Russian Math. Surveys **58**:1 000–000

Uspekhi Mat. Nauk **58**:1 3–32
DOI XXXX

Algebraic relations of multiple zeta values

V. V. Zudilin [W. Zudilin]

Abstract. The subject of the survey is a multidimensional generalization of the Riemann zeta function as a function of natural argument.
Bibliography: 34 items.

1. Introduction

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

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

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

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

is known; the knowledge gives the expression of the values of the zeta function at even integers in terms of the number

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

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

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

Relation (2) yields the coincidence of the rings $\mathbb{Q}[\zeta(2),\zeta(4),\zeta(6),\zeta(8),\dots]$ and $\mathbb{Q}[\pi^2]$, hence, due to Lindemann's theorem [17] on the transcendence of π, we may conclude that each of the rings has transcendence degree 1 over the field of rational numbers. It is much less known on the arithmetic nature of the values of the zeta function at odd integers $s=3,5,7,\dots$: Apéry has proved [1] the irrationality of the number $\zeta(3)$ and, recently, Rivoal settles [22] the infiniteness of the set of irrational numbers among $\zeta(3),\zeta(5),\zeta(7),\dots$. Conjecturally, each of these numbers is transcendent, and the full answer on the above-stated question, about polynomial relations over $\mathbb{Q}$ for the values of series (1) with $s\geqslant 2$ integer, looks very simple.

Conjecture 1. *The numbers*

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

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

This conjecture may be regarded as a mathematical folklore (see, e.g., [7] and [28]). In this survey, a certain generalization of the problem of algebraic independence for the values of the Riemann zeta function at positive integers (*zeta values*) is discussed. Namely, we will speak on the object that is extensively studied during the last decade in connection with problems of not only number theory but also of combinatorics, algebra, analysis, algebraic geometry, quantum physics, and many other branches of mathematics. However, there are no printed works in Russian since nowadays devoted to the subject (mention only the paper [25] in press). By means of the present publication, we hope to attract attention of Russian mathematicians to problems concerning *multiple zeta values.*

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

2. Multiple zeta values

Series (1) enables the following multidimensional generalization. For positive integers $s_1, s_2, \dots, s_l$ with $s_1 > 1$, consider the values of the l-tuple zeta function

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

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

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

which was several times rediscovered after. The quantities (4) are 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 are stated as well as a series of conjectures (that has been partly proved later) on the structure of algebraic relations for the family (4) is indicated. Hoffman also suggests [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 summation. Of course, all relations of series (6) may be rewritten without difficulty for series (4) (see, e.g., [10] and [25]), although several identities possess a compact form by means of just multiple zeta values (6) (see relations (38) in Section 7 below).

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

Note [31] that the series on the right-hand side of (4) converges absolutely in the region $\operatorname{Re} s_1 > 1$, $\sum_{k=1}^{l} \operatorname{Re} s_k > l$; moreover, the multiple zeta function $\zeta(\boldsymbol{s})$ defined in the region by series (4) can be analytically continued to the 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 $m \geqslant 1$ are integer numbers. The questions of existence of a functional equation for $l > 1$ and of localization of non-trivial zeros (the analogue of Riemann's conjecture) for the function $\zeta(\boldsymbol{s})$, remain open.

3. Identities: the partial-fraction method

In this section, we will give examples of identities for multiple zeta values that can be deduced by an elementary analytic method, the *partial-fraction method.*

Theorem 1 (Hoffman's relations [10], Theorem 5.1). *For any admissible multi-index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$, the identity*

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

holds.

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

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

hence

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

$$= \sum_{n_1>\dots>n_k>n_{k+1}>\dots>n_l\geqslant 1} \sum_{m=n_{k+1}+1}^{n_k} \frac{1}{m n_1^{s_1}\cdots n_k^{s_k} n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}$$
$$= \sum_{n_1>n_2>\dots>n_l\geqslant 1} \frac{1}{n_1^{s_1} n_2^{s_2}\cdots n_l^{s_l}} \sum_{m=n_{k+1}+1}^{n_k} \frac{1}{m}\,.$$

Therefore

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

where we introduce the notation $M_k = m_1+m_2+\dots+m_k$ for $k=1,\dots,l+1$ (clearly, $M_k = n_{l+1-k}$ for $k=1,\dots,l$). Notice now the following partial-fraction expansion (in terms of the parameter u):

