Russian Math. Surveys **58**:1 000–000 ©2003 RAS(DoM) and LMS

Uspekhi Mat. Nauk **58**:1 3–32 UDC 511.3+512.573 DOI XXXX

Algebraic relations for multiple zeta values

V. V. Zudilin [W. Zudilin]

Abstract. The survey is devoted to the multidimensional generalization of the Riemann zeta function as a function of positive integral argument.

1. Introduction

In the domain $\operatorname{Re} s > 1$, the *Riemann zeta function* can be defined by the convergent series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} . \tag{1}$$

One of the interesting and still unsolved problems is the problem concerning the polynomial relations over $\mathbb{Q}$ for the numbers $\zeta(s)$, $s = 2, 3, 4, \dots$. Thanks to Euler, we know the formula

$$\zeta(s) = -\frac{(2\pi i)^s B_s}{2s!} \qquad \text{for} \quad s = 2, 4, 6, \dots, \tag{2}$$

which expresses the values of the zeta function at even points in terms of the number

$$\pi = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 3.14159265358979323846\dots$$

and the Bernoulli numbers $B_s \in \mathbb{Q}$ defined by the generating function

$$\frac{t}{e^t - 1} = 1 - \frac{t}{2} + \sum_{s=2}^{\infty} B_s \frac{t^s}{s!} , \qquad B_s = 0 \quad \text{for odd} \quad s \geqslant 3. \tag{3}$$

The relation (2) yields the coincidence of the rings $\mathbb{Q}[\zeta(2), \zeta(4), \zeta(6), \zeta(8), \dots]$ and $\mathbb{Q}[\pi^2]$, and hence, due to Lindemann's theorem [17] on the transcendence of π, we can conclude that each of the rings is of transcendence degree 1 over the field of rational numbers. Much less is known on the arithmetic nature of values of the zeta function at odd integers $s = 3, 5, 7, \dots$, namely, Apéry has proved [1] that the number $\zeta(3)$ is irrational and Rivoal recently showed [22] that there are infinitely many irrational numbers in the list $\zeta(3), \zeta(5), \zeta(7), \dots$. Conjecturally, each of these numbers is transcendent, and the above question on the polynomial relations over $\mathbb{Q}$ for the values of the series (1) at the integers s, $s \geqslant 2$, has the following simple answer.

Conjecture 1. *The numbers*

$$\pi,\ \zeta(3),\ \zeta(5),\ \zeta(7),\ \zeta(9),\ \dots$$

are algebraically independent over $\mathbb{Q}$.

This conjecture can be regarded as a part of mathematical folklore (see, for instance, [7] and [28]). In this survey we discuss a generalization of the problem of algebraic independence for values of the Riemann zeta function at positive integers (the so-called *zeta values*). Namely, we speak of the object extensively studied during the last decade in connection with problems concerning not only number theory but also combinatorics, algebra, analysis, algebraic geometry, quantum physics, and many other branches of mathematics. At the same time, no works in this direction have been published in Russian till now (we only mention the paper [25] in the press). By means of the present publication, we hope to attract attention of Russian mathematicians to problems connected with *multiple zeta values.*

The author is deeply indebted to the referee for several valuable remarks that have essentially improved the presentation.

2. Multiple zeta values

The series (1) admits the following multidimensional generalization. For positive integers $s_1, s_2, \dots, s_l$, where $s_1 > 1$, we consider the values of the l-tuple zeta function

$$\zeta(\boldsymbol{s}) = \zeta(s_1, s_2, \dots, s_l) := \sum_{n_1 > n_2 > \dots > n_l \geqslant 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}; \tag{4}$$

in what follows, the corresponding multi-index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$ is said to be *admissible.* The quantities (4) are called *multiple zeta values* [30] (and abbreviated as MZVs), or *multiple harmonic series* [10], or *Euler sums.* The sums (4) for $l = 2$ originate from Euler [5] who obtained a family of identities connecting double and ordinary zeta values (see the corollary to Theorem 1 below). In particular, Euler had proved the identity

$$\zeta(2, 1) = \zeta(3), \tag{5}$$

which was multiply rediscovered since then. The quantities (4) were introduced by Hoffman in [10] and independently by Zagier in [30] (with the opposite order of summation on the right-hand side of (4)); moreover, in [10] and [30] some $\mathbb{Q}$-linear and $\mathbb{Q}$-polynomial relations were established and several conjectures were stated (some of which were proved later on) concerning the structure of algebraic relations for the family (4). Hoffman also suggested [10] the alternative definition

$$\tilde{\zeta}(\boldsymbol{s}) = \tilde{\zeta}(s_1, s_2, \dots, s_l) := \sum_{n_1 \geqslant n_2 \geqslant \dots \geqslant n_l \geqslant 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}} \tag{6}$$

of the Euler sums with non-strict inequalities in the summation. Clearly, all relations for the series (6) can readily be rewritten for the series (4) (see, for instance, [10] and [25]), although several identities possess a brief form in the very terms of multiple zeta values (6) (see the relations (38) in Section 7 below).

For each number (4) we introduce two characteristics, namely, the *weight* (or the *degree*) $|\boldsymbol{s}| := s_1 + s_2 + \dots + s_l$ and the *length* (or the *depth*) $\ell(\boldsymbol{s}) := l$.

We note [31] that the series on the right-hand side of (4) converges absolutely in the domain given by $\operatorname{Re} s_1 > 1$ and $\sum_{k=1}^{l} \operatorname{Re} s_k > l$; moreover, the multiple zeta function $\zeta(\boldsymbol{s})$ defined in the domain by the series (4) can be continued analytically to a meromorphic function on the whole space $\mathbb{C}^l$ with possible simple poles at the hyperplanes $s_1 = 1$ and $\sum_{k=1}^{j} s_k = j+1-m$, where j, $1 < j \leqslant l$, and m, $m \geqslant 1$, are integers. The problems concerning the existence of a functional equation for $l > 1$ and the localization of non-trivial zeros (an analogue of Riemann's conjecture) for the function $\zeta(\boldsymbol{s})$ remain open.

3. Identities: the method of partial fractions

In this section we present examples of identities (for multiple zeta values) that are proved by an elementary analytic method, namely, the *method of partial fractions.*

Theorem 1 (Hoffman's relations [10], Theorem 5.1). *The identity*

$$\sum_{k=1}^{l} \zeta(s_1, \dots, s_{k-1}, s_k + 1, s_{k+1}, \dots, s_l) = \sum_{\substack{k=1 \\ s_k \geqslant 2}}^{l} \sum_{j=0}^{s_k - 2} \zeta(s_1, \dots, s_{k-1}, s_k - j, j+1, s_{k+1}, \dots, s_l) \tag{7}$$

holds for any admissible multi-index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$.

Proof. For any $k = 1, 2, \dots, l$ we have

$$\begin{aligned} &\sum_{n_k > n_{k+1} > \dots > n_l \geqslant 1} \frac{1}{n_k^{s_k+1} n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}} + \sum_{n_k > m > n_{k+1} > \dots > n_l \geqslant 1} \frac{1}{n_k^{s_k} m n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}} \\ &\quad = \sum_{n_k \geqslant m > n_{k+1} > \dots > n_l \geqslant 1} \frac{1}{n_k^{s_k} m n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}} \\ &\quad = \sum_{n_k > n_{k+1} > \dots > n_l \geqslant 1} \sum_{m = n_{k+1}+1}^{n_k} \frac{1}{m n_k^{s_k} n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}}, \end{aligned}$$

and hence

$$\begin{aligned} &\zeta(s_1, \dots, s_{k-1}, s_k + 1, s_{k+1}, \dots, s_l) + \zeta(s_1, \dots, s_{k-1}, s_k, 1, s_{k+1}, \dots, s_l) \\ &\quad = \sum_{n_1 > \dots > n_k > n_{k+1} > \dots > n_l \geqslant 1} \frac{1}{n_1^{s_1} \cdots n_k^{s_k+1} n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}} \\ &\qquad + \sum_{n_1 > \dots > n_k > m > n_{k+1} > \dots > n_l \geqslant 1} \frac{1}{n_1^{s_1} \cdots n_k^{s_k} m n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}} \\ &\quad = \sum_{n_1 > \dots > n_k > n_{k+1} > \dots > n_l \geqslant 1} \sum_{m = n_{k+1}+1}^{n_k} \frac{1}{m n_1^{s_1} \cdots n_k^{s_k} n_{k+1}^{s_{k+1}} \cdots n_l^{s_l}} \\ &\quad = \sum_{n_1 > n_2 > \dots > n_l \geqslant 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}} \sum_{m = n_{k+1}+1}^{n_k} \frac{1}{m}. \end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{k=1}^{l}\bigl(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\bigr)\\
&\quad=\sum_{n_1>n_2>\dots>n_l\geqslant1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\sum_{m=1}^{n_1}\frac{1}{m}\\
&\quad=\sum_{m_1,m_2,\dots,m_l\geqslant1}\frac{1}{m_1^{s_l}(m_1+m_2)^{s_{l-1}}\cdots(m_1+\dots+m_l)^{s_1}}\sum_{m=1}^{m_1+\dots+m_l}\frac{1}{m}\\
&\quad=\sum_{m_1,m_2,\dots,m_l\geqslant1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_l^{s_1}}\sum_{m_{l+1}\geqslant1}\biggl(\frac{1}{m_{l+1}}-\frac{1}{M_{l+1}}\biggr), \qquad (8)
\end{aligned}$$

where M_k stands for $m_1+m_2+\dots+m_k$ when $k=1,\dots,l+1$ (clearly, $M_k=n_{l+1-k}$ when $k=1,\dots,l$). We now note the following partial-fraction expansion (with respect to the parameter u):

$$\frac{1}{u(u+v)^s}=\frac{1}{v^s u}-\sum_{j=0}^{s-1}\frac{1}{v^{j+1}(u+v)^{s-j}}\,, \qquad u,v\in\mathbb{R}; \qquad (9)$$

for the proof it suffices to use the fact that on the right-hand side we sum a geometric progression. By setting $u=m_{l+1}$, $v=M_l$, and $s=s_1$ in (9), we obtain

$$\frac{1}{m_{l+1}M_{l+1}^{s_1}}=\frac{1}{m_{l+1}(m_{l+1}+M_l)^{s_1}}=\frac{1}{M_l^{s_1}m_{l+1}}-\sum_{j=0}^{s_1-1}\frac{1}{M_l^{j+1}M_{l+1}^{s_1-j}}\,,$$

and hence

$$\frac{1}{M_l^{s_1}}\biggl(\frac{1}{m_{l+1}}-\frac{1}{M_{l+1}}\biggr)=\sum_{j=0}^{s_1-2}\frac{1}{M_l^{j+1}M_{l+1}^{s_1-j}}+\frac{1}{m_{l+1}M_{l+1}^{s_1}}\,.$$

Continuing the equality (8), we see that

$$\begin{aligned}
&\sum_{k=1}^{l}\bigl(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\bigr)\\
&\quad=\sum_{j=0}^{s_1-2}\sum_{m_1,m_2,\dots,m_{l+1}\geqslant1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_{l-1}^{s_2}M_l^{j+1}M_{l+1}^{s_1-j}}\\
&\qquad+\sum_{m_1,m_2,\dots,m_{l+1}\geqslant1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_{l-1}^{s_2}m_{l+1}M_{l+1}^{s_1}}\\
&\quad=\sum_{j=0}^{s_1-2}\zeta(s_1-j,j+1,s_2,\dots,s_l)+\sum_{m_1,m_2,\dots,m_{l+1}\geqslant1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_{l-1}^{s_2}m_lM_{l+1}^{s_1}} \qquad (10)
\end{aligned}$$

(we have transposed the indices m_l and m_{l+1} in the last multiple sum). Using the identity (9) with $u = m_{k+1}$, $v = M_k = M_{k+1} - m_{k+1}$, and $s = s_{l+1-k}$, we conclude that

$$\frac{1}{M_k^{s_{l+1-k}} m_{k+1}} = \sum_{j=0}^{s_{l+1-k}-1} \frac{1}{M_k^{j+1} M_{k+1}^{s_{l+1-k}-j}} + \frac{1}{m_{k+1} M_{k+1}^{s_{l+1-k}}}, \qquad k = 1, 2, \dots, l-1,$$

and therefore

$$\begin{aligned} &\sum_{m_1, m_2, \dots, m_{l+1} \geqslant 1} \frac{1}{M_1^{s_l} \cdots M_k^{s_{l+1-k}} m_{k+1} M_{k+2}^{s_{l-k}} \cdots M_{l+1}^{s_1}} \\ &\quad = \sum_{j=0}^{s_{l+1-k}-1} \sum_{m_1, m_2, \dots, m_{l+1} \geqslant 1} \frac{1}{M_1^{s_l} \cdots M_{k-1}^{s_{l+2-k}} M_k^{j+1} M_{k+1}^{s_{l+1-k}-j} M_{k+2}^{s_{l-k}} \cdots M_{l+1}^{s_1}} \\ &\qquad + \sum_{m_1, m_2, \dots, m_{l+1} \geqslant 1} \frac{1}{M_1^{s_l} \cdots M_{k-1}^{s_{l+2-k}} m_{k+1} M_{k+1}^{s_{l+1-k}} \cdots M_{l+1}^{s_1}} \\ &\quad = \sum_{j=0}^{s_{l+1-k}-1} \zeta(s_1, \dots, s_{l-k}, s_{l+1-k} - j, j+1, s_{l+2-k}, \dots, s_l) \\ &\qquad + \sum_{m_1, m_2, \dots, m_{l+1} \geqslant 1} \frac{1}{M_1^{s_l} \cdots M_{k-1}^{s_{l+2-k}} m_k M_{k+1}^{s_{l+1-k}} \cdots M_{l+1}^{s_1}}, \qquad (11) \\ &\qquad\qquad k = 1, 2, \dots, l-1. \end{aligned}$$

