Advances in Mathematics **164**, 283–300 (2001)
doi:10.1006/aima.2001.2017, available online at http://www.idealibrary.com on IDEAL®

The Algebra of Quasi-Symmetric Functions Is Free over the Integers

Michiel Hazewinkel

CWI, Amsterdam, The Netherlands
E-mail: mich@cwi.nl

Communicated by Alain Lascoux

Received May 21, 2000

Let $\mathscr{Z}$ denote the Leibniz–Hopf algebra, which also turns up as the Solomon descent algebra and the algebra of noncommutative symmetric functions. As an algebra $\mathscr{Z} = \mathbf{Z}\langle Z_1, Z_2, \dots\rangle$, the free associative algebra over the integers in countably many indeterminates. The coalgebra structure is given by $\mu(Z_n) = \sum_{i=0}^{n} Z_i \otimes Z_{n-i}$, $Z_0 = 1$. Let $\mathscr{M}$ be the graded dual of $\mathscr{Z}$. This is the algebra of quasi-symmetric functions. The Ditters conjecture says that this algebra is a free commutative algebra over the integers. In this paper the Ditters conjecture is proved. © 2001 Elsevier Science

Key Words: Leibniz–Hopf algebra; quasi-symmetric functions; Ditters conjecture; Lie–Hopf algebra; Solomon descent algebra; shuffle algebra; overlapping shuffle algebra; noncommutative symmetric functions; divided power sequences; coalgebra; Hopf algebra; free coalgebra; formal group; Lyndon words; symmetric group; symmetric functions; Hecke algebra.

1. INTRODUCTION

Quasi-symmetric functions are a generalization of symmetric functions introduced some 15 years ago to deal with the combinatorics of P-partitions and the counting of permutations with given descent sets [6, 7]; see also [19]. They also appear as the dual algebra over the integers of the Leibniz–Hopf algebra (defined in Section 3 below).

The first statement of the Ditters conjecture dates from 1972 [2], where it was formulated as Proposition 2.2. It states that this dual algebra over the integers, i.e., the algebra of quasi-symmetric functions, is a free commutative algebra over the integers. At that time quasi-symmetric functions had not yet been invented, nor had the Solomon descent algebra.

The fact that this dual algebra is free polynomial over the integers is crucial for the classification theory of noncommutative formal groups, via

0001-8708/01 $35.00

a noncommutative version of p-typification, developed by Ditters and his students; see [2, 18] and the references cited therein. One consequence of the freeness of $\mathscr{M}$ is that the group of curves of a noncommutative formal group has many functorial operations acting on it; see [12].

Perhaps even more importantly, the Leibniz–Hopf algebra is precisely the same as the algebra of noncommutative symmetric functions as defined in [5] and further developed in a slew of subsequent papers (one of which is [13]). The fact that the symmetric functions constitute a free algebra in the elementary symmetric functions is rather important. Thus the fact that the algebra of quasi-symmetric functions is free over the integers is likely to be of some significance.

Let $\mathscr{Z}^c = \mathbf{Z}[z_1, z_2, \dots]$, with the same comultiplication, be the commutative quotient Hopf algebra of $\mathscr{Z}$. The graded dual of this is the subalgebra, *Symm*, of symmetric functions over the integers. This subalgebra of symmetric functions has a well-known representation theoretic interpretation as follows: For each n let S_n be the symmetric group on n letters and $R(S_n)$ be the (Grothendieck) group of (finite-dimensional) complex representations of S_n. Then (as Hopf algebras)

$$\mathit{Symm} \cong \bigoplus_n R(S_n), \tag{1.1}$$

where the multiplication is given by the "induction product"

$$R(S_n) \times R(S_m) \to R(S_{n+m}), \qquad (\rho, \sigma) \mapsto \operatorname{Ind}_{S_n \times S_m}^{S_{n+m}}(\rho \times \sigma). \tag{1.2}$$

The quasi-symmetric functions have a similar representation theoretic interpretation with the symmetric groups replaced with the Hecke algebras at 0, $H_n(0)$, and where two $H_n(0)$-modules are considered equivalent if they have the same composition factors; see [13, 21].

In this paper I prove the Ditters conjecture.

Shortly after the publication of [2] it was remarked and acknowledged (see [3, Chap. II, Section 5, p. 29]) that the proof of Proposition 2.2, i.e., what is now called the Ditters conjecture, had gaps. Since then there have been quite a few purported proofs of the statement, both published and unpublished. All have errors. For detailed remarks on the error in the proofs in [17, 18] see [11]. The latest purported proof in [4] has at least three major errors; the worst one is more or less the same as the one in [17, 18].

The name "Ditters conjecture" for the statement was coined by me a few years back. In [9], I referred to the statement as the Ditters–Scholtens theorem. This was when I still believed the proof in [17, 18] to be correct.

2. THE ALGEBRA OF QUASI-SYMMETRIC FUNCTIONS

Let X be a finite or infinite set (of variables) and consider the ring, of polynomials, $R[X]$, and the ring of power series, $R[[X]]$, over a commutative ring R with unit element, in the commuting variables from X. A polynomial or power series $f(X) \in R[[X]]$ is called symmetric if for any two finite sequences of indeterminates $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ from X and for any sequence of exponents $i_1, i_2, \ldots, i_n \in \mathbf{N}$ the coefficients in $f(X)$ of $X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ and $Y_1^{i_1} Y_2^{i_2} \cdots Y_n^{i_n}$ are the same.

The quasi-symmetric formal power series are a generalization introduced by Gessel ([6]) in connection with the combinatorics of plane partitions. This time one takes a *totally ordered* set of indeterminates, e.g., $V = \{V_1, V_2, \ldots\}$, with the ordering that of the natural numbers, and the condition is that the coefficients of $X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n}$ and $Y_1^{i_1} Y_2^{i_2} \cdots Y_n^{i_n}$ are equal for all totally ordered sets of indeterminates $X_1 < X_2 < \cdots < X_n$ and $Y_1 < Y_2 < \cdots < Y_n$. Thus, for example,

$$X_1 X_2^2 + X_2 X_3^2 + X_1 X_3^2 \tag{2.1}$$

is a quasi-symmetric polynomial in three variables that is not symmetric.