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

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

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

hence

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

Going on equality (8), we find that

$$\sum_{k=1}^{l}\bigl(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\bigr)$$
$$= \sum_{j=0}^{s_1-2} \sum_{m_1,m_2,\dots,m_{l+1}\geqslant 1} \frac{1}{M_1^{s_l} M_2^{s_{l-1}}\cdots M_{l-1}^{s_2} M_l^{j+1} M_{l+1}^{s_1-j}}$$
$$+ \sum_{m_1,m_2,\dots,m_{l+1}\geqslant 1} \frac{1}{M_1^{s_l} M_2^{s_{l-1}}\cdots M_{l-1}^{s_2} m_{l+1} M_{l+1}^{s_1}}$$

$$= \sum_{j=0}^{s_1-2} \zeta(s_1 - j, j+1, s_2, \dots, s_l) + \sum_{m_1, m_2, \dots, m_{l+1} \geqslant 1} \frac{1}{M_1^{s_l} M_2^{s_{l-1}} \cdots M_{l-1}^{s_2} m_l M_{l+1}^{s_1}} \tag{10}$$

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

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

therefore

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

Applying consequently, in inverse order (i.e., starting from $k = l-1$ and ending on $k = 1$), identities (11) for the tuple sum on the right-hand side of equality (10), we obtain

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

Realizing all necessary cancellations of the left-hand and right-hand sides of equality (12), we finally arrive at the desired identity (7).

If $l=1$, the statement of Theorem 1 can be written in the following form.

Corollary (Euler's theorem). *For any integer $s\geqslant 3$, the identity*

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

holds.

Note also that, in the case $s=3$, identity (13) becomes nothing else but relation (5).

In the work [13], the following result is also proved by means of the partial-fraction method.

Theorem 2 (Cyclic sum theorem). *For any admissible multi-index $\boldsymbol{s}=(s_1,s_2,\dots,s_l)$, 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.

Theorem 2 directly yields the result 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 the above mentioned Euler's theorem.

Theorem 3 (Sum theorem). *For any integers $s>1$ and $l\geqslant 1$, 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.

Theorem 1 and 3 are particular 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 all known algebraic relations (i.e., numerical identities) over $\mathbb{Q}$ for the quantities (4), it becomes useful to represent ζ as a linear map of a certain polynomial algebra into the field of real numbers. Consider coding of multi-indices $\boldsymbol{s}$ by words (i.e., by monomials in non-commutative variables) over the alphabet $X=\{x_0,x_1\}$ by the rule

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

Set

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

for all admissible (starting with x_0 and ending on x_1) words; then the weight (or degree) $|x_{\boldsymbol{s}}| := |\boldsymbol{s}|$ coincides with the total degree of the monomial $x_{\boldsymbol{s}}$, while the length $\ell(x_{\boldsymbol{s}}) := \ell(\boldsymbol{s})$ expresses 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 graded by degree $\mathbb{Q}$-algebra (where the degree of each variable x_0 and x_1 is agreed to be 1) of polynomials in the two non-commutative variables; we identify the algebra $\mathbb{Q}\langle X\rangle$ with the graded $\mathbb{Q}$-vector space $\mathfrak{H}$ spanned over monomials in the variables x_0 and x_1. Define as well 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 $\mathfrak{H}^1$ may be regarded as the subalgebra of $\mathbb{Q}\langle X\rangle$ generated by the words $y_s = x_0^{s-1}x_1$, while $\mathfrak{H}^0$ is the $\mathbb{Q}$-vector space spanned over all admissible words. Now, we may view 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).

Define the multiplications $\sqcup\!\sqcup$ (the *shuffle product*) on $\mathfrak{H}$ and $*$ (the *harmonic* or *stuffle product*) on $\mathfrak{H}^1$ by the rules

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

for any word w, and

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

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

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

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

for any word w and integer $j \geqslant 1$, to rule (17). Note that inductive arguments allow to prove commutativity and associativity of each of the multiplications (see Section 8 below for this); the corresponding algebras $\mathfrak{H}_{\sqcup\!\sqcup} := (\mathfrak{H}, \sqcup\!\sqcup)$, $\mathfrak{H}^1_* := (\mathfrak{H}^1, *)$ (and also $\mathfrak{H}_* := (\mathfrak{H}, *)$) are examples of so-called *Hopf algebras*.

The following two statements motivate consideration of the introduced multiplications $\sqcup\!\sqcup$ and $*$; their proofs can be found in [11], [13], and [28].