Successively applying the identities (11) for the multiple sum on the right-hand side of the equality (10) in the inverse order (that is, beginning with $k = l - 1$ and ending with $k = 1$), we obtain

$$\begin{aligned} &\sum_{k=1}^{l} \bigl(\zeta(s_1, \dots, s_{k-1}, s_k + 1, s_{k+1}, \dots, s_l) + \zeta(s_1, \dots, s_{k-1}, s_k, 1, s_{k+1}, \dots, s_l)\bigr) \\ &\quad = \sum_{j=0}^{s_1-2} \zeta(s_1 - j, j+1, s_2, \dots, s_l) \\ &\qquad + \sum_{k=1}^{l-1} \sum_{j=0}^{s_{l+1-k}-1} \zeta(s_1, \dots, s_{l-k}, s_{l+1-k} - j, j+1, s_{l+2-k}, \dots, s_l) \\ &\qquad + \sum_{m_1, m_2, \dots, m_{l+1} \geqslant 1} \frac{1}{m_1 M_2^{s_l} M_3^{s_{l-1}} \cdots M_{l+1}^{s_1}} \\ &\quad = \sum_{k=1}^{l} \sum_{j=0}^{s_k-2} \zeta(s_1, \dots, s_{k-1}, s_k - j, j+1, s_{k+1}, \dots, s_l) \\ &\qquad + \sum_{k=1}^{l} \zeta(s_1, \dots, s_{k-1}, s_k, 1, s_{k+1}, \dots, s_l). \qquad (12) \end{aligned}$$

After necessary reductions on the left- and right-hand sides of the equality (12), we finally arrive at the desired identity (7).

For $l = 1$ the statement of Theorem 1 can be represented in the following form.

Corollary (Euler's theorem). *The identity*

$$\zeta(s)=\sum_{j=1}^{s-2}\zeta(s-j,j) \tag{13}$$

holds for any integer $s\geqslant 3$.

We also note that for $s=3$ the identity (13) is simply the relation (5).

In [13] the following result was also proved by using the method of partial fractions.

Theorem 2 (Cyclic sum theorem). *The identity*

$$\begin{aligned}&\sum_{k=1}^{l}\zeta(s_k+1,s_{k+1},\dots,s_l,s_1,\dots,s_{k-1})\\&\qquad=\sum_{\substack{k=1\\ s_k\geqslant 2}}^{l}\sum_{j=0}^{s_k-2}\zeta(s_k-j,s_{k+1},\dots,s_l,s_1,\dots,s_{k-1},j+1)\end{aligned}$$

holds for any admissible multi-index $\boldsymbol{s}=(s_1,s_2,\dots,s_l)$.

Theorem 2 immediately proves that the sum of all multiple zeta values of fixed length and fixed weight does not depend on the length; this statement, as well as Theorem 1, generalizes Euler's theorem cited above.

Theorem 3 (Sum theorem). *The identity*

$$\sum_{\substack{s_1>1,s_2\geqslant 1,\dots,s_l\geqslant 1\\ s_1+s_2+\dots+s_l=s}}\zeta(s_1,s_2,\dots,s_l)=\zeta(s)$$

holds for any integers $s>1$ *and* $l\geqslant 1$.

Theorems 1 and 3 are special cases of Ohno's relations [21], which will be discussed in Section 12 below.

4. Algebra of multiple zeta values

This section is based on the works [11] and [30]. To describe the known algebraic relations (that is, the numerical identities) over $\mathbb{Q}$ for the quantities (4), it is useful to represent ζ as a linear map of a certain polynomial algebra into the field of real numbers. Let us consider the coding of multi-indices $\boldsymbol{s}$ by words (that is, by monomials in non-commutative variables) over the alphabet $X=\{x_0,x_1\}$ by the rule

$$\boldsymbol{s}\mapsto x_{\boldsymbol{s}}=x_0^{s_1-1}x_1x_0^{s_2-1}x_1\cdots x_0^{s_l-1}x_1.$$

We set

$$\zeta(x_{\boldsymbol{s}}):=\zeta(\boldsymbol{s}) \tag{14}$$

for all admissible words (that is, beginning with x_0 and ending with x_1); then the weight (or the degree) $|x_{\boldsymbol{s}}|:=|\boldsymbol{s}|$ coincides with the total degree of the monomial $x_{\boldsymbol{s}}$, whereas the length $\ell(x_{\boldsymbol{s}}):=\ell(\boldsymbol{s})$ is the degree with respect to the variable x_1.

Let $\mathbb{Q}\langle X\rangle = \mathbb{Q}\langle x_0, x_1\rangle$ be the $\mathbb{Q}$-algebra of polynomials in two non-commutative variables which is graded by the degree (where each of the variables x_0 and x_1 is assumed to be of degree 1); we identify the algebra $\mathbb{Q}\langle X\rangle$ with the graded $\mathbb{Q}$-vector space $\mathfrak{H}$ spanned by the monomials in the variables x_0 and x_1. We also introduce the graded $\mathbb{Q}$-vector spaces $\mathfrak{H}^1 = \mathbb{Q}\mathbf{1} \oplus \mathfrak{H}x_1$ and $\mathfrak{H}^0 = \mathbb{Q}\mathbf{1} \oplus x_0\mathfrak{H}x_1$, where $\mathbf{1}$ denotes the unit (the empty word of weight 0 and length 0) of the algebra $\mathbb{Q}\langle X\rangle$. Then the space $\mathfrak{H}^1$ can be regarded as the subalgebra of $\mathbb{Q}\langle X\rangle$ generated by the words $y_s = x_0^{s-1}x_1$, whereas $\mathfrak{H}^0$ is the $\mathbb{Q}$-vector space spanned by all admissible words. We can now regard the function ζ as the $\mathbb{Q}$-linear map $\zeta\colon \mathfrak{H}^0 \to \mathbb{R}$ defined by the relations $\zeta(\mathbf{1}) = 1$ and (14).

Let us define the products $ш$ (the *shuffle product*) on $\mathfrak{H}$ and $*$ (the *harmonic* or *stuffle product*) on $\mathfrak{H}^1$ by the rules

$$\mathbf{1} ш w = w ш \mathbf{1} = w, \qquad \mathbf{1} * w = w * \mathbf{1} = w \tag{15}$$

for any word w, and

$$x_j u ш x_k v = x_j(u ш x_k v) + x_k(x_j u ш v), \tag{16}$$

$$y_j u * y_k v = y_j(u * y_k v) + y_k(y_j u * v) + y_{j+k}(u * v) \tag{17}$$

for any words u, v, any letters x_j, x_k, and any generators y_j, y_k of the subalgebra $\mathfrak{H}^1$, and then extend the rules (15)–(17) to the whole algebra $\mathfrak{H}$ and the whole subalgebra $\mathfrak{H}^1$ by linearity. Sometimes it becomes useful to consider the stuffle product on the whole algebra $\mathfrak{H}$ by formally adding to (17) the rule

$$x_0^j * w = w * x_0^j = w x_0^j \tag{18}$$

for any word w and any integer $j \geqslant 1$. We note that induction arguments enable us to prove that each of the above products is commutative and associative (for the proof, see Section 8 below); the corresponding algebras $\mathfrak{H}_ш := (\mathfrak{H}, ш)$ and $\mathfrak{H}^1_* := (\mathfrak{H}^1, *)$ (and also $\mathfrak{H}_* := (\mathfrak{H}, *)$) are examples of the so-called *Hopf algebras*.

The following two statements motivate the consideration of the above products $ш$ and $*$; their proofs can be found in [11], [13], and [28].

Theorem 4. *The map ζ is a homomorphism of the shuffle algebra $\mathfrak{H}^0_ш := (\mathfrak{H}^0, ш)$ into $\mathbb{R}$, that is,*

$$\zeta(w_1 ш w_2) = \zeta(w_1)\zeta(w_2) \quad \textit{for all} \quad w_1, w_2 \in \mathfrak{H}^0. \tag{19}$$

Theorem 5. *The map ζ is a homomorphism of the stuffle algebra $\mathfrak{H}^0_* := (\mathfrak{H}^0, *)$ into $\mathbb{R}$, that is,*

$$\zeta(w_1 * w_2) = \zeta(w_1)\zeta(w_2) \quad \textit{for all} \quad w_1, w_2 \in \mathfrak{H}^0. \tag{20}$$

In what follows we present detailed proofs of these two theorems by using the differential-difference origin of the products $ш$ and $*$ in suitable functional models of the algebras $\mathfrak{H}_ш$ and $\mathfrak{H}^0_*$. When proving Theorem 4 (see Section 5), we follow the scheme of the paper [27], whereas our proof of Theorem 5 (in Section 9) is new.

Another family of identities is given by the following statement which is derived from Theorem 1 in Section 11.

Theorem 6. *The map* ζ *satisfies the relations*

$$\zeta(x_1 \sqcup\!\sqcup w - x_1 * w) = 0 \quad \textit{for all} \quad w \in \mathfrak{H}^0 \tag{21}$$

(*in particular, the polynomials* $x_1 \sqcup\!\sqcup w - x_1 * w$ *belong to* $\mathfrak{H}^0$).

All relations (both proved and experimentally obtained) known at present for the multiple zeta values follow from the identities (19)–(21). Therefore, the following conjecture looks quite plausible.

Conjecture 2 ([11], [18], [27]). *All algebraic relations over* $\mathbb{Q}$ *among the multiple zeta values are generated by the identities* (19)–(21); *equivalently,*

$$\ker\zeta = \{u \sqcup\!\sqcup v - u * v : u \in \mathfrak{H}^1,\ v \in \mathfrak{H}^0\}.$$

5. Shuffle algebra of generalized polylogarithms

To prove the shuffle relations (19) for multiple zeta values, we define the *generalized polylogarithms*

$$\mathrm{Li}_{\boldsymbol{s}}(z) := \sum_{n_1>n_2>\dots>n_l\geqslant 1} \frac{z^{n_1}}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}, \qquad |z| < 1, \tag{22}$$

where l is a positive integer, for any l-tuple of positive integers $s_1, s_2, \dots, s_l$. By definition,

$$\mathrm{Li}_{\boldsymbol{s}}(1) = \zeta(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \geqslant 2,\ s_2 \geqslant 1,\ \dots,\ s_l \geqslant 1. \tag{23}$$

By setting

$$\mathrm{Li}_{x_{\boldsymbol{s}}}(z) := \mathrm{Li}_{\boldsymbol{s}}(z), \qquad \mathrm{Li}_{\mathbf{1}}(z) := 1, \tag{24}$$

as above in the case of multiple zeta values, we extend the action of the map $\mathrm{Li}\colon w \mapsto \mathrm{Li}_w(z)$ by linearity to the graded algebra $\mathfrak{H}^1$ (rather than to $\mathfrak{H}$ because the multi-indices are coded by words in $\mathfrak{H}^1$).