Products and sums of quasi-symmetric polynomials and power series are again quasi-symmetric (obviously). Thus one has, for example, the ring of quasi-symmetric power series

$$Qsym_{\mathbf{Z}}(X)^{\wedge} \tag{2.2}$$

in countably many commuting variables over the integers and its subring

$$Qsym_{\mathbf{Z}}(X) \tag{2.3}$$

of quasi-symmetric polynomials in countably many indeterminates, which are the quasi-symmetric power series of bounded degree.

Given a word $w = [a_1, a_2, \ldots, a_n]$ over $\mathbf{N}$, also called a *composition* in this context, consider the quasi-monomial function

$$M_w = \sum_{i_1 < \cdots < i_n} X_{i_1}^{a_1} X_{i_2}^{a_2} \cdots X_{i_n}^{a_n} \tag{2.4}$$

defined by w. These form a basis over the integers of $Qsym_{\mathbf{Z}}(X)$.

The algebra of quasi-symmetric functions is dual to the *Leibniz–Hopf algebra*, see below, or equivalently to the *Solomon descent algebra*, or more precisely, to the direct sum

$$\mathscr{D} = \bigoplus_n D(S_n) \tag{2.5}$$

of the Solomon descent algebras $D(S_n)$ of the symmetric groups, with a new multiplication over which the direct sum of the original multiplications is distributive. See [5, 15].

3. THE LEIBNIZ–HOPF ALGEBRA

The *Leibniz–Hopf algebra* over the integers is the free associative algebra $\mathscr{Z} = \mathbf{Z}\langle Z_1, Z_2, ...\rangle$ over $\mathbf{Z}$ in countably many generators with the comultiplication

$$\mu(Z_n) = \sum_{i+j=n} Z_i \otimes Z_j, \qquad Z_0 = 1. \tag{3.1}$$

Its graded dual over the integers is denoted $\mathscr{M}$. It is not difficult to see that this dual is precisely the algebra of quasi-symmetric functions over the integers. Indeed, for any composition $c = (i_1, ..., i_n)$, define m_c by the dual basis formula

$$\langle m_c, Z^d\rangle = \delta_{c,d}, \tag{3.2}$$

where $Z^d = Z_{j_1} Z_{j_2} \cdots Z_{j_m}$ for a composition $d = (j_1, ..., j_m)$. It is now a simple exercise to check that the m_c multiply exactly as the quasi-symmetric monomials M_c, defined above in Section 2. Explicitly, for instance, writing simply c for the element m_c corresponding to the composition c, one has for the multiplication of the two compositions $[a, b]$ and $[c, d]$;

$$\begin{aligned}[a, b]\,[c, d] = {} & [a, b, c, d] + [a, c, b, d] + [a, c, d, b] + [c, a, b, d] \\ & + [c, a, d, b] + [c, d, a, b] + [a+c, b, d] + [a+c, d, b] \\ & + [c, a+d, b] + [a, b+c, d] + [a, c, b+d] + [c, a, b+d] \\ & + [a+c, b+d].\end{aligned} \tag{3.3}$$

The first six terms of this multiplication are the terms of the more familiar *shuffle algebra*, whose name derives from the familiar rifle shuffle in card playing. I call the multiplication of compositions defined by the multiplication of quasi-symmetric functions the *overlapping shuffle multiplication*. The illustration is that during a rifle shuffle of the two words (=compositions) two cards, one from each word, may stick together; in that case their labels are added. This gives the additional seven terms in the example (3.3) above.

Over the rationals the Leibniz–Hopf algebra is isomorphic to the *Lie–Hopf algebra*

$$\mathcal{U} = \mathbf{Z}\langle U_1, U_2, \ldots\rangle, \qquad \mu(U_n) = 1 \otimes U_n + U_n \otimes 1. \tag{3.4}$$

Let $\mathcal{N}$ be the graded dual of $\mathcal{U}$ over the integers. This is the so-called *shuffle algebra*. An important theorem in the theory of free Lie algebras, for example, states that the algebra $\mathcal{N} \otimes_{\mathbf{Z}} \mathbf{Q}$ is a commutative free polynomial in the Lyndon words; see e.g. [16]. It is not true that $\mathcal{N}$ is a free polynomial over the integers. The Ditters conjecture states that the algebra $\mathcal{M}$, in contrast, is a free polynomial commutative over the integers. This would make it a rather more beautiful version of $\mathcal{N}$, in the sense that $\mathcal{M}$ is a $\mathbf{Z}-\mathbf{Q}$ form of $\mathcal{N}$ (i.e., $\mathcal{M} \otimes_{\mathbf{Z}} \mathbf{Q} \simeq \mathcal{N} \otimes_{\mathbf{Z}} \mathbf{Q}$) with the property that $\mathcal{M}$ is a free polynomial algebra while $\mathcal{N}$ is not.

4. LYNDON WORDS AND FORMULATION OF THE STRONG DITTERS CONJECTURE

Let the elements of $\mathbf{N}^*$, i.e., the words over $\mathbf{N}$, be ordered lexicographically, where any symbol is larger than nothing. Thus $[a_1, a_2, \ldots, a_n] > [b_1, b_2, \ldots, b_m]$ if and only if there is an i such that $a_1 = b_1, \ldots, a_{i-1} = b_{i-1}, a_i > b_i$ (with, necessarily, $1 \leqslant i \leqslant \min\{m, n\}$), or $n > m$ and $a_1 = b_1, \ldots, a_m = b_m$.

A *proper tail* of a word $[a_1, \ldots, a_n]$ is a word of the form $[a_i, \ldots, a_n]$ with $1 < i \leqslant n$. (The empty word and one-symbol words have no proper tails.)