Theorem 4. *The map ζ is a homomorphism of the shuffle algebra $\mathfrak{H}^0_{\sqcup\!\sqcup} := (\mathfrak{H}^0, \sqcup\!\sqcup)$ into $\mathbb{R}$, i.e.,*

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

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

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

Later we will give detailed proofs of the two theorems using the differential-difference origin of the multiplications $\sqcup\!\sqcup$ and $*$ in suitable functional models of the algebras $\mathfrak{H}_{\sqcup\!\sqcup}$ and $\mathfrak{H}^0_*$. Proving Theorem 4 (see Section 5), we follow a scheme of the work [27], while our proof of Theorem 5 (in Section 9) is new.

One more family of identities is given by the following statement that will be deduced in Section 11 from Theorem 1.

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 known to the moment (proved and experimentally derived) identities for the multiple zeta values follow from identities (19)–(21). Therefore the following conjecture looks rather truthful.

Conjecture 2 [11], [18], [27]**.** *All algebraic relations over $\mathbb{Q}$ of multiple zeta values are generated by 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

In order to prove shuffle relations (19) for multiple zeta values, let us define the *generalized polylogaithms*

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

for any collection of positive integers $s_1, s_2, \dots, s_l$. By definition,

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

Taking, as before for multiple zeta values,

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

let us extend action of the map $\mathrm{Li}\colon w \mapsto \mathrm{Li}_w(z)$ by linearity on the graded algebra $\mathfrak{H}^1$ (not $\mathfrak{H}$, since 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 x_j the first letter in its record (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

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

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

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

Therefore, in the case $s_1 > 1$ (corresponding to the letter $x_j = x_0$), we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z} \operatorname{Li}_{x_0 u}(z) &= \frac{1}{z} \sum_{n_1 > n_2 > \dots > n_l \geqslant 1} \frac{z^{n_1}}{n_1^{s_1 - 1} n_2^{s_2} \cdots n_l^{s_l}} \\ &= \frac{1}{z} \operatorname{Li}_{s_1 - 1, s_2, \dots, s_l}(z) = \frac{1}{z} \operatorname{Li}_u(z) \end{aligned}$$

and, in the case $s_1 = 1$ (corresponding to the letter $x_j = x_1$), we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z} \operatorname{Li}_{x_1 u}(z) &= \sum_{n_1 > n_2 > \dots > n_l \geqslant 1} \frac{z^{n_1 - 1}}{n_2^{s_2} \cdots n_l^{s_l}} = \sum_{n_2 > \dots > n_l \geqslant 1} \frac{1}{n_2^{s_2} \cdots n_l^{s_l}} \sum_{n_1 = n_2 + 1}^{\infty} z^{n_1 - 1} \\ &= \frac{1}{1-z} \sum_{n_2 > \dots > n_l \geqslant 1} \frac{z^{n_2}}{n_2^{s_2} \cdots n_l^{s_l}} = \frac{1}{1-z} \operatorname{Li}_{s_2, \dots, s_l}(z) = \frac{1}{1-z} \operatorname{Li}_u(z), \end{aligned}$$

and the result follows.

Lemma 1 motivates another definition of the generalized polylogarithms, now defined for all elements of the algebra $\mathfrak{H}$. As before, it is sufficient to give it for words $w \in \mathfrak{H}$ only, distributing then over all algebra by linearity; set $\operatorname{Li}_{\mathbf{1}}(z) = 1$ and

$$\operatorname{Li}_w(z) = \begin{cases} \dfrac{\log^s z}{s!} & \text{if } w = x_0^s \text{ for some } s \geqslant 1, \\ \displaystyle\int_0^z \omega_j(z) \operatorname{Li}_u(z) \, \mathrm{d}z & \text{if } w = x_j u \text{ contains the letter } x_1. \end{cases} \tag{27}$$

Evidently, Lemma 1 remains true for this extended version (27) of the polylogarithms (the fact yields coincidence of the newly-defined polylogarithms with the "old" ones (24) for words w in $\mathfrak{H}^1$); in addition,

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

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

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

Proof. We have to verify the equalities

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

it is sufficient to do this job for *words* $w_1, w_2 \in \mathfrak{H}$. We will prove equality (28) by induction on the quantity $|w_1| + |w_2|$. If $w_1 = \mathbf{1}$ or $w_2 = \mathbf{1}$, relation (28) becomes tautological by (15). Otherwise, $w_1 = x_j u$ and $w_2 = x_k v$, hence by Lemma 1 and the inductive hypothesis we have

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

Thus

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

and letting $z \to 0+0$ if at least one of the words w_1, w_2 contains letter x_1, or substituting $z = 1$ if the records of w_1, w_2 consist of letter x_0 only, gives the relation $C = 0$. Therefore, equality (29) becomes the required relation (28), and lemma follows.

