \frac{[n]_q!}{[\lambda_1]_q! \cdots [\lambda_\ell]_q!}$$

be the q-analogues of n, $n!$, and $\binom{n}{\lambda}$, respectfully. Let $(x;q)_n = (1-xq^0)\cdots(1-xq^{n-1})$. Finally, define a q-analogue of the exponential function such that

$$\exp_q(x) = \sum_{n\geq 0} \frac{x^n}{[n]_q!} q^{\binom{n}{2}}.$$

1991 *Mathematics Subject Classification.* 05A15, 05E05.
Key words and phrases. generating functions, symmetric group, symmetric functions.

In Gessel's thesis and in a paper by Garsia and Gessel [**Ga, Ge**], it was shown that

$$\sum_{n\geq 0} \frac{t^n}{[n]_q!(x;r)_{n+1}} \sum_{\sigma\in S_n} x^{des(\sigma)} r^{maj(\sigma)} q^{inv(\sigma)} = \sum_{k\geq 0} \frac{x^k}{\exp_q(-tr^0)\cdots\exp_q(-tr^k)}. \tag{1.1}$$

This will be our starting point. First, we give a new, combinatorial proof of 1.1. Then, we show how similar proofs indicate a systematic approach to finding more generating functions for permutation statistics. To this end, we highlight some basic facts about the ring of symmetric functions needed for the journey.

The n^{th} elementary symmetric function e_n and the n^{th} homogeneous symmetric function h_n are polynomials in the variables $x_1, x_2, \ldots$ defined to satisfy

$$\sum_{n\geq 0} h_n t^n = \prod_i \frac{1}{1-x_i t} \quad \text{and} \quad \sum_{n\geq 0} e_n t^n = \prod_i (1+x_i t).$$

Therefore,

$$\sum_{n\geq 0} h_n t^n = \prod_i \frac{1}{1-x_i t} = \left(\prod_i 1 + x_i(-t)\right)^{-1} = \left(\sum_{n\geq 0} e_n(-t)^n\right)^{-1}. \tag{1.2}$$

Multiply both sides of 1.2 by the reciprocal of the right hand side and then compare the coefficient of t^n on both sides of the result to see that

$$\sum_{i=0}^{n} (-1)^i e_i h_{n-i} = 0, \quad \text{or equivalently,} \quad h_n = (-1)^{n-1} e_n + \sum_{i=1}^{n-1} (-1)^{i-1} e_i h_{n-i}. \tag{1.3}$$

For a partition λ, e_λ is defined to be $e_{\lambda_1}\cdots e_{\lambda_\ell}$. It is well known that $\{e_\lambda : \lambda \text{ a partition}\}$ is a basis for the ring of symmetric functions [**S**]. A combinatorial interpretation of the expansion of h_n in terms of this basis was first given by Egecioglu and Remmel [**E**]. It is now described as it will be of great use to us.

A rectangle of height 1 and length n chopped into "bricks" of lengths found in the partition λ is known as a brick tabloid of shape (n) and type λ. For example, Figure 1 shows one brick tabloid of shape (12) and type $(2,3,7)$. Let $B_{\lambda,n}$ be the number of such objects. Note that $(-1)^{n-1}B_{(n),(n)} = (-1)^{n-1}$. Furthermore,

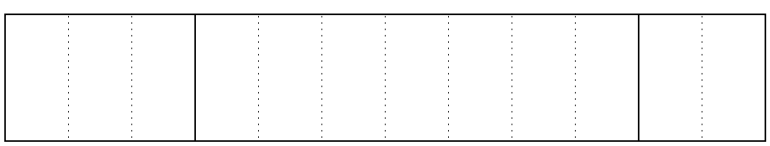


FIGURE 1. A brick tabloid of shape (12) and type $(2,3,7)$.

sorting by the length of the first brick, it may be seen that

$$(-1)^{n-\ell(\lambda)} B_{\lambda,(n)} = \sum_{i=1}^{n-1} (-1)^{i-1} \left((-1)^{(n-i)-(\ell(\lambda)-1)} B_{\lambda\setminus i,(n-i)}\right)$$

where $B_{\lambda\setminus i,(n-i)}$ is defined to be zero if λ does not have a part of size i. These two facts completely determine the numbers $(-1)^{n-\ell(\lambda)}B_{\lambda,(n)}$ recursively. By using 1.3, the coefficient of e_λ in h_n may be shown to satisfy the exact same recursions. Therefore,

$$h_n = \sum_{\lambda\vdash n} (-1)^{n-\ell(\lambda)} B_{\lambda,n} e_\lambda. \tag{1.4}$$

We have now established enough terminology and basic facts to commence our discussion of methods to find generating functions involving the descent, major index, and inversion statistics.

2. A new proof of Gessel's generating function

For $k \geq 0$, define a homomorphism ξ_k on the ring of symmetric functions such that

$$\xi_k(e_n) = \sum_{\substack{i_0,\ldots,i_k \geq 0 \\ i_0+\cdots+i_k=n}} \frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q!\cdots[i_k]_q!} q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}}$$

for indeterminates q and r. Since products of n^{th} elementary symmetric functions form a basis, the definition of ξ_k extends to all other elements in the ring of symmetric functions. In particular, we may apply ξ_k to the n^{th} homogeneous symmetric function.

Theorem 2.1. *For $k, n \geq 0$,*

$$[n]_q!\xi_k(h_n) = \frac{1}{(x;r)_{n+1}} \sum_{\sigma \in S_n} x^{des(\sigma)} r^{maj(\sigma)} q^{inv(\sigma)} \Bigg|_{x^k}$$

where expression$|_x$ denotes the coefficient of x in expression.

Proof. Expand h_n in terms of the elementary symmetric functions by 1.4:

$$\begin{aligned}
[n]_q!\xi_k(h_n) &= [n]_q! \sum_{\lambda \vdash n} (-1)^{n-\ell(\lambda)} B_{\lambda,n} \xi_k(e_\lambda) \\
&= [n]_q! \sum_{\lambda \vdash n} (-1)^{n-\ell(\lambda)} B_{\lambda,n} \prod_{j=1}^{\ell(\lambda)} \sum_{\substack{i_{j,0},\ldots,i_{j,k} \geq 0 \\ i_{j,0}+\cdots+i_{j,k}=\lambda_j}} \frac{r^{0i_{j,0}+\cdots+ki_{j,k}}}{[i_{j,0}]_q!\cdots[i_{j,k}]_q!} q^{\binom{i_{j,0}}{2}+\cdots+\binom{i_{j,k}}{2}}.
\end{aligned}$$

Rewriting q-analogues, the right hand side of the above is equal to

$$(2.1) \qquad \sum_{\lambda \vdash n} \begin{bmatrix} n \\ \lambda \end{bmatrix}_q (-1)^{n-\ell(\lambda)} B_{\lambda,n} \prod_{j=1}^{\ell(\lambda)} \sum_{\substack{i_{j,0},\ldots,i_{j,k} \geq 0 \\ i_{j,0}+\cdots+i_{j,k}=\lambda_j}} \begin{bmatrix} \lambda_j \\ i_{j,0},\ldots,i_{j,k} \end{bmatrix}_q r^{0i_{j,0}+\cdots+ki_{j,k}} q^{\binom{i_{j,0}}{2}+\cdots+\binom{i_{j,k}}{2}}.$$

2.1 may be interpreted as a sum of signed, weighted brick tabloids. After combinatorial objects are described, a sign-reversing, weight-preserving involution will be applied to leave only objects with positive sign. Then, the fixed points will help count the number of permutations with k descents by the major index and inversion statistics.

Start creating combinatorial objects from 2.1 by using the "$\sum_{\lambda \vdash n}$" and the factor of $B_{\lambda,n}$ to give a brick tabloid of shape (n) and type λ for some $\lambda \vdash n$. Let us call this brick tabloid T. The factor of $(-1)^{n-\ell(\lambda)}$ allows for the labeling of each cell not terminating a brick in T with a "-1". In each terminal cell in a brick, place a "1".

For each brick in T, choose nonnegative integers $i_0, \ldots, i_k$ that sum to the total length of the brick. This accounts for the product and second sum in 2.1. Using the power of r, these choices for $i_0, \ldots, i_k$ can be recorded in T. In each brick, place a power of r in each cell such that the powers weakly increase from left to right such that the number of occurrences of r^j will be equal to i_j. At this point, we have constructed T which may look something like Figure 2 below.

-1	-1	1	-1	-1	-1	-1	-1	-1	1	-1	1
r^1	r^1	r^3	r^0	r^0	r^0	r^0	r^2	r^3	r^3	r^1	r^1

Figure 2. One possible T when $k = 3$ and $n = 12$.

The only components in 2.1 which have not been used involve powers of q. We will explain how these powers of q will fill the cells of T with a permutation of n such that a decrease must occur between consecutive cells labeled with the same power of r. Along with this permutation of n, a power of q will be recorded in each cell counting the number of smaller integers in the permutation to the right (given an integer i in a permutation, the number of smaller integers to the right of i is sometimes referred to as the number of inversions caused by i).

In [**C**], Carlitz shows that if $\mathcal{R}(0^{i_0}, \ldots, k^{i_k})$ is the number of rearrangements of i_0 0's, i_1 1's, etc., then

$$\begin{bmatrix} n \\ i_0, \ldots, i_k \end{bmatrix}_q = \sum_{r \in \mathcal{R}(0^{i_0}, \ldots, k^{i_k})} q^{inv(r)}.$$

Thus, the $\begin{bmatrix} n \\ \lambda \end{bmatrix}_q$ term in 2.1 gives a rearrangement of λ_1 0's, λ_2 1's, etc. We will use this to select which integers in a permutation of n will appear in each brick. Start with a brick tabloid T of shape (n) such that the size of the bricks read from left to right are $b_0, \ldots, b_k$. For example, if T is the brick tabloid in Figure 2, then $b_0 = 3$, $b_1 = 7$ and $b_2 = 2$. Then consider a rearrangement r of $0^{b_0}, \ldots, k^{b_k}$ and construct a permutation $\sigma(r)$ by labeling the 0's from left to right with $1, 2, \ldots, b_0$, the 1's from right to left with $b_0 + 1, \ldots, b_0 + b_1$ and in general the i's from right to left with $1 + \sum_{j=1}^{i-1} b_j, \ldots, b_i + \sum_{j=1}^{i-1} b_j$. In this way, $\sigma(r)^{-1}$ starts with the positions of the 0's in r increasing order, followed by the positions of the 1's in r in increasing order, etc. For example, for T pictured in Figure 2, one possible rearrangement to consider is $r = 1\ 0\ 1\ 1\ 2\ 0\ 1\ 2\ 1\ 0\ 1\ 1$. Below we picture $\sigma(r)$ and $\sigma(r)^{-1}$.

$$\begin{array}{lcrrrrrrrrrrrr} r & = & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 2 & 1 & 0 & 1 & 1 \\ \sigma(r) & = & 4 & 1 & 5 & 6 & 11 & 2 & 7 & 12 & 8 & 3 & 9 & 10 \\ \sigma(r)^{-1} & = & 2 & 6 & 10 & 1 & 3 & 4 & 7 & 9 & 11 & 12 & 5 & 8. \end{array}$$

This tells us that when selecting a permutation of 12 to place in T, the integers $2, 6, 10$ should appear in the brick of size 3, the integers $1, 3, 4, 7, 9, 11, 12$ should appear in the brick of size 7, and the integers $5, 8$ should appear in the brick of size 2. It is easy to see that $inv(r) = inv(\sigma(r))$ and $inv(\sigma(r)) = inv(\sigma(r)^{-1})$. Thus, the theorem of Carlitz tells us that $\begin{bmatrix} n \\ \lambda \end{bmatrix}_q$ is the sum of the the number of inversions of all sequences that are the result of placing a permutation of numbers $1, \ldots, n$ in the cells of T such that the numbers in each brick increase from left to right.

For each brick of length λ_j in T, there is an unused term of the form $\begin{bmatrix} \lambda_j \\ i_0, \ldots, i_k \end{bmatrix}_q q^{\binom{i_0}{2} + \cdots + \binom{i_k}{2}}$ where $i_0 + \cdots + i_k = \lambda_j$. The theorem of Carlitz enables us to start with a rearrangement a of i_0 0's, i_1 1's, etc. to use the q-multinomial coefficient. Record from right to left the 0's in a with $1, \ldots, i_0$. Then record the 1's in a from right to left with $i_0 + 1, \ldots, i_1$. Continue this process k times to form a permutation of λ_j from a, τ_a^{-1}. The inverse, τ_a, records the places of the 0's, 1's, etc., and therefore must have decreasing sequences of length $i_0, \ldots, i_k$. Let $\overline{\tau_a}$ be the permutation τ_a where the integers $1, \ldots, \lambda_j$ have been replaced with whatever integers the factor $\begin{bmatrix} n \\ \lambda \end{bmatrix}_q$ dictates should appear in the j^{th} brick.

For example, if $k = 3$ and $i_0 = 4$, $i_1 = 0$, $i_2 = 1$, and $i_3 = 2$ as found in the second brick in Figure 2, a permutation of 7 may be formed from 0 2 0 3 3 0 0. Continuing our example from above, the brick of size 7 should contain the integers $1, 3, 4, 7, 9, 11$, and 12. The permutations τ_a^{-1}, τ_a, and $\overline{\tau_a}$ can be found:

	1	2	3	4	5	6	7
a	0	2	0	3	3	0	0
τ_a^{-1}	4	5	3	7	6	2	1 .
τ_a	7	6	3	1	2	5	4
$\overline{\tau_a}$	12	11	4	1	3	9	7

By construction, we have that

$$inv(\overline{\tau_a}) = inv(\tau_a) = inv(\tau_a^{-1}) = inv(r) + \binom{i_0}{2} + \cdots + \binom{i_k}{2}.$$

Therefore, for each brick of size λ_j, we may associate a permutation of λ_j such that the permutation must have a descent if two consecutive cells have the same power of r. By taking along a power of $q^{inv(\overline{\tau_a})}$, we are able to account for the factors in 2.1 of the form $\left[{\lambda_j \atop i_0,\ldots,i_k}\right]_q q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}}$. Every term in 2.1 has now been used.

Let $\mathfrak{T}$ be the set of all possible brick tabloids decorated in this way. Figure 3 gives one example of such an object. We have shown how each $T \in \mathfrak{T}$ has the following five properties:

−1	−1	1	−1	−1	−1	−1	−1	−1	1	−1	1
r^1	r^1	r^3	r^0	r^0	r^0	r^0	r^2	r^3	r^3	r^1	r^1
q^9	q^1	q^4	q^8	q^7	q^2	q^0	q^0	q^3	q^1	q^1	q^0
10	2	6	12	11	4	1	3	9	7	8	5

FIGURE 3. An object coming from 2.1 when $k = 3$ and $n = 12$.

(1) T is a brick tabloid of shape (n) and type λ for some $\lambda \vdash n$,
(2) the cells in each brick contain -1 except for the final cell which contains 1,
(3) each cell contains a power of r such that the powers weakly increase within each brick,
(4) T contains a permutation of n which must have a decrease between consecutive cells within a brick if the cells are marked with the same power of r, and
(5) each cell contains a power of q recording the number of inversions caused by the integer entry.

Define the sign of $T \in \mathfrak{T}$, $sgn(T)$, as the product of all -1 labels in T. Define the weight of $T \in \mathfrak{T}$, $w(T)$, as the product of all r, and q labels. In this way, the T in Figure 3 has sign $(-1)^9$ and weight $r^{15}q^{36}$. From our development, we have

$$[n]_q!\xi_k(h_n) = \sum_{T \in \mathfrak{T}} sgn(T)w(T).$$

At this point, we will introduce a sign-reversing weight-preserving involution I on $\mathfrak{T}$ to rid ourselves of any $T \in \mathfrak{T}$ with $sgn(T) = -1$. Scan the cells of $T \in \mathfrak{T}$ from left to right looking for the first of two situations:

(1) a cell containing a -1, or
(2) two consecutive cells c_1 and c_2 such that c_1 ends a brick and either the powers of r increase from c_1 to c_2 or the powers of r are the same and the permutation decreases from c_1 to c_2.

If situation 1 is scanned first, let $I(T)$ be T where the brick containing the -1 is broken into two immediately after the violation and the -1 is changed to a 1. If situation 2 is scanned first, let $I(T)$ be T where bricks containing c_1 and c_2 are glued together and the 1 on c_1 is changed to -1. If when scanning from left to right neither case happens, let $I(T) = T$. For example, the image of the element of $\mathfrak{T}$ in Figure 3 under I is displayed in Figure 4.

1	−1	1	−1	−1	−1	−1	−1	−1	1	−1	1
r^1	r^1	r^3	r^0	r^0	r^0	r^0	r^2	r^3	r^3	r^1	r^1
q^9	q^1	q^4	q^8	q^7	q^2	q^0	q^0	q^3	q^1	q^1	q^0
10	2	6	12	11	4	1	3	9	7	8	5

FIGURE 4. The image of Figure 3 under I.

By definition, if $T \neq I(T)$, then $sgn(I(T)) = -sgn(T)$, $w(I(T)) = w(T)$, and $I(I(T)) = T$. Thus I is a sign-reversing weight-preserving involution on $\mathcal{T}$. The fixed points under I have the properties that

(1) there are no bricks with -1 in them,
(2) the powers of r weakly decrease, and
(3) if two consecutive bricks have the same power of r, then the permutation must increase there.

Since every brick of length greater than 1 contains a -1, a fixed point can only have bricks of length 1. One example of a fixed point may be found in Figure 5. We now have

1	1	1	1	1	1	1	1	1	1	1	1
r^3	r^3	r^3	r^2	r^2	r^1	r^1	r^1	r^1	r^1	r^0	r^0
q^3	q^4	q^5	q^1	q^1	q^1	q^1	q^1	q^2	q^2	q^0	q^0
4	6	8	2	3	5	7	9	11	12	1	10

FIGURE 5. A fixed point in $\mathcal{T}$ under I when $k = 3$ and $n = 12$.

$$[n]_q!\xi_k(h_n) = \sum_{T \in \mathcal{T}} sgn(T)w(T) = \sum_{\substack{T \in \mathcal{T} \text{ is a} \\ \text{fixed point under } I}} w(T),$$

so counting fixed points is the only remaining task.

Suppose that the powers of r in a fixed point are $r^{p_1}, \ldots, r^{p_n}$ when read from left to right. It must be the case that $k \geq p_1 \geq \cdots \geq p_n$. Let $a_1, \ldots, a_n$ be the nonnegative numbers defined by $a_i = p_i - p_{i+1}$ for $i = 1, \ldots, n-1$ and $a_n = p_n$. It follows that $p_1 + \cdots + p_n = a_1 + 2a_2 + \cdots + na_n$, $a_1 + \cdots + a_n = p_1 \leq k$, and if σ is the permutation in a fixed point, $a_i \geq \chi(\sigma_i > \sigma_{i+1})$. In this way, we have the sum of weights over all fixed points under I is equal to

$$\begin{aligned}
&\sum_{\substack{\sigma \in S_n \\ k \geq p_1 \geq \cdots \geq p_n \geq 0}} q^{inv(\sigma)} r^{p_1 + \cdots + p_n} \\
&= \sum_{\sigma \in S_n} q^{inv(\sigma)} \sum_{\substack{a_1 + \cdots + a_n \leq k \\ a_i \geq \chi(\sigma_i > \sigma_{i+1})}} r^{a_1 + 2a_2 + \cdots + na_n} \\
&= \sum_{\sigma \in S_n} q^{inv(\sigma)} \sum_{a_1 \geq \chi(\sigma_1 > \sigma_2)} \cdots \sum_{a_n \geq \chi(\sigma_n > n+1)} x^{a_1 + \cdots + a_n} r^{a_1 + 2a_2 + \cdots + na_n} \Bigg|_{x^{\leq k}}
\end{aligned}$$

where the notation $expression|_{x^{\leq k}}$ means to sum the coefficients of x up to and including x^k in *expression*. Rewriting the above equation, we have

$$\begin{aligned}
&\sum_{\sigma \in S_n} q^{inv(\sigma)} \sum_{a_1 \geq \chi(\sigma_1 > \sigma_2)} (xr)^{a_1} \cdots \sum_{a_n \geq \chi(\sigma_n > n+1)} (xr^n)^{a_n} \Bigg|_{x^{\leq k}} \\
&= \sum_{\sigma \in S_n} q^{inv(\sigma)} \frac{(xr)^{\chi(\sigma_1 > \sigma_2)} \cdots (xr^n)^{\chi(\sigma_n > n+1)}}{(1 - xr) \cdots (1 - xr^n)} \Bigg|_{x^{\leq k}} \\
&= \frac{\sum_{\sigma \in S_n} x^{des(\sigma)} r^{maj(\sigma)} q^{inv(\sigma)}}{(1 - xr) \cdots (1 - xr^n)} \Bigg|_{x^{\leq k}}.
\end{aligned}$$

Dividing by $(1 - x)$ allows for $x^{\leq k}$ to be changed to x^k in the above expression, thereby arriving at the statement of the theorem. □

1.1 is a corollary. We have

$$\begin{aligned}
&\sum_{n\geq 0}\frac{t^n}{[n]_q!(x;r)_{n+1}}\sum_{\sigma\in S_n}x^{des(\sigma)}r^{maj(\sigma)}q^{inv(\sigma)} \\
&\qquad = \sum_{k\geq 0}x^k\sum_{n\geq 0}\frac{t^n}{[n]_q!(x;r)_{n+1}}\sum_{\sigma\in S_n}x^{des(\sigma)}r^{maj(\sigma)}q^{inv(\sigma)}\Bigg|_{x^k} \\
&\qquad = \sum_{k\geq 0}x^k\sum_{n\geq 0}t^n\xi_k(h_n),
\end{aligned}$$

which by an application of 1.2 is equal to

$$\begin{aligned}
\sum_{k\geq 0}x^k\xi_k\left(\sum_{n\geq 0}e_n(-t)^n\right)^{-1} &= \sum_{k\geq 0}x^k\left(\sum_{n\geq 0}(-t)^n\sum_{\substack{i_0,\ldots,i_k\geq 0\\ i_0+\cdots+i_k=n}}\frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q!\cdots[i_k]_q!}q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}}\right)^{-1} \\
&= \sum_{k\geq 0}x^k\left(\sum_{n\geq 0}\frac{(-tr^0)^n}{[n]_q!}q^{\binom{n}{2}}\right)^{-1}\cdots\left(\sum_{n\geq 0}\frac{(-tr^k)^n}{[n]_q!}q^{\binom{n}{2}}\right)^{-1} \\
&= \sum_{k\geq 0}\frac{x^k}{\exp_q(-tr^0)\cdots\exp_q(-tr^k)}.
\end{aligned}$$

The proof we have given for 1.1, although elementary and combinatorial, is not any "easier" than that given by Garsia and Gessel. However, there are at least two distinct advantages of our methods. First, the techniques in the proof of theorem 2.1 may be slightly modified to give a wide swath of seemingly unrelated generating functions for the permutation enumeration of the symmetric group, Weyl groups of type B and D, subsets of the symmetric group, and more [**M**]. Second, the ideas in the proof of Theorem 2.1 may be generalized to give new generating functions involving the descent, major index, and inversion statistics.

3. Generating functions for the Weyl groups of type B and D

Let us turn our attention to applying this machinery to the hyperoctahedral group B_n and its subgroup D_n. The hyperoctahedral group B_n may be considered the set of permutations of n where each integer in the permutation is assigned either a $+$ or $-$ sign. For $\sigma \in B_n$, let $neg(\sigma)$ count the total number of negative signs in σ. The subgroup D_n of B_n contains those $\sigma \in B_n$ with $neg(\sigma)$ an even number. These are Weyl groups appearing in the study of root systems and Lie algebras.

Define a linear order Θ on $\{\pm 1, \ldots, \pm n\}$ such that

$$1 <_\Theta \cdots <_\Theta n <_\Theta -n <_\Theta \cdots <_\Theta -1$$

and define $des_B(\sigma)$ on B_n such that

$$des_B(\sigma) = \chi(n <_\Theta \sigma_n) + \sum_i \chi(\sigma_{i+1} <_\Theta \sigma_i).$$

This definition and the linear order Θ arises from an interpretation of B_n as a Coxeter group. For $\sigma \in B_n$, let $maj_B(\sigma)$ and $inv_B(\sigma)$ be the major index and inversion statistics with respect to the linear order Θ.

Using the methods of Garsia and Gessel and the study of upper binomial posets, Reiner found a generalization of 1.1 for B_n involving the statistic counting the number of negative signs in B_n and versions of descents, inversions, and the major index [**R**]. In this Section we will indicate (without a formal proof) how a simple modification of the method given in Section 2 to prove 1.1 can do the same.

Define a homomorphism $\xi_{B,k}$ on the ring of symmetric functions by defining it on e_n such that

$$\xi_{B,k}(e_n) = \sum_{\substack{i_0,\ldots,i_k \geq 0 \\ i_0+\cdots+i_k=n}} \frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q!\cdots[i_k]_q!} q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} [i_0+1]_y \cdots [i_k+1]_y.$$

Then it may be proved that

$$[n]_q!\xi_{B,k}(h_n) = \frac{1}{(x;r)_{n+1}} \sum_{\sigma \in B_n} x^{des_B(\sigma)} r^{maj_B(\sigma)} q^{inv_B(\sigma)} y^{neg(\sigma)} \Bigg|_{x^k}. \tag{3.1}$$

The proof of this fact may be found by first expanding h_n via 1.4 to form combinatorial objects like Figure 6 below. The $[i_0+1]_y \cdots [i_k+1]_y$ term in the definition of the ring homomorphism $\xi_{B,k}$ allows for

−1	−1	−1	−1	1	−1	1	−1	−1	1	1	1
r^0	r^0	r^1	r^1	r^1	r^2	r^3	r^1	r^3	r^3	r^0	r^1
q^{11}	q^0	q^8	q^7	q^0	q^1	q^0	q^1	q^1	q^0	q^1	q^0
12	1	10	9	2	4	3	6	7	5	11	8
y		y	y			y		y		y	

FIGURE 6. An example of a combinatorial object coming from the application of $\xi_{B,k}$ to $[n]_q!h_n$.

the bottom row of T to contain some number of y's in cells marked with the same power of r so that these objects have the exact same properties as the T found in Figures 3, 4, and 5 with the addition of powers of y recorded in the bottom of the object. These powers of y will be interpreted to mean that the integer in the permutation in the cell marked with y is negative.

Suppose we have j consecutive cells within a brick with the same power of r and marked with a y. Instead of writing these integers in decreasing order as prescribed in the proof of Theorem 2.1, let us reverse the order of these j integers in T so that they are in increasing order. An example of this may be found in Figure 7. This is done so that cells in a brick marked with the same power of r are in descending order according

−1	−1	−1	−1	1	−1	1	−1	−1	1	1	1
r^0	r^0	r^1	r^1	r^1	r^2	r^3	r^1	r^3	r^3	r^0	r^1
q^{11}	q^0	q^8	q^7	q^0	q^1	q^0	q^1	q^1	q^0	q^1	q^0
12	1	9	10	2	4	3	6	7	5	11	8
y		y	y			y		y		y	

FIGURE 7. Reversing the order of two integers in Figure 6.

to the linear order Θ. The same brick breaking/combining sign-reversing weight-preserving involution as in Theorem 2.1 may be now applied to leave a set of fixed points which may be counted to yield 3.1.

A generating function may be found employing 1.2. We have

$$\sum_{n\geq 0} \frac{t^n}{[n]_q!(x;r)_{n+1}} \sum_{\sigma\in B_n} x^{des_B(\sigma)} r^{maj_B(\sigma)} q^{inv_B(\sigma)} y^{neg(\sigma)}$$

$$= \sum_{k\geq 0} x^k \left(\sum_{n\geq 0} (-t)^n \sum_{\substack{i_0,\ldots,i_k\geq 0 \\ i_0+\cdots+i_k=n}} \frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q!\cdots[i_k]_q!} q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} [i_0+1]_y\cdots[i_k+1]_y \right)^{-1}$$

which in turn may be simplified to look like

$$\sum_{k\geq 0} \frac{(x-xy)^k}{\left(\exp_q(-tr^0) - y\exp_q(-tyr^0)\right)\cdots\left(\exp_q(-tr^k) - y\exp_q(-tyr^k)\right)}. \tag{3.2}$$

Let $A(t,x,r,q,y)$ denote the generating function in 3.2 above. Notice that $A(t,x,r,q,0)$ is equal to the generating function in 1.1 as it should.

For any series $f(x) = \sum_{n\geq 0} c_n x^n$ for $c_i \in \mathbb{C}$, we have

$$\frac{f(x)+f(-x)}{2} = \sum_{n\geq 0} c_{2n} x^{2n}.$$

Therefore,

$$\sum_{n\geq 0} \frac{t^n}{[n]_q!(x;r)_{n+1}} \sum_{\sigma\in D_n} x^{des_B(\sigma)} r^{maj_B(\sigma)} q^{inv_B(\sigma)} y^{neg(\sigma)} = \frac{A(t,x,r,q,y)+A(t,x,r,q,-y)}{2},$$

giving a multivariate generating function for the Weyl group of type D.

4. A Generating function for pairs of permutations

Let us find a generating function for two copies of the symmetric group S_n. Given $\sigma^1, \sigma^2 \in S_n$, define $comdes(\sigma^1,\sigma^2)$ as the number of times $\sigma_i^j > \sigma_{i+1}^j$ for all $j = 1,2$—this is known as the number of common descents. Let $commaj(\sigma^1,\sigma^2)$ register i for every time $\sigma_i^j > \sigma_{i+1}^j$ for $j = 1,2$. These type of statistics have been studied and a multivariate generating function for descents and inversions was found in [**F**] by Fedou and Rawlings. In this Section, we will apply our methods to finding a generating function involving the descent, major index, and inversion statistics by altering the ring homomorphism ξ_k.

Define $\xi_{k,2}$ as a homomorphism on the ring of symmetric functions by defining it on the n^{th} elementary symmetric functions such that

$$\xi_{k,2}(e_n) = \sum_{\substack{i_0,\ldots,i_k\geq 0\\ i_0+\cdots+i_k=n}} \frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q![i_0]_p!\cdots[i_k]_q![i_k]_p!} q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} p^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}}.$$

The difference between ξ_k and $\xi_{k,2}$ is that all the terms involving q in ξ_k have written down twice in the indeterminates q and p. Using $\xi_{k,2}$, it may be shown that

$$\sum_{n\geq 0} \frac{t^n}{[n]_q![n]_p!(x;r)_{n+1}} \sum_{\sigma^1,\sigma^2\in S_n} x^{comdes(\sigma^1,\sigma^2)} r^{commaj(\sigma^1,\sigma^2)} q^{inv(\sigma^1)} p^{inv(\sigma^2)}$$

$$= \sum_{k\geq 0} x^k \left(\sum_{n\geq 0} (-t)^n \sum_{\substack{i_0,\ldots,i_k\geq 0\\ i_0+\cdots+i_k=n}} \frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q![i_0]_p!\cdots[i_k]_q![i_k]_p!} q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} p^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} \right)^{-1}. \tag{4.1}$$

The proof of 4.1 follows in the same way that we have proved 1.1. First, it may be shown that

$$[n]_q![n]_p!\xi_{k,2}(h_n) = \frac{1}{(x;r)_{n+1}} \sum_{\sigma^1,\sigma^2\in S_n} x^{comdes(\sigma^1,\sigma^2)} r^{commaj(\sigma^1,\sigma^2)} q^{inv(\sigma^1)} p^{inv(\sigma^2)} \Bigg|_{x^k}. \tag{4.2}$$

This is analogous to our Theorem 2.1. The combinatorial objects we are able to create based on brick tabloids are the same as those found in Figures 3, 4, and 5 with one slight change. The q and p analogues give rise to two different permutations in a brick tabloid instead of one. The powers of q and p register the inversions of each permutation.

For example, one such combinatorial object which may be formed starting with 4.2 and using the techniques in the proof of Theorem 2.1 is found in Figure 5 below. The combinatorial objects may be

-1	-1	-1	-1	1	-1	1	-1	-1	1	1	1
r^0	r^0	r^1	r^1	r^1	r^2	r^3	r^1	r^3	r^3	r^0	r^1
q^{11}	q^0	q^8	q^7	q^0	q^1	q^0	q^1	q^1	q^0	q^1	q^0
12	1	10	9	2	4	3	6	7	5	11	8
p^{10}	p^3	p^9	p^8	p^3	p^0	p^0	p^2	p^2	p^1	p^0	p^0
11	4	12	10	5	1	2	7	8	6	3	9

FIGURE 8. An example of T arising from 4.2

constructed to have the following properties:

(1) T is a brick tabloid of shape (n) and type λ for some $\lambda \vdash n$,
(2) the cells not at the end of a brick are marked with -1 and cells at the end a brick are marked with 1,
(3) each cell contains a power of r such that the powers weakly increase within each brick,
(4) T contains *two* permutations of n which both must have a decrease between consecutive cells within a brick if the cells are marked with the same power of r, and
(5) each cell contains a power of q and p recording the number of integers in each of the permutations to the right which are smaller.

The sign of such an object is the product of the -1 signs in the objects and the weight is defined to be the product of all indeterminates in the object. Once these combinatorial objects are defined, a very similar sign-reversing weight-preserving involution I as given in the proof of Theorem 2.1 may be employed. That is, to define I, scan the cells of $T \in \mathcal{T}$ from left to right looking for the first of two situations:

(1) a cell containing a -1, or
(2) two consecutive cells c_1 and c_2 such that c_1 ends a brick and either the powers of r increase from c_1 to c_2 or the powers of r are the same and the both permutations decrease from c_1 to c_2.

If situation 1 is scanned first, let $I(T)$ be T where the brick containing the -1 is broken into two immediately after the violation and the -1 is changed to a 1. If neither (1) or (2) applies, then we define $I(T) = T$. By definition, if $T \neq I(T)$, then $sgn(I(T)) = -sgn(T)$, $w(I(T)) = w(T)$, and $I(I(T)) = T$. Thus I is a sign-reversing weight-preserving involution on $\mathcal{T}$. The fixed points under I have the properties that

(1) there are no bricks with -1 in them,
(2) the powers of r weakly decrease, and
(3) if two consecutive bricks have the same power of r, then at least one of the permutations must increase there.

Using the same techniques as in Section 2, one can show that the fixed points under this involution may then be counted to prove 4.2. 4.1 follows by an application of the simple relationship between the elementary and homogeneous symmetric functions in 1.2.

Instead of defining a "double" version of ξ_k in $\xi_{k,2}$, one may define a "m-tuple" version of ξ_k to help record generating functions for m copies of the symmetric group. The process is no more difficult than the method we have outlined for two copies of the symmetric group except for the fact that there are m indeterminates to keep track of instead of two.

Furthermore, we can keep track of two elements in B_n or D_n using a combination of the ideas behind the definitions of $\xi_{B,k}$ and $\xi_{k,2}$. That is, by defining a homomorphism by mapping e_n to

$$\sum_{\substack{i_0,\dots,i_k\geq 0\\ i_0+\cdots+i_k=n}} \frac{r^{0i_0+\cdots+ki_k}}{[i_0]_q![i_0]_p!\cdots[i_k]_q![i_k]_p!} q^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} p^{\binom{i_0}{2}+\cdots+\binom{i_k}{2}} [i_0+1]_y[i_0+1]_z\cdots[i_k+1]_y[i_k+1]_z,$$

then we can find a generating function registering common B_n descents and major indices together with inversions and the number of negative signs over pairs of signed permutations. This technique, in general, can provide many different generating functions for permutation statistics.

References

[C] L. Carlitz, Sequences and inversions, *Duke Math. J.* **37** (1970) 193–198.

[E] Ö. Egecioglu, J. Remmel, Brick tabloids and the connection matrices between bases of symmetric functions, *Discrete Appl. Math.* **34** (1991) 107–120.

[F] J.-M. Fedou, D. Rawlings, Statistics on pairs of permutations, *Discrete Math.* **143** (1995) 31–45.

[Ga] A. Garsia, I. Gessel, Permutation statistics and partitions, *Adv. in Math.* **31** (1979) 288–305.

[Ge] I. Gessel, Generating functions and enumeration of sequences, *Ph.D. Thesis, MIT* (1977).

[M] A. Mendes, Building generating functions brick by brick, *Ph.D. Thesis, University of California, San Diego* (to appear in 2004).

[S] R. Stanley, Enumerative combinatorics, vol. 2, Cambridge Univ. Press. (1999).

[R] V. Reiner, Signed permutation statistics, *Europ. J. Combinatorics* **14** (1993), 553–567.

Department of Mathematics, University of California, San Diego, USA 92093
E-mail address: amendes@math.ucsd.edu

Department of Mathematics, University of California, San Diego, USA 92093
E-mail address: remmel@math.ucsd.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

A Solution to the Tennis Ball Problem

Anna de Mier and Marc Noy

Abstract. *We present a complete solution to the so-called tennis ball problem, which is equivalent to counting lattice paths in the plane that use North and East steps and lie between certain boundaries. The solution takes the form of explicit expressions for the corresponding generating functions.*

Our method is based on the properties of Tutte polynomials of matroids associated to lattice paths. We also show how the same method provides a solution to a wide generalization of the problem.

RÉSUMÉ. *Nous présentons une solution complète au "problème des balles de tennis", pro-blème qui revient à compter des chemins, formés par des pas Nord et Est, dans une région délimitée de $\mathbb{N}^2$. La solution se présente sous la forme d'expressions explicites pour les séries génératrices correspondantes.*

Notre méthode repose sur certaines propriétés des polynômes de Tutte des matroïdes associés à des chemins de $\mathbb{N}^2$. Nous montrons aussi comment cette méthode permet de résoudre un problème beaucoup plus général.

1. Introduction

The statement of the tennis ball problem is the following. There are $2n$ balls numbered $1, 2, 3, \dots, 2n$. In the first turn balls 1 and 2 are put into a basket and one of them is removed. In the second turn balls 3 and 4 are put into the basket and one of the three remaining balls is removed. Next balls 5 and 6 go in and one of the four remaining balls is removed. The game is played n turns and at the end there are exactly n balls outside the basket. The question is how many different sets of balls may we have at the end outside the basket.

It is easy to reformulate the problem in terms of lattice paths in the plane that use steps $E = (1, 0)$ and $N = (0, 1)$. It amounts to counting lattice paths from $(0, 0)$ to (n, n) that never go above the path $NE \cdots NE = (NE)^n$. Indeed, if $\pi = \pi_1\pi_2 \dots \pi_{2n-1}\pi_{2n}$ is such a path, a moment's thought shows that we can identify the indices i such that π_{2n-i+1} is a N step with the labels of balls that end up outside the basket. The number of such paths is well-known to be a Catalan number, and this is the answer obtained in [**GM**].

The problem can be generalized as follows [**MSV**]. We are given positive integers $t < s$ and sn labelled balls. In the first turn balls $1, \dots, s$ go into the basket and t of them are removed. In the second turn balls $s + 1, \dots, 2s$ go into the basket and t among the remaining ones are removed. After n turns, tn balls lie outside the basket, and again the question is how many different sets of balls may we have at the end. Letting $k = t, l = s - t$, the problem is seen as before to be equivalent to counting lattice paths from $(0, 0)$

Research partially supported by projects BFM2001-2340 and CUR Gen. Cat. 1999SGR00356.

to (ln, kn) that use N and E steps and never go above the path $N^kE^l \cdots N^kE^l = (N^kE^l)^n$. This is the version of the problem we solve in this paper.

From now on we concentrate on lattice paths that use N and E steps. To our knowledge, the only cases solved so far are $k = 1$ and $k = l = 2$. The case $k = 1$ is straightforward, the answer being a generalized Catalan number $\frac{1}{ln+1}\binom{(l+1)n}{n}$. The case $k = l = 2$ (corresponding to the original problem when $s = 4, t = 2$) is solved in [**MSV**] using recurrence equations; here we include a direct solution. We say that a path is a *Catalan path of semilength* n if it goes from $(0, 0)$ to (n, n) and stays below the line $x = y$. The case $k = l = 2$ is illustrated in Fig. 1, to which we refer next. A path π not above $(N^2E^2)^n$ is "almost" a Catalan path, in the sense that it can raise above the dashed diagonal line only through the dotted points. But clearly between two consecutive dotted points hit by π we must have an E step, followed by a path isomorphic to a Catalan path of odd semilength, followed by a N step. Thus, π is essentially a sequence of Catalan paths of odd semilength. If $G(z) = \sum_n \frac{1}{n+1}\binom{2n}{n} z^n$ is the generating function for the Catalan numbers, take the odd part $G_o(z) = (G(z) - G(-z))/2$. Then expand $1/(1 - zG_o(z))$ to obtain the sequence $1, 6, 53, 554, 6363, \ldots$, which agrees with the results in [**MSV**].