A word is *Lyndon* if all its proper tails are larger than the word itself. For example, the words $[1, 1, 3]$, $[1, 2, 1, 3]$, and $[2, 2, 3, 2, 4]$ are all Lyndon and the words $[2, 1]$, $[1, 2, 1, 1, 2]$, and $[1, 3, 1, 3]$ are not Lyndon. The set of Lyndon words is denoted LYN.

Obviously all of these definitions make sense for any totally ordered set and not just for the set of natural numbers.

Now consider $\mathbf{N}^*$ as a semigroup under the concatenation product.

4.1. Theorem (Chen–Fox–Lyndon Factorization [1, 14]). *Every word w in $\mathbf{N}^*$ factors uniquely into a decreasing concatenation product of Lyndon words,*

$$w = v_1 * v_2 * \cdots * v_k, \qquad v_i \in LYN, \quad v_1 \geqslant v_2 \geqslant \cdots \geqslant v_k. \tag{4.2}$$

For example,

$$[2, 3, 1, 3, 1, 4, 1, 3, 1, 1] = [2, 3] * [1, 3, 1, 4] * [1, 3] * [1] * [1]. \tag{4.3}$$

One efficient algorithm for finding the Chen–Fox–Lyndon factorization of a word is the *block decomposition algorithm* from [17].

The Lyndon words are the right kind of thing for the shuffle algebra over the rational numbers, **Q**, and also for the algebra of quasi-symmetric functions (also called the *overlapping shuffle algebra*; see [11]) over **Q**. Indeed, both of these algebras are free polynomial over **Q** with as their generators the words from *LYN*. However, over the integers *LYN* most definitely is not a free generating set for the algebra of quasi-symmetric functions.

A word $w = [a_1, a_2, ..., a_n] \in \mathbf{N}^*$ is called *elementary* if the greatest common divisor of its symbols is 1, $\gcd\{a_1, a_2, ..., a_n\} = 1$. A *concatenation power* of w (or *star power*) is a word of the form

$$w^{*m} = \underbrace{w * w * \cdots * w}_{m \text{ factors}}. \tag{4.4}$$

Let *ESL* denote the set of words which are star powers of elementary Lyndon words. For instance, the words [1, 1, 1, 1], [1, 2, 1, 2], and [1, 2, 1, 4] are in *ESL* (but the first two are not Lyndon), and the words [4], [2, 4] are not in *ESL* but are in *LYN*.

The *strong Ditters conjecture* now states that the elements of *ESL* form a free (communicating) generating set for the overlapping shuffle algebra $\mathscr{M}$ over the integers.

Let the *weight* of a word $w = [a_1, a_2, ..., a_n]$ be equal to $a_1 + a_2 + \cdots + a_n$. The elements of *ESL* of weight $\leqslant 6$ area

$$\begin{gathered}
[1];\\
[1, 1];\\
[1, 1, 1], [1, 2];\\
[1, 1, 1, 1], [1, 1, 2], [1, 3];\\
[1, 1, 1, 1, 1], [1, 1, 1, 2], [1, 1, 3], [1, 2, 2], [1, 4], [2, 3];\\
\text{and}\\
[1, 1, 1, 1, 1, 1], [1, 1, 1, 1, 2], [1, 1, 1, 3], [1, 1, 2, 2], [1, 1, 4], [1, 2, 1, 2],\\
[1, 2, 3], [1, 3, 2], [1, 5].
\end{gathered} \tag{4.5}$$

5. THE SHUFFLE ALGEBRA

There is a second Hopf algebra structure on the free associative algebra in countably many indeterminates over **Z**; i.e., there is a second way to make the ring $\mathbf{Z}\langle Z_1, Z_2, ...\rangle$ into a Hopf algebra. This was briefly mentioned in

Section 3 above. This structure is actually rather better known and it plays a most important role in the theory of free Lie algebras and related matters; see e.g. [16]. In order to avoid notational confusion, let

$$\mathscr{U} = \mathbf{Z}\langle U_1, U_2, \ldots\rangle \tag{5.1}$$

be another copy of the free associative algebra in countably many variables over **Z**, and let the comultiplication be defined by

$$\mu(U_m) = 1 \otimes U_n + U_n \otimes 1. \tag{5.2}$$

Let $\mathscr{N}$ be the graded dual algebra of $\mathscr{U}$. This is the *shuffle algebra*. The shuffle multiplication is the same as the overlapping shuffle multiplication except that overlaps are not allowed. Thus for example

$$\begin{aligned}[][a, b] \times_{sh} [c, d] = {} & [a, b, c, d] + [a, c, b, d] + [a, c, d, b] \\ & + [c, a, b, d] + [c, a, d, b] + [c, d, a, b],\end{aligned}$$

which is just the first six terms of (3.3), and

$$\begin{aligned} [1] \times_{sh} [1] &= 2[1, 1], \\ [1] \times_{sh} [1] \times_{sh} [1] &= 6[1, 1, 1], \end{aligned} \tag{5.3}$$

$$\begin{aligned} [1] \times_{osh} [1] &= 2[1, 1] + [2], \\ [1] \times_{osh} [1] \times_{osh} [1] &= 6[1, 1, 1] + 3[1, 2] + 3[2, 1] + [3]. \end{aligned} \tag{5.4}$$

A well-known theorem says that over the rationals the shuffle algebra is a free polynomial. More precisely, let $\mathbf{Q}[LYN]$ be the free commutative polynomial ring over the set LYN of Lyndon words, then (cf. e.g. [16]) we have

5.5. Theorem (Shuffle Algebra Structure Theorem). $\mathscr{N} \otimes \mathbf{Q} = \mathbf{Q}[LYN]$, *the free commutative algebra over* **Q** *in the symbols from* LYN.