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

Explicit computation of the monodromy group for the system of differential equations (25) allows to Minh, Petitot, and van der Hoeven to prove that the homomorphism $w \mapsto \mathrm{Li}_w(z)$ of the shuffle algebra $\mathfrak{H}_{ш}$ over $\mathbb{C}$ is bijective, i.e., all $\mathbb{C}$-algebraic relations for generalized polylogarithms are originated by shuffle relations (28) only; in particular, generalized polylogarithms are linearly independent over $\mathbb{C}$. A much simpler proof of the linear independence of functions (22), as a consequence of elegant identities for the functions, is due to Ulanskiĭ [25]; the same statement also follows from Sorokin's result [24].

6. Duality theorem

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

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

if $0 < z < 1$. When $x_{\varepsilon_1} \neq x_1$, i.e. $w \in \mathfrak{H}^0$, the integral in (30) converges in the region $0 < z \leqslant 1$, hence, in accordance with (23), we reduce representation [30] for the multiple zeta values

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

in a form of *Chen's iterated integrals.* The following result is evident application of the integral representation (31).

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

Theorem 7 (Duality theorem [30]). *For any word $w \in \mathfrak{H}^0$, the relation*

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

holds.

Proof. To prove the theorem, it is sufficiently to do the change of variable $z_1' = 1 - z_k$, $z_2' = 1 - z_{k-1}$, $\ldots$, $z_k' = 1 - z_1$, and apply relations $\omega_0(z) = \omega_1(1-z)$ followed from (26).

As the simplest consequence of Theorem 7, notice (again) identity (5), which follows for the word $w = x_0^2 x_1$, as well as the general identity

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

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

7. Identities: the generating-function method

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

Let us first remark that, for an admissible multi-index $\boldsymbol{s} = (s_1, \ldots, s_l)$, the corresponding set of *periodic* polylogarithms

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

(see, e.g., [4], [28]), possesses the generating function

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

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

$$\left(\left((1-z)\frac{\mathrm{d}}{\mathrm{d}z}\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)^{s-1} - t^s\right) L_{\boldsymbol{s}}(z, t) = 0,$$

and its solution may be written explicitly by means of a generalized hypergeometric series (see [3], [4], [28]).

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

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

where $F(a,b;c;z)$ denotes the Gauß's hypergeometric function.

Proof. Routine verification (with a help of Lemma 1 for the left-hand side) shows that the both sides of the required equality are annihilated by action of the differential operator

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

in addition, the first terms in z-expansions of the both sides in (33) coincide:

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

Thus the statement of the lemma follows.

Theorem 8 [4], [28]**.** *For any integer $n \geqslant 1$, the identity*

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

holds.

Proof. By the Gauß summation formula [29], Ch. 14,

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

substituting $z=1$ into equality (33) yields

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

Comparison of the coefficients in the same powers of t gives the desired identity.

The statement of Theorem 8 was conjectured in [30]. Identity (34) is not the unique example of application of generating functions. We present more identities from [3], similar to (34), for which the above method is also effective:

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

where $n = 1, 2, \ldots$. Identities

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

may be derived by the generating-function method [10] as well as by application of the stated Theorem 7.

An example of other-type generating functions relates to generalization 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]:

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

Their proofs as well as proofs of several other identities is based on transformation and summation formulae of *generalized hypergeometric functions*, similar to application of formula (35) in deducing Theorem 8. Identities (37) are extraordinary useful in fast computation of the Riemann zeta function at odd integers.

Let us also remark the relations

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

that are obtained by consequent application of the results in [26] (or [33]) and [32]. Equalities (38) also generalize Euler's identity (5) and are deeply related to one way of showing Apéry's theorem [1] and Rivoal's theorem [22], mentioned in Section 1. However, deducing relations (38) from Theorems 4–6 for arbitrary integer n is not yet known.

8. Quasi-shuffle products

Construction, proposed by Hoffman [12], allows to consider each of the algebras $\mathfrak{H}_{\sqcup\!\sqcup}$ and $\mathfrak{H}^1_*$ as a particular case of some general algebraic structure. Description of the structure is the subject of the section.

Consider the non-commutative, graded by degree, polynomial algebra $\mathfrak{A} = \mathcal{K}\langle A\rangle$ over the field $\mathcal{K} \subset \mathbb{C}$; here A denotes a locally finite set of generators (i.e., the set of generators of fixed positive degree is finite). As usual, elements of the set A are said to be letters and monomials in these letters are words. To any word w, assign its length (the number of letters in the record) $\ell(w)$ and its weight (the sum of degrees of the letters) $|w|$. 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 $\mathbf{0}$.