Lemma 1. *Let* $w \in \mathfrak{H}^1$ *be an arbitrary non-empty word and let* x_j *be the first letter in its representation* (*that is,* $w = x_j u$ *for some word* $u \in \mathfrak{H}^1$). *Then*

$$\frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_w(z) = \frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_{x_j u}(z) = \omega_j(z)\,\mathrm{Li}_u(z), \tag{25}$$

where

$$\omega_j(z) = \omega_{x_j}(z) := \begin{cases} \dfrac{1}{z} & \textit{if } x_j = x_0, \\[2mm] \dfrac{1}{1-z} & \textit{if } x_j = x_1. \end{cases} \tag{26}$$

Proof. Assuming that $w = x_j u = x_{\boldsymbol{s}}$ for some multi-index $\boldsymbol{s}$, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_w(z) = \frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_{\boldsymbol{s}}(z) &= \frac{\mathrm{d}}{\mathrm{d}z} \sum_{n_1>n_2>\dots>n_l\geqslant 1} \frac{z^{n_1}}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}, \\ &= \sum_{n_1>n_2>\dots>n_l\geqslant 1} \frac{z^{n_1-1}}{n_1^{s_1-1} n_2^{s_2} \cdots n_l^{s_l}}. \end{aligned}$$

Therefore, if $s_1>1$ (which corresponds to the letter $x_j=x_0$), then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_0u}(z)&=\frac{1}{z}\sum_{n_1>n_2>\dots>n_l\geqslant1}\frac{z^{n_1}}{n_1^{s_1-1}n_2^{s_2}\cdots n_l^{s_l}}\\&=\frac{1}{z}\operatorname{Li}_{s_1-1,s_2,\dots,s_l}(z)=\frac{1}{z}\operatorname{Li}_u(z)\end{aligned}$$

and, if $s_1=1$ (which corresponds to the letter $x_j=x_1$), then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_1u}(z)&=\sum_{n_1>n_2>\dots>n_l\geqslant1}\frac{z^{n_1-1}}{n_2^{s_2}\cdots n_l^{s_l}}=\sum_{n_2>\dots>n_l\geqslant1}\frac{1}{n_2^{s_2}\cdots n_l^{s_l}}\sum_{n_1=n_2+1}^{\infty}z^{n_1-1}\\&=\frac{1}{1-z}\sum_{n_2>\dots>n_l\geqslant1}\frac{z^{n_2}}{n_2^{s_2}\cdots n_l^{s_l}}=\frac{1}{1-z}\operatorname{Li}_{s_2,\dots,s_l}(z)=\frac{1}{1-z}\operatorname{Li}_u(z),\end{aligned}$$

and the result follows.

Lemma 1 motivates another definition of the generalized polylogarithms, which can now be defined for all elements of the algebra $\mathfrak{H}$. As above, it suffices to define the polylogarithms for the words $w\in\mathfrak{H}$ only and then extend the definition to whole algebra by linearity. We set $\operatorname{Li}_{\mathbf{1}}(z)=1$ and

$$\operatorname{Li}_w(z)=\begin{cases}\dfrac{\log^s z}{s!} & \text{if } w=x_0^s \text{ for some } s\geqslant1,\\ \displaystyle\int_0^z\omega_j(z)\operatorname{Li}_u(z)\,\mathrm{d}z & \text{if } w=x_ju \text{ contains the letter } x_1.\end{cases}\tag{27}$$

In this case, Lemma 1 remains valid for the extended version (27) of the polylogarithms (this yields the coincidence of the “new” polylogarithms with “old” ones (24) for the words w in $\mathfrak{H}^1$); moreover,

$$\lim_{z\to0+0}\operatorname{Li}_w(z)=0\quad\text{if the word } w \text{ contains the letter } x_1.$$

An easy verification shows that the generalized polylogarithms are continuous real-valued functions on the interval $(0,1)$.

Lemma 2. *The map* $w\mapsto\operatorname{Li}_w(z)$ *is a homomorphism of the algebra* $\mathfrak{H}_{ш}$ *into* $C((0,1);\mathbb{R})$.

Proof. We must verify the equalities

$$\operatorname{Li}_{w_1ш w_2}(z)=\operatorname{Li}_{w_1}(z)\operatorname{Li}_{w_2}(z)\quad\text{for all}\quad w_1,w_2\in\mathfrak{H};\tag{28}$$

it suffices to do this for any *words* $w_1,w_2\in\mathfrak{H}$ only. Let us prove the equality (28) by induction on the quantity $|w_1|+|w_2|$. If $w_1=\mathbf{1}$ or $w_2=\mathbf{1}$, then relation (28) becomes tautological by (15). Otherwise $w_1=x_ju$ and $w_2=x_kv$, and hence by Lemma 1 and by the induction hypothesis we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\big(\operatorname{Li}_{w_1}(z)\operatorname{Li}_{w_2}(z)\big)&=\frac{\mathrm{d}}{\mathrm{d}z}\big(\operatorname{Li}_{x_ju}(z)\operatorname{Li}_{x_kv}(z)\big)\\&=\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_ju}(z)\cdot\operatorname{Li}_{x_kv}(z)+\operatorname{Li}_{x_ju}(z)\cdot\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_kv}(z)\end{aligned}$$

$$
\begin{aligned}
&= \omega_j(z)\,\mathrm{Li}_u(z)\,\mathrm{Li}_{x_k v}(z) + \omega_k(z)\,\mathrm{Li}_{x_j u}(z)\,\mathrm{Li}_v(z) \\
&= \omega_j(z)\,\mathrm{Li}_{u \,\sqcup\!\sqcup\, x_k v}(z) + \omega_k(z)\,\mathrm{Li}_{x_j u \,\sqcup\!\sqcup\, v}(z) \\
&= \frac{\mathrm{d}}{\mathrm{d}z}\big(\mathrm{Li}_{x_j(u \,\sqcup\!\sqcup\, x_k v)}(z) + \mathrm{Li}_{x_k(x_j u \,\sqcup\!\sqcup\, v)}(z)\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_{x_j u \,\sqcup\!\sqcup\, x_k v}(z) \\
&= \frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_{w_1 \,\sqcup\!\sqcup\, w_2}(z).
\end{aligned}
$$

Thus

$$
\mathrm{Li}_{w_1}(z)\,\mathrm{Li}_{w_2}(z) = \mathrm{Li}_{w_1 \,\sqcup\!\sqcup\, w_2}(z) + C, \tag{29}
$$

and the passage to the limit as $z \to 0+0$ gives the relation $C = 0$ if at least one of the words w_1, w_2 contains the letter x_1, otherwise the substitution $z = 1$ gives the same result if the representations of w_1 and w_2 involve the letter x_0 only. Therefore, the equality (29) becomes the required relation (28), and the lemma follows.

Proof of Theorem 4. Theorem 4 follows from Lemma 2 and the relations (23).

An explicit evaluation of the monodromy group for the system (25) of differential equations enabled Minh, Petitot, and van der Hoeven to prove that the homomorphism $w \mapsto \mathrm{Li}_w(z)$ of the shuffle algebra $\mathfrak{H}_{\sqcup\!\sqcup}$ over $\mathbb{C}$ is bijective, that is, all $\mathbb{C}$-algebraic relations for the generalized polylogarithms are induced by the shuffle relations (28) only; in particular, the generalized polylogarithms are linearly independent over $\mathbb{C}$. The assertion that the functions (22) are linearly independent was obtained in the simplest way (as a consequence of elegant identities for the functions) by Ulanskiĭ [25]; the same assertion follows from the Sorokin result in [24].

6. Duality theorem

By Lemma 1, the following integral representation holds for the word $w = x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k} \in \mathfrak{H}^1$:

$$
\begin{aligned}
\mathrm{Li}_w(z) &= \int_0^z \omega_{\varepsilon_1}(z_1)\,\mathrm{d}z_1 \int_0^{z_1} \omega_{\varepsilon_2}(z_2)\,\mathrm{d}z_2 \cdots \int_0^{z_{k-1}} \omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_k \\
&= \int\!\!\cdots\!\!\int\limits_{z>z_1>z_2>\cdots>z_{k-1}>z_k>0} \omega_{\varepsilon_1}(z_1)\omega_{\varepsilon_2}(z_2)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\,\mathrm{d}z_2\cdots\mathrm{d}z_k
\end{aligned} \tag{30}
$$

in the domain $0 < z < 1$. If $x_{\varepsilon_1} \ne x_1$, that is, $w \in \mathfrak{H}^0$, then the integral in (30) converges in the domain $0 < z \leqslant 1$. Thus, in accordance with (23), we obtain a representation [30] for the multiple zeta values in the form of *Chen's iterated integrals*,

$$
\zeta(w) = \int\!\!\cdots\!\!\int\limits_{1>z_1>\cdots>z_k>0} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k. \tag{31}
$$

The following result is a simple consequence of the integral representation (31).

We denote by τ the anti-automorphism of the algebra $\mathfrak{H} = \mathbb{Q}\langle x_0, x_1\rangle$ transposing x_0 and x_1; for example, $\tau(x_0^2x_1x_0x_1) = x_0x_1x_0x_1^2$. Clearly, τ is an involution preserving the weight. It can readily be seen that τ is also an automorphism of the subalgebra $\mathfrak{H}^0$.

Theorem 7 (Duality theorem [30]). *The relation*

$$\zeta(w)=\zeta(\tau w)$$

holds for any word $w\in\mathfrak{H}^0$.

Proof. To prove the theorem, it suffices to make the change of variables $z_1'=1-z_k$, $z_2'=1-z_{k-1}$, …, $z_k'=1-z_1$, and to apply the relations $\omega_0(z)=\omega_1(1-z)$ following from (26).

As the simplest consequence of Theorem 7 we (again) note the identity (5) which corresponds to the word $w=x_0^2x_1$ and also the general identity

$$\zeta(n+2)=\zeta(2,\underbrace{1,\dots,1}_{n\text{ times}}),\qquad n=1,2,\dots, \tag{32}$$

for the words of the form $w=x_0^{n+1}x_1$.

7. Identities: the generating-function method

Another application of the differential equations proved in Lemma 1 for the generalized polylogarithms is the *generating-function method.*

We first note that for an admissible multi-index $\boldsymbol{s}=(s_1,\dots,s_l)$ the corresponding set of *periodic* polylogarithms

$$\operatorname{Li}_{\{\boldsymbol{s}\}_n}(z),\quad\text{where}\quad \{\boldsymbol{s}\}_n=(\underbrace{\boldsymbol{s},\boldsymbol{s},\dots,\boldsymbol{s}}_{n\text{ times}}),\qquad n=0,1,2,\dots$$

(see, for instance, [4] and [28]), possesses the generating function

$$L_{\boldsymbol{s}}(z,t):=\sum_{n=0}^{\infty}\operatorname{Li}_{\{\boldsymbol{s}\}_n}(z)t^{n|\boldsymbol{s}|},$$

which satisfies an ordinary differential equation with respect to the variable z. For instance, if $\ell(\boldsymbol{s})=1$, that is, $\boldsymbol{s}=(s)$, then, by Lemma 1, the corresponding differential equation is of the form

$$\biggl(\Bigl((1-z)\frac{\mathrm{d}}{\mathrm{d}z}\Bigr)\Bigl(z\frac{\mathrm{d}}{\mathrm{d}z}\Bigr)^{s-1}-t^s\biggr)L_{\boldsymbol{s}}(z,t)=0,$$

and its solution can be given explicitly in terms of generalized hypergeometric series (see [3], [4], and [28]).

Lemma 3 ([4], Theorem 12). *The following equality holds*:

$$L_{(3,1)}(z,t)=F\bigl(\tfrac12(1+i)t,-\tfrac12(1+i)t;1;z\bigr)\cdot F\bigl(\tfrac12(1-i)t,-\tfrac12(1-i)t;1;z\bigr), \tag{33}$$

where $F(a,b;c;z)$ *stands for the Gauss hypergeometric function.*

Proof. A routine verification (using Lemma 1 for the left-hand side) shows that both sides of the required equality are annihilated under the action of the differential operator

$$\Bigl((1-z)\frac{\mathrm{d}}{\mathrm{d}z}\Bigr)^2\Bigl(z\frac{\mathrm{d}}{\mathrm{d}z}\Bigr)^2-t^4;$$

moreover, the first terms in the expansions of both sides of (33) in powers of z coincide,

$$1+\frac{t^4}{8}z^2+\frac{t^4}{18}z^3+\frac{t^8+44t^4}{1536}z^4+\cdots.$$

This implies the assertion of the lemma.

Theorem 8 ([4], [28]). *The identity*

$$\zeta(\{3,1\}_n)=\frac{2\pi^{4n}}{(4n+2)!} \tag{34}$$

holds for any integer $n\geqslant 1$.

Proof. By the Gauss summation formula ([29], Ch. 14) we have

$$F(a,-a;1;1)=\frac{1}{\Gamma(1-a)\Gamma(1+a)}=\frac{\sin\pi a}{\pi a}, \tag{35}$$

and, substituting $z=1$ into the equality (33), we obtain

$$\begin{aligned}\sum_{n=0}^{\infty}\zeta(\{3,1\}_n)t^{4n}=L_{(3,1)}(1,t)&=\frac{\sin\frac12(1+i)\pi t}{\frac12(1+i)\pi t}\cdot\frac{\sin\frac12(1-i)\pi t}{\frac12(1-i)\pi t}\\&=\frac{1}{2\pi^2t^2}\cdot\big(e^{(1+i)\pi t/2}-e^{-(1+i)\pi t/2}\big)\big(e^{(1-i)\pi t/2}-e^{-(1-i)\pi t/2}\big)\\&=\frac{1}{2\pi^2t^2}\cdot\big(e^{\pi t}+e^{-\pi t}-e^{i\pi t}-e^{-i\pi t}\big)\\&=\frac{1}{2\pi^2t^2}\sum_{m=0}^{\infty}(1+(-1)^m-i^m-(-i)^m)\frac{(\pi t)^m}{m!}=\sum_{n=0}^{\infty}\frac{2\pi^{4n}t^{4n}}{(4n+2)!}.\end{aligned}$$

Comparing the coefficients at like powers of t gives the desired identity.

