

# QUASI-SYMMETRIC FUNCTIONS AS POLYNOMIAL FUNCTIONS ON YOUNG DIAGRAMS

JEAN-CHRISTOPHE AVAL, VALENTIN FÉRAY, JEAN-CHRISTOPHE NOVELLI,  
AND JEAN-YVES THIBON

ABSTRACT. We determine the most general form of a smooth function on Young diagrams, that is, a polynomial in the interlacing or multirectangular coordinates whose value depends only on the shape of the diagram. We prove that the algebra of such functions is isomorphic to quasi-symmetric functions, and give a noncommutative analog of this result.

## 1. INTRODUCTION

A central question in this paper is the following problem:

Characterize the polynomials  $f(x_1, x_2, x_3, \dots)$  in infinitely many variables<sup>1</sup> such that

$$(1) \quad f(x_1, x_2, \dots)|_{x_i=x_{i+1}} = f(x_1, \dots, x_{i-1}, x_{i+2}, \dots).$$

**1.1. Motivation: Young diagrams and Equation (1).** Consider a Young diagram  $\lambda$  drawn with the Russian convention, (*i.e.*, draw it with the French convention, rotate it counterclockwise by  $45^\circ$  and scale it by a factor  $\sqrt{2}$ ). Its border can be interpreted as the graph of a piecewise affine function. We denote by  $x_1, x_2, \dots, x_{2m+1}$  the abscissas of its local minima and maxima in decreasing order, see Figure 1.

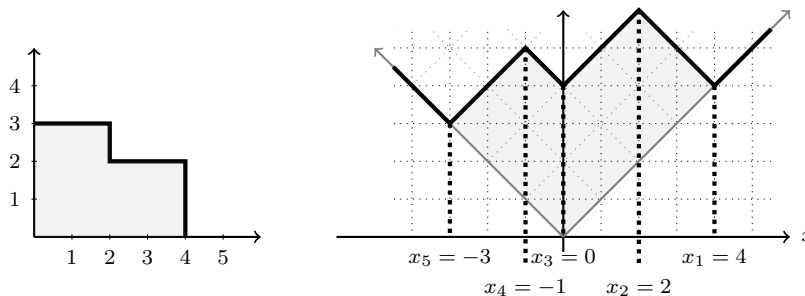


FIGURE 1. Young diagram  $\lambda = (4, 4, 2)$  and the graph of the associated function  $\omega_\lambda$ .

These numbers  $x_1, x_2, \dots, x_{2m+1}$  are called (*Kerov*) *interlacing coordinates*, see, *e.g.*, [21, Section 6 with  $\theta = 1$ ]. They are usually labeled with two different alphabets for minima and maxima, but we shall rather use the same alphabet here and distinguish between odd-indexed and even-indexed variables when necessary.

---

<sup>1</sup>Understood as elements of an inverse limit, see Section 2.

Note that not any decreasing sequence of integers can be obtained in this way, as interlacing coordinates always satisfy the relation  $\sum_i (-1)^i x_i = 0$ . Yet, this construction defines an injective map

$$(2) \quad \mathcal{IC} : \{\text{Young diagrams}\} \rightarrow \left\{ \begin{array}{l} \text{finite sequences} \\ \text{of integers} \end{array} \right\}.$$

A polynomial in infinitely many variables can be evaluated on any finite sequence. By composition with  $\mathcal{IC}$ , one may wish to interpret it as a function on all Young diagrams.

A Young diagram can be easily recovered from its Kerov coordinates  $x_1, \dots, x_{2m+1}$ . To obtain its border, first draw the half-line  $y = -x$  for  $x \leq x_{2m+1}$ , then, without raising the pen, draw line segments of slope alternatively  $+1$  and  $-1$  between points of  $x$ -coordinates  $x_{2m+1}, x_{2m}, \dots, x_1$  and finally a half-line of slope  $+1$  for  $x \geq x_1$ . Starting with a decreasing integral sequence satisfying  $\sum_i (-1)^i x_i = 0$ , the last half-line has equation  $y = x$  and the resulting broken line can be interpreted as the border of a Young diagram drawn with the Russian convention (see [11, Proposition 2.4]).

Apply now the same process to a non-increasing sequence  $x_1, x_2, \dots, x_{2m+1}$  such that  $x_i = x_{i+1}$ . Reaching the  $x$ -coordinate  $x_i = x_{i+1}$ , one has to change twice the sign of the slope, that is, to do nothing. Hence, one obtains the same diagram as for sequence

$$x_1, \dots, x_{i-1}, x_{i+2}, \dots, x_{2m+1}.$$

Therefore, if one wants to interpret a polynomial in infinitely many variables as a function of Young diagrams, it is natural to require that it satisfies Equation (1).

**1.2. Solution.** In Section 3, we shall describe the algebraic structure of the space  $\mathcal{S}$  of solutions of Equation (1).

**Theorem 1.1.** *As an algebra,  $\mathcal{S}$  is isomorphic to  $QSym$ , the algebra of quasi-symmetric functions.*

The algebra of quasi-symmetric functions is a natural extension of that of symmetric functions, widely studied in the literature. Its definition is recalled in Section 2. The isomorphism of Theorem 1.1 is naturally described in terms of Hopf algebra calculus: the solutions to Equation (1) are the quasi-symmetric functions evaluated on the virtual alphabet (this notion is defined in Section 2.3)

$$(3) \quad \mathbb{X} = \ominus(x_1) \oplus (x_2) \ominus (x_3) \oplus (x_4) \cdots$$

We emphasize here the similarity with a result of J.R. Stembridge [25]. He studied the solutions of Equation (1) which are in addition symmetric in the odd-indexed variables  $x_1, x_3, \dots$  and separately in the even-indexed variables  $x_2, x_4, \dots$ . He proved that the space of these functions is algebraically generated by the power sums in the virtual alphabet above, that is

$$(4) \quad p_k(\mathbb{X}) = \sum_i (-1)^i x_i^k.$$

In other words, the *symmetric* solutions to (1) are the symmetric functions evaluated on  $\mathbb{X}$ . From this point of view, our result is the natural quasi-symmetric analogue of Stembridge’s theorem.

**1.3. Back to Young diagrams.** As explained in Section 1.1, the solutions of (1) can be interpreted as functions on Young diagrams.

It turns out that symmetric polynomials evaluated on  $\mathbb{X}$  (which form a subset of  $\mathcal{S}$ ) correspond to a well-known algebra of functions on Young diagrams, denoted here by  $\Lambda$ , and referred to as *symmetric functions on Young diagrams*<sup>2</sup>. This algebra, introduced by Kerov and Olshanski [12], is algebraically generated by the family of functions (4) [11, Corollary 2.8]. Therefore, Equation (1) leads us to a new algebra of functions on Young diagrams, strictly larger than that of Kerov and Olshanski, see Section 4.1 for details.

Our algebra  $\mathcal{S}$  has a rich structure (quasi-symmetric functions), and it contains some non-*symmetric* functions, denoted by  $N_G$ , which have played an important role in the approach to representation theory of the symmetric group developed in some recent papers (see, *e.g.*, [2] and references therein). In Section 4.5, we give a new formula for  $N_G$  in terms of quasi-symmetric functions of  $\mathbb{X}$ .

Besides, in Section 4.4, we describe our new algebra  $\mathcal{S}$  in terms of the so-called multirectangular coordinates of Young diagrams.

**1.4. Generalization to a noncommutative framework.** To avoid confusion, we will always use a set of variables  $\{a_1, a_2, \dots\}$  for polynomials in noncommuting variables. A functional equation, analogous to (1), can be considered:

Characterize the polynomials  $P(a_1, a_2, a_3, \dots)$  in infinitely many *non commuting* variables such that

$$(5) \quad P(a_1, a_2, \dots)|_{a_i=a_{i+1}} = P(a_1, \dots, a_{i-1}, a_{i+2}, \dots).$$

We solve this problem in Section 5.

**Theorem 1.2.** *As an algebra, the space of solutions of (5) is isomorphic to the algebra of word quasi-symmetric functions **WQSym**.*

The definition of **WQSym** is recalled in Section 5.1. This algebra is the natural noncommutative analogue of *QSym*.

As in the commutative setting, the solutions are constructed from the elements of **WQSym** using a virtual alphabet:

$$(6) \quad \mathbb{A} = \ominus(a_1) \oplus (a_2) \ominus (a_3) \oplus (a_4) \ominus \dots$$

However, there was here an extra difficulty. Differences of alphabets for word quasi-symmetric functions cannot be defined by means of the antipode, which is not involutive, so we had to introduce an ad-hoc definition, see Section 5.2.

---

<sup>2</sup>The usual terminology is *polynomial functions on Young diagrams*, which can be confusing as some functions that do not belong to this algebra depend polynomially on interlacing coordinates  $x_1, x_2, \dots$

Evaluating these functions in noncommuting variables on the interlacing coordinates of Young diagrams does not bring any new information, as this operation factors through the commutative version. It is however interesting to study the change of variables between interlacing coordinates and multirectangular coordinates in the noncommutative framework. We study this morphism and describe its kernel in Section 6. This result involves the lifting of a basis of  $QSym$  introduced by K. Luoto (see Remark 6.16) and the computation of the dimension of the smallest two-sided ideal of  $\mathbf{WQSym}$  containing the element of degree 1 and stable by the actions of the symmetric groups (see Theorem 6.17, first item), which might be of interest on their own.

**1.5. Open problem.** The algebra  $\Lambda$  of symmetric functions on Young diagrams considered by Kerov and Olshanski has several equivalent descriptions [11, Sections 1 and 2]:

- as symmetric functions of the interlacing coordinates ;
- as shifted symmetric function of the parts of the partition ;
- as symmetric functions of the modified Frobenius coordinates ;
- as symmetric functions of the multiset of *contents* with coefficients depending polynomially on the size of the diagram.

Our algebra of quasi-symmetric functions on Young diagrams naturally extends the first item. Are there equivalent descriptions of our algebra in terms of these other interpretations?

## 2. DEFINITIONS AND NOTATIONS IN THE COMMUTATIVE FRAMEWORK

**2.1. Stable polynomials.** By “polynomial in infinitely many variables”, we mean an element of an inverse limit in the category of graded rings, *i.e.*, a homogeneous polynomial of degree  $d$  is a sequence  $R = (R_n(x_1, \dots, x_n))_{n \geq 0}$  of homogeneous polynomials of degree  $d$  such that  $R_{n+1}(x_1, \dots, x_n, 0) = R_n(x_1, \dots, x_n)$ . These objects are sometimes called *stable polynomials*. Their set will be denoted  $\mathbb{C}[X]$ , where  $X$  is the infinite variable set  $X = \{x_1, x_2, \dots\}$  (which should not be confused with the virtual alphabet  $\mathbb{X}$  defined by Equation (3)).

This kind of construction is classical in algebraic combinatorics: for instance, symmetric functions (see [15]) and quasi-symmetric functions (see below) are built in this way.

In this context, Equation (1) should be understood as follows. A stable polynomial  $f = (f_n)_{n \geq 0}$  is solution of (1) if for each  $n \geq 2$  and each  $1 \leq i < n$ , one has:

$$f_n(x_1, \dots, x_n)_{|x_{i+1}=x_i} = f_{n-2}(x_1, \dots, x_{i-1}, x_{i+2}, \dots, x_n).$$

The left-hand side means that we substitute  $x_{i+1}$  by  $x_i$ . Then the equality must be understood as an equality between polynomials in  $x_1, \dots, x_{i-1}, x_{i+2}, \dots, x_n$ . In particular, the left-hand side must be independent of  $x_i$ .

We will also need to consider polynomials in two infinite sets of variables. By definition, an element of  $\mathbb{C}[\mathbf{p}, \mathbf{q}]$  is a sequence  $(h_m)_{m \geq 0}$ , where each  $h_m$  is a polynomial

in the  $2m$  variables  $p_1, \dots, p_m, q_1, \dots, q_m$  satisfying the stability property

$$(7) \quad h_{m+1} \begin{pmatrix} p_1 & \cdots & p_m & 0 \\ q_1 & \cdots & q_m & 0 \end{pmatrix} = h_m \begin{pmatrix} p_1 & \cdots & p_m \\ q_1 & \cdots & q_m \end{pmatrix}.$$

Finally, we shall sometimes define stable polynomials by a sequence of polynomials in an odd number of variables  $(R_{2m+1})_{m \geq 0}$  such that

$$R_{2m+1}(x_1, \dots, x_{2m-1}, 0, 0) = R_{2m-1}(x_1, \dots, x_{2m-1}).$$

This is not an issue, as such a sequence can be extended in a unique way to a stable sequence  $(R_n)_{n \geq 0}$  by setting

$$R_{2m}(x_1, \dots, x_{2m}) = R_{2m+1}(x_1, \dots, x_{2m}, 0).$$

**2.2. The Hopf algebra of quasi-symmetric functions.** Quasi-symmetric functions were introduced by I. Gessel [9] and may be seen as a generalization of the notion of symmetric functions.

A *composition* of  $n$  is a sequence  $I = (i_1, i_2, \dots, i_r)$  of positive integers, whose sum is equal to  $n$ . The notation  $I \models n$  means that  $I$  is a composition of  $n$  and  $\ell(I)$  denotes the number of parts of  $I$ . In numerical examples, it is customary to omit the parentheses and the commas. For example, 212 is a composition of 5. Given two compositions  $I = (i_1, i_2, \dots, i_r)$  and  $J = (j_1, j_2, \dots, j_s)$  their *concatenation* is  $I \cdot J = (i_1, \dots, i_r, j_1, \dots, j_s)$ .