Now, define the product $\circ$, additively distributing it over 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 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 it is the 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 particular case $A=\{x_0,x_1\}$, we obtain the shuffle algebra $\mathfrak{A}_\circ=\mathfrak{H}_{\sqcup\mathchoice{}{}{}{}\!\sqcup}$ 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

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

On the algebra $\mathfrak{A}$ with the give functional (41), define the dual product $\bar{\circ}$ by the rules

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

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

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

Proof. It is sufficient to prove the relation

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

for all words $w_1,w_2\in\mathcal{K}\langle A\rangle$. We will proceed the proof by induction on the quantity $\ell(w_1)+\ell(w_2)$. If $\ell(w_1)=0$ or $\ell(w_2)=0$, then relation (42) becomes the evident identity. If $\ell(w_1)=\ell(w_2)=1$, i.e. $w_1=a_1$ and $w_2=a_2$ are letters, then

$$a_1\circ a_2 = a_1a_2 + a_2a_1 + [a_1,a_2] = a_1\mathbin{\bar{\circ}} a_2.$$

If $\ell(w_1) > 1$ and $\ell(w_2) = 1$, then writing $w_1 = a_1ua_2$ and $w_2 = a_3 \in A$ and applying the inductive hypothesis we deduce that

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

In the same vein (but with more cumbersome computations), we proceed in the remaining case $\ell(w_1) > 1$ and $\ell(w_2) > 1$. Namely, writing $w_1 = a_1ua_2$, $w_2 = a_3va_4$ and applying the inductive hypothesis we obtain

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

This concludes the proof.

Remark. If the graded algebras possess property (S3), the above proof may be essentially simplified. Nevertheless, we find the fact of coincidence of the algebras $\mathfrak{A}_\circ$ and $\mathfrak{A}_{\bar\circ}$ in the most general settings, i.e. when the functional (41) satisfies properties (S0)–(S2), to be rather important.

In conclusion of the section, we will proof an auxiliary statement.

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

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

Proof. We will prove the statement by induction on the number of letters in the word u. If the word u is empty, then identity (43) is evident. Otherwise, write the word u as $u = a_1u_1$, where $a_1 \in A$ and the word u_1 consists of less number of letters, hence the identity

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

holds. Then

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

which is the desired result.

9. Functional model of stuffle algebra

The functional model of the stuffle algebra $\mathfrak{H}_*$ cannot be described in the full analogy with the polylogarithmic model of the shuffle algebra $\mathfrak{H}_{\sqcup\!\sqcup}$, since rule (17) has no differential interpretation as (16). Therefore we shall use a difference interpretation of rule (17), namely, the (simplest) difference operator

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

It can be easily verified that

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

and that inverse mapping

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

hence $D(Ig(t)) = g(t)$, is defined up to an additive constant provided some additional restrictions 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 indicated equality is justified by formal application of the Taylor expansion:

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

however the formula is valid for any entire function. Exponentiating derivations (in word algebras), in connection with generalization of Theorem 1, is discussed in Section 12 below.

A natural analogy with Lemmas 1 and 2, by (17) and (44) provides the existence of functions $\omega_j(t)$ satisfying the properties

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

The simplest choice is given by the formulae

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

and yields us to the functions

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

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

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

that, in some sense, is a discrete analogue of formula (25).

Lemma 5. *The following identity holds:*

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

in particular,

$$\mathrm{Ri}_{\boldsymbol{s}}(0) = \zeta(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \geqslant 2,\ s_2 \geqslant 1,\ \dots,\ s_l \geqslant 1, \tag{47}$$

$$\lim_{t \to +\infty} \mathrm{Ri}_{\boldsymbol{s}}(t) = 0, \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \geqslant 2,\ s_2 \geqslant 1,\ \dots,\ s_l \geqslant 1. \tag{48}$$

Proof. By definition, we find that

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

and this implies the required formula (46).