The assertion of Theorem 8 was conjectured in [30]. The identity (34) is far from being the only example using the generating-function method. Let us present some other identities of [3] similar to (34) for which the above method is also effective,

$$\begin{gathered}\zeta(\{2\}_n)=\frac{2(2\pi)^{2n}}{(2n+1)!}\left(\frac12\right)^{2n+1},\quad \zeta(\{4\}_n)=\frac{4(2\pi)^{4n}}{(4n+2)!}\left(\frac12\right)^{2n+1},\\ \zeta(\{6\}_n)=\frac{6(2\pi)^{6n}}{(6n+3)!},\quad \zeta(\{8\}_n)=\frac{8(2\pi)^{8n}}{(8n+4)!}\left(\left(1+\frac{1}{\sqrt2}\right)^{4n+2}+\left(1-\frac{1}{\sqrt2}\right)^{4n+2}\right),\\ \zeta(\{10\}_n)=\frac{10(2\pi)^{10n}}{(10n+5)!}\left(1+\left(\frac{1+\sqrt5}{2}\right)^{10n+5}+\left(\frac{1-\sqrt5}{2}\right)^{10n+5}\right),\end{gathered} \tag{36}$$

where $n=1,2,\dots$. The identities

$$\zeta(m+2,\{1\}_n)=\zeta(n+2,\{1\}_m),\qquad m,n=0,1,2,\dots,$$

can be obtained both by the generating-function method [10] and by applying Theorem 7 proved above.

A somewhat different example of generating functions is related to generalizations of Apéry's identity [1]

$$\zeta(3)=\frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}};$$

namely, the following expansions are valid [16], [2]:

$$\begin{aligned}\sum_{n=0}^{\infty}\zeta(2n+3)t^{2n}&=\sum_{k=1}^{\infty}\frac{1}{k^3(1-t^2/k^2)}\\&=\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}\biggl(\frac{1}{2}+\frac{2}{1-t^2/k^2}\biggr)\prod_{l=1}^{k-1}\biggl(1-\frac{t^2}{l^2}\biggr),\\\sum_{n=0}^{\infty}\zeta(4n+3)t^{4n}&=\sum_{k=1}^{\infty}\frac{1}{k^3(1-t^4/k^4)}=\frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}\,\frac{1}{1-t^4/k^4}\prod_{l=1}^{k-1}\frac{1+4t^4/l^4}{1-t^4/l^4}.\end{aligned}\tag{37}$$

The proofs of these identities and of some other ones are based on the use of transformation and summation formulae for *generalized hypergeometric functions* similar to the way in which the formula (35) was used in the proof of Theorem 8. The identities (37) are very useful for the fast evaluation of values of the Riemann zeta function at odd integers.

We also note the relations

$$\tilde\zeta(\{2\}_n,1)=2\zeta(2n+1),\qquad n=1,2,\dots,\tag{38}$$

obtained by successive application of the results in [26] (or [33]) and [32]. The equalities (38) also generalize Euler's identity (5) and are closely related to one of the ways to prove Apéry's theorem [1] and Rivoal's theorem [22] which were mentioned in Section 1. However, it is still not known how to derive the relations (38) from Theorems 4–6 for an arbitrary integer n.

8. Quasi-shuffle products

A construction suggested by Hoffman [12] enables one to view each of the algebras $\mathfrak{H}_{\sqcup\!\sqcup}$ and $\mathfrak{H}^1_*$ as a special case of some general algebraic structure. The present section is devoted to the description of the structure.

We consider a non-commutative polynomial algebra $\mathfrak{A}=\mathcal{K}\langle A\rangle$ graded by the degree over a field $\mathcal{K}\subset\mathbb{C}$; here A stands for a locally finite set of generators (that is, the set of generators of any given positive degree is finite). As usual, we refer to the elements of the set A as letters, and to the monomials in these letters as words. To any word w we assign its length $\ell(w)$ (the number of letters in the representation) and its weight $|w|$ (the sum of degrees of the letters). The unique word of length 0 and weight 0 is the empty word, which is denoted by $\mathbf{1}$; this word is the unit of the algebra $\mathfrak{A}$. The neutral (zero) element of the algebra $\mathfrak{A}$ is denoted by $\mathbf{0}$.

Let us now define the product $\circ$ (by extending it additively to the whole algebra $\mathfrak{A}$) by the following rules:

$$\mathbf{1}\circ w=w\circ\mathbf{1}=w\tag{39}$$

for any word w, and

$$a_j u \circ a_k v = a_j(u \circ a_k v) + a_k(a_j u \circ v) + [a_j, a_k](u \circ v) \tag{40}$$

for any words u, v and letters $a_j, a_k \in A$, where the functional

$$[\,\cdot\,,\,\cdot\,]\colon \bar{A} \times \bar{A} \to \bar{A} \tag{41}$$

($\bar{A} := A \cup \{\mathbf{0}\}$) satisfies the following properties:

(S0) $[a, \mathbf{0}] = \mathbf{0}$ for any $a \in \bar{A}$;

(S1) $[[a_j, a_k], a_l] = [a_j, [a_k, a_l]]$ for any $a_j, a_k, a_l \in \bar{A}$;

(S2) either $[a_j, a_k] = \mathbf{0}$ or $|[a_k, a_j]| = |a_j| + |a_k|$ for any $a_j, a_k \in A$.

Then $\mathfrak{A}_\circ := (\mathfrak{A}, \circ)$ becomes an associative graded $\mathcal{K}$-algebra and, if the additional property

(S3) $[a_j, a_k] = [a_k, a_j]$ for any $a_j, a_k \in \bar{A}$

holds, then $\mathfrak{A}_\circ$ is a commutative $\mathcal{K}$-algebra ([12], Theorem 2.1).

If $[a_j, a_k] = 0$ for all letters $a_j, a_k \in A$, then $(\mathfrak{A}, \circ)$ is the standard shuffle algebra; in the special case $A = \{x_0, x_1\}$ we obtain the shuffle algebra $\mathfrak{A}_\circ = \mathfrak{H}_{\sqcup\!\sqcup}$ of the multiple zeta values (or of the polylogarithms). The stuffle algebra $\mathfrak{H}^1_*$ corresponds to the choice of the generators $A = \{y_j\}_{j=1}^{\infty}$ and the functional

$$[y_j, y_k] = y_{j+k} \quad \text{for any integers } j \geqslant 1 \text{ and } k \geqslant 1.$$

Let us equip the algebra $\mathfrak{A}$ with a given functional (41), with the dual product $\bar{\circ}$ by the rules

$$\mathbf{1} \,\bar{\circ}\, w = w \,\bar{\circ}\, \mathbf{1} = w,$$
$$u a_j \,\bar{\circ}\, v a_k = (u \,\bar{\circ}\, v a_k) a_j + (u a_j \,\bar{\circ}\, v) a_k + (u \,\bar{\circ}\, v)[a_j, a_k]$$

instead of (39) and (40), respectively. Then $\mathfrak{A}_{\bar{\circ}} := (\mathfrak{A}, \bar{\circ})$ is a graded $\mathcal{K}$-algebra as well (which is commutative if the property (S3) holds).

Theorem 9. *The algebras $\mathfrak{A}_\circ$ and $\mathfrak{A}_{\bar{\circ}}$ coincide.*

Proof. It suffices to prove the relation

$$w_1 \circ w_2 = w_1 \,\bar{\circ}\, w_2 \tag{42}$$

for all words $w_1, w_2 \in \mathcal{K}\langle A \rangle$ only. Let us proceed by induction on the quantity $\ell(w_1) + \ell(w_2)$. If $\ell(w_1) = 0$ or $\ell(w_2) = 0$, then the relation (42) becomes an obvious identity. If $\ell(w_1) = \ell(w_2) = 1$, that is, if $w_1 = a_1$ and $w_2 = a_2$ are letters, then

$$a_1 \circ a_2 = a_1 a_2 + a_2 a_1 + [a_1, a_2] = a_1 \,\bar{\circ}\, a_2.$$

If $\ell(w_1) > 1$ and $\ell(w_2) = 1$, then, writing $w_1 = a_1 u a_2$ and $w_2 = a_3 \in A$ and applying the induction hypothesis, we obtain

$$a_1 u a_2 \circ a_3 = a_1(u a_2 \circ a_3) + a_3 a_1 u a_2 + [a_1, a_3] u a_2$$

$$
\begin{aligned}
&= a_1(ua_2 \mathbin{\bar\circ} a_3) + a_3a_1ua_2 + [a_1, a_3]ua_2 \\
&= a_1((u \mathbin{\bar\circ} a_3)a_2 + ua_2a_3 + u[a_2, a_3]) + a_3a_1ua_2 + [a_1, a_3]ua_2 \\
&= a_1((u \circ a_3)a_2 + ua_2a_3 + u[a_2, a_3]) + a_3a_1ua_2 + [a_1, a_3]ua_2 \\
&= (a_1(u \circ a_3) + a_3a_1u + [a_1, a_3]u)a_2 + a_1ua_2a_3 + a_1u[a_2, a_3] \\
&= (a_1u \circ a_3)a_2 + a_1ua_2a_3 + a_1u[a_2, a_3] \\
&= (a_1u \mathbin{\bar\circ} a_3)a_2 + a_1ua_2a_3 + a_1u[a_2, a_3] \\
&= a_1ua_2 \mathbin{\bar\circ} a_3.
\end{aligned}
$$

We can similarly proceed (with more cumbersome manipulations) with the remaining case $\ell(w_1) > 1$ and $\ell(w_2) > 1$. Namely, writing $w_1 = a_1ua_2$ and $w_2 = a_3va_4$ and applying the induction hypothesis, we see that

$$
\begin{aligned}
a_1ua_2 \circ a_3va_4 &= a_1(ua_2 \circ a_3va_4) + a_3(a_1ua_2 \circ va_4) + [a_1, a_3](ua_2 \circ va_4) \\
&= a_1(ua_2 \mathbin{\bar\circ} a_3va_4) + a_3(a_1ua_2 \mathbin{\bar\circ} va_4) + [a_1, a_3](ua_2 \mathbin{\bar\circ} va_4) \\
&= a_1((u \mathbin{\bar\circ} a_3va_4)a_2 + (ua_2 \mathbin{\bar\circ} a_3v)a_4 + (u \mathbin{\bar\circ} a_3v)[a_2, a_4]) \\
&\qquad + a_3((a_1u \mathbin{\bar\circ} va_4)a_2 + (a_1ua_2 \mathbin{\bar\circ} v)a_4 + (a_1u \mathbin{\bar\circ} v)[a_2, a_4]) \\
&\qquad + [a_1, a_3]((u \mathbin{\bar\circ} va_4)a_2 + (ua_2 \mathbin{\bar\circ} v)a_4 + (u \mathbin{\bar\circ} v)[a_2, a_4]) \\
&= a_1((u \circ a_3va_4)a_2 + (ua_2 \circ a_3v)a_4 + (u \circ a_3v)[a_2, a_4]) \\
&\qquad + a_3((a_1u \circ va_4)a_2 + (a_1ua_2 \circ v)a_4 + (a_1u \circ v)[a_2, a_4]) \\
&\qquad + [a_1, a_3]((u \circ va_4)a_2 + (ua_2 \circ v)a_4 + (u \circ v)[a_2, a_4]) \\
&= (a_1(u \circ a_3va_4) + a_3(a_1u \circ va_4) + [a_1, a_3](u \circ va_4))a_2 \\
&\qquad + (a_1(ua_2 \circ a_3v) + a_3(a_1ua_2 \circ v) + [a_1, a_3](ua_2 \circ v))a_4 \\
&\qquad + (a_1(u \circ a_3v) + a_3(a_1u \circ v) + [a_1, a_3](u \circ v))[a_2, a_4] \\
&= (a_1u \circ a_3va_4)a_2 + (a_1ua_2 \circ a_3v)a_4 + (a_1u \circ a_3v)[a_2, a_4] \\
&= (a_1u \mathbin{\bar\circ} a_3va_4)a_2 + (a_1ua_2 \mathbin{\bar\circ} a_3v)a_4 + (a_1u \mathbin{\bar\circ} a_3v)[a_2, a_4] \\
&= a_1ua_2 \mathbin{\bar\circ} a_3va_4.
\end{aligned}
$$

This completes the proof of the theorem.

Remark. If the property (S3) holds, then the above proof can be simplified significantly. However, in our opinion, it is of importance that the algebras $\mathfrak{A}_\circ$ and $\mathfrak{A}_{\bar\circ}$ coincide in the most general situation, that is, if the functional (41) satisfies the conditions (S0)–(S2).

In conclusion of the section, we prove an auxiliary statement.

Lemma 4. *The following identity holds for each letter $a \in A$ and any words $u, v \in \mathfrak{A}$:*

$$
a \circ uv - (a \circ u)v = u(a \circ v - av). \tag{43}
$$

Proof. Let us prove the statement by induction on the number of letters in the word u. If the word u is empty, then the identity (43) is evident. Otherwise, let us write the word u in the form $u = a_1u_1$, where $a_1 \in A$ and the word u_1 consists of lesser number of letters, and hence satisfies the identity

$$
a \circ u_1v - (a \circ u_1)v = u_1(a \circ v - av).
$$

Then

$$\begin{aligned} a\circ uv-(a\circ u)v &= a\circ a_1u_1v-(a\circ a_1u_1)v \\ &= aa_1u_1v+a_1(a\circ u_1v)+[a,a_1]u_1v \\ &\qquad -(aa_1u_1+a_1(a\circ u_1)+[a,a_1]u_1)v \\ &= a_1(a\circ u_1v-(a\circ u_1)v)=a_1u_1(a\circ v-av) \\ &= u(a\circ v-av), \end{aligned}$$

as was to be proved.

9. Functional model of the stuffle algebra

A functional model of the stuffle algebra $\mathfrak{H}_*$ cannot be described in perfect analogy with the polylogarithmic model of the shuffle algebra $\mathfrak{H}_{⧢}$ because the rule (17) has no differential interpretation in contrast to (16). Therefore we use a difference interpretation of the rule (17), namely, consider the (simplest) difference operator

$$Df(t)=f(t-1)-f(t).$$

It can readily be seen that

$$D\bigl(f_1(t)f_2(t)\bigr)=Df_1(t)\cdot f_2(t)+f_1(t)\cdot Df_2(t)+Df_1(t)\cdot Df_2(t) \tag{44}$$

and that inverse mapping

$$Ig(t)=\sum_{n=1}^{\infty}g(t+n)$$

(such that $D(Ig(t))=g(t)$) is defined up to an additive constant provided that some additional restrictions are imposed on the function $g(t)$ as $t\to+\infty$, for instance, $g(t)=O(t^{-2})$.

Remark. By [6], § 3.1, the operator D is related to the differential operator $\mathrm{d}/\mathrm{d}t$ as follows:

$$D=e^{-\mathrm{d}/\mathrm{d}t}-1=\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\,\frac{\mathrm{d}^n}{\mathrm{d}t^n}\,.$$

The above equality is justified by the formal application of the Taylor expansion,

$$f(t-1)=f(t)+\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\,\frac{\mathrm{d}^n}{\mathrm{d}t^n}f(t);$$

in fact, the last formula is valid for any entire function. The exponentiation of derivations (on algebras of words) is discussed in Section 12 below in connection with a generalization of Theorem 1.

According to (17) and (44), the natural analogy with Lemmas 1 and 2 assumes the existence of functions $\omega_j(t)$ such that

$$\omega_j(t)\omega_k(t)=\omega_{j+k}(t)\quad\text{for the integers } j\geqslant 1 \text{ and } k\geqslant 1.$$

The simplest choice is given by the formulae

$$\omega_j(t)=\frac{1}{t^j}, \qquad j=1,2,\dots,$$

and this leads to the functions

$$\operatorname{Ri}_{\boldsymbol s}(t)=\operatorname{Ri}_{s_1,\dots,s_{l-1},s_l}(t):=I\biggl(\frac{1}{t^{s_l}}\operatorname{Ri}_{s_1,\dots,s_{l-1}}(t)\biggr), \qquad \operatorname{Ri}_{\mathbf 1}(t):=1,$$

defined by induction on the length of the multi-index. By the definition, we have

$$D\operatorname{Ri}_{uy_j}(t)=\frac{1}{t^j}\operatorname{Ri}_u(t), \tag{45}$$

which is in a sense a discrete analogue of the formula (25).

Lemma 5. *The following identity holds*:

$$\operatorname{Ri}_{\boldsymbol s}(t)=\sum_{n_1>\dots>n_{l-1}>n_l\geqslant1}\frac{1}{(t+n_1)^{s_1}\cdots(t+n_{l-1})^{s_{l-1}}(t+n_l)^{s_l}}\,; \tag{46}$$

in particular,

$$\operatorname{Ri}_{\boldsymbol s}(0)=\zeta(\boldsymbol s), \qquad \boldsymbol s\in\mathbb Z^l, \quad s_1\geqslant2,\ s_2\geqslant1,\ \dots,\ s_l\geqslant1, \tag{47}$$

$$\lim_{t\to+\infty}\operatorname{Ri}_{\boldsymbol s}(t)=0, \qquad \boldsymbol s\in\mathbb Z^l, \quad s_1\geqslant2,\ s_2\geqslant1,\ \dots,\ s_l\geqslant1. \tag{48}$$

Proof. By definition,

$$\begin{aligned}
\operatorname{Ri}_{\boldsymbol s}(t)&=I\biggl(\frac{1}{t^{s_l}}\operatorname{Ri}_{s_1,\dots,s_{l-1}}(t)\biggr)\\
&=I\biggl(\frac{1}{t^{s_l}}\sum_{n_1>\dots>n_{l-1}\geqslant1}\frac{1}{(t+n_1)^{s_1}\cdots(t+n_{l-1})^{s_{l-1}}}\biggr)\\
&=\sum_{n=1}^{\infty}\frac{1}{(t+n)^{s_l}}\sum_{n_1>\dots>n_{l-1}\geqslant1}\frac{1}{(t+n_1+n)^{s_1}\cdots(t+n_{l-1}+n)^{s_{l-1}}}\\
&=\sum_{n_1'>\dots>n_{l-1}'>n\geqslant1}\frac{1}{(t+n_1')^{s_1}\cdots(t+n_{l-1}')^{s_{l-1}}(t+n)^{s_l}}\,,
\end{aligned}$$

and this implies the required formula (46).

Let us now define the multiplication $\mathbin{\bar*}$ on the algebra $\mathfrak H^1$ (and, in particular, on the subalgebra $\mathfrak H^0$) by the rules

$$\begin{gathered}
\mathbf 1\mathbin{\bar*}w=w\mathbin{\bar*}\mathbf 1=w,\\
uy_j\mathbin{\bar*}vy_k=(u\mathbin{\bar*}vy_k)y_j+(uy_j\mathbin{\bar*}v)y_k+(u\mathbin{\bar*}v)y_{j+k},
\end{gathered} \tag{49}$$

instead of (15) and (17).

Lemma 6. *The map* $w \mapsto \mathrm{Ri}_w(z)$ *is a homomorphism of the algebra* $(\mathfrak{H}^0, \bar{*})$ *into* $C([0,+\infty);\mathbb{R})$.

Proof. We must verify the relation

$$\mathrm{Ri}_{w_1 \bar{*} w_2}(z) = \mathrm{Ri}_{w_1}(z)\,\mathrm{Ri}_{w_2}(z) \quad \text{for all} \quad w_1, w_2 \in \mathfrak{H}^0. \tag{50}$$

We assume without loss of generality that w_1 and w_2 are words of the algebra $\mathfrak{H}^0$. Let us prove the relation (50) by induction on the quantity $\ell(w_1)+\ell(w_2)$; if $w_1 = \mathbf{1}$ or $w_2 = \mathbf{1}$, then the validity of (50) is evident by (49). Otherwise we write $w_1 = uy_j$, $w_2 = vy_k$ and apply the formulae (44) and (45) and the induction hypothesis,

$$\begin{aligned}
D\big(\mathrm{Ri}_{w_1}(t)\,\mathrm{Ri}_{w_2}(t)\big) &= D\big(\mathrm{Ri}_{uy_j}(t)\,\mathrm{Ri}_{vy_k}(t)\big) \\
&= D\,\mathrm{Ri}_{uy_j}(t)\cdot \mathrm{Ri}_{vy_k}(t) + \mathrm{Ri}_{uy_j}(t)\cdot D\,\mathrm{Ri}_{vy_k}(t) \\
&\qquad + D\,\mathrm{Ri}_{uy_j}(t)\cdot D\,\mathrm{Ri}_{vy_k}(t) \\
&= \frac{1}{t^j}\,\mathrm{Ri}_u(t)\,\mathrm{Ri}_{vy_k}(t) + \frac{1}{t^k}\,\mathrm{Ri}_{uy_j}(t)\,\mathrm{Ri}_v(t) + \frac{1}{t^{j+k}}\,\mathrm{Ri}_u(t)\,\mathrm{Ri}_v(t) \\
&= \frac{1}{t^j}\,\mathrm{Ri}_{u \bar{*} vy_k}(t) + \frac{1}{t^k}\,\mathrm{Ri}_{uy_j \bar{*} v}(t) + \frac{1}{t^{j+k}}\,\mathrm{Ri}_{u \bar{*} v}(t) \\
&= D\big(\mathrm{Ri}_{(u \bar{*} vy_k)y_j}(t) + \mathrm{Ri}_{(uy_j \bar{*} v)y_k}(t) + \mathrm{Ri}_{(u \bar{*} v)y_{j+k}}(t)\big) \\
&= D\,\mathrm{Ri}_{uy_j \bar{*} vy_k}(t) \\
&= D\,\mathrm{Ri}_{w_1 \bar{*} w_2}(t).
\end{aligned}$$

Therefore

$$\mathrm{Ri}_{w_1}(t)\,\mathrm{Ri}_{w_2}(t) = \mathrm{Ri}_{w_1 \bar{*} w_2}(t) + C, \tag{51}$$

and, passing to the limit as t tends to $+\infty$, we obtain $C = 0$ by (48). Thus, relation (51) becomes the required equality (50), and the lemma follows.

Proof of Theorem 5. By (47), Theorem 5 follows from Lemma 6 and Theorem 9.

10. Hoffman's homomorphism for the stuffle algebra

Another way to prove Theorem 5 (and Lemma 6 as well) uses Hoffman's homomorphism $\phi\colon \mathfrak{H}^1 \to \mathbb{Q}[[t_1, t_2, \dots]]$, where $\mathbb{Q}[[t_1, t_2, \dots]]$ is the $\mathbb{Q}$-algebra of formal power series in countably many (commuting) variables $t_1, t_2, \dots$ (see [11] and [13]). Namely, the $\mathbb{Q}$-linear map ϕ is defined by setting $\phi(1) := 1$ and

$$\phi(y_{s_1} y_{s_2} \cdots y_{s_l}) := \sum_{n_1 > n_2 > \dots > n_l \geqslant 1} t_{n_1}^{s_1} t_{n_2}^{s_2} \cdots t_{n_l}^{s_l}, \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \geqslant 1, \ \dots, \ s_l \geqslant 1.$$

The image of the homomorphism ϕ (which is in fact a monomorphism) is the algebra *QSym* of quasi-symmetric functions. Here by a *quasi-symmetric function* we mean a formal power series (of bounded degree) in $t_1, t_2, \dots$ in which the coefficients at $t_{n_1}^{s_1} t_{n_2}^{s_2} \cdots t_{n_l}^{s_l}$ and $t_{n_1'}^{s_1} t_{n_2'}^{s_2} \cdots t_{n_l'}^{s_l}$ coincide whenever $n_1 > n_2 > \dots > n_l$ and $n_1' > n_2' > \dots > n_l'$ (our definition slightly differs from the corresponding definition in [13] but leads to the same algebra *QSym* of quasi-symmetric functions). In these terms, the homomorphism $w \mapsto \mathrm{Ri}_w(t)$ in Lemma 6 is defined as the restriction of the homomorphism ϕ to $\mathfrak{H}^0$ given by the substitution $t_n = 1/(t+n)$, $n = 1, 2, \dots$.

Another approach to the proof of the stuffle relations for multiple zeta values was recently suggested by Cartier (see [28]). Slightly modifying the original scheme of Cartier, we show the main ideas of the approach by the example of proving Euler's identity

$$\zeta(s_1)\zeta(s_2) = \zeta(s_1+s_2) + \zeta(s_1,s_2) + \zeta(s_2,s_1), \qquad s_1 \geqslant 2,\ s_2 \geqslant 2. \tag{52}$$

To this end, we need another integral representation (as compared with (31)) for the admissible multi-indices $\boldsymbol{s}$,

$$\zeta(\boldsymbol{s}) = \underset{[0,1]^{|\boldsymbol{s}|}}{\int \cdots \int} \prod_{j=1}^{l-1} \frac{t_1 t_2 \cdots t_{s_1+\dots+s_j}}{1 - t_1 t_2 \cdots t_{s_1+\dots+s_j}} \cdot \frac{\mathrm{d}t_1\, \mathrm{d}t_2 \cdots \mathrm{d}t_{|\boldsymbol{s}|}}{1 - t_1 t_2 \cdots t_{s_1+s_2+\dots+s_l}}, \qquad l = \ell(\boldsymbol{s}), \tag{53}$$

which was kindly pointed out to us by Nesterenko and can be proved by straightforwardly integrating the series

$$\frac{1}{1-t} = \sum_{n=0}^{\infty} t^n.$$

Substituting $u = t_1 \cdots t_{s_1}$, $v = t_{s_1+1} \cdots t_{s_2}$ into the elementary identity

$$\frac{1}{(1-u)(1-v)} = \frac{1}{1-uv} + \frac{u}{(1-u)(1-uv)} + \frac{v}{(1-v)(1-uv)}$$

and integrating over the hypercube $[0,1]^{s_1+s_2}$ in accordance with (53), we arrive at the identity (52).