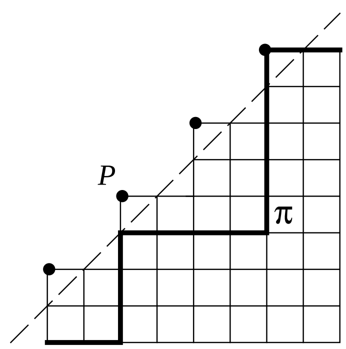


FIGURE 1. The path $\pi = \underline{EE}NN\underline{N}EEEEENNN\underline{NN}EE$ not above $P = (N^2E^2)^4$. It has $i(\pi) = 3$ and $e(\pi) = 2$, corresponding to the steps underlined.

Let P be a lattice path from $(0, 0)$ to (m, r), and let $b(P)$ be the number of paths from $(0, 0)$ to (m, r) that never go above P. If PN denotes the path obtained from P by adding a N step at the end of P, then clearly $b(P) = b(PN)$. However, it is not possible to express $b(PE)$ simply in terms of $b(P)$, where PE has the obvious meaning. As is often the case in counting problems, one has to enrich the objects under enumeration with additional parameters that allow suitable recursive decompositions. This is precisely what is done here: equations (2.2) and (2.3) in the next section contain variables x and y, corresponding to two parameters that we define on lattice paths not above a given path P. These equations are the key to our solution.

The basis of our approach is the connection between lattice paths and matroids established in [**BMN**], where the link with the tennis ball problem was already remarked. For completeness, we recall the basic facts needed from [**BMN**] in the next section. In Section 3 we present our solution to the tennis ball problem, in the form of explicit expressions for the corresponding generating functions; see Theorem 3.1. In Section 4 we show how the same method can be applied to a more general problem. We conclude with some remarks.

2. Preliminaries

The contents of this section are taken mainly from [**BMN**], where the reader can find additional background and references on matroids, Tutte polynomials and lattice path enumeration.

A *matroid* is a pair $(E, \mathcal{B})$ consisting of a finite set E and a nonempty collection $\mathcal{B}$ of subsets of E, called *bases* of the matroid, that satisfy the following conditions: (1) No set in $\mathcal{B}$ properly contains another set in $\mathcal{B}$, and (2) for each pair of distinct sets B, B' in $\mathcal{B}$ and for each element $x \in B - B'$, there is an element $y \in B' - B$ such that $(B - x) \cup y$ is in $\mathcal{B}$.

Let P be a lattice path from $(0, 0)$ to (m, r). Associated to P there is a matroid $M[P]$ on the set $\{1, 2, \ldots, m + r\}$ whose bases are in one-to-one correspondence with the paths from $(0, 0)$ to (m, r) that never go above P. Given such a path $\pi = \pi_1\pi_2 \ldots \pi_{m+r}$, the basis corresponding to π consists of the indices i such that π_i is a N step. Hence, counting bases of $M[P]$ is the same as counting lattice paths that never go above P.

For any matroid M there is a two-variable polynomial with non-negative integer coefficients, the Tutte polynomial $t(M; x, y)$. It was introduced by Tutte [**T1**] and presently plays an important role in combinatorics and related areas (see [**W**]). The key property in this context is that $t(M; 1, 1)$ equals the number of bases of M.

Given a path P as above, there is a direct combinatorial interpretation of the coefficients of $t(M[P]; x, y)$. For a path π not above P, let $i(\pi)$ be the number of N steps that π has in common with P, and let $e(\pi)$ be the number of E steps of π before the first N step, which is 0 if π starts with a N step. This is illustrated in Fig. 1, where the path P corresponds to the upper border of the diagram and hence a path π representing a basis of $M[P]$ corresponds to a path that stays in the region shown.

Then we have (see [**BMN**, Th. 5.4])

$$t(M[P]; x, y) = \sum_{\pi} x^{i(\pi)} y^{e(\pi)}, \tag{2.1}$$

where the sum is over all paths π not above P. A direct consequence is that $t(M[P]; 1, 1)$ is the number of such paths.

Furthermore, for the matroids $M[P]$ there is a rule for computing the Tutte polynomial that we use repeatedly (see [**BMN**, Section 6]). If PN and PE denote the paths obtained from P by adding a N step and an E step at the end of P, respectively, then

$$t(M[PN]; x, y) = x\, t(M[P], x, y), \tag{2.2}$$

$$t(M[PE]; x, y) = \frac{x}{x-1}\, t(M[P], x, y) + \left(y - \frac{x}{x-1}\right) t(M[P]; 1, y). \tag{2.3}$$

The right-hand side of (2.3) is actually a polynomial, since $x - 1$ divides $t(M[P]; x, y) - t(M[P]; 1, y)$ (see the expansion (2.4) below). The key observation here is that we cannot simply set $x = y = 1$ in (2.3) to obtain an equation linking $t(M[PE]; 1, 1)$ and $t(M[P]; 1, 1)$.

For those familiar with matroid theory, we remark that the quantities $i(\pi)$ and $e(\pi)$ correspond to the internal and external activities of the basis associated to π with respect to the order $1 < 2 < \cdots < m + r$ of the ground set of $M[P]$. Also, the matroids $M[PN]$ and $M[PE]$ are obtained from $M[P]$ by adding an isthmus and taking a free extension, respectively; it is known that formulas (2.2) and (2.3) correspond precisely to the effect these two operations have on the Tutte polynomial of an arbitrary matroid.

From (2.1) and the definition of $i(\pi)$ and $e(\pi)$, equation (2.2) is clear, since any path associated to $M[PN]$ has to use the last N step. For completeness, we include a direct proof of equation (2.3).

We first rewrite the right-hand side of (2.3) as

$$
\begin{aligned}
&\frac{x}{x-1}(t(M[P];x,y)-t(M[P];1,y))+yt(M[P];1,y)=\\
&\sum_{\pi}\frac{x}{x-1}y^{e(\pi)}(x^{i(\pi)}-1)+y^{e(\pi)+1}=\\
&\sum_{\pi}y^{e(\pi)}(y+x+x^2+\cdots+x^{i(\pi)}),
\end{aligned}
\tag{2.4}
$$

where the sums are taken over all paths π that do not go above P.

To prove the formula, for each path π not above P we find $i(\pi)+1$ paths not above PE such that their total contribution to $t(M[PE];x,y)$ is $y^{e(\pi)}(y+x+x^2+\cdots+x^{i(\pi)})$. Consider first the path $\pi_0=E\pi$; it clearly does not go above PE and its contribution to the Tutte polynomial is $y^{e(\pi)+1}$. Now for each j with $1\leq j\leq i(\pi)$, define the path π_j as the path obtained from π by inserting an E step after the jth N step that π has in common with P (see Fig. 2). The path π_j has exactly j N steps in common with PE, and begins with $e(\pi)$ E steps. Observe also that, if the j-th N step of π is the k-th step, then π and π_j agree on the first k and on the last $m+r-k$ steps.

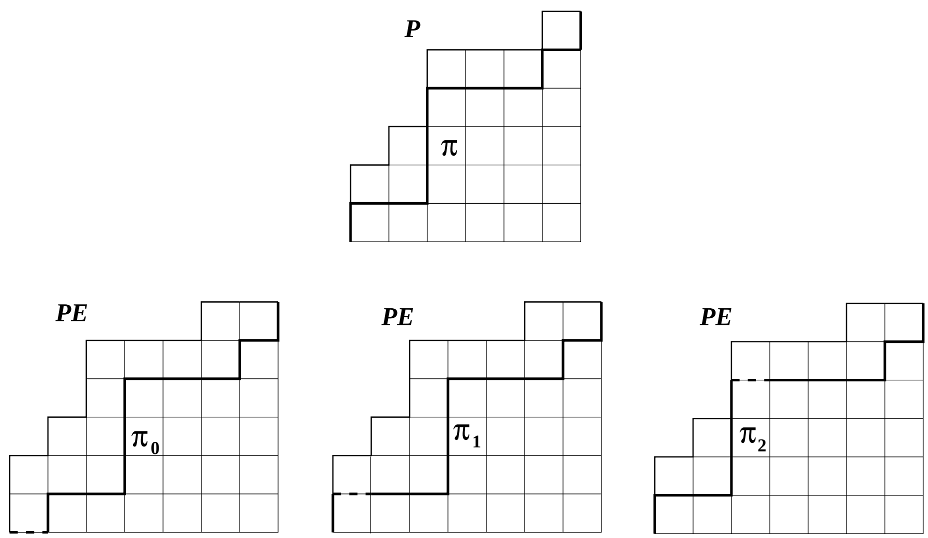


FIGURE 2. Illustrating the combinatorial proof of formula (2.3): from a path π not above P with $i(\pi)=2$ we generate 3 paths not above PE.

It remains only to show that each contribution to the Tutte polynomial of $M[PE]$ arises as described above. Let π' be a path that never goes above PE and consider the last N step that π' has in common with PE; clearly the next step must be E. Let $\widetilde{\pi}$ be the path obtained after removing this E step. Since π' had no N steps in common with PE after the removed E step, the path $\widetilde{\pi}$ does not go above P. Thus the path π' can be obtained from $\widetilde{\pi}$ by adding an E step after the $i(\pi')$-th N step that $\widetilde{\pi}$ has in common with P, and hence π' arises from $\widetilde{\pi}$ as above. By the remarks at the end of the previous paragraph, it is clear that π' cannot be obtained in any other way by applying the procedure described above, and this finishes the proof.

3. Main result

Let k,l be fixed positive integers, and let $P_n=(N^kE^l)^n$. Our goal is to count the number of lattice paths from $(0,0)$ to (ln,kn) that never go above P_n. From the considerations in the previous section, this is the same as computing $t(M[P_n];1,1)$. Let

$$A_n = A_n(x,y) = t(M[P_n];x,y).$$

By convention, P_0 is the empty path and $A_0 = 1$.

In order to simplify the notation we introduce the following operator Φ on two-variable polynomials:

$$\Phi A(x,y) = \frac{x}{x-1}A(x,y) + \left(y - \frac{x}{x-1}\right)A(1,y).$$

Then, by equations (2.2) and (2.3) we have

$$A_{n+1} = \Phi^l(x^k A_n),$$

where Φ^i denotes the operator Φ applied i times.

For each $n \geq 0$ and $i = 1,\dots,l$, we define polynomials $B_{i,n}(x,y)$ and $C_{i,n}(y)$ as

$$\begin{aligned} B_{i,n} &= \Phi^i\left(x^k A_n(x,y)\right), \\ C_{i,n} &= B_{i,n}(1,y). \end{aligned}$$

We also set $C_{0,n}(y) = A_n(1,y)$. Notice that $B_{l,n} = A_{n+1}$, and $C_{0,n}(1) = A_n(1,1)$ is the quantity we wish to compute.

Then, by the definition of Φ, we have:

$$\begin{aligned} B_{1,n} &= \frac{x}{x-1}x^k A_n &+& \left(y - \frac{x}{x-1}\right)C_{0,n}; \\ B_{2,n} &= \frac{x}{x-1}B_{1,n} &+& \left(y - \frac{x}{x-1}\right)C_{1,n}; \\ &\cdots \\ B_{l,n} &= \frac{x}{x-1}B_{l-1,n} &+& \left(y - \frac{x}{x-1}\right)C_{l-1,n}; \\ A_{n+1} &= B_{l,n}. \end{aligned}$$

In order to solve these equations, we introduce the following generating functions in the variable z (but recall the coefficients are polynomials in x and y):

$$A = \sum_{n\geq 0} A_n z^n, \qquad C_i = \sum_{n\geq 0} C_{i,n} z^n, \quad i = 0,\dots,l.$$

We start from the last equation $A_{n+1} = B_{l,n}$ and substitute repeatedly the value of $B_{i,n}$ from the previous equation. Taking into account that $\sum_n A_{n+1}z^n = (A-1)/z$, a simple computation yields

$$\frac{A-1}{z} = \frac{x^{k+l}}{(x-1)^l}A + (yx - y - x)\sum_{i=1}^{l} \frac{x^{i-1}}{(x-1)^i}\,C_{l-i}.$$

We now set $y = 1$ and obtain

$$A\left((x-1)^l - zx^{k+l}\right) = (x-1)^l - z\sum_{i=1}^{l} x^{i-1}(x-1)^{l-i}\,C_{l-i}, \tag{3.1}$$

where it is understood that from now on we have set $y = 1$ in the series A and C_i.

By Puiseux's theorem (see [**S**, Theorem 6.1.5]), the algebraic equation in w

$$(w-1)^l - zw^{k+l} = 0 \tag{3.2}$$

has $k+l$ solutions in the field $\mathbb{C}^{\,\mathrm{fra}}((z)) = \{\sum_{n\geq n_0} a_n z^{n/N}\}$ of fractional Laurent series. Proposition 6.1.8 in [**S**] tells us that exactly l of them are fractional power series (without negative powers of z); let them be $w_1(z),\dots,w_l(z)$.

We substitute $x = w_j$ in (3.1) for $j = 1, \dots, l$, so that the left-hand side vanishes, and obtain a system of l linear equations in $C_0, C_1, \dots, C_{l-1}$, whose coefficients are expressions in the w_j, namely

$$\sum_{i=1}^{l} w_j^{i-1}(w_j - 1)^{l-i} z C_{l-i} = (w_j - 1)^l, \qquad j = 1, \dots, l. \tag{3.3}$$

Notice that, in order of the product in the left hand-side of (3.1) to be defined, the solutions of (3.2) that we substitute cannot have negative powers of z, hence they must be $w_1, \dots, w_l$. We remark that this technique is similar with the one devised by Tutte for counting rooted planar maps (see, for instance, [**T2**]).

It remains only to solve (3.3) to obtain the desired series $C_0 = \sum_n A_n(1,1)z^n$. The system (3.3) can we written as

$$\sum_{i=0}^{l-1} \left(\frac{w_j}{w_j - 1}\right)^i z C_{l-i-1} = w_j - 1, \qquad j = 1, \dots, l.$$

The left-hand sides of the previous equations can be viewed as the result of evaluating the polynomial $\sum_{i=0}^{l-1}(zC_{l-i-1})X^i$ of degree $l-1$ at $X = w_j/(w_j - 1)$, for j with $1 \le j \le l$. Using Lagrange's interpolation formulas, we get that the coefficient of X^{l-1} in this polynomial is

$$zC_0 = \sum_{j=1}^{l} \frac{w_j - 1}{\prod_{i \ne j}\left(\frac{w_j}{w_j - 1} - \frac{w_i}{w_i - 1}\right)}.$$

By straightforward manipulation this last expression is equal to

$$-\prod_{j=1}^{l}(1 - w_j) \sum_{j=1}^{l} \frac{(w_j - 1)^{l-1}}{\prod_{i \ne j}(w_j - w_i)} = -\prod_{j=1}^{l}(1 - w_j),$$

where the last equality follows from an identity on symmetric functions (set $r = 0$ in Exercise 7.4 in [**S**]).

Thus we have proved the following result.

Theorem 3.1. *Let k, l be positive integers. Let q_n be the number of lattice paths from $(0,0)$ to (ln, kn) that never go above the path $(N^k E^l)^n$, and let $w_1, \dots, w_l$ be the unique solutions of the equation*

$$(w - 1)^l - z w^{k+l} = 0$$

that are fractional power series. Then the generating function $Q(z) = \sum_{n \ge 0} q_n z^n$ is given by

$$Q(z) = \frac{-1}{z}(1 - w_1) \cdots (1 - w_l).$$

Note that, by symmetry, the number of paths not above $(N^l E^k)^n$ must be the same as in Theorem 3.1, although the algebraic functions involved in the solution are roots of a different equation.

In the particular case $k = l$ the solution can be expressed directly in terms of the generating function $G(z) = \sum_n \frac{1}{n+1}\binom{2n}{n} z^n$ for the Catalan numbers, which satisfies the quadratic equation $G(z) = 1 + zG(z)^2$. Indeed, (3.2) can be rewritten as

$$w = 1 + z^{1/k} w^2,$$

whose (fractional) power series solutions are $G(\zeta^j z^{1/k})$, $j = 0, \dots, k-1$, where ζ is a primitive k-th root of unity. For instance, for $k = l = 3$ (corresponding to $s = 6, t = 3$ in the original problem), $\zeta = \exp(2\pi i/3)$ and we obtain the solution

$$\frac{-1}{z}(1 - G(z^{1/3}))(1 - G(\zeta z^{1/3}))(1 - G(\zeta^2 z^{1/3})) = \\ 1 + 20z + 662z^2 + 26780z^3 + 1205961z^4 + 58050204z^5 + \cdots.$$

In the same way, if l divides k and we set $p = (k+l)/l$, the solution can be expressed in terms of the generating function $\sum_n \frac{1}{(p-1)n+1}\binom{pn}{n} z^n$ for generalized Catalan numbers; the details are left to the reader. As an example, for $k = 4$, $l = 2$, we obtain the series

$$\frac{-1}{z}(1 - H(z^{1/2}))(1 - H(-z^{1/2})) = \\ 1 + 15z + 360z^2 + 10463z^3 + 337269z^4 + 11599668z^5 + \cdots ,$$

where $H(z) = \sum_n \frac{1}{2n+1}\binom{3n}{n}$ satisfies $H(z) = 1 + zH(z)^3$.

4. A further generalization

In this section we solve a further generalization of the tennis ball problem. Given fixed positive integers $s_1, t_1, \ldots, s_r, t_r$ with $t_i < s_i$ for all i, let $s = \sum s_i, t = \sum t_i$. There are sn labelled balls. In the first turn we do the following: balls $1, \ldots, s_1$ go into the basket and t_1 of them are removed; then balls $s_1 + 1, \ldots, s_1 + s_2$ go into the basket and among the remaining ones t_2 are removed; this goes on until we introduce balls $s - s_r + 1, \ldots, s$, and remove t_r balls. After n turns there are tn balls outside the basket and the question is again how many different sets of tn balls may we have at the end.

The equivalent path counting problem is: given $k_1, l_1, \ldots, k_r, l_r$ positive integers with $k = \sum k_i$, $l = \sum l_i$, count the number of lattice paths from $(0,0)$ to (ln, kn) that never go above the path $P_n = (N^{k_1}E^{l_1} \cdots N^{k_r}E^{l_r})^n$. The solution parallels the one presented in Section 3. We keep the notations and let $A_n = t(M[P_n]; x, y)$, so that

$$A_{n+1} = \Phi^{l_r}(x^{k_r} \cdots \Phi^{l_1}(x^{k_1}A_n) \cdots).$$

As before, we introduce l polynomials $B_{i,n}(x,y)$ and $C_{i,n}(y) = B_{i,n}(1,y)$, but the definition here is a bit more involved:

$$\begin{array}{lll} B_{i,n} & = \Phi^i(x^{k_1}A_n), & i = 1, \ldots, l_1; \\ B_{l_1+i,n} & = \Phi^i(x^{k_2}B_{l_1,n}), & i = 1, \ldots, l_2; \\ B_{l_1+l_2+i,n} & = \Phi^i(x^{k_3}B_{l_1+l_2,n}), & i = 1, \ldots, l_3; \\ & \cdots & \\ B_{l-l_r+i,n} & = \Phi^i(x^{k_r}B_{l-l_r,n}), & i = 1, \ldots, l_r. \end{array} \tag{4.1}$$

We also set $C_{0,n}(y) = A_n(1,y)$. Again, from the definition of Φ, we obtain a set of equations involving A_n, $A_{n+1} = B_{l,n}$, the $B_{i,n}$ and $C_{i,n}$. We define generating functions A and C_i $(i = 0, \ldots, l)$ as in Section 3.

Starting with $A_{n+1} = B_{l,n}$, we substitute repeatedly the values of the $B_{i,n}$ from previous equations and set $y = 1$. After a simple computation we arrive at

$$A\left((x-1)^l - zx^{k+l}\right) = (x-1)^l + z\,U(x, C_0, \ldots, C_{l-1}), \tag{4.2}$$

where U is a polynomial in the variables $x, C_0, \ldots, C_{l-1}$. Observe that the difference between (4.2) and equation (3.1) is that now U is not a concrete expression but a certain polynomial that depends on the particular values of the k_i and l_i.

Let $w_1, \ldots, w_l$ be again the power series solutions of (3.2). Substituting $x = w_j$ in (4.2) for $j = 1, \ldots, l$, we obtain a system of linear equations in the C_i. Since the coefficients are rational functions in the w_j, the solution consists also of rational functions; they are necessarily symmetric since the w_j, being conjugate roots of the same algebraic equation, are indistinguishable.

Thus we have proved the following result.

Theorem 4.1. *Let $k_1, l_1, \ldots, k_r, l_r$ be positive integers, and let $k = \sum k_i$, $l = \sum l_i$. Let q_n be the number of lattice paths from $(0,0)$ to (ln, kn) that never go above the path $(N^{k_1}E^{l_1} \cdots N^{k_r}E^{l_r})^n$, and let $w_1, \ldots, w_l$ be the unique solutions of the equation*

$$(w-1)^l - zw^{k+l} = 0$$

that are fractional power series. Then the generating function $Q(z) = \sum_{n\geq 0} q_n z^n$ *is given by*

$$Q(z) = \frac{1}{z} R(w_1, \ldots, w_l),$$

where R *is a computable symmetric rational function of* $w_1, \ldots, w_l$.

As an example, let $r = 2$ and $(k_1, l_1, k_2, l_2) = (2, 2, 1, 1)$, so that $k = l = 3$. Solving the corresponding linear system we obtain

$$R = \frac{(1-w_1)(1-w_2)(1-w_3)}{2w_1w_2w_3 - (w_1w_2 + w_1w_3 + w_2w_3)},$$

and

$$Q(z) = \frac{1}{z} R = 1 + 16z + 503z^2 + 19904z^3 + 885500z^4 + 42298944z^5 + \cdots.$$

It should be clear that for any values of the k_i and l_i the rational function R can be computed effectively.

5. Concluding Remarks

It is possible to obtain an expression for the generating function of the full Tutte polynomials $A_n(x, y)$ defined in Section 3. We have to find the values of $C_0, C_1, \ldots, C_{l-1}$ satisfying the system (3.3) and substitute back into (3.1). After some algebraic manipulation, the final expression becomes

$$\sum_{n\geq 0} A_n(x, y) z^n = \frac{-(x - w_1) \cdots (x - w_l)}{(zx^{k+l} - (x-1)^l)(y - w_1(y-1)) \cdots (y - w_l(y-1))}.$$

Taking $x = y = 1$ we recover the formula stated in Theorem 3.1.

On the other hand, references [**MS**] and [**MSV**] also study a different question on the tennis ball problem, namely to compute the sum of the labels of the balls outside the basket for all possible configurations. For a given lattice path P_n, this amounts to computing the sum of all elements in all bases of the matroid $M[P_n]$. We remark that this quantity does not appear to be computable from the corresponding Tutte polynomials alone.

Finally, as already mentioned, the technique of forcing an expression to vanish by substituting algebraic functions was introduced by Tutte in his landmark papers on the enumeration of planar maps. Thus the present paper draws in more than one way on the work of the late William Tutte.

References

[BMN] J. E. Bonin, A. de Mier, M. Noy, *Lattice path matroids: enumerative aspects and Tutte polynomials*, J. Combin. Theory Ser. A **104** (2003), 63–94.

[GM] R. Grimaldi, J. Moser, *The Catalan numbers and the tennis ball problem*, Congr. Numer. 125 (1997), 65–71.

[MS] C. L. Mallows, L. W. Shapiro, *Balls in the lawn*, J. Integer Sequences 2 (1999).

[MSV] D. Merlini, R. Sprugnoli, M. C. Verri, *The tennis ball problem*, J. Combin. Theory Ser. A 99 (2002), 307–344.

[S] R. P. Stanley, *Enumerative combinatorics, Vol. 2*, Cambridge University Press, Cambridge, 1999.

[T1] W. T. Tutte, *A contribution to the theory of chromatic polynomials*, Canadian J. Math. 6 (1954), 80–91.

[T2] W. T. Tutte, *On the enumeration of planar maps*, Bull. Amer. Math. Soc. 74 (1968), 64–74.

[W] D. J. A. Welsh, *Complexity: knots, colourings and counting*, Cambridge University Press, Cambridge, 1993.

Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya
Current address: Mathematical Institute, 24–29 St Giles', Oxford OX1 3LB, United Kingdom
E-mail address: demier@maths.ox.ac.uk

Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, Pau Gargallo 5, 08028 Barcelona, Spain
E-mail address: marc.noy@upc.es

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Decomposition of Green polynomials of type A and DeConcini-Procesi algebras of hook partitions

Hideaki Morita and Tatsuhiro Nakajima

Abstract. *A Kraśkiewicz-Weymann type theorem is obtained for the DeConcini-Procesi algebras of hook partitions. The DeConcini-Procesi algebras are graded modules of the symmetric groups, that generalize the coinvariant algebras. Defining the direct sums of the homogeneous components of the algebra in some natural way, we show that these submodules are induced from representations of the corresponding subgroup of the symmetric group. The Green polynomials of type A play an essential role in our argument.*

1. Introduction

Let $R_n = \bigoplus_{d\geq 0} R_n^d$ be the coinvariant algebra of the symmetric group S_n. For each $l \in \{1, 2, \ldots, n\}$ and for each $k = 0, 1, \ldots, l-1$, define a subspace $R_n(k;l)$ of R_n by

$$R_n(k;l) := \bigoplus_{d\equiv k \bmod l} R_n^d \ .$$

In our previous work [MN], we have shown that all $R_n(k;l)$ $(k = 0, \ldots l-1)$ are S_n-submodules of equal dimension and induced by modules of the subgroup $H_n(l)$ of S_n. Here $H_n(l)$ indicates a direct product of a cyclic group of order l generated by

$$(1, \ldots, l)(l+1, \ldots, 2l) \cdots ((d-1)l+1, \ldots, dl) \ .$$

and the symmetric group of order r, where r is a remainder of n divided by l.

In this article we consider the "DeConcini-Procesi algebras" in place of the coinvariant algebras in the preceding result. The DeConcini-Procesi algebra R_μ is a graded S_n-module parameterized by a partition μ of n. The algebra R_μ serves the generalization of R_n, since $R_\mu = R_n$ when $\mu = (1^n)$.

From the geometric point of view, DeConcini-Procesi algebras are isomorphic to the cohomology ring of the fixed point subvarieties of flag varieties. Namely, the coinvariant algebras are isomorphic to the cohomology ring of the flag varieties. Since the fixed point subvarieties are singular, we generally cannot expect nice combinatorial properties of the Betti numbers, such as unimodal symmetry. Hence we may face some difficulty when we try to achieve the results similar to the case of the coinvariant algebras. In fact, even if we collect the homogeneous components of R_μ whose degree is congruent to k modulo n for each $k = 0, \ldots, n-1$, the dimensions of them do not coincide.

However, we find that there exists a positive integer M_μ for a partition μ such that R_μ holds the above mentioned property for the coinvariant algebras for $l = 1, \ldots, M_\mu$. Essential tools to prove our main result

1991 *Mathematics Subject Classification.* 05E05, 05E10, 20E30, 33D52.
Key words and phrases. Green polynomials, DeConcini-Procesi algebras, symmetric functions.

are the "Green polynomials". Green polynomials were introduced by J. A. Green [**Gr**] for the sake of determining the irreducible characters of the general linear groups over finite fields. They also afford the graded characters of the DeConcini-Procesi algebras. In this article, we construct the standard decomposition for the Green polynomials. By applying this decomposition we prove the properties for the homogeneous components of the DeConcini-Procesi algebras associated to the hook type partitions.

2. DeConcini-Procesi algebras

Let $P_n = \mathbb{C}[x_1, \dots, x_n]$ denote the polynomial ring with n variables over $\mathbb{C}$. Then S_n acts on P_n from the left as permutations of variables as follows:

$$(\sigma f)(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)}),$$

where $\sigma \in S_n$ and $f(x_1, \dots, x_n) \in P_n$.

We introduce the homogeneous S_n-stable ideal of P_n associated to a partition $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ of n. Let $\mu' = (\mu'_1, \mu'_2, \dots)$ be the conjugate of μ. Now we designate by I_μ the ideal generated by the collection of symmetric functions

$$\left\{ e_m(x_{i_1}, \dots, x_{i_{n-k+1}}) \middle| \begin{array}{l} k = 1, \dots, \mu_1, \\ n-k+1-(\mu'_k + \mu'_{k+1} + \cdots + \mu'_{\mu_1}) < m \le n-k+1 \end{array} \right\},$$

where $e_m(x_{i_1}, \dots, x_{i_{n-k+1}})$ denotes the m-th elementary symmetric function in the variables $x_{i_1}, \dots, x_{i_{n-k+1}}$. For example, when $\mu = (2,1) \vdash 3$, then

$$I_{(2,1)} = \left\langle \begin{array}{l} e_3(x_1,x_2,x_3), e_2(x_1,x_2,x_3), e_1(x_1,x_2,x_3), \\ e_2(x_1,x_2), e_2(x_1,x_3), e_2(x_2,x_3) \end{array} \right\rangle$$

The DeConcini-Procesi algebra R_μ associated to μ is defined as a quotient algebra

$$R_\mu := P_n / I_\mu$$

[**DP, GP, T**]. It is apparent from the definition of I_μ that R_μ is a graded S_n-module. We write its homogeneous decomposition as

$$R_\mu = \bigoplus_{d \ge 0} R^d_\mu.$$

Let $\operatorname{char}_q R_\mu$ denote the graded S_n-character of R_μ defined by

$$\operatorname{char}_q R_\mu := \sum_{d \ge 0} q^d \operatorname{char} R^d_\mu ,$$

where $\operatorname{char} R^d_\mu$ is the character of S_n-submodule R^d_μ.

For $k = 0, \dots, l-1$, define

$$R_\mu(k; l) := \bigoplus_{d \equiv k \bmod l} R^d_\mu .$$

What we would like to know first is which l makes all the dimensions of $R_\mu(k;l)$ coincide. In order to answer this question, we might consider the Poincaré polynomial of R_μ, which is obtained by evaluating the graded character at the identity. The graded characters are discussed in detail in the next section. In this section we only state the following lemma that is useful for our purpose.

Lemma 2.1. *Let q be an indeterminate and $f(q) = \sum_{i \ge 0} a_i q^i \in \mathbb{C}[q]$ a polynomial in q. Let $\ell \ge 2$ be an integer and ζ_ℓ the primitive ℓ-th root of unity. Then the following conditions are equivalent:*

(a) *$f(\zeta_\ell^k) = 0$ for each $k = 1, \dots, \ell - 1$,*

(b) *The partial sums $c_k = \sum_{i \equiv k \bmod \ell} a_i$ $(k = 0, 1, \dots, \ell - 1)$ of coefficients of the polynomial $f(q)$ are independent of the choice of k.*

3. Green polynomials

The graded characters of the DeConcini-Procesi algebras are known as the Green polynomials (of type A). For $\rho \vdash n$, let X^μ_ρ be the coefficient polynomials of the Hall-Littlewood symmetric functions $P_\mu(x;t)$ in the power sum product $p_\rho(x)$, i.e.,

$$p_\rho(x) = \sum_{\mu \vdash n} X^\mu_\rho(t) P_\mu(x;t).$$

The Green polynomials $Q^\mu_\rho(q)$ [**Gr, Mc**] are defined by

$$Q^\mu_\rho(q) = q^{n(\mu)} X^\mu_\rho(q^{-1}),$$

where $n(\mu) = \sum_{i \ge 1}(i-1)\mu_i$ for $\mu = (\mu_1, \mu_2, \ldots)$.

The Green polynomials have another expression using the modified Kostka polynomials $\tilde{K}_{\lambda\mu}(q)$. [**Mc**]. If we denote by χ^λ_ρ the value of an irreducible character χ^λ for S_n at the element of cycle-type ρ, then

$$Q^\mu_\rho(q) = \sum_{\lambda \vdash n} \chi^\lambda_\rho \tilde{K}_{\lambda\mu}(q).$$

Since the coefficient of q^d in the polynomial $\tilde{K}_{\lambda\rho}$ is also known as the multiplicity of the irreducible component V^λ in R^d_μ (see e.g., [GP]), we find that the Green polynomial $Q^\mu_\rho(q)$ affords the value of the graded character of R_μ at the element of cycle-type ρ, i.e.,

$$Q^\mu_\rho(q) = \mathrm{char}_q\, R_\mu(\rho).$$

In addition, applying the combinatorial expression of the Kostka polynomials [**LS**]

$$\tilde{K}_{\lambda\mu}(q) = \sum_{T \in \mathrm{SSTab}_\mu(\lambda)} q^{\mathrm{coch}(T)} \ ,$$

it immediately follows that

$$[R^d_\mu : V^\lambda] = \sharp\{T \in \mathrm{SSTab}_\mu(\lambda) \mid \mathrm{coch}(T) = d\}.$$

Here $\mathrm{SSTab}_\mu(\lambda)$ denotes the set of semistandard Young tableaux of shape λ with weight μ, and $\mathrm{coch}(T)$ the cocharge of a tableau T.

Some explicit forms of $Q^\mu_\rho(q)$ have been known when μ takes some special partitions. Before we expose them, we give some symbols that appear in the explicit expressions. For each partition $\rho = (1^{m_1}2^{m_2}\cdots n^{m_n})$ of n, we define

$$\begin{aligned} M_\rho &= \max\{m_1, \ldots, m_n\} \ , \\ b_\rho(q) &= \prod_{i \ge 1}(1-q)(1-q^2)\cdots(1-q^{m_i}) \ , \\ e_\rho(q) &= (1-q)^{m_1}(1-q^2)^{m_2}\cdots(1-q^n)^{m_n} \ . \end{aligned}$$

In the case of $\mu = (1^n)$, that corresponds to the graded character of the coinvariant algebra, $Q^{(1^n)}_\rho(q)$ can be expressed as follows (see e.g., [G]).

Proposition 3.1. *For* $\rho \vdash n$

$$\mathrm{char}_q\, R_n(q) = Q^{(1^n)}_\rho(q) = \frac{(1-q)(1-q^2)\cdots(1-q^n)}{e_\rho(q)}.$$

If $\mu = (2, 1^{n-2}) \vdash n$ (a hook type), then A. Morris gives Q^μ_ρ explicitly [**Mr**].

Proposition 3.2 (Morris)**.** *For* $\rho = (1^{r_1}2^{r_2}\cdots n^{r_n})$,

$$Q_\rho^{(2,1^{n-2})}(q) = \frac{(1-q)\cdots(1-q^{n-2})}{e_\rho(q)}\{(r_1-1)q^n - r_1q^{n-1}+1\}.$$

In both of the cases above, we find that Green polynomial $Q_\rho^\mu(q)$ can be decomposed into the rational factor

$$\frac{(1-q)\cdots(1-q^{M_\mu})}{e_\rho(q)}$$

and the polynomial factor. This fact holds for the Green polynomial in general and we expose it in the following theorem. Note that the theorem plays an important role to prove our main result.

Theorem 3.3. *Let μ and ρ be partitions of n. Then there exists a polynomial $G_\rho^\mu(q) \in \mathbf{Z}[q]$ such that*

$$Q_\rho^\mu(q) = \frac{b_{(1^{M_\mu})}(q)}{e_\rho(q)}G_\rho^\mu(q).$$

The theorem is deduced from an expansion of the product of two Hall-Littlewood functions $P_{\bar\mu}P_{(1^r)}$ by P_λ's, where r is the last part of μ' and $\bar\mu = (\mu'\backslash(r))' \vdash n-r$. Applying Lemma 2.1 to Theorem 3.3, we obtain the answer to the question in the previous section.

Corollary 3.4. *Let μ be a partition of n. For any $l \in \{2,\ldots,M_\mu\}$ and for each $k = 0,1,\ldots,l-1$, ζ_l^k is zero of $Q_{(1^n)}^\mu(q)$. Hence the dimensions of $R_\mu(k;l)$ $(k=0,\ldots,l-1)$ are equal each other.*

When the partition $\mu \vdash n$ is of a hook type, we can obtain an explicit expression of $G_\rho^\mu(q)$, that generalizes Morris' formula in Proposition 3.2.

Proposition 3.5. *Let $\mu = (n-h,1^h) \vdash n$ be a hook type partition. Then, for $\rho = (1^{r_1}2^{r_2}\cdots n^{r_n})$,*

$$Q_\rho^\mu(q) = \frac{(1-q)\cdots(1-q^h)}{e_\rho(q)}G_\rho^\mu(q)\ ,$$

where

$$G_\rho^\mu(q) = (1-q^n) - \sum_{k=1}^{n-h-1}\sum_{\tau=(1^{t_1}\cdots k^{t_k})\vdash k}\binom{r_1}{t_1}\cdots\binom{r_k}{t_k}q^{n-k}e_\tau(q).$$

4. DeConcini-Procesi algebras of hook type

In this section we show that, for a hook partition μ, the DeConcini-Proseci algebra R_μ holds the property similar to the coinvariant algebra R_n, i.e., there is a subgroup of S_n and its modules with equal rank such that all $R_\mu(k;l)$ are induced by the modules.

Let $\mu = (n-h,1^h)$ be a hook partition of n. We suppose $n-h>1$ below, since R_μ is the coinvariant algebra if $n-h=1$. In this case $M_\mu = h$ and it follows from Corollary 3.4 that for each $l \in \{(1,)2,\ldots h\}$ the dimensions of $R_\mu(k;l)$ $(k=0,\ldots,l-1)$ coincide. When $\mu=(n-h,1^h)$ and $l \in \{2,\ldots h\}$, we denote by d and r the quotient and the remainder of h divided by l. Set $\bar\mu = (n-h,1^r)$ and $\nu = (1^{dl})$.

Let C_l be a cyclic subgroup of S_n generated by an element

$$a = (1,2,\ldots,l)(l+1,l+2,\ldots,2l)\cdots((d-1)l+1,(d-1)l+2,\ldots,dl)\ ,$$

where $(i,i+1,\ldots,i+l)$ is a cyclic permutation of length l and S_{n-dl} a subgroup of S_n defined by

$$S_{n-dl} := \{\sigma \in S_n | \sigma(i) = i \quad \text{for } i = 1,2,\ldots,dl\}.$$

We should prove the following theorem to obtain the result mentioned above. This is our main result in this article.

Theorem 4.1. *Let $\mu = (n-h, 1^h)$ $(n-h > 1)$ be a hook type partition of n. For an integer $l \in \{2, 3, \ldots, h\}$, let $h = dl + r$ $(0 \geq r \geq l-1)$. We set $\bar{\mu} = (n-h, 1^r)$. Then there is an $S_n \times C_l$-module isomorphism*

$$R_\mu \cong R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} .$$

In the theorem above, the $S_n \times C_l$-module structures are defined as follows. The algebra R_μ is regarded as an $S_n \times C_l$-module. The group C_l acts on R_μ by

$$ax = \zeta_l^d x \quad (x \in R_\mu^d).$$

The induced module

$$R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} = \bigoplus_{\sigma \in S_n / S_{n-dl}} \sigma \otimes R_{\bar{\mu}}$$

is also regarded as an $S_n \times C_l$-module. The action of C_l is defined by

$$a(\sigma \otimes x) = \sigma a^{-1} \otimes ax = \zeta_l^d \sigma a^{-1} \otimes x \quad (x \in R_{\bar{\mu}}^d).$$

Comparing the ζ_l^k-eigenspaces of $a \in C_l$ in both sides, we have the very thing that we would like to prove.

Corollary 4.2. *Let $\mu = (n-h, 1^h)$ $(n-h > 1)$ be a hook type partition of n. If we choose an integer $l \in \{2, 3, \ldots, h\}$, then there is an S_n-module isomorphism*

$$R_\mu(k; l) \cong Z_\mu(k; l) \Big\uparrow_{C_l \times S_{n-dl}}^{S_n}$$

for each $k = 0, 1, \ldots, l-1$. We define here the representation $Z_\mu(k; l)$ of $C_l \times S_{n-dl}$ by

$$Z_\mu(k; l) := \bigoplus_{\lambda \vdash n-dl} \bigoplus_{T \in \mathrm{SSTab}_{\bar{\mu}}(\lambda)} \psi^{(k - \mathrm{coch}(T))} \otimes V^\lambda ,$$

where $\psi^{(s)} : a \mapsto \zeta_l^s$ denotes an irreducible representation of $C_l = \langle a \rangle$, and V^λ a irreducible representation of S_{n-dl} associated to the partition λ of $n - dl$.