Note that nothing like this is true over the integers. Indeed, by the second examples of the shuffle multiplication above (see (5.3)), $\mathscr{N} \otimes \mathbf{Z}/(2)$ has nilpotents and so $\mathscr{N}$ cannot be a free algebra over **Z**. From this point of view the overlapping shuffle algebra $\mathscr{M}$ is a rather nicer "version" of $\mathscr{N}$. Here the word "version" refers to the fact that over the rational numbers, **Q**, $\mathscr{M}$ and $\mathscr{N}$ become isomorphic.

The proof of the shuffle algebra structure theorem is a straightforward application of the following theorem concerning shuffle products in connection with Chen–Fox–Lyndon factorization.

5.6. THEOREM. *Let $w \in \mathbf{N}^*$ be a word on the natural numbers and let $w = v_1 * v_2 * \cdots * v_m$ be its Chen–Fox–Lyndon factorization. Then all words that occur with nonzero coefficient in the shuffle product $v_1 \times_{sh} v_2 \times_{sh} \cdots \times_{sh} v_m$ are lexicographically smaller or equal to w, and w occurs with a nonzero integer coefficient in this product.*

Given this result, the proof of the shuffle algebra theorem proceeds as follows: Order all words lexicographically. Consider some nonempty word w. With induction, $[1]$ being the smallest nonempty word, we can assume that all words lexicographically smaller than w have been written as polynomials in the elements of LYN. Take the Chen–Fox–Lyndon factorization $w = v_1 * v_2 * \cdots * v_m$ of w and consider, using Theorem 5.6,

$$v_1 \times_{sh} v_2 \times_{sh} \cdots \times_{sh} v_m = aw + (\text{remainder}). \tag{5.7}$$

By Theorem 5.6, the coefficient a is nonzero and all the words in (remainder) are lexicographically smaller than w, and hence $\in \mathbf{Q}[LYN]$. It follows that also $w \in \mathbf{Q}[LYN]$. This proves generation; i.e., the surjectivity of the natural map $\mathbf{Q}[LYN] \to \mathscr{N}$. Injectivity follows by counting. The map is homogeneous, both algebras are graded, and $\dim_{\mathbf{Q}}(\mathbf{Q}[LYN]_n) = \dim_{\mathbf{Q}}(\mathscr{N}_n)$. Indeed, these numbers are given by a recursive relation

$$\frac{1-2t}{1-t} = \prod_{n=1}^{\infty} (1-t^n)^{\beta n},$$

where β_n is the number of Lyndon words of weight n. See e.g. [16, 17] for more details.

6. THE OVERLAPPING SHUFFLE ALGEBRA OVER THE RATIONALS

As already stated, the overlapping shuffle algebra and the shuffle algebra become isomorphic over the rationals. Given that the shuffle algebra over the rationals is free polynomial there are of course many possible algebra homomorphisms. There is a particularly nice one which comes from a Hopf algebra isomorphism between $\mathscr{Z} \otimes \mathbf{Q}$ and $\mathscr{U} \otimes \mathbf{Q}$, as follows.

Consider the expression

$$1 + Z_1 t + Z_2 t^2 + Z_3 t^3 + \cdots = \exp(U_1 t + U_2 t^2 + U_3 t^3 + \cdots). \tag{6.1}$$

This gives an expression for each Z_i in terms of the $U_1, \ldots, U_i$ with rational coefficients, and hence defines an algebra homomorphism

$$\beta\colon \mathscr{Z} \otimes \mathbf{Q} \to \mathscr{U} \otimes \mathbf{Q}. \tag{6.2}$$

6.3. THEOREM. *The algebra homomorphism β is an isomorphism of Hopf algebras and hence its dual defines an isomorphism of algebras $\beta^*\colon \mathscr{N} \otimes \mathbf{Q} \to \mathscr{M} \otimes \mathbf{Q}$.*

For details, cf. [8]. This proves of course that $\mathscr{M} \otimes \mathbf{Q}$ is free polynomial and gives a set of generators which is, however, neither the set LYN nor the set ESL.

It is also not difficult to adapt the proof that $\mathscr{N} \otimes \mathbf{Q}$ is free polynomial on LYN to a proof that $\mathscr{M} \otimes \mathbf{Q}$ is free polynomial on LYN. The only modification needed is to change a bit the ordering on words that is used. The ordering which works here is the following:

$$\begin{aligned} w \succ v \Leftrightarrow{} & \operatorname{length}(w) > \operatorname{length}(v) \\ & \text{or } (\operatorname{length}(w) = \operatorname{length}(v) \text{ and } w \geqslant v \text{ (lexicographically)}). \end{aligned} \tag{6.3}$$

7. THE LOCAL VERSION OF THE STRONG DITTERS CONJECTURE

There is a p-adic analogue of the strong Ditters conjecture, and the first step in establishing the Ditters conjecture is to prove these local versions for all prime numbers p.

Let us start with the formulation. A word $w = [a_1, \ldots, a_n]$ on $\mathbf{N}$ is *p-elementary*, where p is a prime number, if the gcd of the $a_1, \ldots, a_n$ is not divisible by p. A p-star-power of a word is a word of the form

$$w = \underbrace{v * \cdots * v}_{p^r \text{ factors}}. \tag{7.1}$$

The set $ESL(p)$ is the set of words which are p-star-powers of p-elementary Lyndon words.

7.2. THEOREM (*p-adic Analogue of the Strong Ditters Conjecture*).

$$\mathscr{M} \otimes \mathbf{Z}_{(p)} = \mathbf{Z}_{(p)}[ESL(p)].$$

That is, $\mathscr{M} \otimes \mathbf{Z}_{(p)}$ is the free commutative algebra on $ESL(p)$ over $\mathbf{Z}_{(p)}$.

To prove this we first need some information on binomial and multinomial coefficients. Extend the usual definition of the binomial coefficients in the standard way:

$$\binom{n}{m}=0 \quad \text{if} \quad m>n, \qquad \binom{n}{0}=1 \quad \text{for} \quad n\geqslant 0.$$

7.3. Proposition. *Consider the p-adic expansion of two natural numbers m and n,*

$$\begin{aligned} n&=a_0+a_1p+\cdots+a_kp^k, \\ m&=b_0+b_1p+\cdots+b_kp^k, \end{aligned} \qquad a_i, b_j\in\{0,1,\ldots,p-1\}. \tag{7.4}$$

Then the value of the binomial coefficient modulo p is equal to

$$\binom{n}{m}\equiv\binom{a_0}{b_0}\binom{a_1}{b_1}\cdots\binom{a_n}{b_n}. \tag{7.5}$$

In particular, if $b_i\leqslant a_i$ for all i, this binomial coefficient is nonzero modulo p.

7.6. Corollary. *The multinomial coefficient*

$$\binom{n}{\underbrace{p^k\cdots p^k}_{a_k \text{ times}}\;\underbrace{p^{k-1}\cdots p^{k-1}}_{a_{k-1} \text{ times}}\;\underbrace{1\cdots 1}_{a_0 \text{ times}}} \tag{7.7}$$

is nonzero modulo p.

Proof of the Proposition. For $0\leqslant n\leqslant p-1$ things are clear. Now let $n\geqslant p$ and write down the p-adic expansion of n and m as in the formulation of the proposition and let

$$n_1=a_0+a_1p+\cdots+a_{k-1}p^{k-1}, \qquad m_1=b_0+b_1p+\cdots+b_{k-1}p^{k-1}. \tag{7.8}$$

We have the modulo p

$$(x+y)^n=(x+y)^{a_kp^k}(x+y)^{n_1}\equiv(x^{p^k}+y^{p^k})^{a_k}(x+y)^{n_1}.$$

Writing things out gives

$$\begin{aligned}(x+y)^n=&\left\{\binom{a_k}{0}(x^{p^k})^{a_k}(y^{p^k})^0+\cdots\right.\\ &\left.+\binom{a_k}{i}(x^{p^k})^{a_k-i}(y^{p^k})^i+\cdots+(x^{p^k})^0(y^{p^k})^{a_k}\right\}\\ &\times\left\{\binom{n_1}{0}x^{n_1}y^0+\cdots+\binom{n_1}{i}x^{n_1-i}y^i+\cdots+\binom{n_1}{0}x^0y^{n_1}\right\}.\end{aligned}$$

It follows that

$$\binom{n}{m} \equiv \binom{a_k}{b_k}\binom{n_1}{m_1},$$

and with induction the desired result follows.

7.9. Lemma (Cardinality of the Sets $ESL(p)$). *The number of elements in $ESL(p)$ of weight n is β_n; i.e., it is the same as that in LYN_n, the set of Lyndon words of weight n.*

Proof. Let $w=[a_1, a_2, ..., a_m]$ be a Lyndon word of weight n. Let p^r be the largest power of the prime number p that divides the greatest common divisor $\gcd(a_1, ..., a_m)$. Now assign to w the word

$$\underbrace{v * v * \cdots * v}_{p^r \text{ factors}}, \qquad v=[a_1/p^r, a_2/p^r, ..., a_m/p^r].$$

This sets up a bijective correspondence between LYN_n and $ESL(p)_n$, the set of words in $ESL(p)$ of weight n.

7.10. *Proof of the p-adic Ditters conjecture.* We use the same ordering of words as at the end of Section 5 above; i.e., length first and then lexicographic ordering on words of equal length. Let $SL(p)$ be the set of all p-star powers of Lyndon words; i.e., words of the form

$$w=v^{*p^k}, \qquad v\in LYN. \tag{7.11}$$

The first step is to prove that all words can be written as polynomials in the elements of $SL(p)$. Let w be a word over $\mathbf{N}$. With induction we can assume that all smaller words can be written as polynomials in $SL(p)$ and by induction on weight that all nontrivial products of weight $\leqslant$ weight(w) can be so written. Let

$$w=v_1^{*n_1} * v_2^{*n_2} * \cdots * v_m^{*n_m}, \qquad v_i\in LYN, \quad v_1>v_2>\cdots>v_m, \tag{7.12}$$

be its Chen–Fox–Lyndon factorization. Consider products of the form

$$\prod_{i=1}^{k_1} v_1^{*n_{1i}} \prod_{i=1}^{k_2} v_2^{*n_{2i}} \cdots \prod_{i=1}^{k_m} v_m^{*n_{mi}}, \tag{7.13}$$

where the products are overlapping shuffle products and where $n_{j1}+\cdots+n_{jk_j}=n_j$, $j=1, ..., m$. The largest word occurring in such a product (in the ordering we are using) will be the word w, independent of how the various star powers are broken up. However, the coefficient of w will depend on how the star powers of the v_j are broken up. Indeed, the coefficient will be the product of multinomial coefficients;

$$\binom{n_1}{n_{11}\cdots n_{1k_1}}\binom{n_2}{n_{21}\cdots n_{2k_2}}\cdots\binom{n_m}{n_{m1}\cdots n_{mk_m}}. \tag{7.14}$$

For instance, if one takes $n_{ij}=1, \forall i, j$ (which is what is done to prove $\mathscr{M}\otimes\mathbf{Q}=\mathbf{Q}[LYN]$; see Section 5 above), the coefficient is $n_1!\,n_2!\,\cdots n_m!$; and if one takes the other extreme, $k_1=k_2=\cdots=k_m=1$, the coefficient is 1. Here, for our present purposes, we break up each n_j according to its p-adic expansion. That is, if $n=a_0+a_1p+\cdots+a_kp^k$, $a_i\in\{0, 1, \ldots, p-1\}$, then it is partitioned (broken up) into

$$\underbrace{p^k, p^k, \ldots, p^k}_{a_k \text{ parts}}, \underbrace{p^{k-1}, \ldots, p^{k-1}}_{a_{k-1} \text{ parts}}, \ldots, \underbrace{p, \ldots, p}_{a_1 \text{ parts}}, \underbrace{1, \ldots, 1}_{a_0 \text{ parts}}. \tag{7.15}$$

Then Corollary 7.6 above says that in this case the coefficient is nonzero modulo p; i.e., it is an invertible element of $\mathbf{Z}_{(p)}$. This proves that also w can be written as a polynomial in $SL(p)$.

Now, for a given weight n, let $w_1, w_2, \ldots, w_m$ be all the words of that weight that are in $SL(p)$ but are not p-elementary. So, if $w_i=[a_{i1}, \ldots, a_{ik_i}]$, $p\,|\,\gcd\{a_{i1}, \ldots, a_{ik_i}\}$. Let

$$b_{ij}=p^{-1}a_{ij}, \qquad v_i=[b_{i1}, \ldots, b_{ik_i}]. \tag{7.16}$$

Now consider the overlapping shuffle powers v_i^p. It is easy to see that these are of the form

$$v_i^p=w_i+p\,(\text{something of weight } n). \tag{7.17}$$

By what has been proved, each of these somethings of weight n can be written as polynomials in the $SL(p)$. Do so. Now calculate modulo nontrivial products and the elements of $ESL(p)$. The result will be m congruence relations;

$$\begin{gathered} a_{11}w_1+\cdots+a_{1m}w_m\equiv 0\\ \vdots\\ a_{m1}w_1+\cdots+a_{mm}w_m\equiv 0, \end{gathered} \tag{7.18}$$

where the matrix $A=(a_{ij})$ has the property $A\equiv I_m \bmod p$. This means that the determinant of the matrix A is invertible in $\mathbf{Z}_{(p)}$, so that the $w_1, \ldots, w_m$ can be eliminated. This proves that the elements from $ESL(p)$ suffice to generate all of $\mathcal{M} \otimes \mathbf{Z}_{(p)}$ over $\mathbf{Z}_{(p)}$. Using Lemma 7.9 above on the cardinality of $ESL(p)$, the same sort of counting argument as that used before in Section 4 finishes the proof.

In more detail, let $\mathcal{F}$ be the free graded algebra over $\mathbf{Z}_{(p)}$ with β_n generators of weight n. Let γ_n be the rank of the free $\mathbf{Z}_{(p)}$ module of elements of weight n. The γ_n are of course recursively determined by the β_n, but the precise formula is not important here. The algebra $\mathbf{Z}_{(p)}[ESL(p)]$ viewed as the free commutative algebra over $\mathbf{Z}_{(p)}$ generated by the symbols from $ESL(p)$ is of course the same thing as $\mathcal{F}$. By what has been proved the natural homomorphism

$$\mathbf{Z}_{(p)}[ESL(p)] \xrightarrow{a} \mathcal{M} \otimes \mathbf{Z}_{(p)}$$

that sends a symbol from $ESL(p)$ to the corresponding element from $\mathcal{M} \otimes \mathbf{Z}_{(p)}$ is surjective.

Both algebras are torsion free, and after tensoring with the rationals the dimensions of their homogeneous parts of weight n are equal by the lemma above and the isomorphism between the overlapping shuffle algebra and the shuffle algebra of Section 5. It follows that α is an isomorphism because surjective homomorphisms between free $\mathbf{Z}_{(p)}$ modules of equal rank are necessarily isomorphisms.

8. PROOFS OF THE MAIN THEOREM

Using the p-adic theorem of Section 6 above, I give in this final section two proofs of the main theorem in the following slightly more precise formulation:

8.1. THEOREM. *The algebra of quasi-symmetric functions over the integers is a free (graded) commutative algebra over the integers with for each $n = 1, 2, \ldots$ precisely β_n generators of weight n.*

8.2. COROLLARY. *Every primitive element in the Leibniz–Hopf algebra $\mathbf{Z}\langle Z\rangle$ can be extended to a divided power sequence of infinite length and every divided power series of finite length n can be extended to one of infinite length.*

For the meanings of the terms in the corollary, see below.

First Proof. Let $\mathcal{M}_n$ be the graded part of weight n of $\mathcal{M}$. By the fact that $\mathcal{M}_n$ is a free Abelian group and by the structure theorem for $\mathcal{M} \otimes_{\mathbf{Z}} \mathbf{Q}$ from Section 5 above, we know that $\mathcal{M}_n$ is a free Abelian group of rank γ_n.

Let G_n be defined by the short exact sequence

$$\bigoplus_{j=1}^{n-1} (\mathscr{M}_j \otimes \mathscr{M}_{n-j}) \to \mathscr{M}_n \to G_n \to 0, \tag{8.3}$$

where the first arrow is given by multiplication. Each G_n is a finitely generated Abelian group. Tensoring with $\mathbf{Z}_{(p)}$ (which is right exact) gives the corresponding exact sequence for $\mathscr{M} \otimes \mathbf{Z}_{(p)}$ and it follows from the p-adic version of the Diners conjecture proved above that $G_n \otimes_{\mathbf{Z}} \mathbf{Z}_{(p)}$ is a free $\mathbf{Z}_{(p)}$ module of rank β_n for each prime number p. This implies that G_n is a free Abelian group of rank β_n and proves that the algebra of symmetric functions can be generated by a set of homogeneous elements $y_{n,1}, y_{n,2}, \ldots, y_{n,\beta_n}$, $n = 1, 2, \ldots$, giving a homogenous surjective ring homomorphism

$$\mathbf{Z}[Y] \xrightarrow{a} \mathscr{M}, \tag{8.4}$$

where $\mathbf{Z}[Y]$ is the graded ring generated by symbols Y_{n,i_n}, $n = 1, 2, \ldots$; $i_n = 1, \ldots, \beta_n$ of weight n. However, the homogenous parts of weight n of $\mathbf{Z}[Y]$ and $\mathscr{M}$ both are free Abelian groups of rank γ_n. It follows immediately that the homogeneous components, $\alpha_n \colon \mathbf{Z}[Y]_n \to \mathscr{M}_n$, of α are isomorphisms and hence that α itself is an isomorphism.

The second proof of the theorem uses the notions of free coalgebras and divided power sequences in coalgebras. This proof is more difficult and perhaps less elegant. I include it here because it seems to offer a better chance to obtain explicit generators; that is, to make progress toward proving the strong Ditters conjecture.

Let B be a free graded module over the integers or over $\mathbf{Z}_{(p)}$ (or over any ring R) whose homogeneous summands are of finite rank, and let B^* be its graded dual. The *graded cofree coalgebra* over the integers, $\mathrm{CoF}(B)$, determined by B is the graded dual of the free associative graded algebra, $\mathrm{Fr}(B^*)$, over the integers generated by B^*. It can be characterized by a universal property that is dual to that of free associative algebras as follows (though this is not important here): It comes with a canonical map $\pi \colon \mathrm{CoF}(B) \to B$, the graded dual of the canonical map $B^* \to \mathrm{Fr}(B^*)$, and satisfies the following property: For every graded map of a graded coalgebra C to the module B, $C \xrightarrow{\varphi} B$, there is a unique morphism of graded coalgebras $C \xrightarrow{\psi} \mathrm{CoF}(B)$ such that $\pi\psi = \varphi$.

The graded cofree coalgebra $\mathrm{CoF}(B)$ can be explicitly described as follows: Take the tensor module

$$T(B) = \bigoplus_{i=0}^{\infty} B^{\otimes i} = \mathbf{Z} \oplus B \oplus B^{\otimes 2} \oplus B^{\otimes 3} \oplus \cdots . \tag{8.5}$$

There are natural isomorphisms

$$\varphi_{i,j}\colon B^{\otimes i} \otimes B^{\otimes j} \to B^{\otimes (i+j)}, \qquad i, j = 0, 1, 2, \ldots . \tag{8.6}$$

Using these, the comultiplication on $T(B)$ is defined by

$$\mu(b_1 \otimes b_2 \otimes \cdots \otimes b_n) = \sum_{i=0}^{n} \varphi^{-1}_{i,n-i}(b_1 \otimes b_2 \otimes \cdots \otimes b_n). \tag{8.7}$$

The cofree coalgebra $\mathrm{CoF}(B)$ has a unique group-like element, viz. $1 \in \mathbf{Z}$ (which is the dual of the augmentation of $\mathrm{Fr}(B^*)$). The *primitives* of $\mathrm{CoF}(B)$, i.e., the elements such that $\mu(x) = 1 \otimes x + x \otimes 1$, are the elements of $B \subset T(B)$.

A *divided power sequence* of length n (resp. of length ∞) over a group-like element 1 in a coalgebra C is a sequence of elements $\delta_1, \delta_2, \ldots, \delta_n$ (resp. $\delta_1, \delta_2, \ldots$) such that

$$\mu(\delta_m) = 1 \otimes \delta_m + \sum_{i=1}^{m-1} \delta_i \otimes \delta_{m-i} + \delta_m \otimes 1, \qquad m = 1, 2, \ldots . \tag{8.8}$$

Note that δ_1 is primitive. In $\mathrm{CoF}(B)$ every primitive element b can be extended to a divided power sequence of infinite length. Indeed, one such sequence is

$$b,\ b \otimes b,\ b \otimes b \otimes b,\ \ldots . \tag{8.9}$$

If H is a Hopf algebra, there is a natural multiplication on the set $DSP(H)$ of divided power sequences of infinite length given by

$$\begin{aligned}&(\delta_1, \delta_2, \delta_3, \ldots), (\delta'_1, \delta'_2, \delta'_3, \ldots)\\ &\quad\mapsto (\delta_1 + \delta'_1, \delta_2 + \delta_1\delta'_1 + \delta'_2, \delta_3 + \delta_2\delta'_1 + \delta_1\delta'_2 + \delta'_3, \ldots),\end{aligned} \tag{8.10}$$

and this turns $DSP(H)$ into a (functorial) group. There are also Verschiebungs operators

$$\mathbf{V}_n\colon (\delta_1, \delta_2, \ldots) \mapsto (\underbrace{0, \ldots, 0}_{n-1},\ \delta_1, \underbrace{0, \ldots, 0}_{n-1},\ \delta_2, \ldots). \tag{8.11}$$

Using these, multiplication, and the fact that $\mathrm{CoF}(B)$ is a Hopf algebra (in many ways; $\mathrm{CoF}(B)$ being the dual of the free associative algebra $\mathrm{Fr}(B^*)$), it now easily follows that any divided power sequence of length n can be extended to one of infinite length. This is seen by induction because if

$$\delta_1, \delta_2, \ldots, \delta_n \qquad \text{and} \qquad \delta_1, \delta_2, \ldots, \delta'_n \tag{8.12}$$

are two different divided power sequences that agree up to degree $n-1$, then the difference of the last terms, $\delta_n-\delta'_n$, is a primitive.

The cofree cocommutative graded coalgebra, CCoF(B), over B, is the subcoalgebra of CoF(B) of symmetric tensors. It is the graded dual of the commutative free algebra generated by B^* as the maximal commutative quotient of Fr(B^*).

8.13. *Remark.* The free commutative coalgebra over B, which satisfies the same universal property for not necessarily graded coalgebras and morphisms is not CoF(B), but a certain recursive completion; see [10] for details.

Given these preparations we can now give a second proof of the main theorem. This proof proceeds by first proving Corollary 8.2 about divided power sequences.

Second proof. Consider the Leibniz–Hopf algebra $\mathscr{Z}=\mathbf{Z}\langle Z\rangle$ and its various localizations $\mathscr{Z}\otimes\mathbf{Z}_{(p)}$. Let $\delta_1,\delta_2,\ldots,\delta_n$ be a divided power sequence of length $n\geqslant 1$ in the Leibniz–Hopf algebra. It suffices to prove that every such sequence can be extended to one of length $n+1$. By what has been said above, for every prime number p this can be done in $\mathscr{Z}\otimes\mathbf{Z}_{(p)}$. This gives the divided power sequences

$$\delta_1,\delta_2,\ldots,\delta_n,\delta^{(p)}_{n+1}=t_p^{-1}\alpha^{(p)},\qquad \alpha^{(p)}\in\mathscr{Z},\quad (p,t_p)=1. \tag{8.14}$$

Let $p_{\max}$ be the largest prime number dividing t_2. Then $(t_2,\ldots,t_{p_{\max}})=1$ and so there exist integers $d_2,\ldots,d_{p_{\max}}$ such that

$$d_2t_2+d_3t_3+\cdots+d_{p_{\max}}t_{p_{\max}}=1. \tag{8.15}$$

Multiply this equation with t_2 to get numbers c_p that are multiples of t_p for all primes from 3 through $p_{\max}$ inclusive such that

$$c_3+c_5+\cdots+c_{p_{\max}}\equiv t_2\mod(t_2^2). \tag{8.16}$$

Now the differences $(t_p^{-1}\alpha^{(p)}-t_2^{-1}\alpha^{(2)})$ are primitives (in the algebra $\mathscr{Z}\otimes\mathbf{Q}$ because of the freeness of $\mathscr{M}\otimes_{\mathbf{Z}}\mathbf{Q}$) and it follows that the sequence

$$\delta_1,\delta_2,\ldots,\delta_n,\alpha \tag{8.17}$$

with

$$\alpha=t_2^{-1}\alpha^{(2)}+\sum_{p\mid t_2}c_p(t_p^{-1}\alpha^{(p)}-t_2^{-1}\alpha^{(2)}) \tag{8.18}$$

is a divided power series of length $n+1$, and because of the properties of the coefficients c_p it is in fact defined over the integers. This concludes the proof of Corollary 8.2.

The theorem itself now follows by a straightforward imitation of the proof given in [20] of the structure of irreducible cocommutative Hopf algebras over fields of characteristic zero. See Theorem 13.0.1 in Chapter 13 of [20]. Given the corollary, one can also repair the original proof of Ditters in [2].

REFERENCES

1. K. T. Chen, R. H. Fox, and R. C. Lyndon, Free differential calculus, IV, *Ann. Math.* **68** (1958), 81–95.
2. E. J. Ditters, Curves and formal (co)groups, *Invent. Math.* **17** (1972), 1–20.
3. E. J. Ditters, "Groupes formels," (course notes), Chap. 2, Sect. 5, p. 29, Université de Paris XI, Orsay, 1974.
4. E. J. Ditters and A. C. J. Scholtens, Free polynomial generators for the Hopf algebra QSym of quasi-symmetric functions, *J. Pure Appl. Algebra* **144** (1999), 213–227.
5. I. M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon, Noncommutative symmetric functions, *Adv. Math.* **122** (1995), 218–348.
6. I. M. Gessel, Multipartite *P*-partitions and inner product of skew Schur functions, *in* "Contemporary Mathematics," Vol. 34, pp. 289–301, American Mathematical Society, Providence, RI, 1984.
7. I. M. Gessel and C. Reutenauer, Counting permutations with given cycle-structure and descent set, *J. Combin. Theory Ser. A* **64** (1993), 189–215.
8. M. Hazewinkel, The Leibniz–Hopf algebra and Lyndon words, preprint, CWI, Amsterdam, 1996.
9. M. Hazewinkel, Leibniz–Hopf algebra, *in* "Encyclopaedia of Mathematics," supplement Vol. I, pp. 349–350, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.
10. M. Hazewinkel, Cofree coalgebras and recursiveness, preprint, CWI, 1999.
11. M. Hazewinkel, Generalized overlapping shuffle algebras, *in* "Proceedings, Pontryagin Memorial Conference, Moscow, 1998," Vol. 8, "Algebra" (S. M. Aseev and S. A. Vakhrameev, Eds.), pp. 193–222, Viniti, 1999 [in Russian]; preprint version, CWI, 1998 [in English].
12. M. Hazewinkel, Quasi-symmetric functions, *in* "Formal Series and Algebraic Combinatorics, Moscow, June 2000" (D. Krob, A. A. Miklalev, and A. V. Miklalev, Eds.), pp. 30–44, Springer, Heidelberg, Germany, 2000.
13. D. Krob and J.-Y. Thibon, Noncommutative symmetric functions. IV. Quantum linear groups and Hecke algebras at $q=0$, *J. Algebraic Combin.* **6** (1997), 339–376.
14. M. Lothaire (Ed.), "Combinatorics on Words," Encyclopedia of Mathematics and Its Applications, Vol. 17, Addison–Wesley, Reading, MA, 1983.
15. C. Malvenuto and C. Reutenauer, Duality between quasi-symmetric functions and the Solomon descent algebra, *J. Algebra* **177** (1994), 967–982.
16. C. Reutenauer, "Free Lie Algebras," Oxford Univ. Press, Oxford, UK, 1993.
17. A. C. J. Scholtens, On the graded dual of the noncommutative universal Leibniz Hopf algebra *Z*, preprint, Math. Seminar of the Free Univ. of Amsterdam, 1994.

18. A. C. J. Scholtens, "*S*-Typical Curves in Non-commutative Hopf Algebras," thesis, Free University of Amsterdam, 1996.
19. R. P. Stanley, On the number of reduced decompositions of elements of Coxeter groups, *European J. Combin.* **5** (1984), 359–372.
20. M. E. Sweedler, "Hopf Algebras," Benjamin, New York, 1969.
21. J.-Y. Thibon and B.-C.-V. Ung, Quantum quasi-symmetric functions and Hecke algebras, *J. Phys. A Math. Gen.* **29** (1996), 7337–7348.