11. Derivations

As in Section 8, let us consider a graded non-commutative polynomial algebra $\mathfrak{A} = \mathcal{K}\langle A \rangle$ over a field $\mathcal{K}$ of characteristic 0 with a locally finite set of generators A. By a *derivation* of the algebra $\mathfrak{A}$ we mean a linear map $\delta\colon \mathfrak{A} \to \mathfrak{A}$ (of the graded $\mathcal{K}$-vector spaces) satisfying the Leibniz rule

$$\delta(uv) = \delta(u)v + u\delta(v) \quad \text{for all} \quad u, v \in \mathfrak{A}. \tag{54}$$

The commutator of two derivations $[\delta_1, \delta_2] := \delta_1\delta_2 - \delta_2\delta_1$ is a derivation, and thus the set of all derivations of the algebra $\mathfrak{A}$ forms a Lie algebra $\mathit{Der}(\mathfrak{A})$ (naturally graded by the degree).

It can readily be seen that it suffices to define a derivation $\delta \in \mathit{Der}(\mathfrak{A})$ on the generators of A only and then to extend it to the whole algebra by linearity and by using the rule (54).

The next assertion gives examples of derivations of the algebra $\mathfrak{A}$ equipped with an additional multiplication $\circ$ having the properties (39) and (40).

Theorem 10. *The map*

$$\delta_a\colon w\mapsto aw-a\circ w \tag{55}$$

is a derivation for any letter $a\in A$.

Proof. It is clear that the map δ_a is linear. By Lemma 4, for any words $u,v\in\mathfrak{A}$ we have

$$\begin{aligned}\delta_a(uv)&=auv-a\circ uv=auv-(a\circ u)v-u(a\circ v-av)\\&=(\delta_a u)v+u(\delta_a v),\end{aligned}$$

and thus (55) is really a derivation.

By Theorem 10, the maps $\delta_{ш}\colon\mathfrak{H}\to\mathfrak{H}$ and $\delta_*\colon\mathfrak{H}^1\to\mathfrak{H}^1$ defined by the formulae

$$\delta_{ш}\colon w\mapsto x_1w-x_1ш w,\qquad \delta_*\colon w\mapsto y_1w-y_1*w=x_1w-x_1*w, \tag{56}$$

are derivations; according to the rule (18), the map δ_* is a derivation on the whole algebra $\mathfrak{H}$. We note that the derivations (56) act on the generators of the algebra according to the rules (15)–(18) as follows:

$$\delta_{ш}x_0=-x_0x_1,\quad \delta_{ш}x_1=-x_1^2,\qquad \delta_*x_0=0,\quad \delta_*x_1=-x_1^2-x_0x_1. \tag{57}$$

For any derivation δ of the algebra $\mathfrak{H}$ (or of the subalgebra $\mathfrak{H}^0$) we define the dual derivation $\overline{\delta}=\tau\delta\tau$, where τ is the anti-automorphism of the algebra $\mathfrak{H}$ (and $\mathfrak{H}^0$) introduced in Section 6. A derivation δ is said to be *symmetric* if $\overline{\delta}=\delta$ and *anti-symmetric* if $\overline{\delta}=-\delta$. Since $\tau x_0=x_1$, an (anti-)symmetric derivation δ is uniquely determined by the image of one of the generators x_0 or x_1, whereas an arbitrary derivation is reconstructed from the images of both generators only.

We now define the derivation D of the algebra $\mathfrak{H}$ by setting $Dx_0=0$ and $Dx_1=x_0x_1$ (that is, $Dy_s=y_{s+1}$ for the generators y_s of the algebra $\mathfrak{H}^1$) and represent the statement of Theorem 1 in the following form.

Theorem 11 (Derivation theorem, [13], Theorem 2.1). *The identity*

$$\zeta(Dw)=\zeta(\overline{D}w) \tag{58}$$

holds for any word $w\in\mathfrak{H}^0$.

Proof. Expressing any word $w\in\mathfrak{H}^0$ in the form $w=y_{s_1}y_{s_2}\cdots y_{s_l}$ (with $s_1>1$), we note that the left-hand side of the equality (7) corresponds to the element

$$Dw=D(y_{s_1}y_{s_2}\cdots y_{s_l})=y_{s_1+1}y_{s_2}\cdots y_{s_l}+y_{s_1}y_{s_2+1}y_{s_3}\cdots y_{s_l}+\cdots+y_{s_1}\cdots y_{s_{l-1}}y_{s_l+1} \tag{59}$$

of the algebra $\mathfrak{H}^0$. On the other hand,

$$\begin{aligned}\overline{D}w&=\tau D\bigl(x_0x_1^{s_l-1}x_0x_1^{s_{l-1}-1}\cdots x_0x_1^{s_2-1}x_0x_1^{s_1-1}\bigr)\\&=\tau\sum_{\substack{k=1\\s_k\geqslant2}}^{l}\sum_{j=0}^{s_k-2}x_0x_1^{s_l-1}\cdots x_0x_1^{s_{k+1}-1}x_0x_1^{j}x_0x_1^{s_k-j-1}x_0x_1^{s_{k-1}-1}\cdots x_0x_1^{s_1-1}\\&=\sum_{\substack{k=1\\s_k\geqslant2}}^{l}\sum_{j=0}^{s_k-2}x_0^{s_1-1}x_1\cdots x_0^{s_{k-1}-1}x_1x_0^{s_k-j-1}x_1x_0^{j}x_1x_0^{s_{k+1}-1}x_1\cdots x_0^{s_l-1}x_1,\end{aligned} \tag{60}$$

which corresponds to the right-hand side of (7). Applying the map ζ to the resulting equalities (59) and (60), we obtain the required identity (58).

Remark. The condition $w \in \mathfrak{H}^0$ in Theorem 11 cannot be weakened. The equality (58) fails for the word $w = x_1$,

$$\zeta(Dx_1) = \zeta(x_0x_1) \neq 0 = \zeta(\overline{D}x_1).$$

Proof of Theorem 6. Comparing the action (57) of the derivations (56) with the action of D and $\overline{D}$ on the generators of the algebra $\mathfrak{H}$,

$$Dx_0 = 0, \quad Dx_1 = x_0x_1, \qquad \overline{D}x_0 = x_0x_1, \quad \overline{D}x_1 = 0,$$

we see that $\delta_* - \delta_{ш} = \overline{D} - D$. Therefore, the application of Theorem 11 to the word $w \in \mathfrak{H}^0$ leads to the required equality,

$$\zeta(x_1 ш w - x_1 * w) = \zeta\big((\delta_* - \delta_{ш})w\big) = \zeta\big((\overline{D} - D)w\big) = \zeta(\overline{D}w) - \zeta(Dw) = 0.$$

This completes the proof.

Remark. Another proof of Theorem 6, based on the shuffle and stuffle relations for so-called *coloured* polylogarithms

$$\mathrm{Li}_{\boldsymbol{s}}(\boldsymbol{z}) = \mathrm{Li}_{(s_1,s_2,\dots,s_l)}(z_1,z_2,\dots,z_l) := \sum_{n_1>n_2>\dots>n_l\geqslant 1} \frac{z_1^{n_1}z_2^{n_2}\cdots z_l^{n_l}}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\,, \tag{61}$$

can be found in [28]. (It is clear that the specialization $z_2 = \dots = z_l = 1$ transforms the functions (61) to the generalized polylogarithms (22).) We do not intend to present the properties of the functional model (61) in this survey and refer the interested reader to [4], [7], and [28].

12. Ihara–Kaneko derivations and Ohno's relations

Theorem 11 has a natural generalization. For any $n \geqslant 1$ we define an antisymmetric derivation $\partial_n \in \mathit{Der}(\mathfrak{H})$ by the rule $\partial_n x_0 = x_0(x_0+x_1)^{n-1}x_1$. As was mentioned in the proof of Theorem 6, we have $\partial_1 = \overline{D} - D = \delta_* - \delta_{ш}$. The following assertion holds.

Theorem 12 [14] (see also [13]). *The identity*

$$\zeta(\partial_n w) = 0 \tag{62}$$

holds for any $n \geqslant 1$ and any word $w \in \mathfrak{H}^0$.

Below we sketch the proof of Theorem 12 presented in the preprint [14] (the proof in [13] uses other ideas).

The following result, which was proved in [21] by the generating-function method, contains Theorems 1, 3, and 7 as special cases (the corresponding implications are also given in [21]).

Theorem 13 (Ohno's relations). *Let a word $w \in \mathfrak{H}^0$ and its dual $w' = \tau w \in \mathfrak{H}^0$ have the following representations in terms of the generators of the algebra $\mathfrak{H}^1$:*

$$w = y_{s_1} y_{s_2} \cdots y_{s_l}, \qquad w' = y_{s'_1} y_{s'_2} \cdots y_{s'_k}.$$

Then the identity

$$\sum_{\substack{e_1,e_2,\dots,e_l \geqslant 0 \\ e_1+e_2+\dots+e_l=n}} \zeta(y_{s_1+e_1} y_{s_2+e_2} \cdots y_{s_l+e_l}) = \sum_{\substack{e_1,e_2,\dots,e_k \geqslant 0 \\ e_1+e_2+\dots+e_k=n}} \zeta(y_{s'_1+e_1} y_{s'_2+e_2} \cdots y_{s'_k+e_k})$$

holds for any integer $n \geqslant 0$.

Following [14], for each integer $n \geqslant 1$ we define the derivation $D_n \in Der(\mathfrak{H})$ by setting $D_n x_0 = 0$ and $D_n x_1 = x_0^n x_1$. One can readily see that the derivations $D_1, D_2, \dots$ commute; this fact holds for the dual derivations $\overline{D}_1, \overline{D}_2, \dots$ as well. Let us consider the completion of $\mathfrak{H}$, namely, the algebra $\widehat{\mathfrak{H}} = \mathbb{Q}\langle\langle x_0, x_1 \rangle\rangle$ of the formal power series in non-commutative variables x_0, x_1 over the field $\mathbb{Q}$. The action of the anti-automorphism τ and of the derivations $\delta \in Der(\mathfrak{H})$ can naturally be extended to the whole algebra $\widehat{\mathfrak{H}}$. For simplicity, let us write $w \in \ker\zeta$ if all homogeneous components of the element $w \in \widehat{\mathfrak{H}}$ belong to $\ker\zeta$. The maps

$$\mathcal{D} = \sum_{n=1}^{\infty} \frac{D_n}{n}, \qquad \overline{\mathcal{D}} = \sum_{n=1}^{\infty} \frac{\overline{D}_n}{n}$$

are derivations of the algebra $\widehat{\mathfrak{H}}$, and it follows from the standard relation between derivations and homomorphisms that the maps

$$\sigma = \exp(\mathcal{D}), \qquad \overline{\sigma} = \tau\sigma\tau = \exp(\overline{\mathcal{D}})$$

are automorphisms of the algebra $\widehat{\mathfrak{H}}$. In these terms, Ohno's relations can be stated as follows.

Theorem 14 [14]**.** *The inclusion*

$$(\sigma - \overline{\sigma})w \in \ker\zeta \tag{63}$$

holds for any word $w \in \mathfrak{H}^0$.

Proof. Since $\mathcal{D}x_0 = 0$ and

$$\mathcal{D}x_1 = \left(x_0 + \frac{x_0^2}{2} + \frac{x_0^3}{3} + \cdots\right)x_1 = (-\log(1-x_0))x_1,$$

it follows that $\mathcal{D}^n x_0 = 0$ and $\mathcal{D}^n x_1 = (-\log(1-x_0))^n x_1$, and hence $\sigma x_0 = x_0$ and

$$\sigma x_1 = \sum_{n=0}^{\infty} \frac{1}{n!}(-\log(1-x_0))^n x_1 = (1-x_0)^{-1}x_1 = (1 + x_0 + x_0^2 + x_0^3 + \cdots)x_1.$$

Therefore, for the word $w=y_{s_1}y_{s_2}\cdots y_{s_l}\in\mathfrak{H}^0$ we have

$$\begin{aligned}\sigma w&=\sigma(x_0^{s_1-1}x_1x_0^{s_2-1}x_1\cdots x_0^{s_l-1}x_1)\\&=x_0^{s_1-1}(1+x_0+x_0^2+\cdots)x_1x_0^{s_2-1}(1+x_0+x_0^2+\cdots)x_1\cdots\\&\qquad\cdots x_0^{s_l-1}(1+x_0+x_0^2+\cdots)x_1\\&=\sum_{n=0}^{\infty}\sum_{\substack{e_1,e_2,\dots,e_l\geqslant0\\e_1+e_2+\dots+e_l=n}}x_0^{s_1-1+e_1}x_1x_0^{s_2-1+e_2}x_1\cdots x_0^{s_l-1+e_l}x_1;\end{aligned}$$

thus, $\sigma w-\sigma\tau w\in\ker\zeta$ by Theorem 13. Applying now Theorem 7, we obtain the desired inclusion (63).