5. Outline of the proof of Theorem 4.1

We will show that

$$\mathrm{char}\, R_\mu(w, a^j) = \mathrm{char}\, R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} (w, a^j)$$

for each $(w, a^j) \in S_n \times C_l$.

First we calculate the right-hand side of the above identity. If $\mathrm{char}\, R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} (w, a^j) \neq 0$, then there exists a basis element $\sigma \otimes x$ of $\mathrm{char}\, R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} (w, a^j)$ such that $(w, a^j)(\sigma \otimes x)|_{\sigma \otimes x} \neq 0$. Since $(w, a^j)(\sigma \otimes x) = w\sigma a^{-j} \otimes a^j x = \zeta_l^{dj} w \sigma a^{-j} \otimes x$, it should follow that $w\sigma a^{-j} \equiv \sigma \bmod S_{n-dl}$ if $\mathrm{char}\, R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} (w, a^j) \neq 0$. Thus we have the following lemma:

Lemma 5.1. *If $\mathrm{char}\, R_{\bar{\mu}} \Big\uparrow_{S_{n-dl}}^{S_n} (w, a^j) \neq 0$, then $w \sim a^j v$ for some $v \in S_{n-dl}$*

If $w\sigma a^{-j} \equiv \sigma \bmod S_{n-dl}$, then $w\sigma a^{-j} = \sigma\tau$ for some $\tau \in S_{n-dl}$. Setting

$$\mathcal{S}_\tau^{(j)}(w) := \{\sigma \in S_n / S_{n-dl} | w\sigma a^{-j} = \sigma\tau\}$$
$$\mathcal{S}^{(j)}(w) := \{\sigma \in S_n / S_{n-dl} | w\sigma a^{-j} \equiv \sigma \bmod S_{n-dl}\},$$

we have

$$\begin{aligned}\operatorname{char} R_{\bar\mu}\Big\uparrow_{S_{n-dl}}^{S_n}(w,a^j) &= \sum_{\tau\in S_{n-dl}} \sharp\mathcal{S}_\tau^{(j)}(w)\operatorname{char} R_{\bar\mu}(\tau)\big|_{q=\zeta_l^j} \\ &= \sum_{\tau\in S_{n-dl}} \sharp\mathcal{S}_\tau^{(j)}(w)\operatorname{char} R_{\bar\mu}(v)\big|_{q=\zeta_l^j} \\ &= \sharp\mathcal{S}^{(j)}(w)\operatorname{char} R_{\bar\mu}(v)\big|_{q=\zeta_l^j}.\end{aligned}$$

(Note that $v\sim\tau$ if $w=a^jv$ and $w\sigma a^{-j}=\sigma\tau$.) We can enumerate the permutations belonging to the set $\mathcal{S}^{(j)}(w)$:

Proposition 5.2. *For $w=a^jv\in S_n$,*

$$\sharp\mathcal{S}^{(j)}(w)=p^e e!\binom{z_p+e}{e},$$

where $\lambda(a^j)=(p^e)$, $\lambda(v)=(1^{z_1}2^{z_2}\cdots)$.

Let us summarize the above arguments:

Proposition 5.3. *Suppose $\lambda(a^j)=(p^e)$. Then*

$$\operatorname{char} R_{\bar\mu}\Big\uparrow_{S_{n-dl}}^{S_n}(w,a^j)=\begin{cases} p^e e!\binom{z_p+e}{e}\operatorname{char}_q R_{\bar\mu}(v)\big|_{q=\zeta_l^j}, & \text{if } w\sim a^jv \text{ for some } v\in S_{n-dl}\\ 0, & \text{otherwise.}\end{cases}$$

On the other hand, for the left-hand side of the identity that we are now proving, it follows from Theorem 3.3 the following lemma:

Lemma 5.4. *If $Q^\mu_{\lambda(w)}(\zeta_l^j)\neq 0$ for some $w\in S_n$, then $w\sim a^jv$ for some $v\in S_{n-dl}$.*

Applying decomposition of the Green polynomials in Theorem 3.3 to some recursive relations of them, we can obtain

$$\operatorname{char}_q R_\mu(w)\Big|_{q=\zeta_l^j}=Q^\mu_{\lambda(w)}(\zeta_l^j)=p^e e!\binom{z_p+e}{e}Q^{\bar\mu}_{\lambda(v)}(\zeta_l^j).$$

We have thus shown the identity of two characters, thereby completing the proof of our main theorem.

References

[DP] C. DeConcini and C. Procesi, Symmetric functions, conjugacy classes, and the flag variety, *Inv. Math.* **64** (1981), 203-230.

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

[G] A. M. Garsia, Combinatorics of the free Lie algebra and the symmetric group, in *Analysis, et cetera..., Jürgen Moser Festschrift.* Academic Press, New York, 1990, pp. 309–382.

[Gr] J. A. Green, The character of the finite general linear groups, *Trans. Amer. Math. Soc.*, **80** (1955), 402-447.

[KW] W. Kraśkiewicz and J. Weymann, Algebra of coinvariants and the action of Coxeter elements, *Bayreuth. Math. Schr.* **63** (2001), 265–284.

[LS] A. Lascoux and M. P. Schützenberger, Sur une conjecture de H. O. Foulkes, *C. R. Acad. Sci. Paris.* **286** (1978), 323-324.

[Mc] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford University Press, 1995.

[Mr] A. O. Morris, The characters of the group $GL(n,q)$, *Math. Z.* **1** (1963), 112-123.

[MN] H. Morita and T. Nakajima, The coinvariant algebra of the symmetric group as a direct sum of induced modules, submitted to *Osaka J. Math.*

[R] C. Reutenauer, *Free Lie Algebras*, Oxford University Press, Oxford, 1993.

[T] T. Tanisaki, Defining ideals of the closures of conjugacy classes and representations of the Weyl groups, *Tohoku J. Math.* **34** (1982), 575-585.

School of Science, Tokai University, Hiratsuka 259-1292, Japan
E-mail address: moritah@msg.biglobe.ne.jp

Faculty of Economics, Meikai University, Urayasu 279-8550, Japan
E-mail address: t.nkjm@ny.thn.ne.jp

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

On Computing the Coefficients of Rational Formal Series

Paolo Massazza and Roberto Radicioni

Abstract. *In this work we study the problem of computing the coefficients of rational formal series in two commuting variables. Given a rational formal series* $\phi(x,y) = \sum_{n,k\geq 0} c_{nk}x^n y^k = P(x,y)/Q(x,y)$ *with* $P, Q \in \mathbb{Q}[\{x,y\}]$ *and* $Q(0,0) \neq 0$*, we show that the coefficient* $[x^i y^j]\phi(x,y)$ *can be computed in time* $O(i+j)$ *under the uniform cost criterion.*

RÉSUMÉ. *Dans cet article, nous étudions le problème du calcul des coefficients de séries formelles rationelles en deux variables commutatives. Etant donné une série formelle rationnelle* $\phi(x,y) = \sum_{n,k\geq 0} c_{nk}x^n y^k = P(x,y)/Q(x,y)$ *ou* $P, Q \in \mathbb{Q}[\{x,y\}]$ *et* $Q(0,0) \neq 0$*, nous montrons que le coefficient* $[x^i y^j]\phi(x,y)$ *peut être calculé en un temps* $O(i+j)$ *sous le critère de coût uniforme.*

1. Introduction

The problem of computing the coefficients of formal power series (known as the *Coefficient Problem*) is of primarly interest in many different areas such as combinatorics and theory of languages. For example, the problem of counting objects with a given property that belong to a combinatorial structure S can be easily reduced to computing the coefficients of suitable formal power series: the property is codified into a weight function $w : S \to \mathbb{N}$ and the formal series $\sum_{s\in S} w(s)\, s$ is considered. Then, the counting problem associated with S and w consists of computing the function $f(n) = \sharp\{s \in S | w(s) = n\}$.

Another setting where the Coefficient Problem arises is the random generation of combinatorial structures (see, for instance, [**10**]). Efficient algorithms for the random generation of strings in a language can also be derived by exploiting the generating function associated with the language (see, for example, [**7**]).

A likewise important and (intuitively) more general version of the Coefficient Problem can be stated considering a multivariate formal series in commutative variables. More precisely, the Coefficient Problem for a class $\mathcal{A}$ of commutative formal series consists of computing, given a series in k variables $f = \sum_{\underline{n}\in\mathbb{N}^k} c_{\underline{n}}\underline{x}^{\underline{n}} \in \mathcal{A}$ and a multi-index $\underline{i} \in \mathbb{N}^k$, the coefficient $c_{\underline{i}}$ of f. When dealing with counting and random generation, this generalization appears whenever a multiple output weight function $w : S \to \mathbb{N}^k$ is considered. Some examples are the problem of counting and random sampling words with fixed occurrences of each letter of the alphabet ([**2**],[**8**]) or the random generation through object grammars ([**9**]).

In this paper we consider the Coefficient Problem for the class $\mathbb{Q}[[\{x,y\}]]_r$ of the rational formal series in two commuting variables. These are power series expansions of functions of the form $P(x,y)/Q(x,y)$ where P, Q are polynomials with rational coefficients and $Q(0,0) \neq 0$.

We show that given a couple of integers (i,j) and a couple of polynomials $P, Q \in \mathbb{Q}[\{x,y\}]$, with $Q(0,0) \neq 0$, the coefficient $[x^i y^j]\phi(x,y)$ of $\phi(x,y) = P(x,y)/Q(x,y)$ can be computed in time $O(i+j)$ under

Key words and phrases. holonomic functions, shift algebra.

This work has been supported by the project M.I.U.R. COFIN "Formal Languages and Automata: Theory and Applications".

the uniform cost criterion. Our method is based on the theory of holonomic power series. We derive suitable recurrence equations with polynomial coefficients from the holonomic system associated with $\phi(x,y)$, then we use them, together with a recurrence with constant coefficients, in order to compute $[x^i y^j]\phi(x,y)$ in an efficient way. This is a significant improvement on a more general algorithm presented in [**11**], that let us to compute $[x^i y^j]\phi(x,y)$ in time $O(i \cdot j)$.

2. Preliminaries

We denote by $\mathbb{N}$ ($\mathbb{Q}$) the set of the natural (rational) numbers. A 2-dimensional sequence c with values in $\mathbb{Q}$ is a function $c : \mathbb{N}^2 \mapsto \mathbb{Q}$, usually denoted by $\{c_{nk}\}$. We denote by $\mathbb{Q}^{(2)}$ the ring of 2-dimensional sequences on $\mathbb{Q}$ with the operations of sum, $\{a_{nk}\} + \{b_{nk}\} = \{a_{nk} + b_{nk}\}$ and product, $\{a_{nk}\} \cdot \{b_{nk}\} = \{c_{nk}\}$ where $c_{nk} = \sum_{\substack{l+m=n \\ i+j=k}} a_{li} b_{mj}$.

Moreover, we consider the following operators from $\mathbb{Q}^{(2)}$ into $\mathbb{Q}^{(2)}$:

- External product by $e \in \mathbb{Q}$: $e \cdot \{a_{nk}\} = \{e a_{nk}\}, \quad e \in \mathbb{Q}$
- Shift : $E_n\{a_{nk}\} = \{a_{n-1\,k}\}, \quad E_k\{a_{nk}\} = \{a_{n\,k-1}\}$
- Multiplication by n, k : $n\{a_{nk}\} = \{n a_{nk}\}, \quad k\{a_{nk}\} = \{k a_{nk}\}$

Then, the so called *shift algebra* $\mathbb{Q}\langle n, k, E_n, E_k\rangle$ is a particular Ore algebra (see, for instance, [**6**]) and can be interpreted as a (noncommutative) ring of linear operators on $\mathbb{Q}^{(2)}$, with pseudo-commutative rules given by:

$$nk = kn, \quad nE_k = E_k n, \quad kE_n = E_n k,$$
$$nE_n = E_n n + E_n, \quad kE_k = E_k k + E_k.$$

More simply, a polynomial in $\mathbb{Q}\langle n, k, E_n, E_k\rangle$ represents a linear recurrence with polynomial coefficients.

2.1. Rational formal series and Holonomic functions.

Let Σ^c be the commutative free monoid generated by a finite alphabet Σ. Given a commutative ring $\mathbb{K}$, a *formal series* ψ in commutative variables Σ is a function $\psi : \Sigma^c \mapsto \mathbb{K}$, usually indicated by $\sum_{\underline{x} \in \Sigma^c} \psi(\underline{x})\underline{x}$; the *support* of ψ is the set of monomials $\{\underline{x} \in \Sigma^c \,|\, \psi(\underline{x}) \neq 0\}$. We denote by $\mathbb{K}[[\Sigma]]$ the ring of commutative formal series with coefficients in $\mathbb{K}$ equipped with the usual operations of sum $(+)$ and product $(\cdot)$. Formal series with finite support belong to the ring of polynomials $\mathbb{K}[\Sigma]$. The ring of rational formal series $\mathbb{K}[[\Sigma]]_r$ can be defined as the smallest subring of $\mathbb{K}[[\Sigma]]$ containing $\mathbb{K}[\Sigma]$ and rationally closed (i.e. closed with respect to $\star$, $+$, $\cdot$ and the two external products of $\mathbb{K}$ on $\mathbb{K}[[\Sigma]]$ — where $\star$ is the usual closure operation that is defined for proper series, i.e. series ψ s.t. $\psi(\epsilon) = 0$).

In the sequel we will consider the alphabet $X = \{x, y\}$ and $\mathbb{K} = \mathbb{Q}$. A rational formal series $\phi \in \mathbb{Q}[[X]]_r$ is then the power series expansion of a suitable rational function,

$$\phi(x,y) = \sum_{n,k \in \mathbb{N}} c_{nk} x^n y^k = \frac{P(x,y)}{Q(x,y)} \quad P, Q \in \mathbb{Q}[X],\ Q(0,0) \neq 0.$$

We often use the notation $[x^n y^k]\phi(x,y)$ to indicate the coefficient c_{nk} of a formal series ϕ. We refer to [**3**] for a detailed analysis of the class of the rational series.

It is well known (see, for example, [**14**]) that the class of the rational functions is properly contained in the class of the holonomic functions defined as follows.

Definition 2.1. A function $\phi(x,y)$ is *holonomic* iff there exist some polynomials

$$p_{ij} \in \mathbb{Q}[X], \qquad 1 \leq i \leq 2, \quad 0 \leq j \leq d_i, \quad p_{id_i} \neq 0$$

such that

$$\sum_{j=0}^{d_1} p_{1j} \partial_x^j \phi = 0, \qquad \sum_{j=0}^{d_2} p_{2j} \partial_y^j \phi = 0.$$

The above equations are said to be a *holonomic system* for ϕ.

Holonomic systems were first introduced by I.N. Bernstein in the 1970s ([**1**]) and deeply investigated by Stanley, Lipshitz, Zeilberger et al. (see [**5**], [**11**], [**13**] and [**14**]). In this setting, we are interested in the following result:

Theorem 2.2. *Let $\phi(x,y)=\sum_{n,k\geq 0} c_{nk}x^n y^k$ be a holonomic function. Then the sequence of coefficients $\{c_{nk}\}$ satisfies a system of linear recurrence equations with polynomial coefficients*

$$S=\begin{cases} P(n,k,E_n)\{c_{nk}\}=0 \\ Q(n,k,E_k)\{c_{nk}\}=0 \end{cases} \tag{2.1}$$

where $P(n,k,E_n)=\sum_{i=0}^{r} p_i(n,k){E_n}^i$ and $Q(n,k,E_k)=\sum_{j=0}^{s} q_j(n,k){E_k}^j$ belong to $\mathbb{Q}\langle n,k,E_n,E_k\rangle$.

Proof. See, for instance, [**11**]. □

A direct consequence of the previous theorem is that the sequence of the coefficients of a rational formal series satisfies a system of recurrence equations of type (2.1): we give here an outline of how to compute such a system.

In [**14**] it is shown how to compute a holonomic system $\{D_1,D_2\}$ associated with a rational function $\phi(x,y)$. Then, given $\{D_1,D_2\}$, we can obtain in two steps a system of recurrences $S=\{P,Q\}$ of type (2.1) satisfied by the sequence $\{c_{nk}\}$. First, we compute two operators $w_1,w_2\in\mathbb{Q}\langle n,k,E_n,E_k\rangle$ such that $w_1(n,E_n,E_k)\{c_{n,k}\}=w_2(k,E_n,E_k)\{c_{n,k}\}=0$. This is easily done by observing the following correspondence between operators

$$x^r y^s \partial_x^i \partial_y^j \equiv \left(\prod_{h=1}^{i}(n-r+h)\prod_{h=1}^{j}(k-s+h)\right)E_n^{r-i}E_k^{s-j}.$$

Then, we get the first recurrence $P(n,k,E_n)$ by solving an elimination problem in $\mathbb{Q}\langle n,k,E_n,E_k\rangle$. This can be done, for example, by computing the Gröbner basis associated with w_1, w_2 with respect to a suitable ordering on $\mathbb{Q}\langle n,k,E_n,E_k\rangle$. We proceed similarly in order to get the second recurrence $Q(n,k,E_k)$. Useful packages for such computations have been recently developed (see [**6**], [**12**]).

We recall that the coefficients of a rational series satisfy a linear recurrence with constant coefficients. More formally, we have the following:

Theorem 2.3. *Let be $\phi(x,y)=\sum_{n,k\geq 0} c_{nk}x^n y^k = P(x,y)/Q(x,y)$ with $P,Q\in\mathbb{Q}[X]$ and $Q(0,0)\neq 0$. Then the sequence of coefficients $\{c_{nk}\}$ satisfies a linear recurrence equation with constant coefficients*

$$B(E_n,E_k)\{c_{nk}\}=0 \tag{2.2}$$

where $B(E_n,E_k)=\sum_{i,j=0}^{r\,s} q_{ij}{E_n}^i{E_k}^j$, $q_{ij}\in\mathbb{Q}$ and $q_{00}\neq 0$.

Proof. Let be $P(x,y)=\sum_{i,j=0}^{r_1,s_1} p_{ij}x^i y^j$ and $Q(x,y)=\sum_{i,j=0}^{r_2,s_2} q_{ij}x^i y^j$. Then we have

$$Q(x,y)\phi(x,y)=\sum_{n,k\geq 0}\left(\sum_{\substack{i_1+i_2=n\\ j_1+j_2=k}} q_{i_1j_1}c_{i_2j_2}\right)x^n y^k=\sum_{i,j=0}^{r_1,s_1} p_{ij}x^i y^j.$$

So, for $n>r_1$ and $k>s_1$ it holds $\displaystyle\sum_{i,j=0}^{r_2,s_2} q_{ij}c_{n-ik-j}=\sum_{i,j=0}^{r_2,s_2} q_{ij}{E_n}^i{E_k}^j c_{nk}=0.$ □

A naive method for computing the coefficient c_{ij} of a rational series in time $O(ij)$ can be obtained as an immediate application of the theorem above. Linear recurrences with constant coefficients have been deeply studied in [**4**], where it is shown a necessary and sufficient condition that let us to define a well ordering on the elements of the solution. Moreover, it is also shown that the solutions of such recurrences can have generating functions that are not rational (they can be algebraic, holonomic or even non-holonomic).

3. Computing the coefficient

Given a rational formal series $\phi = \sum_{n,k\in\mathbb{N}} c_{nk}x^n y^k$, by applying Theorem 2.3 we can easily obtain a linear recurrence equation with constant coefficients satisfied by $\{c_{nk}\}$. Then, we can use it for computing an arbitrary coefficient c_{ij} once a suitable set of initial conditions is known. On the other hand, because in general both E_n and E_k appear in the recurrence, this technique requires $O(ij)$ coefficients in order to determine c_{ij}.

As shown before, the theory of holonomic systems let us to obtain particular linear recurrence equations with polynomial coefficients that are more suitable for computing coefficients. More precisely, we get two operators in the shift algebra $\mathbb{Q}\langle n, k, E_n, E_k\rangle$ that depend on n, k and either E_n or E_k. So, we can efficiently compute all the coefficients along a line $n = \overline{n}$ or $k = \overline{k}$ if the leading and the least coefficients of the recurrences do not vanish on that line.

Our approach takes advantage of both types of recurrences in order to get a method that efficiently computes the coefficient c_{ij} by starting with a suitable set of initial conditions and proceeding by choosing at each step the "right" recurrence to use.

More formally, we consider the Coefficient Problem for rational series defined as follows.

Problem: to determine the coefficient c_{ij} in the power series expansion of a rational series $\phi(x, y)$.

Input: A tuple $\langle \mathcal{N}, \mathcal{K}, \mathcal{B}, I, i, j\rangle$ where:

-: $\mathcal{N}$, $\mathcal{K}$ and $\mathcal{B}$ are three recurrence equations of type

$$\mathcal{N}(n,k,E_n) = \sum_{i=0}^{r} p_i(n,k){E_n}^i \qquad p_i(n,k) \in \mathbb{Q}[\{n,k\}], \quad p_r(n,k) \neq 0 \tag{3.1}$$

$$\mathcal{K}(n,k,E_k) = \sum_{j=0}^{s} q_j(n,k){E_k}^j \qquad q_j(n,k) \in \mathbb{Q}[\{n,k\}], \quad q_s(n,k) \neq 0 \tag{3.2}$$

$$\mathcal{B}(E_n,E_k) = \sum_{i,j=0}^{r_2,s_2} h_{ij}{E_n}^i{E_k}^j \qquad h_{ij} \in \mathbb{Q}, \quad h_{00} \neq 0 \tag{3.3}$$

satisfied by the sequence $\{c_{nk}\}$.

-: I is a suitable set of initial conditions for $\mathcal{N}$, $\mathcal{K}$, $\mathcal{B}$, i.e. the set of coefficients

$$I = \{c_{nk} \mid (0 \le n \le \alpha \wedge 0 \le k \le e) \vee (0 \le n \le e \wedge 0 \le k \le \beta)\}$$

with
- $e = \max\{r, s, r_2, s_2\}$
- $\alpha = \max\{n \in \mathbb{N} \mid p_r(n,k) = 0 \vee p_0(n,k) = 0, 0 \le k \le e\}$
- $\beta = \max\{k \in \mathbb{N} \mid q_s(n,k) = 0 \vee q_0(n,k) = 0, 0 \le n \le e\}$

-: $(i,j) \in \mathbb{N}^2$.

Output: the coefficient c_{ij}.

3.1. Clusters and coefficients.

In this section we give some definitions and prove some basic results that are useful to describe the behaviour of the algorithm.

Definition 3.1 ($R^{(e)}(x,y)$)**.** Let be $e, x, y \in \mathbb{N}$. The *square* $R^{(e)}(x,y)$ is the set of points

$$R^{(e)}(x,y) = \{(x',y') \mid (x',y') \in \mathbb{N}^2, x - e < x' \le x \ \wedge \ y - e < y' \le y\}.$$

Note that each square is identified by the point on the right upper corner.

Definition 3.2 ($\mathrm{SQ}^{(e)}$)**.** Let be $e \in \mathbb{N}$. Then

$$\mathrm{SQ}^{(e)} = \{R^{(e)}(x,y) \mid (x,y) \in \mathbb{N}^2, \ x \ge e, \ y \ge e\}.$$

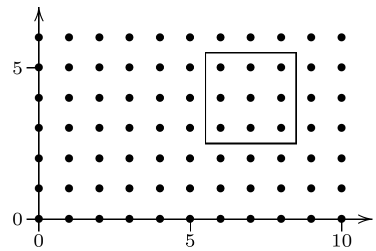


Figure 1. The square $R^{(3)}(8,5)$.

We give a notion of neighbor of a square introducing the following partial functions from $\mathrm{SQ}^{(e)}$ to $\mathrm{SQ}^{(e)}$:

$$\begin{array}{rcll} N(R^{(e)}(x,y)) & = & R^{(e)}(x,y+e) & \\ S(R^{(e)}(x,y)) & = & R^{(e)}(x,y-e) & \text{(defined if } y \geq 2e) \\ E(R^{(e)}(x,y)) & = & R^{(e)}(x+e,y) & \\ W(R^{(e)}(x,y)) & = & R^{(e)}(x-e,y) & \text{(defined if } x \geq 2e) \end{array}$$

Let be $T \in \{N,E,S,W\}$, then we often write $T^i(R)$ for $T(T^{i-1}(R))$, $T^0(R) = R$. Moreover, we also consider the shortcuts $SW(R) = S(W(R))$, $NW(R) = N(W(R))$, $SE(R) = S(E(R))$ and $NE(R) = N(E(R))$.

Definition 3.3 ($\mathrm{SQ}^{(e)}(\overline{R})$, $\mathrm{SQ}_V^{(e)}(\overline{R})$)**.** Let be $e \in \mathbb{N}$ and $V \subseteq \mathbb{N}^2$. Given $\overline{R} \in \mathrm{SQ}^{(e)}$ such that $\overline{R} = N^c(W^d(R^{(e)}(e,e)))$ $(c,d \in \mathbb{N})$ we define

$$\begin{array}{rcl} \mathrm{SQ}^{(e)}(\overline{R}) & = & \left\{R \in \mathrm{SQ}^{(e)} \mid \exists u,v \in \mathbb{N}^2, R = W^u(S^v(\overline{R}))\right\} \\ \mathrm{SQ}_V^{(e)}(\overline{R}) & = & \{R \in \mathrm{SQ}^{(e)}(\overline{R}) \mid R \cap V \neq \emptyset\} \end{array}$$

We introduce a reflexive and symmetric relation $\diamond \subset \mathrm{SQ}^{(e)} \times \mathrm{SQ}^{(e)}$:

Definition 3.4 ($\diamond$)**.** $R_1^{(e)}$ is a *neighbor* of $R_2^{(e)}$, $R_1^{(e)} \diamond R_2^{(e)}$, if and only if

$$\exists T \in \{N, NE, E, SE, S, SW, W, NW\} \quad \text{s.t.} \quad R_1^{(e)} = T(R_2^{(e)}).$$

Particular sequences of squares will be of interest when considering the behaviour of the algorithm.

Definition 3.5. Let $\mathrm{Seq} = R_1, \ldots, R_k$ be a sequence of squares in $\mathrm{SQ}^{(e)}$. Then, Seq is

- *8-connected* iff for $1 \leq i < k$ it holds $R_i \diamond R_{i+1}$
- *4-connected* iff for $1 \leq i < k$ it holds $R_{i+1} = T_i(R_i)$ with $T_i \in \{N,E,S,W\}$
- *descending* iff Seq is 8-connected or 4-connected and for $1 \leq i < k$ it holds $R_{i+1} = T_i(R_i)$ with $T_i \in \{E, SE, S, SW, W\}$
- *ascending* iff Seq is 8-connected or 4-connected and for $1 \leq i < k$ it holds $R_{i+1} = T_i(R_i)$ with $T_i \in \{W, NW, N, NE, E\}$

Henceforward, we fix an instance $\langle \mathcal{N}, \mathcal{K}, \mathcal{B}, I, i, j\rangle$ of the Coefficient Problem for a series $\phi(x,y) \in \mathbb{Q}[[X]]_r$ and we associate with it the following values:

- $Z = Z(\mathcal{N},\mathcal{K}) = \{(x,y) \in \mathbb{N}^2 \mid p_r(x,y) = 0 \vee p_0(x,y) = 0 \vee q_s(x,y) = 0 \vee q_0(x,y) = 0\}$.
- $e = e(\mathcal{N},\mathcal{K}) = \max\{r, s, r_2, s_2\}$
- $R_0 = R^{(e)}(e-1, e-1)$
- $\overline{R} = \overline{R}(\bar{\imath},\bar{\jmath}) = N^c(E^d(R_0))$ with $c = \lfloor j/e \rfloor$ $d = \lfloor i/e \rfloor$.

Once we have fixed the values above, we can write R for $R^{(e)}$ and SQ for $\mathrm{SQ}^{(e)}$ whenever the context is clear. Note that $\mathrm{SQ}_Z(\overline{R})$ consists of those squares in $\mathrm{SQ}(\overline{R})$ that contain at least one point (n,k) such that at least one of the following methods fails:

Compute$_{\mathcal{N}}$$(n,k)$: use $\mathcal{N}$ to compute $[x^n y^k]\phi$ from the values $[x^{n-l}y^k]\phi$ or $[x^{n+l}y^k]\phi$, $1 \le l \le r$

Compute$_{\mathcal{K}}$$(n,k)$: use $\mathcal{K}$ to compute $[x^n y^k]\phi$ from the values $[x^n y^{k-l}]\phi$ or $[x^n y^{k+l}]\phi$, $1 \le l \le s$

The next lemma will be frequently used in the sequel.

Lemma 3.6. *Let be $\overline{R} = \overline{R}(\overline{\imath}, \overline{\jmath})$. Then the number of squares in $SQ(\overline{R})$ containing at least one point for which Compute$_{\mathcal{N}}(n,k)$ or Compute$_{\mathcal{K}}(n,k)$ fail is $\sharp SQ_Z(\overline{R}) = O(\overline{\imath} + \overline{\jmath})$.*

Proof. We first note that if $(x,y) \in R$ with $R \in \mathrm{SQ}_Z(\overline{R})$ then $0 \le x \le \overline{\imath}$ and $0 \le y \le \overline{\jmath}$. Then, consider the set $Z_{\overline{\jmath}} = \{(x,y) \in Z \,|\, 0 \le y \le \overline{\jmath}\}$ and observe that $\sharp Z_{\overline{\jmath}} \le [\deg_n(p_0(n,k)) + \deg_n(p_r(n,k)) + (\deg_n(q_0(n,k)) + \deg_n(q_s(n,k))](\overline{\jmath}+1) = O(\overline{\jmath}) = O(\overline{\imath} + \overline{\jmath})$.
Since each square in $\mathrm{SQ}_Z(\overline{R})$ contains at least one point in $Z_{\overline{\jmath}}$, we have $\sharp\mathrm{SQ}_Z(\overline{R}) = O(\overline{\imath} + \overline{\jmath})$. □

In the sequel, we will denote by *Coeff$_\phi(R)$* the set of the coefficients of ϕ associated with R, that is,

$$\mathrm{Coeff}_\phi(R) = \{[x^a y^b]\phi(x,y) \,|\, (a,b) \in R\}.$$

The following lemmas tell us how to compute the coefficients in $\mathrm{Coeff}_\phi(R)$ from the knowledge of the coefficients in the neighborhood.

Lemma 3.7. *Let be $R \in SQ(\overline{R}) \setminus SQ_Z(\overline{R})$. If there exists $T \in \{N, W, S, E\}$ such that Coeff$_\phi(T(R))$ is known, then Coeff$_\phi(R)$ can be computed in time $O(1)$.*

Proof. Suppose that we know $\mathrm{Coeff}_\phi(E(R))$, that is, the set $\mathrm{Coeff}_\phi(R(l-e,m))$. Then it is immediate to obtain $\mathrm{Coeff}_\phi(R(l-e+1,m))$ by computing only e coefficients in $\mathrm{Coeff}_\phi(R(l-e+1,m)) \setminus \mathrm{Coeff}_\phi(R(l-e,m))$: this can be easily done with e^2 arithmetical operations (in order to get $c_{l-e+1\,m-i+1}$ we use the recurrence $\mathcal{N}$ and the values of the i-th row of $\mathrm{Coeff}_\phi(R(l-e,m))$). Then, for $i = 2 \ldots e$, we compute $\mathrm{Coeff}_\phi(R(l-e+i,m))$ from $\mathrm{Coeff}_\phi(R(l-e+i-1,m))$. Since there are e steps with cost $O(e^2)$, the overall computation requires time $O(e^3) = O(1)$. The other cases are similar. □

Lemma 3.8. *Let $R = R(l,m) \in SQ_Z(\overline{R})$. If Coeff$_\phi(W(R))$, Coeff$_\phi(SW(R))$ and Coeff$_\phi(S(R))$ are known, then Coeff$_\phi(R)$ can be computed in time $O(1)$.*

Proof. We define an ordering $\prec$ on the set $\mathrm{Coeff}_\phi(R)$ as follows: $c_{\alpha\beta} \prec c_{\gamma\delta}$ iff $\beta < \delta$ or $\beta = \delta$ and $\alpha \le \gamma$. Then, we can compute the coefficients according to the $\prec$ ordering, starting with $\min(\mathrm{Coeff}_\phi(R))$ and using equation $\mathcal{B}$ of the instance (see Equation (3.3) in the Coefficient Problem definition). At each step we compute one coefficient with e^2 arithmetical operations. Since we have e^2 coefficients, the total time is $O(e^4) = O(1)$. □

The transitive closure of $\diamond$ defines an equivalence relation $\diamond^\star \subseteq \mathrm{SQ}_Z(\overline{R})^2$, i.e. it defines a partition of $\mathrm{SQ}_Z(\overline{R})$ into equivalence classes that we call *clusters*. More precisely, let be

$$\diamond_{\overline{R}} \quad = \quad \diamond \,\cap\, \left(\mathrm{SQ}_Z(\overline{R}) \times \mathrm{SQ}_Z(\overline{R})\right)$$

and consider the following:

Definition 3.9 (Cluster)**.** The cluster generated by $R \in \mathrm{SQ}_Z(\overline{R}))$ is

$$\mathrm{Cl}_R^{\overline{R}} = [R]_{\diamond^\star_{\overline{R}}} = \left\{Q \in \mathrm{SQ}_Z(\overline{R}) \,|\, Q \diamond^\star_{\overline{R}} R\right\}.$$

It is then immediate to observe that it holds the following partition

$$\mathrm{SQ}_Z(\overline{R}) = \dot{\bigcup_{h=1}^{k}} \ \mathrm{Cl}_{R_h}^{\overline{R}}$$

with $R_h \in \mathrm{SQ}_Z(\overline{R})$ and $R_{h_1} \not\diamond^\star_{\overline{R}} R_{h_2}$.

Example 3.10. Let us consider the function $\phi(x,y) = \frac{1}{1-x^2y-xy^3}$ and the recurrences

$$\begin{aligned} N &= \left(4\,n^3 - 4\,n^2k - 30\,n^2 - 7\,nk^2 + 50\,n - 5\,nk + 25\,k + 5\,k^2 - 2\,k^3\right) {E_n}^5 + \\ &\quad (27\,n^3 - 135\,n^2 - 27\,n^2k + 90\,nk + 150\,n + 9\,nk^2 - k^3 - 50\,k - 15\,k^2)\,{E_n}^0 \\ K &= \left(10\,n + nk + 6\,n^2 - k^2 + 5\,k\right) {E_k}^5 - (4\,k^2 + 4\,nk - 5\,n - n^2 + 10\,k)\,{E_k}^0 \end{aligned}$$

associated with it. Let be $\overline{R} = R^{(5)}(59,59)$ and $R = R^{(5)}(4,4)$. The graphical representation of the cluster $\mathrm{Cl}_R^{\overline{R}}$ is given in Figure 2.

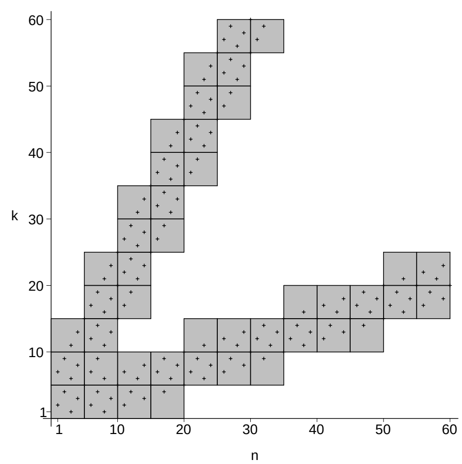


FIGURE 2. A cluster associated with $\phi(x,y) = \frac{1}{1-x^2y-xy^3}$. Dots are elements of $Z(\mathcal{N},\mathcal{K})$.

Given a cluster $\mathrm{Cl}_R^{\overline{R}}$ we define its *border* as the set

$$B(\mathrm{Cl}_R^{\overline{R}}) = \{R' \in \mathrm{SQ}(\overline{R}) \setminus \mathrm{SQ}_Z(\overline{R}) \,|\, \exists R'' \in \mathrm{Cl}_R^{\overline{R}} \text{ s.t. } R' \diamond_{\overline{R}} R''\}.$$

It is immediate to observe that $\sharp B(\mathrm{Cl}_R^{\overline{R}}) = O(\sharp \mathrm{Cl}_R^{\overline{R}}) = O(\overline{\imath} + \overline{\jmath})$.

3.2. The algorithm. As shown in the previous section, an instance $\langle \mathcal{N}, \mathcal{K}, \mathcal{B}, I, i, j\rangle$ univocally identifies a set Z, an integer e, a square R_0 at the origin and a square $\overline{R} = N^{\lfloor j/e \rfloor}(E^{\lfloor i/e \rfloor}(R_0))$ containing the point (i,j). We compute the coefficient $[x^iy^j]\phi(x,y)$ through a procedure that starts with $\mathrm{Coeff}_\phi(R_0) \subseteq I$ and halts having computed $\mathrm{Coeff}_\phi(\overline{R})$ after $O(i+j)$ steps.
Informally, the procedure works by computing a main sequence

$$\begin{aligned} &\mathrm{Coeff}_\phi(E^0(R_0)), \mathrm{Coeff}_\phi(E^1(R_0)), \ldots, \mathrm{Coeff}_\phi(E^{\lfloor i/e \rfloor}(R_0)), \\ &\mathrm{Coeff}_\phi(N^1(E^{\lfloor i/e \rfloor}(R_0))), \mathrm{Coeff}_\phi(N^2(E^{\lfloor i/e \rfloor}(R_0)))\ldots, \mathrm{Coeff}_\phi(\overline{R}). \end{aligned}$$

The first $\lfloor i/e \rfloor + 1$ sets of coefficients are easily computed in time $O(i)$. Then, we compute each set $\mathrm{Coeff}_\phi(N^k(E^{\lfloor i/e \rfloor}(R_0)))$ having as input $\mathrm{Coeff}_\phi(N^{k-1}(E^{\lfloor i/e \rfloor}(R_0)))$ according to the following rule: if $N^k(E^{\lfloor i/e \rfloor}(R_0)) \in \mathrm{SQ}(\overline{R}) \setminus \mathrm{SQ}_Z(\overline{R})$ then $\mathrm{Coeff}_\phi(N^k(E^{\lfloor i/e \rfloor}(R_0)))$ is computed as shown in Lemma 3.7, otherwise all the coefficients associated with the cluster $\mathrm{Cl}^{\overline{R}}_{N^k(E^{\lfloor i/e \rfloor}(R_0))}$ are computed in a suitable order.

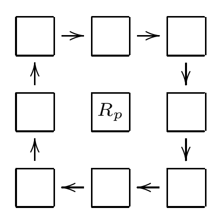
In Figure 3 we define a procedure COEFF(i,j) that has as input two positive integers i, j and returns the value $[x^iy^j]\phi(x,y)$. In the code a procedure COMPUTE$_\circlearrowleft(R_{out}, R_{in})$ is called. It takes as input two squares such that $R_{out} = T(R_{in})$, with $T \in \{N, E, S, W\}$, and computes the set $\mathrm{Coeff}_\phi(R_{out})$ under the assumption that $\mathrm{Coeff}_\phi(R_{in})$ has been previously computed.

