# Deformations of shuffles and quasi-shuffles 

Loïc Foissy, Frédéric Patras, Jean-Yves Thibon

## To cite this version:

Loïc Foissy, Frédéric Patras, Jean-Yves Thibon. Deformations of shuffles and quasi-shuffles. Annales de l’Institut Fourier, 2016, 66 (1), pp.209-237. 10.5802/aif.3010 . hal-00880720

HAL Id: hal-00880720

## https://hal.science/hal-00880720

Submitted on 6 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# DEFORMATIONS OF SHUFFLES AND QUASI-SHUFFLES 

LOÏC FOISSY, FRÉDÉRIC PATRAS, JEAN-YVES THIBON


#### Abstract

We investigate deformations of the shuffle Hopf algebra structure $\operatorname{Sh}(A)$ which can be defined on the tensor algebra over a commutative algebra $A$. Such deformations, leading for example to the quasi-shuffle algebra $\operatorname{QSh}(A)$, can be interpreted as natural transformations of the functor Sh, regarded as a functor from commutative nonunital algebras to coalgebras. We prove that the monoid of natural endomophisms of the functor 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 functionsWQSym, and in turn define deformations of its struture. This leads to a new embedding of free quasi-symmetric functions into WQSym, whose relevance is illustrated by a simple and transparent proof of Goldberg's formula for the coefficients of the Hausdorff series.


## 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 quasishuffles. 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., [11]). 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].

[^0]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 naturallity 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.

The authors acknowledge support from the grant CARMA ANR-12-BS01-0017.

## 1. The shuffle algebra over a commutative algebra

Let $A$ be a commutative algebra over the rationals ${ }^{1}$, (not necessarily with a unit) and let $T(A)$ be its tensor algebra

$$
T(A)=\bigoplus_{n \in \mathbb{N}} T_{n}(A)=\bigoplus_{n \in \mathbb{N}} A^{\otimes n}
$$

with $A^{\otimes 0}:=\mathbf{Q}$. Tensors $a_{1} \otimes \ldots \otimes a_{n}$ in $T_{n}(A)$ will be written as words $a_{1} \ldots a_{n}$, the $\otimes$ sign being reserved for tensor products of elements of $T(A)$.

The concatenation product in $T(A)$ is written $\times$ to distinguish it from the internal product in $A$

$$
a_{1} \ldots a_{n} \times b_{1} \ldots b_{m}:=a_{1} \ldots a_{n} b_{1} \ldots b_{m}
$$

The (internal) product of $a$ and $b$ in $A$ will be always written $a \cdot b \in A$ (and not $a b$, which represents an element in $A^{\otimes 2}$ ).

Definition 1. The shuffle bialgebra $\operatorname{Sh}(A)=\bigoplus_{n \in \mathbb{N}} \operatorname{Sh}_{n}(A)$ is the graded connected (i.e. $\mathrm{Sh}_{0}(A)=\mathbf{Q}$ ) commutative Hopf algebra such that

- As vector spaces $\mathrm{Sh}_{n}(A)=T_{n}(A)$.
- Its product $\boldsymbol{\amalg}$ is defined recursively as the sum of the two half-shuffle products $\prec, \succ$

$$
\begin{aligned}
a_{1} \ldots a_{n} & \prec b_{1} \ldots b_{p}:=a_{1} \times\left(a_{2} \ldots a_{n} ш b_{1} \ldots b_{p}\right) \\
a_{1} \ldots a_{n} \succ b_{1} \ldots b_{p}:=b_{1} \ldots b_{p} & \prec a_{1} \ldots a_{n}=b_{1} \times\left(a_{1} \ldots a_{n} ш b_{2} \ldots b_{p}\right),
\end{aligned}
$$

and $\boldsymbol{\omega}=\prec+\succ$.

- Its coalgebra structure is defined by the deconcatenation coproduct

$$
\Delta\left(a_{1} \ldots a_{n}\right):=\sum_{0 \leq k \leq n} a_{1} \ldots a_{k} \otimes a_{k+1} \ldots a_{n}
$$

[^1]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 [13].

The above construction is functorial: $\mathrm{Sh}: A \rightarrow \mathrm{Sh}(A)$ is a functor from the category of commutative algebras without a unit to the category of graded connected commutative Hopf algebras. Since a vector space can be viewed as a commutative algebra with the null product, our definition of $\operatorname{Sh}(A)$ encompasses the contruction of the shuffle algebra over a vector space. Although it may seem artificial at the moment (since we did not make use of the internal product • to define $\operatorname{Sh}(A)$ ), viewing $\operatorname{Sh}$ as a functor from commutative algebras to graded connected commutative Hopf algebras is better suited to the forthcoming developments.

Equivalently, for all $a_{1}, \ldots, a_{k+l} \in A$,

$$
\begin{aligned}
& a_{1} \ldots a_{k} \prec a_{k+1} \ldots a_{k+l}=\sum_{\alpha \in \operatorname{Des}_{\subseteq\{k\}}, \alpha^{-1}(1)=1} a_{\alpha^{-1}(1)} \ldots a_{\alpha^{-1}(k+l)}, \\
& a_{1} \ldots a_{k} \succ a_{k+1} \ldots a_{k+l}=\sum_{\alpha \in \operatorname{Des}_{\subseteq\{k\}}, \alpha^{-1}(1)=k+1} a_{\alpha^{-1}(1)} \ldots a_{\alpha^{-1}(k+l)},
\end{aligned}
$$

where the $\alpha$ are permutations of $[k+l]=\{1, \ldots, k+l\}$.
The notation $\alpha \in \operatorname{Des}_{\subseteq\{k\}}$ 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 $\sigma, \operatorname{desc}(\alpha)$ is the set of all descents of $\alpha$,

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

For $I \subset[n]$, we write $\operatorname{Des}_{I}:=\{\sigma, \operatorname{desc}(\sigma)=I\}$ and $\operatorname{Des}_{\subseteq I}:=\{\sigma, \operatorname{desc}(\sigma) \subseteq I\}$.
A Theorem due to Schützenberger [20] characterizes abstractly the shuffle algebras:
Proposition 2. As a commutative algebra, $\operatorname{Sh}(A)$ is the free algebra over the vector space $A$ for the relation

$$
\begin{equation*}
(a \prec b) \prec c=a \prec(b \prec c+c \prec b) . \tag{1}
\end{equation*}
$$

An algebra equipped with a product $\prec$ satisfying this relation is sometimes refered 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 [6] for historical details.

A last ingredient of the theory of tensors and shuffle algebras over commutative algebras will be useful: 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 $\mathbf{W Q S y m}_{n}$ (resp. WQSym ${ }_{n, p}$ ). We write $\widehat{\text { WQSym }}$ for $\prod_{n, p} \mathbf{W Q S y m}_{n, p}$.

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 $B$,

$$
\begin{equation*}
T_{m}(f) \circ \mu_{A}=\mu_{B} \circ T_{n}(f) \tag{2}
\end{equation*}
$$

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

Proposition 3. The vector space of linear maps that satisfy the nonlinear SchurWeyl duality with parameters $n, m$ is canonically isomorphic to $\mathbf{W Q S y m}_{n, m}$, the linear span of surjections from $[n]$ to $[m]$. Equivalently, the vector space of natural endomorphisms of the functor $T$ is canonically isomorphic to WQSym.

## 2. Natural coalgebra endomorphisms

The nonlinear Schur-Weyl duality shows that $\widehat{\text { WQSym }}$ 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 $\widehat{\mathbf{W Q S y m}}$, 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 $\mathrm{id}_{C}$ 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

$$
\left(\eta_{C} \otimes \operatorname{id}_{C}\right) \circ \Delta=\mathrm{id}_{C}=\left(\mathrm{id}_{C} \otimes \eta_{C}\right) \circ \Delta .
$$

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 $\phi: C \rightarrow C^{\prime}$ are required to preserve the coaugmentation and the counit, so that $\phi\left(C^{+}\right) \subset C^{\prime+}$.

Recall that $\operatorname{Sh}(A)$ is equipped with a coalgebra structure

$$
\Delta\left(a_{1} \ldots a_{n}\right):=\sum_{i=0}^{n} a_{1} \ldots a_{i} \otimes a_{i+1} \ldots a_{n} \in \operatorname{Sh}(A) \otimes \operatorname{Sh}(A)
$$

Its counit is the projection onto $\mathrm{Sh}_{0}(A)=\mathbf{Q}$ and we write

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

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

$$
\operatorname{Hom}_{\text {lin }}\left(C^{+}, A\right) \cong \operatorname{Hom}_{\text {coalg }}(C, \operatorname{Sh}(A))
$$

where $C$ runs over coalgebras and $H_{\text {lin }}$, resp. Hom ${ }_{\text {coalg }}$, 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 4. There is a canonical bijection

$$
\operatorname{Hom}_{\operatorname{lin}}\left(\operatorname{Sh}^{+}(A), A\right) \cong \operatorname{Hom}_{\text {coalg }}(\operatorname{Sh}(A), \operatorname{Sh}(A))
$$

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

$$
\phi\left(a_{1} \ldots a_{n}\right):=\sum_{i_{1}+\ldots+i_{k}=n} f\left(a_{1} \ldots a_{i_{1}}\right) \otimes \cdots \otimes f\left(a_{i_{1}+\ldots+i_{k-1}+1} \ldots a_{n}\right) .
$$

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 $f \in \operatorname{Hom}_{\text {lin }}\left(\operatorname{Sh}^{+}(A), A\right)$ 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 endomophism of the functor Sh viewed as a functor from commutative nonunital algebras to coalgebras is meant a familly $\mu_{A}$ (indexed by commutative nonunital algebras $A$ ) of coalgebra endomorphisms of the $\operatorname{Sh}(A)$ commuting with the morphisms (from $\operatorname{Sh}(A)$ to $\operatorname{Sh}(B)$ ) induced by algebra maps (from $A$ to $B$ ).

Theorem 5. Let Coalg be the monoid of natural endomophisms of the functor Sh viewed as a functor from commutative nonunital algebras to coalgebras. Then, there is an isomorphism between Coalg and the monoid Diff of formal power series without constant term, 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 Coalg ${ }_{1}$ of tangent-to-identity natural endomorphisms of the functor Sh is a group canonically in bijection with the group $\mathrm{Diff}_{1}=X+X^{2} \mathbf{Q}[[X]]$ of tangent-to-identity formal diffeomorphisms.

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

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

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 $a_{1} \ldots a_{n} \in \operatorname{Sh}(A)$ can be described explicitely

$$
\begin{gather*}
f_{P}\left(a_{1} \ldots a_{n}\right)=p_{n} \cdot\left(a_{1} \cdot \ldots \cdot a_{n}\right) \in A  \tag{3}\\
\phi_{P}\left(a_{1} \ldots a_{n}\right)=\sum_{k=1}^{n} \sum_{i_{1}+\ldots+i_{k}=n} p_{i_{1} \ldots} \ldots p_{i_{k}}\left(a_{1} \cdot \ldots \cdot a_{i_{1}}\right) \otimes \ldots \otimes\left(a_{i_{1}+\ldots+i_{k-1}+1} \cdot \ldots \cdot a_{n}\right)
\end{gather*}
$$

This last formula describes the embedding of Coalg into $\widehat{\text { WQSym }}$ induced by SchurWeyl duality (the tensor product $\left(a_{1} \cdot \ldots \cdot a_{i_{1}}\right) \otimes \ldots \otimes\left(a_{i_{1}+\ldots+i_{k-1}+1} \cdot \ldots \cdot a_{n}\right)$ corresponding to the nondecreasing surjection from $[n]$ to $[k]$ sending the first $i_{1}$ integers to $1, \ldots$, the last $i_{k}$ integers to $k$ ). This embedding is of course different from the one induced by the bijection with $\prod_{n \geq 1} \mathbf{W Q S y m}_{n, 1}$ 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 $\prod_{n \geq 1} \mathbf{W Q S y m}_{n, 1}$ (Fla (3)), whereas the $\phi_{P}$ S are most naturally encoded by elements in WQSym (Fla (4)).

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

$$
\begin{equation*}
\phi_{P} \circ \phi_{Q}=\phi_{P(Q)}, \tag{5}
\end{equation*}
$$

where $P(Q)(X):=P(Q(X))$. For an arbitrary commutative algebra $A$ and $a_{1}, \ldots, a_{n} \in$ $A$, we have indeed (with self-explaining notations for the coefficients of $P$ and $Q$ )

$$
\begin{gathered}
\pi \circ \phi_{P} \circ \phi_{Q}\left(a_{1} \ldots a_{n}\right)= \\
=f_{P}\left(\sum_{k=1}^{n} \sum_{i_{1}+\ldots+i_{k}=n} q_{\left.i_{1} \ldots q_{i_{k}}\left(a_{1} \cdot \ldots \cdot a_{i_{1}}\right) \otimes \ldots \otimes\left(a_{i_{1}+\ldots+i_{k-1}+1} \cdot \ldots \cdot a_{n}\right)\right)}^{=} \sum_{k=1}^{n} \sum_{i_{1}+\ldots+i_{k}=n} p_{k} q_{i_{1} \ldots} \ldots q_{i_{k}}\left(a_{1} \cdot \ldots \cdot a_{n}\right)=f_{P(Q)}\left(a_{1} \ldots a_{n}\right)=\pi \circ \phi_{P(Q)}\left(a_{1} \ldots a_{n}\right) .\right.
\end{gathered}
$$

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

## 3. Formal diffeomorphisms and WQSym

We have shown that Coalg and Diff embed naturally into $\widehat{\widehat{\text { QSym}}}$ (Fla (4)). We have already noticed that they are actually embedded in 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 $f_{i}:=\left|f^{-1}(i)\right|$ 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 $f_{1}, \ldots, f_{k}$ 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 quasisymmetric functions, see [7], although the linear embeddings of Sym and QSym into 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 ${ }_{1}$ 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 $A=\left\{a_{1}<a_{2}<\ldots\right\}$ an infinite linearly ordered alphabet and by $A^{*}$ the corresponding set of words.

The packed word $u=\operatorname{pack}(w)$ associated with a word $w \in A^{*}$ is obtained by the following process. If $b_{1}<b_{2}<\cdots<b_{r}$ are the letters occuring in $w, u$ is the image of $w$ by the homomorphism $b_{i} \mapsto a_{i}$. For example, if $A=\mathbb{N}^{*}$, $\operatorname{pack}(35381)=23241$. A word $u$ is said to be packed if $\operatorname{pack}(u)=u$. We denote by PW the set of packed words. With a word $u \in \mathrm{PW}$, we associate the noncommutative polynomial ${ }^{2}$

$$
\begin{equation*}
\mathbf{M}_{u}(A):=\sum_{\operatorname{pack}(w)=u} w \tag{6}
\end{equation*}
$$

For example, restricting $A$ to the first five integers,

$$
\begin{align*}
\mathbf{M}_{13132}(A)= & 13132+14142+14143+24243 \\
& +15152+15153+25253+15154+25254+35354 \tag{7}
\end{align*}
$$

Packed words $u=u_{1} \ldots u_{n}$ are in bijection with surjections: taking for $A$ the set of integers, if $1, \ldots, p$ are the letters occuring 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 $\mathbf{M}_{u} \mathrm{~s}$ is a linear combination of $\mathbf{M}_{u} \mathrm{~s}$, this presentation induces an algebra structure on WQSym and an algebra embedding into $\mathbf{Q}\langle A\rangle$. This algebra structure on WQSym is closely related to the Hopf algebra structure of $\operatorname{QSh}(A)$, 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 $A$ chosen to define word quasi-symmetric functions $\mathbf{M}_{u}(A)$ is largely irrelevant provided it has enough elements. We will therefore often omit the $A$-dependency and write simply $\mathbf{M}_{u}$ for $\mathbf{M}_{u}(A)$, except when we want to emphasize this dependency (and similarly for the other types of generalized symmetric functions we will have to deal with).

[^2]The important point for us now is that WQSym carries naturally a Hopf algebra structure for the coproduct:

$$
\begin{equation*}
\Delta\left(\mathbf{M}_{u}\right):=\sum_{i=0}^{n} \mathbf{M}_{u_{[11, i]}} \otimes \mathbf{M}_{\left.\mathrm{pack}\left(u_{[i+1, n]}\right)\right)} \tag{8}
\end{equation*}
$$

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

Lemma 6. 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 $[8,12,21]$ 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 free 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 $\gamma_{i_{1} \ldots} \gamma_{i_{k}}$ for the unique surjection $f$ from $\left[i_{1}+\ldots+i_{k}\right]$ to $[k]$ such that $f(1)=\ldots=f\left(i_{1}\right):=1, \ldots, f\left(i_{1}+\ldots+i_{j-1}+1\right)=\ldots=$ $f\left(i_{1}+\ldots+i_{k}\right):=k$, then,

$$
\Delta\left(\gamma_{i_{1}} \ldots \gamma_{i_{k}}\right)=\sum_{j=0}^{k} \gamma_{i_{1} \ldots} \gamma_{i_{j}} \otimes \gamma_{i_{j+1}} \ldots \gamma_{i_{k}}
$$

Proposition 7. The embedding of Diff $_{1}$ into $\mathbf{I} \widehat{\mathbf{Q S y m}}$ factorizes through the set of grouplike elements in I $\widehat{\mathbf{Q S y m}}$ (the same statement holds if we replace IQSym by the isomorphic coalgebra QSym). In other terms, the group structure of $\mathbf{D i f f}_{1}$ 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 $\mathbf{M}_{u}$,

$$
\phi_{P}=\sum_{n \geq 0} \sum_{k=1}^{n} \sum_{i_{1}+\ldots+i_{k}=n} p_{i_{1} \ldots p_{i_{k}}} M_{1^{i_{1}} \ldots k^{i_{k}}}
$$

where $1^{i_{1}} \ldots k^{i_{k}}$ stands for the nondecreasing packed word with $i_{1}$ copies of $1, \ldots, i_{k}$ 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

$$
\begin{gathered}
\Delta\left(\phi_{P}\right)=\sum_{n \geq 0} \sum_{k=1}^{n} \sum_{a+b=k} \sum_{i_{1}+\ldots+i_{k}=n} p_{i_{1} \ldots p_{i_{a}}} M_{1^{i_{1}} \ldots a^{i_{a}}} \otimes p_{i_{a+1} \ldots} p_{i_{k}} M_{1^{i_{a+1} \ldots(b-a)^{i_{k}}}} \\
=\phi_{P} \otimes \phi_{P} .
\end{gathered}
$$

Notice that, as a by-product, we have characterized the grouplike elements $\phi$ in WQSym which satisfy the compatibility property

$$
\begin{equation*}
\Delta(\phi) \circ \Delta\left(a_{1} \ldots a_{n}\right)=\Delta \circ \phi\left(a_{1} \ldots a_{n}\right) . \tag{9}
\end{equation*}
$$

In general, for $\phi$ in WQSym (not necessarily grouplike), equation (9) cannot hold if $\phi \notin \mathbf{I} \widehat{\text { QSym }}$ (this is because the deconcatenation coproduct $\Delta$ preserves the relative ordering of the $a_{i} \mathrm{~s}$ ). The converse statement is true and follows from the definition of the coproduct on WQSym (its proof is left to the reader):

Lemma 8. Equation (9) holds for all $\phi \in \widehat{\mathbf{I S S y m}}$, This property characterizes $\widehat{\mathrm{IQSym}}$ as a subspace of WQSym.

Proposition 9. The internal product (defined as the composition of surjections on $\mathbf{I Q S y m}^{+}$, the usual product on $\mathbf{Q}=\mathbf{I Q S y m}_{0}=\mathbf{I Q S y m} \cap \mathbf{W Q S y m}_{0}$-the product of elements in $\mathbf{I Q S y m}^{+}$and $\mathbf{Q}$ is defined to be the null product) provides IQSym with a bialgebra ${ }^{3}$ structure. That is, the coproduct $\Delta$ and the composition of surjections $\circ$ satisfy

$$
\begin{equation*}
\Delta(f \circ g)=\Delta(f) \circ \Delta(g) \tag{10}
\end{equation*}
$$

Indeed, by nonlinear Schur-Weyl duality, the identity is true if and only if the following identity holds for arbitrary $A$, and $a_{1}, \ldots, a_{n}$ in $A$

$$
\Delta(f \circ g) \circ \Delta\left(a_{1} \ldots a_{n}\right)=\Delta(f) \circ \Delta(g) \circ \Delta\left(a_{1} \ldots a_{n}\right),
$$

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

The identity map $I \in \widehat{\mathbf{Q S y m}}$ (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, $I \notin$ 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 IQSym. The next section aims at clarifying these questions.

[^3]
## 4. Natural coderivations and the Faì di Bruno algebra

In the previous section, we have characterized the natural coalgebra endomorphisms of shuffle algebras, or, equivalently, grouplike elements in 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 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 SIQSym be the corresponding subspace of IQSym. By construction, the embedding of Diff into IQSym factorizes through SIQSym (see Eq. (4)). 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 10. 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 $\operatorname{Sh}(A)$ (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 $A \subset \operatorname{Sh}(A)$ is the null map. As usual, a natural coderivation of the shuffle algebras is a familly of coderivations (of the $\operatorname{Sh}(A)$ ) commuting with the morphisms induced by algebra maps (from $A$ to $B$, where $A$ and $B$ run over commutative nonunital algebras).

Lemma 11. Natural tangent-to-identity coalgebra endomorphisms of shuffle algebras identify with tangent-to-identity grouplike elements in 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 $\left[X^{m}, X^{n}\right]:=(m-n) X^{m+n-1}$.

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 $\log (a b)=\log (a)+\log (b)$ 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-toidentity 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 $\operatorname{Coder}^{+}(\operatorname{Sh}(A))$ for the coderivations of $\operatorname{Sh}(A)$ vanishing on $\mathbf{Q}$, we have $\operatorname{Coder}^{+}(\operatorname{Sh}(A)) \cong \operatorname{Lin}\left(\operatorname{Sh}^{+}(A), A\right)$, 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 $f: V \longrightarrow T(V)$

$$
f\left(v_{1} \ldots v_{n}\right):=\sum_{i=1}^{n} v_{1} \ldots v_{i-1} f\left(v_{i}\right) v_{i+1} \ldots v_{n}
$$

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

$$
D_{P}\left(a_{1} \ldots a_{n}\right)=\sum_{i=1}^{n} \sum_{j=1}^{n-i+1} p_{i} a_{1} \ldots a_{j-1}\left(a_{j} \cdot \ldots \cdot a_{j+i-1}\right) a_{j+i} \ldots a_{n}
$$

In particular, the restriction of $\phi_{P}$ and $D_{P}$ to maps from $\operatorname{Sh}(A)$ to $A$ agree and are both given by

$$
\phi_{P}\left(a_{1} \ldots a_{n}\right)=D_{P}\left(a_{1} \ldots a_{n}\right)=p_{n} a_{1} \cdot \ldots \cdot a_{n} .
$$

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

$$
Y\left(a_{1} \ldots a_{n}\right)=\sum_{i=1}^{n} a_{1} \ldots a_{i-1} I\left(a_{i}\right) a_{i+1} \ldots a_{n}=n a_{1} \ldots a_{n}
$$

Similarly, $D_{\lambda X}\left(a_{1} \ldots a_{n}\right)=n \cdot \lambda\left(a_{1} \ldots a_{n}\right)$. 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 $\left(\left[D_{X^{m}}, D_{X^{n}}\right]=(m-n) D_{X^{m+n-1}}\right)$ 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. [5].

Summarizing our previous results, we get:
Theorem 12. 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 $\left[X^{m}, X^{n}\right]:=(m-n) X^{m+n-1}$. 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 [5] for details).

## 5. Deformations of shuffles

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

$$
\operatorname{Sh}_{P}(A)=\left(\operatorname{Sh}(A), \Delta, \omega_{P}\right),
$$

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

$$
x \boldsymbol{w}_{P} y:=\phi_{P}\left(\phi_{P}^{-1}(x) ш \phi_{P}^{-1}(y)\right) .
$$

For the sake of completeness, let us check explicitely the compatibity relation of the product $\omega_{P}$ and the coproduct: since $\phi_{P}^{-1}$ is a coalgebra automorphism,

$$
\begin{gathered}
\Delta\left(x 山_{P} y\right)=\Delta \circ \phi_{P}\left(\phi_{P}^{-1}(x) ш \phi_{P}^{-1}(y)\right)=\left(\phi_{P} \otimes \phi_{P}\right) \circ \Delta\left(\phi_{P}^{-1}(x) ш \phi_{P}^{-1}(y)\right) \\
=\left(\phi_{P} \otimes \phi_{P}\right) \circ\left(\Delta\left(\phi_{P}^{-1}(x)\right) ш \Delta\left(\phi_{P}^{-1}(y)\right)\right) \\
=\left(\phi_{P} \otimes \phi_{P}\right) \circ\left(\left(\phi_{P}^{-1} \otimes \phi_{P}^{-1}\right) \circ \Delta(x) ш\left(\phi_{P}^{-1} \otimes \phi_{P}^{-1}\right) \circ \Delta(y)\right) \\
=\Delta(x) \boldsymbol{w}_{P} \Delta(y) .
\end{gathered}
$$

Definition 13. The Hopf algebra $\operatorname{Sh}_{P}(A)$ is called the $P$-twisted shuffle algebra. It is isomorphic to $\mathrm{Sh}(A)$ as a Hopf algebra.

It inherits therefore all the properties of $\operatorname{Sh}(A)$. The reader is referred to Reutenauer's book [19] for a systematic study of shuffle algebras. As an algebra, $\operatorname{Sh}(A)$ 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 $\operatorname{Sh}(A)$ and $\operatorname{Sh}_{E_{q}}(A)$ is then given by

$$
\phi_{E_{q}}\left(a_{1} \ldots a_{n}\right)=\sum_{\mathcal{P}} \frac{q^{n-k}}{P_{1}!\ldots P_{k}!} a_{P_{1}} \ldots a_{P_{k}}
$$

where $\mathcal{P}=\left(P_{1}, \ldots, P_{k}\right)$ runs over the nondecreasing partitions of $[n]\left(P_{1} \amalg \ldots \amalg P_{k}=\right.$ $[n]$ 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 $\left|P_{1}\right|!$.

Lemma 14. When $q=1, \operatorname{Sh}_{E_{1}}(A)$ identifies with $\operatorname{QSh}(A)$, the quasi-shuffle algebra over $A$, whose product, written $\uplus$, is defined recursively by

$$
\begin{gathered}
a_{1} \ldots a_{n} \uplus b_{1} \ldots b_{m}:= \\
a_{1}\left(a_{2} \ldots a_{n} \uplus b_{1} \ldots b_{m}\right)+b_{1}\left(a_{1} \ldots a_{n} \uplus b_{2} \ldots b_{m}\right)+\left(a_{1} \cdot b_{1}\right)\left(a_{2} \ldots a_{n} \uplus b_{2} \ldots b_{m}\right) .
\end{gathered}
$$

The exponential map $\phi_{E_{1}}$ generalizes the Hoffman isomorphism between $\operatorname{Sh}(A)$ and $\operatorname{QSh}(A)$ introduced and studied in [11] 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 $\operatorname{QSh}(A)$ 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 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 $[8,7,11]$ for details.

The proof amounts to showing that

$$
\phi_{E_{1}}\left(a_{1} \ldots a_{n} ш a_{n+1} \ldots a_{n+m}\right)=\phi_{E_{1}}\left(a_{1} \ldots a_{n}\right) \uplus \phi_{E_{1}}\left(a_{n+1} \ldots a_{n+m}\right) .
$$

Let us write $R_{1} \ldots R_{k}$ 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 $a_{R_{1}} \ldots a_{R_{k}}$ 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 $\frac{1}{P_{1}!\ldots P_{k}!} \frac{1}{Q_{1}!\ldots Q_{k}!}$, where $P_{i}:=R_{i} \cap[n]$, $Q_{i}:=R_{i} \cap\{n+1, \ldots, n+m\}$. 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 $\pm$ ). For the left-hand side, it follows from the identity $\left(\begin{array}{l}\left.\left\lvert\, \begin{array}{l}\left|R_{i}\right| \\ \mid P_{i}\end{array}\right.\right) \times \frac{1}{R_{i}!}=\frac{1}{P_{i}!Q_{i}!} \text { and the fact that }{ }^{2} \mid\end{array}\right.$ the number of words in the expansion of a shuffle product $x_{1} \ldots x_{l} ш y_{1} \ldots y_{k}$ 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 SchurWeyl duality, they belong to WQSym, which inherits an associative (convolution) product from the Hopf algebra structure of the $\operatorname{Sh}(A)$ : for $f \in \mathbf{W Q S y m}_{n}, g \in$ WQSym $_{m}$ and arbitrary $a_{1}, \ldots, a_{n+m} \in A$, where $A$ is an arbitrary commutative algebra,

$$
f ш g\left(a_{1} \ldots a_{n+m}\right):=f\left(a_{1} \ldots a_{n}\right) ш g\left(a_{n+1} \ldots a_{n+m}\right) .
$$

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. [12, 2]. Writing $\mathrm{Sh}_{n, m}$ 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 $\sigma(1)<\ldots<\sigma(n)$
and $\sigma(n+1)<\ldots<\sigma(n+m))$, we get

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

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

$$
\Delta(f):=\sum_{i=0}^{k} f_{\mid\{1, \ldots, i\}} \otimes \operatorname{pack}\left(f_{\mid\{i+1, \ldots, k\}}\right),
$$

this coproduct together with $\boldsymbol{m}$ 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 FQSym $\subset$ WQSym is an embedding of Hopf algebras for these structures. Recall however that $\boldsymbol{\mathrm { w }}$ 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 15. For $f$ a nondecreasing surjection ( $f \in \mathbf{I Q S y m}$ ) and $g \in$ 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 $\operatorname{Sh}_{P}(A)$. This new product, written $\boldsymbol{w}_{P}$ (resp. $\pm$ when $P=E_{1}$, resp. $\omega_{q}$ when $P=E_{q}$ ) is defined by

$$
\begin{gathered}
\forall f \in W Q S y m_{n}, g \in W Q \operatorname{Sym}_{m} \\
f \uplus_{P} g\left(a_{1} \ldots a_{m+n}\right):=f\left(a_{1} \ldots a_{n}\right) \uplus_{P} g\left(a_{n+1} \ldots a_{n+m}\right),
\end{gathered}
$$

(the product $f \omega_{P} g$ 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 16. 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 (WQSym, $\Delta$, ш) to (WQSym, $\Delta$, $\boldsymbol{\omega}_{P}$ ). For $P=E_{1}$ (resp. $P=E_{q}$ ), we will write simply $\Phi_{1}(f)$ (resp. $\Phi_{p}(f)$ ) 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 15, from the definition of $f_{P}$ as the composition of $f$ with a sum of nondecreasing surjections, and from the
fact that $\Delta\left(\phi_{P}\right)=\phi_{P} \otimes \phi_{P}$. On an other hand, the definition of the twisted product $\omega_{P}$ implies

$$
f_{P} \boldsymbol{山}_{P} g_{P}=\phi_{P} \circ(f ш g)=(f ш g)_{P} .
$$

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

Corollary 17. Any $P \in X+X^{2} \mathbf{Q}[[X]]$ induces a Hopf algebra embedding of (FQSym, $\Delta$, ш) into (WQSym, $\Delta, \boldsymbol{\omega}_{P}$ ). This embedding is $S_{n}$-equivariant: for $\sigma, \beta \in S_{n}=$ FQSym $_{n}$, we have:

$$
(\sigma \circ \beta)_{P}=\sigma_{P} \circ \beta
$$

Corollary 18. For $P=E_{1}$, we get that (FQSym, $\Delta$, w) embeds naturally into (WQSym, $\Delta, \pm$ ). 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 $S_{n}$ )

$$
\sigma_{P}=\left(1_{n}\right)_{P} \circ \sigma .
$$

Since $\left(1_{n}\right)_{P}$ is simply the component of $\phi_{P}$ in $\mathbf{W Q S y m}_{n}$, we get finally

$$
\sigma_{P}=\sum_{k=1}^{n} \sum_{i_{1}+\ldots+i_{k}=n} p_{i_{1} \ldots} \ldots p_{i_{k}} 1^{i_{1}} \ldots k^{i_{k}} \circ \sigma
$$

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

Lemma 19. Let $\sigma \in S_{n}$ and $\tau \in \mathbf{W Q S y m}_{n}$. 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 $\tau!:=\prod_{i=1}^{\max (\tau)}\left|\tau^{-1}(\{i\})\right|!$ and $r(\tau)$ for the relative degree of $\tau$. Then, the Hopf algebra embedding $\Phi_{p}$ from FQSym into 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 $\operatorname{Sh}(A)$ to $\operatorname{Sh}_{P}(A)$ for an arbitrary commutative algebra $A$ and an isomorphism between WQSym equipped with the shuffle product $\boldsymbol{\omega}$ to WQSym equipped with the twisted shuffle product $\boldsymbol{\omega}_{P}$.

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 QSh $(A)$.

Recall from $[17,18]$ that $e_{1}:=\log ^{\boldsymbol{\omega}}(I d)$ (the logarithm of the identity map of $\operatorname{Sh}(A)$ computed using the shuffle product $\boldsymbol{\operatorname { m }}$ ) is a canonical section of the projection from $\operatorname{Sh}(A)$ to the indecomposables $\operatorname{Sh}(A)^{+} /\left(\operatorname{Sh}(A)^{+}\right)^{2}$. In particular, due to the
structure theorems for graded connected Hopf algebras over a field of characteristic 0 (Leray, in that particular case), $\operatorname{Sh}(A)$ 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 $(I d)^{w k}$.

These properties are clearly invariant by conjugacy and we get, since $\phi_{P} \circ I d \circ \phi_{P}^{-1}=$ $I d$, the following description of $\mathrm{Sh}_{P}(A)$ as a free commutative algebra:

Proposition 20. For an arbitrary tangent-to-identity $P, e_{1}:=\log ^{\omega_{P}}(I d)$ (the logarithm of the identity for the $\omega_{P}$ product) is a section of the projection from $\operatorname{Sh}_{P}(A)$ to the indecomposables $\operatorname{Sh}_{P}(A)^{+} /\left(\operatorname{Sh}_{P}(A)^{+}\right)^{2}$. In particular, $\operatorname{Sh}_{P}(A)$ 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 $(I d)^{\boldsymbol{\omega}_{P} k}$.

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 projection and write $\tilde{f}$ for the projection from $[n]$ to $[p]$ defined by: $\tilde{f}(i):=f(n-i)$. We have:

Proposition 21. In WQSym equipped with the quasi-shuffle $\uplus$,

$$
e_{1}:=\log { }^{\boldsymbol{\omega}}(I d)=\sum_{n \geq 1} \frac{1}{n} \sum_{I=n} \frac{(-1)^{l(I)-1}}{\binom{n-1}{l(I)-1}} \sum_{\operatorname{Des}(f)=[n]-\left\{i_{l(I)}, \ldots, i_{l(I)}+\ldots+i_{1}\right\}} \tilde{f},
$$

where $I \models n$ means that $I=\left(i_{1}, \ldots, i_{l(I)}\right)$ 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 [3] 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 Id $\in \widehat{\text { WQSym }}$ 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,

$$
F(k)=1+\sum_{i=1}^{k-1} \sum_{1 \leq a_{1}<\ldots<a_{k} \leq k-1} f\left(a_{1}\right) \ldots f\left(a_{k}\right) .
$$

Since the concatenation product of words induces an associative product on WQSym that identifies with $\uplus[2]$, we get finally that $e_{1}$ computes in that case $\log (F)(n)$. This phenomenon was studied recently in detail in [4], to which we refer for details.

## 7. Gradations

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 $\operatorname{Sh}(A)$ to a degree operator $Y_{P}:=\phi_{P} \circ Y \circ \phi_{P}^{-1}$ on $\operatorname{Sh}_{P}(A)$. That is, more explicitly:

Proposition 22. The operator $Y_{P}:=\phi_{P} \circ Y \circ \phi_{P}^{-1}$ acting on $\operatorname{Sh}_{P}(A)$ is a derivation and a coderivation which leaves invariant the subspaces $\operatorname{Sh}_{P}^{\leq n}(A):=\bigoplus_{i \leq n} A^{\otimes n}$. 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 $\operatorname{Sh}_{\bar{P}}^{\leq n}(A)$ and its intersection with $\operatorname{Sh}_{\bar{P}}^{\leq n-1}(A)$ is the null vector space, more precisely:

$$
\operatorname{Sh}_{\bar{P}}^{\leq n}(A)=\operatorname{Sh}_{\bar{P}}^{\leq n-1}(A) \oplus \operatorname{Ker}\left(Y_{p}-n I d\right) .
$$

Concretely, conjugacy by $\phi_{P}$ defines an isomorphism between $A^{\otimes n} \subset \operatorname{Sh}(A)$ and the eigenspace associated with the eigenvalue $n$ of $Y_{P}$ in $\operatorname{Sh}_{P}(A)$. The conjugacy map can be described explicitely as follows:

Proposition 23. 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,

$$
\phi_{U}^{-1} \circ D_{V} \circ \phi_{U}=D_{\frac{V_{0 U}}{U^{\prime}}} .
$$

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,

$$
\begin{gathered}
\pi \circ \phi_{W} \circ D_{X^{p}} \circ \phi_{U}\left(a_{1} \ldots a_{n}\right) \\
=f_{W} \circ D_{X^{p}}\left(\sum_{k=1}^{n} \sum_{i_{1}+\ldots+i_{k}=n} u_{i_{1} \ldots u_{i_{k}}}\left(a_{1} \cdot \ldots \cdot a_{i_{1}}\right) \ldots\left(a_{i_{1}+\ldots+i_{k-1}+1} \cdot \ldots \cdot a_{n}\right)\right) \\
=\sum_{k=p}^{n} \sum_{i_{1}+\ldots+i_{k}=n}(k-p+1) w_{k-p+1} u_{i_{1} \ldots} \ldots u_{i_{k}}\left(a_{1} \cdot \ldots \cdot a_{n}\right),
\end{gathered}
$$

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

$$
\begin{aligned}
& \sum_{k=p}^{\infty}(k-p+1) w_{k-p+1} U^{k}=\left(\sum_{k=p}^{\infty}(k-p+1) w_{k-p+1} X^{k}\right) \circ U \\
= & \left(\sum_{i=1}^{\infty} i w_{i} X^{i-1+p}\right) \circ U=\left(X^{p} W^{\prime}\right) \circ U=U^{p} \cdot\left(W^{\prime} \circ U\right)=\frac{U^{p}}{U^{\prime}} .
\end{aligned}
$$

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 24. The eigenspaces of the coderivation $D_{(1+X) \ln (1+X)}$ provide the quasi-shuffle algebras $\operatorname{QSh}(A)$ with a grading and, more precisely, provide the triple $(\operatorname{QSh}(A), \pm, \Delta)$ 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 $(1+X) \ln (1+X)=1+\sum_{k+2}^{\infty} \frac{(-1)^{k}}{k(k+1)} X^{k}$,
$D_{(1+X) \ln (1+X)}\left(a_{1} \ldots a_{n}\right)=n \cdot a_{1} \ldots a_{n}+\sum_{i=2}^{n} \sum_{j=1}^{n-i+1} \frac{(-1)^{i}}{i(i-1)} a_{1} \ldots a_{j-1}\left(a_{j} \cdot \ldots \cdot a_{j+i-1}\right) a_{j+i} \ldots a_{n}$.

## 8. A NEW REALIZATION OF Sym

In the last sections, we develop some combinatorial applications of our previous results. We will focus mainly on the consequences of Lemma 19, that is, the existence of a new isomorphic embedding of FQSym into 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 Sym and FQSym that will appear in the forthcoming section to be meaningful for the combinatorial study of the Hausdorff series.

Recall that 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 $\left(\Delta\left(S_{n}\right):=\sum_{i=0}^{n} S_{i} \otimes S_{n-i}\right.$, where $\left.S_{0}:=1 \in \mathbf{Q}\right)$. In [7], it has been shown that this Hopf algebra could bring considerable simplifications in the analysis of the socalled 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 $X^{\prime}(t)=X(t) A(t)(X(0)=1)$ as iterated integrals of products of factors $A\left(t_{i}\right)$. 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 [7] (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 19 and define a new polynomial realization of Sym using the existence of an isomorphic embedding $\Phi_{1}$ of FQSym into WQSym:

$$
\begin{equation*}
\hat{S}_{n}=\sum_{u \text { nondecreasing, }|u|=n} \frac{1}{u!} \mathbf{M}_{u} \tag{11}
\end{equation*}
$$

where for notational convenience we write $\hat{S_{n}}$ for $\Phi_{1}\left(S_{n}\right)$ (and similarly for the images by $\Phi_{1}$ of the other elements of Sym and FQSym in WQSym). In view of Fla (6), we have equivalently

$$
\begin{equation*}
\hat{\sigma}_{t}:=\sum_{n \geq 0} t^{n} \hat{S}_{n}=e^{t x_{1}} e^{t x_{2}} \cdots=\prod_{i \geq 1}^{\vec{~}} e^{t x_{i}} \tag{12}
\end{equation*}
$$

Notice that $\hat{\sigma}_{t}$ is (as expected) a grouplike element for the standard coproduct 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

$$
\begin{equation*}
\Phi=\log \left(e^{x_{1}} e^{x_{2}} \cdots\right)=H\left(x_{1}, x_{2}, x_{3}, \ldots\right) \tag{13}
\end{equation*}
$$

Moreover, two nondecreasing words $v$ and $w$ such that $\operatorname{pack}(v)=\operatorname{pack}(w)=u$ have the same coefficient in $\hat{\sigma}_{1}$, that is,

$$
\begin{equation*}
\frac{1}{u!}, \quad \text { where } u!:=\prod_{i} m_{i}(u)! \tag{14}
\end{equation*}
$$

and $m_{i}(u)$ is the number of occurences of $i$ in $u$.
The Hausdorff series can now be expanded in the basis $\mathbf{M}_{u}$ of WQSym as

$$
\begin{equation*}
\Phi=\sum_{u} c_{u} \mathbf{M}_{u} \tag{15}
\end{equation*}
$$

and one may ask whether the previous formalism can shed any light on the coeffients $c_{u}$. There is actually a formula for $c_{u}$, due to Goldberg [9], and reproduced in Reutenauer's book [19, 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 $I=\left(i_{1}, \ldots, i_{r}\right)$, we set $S^{I}:=S_{i_{1}} \ldots S_{i_{r}}$; we also write $\ell(I)=r$ and $|I|=i_{1}+\ldots+i_{r}$ (so that $I \vDash|I|$ ). By definition,

$$
\begin{equation*}
\Phi=\log \left(1+\hat{S}_{1}+\hat{S}_{2}+\cdots\right)=\sum_{r \geq 1} \frac{(-1)^{r-1}}{r} \sum_{\ell(I)=r} \hat{S}^{I}=\int_{-1}^{0}\left(\sum_{I} t^{\ell(I)} \hat{S}^{I}\right) \frac{d t}{t} \tag{16}
\end{equation*}
$$

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

$$
\begin{equation*}
c_{u}=\int_{-1}^{0}\left\langle\mathbf{N}_{u}, \sum_{I} t^{\ell(I)} \hat{S}^{I}\right\rangle \frac{d t}{t} \tag{17}
\end{equation*}
$$

For $u$ a word of length $n$, we have therefore to evaluate $\left\langle\mathbf{N}_{u}, \hat{A}_{n}(t)\right\rangle$ with $A_{n}(t):=$ $\sum_{|I|=n} t^{l(I)} S^{I}$. This last sum is related to a well-known series. The noncommutative

Eulerian polynomials are defined by [7, Section 5.4]

$$
\begin{equation*}
\mathcal{A}_{n}(t)=\sum_{k=1}^{n}\left(\sum_{\substack{|I|=n \\ \ell(I)=k}} R_{I}\right) t^{k}=\sum_{k=1}^{n} \mathbf{A}(n, k) t^{k} . \tag{18}
\end{equation*}
$$

where $R_{I}$ is the ribbon basis (the basis of $\mathbf{S y m}$ obtained from the $S^{I}$ basis by Möbius inversion in the boolean lattice) [7, Section 3.2]. The generating series of the $\mathcal{A}_{n}(t)$ is given by

$$
\begin{equation*}
\mathcal{A}(t):=\sum_{n \geq 0} \mathcal{A}_{n}(t)=(1-t)\left(1-t \sigma_{1-t}\right)^{-1} \tag{19}
\end{equation*}
$$

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

$$
\begin{equation*}
\mathcal{A}^{*}(t):=\sum_{n \geq 0} \mathcal{A}_{n}^{*}(t)=\sum_{I}\left(\frac{t}{1-t}\right)^{\ell(I)} S^{I} \tag{20}
\end{equation*}
$$

and

$$
\begin{equation*}
\sum_{I \vDash n} t^{\ell(I)} S^{I}=A_{n}(t)=\mathcal{A}_{n}^{*}\left(\frac{t}{1+t}\right)=(1+t)^{n} \mathcal{A}_{n}\left(\frac{t}{1+t}\right) . \tag{21}
\end{equation*}
$$

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

$$
\begin{equation*}
\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right)=\frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_{n}} \mathcal{F}_{\sigma}, \tag{22}
\end{equation*}
$$

where $\Phi_{1}^{\dagger}$ is the adjoint map, so that in this case,

$$
\begin{equation*}
\left\langle\mathbf{N}_{u}, \hat{A}_{n}(t)\right\rangle=\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), A_{n}(t)\right\rangle=\frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_{n}} t^{d(\sigma)+1}(1+t)^{r(\sigma)}=\frac{1}{n!} t E_{n}(t, t+1) \tag{23}
\end{equation*}
$$

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 [19]

$$
\begin{equation*}
E_{n}(x, y)=\sum_{\sigma \in \mathfrak{S}_{n}} x^{d(\sigma)} y^{r(\sigma)} \tag{24}
\end{equation*}
$$

Now, recall that the coproduct of $\mathbf{N}_{u}$ dual to the product of the $\mathbf{M}_{u}$ is [16]

$$
\begin{equation*}
\Delta \mathbf{N}_{u}=\sum_{u=u_{1} u_{2}} \mathbf{N}_{\operatorname{pack}\left(u_{1}\right)} \otimes \mathbf{N}_{\operatorname{pack}\left(u_{2}\right)} \tag{25}
\end{equation*}
$$

(deconcatenation). We can omit the packing operation in this formula if we make the convention that $\mathbf{N}_{w}=\mathbf{N}_{u}$ if $u=\operatorname{pack}(w)$. Then, since $\Phi_{1}$, and hence also $\Phi_{1}^{\dagger}$ are morphisms of Hopf algebras, for a composition $L=\left(l_{1}, \ldots, l_{p}\right)$,

$$
\begin{equation*}
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), S^{L}\right\rangle=\left\langle\Delta^{[k]}\left(\mathbf{N}_{u}\right), S_{l_{1}} \otimes \ldots \otimes S_{l_{p}}\right\rangle=\prod_{k=1}^{p}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u_{k}}\right), S_{l_{k}}\right\rangle \tag{26}
\end{equation*}
$$

where $\Delta^{[k]}$ is the $k$-th iterated coproduct and $u=u_{1} u_{2} \cdots u_{p}$ with $\left|u_{k}\right|=l_{k}$ for all $k$. Moreover, this is nonzero only if all the $u_{k}$ are nondecreasing, in which case the result is $1 /\left(u_{1}!\cdots u_{p}!\right)$.

Thus, if

$$
\begin{equation*}
u=w_{1} \cdots w_{m} \tag{27}
\end{equation*}
$$

is the factorization of $u$ into maximal nondecreasing words, with $\left|w_{k}\right|=n_{k}$, we have

$$
\begin{equation*}
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), A_{n}(t)\right\rangle=\prod_{k=1}^{m}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{w_{k}}\right), A_{n_{k}}(t)\right\rangle \tag{28}
\end{equation*}
$$

since

$$
\begin{equation*}
\prod_{k=1}^{m} A_{n_{k}}(t)=\sum_{I \in C_{u}} t^{\ell(I)} S^{I} \tag{29}
\end{equation*}
$$

where $C_{u}$ is the set of compositions which are a refinement of $\left(n_{1}, \ldots, n_{m}\right)$ and are the ones such that $\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), S^{I}\right\rangle \neq 0$.

Next, if $v=1^{l_{1}} 2^{l_{2}} \cdots p^{l_{p}}$,

$$
\begin{equation*}
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{v}\right), S^{L}\right\rangle=\prod_{k=1}^{p}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{k^{l_{k}}}\right), S^{l_{k}}\right\rangle \tag{30}
\end{equation*}
$$

(both sides are equal to $1 /\left(l_{1}!\cdots l_{p}!\right)$ ), so that

$$
\begin{equation*}
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{v}\right), A_{|L|}(t)\right\rangle=\left(1+\frac{1}{t}\right)^{r(v)} \prod_{k=1}^{p}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{k^{l_{k}}}\right), A_{l_{k}}(t)\right\rangle . \tag{31}
\end{equation*}
$$

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

$$
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), S^{I}\right\rangle=\prod_{k=1}^{p}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{k}^{l_{k}}\right), S^{I \mid k}\right\rangle
$$

where $I \mid k$ is the partition of $\left\{l_{1}+\ldots+l_{k-1}+1, \ldots, l_{1}+\ldots+l_{k}\right\}$ 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 $I \mid k)$, using

$$
\begin{equation*}
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), t^{l(I)} S^{I}\right\rangle=t^{l(I)-l(I \cap L)}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), t^{l(I \cap L)} S^{I \cap L}\right\rangle \tag{32}
\end{equation*}
$$

and noting that $|\{I \models|L|, I \cap L=K, l(I)-l(K)=k<r(u)\}|=\binom{r(u)}{k}$, we get (31).
We can now see that if we decompose $u$ into maximal blocks of identical letters,

$$
\begin{equation*}
u=i_{1}^{j_{1}} i_{2}^{j_{2}} \cdots i_{s}^{j_{s}} \tag{33}
\end{equation*}
$$

we have finally

$$
\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{u}\right), A_{n}(t)\right\rangle=\left(1+\frac{1}{t}\right)^{r(u)} \prod_{k=1}^{s}\left\langle\Phi_{1}^{\dagger}\left(\mathbf{N}_{i_{k}^{j_{k}}}\right), A_{j_{k}}(t)\right\rangle
$$

$$
=t^{d(u)+1}(1+t)^{r(u)} \prod_{k=1}^{s} \frac{E_{j_{k}}(t, 1+t)}{j_{k}!}
$$

which implies Goldberg's formula:
Theorem 25. The coefficient $c_{u}$ of $\mathbf{M}_{u}$ in the Hausdorff series $\Phi$ is given by:

$$
\begin{equation*}
c_{u}=\int_{-1}^{0} t^{d(u)+1}(1+t)^{r(u)} \prod_{k=1}^{s} \frac{E_{j_{k}}(t, 1+t)}{j_{k}!} \frac{d t}{t} \tag{34}
\end{equation*}
$$

More generally, for an arbitrary moment generating function

$$
\begin{equation*}
f(z)=\sum_{n \geq 1} f_{n} z^{n} \tag{35}
\end{equation*}
$$

with

$$
\begin{equation*}
f_{n}=\int_{\mathcal{R}} t^{n} d \mu(t) \tag{36}
\end{equation*}
$$

the coefficient of $\mathbf{M}_{u}$ in $f\left(\hat{\sigma}_{1}\right)$ is

$$
\begin{equation*}
\int_{\mathcal{R}} t^{d(u)+1}(1+t)^{r(u)} \prod_{k=1}^{s} \frac{E_{j_{k}}(t, 1+t)}{j_{k}!} d \mu(t) \tag{37}
\end{equation*}
$$

## References

[1] R. Block and P. Leroux, Generalized dual coalgebras of algebras, with applications to cofree coalgebras, J. Pure Appl. Algebra, 36 (1985), 15-21.
[2] G. Duchamp, F. Hivert and J.-Y. Thibon, Noncommutative symmetric functions VI: free quasi-symmetric functions and related algebras, International Journal of Algebra and Computation 12 (2002), 671-717.
[3] K. Ebrahimi-Fard, J. Gracia-Bondia and F. Patras. Rota-Baxter algebras and new combinatorial identities. Letters in Math. Physics 81, (1), (2007), 61-75.
[4] K. Ebrahimi-Fard and D. Manchon, The tridendriform structure of a Magnus expansion. ArXiv:1306.6439.
[5] H. Figueroa, J. M. Gracia-Bonda, Combinatorial Hopf algebras in quantum field theory I. Reviews in Mathematical Physics, 17 (08), (2005), 881-976.
[6] L. Foissy and F. Patras, Natural endomorphisms of shuffle algebras, International Journal of Algebra and Computation, 23 (4) (2013), 989-1009.
[7] 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.
[8] I. Gessel, Multipartite P-partitions and inner products of skew Schur functions, Contemp. Math. 34 (1984), 289-301.
[9] K. Goldberg, The formal power series for $\log e^{x} e^{y}$, Duke Math. J. 23 (1956), 13-21.
[10] F. Hivert, Combinatoire des fonctions quasi-symétriques, Thèse de doctorat, 1999.
[11] M. E. Hoffman, Quasi-shuffle products, J. Algebraic Combin., 11 (1) (2000) 49-68.
[12] C. Malvenuto and C. Reutenauer Duality between Solomon's algebra and quasi-symmetric functions, J. Algebra, 177 (1995), 967-982.
[13] J.W. Milnor, J.C. Moore, On the structure of Hopf algebras, Ann. of Math. 81 (2) (1965) 211-264.
[14] J.C. Novelli, F. Patras and J.Y. Thibon. Natural endomorphisms of quasi-shuffle Hopf algebras. Bull. Soc. math. France 141 (1), 2013, 107-130.
[15] J.-C. Novelli and J.-Y. Thibon, Free quasi-symmetric functions and descent algebras for wreath products, and noncommutative multi-symmetric functions, Discrete Mathematics 310 (2010), 3584-3606.
[16] J.-C. Novelli and J.-Y. Thibon, Polynomial realizations of some trialgebras, Proc. FPSAC'06, San Diego, 2006.
[17] F. Patras, La décomposition en poids des algèbres de Hopf, Ann. Inst. Fourier, 43 (1993), 1067-1087.
[18] F. Patras, L'algèbre des descentes d'une bigèbre graduée. J. Algebra 170, 2 (1994), 547-566.
[19] C. Reutenauer. Free Lie algebras. Oxford University Press, 1993.
[20] 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), Paris, Année 1958/59.
[21] R. P. Stanley, Enumerative Combinatorics, Vol. 2, Cambridge University Press, 1999.
(Foissy) Laboratoire de Mathématiques Pures et Appliquées Joseph Liouville, Université du Littoral Côte dOpale, Centre Universitaire de la Mi-Voix, 50, rue Ferdinand Buisson, CS 80699, 62228 Calais Cedex, France
(Patras) Laboratoire de Mathématiques J.A. Dieudonné, Université de Nice - Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 02, France
(Thibon) Laboratoire d'Informatique Gaspard-Monge, Université Paris-Est, 5 , boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, France

E-mail address, L. Foissy: foissy@lmpa.univ-littoral.fr
E-mail address, F. Patras: patras@unice.fr
E-mail address, J.-Y. Thibon: jyt@univ-mlv.fr


[^0]:    1991 Mathematics Subject Classification. 05E05,16T30.
    Key words and phrases. Shuffle algebras, Combinatorial Hopf algebras, Hausdorff series.

[^1]:    ${ }^{1}$ Any field of characteristic zero would be suitable as well.

[^2]:    ${ }^{2}$ 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 $Q\langle A\rangle$ the corresponding algebra. Since these are elements of a projective limit of polynomial rings, purists may want to call these objects (noncommutative) prolynomials.

[^3]:    ${ }^{3}$ By "bialgebra", we simply mean in the present article a compatible product and coproduct as in identity (10), 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).