Recall that the *multiset* of *shuffles* of  $I$  and  $J$  is defined recursively by:

$$(8) \quad I \sqcup J = (i_1) \cdot \left( (i_2, \dots, i_r) \sqcup J \right) \sqcup (j_1) \cdot \left( I \sqcup (j_2, \dots, j_s) \right).$$

In quasi-symmetric function theory, we use a slight modification of the shuffle, called *quasi-shuffle*, defined recursively by:

$$(9) \quad \begin{aligned} I \star J &= (i_1) \cdot \left( (i_2, \dots, i_r) \star J \right) \sqcup (j_1) \cdot \left( I \star (j_2, \dots, j_s) \right) \\ &\quad \sqcup (i_1 + j_1) \left( (i_2, \dots, i_r) \star (j_2, \dots, j_s) \right). \end{aligned}$$

For a composition  $I = (i_1, \dots, i_r)$ , we denote by  $\bar{I}$  the composition  $(i_r, \dots, i_1)$  *mirror* to  $I$ . For two compositions  $I = (i_1, \dots, i_r)$  and  $J = (j_1, \dots, j_s)$  we say that  $J$  is a *refinement* of  $I$  if

$$\{i_1, i_1 + i_2, \dots, i_1 + \dots + i_r\} \subseteq \{j_1, j_1 + j_2, \dots, j_1 + \dots + j_s\},$$

which will be denoted by  $I \leq J$ .

The algebra  $QSym$  of quasi-symmetric functions is a subalgebra of the algebra  $\mathbb{C}[X]$  of polynomials in the totally ordered commutative alphabet  $X = \{x_1, x_2, \dots\}$ . Monomials  $X^{\mathbf{v}}$  correspond to vectors  $\mathbf{v} = v_1, v_2, \dots$  with finitely many non-zero entries. For such a vector, we denote by  $\mathbf{v}_{\leftarrow}$  the vector obtained by omitting the zero entries. A polynomial  $P \in \mathbb{C}[X]$  is said to be *quasi-symmetric* if and only if for any  $\mathbf{v}$  and  $\mathbf{w}$  such that  $\mathbf{v}_{\leftarrow} = \mathbf{w}_{\leftarrow}$ , the coefficients of  $X^{\mathbf{v}}$  and  $X^{\mathbf{w}}$  in  $P$  are equal. It should be clear that any symmetric polynomial is quasi-symmetric. The algebra  $QSym$  of

quasi-symmetric functions has a basis of monomial quasi-symmetric functions  $M$  indexed by compositions  $I = (i_1, \dots, i_r)$ , where

$$(10) \quad M_I = \sum_{a_1 < \dots < a_r} x_{a_1}^{i_1} \cdots x_{a_r}^{i_r}.$$

By convention,  $M_{\emptyset} = 1$ , where  $\emptyset$  designs the empty composition.

The product of monomial functions is given by the quasi-shuffle of their indices:  $M_I M_J = \sum_{K \in I \star J} M_K$ . For instance,

$$M_2 M_{11} = M_{112} + M_{121} + M_{211} + M_{13} + M_{31}.$$

This given, the set  $QSym$  is an algebra with unit  $M_{\emptyset} = 1$ . Moreover it is graded by the usual degree. The coproduct of  $QSym$  may be defined on monomial functions through the deconcatenation of compositions:

$$(11) \quad \Delta(M_I) = \sum_{k=0}^r M_{(i_1, \dots, i_k)} \otimes M_{(i_{k+1}, \dots, i_r)}.$$

As an example, one has

$$\Delta M_{21} = 1 \otimes M_{21} + M_2 \otimes M_1 + M_{21} \otimes 1.$$

This operation endows  $QSym$  with a bialgebra structure, whence a Hopf algebra structure since it is graded. The antipode, denoted as usual by  $S$  has been explicitly computed by C. Malvenuto [16] and R. Ehrenborg [3]:

$$(12) \quad S(M_I) = (-1)^{\ell(I)} \sum_{J \leq I} M_J.$$

For example, we have:

$$S(M_{122}) = -(M_{221} + M_{41} + M_{23} + M_5).$$

**2.3. Evaluation of quasi-symmetric functions on sums and differences of alphabets.** We explain now how to evaluate quasi-symmetric functions on sums or differences of alphabets. To start with, we shall give another interpretation of the coproduct  $\Delta$ . To do this, we consider two ordered alphabets  $X$  and  $Y$  and we denote by  $X \oplus Y$  their ordinal sum, that is, their disjoint union seen as an ordered alphabet with  $x < y$  for  $x \in X$  and  $y \in Y$ . From (11), we may check that

$$(13) \quad \Delta P = \sum_k F_k \otimes G_k$$

implies

$$(14) \quad P(X \oplus Y) = \sum_k F_k(X) G_k(Y).$$

This defines sums of alphabets and the evaluation of quasi-symmetric functions on these.

Let us now introduce the formal inverse of  $Y$  for operation  $\oplus$ , denoted by  $\ominus Y$ . Of course,  $\ominus Y$  does not exist as an alphabet; we refer to it as a *virtual alphabet*. We may write

$$(15) \quad P(X \ominus Y) = \sum_k F_k(X) G_k(\ominus Y)$$

with the same notations as above. For this to make sense, we need to define the evaluation  $P(\ominus Y)$  of a quasi-symmetric function  $P$  on the opposite of  $Y$ . We set

$$(16) \quad P(\ominus Y) = S(P)(Y).$$

Let us explain this choice. Using the axiom of the antipode in a Hopf algebra, we observe that (15) evaluated at  $X = Y$  gives, for any homogeneous quasi-symmetric function  $P$  of positive degree,

$$(17) \quad P(Y \ominus Y) = 0.$$

Similarly,

$$(18) \quad P(\ominus Y \oplus Y) = 0.$$

This is consistent with the fact that  $\ominus Y$  is the inverse of  $Y$  for  $\oplus$ , that is that  $Y \ominus Y = \ominus Y \oplus Y = \emptyset$  ( $\emptyset$  is here the empty alphabet).

Equations (14) and (16) enable us to evaluate quasi-symmetric functions on differences of alphabets. As an example, we have:

$$\begin{aligned} M_{21}(X \ominus Y) &= M_{21}(X) + M_2(X)S(M_1)(Y) + S(M_{21})(Y) \\ &= M_{21}(X) - M_2(X)M_1(Y) + M_{12}(Y) + M_3(Y). \end{aligned}$$

It is also possible to have more complicated linear combinations of alphabets (beware that  $\oplus$  is not a commutative operator), for instance,

$$\begin{aligned} M_{21}(X \ominus Y \oplus Z) &= M_{21}(X) + M_2(X)S(M_1)(Y) + M_2(X)M_1(Z) \\ &\quad + S(M_{21})(Y) + S(M_2)(Y)M_1(Z) + M_{21}(Z); \\ &= M_{21}(X) - M_2(X)M_1(Y) + M_2(X)M_1(Z) \\ &\quad + M_{12}(Y) + M_3(Y) - M_2(Y)M_1(Z) + M_{21}(Z). \end{aligned}$$

### 3. SOLUTION OF THE PROBLEM

Consider the virtual ordered alphabet for quasi-symmetric functions

$$(19) \quad \mathbb{X} = \ominus(x_1) \oplus (x_2) \ominus (x_3) \oplus (x_4) \cdots$$

If  $F$  is a quasi-symmetric function, one can define a stable polynomial  $F(\mathbb{X})$  as follows: for  $n \geq 0$ ,

$$(F(\mathbb{X}))_n(x_1, \dots, x_n) = F(\ominus(x_1) \oplus (x_2) \ominus \dots (x_n)).$$

Using Equations (17) and (18), we see that setting  $x_i = x_{i+1}$  in this alphabet cancels these variables, so that  $F(\mathbb{X})$  satisfies (1). The converse is also true:

**Theorem 3.1.** *A function  $f$  satisfies the functional equation (1) if and only if  $f \in QSym(\mathbb{X})$ .*

*Proof.* Note that a polynomial  $f$  satisfies Equation (1) if and only if all its homogeneous components do. Therefore it is enough to prove the statement for a homogeneous function  $f$ .

Let us first prove that the dimension of the space of homogeneous polynomials in  $\mathbb{C}[X]$  of degree  $n$  satisfying (1) is at most equal to  $2^{n-1}$ . We say that a monomial

$$X^{\mathbf{v}} = x_1^{v_1} x_2^{v_2} \cdots$$

in  $\mathbb{C}[X]$  is *packed* if  $\mathbf{v}$  can be written as  $\mathbf{c}, 0, 0, 0, \dots$  with  $\mathbf{c}$  a composition (*i.e.* a vector whose entries are positive integers). Thus, the number of packed monomials of degree  $n$  is  $2^{n-1}$ . Let

$$P = \sum_{\mathbf{v}} c_{\mathbf{v}} X^{\mathbf{v}}$$

be a homogeneous polynomial of degree  $n$ , which is solution of (1) (the sum runs over sequences of non-negative integers of sum  $n$ ). Associate with each monomial  $X^{\mathbf{v}}$  an integer

$$(20) \quad \ell(X^{\mathbf{v}}) = \sum_{i \in \mathbb{N}} i v_i.$$

Then, all the coefficients of  $P$  are determined by those of packed monomials. To see this, consider a non-packed monomial  $X^{\mathbf{w}} = x_1^{w_1} x_2^{w_2} \cdots$  with  $w_i = 0$  and  $w_{i+1} \neq 0$ . We substitute  $x_i = x_{i+1} = x$  in  $P$ . Looking at the monomial

$$x_1^{w_1} \cdots x_{i-1}^{w_{i-1}} x^{w_{i+1}} x_{i+2}^{w_{i+2}} \cdots$$

that does not appear on the right-hand side of (1), we get a linear relation between  $c_{\mathbf{w}}$  and the coefficients  $c_{\mathbf{v}}$  of the monomials  $X^{\mathbf{v}}$  such that  $\ell(X^{\mathbf{v}}) < \ell(X^{\mathbf{w}})$ , whence the upper bound on the dimension.

Now, for each composition  $I$  of  $n$ , the stable polynomial  $M_I(\mathbb{X})$  satisfies Equation (1). Besides, all  $M_I(\mathbb{X})$  are linearly independent since, setting  $x_{2i+1} = 0$  in  $\mathbb{X}$  transforms  $M_I(\mathbb{X})$  into the usual monomial quasi-symmetric functions in even-indexed variables  $M_I(x_2, x_4, x_6, \dots)$ . ■

**Example 3.2.** Here are the functions  $M_I(\mathbb{X})$  for compositions  $I$  of length at most 3:

$$(21) \quad M_{(k)}(\mathbb{X}) = -x_1^k + x_2^k - x_3^k + x_4^k - \dots;$$

$$(22) \quad M_{(k,\ell)}(\mathbb{X}) = \sum_i x_{2i+1}^{k+\ell} + \sum_{i < j} (-1)^{i+j} x_i^k x_j^\ell;$$

$$(23) \quad M_{(k,\ell,m)}(\mathbb{X}) = - \sum_i x_{2i+1}^{k+\ell+m} + \sum_{\substack{i,j \\ i < 2j+1}} (-1)^i x_i^k x_{2j+1}^{\ell+m} \\ + \sum_{\substack{i,j \\ 2i+1 < j}} (-1)^j x_{2i+1}^{k+\ell} x_j^m + \sum_{h < i < j} (-1)^{h+i+j} x_h^k x_i^\ell x_j^m$$

At least in the first two formulas, it is not hard to check that putting  $x_p = x_{p+1}$  eliminates these variables.



4. A NEW ALGEBRA OF FUNCTIONS ON YOUNG DIAGRAMS

4.1. **Interpreting elements of  $\mathcal{S}$  as functions on Young diagrams.** Recall from the introduction that the interlacing coordinates define a map  $\mathcal{IC}$  from Young diagrams to finite sequences of integers. Besides, a polynomial in infinitely many variables (and in particular the elements of  $\mathcal{S}$ ) can be evaluated on any finite sequence. Therefore, if  $\mathcal{F}(\mathcal{Y}, \mathbb{C})$  denotes the algebra of complex-valued functions on the set  $\mathcal{Y}$  of all Young diagrams, one has a mapping:

$$\text{eval}_{\mathcal{Y}} : \begin{array}{l} \mathcal{S} \rightarrow \mathcal{F}(\mathcal{Y}, \mathbb{C}) \\ f \mapsto f \circ \mathcal{IC} \end{array} .$$

**Proposition 4.1.** *The kernel of  $\text{eval}_{\mathcal{Y}}$  is the ideal  $\langle M_1(\mathbb{X}) \rangle$  generated by  $M_1(\mathbb{X}) = \sum_i (-1)^{i+1} x_i$ .*

*Proof.* It is known that the alternating sum of the interlacing coordinates is always zero (see, e.g., [11, Proposition 2.4]), which means that the function  $M_1(\mathbb{X})$  is in the kernel of  $\text{eval}_{\mathcal{Y}}$ .

For the converse statement, we need the following lemma.

**Lemma 4.2.** *Let  $f = (f_n)_{n \geq 0}$  be a stable polynomial in  $\mathbb{C}[X]$ . There exists a unique stable polynomial  $f'$  such that for any  $n$ ,*

$$(24) \quad f_n(x_1, \dots, x_n) = f'_n \left( \sum_{i=1}^n (-1)^i x_i, x_2, \dots, x_n \right) .$$

We denote by  $\Phi_{x \rightarrow \Sigma}$  the map  $f \mapsto f'$ .

