



# ANNALES

DE

# L'INSTITUT FOURIER

Loïc FOISSY, Frédéric PATRAS & Jean-Yves THIBON

**Deformations of shuffles and quasi-shuffles**

Tome 66, n° 1 (2016), p. 209-237.

[http://aif.cedram.org/item?id=AIF\\_2016\\_\\_66\\_1\\_209\\_0](http://aif.cedram.org/item?id=AIF_2016__66_1_209_0)



© Association des Annales de l'institut Fourier, 2016,

*Certains droits réservés.*



Cet article est mis à disposition selon les termes de la licence  
CREATIVE COMMONS ATTRIBUTION – PAS DE MODIFICATION 3.0 FRANCE.  
<http://creativecommons.org/licenses/by-nd/3.0/fr/>

L'accès aux articles de la revue « Annales de l'institut Fourier »  
(<http://aif.cedram.org/>), implique l'accord avec les conditions générales  
d'utilisation (<http://aif.cedram.org/legal/>).

cedram

Article mis en ligne dans le cadre du  
Centre de diffusion des revues académiques de mathématiques  
<http://www.cedram.org/>

## DEFORMATIONS OF SHUFFLES AND QUASI-SHUFFLES

by Loïc FOISSY, Frédéric PATRAS & Jean-Yves THIBON (\*)

---

ABSTRACT. — We investigate deformations of the shuffle Hopf algebra structure  $\text{Sh}(A)$  which can be defined on the tensor algebra over a commutative algebra  $A$ . Such deformations, leading for example to the quasi-shuffle algebra  $\text{QSh}(A)$ , can be interpreted as natural transformations of the functor  $\text{Sh}$ , regarded as a functor from commutative nonunital algebras to coalgebras. We prove that the monoid of natural endomorphisms of the functor  $\text{Sh}$  is isomorphic to the monoid of formal power series in one variable without constant term under composition, so that in particular its natural automorphisms are in bijection with formal diffeomorphisms of the line.

These transformations can be interpreted as elements of the Hopf algebra of word quasi-symmetric functions (or, equivalently, of surjections)  $\mathbf{WQSym}$ , and in turn define deformations of its structure. This leads to a new embedding of free quasi-symmetric functions into  $\mathbf{WQSym}$ , whose relevance is illustrated by a simple and transparent proof of Goldberg's formula for the coefficients of the Hausdorff series.

RÉSUMÉ. — On s'intéresse aux déformations de la structure d'algèbre de Hopf des battages (ou shuffles)  $\text{Sh}(A)$  définie sur l'algèbre tensorielle sur une algèbre commutative  $A$ . Ces déformations, dont un cas remarquable est donné par l'algèbre de Hopf des quasi-shuffles  $\text{QSh}(A)$ , s'interprètent comme transformations naturelles du foncteur  $\text{Sh}$  vu comme foncteur des algèbres commutatives non unitaires vers les coalgèbres. On montre en particulier que le monoïde des endomorphismes naturels du foncteur  $\text{Sh}$  est isomorphe au monoïde des séries formelles en une variable sans terme constant pour la loi de composition des séries. Les automorphismes naturels du foncteur sont donc en bijection avec les difféomorphismes formels de la droite.

Ces transformations s'interprètent aussi comme des éléments de l'algèbre de Hopf des surjections (ou, de façon équivalente des fonctions quasi-symétriques en mots)  $\mathbf{WQSym}$ , et en définissent à leur tour des déformations. Cette remarque conduit entre autres à un nouveau plongement des fonctions quasi-symétriques libres dans  $\mathbf{WQSym}$  dont la pertinence est illustrée par une preuve simple de la formule de Goldberg pour les coefficients de la série de Hausdorff.

---

*Keywords:* Shuffle algebras, Combinatorial Hopf algebras, Hausdorff series.

*Math. classification:* 05E05, 16T30.

(\*) The authors acknowledge support from the grant CARMA ANR-12-BS01-0017.

## Introduction

The tensor algebra over a commutative algebra  $A$  is provided by the shuffle product with a commutative (Hopf) algebra structure. However, other products (resp. Hopf algebra) structures can be defined, the best known one being the quasi-shuffle product (resp. quasi-shuffle Hopf algebra). These products arise in many contexts, e.g., Rota–Baxter algebras, multiple zeta values (MZVs), noncommutative symmetric functions, operads ... Moreover, they are natural: they commute with morphisms of commutative algebras; in other terms, they define functors from the category of commutative algebras to that of commutative Hopf algebras.

The present paper aims at studying and classifying these products and Hopf algebra structures. The Faà di Bruno Hopf algebra plays a key role in this classification: its characters (in bijection with formal diffeomorphisms tangent to the identity) happen to classify the natural deformations of shuffle algebras considered in this article.

Our approach also sheds a new light on classical constructions such as quasi-shuffles. Most structure properties of quasi-shuffle algebras appear, from our point of view, as straightforward consequences of their definition as deformations by conjugacy of shuffle algebras. This allows to transport automatically *all* known results on shuffle algebras to quasi-shuffles and does not require the algebra  $A$  to be graded (compare, e.g., [9]). We may quote the existence of a “natural” (but not straightforward !) gradation, and of fine Hopf algebraic properties such as the ones studied in [14].

Most of these results can be phrased in terms of combinatorial properties of Hopf algebras based on permutations (free quasi-symmetric functions) and surjections (word quasi-symmetric functions). In particular, the analysis of the relations between shuffles and quasi-shuffles leads to a new polynomial realization of noncommutative symmetric functions, which can be extended to free quasi-symmetric functions. Beyond its naturalness from the point of view of quasi-shuffle algebras, this new realization has an interest on its own: as an application, a simple (and, in our opinion, enlightening) proof of Goldberg’s formula for the coefficients of the Haudorff series and a generalization thereof to other, similar, series is obtained.

### 1. The shuffle algebra over a commutative algebra

Let  $V$  be a vector space over the rationals<sup>(1)</sup>, and let  $T(V)$  be its tensor algebra

$$T(V) = \bigoplus_{n \in \mathbb{N}} T_n(V) = \bigoplus_{n \in \mathbb{N}} V^{\otimes n}$$

with  $V^{\otimes 0} := \mathbf{Q}$ . Tensors  $v_1 \otimes \dots \otimes v_n$  in  $T_n(V)$  will be written as words  $v_1 \dots v_n$ , the  $\otimes$  sign being reserved for tensor products of elements of  $T(V)$ .

The concatenation product in  $T(V)$  is written  $\times$

$$v_1 \dots v_n \times w_1 \dots w_m := v_1 \dots v_n w_1 \dots w_m.$$

DEFINITION 1.1. — *The shuffle bialgebra  $\text{Sh}(V) = \bigoplus_{n \in \mathbb{N}} \text{Sh}_n(V)$  is the graded connected (i.e.  $\text{Sh}_0(V) = \mathbf{Q}$ ) commutative Hopf algebra such that*

- As vector spaces  $\text{Sh}_n(V) = T_n(V)$ .
- Its product  $\sqcup$  is defined recursively as the sum of the two half-shuffle products  $\prec, \succ$

$$v_1 \dots v_n \prec w_1 \dots w_p := v_1 \times (v_2 \dots v_n \sqcup w_1 \dots w_p)$$

$$v_1 \dots v_n \succ w_1 \dots w_p := v_1 \dots v_p \prec w_1 \dots w_n = w_1 \times (v_1 \dots v_n \sqcup w_2 \dots w_p),$$

and  $\sqcup = \prec + \succ$ .

- Its coalgebra structure is defined by the deconcatenation coproduct

$$\Delta(v_1 \dots v_n) := \sum_{0 \leq k \leq n} v_1 \dots v_k \otimes v_{k+1} \dots v_n.$$

Recall that the notions of connected commutative Hopf algebra and connected commutative bialgebra are equivalent since a graded connected commutative bialgebra always has an antipode [12].

The above construction is functorial:  $\text{Sh} : V \rightarrow \text{Sh}(V)$  is a functor from the category of vector spaces to the category of graded connected commutative Hopf algebras.

Equivalently, for all  $v_1, \dots, v_{k+l} \in V$ ,

$$v_1 \dots v_k \prec v_{k+1} \dots v_{k+l} = \sum_{\alpha \in \text{Des}_{\subseteq \{k\}}, \alpha^{-1}(1)=1} v_{\alpha^{-1}(1)} \dots v_{\alpha^{-1}(k+l)},$$

$$v_1 \dots v_k \succ v_{k+1} \dots v_{k+l} = \sum_{\alpha \in \text{Des}_{\subseteq \{k\}}, \alpha^{-1}(1)=k+1} v_{\alpha^{-1}(1)} \dots v_{\alpha^{-1}(k+l)},$$

where the  $\alpha$  are permutations of  $[k+l] = \{1, \dots, k+l\}$ .

---

<sup>(1)</sup>Any field of characteristic zero would be suitable as well in all our forthcoming developments.

The notation  $\alpha \in \text{Des}_{\subseteq\{k\}}$  means that  $\alpha$  has at most one descent in position  $k$ . Recall that a permutation  $\sigma$  of  $[n]$  is said to have a descent in position  $i < n$  if  $\sigma(i) > \sigma(i+1)$ . The descent set of  $\sigma$ ,  $\text{desc}(\sigma)$  is the set of all descents of  $\alpha$ ,

$$\text{desc}(\sigma) := \{i < n, \sigma(i) > \sigma(i+1)\}.$$

For  $I \subset [n]$ , we write  $\text{Des}_I := \{\sigma, \text{desc}(\sigma) = I\}$  and  $\text{Des}_{\subseteq I} := \{\sigma, \text{desc}(\sigma) \subseteq I\}$ .

A Theorem due to Schützenberger [19] characterizes abstractly the shuffle algebras:

**PROPOSITION 1.2.** — *As a commutative algebra,  $\text{Sh}(V)$  is the free algebra over the vector space  $V$  for the relation*

$$(1.1) \quad (a \prec b) \prec c = a \prec (b \prec c + c \prec b).$$

An algebra equipped with a product  $\prec$  satisfying this relation is sometimes referred to as a chronological algebra (although the term has also other meanings) or as a Zinbiel algebra (because it is dual to Cuvier's notion of Leibniz algebra), we refer to [5] for historical details.

Let now  $A$  be a commutative algebra over the rationals (not necessarily with a unit). The (internal) product of  $a$  and  $b$  in  $A$  will be always written  $a \cdot b \in A$  to distinguish it from  $ab$ , which will represent as previously an element in  $A^{\otimes 2}$ .