Procedure COEFF(i, j)
Begin
  $R_0 := R(e-1, e-1)$; $c_1 := \lfloor i/e \rfloor$; $c_2 := \lfloor j/e \rfloor$;
  For k **from** 1 **to** c_1 **do**
    compute $\mathrm{Coeff}_\phi[E^k(R_0)]$ *from* $\mathrm{Coeff}_\phi[E^{k-1}(R_0)]$
    by using the equation $\mathcal{N}$ *and the initial conditions* I;
  For k **from** 1 **to** c_2 **do**
    if $N^k(E^{c_1}(R_0)) \notin \mathrm{SQ}_Z(\overline{R})$
      then *compute* $\mathrm{Coeff}_\phi[N^k(E^{c_1}(R_0))]$ *by using equation* $\mathcal{K}$ *and* $\mathrm{Coeff}_\phi[N^{k-1}(E^{c_1}(R_0))]$;
      else $\mathrm{COMPUTE}_{\circlearrowright}(N^k(E^{c_1}(R_0)), N^{k-1}(E^{c_1}(R_0)))$;
  return $[x^i y^j]\phi(x, y)$ *from* $\mathrm{Coeff}_\phi[N^{c_2}(E^{c_1}(R_0))]$;
End;

FIGURE 3. Procedure COEFF

Both procedures are supposed to use two global variables: a suitable data structure for the sets $\mathrm{Coeff}_\phi(R)$ and an integer variable e (the size of the edge of a square). As we note, the core of the algorithm consists of the procedure $\mathrm{COMPUTE}_{\circlearrowright}(R_{out}, R_{in})$. This procedure computes $\mathrm{Coeff}_\phi(R_{out})$ starting from $\mathrm{Coeff}_\phi(R_{in})$ and moving clockwise by using coefficients previously computed. In the code we find an indexed function $\mathrm{next}_{R_p}(R)$: this is used to identify the square R' that is neighbor to R_p and follows R (clockwise). More formally:

$$next_{R_p}(R) \begin{cases} NE(R_p) & if \quad R = N(R_p) \\ E(R_p) & if \quad R = NE(R_p) \\ SE(R_p) & if \quad R = E(R_p) \\ S(R_p) & if \quad R = SE(R_p) \\ SW(R_p) & if \quad R = S(R_p) \\ W(R_p) & if \quad R = SW(R_p) \\ NW(R_p) & if \quad R = W(R_p) \\ N(R_p) & if \quad R = NW(R_p) \end{cases}$$



Function *next*

As an example, suppose we call $\mathrm{COMPUTE}_{\circlearrowright}(R, S(R))$, that is, we want to compute $\mathrm{Coeff}_\phi(R)$ knowing $\mathrm{Coeff}_\phi(S(R))$. So, if $R \in \mathrm{SQ}_Z(\overline{R})$ then $\mathrm{Coeff}_\phi(SW(R))$ and $\mathrm{Coeff}_\phi(W(R))$ are needed, as shown in Lemma 3.8. Hence, the procedure advances clockwise around R, in order to get (recursively) $\mathrm{Coeff}_\phi(SW(R))$ from $\mathrm{Coeff}_\phi(S(R))$ and then $\mathrm{Coeff}_\phi(W(R))$ from $\mathrm{Coeff}_\phi(SW(R))$.
Figure 4 shows the procedure $\mathrm{COMPUTE}_{\circlearrowright}$; a simple example of computation is sketched in Figure 5, while in Figure 6 a real computation is shown.

4. Complexity

It is straightforward to see that $\mathrm{COEFF}(i, j)$ computes $[x^i y^j]\phi(x, y)$ if and only if every call $\mathrm{COMPUTE}_{\circlearrowright}(N^k(E^{c_1}(R_0)), N^{k-1}$ terminates and computes $\mathrm{Coeff}_\phi[N^k(E^{c_1}(R_0))]$.
Hence, the problem is to analyse which sets of coefficients are computed by the recursive procedure $\mathrm{COMPUTE}_{\circlearrowright}$.
Note that a call $\mathrm{COMPUTE}_{\circlearrowright}(R_{out}, R_{in})$ recursively calls itself if and only if $R_{out} \in \mathrm{SQ}_Z(\overline{R})$. So, let be $\mathrm{Out}_0 = N^k(E^{c_1}(R_0))$, $\mathrm{In}_0 = N^{k-1}(E^{c_1}(R_0))$ for a suitable integer $k \le c_2$ and consider the sequence of calls

$$\mathrm{COMPUTE}_{\circlearrowright}(\mathrm{Out}_0, \mathrm{In}_0), \ldots, \mathrm{COMPUTE}_{\circlearrowright}(\mathrm{Out}_l, \mathrm{In}_l)$$

Procedure COMPUTE$_\circlearrowright(R_{out}, R_{in})$
Begin
  While Undef(Coeff$_\phi[S(R_{out})]$) or Undef(Coeff$_\phi[SW(R_{out})]$) or Undef(Coeff$_\phi[W(R_{out})]$) **do**
    $R' := next_{R_{out}}(R_{in})$;
    if Undef(Coeff$_\phi[R']$) **then**
      if $R' \notin \text{SQ}_Z(\overline{R})$ **then** *compute* Coeff$_\phi[R']$ *by using* $\mathcal{N}$ *or* $\mathcal{K}$ *and* Coeff$_\phi[R_{in}]$;
        else COMPUTE$_\circlearrowright(R', R_{in})$;
    $R_{in} := R'$;
  EndWhile
  compute Coeff$_\phi[R_{out}]$ *by using* $\mathcal{B}$ *and* Coeff$_\phi[S(R_{out})]$, Coeff$_\phi[SW(R_{out})]$, Coeff$_\phi[W(R_{out})]$;
End;

FIGURE 4. Procedure COMPUTE$_\circlearrowright$.

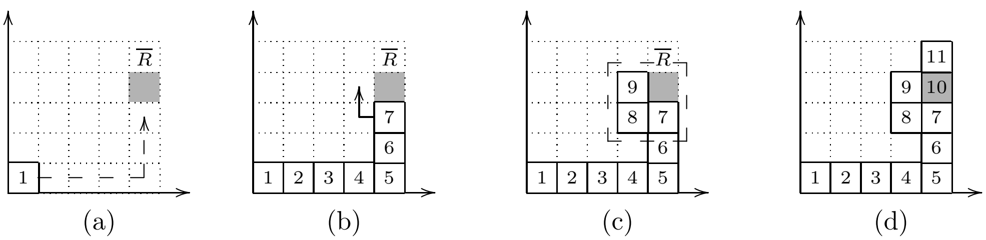


FIGURE 5. A run of COEFF. Squares are numbered with respect to the order of computation. The gray square is in $\text{SQ}_Z(\overline{R})$; in order to compute it, COMPUTE$_\circlearrowright$ moves clockwise until South, West and South-West neighbors have been computed.

contained in the stack associated with the call COMPUTE$_\circlearrowright(\text{Out}_0,\text{In}_0)$ (at the bottom).
For each $0 \le p < l$, let $\text{Step}_p = R_{p_1}, \ldots, R_{p_h}$ be the 4-connected sequence of squares adjacent to Out_p such that

$$R_{p_i} = \begin{cases} \text{In}_p & : \quad i = 1 \\ \text{next}_{\text{Out}_p}(R_{p_{i-1}}) & : \quad i > 1 \end{cases}$$

and $h = \min\{j \mid \text{next}_{\text{Out}_p}(R_{p_j}) = \text{In}_{p+1}\}$.
In the sequel, we consider three sequences Seq, $\widetilde{\text{Seq}}$ and $\widehat{\text{Seq}}$, associated with the stack and defined as follows:

- $\text{Seq} = \text{Out}_0, \ldots, \text{Out}_l$ is the 8-connected sequence of squares in $\text{SQ}_Z(\overline{R})$ such that $\text{Coeff}_\phi(\text{Out}_i)$ is not known $(0 \le i \le l)$.
- $\widetilde{\text{Seq}} = R_0, E(R_0), \ldots, E^{c_1}(R_0), N(E^{c_1}(R_0)), \ldots, N^{k-1}(E^{c_1}(R_0))$ is the 4-connected ascending sequence of $c_1 + k$ squares such that $\text{Coeff}_\phi(R)$ is known, $R \in \widetilde{\text{Seq}}$.
- $\widehat{\text{Seq}} = \text{Step}_0, \ldots, \text{Step}_{l-1}, \text{In}_l$ is the 4-connected sequence such that for all $R \in \widehat{\text{Seq}}$, $\text{Coeff}_\phi(R)$ has been computed by recursive calls to COMPUTE$_\circlearrowright$.

We analyse which sets of coefficients are computed by COMPUTE$_\circlearrowright$ by proving the following:
Lemma 4.1. *Let* COMPUTE$_\circlearrowright(N^k(E^{c_1}(R_0)), N^{k-1}(E^{c_1}(R_0)))$ *be a call occurring in* COEFF. *Then, for all the calls* COMPUTE$_\circlearrowright(R_{out}, R_{in})$ *that are pushed onto the stack we have*

$$R_{out} \in Cl^{N^k(E^{c_1}(R_0))}_{N^k(E^{c_1}(R_0))}.$$

PROOF. Let be $\text{Out}_0 = N^k(E^{c_1}(R_0))$ and let Seq, $\widetilde{\text{Seq}}$ and $\widehat{\text{Seq}}$ be the sequences associated with the stack having the call COMPUTE$_\circlearrowright(\text{Out}_0, S(\text{Out}_0)))$ at the bottom. We show that for all Out_i in Seq we

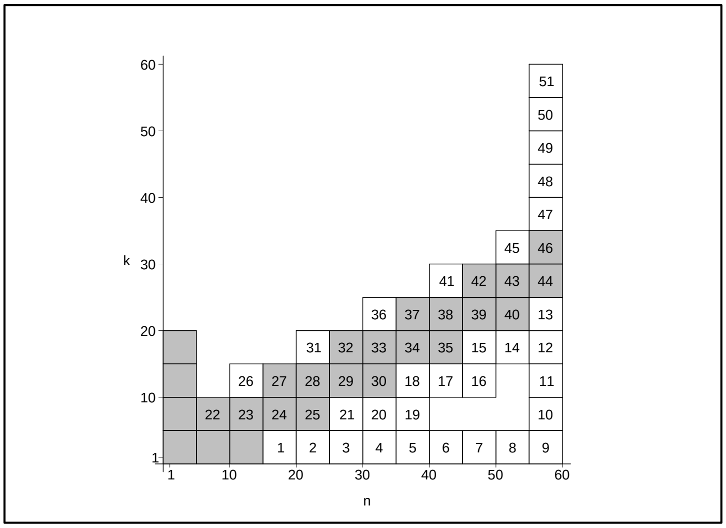


FIGURE 6. Running COEFF(59,59) for the function $\phi(x,y) = \frac{1}{1-x^2y-xy^3}$ (Example 3.10). Squares that are not numbered belong to the set of initial conditions.

have $\mathrm{Out}_i \in \mathrm{Cl}^{\mathrm{Out}_0}_{\mathrm{Out}_0}$, that is, $\mathrm{Out}_i \diamond^{\star}_{\mathrm{Out}_0} \mathrm{Out}_0$. Hence, since for $0 \le i < l$ it holds $\mathrm{Out}_i \diamond \mathrm{Out}_{i+1}$, it is sufficient to prove that

$$\mathrm{Out}_i \in \mathrm{SQ}_Z(\mathrm{Out}_0) \tag{4.1}$$

Observe that $\{\mathrm{Seq}\} \cap \{\widehat{\mathrm{Seq}}\} = \{\mathrm{Seq}\} \cap \{\widetilde{\mathrm{Seq}}\} = \emptyset$ and note that we can univocally identify g sequences Seq_i $(1 \le i \le g)$ such that $\mathrm{Seq} = \mathrm{Seq}_1, \mathrm{Seq}_2, \dots, \mathrm{Seq}_g$ with

- Seq_1 is the longest descending sequence that appears at the beginning of Seq
- Seq_{2i} is the longest ascending sequence after $\mathrm{Seq}_1, \mathrm{Seq}_2, \dots, \mathrm{Seq}_{2i-1}$, $1 < 2i \le g$
- Seq_{2i+1} is the longest descending sequence after $\mathrm{Seq}_1, \mathrm{Seq}_2, \dots, \mathrm{Seq}_{2i}$, $1 < 2i+1 \le g$

We prove Property (4.1) by induction on the number g of ascending or descending sequences in the decomposition of Seq shown above.

BASIS: Seq consists of one descending sequence $\mathrm{Out}_0, \dots, \mathrm{Out}_l$, where $\mathrm{Out}_i = E^{w_i}(W^{v_i}(S^{u_i}(\mathrm{Out}_0)))$ and either $\mathrm{Out}_1 = W(\mathrm{Out}_0)$ or $\mathrm{Out}_1 = SW(\mathrm{Out}_0)$. So, it trivially holds that $\mathrm{Out}_i \in \mathrm{SQ}_Z(\mathrm{Out}_0)$ if $v_i > w_i$, $1 \le i \le l$. By absurd, let be $\bar{\imath} = \min\{i \,|\, v_i < w_i\}$ (note that it must be $v_{\bar{\imath}} \ne w_{\bar{\imath}}$ since $\{\mathrm{Seq}\} \cap \{\widetilde{\mathrm{Seq}}\} = \emptyset$). So, we would have $v_{\bar{\imath}-1} > w_{\bar{\imath}-1}$ and $v_{\bar{\imath}} < w_{\bar{\imath}}$: this would imply that $\mathrm{Out}_{\bar{\imath}-1} \not\diamond \mathrm{Out}_{\bar{\imath}}$.

INDUCTION: $\mathrm{Seq} = \mathrm{Seq}_1, \mathrm{Seq}_2, \cdots, \mathrm{Seq}_{n-1}, \mathrm{Seq}_n$. By induction we know that all the squares in $\mathrm{Seq}_1, \mathrm{Seq}_2, \cdots, \mathrm{Seq}_{n-1}$ satisfy Property (4.1). Let be $\mathrm{Seq}_n = \mathrm{Out}_s, \dots, \mathrm{Out}_l$ and let Out_{s-1} be the last square of Seq_{n-1}. We distinguish two cases.

n IS ODD: Seq_n is a descending sequence. By induction we know that $\mathrm{Out}_{s-1} \in \mathrm{SQ}_Z(\mathrm{Out}_0)$, that is, $\mathrm{Out}_{s-1} = W^{\alpha_{s-1}}(S^{u_{s-1}}(\mathrm{Out}_0))$ with $\alpha_{s-1}, u_{s-1} \in \mathbb{N}$, $\alpha_{s-1} > 0$. Since $\mathrm{Out}_s = T(\mathrm{Out}_{s-1})$ with $T \in \{SW, S, SE\}$ it holds $\mathrm{Out}_s = W^{\alpha_s}(S^{u_{s-1}+1}(\mathrm{Out}_0))$ with $\alpha_s \ge 0$. Again, $\alpha_s \ne 0$ since $\{\mathrm{Seq}\} \cap \{\widehat{\mathrm{Seq}}\} = \emptyset$. Now, the same analysis done for the basis shows that Property 4.1 holds for the squares of Seq_n.

n IS EVEN: Seq_n is an ascending sequence, that is, $\mathrm{Out}_s = T(\mathrm{Out}_{s-1})$, $T \in \{NW, N, NE\}$. We claim that $\mathrm{Out}_s = NE(\mathrm{Out}_{s-1})$. In fact, recall that each sequence Step_i of squares examined by $\mathrm{COMPUTE}_{\circlearrowleft}(\mathrm{Out}_i, \mathrm{In}_i)$

before calling $\text{COMPUTE}_{\circlearrowleft}(\text{Out}_{i+1}, \text{In}_{i+1})$ is 4-connected. This means that the sequence $\text{Out}_{s-1}, \text{In}_{s-1}$ is 4-connected, that is, $\text{In}_{s-1} = T(\text{Out}_{s-1})$ with $T = \{N, E, S, W\}$. In particular, note that $\text{In}_{s-1} = N(\text{Out}_{s-1})$ since in the other three cases the call $\text{COMPUTE}_{\circlearrowleft}(\text{Out}_{s-1}, \text{In}_{s-1})$ would compute $\text{Coeff}_{\phi}(\text{Out}_{s-1})$ without any recursion.

Therefore, $\text{COMPUTE}_{\circlearrowleft}(\text{Out}_{s-1}, \text{In}_{s-1})$ recursively calls $\text{COMPUTE}_{\circlearrowleft}(NE(\text{Out}_{s-1}), \text{In}_s)$ with $\text{In}_s = N(\text{Out}_{s-1}) \in \widehat{\text{Seq}}$.

Now, consider the 4-connected sequence $\widehat{\widehat{\text{Seq}}}$ obtained by joining $\widetilde{Seq}$ to $\widehat{Seq}$,

$$\widehat{\widehat{Seq}} = R_0, E(R_0), \ldots, E^{c_1}(R_0), N(E^{c_1}(R_0)), \ldots, N^{k-1}(E^{c_1}(R_0)), \text{Step}_0, \ldots, \text{Step}_{s-1}, \text{In}_s.$$

We trivially have $\{\widehat{\widehat{Seq}}\} \cap \{\text{Seq}_n\} = \emptyset$. In fact, the value $\text{Coeff}_{\phi}[R]$ is defined if $R \in \widehat{\widehat{\text{Seq}}}$ and undefined if $R \in \text{Seq}_n$. Informally, this means that the squares of the ascending sequence Seq_n are restricted to lie in a closed area (delimited by $\widehat{\widehat{Seq}}$) consisting of squares that satisfy Property (4.1).

□

An immediate consequence of the previous lemma is:

Corollary 4.2. *Let be $R_k = N^k(E^{c_1}(R_0)) \in SQ_Z(\overline{R})$. If $Coeff_{\phi}(R)$ is computed by a call* $\text{COMPUTE}_{\circlearrowleft}(R_k, S(R_k))$ *occurring in COEFF then*

$$R \in B(Cl^{R_k}_{R_k}) \cup Cl^{R_k}_{R_k}.$$

Proof. By inspecting the code of $\text{COMPUTE}_{\circlearrowleft}$ we note that for each computed set $\text{Coeff}_{\phi}(R)$, either $R \in \text{SQ}_Z(\overline{R})$ (and $\text{COMPUTE}_{\circlearrowleft}(R, R_{in})$ is a call generated by $\text{COMPUTE}_{\circlearrowleft}(R_k, S(R_k))$) or $R \notin \text{SQ}_Z(\overline{R})$ (and $\text{Coeff}_{\phi}(R)$ is computed by a recursive call $\text{COMPUTE}_{\circlearrowleft}(R_{out}, Q)$ generated by $\text{COMPUTE}_{\circlearrowleft}(R_k, S(R_k))$ such that $R \diamond R_{out}$).

In the first case Lemma 4.1 states that $R \in \text{Cl}^{R_k}_{R_k}$ while in the second we have $R \in B(\text{Cl}^{R_k}_{R_k})$. □

Lemma 4.3. *Let St be the stack associated with a call* $\text{COMPUTE}_{\circlearrowleft}(R, S(R))$ *occurring in* COEFF. *Then, St does not contain two identical calls.*

Proof. (By contradiction) Let $\text{COMPUTE}_{\circlearrowleft}(\text{Out}_h, \text{In}_h)$ be the first repeated occurrence of a call, that is, $h = \min\{0 \le i \le l \,|\, \exists\, \delta > 0,\, \text{Out}_i = \text{Out}_{i-\delta} \wedge \text{In}_i = \text{In}_{i-\delta}\}$. Without loss of generality, we suppose that $\text{In}_h = W(\text{Out}_h)$. Consider the 8-connected sequence S that is a subsequence of Seq,

$$\text{S} = \text{Out}_{h-\delta}, \text{Out}_{h-\delta+1}, \ldots, \text{Out}_h,$$

together with the 4-connected subsequence of $\widehat{\text{Seq}}$,

$$\widehat{\text{S}} = \text{Step}_{h-\delta}, \text{Step}_{h-\delta+1}, \ldots, \text{Step}_{h-1}, \text{In}_h.$$

We recall that for $R \in \widehat{\text{S}}$ the set $\text{Coeff}_{\phi}(R)$ is known and that for each $R \in \widehat{\text{S}}$ ($R \in \text{S}$) there exists $Q \in \text{S}$ ($Q \in \widehat{\text{S}}$) such that $R \diamond Q$. Note that both sequences are "closed", that is, their first and last squares coincide. Moreover, we have $\{\text{S}\} \cap \{\widehat{\text{S}}\} = \emptyset$.

Let be $R_0 = R(e-1, e-1)$ and for each closed sequence S denote by $\text{Inside}(S)$ the set of all the squares in $\text{SQ}(\overline{R})$ that lie in the area surrounded by S. Then, it is immediate to observe that we have only two cases:

$\widehat{\text{S}} \subseteq \text{Inside(S)}$: This means that if $R \in \widehat{\text{S}}$ it is impossible to find a 4-connected sequence $T_R = R_0, \ldots, R$ such that $\{T_R\} \cap \{\text{S}\} = \emptyset$. On the other hand, we know that for every $R \in \widehat{\text{S}}$ there exists a 4-connected sequence T_R from R_0 to R, consisting of squares in $\text{SQ}(\overline{R})$, such that for Q in T_R the set $\text{Coeff}_{\phi}(Q)$ has been computed (see the sequence $\widehat{\widehat{\text{Seq}}}$ in the proof of Lemma 4.1). Therefore, we have $\widehat{\text{S}} \not\subseteq \text{Inside(S)}$.

$\mathrm{S} \subseteq$ Inside$(\widehat{\mathrm{S}})$: Let be $k_1, k_2 \in \mathbb{N}$ such that

$$\begin{cases} N^{k_1}(E^{k_2}(R_0)) \in \mathrm{S} \\ N^{h_1}(E^{h_2}(R_0)) \in \mathrm{S} \Rightarrow k_1 + k_2 \leq h_1 + h_2 \end{cases}$$

Let be $\mathrm{Out}_{\overline{h}} = N^{k_1}(E^{k_2}(R_0))$. Since $\mathrm{S} \subseteq \mathrm{Inside}(\widehat{\mathrm{S}})$, it is immediate to prove that $S(\mathrm{Out}_{\overline{h}})$ and $W(\mathrm{Out}_{\overline{h}})$ belong to $\widehat{\mathrm{S}}$. More precisely, because $\widehat{\mathrm{S}}$ is 4-connected, it follows that

$$\widehat{\mathrm{S}} = \mathrm{In}_h, \ldots, S(\mathrm{Out}_{\overline{h}}), SW(\mathrm{Out}_{\overline{h}}), W(\mathrm{Out}_{\overline{h}}), \ldots, \mathrm{In}_h.$$

Then, the call that computes $\mathrm{Coeff}_\phi(SW(\mathrm{Out}_{\overline{h}}))$ must be $\mathrm{COMPUTE}_\circlearrowleft(\mathrm{Out}_{\overline{h}}, \mathrm{In}_{\overline{h}})$. By observing the code of $\mathrm{COMPUTE}_\circlearrowleft$, we note that if $\mathrm{COMPUTE}_\circlearrowleft(\mathrm{Out}_{\overline{h}}, \mathrm{In}_{\overline{h}})$ computes $\mathrm{Coeff}_\phi(SW(\mathrm{Out}_{\overline{h}}))$ then it has previously computed $\mathrm{Coeff}_\phi(S(\mathrm{Out}_{\overline{h}}))$ and it necessarely computes also $\mathrm{Coeff}_\phi(W(\mathrm{Out}_{\overline{h}}))$. So, this call would terminate without any recursion.

□

The following lemma states that we can develop a suitable data structure for storing all the coefficients that are computed by the algorithm. More precisely, we have:

Lemma 4.4. *The data structure $\mathit{Coeff}_\phi[]$ can be implemented in space $O(i+j)$ and accessed in time $O(1)$.*

Proof. $\mathrm{Coeff}_\phi[]$ can be easily implemented as a dynamic data structure representing a graph. We give here an outline for such implementation.

Let be

$$d = \max\{\deg_n(p_r(n,k)), \deg_n(p_0(n,k)), \deg_n(q_s(n,k)), \deg_n(q_0(n,k))\}$$

the integer univocally associated with an instance $\langle \mathcal{N}, \mathcal{K}, \mathcal{B}, I, i, j \rangle$ and let be $\zeta = 4d(e+1)$. Note that for every integer k we have $\sharp\{N^k(E^h(R(e-1,e-1))) \in \mathrm{SQ}_Z(\overline{R})\} \leq \zeta$. Since we have to consider also squares that belong to the border of a cluster, for each k we have at most 9ζ squares $R_{kh} = N^k(E^h(R(e-1,e-1)))$ such that $\mathrm{Coeff}_\phi(R_{kh})$ is computed. So, an immediate implementation for the sets $\mathrm{Coeff}_\phi()$ is based on a list of lists. More precisely, we have a primary double linked list whose length is $\lfloor j/e \rfloor + 1$. The k-th node of this list contains a link to the list for the sets $\mathrm{Coeff}_\phi(R_{kh})$: this is a list whose length is less or equal to 9ζ. Then, it is immediate to note that we access to $\mathrm{Coeff}_\phi[R_{kh}]$ in constant time if the procedure $\mathrm{COMPUTE}_\circlearrowleft(R_{kh}, In)$ is equipped with a suitable link to the k-th node of the main list. □

Theorem 4.5. *The total number of calls to* $\mathrm{COMPUTE}_\circlearrowleft$ *during the execution of* $\mathrm{COEFF}(i,j)$ *is* $O(i+j)$.

Proof. Recall that $\overline{R} = R(\overline{\imath}, \overline{\jmath}) = N^{\lfloor j/e \rfloor}(E^{\lfloor i/e \rfloor}(R(e-1,e-1)))$ and let

$$\mathrm{COMPUTE}_\circlearrowleft(R_1, S(R_1)), \ldots, \mathrm{COMPUTE}_\circlearrowleft(R_t, S(R_t))$$

be the sequence of calls observed in $\mathrm{COEFF}(i,j)$ $(t = O(j))$. Moreover, let be

$$\mathrm{TOT} = \left\{ \mathrm{COMPUTE}_\circlearrowleft(R, T(R)) \,|\, R \in \mathrm{SQ}_Z(\overline{R}),\ T \in \{N, E, S, W\} \right\},$$

and, for $1 \leq k \leq t$,

$$\mathrm{TOT}_k = \{C \in \mathrm{TOT} \,|\, C \text{ is a recursive call originated by } \mathrm{COMPUTE}_\circlearrowleft(R_k, S(R_k))\}.$$

Note that $\mathrm{COMPUTE}_\circlearrowleft(R_k, S(R_k))$ recursively generates calls of type $\mathrm{COMPUTE}_\circlearrowleft(R, Q)$ with $R \in \mathrm{Cl}^{R_k}_{R_k}$, such that $\mathrm{Coeff}_\phi(R)$ has not been previously computed by $\mathrm{COMPUTE}_\circlearrowleft(R_l, S(R_l))$ with $1 \leq l < k$. In other words, $\mathrm{TOT}_l \cap \mathrm{TOT}_m = \emptyset$, for $l \neq m$.

Lemma 4.3 guarantees that the number of recursive calls generated by $\mathrm{COMPUTE}_\circlearrowleft(R_k, S(R_k))$ is exactly $\sharp\mathrm{TOT}_k$. Hence, recalling Lemma 3.6, the total number of calls is

$$\sum_{k=1}^{t} \sharp\mathrm{TOT}_k \;=\; \sharp \bigcup_{k=1}^{t} \mathrm{TOT}_k \;\leq\; \sharp\mathrm{TOT} \;=\; 4 \cdot \sharp\mathrm{SQ}_Z(\overline{R}) \;=\; O(\overline{\imath} + \overline{\jmath}) \;=\; O(i+j)$$

□

At last, we have:
Theorem 4.6. COEFF(i, j) *runs in time* $O(i + j)$ *and in space* $O(i + j)$.

PROOF. By Th. 5.4 we know that procedure COMPUTE$_{\circlearrowleft}$ is called $O(i + j)$ times. By inspecting the code we note that each call consists of a constant number of operations because the cost of accessing Coeff$_\phi$[] is $O(1)$ (see Lemma 5.3). Moreover, the space requirement is bounded by the sum of the maximum stack size and the size of the data structure for Coeff$_\phi$[]. So, we conclude that COEFF(i, j) runs in time $O(i + j)$ using $O(i + j)$ space. □

5. Conclusions

In this paper we have presented an algorithm that computes the coefficient $[x^i y^j]\phi(x, y)$ of a rational formal series $\phi(x, y)$ working in time and space $O(i + j)$ under the uniform cost criterion. If we adopt the logarithmic cost criterion, we expect that the complexity of the algorithm becomes $O((i + j)^2)$, since the growing of the coefficients $[x^n y^k]\phi(x, y)$ is at most exponential (i.e. the cost of a single arithmetical operation is at most linear).

Two remarks are worthwhile with respect to such algorithm: the first is related to the computation of the recurrences, the second deals with the size e of the squares. We pointed out that the recurrences can be obtained through an elimination process in a noncommutative algebra. Actually, in order to compute a Gröebner Basis in the shift algebra $\mathbb{Q}\langle n, k, E_n, E_k\rangle$, we took advantage of the package 'Mgfun', running under Maple and implemented by Chyzak ([**5**]). This step can be quite expensive and so it would be interesting to look for a method that directly computes the recurrences from the linear representation of a rational series.

With respect to the size e, let us show an upper bound for its value. Consider a rational function $\phi = \frac{P(x,y)}{Q(x,y)}$ and let d_x (d_y) be the degree of $P \cdot Q$ in x (y). A system of independent recurrence equations is obtained by converting the holonomic system

$$\left\{ \begin{array}{rcl} (\partial_x P\ Q - P\ \partial_x Q)\ \partial_x \phi & = & 0 \\ (\partial_y P\ Q - P\ \partial_y Q)\ \partial_x \phi \ + \ (\partial_x P\ Q - P\ \partial_x Q)\ \partial_y \phi & = & 0 \end{array} \right.$$

into operators of the shift algebra. The degree in E_n and E_k of such operators is respectively d_x and d_y. By applying the Zeilberger's elimination algorithm ([**14**]), we obtain operators depending either on E_n or on E_k. Their degrees in E_n and E_k are at most quadratic with respect to d_x and d_y (see Section 5.3 in [**14**]). Then, we have $e = O((\max\{d_x, d_y\})^2)$.

Last but not least, it might be interesting to study whether the technique we have presented can be modified in order to deal with a number of variables greater that two. We point out that the straightforward extension of this method to the 3-D case does not work (in the 3-D case the set of the values that the algorithm considers as initial conditions has not size $O(1)$). Moreover, as our method is related to the theory of the holonomic series, it would be useful to generalize it in order to get an efficient algorithm for the Coefficient Problem for holonomic series.

References

[1] I. N. Bernstein, *Modules over a ring of differential operators, study of the fundamental solutions of equations with constant coefficients.* Functional Anal. Appl., vol. 5 (1971), pp. 1–16 (russian); pp. 89–101 (english).

[2] A. Bertoni, P. Massazza, R. Radicioni, *Random generation of words in regular languages with fixed occurrences of symbols.* Proceedings of WORDS'03, 4^{th} International Conference on Combinatorics of Words, Turku, Finland, 2003, pp. 332–343.

[3] J. Berstel, C. Reutenauer, *Rational Series and Their Languages.* Sprinter-Verlag, EATCS Monographs on Theoretical Computer Science, vol. 12 (1984).

[4] M. Bousquet-Mélou, M. Petkovšek,*Linear recurrences with constant coefficients: the multivariate case.* Discrete Mathematics **225** (2000), pp. 51–75.

[5] F. Chyzak, *An Extension of Zeilberger's Fast Algorithm to General Holonomic Functions.* Proceedings of FPSAC'97, Universität Wien, 1997, pp. 172–183.
[6] F. Chyzak, B. Salvy, *Non-commutative Elimination in Ore Algebras Proves Multivariate Identities.* Journal of Symbolic Computation **2**, vol.26 (1998), pp. 187–227.
[7] A. Denise,*Génération aléatoire uniforme de mots de langages rationnels.* Theoretical Computer Science **159** (1996), pp. 43–63.
[8] A. Denise, O. Roques, M. Termier, *Random generation of words of context-free languages according to the frequencies of letters.* Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities, Trends in Mathematics, Birkhaüser (2000), pp. 113–125. (Proceedings of Colloquium on Mathematics and Computer Science, Versailles, September 2000).
[9] I. Dutour, J. M. Fédou, *Object grammars and random generation.* Discrete Mathematics and Theoretical Computer Science **2** (1998), pp. 47–61.
[10] P. Flajolet, P. Zimmermann and B. Van Cutsem, *A calculus for the random generation of labelled combinatorial structures.* Theoretical Computer Science **132** (1994), pp. 1–35.
[11] L. Lipshitz, *D−Finite Power Series.* Journal of Algebra **122** (1989), pp. 353–373.
[12] E. Mansfield, *Differential Gröbner Bases.* Ph.D Thesis, University of Sidney (1991).
[13] R.P. Stanley, *Differentiably Finite Power Series.* European Journal of Combinatorics **1** (1980), pp. 175–188.
[14] D. Zeilberger, *A holonomic systems approach to special functions identities.* Journal of Computational and Applied Mathematics **32** (1990), pp. 321–368.

Dipartimento di Informatica e Comunicazione, Università degli Studi dell'Insubria, Via Mazzini 5, 21100 Varese, Italy
E-mail address: paolo.massazza@uninsubria.it

Dipartimento di Scienze dell'Informazione, Università degli Studi di Milano, via Comelico 39, 20135 Milano, Italy
E-mail address: radicioni@dsi.unimi.it

An immediate consequence of the previous lemma is:

Corollary 5.1. *Let be* $R(\overline{\imath}, ke) \in SQ_Z(\overline{R})$ *and* $x, y \in \mathbb{N}$ *such that the set* $Coeff_\phi(x, y)$ *is computed by a call* $\mathrm{COMPUTE}_{\circlearrowleft}(\overline{\imath}, ke, \overline{\imath}, (k-1)e)$ *occurring in COEFF. Then*

$$R(x, y) \in B(Cl^{(\overline{\imath},ke)}_{R(\overline{\imath},ke)}) \cup Cl^{(\overline{\imath},ke)}_{R(\overline{\imath},ke)}.$$

Proof. By inspecting the code of $\mathrm{COMPUTE}_{\circlearrowleft}$ we note that each recursive call occurs on input (x, y, z, t) such that $R(x, y)$ is either in the equivalence class of $R(\overline{\imath}, ke)$ (with respect to relation $\diamond^\star$) or in its border. Lemma 4.1 bounds x and y properly. □

Lemma 5.2. *Let* $\mathrm{COMPUTE}_{\circlearrowleft}(l, m, l, m-e)$ *be a call in* COEFF*. Then, the stack that has* $\mathrm{COMPUTE}_{\circlearrowleft}(l, m, l, m-e)$ *at the bottom does not contain two identical calls.*

Proof. (By contradiction) Let us denote by C_h the call at level h in the stack. Without loss of generality, let $C_{\hat{h}} = COMPUTE_{\circlearrowleft}(\hat{x}, \hat{y}, \hat{x} - e, \hat{y})$ be the first occurrence of a call that appears twice in the stack. Then, there exists $\delta > 0$ such that $C_{\hat{h}-\delta} = COMPUTE_{\circlearrowleft}(\hat{x}, \hat{y}, \hat{x} - e, \hat{y})$. The portion of stack $C_{\hat{h}-\delta}, C_{\hat{h}-\delta+1}, \ldots, C_{\hat{h}}$ identifies an 8-connected sequence

$$\mathrm{Seq} = R(x_0, y_0), \ldots, R(x_\delta, y_\delta) \qquad (\text{with } R(x_0, y_0) = R(x_\delta, y_\delta) = R(\hat{x}, \hat{y}))$$

that has an associated 4-connected sequence

$$\widetilde{\mathrm{Seq}} = R(\tilde{x}_0, \tilde{y}_0), \ldots, R(\tilde{x}_\gamma, \tilde{y}_\gamma) \qquad (\text{with } R(\tilde{x}_0, \tilde{y}_0) = R(\tilde{x}_\gamma, \tilde{y}_\gamma) = R(\hat{x} - e, \hat{y})))$$

such that $\mathrm{Coeff}(\tilde{x}_j, \tilde{y}_j)$ is known, $0 \le j \le \gamma$. Note that Seq and $\widetilde{\mathrm{Seq}}$ are closed path and that $\{\mathrm{Seq}\} \cap \{\widetilde{\mathrm{Seq}}\} = \emptyset$. Let Inside($P$) denote the set of all the squares of SQ that lie in the area surrounded by a closed path P. Then, it is immediate to observe that we have only two cases:

$\widetilde{\text{Seq}} \subseteq$ Inside(Seq): This means that for each $R(x,y) \in \widetilde{\text{Seq}}$ it is impossible to find a 4-connected sequence $T_{(x,y)} = R(e,e), \ldots, R(x,y)$ such that $\{T_{(x,y)}\} \cap \{\text{Seq}\} = \emptyset$. On the other hand, we know that for every $R(x,y) \in \widetilde{\text{Seq}}$ there is a 4-connected sequence of squares $R(e,e), \ldots, R(x,y)$ whose associated coefficients have been computed (see the definition of the sequence $\widetilde{\widetilde{\text{Seq}}}$ in the proof of Lemma 4.1). We have then the contradiction $\widetilde{\text{Seq}} \not\subseteq \text{Inside(Seq)}$.

Seq $\subseteq$ Inside($\widetilde{\text{Seq}}$): Since $\widetilde{\text{Seq}}$ is a closed 4-connected sequence, there must be $R(a,b) \in \widetilde{\text{Seq}}$ such that $\widetilde{\text{Seq}} = R(\tilde{x}_0, \tilde{y}_0), \ldots, R(a,b), R(a-e,b), R(a-e,b+e), \ldots R(\tilde{x}_\gamma, \tilde{y}_\gamma)$ and $R(a,b+e) \in \text{Seq}$, $R(a,b+e) \in \text{SQ}_S$. This is impossible, because there would be in the stack a call $COMPUTE_{\circlearrowleft}(a, b+e, \alpha, \beta)$ that would halt the recursion.

□

The following lemma states that we can develop a suitable data structure for storing all the coefficients that are computed by the algorithm. More precisely, we have:

Lemma 5.3. *Let $\langle \mathcal{N}, \mathcal{K}, \mathcal{B}, I, i, j\rangle$ be an instance of the Coefficient Problem associated with a rational series $\phi(x,y)$. Then, there exists a data structure G for the sets $Coeff_\phi(x,y)$ that are computed by $COMPUTE(i,j)$ such that G can be implemented in space $O(i+j)$ and access time $O(1)$.*

Proof. G can be easily implemented as a dynamic data structure representing a graph. We give here an outline for such implementation.
Let us consider the system of recurrences $\{N, K\}$ and let be

$$d = \max\{\deg_n(p_r(n,k)), \deg_n(p_0(n,k)), \deg_n(q_s(n,k)), \deg_n(q_0(n,k))\}.$$

Note that for every integer $k \geq e, k \equiv_e 0$, there exists at most $\zeta = 4d(e+1)$ values x_{k_l} such that $R(x_{k_l}, k) \in \text{SQ}_S(i,j)$, and $\text{Coeff}_\phi(x_{k_l}, k)$ is computed by the algorithm. Since we have to consider also squares that belong to borders, we have at most $\xi = \zeta + 2\zeta + 3 \cdot 2\zeta$ squares $R(x,k)$ such that $\text{Coeff}_\phi(x,k)$ is computed. So, an immediate implementation for the sets $\text{Coeff}_\phi(x,y)$ is based on a list of lists. More precisely, we have a primary double linked list whose length is $\lceil j/e \rceil$. The h-th node of this list contain a link to the list for the sets $\text{Coeff}_\phi(x, he)$: this is a list whose length is bounded by the constant ξ. Then, it is immediate to note that the access to the sets $\text{Coeff}_\phi(x, be)$ requires constant time if the procedure $\text{COMPUTE}_{\circlearrowleft}(a,b,c,d)$ is equipped with a suitable link to the b-th node of the main list. □

Theorem 5.4. *The total number of calls to $COMPUTE_{\circlearrowleft}$ during the execution of $COEFF(i,j)$ is $O(i+j)$.*

Proof. Let $\text{COMP}(x,y,z,t)$ be the number of recursive calls generated by $\text{COMPUTE}_{\circlearrowleft}(x,y,z,t)$. Define $\text{CALL} = \{(x,y,z,t) \in \mathbb{N}^4 \mid \text{COEFF}(i,j) \text{ calls } \text{COMPUTE}_{\circlearrowleft}(x,y,z,t)\}$. We obviously have $\sharp\text{CALL} = O(i+j)$. Moreover, let be $\text{CALL}_1 = \{(x,y,z,t) \in \text{CALL} \mid R(x,y) \in \text{SQ}_S(i,j)\}$ and $\text{CALL}_2 = \{(x,y,z,t) \in \text{CALL} \mid R(x,y) \notin \text{SQ}_S(i,j)\}$.
We need to estimate

$$\sum_{\alpha \in \text{CALL}} \text{COMP}(\alpha) = \sum_{\alpha \in \text{CALL}_1} \text{COMP}(\alpha) \quad + \sum_{\alpha \in \text{CALL}_2} \text{COMP}(\alpha).$$

We first note that $\sum_{\alpha \in \text{CALL}_2} \text{COMP}(\alpha) = O(i+j)$, then we partition CALL_1 according to the partition of $\text{SQ}_S(i,j)$ in clusters, obtaining

$$\text{CALL}_1 = \dot{\bigcup_{h=1}^{k}} \text{CALL}(x_h, y_h, z_h, t_h)$$

where $\text{CALL}(x_h, y_h, z_h, t_h) = \{(x,y,z,t) \in \text{CALL}_1 \mid R(x,y) \diamond^\star_{(i,j)} R(x_h, y_h)\}$ and for each $1 \leq r < s \leq h$, $R(x_r, y_r) \not\diamond^\star_{(i,j)} R(x_s, y_s)$.

Since the sets $\mathrm{Coeff}_\phi(a,b)$ are computed once, recalling Lemma 4.3 and Corollary 4.2 we have

$$\sum_{\alpha\in \mathrm{CALL}_1} \mathrm{COMP}(\alpha) = \sum_{h=1}^{k} \sum_{\alpha\in \mathrm{CALL}(x_h,y_h,z_h,t_h)} \mathrm{COMP}(\alpha) \quad = \quad O\left(\sum_{h=1}^{k} \sharp \mathrm{Cl}_{R(x_h,y_h)}^{(i,j)}\right) \quad = \quad O(i+j).$$

Hence, we conclude that

$$\sum_{\alpha\in \mathrm{CALL}} \mathrm{COMP}(\alpha) = O(i+j).$$

□

At last, we have:
Theorem 5.5. *COEFF(i,j) runs in time $O(i+j)$ and in space $O(i+j)$.*

PROOF. By Th. 5.4 we know that there are $O(i+j)$ calls to the procedure $\mathrm{COMPUTE}_{\circlearrowleft}$. By inspecting the code we note that each instance consists of a constant number of operations if accessing $\mathrm{Coeff}_\phi(x,y)$ costs $O(1)$. Moreover, the space requirement is bounded by the sum of the maximum stack size and the size of the data structure for $\mathrm{Coeff}_\phi(x,y)$. Recalling Theoren **??** and Lemma 5.3 we conclude that COEFF(i,j) runs in time $O(i+j)$ using $O(i+j)$ space. □

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Ribbon tilings of Ferrers diagrams, flips and the 0-Hecke algebra

Gilles Radenne

Abstract.

In this article we study how the 0-Hecke algebra $H_m(0)$ can be used to give an algebraic structure to ribbon tilings of Ferrers diagrams. The key to this structure is the Stanton-White bijection, which gives a bijection between n-ribbon tableaux and n-upplets of Young tableaux. Restricting to standard ribbon tableaux, we can define a natural action of $H_m(0)$. Thus we can define local actions on ribbon tableaux, which we call flips or pseudo-flips, and which are generalisations of domino flips. Then with some help from the Yang-Baxter relations we prove some properties about minimal flip chains, properties which remain true for ribbon tilings.

Résumé. *Le but de cet article est de montrer comment on peut utiliser la 0-algèbre de Hecke $H_m(0)$ afin de donner une structure algébrique aux pavages par rubans d'un diagramme de Ferrers donné. Cette structure découle de la bijection de Stanton-White entre les tableaux de n-rubans et les n-uplets de tableaux de Young. Si on se limite aux tableaux standards de rubans, cela nous donne une action naturelle de $H_m(0)$, qui nous permet alors de définir des modifications locales sur les tableaux de rubans, que nous appellons flips et pseudo-flips. Ce sont des généralisations du flip classique de dominos. Grâce aux relations de Yang-Baxter on peut alors donner des invariants sur les chaînes minimales de flips, qui se conservent quand on passe aux pavages par rubans.*

1. Introduction

The goal of this article is to study a class of tilings called ribbon tilings of a Ferrers diagram. The main motivation for this study is to give a general, algebraic generalisation of domino tilings. In order to try and give some order and algebraic structures to these tilings, we will use ribbon tableaux.

Ribbon tableaux originate from rim hook tableaux, introduced to study the representations of the symetric group [**dBR61, GJ81**]. These tableaux have been studied from an algebraic point of view (see [**SW85, CL95, LLT97**]). In particular, [**SW85**] gives us a bijection between Ferrers diagrams and n-tuples of Ferrers diagrams, with interesting properties regarding ribbons. This bijection can be extended to a bijection between ribbons tableaux and n-tuples of Young tableaux [**Pak90**].

Section 2 gives some definitions and notation, then recalls some existing results :

We first define ribbon tilings and tableaux, recall basic facts about them and define the Stanton-White bijection. If we restrict our view to standard ribbon tableaux, we obtain from this bijection standard n-tuples of Young tableaux, or equivalently skew standard Young tableaux, upon which there are classical algebraic structures.

We then consider one such algebraic structure, the Hecke algebra for $q = 0$, $H_m(0)$, which is related to posets (that is partially ordered sets). We recall its classical actions on permutations and Young tableaux.

In Section 3, we extend this action to standard ribbon tableaux, using the Stanton-White bijection. By giving a $H_m(0)$-module structure to standard ribbon tableaux, we prove that it has a lattice structure, which we study.

In Section 4 Having thus given a poset structure to standard ribbon tableaux, we study the covering relation, using elementary generators of $H_m(0)$. We obtain local actions called pseudo-flips and flips, the latter being a generalisation of the flips encountered with domino tiling. After giving a geometric description and classification of these flips, we prove an invariant results concerning minimum flip paths, the key of the proof being the Yang-Baxter relations met by $H_m(0)$.

2. Definitions and existing results

In this section we first define ribbon tilings, ribbon tableaux and define some conventions and notation. We then give some basics facts about these objects, before recalling the Stanton-White bijection and the induced bijection for ribbon tableaux.

2.1. Presentation of ribbon tilings and notation. All the geometric objects which we define will be placed in the discrete plane $\mathbb{N}^2$, and we identify a unit square with its lower left corner.

A *n-ribbon* is a polyomino (that is a finite part of the discrete plane) formed by n squares, defining a path composed only of left or up steps. (It is a simply connected polyomino.) Therefore we can define the *head* of a ribbon as its bottom right square. An n-ribbon can thus be given be the coordinate of its head in the discrete plane, and by a word in $\{0,1\}^{n-1}$ coding its shape, each 0 representing a left step, and each 1 a up step. Figure 1 gives two examples of 8-ribbons.

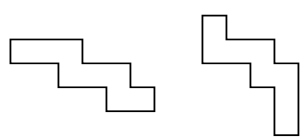


Figure 1. Two 8-ribbons

Two particular cases of ribbons are 1-ribbons, which are elementary squares in the discrete planes, and 2-ribbons, which are the classical dominoes, upon which much litterature exists [**CL95**].

Given a partition λ, that is a decreasing integer sequence of finite length, its *Ferrers* diagram is the shape in $\mathbb{N}^2$ whose length rows are givens by the terms of λ. (We use the cartesian convention, and thus rows lengths are decreasing upward.) We identify a partition and its Ferrers diagram, and call the set of all partitions (or equivalently of all Ferrers diagrams) Π.

Π is the set of all partitions, considering two partitions equals if we can obtain one from the other by adding some zeros at its right. For a partition λ its *length* $l(\lambda)$, is its length as a finite sequence (the tail zeros do not count in the length). The *weight* of λ, denoted by $|\lambda|$, is the sum of its terms.

Since we are in the discrete plane, we can define the diagonal d as $\Delta_d = \left\{(x,y) \in \mathbb{N}^2, x-y=d\right\}$. The *content* of a cell is the diagonal which it belongs to. The content of a ribbon is the content of its head.

Let us now define a *ribbon tiling* : We fix an integer n, and we tile λ by removing n-ribbons from its rim, in such a way that the remaining polyomino is still the Ferrers diagram associated with a partition, and then we go on, until we cannot remove ribbons anymore. Is is a classical result that the remaining partition, $\lambda_{(n)}$ does not depend on how the ribbons were removed (see [**dBR61, GJ81**]), and is called the n-core of λ. Thus all the tilings we can define this way are tilings of the same part of the discrete plane, $\lambda \setminus \lambda_{(n)}$. Figure 2 gives two ribbon tilings of the same Ferrers diagram.

General ribbon tiling have been studied in [**She**], where it was proved that the set of all tilings of a given Ferrers diagram is in bijection with the set of all acyclics orientation of a particular graph. Reversing

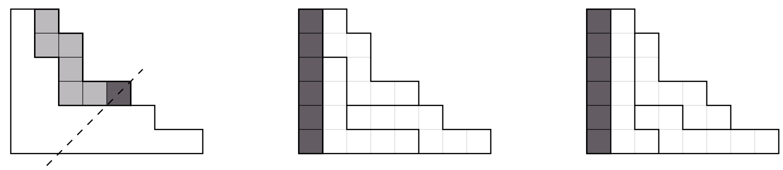


FIGURE 2. An example of rim 7-ribbon with its head and diagonal, and two 7-tiling of $(8, 6, 5, 3, 3, 2)$ with the 7-core shaded.

an edge in a given orientation, when it is possible then gives an action on ribbon tilings. This action is a local one called flip, for which the set of all tilings of a diagram is connex. But it does not give additional structure. We will give an algebraic interpretation for these flips.

Ribbon tilings of Ferrers diagrams have been studied in [**Pak90**], which defines a family of functions on ribbons. It is then proved that these function form a basis for the tiling invariant. That is to say that for a giver Ferrers diagram λ these functions are all invariant on the ribbon tilings of λ, and every such invariant can be obtained from these functions.

Given a Ferrers diagram λ, a *Young tableau* of shape λ is a filling of the cells of λ with strictly positive integers, in such a way that in each row the numbers are weakly increasing, and in each column they are strictly increasing upward.

It is natural to extend this notion to ribbons, which gives ribbon tableaux. A *n-ribbon tableau* of shape $\lambda \setminus \lambda_{(n)}$ is a tiling of $\lambda \setminus \lambda_{(n)}$ by n-ribbons to which we give integer numbers with the following growth conditions : On each row and each colon, the numbers of the encoutered ribbons must be weakly increasing, and the head of a ribbon cannot be above another ribbon with the same number. A young tableau is actually a 1-ribbon tableau, therefore definitions which apply to both will be given for ribbon tableaux.

We can define the *weight* of a ribbon tableau (thus of a Young tableau) as the sequence $(w_i)_i$ where w_i is the number of ribbon whose number is i, and a *standard ribbon tableau* will be a ribbon tableau of weigth $(1, 1, 1, 1, 0, 0, \dots)$. A standard ribbon tableau can be seen as a ribbon tiling, together with in which order the ribbons are added from the core. An example of standard ribbon tableau is given in figure 3. From now on, all the ribbon tableaux considered will be standard.

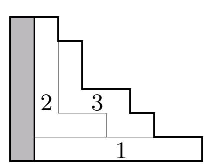


FIGURE 3. A standard ribbon tableau of shape $(8, 6, 5, 3, 3, 2) \setminus (1^6)$

The set of ribbon tableau and the set of standard ribbon tableaux of shape $\lambda \setminus \lambda_{(n)}$ will be denoted respectively by $Tab_n(\lambda \setminus \lambda_{(n)})$, and $STab_n(\lambda \setminus \lambda_{(n)})$.

2.2. The Stanton-White bijection.

Let us now recall here the classical Stanton-White bijection, which occurs for a given n-core μ between all partitions whose n-core is μ and all the n-tuples of partitions. This bijection can thus be considered as a bijection between the set all partitions and the outer product of the set of all n-tuples of partitions and the set of all n-cores. Let us define it algorithmically before giving an interpretation :

Given a partition $\lambda \in \Pi$, first add to it the stair $\Lambda_m = (m-1, m-2, \dots, 2, 1, 0)$ with m such that $m \geq l(\lambda)$ and $m = \alpha n$ with $\alpha \in \mathbb{N}^*$. We thus obtain a strict partition λ^Λ of length m, which we decompose modulo n into n partitions $\lambda^\Lambda(0),\dots,\lambda^\Lambda(n-1)$ in the following way : For each term x of λ, we make the integer division by n, $x = nq + r$, and add the term q (even if $q = 0$) to $\lambda^\Lambda(r)$. (By separating λ^Λ modulo n, we actually separate λ along its diagonals modulo n.) We set l_i as the length of $\lambda^\Lambda(i)$ (including a possible 0), and then substract Λ_{l_i} to $\lambda^\Lambda(i)$, to obtain a partition $\lambda^{(n)}(i)$. $(\lambda^{(n)}(1),\dots,\lambda^{(n)}(n-1))$ is then an n-tuple of partitions corresponding to λ, called the n-quotient of λ, and is denoted by $\lambda^{(n)}$.

This function is surjective, as for any n-tuple of partitions $\lambda^{(n)}$, the $\lambda^\Lambda(i)$ are simply obtained by adding stairs, and then λ^Λ by multiplying the terms $\lambda^\Lambda(i)$ by n and adding i to them, and merging the resulting partition into a strict partition λ^Λ. We then just have to substract a stair to λ^Λ to obtain a partition whose image is $\lambda^{(n)}$.

Let us illustrate this bijection by taking $\lambda = (5,5,3,2,1)$ and $n = 3$, we add $(5,4,3,2,1,0)$ to λ to get $\lambda^\Lambda = (10,9,6,4,2,0)$, which we divide modulo 3 to obtain $\lambda^\Lambda(0) = (3,2,0)$, $\lambda^\Lambda(1) = (3,1)$ and $\lambda^\Lambda(2) = (0)$. Substracting the corresponding stairs, we get $\lambda^{(3)}(0) = (1,1,0)$, $\lambda^{(3)}(1) = (2,1)$ and $\lambda^{(3)}(2) = (0)$. This is illustred in figure 4

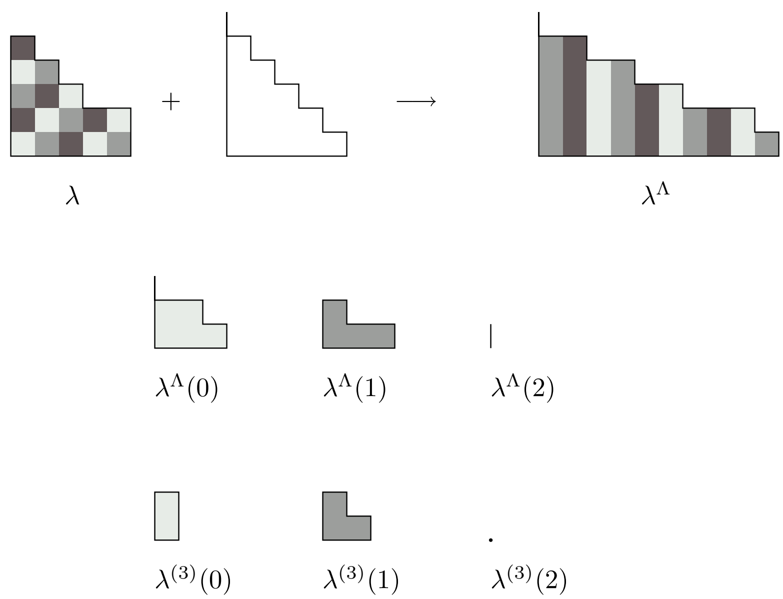


FIGURE 4

This transformation also gives us a coding of $\lambda_{(n)}$, by setting $w_i = l_i - l(\lambda)$. w is then an n-dimensional vector with sum equal to 0, which depends only of $\lambda_{(n)}$.

If we call $\mathbb{Z}_0^n$ the set of all n-dimensionals vectors with sum 0. We then can state the following result ([**GJ81, SW85**]) :

Theorem 2.1. *The function* $\lambda \mapsto (\lambda^{(n)}, \lambda_{(n)})$ *is a bijection between* Π *and* $\Pi^n \times \mathbb{Z}_0^n$

For $n = 1$, $\lambda^{(1)} = (\lambda)$, and $\lambda_{(1)}$ is empty, so the Stanton-White bijection is the canonical bijection between Π and $\Pi \times \{0\}$.

Note that if μ is obtained by removing an n-ribbon from λ's rim, then μ^Λ is obtained from λ^Λ by substracting n to a term and sorting the remaining terms. This term is equal modulo n to the content of the ribbon removed. Furthermore, each term of λ^Λ to which we can substract n (in such a way that μ^Λ remains a strict partition) corresponds to a ribbon which can be removed from λ's rim, as can be seen in figure 5

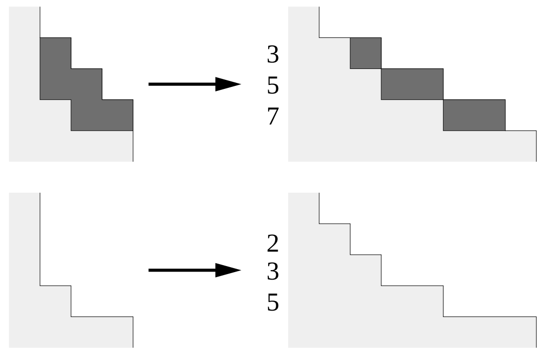


FIGURE 5. Removing a 5-ribbon from λ^{Λ}

We then obtain $\mu^{(n)}$ from $\lambda^{(n)}$ by removing a square of diagonal $d' + c_i$ from λ^i, where $d = nd' + i$ is the integer division of d by n (the c_i are offset parameters which depend on $\lambda_{(n)}$). So the squares of $\lambda^{(n)}$ correspond to ribbons, and the number of ribbons in a tiling of $\lambda \setminus \lambda_{(n)}$ is given by $\sum_{i=0}^{n-1} |\lambda^i|$.

This bijection translates to ribbon tableaux. The corresponding objects are then n-tuples of Young tableaux, each square corresponding to a precise ribbon. (The growth condition on ribbon tableau gives precisely classical growth condition of Young tableaux.) So, for a partition λ we have a bijection between $Tab_n(\lambda \setminus \lambda_{(n)})$ and all n-tuples of Young tableaux of shape $\lambda^{(n)}$, and this bijection preserves the weight. Figure 6 gives an example for this bijection.

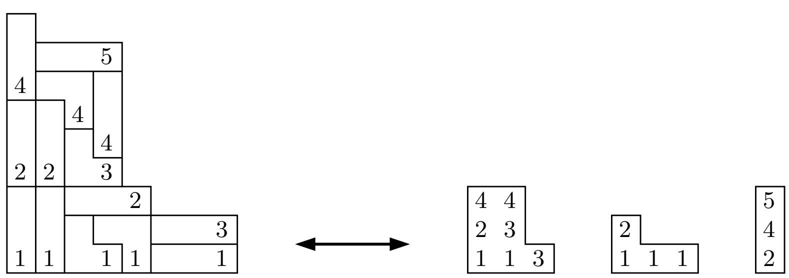


FIGURE 6

Moreover, we can arrange an n-tuple of Young tableau in such a way as to obtain a skew Young tableau. Figure 7 shows how it is done for the example we took in figure 6. (There are $2^n n!$ skew different Young

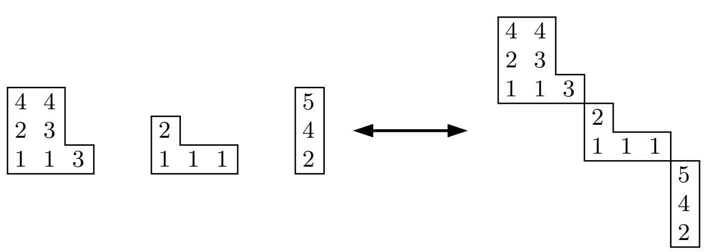


FIGURE 7

tableaux we can obtain this way, by taking the n Young tableaux in different orders and by transposing some of them.)

So for a given shape λ, we can obtain a skew Ferrers diagram $\mu \setminus \nu$ such that we have a bijection between $Tab_n(\lambda \setminus \lambda_{(n)})$ and skew Young tableaux of shape $\mu \setminus \nu$, which leaves the weight invariant. If we restrict ourselves to standard ribbon tableaux, we then have a bijection between standard ribbon tableaux and standard Young tableaux, upon which we can define algebraic structures, which we are now going to do.

2.3. The 0-Hecke algebra $H_m(0)$ and its actions.

In this section we define a classical combinatorics algebra, the 0-Hecke algebra ([**KT97**]) and explain how it can be used to give en ordering structure on permutations and Young tableaux.

We define $H_m(0)$, the *Hecke algebra* for $q = 0$ (over the field of complex numvers), as the $\mathbb{C}$-algebra spanned by $n-1$ generators $T_1, \ldots, T_{m-1}$ with the following relations* :

$$\begin{cases} {T_i}^2 = T_i & \forall 1 \le i \le m-1 \quad (1) \\ T_iT_j = T_jT_i & \text{if } |i-j| > 1 \quad (2) \\ T_iT_{i+1}T_i = T_{i+1}T_iT_{i+1} & \forall 1 \le i \le m-2 \quad (3) \end{cases}$$

Actually, $H_m(0)$ (and more generally $H_m(q)$) is a deformation of the symetric group, and such algebras can be defined for all Coxeter groups. If we take $q = 1$, $H_m(1)$ is $\mathbb{K}\mathfrak{S}_m$, the symetric group algebra.

Let us first remark that (1) can can be rewritten as $T_i(T_i - 1) = 0$, and thus 0 and 1 are possible eigenvalues for T_i seen as an operator. The fact that 0 is a possible eigenvalue is an important aspect of $H_m(0)$, as we will see later.

Given a permutation $\sigma \in \mathfrak{S}_m$, let us define T_σ in the following way : starting from a elementary decomposition of σ, $\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_k}$, where $\sigma_i = (i, i+1)$, we set $T_\sigma = T_{i_1}T_{i_2}\ldots T_{i_k}$. T_σ does not depend on the elementary decomposition chosen, as all such decomposition are congruent by the braid relation in the symetric group, and thus the corresponding products of T_i are congruent by (3).

Let us now define two natural actions of $H_m(0)$ on the symetric group algebra $\mathcal{K}\mathfrak{S}_m$, the left action L and the right action R. L is also called the *action on values* and R the *action on places.*

$$\begin{cases} R(T_i)\sigma = \sigma & \text{if } \sigma(i) < \sigma(i+1) \\ R(T_i)\sigma = \sigma\sigma_i & \text{otherwise} \end{cases}$$

$$\begin{cases} L(T_i)\sigma = \sigma & \text{if } \sigma^{-1}(i) < \sigma^{-1}(i+1) \\ L(T_i)\sigma = \sigma_i\sigma & \text{otherwise} \end{cases}$$

These action are in duality, by $L(T_i)\sigma = 132R(T_i){\sigma^{-1}}^{-1}$. From now on, we will only consider the right action of $H_m(0)$, which is shown in figure 8 for $n = 3$.

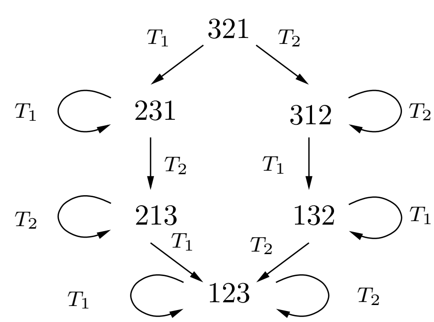


FIGURE 8. The right action of $H_3(0)$ on $\mathfrak{S}_3$

*Note that this is not the usual presentation of $H_m(0)$. The Hecke algebra for q given is usually defined by the quadratic relation $(T_i - 1)(T_i + q) = 0$, or equivalently ${T_i}^2 = (1-q)T_1 + q$, hence for $q = 0$ ${T_i}^2 = -T_i$. What we define here is a renormalisation obtained by replacing T_i with $-T_i$, which is more convenient in the present case.

Actually, $H_m(0)$ is linked with bubble sort : The action of a T_i is exactly an elementary step of bubble sort, and if we take the maximal permutation $\omega_0 = (m, m-1, \ldots, 2, 1) = \sigma_1\sigma_2\sigma_1\sigma_3\sigma_2\sigma_1 \ldots$ $\sigma_{m-1}\sigma_{m-2}\ldots\sigma_2\sigma_1$, then T_{ω_0} corresponds to the whole bubble sort.

Moreover, if we define a partial ordering on $\mathfrak{S}_n$ by setting $\sigma \preccurlyeq \sigma'$ when there exists a permutation ω such that $\sigma = T_\omega \sigma'$ or $\sigma = \sigma'$, then we find the classical weak order of $\mathfrak{S}_m$. The graph whose vertices are the permutations of $\mathfrak{S}_n$ and the edges are the action of the T_i minus the loops is the right n-permutaedron. The poset whose Hasse diagram is the permutaedron is a graded lattice ([**GR63, Bjö84**]), its maximum element is the trivial permutation 1 and its minimum element the permutation ω_0. (If we take the left action instead of the right action, we obtain another poset structure on $\mathfrak{S}_n$, isomorph to the one defined by the right action.) Note that this lattice is not distributive for $n \geq 3$.

(We recall that a *lattice* is a poset in which every pair of elements admits a supremum and infimum, noted by $\vee$ and $\wedge$. A lattice is *distributive* if $\vee$ and $\wedge$ are each distributive over the other, and a lattice P is *graded* if it admits a graduation, that is a function $f : P \to \mathbb{Z}$ such that if a covers b in P, then $f(a) = f(b) + 1$. See [**DP90**])

This action can easily be extended to space spanned by all the standard Young tableau of a given shape λ with $|\lambda| = m$: Given a standard Young tableau t, we call σ_t the permutation given by its line reading, and inversely t_σ is the tableau of shape λ whose line reading is σ.

$$\begin{cases} T_i t = t_{L(T_i)\sigma_t} & \text{if } i \text{ and } i+1 \text{ are not adjacent in } t \\ T_i t = 0 & \text{otherwise} \end{cases}$$

If i and $i+1$ are not adjacent, they cannot be on the same line as t is standard, and so the action of T_i is to reorder them so that $i+1$ is higher than i. Thus T_{ω_0} reorders t into the line filling of λ. This is illustrated in figure 9 for the tableaux of shape $(3, 2)$.

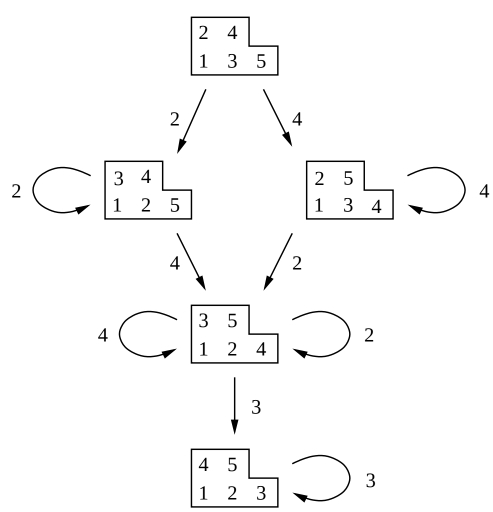


FIGURE 9. The action of $H_5(0)$ on standard Young tableaux of shape $(3, 2)$

This defines what is called a Specht module structure on the vector space spanned by standard Young tableaux of shape λ. It also gives an ordered structure on standard Young tableaux : Given two standard Young tableaux of same shape, t and t', we set $t \preccurlyeq t'$ if there exists a permutation ω such that $t = T_\omega t'$ (including the cas $t = t'$). The same kind of results hold as for the permutation group :

Theorem 2.2. *Take $\lambda \in \Pi$, and $m = |\lambda|$. The poset defined by the action of $H_m(0)$ on the set of all standard Young tableaux of shape λ is a lattice. The maximum element of the lattice is the row filling of λ and the minimum element is the column filing of λ.*

This lattice is usually not distributive. (For example the lattice of all standard Young tableaux of shape $(3, 2, 1)$ admits the 3-permutaedron as a sub-lattice.)

3. The lattice structure on standard ribbon tableaux

Having recalled all these results, we can now extend them to standard ribbon tableaux. This action of $H_m(0)$ on standard Young tableaux naturally extends in the same way to skew Young tableaux, and it keeps all its properties, including the lattice structure it defines.

As the Stanton White bijection induces a bijection between standard ribbon tableaux and standard skew Young tableaux, we naturally have an action of $H_m(0)$ on the vector space spanned by $STab_n(\lambda \setminus \lambda_{(n)})$ (where m is the number of n-ribbons in a tiling of $\lambda \setminus \lambda_{(n)}$, that is $\sum |\lambda^{(n)}(i)|$), and the following result :
Theorem 3.1. *$H_m(0)$ induces a graded structure lattice on $STab_n(\lambda \setminus \lambda_{(n)})$*

Figure 10 gives this structure for $STab_3 132(4^3)$, and figure 11 gives the isomorphic structure for the corresponding triplet of Ferrers diagrams, which is $132(2), (1), (1)$. Note that in this case different standard ribbons tableaux give different ribbons tilings, but this is not always the case.

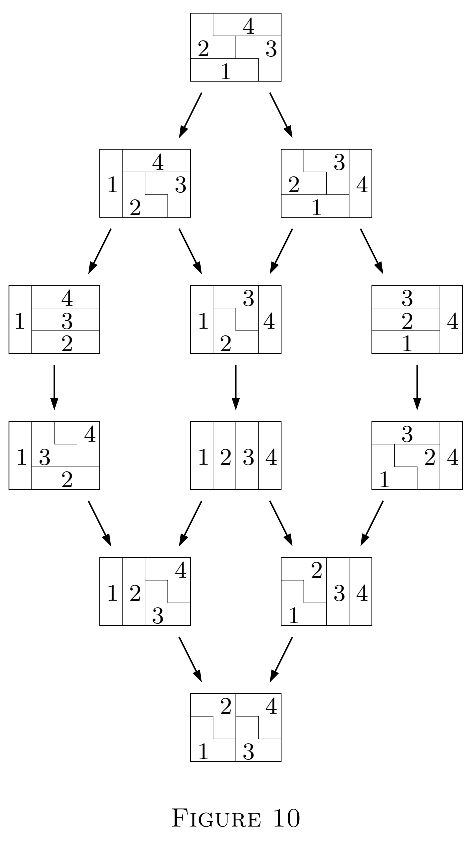


FIGURE 10

Depending on the way we build the skew Ferrers diagram on which $H_m(0)$ acts, we can obtain up to $2^n n!$ different action of $H_m(0)$, which will define as many different lattice structures (some of them will often coincide), but the non-oriented graph underlying the Hasse diagram will remain the same, as the edge

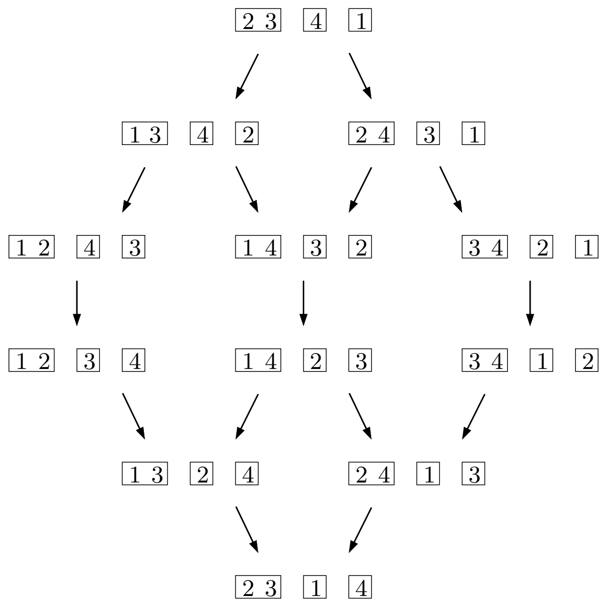


FIGURE 11

are given by switchs between consecutive numbers in the n-tuples of Young diagram corresponding to the standard ribbon tableaux.

4. Flips, pseudo-flips and the Yang-Baxter relation

In a first part we define flips and pseudo-flips, which are local actions given by the action of the generators of $H_m(0)$. This definition allows us in a second subpart to classify and count the possible flips. We then use the classical Yang-Baxter relations to give invariants on minimum flip paths.

4.1. Local action of the T_i.

We have in the previous section defined a bipartite graph whose vertices set is $STab_n(\lambda \setminus \lambda_{(n)})$, and whose edges are given by the covering relation of one the lattice defined by the actions of $H_m(0)$, the graph itself being independant of the chosen action of $H_m(0)$. It is natural to study this covering relation, more specifically the action of a T_i on the ribbon tableaux themselves.

This covering relation is given by the action of a T_i, that is the switching of i and $i+1$ in the n-tuple of Young tableaux. A standard ribbon tableau can be constructed from the corresponding Young tableaux by adding ribbons with increasing or decreasing number, or both. Thus if we have two standard n-tuples of Young tableaux, $(t_j)_{0 \geq j \geq k-1}$ and $(t'_j)_{0 \geq j \geq k-1}$ which can be obtained from the other by switching i and $i+1$, we can construct the associated ribbon tableaux by placing the ribbons 1 to $i-1$ and then the ribbons n downto $i+2$ on the border of λ, and these will be the same in the two ribbons tableaux.

So the action of T_i is a local one, which only acts on the ribbons i and $i+1$, leaving all the others unchanged. Its effects depend on the difference Δd between the diagonals of the two ribbons, d_i and d_{i+1}, and whether it's bigger or lesser than n. Δ_d cannot be equal to n, because then i and $i+1$ would be in the same t_j, and they would be on adjacent diagonals, thus they would be adjacents, in which case they could not be switched.

If $\Delta d > n$, then the two ribbons are disjoint, so switching i and $i+1$ in t amounts to switching the numbers of the ribbons. We call this transformation a *n-pseudo-flip* (or just pseudo-flip) of *genus* Δd. A pseudo-flip changes the standard ribbon tableau, but leaves the underlying ribbon tiling invariant.

If $\Delta d < n$, then the two ribbons overlap, and by switching i and $i+1$ we change the ribbon which overlaps the other, and then we have a *n-flip* of *genus* Δd.

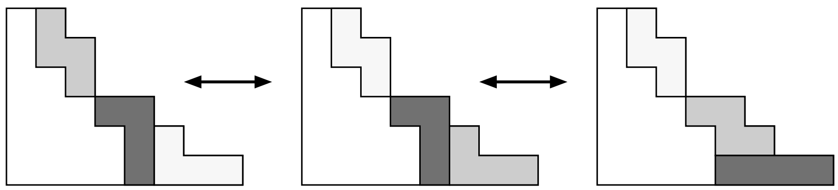


FIGURE 12. A pseudo-flip and a flip of 4-ribbons

Figure 12 gives examples of flip and pseudo-flip for 4-ribbons. If we only consider the underlaying ribbon tiling, a flip between standard ribbon tableaux defines a flip between two tilings. This is the same flip as was defined in [**She**], even tough it was apparently differently defined. For $n = 2$, we find the usual domino flip.

4.2. Geometric classification and enumeration of flips.

The genus of a flip describes the geometry of this flip it is the difference between the diagnoals of the two ribbons involved, and so it tells the length along which the two ribbons overlap. In a flip of genus d between n-ribbons, the ribbons overlap on $n-d+1$ cells. This gives a geometric classification of flips, and will allow us to count the possible flips :

Theorem 4.1. *There is $(2^{n-1}-1)2^{n-2}$ different possible geometries for flips of n-ribbons.*

PROOF. We will call $F_d(n)$ the number of geometrically differents n-flips of genus d, and $F(n)$ the number of all n-flips.

When $d > 1$, as only these $n-d+1$ cells are involved in the flip, the geometry of this flip is the same as for a flip of genus 1 between ribbons of length $n-d+1$. So an n-flip of genus d geometrically consis tof two parts : First, the overlapping part, where the flip occur, which is a flip of genus 1 of $(n-d+1)$-ribbons, and then the remaining part of the ribbons, which can be seen as two $d-1$ ribbons starting from the overlapping part. From this we can derive the relation $F_d(n) = 2^{2(d-1)}F_1(n-d+1)$.

In order to compute $F_1(n)$ let us see what exactly is a flip of genus 1, which we will call a maximal flip.

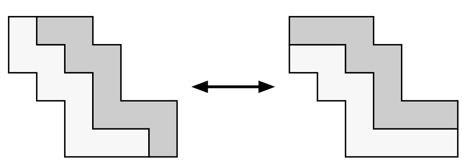


FIGURE 13

In such a flip, the ribbons' heads are adjacent, and the ribbons overlap on their whole length. In one of the two configurations involved in the flip, the ribbon's head will be side to side, and so the one with the greater diagonal (that is the one whose head is on the right) will have a "up" as its first step, as can be seen in figure 13. Now, if we take such a ribbon with an upper first step, we can put a ribbon immediatly to its left, such that these two ribbons overlap on their whole length, and it is possible to do a maximal flip on these ribbons. So a flip of genus 1 can be coded by a ribbon whose first step is imposed, thus by a word of

$\{0,1\}^{n-1}$. This word can also be seen as coding the intersection of the common boundaries of the ribbons in the two configurations of the flip, as shown in figure 14

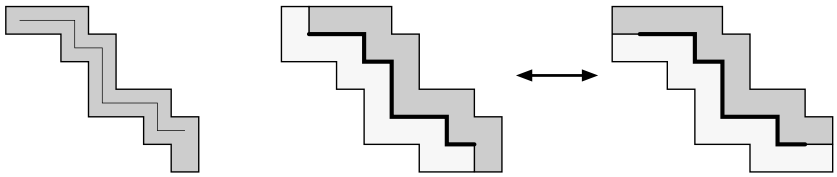


FIGURE 14. A 12-flip of genus 1, whose coding is $(0,1,0,0,1,1,0,1,0,0)$

This gives us $F_1(n) = 2^{n-2}$, and so :

$$F_d(n) = 2^{2(d-1)}2^{n-d-1} = 2^{n+d-3}$$

$$F(n) = 2^{n-3}\sum_{d=1}^{k-1} 2^d = (2^{n-1}-1)2^{n-2}$$

□

For $n = 3$, we have $F_1(3) = 2$, $F_2(3) = 4$ and $F(3) = 6$. These six flips of 3-ribbons are given in figure 15 and 16.

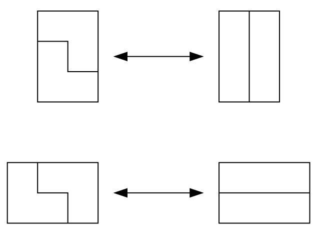


FIGURE 15. The 3-flips with genus 1

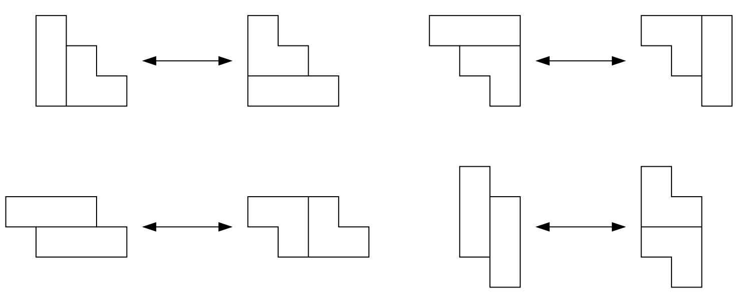


FIGURE 16. The 3-flips with genus 2

Remark that taking $n = 2$, we have $F(2) = 1$, so this is consistent with the trivial fact on domino tilings that there exists only one kind of flip.

4.3. Minimal flip paths and Yang-Baxter relations.

We will now prove the following result :

Theorem 4.2. *Given t and t' two standard ribbon tableaux of same shape, the set of the genus (with multiplicity) of the flips and pseudo-flips in a path between t and t' is invariant for minimal length paths.*

Proof. Let's consider two standard ribbons tableaux t and t' of same shape $\lambda \setminus \lambda_{(n)}$, and minimal length paths of flips and pseudo-flips between these paths. By chosing a way to rearrange $\lambda^{(n)}$ in a skew partition, we can see t and t' as the permutations σ_t and $\sigma_{t'}$ corresponding to the line-readings of the skew Young tableaux associated. Flips and pseudo-flips paths then are chains of elementary transpoition from σ_t to σ_t'. As these paths are minimal, the corresponding chains of transpositions are congruent by the braid relation, and so we can use following result for Coxeter groups :

An elementary transposition, σ_i, switches i and $i+1$ in σ. We will associate to this transposition a factor $(x_{\sigma^{-1}(i+1)} - x_{\sigma^{-1}(i)})$, where $x_1, \ldots, x_n$ is a set of independant variables. This gives the places of the switched terms in σ. For a chain of transposition $\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_m}$ from σ_t to $\sigma_{t'}$, let us take the product

$$(x_{\sigma_t{}^{-1}(i_1+1)} - x_{\sigma_t{}^{-1}(i_1)})(x_{(\sigma_t\sigma_{i_1})^{-1}(i_2)} - x_{(\sigma_t\sigma_{i_1})^{-1}(i_2+1)})\ldots$$

which gives all the places of the switched elements in the path from σ_t to σ_t'. Then this product is invariant for all chains of transposition from t to t' which are congruent by the braid relation.

Now let us specialize the x_j into another set y_l by setting $x_j = y_{d(j)}$ where $d(j)$ is the diagonal of a ribbon corresponding to the element at place j in σ_t. $(x_{j_1} - x_{j_2})$ becomes $(y_{l_1} - y_{l_2})$ where l_1 and l_2 are the diagonals of the ribbon involved in the flip or pseudo-flip. $(y_{l_1} - yl_2)$ then gives the genus of the flip or pseudo-flip, and so the multiset of genus is invariant for all minimal paths between t and t'. □

Two such minimal paths are given in figure 17 for 4-ribbons.

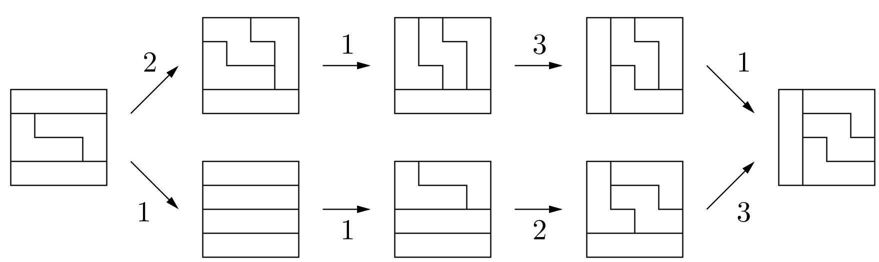


Figure 17. Two minimal flip paths between two 4-tilings, with the genuses of the flips shown

5. Conclusion and perspective

Up to this day, ribbon tilings and ribbon tableaux have been considered two completly different kinds of combinatorics objects, studied by different people with different methods. This paper is intended as an attempt to link these domains, altough the link is somewhat tenuous.

One important question relative to this topic is to know whether the different lattice structures thus defined on $STab_n(\lambda \setminus \lambda_{(n)})$ induce lattice structures on the set of n-ribbon tillings of $\lambda \setminus \lambda_{(n)}$. Actually, the present study was motivated by this problem. Alas several unsolved related questions remain unsolved.

In particular, we meet the following problem : given two standard n-tuples of Young tableaux of same shape, is there a simple way to determine if the associated ribbon tableaux have the same underlying ribbon tilings, without having to compute these ribbon tableaux ?

References

[Bjö84] A. Björner, *Orderings of coxeter groups*, Combinatorics and Algebra (Providence, R. I.) (C. Greene, ed.), vol. 34, Am. Math. Society, 1984, pp. 175–195.

[BW88] A. Björner and M. Wachs, *Generalized quotients in coxeter groups*, Trans. of the Am. Math. Society (1988).

[CL95] C. Carré and B. Leclerc, *Splitting the square of a schur function i nto its symetric and antisymetric parts*, Journal of algebraic combinatorics **4** (1995), 201–231.

[dBR61] Gilbert de B. Robinson, *Representation theory of the symetric group*, 1961.

[DP90] P. A. Davey and H. A. Priestley, *An introduction to lattice and order s*, Cambridge University Press, 1990.

[GJ81] A. Kerber G. James, *The representation theory of the symetric group*, 1981.

[GR63] G. Guilbaud and P. Rosenstiehl, *Analyse algébrique d'un scrutin*, M ath. Sci. Hum (1963), 9–33.

[KT97] D. Krob and J.-Y. Thibon, *Noncommutative symetric functions IV: Qua ntum linear groups and hecke algebras at* $q = 0$, Journal of algebraic cominato rics **6** (1997), 339–376.

[Las] A. Lascoux, *Ordering the affine symmetric group*, http://www-igm.univ-mlv.fr/ al/ARTICLES/OrderingAffine.ps.gz.

[LLT96] A. Lascoux, B. Leclerc, and Jean-Yves Thibon, *Hecke algebras at roots of unity and canonical bases of quantum affine algebras*, Communications in Math. Physics **181** (1996), 205–263.

[LLT97] A. Lascoux, B. Leclerc, and J.-Y. Thibon, *Ribbon tableaux, hall-littlewood functions, quantum, affine algebras, and unipotent varieties*, Journal of math. physics **38 (2)** (1997).

[LT00] B. Leclerc and J.-Y. Thibon, *Littlewood-richardson coefficients and kazhdan-lusztig polynomials*, Advanced Studies in Pure Mathematics **28** (2000), 155–200.

[MP] C. Moore and I. Pak, *Ribbon tile invariants from signed area*, Journa l of combinatorial theory **Series A**.

[Pak90] I. Pak, *Ribbon tile invariants*, Trans. of the Am. Math. Society (199 0).

[Rém99] É. Rémila, *On the lattice structure of the set of tilings of a simply connected figure with dominoes*, Proccedings of the 3rd Internatio nal Conference on Orders, Algorithms and Applications (ORDAL), 1999.

[She] S. Sheffield, *Ribbon tilings and multidimensional height functions*, arxiv.org:math.CO/0107095.

[SW85] D. W. Stanton and D. E. White, *A Schensted algorithm for rim hook tableaux*, Journal of combinatorial theory **Series A** (1985), 211–247.

École normale supérieure de Lyon (LIP), 46 allée d'Italie, 69364 Lyon Cedex 07, France
E-mail address: `gilles.radenne@ens-lyon.fr`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

On Inversions in Standard Young Tableaux

Michael Shynar

Abstract. *In this work, we present the inversion number of a standard Young tableau, and determine its distribution over certain sets of standard Young tableaux. Specifically, the work determines the distribution of the inversion number over hook-shaped tableaux and over tableaux of shape (n, n). We also study the parity (also referred to as 'sign balance') of the inversion number over hook-shaped tableaux and over $(n-k, k)$-shaped tableaux. The latter results resemble results in the field of pattern-avoiding permutations, achieved by Adin, Roichman and Reifegerste.*

1. Preliminaries

Definition 1.1. Let $n \in \mathbb{N}$ ($\mathbb{N}$ denotes the set of positive integers). A *partition* of n is a vector of positive integer numbers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_k = n$. We write $\lambda \vdash n$. We denote by $\lambda' = (\lambda'_1, \ldots, \lambda'_t)$ the *conjugate* partition, where λ'_i is the number of parts in λ greater or equal to i. We define $\mid \lambda \mid= n$.
Definition 1.2. The set $\{(i,j) \mid i,j \in \mathbb{N},\ i \leq k,\ j \leq \lambda_i\}$ is called the *Young diagram* of shape λ (notice that 'English notation' is used).
Definition 1.3. A *standard Young tableau* of shape λ consists of inserting the integers $1, 2, \ldots, n$ as *entries* in the cells of the Young diagram of λ, allowing no repetitions and having entries increase along rows and columns. λ is normally denoted $Sh(T)$.
Definition 1.4. A *descent* in a standard Young tableau T, is an entry i, such that i+1 is strictly south (and weakly west) of i. Denote the set of all descents in T by $D(T)$. We define two statistics on a standard Young tableau:

(a) The *descent number* of T. $des(T) = \sum_{i \in D(T)} 1$.
(b) The *major index* of T. $maj(T) = \sum_{i \in D(T)} i$.

Stanley has found a generalization of the hook formula, giving the generating function for $maj(T)$ when T is of shape λ.
Theorem 1.5 (Stanley's q-analogue of the hook formula, see [**ST2**]).

$$\sum_{shape(T)=\lambda} q^{maj(T)} = \frac{\prod_{k=1}^{n} [k]_q}{\prod_{(i,j)\in\lambda} [h_{i,j}]_q} \tag{1.1}$$

where $[m]_q = 1 + q + q^2 + \cdots + q^{m-1}$.

1991 *Mathematics Subject Classification.* Primary 05E10, 05E99; Secondary 20B99.
Key words and phrases. algebraic combinatorics, standard young tableau, inversion.
The author was supported in part by the internal research grants from Bar-Ilan University.

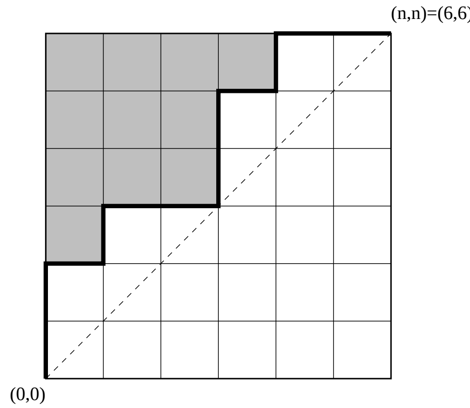


FIGURE 1. A Dyck path. The area of this Dyck path is shaded in gray.

Seeing how natural the generating function of the major index turns out to be, it is surprising that there is no similar result for the descent number. However, Adin and Roichman ina joint study, and Hästö in a parallel study, managed to establish the expected value and variance of $des(T)$ for a random standard Young tableau of a given shape (see [**AR1, H**]).

One can also think of defining the inversion number of a tableau. Not much is known regarding the distribution of the inversion number over tableaux of a fixed shape, and that is in fact the primary goal of this research.

Remark 1.6. The convention within this paper is that $\binom{n}{k} = 0$ when $k < 0$.

Definition 1.7. A *lattice path* in the plane is defined to be a sequence $L = (v_1, \dots v_k)$ where $v_i \in \mathbb{N}^2$ and $v_{i+1} - v_i \in \{(1,0),(0,1)\}$. The last condition indicates that when moving from v_i to v_{i+1}, we move either one unit north, or one unit east.

Definition 1.8. A *Dyck path* of order n is a lattice paths starting at $(0,0)$ and ending at (n,n), which always remain above or on the line $x = y$. A Dyck path can be encoded by a sequence $(a_1, \dots a_{2n})$ where $a_i \in \{1,-1\}$ with $a_i = 1$ indicating a north move at the i-th step, and $a_i = -1$ indicating an east move at the i-th step.

The *area above a dyck path D* (denoted: $area(D)$) is the area between D and the dyck path encoded by $\{\underbrace{1,1,\dots,1}_{n},\underbrace{-1,-1,\dots,-1}_{n}\}$.

Example 1.9. The Dyck path corresponding to the series $\{1,1,-1,1,-1,-1,1,1,-1,1,-1,-1\}$ is drawn in figure 1 with its area shaded in gray.

Recall that the number of Dyck paths of order n is called the *n-th Catalan number*, and is denoted C_n.

Recall the following well-known corollary of the q-binomial theorem (see [**GR**, page 7]):

Theorem 1.10. *(The Cauchy binomial theorem*[**GR**, page 7]*)*

$$\prod_{k=1}^{n}(1+yq^k) = \sum_{m=0}^{n} y^m q^{\binom{m+1}{2}} \begin{bmatrix} n \\ m \end{bmatrix}_q$$

Definition 1.11. The n-th Carlitz-Riordan q-Catalan number is defined as follows: $C_n(q) = \sum_{D \in Dyck(n)} q^{area(D)}$, where $Dyck(n)$ is the set of all Dyck paths of order n.

This q-Catalan number was studied by Carlitz and Riordan (see [**C, CR**]), and further studied by Fürlinger and Hofbauer in 1985 (see [**FH**], which also includes further references within). There is no known generating function for this q-Catalan number, however Fürlinger and Hofbauer expressed it as a term within a generating function, and several determinant formulas were provided, the most recent one by Loehr (see [**L**, Theorem 16]).

Lemma 1.12. [**FH**, Eq. 2.2] *The Carlitz-Riordan q-Catalan numbers abide to the recursion:*

$$C_{n+1}(q) = \sum_{k=0}^{n} C_k(q) C_{n-k}(q) \cdot q^{(k+1)(n-k)}$$

with starting condition $C_0(q) = 1$.

Remark 1.13. Some authors define the Carlitz-Riordan q-Catalan as $\tilde{C}_n(q) = q^{\binom{n}{2}} C_n(q)$. These numbers describe the distribution of the area between Dyck paths of order n, and the "diagonal Dyck path" $(1, -1, 1, -1, 1, -1, \ldots, 1, -1)$.

We cite the following "common knowledge" result:

Lemma 1.14. *For any two positive integers* $k \leq n$,

$$\begin{bmatrix} n \\ k \end{bmatrix}_{q=-1} \begin{cases} 0 & \begin{matrix} n\, even \\ k\, odd \end{matrix} \\ \binom{\lfloor \frac{n}{2} \rfloor}{\lfloor \frac{k}{2} \rfloor} & otherwise \end{cases}$$

Corollary 1.15.

$$\sum_{k=0}^{2n+1} \begin{bmatrix} 2n+1 \\ k \end{bmatrix}_{q=-1} q^k = (1+q) \sum_{k=0}^{n} \binom{n}{k} q^{2 \cdot k}$$

$$\sum_{k=0}^{2n} \begin{bmatrix} 2n \\ k \end{bmatrix}_{q=-1} q^k = \sum_{k=0}^{n} \binom{n}{k} q^{2 \cdot k}$$

2. Inversions in Tableaux and Signs of Tableaux

This chapter presents the most fundamental concept of the work.

As we saw in definition , there is a meaningful way to define the descent set of a tableau. The definitions of the descent number and the major index follow naturally. The following definition of an inversion in a standard Young tableau is natural as an extension of the descent definition. It is a variant of the definition given by Stanley (see [**ST3**, page 15]).

Definition 2.1. An *inversion* in a standard Young tableau is a pair (i, j) such that $i < j$ and the entry for j appears strictly south and strictly west of the entry for i. The *inversion number* of a standard Young tableau T (denoted: $inv(T)$) is the number of inversions in this standard Young tableau.

Definition 2.2. A *weak* inversion in a standard Young tableau T is a pair of integers (i, j) such that $i < j$ and j is weakly south and weakly west of i. The number of weak inversions in T is called the *weak inversion number* of T and denoted $winv(T)$.

There is a simple connection between the inversion and the weak inversion numbers: Let T be a standard Young tableau with $sh(T) = \lambda = (\lambda_1, \ldots, \lambda_k)$, and denote $\lambda' = (\lambda'_1, \ldots, \lambda'_{\lambda_1})$ to be the conjugate partition, then $winv(T) = inv(T) + \sum_{i=1}^{\lambda_1} \binom{\lambda'_i}{2}$.

Definition 2.3. Let T be a standard Young tableau. The *sign* of T is defined: $sign(T) = (-1)^{inv(T)}$.

3. Hook Shaped Tableaux

Definition 3.1. A *hook-shaped tableaux* is a tableaux with one row and one column. Alternatively, it is a tableaux T with shape $\lambda = (k, 1, 1, \ldots, 1)$ with $k \geq 1$. The *column length* of T (denoted $col(T)$) is defined as $\lambda'_1 - 1$, or equivalently, the number of parts in λ, reducing 1. The *row length* of T is defined as $\lambda_1 - 1$.

Definition 3.2. Write $sh(T) \in hook(n)$ if T is a hook-shaped standard Young tableau of order n. Write $sh(T) \in hook(n, k)$ if T is a hook-shaped standard Young tableau of order n with column length k.

Lemma 3.3.

$$F_{n,k}(q) = \sum_{sh(T)\in hook(n+1,k)} q^{inv(T)} = \begin{bmatrix} n \\ k \end{bmatrix}_q \tag{3.1}$$

PROOF. This proposition may be proved using the recursion $F_{n,k}(q) = F_{n-1,k}(q) + q^{n-k}F_{n-1,k-1}(q)$. □

Using Cauchy's binomial theorem (Theorem 1.10) we deduce that

$$\sum_{sh(T)\in hook(n+1)} q^{winv(T)} = \prod_{k=1}^{n}(1+q^k)$$

for a detailed proof see [**SH**].

4. Tableaux of Two Rows

4.1. Counting Inversions.

Definition 4.1. Let T be a standard Young tableau. If $sh(T) = (n-k,k)$ with $n-k, k \geq 0$, we say T is a *two-rowed tableau*, and write $T \in tworows(n)$. If $n-k = k$ we say T is *equal-rowed.*

Lemma 4.2. *Let $(x_1, x_2, \ldots, x_{2n})$ be an encoding of a Dyck path (see definition 1.8). In each Dyck path of order n there are exactly n 1's, call them $x_{a_1}, \ldots, x_{a_n}$. Then $area(D) = \sum_{i=1}^{n}(a_i - i)$.*

The proof of this proposition is left to the reader.

Theorem 4.3. *Recall the definition of $\tilde{C}_n(q)$ in note 1.13.*

$$\sum_{sh(T)=(n,n)} q^{inv(T)} = \tilde{C}_n(q) \tag{4.1}$$

PROOF. There is a well known bijection between Dyck paths of order n, and standard Young tableaux of shape (n,n): Take a standard Young tableaux T of shape (n,n). The corresponding Dyck path encoding is given by $a_i = 1$ if the entry i lies within the first row of T, and $a_i = -1$ if the entry i lies within the second row of T.

Now, observe that the entry values in T are uniquely determined by choosing the entries in the first row, since there is only one unique way to arrange the remainder "unused" entries in the second row. Moreover, the sum of entries in the first row uniquely determines the number of inversions.

To prove this, write:

$$T = \begin{matrix} a_1 & a_2 & \ldots & a_n \\ b_1 & b_2 & \ldots & b_n \end{matrix}$$

Notice that from the definition of an inversion, it must follow that any two entries i, j creating an inversion, must reside in two different rows. Thus, to calculate the number of inversions for our given tableau, it is sufficient to determine the number of inversions involving one element a_i and one element b_j $(j < i)$. Thus, we need to determine the number of elements $b_j < a_i$ with $j < i$. We know there are $a_i - 1$ values smaller than a_i, and $i-1$ of them are in the first row (all the entries to before the i-th entry), so there are $a_i - i$ entries smaller than a_i in the second row, and they all must appear in smaller column indices than i. That leaves room for exactly $i - 1 - (a_i - i) = 2i - a_i - 1$ entries larger than a_i in the second row, with a smaller column index, and hence that is also the number of inversions in which a_i participates. The number of total inversions in the tableau would be $\sum_{i=1}^{n}(2i - a_i - 1) = \binom{n+1}{2} - n + \sum_{i=1}^{n}(i - a_i)$. By proposition 4.2 we get $\sum_{i=1}^{n}(i - a_i) = -area(D)$. Thus, $\sum_{sh(T)=(k,k)} q^{inv(T)} = q^{\binom{n+1}{2}-n} C_n(\frac{1}{q}) = q^{\binom{n}{2}} C_n(\frac{1}{q}) = \tilde{C}_n(q)$. □

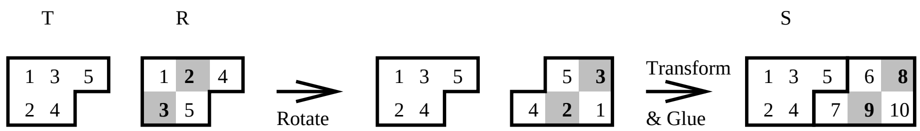


FIGURE 2. Gluing together equal-shaped tableaux. Notice that all inversions in the tableau R are preserved. The entry couple $(2,3)$ is an inversion. It is highlighted throughout the process. At the end it corresponds with the entry couple $(8,9)$ which is an inversion in S.

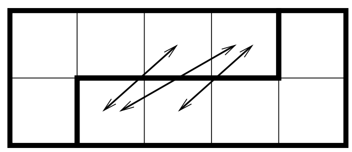


FIGURE 3. In this illustration, there are exactly 3 inversions involving exactly one entry from each of the two merged tableaux. The specific values within the tableaux do not make any difference here.

Corollary 4.4. *Let* $G_{n-k,k}(q) = \sum_{sh(T)=(n-k,k)} q^{inv(T)}$. *Then*

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} q^{\binom{n-2k}{2}} G_{n-k,k}(q)^2 = \tilde{C}_n(q)$$

PROOF. Let T and R be tableaux of shape $(n-k,k)$ $(0 \leq 2k \leq n$ with $n > 0)$. Any two tableaux of the same shape may be "glued" together in a certain fashion we shall describe, to obtain a standard Young tableau of shape (n,n), which we denote S. From there we use 4.1 to conclude the result.

Let T, R be standard Young tableaux of shape $(n-k,k)$. Take the second row of R. Reverse the order of elements in it, and then replace each element a by $2n-a-1$. Add the result to the end of T's first row. This is the first row of S. The second row is acquired from applying the same transformation on R's first row, and adding it to T's second row. It is required to verify that S is indeed a standard Young tableau, which is left as an exercise for the reader. Notice that all elements originating from R are bigger than all elements in T. See Figure 2 for an illustration of the process.

Now we look at the relation between inversions of T and R, and those of S. First notice that for any two entries $i < j$ in R, with j strictly southwest of i the corresponding entries in S would be $2n-i-1 > 2n-j-1$ and $2n-i-1$would be strictly southwest of $2n-j-1$. Thus, these entries would form an inversion in S. All inversions in T are obviously preserved in S. Furthermore: any inversion in S not derived from R or T would have to consist of one element from T and one element from R (else the inversion would have to appear either in R or S). The only inversions in S consisting of one element from T and one element from R could be found in the "middle region" of S, i.e. in column indices ranging from $k+1$ to $n-k$. All elements in R are bigger than those in T, thus we can calculate exactly this number of inversions: it is the number of matchings of elements from the first row within this "middle region", and elements of the second row, also in this "middle region". This gives us exactly $\binom{n-2k}{2}$ extra inversions. See Figure 3 for an example.

This transformation is a bijection from standard Young tableaux of shapes $(n-k,k)$ with $0 \leq k \leq \lfloor \frac{n}{2} \rfloor$ to (n,n)-shaped tableaux. Thus, the inversions over $q^{\binom{n-2k}{2}} G_{n-k,k}(q)^2$ distribute exactly as they do over (n,n)-shaped tableaux. □

4.2. Sign Balance. When addressing standard Young tableaux of shape $(n-k,k)$, we can give an explicit formula for the sign distribution.

Definition 4.5. $row_2(T)$ will denote the length of the second row of T. i.e. if T is of shape $(n-k,k)$ then $row_2(T)=k$.

Theorem 4.6. *Recall that we defined $\binom{n}{k}=0$ whenever k is negative. Then:*

$$\sum_{T\in tworows(2n+1)} sign(T)q^{row_2(T)} = \sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}(-1)^k\left[\binom{n}{k}-\binom{n}{k-1}\right]q^{2k}$$

$$\sum_{T\in tworows(2n)} sign(T)q^{row_2(T)} = (1+q)\sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}(-1)^k\left[\binom{n-1}{k}-\binom{n-1}{k-1}\right]q^{2k}$$

Proof. Denote $Sum(n,k)=\sum_{sh(T)=(n-k,k)} sign(T)$. Then it is sufficient to prove that for $0\le 2k\le n$:

(a) $Sum(2n+1,2k)=(-1)^k\left[\binom{n}{k}-\binom{n}{k-1}\right]$.
(b) $Sum(2n+1,2k+1)=0$. $(2k\ne n)$
(c) $Sum(2n,2k)=(-1)^k\left[\binom{n-1}{k}-\binom{n-1}{k-1}\right]$.
(d) $Sum(2n,2k+1)=(-1)^k\left[\binom{n-1}{k}-\binom{n-1}{k-1}\right]$. $(2k\ne n)$

The proof is done by induction on n. It is clear that $Sum(1,0)=1$. We give the induction step for the first case. The other three cases are very similar.

Take a tableau of shape $(2n+1-2k,2k)$ with $n\ge 2k>0$. Each such tableaux is uniquely achieved either by a $(2n+1-2k,2k-1)$-shaped tableaux with the entry $2n+1$ added to its second row, or by a $(2n-2k,2k)$-shaped tableaux with the entry $2n+1$ added to its first row. If the entry $2n+1$ resides in the first row, it would participate in no inversions, and thus its removal would not alter the sign. If it resides in the second row, it would participate in $2n+1-4k$ inversions (the number of elements in the first row with greater column index), and its removal would flip the sign. Thus,

$$\begin{aligned}
Sum(2n+1,2k) &= Sum(2n,2k)-Sum(2n,2(k-1)+1)= \\
&= (-1)^k\binom{n-1}{k}-\binom{n-1}{k-1}+(-1)^k\binom{n-1}{k-1}-\binom{n-1}{k-2}= \\
&= (-1)^k\left[\binom{n-1}{k}+\binom{n-1}{k-1}\right]-(-1)^k\left[\binom{n-1}{k-1}+\binom{n-1}{k-2}\right]= \\
&= (-1)^k\left[\binom{n}{k}-\binom{n}{k-1}\right]
\end{aligned}$$

Notice that this result would be true also for $k=0$ since then we would look only at $Sum(2n,2\cdot 0)=(-1)^0\left[\binom{n}{0}-\binom{n}{0-1}\right]=\binom{n}{0}$, which is also what we would get by substituting $k=0$ in the formula for $Sum(2n+1,0)$. ☐

5. The "$\frac{n}{2}$ Phenomenon"

The following results are corollaries of previous theorems in this work.

Theorem 5.1. *Let $sh(T)'=(\lambda_1',\dots,\lambda_l')$. Denote $col(T)=\lambda_1'-1$. Then*

$$\sum_{T\in hook(2n-1)} sign(T)q^{col(T)} = \sum_{T\in hook(n)} q^{2\cdot col(T)}$$

$$\sum_{T\in hook(2n)} sign(T)q^{col(T)} = (1+q)\sum_{T\in hook(n)} q^{2\cdot col(T)}$$

Theorem 5.2.

$$\sum_{T\in tworows(2n+1)} sign(T)q^{row_2(T)} = \sum_{T\in tworows(n)} (-q^2)^{row_2(T)}$$

$$\sum_{T\in tworows(2n+2)} sign(T)q^{row_2(T)} = (1+q)\sum_{T\in tworows(n)} (-q^2)^{row_2(T)}$$

Remark 5.3. As a special case of Theorem 5.2, we see that for the Carlitz-Riordan q-Catalan numbers:

$$\sum_{n=1}^{\infty} q^n \cdot \tilde{C}_n(-1) = \sum_{n=1}^{\infty} q^{2n+1} \cdot \tilde{C}_n$$

These results resemble recent results of Adin and Roichman (see [**AR2**]) and Reifegerste (see [**R**]) regarding 321-avoiding permutations, which are brought hereby.

Definition 5.4. Let $T_n := \{\pi \in S_n \mid \nexists i < j < k \text{ such that } \pi(i) > \pi(j) > \pi(k)\}$ be the set of all 321-avoiding permutations. Define $ldes(\pi) := \max\{1 \le i \le n-1 \mid \pi(i) > \pi(i+1)\}$ and define $ldes(id) = 0$.

Theorem 5.5. [**AR2**]

$$\sum_{\pi\in T_{2n+1}} sign(\pi)\cdot q^{ldes(\pi)} = \sum_{\pi\in T_n} q^{2\cdot ldes(\pi)}$$

$$\sum_{\pi\in T_{2n}} sign(\pi)\cdot q^{ldes(\pi)} = (1-q)\sum_{\pi\in T_n} q^{2\cdot ldes(\pi)}$$

Definition 5.6. Define $lis(\pi)$ as the longest increasing subsequence in π.

Theorem 5.7. [**R**]

$$\sum_{\pi\in T_{2n+1}} sign(\pi)\cdot q^{lis(\pi)} = \sum_{\pi\in T_n} q^{2\cdot lis(\pi)+1}$$

$$\sum_{\pi\in T_{2n+2}} sign(\pi)\cdot q^{lis(\pi)} = (q-1)\sum_{\pi\in T_n} q^{2\cdot lis(\pi)+1}$$

Theorem 5.8. [**R**]

$$\sum_{\pi\in T^*_{2n+1}} sign(\pi)\cdot q^{lis(\pi)}t^{ldes(\pi)} = \sum_{\pi\in T_n} q^{2\cdot lis(\pi)+1}t^{2\cdot ldes(\pi)}$$

$$(1+q)\sum_{\pi\in T^*_{2n}} sign(\pi)q^{lis(\pi)}t^{ldes(\pi)} = \sum_{\pi\in T_n} q^{2\cdot lis(\pi)+1}t^{2\cdot ldes(\pi)}$$

A fuller understanding of such results would require additional research.

Acknowledgements

This work is part of an M.Sc. Thesis in mathematics, under supervision of Prof. Ron Adin and Prof. Yuval Roichman of Bar Ilan University, Israel.

I would also like to thank Nick Loehr for his help.

References

[AR1] R. M. Adin and Y. Roichman, Descent Functions and Random Young Tableaux. Combin., Prob., and Computing **10** (2001), 187-201.

[AR2] R. M. Adin and Y. Roichman, Equidistribution and Sign-Balance on 321-Avoiding Permutations, preprint, 2003, math.CO/0304429.

[BR] H. Barcelo and A. Ram, Combinatorial Representation Theory. New Perspectives in Algebraic Combinatorics, Cambridge Univ. Press (1999), 23-90.

[C] L. Carlitz, Sequences, paths, ballot numbers. Fibonacci Quart. **10** (1972), 531-549.

[CR] L. Carlitz and J. Riordan, two element lattice permutation numbers and their q-generalizations. Duke J. Math **31** (1964), 371-388.

[FH] J. Fürlinger and J. Hofbauer, q-Catalan Numbers. J. Combin. Theory Ser. A **40** (1985), 248-264.

[FS] D. Foata and M. P. Schüstenberger, Major index and inversion number of permutations, Math. Nachr. **86** (1978), 143-159.

[FULM] J. Fulman, The distribution of descents in fixed conjugacy classes of the symmetric groups. J. Combin. Theory Ser. A **84** (1998), 171-180.

[FULT] W. Fulton, Young Tableaux: With Applications To Representation Theory And Geometry. Cambridge Univ. Press (1996).

[GG] A. M. Garsia and I. Gessel, Permutation statistics and partitions. Adv. in Math. **31** (1979), 288-305.

[GR] Gaspar G. and Rahman, M., Basic Hypergeometric Series. Cambridge, England: Cambridge University Press, 1990.

[H] P. A. Hästö, On Descents in Standard Young Tableaux. Electron. J. Combin. **7** (2000) #59.

[L] N. Loehr, Combinatorial Proofs of Lattice Path Identities, currently in submission. http://math.ucsd.edu/~nloehr/paths.pdf.

[M] P. A. MacMahon, Combinatory Analysis, Chelsea, New York (1960). (originally published in 2 vols., Cambridge Univ. Press, Cambridge, UK, 1915-1916).

[R] A. Reifegerste, Refined Sign-Balance on 321-Avoiding Permutations, preprint, 2003, math.CO/0305327.

[SA] B. E. Sagan, The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions. Wadsworth and Brooks/Cole, CA (1991).

[SH] M. Shynar, Statistics in Standard Young Tableaux, M.Sc. Thesis, Bar Ilan University (2004).

[ST1] R. P. Stanley, Enumerative Combinatorics, Vol. 1, Wadsworth and Brooks/Cole, Monterey, CA (1986).

[ST2] R. P. Stanley, Enumerative Combinatorics, Vol. 2, Wadsworth and Brooks/Cole, Monterey, CA (1986).

[ST3] R. P. Stanley, Some Remarks on Sign-Balanced and Maj-Balanced Posets, preprint, 2003, math.CO/0211113.

Department of Mathematics, Bar Ilan University, Ramat Gan 52900, Israel
E-mail address: `shyarm@math.biu.ac.il`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Finite-Dimensional Crystals for Quantum Affine Algebras of type $D_n^{(1)}$

Philip Sternberg

Abstract. *The combinatorial structure of the crystal basis $B^{(2,2)}$ for the $U_q'(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,2)}$ is given, and a conjecture is presented for the combinatorial structure of the crystal basis $B^{(2,s)}$ for the $U_q'(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,s)}$.*

Résumé. *Nous donnons la structure combinatoire de la base cristalline $B^{(2,2)}$ pour le $U_q'(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,2)}$, et nous conjecturons la structure combinatoire de la base cristalline $B^{(2,s)}$ pour le $U_q'(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,s)}$.*

1. Introduction

While studying representations of quantum groups, Kashiwara developed the theory of crystal bases, which allow modules over quantum groups to be studied in terms of a crystal graph, a purely combinatorial object [**5**]. An open question in the area of crystal basis theory is to determine for which irreducible representations of quantum affine algebras a crystal basis exists, and when they exist, what combinorial structure the crystals have. It is conjectured [**3, 4**] that there is a family of irreducible finite-dimensional $U_q'(\mathfrak{g})$-modules $W^{(k,s)}$, called Kirillov-Reshetikhin modules, which have crystal bases $B^{(k,s)}$, where k is a Dynkin node and s is a positive integer; furthermore, it is expected that all irreducible finite-dimensional $U_q'(\mathfrak{g})$-modules which have crystal bases are tensor products of the modules $W^{(k,s)}$. A first step towards understanding these crystals is to determine their combinatorial structure.

For type $A_n^{(1)}$, the existence of the modules $W^{(k,s)}$ has been established [**8**], and the explicit combinatorial structure is also well-known [**14**]. For non-simply laced types, the following well-known algebra embeddings are conjectured to apply to crystals as well [**12**], which would yield the combinatorial structure of the corresponding crystals in terms of the crystal structure for the simply-laced types.:

$$\begin{aligned} C_n^{(1)}, A_{2n}^{(2)}, A_{2n}^{(2)\dagger}, D_{n+1}^{(2)} &\hookrightarrow A_{2n-1}^{(1)} \\ A_{2n-1}^{(2)}, B_n^{(1)} &\hookrightarrow D_{n+1}^{(1)} \\ E_6^{(2)}, F_4^{(1)} &\hookrightarrow E_6^{(1)} \\ D_4^{(3)}, G_2^{(1)} &\hookrightarrow D_4^{(1)}. \end{aligned}$$

Therefore, the next step in developing a general theory of affine crystals is to explore crystals of types $D_n^{(1)}$ ($n \geq 4$) and $E_n^{(1)}$ ($n = 6, 7, 8$). In this paper, we concentrate on type $D_n^{(1)}$. For irreducible representations corresponding to multiples of the first fundamental weight (indexed by a one-row Young diagram) or

2000 *Mathematics Subject Classification.* Primary 17B37; Secondary 81R10.
Key words and phrases. crystal bases, algebraic combinatorics.
The author was supported in part by NSF Grants DMS-0135345 and DMS-0200774.

any single fundamental weight (indexed by a one-column Young diagram) the crystals are known to exist and the structure is known [**8, 7**]. It is conjectured in [**3, 4**] that as $U_q(\mathfrak{g}_{I\setminus\{0\}})$-crystals, the crystal $B^{(k,s)}$ decomposes as

$$B^{(k,s)} = \bigoplus_{\Lambda} B(\Lambda),$$

where the direct sum is taken over all partitions which result from removing any number of 2×1 vertical dominoes from a $k \times s$ rectangle, whenever $k \leq n-2$. In the sequel, we consider the case $k = 2$, for which the above direct sum specializes to

$$B^{(2,s)} = \bigoplus_{i=0}^{s} B(i\Lambda_2).$$

First, we will use tableaux of shape (s, s) to define a $U_q(\mathfrak{so}_{2n})$-crystal whose vertices are in bijection with the classical tableaux from the above direct sum decomposition. Because of the way we define our tableaux, we will be able to combinatorially define the unique action of $\tilde{f}_0$ which makes this crystal into a connected perfect crystal of level s. Finally, we present a conjecture for an explicit construction of the representation $W^{(2,s)}$ which is compatible with the crystal basis $B^{(2,s)}$ as constructed. Full details of our results will be forthcoming [**13**].

2. Review of quantum groups and crystal bases

For $n \in \mathbb{Z}$ and a formal parameter q, we use the notations

$$[n]_q = \frac{q^n - q^{-n}}{q - q^{-1}}, \quad [n]_q! = \prod_{k=1}^{n} [k]_q, \text{ and } \begin{bmatrix} m \\ n \end{bmatrix}_q = \frac{[m]_q!}{[n]_q![m-n]_q!}.$$

These are all elements of $\mathbb{Q}(q)$, called the q-integers, q-factorials, and q-binomial coefficients, respectively.

Let $\mathfrak{g}$ be a Lie algebra with Cartan datum $(A, \Pi, \Pi^\vee, P, P^\vee)$, a Dynkin diagram indexed by I, and let $\{s_i | i \in I\}$ be the entries of the diagonal symmetrizing matrix of A. Let $q_i = q^{s_i}$ and $K_i = q^{s_i h_i}$. We may then construct the quantum enveloping algebra $U_q(\mathfrak{g})$ as the associative $\mathbb{Q}(q)$-algebra generated by e_i and f_i for $i \in I$, and q^h for $h \in P^\vee$, with the following relations:

(a) $q^0 = 1$, $q^h q^{h'} = q^{h+h'}$ for all $h, h' \in P^\vee$,
(b) $q^h e_i q^{-h} = q^{\alpha_i(h)} e_i$ for all $h \in P^\vee$,
(c) $q^h f_i q^{-h} = q^{\alpha_i(h)} f_i$ for all $h \in P^\vee$,
(d) $e_i f_j - f_j e_i = \delta_{ij} \frac{K_i - K_i^{-1}}{q_i - q_i^{-1}}$ for $i, j \in I$,
(e) $\sum_{k=0}^{1-a_{ij}} (-1)^k \begin{bmatrix} 1-a_{ij} \\ k \end{bmatrix}_{q_i} e_i^{1-a_{ij}-k} e_j e_i^k = 0$ for all $i \neq j$,
(f) $\sum_{k=0}^{1-a_{ij}} (-1)^k \begin{bmatrix} 1-a_{ij} \\ k \end{bmatrix}_{q_i} f_i^{1-a_{ij}-k} f_j f_i^k = 0$ for all $i \neq j$.

We can view $U_q(\mathfrak{g})$ as a q-deformation of $U(\mathfrak{g})$. Similarly, a $U_q(\mathfrak{g})$-module V may be seen as a q-deformation of a $U(\mathfrak{g})$-module. The representation theory of $U_q(\mathfrak{g})$ does not depend on q, provided $q \neq 0$ and $q^k \neq 1$ for all $k \in \mathbb{Z}$. Furthermore, through appropriate tensoring and factoring, we may "take the limit as q goes to zero" in $U_q(\mathfrak{g})$ and V. This process makes V very simple, so that we may study it using a colored directed graph whose vertices correspond to a canonical basis of V. In the solvable lattice models which provided the original motivation for quantum groups, q parameterized temperature, so letting q approach 0 in the quantum group corresponds to the temperature approaching absolute zero in the physical models. Thus, the graph described above is called a crystal graph, and its vertices are a crystal basis B for V [**5**]. The edges are colored by the index set I, which indicates the action of the Kashiwara operators $\tilde{e}_i$ and $\tilde{f}_i$ on B. The Kashiwara operators are a "crystal version" of the Chevalley generators of $\mathfrak{g}$.

We are particularly interested in a class of crystals called perfect crystals, since they allow us to construct infinite-dimensional highest weight modules over $U_q(\mathfrak{g})$, where $\mathfrak{g}$ is of affine type [**9**]. To define them, we need a few preliminary definitions.

Let P be the weight lattice of an affine Lie algebra $\mathfrak{g}$. Define $P_{cl} = P/\mathbb{Z}\delta$, $P_{cl}^+ = \{\lambda \in P_{cl} | \langle h_i, \lambda\rangle \geq 0 \text{ for all } i \in I\}$, and $U_q'(\mathfrak{g})$ to be the quantum enveloping algebra with the Cartan datum $(A, \Pi, \Pi^\vee, P_{cl}, P_{cl}^\vee)$.

A crystal pseudobase for a module V is a set B such that there is a crystal base B' for V such that $B = B' \cup -B'$.