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

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

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. It is sufficient to verify the relations

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

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

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

Therefore

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

and letting t tend to $+\infty$, by (48) we obtain $C = 0$. 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 stuffle algebra

Another way to prove Theorem 5 (and Lemma 6 as well) is due to 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 the countable set of (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 (actually, the monomorphism) ϕ is the algebra *QSym* of quasi-symmetric functions. A formal power series (of bounded degree) in $t_1, t_2, \dots$ is called here a *quasi-symmetric function* if the coefficients of $t_{n_1}^{s_1} t_{n_2}^{s_2} \cdots t_{n_l}^{s_l}$ and $t_{n_1'}^{s_1} t_{n_2'}^{s_2} \cdots t_{n_l'}^{s_l}$ are the same whenever $n_1 > n_2 > \dots > n_l$ and $n_1' > n_2' > \dots > n_l'$ (our definition slightly differs from the corresponding version of [13] but leads to the same algebra *QSym* of quasi-symmetric functions). By the above means the homomorphism $w \mapsto \mathrm{Ri}_w(t)$ in Lemma 6 is defined as restriction of the homomorphism ϕ on $\mathfrak{H}^0$ by setting $t_n = 1/(t+n)$, $n = 1, 2, \dots$.

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

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

as an example. In order to do this, we require the integral representation

$$\zeta(\boldsymbol{s})=\underset{[0,1]^{|\boldsymbol{s}|}}{\int\cdots\int}\prod_{j=1}^{l-1}\frac{t_1t_2\cdots t_{s_1+\dots+s_j}}{1-t_1t_2\cdots t_{s_1+\dots+s_j}}\cdot\frac{\mathrm{d}t_1\,\mathrm{d}t_2\cdots\mathrm{d}t_{|\boldsymbol{s}|}}{1-t_1t_2\cdots t_{s_1+s_2+\dots+s_l}}, \quad l=\ell(\boldsymbol{s}), \tag{53}$$

for admissible multi-indices $\boldsymbol{s}$, which differs from that in (31). This representation is kindly pointed out to us by Nesterenko; it may be proved by straightforward integrating the series

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

Substituting $u=t_1\cdots t_{s_1}$, $v=t_{s_1+1}\cdots t_{s_2}$ in 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 identity (52).

11. Derivations

As in Section 8, consider the graded non-commutative polynomial algebra $\mathfrak{A}=\mathcal{K}\langle A\rangle$ over the field $\mathcal{K}$ of characteristic 0 with the 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) that satisfies 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, hence the set of all derivations of the algebra $\mathfrak{A}$ forms the Lie algebra $\mathit{Der}(\mathfrak{A})$ (naturally graded by degree).

It can be easily seen that, for defining a derivation $\delta\in\mathit{Der}(\mathfrak{A})$, it is sufficient to give its image on the generators A and distribute then over the whole algebra by linearity and in accordance with rule (54).

The nest assertion gives examples of derivations of $\mathfrak{A}$, when the algebra possesses an additive multiplication $\circ$ with the properties (39) and (40).

Theorem 10. *For any letter $a\in A$, the map*

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

is a derivation.

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

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

thus (55) is actually a derivation.

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

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

are derivations; thanks to rule (18), the map δ_* is a derivation on the whole algebra $\mathfrak{H}$. We mention the action of derivations (56), obtained in accordance with (15)–(18), on the generators of the algebra:

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

For any derivation δ of the algebra $\mathfrak{H}$ (or of the subalgebra $\mathfrak{H}^0$), define the dual derivation $\overline{\delta}=\tau\delta\tau$, where τ is the anti-automorphism of the algebra $\mathfrak{H}$ (and $\mathfrak{H}^0$) 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 its value on one of the generators x_0 or x_1, while an arbitrary derivation requires its values on the both generators.

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

Theorem 11 (Derivation theorem [13], Theorem 2.1). *For any word $w\in\mathfrak{H}^0$, the identity*

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

holds.

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

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

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

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

that corresponds to the right-hand side in (7). Applying now the map ζ to the both sides of obtained equalities (59) and (60), by Theorem 1 we deduce the required identity (58).

Remark. The condition $w \in \mathfrak{H}^0$ in Theorem 11 cannot be weakened; equality (58) is false 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 action (57) of derivations (56) with those of $D, \overline{D}$ 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 $\delta_* - \delta_{\sqcup\!\sqcup} = \overline{D} - D$. Therefore application of Theorem 11 to the word $w \in \mathfrak{H}^0$ leads to the required equality:

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

This completes the proof.

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

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

can be found in [28]. (As it is easily seen, specializing $z_2 = \dots = z_l = 1$ functions (61) become generalized polylogarithms (22).) We do not have a goal to expose properties of the functional model (61) in this survey, and refer the interested reader to the works [4], [7], and [28].