The shuffle bialgebra over  $A$ ,  $Sh(A)$  is defined as previously (the definition of  $Sh(A)$  does not make use of the product in  $A$ : in functorial terms, we simply use the forgetful functor from commutative algebras to vector spaces to define  $Sh(A)$ ). From now on we will consider accordingly  $Sh$  as a functor from the category of commutative algebras (without a unit) to the category of connected bialgebras and also as a functor to the category of connected coalgebras.

An ingredient of the theory of tensors and shuffle algebras over commutative algebras will be useful in our forthcoming developments: namely, the nonlinear Schur-Weyl duality established in [14]. Let us recall that **WQSym** stands for the Hopf algebra of word quasi-symmetric functions. This algebra can be given various equivalent realizations (that is, its elements can be encoded by means of surjections, packed words, set compositions or faces of permutohedra) and carries various algebraic structures on which we will come back later. We refer, e.g., to [14] for details. It will appear later in this article that, whereas the realization of **WQSym** in terms of surjections is the one suited for studying deformations of shuffle algebras, that in terms of words is the one suited to the analysis of Hausdorff series.

For the time being, we simply recall that **WQSym** can be realized as the linear span of all surjections  $f$  from  $[n]$  to  $[p]$ , where  $n$  runs over the integers and  $1 \leq p \leq n$  and postpone the definition of the product and the coproduct. We say that such a map  $f$  is of degree  $n$ , relative degree  $n - p$ , and bidegree  $(n, p)$ . The linear span of degree  $n$  (resp. bidegree  $(n, p)$ ) elements is written  $\widehat{\mathbf{WQSym}}_n$  (resp.  $\widehat{\mathbf{WQSym}}_{n,p}$ ). We write  $\widehat{\mathbf{WQSym}}$  for  $\prod_{n,p} \widehat{\mathbf{WQSym}}_{n,p}$ .

We consider now natural endomorphisms of the functor  $T$ , viewed as a functor from nonunital commutative algebras to vector spaces. Concretely, we look for families of linear maps  $\mu_A$  from  $T_n(A)$  to  $T_m(A)$  (where  $A$  runs over nonunital commutative algebras and  $m$  and  $n$  run over the nonzero integers) such that, for any map  $f$  of nonunital commutative algebras from  $A$  to  $B$ ,

$$(1.2) \quad T_m(f) \circ \mu_A = \mu_B \circ T_n(f).$$

Let us say that such a family  $\mu_A$  satisfies the nonlinear Schur-Weyl duality (with parameters  $n, m$ ). We have [14]:

**PROPOSITION 1.3.** — *The vector space of linear maps that satisfy the nonlinear Schur-Weyl duality with parameters  $n, m$  is canonically isomorphic to  $\widehat{\mathbf{WQSym}}_{n,m}$ , the linear span of surjections from  $[n]$  to  $[m]$ . Equivalently, the vector space of natural endomorphisms of the functor  $T$  is canonically isomorphic to  $\widehat{\mathbf{WQSym}}$ .*

## 2. Natural coalgebra endomorphisms

The nonlinear Schur-Weyl duality shows that  $\widehat{\mathbf{WQSym}}$  is the natural object for investigating the linear structure of the tensor and shuffle algebras over a commutative algebra. In this section and the following ones, we study shuffle algebras from a refined point of view. Namely, we aim at characterizing, inside  $\widehat{\mathbf{WQSym}}$ , the linear endomorphisms that preserve some extra structure. Particularly important from this point of view are the coalgebra endomorphisms, whose classification is the object of the present section.

From now on, coalgebras  $C$  are coaugmented, counital and conilpotent. That is, writing  $\Delta$  the coproduct and  $\text{id}_C$  the identity map of  $C$ : the coalgebra is equipped with a map  $\eta_C$  (the counit) from  $C$  to the ground field  $\mathbf{Q}$  which satisfies

$$(\eta_C \otimes \text{id}_C) \circ \Delta = \text{id}_C = (\text{id}_C \otimes \eta_C) \circ \Delta.$$

The ground field embeds (as a coalgebra) into  $C$  so that, canonically,  $C = C^+ \oplus \mathbf{Q}$  with  $C^+ := Ker \eta_C$  (coaugmentation) and, finally, writing  $\bar{\Delta}$  the reduced coproduct ( $\bar{\Delta}(x) := \Delta(x) - x \otimes 1 - 1 \otimes x$ ), for all  $x \in C^+$ , there exists an integer  $n$  (depending on  $x$ ) such that  $\bar{\Delta}^n(x) = 0$  (conilpotency). Coalgebra maps  $\phi : C \rightarrow C'$  are required to preserve the coaugmentation and the counit, so that  $\phi(C^+) \subset C'^+$ .

Recall that  $Sh(A)$  is equipped with a coalgebra structure

$$\Delta(a_1 \dots a_n) := \sum_{i=0}^n a_1 \dots a_i \otimes a_{i+1} \dots a_n \in Sh(A) \otimes Sh(A).$$

Its counit is the projection onto  $Sh_0(A) = \mathbf{Q}$  and we write

$$Sh(A)^+ = \bigoplus_{n \geq 1} A^{\otimes n}.$$

The usual universal properties of the tensor algebra  $T(A)$  as a free associative algebra over  $A$  (viewed as a vector space) dualize, and we have the adjunction property

$$Hom_{lin}(C^+, A) \cong Hom_{coalg}(C, Sh(A)),$$

where  $C$  runs over coalgebras and  $Hom_{lin}$ , resp.  $Hom_{coalg}$ , stand for the set of linear maps, resp. coalgebra maps. In other terms, the coaugmentation coideal functor from coalgebras to vector spaces is left adjoint to the free coalgebra functor. Recall that in this statement coalgebra means coaugmented, counital and conilpotent coalgebra. The fact that we consider only conilpotent coalgebras is essential for the adjunction to hold, see e.g. [1] for details on the structure of cofree coalgebras in general. In particular, we get:

COROLLARY 2.1. — *There is a canonical bijection*

$$Hom_{lin}(Sh^+(A), A) \cong Hom_{coalg}(Sh(A), Sh(A)).$$

That is, a coalgebra endomorphism  $\phi$  of  $Sh(A)$  is entirely determined by the knowledge of  $f := \pi \circ \phi$ , where we write  $\pi$  for the projection from  $Sh(A)$  to  $A$  orthogonally to  $A^0 = \mathbf{Q}$  and to the  $A^{\otimes n}$ ,  $n > 1$ . Conversely, any map  $f \in Hom_{lin}(Sh^+(A), A)$  determines a unique coalgebra endomorphism  $\phi$  of  $Sh(A)$  by

$$\phi(a_1 \dots a_n) := \sum_{i_1 + \dots + i_k = n} f(a_1 \dots a_{i_1}) \otimes \dots \otimes f(a_{i_1 + \dots + i_{k-1} + 1} \dots a_n).$$

A triangularity argument that we omit shows that  $\phi$  is a coalgebra automorphism if and only if the restriction of  $f$  to  $A$  is a linear isomorphism. For reasons that will become clear soon, we say that  $f \in Hom_{lin}(Sh^+(A), A)$  is

tangent to identity if its restriction to  $A$  (that is to a linear endomorphism of  $A$ ) is the identity map.

Recall that by natural endomorphism of the functor  $\text{Sh}$  viewed as a functor from commutative nonunital algebras to coalgebras is meant a family  $\mu_A$  (indexed by commutative nonunital algebras  $A$ ) of coalgebra endomorphisms of the  $\text{Sh}(A)$  commuting with the morphisms (from  $\text{Sh}(A)$  to  $\text{Sh}(B)$ ) induced by algebra maps (from  $A$  to  $B$ ).

**THEOREM 2.2.** — *Let  $\mathbf{Coalg}$  be the monoid of natural endomorphisms of the functor  $\text{Sh}$  viewed as a functor from commutative nonunital algebras to coalgebras. Then, there is an isomorphism between  $\mathbf{Coalg}$  and the monoid  $\mathbf{Diff}$  of formal power series without constant term,  $\mathbf{Coalg} \cong X\mathbf{Q}[[X]]$  equipped with the substitution product (for  $P(X), Q(X) \in X\mathbf{Q}[[X]]$ ,  $P \circ Q(X) := P(Q(X))$ ).*

*In particular, the set  $\mathbf{Coalg}_1$  of tangent-to-identity natural endomorphisms of the functor  $\text{Sh}$  is a group canonically in bijection with the group  $\mathbf{Diff}_1 = X + X^2\mathbf{Q}[[X]]$  of tangent-to-identity formal diffeomorphisms.*

Let us prove first that  $\mathbf{Coalg} \cong X\mathbf{Q}[[X]]$ . Since we have a natural isomorphism  $\text{Hom}_{\text{lin}}(\text{Sh}^+(A), A) \cong \text{Hom}_{\text{coalg}}(\text{Sh}(A), \text{Sh}(A))$ ,  $\mathbf{Coalg}$  is canonically in bijection with natural transformations from  $\text{Sh}^+$  to the identity functor (viewed now as functors from commutative algebras to vector spaces). By Schur-Weyl duality, we get  $\mathbf{Coalg} \cong \prod_{n \geq 1} \mathbf{WQSym}_{n,1}$ . Identifying the unique surjection from  $[n]$  to  $[1]$  with the monomial  $X^n$  yields the bijection  $\mathbf{Coalg} \cong X\mathbf{Q}[[X]]$ .

We will write from now on  $\phi_P$  for the element in  $\text{Hom}_{\text{coalg}}(\text{Sh}(A), \text{Sh}(A))$  associated with a given formal power series  $P(X) \in X\mathbf{Q}[[X]]$  and  $f_P$  for the corresponding element in  $\text{Hom}_{\text{lin}}(\text{Sh}^+(A), A)$ .

Notice that for  $P(X) = \sum_{i=1}^{\infty} p_i X^i$ , the action of  $f_P$  and  $\phi_P$  on an arbitrary tensor  $a_1 \dots a_n \in \text{Sh}(A)$  can be described explicitly

$$(2.1) \quad f_P(a_1 \dots a_n) = p_n \cdot (a_1 \cdot \dots \cdot a_n) \in A,$$

$$(2.2) \quad \begin{aligned} &\phi_P(a_1 \dots a_n) \\ &= \sum_{k=1}^n \sum_{i_1 + \dots + i_k = n} p_{i_1} \dots p_{i_k} (a_1 \cdot \dots \cdot a_{i_1}) \otimes \dots \otimes (a_{i_1 + \dots + i_{k-1} + 1} \cdot \dots \cdot a_n). \end{aligned}$$

This last formula describes the embedding of  $\mathbf{Coalg}$  into  $\widehat{\mathbf{WQSym}}$  induced by Schur-Weyl duality (the tensor product  $(a_1 \cdot \dots \cdot a_{i_1}) \otimes \dots \otimes (a_{i_1 + \dots + i_{k-1} + 1} \cdot \dots \cdot a_n)$  corresponding to the nondecreasing surjection from  $[n]$  to  $[k]$  sending

the first  $i_1$  integers to 1, ..., the last  $i_k$  integers to  $k$ ). This embedding is of course different from the one induced by the bijection with  $\prod_{n \geq 1} \mathbf{WQSym}_{n,1}$  and corresponds to the fact that elements in **Coalg** can be represented equivalently by a  $f_P$  or a  $\phi_P$ : the  $f_P$ s are naturally encoded by elements in  $\prod_{n \geq 1} \mathbf{WQSym}_{n,1}$  (Formula (2.1)), whereas the  $\phi_P$ s are most naturally encoded by elements in  $\widehat{\mathbf{WQSym}}$  (Formula (2.2)).

Let us show now that, for arbitrary  $P(X), Q(X) \in X\mathbf{Q}[[X]]$ ,

$$(2.3) \quad \phi_P \circ \phi_Q = \phi_{P \circ Q},$$

where  $(P \circ Q)(X) := P(Q(X))$ . For an arbitrary commutative algebra  $A$  and  $a_1, \dots, a_n \in A$ , we have indeed (with self-explaining notations for the coefficients of  $P$  and  $Q$ )

$$\begin{aligned} & \pi \circ \phi_P \circ \phi_Q(a_1 \dots a_n) \\ &= f_P\left(\sum_{k=1}^n \sum_{i_1 + \dots + i_k = n} q_{i_1} \dots q_{i_k}(a_1 \cdot \dots \cdot a_{i_1}) \otimes \dots \otimes (a_{i_1 + \dots + i_{k-1} + 1} \cdot \dots \cdot a_n)\right) \\ &= \sum_{k=1}^n \sum_{i_1 + \dots + i_k = n} p_k q_{i_1} \dots q_{i_k}(a_1 \cdot \dots \cdot a_n) \\ &= f_{P(Q)}(a_1 \dots a_n) \\ &= \pi \circ \phi_{P(Q)}(a_1 \dots a_n). \end{aligned}$$

Thus,  $\phi_P \circ \phi_Q = \phi_{P \circ Q}$  and the theorem follows.

### 3. Formal diffeomorphisms and $\mathbf{WQSym}$

We have shown that **Coalg** and **Diff** embed naturally into  $\widehat{\mathbf{WQSym}}$  (Formula (2.2)). We have already noticed that they are actually embedded in  $\widehat{\mathbf{IQSym}}$ , where **IQSym** stands for the linear span of nondecreasing surjections. Since a nondecreasing surjection  $f$  from  $[n]$  to  $[k]$  is characterized by the number of elements  $f_i := |f^{-1}(i)|$  in the inverse images of the elements of  $i \in [k]$ , nondecreasing surjections from  $[n]$  to  $[k]$  are in bijection with compositions of  $n$  of length  $k$ , that is, ordered sequences of integers  $f_1, \dots, f_k$  adding up to  $n$ . Said otherwise, nondecreasing surjections are naturally in bijection with a linear basis of **Sym** and **QSym**, respectively the Hopf algebra of noncommutative symmetric functions and the dual Hopf algebra of quasi-symmetric functions, see [6], although the linear embeddings of **Sym** and **QSym** into  $\mathbf{WQSym}$  induced by the bijection between

the basis and the embedding of **IQSym** into **WQSym** are not standard ones.

We shall return later on these various embeddings into **WQSym**; the present section studies the compatibility relations between the group structure of **Diff**<sub>1</sub> and the coalgebra structure existing on **WQSym**.

Let us recall the relevant definitions. The word realization of **WQSym** to be introduced now will be useful when we will discuss later some of its combinatorial properties. We denote by  $\mathcal{A} = \{a_1 < a_2 < \dots\}$  a denumerable infinite linearly ordered alphabet and by  $\mathcal{A}^*$  the corresponding set of words.

The *packed word*  $u = \text{pack}(w)$  associated with a word  $w \in \mathcal{A}^*$  is obtained by the following process. If  $b_1 < b_2 < \dots < b_r$  are the letters occurring in  $w$ ,  $u$  is the image of  $w$  by the homomorphism  $b_i \mapsto a_i$ . For example, if  $\mathcal{A} = \mathbb{N}^*$ ,  $\text{pack}(3\ 5\ 3\ 8\ 1) = 2\ 3\ 2\ 4\ 1$ . A word  $u$  is said to be *packed* if  $\text{pack}(u) = u$ . We denote by PW the set of packed words. With a word  $u \in \text{PW}$ , we associate the noncommutative polynomial<sup>(2)</sup>

$$(3.1) \quad \mathbf{M}_u(\mathcal{A}) := \sum_{\text{pack}(w)=u} w.$$

For example, restricting  $\mathcal{A}$  to the first five integers,

$$(3.2) \quad \begin{aligned} \mathbf{M}_{13132}(A) = & 13132 + 14142 + 14143 + 24243 + 15152 \\ & + 15153 + 25253 + 15154 + 25254 + 35354. \end{aligned}$$

Packed words  $u = u_1 \dots u_n$  are in bijection with surjections: taking for  $\mathcal{A}$  the set of integers, if  $1, \dots, p$  are the letters occurring in  $u$ , the surjection associated with  $u$  is simply the map from  $[n]$  to  $[p]$  defined by  $f(i) := u_i$ . The  $\mathbf{M}_u$  can therefore be chosen as linear generators of **WQSym**. Since the product of two  $\mathbf{M}_u$ s is a linear combination of  $\mathbf{M}_u$ s, this presentation induces an algebra structure on **WQSym** and an algebra embedding into  $\mathbf{Q}\langle \mathcal{A} \rangle$ . This algebra structure on **WQSym** is closely related to the Hopf algebra structure of  $\text{QSh}(A)$ , see [14] for details. Since both presentations of **WQSym** ( $\mathbf{M}_u$  or surjections) are equivalent, we will not distinguish between them, except notationally.

As for classical symmetric functions, the nature of the ordered alphabet  $\mathcal{A}$  chosen to define word quasi-symmetric functions  $\mathbf{M}_u(\mathcal{A})$  is largely irrelevant provided it is denumerable infinite. We will therefore often omit the

---

<sup>(2)</sup> As is customary in this theory, “polynomial” means a formal series of bounded degree in an infinite number of noncommuting variables, where each monomial of finite degree may carry a non zero scalar coefficient. We still denote by  $\mathbf{Q}\langle \mathcal{A} \rangle$  the corresponding algebra. Since these are elements of a projective limit of polynomial rings, purists may want to call these objects (*noncommutative*) *polynomials*.

$\mathcal{A}$ -dependency and write simply  $\mathbf{M}_u$  for  $\mathbf{M}_u(\mathcal{A})$ , except when we want to emphasize this dependency (and similarly for the other types of generalized symmetric functions we will have to deal with).

The important point for us now is that **WQSym** carries naturally a Hopf algebra structure for the coproduct:

$$(3.3) \quad \Delta(\mathbf{M}_u) := \sum_{i=0}^n \mathbf{M}_{u_{|[1,i]}} \otimes \mathbf{M}_{\text{pack}(u_{|[i+1,n]})}.$$

Here,  $u$  is a packed word over the letters  $1, \dots, n$  and, for an arbitrary subset  $S$  of  $[n]$ ,  $u_{|_S}$  stands for the word obtained from  $u$  by erasing all the letters that do not belong to  $S$ .

LEMMA 3.1. — *The direct sum of the spaces of nondecreasing surjections, and of the scalars, **IQSym**, is a subcoalgebra of **WQSym**. It is a cofree coalgebra, cogenerated by the set  $\Gamma \cong \mathbb{N}^+$  of “elementary” surjections  $\gamma_n$  from  $[n]$  to  $[1]$ ,  $n \geq 1$ . That is, as a coalgebra, **IQSym** identifies with  $T(\Gamma)$  (the linear span of words over the alphabet of elementary surjections) equipped with the deconcatenation coproduct.*

In particular, **IQSym** is isomorphic as a coalgebra to **QSym**, the coalgebra of quasi-symmetric functions [7, 11, 20] and the embedding of **IQSym** in **WQSym** induces a coalgebra embedding of **QSym** in **WQSym**.

Recall that coalgebra means here coaugmented counital conilpotent coalgebra, so that the cofree coalgebra over a generating set  $X$  identifies with the tensor algebra over  $\mathbf{Q}X$  equipped with the deconcatenation coproduct. The Lemma follows then from the definition of  $\Delta$ . Indeed, if we write  $\gamma_{i_1} \dots \gamma_{i_k}$  for the unique surjection  $f$  from  $[i_1 + \dots + i_k]$  to  $[k]$  such that  $f(1) = \dots = f(i_1) := 1, \dots, f(i_1 + \dots + i_{j-1} + 1) = \dots = f(i_1 + \dots + i_k) := k$ , then,

$$\Delta(\gamma_{i_1} \dots \gamma_{i_k}) = \sum_{j=0}^k \gamma_{i_1} \dots \gamma_{i_j} \otimes \gamma_{i_{j+1}} \dots \gamma_{i_k}.$$

PROPOSITION 3.2. — *The embedding of  $\widehat{\mathbf{Diff}}_1$  into  $\widehat{\mathbf{IQSym}}$  factorizes through the set of grouplike elements in  $\widehat{\mathbf{IQSym}}$  (the same statement holds if we replace **IQSym** by the isomorphic coalgebra **QSym**). In other terms, the group structure of  $\widehat{\mathbf{Diff}}_1$  is compatible with the coalgebra structure of **WQSym** (resp. **QSym**).*

Let  $P(X) = X + \sum_{i>1} p_i X^i$  and  $p_1 := 1$ . Then, in the basis  $\mathbf{M}_u$ ,

$$\phi_P = \sum_{n \geq 0} \sum_{k=1}^n \sum_{i_1 + \dots + i_k = n} p_{i_1} \dots p_{i_k} M_{1^{i_1} \dots k^{i_k}},$$

where  $1^{i_1} \dots k^{i_k}$  stands for the nondecreasing packed word with  $i_1$  copies of 1, ...,  $i_k$  copies of  $k$  (so that e.g.  $1^3 2^2 3^2 = 1112233$ ) with the convention that the component  $n = 0$  of the sum contributes  $1 \in \mathbf{Q}$  (this corresponds to  $\phi_P(1) = 1$ ). We get

$$\begin{aligned} \Delta(\phi_P) &= \sum_{n \geq 0} \sum_{k=1}^n \sum_{a+b=k} \sum_{i_1+\dots+i_k=n} p_{i_1} \dots p_{i_a} M_{1^{i_1} \dots a^{i_a}} \\ &\quad \otimes p_{i_{a+1}} \dots p_{i_k} M_{1^{i_{a+1}} \dots (b-a)^{i_k}} \\ &= \phi_P \otimes \phi_P. \end{aligned}$$

Notice that, as a by-product of Thm. 2.2 and Prop. 3.2, we have characterized the grouplike elements  $\phi$  in  $\widehat{\mathbf{WQSym}}$  which are natural coalgebra endomorphisms of  $Sh$ , that is, which satisfy for an arbitrary commutative algebra  $A$  and elements  $a_1, \dots, a_n \in A$ , the compatibility property

$$(3.4) \quad (\Delta(\phi) \circ \Delta)(a_1 \dots a_n) = (\Delta \circ \phi)(a_1 \dots a_n).$$

where  $\Delta(\phi) = \phi \otimes \phi$ . In general, for  $\phi$  in  $\widehat{\mathbf{WQSym}}$  (not necessarily group-like), equation (3.4) cannot hold if  $\phi \notin \widehat{\mathbf{IQSym}}$  (this is because the deconcatenation coproduct  $\Delta$  preserves the relative ordering of the  $a_i$ s). The converse statement is true and follows from the definition of the coproduct on  $\mathbf{WQSym}$  (its proof is left to the reader):

LEMMA 3.3. — Equation (3.4) holds for all  $\phi \in \widehat{\mathbf{IQSym}}$ . This property characterizes  $\widehat{\mathbf{IQSym}}$  as a subspace of  $\widehat{\mathbf{WQSym}}$ .