Let us return to the derivations $\partial_1,\partial_2,\dots$ and consider the derivation

$$\partial=\sum_{n=1}^{\infty}\frac{\partial_n}{n}\in Der(\widehat{\mathfrak{H}}).$$

Lemma 7. *The following equality holds*:

$$\exp(\partial)=\overline{\sigma}\cdot\sigma^{-1}. \tag{64}$$

Proof. We first note that the operators ∂_n, $n=1,2,\dots$, commute. Really, since $\partial_n(x_0+x_1)=0$ for any $n\geqslant1$, it suffices to prove the equality $\partial_n\partial_mx_0=\partial_m\partial_nx_0$ for $n,m\geqslant1$. Since $\partial_n(x_0+x_1)^k=0$ for any $n\geqslant1$ and $k\geqslant0$, we see that

$$\begin{aligned}\partial_n\partial_mx_0&=\partial_n(x_0(x_0+x_1)^{m-1}x_1)\\&=x_0(x_0+x_1)^{n-1}x_1(x_0+x_1)^{m-1}x_1-x_0(x_0+x_1)^{m-1}x_0(x_0+x_1)^{n-1}x_1\\&=x_0(x_0+x_1)^{n-1}(x_0+x_1-x_0)(x_0+x_1)^{m-1}x_1\\&\qquad-x_0(x_0+x_1)^{m-1}(x_0+x_1-x_1)(x_0+x_1)^{n-1}x_1\\&=-x_0(x_0+x_1)^{n-1}x_0(x_0+x_1)^{m-1}x_1+x_0(x_0+x_1)^{m-1}x_1(x_0+x_1)^{n-1}x_1\\&=\partial_m\partial_nx_0,\end{aligned}$$

as was to be proved.

Let us consider the family $\varphi(t)$, $t\in\mathbb{R}$, of automorphisms of the algebra $\widehat{\mathfrak{H}}_{\mathbb{R}}=\mathbb{R}\langle\langle x_0,x_1\rangle\rangle$ that are defined on the generators $x_0'=x_0+x_1$ and x_1 by the rules

$$\varphi(t)\colon x_0'\mapsto x_0',\quad \varphi(t)\colon x_1\mapsto(1-x_0')^tx_1\biggl(1-\frac{1-(1-x_0')^t}{x_0'}x_1\biggr)^{-1},\qquad t\in\mathbb{R}.$$

The routine verification [14] shows that

$$\varphi(t_1)\varphi(t_2)=\varphi(t_1+t_2),\quad\varphi(0)=\mathrm{id},\quad\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t)\Big|_{t=0}=\partial,\quad\varphi(1)=\overline{\sigma}\cdot\sigma^{-1};$$

hence $\varphi(t)=\exp(t\partial)$, and the substitution $t=1$ leads to the required result (64).

Proof of Theorem 12. Let us show how Theorem 12 follows from Theorem 14 and Lemma 7. On the one hand, we have

$$\partial = \log(\overline{\sigma} \cdot \sigma^{-1}) = \log(1 - (\sigma - \overline{\sigma})\sigma^{-1}) = -(\sigma - \overline{\sigma}) \sum_{n=1}^{\infty} \frac{((\sigma - \overline{\sigma})\sigma^{-1})^{n-1}}{n} \sigma^{-1},$$

and on the other hand

$$\sigma - \overline{\sigma} = (1 - \overline{\sigma} \cdot \sigma^{-1})\sigma = (1 - \exp(\partial))\sigma = -\partial \sum_{n=1}^{\infty} \frac{\partial^{n-1}}{n!} \sigma;$$

hence, $\partial\mathfrak{H}^0 = (\sigma - \overline{\sigma})\mathfrak{H}^0$, and Theorem 14 yields the required identities (62).

Does there exist a simpler way to prove relations (62)? The explicit computations in [14] show that $\partial_1 = \delta_* - \delta_{\sqcup\sqcup}$,

$$\partial_2 = [\delta_*, \overline{\delta}_*],$$
$$\partial_3 = \frac{1}{2}[\delta_*, [\partial_1, \overline{\delta}_*]] - \frac{1}{2}[\delta_*, \partial_2] - \frac{1}{2}[\overline{\delta}_*, \partial_2],$$
$$\partial_4 = \frac{1}{6}[\delta_*, [\partial_1, [\partial_1, \overline{\delta}_*]]] - \frac{1}{6}[\overline{\delta}_*, [\delta_*, [\partial_1, \overline{\delta}_*]]] + \frac{1}{6}[\partial_1, [\partial_2, \overline{\delta}_*]] + \frac{1}{3}[\partial_3, \delta_*] + \frac{1}{3}[\partial_3, \overline{\delta}_*],$$

and, moreover, $\delta_* + \overline{\delta}_* = \delta_{\sqcup\sqcup} + \overline{\delta}_{\sqcup\sqcup}$; therefore, the cases $n = 1, 2, 3, 4$ in Theorem 12 can be processed by induction (with Theorem 11 as the base of induction). This motivates the following conjecture.

Conjecture 3 [14]**.** *For any $n \geqslant 1$ the above anti-symmetric derivation ∂_n is contained in the Lie subalgebra of* Der$(\mathfrak{H})$ *generated by the derivations δ_*, $\overline{\delta}_*$, $\delta_{\sqcup\sqcup}$, and $\overline{\delta}_{\sqcup\sqcup}$.*

We also note that the preprint [14] contains some other ideas (as compared with Conjecture 2) concerning the complete description of the identities for multiple zeta values in terms of regularized shuffle-stuffle relations.

13. Open questions

Along with the above Conjectures 1–3, let us also mention some other important conjectures concerning the structure of the subspace $\ker\zeta \subset \mathfrak{H}$. We denote by $\mathcal{Z}_k$ the $\mathbb{Q}$-vector subspace in $\mathbb{R}$ spanned by the multiple zeta values of weight k and set $\mathcal{Z}_0 = \mathbb{Q}$ and $\mathcal{Z}_1 = \{0\}$. Then the $\mathbb{Q}$-subspace $\mathcal{Z} \in \mathbb{R}$ spanned by all multiple zeta values is a subalgebra of $\mathbb{R}$ over $\mathbb{Q}$ graded by the weight.

Conjecture 4 ([8], [28]). *When regarded as a $\mathbb{Q}$-algebra, the algebra $\mathcal{Z}$ is the direct sum of the subspaces $\mathcal{Z}_k$, $k = 0, 1, 2, \dots$.*

We can readily see that the relations (19)–(21) for multiple zeta values are homogeneous with respect to the weight, and hence Conjecture 4 follows from Conjecture 2.

Let d_k be the dimension of the $\mathbb{Q}$-space $\mathcal{Z}_k$, $k = 0, 1, 2, \dots$. We note that $d_0 = 1$, $d_1 = 0$, $d_2 = 1$ (since $\zeta(2) \neq 0$), $d_3 = 1$ (since $\zeta(3) = \zeta(2,1) \neq 0$), and $d_4 = 1$ (since $\mathcal{Z}_4 = \mathbb{Q}\pi^4$ by (32), (34), and (36)). For $k \geqslant 5$ the above identities enable one to write out the upper bounds; for instance, $d_5 \leqslant 2$, $d_6 \leqslant 2$, and so on.

Conjecture 5 [30]**.** *For $k \geqslant 3$ we have the recurrence relations*

$$d_k = d_{k-2} + d_{k-3};$$

in other words,

$$\sum_{k=0}^{\infty} d_k t^k = \frac{1}{1 - t^2 - t^3}\,.$$

Even if the answer to Conjectures 4 and 5 is positive, the question of choosing a transcendence basis of the algebra $\mathcal{Z}$ and (or) a rational basis of the $\mathbb{Q}$-spaces $\mathcal{Z}_k$, $k = 0, 1, 2, \dots$, would be still open. In this connection, the following Hoffman's conjecture is of interest.

Conjecture 6 [11]**.** *For any $k = 0, 1, 2, \dots$ the number set*

$$\bigl\{\zeta(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2, 3\},\ j = 1, \dots, \ell(\boldsymbol{s})\bigr\} \tag{65}$$

is a basis of the $\mathbb{Q}$-space $\mathcal{Z}_k$.

Not only the experimental confirmation for $k \leqslant 16$ (under the assumption that Conjecture 2 is true) but also the coincidence of the dimension of the $\mathbb{Q}$-space spanned by the numbers (65) with the dimension d_k of the space $\mathcal{Z}_k$ in Conjecture 5 (the last fact was proved by Hoffman in [11]) is an argument in favour of Conjecture 6.

14. q-Analogues of multiple zeta values

Thirty three years after the Gauss work on hypergeometric series, Heine [9] considered series depending on an additional parameter q and possessing properties similar to those of the Gauss series. Moreover, as q tends to 1 (at least term-wise), the Heine q-series become hypergeometric series, and thus the Gauss results can be obtained from the corresponding results for q-series by this passage to the limit and the theorem on analytic continuation.

Similar q-extensions of classical objects are possible not only in analysis; we refer the interested reader to the Hoffman paper [12] in which a possible q-deformation of the stuffle algebra $\mathfrak{H}_*$ is discussed. The objective of the present section is to discuss problems of q-extension for multiple zeta values.

The simplest (and rather obvious) way is as follows: for positive integers s_1, s_2, $\dots$, s_l we set

$$\begin{aligned}\zeta_q^*(x_{\boldsymbol{s}}) &= \zeta_q^*(\boldsymbol{s}) = \zeta_q^*(s_1, s_2, \dots, s_l)\\ &:= \sum_{n_1 > n_2 > \dots > n_l \geqslant 1} \frac{q^{n_1 s_1 + n_2 s_2 + \dots + n_l s_l}}{(1 - q^{n_1})^{s_1} (1 - q^{n_2})^{s_2} \cdots (1 - q^{n_l})^{s_l}}\,, \qquad |q| < 1,\end{aligned} \tag{66}$$

and additively extend the $\mathbb{Q}$-linear map ζ_q^* to the whole algebra $\mathfrak{H}^1$. An easy verification shows that, if $s_1 > 1$, then

$$\lim_{\substack{q \to 1\\ 0<q<1}} (1 - q)^{|\boldsymbol{s}|} \zeta_q^*(\boldsymbol{s}) = \zeta(\boldsymbol{s}),$$

that is, the series in (66) are really q-extensions of the series in (4). Moreover, ζ_q^* is a (q-parametric) homomorphism of the stuffle algebra $\mathfrak{H}^1_*$; to prove this fact, it suffices to consider the specialization $t_n = q^n/(1-q^n)$ of the Hoffman homomorphism ϕ defined in Section 10. Hence,

$$\zeta_q^*(w_1 * w_2) = \zeta_q^*(w_1)\zeta_q^*(w_2) \quad \text{for all} \quad w_1, w_2 \in \mathfrak{H}^1.$$

This model of multiple q-zeta values (and also of generalized q-polylogarithms) is described in [23]; the main demerit of the model is the absence of any description of other linear and polynomial relations over $\mathbb{Q}$, in other words, the absence of a suitable q-shuffle product.

Another way to q-extend (non-multiple) zeta values was suggested simultaneously and independently in [15] and [34],

$$\zeta_q(s) = \sum_{n=1}^{\infty} \sigma_{s-1}(n) q^n = \sum_{n=1}^{\infty} \frac{n^{s-1} q^n}{1-q^n}\,, \qquad s = 1, 2, \dots, \tag{67}$$

where $\sigma_{s-1}(n) = \sum_{d|n} d^{s-1}$ stands for the sum of powers of the divisors; the limit relations

$$\lim_{\substack{q\to 1\\ 0<q<1}} (1-q)^s \zeta_q(s) = (s-1)! \cdot \zeta(s), \qquad s = 2, 3, \dots,$$

are also proved in these papers. The q-zeta values (67) can readily be recalculated in terms of (66) with $l = 1$, namely,

$$\zeta_q(1) = \sum_{n=1}^{\infty} \frac{q^n}{1-q^n}\,, \quad \zeta_q(2) = \sum_{n=1}^{\infty} \frac{q^n}{(1-q^n)^2}\,, \quad \zeta_q(3) = \sum_{n=1}^{\infty} \frac{q^n(1+q^n)}{(1-q^n)^3}\,,$$

$$\zeta_q(4) = \sum_{n=1}^{\infty} \frac{q^n(1+4q^n+q^{2n})}{(1-q^n)^4}\,, \quad \zeta_q(5) = \sum_{n=1}^{\infty} \frac{q^n(1+11q^n+11q^{2n}+q^{3n})}{(1-q^n)^5}\,,$$

and, generally,

$$\zeta_q(k) = \sum_{n=1}^{\infty} \frac{q^n \rho_k(q^n)}{(1-q^n)^k}\,, \qquad k = 1, 2, 3, \dots,$$

where the polynomials $\rho_k(x) \in \mathbb{Z}[x]$ are determined recursively by the formulae

$$\rho_1 = 1, \qquad \rho_{k+1} = (1+(k-1)x)\rho_k + x(1-x)\rho_k' \quad \text{for} \quad k = 1, 2, \dots$$

(see [34]).

If $s \geqslant 2$ is even, then the series $E_s(q) = 1 - 2s\zeta_q(s)/B_s$, where the Bernoulli numbers $B_s \in \mathbb{Q}$ are already defined in (3), are known as the *Eisenstein series*. This fact enables one to prove the coincidence of the rings $\mathbb{Q}[q, \zeta_q(2), \zeta_q(4), \zeta_q(6), \zeta_q(8), \zeta_q(10), \dots]$ and $\mathbb{Q}[q, \zeta_q(2), \zeta_q(4), \zeta_q(6)]$ (cf. the corresponding result in Section 1 for ordinary zeta values). However, the problem to construct a model of multiple q-zeta values that includes the ordinary multiplicity-free model (67) remains open. The natural requirement concerning such a model is the existence of q-analogues of the shuffle and stuffle product relations. In conclusion we present a possible q-extension of Euler's formula (5) for the quantity

$$\zeta_q(2,1) = \sum_{n_1 > n_2 \geqslant 1} \frac{q^{n_1}}{(1-q^{n_1})^2(1-q^{n_2})}\,.$$

Theorem 15. *The following identity holds*:

$$2\zeta_q(2,1) = \zeta_q(3).$$

Proof. As in the proof of Theorem 1, we use the method of partial fractions, namely, the expansion

$$\frac{1}{(1-u)(1-uv)^s} = \frac{1}{(1-v)^s(1-u)} - \sum_{j=0}^{s-1} \frac{v}{(1-v)^{j+1}(1-uv)^{s-j}}, \qquad s=1,2,3,\dots. \tag{68}$$

This identity can be proved in the same way as (9), by summing the geometric progression on the right-hand side. For $s=2$ we multiply the identity (68) by $u(1+v)$,

$$\frac{u(1+v)}{(1-u)(1-uv)^2} = \frac{u(1+v)}{(1-v)^2(1-u)} - \frac{uv(1+v)}{(1-v)(1-uv)^2} - \frac{uv(1+v)}{(1-v)^2(1-uv)},$$

set $u=q^m$ and $v=q^n$, and sum over all positive integers m and n. This results in the equality whose left-hand side contains the double sum

$$\sum_{m=1}^\infty\sum_{n=1}^\infty \frac{q^m(1+q^n)}{(1-q^m)(1-q^{n+m})^2} = \sum_{n=1}^\infty\sum_{m=1}^\infty \frac{q^n(1+q^m)}{(1-q^n)(1-q^{n+m})^2}$$

and the right-hand side is the sum

$$\begin{aligned}
&\sum_{n=1}^\infty\sum_{m=1}^\infty \biggl(\frac{q^m(1+q^n)}{(1-q^n)^2(1-q^m)} - \frac{q^{n+m}(1+q^n)}{(1-q^n)(1-q^{n+m})^2} - \frac{q^{n+m}(1+q^n)}{(1-q^n)^2(1-q^{n+m})}\biggr) \\
&\qquad = \sum_{n=1}^\infty \frac{1+q^n}{(1-q^n)^2} \sum_{m=1}^\infty \biggl(\frac{q^m}{1-q^m} - \frac{q^{n+m}}{1-q^{n+m}}\biggr) - \sum_{n=1}^\infty\sum_{m=1}^\infty \frac{q^{n+m}(1+q^n)}{(1-q^n)(1-q^{n+m})^2}.
\end{aligned}$$

Carrying the last sum to the left-hand side, we obtain

$$\begin{aligned}
&\sum_{n=1}^\infty\sum_{m=1}^\infty \frac{q^n(1+q^m)+q^{n+m}(1+q^n)}{(1-q^n)(1-q^{n+m})^2} \\
&\qquad = \sum_{n=1}^\infty \frac{1+q^n}{(1-q^n)^2} \sum_{m=1}^\infty \biggl(\frac{q^m}{1-q^m} - \frac{q^{n+m}}{1-q^{n+m}}\biggr) = \sum_{n=1}^\infty \frac{1+q^n}{(1-q^n)^2} \sum_{m=1}^n \frac{q^m}{1-q^m} \\
&\qquad = \sum_{n=1}^\infty \frac{1+q^n}{(1-q^n)^2} \biggl(\frac{q^n}{1-q^n} + \sum_{m=1}^{n-1} \frac{q^m}{1-q^m}\biggr) = \zeta_q(3) + \sum_{n>m\geqslant1} \frac{(1+q^n)q^m}{(1-q^n)^2(1-q^m)}.
\end{aligned} \tag{69}$$

On the other hand, the left-hand side of the last equality can be represented in the form $(n+m=l)$

$$\sum_{n=1}^\infty\sum_{l=n+1}^\infty \frac{q^n+2q^l+q^{l+n}}{(1-q^n)(1-q^l)^2} = \sum_{l>n\geqslant1} \frac{q^n+2q^l+q^{l+n}}{(1-q^l)^2(1-q^n)}, \tag{70}$$

and hence, setting $n_1 = n$ and $n_2 = m$ on the right-hand side of (69) and $n_1 = l$ and $n_2 = n$ in (70), we finally obtain the desired identity

$$\begin{aligned}\zeta_q(3) &= \sum_{n_1>n_2\geqslant 1} \frac{q^{n_2}+2q^{n_1}+q^{n_1+n_2}}{(1-q^{n_1})^2(1-q^{n_2})} - \sum_{n_1>n_2\geqslant 1} \frac{(1+q^{n_1})q^{n_2}}{(1-q^{n_1})^2(1-q^{n_2})} \\ &= \sum_{n_1>n_2\geqslant 1} \frac{2q^{n_1}}{(1-q^{n_1})^2(1-q^{n_2})}.\end{aligned}$$

Bibliography

[1] R. Apéry, "Irrationalité de $\zeta(2)$ et $\zeta(3)$", *Astérisque* **61** (1979), 11–13.
[2] J. Borwein and D. Bradley, "Empirically determined Apéry-like formulae for $\zeta(4n+3)$", *Experiment. Math.* **6**:3 (1997), 181–194.
[3] J. M. Borwein, D. M. Bradley, and D. J. Broadhurst, "Evaluations of k-fold Euler/Zagier sums: A compendium of results for arbitrary k", *Electron. J. Combin.* **4** (1997), #R5; Printed version, *J. Combin.* **4**:2 (1997), 31–49.
[4] J. M. Borwein, D. M. Bradley, D. J. Broadhurst, and P. Lisoněk, "Special values of multiple polylogarithms", *Trans. Amer. Math. Soc.* **353**:3 (2001), 907–941.
[5] L. Euler, "Meditationes circa singulare serierum genus", *Novi Comm. Acad. Sci. Petropol.* **20** (1775), 140–186; Reprinted, Opera Omnia Ser. I, vol. 15, Teubner, Berlin 1927, pp. 217–267.
[6] A. O. Gel'fond, *Calculus of finite differences*, Nauka, Moscow 1967; English transl., Intern. Monographs Adv. Math. Phys., Hindustan Publ. Corp., Delhi 1971.
[7] A. B. Goncharov, "Polylogarithms in arithmetic and geometry", Proceedings of the International Congress of Mathematicians (ICM'94, Zürich, August 3–11, 1994) (S. D. Chatterji, ed.), vol. I, Birkhäuser, Basel 1995, pp. 374–387.
[8] A. B. Goncharov, "The double logarithm and Manin's complex for modular curves", *Math. Res. Lett.* **4**:5 (1997), 617–636.
[9] E. Heine, "Über die Reihe …", *J. Reine Angew. Math.* **32** (1846), 210–212.
[10] M. E. Hoffman, "Multiple harmonic series", *Pacific J. Math.* **152**:2 (1992), 275–290.
[11] M. E. Hoffman, "The algebra of multiple harmonic series", *J. Algebra* **194**:2 (1997), 477–495.
[12] M. E. Hoffman, "Quasi-shuffle products", *J. Algebraic Combin.* **11**:1 (2000), 49–68.
[13] M. E. Hoffman and Y. Ohno, "Relations of multiple zeta values and their algebraic expression", Preprint, 2000; E-print `math.QA/0010140`.
[14] K. Ihara and M. Kaneko, "Derivation relations and regularized double shuffle relations of multiple zeta values", Preprint, 2000.
[15] M. Kaneko, N. Kurokawa, and M. Wakayama, "A variation of Euler's approach to values of the Riemann zeta function", E-print `math.QA/0206171`, 2002.
[16] M. Koecher, "Letter to the editor", *Math. Intelligencer* **2**:2 (1979/1980), 62–64.
[17] F. Lindemann, "Über die Zahl π", *Math. Ann.* **20** (1882), 213–225.
[18] H. M. Minh, G. Jacob, M. Petitot, and N. E. Oussous, "Aspects combinatoires des polylogarithmes et des sommes d'Euler–Zagier", *Sém. Lothar. Combin.* **43** (1999), Art. B43e (electronic); `http://www.mat.univie.ac.at/~slc/wpapers/s43minh.html`.
[19] H. M. Minh and M. Petitot, "Lyndon words, polylogarithms and the Riemann ζ function", Formal Power Series and Algebraic Combinatorics (Vienna 1997), *Discrete Math.* **217**:1–3 (2000), 273–292.
[20] H. M. Minh, M. Petitot, and J. van der Hoeven, "Shuffle algebra and polylogarithms", Formal Power Series and Algebraic Combinatorics (Toronto, ON, June 1998), *Discrete Math.* **225**:1–3 (2000), 217–230.
[21] Y. Ohno, "A generalization of the duality and sum formulas on the multiple zeta values", *J. Number Theory* **74**:1 (1999), 39–43.
[22] T. Rivoal, "La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs", *C. R. Acad. Sci. Paris Sér. I Math.* **331**:4 (2000), 267–270.

[23] K.-G. Schlesinger, "Some remarks on q-deformed multiple polylogarithms", E-print `math.QA/0111022`, 2001.
[24] V. N. Sorokin, "On the linear independence of the values of generalized polylogarithms", *Mat. Sb.* **192**:8 (2001), 139–154; English transl., *Sb. Math.* **192** (2001), 1225–1239.
[25] E. A. Ulanskiĭ [Ulansky], "Identities for generalized polylogarithms", *Mat. Zametki* **73**:4 (2003), 613–624; English transl., *Math. Notes* **73** (2003).
[26] D. V. Vasil'ev [Vasilyev], "Some formulae for the zeta function at integer points", *Vestnik Moscov. Univ. Ser. I Mat. Mekh.*:1 (1996), 81–84; English transl., *Moscow Univ. Math. Bull.* **51** (1996), 41–43.
[27] M. Waldschmidt, "Valeurs zêta multiples: une introduction", *J. Théor. Nombres Bordeaux* **12**:2 (2000), 581–595.
[28] M. Waldschmidt, "Multiple polylogarithms", Lectures at Institute of Math. Sciences (Chennai, November 2000); Updated version, `http://www.math.jussieu.fr/~miw/articles/ps/mpl.ps`.
[29] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge University Press, Cambridge 1927.
[30] D. Zagier, "Values of zeta functions and their applications", First European Congress of Mathematics (Paris 1992), V. II (A. Joseph et al, ed.), Progr. Math., vol. 120, Birkhäuser, Boston 1994, pp. 497–512.
[31] J. Zhao, "Analytic continuation of multiple zeta functions", *Proc. Amer. Math. Soc.* **128**:5 (2000), 1275–1283.
[32] S. A. Zlobin, "Integrals expressible as linear forms in generalized polylogarithms", *Mat. Zametki* **71**:5 (2002), 782–787; English transl., *Math. Notes* **71** (2002), 711–716.
[33] V. V. Zudilin [W. Zudilin], "Very well-poised hypergeometric series and multiple integrals", *Uspekhi Mat. Nauk* **57**:4 (2002), 177–178; English transl., *Russian Math. Surveys* **57** (2002), 824–826.
[34] V. V. Zudilin [W. Zudilin], "Diophantine problems for q-zeta values", *Mat. Zametki* **72**:6 (2002), 936–940; English transl., *Math. Notes* **72** (2002), 858–862.

M. V. Lomonosov Moscow State University
E-mail address: wadim@ips.ras.ru

Received 30/OCT/2001