Denote by c the canonical central element of $\mathfrak{g}$. In the sequel, we only consider $\mathfrak{g}$ of type $D_n^{(1)}$, in which case

$$c = \Lambda_0 + \Lambda_1 + 2\Lambda_2 + \cdots + 2\Lambda_{n-2} + \Lambda_{n-1} + \Lambda_n.$$

Define the set of level ℓ weights to be $(P_{cl}^+)_\ell = \{\lambda \in P_{cl}^+ | \langle c, \lambda\rangle = \ell\}$. For a crystal basis element $b \in B$, define $\varepsilon_i(b) = \max\{n \geq 0 | \tilde{e}_i^n(b) \in B\}$, and $\varepsilon(b) = \sum_{i\in I} \varepsilon_i(b)\Lambda_i$, and similarly, $\varphi_i(b) = \max\{n \geq 0 | \tilde{f}_i^n(b) \in B\}$, and $\varphi(b) = \sum_{i\in I} \varphi_i(b)\Lambda_i$. Finally, for a crystal basis B, we define B_{min} to be the set of crystal basis elements b such that $\langle c, \varepsilon(b)\rangle$ is minimal over $b \in B$.

A crystal B is a perfect crystal of level ℓ if:

(a) $B \otimes B$ is connected;
(b) there exists $\lambda \in P_{cl}$ such that $wt(B) \subset \lambda + \sum_{i\neq 0} \mathbb{Z}_{\leq 0}\alpha_i$ and $\#(B_\lambda) = 1$;
(c) there is a finite-dimensional irreducible $U_q'(\mathfrak{g})$-module V with a crystal pseudobase of which B is an associated crystal;
(d) for any $b \in B$, we have $\langle c, \varepsilon(b)\rangle \geq \ell$;
(e) the maps ε and φ from B_{min} to $(P_{cl}^+)_\ell$ are bijective.

We may now state the main result of this paper.

Theorem 2.1. *Suppose that the $U_q'(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,2)}$ has a crystal basis $B^{(2,2)}$ as conjectured in* [**3**]. *Then $B^{(2,2)} \cong \tilde{B}^{(2,2)}$, where $\tilde{B}^{(2,2)}$ is the affine crystal given explicitly by the construction below. Furthermore, we conjecture that the construction of $\tilde{B}^{(2,s)}$ below explicitly gives the crystal graph associated to the $U_q'(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,s)}$.*

Specifically, we will construct a $U_q'(\widehat{\mathfrak{so}}_{2n})$-crystal $\tilde{B}^{(2,s)}$ with the conjectured classical decomposition, and then show that it is the only perfect crystal which can admit such a decomposition. This is the first step in confirming Conjecture 2.1 of [**4**], which states that as modules over the embedded classical quantum group, $W^{(2,s)}$ decomposes as $\bigoplus_{i=0}^{s} V(i\Lambda_2)$, where $V(\Lambda)$ is the classical module with highest weight Λ, $W^{(2,s)}$ has a crystal basis, and this is a perfect crystal of level s.

3. Decomposition of $\tilde{B}^{(2,s)}$

Let $B(k\Lambda_2)$ denote the crystal basis of the irreducible representation of $U_q(\mathfrak{so}_{2n})$ with highest weight $k\Lambda_2$ for $k \in \mathbb{Z}_{\geq 0}$. We may associate with each crystal element a tableau of shape $\lambda = (k, k)$ on the partially ordered alphabet

$$1 < 2 < \cdots < n-1 < \begin{matrix} n \\ \bar{n} \end{matrix} < \overline{n-1} < \cdots \bar{2} < \bar{1}$$

such that [**2**, page 202]

(a) if ab is in the filling, then $a \leq b$;
(b) if $\begin{smallmatrix} a \\ b \end{smallmatrix}$ is in the filling, then $b \not\leq a$;
(c) no configuration of the form $\begin{smallmatrix} a & a \\ & \bar{a} \end{smallmatrix}$ or $\begin{smallmatrix} a & \\ \bar{a} & \bar{a} \end{smallmatrix}$ appears;
(d) no configuration of the form $\begin{smallmatrix} n-1 & n \\ n & n-1 \end{smallmatrix}$ or $\begin{smallmatrix} n-1 & \bar{n} \\ \bar{n} & n-1 \end{smallmatrix}$ appears;
(e) no configuration of the form $\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}$ appears.

Note that for $k \geq 2$, condition 5 follows from conditions 1 and 3.

Consider the set $\mathcal{T}$ of tableaux of shape (s,s) which violate only condition 3. We will construct a bijection between $\mathcal{T}$ and the vertices of $\bigoplus_{i=0}^{s-1} B(i\Lambda_2)$, so that $\mathcal{T} \cup B(s\Lambda_2)$ may be viewed as a $U_q(\mathfrak{so}_{2n})$-crystal with the conjectured classical decomposition of $B^{(2,s)}$. In section 4 we will define $\tilde{f}_0$ on $\mathcal{T} \cup B(s\Lambda_2)$ to give it the structure of a perfect $U_q'(\widehat{\mathfrak{so}}_{2n})$-crystal. This will be the crystal $\tilde{B}^{(2,s)}$ mentioned in Theorem 2.1. For proofs of all claims, see [**13**].

Let $T \in \mathcal{T}$, and define $\bar{\bar{i}} = i$ for $1 \le i \le n$. Then there is a unique $a \in \{1, \dots, n, \bar{n}\}$, $m \in \mathbb{Z}_{>0}$ such that T has exactly one configuration of one of the following forms:

$$\begin{array}{c} a \\ b_1 \end{array} \underbrace{\begin{array}{ccc} a & & a \\ & \cdots & \\ \bar{a} & & \bar{a} \end{array}}_{m} \begin{array}{c} c_1 \\ d_1 \end{array}, \qquad \text{where } b_1 \neq \bar{a}, \text{ and } c_1 \neq a \text{ or } d_1 \neq \bar{a};$$

$$\begin{array}{c} b_2 \\ c_2 \end{array} \underbrace{\begin{array}{ccc} a & & a \\ & \cdots & \\ \bar{a} & & \bar{a} \end{array}}_{m} \begin{array}{c} d_2 \\ \bar{a} \end{array}, \qquad \text{where } d_2 \neq a, \text{ and } b_2 \neq a \text{ or } c_2 \neq \bar{a};$$

$$\begin{array}{c} b_3 \\ c_3 \end{array} \underbrace{\begin{array}{ccc} a & & a \\ & \cdots & \\ \bar{a} & & \bar{a} \end{array}}_{m+1} \begin{array}{c} d_3 \\ e_3 \end{array}, \qquad \text{where } b_3 \neq a \text{ and } e_3 \neq \bar{a}.$$

To find the corresponding element of $\bigoplus_{i=0}^{s-1} B(i\Lambda_2)$, remove $\underbrace{\begin{array}{ccc} a & & a \\ & \cdots & \\ \bar{a} & & \bar{a} \end{array}}_{m}$ from T. The result will be a tableau in $B((s-m)\Lambda_2)$. Denote the image of T under this map by $D_2(T)$. We call D_2 the height-two drop map. For example, we have

$$T = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 3 \\ \hline \bar{4} & \bar{2} & \bar{2} & \bar{1} \\ \hline \end{array}, \quad D_2(T) = \begin{array}{|c|c|c|} \hline 1 & 3 & 3 \\ \hline \bar{4} & \bar{2} & \bar{1} \\ \hline \end{array}.$$

Let $t \in B(i\Lambda_2)$. The map F_2 (the height-two fill map) which inverts D_2 is given by finding a configuration $\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}$ in t such that either $c \le \bar{a} \le d$ or $a \le \bar{d} \le b$, and filling with $\underbrace{\begin{array}{ccc} a & & a \\ & \cdots & \\ \bar{a} & & \bar{a} \end{array}}_{s-i}$ or $\underbrace{\begin{array}{ccc} \bar{d} & & \bar{d} \\ & \cdots & \\ d & & d \end{array}}_{s-i}$, respectively. If more than one such configuration exists, or if both pairs of inequalities are satisfied, then $F_2(t)$ is independent of any of these choices. For example,

$$t = \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline \bar{4} & \bar{2} & \bar{1} \\ \hline \end{array}, \quad F_2(t) = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 2 & 3 \\ \hline \bar{4} & \bar{2} & \bar{2} & \bar{1} \\ \hline \end{array}.$$

While we could choose either column two or column three as the filling location, either choice results in the same tableau.

The action of the Kashiwara operators $\tilde{e}_i$, $\tilde{f}_i$ for $i \in \{1, \dots, n\}$ on $\mathcal{T} \cup B(s\Lambda_2)$ can be defined by direct combinatorial construction, but for the sake of simplicity, we describe them in terms of the above bijection. Let $T \in \mathcal{T} \cup B(s\Lambda_2)$. We define

$$\begin{aligned} \tilde{e}_i(T) &= F_2(\tilde{e}_i(D_2(T))) \\ \tilde{f}_i(T) &= F_2(\tilde{f}_i(D_2(T))), \end{aligned}$$

where the $\tilde{e}_i$ and $\tilde{f}_i$ on the right are the standard Kashiwara operators on $U_q(\mathfrak{so}_{2n})$-crystals [**10**].

4. Affine Kashiwara operators

We know that once $B^{(2,s)}$ is determined, there will be a map $\sigma : B^{(2,s)} \to B^{(2,s)}$ such that $\tilde{e}_0 = \sigma \tilde{e}_1 \sigma$ and $\tilde{f}_0 = \sigma \tilde{f}_1 \sigma$, corresponding to the automorphism of $U_q'(\widehat{\mathfrak{so}}_{2n})$ which interchanges nodes 0 and 1 of the Dynkin

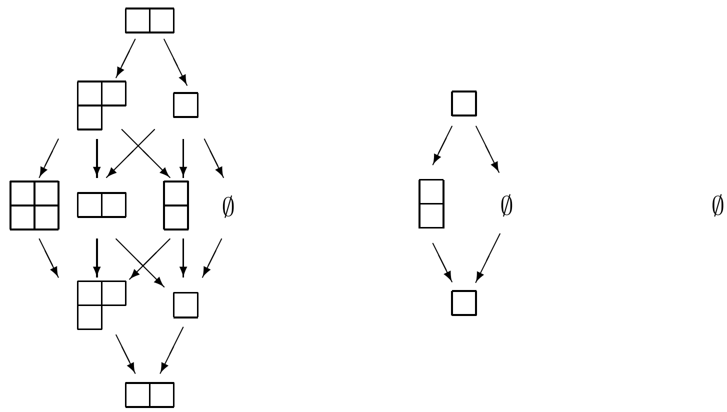


FIGURE 1. The branching component graph $\mathcal{BC}(\tilde{B}^{(2,2)})$

diagram. With this in mind, suppose we have defined $\tilde{f}_0$ on $\mathcal{T} \cup B(s\Lambda_2)$ to produce $\tilde{B}^{(2,s)}$, and consider the following operations on $\tilde{B}^{(2,s)}$; let $J \subset I$, and denote by B_J the graph which results from removing all j-colored edges from $\tilde{B}^{(2,s)}$ for $j \in J$. Then as directed graphs, we expect $B_{\{0\}}$ to be isomorphic to $B_{\{1\}}$, otherwise, $\tilde{B}^{(2,s)}$ and $B^{(2,s)}$ will not be isomorphic. We can gain some information about σ by considering $B_{\{0,1\}}$.

It is easy to see that the connected components of $B_{\{0,1\}}$ will be $U_q(\mathfrak{so}_{2n-2})$-crystals, indexed by partitions as described below. This decomposition produces a "branching component graph" for $\tilde{B}^{(2,s)}$, which we denote $\mathcal{BC}(\tilde{B}^{(2,s)})$. It suffices to describe the decomposition of the component of $\tilde{B}^{(2,s)}$ with classical highest weight $k\Lambda_2$ into $U_q(\mathfrak{so}_{2n-2})$-crystals. Denote this branching component subgraph by $\mathcal{BC}(k\Lambda_2)$. Each vertex v_λ of this graph will be labeled by a partition indicating the classical highest weight λ of the corresponding $U_q(\mathfrak{so}_{2n-2})$-crystal. Let $B(v_\lambda)$ denote the set of tableaux in $B(k\Lambda_2)$ contained in the $U_q(\mathfrak{so}_{2n-2})$-crystal indexed by v_λ. Then $\mathcal{BC}(k\Lambda_2)$ has a 1-colored edge from v_λ to v_μ if there is a tableau $b \in B(v_\lambda)$ such that $\tilde{f}_1(b) \in B(v_\mu)$.

We can give an explicit combinatorial description of $\mathcal{BC}(k\Lambda_2)$ as follows. The "highest weight" component of $\mathcal{BC}(k\Lambda_2)$ is a $1 \times k$ rectangle; call this vertex v_k. The function

$$r_k(v) = d(v, v_k) = \min_{P(v,v_k)} (\text{number of edges in } P(v, v_k))$$

is a rank function on $\mathcal{BC}(k\Lambda_2)$, where $P(v, v_k)$ is the set of all paths from v to v_k in $\mathcal{BC}(k\Lambda_2)$. For any partition λ, in each rank no more than one vertex may be indexed by λ. Let $v_\lambda \in \mathcal{BC}(k\Lambda_2)$ have rank less than k; then there is a 1-edge from v_λ to v_μ, where $r_k(v_\mu) = r_k(v_\lambda) + 1$ and there is an edge between λ and μ in the Young lattice. Also note that if $v_\lambda \in \mathcal{BC}(k\Lambda_2)$, then $\lambda \subset (k, k)$.

If $v_\lambda \in \mathcal{BC}(k\Lambda_2)$ has rank p, there is a vertex v'_λ, called the complementary vertex of v_λ, with rank $2k-p$. Let v_λ have a 1-edge to v_μ. Then there is also a 1-edge from v'_μ to v'_λ. Combined with the above description of the first $k+1$ ranks, this completely characterizes $\mathcal{BC}(k\Lambda_2)$.

Observe that $\mathcal{BC}(\tilde{B}^{(2,s)}) = \bigcup_{i=0}^{s} \mathcal{BC}(i\Lambda_2)$. Let $v_\lambda \in \mathcal{BC}(i\Lambda_2) \subset \mathcal{BC}(\tilde{B}^{(2,s)})$. Define $R(v_\lambda) = r_i(v_\lambda) + s - i$. This puts a rank on all of $\mathcal{BC}(B^{(2,s)})$. Note that $\mathcal{BC}(i\Lambda_2) \subset \mathcal{BC}((i+1)\Lambda_2)$, and this inclusion is compatable with R. Also note that if $R(v_\lambda) = p$, then v'_λ, the complementary vertex to v_λ, is now defined to be the vertex of rank $2s - p$ with the same shape and in the same component as v_λ.

To illustrate, $\mathcal{BC}(\tilde{B}^{(2,2)})$ is given in Figure 1, with rank 0 in the first line, rank 1 in the second, etc.

Since we know that $B_{\{0\}}$ and $B_{\{1\}}$ are isomorphic as directed graphs, it is clear that we can put 0-colored edges in the branching component graph in such a way that interchanging the 1-edges and the 0-edges

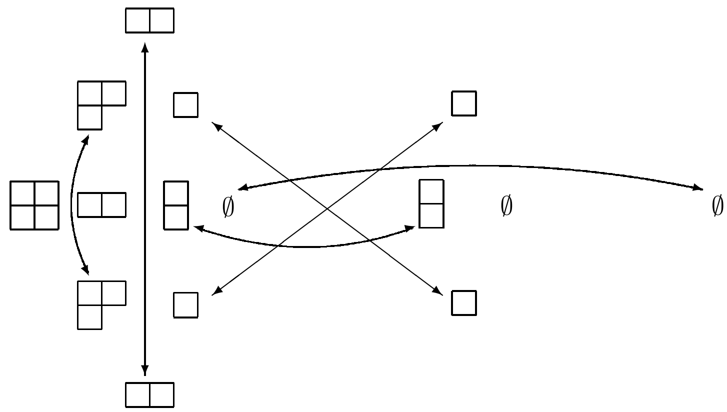


FIGURE 2. Definition of $\hat{\sigma}$ on $\mathcal{BC}(\tilde{B}^{(2,2)})$

and applying some shape-preserving bijection $\hat{\sigma}$ (defined below) to the vertices will produce an isomorphic colored directed graph. Such a bijection can be naturally extended to $\sigma : \tilde{B}^{(2,s)} \to \tilde{B}^{(2,s)}$ as follows. Let $b \in B(v_\lambda) \subset \tilde{B}^{(2,s)}$ for some supercrystal vertex v_λ, and let u_λ denote the $U_q(\mathfrak{so}_{2n-2})$ highest weight vector of $B(v_\lambda)$. Then for some finite sequence $i_1, \ldots, i_k$ of integers in $\{2, \ldots, n\}$, we know that $\tilde{f}_{i_1} \cdots \tilde{f}_{i_k} u_\lambda = b$. Let $v_\lambda^* = \hat{\sigma}(v_\lambda)$, and let u_λ^* be the highest weight vector of $B(v_\lambda^*)$. We may define $\sigma(b) = \tilde{f}_{i_1} \cdots \tilde{f}_{i_k} u_\lambda^*$. This involution of $\tilde{B}^{(2,s)}$ will satisfy $\tilde{f}_0 = \sigma \tilde{f}_1 \sigma$.

We will define $\hat{\sigma}(v_\lambda)$ for $R(v_\lambda) \leq s$, and observe that $\hat{\sigma}(v_\lambda') = \hat{\sigma}(v_\lambda)'$, where v' denotes the complementary vertex of v. Let $v_\lambda \in \mathcal{BC}(k\Lambda_2)$, and $R(v_\lambda) = p$. Then by the inclusion $\mathcal{BC}(i\Lambda_2) \subset \mathcal{BC}((i+1)\Lambda_2)$, there are $p+1$ vertices of the same shape as v_λ of rank p in $\mathcal{BC}(\tilde{B}^{(2,s)})$, one in each $\mathcal{BC}(j\Lambda_2)$ for $j = \{s-p, \ldots, s\}$. We define $\hat{\sigma}(v_\lambda)$ to be the vertex of the same shape as v_λ of rank $2s-p$ in $\mathcal{BC}((2s-p-k)\Lambda_2)$.

The action of $\hat{\sigma}$ on $\mathcal{BC}(\tilde{B}^{(2,2)})$ is given in Figure 2.

The observant reader will note that there are other permutations of the set of vertices of $\mathcal{BC}(\tilde{B}^{(2,s)})$ which respect the shape of the associated partitions. First, note that if a tableau T is in a vertex of rank p, we expect $\tilde{f}_0(T) = \sigma \tilde{f}_1 \sigma(T)$ to be in a vertex with rank $p-1$; otherwise there will be some T for which $\tilde{f}_0(T) = \tilde{f}_1(T)$, which must not be the case. Even this does not completely specify $\hat{\sigma}$, since (for instance) we might permute the three empty partitions in any manner and still satisfy all the above requirements. Note, however, that ε_0 depends entirely on the definition of σ, and the perfectness of a crystal depends on the function ε_0. (Recall the definitions from section 2.) For a detailed proof of the following theorem, see [**13**].

Theorem 4.1. *The above definition of σ, interpreted as a permutation of the vertices of $\bigoplus_{i=0}^{s} B(i\Lambda_2)$, is the only map such that defining $\tilde{f}_0 = \sigma \tilde{f}_1 \sigma$ produces a perfect crystal of level s for $s = 2$. We conjecture that this is true for all s.*

5. Perfectness of $\tilde{B}^{(2,s)}$, Part 1

We must show that $\tilde{B}^{(2,s)}$ satisfies conditions 1-5 from Section 2 with $\ell = s$. Condition 1 is verified by showing that each vertex of $\tilde{B}^{(2,s)} \otimes \tilde{B}^{(2,s)}$ is connected to $u_\emptyset \otimes u_\emptyset$, where $u_\emptyset \in \tilde{B}^{(2,s)}$ is the unique vector of the $U_q(\mathfrak{so}_{2n})$-crystal $B(0)$ [**13**]. Condition 2 is satisfied by $\lambda = s\Lambda_2 - 2s\Lambda_0$. We discuss a conjecture which satisfies Condition 3 in section 7. Conditions 4 and 5 can be dealt with simultaneously, and have been proved for $s = 2$ as described below. For proofs of all claims, see [**13**].

Given a weight $\lambda \in (P_{cl}^+)_s$, we can construct a tableau $T_\lambda \in \tilde{B}^{(2,s)}$ such that $\varepsilon(T_\lambda) = \varphi(T_\lambda) = \lambda$. First, observe the following. Let $T \in B(k\Lambda_2) \subset \tilde{B}^{(2,s)}$, and let $T^* = \iota_k^s(T)$, where $\iota_i^j : B(i\Lambda_2) \hookrightarrow B(j\Lambda_2)$ is the natural inclusion map which is compatible with the inclusion $\mathcal{BC}(i\Lambda_2) \hookrightarrow \mathcal{BC}(j\Lambda_2)$. Assume T to be such that $T^* \in \iota^s_{s-1}(B((s-1)\Lambda_2))$. Let $T_m = (\iota^s_m)^{-1}(T^*)$ for $m = s, s-1, \ldots, k$, where k is the smallest number such that $T_k \notin \iota^k_{k-1}(B((k-1)\Lambda_2))$. Then we have

$$\langle \varepsilon(T_s), \Lambda_0 + \Lambda_1 \rangle = \langle \varepsilon(T_{s-1}), \Lambda_0 + \Lambda_1 \rangle = \quad \cdots \quad = \langle \varepsilon(T), \Lambda_0 + \Lambda_1 \rangle \neq 0,$$

and for $i = 2, \ldots, n$,

$$\langle \varepsilon(T_s), \Lambda_i \rangle = \langle \varepsilon(T_{s-1}), \Lambda_i \rangle = \quad \cdots \quad = \langle \varepsilon(T_k), \Lambda_i \rangle.$$

This allows us to temporarily restrict our attention to those level s weights λ which satisfy $\langle \lambda, \Lambda_0 \rangle = \langle \lambda, \Lambda_1 \rangle = 0$; i.e., which can be expressed as $\lambda = \sum_{i=2}^n a_i \Lambda_i$. These weights correspond to tableaux $T_\lambda \in B_{min} \cap B(s\Lambda_2) \setminus \iota^s_{s-1}(B((s-1)\Lambda_2))$. We will later recursively construct the tableaux corresponding to all other level s weights.

First, let $\lambda = k\Lambda_{n-1} + (s-k)\Lambda_n$. If s is even and $k \geq s/2$, the corresponding tableau is

$$T_\lambda = \underbrace{\begin{array}{c} n-2 \\ n-1 \end{array} \cdots \begin{array}{c} n-2 \\ n-1 \end{array}}_{s-k} \underbrace{\begin{array}{c} n-1 \\ \bar{n} \end{array} \cdots \begin{array}{c} n-1 \\ \bar{n} \end{array}}_{k-s/2} \underbrace{\begin{array}{c} n \\ \overline{n-1} \end{array} \cdots \begin{array}{c} n \\ \overline{n-1} \end{array}}_{k-s/2} \underbrace{\begin{array}{c} \overline{n-2} \\ \overline{n-1} \end{array} \cdots \begin{array}{c} \overline{n-2} \\ \overline{n-1} \end{array}}_{s-k}$$

If s is odd and $k \geq s/2$, we have

$$T_\lambda = \underbrace{\begin{array}{c} n-2 \\ n-1 \end{array} \cdots \begin{array}{c} n-2 \\ n-1 \end{array}}_{s-k} \underbrace{\begin{array}{c} n-1 \\ \bar{n} \end{array} \cdots \begin{array}{c} n-1 \\ \bar{n} \end{array}}_{k-s/2} \begin{array}{c} n \\ \bar{n} \end{array} \underbrace{\begin{array}{c} n \\ \overline{n-1} \end{array} \cdots \begin{array}{c} n \\ \overline{n-1} \end{array}}_{k-s/2} \underbrace{\begin{array}{c} \overline{n-2} \\ \overline{n-1} \end{array} \cdots \begin{array}{c} \overline{n-2} \\ \overline{n-1} \end{array}}_{s-k}$$

In either case, if $k < s/2$, interchange n and $\bar{n}$ in T_λ.

Next, we describe how to construct T_λ recursively when $\lambda = \sum_{i=2}^n a_i \Lambda_i$ and $\langle \lambda, \Lambda_{n-1} + \Lambda_n \rangle < s$. Let j be the minimal index for which $\langle \lambda, \Lambda_j \rangle = k \neq 0$, let $\lambda' = \lambda - k\Lambda_j$, and let $T_{\lambda'}$ be the tableaux such that $\varepsilon(T_{\lambda'}) = \lambda'$. We then set

$$T_\lambda = \underbrace{\begin{array}{c} j-1 \\ j \end{array} \cdots \begin{array}{c} j-1 \\ j \end{array}}_{k} \boxed{T_{\lambda'}} \underbrace{\begin{array}{c} \bar{j} \\ \overline{j-1} \end{array} \cdots \begin{array}{c} \bar{j} \\ \overline{j-1} \end{array}}_{k},$$

which is simply the result of inserting $T_{\lambda'}$ between the two $2 \times k$ tableaux on either side.

We now consider level s weights λ such that $\langle \lambda, \Lambda_1 \rangle = k_1 \neq 0$ or $\langle \lambda, \Lambda_0 \rangle = k_0 \neq 0$ (or both). Let $\lambda' = \lambda - k_1\Lambda_1 - k_0\Lambda_0$, let $k_{\lambda'} = \langle c, \lambda' \rangle$, and once again, let $T_{\lambda'}$ be such that $\varepsilon(T_{\lambda'}) = \lambda'$. It follows that $T_{\lambda'}$ is a tableau of shape $(k_{\lambda'}, k_{\lambda'})$. If $k_1 \leq k_{\lambda'}$, then change $T_{\lambda'}$ into a skew tableau $S_{\lambda'}$ of shape $(k_{\lambda'} + k_1, k_{\lambda'})/(k_1)$ by Lecouvey D-equivalence [**11**], then fill the northwest boxes with 1's and the southeast boxes with $\bar{1}$'s to get a tableau of shape $(k_{\lambda'} + k_1, k_{\lambda'} + k_1)$. If $k_1 > k_{\lambda'}$, change $T_{\lambda'}$ into a skew tableau $S_{\lambda'}$ of shape $(2k_{\lambda'}, k_{\lambda'})/(k_{\lambda'})$ by Lecouvey D-equivalence, fill the northwest and southwest boxes as above, and insert a tableau of the form

$$\underbrace{\begin{array}{c} 1 \\ \bar{2} \end{array} \cdots \begin{array}{c} 1 \\ \bar{2} \end{array} \begin{array}{c} 2 \\ \bar{1} \end{array} \cdots \begin{array}{c} 2 \\ \bar{1} \end{array}}_{k_1 - k_{\lambda'}} \qquad \text{if } k_1 - k_{\lambda'} \text{ is even;}$$

$$\underbrace{\begin{array}{c} 1 \\ \bar{2} \end{array} \cdots \begin{array}{c} 1 \\ \bar{2} \end{array} \begin{array}{c} 2 \\ \bar{2} \end{array} \begin{array}{c} 2 \\ \bar{1} \end{array} \cdots \begin{array}{c} 2 \\ \bar{1} \end{array}}_{k_1 - k_{\lambda'}} \qquad \text{if } k_1 - k_{\lambda'} \text{ is odd;}$$

between the first $k_{\lambda'}$ columns and the last $k_{\lambda'}$ columns to get a tableau $T_{\lambda''}$ of shape $(k_{\lambda'} + k_1, k_{\lambda'} + k_1)$. Observe that $\varepsilon(T_{\lambda''}) = \lambda'' = \lambda - k_0\Lambda_0$.

Finally, use the filling map of section 3 to add k_0 columns to $T_{\lambda''}$, yielding T_λ with $\varepsilon(T_\lambda) = \lambda$.

6. Perfectness of $\tilde{B}^{(2,s)}$, Part II

We must now show that the tableaux constructed in section 5 are in bijection with $(P_{cl}^+)_s$. Once again, for proofs of the following Lemmas, see [**13**].

Lemma 6.1. *Let ι be the crystal endomorphism of $\tilde{B}^{(2,s)}$ defined by $\iota = \bigoplus_{i=0}^{s-1} \iota_i^{i+1}$, and let $T \in \tilde{B}^{(2,s)}$ be a tableau in the range of ι. Then $\varepsilon(\iota(T)) = \varepsilon(T) + \Lambda_1 - \Lambda_0$.*

This means that given a weight $\Lambda = k_0\Lambda_0 + k_1\Lambda_1 + \Lambda'$, where $\langle \Lambda', \Lambda_0 \rangle = \langle \Lambda', \Lambda_1 \rangle = 0$, it suffices to search for tableaux which correspond to the weight Λ'. Furthermore, such a tableau will appear in the "new" part of $B(s\Lambda_2)$, where s is the level of Λ'. We may thus restrict our attention to tableaux $T \in B(s\Lambda_2) \setminus \iota_{s-1}^s(B((s-1)\Lambda_2))$.

Lemma 6.2. *Let $v_\lambda \in \mathcal{BC}(\tilde{B}^{(2,s)})$ with complimentary vertex v'_λ. (Recall the definitions of the complimentary vertex of v_λ and $B(v_\lambda)$ from section 4.) If $B(v_\lambda)$ contains no minimal tableaux, then neither does $B(v'_\lambda)$.*

Therefore, we need only consider tableaux in the upper half (including the middle row) of the branching component graph.

Lemma 6.3. *Let $k \geq s/2$. If T has k or more 1's in the first row and no $\bar{1}$'s, then T is not minimal.*

This eliminates many tableaux. In particular, in $\tilde{B}^{(2,2)}$, we only need to check the middle vertices of the branching component graph with shape $(2,2)$ and (2). Exhaustion shows the conjectured tableaux to be the only tableaux of level 2 in those sets.

7. Construction of $W^{(2,s)}$

In [**9**], Kang et al. discuss the relationship between an arbitrary finite-dimen-sional $U'_q(\mathfrak{g})$-module M (where $\mathfrak{g}$ is of affine type) and $\mathrm{Aff}(M)$, the infinite-dimensional $U_q(\mathfrak{g})$-module constructed by "affinizing" M. Loosely speaking, $\mathrm{Aff}(M) \simeq \bigoplus_{n\in\mathbb{Z}} T^n M$, where e_0 and f_0 respectively raise and lower the degree of T in addition to their ordinary action on M. To make the weight spaces of $\mathrm{Aff}(M)$ finite-dimensional, we add $n\delta$ to the weight of a vector in T^nM, where δ is the null root of $\mathfrak{g}$. Kang et al. also construct $\mathrm{Aff}(B)$ for any $U'_q(\mathfrak{g})$-crystal B, and state that if (L, B) is a crystal base of M, then $(\mathrm{Aff}(L), \mathrm{Aff}(B))$ is a crystal base of $\mathrm{Aff}(M)$.

The inverse of this process for level zero extremal weight modules generated by a basic weight vector is given in [**6**] as follows: given a fundamental infinite-dimensional $U_q(\mathfrak{g})$-module $V(\varpi_i)$, there is a $U'_q(\mathfrak{g})$-linear automorphism z_i of $V(\varpi_i)$ of weight $d_i\delta$, where d_i is an integer constant determined by the root system of $\mathfrak{g}$. The finite-dimensional $U'_q(\mathfrak{g})$-module $W(\varpi_i)$ is given by $W(\varpi_i) = V(\varpi_i)/(z_i - 1)V(\varpi_i)$, and $V(\varpi_i)$ can be naturally embedded in $\mathrm{Aff}(W(\varpi_i))$.

Later, Kashiwara also conjecturally gives an embedding for $V(\lambda) \subset \bigotimes V(\varpi_i)^{\otimes m_i}$, where $\lambda = \sum m_i\varpi_i$ is a level zero extremal weight. This conjecture is verified in [**1**] for symmetric untwisted affine Lie algebras, using Schur functions in the operators $z_{i,\nu}$, which correspond to z_i as above acting on the i,ν-th component of the tensor product.

We conjecture that the $U'_q(\widehat{\mathfrak{so}}_{2n})$-module $W^{(2,s)} = W(s\varpi_2)$ can be constructed as the quotient $V(s\varpi_2)/(z_{2s}-1)V(s\varpi_2)$, where z_{2s} is the $U'_q(\widehat{\mathfrak{so}}_{2n})$-linear automorphism of $V(s\varpi_2)$ of weight $2s\delta$. Such a construction would be compatible with $B^{(2,s)}$ as constructed here, and would give an embedding of $V(s\varpi_2)$ in $\mathrm{Aff}(W(s\varpi_2))$ similar to the embedding in [**6**] for fundamental representations.

Acknowledgements. The author would like to thank Anne Schilling for suggesting this problem and for many helpful conversations and suggestions; the author is also grateful to the referees for their helpful comments.

References

[1] J. Beck, *Crystal structure of level zero extremal weight modules*, Lett. Math. Phys., **61** (2002), no. 3, 221–229.

[2] J. Hong, S.-J. Kang, *Introduction to quantum groups and crystal bases*, Graduate Studies in Mathematics, 42. American Mathematical Society, Providence, RI, 2002. xviii+307 pp. ISBN: 0-8218-2874-6.

[3] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, Y. Yamada, *Remarks on fermionic formula*, Recent developments in quantum affine algebras and related topics (Raleigh, NC, 1998), 243–291, Contemp. Math., **248**, Amer. Math. Soc., Providence, RI, 1999.
[4] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, Z. Tsuboi, *Paths, crystals and fermionic formulae*, MathPhys odyssey, 2001, 205–272, Prog. Math. Phys., 23, Birkhäuser Boston, Boston, MA, 2002.
[5] M. Kashiwara, *Crystalizing the q-analogue of universal enveloping algebras*, Comm. Math. Phys. **133** (1990), no. 2, 249–260.
[6] M. Kashiwara, *On level-zero representation of quantized affine algebras*, Duke Math. J. **112** (2002) no. 1, 117–195.
[7] Y. Koga, *Level one perfect crystals for* $BSp_{2n}(\mathbb{F})(1)_n, CSp_{2n}(\mathbb{F})(1)_n$, *and* $DSp_{2n}(\mathbb{F})(1)_n$, J. Algebra **217** (1999), no. 1, 312–334.
[8] S.-J. Kang, M. Kashiwara, K.C. Misra, T. Miwa, T. Nakashima, A. Nakayashiki, *Perfect crystals of quantum affine Lie algebras*, Duke Math. J. **68** (1992), no. 3, 499–607.
[9] S.-J. Kang, M. Kashiwara, K.C. Misra, T. Miwa, T. Nakashima, A. Nakayashiki, *Affine crystals and vertex models*, Infinite analysis, Part A, B (Kyoto, 1991), 449–484, Adv. Ser. Math. Phys., **16**, World Sci. Publishing, River Edge, NJ, 1992.
[10] M. Kashiwara, T. Nakashima, *Crystal graphs for representations of the q-analogue of classical Lie algebras*, J. Algebra **165** (1994), no. 2, 295–345.
[11] C. Lecouvey, *Schensted-type correspondences and plactic monoid for types* B_n *and* D_n, arXiv:math.CO/0211444.
[12] M. Okado, A. Schilling, M. Shimozono, *Virtual crystals and Kleber's algorithm* Commun. Math. Phys. **238** (2003) 187–209.
[13] A. Schilling, P. Sternberg, *Finite-Dimensional Crystals for Quantum Affine Algebras of type* $D_n^{(1)}$, in preparation.
[14] M. Shimozono, *Affine type A crystal structure on tensor products of rectangles, Demazure characters, and nilpotent varieties*, J. Algebraic Combin. **15** (2002), no. 2, 151–187.

Department of Mathematics, University of California, One Shields Ave, Davis, CA, 95616-8633, USA
E-mail address: `sternberg@math.ucdavis.edu`
URL: `http://www.math.ucdavis.edu/~sternberg/`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Rook Numbers and the Normal Ordering Problem

Anna Varvak

Abstract. *For an element w in the Weyl algebra generated by D and U with relation $DU = UD+1$, the normally ordered form is $w = \sum c_{i,j} U^i D^j$. We demonstrate that the normal order coefficients $c_{i,j}$ of a word w are rook numbers on a Ferrers board. We use this interpretation to give a new proof of the rook factorization theorem, which we use to provide an explicit formula for the coefficients $c_{i,j}$. We calculate the Weyl binomial coefficients: normal order coefficients of the element $(D+U)^n$ in the Weyl algebra. We extend some of these results to the q-analogue of the Weyl algebra.*

Résumé. *Pour un élément dans l'algèbre du Weyl généré par D et U avec de relation $DU = UD+1$, la forme d'ordre normal est $w = \sum c_{i,j} U^i D^j$. Nous démontrons que les coefficients d'ordre normal $c_{i,j}$ sont des nombres de tours sur d'amier de Ferrers. Nous employons cette interprétation pour fournir une nouvelle preuve de théorème de factorisation de tours, à la laquelle nous mène une formule explicite pour les coefficients $c_{i,j}$. Nous calculons les coefficients binomiaux du Weyl: les coefficients d'ordre normal d'élément $(D+U)^n$ de l'algèbre du Weyl. Nous prolongeons quelque de ces résultats à q-analogue d'algèbre du Weyl.*

1. Introduction

For an element w in the Weyl algebra generated by D and U with relation $DU = UD + 1$, the normally ordered form is $w = \sum c_{i,j} U^i D^j$. For example, in the algebra of differential operators where $D = \frac{d}{dx}$ and U acts as multiplication by x, the operator w, applied to the polynomial $f(x)$, is expressed in the normally ordered form as

$$w\big(f(x)\big) = \sum c_{i,j}\ x^i\ \frac{d^j f}{dx^j}(x)\ .$$

The problem of finding explicit formulae for the normal order coefficients $c_{i,j}$ appears more frequently in the context where the Weyl algebra is the algebra of boson operators [**BPS, Ka, Ma, Mi1, Mi2, Sc**], generated by the creation and annihilation operators typically denoted as $a^\dagger$ and a. A boson is a type of particle like the light particle, the photon. According to the theory of quantum mechanics, the possible amount of energy that a particle can have is not continuous but quantized, so there is the smallest amount of energy—the zeroth state, frequently referred to as the ground state; there is the first state, which is the next smallest amount of energy allowed, and so on. The boson operators change the energy state of the particle like the differential operators change the power of x. While the machanics of it are fascinating, for our purposes the assurance of the commutation relation $aa^\dagger - a^\dagger a = 1$ is sufficient.

As a hint of combinatorial interest in the problem of normal ordering, it has long been known that the normal order coefficients of $(UD)^n$ are the Stirling numbers $S(n,k)$ of the second kind. These numbers can

1991 *Mathematics Subject Classification.* Primary 05A10, 05A30; Secondary 12H05.
Key words and phrases. rook numbers, normal ordering problem, differential operators, binomial coefficients.

be defined algebraically by the formula

$$x^n = \sum_{k=0}^{n} S(n,k)\, x(x-1)\cdots(x-k+1),$$

and they also have a combinatorial interpretation of counting the number of ways to partition a set of n elements into k subsets.

Mikhailov and Katriel [**Ka, Mi1, Mi2**] have extended the definition of the Stirling numbers, finding explicit formulas for the normal ordered form of operators such as $(D+U^r)^n$, and $(UD+U^r)^n$. Recently, Blasiak, Penson, and Solomon [**BPS**] have generalized the Stirling numbers even further to address the normal ordering problem for operators of the form $(U^rD^s)^n$.

The interpretation of the normal order coefficients of a word as rook numbers on a Ferrers board was given by Navon [**Na**] in 1973, but it requires the power of the rook factorization theorem, presented two years later by Goldman, Joichi, and White [**GJW**] to give an explicit formula. Interestingly enough, the interpretation provides a proof of the rook factorization theorem.

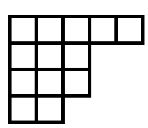
Here is an outline of the article. In section 2, we cover the basic definitions concerning Ferrers boards and rook numbers. We demonstrate that the normal order coefficents of a word are rook numbers on a Ferrers board in section 3, and give an explicit formula for them in section 4, together with a new proof of the rook factorization theorem. We discuss the Weyl binomial coefficients in section 5. Finally, we extend the interpretation of the normal order coefficients to the q-analogue in section 6.

2. Definitions concerning rook numbers and Ferrers boards

For n a positive integer, we denote by $[n]$ the set $\{1,2,\ldots,n\}$. A *board* is a subset of $[n]\times[m]$, where n and m are positive integers. Intuitively, we think of a board as an array of squares arranged in rows and columns. An element $(i,j)\in B$ is then represented by a square in the ith column and jth row. It will be convenient to consider columns numbered from left to right, and rows numbered from top to bottom, so that the square $(1,1)$ appears in the top left corner.

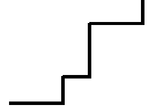
A board B is a *Ferrers* board if there is a non-increasing sequence of positive integers $h(B)=(h_1,h_2,\ldots,h_n)$ such that $B=\{(i,j)\mid i\le n \text{ and } j\le h_i\}$. Intuitively, a Ferrers board is a board made up of adjacent solid columns with a common upper edge, such that the heights of the columns from left to right form a non-increasing sequence.

Example 2.1. A Ferrers board with height sequence $(4,4,3,1,1)$ can be visually represented as



The connection between Ferrers boards and words composed of two letters is as follows. We represent the letter D as a step to the right, and the letter U as a step up. The resulting path outlines a Ferrers board.

Example 2.2. The word $w=DDUDUUDDU$ outlines the Ferrers board in example 2.1. This is easy to see from the path representing w:



Note that a word $w'=U^iwD^j$ outlines the same Ferrers board for any nonnegative integers i and j. If the Ferrers board B is contained in a rectangle with n columns and m rows, then there is a unique word with n D's and m U's that outlines B.

We denote the Ferrers board outlined by the word w by B_w.

For a board B, let $r_k(B)$ denote the number of ways of marking k squares of the board B, no two in the same row or column. In chess terminology, we are placing k rook pieces on the board B in non-attacking positions. The number $r_k(B)$ is called the kth *rook number* of B.

3. Normal order coefficients of a word

Recall that the Weyl algebra is the algebra generated by D and U, with the commutation relation $DU = UD + 1$.

Definition 3.1. For w an element in the Weyl algebra, the *normally ordered form* of w is the sum

$$w = \sum_{i,j} c_{i,j} U^i D^j,$$

where in each term the D's appear to the right of the U's. The numbers $c_{i,j}$ are the *normal order coefficients* of w.

We call w a *word* if w has a representation $w = w_1 w_2 \dots w_n$, where $w_i \in \{D, U\}$. We demonstrate that the normal order coefficients of a word w are rook numbers on the Ferrers board outlined by w. This combinatorial interpretation was originally given by Navon [**Na**].

Theorem 3.2. Normally Ordered Word

Let the element w in the Weyl algebra be a word composed of n D's and m U's. Then

$$w = \sum_{k=0}^{n} r_k(B_w) U^{m-k} D^{n-k}.$$

Proof. It is easy to see that the terms in the normally ordered form of w are $U^{m-k}D^{n-k}$, where $k = 0, 1, \dots, \min(m, n)$. Every time we replace DU with $UD + 1$ and expand the result, one of the new terms retains the same number of D's and U's as before, and the other term has one fewer of each.

By definition, the normal order coefficient $c_{m-k,n-k}$ is the number of terms $U^{m-k}D^{n-k}$ in the normally ordered form of w, which is obtained by successively replacing DU with $UD + 1$ and expanding. For the sake of consistency, we always choose to replace the rightmost DU.

We can regard the terms as words. Then the normal order coefficient $c_{m-k,n-k}$ is the number of ways to get the word $U^{m-k}D^{n-k}$ from the word w, by successively replacing the rightmost DU with either UD or 1 (that is, deleting it), choosing to do the latter k times.

We now can consider the substitutions in terms of the outlined Ferrers boards. The rightmost DU outlines the rightmost inner corner square of the board. Replacing DU with UD amounts to deleting that square, whereas deleting the DU amounts to deleting the square together with its row and column. Therefore the normal order coefficient $c_{m-k,n-k}$ is the number of ways to reduce the Ferrers board B_w outlined by w to the trivial board by successively deleting the rightmost inner corner square either alone, or together with its row and column, choosing to do the latter k times.

The k squares that are deleted together with their rows and columns cannot share either a row or a column. So the normal order coefficient $c_{m-k,n-k}$ is the number of ways to mark k squares on the Ferrers board B_w outlined by the word w, no two in the same row or column. This is exactly the kth rook number $r_k(B_w)$.

Finally, since $r_k(B_w) = 0$ for $k > \min(m, n)$, we let the sum range from 1 to n. □

Remark 3.3. As mentioned in the introduction, the Stirling numbers $S(n, k)$ of the second kind are the normal order coefficients of the word $(UD)^n$. Mikhailov [**Mi1**] defined, in a purely algebraic way, a more generalized version of the Stirling numbers to find the normal ordered form of operators of the form $(U^r+D)^n$. In a recent paper unrelated to the normal ordering problem, Lang [**La**] studied a similar generalization of the Stirling numbers, finding combinatorial interpretations for certain particular cases. Recently Blasiak, Penson, and Solomon [**BPS**] introduced the generalized Stirling numbers of the second kind, denoted $S_{r,s}(n,k)$ for $r \geq s \geq 0$, defined by the relation

$$(U^r\, D^s)^n = U^{n(r-s)} \sum_{k=s}^{ns} S_{r,s}(n,k) U^k\, D^k.$$

The standard Stirling numbers of the second kind are $S_{1,1}(n,k)$, and the generalized Stirling numbers of Mikhailov are $S_{r,1}(n,k)$.

We define the *staircase board* $J_{r,s,n}$ to be the Ferrers board outlined by the word $(U^r D^s)^n$.

Corollary 3.4.

$$S_{r,s}(n,k) = r_{ns-k}\left(J_{r,s,n}\right).$$

Remark 3.5. We can easily adapt the proof of Theorem 3.2 to work in the case where the algebra generated by D and U has the commutation relation $DU = UD + c$. For a word w, we get the normal ordering form of w by successively replacing DU by $UD + c$, and expanding. Just as before, we consider w as a word in the letters D, U, and the substitution as a choice of either replacing the rightmost DU by UD or deleting it, but the choice of deleting is weighted by c. In terms of the associated Ferrers board, we weight each placement of a rook by c. So we get

$$w = \sum_{k=0}^{n} c^k r_k(B_w) U^{m-k} D^{n-k}.$$

We should note that this algebra is isomorphic to the Weyl algebra, because if $DU - UD = 1$, then $D(cU) - (cU)D = c$.

We can also assign a weight to the choice of replacing DU by UD, and thus extend the result to algebra with the relation $DU = qUD + 1$. The algebra with this relation is know as the q-Weyl algebra, and is of interest both to combinatorialists and to physicists. To the latter, because such algebras are models for q-degenerate bosonic operators [**Sc**]. To the former, because it involves the q-analogue of rook numbers [**GR**]. We discuss this case in detail in section 6.

First, we show how Theorem 3.2 allows for a new proof of the Rook Factorization Theorem [**GJW**], which in turn leads to an explicit formula for computing the normal order coefficients of a word.

4. Computing the normal order coefficients

For a general board B, rook numbers can be computed recursively [**Ri**]. Choose a square of B, and let B_1 be the board obtained from B by deleting that square, and let B_2 be the board obtained from B by deleting the square together with its row and column. Then $r_k(B) = r_k(B_1) + r_{k-1}(B_2)$, reflecting the fact we may or may not mark the square in question.

There are better methods for calculating rook numbers on Ferrers boards, owing to the fact that the generating function of rook numbers on a Ferrers board, expressed in terms of falling factorials, completely factors.

We define the kth *falling factorial* of x by

$$x^{\underline{k}} = x(x-1)\cdots(x-k+1).$$

Goldman et al. [**GJW**] show that the factorial rook polynomial $r(B,x) := \sum_{k=0}^{n} r_k(B)\, x^{\underline{n-k}}$ of a Ferrers board is a product of linear factors.

Theorem 4.1. Rook Factorization Theorem

For a Ferrers board B with column heights $h(B) = (h_1, \ldots, h_n)$,

$$\sum_{k=0}^{n} r_k(B) x^{\underline{n-k}} = \prod_{i=1}^{n} (x + h_i - n + i)$$

We provide a new proof the Rook Factorization Theorem, using Theorem 3.2.

Proof. Let w be the word with n D's and h_1 U's that outlines the Ferrers board B. By Theorem 3.2,

$$w = \sum_{k=0}^{n} r_k(B_w) U^{h_1-k} D^{n-k}.$$

as an element in the Weyl algebra.

We consider a particular manifestation of the Weyl algebra as the algebra of operators generated by $D = \frac{d}{dt}$ and $U=$ multiplication by t, acting on functions in the variable t. So

$$w = \sum_{k=0}^{n} r_k(B_w) t^{h_1-k} \left(\frac{d}{dt}\right)^{n-k}.$$

We apply both sides of the equation to t^x, where x is a real number.

Since $\left(\frac{d}{dt}\right)^{n-k}(t^x) = x(x-1)\cdots(x-(n-k)+1)t^{x-n+k} = x^{\underline{n-k}}\, t^{x-n+k}$, the right hand side is

$$\sum_{k=0}^{n} r_k(B)\, x^{\underline{n-k}}\, t^{x-n+h_1}.$$

On the left-hand side we get the product of the following factors. The jth application of D gives the factor of $x + a_U - a_D$, where a_U the number of times U was previously applied, and a_D the number of times D was previously applied. There are $j-1$ D's to the right of the jth D, so $a_D = j-1$. The jth D from the right is the $(n-j+1)$st D from the left, so $a_U = h_{n-j+1}$. Therefore the left-hand side is

$$t^{x-n+h_1} \prod_{j=1}^{n} (x + h_{n-j+1} - j + 1).$$

If we let $i = n-j+1$, then the left-hand side is

$$t^{x-n+h_1} \prod_{i=1}^{n} (x + h_i - n + i).$$

Now we set $t = 1$ to get the desired result. □

Example 4.2. For $w = DDUUUDDUD$, by Theorem 3.2,

$$\frac{d}{dt}\frac{d}{dt}\, t\cdot t\cdot t\cdot \frac{d}{dt}\frac{d}{dt}\, t\cdot \frac{d}{dt}\,(t^x) = \sum_{k=0}^{5} r_k(B_w)\, t^{4-k} \left(\frac{d}{dt}\right)^{5-k} (t^x)$$

$$x\cdot(x+1)\cdot(x-1)\cdot x\cdot x\cdot\left(t^{x-1}\right) = \sum_{k=0}^{5} r_k(B_w)\, x^{\underline{5-k}}\cdot\left(t^{x-1}\right)$$

$$x\cdot(x+1)\cdot(x-1)\cdot x\cdot x = \sum_{k=0}^{5} r_k(B_w)\, x^{\underline{5-k}}\,.$$

The left hand side is the complete factorization of the factorial rook polynomial of B_w.

The falling factorials $1, x, x(x-1), \dots$ form a basis of polynomials in x. If $P(x) = \sum_{k=0}^{n} p_k\, x^{\underline{k}}$, it is well known that the coefficients are $p_k = \frac{1}{k!}\Delta^k P(x)\Big|_{x=0}$, where Δ is the difference operator defined by $\Delta P(x) = P(x+1) - P(x)$. Explicitly [**St**], the coefficients are

$$p_k = \frac{1}{k!}\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i} P(i)\,.$$

Corollary 4.3. Computing the Normal Order Coefficients of a Word

For w a word in the Weyl algebra composed of n D's and m U's, let $P(x) = \prod_{i=1}^{n}(x + h_i - n + i)$, where $h_1, h_2, \ldots, h_n$ are the column heights of the Ferrers board outlined by w. Then

$$w = \sum_{k=0}^{n} r_k U^{m-k} D^{n-k},$$

where

$$r_k = \frac{1}{(n-k)!} \sum_{i=0}^{n-k} (-1)^{n-k-i} \binom{n-k}{i} P(i) \,.$$

5. Weyl binomial coefficients

The Weyl binomial coefficient $\binom{n}{m}_k$ is the coefficient of the term $U^{n-m-k}D^{m-k}$ in $(D+U)^n$, where the commutation relation is $DU = UD + 1$. The binomial product $(D+U)^n$ is the sum of all words in letters D and U of length n, and the normally ordered term $U^{n-m-k}D^{m-k}$ comes from all words with $n-m$ U's and m D's, where k paris of D and U are deleted during the normal ordering. Each of these words outlines a unique Ferrers board with at most m columns of height at most $(n-m)$. Therefore the Weyl binomial coefficient $\binom{n}{m}_k$ can be expressed as the sum of the kth rook numbers over all Ferrers boards contained in the m-by-$(n-m)$ rectangle.

The classical binomial coefficient $\binom{n}{m}$, and its q-analogue $\left[{n \atop m}\right]$, can be similarly expressed in terms of Ferrers boards. The binomial coefficient $\binom{n}{m}$ is the coefficient of the term $U^{n-m}D^m$ in $(D+U)^n$, where the commutation relation is $DU = UD$. The term $U^{n-m}D^m$ is the normally ordered form of any word w with $n-m$ U's and m D's. Since all the letters commute, there is only one way to normally order a word, so the binomial coefficient $\binom{n}{m}$ counts the words with $n-m$ U's and m D's. Since there is a bijection between the set of such words and the set of Ferrers boards contained in m-by-$(n-m)$ rectange, $\binom{n}{m}$ counts such Ferrers boards. Therefore

$$\binom{n}{m} = \sum_{B \subseteq [m]\times[n-m]} 1.$$

Similarly, the q-binomial coefficient $\left[{n \atop m}\right]$ is the coefficient of the term $U^{n-m}D^m$ in $(D+U)^n$, where the commutation relation is $DU = qUD$. Again, the term $U^{n-m}D^m$ comes from any word w with $n-m$ U's and m D's, which outlines the Ferrers board B_w contained in the rectangle $[m]\times[n-m]$. The relation $DU = qUD$ specifies that each square of B_w has weight q, so $w = q^{|B_w|}U^{n-m}D^m$. In other words,

$$\begin{bmatrix} n \\ m \end{bmatrix} = \sum_{B \subseteq [m]\times[n-m]} q^{|B|}.$$

Theorem 5.1. *Let $k \leq m$ be an integer. Then*

$$\binom{n}{m}_k = \sum_{B \subseteq [m]\times[n-m]} r_k(B) = \frac{n!}{2^k\, k!\,(m-k)!\,(n-m-k)!}$$

Proof. Any Ferrers board B in $[m]\times[n-m]$ is B_w for a particular word w with $n-m$ U's and m D's. From the proof of Theorem 3.2, the number $r_k(B_w)$ is the number of ways to get the word $U^{n-m-k}D^{m-k}$ from the word w, by successively either replacing DU with UD or deleting it, choosing to do the latter k times. Equivalently, $r_k(B_w)$ is the number of ways to mark k pairs of the letters D and U in the word w, such that in each pair the D appears to the left of the U. The marked pairs are deleted, and the rest of the letters commute into the normally ordered form.

We therefore count the number of ways to construct words with $n-m$ U's and m D's, with k marked pairs of the letters D and U, the former to the left of the latter. We begin with n spaces for the letters in w. Choose $n-m-k$ of these to be U, and $m-k$ to be D. There are

$$\binom{n}{n-m-k,\ m-k} = \frac{n!}{(n-m-k)!\,(m-k)!\,(2k)!}$$

ways to do so. For the $2k$ remaining spaces, we pair them, forming k pairs. There are $(2k-1)\cdot(2k-3)\cdots 5\cdot 3\cdot 1$ ways to do this. For each pair, let the space on the left be D, and the space on the right be U.

By this construction, the sum of the kth rook numbers over all Ferrers boards contained in the m-by-$(n-m)$ rectangle is

$$\frac{(2k-1)\cdot(2k-3)\cdots 5\cdot 3\cdot 1}{(2k)!}\frac{n!}{(n-m-k)!\,(m-k)!} = \frac{n!}{2^k\,k!\,(m-k)!\,(n-m-k)!}.$$

□

Since $(D+U)^n$ is the sum of all words composed of letters D and U of length n, we have a formula for normal ordering of $(D+U)^n$, as shown by Mikhailov in [**Mi1**].

Corollary 5.2.

$$(D+U)^n = \sum_{m=0}^{n} \sum_{k=0}^{\min(m,n-m)} \frac{n!}{2^k\,k!\,(m-k)!\,(n-m-k)!} U^{m-k} D^{n-m-k}$$

Remark 5.3. The Weyl binomial coefficients obey the recursive formula

$$\binom{n}{m}_k = \binom{n-1}{m}_k + \binom{n-1}{m-1}_k + m\binom{n-2}{m-1}_{k-1},$$

with boundary conditions $\binom{1}{0}_0 = \binom{1}{1}_0 = 1$, $\binom{n}{m}_{k-1} = 0$. Consider the pairs (B, C), where B is a Ferrers board in $[m]\times[n-m]$ and C is a placement of k rooks on B. The Weyl binomial coefficient counts the number of such pairs. The set of such pairs is a disjoint union of three sets: one where the height h_1 of the first column B is strictly less than $n-m$, one where $h_1 = m$ and C doesn't place a rook in the first column, and one where $h_1 = m$ and C places a rook in the first column. The recursive formula follows.

The first two terms in the recursive formula are the same as for the classical binomial coefficients. In fact, the Weyl binomial coefficients can be expressed in terms of classical coefficients as follows.

Corollary 5.4. *Let* $C(y) = \sum_{k\geq 0} \binom{n}{k}\frac{y^k}{k!}$ *be the exponential generating function of the binomial coefficients. Then the ordinary generating function of the Weyl binomial coefficients is*

$$\sum_{k\geq 0}\binom{n}{m}_k x^k = \left(\frac{d}{dy}\right)^{n-m} C(y)\,\Bigg|_{y=\frac{x}{2}}.$$

Proof.

$$\begin{aligned} C(y) &= \sum_{k\geq 0}\binom{n}{k}\frac{y^k}{k!} \\ &= \sum_{k\geq 0}\binom{n}{n-k}\frac{y^k}{k!}, \end{aligned}$$

so

$$\left(\frac{d}{dy}\right)^{n-m} C(y) = \sum_{k\geq 0} \binom{n}{m-k} \frac{y^k}{k!} ,$$

therefore

$$\left(\frac{d}{dy}\right)^{n-m} C(y) \Bigg|_{y=\frac{x}{2}} = \sum_{k\geq 0} \binom{n}{m-k} \frac{1}{k!} \frac{x^k}{2^k}$$

$$= \sum_{k\geq 0} \frac{n!}{2^k\, k!\, (m-k)!\, (n-m-k)!}\, x^k .$$

□

6. The q-analogue

We extend the combinatorial interpretation of the normal order coefficients to the q-Weyl algebra: the algebra with two generators D and U, and the relation $DU = qUD + 1$.

The commutation relation twisted by q comes up in physics as the relation obeyed by the creation and annihilation operators of q-deformed bosons [**Sc**]. The problem of normally ordering these operators has been studied by Katriel [**Ka**], and recently by Schork [**Sc**].

The basic idea of the q-analogue of numbers is that the polynomial $q^0 + q^1 + q^2 + \cdots + q^{n-1}$ plays the role of the positive integer n. We denote the q-analogue of n by $[n]$. Since $1 + q + q^2 + \cdots + q^{n-1} = \frac{1-q^n}{1-q}$, we can extend the definition of the q-analogue to all numbers t by defining

$$[t] := \frac{1-q^t}{1-q}.$$

The q-analogue of the derivative $\frac{d}{dx}$ acting on the ring of polynomials in x is defined as

$$D_q\, f(x) := \frac{f(qx) - f(x)}{(q-1)x} ,$$

and it is easy to check that $D_q(x^n) = [n]x^{n-1}$. For a good exposition of the q-analogue of the derivative, we refer the reader to "Quantum Calculus" by Kac and Cheung [**KC**].

If we let $D = D_q$, and let U be the operator acting by multiplication by x, then the algebra generated by D and U has the relation $DU = qUD + 1$, and is therefore the q-Weyl algebra.

Let the element w in the q-Weyl algebra be a word composed of the letters D and U. We adapt the proof of Theorem 3.2 to find the normal order coefficients of w.

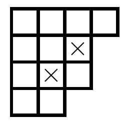
In terms of algebraic operations, we get the normal ordering form of w by successively replacing DU by $qUD + 1$, and expanding. As before, we can consider w as a formal word in the letters D and U, and the substitution as a choice of either replacing the rightmost DU by UD, weighting this choice by q, or deleting the rightmost DU. In terms of the Ferrers board B_w outlined by w, we assign the weight q to each square that doesn't have a rook either on it, below it in the same column, or to the right of it in the same row. If we consider the weight of a rook placement to be the product of the weights of all squares of the board, such weights of rook placements describe exactly the q-rook numbers of Garsia and Remmel [**GR**].

Definition 6.1. Let B be a board, and denote by $C_k(B)$ the collection of all placements of k marked squares (rooks) on B, no two in the same row or column. We define the kth q-rook number of B to be

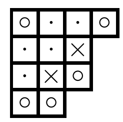
$$R_k(B, q) = \sum_{C\in C_k(B)} q^{inv(\sigma)(C)},$$

where $inv(\sigma)(C)$ is the number of squares in the placement C that do not have a rook either on them, below them in the same column, or to the right of them in the same row.

Remark 6.2. To clarify the statistic $inv(\sigma)(C)$, we demonstrate by an example. Suppose B is a Ferrers board with column heights $h(B) = (4, 4, 3, 1)$, and we have the following placement C of two rooks



If we mark with a dot the squares above or to the left of a rook, and with a circle the rest of the squares, we get



Then $inv(\sigma)(C)$ is the number of squares marked with a circle, which in this example is 5.

The statistic $inv(\sigma)(C)$ is a generalization of the inversion statistic on permutations. Given a permutation $\sigma = (\sigma_1, \ldots, \sigma_n)$, we get a placement C of n rooks on an n-by-n board where the rook in column i is placed in row σ_i. Then each square marked with a circle has a rook to the left of it and a rook above it, so the square corresponds to an inversion pair $i < j$ such that $\sigma_i > \sigma_j$. So in this case, $inv(\sigma)(C)$ is the number of inversions of σ.

Analogous to Theorem 3.2, the normal order coefficients of a word w are the q-rook numbers of the Ferrers board outlined by w.

Theorem 6.3. *Let the element w in the q-Weyl algebra be a word composed of n D's and m U's. Then*

$$w = \sum_{k=0}^{n} R_k(B_w, q) U^{m-k} D^{n-k}.$$

The Factorization Theorem for q-rook numbers [**GR**] can be proved as a corollary.

Theorem 6.4. Factorization Theorem for q-Rook Numbers

For a Ferrers board B with column heights $h(B) = (h_1, \ldots, h_n)$,

$$\sum_{k=0}^{n} R_k(B, q)[x][x-1]\cdots[x-(n-k)+1] = \prod_{i=1}^{n} [x + h_i - n + i]$$

The proof is exactly the same as for Theorem 4.1, replacing the real numbers, the falling factorials, and the derivative with their q-analogue.

References

[BPS] P. Blasiak, K. A. Penson, and A. I. Solomon, *The general boson normal ordering problem*, Physics Letters A **309** (2003), 198–205.

[GR] A. M. Garsia and J. B. Remmel, *q-Counting rook configurations and a formula of Frobenius*, J. Comb. Theory, Series A **41** (1986), 246–275.

[GJW] J. R. Goldman, J. T. Joichi, and D. E. White, *Rook theory. I. Rook equivalence of Ferrers boards*, Proc. Amer. Math. Soc. **52** (1975), 485–492.

[KC] V. Kac and P. Cheung, *Quantum calculus*, Springer-Verlag, NY (2002).

[Ka] J. Katriel, *Normal ordering formulae for some boson operators*, J. Phys. A **16** (1983), 4171–4173.

[La] W. Lang, *On generalizations of the Stirling number triangles,* J. Int. Seq. **3** (2000), Article 00.2.4.

[Ma] J. H. Marburger, *Conditions for the existence of closed solutions by the normal ordering method*, J. Mathematical Phys. **7** (1966), 829–832.

[Mi1] V. V. Mikhailov, *Ordering of some boson operator functions*, J. Phys. A **16** (1983), 3817–3827.

[Mi2] V. V. Mikhailov, *Normal ordering and generalized Stirling numbers*, J. Phys. A **18** (1985), 231–235.

[Na] A. M. Navon, *Combinatorics and fermion algebra*, Nuovo Cimento **16** (1973), 324–330.

[Ri] J. Riordan, *An introduction to combinatorial analysis*, Wiley, New York 1958.

[Sc] M. Schork, *On the combinatorics of normal ordering bosonic operators and deformations of it*, J. Phys. A **36** (2003), 4651–4665.
[St] R. P. Stanley, *Enumerative combinatorics*, Vol. 1, Cambridge University Press, Cambridge (1997).

Department of Mathematics, Brandeis University, Waltham, MA, USA, 02454
E-mail address: anka@brandeis.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
Vancouver 2004

Sheared Tableaux and bases for the symmetric functions

Benjamin Young

Abstract.
We study the operation of shearing *Schur functions, which yields a new family of bases for the space of symmetric functions. In the course of this study, we derive some interesting combinatorial results and inequalities on the Littlewood-Richardson coefficients of sheared Schur functions.*
Résumé. *Nous étudions l'opération de trancher les fonctions de Schur, qui nous donne une nouvelle famille de bases pour l'espace des fonctions symétriques. Pendent cette étude, nous dérivons des résultats et des inégalités combinatoires intéressants. Ces résultats décrivant les coefficients Littlewood-Richardson des fonctions Schur tranchées.*

1. Introduction

Suppose n is a positive integer and $\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0)$ is a partition of n, denoted $\lambda \vdash n$. The *(Young) diagram* associated to λ is a diagram made of rows of boxes, in which the kth row has λ_k boxes. If $m \leq n$ and μ is a partition of m such that the Young diagram of μ is contained within the Young diagram of λ, we define the *skew diagram* λ/μ to be the set of all boxes contained in the diagram of λ, but not contained in the diagram of μ. For emphasis, we shall refer to Young diagrams which are not skew as *normal* diagrams.

A *Young tableau* T is a Young diagram in which the boxes have been filled with positive integers. If the rows of T are weakly increasing and the columns of T are strictly increasing, then T is said to be *semistandard*. The *content* $c(T)$ of T is the weak composition of nonnegative integers $(\gamma_1, \ldots, \gamma_m)$ for which γ_i is the number of i's in T.

Let Λ denote the graded algebra of symmetric functions. Λ has a well-known basis consisting of *Schur functions*, denoted s_λ and indexed by normal Young diagrams $\lambda \vdash n$, for all positive integers n. We define s_λ as

$$s_\lambda = \sum_T x^{\mathrm{c(T)}}$$

where T runs over all semistandard tableaux with shape λ, and $x^\gamma = x_1^{\gamma_1} x_2^{\gamma_2} \cdots$. We can also define the *skew Schur function* $s_{\lambda/\mu}$ in precisely the same fashion.

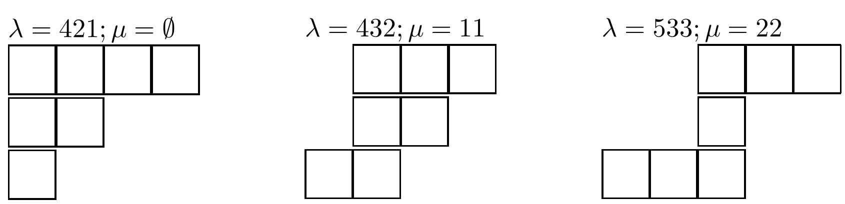


FIGURE 1. Normal, skew, and ribbon diagrams

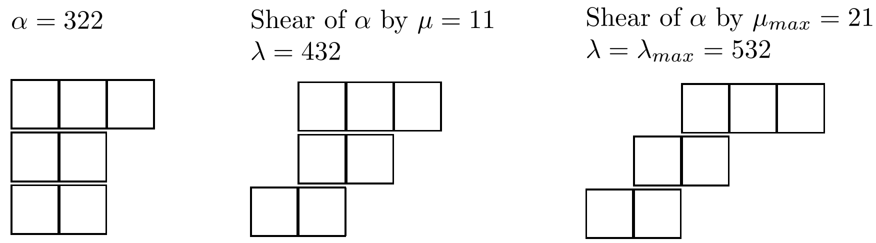


FIGURE 2. Shearings of a partition α

A diagram λ is said to be *connected* if any two adjacent rows of λ share a common horizontal edge. We define a *ribbon diagram* to be a connected diagram which does not contain a two-by-two sub-diagram, and define *ribbon tableaux* and *ribbon Schur functions* accordingly. Please refer to Figure 1 for examples of normal, skew, and ribbon diagrams.

Ribbons are of particular interest as they play a fundamental rule in the Murnaghan-Nakayama rule (see [**4**], chapter 7). In addition, the recent papers [**1**] and [**2**] link ribbons to the Fock space representation of $U_q(\widehat{\mathfrak{sl}_n})$.

In this paper, we provide a new basis of the symmetric functions which consists of ribbon Schur functions. This basis is obtained from the normal Schur functions by the process of *shearing*. Moreover, the change-of-basis matrix from the normal Schur functions to the shear basis has some interesting combinatorial properties. In particular, the entries of this matrix are Littlewood-Richardson coefficients, some of which are explicitly computed to be 0 or 1. Further results on symmetric functions and Littlewood-Richardson coefficients can be found in [**4**].

2. Shearing

Suppose $\alpha = (\alpha_1, \ldots, \alpha_k) \vdash n$. Define the *maximal shear of* α to be the ribbon diagram with rows of length $\alpha_1, \ldots, \alpha_n$.

If the maximal shear of α is the skew diagram λ_{max}/μ_{max}, and μ is any diagram contained in μ_{max}, consider the skew diagram λ/μ, where $\lambda = \mu + \alpha = (\alpha_1 + \mu_1, \ldots, \alpha_k - 1 + \mu_k - 1, \alpha_k)$. If λ/μ is a connected diagram, we say that μ is a *shearing diagram* for λ. We define the *shear of* α *by* μ, denoted $\mathrm{Shear}_\mu(\alpha)$, to be the skew diagram λ/μ.

While the maximal shears of normal diagrams α are the primary objects of interest, our main theorem works for any shearing of α. For an example of sheared Young diagrams, see Figure 2.

Let $P(n)$ be the set of all partitions of n. For the remainder of this section, fix a function $M : P(n) \to \bigcup_{i=0}^{\infty} P(i)$ which maps each $\alpha \vdash n$ to a shearing diagram μ for α. To simplify the notation, for a fixed partition α of n, we will abuse notation and write $\mathrm{Shear}(\alpha)$ or $s_{\lambda/\mu}$ in place of $s_{\mathrm{Shear}_M(\alpha)}$.

We will prove the following:

Theorem 1. The set of skew Schur functions

$$\left\{\mathrm{Shear}(\lambda) \middle| n \in \mathbb{Z}^+, \lambda \vdash n\right\},$$

forms a basis for Λ.

Applying Theorem 1 using maximal shears gives the following:

Corollary. The set of ribbon Schur functions in which the rows are weakly decreasing in length form a basis of Λ.

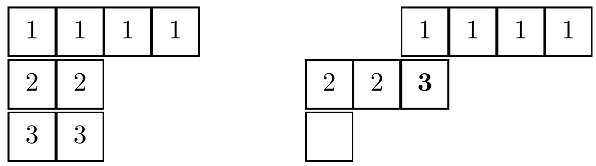


FIGURE 3. An ill-fated attempt to create a tableau T of shape 521/21 and content 43

To prove this theorem, we shall restrict ourselves to the set of symmetric functions of a fixed degree n, denoted Λ^n. The Schur functions of degree n, $\{s_\lambda | \lambda \vdash n\}$, form a basis for Λ^n, and represent the trivial shear $M = 0$. We shall demonstrate that the set of Schur functions

$$\{\text{Shear}(\lambda) \big| \lambda \vdash n\}$$

also form a basis for Λ^n, by creating a change-of-basis matrix from the Schur function basis to $\{Shear(\lambda)\}$. The theorem follows immediately, since $\Lambda = \bigoplus_{i=1}^{\infty} \Lambda^i$.

We express λ/μ in terms of the Schur function basis using Littlewood-Richardson coefficients:

$$\text{Shear}(\alpha) = \sum_{\nu} c^{\lambda}_{\mu\nu} s_\nu$$

Following [**3**], we say that a word $a_1 a_2 \ldots a_n$ is a reverse lattice partition if, in any of the suffixes $a_k a_{k+1} \ldots a_n$, the number of l's is at least as large as the number of $(l+1)$'s. By the Littlewood-Richardson rule, the coefficient $c^{\lambda}_{\mu\nu}$ counts the number of semistandard tableaux T for which

(a) T has shape λ/μ and content ν,
(b) the row word of T is a reverse lattice partition.

In particular, if $\nu = \alpha$, the first row of T must end with a 1 by condition (2), and must be weakly increasing by condition (1). Therefore, the first row of T is made up of α_1 ones. However, the content of T is equal to α – that is, there are only α_1 ones in T. Hence, the rest of T must be filled with integers greater or equal to 2. Repeating this argument for the rest of the rows in λ/μ, we find that the only way to construct T is to fill the ith row with the number i. We have proven

Lemma 1. $c^{\lambda}_{\mu\alpha} = 1$.

Now, suppose that $\alpha \not\preceq \nu$, where $\preceq$ is the dominance ordering of partitions: $(\beta_1, \ldots, \beta_k) \preceq (\gamma_1, \ldots, \gamma_j)$ means $\sum_{i=1}^{t} \beta_i \le \sum_{i=1}^{t} \gamma_i$ for each $t \le \max\{k, j\}$. Suppose further that we have constructed a tableau T of shape Shear(α) satisfying conditions (1) and (2) above.

Observe that row t of T contains only numbers which are less than or equal to t. This can be seen by induction on t. The base case asserts that the first row of T contains only ones, which has already been shown in the context of Lemma 1. Now suppose that the first $t-1$ rows contain only numbers which are less than or equal to $t-1$. The largest element t_1 of row t must occur at the right end of row t; in order to satisfy the requirement that the row word of T be a reverse lattice partition, $t_1 = t$ (otherwise, the number of t_1's leads the number of t's at the right end of row t). See Figure 3 for an example.

Let t_0 be the first t for which $\sum_{i=1}^{t} \alpha_i > \sum_{i=1}^{t} \nu_i$. The sum on the left side of the inequality counts the boxes in the first t_0 rows of λ/μ, whereas the sum on the right side of the inequality counts the boxes in the first t_0 rows of ν.

Let t_1 be the largest element of row t_0 in the diagram λ/μ. Since T is to be filled with the content of ν, it follows that $t_1 > t_0$, a contradiction. So such a tableau T cannot exist after all. We have proven

Lemma 2. If $\alpha \not\preceq \nu$, then $c^{\lambda}_{\mu\nu} = 0$.

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 2 & 2 & 2 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 2 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

FIGURE 4. The change–of–basis matrix M for maximal shears of s_λ, $n = 7$

Proof of Theorem 1: Let us write $c_\nu^{\mathrm{Shear}(\alpha)}$ for $c_{\mu\nu}^\lambda$, in order to emphasize the connection between α and ν. Let M be the matrix $[c_\nu^{\mathrm{Shear}(\alpha)}]$ in which the rows are ordered lexicographically by α, and the columns are ordered lexicographically by ν. Observe that M gives the coordinates of each $\mathrm{Shear}(\alpha)$ in terms of the normal Schur function basis s_ν.

Since the dominance ordering is a strengthening of the lexicographic ordering, Lemma 2 implies that M is upper triangular, while Lemma 1 implies that M has ones on its main diagonal (for a concrete example of M, see Figure 4). Therefore, M is invertible, so the chosen set of sheared Schur functions must form a basis for Λ^n. □

3. Shears of a single diagram α

Let us now fix a partition α of n. Suppose μ and η are two shearing diagrams for α such that μ is contained in η. We say that $\mathrm{Shear}_\eta(\alpha)$ is a *relative shearing* of $\mathrm{Shear}_\mu(\lambda)$ if the skew diagram η/μ is a shearing of some normal diagram – that is, if the rows of η/μ are weakly decreasing in length. We have the following theorem:

Theorem 2. If $\mathrm{Shear}_\eta(\alpha)$ is a relative shearing of $\mathrm{Shear}_\mu(\alpha)$, then

$$c_\nu^{\mathrm{Shear}_\mu(\alpha)} \leq c_\nu^{\mathrm{Shear}_\eta(\alpha)}.$$

Proof. In order to prove the theorem, suppose we have constructed a Young tableau T_μ with shape $\mathrm{Shear}_\mu(\alpha)$ and content ν which meets conditions (1) and (2) of the Littlewood-Richardson rule. Let T_η be the Young tableau with shape $\mathrm{Shear}_\eta(\alpha)$ such that the ith row of T_η is equal to the ith row of T_μ. If we show that T_η also satisfies conditions (1) and (2), then the Littlewood-Richardson rule will imply the theorem.

Observe that the row word of T_μ is equal to the row word of T_η, and thus the rows of T_η are weakly increasing. So we need only check that T_η has strictly increasing columns.

Let $\Delta = (\eta_r - \mu_r) - (\eta_{r+1} - \mu_{r+1})$ be the number of boxes that row r moves with respect to row $r+1$ as we pass from $\mathrm{Shear}_\mu(\alpha)$ to $\mathrm{Shear}_\eta(\alpha)$. Because $\mathrm{Shear}_\eta(\alpha)$ is a relative shearing of $\mathrm{Shear}_\mu(\alpha)$, we know that $\Delta \geq 0$.

Suppose the element a lies in row r of T_η, and b lies directly below a in row $r+1$ of T_η. There are at least Δ elements to the left of b; take the rightmost of these and label them $b_1, \ldots, b_\Delta$. Observe that a lies

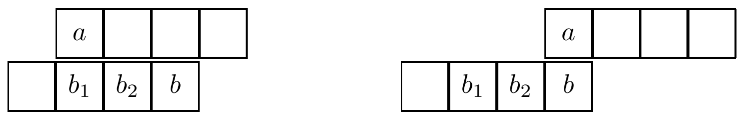


FIGURE 5. Two adjacent rows of T_μ and T_η; $\Delta = 2$

above b_1 in T_μ. Because T_μ is semistandard, we have that $a < b_1 \leq \cdots \leq b_\Delta \leq b$ (see Figure 5). Thus the columns of T_η are strictly increasing. □

4. Vertical shearing

In the preceeding development, all of our shears have been *horizontal*, in the sense that we have obtained a shear of the tableau α by shifting some of the rows of α to the right. We could equally well shift some of the *columns* of α downward, to obtain a *vertical shearing* of α.

A proof similar to that of Theorem 1 can be employed to show that any set of vertical shears of the normal Schur functions also yields a new basis for λ. In the proof, the change of basis matrix M becomes lower triangular. Likewise, Theorem 2 also holds with relative shears being replaced with relative vertical shears.

5. Further work

Aside from exploring connections to the active ribbon-based research areas mentioned, the most pressing issue which arises from this work is to study the change-of-basis matrix M in greater detail. The author has computed M for $n \leq 8$, using an algorithm for computing Littlewood-Richardson coefficents. These small matrices M are fairly sparse and have small entries. It seems likely that the best way to obtain further information about shearing is to sharpen these observations.

Computations of M grow quickly intractable as n increases, so it would also be worthwhile to look for shear-specific algorithms for computing Littlewood-Richardson coefficients of sheared Schur functions. Theorem 2, in particular, suggests that to compute M for $\mathrm{Shear}_\eta(\alpha)$, we could proceed iteratively. First, one would find shearing functions $\mu_0 = \emptyset \subseteq \mu_1 \subseteq \ldots \subseteq \mu_k = \eta$, where each μ_i is a relative shearing of μ_{i-1}. Then, one would compute the corresponding matrices M_i for $\mathrm{Shear}_{\mu_i}(\alpha)$. Hopefully, if enough intermediary μ_i are used, the changes between M_i and M_{i+1} will be small. Of course, in order for this to work, we would need to sharpen Theorem 2 considerably, providing at least an *upper* bound for the Littlewood-Richardson coefficients in question.

References

[1] Thomas Lam, *Ribbon tableaux and the Heisenberg algebra*, preprint, www.arXiv.org, 2003.

[2] Alain Lascoux, Bernard Leclerc, and Jean-Yves Thibon, *Ribbon tableaux, Hall-Littlewood functions, quantum affine algebras, and unipotent varieties*, Journal of Mathematical Physics **38** (1997), 1041–1068.

[3] Bruce E. Sagan, *The symmetric group: Representations, combinatorial algorithms, and symmetric functions*, Springer-Verlag New York Inc., New York, 2001.

[4] Richard P. Stanley, *Enumerative combinatorics*, vol. 2, Cambridge University Press, Cambridge, UK, 1999.

Index of Authors and Keywords