*Proof of the Lemma.* It is clear that, for a fixed integer  $n$ , the polynomial  $f'_n$  is uniquely determined by Equation (24):

$$f'_n(x'_1, \dots, x'_n) = f \left( \sum_{i=1}^n (-1)^i x'_i - x'_1, x'_2, \dots, x'_n \right) .$$

It remains to check that the sequence  $(f'_n)_{n \geq 0}$  defines a stable polynomial, which is straightforward. ■

*End of the proof of the proposition.* Let  $f$  be a stable polynomial in the kernel of  $\text{eval}_{\mathcal{Y}}$ . Fix some integer  $m$  and a decreasing sequence  $x_2 > x_3 > \dots > x_{2m+1}$  of negative integers. Set

$$x_1 = x_2 - x_3 + x_4 - \dots + x_{2m} - x_{2m+1} .$$

Then  $x_1$  is a positive integer, hence  $x_1 > x_2$  and, as its alternating sum vanishes,  $(x_1, \dots, x_{2m+1})$  is the list of interlacing coordinates of some Young diagram. Thus,

$$f(x_1, x_2, x_3, \dots, x_{2m+1}) = 0 .$$

Let  $f' = \Phi_{x \rightarrow \Sigma}(f)$ . By Proposition 4.1,

$$f'(0, x_2, x_3, \dots, x_{2m+1}) = f(x_1, x_2, x_3, \dots, x_{2m+1})$$

and thus, it vanishes. This is true for all decreasing lists  $(x_2, \dots, x_{2m+1})$  of negative integers, so the polynomial  $f'(0, x_2, x_3, \dots, x_{2m+1})$  is identically zero. In other terms, there exists some polynomial  $g'_{2m+1} \in \mathbb{C}[x_1, \dots, x_{2m+1}]$  such that

$$f'_{2m+1} = x'_1 \cdot g'_{2m+1}, \text{ or equivalently } f_{2m+1} = \left( \sum_{i=1}^n (-1)^i x_i \right) \cdot g_{2m+1}$$

for some polynomial  $g_{2m+1} \in \mathbb{C}[x_1, \dots, x_{2m+1}]$ . It is straightforward to check that  $(g_{2m+1})_{m \geq 0}$  defines a stable polynomial and hence

$$f = \left( \sum_{i=1}^n (-1)^i x_i \right) \cdot g$$

for some  $g \in \mathbb{C}[X]$ , which is what we wanted to prove. ■

The image of  $\text{eval}_y$  is a subalgebra of functions on Young diagrams, that we shall call from now on *quasi-symmetric functions on Young diagrams* and denote by  $Q\Lambda$ . By definition, it is isomorphic to  $QSym/\langle M_1 \rangle$  and contains the algebra  $\Lambda = \text{eval}_y(\text{Sym}(\mathbb{X}))$  of *symmetric functions on Young diagrams* considered by Kerov and Olshanski.

**Corollary 4.3.** *The homogeneous component of degree  $n \geq 2$  of  $Q\Lambda$  has dimension  $2^{n-2}$ .*

**4.2. Multirectangular coordinates.** We consider here a different set of coordinates for Young diagrams, called *multirectangular coordinates* and introduced<sup>3</sup> by R. Stanley in [23].

Consider two sequences  $\mathbf{p}$  and  $\mathbf{q}$  of non-negative integers of the same length  $m$ . We associate with these the Young diagram drawn on Figure 2. Note that we allow some  $p_i$  or some  $q_i$  to be zero, so that the same diagram can correspond to several sequences (see again Figure 2). Nevertheless, if we require the variables  $p_i$  and  $q_i$  to be positive, there is a unique way to associate with a diagram some multirectangular coordinates. This defines a mapping:

$$\mathcal{MC} : \{\text{Young diagrams}\} \rightarrow \left\{ \begin{array}{l} \text{pairs of integer sequences} \\ \text{of the same length} \end{array} \right\}.$$

The multirectangular coordinates are related to interlacing coordinates by the following changes of variables: for all  $i \leq m$ ,

$$(25) \quad \begin{cases} p_i = x_{2i-1} - x_{2i}; \\ q_i = x_{2i} - x_{2i+1}; \end{cases} \quad \begin{cases} x_{2i+1} = (q_{i+1} + \dots + q_m) - (p_1 + \dots + p_i); \\ x_{2i} = (q_i + \dots + q_m) - (p_1 + \dots + p_i). \end{cases}$$

It should not be surprising that there are only  $2m$  rectangular coordinates while there are  $2m + 1$  interlacing coordinates because the latter must satisfy the linear relation  $\sum_i (-1)^i x_i = 0$ .

---

<sup>3</sup>In fact, R. Stanley considered coordinates  $\mathbf{p}'$  and  $\mathbf{q}'$  related to ours by  $p'_i = p_i$  and  $q'_i = q_1 + \dots + q_i$ . However, for our purpose, we prefer the more symmetric version presented here.

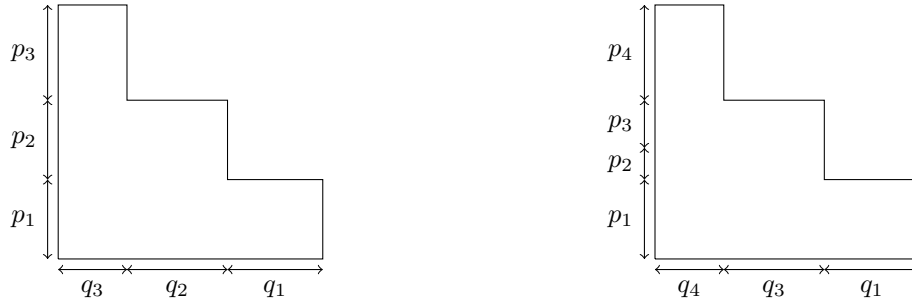


FIGURE 2. The same Young diagram given by two sets of multirectangular coordinates (in the right-hand side,  $q_2 = 0$ ).

Let us now consider functions on Young diagrams which are polynomials in its rectangular coordinates. This amounts to looking for polynomials in two infinite sets of variables

$$h \begin{pmatrix} p_1 & p_2 & \cdots \\ q_1 & q_2 & \cdots \end{pmatrix} \in \mathbb{C}[\mathbf{p}, \mathbf{q}]$$

satisfying the following two equations: for all  $1 \leq i \leq m$ ,

$$(26) \quad h_m \begin{pmatrix} p_1 & \cdots & p_m \\ q_1 & \cdots & q_m \end{pmatrix} \Big|_{q_i=0} = h_{m-1} \begin{pmatrix} p_1 & \cdots & p_{i-1} & p_i + p_{i+1} & \cdots & p_m \\ q_1 & \cdots & q_{i-1} & q_{i+1} & \cdots & q_m \end{pmatrix};$$

$$(27) \quad h_m \begin{pmatrix} p_1 & \cdots & p_m \\ q_1 & \cdots & q_m \end{pmatrix} \Big|_{p_i=0} = h_{m-1} \begin{pmatrix} p_1 & \cdots & p_{i-1} & p_{i+1} & \cdots & p_m \\ q_1 & \cdots & q_{i-1} + q_i & q_{i+1} & \cdots & q_m \end{pmatrix}.$$

In both cases  $i = 1$  and  $i = m$ , erase the column containing non-defined variables.

All these equations express the fact that the images of two sequences corresponding to the same Young diagram (as the ones of Figure 2) have the same image by  $h$ . We denote by  $\mathcal{S}'$  the space of solutions of this system.

**4.3. Link between the two systems.** In this Section, we study the relation between the two spaces of solutions  $\mathcal{S}$  and  $\mathcal{S}'$ .

Consider an element  $f = (f_n)_{n \geq 0}$  in  $\mathcal{S}$ . Let  $m \geq 0$ . Replace all variables  $x_1, \dots, x_{2m+1}$  in  $f_{2m+1}$  according to (25), and set

$$(28) \quad h_m \begin{pmatrix} p_1 & \cdots & p_m \\ q_1 & \cdots & q_m \end{pmatrix} = f_{2m+1}(x_1, \dots, x_{2m+1}).$$

Clearly,  $h_m$  is a polynomial in  $p_1, \dots, p_m, q_1, \dots, q_m$ .

Besides, by definition,

$$h_{m+1} \begin{pmatrix} p_1 & \cdots & p_m & 0 \\ q_1 & \cdots & q_m & 0 \end{pmatrix} = f_{2m+3}(x_1, \dots, x_{2m+1}, x_{2m+1}, x_{2m+1}).$$

But, as  $f$  is an element of  $\mathcal{S}$ , the right-hand side is equal to  $f_{2m+1}(x_1, \dots, x_{2m+1})$ . This implies that  $(h_m)_{m \geq 0}$  is an element of  $\mathbb{C}[\mathbf{p}, \mathbf{q}]$ .

We will now show that  $h$  satisfies Equation (26). Let us now consider an integer  $m \geq 1$  and variables  $p_1, \dots, p_m, q_1, \dots, q_m$ . Assume additionally that  $q_i = 0$  for some  $i$ , which implies  $x_{2i} = x_{2i+1}$ . Thus, as  $f$  is an element of  $\mathcal{S}$ ,

$$f_{2m+1}(x_1, \dots, x_{2m+1}) = f_{2m-1}(x_1, \dots, x_{2i-1}, x_{2i+2}, \dots, x_{2m+1}).$$

Observe that the right-hand side corresponds to the definition of

$$h_{m-1} \begin{pmatrix} p_1 & \cdots & p_{i-1} & p_i + p_{i+1} & \cdots & p_m \\ q_1 & \cdots & q_{i-1} & q_{i+1} & \cdots & q_m \end{pmatrix},$$

which ends the proof of Equation (26). Equation (27) can be proved in a similar way.

Finally, from a stable polynomial  $f$  in  $\mathcal{S}$ , we have constructed an element  $h = (h_m)_{m \geq 0}$  in  $\mathcal{S}'$ . We denote by  $\Phi_{x \rightarrow p, q}$  this map (from  $\mathcal{S}$  to  $\mathcal{S}'$ ).

A map  $\Phi_{p, q \rightarrow x}$  from  $\mathcal{S}'$  to  $\mathcal{S}$  can be constructed in a similar way using the first part of Equation (25).

**Lemma 4.4.** *One has  $\Phi_{x \rightarrow p, q} \circ \Phi_{p, q \rightarrow x} = \text{Id}_{\mathcal{S}'}$ . Moreover,  $\Phi_{x \rightarrow p, q}$  has the same kernel as  $\text{eval}_y$ , that is  $\langle M_1(\mathbb{X}) \rangle$ .*

*Proof.* Fix  $h \in \mathcal{S}'$ . Let  $f = \Phi_{p, q \rightarrow x}(h)$  and  $\tilde{h} = \Phi_{x \rightarrow p, q}(f)$ . By definition,

$$\tilde{h}_m \begin{pmatrix} \tilde{p}_1 & \cdots & \tilde{p}_m \\ \tilde{q}_1 & \cdots & \tilde{q}_m \end{pmatrix} = f(x_1, \dots, x_{2m+1}),$$

where  $x_1, \dots, x_{2m+1}$  are defined in terms of  $\tilde{p}_1, \dots, \tilde{p}_m, \tilde{q}_1, \dots, \tilde{q}_m$  using (25). But

$$f(x_1, \dots, x_{2m+1}) = h_m \begin{pmatrix} p_1 & \cdots & p_m \\ q_1 & \cdots & q_m \end{pmatrix},$$

where  $p_1, \dots, p_m, q_1, \dots, q_m$  are defined in terms of  $x_1, \dots, x_{2m+1}$  using (25). Applying both relations (25) *in that order*, one sees directly that for all  $i \leq m$ , one has  $p_i = \tilde{p}_i$  and  $q_i = \tilde{q}_i$ . Hence  $\tilde{h}_m$  and  $h_m$  are the same polynomial, which proves the first part of the Lemma.

For the second one, first observe that  $\Phi_{x \rightarrow p, q}(M_1) = 0$ , thus  $\langle M_1(\mathbb{X}) \rangle \subset \text{Ker}(\text{eval}_y)$ . Let us prove the opposite inclusion. Consider a stable polynomial  $f$  such that  $\Phi_{x \rightarrow p, q}(f) = 0$ . This implies that  $f$  vanishes on all lists of interlacing coordinates of Young diagrams, that is  $\text{eval}_y(f) = 0$ . In other terms,  $\text{Ker}(\Phi_{x \rightarrow p, q}) \subset \text{Ker}(\text{eval}_y)$ . The other implication is easy, as we already know by Proposition 4.1 that  $\text{Ker}(\text{eval}_y) = \langle M_1(\mathbb{X}) \rangle$ .  $\blacksquare$

**Corollary 4.5.** *The composition  $\text{eval}_y' = \text{eval}_y \circ \Phi_{p, q \rightarrow x}$  defines an isomorphism from  $\mathcal{S}'$  to  $Q\Lambda$ .*

*Proof.* Thanks to the previous Lemma,  $\Phi_{p, q \rightarrow x}$  is injective and by definition of  $Q\Lambda$ , the map  $\text{eval}_y$  is surjective. It remains to prove that  $\text{Im}(\Phi_{p, q \rightarrow x})$  and  $\text{Ker}(\text{eval}_y)$  are complementary subspaces.

As  $\Phi_{x \rightarrow p, q} \circ \Phi_{p, q \rightarrow x} = \text{Id}_{\mathcal{S}'}$ , the composition  $\Phi_{p, q \rightarrow x} \circ \Phi_{x \rightarrow p, q}$  is a projector. Besides,  $\text{Im}(\Phi_{p, q \rightarrow x} \circ \Phi_{x \rightarrow p, q}) = \text{Im}(\Phi_{p, q \rightarrow x})$  and  $\text{Ker}(\Phi_{p, q \rightarrow x} \circ \Phi_{x \rightarrow p, q}) = \text{Ker}(\Phi_{x \rightarrow p, q}) = \text{Ker}(\text{eval}_y)$  (the last equality is the second statement of the previous Lemma). This

ends the proof, as the image and kernel of a projector are complementary subspaces. ■

The following diagram summarizes the morphisms considered so far:

$$(29) \quad \begin{array}{ccc} \mathcal{S} & \xrightarrow{\text{evaly}} & Q\Lambda \\ & \searrow & \nearrow \text{evaly}' \\ & \mathcal{S}' & \end{array}$$

The two arrows between  $\mathcal{S}$  and  $\mathcal{S}'$  correspond respectively to  $\Phi_{x \rightarrow p, q}$  and  $\Phi_{p, q \rightarrow x}$ .

**4.4. The basis  $H_I$ .** In this Section, we exhibit a basis of  $\mathcal{S}'$  that has an explicit expression in terms of multirectangular coordinates. It is also related to some basis of  $QSym$ , whose product is given by the shuffle of the parts of the compositions indexing the elements, and which can be recognized as the one introduced by C. Malvenuto and C. Reutenauer in [17].

It is here more convenient to work with the original multirectangular coordinates, as defined by R. Stanley: set  $q'_i = q_i + q_{i+1} + \dots$  for all  $i \geq 1$ . With this change of variables, stable polynomials in the  $\mathbf{q}$  and  $\mathbf{q}'$  are the same, (26) remains the same too, and (27) becomes

$$(30) \quad h_m \left( \begin{array}{ccc} p_1 & \cdots & p_m \\ q'_1 & \cdots & q'_m \end{array} \right) \Big|_{p_i=0} = h_{m-1} \left( \begin{array}{ccccc} p_1 & \cdots & p_{i-1} & p_{i+1} & \cdots & p_m \\ q'_1 & \cdots & q'_{i-1} & q'_{i+1} & \cdots & q'_m \end{array} \right)$$

Let  $I$  be a composition of  $n$  with last part greater than 1. We define

$$(31) \quad H_I \left( \begin{array}{ccc} p_1 & p_2 & \cdots \\ q_1 & q_2 & \cdots \end{array} \right) = \sum_{s \geq 1} \sum_{\substack{I=I_1 \cdot I_2 \cdots I_s \\ k_1 < \cdots < k_s}} \prod_t \frac{p_{k_t}^{\ell(I_t)} (q'_{k_t})^{|I_t| - \ell(I_t)}}{\ell(I_t)!}.$$

The summation index  $I = I_1 \cdot I_2 \cdots I_s$  means that we consider all ways of writing  $I$  as a concatenation of  $s$  non-empty compositions. Here are a few examples (the arguments are omitted for readability): for  $v \geq 2$ ,

$$(32) \quad H_{(v)} = \sum_i p_i (q'_i)^{v-1}; \quad H_{(u,v)} = \sum_{i < j} p_i (q'_i)^{u-1} p_j (q'_j)^{v-1} + \frac{1}{2} \sum_i p_i^2 (q'_i)^{u+v-2}.$$

**4.4.1. The  $H_I$  belong to  $\mathcal{S}'$ .** Note that, for any composition  $I$  of length  $\ell$ ,

$$H_I = \sum_{k_1 < \cdots < k_\ell} p_{k_1} (q'_{k_1})^{i_1-1} \cdots p_{k_\ell} (q'_{k_\ell})^{i_\ell-1} + \text{non } p\text{-square free terms.}$$

In particular, the  $H_I$  are linearly independent.

**Proposition 4.6.** *The functions  $(H_I)$ , where  $I$  runs over compositions with last part greater than 1, form a basis of  $\mathcal{S}'$ .*

**Remark 4.7.** The proof given below uses the material of the previous Section. However, it would be possible to give a direct proof, by translating Equations (26) and (27) into equations on the coefficients of the unknown polynomial  $h$ . This approach would allow to guess the definition of the functions  $\mathbf{H}_I$  in a very natural way.

*Proof.* By Corollaries 4.3 and 4.5, the dimension of the homogeneous component of degree  $n$  is  $2^{n-2}$ , which is exactly the number of functions  $\mathbf{H}_I$  of degree  $n$  (they are indexed by compositions with last part greater than 1). As the functions  $(\mathbf{H}_I)_{I \neq n}$  are linearly independent, it is enough to prove that they indeed belong to  $\mathcal{S}'$ .

Fix a composition  $I$ . It is straightforward to see that  $(\mathbf{H}_I)$  satisfies (27). Indeed, all monomials that contain  $q'_i$  also contain  $p_i$ . Let us prove that  $(\mathbf{H}_I)$  satisfies (26).

Let  $\ell$  be the length of  $I$ . Assume that  $q_j = 0$  for some  $j < m$ , that is  $q'_j = q'_{j+1}$ . Consider Equation (26) for  $h = \mathbf{H}_I$  and rewrite both sides using Equation (31). The summands corresponding to a sequence of indices  $\mathbf{k}$  that does not contain  $j$  or  $j+1$  are the same on both sides. So let us consider some factorization  $f = (I_1, \dots, I_s)$  of  $I$  and some sequence  $\mathbf{k}$  such that  $k_t = j$  or  $k_t = j+1$  for some  $t$ .

When  $k_t = j$ , it may happen that in addition  $k_{t+1} = j+1$ . In this case, denote by  $\bar{f}$  the factorization  $(I_1, \dots, I_{t-1}, I_t \cdot I_{t+1}, \dots, I_s)$ , otherwise set  $\bar{f} = f$ . We consider in  $\mathbf{H}_I$  the summands corresponding to factorizations  $f$  sent to a given factorization  $\bar{f} = (\bar{I}_1, \dots, \bar{I}_s)$ . We get (recall that we set  $q'_{j+1} = q'_j$ ):

$$M(q'_j)^{|\bar{I}_t| - \ell(\bar{I}_t)} \cdot \left( \frac{p_j^{\ell(\bar{I}_t)}}{\ell(\bar{I}_t)!} + \sum_{\substack{r_1+r_2=\ell(\bar{I}_t) \\ r_1, r_2 \geq 1}} \frac{p_j^{r_1} p_{j+1}^{r_2}}{r_1! r_2!} + \frac{p_{j+1}^{\ell(\bar{I}_t)}}{\ell(\bar{I}_{t+1})!} \right),$$

where  $M$  is some monomial in  $p_k$  and  $q'_k$  ( $k \neq j, j+1$ ). The first (resp. last) term corresponds to the case where  $f = \bar{f}$  and  $k_t = j$  (resp.  $k_t = j+1$ ). The sum in the middle corresponds to the cases where  $k_t = j$  and  $k_{t+1} = j+1$  and where  $\bar{f}$  is obtained from  $f$  by gluing two non-trivial factors of respective lengths  $r_1$  and  $r_2$ .

This expression simplifies *via* Newton's binomial formula to

$$M(q'_j)^{|\bar{I}_t| - \ell(\bar{I}_t)} \frac{(p_j + p_{j+1})^{\ell(\bar{I}_t)}}{\ell(\bar{I}_{t+1})!},$$

which is the term corresponding to the factorization  $\bar{f}$  in the right-hand side of (26).

Finally, we get that  $\mathbf{H}_I$  satisfies also (26) for  $j < m$ . The case  $j = m$  must be treated separately, as the column containing  $p_j + p_{j+1}$  in (26) does not exist. However, in this case, it is immediate to check that  $\mathbf{H}_I$  satisfies (26) if the last part of  $I$  is greater than 1. So  $\mathbf{H}_I$  lies in  $\mathcal{S}'$ . ■

4.4.2. *Some expansions on the basis  $\mathbf{H}_I$ .* The fact that the  $p$ -square free terms of the  $\mathbf{H}_I$  are distinct monomials helps to compute the expansion of a given function on the  $\mathbf{H}$ -basis. Here are a few examples.

As  $\mathcal{S}'$  is an algebra (the product of two solutions of (26) and (27) is still a solution of these equations), the product of two  $H_I$  is a linear combination of  $H_J$ . It turns out to have a very simple description.

**Proposition 4.8.** *For any two compositions  $I$  and  $J$  (with their last part greater than 1),*

$$H_I \cdot H_J = \sum_{K \in I \sqcup J} H_K,$$

where  $I \sqcup J$  denotes the multiset of compositions obtained by shuffling the parts of  $I$  and  $J$  (see Section 2.2).

As this result has already appeared (under a different form) in the literature – see Section 4.4.3 – and as we do not use it in this paper, we only sketch its proof.

*Sketch of proof.* Both sides of the equality lie in  $\mathcal{S}'$  and share the same  $p$ -square free terms. As the  $H_I$  form a basis of  $\mathcal{S}'$  and have linearly independent  $p$ -square free terms, this is enough to conclude. ■

It is also possible to obtain the  $H$ -expansions of the functions  $\Phi_{x \rightarrow p, q}(M_k(\mathbb{X}))$ , which are generators of  $\Lambda$ .

**Proposition 4.9.** *For any  $k \geq 2$ , one has*

$$\Phi_{x \rightarrow p, q}(M_k(\mathbb{X})) = \sum_{i=0}^{k-2} (-1)^{i+1} k(k-1) \cdots (k-i+1) H_{1^i, k-i}.$$

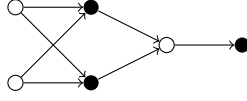
*Sketch of proof.* A lengthy but straightforward computation shows that both sides share the same  $p$ -square free terms. As both sides lie in  $\mathcal{S}'$ , this is enough to conclude. ■

Another interesting family of functions on Young diagrams is the one that motivated the study of functions of Young diagram: normalized character values, cf. [12]. A combinatorial interpretation for the coefficients of these functions written in terms of coordinates  $\mathbf{p}$  and  $\mathbf{q}'$  is given in [5]. Using the material here, we can directly deduce from it a combinatorial interpretation for the coefficients of the  $H$ -expansion of these renormalized character values.

4.4.3. *Collapsing the alphabets.* Consider now the morphism which sends  $p_j$  and  $q'_j$  to the same variable  $y_j$ . The image of  $H_I$  is the following quasi-symmetric function of the alphabet  $Y$

$$(33) \quad \sum_{s \geq 1} \sum_{I=I_1 \cdot I_2 \cdots I_s} \prod_t \frac{1}{\ell(I_t)!} M_{|I_1|, |I_2|, \dots, |I_s|}(Y).$$

Call  $H_I(Y)$  this function. Then,  $H_I$ , when  $I$  runs over all compositions, form a basis of  $QSym$  since the transition matrix with the monomial basis is triangular. Proposition 4.8 implies that their product is given by the shuffle operation on the parts of the compositions.

FIGURE 3. Example of an unlabelled bipartite graph  $G$ .

These functions already appear in [17], Equation (2.12), where they are defined as the dual basis of some noncommutative symmetric functions denoted by  $\Phi^I$  in [8] (analogues of power sums). The shuffle property is clear in this context, as the  $\Phi^I$  form a multiplicative basis on primitive generators.

**4.5. The functions  $N_G$ .** In this Section, we study a family of functions on Young diagrams indexed by bipartite graphs with two types of edges.

Let  $G$  be an unlabelled bipartite graph with vertex set  $V = V_1 \sqcup V_2$  and edge set  $E = E_{1,2} \sqcup E_{2,1} \subset V_1 \times V_2$ . We consider the polynomial  $N_G$  in the variables  $p_i$  and  $q_i$  defined as follows:

$$N_G \begin{pmatrix} p_1 & p_2 & \cdots \\ q_1 & q_2 & \cdots \end{pmatrix} = \sum_r \left( \prod_{v_1 \in V_1} p_{r(v_1)} \prod_{v_2 \in V_2} q_{r(v_2)} \right),$$

where the sum runs over functions  $r : V \rightarrow \mathbb{N}$  satisfying the following *order condition* (the large and strict inequalities are important!):

- for each edge  $e = (v_1, v_2)$  in  $E_{1,2}$ , one has  $r(v_1) \leq r(v_2)$  ;
- for each edge  $e = (v_1, v_2)$  in  $E_{2,1}$ , one has  $r(v_2) < r(v_1)$ .

**Example 4.10.** Consider the graph  $G_{\text{ex}}$  drawn on Figure 3. Vertices in  $V_1$  (resp.  $V_2$ ) are drawn in white (resp. black). edges in  $E_{1,2}$  (resp.  $E_{2,1}$ ) are represented directed from their extremity in  $V_1$  to their extremity in  $V_2$  (resp. from their extremity in  $V_2$  to their extremity in  $V_1$ ).

Let  $r$  be a function from its vertex set to  $\mathbb{N}$ . Denote by  $e$  and  $f$  the images of the leftmost white vertices, by  $g$  and  $h$  the images of the black vertices just to their right, then by  $i$  the image of the white vertex to their right and finally by  $j$  the image of the rightmost black vertex.

Then, by definition,  $r$  satisfies the order condition if and only if

$$e, f \leq g, h < i \leq j.$$

Note the alternating large and strict inequalities. Finally, one has

$$N_{G_{\text{ex}}} \begin{pmatrix} p_1 & p_2 & \cdots \\ q_1 & q_2 & \cdots \end{pmatrix} = \sum_{e, f \leq g, h < i \leq j} p_e p_f p_i q_g q_h q_j.$$

**Remark 4.11** (Why are we considering this family ?). In the case where we have only edges of the first type, that is  $E_{2,1} = \emptyset$ , the functions have played an important role in some recent works on irreducible character values of symmetric groups; see [6, 4, 2] and in particular [4, Definition 2.2.1]. The extension to the case  $E_{2,1} \neq \emptyset$  will be useful in Section 6.4.



**Lemma 4.12.** *Let  $G$  be a bipartite graph as above. Assume that each element in  $V$  is the extremity of at least one edge in  $E_{1,2}$ . Then the polynomial  $N_G$  belongs to  $\mathcal{S}'$ .*

*Proof.* Let us check that  $N_G$  satisfies equation (26). We define

$$\begin{cases} p'_j = p_j & \text{if } j < i; \\ p'_i = p_i + p_{i+1}; \\ p'_j = p_{j+1} & \text{if } j > i; \end{cases} \quad \begin{cases} q'_j = q_j & \text{if } j < i; \\ q'_j = q_{j+1} & \text{if } j \geq i. \end{cases}$$

These are the variables in the right-hand side of (26). Consider first the left-hand side:

$$N_G \left( \begin{array}{ccc} p_1 & \cdots & p_m \\ q_1 & \cdots & q_m \end{array} \right) \Big|_{q_i=0} = \sum_r \left( \prod_{v_1 \in V_1} p_{r(v_1)} \prod_{v_2 \in V_2} q_{r(v_2)} \right),$$

where the sum runs over functions  $r : V \rightarrow \{1, \dots, m\}$  satisfying the order condition. Besides, one can restrict the sum to functions  $r$  such that  $r(v_2) \neq i$  for any  $v_2 \in V_2$  (we call them  $i$ -avoiding functions).

If  $i = m$ , an  $i$ -avoiding function  $r$  which satisfies the order condition cannot associate  $m$  with a vertex in  $V_1$  (indeed, as each vertex  $v_1$  in  $V_1$  is the extremity of at least one edge  $(v_1, v_2)$  in  $E_{1,2}$ , so that  $r(v_1) \leq r(v_2) < m$ ). Therefore equation (26) is satisfied for  $i = m$ .

Let us consider now the case  $i < m$ . With any function  $r$ , we associate a function  $r' = \Phi(r) : V \rightarrow \{1, \dots, m-1\}$  defined as follows:

$$r'(v) = r(v) \text{ if } r(v) \leq i \text{ and } r'(v) = r(v) - 1 \text{ if } r(v) > i.$$

It is straightforward to check that, if  $r$  is  $i$ -avoiding,  $r'$  satisfies the order condition. Indeed, the only problem which could occur is that  $r(v_2) = i$  and  $r(v_1) = i+1$  for an edge  $(v_1, v_2)$  in  $E_{2,1}$ , but we forbid  $r(v_2) = i$ .

The preimage of a given function  $r'$  is obvious: it is the set of functions  $r$  with

$$\begin{cases} r(v) = r'(v) & \text{if } r'(v) < i; \\ r(v) \in \{i, i+1\} & \text{if } r'(v) = i; \\ r(v) = r'(v) + 1 & \text{if } r'(v) > i. \end{cases}$$

If  $r'$  satisfies the order condition, all its  $i$ -avoiding pre-images  $r$  also satisfy the order condition. Once again, the only obstruction to this would be in the case when  $r'(v_1) = r'(v_2) = i$  for some edge  $(v_1, v_2)$  in  $E_{1,2}$ . In this case, preimages  $r$  with  $r(v_1) = i+1$  and  $r(v_2) = i$  would not satisfy the order condition but we forbid  $r(v_2) = i$ .

Now, using the above description of the preimage, for any function  $r'$ , one has:

$$\sum_{r \in \Phi^{-1}(r')} \left( \prod_{v_1 \in V_1} p_{r(v_1)} \prod_{v_2 \in V_2} q_{r(v_2)} \right) = \left( \prod_{v_1 \in V_1} p'_{r'(v_1)} \prod_{v_2 \in V_2} q'_{r'(v_2)} \right),$$

Summing over all functions  $r' : V \rightarrow \{1, \dots, m-1\}$  with the order condition, we get equality (26).

The proof of (27) is similar. ■

**Remark 4.13.** In [6, Section 1.5], an equivalent definition of  $N_G$  as a function on Young diagrams is given in the case  $E_{2,1} = \emptyset$ . The sole fact that  $N_G$  can be defined using only the Young diagram and not its rectangular coordinates explains that it belongs to  $\mathcal{S}'$ .

As  $N_G$  belongs to  $\mathcal{S}'$ , its image  $\Phi_{p,q \rightarrow x}(N_G)$  is an element of  $\mathcal{S}$ , that is, some quasi-symmetric function evaluated on the virtual alphabet  $\mathbb{X}$ . We shall now determine this quasi-symmetric function.

Define  $F_G$  as the generating series of functions on  $G$  satisfying the order condition:

$$F_G(u_1, u_2, \dots) = \sum_{\substack{r: V \rightarrow \mathbb{N} \\ \text{with order condition}}} \prod_{v \in V} u_{r(v)}.$$

Note that  $F_G$  is a quasi-symmetric function.

**Proposition 4.14.** *For any bipartite graph  $G$  with vertex set  $V = V_1 \sqcup V_2$ ,*

$$N_G \begin{pmatrix} p_1 & p_2 & \cdots \\ q_1 & q_2 & \cdots \end{pmatrix} = (-1)^{|V_1|} \Phi_{x \rightarrow p, q}(F_G(\mathbb{X})).$$

*Proof.* The stable polynomial  $\Phi_{p,q \rightarrow x}(N_G)$  is an element of  $\mathcal{S}$ , that is  $F(\mathbb{X})$  for some quasi-symmetric function  $F$ . To identify  $F$ , we shall send all odd-indexed variables  $x_{2i+1}$  to 0. This amounts to sending  $p_i$  to  $-x_{2i}$  and  $q_i$  to  $x_{2i}$ . Thus we get that

$$F(x_2, x_4, \dots) = (-1)^{|V_1|} F_G(x_2, x_4, \dots).$$

This implies  $F = (-1)^{|V_1|} F_G$  and the proposition follows by applying  $\Phi_{x \rightarrow p, q}$ .  $\blacksquare$

**Remark 4.15.** We have just shown that the series  $N_G$  in two sets of variables can be recovered from the corresponding series  $F_G$  in one set of variables. This is a bit surprising and it is quite remarkable that Young diagrams give a natural way of doing this.

**Remark 4.16.** In the above proof, to identify  $F$ , we could also have sent all even-indexed variables  $x_{2i}$  to zero. Under this specialization,

$$F(\mathbb{X}) = F(\ominus(x_1) \ominus (x_3) \ominus \dots) = F(\ominus(\dots, x_3, x_1)) = S(F)(\dots, x_3, x_1),$$

*i.e.* the antipode of  $F$  evaluated in the *reverted* alphabet  $(\dots, x_3, x_1)$ .

On the other side  $\Phi_{p,q \rightarrow x}(N_G)$ , when  $x_{2i} = 0$ , consists in plugging  $p_{i+1} = -q_i = x_{2i+1}$  in  $N_G$ . Thus we get, up to a sign  $(-1)^{|V_2|}$ , the generating series of functions  $r$  on the graph  $G$  verifying some modified order condition: just take the order condition above and exchange large and strict inequalities (this modification of type of equalities comes from the above index shift).

Finally, as  $F = (-1)^{|V_1|} F_G$ , we get that  $S(F_G)$  is, up to a sign  $(-1)^{|V|}$ , the generating series  $F_{G'}$  where  $G'$  is obtained from  $G$  by exchanging the sets  $E_{1,2}$  and  $E_{2,1}$ .

This result can be extended to general directed graphs with two types of edges (here, we only considered the "bipartite" case) and is a direct consequence of the fundamental lemma on  $P$ -partitions [24, Theorem 6.2] and of the image of a fundamental quasi-symmetric function by the antipode [17, Corollary 2.3].

5. NONCOMMUTATIVE GENERALIZATION

Let now  $A$  be a noncommutative ordered alphabet. We shall find all noncommutative polynomials  $P(A)$  satisfying the noncommutative version of Equation (1), that is, Equation (5), reproduced here for convenience:

$$P(a_1, a_2, \dots)|_{a_i=a_{i+1}} = P(a_1, \dots, a_{i-1}, a_{i+2}, \dots).$$

As in the commutative framework, homogeneous polynomials in infinitely many variables are formally sequences of homogeneous polynomials  $(R_n)_{n \geq 1}$ , where  $R_n$  lies in the algebra  $\mathbb{C}\langle a_1, \dots, a_n \rangle$  (polynomials in  $n$  noncommuting variables) such that  $R_{n+1}(a_1, \dots, a_n, 0) = R_n(a_1, \dots, a_n)$ . The vector space of these stable noncommutative polynomials will be denoted  $\mathbb{C}\langle A \rangle$ .

**5.1. Word quasi symmetric functions.** The natural noncommutative analogue of  $QSym$  is the algebra of *word quasi symmetric functions*, denoted by **WQSym**. We recall here its construction (see, e.g., [18]).

Monomials in the noncommutative framework are canonically indexed by finite words  $w$  on the alphabet  $\mathbb{N}$  as follows

$$\mathbf{a}_w = a_{w_1} a_{w_2} \dots a_{w_{|w|}}.$$

The *evaluation* of a word  $w$  is the integer sequence  $v = (v_1, v_2, \dots)$ , where  $v_i$  is the number of letters  $i$  in  $w$ . Then the commutative image of  $\mathbf{a}_w$  is  $\mathbf{X}^{\text{eval}(w)}$ .

In the noncommutative framework, set compositions play the role of compositions. A *set composition* of  $n$  is an (*ordered*) list  $(I_1, \dots, I_p)$  of pairwise disjoint non-empty subsets of  $\{1, \dots, n\}$ , whose union is  $\{1, \dots, n\}$ .

Such an object can be encoded by a word  $w$  of length  $n$  defined as follows:  $w_i = j$  if  $i \in I_j$ . This encoding is injective. Words obtained that way are exactly those satisfying the following property: for  $j \geq 1$ , if the letter  $j + 1$  appears in  $w$ , then  $j$  appears in  $w$ . Such words are called *packed*. Equivalently, a word  $w$  is packed if and only if its evaluation  $\text{eval}(w)$  can be written as  $\mathbf{c}, 0, 0, \dots$ , where  $\mathbf{c}$  is a composition (that is, a vector of positive integers).

With each word  $w$  is associated a packed word, called *packing* of  $w$  and denoted  $\text{pack}(w)$ : replace all occurrences of the smallest letter appearing in  $w$  by 1 then, occurrences of the second smallest letter by 2 and so on. For example the packing of 364434 is 132212.

Depending on the point of view, it may be more convenient to use set compositions or packed words. In this section, we use packed words while the next one deals with set compositions.

By definition, **WQSym** is a subalgebra of the algebra of stable polynomials in noncommuting variables  $a_1, a_2, \dots$ . A basis of **WQSym** is given as follows:

$$P_u = \sum_{\substack{w \text{ s.t.} \\ \text{pack}(w)=u}} \mathbf{a}_w$$

Note that the commutative image of  $P_u$  is  $M_{\text{eval}(u)}$ .

Conversely, if  $u$  is a non-decreasing packed word,  $P_u$  is the unique noncommutative stable polynomial whose commutative image is  $M_{\text{eval}(u)}$  and in which all monomials

have letters in nondecreasing order. Now, for any packed word  $u$ , consider its non-decreasing rearrangement  $u^\uparrow$  and the smallest permutation  $\sigma_u$  in  $\mathfrak{S}_{|u|}$  sending  $u^\uparrow$  to  $u$ . Then, the polynomial  $P_u$  is obtained by letting  $\sigma_u$  act on all monomials of  $P_{u^\uparrow}$  (note that all monomials in  $P_{u^\uparrow}$  have  $|u|$  letters). These two properties characterize the elements  $P_u$ .

To finish, let us mention that the ordered Bell numbers  $OB(n)$  [22, A000670] count packed words, thus give the dimension of the homogeneous subspace of degree  $n$  of  $\mathbf{WQSym}$ .

**5.2.  $\mathbf{WQSym}$  and a virtual alphabet.** We want to make sense of the virtual alphabet

$$\mathbb{A} = \ominus(a_1) \oplus (a_2) \ominus (a_3) \oplus (a_4) \ominus \dots$$

We cannot define  $\ominus$  by means of the antipode of  $\mathbf{WQSym}$ , since it is not involutive.

Rather, we define  $\mathbf{WQSym}(\mathbb{A})$  as follows: if  $u$  is a nondecreasing packed word,  $P_u(\mathbb{A})$  is the noncommutative analogue of  $M_{\text{eval}(u)}(\mathbb{X})$  where  $x_k$  is replaced by  $a_k$  and all letters in any monomial of  $P_u(\mathbb{A})$  are in nondecreasing order. Now, for any  $u$ , the (stable) polynomial  $P_u(\mathbb{A})$  is obtained by letting  $\sigma_u$  act on all monomials of  $P_{u^\uparrow}(\mathbb{A})$ , where  $\sigma_u$  and  $u^\uparrow$  are defined as above.

Finally, for  $F$  in  $\mathbf{WQSym}$ , we define  $F(\mathbb{A})$  by linearity. For example,

$$(34) \quad P_{1^k}(\mathbb{A}) = -a_1^k + a_2^k - a_3^k + a_4^k - \dots;$$

$$(35) \quad P_{1^k 2^\ell}(\mathbb{A}) = \sum_i a_{2i+1}^{k+\ell} + \sum_{i<j} (-1)^{i+j} a_i^k a_j^\ell;$$

$$(36) \quad P_{112}(\mathbb{A}) = \sum_i a_{2i+1}^3 + \sum_{i<j} (-1)^{i+j} a_i a_i a_j;$$

$$(37) \quad P_{121}(\mathbb{A}) = \sum_i a_{2i+1}^3 + \sum_{i<j} (-1)^{i+j} a_i a_j a_i;$$

$$(38) \quad P_{211}(\mathbb{A}) = \sum_i a_{2i+1}^3 + \sum_{i<j} (-1)^{i+j} a_j a_i a_i;$$

It is not hard to check that plugging  $a_p = a_{p+1}$  for some  $p$  in the noncommutative polynomials above eliminates these variables. This property will be established for any packed word  $u$  in the next section.

Finally, we denote by  $\mathbf{WQSym}(\mathbb{A})$  the following subspace of  $\mathbb{C}\langle A \rangle$ :

$$\mathbf{WQSym}(\mathbb{A}) = \{P(\mathbb{A}), P \in \mathbf{WQSym}\}$$

We will see in next section that  $\mathbf{WQSym}(\mathbb{A})$  is an algebra and that  $F \mapsto F(\mathbb{A})$  is an isomorphism of algebras from  $\mathbf{WQSym}$  to  $\mathbf{WQSym}(\mathbb{A})$ .

### 5.3. Solution of Equation (5).

**Theorem 5.1.** *A polynomial  $P$  satisfies Equation (5) if and only if  $P$  belongs to  $\mathbf{WQSym}(\mathbb{A})$ .*

*Proof.* We introduce the following ring homomorphism

$$\begin{aligned} \phi : \mathbb{C}\langle A \rangle &\rightarrow \mathbb{C}[X]; \\ a_i &\mapsto x_i. \end{aligned}$$

We first prove that the dimension of the space of homogeneous polynomials in  $\mathbb{C}[A]$  of degree  $n$  which satisfy (5) is at most equal to  $OB(n)$  (ordered Bell number). We may observe that a word  $w \in A^*$  is packed iff  $\phi(w)$  is packed as a monomial in  $\mathbb{C}[X]$ . Let

$$P = \sum_{w \in A^*} c_w w$$

be a homogeneous polynomial of degree  $n$ , solution of (5). The goal is to prove that all the coefficients of  $P$  are determined by its coefficients on packed words. Let  $w$  be a non-packed word and  $(a_i, a_{i+1})$  be a pair of letters such that:  $a_i$  does not appear in  $w$ , and  $a_{i+1}$  does. We let  $a_i = a_{i+1} = a$  in  $P$ . By looking at the coefficient of the word obtained by replacing  $a_{i+1}$  by  $a$  in  $w$ , which does not appear on the right-hand side of (5), we get a linear relation between  $c_w$  and coefficients  $c_v$  of words  $v$  such that  $\ell(\phi(v)) < \ell(\phi(w))$  ( $\ell$  is defined in (20)), whence the upper bound on the dimension.

Let us now prove that, if  $P \in \mathbf{WQSym}(\mathbb{A})$ , it satisfies the functional Equation (5). First consider the case  $P = P_u(\mathbb{A})$ , with  $u$  nondecreasing. By definition,  $P_u$  contains only monomials with variables with *non-decreasing indices*. This is still true after substitution  $a_{i+1} = a_i$  as we equate two *consecutive* variables. Therefore any cancellation occurring in the commutative image, is mimicked in the noncommutative version. As  $M_{\text{eval}(u)}(\mathbb{X})$ , which is the commutative image of  $P_u(\mathbb{A})$ , is a solution of Equation (1), the noncommutative polynomial  $P_u(A)$  is a solution of (5).

Moreover, if it is true for a polynomial  $P$ , then it is true for any polynomial obtained by action of a permutation on it, so any element of  $\mathbf{WQSym}(\mathbb{A})$  satisfies (5). Besides, the  $P_u(\mathbb{A})$  are linearly independent since, setting  $a_{2i+1} = 0$ , one transforms  $P_u(\mathbb{A})$  into the usual word quasi-symmetric function in even-indexed variables  $P_u(a_2, a_4, \dots)$ . This gives a lower bound on the dimension of the space of solutions of Equation (5), corresponding to the upper bound found above, which ends the proof. ■

We can now prove the following.

**Corollary 5.2.** *The space  $\mathbf{WQSym}(\mathbb{A})$  is a subalgebra of  $\mathbb{C}\langle A \rangle$  and  $F \mapsto F(\mathbb{A})$  is an isomorphism of algebras from  $\mathbf{WQSym}$  to  $\mathbf{WQSym}(\mathbb{A})$ .*

*Proof.* The space of solution of Equation (5) is clearly a subalgebra of  $\mathbb{C}\langle A \rangle$ . But we have just proved that it is  $\mathbf{WQSym}(\mathbb{A})$ , whence our first claim.

Surjectivity of  $F \mapsto F(\mathbb{A})$  comes directly from the definition and injectivity corresponds to the fact the  $P_u(\mathbb{A})$  are linearly independent, proved above. Thus the only thing to prove is that it indeed defines an algebra morphism.

Let  $F$  and  $G$  be two elements of  $\mathbf{WQSym}$  and, to avoid confusion, denote by  $F \star G$  their product in  $\mathbf{WQSym}$ . As  $\mathbf{WQSym}(\mathbb{A})$  is an algebra, the stable polynomial  $F(\mathbb{A}) \cdot G(\mathbb{A})$  (here, the product is the product in  $\mathbb{C}\langle A \rangle$ ) can be written as  $H(\mathbb{A})$ , for some element  $H \in \mathbf{WQSym}$ . But setting  $a_{2i+1} = 0$  sends  $F(\mathbb{A})$ ,  $G(\mathbb{A})$  and  $H(\mathbb{A})$  to the *usual* word quasi symmetric functions  $F$ ,  $G$  and  $H$  in even indexed variables.

Therefore this specialization also sends  $F(\mathbb{A}) \cdot G(\mathbb{A})$  to  $F \star G$  in even indexed variables. Thus, necessarily,  $H = F \star G$ , which concludes the proof.  $\blacksquare$

**Remark 5.3.** This morphism property is natural to look for when defining evaluation on virtual alphabets. It is quite remarkable that our functional equation helps to prove it.

## 6. NON COMMUTATIVE MULTIRECTANGULAR COORDINATES

**6.1. What can be extended to the noncommutative framework?** We would like to lift the sets and maps of the commutative diagram (29) to the noncommutative framework.

In the previous Section, we have studied the structure of  $\mathcal{S}_{\text{nc}}$ , the noncommutative analog of  $\mathcal{S}$ . The algebra  $Q\Lambda$ , defined as a subalgebra of functions on Young diagrams, has no noncommutative analog. A noncommutative analog of  $\mathcal{S}'$  is easily defined as follows.

Consider the space  $\mathcal{S}'_{\text{nc}}$  of polynomials in two infinite sets of non commutative variables satisfying

$$(39) \quad h_m \left( \begin{array}{cccc} b_1 & \cdots & b_m \\ d_1 & \cdots & d_m \end{array} \right) \Big|_{d_i=0} = h_{m-1} \left( \begin{array}{cccc} b_1 & \cdots & b_{i-1} & b_i + b_{i+1} & \cdots & b_m \\ d_1 & \cdots & d_{i-1} & d_{i+1} & \cdots & d_m \end{array} \right)$$

$$(40) \quad h_m \left( \begin{array}{cccc} b_1 & \cdots & b_m \\ d_1 & \cdots & d_m \end{array} \right) \Big|_{b_i=0} = h_{m-1} \left( \begin{array}{cccc} b_1 & \cdots & b_{i-1} & b_{i+1} & \cdots & b_m \\ d_1 & \cdots & d_{i-1} + d_i & d_{i+1} & \cdots & d_m \end{array} \right)$$

We also consider the substitutions

$$(41) \quad \begin{cases} b_i = a_{2i-1} - a_{2i} \\ d_i = a_{2i} - a_{2i+1} \end{cases} \quad \begin{cases} a_{2i+1} = (d_{i+1} + \cdots + d_m) - (b_1 + \cdots + b_i) \\ a_{2i} = (d_i + \cdots + d_m) - (b_1 + \cdots + b_i) \end{cases}$$

Then we have the following:

**Proposition 6.1.** *The substitutions (41) define morphisms*

$$\begin{aligned} \Phi_{a \rightarrow b, d} : \mathcal{S}_{\text{nc}} &\rightarrow \mathcal{S}'_{\text{nc}} \\ \Phi_{b, d \rightarrow a} : \mathcal{S}'_{\text{nc}} &\rightarrow \mathcal{S}_{\text{nc}}, \end{aligned}$$

with  $\Phi_{a \rightarrow b, d} \circ \Phi_{b, d \rightarrow a} = \text{Id}_{\mathcal{S}'_{\text{nc}}}$ . In particular  $\Phi_{a \rightarrow b, d}$  is surjective.

*Proof.* The arguments given in the commutative setting in Section 4.3 work without any changes in the noncommutative setting.  $\blacksquare$

A natural question is to determine the kernel of  $\Phi_{a \rightarrow b, d}$ . This kernel is a two-sided element of  $\mathbf{WQSym}(\mathbb{A})$  and thus, by Corollary 5.2, can be identified to a two-sided element of  $\mathbf{WQSym}$ . Clearly  $P_1(\mathbb{A}) = \sum (-1)^i a_i$  lies in this kernel, and hence, the two-sided ideal generated by  $P_1$  is included in it.

Recall that the symmetric group  $\mathfrak{S}_n$  acts on polynomials of degree  $n$  in non-commuting variables by permuting the letters inside words. This action stabilizes the homogeneous component  $\mathbf{WQSym}_n$  of degree  $n$  of  $\mathbf{WQSym}$ . Note that  $\Phi_{a \rightarrow b, d}$  is compatible with the reordering of variables inside a noncommutative monomial

and, hence, the homogeneous part of degree  $n$  of its kernel is invariant by the action of  $\mathfrak{S}_n$ .

We shall prove in this Section that the kernel of  $\Phi_{a \rightarrow b, d}$  is indeed the smallest two-sided ideal containing  $P_1$  and stable by reordering variables.

**Notation.** Let us fix some notations and conventions for the symmetric group and its action.

Applying the product  $\sigma\tau$  of two permutations  $\sigma$  and  $\tau$  means applying  $\sigma$  and *then*  $\tau$ . Then the symmetric group  $S_n$  has natural right actions on positive integers smaller or equal to  $n$ , subsets of  $\{1, \dots, n\}$ , words of length  $n$ , noncommutative polynomials of degree  $n$ . We shall denote the image of an object  $O$  by a permutation  $\sigma$  by  $O \cdot \sigma$ . By example

$$\begin{aligned} \{1, 2\} \cdot 3124 &= \{1, 3\} \\ (a_{i_1} a_{i_2} a_{i_3} a_{i_4}) \cdot 3124 &= a_{i_3} a_{i_1} a_{i_2} a_{i_4}. \end{aligned}$$

**6.2. A lower bound on the dimension of the kernel.** Let us denote by  $\mathcal{K}_n$  the subspace of  $\mathbf{WQSym}_n$  obtained by taking:

- elements of degree  $n$  of the *left* ideal generated by  $P_1$ ;
- their images by the action of permutations in  $\mathfrak{S}_n$ .

Clearly,  $\mathcal{K}_n$  is included in the kernel of  $\Phi_{a \rightarrow b, d}$  (it is a priori smaller than the degree  $n$  component of the smallest two-sided ideal containing 1 and stable by the action of  $\mathfrak{S}_n$ ). We shall determine the dimension of  $\mathcal{K}_n$  and, hence, a lower bound on the dimension of the kernel of  $\Phi_{a \rightarrow b, d}$  in degree  $n$ .

Let us first introduce a few more notations. For each  $n$ , choose a complementary subspace  $U_n$  of  $\mathcal{K}_n$  in  $\mathbf{WQSym}_n$ , stable by the action of  $\mathfrak{S}_n$  (as  $\mathcal{K}_n$  is stable by the action of the *finite* group  $\mathfrak{S}_n$ , the existence of a stable complementary subspace is a classical lemma in representation theory, see, *e.g.*, [7, Proposition 1.5]). Then construct subspaces  $M_E^n \subset \mathbf{WQSym}_n$  indexed by subsets  $E$  of  $\{1, \dots, n\}$  as follows.

- If  $E = \{n - k + 1, \dots, n\}$  for some  $k \in \{0, \dots, n\}$ , then set  $M_E^n = U_{n-k} \cdot P_1^k$ .
- Otherwise, set  $k = |E|$  and choose a permutation  $\sigma_E \in \mathfrak{S}_n$  such that one has  $\{n - k + 1, \dots, n\} \cdot \sigma_E = E$  and define  $M_E^n = M_{\{n-k+1, \dots, n\}}^n \cdot \sigma_E$ .

**Lemma 6.2.** *The space  $M_E^n$  is well-defined. Moreover, if  $\sigma \in \mathfrak{S}_n$  and  $E$  is a subset of  $\{1, \dots, n\}$ , then  $M_E^n \cdot \sigma = M_{E \cdot \sigma}^n$ .*

*Proof.* We need to check that  $M_{\{n-k+1, \dots, n\}}^n \cdot \sigma_E$  does not depend of the chosen permutation  $\sigma_E$ . Let  $\sigma_E$  and  $\sigma'_E$  be permutations of size  $n$  with

$$\{n - k + 1, \dots, n\} \cdot \sigma_E = \{n - k + 1, \dots, n\} \cdot \sigma'_E = E.$$

Then  $\sigma_E = \tau\sigma'_E$  with  $\{n - k + 1, \dots, n\} \cdot \tau = \{n - k + 1, \dots, n\}$ , that is  $\tau \in \mathfrak{S}_{n-k} \times S_k$ . As  $\{U_{n-k}\}$  and  $P_1^k$  are respectively stable by the actions of  $\mathfrak{S}_{n-k}$  and  $\mathfrak{S}_k$ , the space  $M_{\{n-k+1, \dots, n\}}^n = U_{n-k} P_1^k$  is stable by  $\tau$ . Hence

$$M_{\{n-k+1, \dots, n\}}^n \cdot \sigma_E = M_{\{n-k+1, \dots, n\}}^n \cdot \tau\sigma'_E = M_{\{n-k+1, \dots, n\}}^n \cdot \sigma'_E,$$

and  $M_E^n$  is well-defined.

To prove the second claim, let  $k = |E|$ . Note that if  $\{n - k + 1, \dots, n\} \cdot \sigma_E = E$  then  $\{n - k + 1, \dots, n\} \cdot \sigma_E \sigma = E \cdot \sigma$  and thus

$$M_{E \cdot \sigma}^n = M_{\{n-k+1, \dots, n\}}^n \cdot \sigma_E \sigma = M_E^n \sigma. \quad \blacksquare$$

Finally, denote by  $r_{P_1}$  (resp.  $l_{P_1}$ ) the right- (resp. left-) multiplication by  $P_1$ . We also consider the operator  $\delta$  on word quasi-symmetric functions, defined as follows:

- if, for some packed word  $v$ ,  $u = v \cdot (m + 1)$ , where  $m$  is the biggest letter in  $v$ , then  $\delta(P_u) = P_v$ ;
- otherwise,  $\delta(P_u) = 0$ .

We give a few trivial computational rules relating  $r_{P_1}$ ,  $l_{P_1}$ ,  $\delta$ , and the symmetric group action. In the following equation,  $w$  is an element of  $\mathbf{WQSym}_n$  or  $\mathbf{WQSym}_{n-1}$  and  $\sigma$  a permutation of  $\mathfrak{S}_{n-1}$ . Then  $\bar{\sigma}$  denotes the trivial extension of  $\sigma$  to  $\{1, \dots, n\}$  ( $\bar{\sigma}(n) = n$ ) and  $\vec{\sigma}$  the permutation of  $\mathfrak{S}_n$  defined by  $\vec{\sigma}(1) = 1$  and  $\vec{\sigma}(i + 1) = \sigma(i) + 1$  for  $i$  in  $\{1, \dots, n - 1\}$ . We have:

$$(42) \quad \delta(wP_1) = w,$$

$$(43) \quad \text{if } \deg(w) = n - 1 > 0 ; \delta(P_1w) = P_1\delta(w);$$

$$(44) \quad \delta(w \cdot \bar{\sigma}) = (\delta(w)) \cdot \sigma;$$

$$(45) \quad P_1(w \cdot \sigma) = (P_1w) \cdot \vec{\sigma};$$

$$(46) \quad (w \cdot \sigma)P_1 = (wP_1) \cdot \vec{\sigma}.$$

**Lemma 6.3.** *We have the following compatibility properties between spaces  $M_E^n$  and operators  $\delta$  and  $r_{P_1}$ .*

- If  $E$  is a subset of  $\{1, \dots, n - 1\}$ , then

$$r_{P_1}(M_E^{n-1}) = M_{E \sqcup \{n\}}^n;$$

$$l_{P_1}(M_E^{n-1}) = M_{\{1\} \sqcup \vec{E}}^n,$$

where  $\vec{E} = \{i + 1, i \in E\}$ .

- Let  $r < n$  be a non-negative integer. If  $E \not\subseteq \{n - r + 1, \dots, n\}$ , then

$$\delta^r(M_E^n) \subset \mathcal{K}_{n-r}.$$

*Proof.* The first item follows directly from the relevant definitions and Equations (45) and (46).

Let us prove the second item. We begin by the case where  $1 \in E$ , that is  $E = 1 \sqcup \vec{E}'$  for some set  $E'$ . Using the first item,  $M_E^n = l_{P_1}(M_{E'}^{n-1})$ . In particular, every element  $m_E$  of  $M_E^n$  can be written as  $P_1 \cdot w_{n-1}$  for some element  $w \in \mathbf{WQSym}_{n-1}$ . Then, using (43),

$$\delta^r(m_E) = \delta^r(P_1w_{n-1}) = P_1\delta^r(w_{n-1}) = (\delta^r(w_{n-1})P_1) \cdot \sigma,$$

where  $\sigma \in \mathfrak{S}_{n-r}$  has word notation  $23 \dots (n - r)1$ . This shows that  $\delta^r(m_E)$  belongs to  $\mathcal{K}_{n-r}$ , as required.



Now, let  $E$  be some set not contained in  $\{n-r+1, \dots, n\}$ . Choose a permutation  $\tau$  in  $S_n$  fixing  $n-r+1, \dots, n$  and such that  $E \cdot \tau$  contains 1. Denote by  $\tilde{\tau} \in \mathfrak{S}_{n-r}$  its restriction to  $\{1, \dots, n-r\}$ . Iterating (44), we get that, for  $m_E$  in  $M_E^n$ ,

$$\delta^r(m_E) = [\delta^r(m_E \cdot \tau)] \cdot \tilde{\tau}^{-1}.$$

But  $m_E \tau \in M_{\tau(E)}^n$  and hence, using the case considered above,  $\delta^r(m_E \tau)$  is in  $\mathcal{K}_{n-r}$ . As  $\mathcal{K}_{n-r}$  is stable by the action of  $\mathfrak{S}_{n-r}$ ,  $\delta^r(m_E)$  is also in  $\mathcal{K}_{n-r}$ , which is the desired result.  $\blacksquare$

**Proposition 6.4.** *With the notation above,*

$$\mathbf{WQSym}_n = \bigoplus_{E \subseteq \{1, \dots, n\}} M_E^n.$$

*Proof.* Let us first prove that the sum is indeed a direct sum. Consider a vanishing linear combination  $\sum_{E \subseteq \{1, \dots, n\}} m_E = 0$  with  $m_E = u_E P_1^{|E|} \cdot \sigma_E \in M_E^n$  (here,  $u_E$  lies in  $U_{n-|E|}$ ). First, write

$$m_\emptyset = - \sum_{\substack{E \subseteq \{1, \dots, n\} \\ E \neq \emptyset}} m_E$$

The left-hand side is in  $U_n$  by definition, while the right-hand side is in  $\mathcal{K}_n$  (case  $r = 0$  of the second item of Lemma 6.3). Thus  $m_\emptyset$  necessarily vanishes.

Note that  $\delta(m_{\{n\}}) = u_{\{n\}}$  lies in  $U_{n-1}$ . But by linearity of  $\delta$ ,

$$\delta(m_{\{n\}}) = - \sum_{\substack{E \subseteq \{1, \dots, n\} \\ E \neq \emptyset, \{n\}}} \delta(m_E).$$

But, using the second item of Lemma 6.3 for  $r = 1$ , we get that  $\delta(m_E)$  lies in  $\mathcal{K}_{n-1}$  for  $E \neq \emptyset, \{n\}$ . Hence,  $\delta(m_{\{n\}}) = u_{\{n\}}$  lies also in  $\mathcal{K}_{n-1}$  and vanishes. The same arguments applied to  $\sum_E m_E \cdot (i, n)$  imply that  $m_{\{i\}} = 0$  (here,  $i$  is an positive integer smaller than  $n$  and  $(i, n)$  is the transposition in  $\mathfrak{S}_n$  exchanging  $i$  and  $n$ ).

Then we write that

$$\delta^2(m_{\{n-1, n\}}) = - \sum_{\substack{E \subseteq \{1, \dots, n\} \\ |E| \geq 2, E \neq \{n-1, n\}}} \delta^2(m_E).$$

The condition  $|E| \geq 2$  arises because we have already proved that  $m_E = 0$  for  $|E| \leq 1$ . Using the second item of Lemma 6.3 for  $r = 2$ , we immediately see that the right-hand side lies in  $\mathcal{K}_{n-2}$ . But  $\delta^2(m_{\{n-1, n\}}) = u_{\{n-1, n\}}$  also lies in  $U_{n-2}$  and thus vanishes.

Applying the same arguments to  $\sum_E m_E \cdot \sigma$  for well-chosen permutations  $\sigma$ , we prove that  $m_E = 0$  for any pair  $E$  included in  $\{1, \dots, n\}$ .

By iterating the same arguments, we conclude that  $m_E$  vanishes for any subset  $E$  of  $\{1, \dots, n\}$ . This proves that the sum in the statement is direct. Let us prove that it is indeed  $\mathbf{WQSym}_n$ .

We proceed by induction on  $n$ . For  $n = 0$ ,  $\mathbf{WQSym}_0 \simeq \mathbb{C}$  while  $\mathcal{K}_0 = \{0\}$ . Hence  $U_0 \simeq \mathbb{C}$  and  $M_\emptyset^0 = U_0 \simeq \mathbb{C}$ . Therefore, the result is true in this case.

Assume that it is true for  $n - 1$ . Consider an element  $w_{n-1}P_1$ , where  $w_{n-1}$  lies in  $\mathbf{WQSym}_{n-1}$ . By induction hypothesis

$$w_{n-1} = \sum_{E \subset \{1, \dots, n-1\}} m'_E,$$

where  $m'_E \in M_E^{n-1}$ .

Using the first item of Lemma 6.3, for any  $E \subset \{1, \dots, n-1\}$ , the product  $m'_E P_1$  lies in  $M_{E \sqcup \{n\}}^n$ . This proves that  $w_{n-1}P_1$  lies in

$$\bigoplus_{E \subset \{1, \dots, n\}} M_E^n.$$

Besides, this sum is invariant by the action of  $\mathfrak{S}_n$ , because of Lemma 6.2. As it contains elements of the form  $w_{n-1}P_1$ , it contains  $\mathcal{K}_n$ . Furthermore, it contains  $M_\emptyset^n = U_n$ . Finally, it is equal to  $\mathbf{WQSym}_n$ .  $\blacksquare$

Recall that  $\dim(\mathbf{WQSym}_n) = OB(n)$  the  $n$ -th ordered Bell number. Denote by  $k_n$  the dimension of  $\mathcal{K}_n$ . The spaces  $M_E^n$  have the same dimension as  $U_{n-|E|}$ , that is  $OB(n-|E|) - k_{n-|E|}$ . Therefore, we have the following immediate numerical corollary of Proposition 6.4.

**Corollary 6.5.** *For  $n \geq 1$ ,*

$$OB(n) = \sum_{j=0}^n \binom{n}{j} (OB(n-j) - k_{n-j}).$$

With this relation and the base case  $k_0 = 0$ , the numbers  $k_j$  can be computed inductively. The first values are 0, 1, 1, 7, 37, 271. This sequence appears in the Online Encyclopedia of Integer Sequences [22, A089677]. We shall give a simple combinatorial interpretation of it, not mentioned in [22].

Recall that  $OB(n)$  counts the set-compositions (*i.e.*, ordered set partitions) of  $\{1, \dots, n\}$ . Denote by  $OB_{\text{odd}}(n)$  the number of set compositions with an odd number of parts. A set-composition with an odd number of parts can be specified as follows

- the first set  $I_1$  of the set composition;
- a set composition of  $\{1, \dots, n\} \setminus I_1$  in an even number of parts.

For a given  $j$ , there are  $\binom{n}{j}$  sets  $I_1$  of size  $j$ , and, for each of them, there are exactly  $OB(n-j) - OB_{\text{odd}}(n-j)$  set compositions with an even number of parts of  $\{1, \dots, n\} \setminus I_1$ . Thus, the sequence  $(OB_{\text{odd}}(n))_{n \geq 0}$  satisfies the following induction

$$OB_{\text{odd}}(n) = \sum_{h=1}^n \binom{n}{h} (OB(n-h) - OB_{\text{odd}}(n-h)),$$

together with base case  $OB_{\text{odd}}(0) = 0$ . It is the same induction and base case as for  $k_n$ . Hence one has the following combinatorial interpretation for  $k_n$

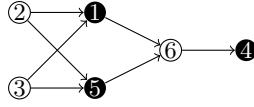


FIGURE 4. Example of a labelled bipartite graph  $G$ .

**Proposition 6.6.**  $k_n$  counts the set-compositions of  $\{1, \dots, n\}$  with an odd number of parts.

**Corollary 6.7.** The dimension of the kernel of  $\Phi_{a \rightarrow b, d}$  in degree  $n$  is at least the number of set-compositions of  $\{1, \dots, n\}$  with an odd number of parts.

**Open problem 6.8.** Find a basis of  $\mathcal{K}_n$  indexed by set-compositions of  $\{1, \dots, n\}$  with an odd number of parts.

**Remark 6.9.** One also has the remarkable relation

$$k_n = \frac{1}{2}(OB(n) - (-1)^n).$$

This identity, suggested by V. Jovovic [22, A089677], can be proved easily from our combinatorial interpretation, but is not useful in this paper.

**6.3. Functions  $N_G$  in noncommuting variables.** In the next Section, we shall exhibit an explicit independent homogeneous family in  $\mathcal{S}'_{nc}$ . As  $\Phi_{a \rightarrow b, d}$  is surjective, this will give us, for each  $n \geq 1$ , a lower bound on the dimension of the image of  $\Phi_{a \rightarrow b, d}$  in degree  $n$ . This will be done by lifting the construction of  $N_G$ , done in Section 4.5, to the noncommutative world.

Let us first lift the one-alphabet function  $F_G$ . Take as data a *labelled* bipartite graph  $\mathbf{G}$  (with two types of edges) with vertex set  $V = V_1 \sqcup V_2 = \{1, \dots, n\}$ . Then we define the noncommutative analog  $\mathbf{F}_G$  of  $F_G$  as follows:

$$\mathbf{F}_G(a_1, a_2, \dots) = \sum_{\substack{r: V \rightarrow \mathbb{N} \\ \text{with order condition}}} a_{r(1)} a_{r(2)} \dots a_{r(n)}.$$

Here, as usual, the  $a_i$  are noncommuting variables. Clearly,  $\mathbf{F}_G$  is a word quasi-symmetric function.

In the same way, we can define a noncommutative analog  $\mathbf{N}_G$  of  $N_G$ :

$$\mathbf{N}_G \begin{pmatrix} b_1 & b_2 & \dots \\ d_1 & d_2 & \dots \end{pmatrix} = \sum_{\substack{r: V \rightarrow \mathbb{N} \\ \text{with order condition}}} \gamma_{r(1)} \gamma_{r(2)} \dots \gamma_{r(n)},$$

where we use the abusive shorthand notation  $\gamma_{r(i)} = b_{r(i)}$  for  $i \in V_1$  and  $\gamma_{r(i)} = d_{r(i)}$  for  $i \in V_2$ .

**Example 6.10.** Consider the graph  $\mathbf{G}_{ex}$  drawn on Figure 4. This is a labelled version of the graph  $G_{ex}$  of Figure 3. As on this figure, vertices in  $V_1$  (resp.  $V_2$ ) are drawn in white (resp. black). Edges in  $E_{1,2}$  (resp.  $E_{2,1}$ ) are represented directed from their extremity in  $V_1$  to their extremity in  $V_2$  (resp. from their extremity in  $V_2$  to their extremity in  $V_1$ ).

Let  $r$  be a function from its vertex set to  $\mathbb{N}$ . Define

$$e := r(2), \quad f := r(3), \quad g := r(1), \quad h := r(5), \quad i := r(6), \quad j := r(4).$$

Then, by definition,  $r$  satisfy the order condition if and only if

$$e, f \leq g, h < i \leq j,$$

so one has,

$$\mathbf{N}_{\mathbf{G}_{\text{ex}}} \begin{pmatrix} b_1 & b_2 & \cdots \\ d_1 & d_2 & \cdots \end{pmatrix} = \sum_{e, f \leq g, h < i \leq j} d_g b_e b_f d_j d_h b_i,$$

which is a noncommutative version of the function  $N_{\mathbf{G}_{\text{ex}}}$  given in Example 6.10.

The following extension of Lemma 4.12 has the very same proof.

**Lemma 6.11.** *Let  $G$  be bipartite graph as above. Assume that each element in  $V$  is the extremity of at least one edge in  $E_{1,2}$ . Then the function  $\mathbf{N}_{\mathbf{G}}$  belongs to  $\mathcal{S}'_{\text{nc}}$ .*

Proposition 4.14 can also be lifted directly.

**Proposition 6.12.** *For any labelled bipartite graph  $G$  with vertex set  $V = V_1 \sqcup V_2 = \{1, \dots, n\}$ ,*

$$\mathbf{N}_{\mathbf{G}} \begin{pmatrix} b_1 & b_2 & \cdots \\ d_1 & d_2 & \cdots \end{pmatrix} = (-1)^{|V_1|} \Phi_{a \rightarrow b, d}(\mathbf{F}_{\mathbf{G}}(\mathbb{A})).$$

**6.4. A large linearly independent explicit family in  $(\mathcal{S}'_{\text{nc}})_n$ .** We shall now construct a large enough family of labelled graphs  $\mathbf{G}$ , such that the corresponding  $\mathbf{N}_{\mathbf{G}}$  are linearly independent.

Consider a set composition  $\mathbf{K}$  of  $n$  with an even number of parts

$$\mathbf{K} = (K_1, \dots, K_{2\ell}) \text{ with } K_1 \sqcup \cdots \sqcup K_{2\ell} = \{1, \dots, n\}$$

Then we define  $\mathbf{G}_{\mathbf{K}}$  as follows:

- Its vertex set  $\{1, \dots, n\}$  is split into two parts:

$$V_1 = K_1 \sqcup K_3 \sqcup \cdots \sqcup K_{2\ell-1};$$

$$V_2 = K_2 \sqcup K_4 \sqcup \cdots \sqcup K_{2\ell}.$$

- Its edge sets are given by

$$E_{1,2} = (K_1 \times K_2) \sqcup (K_3 \times K_4) \sqcup \cdots \sqcup (K_{2\ell-1} \times K_{2\ell});$$

$$E_{2,1} = (K_2 \times K_3) \sqcup (K_4 \times K_5) \sqcup \cdots \sqcup (K_{2\ell-2} \times K_{2\ell-1}).$$

**Example 6.13.** If  $\mathbf{K} = (\{2, 3\}, \{1, 5\}, \{6\}, \{4\})$ , then  $\mathbf{G}_{\mathbf{K}}$  is the graph of figure 4.

By Lemma 6.11, the associated functions  $\mathbf{N}_{\mathbf{G}_{\mathbf{K}}}$  belong to  $\mathcal{S}'_{\text{nc}}$ .

**Lemma 6.14.** *The functions  $\mathbf{N}_{\mathbf{G}_{\mathbf{K}}}$ , where  $\mathbf{K}$  runs over set compositions of  $\{1, \dots, n\}$  (for  $n \geq 1$ ) with an even number of parts are linearly independent.*

*Proof.* With a noncommutative monomial (a word) in  $b_i$  and  $d_i$ , we can associate its *evaluation*, which we define as the integer sequence

$$(\text{number of } b_1, \text{number of } d_1, \text{number of } b_2, \dots).$$

It is immediate to see that the monomial in  $\mathbf{N}_{\mathbf{G}_K}$  with the lexicographically largest evaluation is obtained as follows: it has letters  $b_1$  in positions given by  $K_1$ , letters  $d_1$  in position given by  $K_2$ , letters  $b_2$  in positions given by  $K_3, \dots$ . It follows that the set-composition  $\mathbf{K}$  can be recovered from the monomial of lexicographically largest evaluation in  $\mathbf{N}_{\mathbf{G}_K}$ , which implies the linear independence of the  $\mathbf{N}_{\mathbf{G}_K}$ . ■

**Corollary 6.15.** *The dimension of  $(\mathcal{S}'_{nc})_n$ , that is of the image of  $\Phi_{a \rightarrow b, d}$  in degree  $n$  is at least the number of set-compositions of  $\{1, \dots, n\}$  with an even number of parts.*

**Remark 6.16.** The corresponding family  $\mathbf{F}_{\mathbf{G}_K}$  is a natural noncommutative lifting of the basis introduced by K. Luoto in [14] (here, we only lift elements indexed by even-length composition, but it would not be hard to lift all elements).

An interesting feature is that the above argument together with the fact that  $\mathbf{N}_{\mathbf{G}_K} = (-1)^{|V_1|} \Phi_{a \rightarrow b, d}(\mathbf{F}_{\mathbf{G}_K})$  implies that the  $\mathbf{F}_{\mathbf{G}_K}$  are linearly independent, which would have been difficult to prove directly.

**6.5. Conclusion.** As the dimension of  $(\mathcal{S}_{nc})_n \simeq \mathbf{WQSym}_n$  is the number of set-compositions of  $[n]$ , Corollaries 6.7 and 6.15 imply:

**Theorem 6.17.** (i) *The kernel of  $\Phi_{a \rightarrow b, d}$  is  $\bigoplus_{n \geq 1} \mathcal{K}_n$  and its degree  $n$  component has a dimension equal to the number of set-compositions of  $\{1, \dots, n\}$  in an odd number of parts. Besides, it is the smallest homogeneous two-sided ideal containing  $P_1$  and whose homogeneous components are stable by the action of the symmetric groups.*

(ii) *The dimension of the degree  $n$  component of the image of  $\Phi_{a \rightarrow b, d}$ , that is of  $(\mathcal{S}'_{nc})_n$ , is exactly the number of set-compositions of  $\{1, \dots, n\}$  in an even number of parts.*

## 7. APPENDIX

The definitions of the virtual alphabet  $\mathbb{X}$  and of the functions  $\mathbf{H}_I$  can be made more transparent if one introduces the formalism of noncommutative symmetric functions [8].

For a totally ordered alphabet  $A = \{a_i | i \geq 1\}$  of noncommuting variables, and  $t$  an indeterminate, one sets

$$(47) \quad \sigma_t(A) = \prod_{i \geq 1}^{\rightarrow} (1 - ta_i)^{-1} = \sum_{n \geq 0} S_n(A) t^n.$$

The complete symmetric functions  $S_n(A)$  generate a free associative algebra  $\mathbf{Sym}(A)$  (noncommutative symmetric functions). One denotes by

$$(48) \quad S^I(A) := S_{i_1} S_{i_2} \cdots S_{i_r}$$

its natural basis. Actually,  $\mathbf{Sym}(A)$  is a Hopf algebra, and its graded dual is  $QSym$ . This can be deduced from the noncommutative Cauchy formula

$$(49) \quad \sigma_1(XA) := \prod_{i \geq 1}^{\rightarrow} \sigma_{x_i}(A) = \sum_I M_I(X) S^I(A)$$

which allows to identify  $M_I$  with the dual basis of  $S^I$  [8, 17].

Now, the virtual alphabet  $\mathbb{X}$  can be defined by

$$(50) \quad \sigma_1(\mathbb{X}A) = \sum_I M_I(\mathbb{X}) S^I(A) = \prod_{i \geq 1}^{\rightarrow} \sigma_{x_i}(A)^{(-1)^i}.$$

It is clear that the right-hand side (seen as a polynomial in infinitely many variables and coefficients in  $\mathbf{Sym}(A)$ ) is a solution of (1). Hence all the coefficients in its  $S^I$  expansion, that is the  $M_I(\mathbb{X})$ , are also solution of (1).

The set of polynomials in  $p_1, p_2, \dots, q_1, q_2, \dots$ , that are solutions of equations (26) and (27) can also be described with the language of virtual alphabets.

Consider the virtual alphabet  $\mathbb{Y}$  defined by

$$(51) \quad \sigma_1(\mathbb{Y}A) = \prod_{i \geq 1}^{\rightarrow} \sigma_{q'_i}(A)^{\frac{p_i}{q'_i}}.$$

Then we have:

**Theorem 7.1.** *A linear basis for the space  $\mathcal{S}'$  of solutions of (26) and (27) is given by the function  $M_I(\mathbb{Y})$ , where  $I$  runs over compositions with the last part bigger than 1.*

*Proof.* On the series (51), it is immediate that quasi-symmetric functions in  $\mathbb{Y}$  satisfy conditions (26) and (27), except equation (26) for  $i = m$ . Indeed, when we restrict to  $m$  variables and set  $q_m = 0$ , then  $q'_m = 0$  and

$$\sigma_{q'_m}(A)^{\frac{p_m}{q'_m}} = \exp(p_m S_1(A)),$$

and the variable  $p_m$  does not disappear.

This can be corrected as follows: define  $\tilde{M}_I$  by

$$(52) \quad \sum_I \tilde{M}_I S^I := \sigma_1(\mathbb{Y}A) e^{-(\sum_i p_i) S_1}.$$

Then  $\tilde{M}_I$  satisfies equations (26) and (27). But the second factor in the right-hand side of (52) contains only  $S_1$ , thus, if the last part of a composition  $I$  is bigger than 1, we can forget this factor when extracting the coefficient of  $S^I$ . In other terms, for such a composition  $I$

$$\tilde{M}_I = M_I(\mathbb{Y}).$$

This shows that the  $M_I(\mathbb{Y})$  are indeed solutions of (26) and (27). If we substitute  $p_i = q'_i = y_i$ , we recover  $M_I(Y)$ , and hence they are linearly independent. A dimension argument finishes the proof. ■

This implies that  $H_I = F(\mathbb{Y})$  for some quasi-symmetric function  $F$  and we shall now identify  $F$ .

Let  $H_I$  be the function defined in equation (33). Alternatively, let  $\phi_n$  be the noncommutative symmetric functions defined by

$$(53) \quad \log \sigma_t(A) = \sum_{n \geq 1} \phi_n(A)t^n,$$

then  $H_I \in QSym$  is the dual basis of the basis  $\phi^I$  of **Sym**.

**Proposition 7.2.**  $H_I = H_I(\mathbb{Y})$ .

*Proof.* After substitution  $p_i = q'_i = y_i$ ,  $F(\mathbb{Y})$  is sent to  $F(Y)$  and  $H_I$  to  $H_I(Y)$  (see Section 4.4.3). This yields  $F = H_I$ . ■

**Remark 7.3.** Equation (52) is an avatar of the  $(1 - \mathbb{E})$ -transform investigated in [10], another example being the so-called quasi-shuffle regularization of Multiple Zeta Values (see, e.g., [1]).

REFERENCES

[1] P. CARTIER, *Fonctions polylogarithmes, nombres polyzetas et groupes pro-unipotents*, Astérisque, 282, (2002), 137–173, (Sem. Bourbaki no. 885).

[2] M. DOŁĘGA, V. FÉRAY, and P. ŚNIADY, *Explicit combinatorial interpretation of Kerov character polynomials as numbers of permutation factorizations*. Adv. Math., **225** (1) (2010), 81- 120.

[3] R. EHRENBORG, *On Posets and Hopf Algebras*, Adv. Math **119** (1996), 1–25.

[4] V. FÉRAY, *Combinatorial interpretation and positivity of Kerov’s character polynomials*, J. Alg. Comb., **29** (4) (2009), 473–507.

[5] V. FÉRAY, *Stanley’s formula for characters of the symmetric group*, Ann. Comb., **13** (4) (2010), 453–461.

[6] V. FÉRAY and P. ŚNIADY, *Asymptotics of characters of symmetric groups related to Stanley character formula*. Ann. Math., **173** (2) (2011), 887–906.

[7] W. FULTON and J. HARRIS, *J. Representation theory: a first course*, volume **129** of *Graduate Texts in Mathematics*, Springer, 1991.

[8] I.M. GELFAND, D. KROB, A. LASCoux, B. LECLERC, V.S. RETAKH, and J.-Y. THIBON, *Noncommutative symmetric functions*, Adv. Math. **112** (1995), 218–348.

[9] I.M. GESSEL, *Multipartite P-partitions and inner products of skew Schur functions*, Contemp. Math. **34** (1984), 289–301.

[10] F. HIVERT, J.-G. LUQUE, J.-C. NOVELLI, and J.-Y. THIBON, *The (1-E)-transform in combinatorial Hopf algebras*, J. Algebraic Combin. **33** (2011), 277–312.

[11] V. IVANOV and G. OLSHANSKI, *Kerov’s central limit theorem for the Plancherel measure on Young diagrams*. In *Symmetric Functions 2001: Surveys of Developments and Perspectives*, volume 74 of *NATO Science Series II. Mathematics, Physics and Chemistry* (2002), 93–151.

[12] S. KEROV and G. OLSHANSKI, *Polynomial functions on the set of Young diagrams*, C. R. Acad. Sci. Paris Sér. I Math. 319 (1994), no. 2, 121–126

[13] D. KROB, B. LECLERC and J.-Y. THIBON, *Noncommutative symmetric functions II: Transformations of alphabets*, Internat. J. Alg. Comp. **7** (1997), 181–264.

[14] K.W. LUOTO, *A matroid-friendly basis for the quasisymmetric functions*, J. Combin. Th., Series A, **115** (2008), no. 5, 777–798.

[15] I.G. MACDONALD, *Symmetric functions and Hall polynomials*, 2nd edition, Oxford, 1995.

- [16] C. MALVENUTO, *Produits et coproduits des fonctions quasi-symétriques et de l'algèbre des descentes*, Thèse de doctorat, Université du Québec à Montréal, 1993 (Publications du LaCIM, **16**).
- [17] C. MALVENUTO and C. REUTENAUER, *Duality between quasi-symmetric functions and the Solomon descent algebra*, Journal of Algebra **177** (1995), 67-982.
- [18] J.-C. NOVELLI and J.-Y. THIBON, *Polynomial realizations of some trialgebras*, FPSAC'06, San-Diego.
- [19] J.-C. NOVELLI and J.-Y. THIBON, *Superization and  $(q,t)$ -specialization in combinatorial Hopf algebras*, Electronic J. Combin. **16** (2) (2009), R21.
- [20] A. OKOUNKOV and G. OLSHANSKI, *Shifted Schur functions*. (Russian) Algebra i Analiz **9** (1997), no. 2, 73–146; translation in St. Petersburg Math. J. **9** (1998), no. 2, 239–300.
- [21] G. OLSHANSKI, *Plancherel averages: Remarks on a paper by Stanley*, Elec. J. Combin. **17** (2010), R43.
- [22] *The On-Line Encyclopedia of Integer Sequences*, published electronically at <http://oeis.org>, 2013.
- [23] R.P. STANLEY, *Irreducible symmetric group characters of rectangular shape*, Sémin. Loth. Comb. (elec) **50**, B50d (2003).
- [24] R.P. STANLEY, *Ordered structures and partitions*, Mem. Amer. Math. Soc., **119**, 1972.
- [25] J.R. STEMBRIDGE, *A characterization of supersymmetric polynomials*. J. Alg. **95** (2) (1985), 439–444.

(Aval) LABRI, UNIVERSITÉ BORDEAUX I, 351 COURS DE LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE

(Féray) INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH, WINTERTHURERSTRASSE 190, 8057 ZÜRICH, SWITZERLAND

(Novelli and Thibon) INSTITUT GASPARD MONGE, UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE, 5 BOULEVARD DESCARTES, CHAMPS-SUR-MARNE, 77454 MARNE-LA-VALLÉE CEDEX 2, FRANCE

*E-mail address*, Jean-Christophe Aval: [aval@labri.fr](mailto:aval@labri.fr)

*E-mail address*, Valentin Féray: [valentin.feray@math.uzh.ch](mailto:valentin.feray@math.uzh.ch)

*E-mail address*, Jean-Christophe Novelli: [novelli@univ-mlv.fr](mailto:novelli@univ-mlv.fr)

*E-mail address*, Jean-Yves Thibon: [jyt@univ-mlv.fr](mailto:jyt@univ-mlv.fr)