12. Derivations of Ihara–Kaneko and Ohno's relations

Theorem 11 has a natural generalization. For any $n \geqslant 1$, define the anti-symmetric 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 mentioned in the proof of Theorem 6, we have $\partial_1 = \overline{D} - D = \delta_* - \delta_{\sqcup\!\sqcup}$. The following result is valid.

Theorem 12 [14] (see also [13]). *For any $n \geqslant 1$ and any word $w \in \mathfrak{H}^0$, the identity*

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

holds.

Further, we describe a scheme of the proof of the theorem given in the preprint [14] (the proof in [13] lies on other ideas).

The following result, proved in the paper [21] by means of the partial-fraction method,[1] contains as particular cases Theorems 1, 3, and 7 (corresponding implications are also given in [21]).

[1]**NB.** The Russian version contains the reference on the generating-function method that is completely wrong! ??? CHECK !!!

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 records 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, for any integer $n \geqslant 0$, 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.

Following [14], for each integer $n \geqslant 1$ define the derivation $D_n \in \mathit{Der}(\mathfrak{H})$ setting $D_n x_0 = 0$ and $D_n x_1 = x_0^n x_1$. It may be easily justified that the derivations $D_1, D_2, \dots$ pairwise commute; this holds for the dual derivations $\overline{D}_1, \overline{D}_2, \dots$ as well. Consider a completion of $\mathfrak{H}$, namely the algebra $\widehat{\mathfrak{H}} = \mathbb{Q}\langle\langle x_0, x_1 \rangle\rangle$ of formal power series in non-commutative variables x_0, x_1 over the field $\mathbb{Q}$. Action of the anti-automorphism τ and of derivations $\delta \in \mathit{Der}(\mathfrak{H})$ is naturally extended to the whole algebra $\widehat{\mathfrak{H}}$. For simplicity, the record $w \in \ker\zeta$ will mean that all homogeneous components of the element $w \in \widehat{\mathfrak{H}}$ belongs 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 the standard relation of a derivation and homomorphism implies that the maps

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

are automorphisms of the algebra $\widehat{\mathfrak{H}}$. By the above means, Ohno's relations may be stated as follows.

Theorem 14 [14]**.** *For any word $w \in \mathfrak{H}^0$, the inclusion*

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

holds.

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

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

we may conclude that $\mathcal{D}^n x_0 = 0$ and $\mathcal{D}^n x_1 = (-\log(1-x_0))^n x_1$, 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

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

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

Reminding $\partial_1, \partial_2, \dots$, 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. First of all, let us note pairwise commutativity of the operators ∂_n, $n = 1, 2, \dots$. Indeed, since $\partial_n(x_0+x_1) = 0$ for any $n \geqslant 1$, it is sufficient to verify the equality $\partial_n\partial_m x_0 = \partial_m\partial_n x_0$ for $n, m \geqslant 1$. Taking in mind $\partial_n(x_0+x_1)^k = 0$, for any $n \geqslant 1$ and $k \geqslant 0$ we obtain the desired property:

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

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

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

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 substitution $t = 1$ leads to the required result (64).

Proof of Theorem 12. Now let us show how Theorem 12 follows from Theorem 14 and Lemma 7. First 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 secondly

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

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 of proving relations (62)? Explicit computations in [14] show that $\partial_1 = \delta_* - \delta_{\sqcup\sqcup}$,

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

and, in addition, $\delta_* + \overline{\delta}_* = \delta_{\sqcup\sqcup} + \overline{\delta}_{\sqcup\sqcup}$; therefore cases $n = 1, 2, 3, 4$ in Theorem 12 are served by induction (with Theorem 11 as inductive base). This circumstance motivates the following hypothesis.

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

Note also that the preprint [14] includes some other (in comparison with Conjecture 2) ideas of total description of identities for multiple zeta values in terms of shuffle-stuffle relations.

13. Open questions

In addition to the above-indicated Conjectures 1–3, we mention a series of other important conjectures concerning the structure of the subspace $\ker\zeta \subset \mathfrak{H}$. Denote by $\mathcal{Z}_k$ the $\mathbb{Q}$-vector space in $\mathbb{R}$ spanned by multiple zeta values of weight k; in particular, $\mathcal{Z}_0 = \mathbb{Q}$ and $\mathcal{Z}_1 = \{0\}$. Then the $\mathbb{Q}$-subspace $\mathcal{Z} \in \mathbb{R}$ spanned by all multiple zeta values is the subalgebra of $\mathbb{R}$ over $\mathbb{Q}$ graded by weight.

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

It can be easily seen that relations (19)–(21) for multiple zeta values are homogeneous in weight, hence Conjecture 4 follows from Conjecture 2.