PROPOSITION 3.4. — We define an internal product  $\circ$  on  $\mathbf{IQSym}$  in the following way:

- On  $\mathbf{IQSym}_+$ , this is the composition of surjections.
- On  $\mathbf{Q} = \mathbf{IQSym}_0 = \mathbf{IQSym} \cap \mathbf{WQSym}_0$ , this is the usual product.
- The product of elements in  $\mathbf{IQSym}^+$  and  $\mathbf{Q}$  is null.

This internal product provides  $\mathbf{IQSym}$  with a bialgebra<sup>(3)</sup> structure. That is, the coproduct  $\Delta$  and the composition of surjections  $\circ$  satisfy

$$(3.5) \quad \Delta(f \circ g) = \Delta(f) \circ \Delta(g).$$

Indeed, by nonlinear Schur-Weyl duality, the identity is true if and only if the following identity holds for arbitrary  $A$ , and  $a_1, \dots, a_n$  in  $A$

$$\Delta(f \circ g) \circ \Delta(a_1 \dots a_n) = \Delta(f) \circ \Delta(g) \circ \Delta(a_1 \dots a_n),$$

---

<sup>(3)</sup>By “bialgebra”, we simply mean in the present article a compatible product and coproduct as in identity (3.5), without requiring extra properties (very often one requires the coproduct and the product to have also compatibility properties with the unit of the algebra and the counit of the coalgebra).

which follows from the previous Lemma and from the stability of **IQSym** by composition (the composition of two nondecreasing surjections is a nondecreasing surjection).

The identity map  $I \in \widehat{\mathbf{IQSym}}$  (the sum of all identity maps on the finite sets  $[n]$ ) is a unit for  $\circ$  and behaves as a grouplike element for  $\Delta$ . However,  $I \notin \mathbf{IQSym}$  and problems arise if one tries to provide **IQSym** with a classical *graded* unital bialgebra structure. A natural gradation of **IQSym** would be the relative degree of surjections (since the relative degree of the composition of two surjections  $f$  and  $g$  is the sum of their relative degrees). However, for this gradation, each graded component of **IQSym** is infinite dimensional (even in degree 0), so that many of the usual arguments regarding graded Hopf algebras do not apply directly to **IQSym**.

Since we aim at providing a group-theoretical picture of the theory of shuffle algebras, our final goal is to understand the fine structure of the image of **Diff** in  $\widehat{\mathbf{IQSym}}$ . The next section aims at clarifying these questions.

#### 4. Natural coderivations and the Faà di Bruno algebra

In the previous section, we have characterized the natural coalgebra endomorphisms of shuffle algebras, or, equivalently, grouplike elements in  $\widehat{\mathbf{IQSym}}$ . We have also shown that the tangent-to-identity elements form a group isomorphic to the group of tangent-to-identity formal diffeomorphisms. We are going to study now the corresponding Lie algebra  $L$  of natural coderivations of shuffle algebras.

Recall the canonical isomorphism  $\mathbf{IQSym} \cong T(\mathbf{Q}\Gamma)$  (with  $\Gamma \cong \mathbb{N}^*$ ). Let  $S(\Gamma)$  stand for the subspace of symmetric tensors in  $T(\mathbf{Q}\Gamma)$  and  $\mathbf{SIQSym}$  be the corresponding subspace of **IQSym**. By construction, the embedding of **Diff** into **IQSym** factorizes through **SIQSym** (see Eq. (2.2)). Since symmetric tensors form a subcoalgebra of the tensor algebra for the deconcatenation coproduct and since the composition of two elements in **SIQSym** is still in **SIQSym**, the following Lemma is a consequence of our previous results.

LEMMA 4.1. — *The embedding of **SIQSym** into **IQSym** is an embedding of bialgebras (for the composition product).*

Let us call tangent-to-identity an element  $\mu$  in  $\widehat{S(\Gamma)}$  if  $\mu - I$  is a (possibly infinite) linear combination of nondecreasing strict surjections (nondecreasing surjections from  $[n]$  to  $[m]$  with  $n > m$ ).

A coderivation  $D$  in  $\text{Sh}(A)$  (that is, a linear endomorphism such that  $\Delta \circ D = (D \otimes I + I \otimes D) \circ \Delta$ , where  $I$  stands for the identity map) is called infinitesimal if its restriction to  $A \subset \text{Sh}(A)$  is the null map. As usual, a natural coderivation of the shuffle algebras is a family of coderivations (of the  $\text{Sh}(A)$ ) commuting with the morphisms induced by algebra maps (from  $A$  to  $B$ , where  $A$  and  $B$  run over commutative nonunital algebras).

LEMMA 4.2. — *Natural tangent-to-identity coalgebra endomorphisms of shuffle algebras identify with tangent-to-identity grouplike elements in  $\widehat{\text{SIQSym}} \cong \widehat{S(\Gamma)}$ . The corresponding Lie algebra of primitive elements in  $\widehat{S(\Gamma)}$  is the Lie algebra of natural infinitesimal coalgebra coderivations of the shuffle algebras. It is canonically in bijection with the Lie algebra of formal power series  $X^2\mathbf{Q}[[X]]$  equipped with the Lie bracket  $[X^m, X^n] := (m - n)X^{m+n-1}$ .*

The Lemma follows from the general property according to which, in a bialgebra, grouplike elements and primitive elements are in bijection through the logarithm and exponential maps, provided these maps make sense (that is, provided no convergence issue of the series arises). This is because, formally, for  $\phi$  a grouplike element,

$$\Delta(\log(\phi)) = \log(\Delta(\phi)) = \log(\phi \otimes \phi) = \log(\phi) \otimes I + I \otimes \log(\phi),$$

since  $\log(ab) = \log(a) + \log(b)$  when  $a$  and  $b$  commute and since  $I$  is the unit element for the composition product.

The formal convergence of the series under consideration in the present case is ensured by the fact that a surjection from  $n$  to  $p < n$  can be written as the product of at most  $n - p$  strict surjections (so that the coefficient of such a surjection in the expansion of  $\log(\phi)$  is necessarily finite and equal to its coefficient in the expansion of the truncation of the logarithmic series at order  $n - p$ ).

The last statement of the Lemma follows from the isomorphism between tangent-to-identity coalgebra endomorphisms of shuffle algebras and tangent-to-identity formal diffeomorphisms. It can also be deduced directly from the adjunction property (dual to the one according to which derivations in the tensor algebra are in bijection with linear morphisms from  $V$  to  $T(V)$ ): writing  $\text{Coder}^+(\text{Sh}(A))$  for the coderivations of  $\text{Sh}(A)$  vanishing on  $\mathbf{Q}$ , we have  $\text{Coder}^+(\text{Sh}(A)) \cong \text{Lin}(\text{Sh}^+(A), A)$ , which implies, by nonlinear Schur-Weyl duality that natural coalgebra coderivations of the shuffle algebras are canonically in bijection with  $X\mathbf{Q}[[X]]$ .

This bijection with  $X\mathbf{Q}[[X]]$  can be made explicit: dualizing the formula for the derivation associated with a map  $f : V \rightarrow T(V)$

$$f(v_1 \dots v_n) := \sum_{i=1}^n v_1 \dots v_{i-1} f(v_i) v_{i+1} \dots v_n,$$

we get, for  $P = \sum_{i \geq 1} p_i X^i$

$$D_P(a_1 \dots a_n) = \sum_{i=1}^n \sum_{j=1}^{n-i+1} p_i a_1 \dots a_{j-1} (a_j \cdot \dots \cdot a_{j+i-1}) a_{j+i} \dots a_n.$$

In particular, the restriction of  $\phi_P$  and  $D_P$  to maps from  $\text{Sh}(A)$  to  $A$  agree and are both given by

$$\phi_P(a_1 \dots a_n) = D_P(a_1 \dots a_n) = p_n a_1 \cdot \dots \cdot a_n.$$

The simplest example of a coderivation is for  $P(X) = X$ : it is the degree operator,  $Y := D_X$ ,

$$Y(a_1 \dots a_n) = \sum_{i=1}^n a_1 \dots a_{i-1} I(a_i) a_{i+1} \dots a_n = n a_1 \dots a_n.$$

Similarly,  $D_{\lambda X}(a_1 \dots a_n) = n \cdot \lambda (a_1 \dots a_n)$ . In general we have, for arbitrary polynomials  $P, Q$  and  $\lambda \in \mathbf{Q}$ ,

$$D_P + \lambda D_Q = D_{P+\lambda Q}.$$

The description of the Lie algebra structure on  $X\mathbf{Q}[[X]]$  induced by the isomorphism with the Lie algebra of natural coalgebra coderivations follows from the explicit formula for the action of coderivations ( $[D_{X^m}, D_{X^n}] = (m-n)D_{X^{m+n-1}}$ ) but can also be deduced from the fact that composition of coalgebra endomorphisms is reflected in the composition of formal power series. Recall indeed that the set of formal power series  $X + X^2\mathbf{Q}[[X]]$  equipped with the composition product is the group of characters of the Faà di Bruno Hopf algebra, see e.g. [4].

Summarizing our previous results, we get:

**THEOREM 4.3.** — *The Lie algebra of natural infinitesimal coalgebra coderivations of shuffle algebras is naturally isomorphic to (the completion of) the Lie algebra generated by the  $X^n$ ,  $n > 1$ , with Lie bracket  $[X^m, X^n] := (m - n)X^{m+n-1}$ . Equivalently, it is isomorphic to (the completion of) the Lie algebra of primitive elements in the Hopf algebra dual to the Faà di Bruno Hopf algebra.*

Here, “completion” is understood with respect to the underlying implicit grading of these Lie algebras and Hopf algebras (e.g.  $X^n$  is naturally of degree  $n - 1$ , see [4] for details).

### 5. Deformations of shuffles

Let us restrict again our attention to tangent-to-identity coalgebra automorphisms of the  $\text{Sh}(A)$ . Any such  $\Phi_P$  (with  $P(X) - X \in X^2\mathbf{Q}[[X]]$ ) defines a natural deformation of  $Sh$ , that is, a new functor from commutative algebras to Hopf algebras

$$\text{Sh}_P(A) = (\text{Sh}(A), \Delta, \sqcup_P),$$

that is,  $\text{Sh}_P(A)$  identifies with  $\text{Sh}(A)$  (and with  $T(A)$  equipped with the deconcatenation coproduct) as a coalgebra, but carries a new product defined by conjugacy

$$x \sqcup_P y := \phi_P(\phi_P^{-1}(x) \sqcup \phi_P^{-1}(y)).$$

DEFINITION 5.1. — *The Hopf algebra  $\text{Sh}_P(A)$  is called the  $P$ -twisted shuffle algebra. It is isomorphic to  $\text{Sh}(A)$  as a Hopf algebra.*

It inherits therefore all the properties of  $\text{Sh}(A)$ . The reader is referred to Reutenauer’s book [18] for a systematic study of shuffle algebras. As an algebra,  $\text{Sh}(A)$  is, for example, a free commutative algebra over a set of generators parametrized by Lyndon words.

The fundamental example of a deformation is provided by the “ $q$ -exponential” map

$$E_q := \sum_{n \in \mathbb{N}^*} \frac{q^{n-1} x^n}{n!}$$

which interpolates between the identity  $E_0 = x$  and  $E_1 = e^x - 1$ . The corresponding isomorphism between  $\text{Sh}(A)$  and  $\text{Sh}_{E_q}(A)$  is then given by

$$\phi_{E_q}(a_1 \dots a_n) = \sum_{\mathcal{P}} \frac{q^{n-k}}{P_1! \dots P_k!} a_{P_1} \dots a_{P_k},$$

where  $\mathcal{P} = (P_1, \dots, P_k)$  runs over the nondecreasing partitions of  $[n]$  ( $P_1 \amalg \dots \amalg P_k = [n]$  and  $P_i < P_j$  if  $i < j$ ;  $a_{P_i} := \prod_{j \in P_i} a_j$  and  $P_1!$  is a shortcut for  $|P_1|!$ ).

LEMMA 5.2. — When  $q = 1$ ,  $\text{Sh}_{E_1}(A)$  identifies with  $\text{QSh}(A)$ , the quasi-shuffle algebra over  $A$ , whose product, written  $\boxplus$ , is defined recursively by

$$\begin{aligned} a_1 \dots a_n \boxplus b_1 \dots b_m &:= a_1(a_2 \dots a_n \boxplus b_1 \dots b_m) + b_1(a_1 \dots a_n \boxplus b_2 \dots b_m) \\ &+ (a_1 \cdot b_1)(a_2 \dots a_n \boxplus b_2 \dots b_m). \end{aligned}$$

The exponential map  $\phi_{E_1}$  generalizes the Hoffman isomorphism between  $\text{Sh}(A)$  and  $\text{QSh}(A)$  introduced and studied in [9] in the case where  $A$  is a locally finite dimensional graded connected algebra. Using a graded connected commutative algebra  $A$  (instead of an arbitrary commutative algebra as in the present article), although a strong restriction in view of applications, has some technical advantages: it allows, for example, to treat directly  $\text{QSh}(A)$  as a graded connected Hopf algebra, making possible the use of structure theorems for such algebras (Cartier-Milnor-Moore, Leray...). The classical illustration of these phenomena is provided by the algebra of quasi-symmetric functions (the quasi-shuffle algebra over the monoid algebra of the positive integers) and the dual algebra of noncommutative symmetric functions: using the gradation on  $\mathbf{QSym}$  induced by the one of the integers,  $\mathbb{N}^*$ , the exponential/logarithm transform amounts then to a mere change of basis (between a family of grouplike vs primitive generators) see [7, 6, 9] for details.

The proof amounts to showing that

$$\phi_{E_1}(a_1 \dots a_n \sqcup a_{n+1} \dots a_{n+m}) = \phi_{E_1}(a_1 \dots a_n) \boxplus \phi_{E_1}(a_{n+1} \dots a_{n+m}).$$

Let us write  $R_1 \dots R_k$  for an arbitrary partition of  $[n + m]$  such that for  $1 \leq i < j \leq n$  or  $n + 1 \leq i < j \leq n + m$   $i \in R_p$ ,  $j \in R_q \Rightarrow p \leq q$ . The problem amounts to computing the coefficient of  $a_{R_1} \dots a_{R_k}$  in the expansion of the left and right-hand sides of the equation. We leave to the reader the verification that only such tensors appear in these expansions. The coefficient is in both cases  $\frac{1}{P_1! \dots P_k!} \frac{1}{Q_1! \dots Q_k!}$ , where  $P_i := R_i \cap [n]$ ,  $Q_i := R_i \cap \{n + 1, \dots, n + m\}$ . This is straightforward for the right-hand side (the reader not familiar with quasi-shuffle products is encouraged to write down the tedious but straightforward details of the proof -using e.g. the recursive definition of  $\boxplus$ ). For the left-hand side, it follows from the identity  $\binom{|R_i|}{|P_i|} \times \frac{1}{R_i!} = \frac{1}{P_i! Q_i!}$  and the fact that the number of words in the expansion of a shuffle product  $x_1 \dots x_l \sqcup y_1 \dots y_k$  is  $\binom{l+k}{k}$ .

As we shall see in the sequel, even in the well-known special case of a graded algebra, the point of view developed in the present article is not without interest: being more general and conceptual than usual approaches to quasi-shuffle algebras, it allows the derivation of new insights on the fine structure of their operations, refining the results already obtained in [14].

Interesting new phenomena do actually occur as soon as one considers natural linear endomorphisms of shuffle algebras. We have already recalled from [14] that, by Schur-Weyl duality, they belong to **WQSym**, which inherits an associative (convolution) product from the Hopf algebra structure of the  $\text{Sh}(A)$ : for  $f \in \mathbf{WQSym}_n$ ,  $g \in \mathbf{WQSym}_m$  and arbitrary  $a_1, \dots, a_{n+m} \in A$ , where  $A$  is an arbitrary commutative algebra,

$$f \sqcup g(a_1 \dots a_{n+m}) := f(a_1 \dots a_n) \sqcup g(a_{n+1} \dots a_{n+m}).$$

When  $f$  and  $g$  belong to **FQSym**, the subset of permutations in **WQSym**, this convolution product has a simple expression and defines the “usual” product on **FQSym**, see e.g. [11, 2]. Writing  $\text{Sh}_{n,m}$  for the set of  $(n, m)$ -shuffles (that is the elements  $\sigma$  in the symmetric group  $S_{n+m}$  of order  $n + m$  such that  $\sigma(1) < \dots < \sigma(n)$  and  $\sigma(n + 1) < \dots < \sigma(n + m)$ ), we get

$$f \sqcup g = \sum_{\zeta \in \text{Sh}_{n,m}} \zeta \circ (f \cdot g),$$

where  $f \cdot g$  stands for the “concatenation” of permutations:  $f \cdot g(i) := f(i)$  for  $i \leq n$  and  $f \cdot g(i) := n + g(i - n)$  else.

Recall that word and free quasisymmetric functions (**WQSym** and **FQSym**) carry a coproduct, defined on packed words  $f$  over  $k$  letters by

$$\Delta(f) := \sum_{i=0}^k f|_{\{1, \dots, i\}} \otimes \text{pack}(f|_{\{i+1, \dots, k\}}),$$

this coproduct together with  $\sqcup$  defines a graded Hopf algebra structure on both **WQSym** and **FQSym** (the grading is then defined by requiring a surjection from  $[n]$  to  $[p]$  to be of degree  $n$ ). In particular, the embedding  $\mathbf{FQSym} \subset \mathbf{WQSym}$  is an embedding of Hopf algebras for these structures. Recall however that  $\sqcup$  is not the usual product used when studying **WQSym**, see below for details.

The following lemma is instrumental and will prove quite useful. Its proof is left to the reader.

LEMMA 5.3. — For  $f$  a nondecreasing surjection ( $f \in \mathbf{IQSym}$ ) and  $g \in \mathbf{WQSym}$ ,

$$\Delta(f \circ g) = \Delta(f) \circ \Delta(g).$$

Notice that the Lemma would not hold with  $f$  arbitrary in **WQSym**.

From our previous considerations, any formal power series in  $X\mathbf{Q}[[X]]$  will allow us to define a new convolution product on **WQSym** associated

with the P-twisted Hopf algebra structure of the  $\text{Sh}_P(A)$ . This new product, written  $\sqcup_P$  (resp.  $\boxplus$  when  $P = E_1$ , resp.  $\sqcup_q$  when  $P = E_q$ ) is defined by

$$\forall f \in WQSym_n, g \in WQSym_m,$$

$$f \sqcup_P g(a_1 \dots a_{m+n}) := f(a_1 \dots a_n) \sqcup_P g(a_{n+1} \dots a_{m+n}),$$

(the product  $f \sqcup_P g$  acts as the null map on tensors of length different from  $n + m$ ).

The following result, although elementary, is stated as a theorem in view of its importance:

**THEOREM 5.4.** — *For an arbitrary  $P \in X + X^2\mathbf{Q}[[X]]$ , the composition map*

$$f \longmapsto \Phi_P(f) = f_P := \phi_P \circ f$$

*induces an isomorphism of bialgebras from  $(\mathbf{WQSym}, \Delta, \sqcup)$  to  $(\mathbf{WQSym}, \Delta, \sqcup_P)$ . For  $P = E_1$  (resp.  $P = E_q$ ), we will write simply  $\Phi_1(f)$  (resp.  $\Phi_q(f)$ ) for  $\Phi_P(f)$ . This isomorphism is equivariant with respect to the composition product:*

$$\phi_P(f \circ g) = \phi_P(f) \circ g.$$

The compatibility with the coproduct follows from Lemma 5.3, from the definition of  $f_P$  as the composition of  $f$  with a sum of nondecreasing surjections, and from the fact that  $\Delta(\phi_P) = \phi_P \otimes \phi_P$ . On an other hand, the definition of the twisted product  $\sqcup_P$  implies

$$f_P \sqcup_P g_P = \phi_P \circ (f \sqcup g) = (f \sqcup g)_P.$$

The following corollaries are motivated by the key role of  $\mathbf{FQSym}$  and  $\mathbf{WQSym}$  in the theory of noncommutative symmetric functions and their various application fields:

**COROLLARY 5.5.** — *Any  $P \in X + X^2\mathbf{Q}[[X]]$  induces a Hopf algebra embedding of  $(\mathbf{FQSym}, \Delta, \sqcup)$  into  $(\mathbf{WQSym}, \Delta, \sqcup_P)$ . This embedding is  $S_n$ -equivariant: for  $\sigma, \beta \in S_n = \mathbf{FQSym}_n$ , we have:*

$$(\sigma \circ \beta)_P = \sigma_P \circ \beta.$$

**COROLLARY 5.6.** — *For  $P = E_1$ , we get that  $(\mathbf{FQSym}, \Delta, \sqcup)$  embeds naturally into  $(\mathbf{WQSym}, \Delta, \boxplus)$ . As in the previous corollary, this embedding is  $S_n$ -equivariant.*

These corollaries allow to give an explicit formula for the embeddings. Indeed, for an arbitrary  $\sigma \in S_n$ , we get (writing  $1_n$  the identity permutation in  $S_n$ )

$$\sigma_P = (1_n)_P \circ \sigma.$$

Since  $(1_n)_P$  is simply the component of  $\phi_P$  in  $\mathbf{WQSym}_n$ , we get finally

$$\sigma_P = \sum_{k=1}^n \sum_{i_1+\dots+i_k=n} p_{i_1}\dots p_{i_k} 1^{i_1}\dots k^{i_k} \circ \sigma,$$

where we write  $1^{i_1}\dots k^{i_k}$  for the surjection sending the first  $i_1$  integers to 1, ..., the integers from  $i_1 + \dots + i_{k-1} + 1$  to  $n$  to  $k$ . For  $E_q$  this formula simplifies:

LEMMA 5.7. — *Let  $\sigma \in S_n$  and  $\tau \in \mathbf{WQSym}_n$ . We shall say that  $\tau \propto \sigma$  if for all  $1 \leq i, j \leq n$ ,  $\sigma(i) < \sigma(j) \Rightarrow \tau(i) \leq \tau(j)$ . We also set  $\tau! := \prod_{i=1}^{\max(\tau)} |\tau^{-1}(\{i\})|!$  and  $r(\tau)$  for the relative degree of  $\tau$ . Then, the Hopf algebra embedding  $\Phi_P$  from  $\mathbf{FQSym}$  into  $\mathbf{WQSym}$  is given by*

$$\Phi_P(\sigma) = \sum_{\tau \propto \sigma} \frac{q^{r(\tau)} \tau}{\tau!}.$$

Other consequences of the existence of such embeddings will be drawn in the sequel.

### 6. Structure of twisted shuffle algebras

The map  $\phi_P$  defines an isomorphism from  $\text{Sh}(A)$  to  $\text{Sh}_P(A)$  for an arbitrary commutative algebra  $A$  and an isomorphism between  $\mathbf{WQSym}$  equipped with the shuffle product  $\sqcup$  to  $\mathbf{WQSym}$  equipped with the twisted shuffle product  $\sqcup_P$ .

In this section, we briefly develop the consequences of these isomorphisms and recover, among others, the results of [14] on projections onto the indecomposables in  $\text{QSh}(A)$ .

Recall from [17, 16] that  $e_1 := \log^{\sqcup}(Id)$  (the logarithm of the identity map of  $\text{Sh}(A)$  computed using the shuffle product  $\sqcup$ ) is a canonical section of the projection from  $\text{Sh}(A)$  to the indecomposables  $\text{Sh}(A)^+ / (\text{Sh}(A)^+)^2$ . In particular, due to the structure theorems for graded connected Hopf algebras over a field of characteristic 0 (Leray, in that particular case),  $\text{Sh}(A)$  is a free commutative algebra over the image of  $e_1$ . Equivalently,  $e_1$  is the projection on the eigenspace of eigenvalue  $k$  of the  $k$ -th Adams operation, that is, the  $k$ -th power of the identity  $(Id)^{\sqcup k}$ .

These properties are clearly invariant by conjugacy and we get, since  $\phi_P \circ Id \circ \phi_P^{-1} = Id$ , the following description of  $\text{Sh}_P(A)$  as a free commutative algebra:

PROPOSITION 6.1. — For an arbitrary tangent-to-identity  $P$ ,  $e_1 := \log^{\sqcup^P}(Id)$  (the logarithm of the identity for the  $\sqcup_P$  product) is a section of the projection from  $\text{Sh}_P(A)$  to the indecomposables  $\text{Sh}_P(A)^+ / (\text{Sh}_P(A)^+)^2$ . In particular,  $\text{Sh}_P(A)$  is a free commutative algebra over the image of  $e_1$ . Equivalently,  $e_1$  is the projection on the eigenspace associated with the eigenvalue  $k$  of the  $k$ -th Adams operation, that is, the  $k$ -th power of the identity  $(Id)^{\sqcup_P^k}$ .

The particular case of the quasi-shuffle algebra was investigated in [14]. The projection  $e_1$  can then be computed explicitly. Recall that a surjection  $f$  from  $[n]$  to  $[p]$  has a descent in position  $i$  if and only if  $f(i) > f(i + 1)$ . Let us call conjugate surjection and write  $\tilde{f}$  for the surjection from  $[n]$  to  $[p]$  defined by:  $\tilde{f}(i) := f(n - i)$ . We have:

PROPOSITION 6.2. — In **WQSym** equipped with the quasi-shuffle  $\boxplus$ ,

$$e_1 := \log^{\boxplus}(Id) = \sum_{n \geq 1} \frac{1}{n} \sum_{I \models n} \frac{(-1)^{l(I)-1}}{\binom{n-1}{l(I)-1}} \sum_{\text{Des}(f)=[n]-\{i_{l(I)}, \dots, i_{l(I)+\dots+i_1\}} \tilde{f},$$

where  $I \models n$  means that  $I = (i_1, \dots, i_{l(I)})$  is a composition of  $n$ .

To start investigating the word interpretation of **WQSym** and the combinatorial meaning of the formulas and results obtained so far, recall that, as any formula regarding **WQSym**, this proposition can be translated into a result on words (and actually also into a result on Rota–Baxter algebras, due to the relationship established in [10] between **WQSym** and free Rota–Baxter algebras).

Recall that the elements of **WQSym** can be realized as formal sums of words over a totally ordered alphabet  $X$ . For example, the identity map  $\text{Id} \in \widehat{\mathbf{WQSym}}$  identifies under this correspondence with the formal sum of all nondecreasing words over  $X$ .

When the alphabet is taken to be the sequence of values of a function from  $[n - 1]$  into an associative algebra, this formal sum identifies with the  $n$ -th value of the function (unique, formal) solution to the recursion  $F = 1 + S(F \cdot f)$ , where  $S$  is the summation operator  $S(f)(j) := \sum_{i=1}^{j-1} f(i)$ ,  $S(f)(1) := 0$ . That is,

$$F(k) = 1 + \sum_{i=1}^{k-1} \sum_{1 \leq a_1 < \dots < a_k \leq k-1} f(a_1) \dots f(a_k).$$

Since the concatenation product of words induces an associative product on **WQSym** that identifies with  $\boxplus$  [2], we get finally that  $e_1$  computes in

that case  $\log(F)(n)$ . This phenomenon was studied recently in [3], to which we refer for details.

### 7. Graduations

In the context of shuffle algebras, the conjugacy map by  $\phi_P$ , for an arbitrary tangent-to-identity  $P$ , maps the degree operator  $Y$  on  $\text{Sh}(A)$  to a degree operator  $Y_P := \phi_P \circ Y \circ \phi_P^{-1}$  on  $\text{Sh}_P(A)$ . That is, more explicitly:

PROPOSITION 7.1. — *The operator  $Y_P := \phi_P \circ Y \circ \phi_P^{-1}$  acting on  $\text{Sh}_P(A)$  is a derivation and a coderivation which leaves invariant the subspaces  $\text{Sh}_P^{\leq n}(A) := \bigoplus_{i \leq n} A^{\otimes i}$ . Its action is diagonalizable, with eigenvalues  $i \in \mathbb{N}$ .*

*The eigenspaces for the eigenvalues 0 and 1 are the scalars, resp.  $A$ . In general, the eigenspace for the eigenvalue  $n$  is contained in  $\text{Sh}_P^{\leq n}(A)$  and its intersection with  $\text{Sh}_P^{\leq n-1}(A)$  is the null vector space, more precisely:*

$$\text{Sh}_P^{\leq n}(A) = \text{Sh}_P^{\leq n-1}(A) \oplus \text{Ker}(Y_P - nId).$$

Concretely, conjugacy by  $\phi_P$  defines an isomorphism between  $A^{\otimes n} \subset \text{Sh}(A)$  and the eigenspace associated with the eigenvalue  $n$  of  $Y_P$  in  $\text{Sh}_P(A)$ . The conjugacy map can be described explicitly as follows:

PROPOSITION 7.2. — *For  $U \in X + X^2\mathbf{Q}[[X]]$  (or more generally in  $\mathbf{Q}^*X + X^2\mathbf{Q}[[X]]$ ) and  $V$  an arbitrary formal power series without constant term,*

$$\phi_U^{-1} \circ D_V \circ \phi_U = D_{\frac{V \circ U}{U'}}.$$

By linearity of  $D$  and (formal) continuity of the action by conjugacy, it is enough to prove the formula when  $V = X^p$ . We denote by  $W$  the inverse of  $U$  for the composition and write  $u_i$  the coefficients of  $U$ , and similarly for  $V$  and  $W$ . Then,

$$\begin{aligned} & \pi \circ \phi_W \circ D_{X^p} \circ \phi_U(a_1 \dots a_n) \\ &= f_W \circ D_{X^p} \left( \sum_{k=1}^n \sum_{i_1 + \dots + i_k = n} u_{i_1} \dots u_{i_k} (a_{i_1} \dots a_{i_1}) \dots (a_{i_1 + \dots + i_{k-1} + 1} \dots a_n) \right) \\ &= \sum_{k=p}^n \sum_{i_1 + \dots + i_k = n} (k - p + 1) w_{k-p+1} u_{i_1} \dots u_{i_k} (a_1 \dots a_n), \end{aligned}$$

so that  $\pi \circ \phi_W \circ D_{X^p} \circ \phi_U$  is the linear map associated to the formal power series:

$$\sum_{k=p}^{\infty} (k - p + 1) w_{k-p+1} U^k = \left( \sum_{k=p}^{\infty} (k - p + 1) w_{k-p+1} X^k \right) \circ U$$

$$= \left( \sum_{i=1}^{\infty} i w_i X^{i-1+p} \right) \circ U = (X^p W') \circ U = U^p \cdot (W' \circ U) = \frac{U^p}{U'}.$$

Hence,

$$\phi_U^{-1} \circ D_{X^p} \circ \phi_U = D_{\frac{U^p}{U'}},$$

and the proposition follows.

In particular, taking  $P = E_1$ , we get:

PROPOSITION 7.3. — *The eigenspaces of the coderivation  $D_{(1+X)\ln(1+X)}$  provide the quasi-shuffle algebras  $\text{QSh}(A)$  with a grading and, more precisely, provide the triple  $(\text{QSh}(A), \uplus, \Delta)$  with the structure of graded connected commutative Hopf algebra.*

Indeed, with  $D = \psi \circ D_X \circ \psi^{-1} = \psi \circ Y \circ \psi^{-1}$  and  $\psi := \phi_{\exp(X)-1}$ ,

$$D = \phi_{\ln(1+X)}^{-1} \circ D_X \circ \phi_{\ln(1+X)} = D_{(1+X)\ln(1+X)}.$$

Notice that, since  $(1 + X) \ln(1 + X) = 1 + \sum_{k=2}^{\infty} \frac{(-1)^k}{k(k+1)} X^k$ ,

$$\begin{aligned} & D_{(1+X)\ln(1+X)}(a_1 \dots a_n) \\ &= n \cdot a_1 \dots a_n + \sum_{i=2}^n \sum_{j=1}^{n-i+1} \frac{(-1)^i}{i(i-1)} a_1 \dots a_{j-1} (a_j \cdot \dots \cdot a_{j+i-1}) a_{j+i} \dots a_n. \end{aligned}$$

### 8. A new realization of $\text{Sym}$

In the last sections, we develop some combinatorial applications of our previous results. We will focus mainly on the consequences of Lemma 5.7, that is, the existence of a new isomorphic embedding of  $\mathbf{FQSym}$  into  $\mathbf{WQSym}$ .

The present section explains briefly how these results translate in terms of *polynomial* realizations of these algebras. In particular, we emphasize that the previous embedding gives rise to *new* realizations of  $\mathbf{Sym}$  and  $\mathbf{FQSym}$  that will appear in the forthcoming section to be meaningful for the combinatorial study of the Hausdorff series.

Recall that  $\mathbf{Sym}$ , the Hopf algebra of noncommutative symmetric functions, is the free associative algebra generated by a sequence  $S_i$ ,  $i \in \mathbb{N}^*$  of divided powers  $(\Delta(S_n) := \sum_{i=0}^n S_i \otimes S_{n-i}$ , where  $S_0 := 1 \in \mathbf{Q}$ ). In [6], it

has been shown that this Hopf algebra could bring considerable simplifications in the analysis of the so-called continuous BCH (Baker-Campbell-Hausdorff) series, the formal series  $\Omega(t) = \log X(t)$  expressing the logarithm of the solution of the (noncommutative) differential equation  $X'(t) = X(t)A(t)$  ( $X(0) = 1$ ) as iterated integrals of products of factors  $A(t_i)$ . The new polynomial realization of **Sym** to be introduced will lead instead to a straightforward proof of Goldberg's formula for the coefficients of the (usual) Hausdorff series.

This algebra **Sym** can be embedded as a Hopf subalgebra into **FQSym** by sending  $S_n$  to the identity element in the symmetric group of order  $n$ . This results into the usual polynomial realization of **Sym** introduced in [6] (that is, its realization through an embedding into the algebra of noncommutative polynomials  $\mathbf{Q}\langle X \rangle$ ) sending  $S_n$  to the sum of all nondecreasing words of length  $n$ . We refer e.g. to [15] for details.

Instead of doing so, we may now take advantage of Lemma 5.7 and define a new polynomial realization of **Sym** using the existence of an isomorphic embedding  $\Phi_1$  of **FQSym** into **WQSym**:

$$(8.1) \quad \hat{S}_n = \sum_{u \text{ nondecreasing, } |u|=n} \frac{1}{u!} \mathbf{M}_u,$$

where for notational convenience we write  $\hat{S}_n$  for  $\Phi_1(S_n)$  (and similarly for the images by  $\Phi_1$  of the other elements of **Sym** and **FQSym** in **WQSym**). In view of Formula (3.1), we have equivalently

$$(8.2) \quad \hat{\sigma}_t := \sum_{n \geq 0} t^n \hat{S}_n = e^{tx_1} e^{tx_2} \dots = \prod_{i \geq 1}^{\rightarrow} e^{tx_i}.$$

Notice that  $\hat{\sigma}_t$  is (as expected) a grouplike element for the standard co-product of noncommutative polynomials for which letters  $x_i$  are primitive.

An interesting feature of this realization is that  $\Phi := \log \hat{\sigma}_1$  is now the Hausdorff series

$$(8.3) \quad \Phi = \log(e^{x_1} e^{x_2} \dots) = H(x_1, x_2, x_3, \dots).$$

Moreover, two nondecreasing words  $v$  and  $w$  such that  $\text{pack}(v) = \text{pack}(w) = u$  have the same coefficient in  $\hat{\sigma}_1$ , that is,

$$(8.4) \quad \frac{1}{u!}, \quad \text{where } u! := \prod_i m_i(u)!$$

and  $m_i(u)$  is the number of occurrences of  $i$  in  $u$ .

The Hausdorff series can now be expanded in the basis  $\mathbf{M}_u$  of  $\mathbf{WQSym}$  as

$$(8.5) \quad \Phi = \sum_u c_u \mathbf{M}_u$$

and one may ask whether the previous formalism can shed any light on the coefficients  $c_u$ . There is actually a formula for  $c_u$ , due to Goldberg [8], and reproduced in Reutenauer’s book [18, Th. 3.11 p. 63]. This formula, which was obtained as a combinatorial *tour de force*, will be shown in the sequel to be a direct consequence of our previous results.

### 9. Goldberg’s formula revisited

Let us fix first some notations. For  $I = (i_1, \dots, i_r)$ , we set  $S^I := S_{i_1} \dots S_{i_r}$ ; we also write  $\ell(I) = r$  and  $|I| = i_1 + \dots + i_r$  (so that  $I \models |I|$ ). By definition,

$$(9.1) \quad \begin{aligned} \Phi &= \log(1 + \hat{S}_1 + \hat{S}_2 + \dots) \\ &= \sum_{r \geq 1} \frac{(-1)^{r-1}}{r} \sum_{\ell(I)=r} \hat{S}^I \\ &= \int_{-1}^0 \left( \sum_I t^{\ell(I)} \hat{S}^I \right) \frac{dt}{t} \end{aligned}$$

so that the coefficient  $c_u$  of  $\mathbf{M}_u$  in the Hausdorff series is, denoting by  $\mathbf{N}_u$  the dual basis of  $\mathbf{M}_u$ ,

$$(9.2) \quad c_u = \int_{-1}^0 \left\langle \mathbf{N}_u, \sum_I t^{\ell(I)} \hat{S}^I \right\rangle \frac{dt}{t}.$$

For  $u$  a word of length  $n$ , we have therefore to evaluate  $\langle \mathbf{N}_u, \hat{A}_n(t) \rangle$  with  $A_n(t) := \sum_{|I|=n} t^{\ell(I)} S^I$ . This last sum is related to a well-known series. The *noncommutative Eulerian polynomials* are defined by [6, Section 5.4]

$$(9.3) \quad A_n(t) = \sum_{k=1}^n \left( \sum_{\substack{|I|=n \\ \ell(I)=k}} R_I \right) t^k = \sum_{k=1}^n \mathbf{A}(n, k) t^k.$$

where  $R_I$  is the ribbon basis (the basis of  $\mathbf{Sym}$  obtained from the  $S^I$  basis by Möbius inversion in the boolean lattice) [6, Section 3.2]. The generating series of the  $\mathcal{A}_n(t)$  is given by

$$(9.4) \quad \mathcal{A}(t) := \sum_{n \geq 0} \mathcal{A}_n(t) = (1-t) (1-t\sigma_{1-t})^{-1},$$

where  $\sigma_{1-t} = \sum (1-t)^n S_n$ . Let  $\mathcal{A}_n^*(t) = (1-t)^{-n} \mathcal{A}_n(t)$ . Then,

$$(9.5) \quad \mathcal{A}^*(t) := \sum_{n \geq 0} \mathcal{A}_n^*(t) = \sum_I \left( \frac{t}{1-t} \right)^{\ell(I)} S^I .$$

and

$$(9.6) \quad \sum_{I \models n} t^{\ell(I)} S^I = A_n(t) = \mathcal{A}_n^* \left( \frac{t}{1+t} \right) = (1+t)^n \mathcal{A}_n \left( \frac{t}{1+t} \right) .$$

To evaluate, for a packed word  $u$  of length  $n$ , the pairing  $\langle \mathbf{N}_u, \hat{A}_n(t) \rangle$ , let us start with the observation that, if  $u = 1^n$ , then, writing  $\mathcal{F}_\sigma$  for the dual basis to the  $\sigma \in S_n = \mathbf{FQSym}_n$ ,

$$(9.7) \quad \Phi_1^\dagger(\mathbf{N}_u) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \mathcal{F}_\sigma ,$$

where  $\Phi_1^\dagger$  is the adjoint map, so that in this case,

$$(9.8) \quad \begin{aligned} \langle \mathbf{N}_u, \hat{A}_n(t) \rangle &= \langle \Phi_1^\dagger(\mathbf{N}_u), A_n(t) \rangle \\ &= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} t^{d(\sigma)+1} (1+t)^{r(\sigma)} \\ &= \frac{1}{n!} t E_n(t, t+1) \end{aligned}$$

where  $d(\sigma)$  is the number of descents of  $\sigma$ ,  $r(\sigma) = n - d(\sigma)$  the number of rises, and  $E_n$  is the homogeneous Eulerian polynomial normalized as in [18]

$$(9.9) \quad E_n(x, y) = \sum_{\sigma \in \mathfrak{S}_n} x^{d(\sigma)} y^{r(\sigma)} .$$

Now, recall that the coproduct of  $\mathbf{N}_u$  dual to the product of the  $\mathbf{M}_u$  is [13]

$$(9.10) \quad \Delta \mathbf{N}_u = \sum_{u=u_1 u_2} \mathbf{N}_{\text{pack}(u_1)} \otimes \mathbf{N}_{\text{pack}(u_2)}$$

(deconcatenation). We can omit the packing operation in this formula if we make the convention that  $\mathbf{N}_w = \mathbf{N}_u$  if  $u = \text{pack}(w)$ . Then, since  $\Phi_1$ , and hence also  $\Phi_1^\dagger$  are morphisms of Hopf algebras, for a composition  $L = (l_1, \dots, l_p)$ ,

$$(9.11) \quad \begin{aligned} \langle \Phi_1^\dagger(\mathbf{N}_u), S^L \rangle &= \langle \Delta^{[k]}(\mathbf{N}_u), S_{l_1} \otimes \dots \otimes S_{l_p} \rangle \\ &= \prod_{k=1}^p \langle \Phi_1^\dagger(\mathbf{N}_{u_k}), S_{l_k} \rangle \end{aligned}$$

where  $\Delta^{[k]}$  is the  $k$ -th iterated coproduct and  $u = u_1 u_2 \cdots u_p$  with  $|u_k| = l_k$  for all  $k$ . Moreover, this is nonzero only if all the  $u_k$  are nondecreasing, in which case the result is  $1/(u_1! \cdots u_p!)$ .

Thus, if

$$(9.12) \quad u = w_1 \cdots w_m$$

is the factorization of  $u$  into maximal nondecreasing words, with  $|w_k| = n_k$ , we have

$$(9.13) \quad \langle \Phi_1^\dagger(\mathbf{N}_u), A_n(t) \rangle = \prod_{k=1}^m \langle \Phi_1^\dagger(\mathbf{N}_{w_k}), A_{n_k}(t) \rangle$$

since

$$(9.14) \quad \prod_{k=1}^m A_{n_k}(t) = \sum_{I \in C_u} t^{\ell(I)} S^I$$

where  $C_u$  is the set of compositions which are a refinement of  $(n_1, \dots, n_m)$  and are the ones such that  $\langle \Phi_1^\dagger(\mathbf{N}_u), S^I \rangle \neq 0$ .

Next, if  $v = 1^{l_1} 2^{l_2} \cdots p^{l_p}$ ,

$$(9.15) \quad \langle \Phi_1^\dagger(\mathbf{N}_v), S^L \rangle = \prod_{k=1}^p \langle \Phi_1^\dagger(\mathbf{N}_{k^{l_k}}), S^{l_k} \rangle$$

(both sides are equal to  $1/(l_1! \cdots l_p!)$ ), so that

$$(9.16) \quad \langle \Phi_1^\dagger(\mathbf{N}_v), A_{|L|}(t) \rangle = \left(1 + \frac{1}{t}\right)^{r(v)} \prod_{k=1}^p \langle \Phi_1^\dagger(\mathbf{N}_{k^{l_k}}), A_{l_k}(t) \rangle.$$

where  $r(v)$  is the number of different letters (or of strict rises) of  $v$ . Indeed,  $A_{|L|}(t) = \sum_{|I|=|L|} t^{\ell(I)} S^I$  and, since the  $S_I$  are grouplike,

$$\langle \Phi_1^\dagger(\mathbf{N}_u), S^I \rangle = \prod_{k=1}^p \langle \Phi_1^\dagger(\mathbf{N}_k^{l_k}), S^{I|k} \rangle,$$

where  $I|k$  is the partition of  $\{l_1 + \dots + l_{k-1} + 1, \dots, l_1 + \dots + l_k\}$  induced by the partition  $I$  of  $[|L|]$ . Finally, writing  $I \cup L$  for the partition refining  $I$  and  $L$  (obtained, e.g., by gluing the  $I|k$ ), using

$$(9.17) \quad \langle \Phi_1^\dagger(\mathbf{N}_u), t^{\ell(I)} S^I \rangle = t^{\ell(I) - \ell(I \cap L)} \langle \Phi_1^\dagger(\mathbf{N}_u), t^{\ell(I \cap L)} S^{I \cap L} \rangle$$

and noting that  $|\{I \models |L|, I \cap L = K, \ell(I) - \ell(K) = k < r(u)\}| = \binom{r(u)}{k}$ , we get (9.16).

We can now see that if we decompose  $u$  into maximal blocks of identical letters,

$$(9.18) \quad u = i_1^{j_1} i_2^{j_2} \cdots i_s^{j_s}$$

we have finally

$$\begin{aligned} \langle \Phi_1^\dagger(\mathbf{N}_u), A_n(t) \rangle &= \left(1 + \frac{1}{t}\right)^{r(u)} \prod_{k=1}^s \langle \Phi_1^\dagger(\mathbf{N}_{i_k^{j_k}}), A_{j_k}(t) \rangle \\ &= t^{d(u)+1} (1+t)^{r(u)} \prod_{k=1}^s \frac{E_{j_k}(t, 1+t)}{j_k!} \end{aligned}$$

which implies Goldberg’s formula:

THEOREM 9.1. — *The coefficient  $c_u$  of  $\mathbf{M}_u$  in the Hausdorff series  $\Phi$  is given by:*

$$(9.19) \quad c_u = \int_{-1}^0 t^{d(u)+1} (1+t)^{r(u)} \prod_{k=1}^s \frac{E_{j_k}(t, 1+t)}{j_k!} \frac{dt}{t}.$$

More generally, for an arbitrary moment generating function

$$(9.20) \quad f(z) = \sum_{n \geq 1} f_n z^n.$$

with

$$(9.21) \quad f_n = \int_{\mathcal{R}} t^n d\mu(t)$$

the coefficient of  $\mathbf{M}_u$  in  $f(\hat{\sigma}_1)$  is

$$(9.22) \quad \int_{\mathcal{R}} t^{d(u)+1} (1+t)^{r(u)} \prod_{k=1}^s \frac{E_{j_k}(t, 1+t)}{j_k!} d\mu(t).$$

### BIBLIOGRAPHY

- [1] R. E. BLOCK & P. LEROUX, “Generalized dual coalgebras of algebras, with applications to cofree coalgebras”, *J. Pure Appl. Algebra* **36** (1985), no. 1, p. 15-21.
- [2] G. DUCHAMP, F. HIVERT & J.-Y. THIBON, “Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras”, *Internat. J. Algebra Comput.* **12** (2002), no. 5, p. 671-717.
- [3] K. EBRAHIMI-FARD & D. MANCHON, “The tridendriform structure of a discrete magnus expansion”, *Discrete Contin. Dyn. Syst.* **34** (2014), no. 3, p. 1021-1040.
- [4] H. FIGUEROA & J. M. GRACIA-BONDÍA, “Combinatorial Hopf algebras in quantum field theory. I”, *Rev. Math. Phys.* **17** (2005), no. 8, p. 881-976.
- [5] L. FOISSY & F. PATRAS, “Natural endomorphisms of shuffle algebras”, *Internat. J. Algebra Comput.* **23** (2013), no. 4, p. 989-1009.
- [6] I. M. GELFAND, D. KROB, A. LASCoux, B. LECLERC, V. S. RETAKH & J.-Y. THIBON, “Noncommutative symmetric functions”, *Adv. Math.* **112** (1995), no. 2, p. 218-348.
- [7] I. M. GESSEL, “Multipartite  $P$ -partitions and inner products of skew Schur functions”, in *Combinatorics and algebra (Boulder, Colo., 1983)*, Contemp. Math., vol. 34, Amer. Math. Soc., Providence, RI, 1984, p. 289-317.

- [8] K. GOLDBERG, “The formal power series for  $\log e^x e^y$ ”, *Duke Math. J.* **23** (1956), p. 13-21.
- [9] M. E. HOFFMAN, “Quasi-shuffle products”, *J. Algebraic Combin.* **11** (2000), no. 1, p. 49-68.
- [10] E.-F. KURUSCH, J. M. GRACIA-BONDÍA & F. PATRAS, “Rota-Baxter algebras and new combinatorial identities”, *Lett. Math. Phys.* **81** (2007), no. 1, p. 61-75.
- [11] C. MALVENUTO & C. REUTENAUER, “Duality between quasi-symmetric functions and the Solomon descent algebra”, *J. Algebra* **177** (1995), no. 3, p. 967-982.
- [12] J. W. MILNOR & J. C. MOORE, “On the structure of Hopf algebras”, *Ann. of Math.* (2) **81** (1965), p. 211-264.
- [13] J.-C. NOVELLI & J.-Y. THIBON, “Polynomial realizations of some trialgebras”, <http://arxiv.org/abs/math/0605061>, presented at FPSAC’06, 2006.
- [14] J.-C. NOVELLI, F. PATRAS & J.-Y. THIBON, “Natural endomorphisms of quasi-shuffle Hopf algebras”, *Bull. Soc. Math. France* **141** (2013), no. 1, p. 107-130.
- [15] J.-C. NOVELLI & J.-Y. THIBON, “Free quasi-symmetric functions and descent algebras for wreath products, and noncommutative multi-symmetric functions”, *Discrete Math.* **310** (2010), no. 24, p. 3584-3606.
- [16] F. PATRAS, “L’algèbre des descentes d’une bigèbre graduée”, *J. Algebra* **170** (1994), no. 2, p. 547-566.
- [17] F. PATRAS, “La décomposition en poids des algèbres de Hopf”, *Ann. Inst. Fourier (Grenoble)* **43** (1993), no. 4, p. 1067-1087.
- [18] C. REUTENAUER, *Free Lie algebras*, London Mathematical Society Monographs. New Series, vol. 7, The Clarendon Press, Oxford University Press, New York, 1993, Oxford Science Publications, xviii+269 pages.
- [19] M. P. SCHÜTZENBERGER, “Sur une propriété combinatoire des algèbres de Lie libres pouvant être utilisée dans un problème de mathématiques appliquées”, *Séminaire Dubreil–Jacotin Pisot (Algèbre et théorie des nombres)* **12** (1958–1959), no. 1, p. 1-23.
- [20] R. P. STANLEY, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999, With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin, xii+581 pages.

Manuscrit reçu le 6 novembre 2013,  
révisé le 17 mars 2015,  
accepté le 20 mai 2015.

Loïc FOISSY  
Laboratoire de Mathématiques Pures et Appliquées  
Joseph Liouville  
Université du Littoral Côte d’Opale, Centre  
Universitaire de la Mi-Voix  
50, rue Ferdinand Buisson, CS 80699  
62228 Calais Cedex (France)  
foissy@lmpa.univ-littoral.fr

Frédéric PATRAS  
Laboratoire de Mathématiques J.A. Dieudonné  
Université de Nice - Sophia Antipolis  
Parc Valrose  
06108 Nice Cedex 02 (France)  
patras@unice.fr

Jean-Yves THIBON  
Laboratoire d'Informatique Gaspard-Monge,  
Université Paris-Est, 5, boulevard Descartes  
Champs-sur-Marne  
77454 Marne-la-Vallée cedex 2 (France)  
jyt@univ-mlv.fr