Denoting by d_k the dimension of the $\mathbb{Q}$-space $\mathcal{Z}_k$, $k = 0, 1, 2, \dots$, 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 $k \geqslant 5$, above-deduced identities allow to compute the upper bounds; for instance, $d_5 \leqslant 2$, $d_6 \leqslant 2$, and so on.

Conjecture 5 [30]**.** *For $k \geqslant 3$, the recurrent relations*

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

hold; equivalently,

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

Even if Conjectures 4 and 5 are positively solved, the question of choosing a transcendence basis of the algebra $\mathfrak{Z}$ and (or) a rational basis of the $\mathbb{Q}$-spaces $\mathfrak{Z}_k$, $k = 0, 1, 2, \ldots$, is still open. Concerning this problem, we find the next conjecture of Hoffman rather curious.

Conjecture 6 [11]**.** *For any $k = 0, 1, 2, \ldots$, a basis of the $\mathbb{Q}$-spaces $\mathfrak{Z}_k$ is given by the set of numbers*

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

A serious argument for Conjecture 6 to be valid, is not only experimental confirmation for $k \leqslant 16$ (under the hypothesis of Conjecture 2) but also agreement of the dimension of the $\mathbb{Q}$-space spanned by the numbers (65) with the dimension d_k of the spaces $\mathfrak{Z}_k$ in Conjecture 5. The last fact is proved by Hoffman in [11].

14. q-analogues of multiple zeta values

Thirty three years after Gauß's work on hypergeometric series, Heine considered [9] series depending on the additional parameter q and possessing properties similar to those for Gauß's series. Moreover, when q tends to 0 (at least term-wise), Heine's q-series become hypergeometric series so that Gauß's results may be derived from the corresponding results for q-series by this limit procedure and the theory of analytic continuation.

Similar q-extensions of classical objects are possible not only in analysis: the interested reader is referred to Hoffman's work [12], where possible q-deformation of the stuffle algebra $\mathfrak{H}_*$ are discussed. The aim of this section is to consider problems of q-extending multiple zeta values.

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

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

and distribute the $\mathbb{Q}$-linear map ζ_q^* over the whole algebra $\mathfrak{H}^1$ by addition. Easy verification shows that, when $s_1 > 1$, we have

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

i.e., the series in (66) are indeed q-extensions of the series in (4). In addition, ζ_q^* is a (q-parametric) homomorphism of the stuffle algebra $\mathfrak{H}_*^1$; for the proof of this fact, it is sufficient to apply specialization $t_n = q^n/(1-q^n)$ of Hoffman's homomorphism ϕ in Section 10. Thus

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

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

Another way to q-extend (not multiple) zeta values is proposed simultaneously and independently in the works [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 $\sigma_{s-1}(n) = \sum_{d|n} d^{s-1}$ denotes the sum of powers of the divisors; there 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. The q-zeta values (67) can be easily expressed 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, in general,

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

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

For $s \geqslant 2$ even, 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 circumstance allows 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 question of constructing a model of multiple q-zeta values, which contains the ordinary model (67), remains open. A natural requirement to such a model is possession of q-analogues of the shuffle and stuffle product-relations. We conclude by giving 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 will use the partial-fraction method, namely, the expansion

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

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

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

put then $u=q^m$, $v=q^n$, and sum over all positive integers m,n. Finally, we obtain the equality with the double sum

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

on the right-hand side, and

$$\begin{aligned}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}&\left(\frac{q^m(1+q^n)}{(1-q^n)^2(1-q^m)}-\frac{q^{n+m}(1+q^n)}{(1-q^n)(1-q^{n+m})^2}-\frac{q^{n+m}(1+q^n)}{(1-q^n)^2(1-q^{n+m})}\right)\\&=\sum_{n=1}^{\infty}\frac{1+q^n}{(1-q^n)^2}\sum_{m=1}^{\infty}\left(\frac{q^m}{1-q^m}-\frac{q^{n+m}}{1-q^{n+m}}\right)-\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{q^{n+m}(1+q^n)}{(1-q^n)(1-q^{n+m})^2}\end{aligned}$$

on the left-hand side. Moving the last sum from the right-hand side to the left-hand side, we deduce that

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

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

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

hence, taking $n_1 = n$, $n_2 = m$ on the right-hand side of (69) and $n_1 = l$, $n_2 = n$ in (70), we finally arrive at 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.

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

Received 30.10.2001
Translated by W. ZUDILIN

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX