MATH HONOURS: MULTIPLE ZETA VALUES

JONATHAN M. BORWEIN AND WADIM ZUDILIN

Preliminary contents

Resources and Links

[1] EZ-Face
[2] Michael Hoffman's site contains some basic information about the MZVs. Hoffman also has a comprehensive list of references on MZVs and related stuff
[3] Jonathan M. Borwein, David M. Bradley, David J. Broadhurst, and Petr Lisonek, "Special values of multidimensional polylogarithms," *Trans. Amer. Math. Soc.* **353** (2001), 907–941
[4] Wadim Zudilin, "Algebraic relations for multiple zeta values," *Russian Math. Surveys* **58**:1 (2003), 1–29
[5] Jonathan M. Borwein and David M. Bradley, "Thirty Two Goldbach Variations," *Int. J. Number Theory* **2**:1 (2006), 65–103
[6] David M. Bradley, "Multiple q-Zeta Values," *Journal of Algebra* **283**:2 (2005), 752–798

Date: 2nd semester 2011.

1. Riemann's zeta function, Bernoulli numbers, and multiple zeta values

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

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}\,. \tag{1.1}$$

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

The first breakthrough in this direction is due to Euler, who showed that $\zeta(2k)$ is always a rational multiple of π^{2k}, where

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

Although we do not follow Euler's original method (but its variation), the derivation is worth reproducing here.

The *Bernoulli numbers* $B_s \in \mathbb{Q}$, $s = 2, 3, 4, \dots$, are defined by the generating function

$$\frac{t}{e^t - 1} = 1 - \frac{t}{2} + \sum_{s=2}^{\infty} B_s \frac{t^s}{s!}\,.$$

Using

$$\frac{t}{e^t - 1} + \frac{t}{2} = \frac{t}{2} \cdot \frac{e^t + 1}{e^t - 1} = \frac{t}{2} \cdot \frac{e^{t/2} + e^{-t/2}}{e^{t/2} - e^{-t/2}}$$

and taking $t = 2iz$, where, of course, $i = \sqrt{-1}$, we obtain

$$z \frac{\cos z}{\sin z} = 1 + \sum_{s=2}^{\infty} B_s \frac{(2iz)^s}{s!}. \tag{1.2}$$

Since the function on the left-hand side of (1.2) is even, we arrive at the following result.

Lemma 1.1. *For each integer $k \ge 1$, we have $B_{2k+1} = 0$.*

Recall now the infinite product expansion of the sine function,

$$\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2 \pi^2}\right). \tag{1.3}$$

Computing its logarithmic derivative we find that

$$z \frac{\cos z}{\sin z} = 1 + 2 \sum_{n=1}^{\infty} \frac{z^2}{z^2 - n^2 \pi^2} = 1 - 2 \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \left(\frac{z^2}{n^2 \pi^2}\right)^k = 1 - 2 \sum_{k=1}^{\infty} \frac{z^{2k} \zeta(2k)}{\pi^{2k}}. \tag{1.4}$$

Comparing the right-hand sides of (1.2) and (1.4) results in the following statement.

Lemma 1.2. *For each integer $k \ge 1$, we have*

$$2\zeta(2k) = (-1)^{k+1} \frac{B_{2k}}{(2k)!} (2\pi)^{2k}.$$

In particular,

$$\zeta(2)=\frac{\pi^2}{2\cdot 3},\quad \zeta(4)=\frac{\pi^4}{2\cdot 3^2\cdot 5},\quad \zeta(6)=\frac{\pi^6}{3^3\cdot 5\cdot 7},\quad \zeta(8)=\frac{\pi^8}{2\cdot 3^3\cdot 5^2\cdot 7},$$
$$\zeta(10)=\frac{\pi^{10}}{3^5\cdot 5\cdot 7\cdot 11},\quad \zeta(12)=\frac{691\pi^{12}}{3^6\cdot 5^3\cdot 7^2\cdot 11\cdot 13},\quad \zeta(14)=\frac{2\pi^{14}}{3^6\cdot 5^2\cdot 7\cdot 11\cdot 13},$$

and so on.

Lemma 1.2 gives us the expression of the values of the zeta function at even integers in terms of π and the (rational) Bernoulli numbers. It implies the coincidence of the rings $\mathbb{Q}[\zeta(2),\zeta(4),\zeta(6),\zeta(8),\dots]$ and $\mathbb{Q}[\pi^2]$. Lindemann's theorem from 1882 asserts the transcendence of π, therefore we may conclude that each of the rings has transcendence degree 1 over the field of rational numbers.

Much less is known on the arithmetic nature of the values of the zeta function at odd integers $s=3,5,7,\dots$: in 1978, Apéry has proved the irrationality of the number $\zeta(3)$ and there are more recent but partial linear independence results of Rivoal and Zudilin. Rivoal's theorem settles the infiniteness of the set of irrational numbers among $\zeta(3),\zeta(5),\zeta(7),\dots$. Conjecturally, each of these numbers is transcendental, and the full answer on the above-stated question, about polynomial relations over $\mathbb{Q}$ for the values of series (1.1) with $s\ge 2$ 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. It seems to be unattainable by the present methods. In this course, 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.

Series (1.1) 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

$$\zeta(\boldsymbol{s})=\zeta(s_1,s_2,\dots,s_l):=\sum_{n_1>n_2>\dots>n_l\ge 1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}};\tag{1.5}$$

the corresponding multi-index $\boldsymbol{s}=(s_1,s_2,\dots,s_l)$ will be further regarded as *admissible*. The quantities (1.5) are called the *multiple zeta values* (and abbreviated MZVs), or the *multiple harmonic series*, or the *Euler sums*. The sums (1.5) for $l=2$ rise from Euler, who has obtained a family of identities connecting double and ordinary zeta values (which we discuss below). In particular, had Euler proved the identity

$$\zeta(2,1)=\zeta(3),\tag{1.6}$$

which was several times rediscovered after.

Exercise 1.1. Find your own (elementary) proof of (1.6).

The quantities (1.5) were introduced in the 1990s by Hoffman and, independently, by Zagier (with the opposite order of summation on the right-hand side of (1.5)). Those very first papers produced some $\mathbb{Q}$-linear and $\mathbb{Q}$-polynomial relations as well as indicated a series of conjectures (that has been partly resolved later) on the structure of algebraic relations for the family (1.5). Hoffman also suggests the alternative definition

$$\zeta^\star(\boldsymbol{s}) = \zeta^\star(s_1, s_2, \dots, s_l) := \sum_{n_1 \ge n_2 \ge \dots \ge n_l \ge 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}} \tag{1.7}$$

of the Euler sums, with non-strict inequalities in summation. These are nowadays known as *multiple zeta star values*.

Exercise 1.2. For any admissible index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$, we have the (dual) relations

$$\zeta^\star(\boldsymbol{s}) = \sum_{\boldsymbol{p}} \zeta(\boldsymbol{p}) \qquad \text{and} \qquad \zeta(\boldsymbol{s}) = \sum_{\boldsymbol{p}} (-1)^{\sigma(\boldsymbol{p})} \zeta^\star(\boldsymbol{p}),$$

where $\boldsymbol{p}$ runs through all indices of the form $(s_1 \circ s_2 \circ \dots \circ s_l)$ with '$\circ$' being either the symbol ',' or the sign '+', and the exponent $\sigma(\boldsymbol{p})$ denotes the number of signs '+' in $\boldsymbol{p}$. (The total number of such indices $\boldsymbol{p}$ is 2^{l-1}.)

Although all relations of series (1.7) may be rewritten with the help of Exercise 1.2 for series (1.5), several identities possess a compact form by means of (1.7); for example,

$$\zeta^\star(\{2\}^k, 1) := \zeta^\star(\underbrace{2, \dots, 2}_{k \text{ times}}, 1) = 2\zeta(2k+1), \qquad k = 1, 2, \dots . \tag{1.8}$$

To each number (1.5) (or (1.7)), assign the two characteristics: the *weight* (or *degree*) $|\boldsymbol{s}| := s_1 + s_2 + \dots + s_l$ and the *length* (or *depth*) $\ell(\boldsymbol{s}) := l$. Note that all relations known so far for the MZVs (1.5) and (1.7) are weight-preserving.

2. The partial-fraction method

This elementary analytic method is a powerful source of identities for multiple zeta values.

Theorem 2.1 (Hoffman's relations)**.** *For any admissible multi-index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$, the identity*

$$\begin{aligned} &\sum_{k=1}^{l} \zeta(s_1, \dots, s_{k-1}, s_k + 1, s_{k+1}, \dots, s_l) \\ &\qquad = \sum_{\substack{k=1 \\ s_k \ge 2}}^{l} \sum_{j=0}^{s_k - 2} \zeta(s_1, \dots, s_{k-1}, s_k - j, j+1, s_{k+1}, \dots, s_l) \end{aligned} \tag{2.1}$$

holds.

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

$$\begin{aligned}
&\sum_{n_k>n_{k+1}>\dots>n_l\ge1}\frac{1}{n_k^{s_k+1}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}+\sum_{n_k>m>n_{k+1}>\dots>n_l\ge1}\frac{1}{n_k^{s_k}mn_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\
&\quad=\sum_{n_k\ge m>n_{k+1}>\dots>n_l\ge1}\frac{1}{n_k^{s_k}mn_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\
&\quad=\sum_{n_k>n_{k+1}>\dots>n_l\ge1}\sum_{m=n_{k+1}+1}^{n_k}\frac{1}{mn_k^{s_k}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\,,
\end{aligned}$$

hence

$$\begin{aligned}
&\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\\
&\quad=\sum_{n_1>\dots>n_k>n_{k+1}>\dots>n_l\ge1}\frac{1}{n_1^{s_1}\cdots n_k^{s_k+1}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\
&\qquad+\sum_{n_1>\dots>n_k>m>n_{k+1}>\dots>n_l\ge1}\frac{1}{n_1^{s_1}\cdots n_k^{s_k}mn_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\
&\quad=\sum_{n_1>\dots>n_k>n_{k+1}>\dots>n_l\ge1}\sum_{m=n_{k+1}+1}^{n_k}\frac{1}{mn_1^{s_1}\cdots n_k^{s_k}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\
&\quad=\sum_{n_1>n_2>\dots>n_l\ge1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\sum_{m=n_{k+1}+1}^{n_k}\frac{1}{m}\,.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{k=1}^{l}\bigl(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\bigr)\\
&\quad=\sum_{n_1>n_2>\dots>n_l\ge1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\sum_{m=1}^{n_1}\frac{1}{m}\\
&\quad=\sum_{m_1,m_2,\dots,m_l\ge1}\frac{1}{m_1^{s_l}(m_1+m_2)^{s_{l-1}}\cdots(m_1+\dots+m_l)^{s_1}}\sum_{m=1}^{m_1+\dots+m_l}\frac{1}{m}\\
&\quad=\sum_{m_1,m_2,\dots,m_l\ge1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_l^{s_1}}\sum_{m_{l+1}\ge1}\biggl(\frac{1}{m_{l+1}}-\frac{1}{M_{l+1}}\biggr),
\end{aligned}\tag{2.2}$$

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):

$$\frac{1}{u(u+v)^s}=\frac{1}{v^su}-\sum_{j=0}^{s-1}\frac{1}{v^{j+1}(u+v)^{s-j}}\,,\qquad u,v\in\mathbb{R};\tag{2.3}$$

for the proof, it is sufficient 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 (2.3), 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 (2.2), we find that

$$\begin{aligned}
&\sum_{k=1}^{l}\bigl(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\bigr)\\
&\quad=\sum_{j=0}^{s_1-2}\sum_{m_1,m_2,\dots,m_{l+1}\ge1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_{l-1}^{s_2}M_l^{j+1}M_{l+1}^{s_1-j}}\\
&\qquad\quad+\sum_{m_1,m_2,\dots,m_{l+1}\ge1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_{l-1}^{s_2}m_{l+1}M_{l+1}^{s_1}}\\
&\quad=\sum_{j=0}^{s_1-2}\zeta(s_1-j,j+1,s_2,\dots,s_l)+\sum_{m_1,m_2,\dots,m_{l+1}\ge1}\frac{1}{M_1^{s_l}M_2^{s_{l-1}}\cdots M_{l-1}^{s_2}m_lM_{l+1}^{s_1}}
\end{aligned}\tag{2.4}$$

(in the last tuple sum we interchange the indices m_l and m_{l+1}). Using now identity (2.3) with $u = m_{k+1}$, $v = M_k = M_{k+1} - m_{k+1}$, and $s = s_{l+1-k}$, we find out 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

$$\begin{aligned}
&\sum_{m_1,m_2,\dots,m_{l+1}\ge1}\frac{1}{M_1^{s_l}\cdots M_k^{s_{l+1-k}}m_{k+1}M_{k+2}^{s_{l-k}}\cdots M_{l+1}^{s_1}}\\
&\quad=\sum_{j=0}^{s_{l+1-k}-1}\sum_{m_1,m_2,\dots,m_{l+1}\ge1}\frac{1}{M_1^{s_l}\cdots M_{k-1}^{s_{l+2-k}}M_k^{j+1}M_{k+1}^{s_{l+1-k}-j}M_{k+2}^{s_{l-k}}\cdots M_{l+1}^{s_1}}\\
&\qquad\quad+\sum_{m_1,m_2,\dots,m_{l+1}\ge1}\frac{1}{M_1^{s_l}\cdots M_{k-1}^{s_{l+2-k}}m_{k+1}M_{k+1}^{s_{l+1-k}}\cdots M_{l+1}^{s_1}}
\end{aligned}$$

$$= \sum_{j=0}^{s_{l+1-k}-1} \zeta(s_1, \dots, s_{l-k}, s_{l+1-k} - j, j+1, s_{l+2-k}, \dots, s_l) + \sum_{m_1, m_2, \dots, m_{l+1} \ge 1} \frac{1}{M_1^{s_l} \cdots M_{k-1}^{s_{l+2-k}} m_k M_{k+1}^{s_{l+1-k}} \cdots M_{l+1}^{s_1}}, \tag{2.5}$$

for $k = 1, 2, \dots, l-1$. Applying consequently, in inverse order (i.e., starting from $k = l-1$ and ending on $k = 1$), identities (2.5) for the tuple sum on the right-hand side of equality (2.4), we obtain

$$\begin{aligned}
&\sum_{k=1}^{l} \big(\zeta(s_1, \dots, s_{k-1}, s_k + 1, s_{k+1}, \dots, s_l) + \zeta(s_1, \dots, s_{k-1}, s_k, 1, s_{k+1}, \dots, s_l)\big) \\
&\quad = \sum_{j=0}^{s_1-2} \zeta(s_1 - j, j+1, s_2, \dots, s_l) \\
&\qquad + \sum_{k=1}^{l-1} \sum_{j=0}^{s_{l+1-k}-1} \zeta(s_1, \dots, s_{l-k}, s_{l+1-k} - j, j+1, s_{l+2-k}, \dots, s_l) \\
&\qquad + \sum_{m_1, m_2, \dots, m_{l+1} \ge 1} \frac{1}{m_1 M_2^{s_l} M_3^{s_{l-1}} \cdots M_{l+1}^{s_1}} \\
&\quad = \sum_{k=1}^{l} \sum_{j=0}^{s_k-2} \zeta(s_1, \dots, s_{k-1}, s_k - j, j+1, s_{k+1}, \dots, s_l) \\
&\qquad + \sum_{k=1}^{l} \zeta(s_1, \dots, s_{k-1}, s_k, 1, s_{k+1}, \dots, s_l).
\end{aligned} \tag{2.6}$$

Realizing all necessary reductions of the left-hand and right-hand sides of equality (2.6), we finally arrive at the desired identity (2.1). □

If $l = 1$, the statement of Theorem 2.1 can be written in the following form.

Theorem 2.2 (Euler's theorem)**.** *For any integer $s \ge 3$, the identity*

$$\zeta(s) = \sum_{j=2}^{s-1} \zeta(j, s-j) \tag{2.7}$$

holds.

Note also that, in the case $s = 3$, identity (2.7) becomes nothing else but relation (1.6).

Exercise 2.1 (Weighted analogue of Euler's formula). (a) Show that, for any $s \ge 3$,

$$\sum_{j=2}^{s-1} 2^j \zeta(j, s-j) = (s+1)\zeta(s). \tag{2.8}$$

(b) Generalize identity (2.8) to higher lengths.

Hint for part (a). Use the double-shuffle relations given below in the form (13.2), (13.3). □

In 2000, Hoffman and Ohno proved the following result also by means of the partial-fraction method. A somewhat simpler proof was later given by Ohno and Wakabayashi.

Theorem 2.3 (Cyclic sum theorem). *For any admissible multi-index* $\boldsymbol{s} = (s_1, s_2, \dots,$ $[2]s_l)$*, the identity*

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

holds.

Theorem 2.3 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 2.1, generalizes Euler's theorem.

Theorem 2.4 (Sum theorem). *For any integers* $s > 1$ *and* $l \ge 1$*, the identity*

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

holds.

Theorem 2.1 and 2.4 are particular instances of Ohno's relations, which will be discussed later (see Theorem 8.4 and its proof in Section 9).

3. Algebra of multiple zeta values

In this part, we expose the standard algebraic setup of the MZVs. It is expected that all known algebraic relations (i.e., numerical identities) over $\mathbb{Q}$ for the quantities (1.5) are produced by the so-called *double shuffle relations* which we describe below.

It is 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

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

Set

$$\zeta(x_{\boldsymbol{s}}) := \zeta(\boldsymbol{s}) \tag{3.1}$$

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 (3.1).

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

$$\mathbf{1} ⧢ w = w ⧢ \mathbf{1} = w, \qquad \mathbf{1} * w = w * \mathbf{1} = w \tag{3.2}$$

for any word w, and

$$x_j u ⧢ x_k v = x_j(u ⧢ x_k v) + x_k(x_j u ⧢ v), \tag{3.3}$$

$$y_j u * y_k v = y_j(u * y_k v) + y_k(y_j u * v) + y_{j+k}(u * v) \tag{3.4}$$

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 (3.2)–(3.4) 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

$$x_0^j * w = w * x_0^j = w x_0^j \tag{3.5}$$

for any word w and integer $j \ge 1$, to rule (3.4).

Exercise 3.1. Compute $x_0x_1 ⧢ x_0x_1$ and $x_0x_1 * x_0x_1$.

Exercise 3.2. Use the inductive argument to prove commutativity and associativity of each of the multiplications.

The corresponding algebras $\mathfrak{H}_{⧢} := (\mathfrak{H}, ⧢)$, $\mathfrak{H}^1_* := (\mathfrak{H}^1, *)$ (and also $\mathfrak{H}_* := (\mathfrak{H}, *)$) are examples of so-called *Hopf algebras*.

The following two statements motivate consideration of the introduced multiplications $⧢$ and $*$.

Theorem 3.1. *The map ζ is a homomorphism of the shuffle algebra $\mathfrak{H}^0_{⧢} := (\mathfrak{H}^0, ⧢)$ into $\mathbb{R}$, i.e.,*

$$\zeta(w_1 ⧢ w_2) = \zeta(w_1)\zeta(w_2) \quad \textit{for all} \quad w_1, w_2 \in \mathfrak{H}^0. \tag{3.6}$$

Theorem 3.2. *The map ζ is a homomorphism of the stuffle algebra $\mathfrak{H}^0_* := (\mathfrak{H}^0, *)$ into $\mathbb{R}$, i.e.,*

$$\zeta(w_1 * w_2) = \zeta(w_1)\zeta(w_2) \quad \textit{for all} \quad w_1, w_2 \in \mathfrak{H}^0. \tag{3.7}$$

Later we give detailed proofs of the two theorems using the differential-difference origin of the multiplications $⧢$ and $*$ in suitable functional models of the algebras $\mathfrak{H}_{⧢}$ and $\mathfrak{H}^0_*$.

One more family of identities is given by the following statement whose proof is deduced later.

Theorem 3.3. *The map ζ satisfies the relations*

$$\zeta(x_1 ш w - x_1 * w) = 0 \quad \textit{for all} \quad w \in \mathfrak{H}^0 \tag{3.8}$$

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

All (rigorously and experimentally) known identities for the multiple zeta values (are expected to) "follow" from identities (3.6)–(3.8) — the double shuffle relations. Therefore the following conjecture looks rather truthful.

Conjecture 2. *All linear relations over $\mathbb{Q}$ of multiple zeta values are generated by identities (3.6)–(3.8); equivalently,*

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

4. Shuffle algebra of generalized polylogarithms. Duality theorem

In order to prove shuffle relations (3.6) for multiple zeta values, let us define the *generalized polylogaithms*

$$\operatorname{Li}_{\boldsymbol{s}}(z) := \sum_{n_1>n_2>\cdots>n_l\ge1} \frac{z^{n_1}}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}\,, \qquad |z|<1, \tag{4.1}$$

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

$$\operatorname{Li}_{\boldsymbol{s}}(1) = \zeta(\boldsymbol{s}), \qquad \boldsymbol{s}\in\mathbb{Z}^l, \quad s_1\ge2,\ s_2\ge1,\ \dots,\ s_l\ge1. \tag{4.2}$$

Taking, as before for multiple zeta values,

$$\operatorname{Li}_{x_{\boldsymbol{s}}}(z) := \operatorname{Li}_{\boldsymbol{s}}(z), \qquad \operatorname{Li}_{\boldsymbol{1}}(z) := 1, \tag{4.3}$$

let us extend action of the map $\operatorname{Li}\colon w\mapsto \operatorname{Li}_w(z)$ by linearity on the graded algebra $\mathfrak{H}^1$ (not $\mathfrak{H}$, since multi-indices are coded by words in $\mathfrak{H}^1$).

Lemma 4.1. *Let $w\in\mathfrak{H}^1$ be an arbitrary non-empty word and x_j the first letter in its record (that is, $w=x_ju$ for some word $u\in\mathfrak{H}^1$). Then*

$$\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_w(z) = \frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_ju}(z) = \omega_j(z)\operatorname{Li}_u(z), \tag{4.4}$$

where

$$\omega_j(z) = \omega_{x_j}(z) := \begin{cases} \dfrac{1}{z} & \text{if } x_j = x_0, \\ \dfrac{1}{1-z} & \text{if } x_j = x_1. \end{cases} \tag{4.5}$$

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

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_w(z) = \frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{\boldsymbol{s}}(z) &= \frac{\mathrm{d}}{\mathrm{d}z}\sum_{n_1>n_2>\cdots>n_l\ge1} \frac{z^{n_1}}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}\,, \\ &= \sum_{n_1>n_2>\cdots>n_l\ge1} \frac{z^{n_1-1}}{n_1^{s_1-1} n_2^{s_2} \cdots n_l^{s_l}}\,. \end{aligned}$$

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

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_0u}(z) &= \frac{1}{z}\sum_{n_1>n_2>\cdots>n_l\ge1} \frac{z^{n_1}}{n_1^{s_1-1} n_2^{s_2} \cdots n_l^{s_l}} \\ &= \frac{1}{z}\operatorname{Li}_{s_1-1,s_2,\dots,s_l}(z) = \frac{1}{z}\operatorname{Li}_u(z) \end{aligned}$$

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

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_1u}(z) &= \sum_{n_1>n_2>\cdots>n_l\ge1} \frac{z^{n_1-1}}{n_2^{s_2} \cdots n_l^{s_l}} = \sum_{n_2>\cdots>n_l\ge1} \frac{1}{n_2^{s_2} \cdots n_l^{s_l}} \sum_{n_1=n_2+1}^{\infty} z^{n_1-1} \\ &= \frac{1}{1-z}\sum_{n_2>\cdots>n_l\ge1} \frac{z^{n_2}}{n_2^{s_2} \cdots n_l^{s_l}} = \frac{1}{1-z}\operatorname{Li}_{s_2,\dots,s_l}(z) = \frac{1}{1-z}\operatorname{Li}_u(z), \end{aligned}$$

and the result follows. ∎

Lemma 4.1 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 $\mathrm{Li}_{\mathbf{1}}(z) = 1$ and

$$\mathrm{Li}_w(z) = \begin{cases} \dfrac{\log^s z}{s!} & \text{if } w = x_0^s \text{ for some } s \geq 1, \\ \displaystyle\int_0^z \omega_j(z)\,\mathrm{Li}_u(z)\,\mathrm{d}z & \text{if } w = x_j u \text{ contains the letter } x_1. \end{cases} \tag{4.6}$$

Evidently, Lemma 4.1 remains true for this extended version (4.6) of the polylogarithms (the fact yields coincidence of the newly-defined polylogarithms with the 'old' ones (4.3) for words w in $\mathfrak{H}^1$); in addition,

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

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

Lemma 4.2. *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

$$\mathrm{Li}_{w_1 ш w_2}(z) = \mathrm{Li}_{w_1}(z)\,\mathrm{Li}_{w_2}(z) \quad \text{for all} \quad w_1, w_2 \in \mathfrak{H}; \tag{4.7}$$

it is sufficient to do this job for *words* $w_1, w_2 \in \mathfrak{H}$. We will prove equality (4.7) by induction on the quantity $|w_1| + |w_2|$. If $w_1 = \mathbf{1}$ or $w_2 = \mathbf{1}$, relation (4.7) becomes tautological by (3.2). Otherwise, $w_1 = x_j u$ and $w_2 = x_k v$, hence by Lemma 4.1 and the inductive hypothesis we have

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

Thus

$$\mathrm{Li}_{w_1}(z)\,\mathrm{Li}_{w_2}(z) = \mathrm{Li}_{w_1 ш w_2}(z) + C, \tag{4.8}$$

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 (4.8) becomes the required relation (4.7), and the lemma follows. ∎

Proof of Theorem 3.1. Theorem 3.1 follows from Lemma 4.2 and relations (4.2). □

Explicit computation of the monodromy group for the system of differential equations (4.4) allows to Minh, Petitot, and van der Hoeven to prove that the homomorphism $w \mapsto \mathrm{Li}_w(z)$ of the shuffle algebra $\mathfrak{H}_{⧢}$ over $\mathbb{C}$ is bijective, i.e., all $\mathbb{C}$-algebraic relations for generalized polylogarithms are originated by shuffle relations (4.7) only; in particular, generalized polylogarithms are linearly independent over $\mathbb{C}$. A much simpler proof of the linear independence of functions (4.1), as a consequence of elegant identities for the functions, is due to Ulanskiĭ. On the other hand, Sorokin proved the linear independence result for the *values* of generalized polylogarithms (for rational z from a neighbourhood of the origin); this also implies the independence of the functions themselves.

By Lemma 4.1, the following integral representation is valid for the word $w = x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k} \in \mathfrak{H}^1$:

$$\begin{aligned}\mathrm{Li}_w(z) &= \int_0^z \omega_{\varepsilon_1}(z_1)\,\mathrm{d}z_1 \int_0^{z_1} \omega_{\varepsilon_2}(z_2)\,\mathrm{d}z_2 \dots \int_0^{z_{k-1}} \omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_k \\ &= \int\cdots\int_{z>z_1>z_2>\cdots>z_{k-1}>z_k>0} \omega_{\varepsilon_1}(z_1)\omega_{\varepsilon_2}(z_2)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\,\mathrm{d}z_2\cdots\mathrm{d}z_k \qquad (4.9)\end{aligned}$$

if $0 < z < 1$. When $x_{\varepsilon_1} \ne x_1$, i.e., $w \in \mathfrak{H}^0$, the integral in (4.9) converges in the region $0 < z \le 1$, hence, in accordance with (4.2), we reduce representation for the multiple zeta values

$$\zeta(w) = \int\cdots\int_{1>z_1>\cdots>z_k>0} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k \qquad (4.10)$$

in a form of *Chen's iterated integrals*. The following result is evident application of the integral representation (4.10).

Remark. There is a simple mnemonic way to write down the integral representation (4.10):

$$\zeta(x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k}) = \int_0^1 x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k}, \qquad (4.11)$$

where (with a definite ambiguity!) x_0 and x_1 denote the corresponding differential forms $\omega_0(z)\,\mathrm{d}z$ and $\omega_1(z)\,\mathrm{d}z$.

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, $\tau(x_0^2x_1x_0x_1) = x_0x_1x_0x_1^2$. 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 4.1 (Duality theorem). *For any word $w \in \mathfrak{H}^0$, the relation*

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

holds.

Proof. To prove the theorem, it is sufficient to do the change of variable $z_1' = 1 - z_k$, $z_2' = 1 - z_{k-1}$, $\dots$, $z_k' = 1 - z_1$, and apply relations $\omega_0(z) = \omega_1(1-z)$ followed from (4.5). ☐

As the simplest consequence of Theorem 4.1, notice (again) identity (1.6), which follows for the word $w = x_0^2 x_1$, as well as the general identity

$$\zeta(n+2) = \zeta(2, \underbrace{1, \dots, 1}_{n \text{ times}}), \qquad n = 1, 2, \dots, \tag{4.12}$$

for the words $w = x_0^{n+1} x_1$. In the remaining part of the course, we will keep the compact notation $\{\boldsymbol{s}\}^n$ for the multi-index consisting of n copies of the multi-index s. Then (4.12) can be written as $\zeta(n+2) = \zeta(2, \{1\}^n)$, and the notation has been already used in (1.8).

Exercise 4.1. Show that

$$\zeta(\{2,1\}^n) = \zeta(\{3\}^n), \qquad n = 1, 2, \dots. \tag{4.13}$$

The iterated integral representations of MZVs and generalized polylogarithms motivate considering a slightly general than (1.5) version of MZVs, namely, the (*alternating* or *alternative*) *Euler sums*

$$\zeta(s_1, \dots, s_l; \sigma_1, \dots, \sigma_l) := \sum_{n_1 > n_2 > \dots > n_l \ge 1} \frac{\sigma_1^{n_1} \sigma_2^{n_2} \cdots \sigma_l^{n_l}}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}, \tag{4.14}$$

where $\sigma_j \in \{\pm 1\}$ are 'signs' and s_j, as before, are positive integers. It is customary to shortcut the notation by combining strings of exponents and signs and replacing s_j by $\overline{s}_j$ in the multi-index string if and only if the corresponding $\sigma_j = -1$. Thus, $\zeta(\overline{1}) = \zeta(1; -1) = \mathrm{Li}_1(-1) = -\log 2$ and $\zeta(\overline{2}, 1) = \zeta(2, 1; -1, 1)$.

Exercise 4.2. Show that

$$\text{(a) } \zeta(\overline{1}, \{1\}^{n-1}) = \mathrm{Li}_{\{1\}^n}(-1) = \frac{(-\log 2)^n}{n!}, \quad n = 1, 2, \dots; \qquad \text{(b) } \zeta(\overline{2}, 1) = \frac{\zeta(3)}{8}.$$

In what follows we will see that the standard algebraic setup for the Euler sums is an extension of the non-commutative algebra $\mathbb{Q}\langle x_0, x_1 \rangle$ to $\mathbb{Q}\langle x_0, x_1, \overline{x}_1 \rangle$, and generalization of the integral in (4.11) by allowing the three differential forms

$$\begin{gathered} x_0 \mapsto a = \omega_0(z)\,\mathrm{d}z = \frac{\mathrm{d}z}{z}, \quad x_1 \mapsto b = \omega_1(z)\,\mathrm{d}z = \frac{\mathrm{d}z}{1-z}, \\ \text{and} \quad \overline{x}_1 \mapsto c = \overline{\omega}_1(z)\,\mathrm{d}z = \frac{-\mathrm{d}z}{1+z}. \end{gathered} \tag{4.15}$$

5. The generating-function method

Another application of differential equations for generalized polylogarithms, deduced in Lemma 4.1, is the *generating-function method*.

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

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

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 4.1, 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 *generalized hypergeometric series*.

In order to show any reasonable result for MZVs using generating functions, we have to be familiar Euler's *gamma function* $\Gamma(s)$ and its properties as well as with the Euler–Gauss *hypergeometric function* (or *hypergeometric series*)

$$\begin{aligned} F(a,b;c;z) &= {}_2F_1\left(\begin{matrix} a,\ b \\ c \end{matrix} \,\middle|\, z\right) \\ &= \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n!(c)_n} z^n \\ &= 1 + \frac{a\cdot b}{1\cdot c} z + \frac{a(a+1)\cdot b(b+1)}{1\cdot 2\cdot c(c+1)} z^2 \\ &\qquad + \frac{a(a+1)(a+2)\cdot b(b+1)(b+2)}{1\cdot 2\cdot 3\cdot c(c+1)(c+2)} z^3 + \cdots , \end{aligned}$$

where

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = \begin{cases} 1 & \text{if } n = 0, \\ a(a+1)\cdots(a+n-1) & \text{if } n \ge 1, \end{cases}$$

abbreviates the product of n consecutive numbers starting from a, the so-called *Pochhammer's symbol* or *shifted factorial* (because in case $a = 1$ we clearly get the standard factorial $(1)_n = n!$).

The convergence of the series can be determined by the ratio test. If we denote

$$a_n = \frac{(a)_n (b)_n}{n!(c)_n}$$

the nth coefficient of the hypergeometric series $F(a,b;c;z)$, then

$$\frac{a_{n+1}}{a_n} = \frac{(a+n)(b+n)}{(1+n)(c+n)} \to 1 \quad \text{as } n \to \infty,$$

hence the series converges in the unit disc, $|z| < 1$. In several cases, depending on the parameters a, b, c, the series may converge on the boundary of the disc, for example, at $z = 1$. We will examine the latter situation.

Because of the relation

$$(1+n)(c+n)\cdot a_{n+1} = (a+n)(b+n)\cdot a_n \quad \text{for } n = 0, 1, 2, \ldots,$$

we have

$$\begin{aligned}
z\left(z\frac{\mathrm{d}}{\mathrm{d}z}+a\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}+b\right)F(a,b;c;z) &= z\left(z\frac{\mathrm{d}}{\mathrm{d}z}+a\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}+b\right)\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{n!(c)_n}z^n \\
&= z\sum_{n=0}^{\infty}\frac{(a)_n(a+n)\cdot(b)_n(b+n)}{n!(c)_n}z^n = \sum_{n=0}^{\infty}\frac{(a)_{n+1}(b)_{n+1}}{n!(c)_n}z^{n+1} \\
&= \sum_{n=1}^{\infty}\frac{(a)_n(b)_n}{(n-1)!(c)_{n-1}}z^n = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n\cdot n(c+n)}{n!(c)_n}z^n \\
&= \left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}+c-1\right)\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{n!(c)_n}z^n \\
&= \left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}+c-1\right)F(a,b;c;z).
\end{aligned}$$

Lemma 5.1. *The hypergeometric function $F(a,b;c;z)$ satisfies the differential equation*

$$\left(z\left(z\frac{\mathrm{d}}{\mathrm{d}z}+a\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}+b\right)-\left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)\left(z\frac{\mathrm{d}}{\mathrm{d}z}+c-1\right)\right)y=0;$$

in equivalent form,

$$z(1-z)\frac{\mathrm{d}^2y}{\mathrm{d}z^2}+(c-(a+b+1)z)\frac{\mathrm{d}y}{\mathrm{d}z}-aby=0.$$

Lemma 5.2 (Pochhammer's integral)**.** *If $\operatorname{Re}c>\operatorname{Re}b>0$ and $|z|<1$, then*

$$F(a,b;c;z)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 x^{b-1}(1-x)^{c-b-1}(1-zx)^{-a}\,\mathrm{d}x.$$

Note that for $a=0$ the integral on the right-hand side reduces to Euler's integral of the first kind $\mathrm{B}(b,c-b)$.

Proof. The conditions $\operatorname{Re}b>0$ and $\operatorname{Re}(c-b)>0$ ensure convergence of the integral

$$I(a,b;c;z)=\int_0^1 x^{b-1}(1-x)^{c-b-1}(1-zx)^{-a}\,\mathrm{d}x.$$

Furthermore, for $|z|<1$,

$$(1-zx)^{-a}=\sum_{n=0}^{\infty}\frac{(a)_n}{n!}z^nx^n.$$

Therefore,

$$\begin{aligned} I(a,b;c;z) &= \int_0^1 \sum_{n=0}^{\infty} \frac{(a)_n z^n}{n!} x^{b+n-1}(1-x)^{c-b-1}\,\mathrm{d}x \\ &= \sum_{n=0}^{\infty} \frac{(a)_n z^n}{n!} \int_0^1 x^{b+n-1}(1-x)^{c-b-1}\,\mathrm{d}x \\ &= \sum_{n=0}^{\infty} \frac{(a)_n z^n}{n!} \frac{\Gamma(b+n)\Gamma(c-b)}{\Gamma(c+n)} \\ &= \frac{\Gamma(b)\Gamma(c-b)}{\Gamma(c)} F(a,b;c;z), \end{aligned}$$

and the result follows. □

As a corollary of this result and Abel's theorem on power series, we deduce

Lemma 5.3 (Gauss' summation formula)**.** *If* $\operatorname{Re} c > \operatorname{Re}(a+b)$, *then*

$$F(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}.$$

Proof. The result follows, whenever $\operatorname{Re} c > \operatorname{Re} b > 0$ and $\operatorname{Re}(c-a-b) > 0$, by taking the limit $z \to 1$ in Lemma 5.2 and using the beta integral evaluation of the resulted definite integral:

$$\begin{aligned} F(a,b;c;1) &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 x^{b-1}(1-x)^{c-b-a-1}\,\mathrm{d}x \\ &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \cdot \frac{\Gamma(b)\Gamma(c-a-b)}{\Gamma(c-a)}. \end{aligned}$$

To get rid of restriction $\operatorname{Re} c > \operatorname{Re} b > 0$, note that the formula is valid for $\operatorname{Re}(c-a-b) > 0$ and use the theory of analytic continuation. □

Remark. When a is a negative integer $-m$, the theorem becomes

$$\sum_{n=0}^{m} \binom{m}{n} \frac{(b)_n}{(c)_n}(-1)^n = F(-m,b;c;1) = \frac{(c-b)_m}{(c)_m},$$

the result known as the Chu–Vandermonde summation. With the help of the latter formula one can show the following binomial evaluation:

$$\sum_{n=0}^{m} \binom{p}{n}\binom{q}{m-n} = \binom{p+q}{m}.$$

Exercise 5.1. (a) Show that

$$F(a,b;1+b-a;-1) = \frac{\Gamma(1+b-a)\Gamma(1+\frac{1}{2}b)}{\Gamma(1+b)\Gamma(1+\frac{1}{2}b-a)}.$$

(b) Give a gamma-function evaluation of the hypergeometric series $F(a,1-a;c;\frac{1}{2})$.

It is now a good time to go back to the MZV story.

Lemma 5.4. *The following equality holds:*

$$L_{(3,1)}(z,t) = F\big(\tfrac{1}{2}(1+i)t, -\tfrac{1}{2}(1+i)t; 1; z\big) \cdot F\big(\tfrac{1}{2}(1-i)t, -\tfrac{1}{2}(1-i)t; 1; z\big), \qquad (5.1)$$

where $F(a,b;c;z)$ denotes the hypergeometric function and $i=\sqrt{-1}$.

Proof. Routine verification (with a help of Lemma 4.1 for the left-hand side) shows that the both sides of the required equality are annihilated by action of the differential operator

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

in addition, the first terms in z-expansions of the both sides in (5.1) coincide:

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

Thus the statement of the lemma follows. □

Exercise 5.2. Fill in the missing details.

The following result was conjectured by Zagier in his pioneering talk at the European Congress of Mathematics in 1994. The proof was given some years later in joint work of Borwein, Bradley, Broadhurst and Lisoněk.

Theorem 5.1. *For any integer $n \ge 1$, the identity*

$$\zeta(\{3,1\}^n) = \frac{2\pi^{4n}}{(4n+2)!} \qquad (5.2)$$

holds.

Proof. By Lemma 5.3 (Gauss' summation formula),

$$F(a,-a;1;1) = \frac{1}{\Gamma(1-a)\Gamma(1+a)} = \frac{\sin \pi a}{\pi a}, \qquad (5.3)$$

substituting $z=1$ into equality (5.1) yields

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

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

Identity (5.2) is not the unique example of application of generating functions. We present more identities of Borwein, Bradley and Broadhurst, similar to (5.2), for which the above method is also effective:

$$\zeta(\{2\}^n)=\frac{2(2\pi)^{2n}}{(2n+1)!}\left(\frac{1}{2}\right)^{2n+1}, \quad \zeta(\{4\}^n)=\frac{4(2\pi)^{4n}}{(4n+2)!}\left(\frac{1}{2}\right)^{2n+1},$$
$$\zeta(\{6\}^n)=\frac{6(2\pi)^{6n}}{(6n+3)!}, \quad \zeta(\{8\}^n)=\frac{8(2\pi)^{8n}}{(8n+4)!}\left(\left(1+\frac{1}{\sqrt{2}}\right)^{4n+2}+\left(1-\frac{1}{\sqrt{2}}\right)^{4n+2}\right),$$
$$\zeta(\{10\}^n)=\frac{10(2\pi)^{10n}}{(10n+5)!}\left(1+\left(\frac{1+\sqrt{5}}{2}\right)^{10n+5}+\left(\frac{1-\sqrt{5}}{2}\right)^{10n+5}\right), \tag{5.4}$$

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

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

may be derived by the generating-function method (as well as by straightforward application of Theorem 4.1).

Exercise 5.3. Prove (some) identities in (5.4).

Exercise 5.4. Show that

$$\zeta(\{3,1\}^n)=\frac{1}{2n+1}\,\zeta(\{2\}^{2n}).$$

The family of identities

$$\zeta(\{\overline{2},1\}^n)=\frac{1}{8^n}\,\zeta(\{3\}^n), \qquad n=1,2,\dots, \tag{5.5}$$

conjectured by Borwein, Bradley and Broadhurst in 1996, generalises Exercises 4.1 and 4.2 (b) and looks very similar to that in Theorem 5.1. It was proven only recently by Zhao using the standard relations for the alternating Euler sums; a proof by generating functions is still wanted.

Another family

$$\zeta(\{2\}^{n+3})+2\zeta(\{2\}^n,3,3)=\zeta(2,1,\{2\}^n,3), \qquad n=1,2,\dots, \tag{5.6}$$

conjectured by Hoffman, is shown to be true by Vermaseren for $n\le 8$. The general case remains a conjecture.

An example of other-type generating functions relates to generalization of Apéry's identity

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

namely, the following expansions are valid:

$$\begin{aligned}\sum_{n=0}^{\infty}\zeta(2n+3)t^{2n} &= \sum_{k=1}^{\infty}\frac{1}{k^3(1-t^2/k^2)}\\ &= \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}\biggl(\frac{1}{2}+\frac{2}{1-t^2/k^2}\biggr)\prod_{l=1}^{k-1}\biggl(1-\frac{t^2}{l^2}\biggr),\\ \sum_{n=0}^{\infty}\zeta(4n+3)t^{4n} &= \sum_{k=1}^{\infty}\frac{1}{k^3(1-t^4/k^4)} = \frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}\,\frac{1}{1-t^4/k^4}\prod_{l=1}^{k-1}\frac{1+4t^4/l^4}{1-t^4/l^4}.\end{aligned} \tag{5.7}$$

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 (5.3) in deducing Theorem 5.1.

Identities (5.7) can be used in fast computation of the Riemann zeta function at odd integers. To see that note that they both come as special cases ($s = 0$ and $t = 0$) of the bivariate generating function identity

$$\begin{aligned}&\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\binom{n+m}{n}\zeta(2n+4m+3)s^{2n}t^{4m}\\ &\quad= \sum_{k=1}^{\infty}\frac{k}{k^4-s^2k^2-t^4} = \frac{1}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k\binom{2k}{k}}\,\frac{5k^2-s^2}{k^4-s^2k^2-t^4}\prod_{m=1}^{k-1}\frac{(m^2-s^2)^2+4t^4}{m^4-s^2m^2-t^4},\end{aligned}$$

which was conjectured by Cohen and proved independently by Bradley and Rivoal. Recently, applying the so-called Markov–WZ algorithm, Hessami Pilehroods gave a different identity

$$\sum_{k=1}^{\infty}\frac{k}{k^4-s^2k^2-t^4} = \frac{1}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}r(n)}{n\binom{2n}{n}}\,\frac{\prod_{m=1}^{n-1}((m^2-s^2)^2+4t^4)}{\prod_{m=n}^{2n}(m^4-s^2m^2-t^4)}, \tag{5.8}$$

where

$$r(n) = 205n^6-160n^5+(32-62s^2)n^4+40s^2n^3+(s^4-8s^2-25t^4)n^2+10t^4n+t^4(s^2-2).$$

Formula (5.8) generates (Apéry-like) series for all $\zeta(2n+4m+3)$, $n, m \ge 0$, convergent at the geometric rate with ratio 2^{-10}. For example, if $s = t = 0$ one gets the Amdeberhan–Zeilberger series for $\zeta(3)$,

$$\zeta(3) = \frac{1}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}(205n^2-160n+32)}{n^5\binom{2n}{n}^5}.$$

Exercise 5.5. Using (5.8), find fast converging series for $\zeta(5)$ and $\zeta(7)$.

6. Quasi-shuffle products

The following construction, due to Hoffman, allows one 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:

$$\mathbf{1} \circ w = w \circ \mathbf{1} = w \tag{6.1}$$

for any word w, and

$$a_j u \circ a_k v = a_j(u \circ a_k v) + a_k(a_j u \circ v) + [a_j, a_k](u \circ v) \tag{6.2}$$

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

$$[\,\cdot\,,\,\cdot\,]\colon \bar{A} \times \bar{A} \to \bar{A} \tag{6.3}$$

($\bar{A} := A \cup \{\mathbf{0}\}$) 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 (the result of Hoffman).

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 $\mathfrak{A}_\circ = \mathfrak{H}_{⧢}$ 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

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

On the algebra $\mathfrak{A}$ with the given functional (6.3), define the dual product $\bar{\circ}$ by the rules

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

in place of (6.1) and (6.2), respectively. Then $\mathfrak{A}_{\bar{\circ}} := (\mathfrak{A}, \bar{\circ})$ is a (commutative, if property (S3) holds) graded $\mathcal{K}$-algebra as well.

Theorem 6.1. *The algebras $\mathfrak{A}_\circ$ and $\mathfrak{A}_{\bar{\circ}}$ coincide.*

Proof. It is sufficient to prove the relation

$$w_1 \circ w_2 = w_1 \bar{\circ} w_2 \tag{6.4}$$

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 (6.4) 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

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

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

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

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

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

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 (6.3) satisfies properties (S0)–(S2), to be rather important.

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

Lemma 6.1. *For any letter $a \in A$ and any words $u, v \in \mathfrak{A}$, the following identity holds:*

$$a \circ uv - (a \circ u)v = u(a \circ v - av). \tag{6.5}$$

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 (6.5) is evident. Otherwise, write the word u as $u = a_1 u_1$, where $a_1 \in A$ and the word u_1 consists of less number of letters, hence the identity

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

holds. Then

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

which is the desired result. □

7. 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 $\mathfrak{H}_{⧢}$, since rule (3.4) has no differential interpretation as (3.3). Therefore we shall use a difference interpretation of rule (3.4), namely, the (simplest) difference operator

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

It can be easily verified that

$$D\big(f_1(t)f_2(t)\big) = Df_1(t) \cdot f_2(t) + f_1(t) \cdot Df_2(t) + Df_1(t) \cdot Df_2(t), \tag{7.1}$$

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. The operator D can be 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 an entire function. Exponentiating derivations (in word algebras), in connection with generalization of Theorem 3.1, is discussed below.

A natural analogy with Lemmas 4.1 and 4.2, by (3.4) and (7.1) provides the existence of functions $\omega_j(t)$ satisfying the properties

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

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

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

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

$$D\,\mathrm{Ri}_{uy_j}(t) = \frac{1}{t^j} \mathrm{Ri}_u(t) \tag{7.2}$$

that, in some sense, is a discrete analogue of formula (4.4).

Lemma 7.1. *The following identity holds:*

$$\mathrm{Ri}_{\boldsymbol{s}}(t) = \sum_{n_1 > \dots > n_{l-1} > n_l \ge 1} \frac{1}{(t+n_1)^{s_1} \cdots (t+n_{l-1})^{s_{l-1}} (t+n_l)^{s_l}} ; \tag{7.3}$$

in particular,

$$\mathrm{Ri}_{\boldsymbol{s}}(0) = \zeta(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \ge 2,\ s_2 \ge 1,\ \dots,\ s_l \ge 1, \tag{7.4}$$

$$\lim_{t\to+\infty} \mathrm{Ri}_{\boldsymbol{s}}(t) = 0, \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \ge 2,\ s_2 \ge 1,\ \dots,\ s_l \ge 1. \tag{7.5}$$

Proof. By definition, we find that

$$\begin{aligned}
\mathrm{Ri}_{\boldsymbol{s}}(t) &= I\biggl(\frac{1}{t^{s_l}}\,\mathrm{Ri}_{s_1,\dots,s_{l-1}}(t)\biggr) \\
&= I\biggl(\frac{1}{t^{s_l}}\sum_{n_1>\dots>n_{l-1}\ge1}\frac{1}{(t+n_1)^{s_1}\cdots(t+n_{l-1})^{s_{l-1}}}\biggr) \\
&= \sum_{n=1}^{\infty}\frac{1}{(t+n)^{s_l}}\sum_{n_1>\dots>n_{l-1}\ge1}\frac{1}{(t+n_1+n)^{s_1}\cdots(t+n_{l-1}+n)^{s_{l-1}}} \\
&= \sum_{n_1'>\dots>n_{l-1}'>n\ge1}\frac{1}{(t+n_1')^{s_1}\cdots(t+n_{l-1}')^{s_{l-1}}(t+n)^{s_l}}\,,
\end{aligned}$$

and this implies the required formula (7.3). □

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

$$\begin{gathered}
\mathbf{1}\bar{*}w = w\bar{*}\mathbf{1} = w, \\
uy_j\bar{*}vy_k = (u\bar{*}vy_k)y_j + (uy_j\bar{*}v)y_k + (u\bar{*}v)y_{j+k}
\end{gathered} \tag{7.6}$$

instead of (3.2) and (3.4).

Lemma 7.2. *The map $w \mapsto \mathrm{Ri}_w(t)$ is a homomorphism of the algebra $(\mathfrak{H}^0, \bar{*})$ into $C([0,+\infty);\mathbb{R})$.*

Proof. It is sufficient to verify the relations

$$\mathrm{Ri}_{w_1\bar{*}w_2}(t) = \mathrm{Ri}_{w_1}(t)\,\mathrm{Ri}_{w_2}(t) \quad \text{for all} \quad w_1, w_2 \in \mathfrak{H}^0; \tag{7.7}$$

without loss of generality we may assume that w_1, w_2 are polynomials of the algebra $\mathfrak{H}^0$. We will prove relation (7.7) by induction on the quantity $\ell(w_1) + \ell(w_2)$; if $w_1 = \mathbf{1}$ or $w_2 = \mathbf{1}$, then validity of (7.7) is evident due to (7.6). Otherwise, write $w_1 = uy_j$, $w_2 = vy_k$ and apply formulae (7.1), (7.2) 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

$$\mathrm{Ri}_{w_1}(t)\,\mathrm{Ri}_{w_2}(t) = \mathrm{Ri}_{w_1\bar{*}w_2}(t) + C, \tag{7.8}$$

and letting t tend to $+\infty$, by (7.5) we obtain $C = 0$. Thus, relation (7.8) becomes the required equality (7.7), and the lemma follows. □

Proof of Theorem 3.2. By (7.4), Theorem 3.2 follows from Lemma 7.2 and Theorem 6.1. □

Another way to prove Theorem 3.2 (and Lemma 7.2 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 $t_1, t_2, \dots$. Namely, the $\mathbb{Q}$-linear map ϕ is defined by setting $\phi(1) := 1$ and

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

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 $n_1' > n_2' > \dots > n_l'$. By the above means the homomorphism $w \mapsto \mathrm{Ri}_w(t)$ in Lemma 7.2 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. Slightly modifying the original scheme of Cartier, we will expose main ideas of the approach for proving Euler's identity

$$\zeta(s_1)\zeta(s_2) = \zeta(s_1 + s_2) + \zeta(s_1, s_2) + \zeta(s_2, s_1), \qquad s_1 \ge 2, \ s_2 \ge 2, \tag{7.9}$$

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

$$\zeta(\boldsymbol{s}) = \underset{[0,1]^{|\boldsymbol{s}|}}{\int \cdots \int} \prod_{j=1}^{l-1} \frac{t_1 t_2 \cdots t_{s_1 + \dots + s_j}}{1 - t_1 t_2 \cdots t_{s_1 + \dots + s_j}} \cdot \frac{\mathrm{d}t_1 \, \mathrm{d}t_2 \cdots \mathrm{d}t_{|\boldsymbol{s}|}}{1 - t_1 t_2 \cdots t_{s_1 + s_2 + \dots + s_l}}, \quad l = \ell(\boldsymbol{s}), \tag{7.10}$$

for admissible multi-indices $\boldsymbol{s}$, which differs from that in (4.10). This representation may be proved by straightforward integrating the series

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

Substituting $u = t_1 \cdots t_{s_1}$, $v = t_{s_1+1} \cdots t_{s_1+s_2}$ 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 (7.10), we arrive at identity (7.9).

8. Ihara–Kaneko derivations and Ohno's relations

As in Section 6, 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

$$\delta(uv)=\delta(u)v+u\delta(v) \quad \text{for all} \quad u,v\in\mathfrak{A}. \tag{8.1}$$

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 $\operatorname{Der}(\mathfrak{A})$ (naturally graded by degree).

It can be easily seen that, for defining a derivation $\delta\in\operatorname{Der}(\mathfrak{A})$, 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 (8.1).

The nest assertion gives examples of derivations of $\mathfrak{A}$, when the algebra possesses an additive multiplication $\circ$ with the properties (6.1) and (6.2).

Theorem 8.1. *For any letter $a\in A$, the map*

$$\delta_a\colon w\mapsto aw-a\circ w \tag{8.2}$$

is a derivation.

Proof. Linearity of the map δ_a is clear. By Lemma 6.1, 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 (8.2) is actually a derivation. □

Theorem 8.1 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

$$\delta_{\sqcup\!\sqcup}\colon w\mapsto x_1w-x_1\sqcup\!\sqcup w, \qquad \delta_*\colon w\mapsto y_1w-y_1*w=x_1w-x_1*w, \tag{8.3}$$

are derivations; thanks to rule (3.5), the map δ_* is a derivation on the whole algebra $\mathfrak{H}$. We mention the action of derivations (8.3), obtained in accordance with (3.2)–(3.5), on the generators of the algebra:

$$\delta_{\sqcup\!\sqcup}x_0=-x_0x_1, \quad \delta_{\sqcup\!\sqcup}x_1=-x_1^2, \qquad \delta_*x_0=0, \quad \delta_*x_1=-x_1^2-x_0x_1. \tag{8.4}$$

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 4. 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 2.1 (Hoffman's relations) in the following form.

Theorem 8.2 (Derivation theorem)**.** *For any word $w\in\mathfrak{H}^0$, the identity*

$$\zeta(Dw)=\zeta(\overline{D}w) \tag{8.5}$$

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 (2.1) corresponds to the element

$$Dw = D(y_{s_1}y_{s_2}\cdots y_{s_l}) = y_{s_1+1}y_{s_2}\cdots y_{s_l} + y_{s_1}y_{s_2+1}y_{s_3}\cdots y_{s_l} + \cdots + y_{s_1}\cdots y_{s_{l-1}}y_{s_l+1} \tag{8.6}$$

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

$$\begin{aligned}\overline{D}w &= \tau D\bigl(x_0x_1^{s_l-1}x_0x_1^{s_{l-1}-1}\cdots x_0x_1^{s_2-1}x_0x_1^{s_1-1}\bigr)\\ &= \tau\sum_{\substack{k=1\\ s_k\ge2}}^{l}\sum_{j=0}^{s_k-2} x_0x_1^{s_l-1}\cdots x_0x_1^{s_{k+1}-1}x_0x_1^{j}x_0x_1^{s_k-j-1}x_0x_1^{s_{k-1}-1}\cdots x_0x_1^{s_1-1}\\ &= \sum_{\substack{k=1\\ s_k\ge2}}^{l}\sum_{j=0}^{s_k-2} x_0^{s_1-1}x_1\cdots x_0^{s_{k-1}-1}x_1x_0^{s_k-j-1}x_1x_0^{j}x_1x_0^{s_{k+1}-1}x_1\cdots x_0^{s_l-1}x_1 \end{aligned}\tag{8.7}$$

that corresponds to the right-hand side in (2.1). Applying now the map ζ to the both sides of obtained equalities (8.6) and (8.7), by Theorem 2.1 we deduce the required identity (8.5). □

Remark. The condition $w \in \mathfrak{H}^0$ in Theorem 8.2 cannot be weakened; equality (8.5) 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 3.3. Comparing action (8.4) of derivations (8.3) 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 $\delta_* - \delta_ш = \overline{D} - D$. Therefore application of Theorem 8.2 to the word $w \in \mathfrak{H}^0$ leads to the required equality:

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

This completes the proof. □

Remark. Another proof of Theorem 3.3, based on the shuffle and stuffle relations for the so-called *coloured* polylogarithms

$$\mathrm{Li}_{\boldsymbol{s}}(\boldsymbol{z}) = \mathrm{Li}_{(s_1,s_2,\dots,s_l)}(z_1, z_2, \dots, z_l) := \sum_{n_1>n_2>\cdots>n_l\ge1} \frac{z_1^{n_1}z_2^{n_2}\cdots z_l^{n_l}}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}, \tag{8.8}$$

is given by Waldschmidt. (As it is easily seen, specializing $z_2 = \cdots = z_l = 1$ functions (8.8) become generalized polylogarithms (4.1).) We do not however have a goal to expose properties of the functional model (8.8) in these lectures.

Theorem 8.2 has a natural generalization. For any $n \ge 1$, define the anti-symmetric derivation $\partial_n \in \mathrm{Der}(\mathfrak{H})$ by the rule $\partial_n x_0 = x_0(x_0+x_1)^{n-1}x_1$; as mentioned in the proof of Theorem 3.3, we have $\partial_1 = \overline{D} - D = \delta_* - \delta_ш$. The following result is valid.

Theorem 8.3. *For any $n \ge 1$ and any word $w \in \mathfrak{H}^0$, the identity*

$$\zeta(\partial_n w) = 0 \tag{8.9}$$

holds.

In what follows, we describe a scheme of the proof of the theorem given in a preprint by Kaneko and Ihara (the proof of Hoffman and Ohno is based on a different method).

The following result, proved by Ohno by means of the partial-fraction method, contains as particular cases Theorems 2.1, 2.4, and 4.1 (corresponding implications are also given by Ohno).

Theorem 8.4 (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 $n \ge 0$, the identity

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

holds.

For each integer $n \ge 1$ define the derivation $D_n \in \operatorname{Der}(\mathfrak{H})$ 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 $\delta \in \operatorname{Der}(\mathfrak{H})$ 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 8.5. *For any word $w \in \mathfrak{H}^0$, the inclusion*

$$(\sigma - \overline{\sigma}) w \in \ker\zeta \tag{8.10}$$

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

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

thus $\sigma w - \sigma\tau w \in \ker\zeta$ by Theorem 8.4. Applying now Theorem 4.1, we arrive at the desired inclusion (8.10). ☐

Recalling $\partial_1, \partial_2, \dots$, consider the derivation

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

Lemma 8.1. *The following equality holds:*

$$\exp(\partial) = \overline{\sigma} \cdot \sigma^{-1}. \tag{8.11}$$

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 $n \ge 1$, it is sufficient to verify the equality $\partial_n\partial_m x_0 = \partial_m\partial_n x_0$ for $n, m \ge 1$. Taking in mind $\partial_n(x_0 + x_1)^k = 0$, for any $n \ge 1$ and $k \ge 0$ 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

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

Routine verification 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 (8.11). □

Proof of Theorem 8.3. Now let us show how Theorem 8.3 follows from Theorem 8.5 and Lemma 8.1. 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 8.5 yields the required identities (8.9). □

Does there exist a simpler way of proving relations (8.9)? Explicit computations show that $\partial_1 = \delta_* - \delta_{ш}$,

$$\partial_2 = [\delta_*, \overline{\delta}_*],$$

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

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

and, in addition, $\delta_* + \overline{\delta}_* = \delta_{ш} + \overline{\delta}_{ш}$; therefore cases $n = 1, 2, 3, 4$ in Theorem 8.3 are served by induction (with Theorem 8.2 as inductive base). This circumstance motivates the following hypothesis.

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

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

9. Proof of Ohno's relations

In this section, we will discuss the original proof of Theorem 8.4 given by Y. Ohno. For an admissible multi-index $\boldsymbol{s}=(s_1,s_2,\dots,s_l)$ and an integer $n\ge 0$, denote

$$\begin{aligned} Z(\boldsymbol{s};n) &:= \sum_{\substack{e_1,e_2,\dots,e_l\ge 0\\ e_1+e_2+\dots+e_l=n}} \zeta(s_1+e_1,s_2+e_2,\dots,s_l+e_l) \\ &= \sum_{\substack{e_1,e_2,\dots,e_l\ge 0\\ e_1+e_2+\dots+e_l=n}} \zeta(x_0^{s_1+e_1-1}x_1x_0^{s_2+e_2-1}x_1\cdots x_0^{s_l+e_l-1}x_1), \end{aligned}$$

the sum which occurs on the both sides of Ohno's relations. If we express

$$x_0^{s_1-1}x_1x_0^{s_2-1}x_1\cdots x_0^{s_l-1}x_1 = x_0^{\mu_1}x_1^{\nu_1}x_0^{\mu_2}x_1^{\nu_2}\cdots x_0^{\mu_k}x_1^{\nu_k}, \tag{9.1}$$

where all the exponents are positive integers, then

$$\begin{aligned} Z(\boldsymbol{s};n) = \sum_{\substack{q_1,\dots,q_k\ge 0\\ q_1+\dots+q_k=n}} \sum_{\substack{\varepsilon_{i,1}+\dots+\varepsilon_{i,\nu_i+q_i}=\nu_i\\ \varepsilon_{i,j}\in\{0,1\},\ \varepsilon_{i,\nu_i+q_i}=1,\ i=1,\dots,k}} &\zeta\big(x_0^{\mu_1}x_{\varepsilon_{1,1}}\cdots x_{\varepsilon_{1,\nu_1+q_1}}x_0^{\mu_2}x_{\varepsilon_{2,1}}\cdots x_{\varepsilon_{2,\nu_2+e_2}} \\ &\qquad\cdots x_0^{\mu_k}x_{\varepsilon_{k,1}}\cdots x_{\varepsilon_{k,\nu_k+q_k}}\big) \\ = \sum_{\substack{q_1,\dots,q_k\ge 0\\ q_1+\dots+q_k=n}} \Sigma_{\boldsymbol{s}}(\nu_1,\dots,\nu_k;q_1,\dots,q_k) \end{aligned}$$

in notation

$$\begin{aligned} &\Sigma_{\boldsymbol{s}}(\lambda_1,\dots,\lambda_k;q_1,\dots,q_k) \\ &\quad= \sum_{\substack{\varepsilon_{i,1}+\dots+\varepsilon_{i,\nu_i+q_i}=\lambda_i\\ \varepsilon_{i,j}\in\{0,1\},\ \varepsilon_{i,\nu_i+q_i}=1,\ i=1,\dots,k}} \zeta\big(x_0^{\mu_1}x_{\varepsilon_{1,1}}\cdots x_{\varepsilon_{1,\nu_1+q_1}}x_0^{\mu_2}x_{\varepsilon_{2,1}}\cdots x_{\varepsilon_{2,\nu_2+e_2}} \\ &\qquad\qquad\cdots x_0^{\mu_k}x_{\varepsilon_{k,1}}\cdots x_{\varepsilon_{k,\nu_k+q_k}}\big). \end{aligned}$$

Furthermore, we have

$$\begin{aligned} &\sum_{\substack{\lambda_i=1\\ i=1,\dots,k}}^{\nu_i+q_i} \Sigma_{\boldsymbol{s}}(\lambda_1,\dots,\lambda_k;q_1,\dots,q_k)T_1^{\lambda_1-1}\cdots T_k^{\lambda_k-1} \\ &\qquad= \zeta\big(x_0^{\mu_1}(x_0+T_1x_1)^{\nu_1+q_1-1}x_1\cdots x_0^{\mu_k}(x_0+T_kx_1)^{\nu_k+q_k-1}x_1\big). \end{aligned}$$

To compute the latter expression, we use the integral representation from (4.10); performing the integration for each subword $x_0^{\mu_i-1}x_0(x_0+T_ix_1)^{\nu_i+q_i-1}x_1$, $i=1,\dots,k$, we obtain

$$\begin{aligned}
&\int\cdots\int\limits_{t_{2i-2}>z_1>\dots>z_{\mu_i-1}>t_{2i-1}>z'_1>\dots>z'_{\nu_i+q_i-1}>t_{2i}} \frac{\mathrm{d}z_1}{z_1}\cdots\frac{\mathrm{d}z_{\mu_i-1}}{z_{\mu_i-1}}\,\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}} \\
&\qquad\times\biggl(\frac{\mathrm{d}z'_1}{z'_1}+T_i\frac{\mathrm{d}z'_1}{1-z'_1}\biggr)\cdots\biggl(\frac{\mathrm{d}z'_{\nu_i+q_i-1}}{z'_{\nu_i+q_i-1}}+T_i\frac{\mathrm{d}z'_{\nu_i+q_i-1}}{1-z'_{\nu_i+q_i-1}}\biggr)\frac{\mathrm{d}t_{2i}}{1-t_{2i}} \\
&\quad=\frac{1}{(\mu_i-1)!\,(\nu_i+q_i-1)!}\iint\limits_{t_{2i-2}>t_{2i-1}>t_{2i}}\biggl(\log\frac{t_{2i-2}}{t_{2i-1}}\biggr)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}} \\
&\qquad\times\biggl(\log\frac{t_{2i-1}}{t_{2i}}-T_i\log\frac{1-t_{2i-1}}{1-t_{2i}}\biggr)^{\nu_i+q_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}},
\end{aligned}$$

so that, with the help of the binomial theorem, the coefficient of $T_i^{\nu_i-1}$ in the latter expression is equal to

$$\begin{aligned}
&\frac{1}{(\mu_i-1)!\,q_i!\,(\nu_i-1)!}\iint\limits_{t_{2i-2}>t_{2i-1}>t_{2i}}\biggl(\log\frac{t_{2i-2}}{t_{2i-1}}\biggr)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}}\biggl(\log\frac{t_{2i-1}}{t_{2i}}\biggr)^{q_i} \\
&\qquad\times\biggl(-\log\frac{1-t_{2i-1}}{1-t_{2i}}\biggr)^{\nu_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}}
\end{aligned}$$

and we finally arrive at

$$\begin{aligned}
\Sigma_{\boldsymbol{s}}(\nu_1,\dots,\nu_k;q_1,\dots,q_k)&=\frac{1}{\prod_{i=1}^k(\mu_i-1)!\,q_i!\,(\nu_i-1)!}\int\cdots\int\limits_{1>t_1>\dots>t_{2k}>0}\prod_{i=1}^k\biggl(\log\frac{t_{2i-2}}{t_{2i-1}}\biggr)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}} \\
&\qquad\times\biggl(\log\frac{t_{2i-1}}{t_{2i}}\biggr)^{q_i}\biggl(\log\frac{1-t_{2i}}{1-t_{2i-1}}\biggr)^{\nu_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}}
\end{aligned}$$

and

$$\begin{aligned}
Z(\boldsymbol{s};n)&=\sum_{\substack{q_1,\dots,q_k\ge0\\ q_1+\dots+q_k=n}}\Sigma_{\boldsymbol{s}}(\nu_1,\dots,\nu_k;q_1,\dots,q_k) \\
&=\frac{1}{n!\prod_{i=1}^k(\mu_i-1)!\,(\nu_i-1)!}\int\cdots\int\limits_{1>t_1>\dots>t_{2k}>0}\biggl(\sum_{i=1}^k\log\frac{t_{2i-1}}{t_{2i}}\biggr)^n \\
&\qquad\times\prod_{i=1}^k\biggl(\log\frac{t_{2i-2}}{t_{2i-1}}\biggr)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}}\biggl(\log\frac{1-t_{2i}}{1-t_{2i-1}}\biggr)^{\nu_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}}
\end{aligned}$$

with the convention $t_0=1$.

In the latter integral we introduce the change of variables

$$u_{2i-1}=\log\frac{t_{2i-2}}{t_{2i-1}},\quad u_{2i}=\log\frac{1-t_{2i}}{1-t_{2i-1}},\qquad i=1,\dots,k,$$

so that

$$\prod_{i=1}^{k} \frac{\mathrm{d}t_{2i-1}}{t_{2i-1}} \, \frac{\mathrm{d}t_{2i}}{1-t_{2i}} = \mathrm{d}u_1 \, \mathrm{d}u_2 \cdots \mathrm{d}u_{2k}$$

and the expression

$$\begin{aligned}\biggl(\prod_{i=1}^{k} \frac{t_{2i-1}}{t_{2i}}\biggr)^{-1} &= \prod_{i=1}^{k} \frac{t_{2i-2}}{t_{2i-1}} \cdot t_{2k} \\ &= \exp\biggl(\sum_{i=1}^{k} u_{2i-1}\biggr) \cdot \sum_{j=0}^{2k} (-1)^j \exp\biggl(\sum_{i=j+1}^{2k} (-1)^i u_i\biggr) \\ &=: f(u_1, u_2, \dots, u_{2k-1}, u_{2k})\end{aligned}$$

satisfies the symmetry relation

$$f(u_1, u_2, \dots, u_{2k-1}, u_{2k}) = f(u_{2k}, u_{2k-1}, \dots, u_2, u_1).$$

Then

$$\begin{aligned} Z(\boldsymbol{s}; n) = \frac{1}{n! \prod_{i=1}^{k} (\mu_i - 1)! \, (\nu_i - 1)!} \underset{\substack{u_i > 0, \ i=1,\dots,2k \\ f(u_1,\dots,u_{2k}) > 0}}{\int \cdots \int} \bigl(-\log f(u_1, \dots, u_{2k})\bigr)^n \\ \times \prod_{i=1}^{k} u_{2i-1}^{\mu_i - 1} u_{2i}^{\nu_i - 1} \, \mathrm{d}u_1 \, \mathrm{d}u_2 \cdots \mathrm{d}u_{2k}, \end{aligned}$$

and the change of variables

$$(u_1, u_2, \dots, u_{2k-1}, u_{2k}) \leftrightarrow (u_{2k}, u_{2k-1}, \dots, u_2, u_1)$$

swaps the roles of μ_i and ν_i, $i = 1, \dots, k$, and reverses them; in other words, as the record (9.1) shows, it reduces the resulting expression to $Z(\boldsymbol{s}'; n)$. This completes the proof of Theorem 8.4.

It is straightforward that case $n = 0$ in Theorem 8.4 is the duality theorem (Theorem 4.1).

Exercise 9.1. (a) Show that the choice $n = 1$ in Theorem 8.4 corresponds to Hoffman's relations (Theorem 2.1).

(b) Show that, if multi-index $\boldsymbol{s}$ in Theorem 8.4 is one-component (that is, $\boldsymbol{s} = (s)$), then the theorem reduces to the sum theorem (Theorem 2.4).

10. The identity of Borwein, Bradley and Broadhurst

In this section we will sketch Zhao's proof of identity (5.5); this is the only proof known so far.

We have already settled standard setup for the (alternating) Euler sums: the non-commutative algebra $\mathbb{Q}\langle x_0, x_1\rangle$ is extended to the algebra $\mathfrak{H} = \mathbb{Q}\langle x_0, x_1, \overline{x}_1\rangle$, and the subalgebra $\mathfrak{H}^1 = \mathbb{Q}\mathbf{1} \oplus \mathfrak{H}x_1$ is generated by words $y_s := x_0^{s-1}x_1$ and $\overline{y}_s := x_0^{s-1}\overline{x}_1$ (those not ending with x_0). The subalgebra $\mathfrak{H}^0$ of *admissible* words is generated by words not beginning with x_1 and not ending with x_0.

By assigning the three differential forms

$$x_0 \mapsto \omega_0(z)\,\mathrm{d}z = \frac{\mathrm{d}z}{z}, \quad x_1 \mapsto \omega_1(z)\,\mathrm{d}z = \frac{\mathrm{d}z}{1-z},$$
$$\text{and} \quad \overline{x}_1 \mapsto \overline{\omega}_1(z)\,\mathrm{d}z = \frac{-\mathrm{d}z}{1+z}.$$

(cf. (4.15)) to the three letters, for a word $w \in \mathfrak{H}^1$ we define the evaluation zeta map by

$$\zeta(w) := \int_0^1 w$$

(with the convention used in (4.11)). Then, of course, $\zeta(\boldsymbol{s}) = \zeta(y_{s_1}\cdots y_{s_l})$ if the multi-index $\boldsymbol{s} = (s_1, \dots, s_l)$ does not involve bars (so that the corresponding word does not contain letter $\overline{x}_1$). For example,

$$(\{3\}^n) \mapsto y_3^n = (x_0^2 x_1)^n.$$

If however the multi-index $\boldsymbol{s}$ involves bars, then the rule of assigning the word is as follows. Going for s_1 to s_l, as soon as we see the first signed entry $\overline{s}_i$ we change every y after y_{s_i} (inclusive) to $\overline{y}$ until the next signed entry $\overline{s}_j$ occur. We then leave all the y's after y_{s_j} (again inclusive) until we see the next signed entry when we start toggling again, and so on. In other words, we can thin of the bars as of switches between y and $\overline{y}$.

Exercise 10.1. Write the word which corresponds to the multi-index $(2, 1, \overline{2}, 3, \overline{4}, \overline{5})$.

Exercise 10.2. Prove the following correspondence:

$$(\{\overline{2}, 1\}^n) \mapsto (x_0\overline{x}_1^2 x_0 x_1^2)^{\lfloor n/2 \rfloor}(x_0\overline{x}_1^2)^{2\{n/2\}} = \begin{cases} (x_0\overline{x}_1^2 x_0 x_1^2)^k (x_0\overline{x}_1^2) & \text{if } n = 2k+1, \\ (x_0\overline{x}_1^2 x_0 x_1^2)^k & \text{if } n = 2k. \end{cases}$$

The shuffle and stuffle products in (3.2)–(3.4) are extended to the algebra $\mathfrak{H}^0$ as well. In fact, the shuffle product uses the old rules, now allowing one extra letter $\overline{x}_1$ for either x_j or x_k in (3.3). As for the stuffle product, to complement rule (3.4) we use

$$y_j u * y_k v = y_j \gamma_{y_j}(\gamma_{y_j} u * y_k v) + y_k \gamma_{y_k}(y_j u * \gamma_{y_k} v) + [y_j, y_k]\gamma_{[y_j, y_k]}(\gamma_{y_j} u * \gamma_{y_k} v), \quad (10.1)$$

where $\gamma_{y_j} w = w$ for $y_j = x_0^{j-1} x_1$ and $\gamma_{\overline{y}_j} w$ is the word with all y and $\overline{y}$ toggled,

$$[y_j, y_k] = [\overline{y}_j, \overline{y}_k] = y_{j+k}, \quad [y_j, \overline{y}_k] = [\overline{y}_j, y_k] = \overline{y}_{j+k}.$$

Then

$$\zeta(w_1 \sqcup\!\sqcup w_2) = \zeta(w_1 * w_2) = \zeta(w_1)\zeta(w_2).$$

For a word $w = a_1 a_2 \cdots a_m$ over the alphabet $\{x_0, x_1, \overline{x}_1\}$, define the ith shuffle iteration by

$$\sqcup\!\sqcup_i w := \begin{cases} a_1 a_2 \cdots a_i \sqcup\!\sqcup a_{i+1} \cdots a_m & \text{if } i \text{ is odd}, \\ a_i \cdots a_2 a_1 \sqcup\!\sqcup a_{i+1} \cdots a_m & \text{if } i \text{ is even}, \end{cases} \qquad i = 0, 1, \dots, m.$$

Similarly, but considering a word over the infinite alphabet $\{\overline{y}_0, y_1, \overline{y}_1, \dots\}$, define the ith harmonic (stuffle) iteration $*_i$. Finally, define the $\star$-concatenation by settling $w_1 \star w_2 = w_1 w_2$ except that

$$x_1 \star x_1 = x_1 \overline{x}_1 \quad \text{and} \quad \overline{x}_1 \star \overline{x}_1 = \overline{x}_1 x_1.$$

Exercise 10.3. Prove by induction that for every positive n,

$$\sum_{i=0}^{2n} (-1)^i *_i \big((\overline{x}_1 z)^{\star n}\big) = (-1)^n \big(x_0^2(x_1 + \overline{x}_1)\big)^n,$$

where $z = x_0(x_1 + \overline{x}_1)$ is regarded as one letter when the ith harmonic iteration is preformed, retaining the $\star$-concatenation. Note that $z \star x_1 = z \star \overline{x}_1 = x_0(x_1\overline{x}_1 + \overline{x}_1 x_1)$.

Exercise 10.4. Prove by induction that for every positive n,

$$\sum_{i=0}^{2n} (-1)^i ш_i \big((\overline{x}_1 z)^{\star n}\big) = (-2)^n (x_0 \overline{x}_1^2 x_0 x_1^2)^{\lfloor n/2 \rfloor} (x_0 \overline{x}_1^2)^{2\{n/2\}}$$

and

$$\sum_{i=0}^{2n} (-1)^i ш_i \big((x_1 z)^{\star n}\big) = (-2)^n (x_0 x_1^2 x_0 \overline{x}_1^2)^{\lfloor n/2 \rfloor} (x_0 x_1^2)^{2\{n/2\}}.$$

Exercise 10.5 (Distribution relation). Show that for every positive n,

$$\zeta\Big(\big(x_0^2(x_1 + \overline{x}_1)\big)^n\Big) = \frac{1}{4^n} \zeta\big((x_0^2 x_1)^n\big) = \frac{1}{4^n} \zeta(\{3\}^n).$$

Hint. Perform the substitution $z \mapsto z^2$ into Chen's iterated integral for $\zeta\big((x_0^2 x_1)^n\big)$. □

Using Exercises 10.2–10.5, we deduce identity (5.5).

11. Open questions

In addition to Conjectures 1–3 indicated in Section 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. *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 (3.6)–(3.8) 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 (4.12), (5.2), and (5.4)). For $k \geq 5$, above-deduced identities allow to compute the upper bounds; for instance, $d_5 \leq 2$, $d_6 \leq 2$, and so on.

Conjecture 5. *For $k \ge 3$, the recurrent relations*

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

hold; equivalently,

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

Note that the estimates $\dim_{\mathbb{Q}} \mathcal{Z}_k \le d_k$ for all $k \ge 2$ have been recently shown by Terasoma and, independently, by Deligne and Goncharov.

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

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

$$\bigl\{\zeta(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2, 3\},\ j = 1, \dots, \ell(\boldsymbol{s})\bigr\}. \tag{11.1}$$

A serious argument for Conjecture 6 to be valid, is not only experimental confirmation for $k \le 16$ (under the hypothesis of Conjecture 2) but also agreement of the dimension of the $\mathbb{Q}$-space spanned by the numbers (11.1) with the dimension d_k of the spaces $\mathcal{Z}_k$ in Conjecture 5. The last fact is proved by Hoffman. In his recent work, F. Brown shows that Conjecture 6 is true for the 'motivic' version of MZVs; in particular, that all usual MZVs can be expressed by means of the elements (11.1) of Hoffman's basis. In the heart of the proof, there is a remarkable identity of MZVs which was shown to be true by D. Zagier. It is this identity which we discuss in the next section.

Exercise 11.1. (a) How many different MZVs of given weight k exists?

(b) Compute the limit of $d_k^{1/k}$ as $k \to \infty$ for the sequence d_k constructed in Conjecture 5.

(c) Any polynomial in single zeta values,

$$(\pi^2)^{s_0}\zeta(3)^{s_1}\zeta(5)^{s_2}\cdots\zeta(2l+1)^{s_l}, \qquad s_0, s_1, s_2, \dots, s_l \in \mathbb{Z}_{\ge 0},$$

belongs to the linear space $\mathcal{Z}_k$ of MZVs of weight

$$k = 2s_0 + 3s_1 + 5s_2 + \cdots + (2l+1)s_l.$$

Assuming Conjecture 1, all these polynomials are linearly independent over $\mathbb{Q}$. Denote by c_k the total number of such polynomials of given weight k. Compute c_k for small values of k (namely, for $k \le 12$) and show that $c_k < d_k$ for $k \ge 8$. (In other words, the algebra of MZVs cannot be fully generated by single zeta values.)

(d) For the sequence c_k from part (c), find a general analytic formula and compute the limit of $c_k^{1/k}$ as $k \to \infty$.

Although proving Conjectures 4–6 in the form they are given is hopeless at the present time, the 'true' MZVs in $\mathbb{R}$ are the images under a $\mathbb{Q}$-linear map of certain 'motivic' MZVs which are defined *purely algebraically*. The Terasoma and Goncharov–Deligne bound $\dim_{\mathbb{Q}} \mathcal{Z}_k \le d_k$, as well as Conjecture 4 about disjointness of the subspaces $\mathcal{Z}_k$, are shown to be true for this algebraic version of MZVs. Terasoma and Goncharov established the bound by showing that all MZVs are periods of so-called *mixed Tate motives* that are unramified over $\mathbb{Z}$. Another well-known conjecture in the area states the converse, that is, that all periods of mixed Tate motives over $\mathbb{Z}$ can be expressed as linear combinations (over $\mathbb{Q}[(2\pi i)^{\pm 1}]$) of MZVs. Equivalently, this says that the dimension of the space of motivic MZVs of weight k is exactly d_k.

The result obtained by Brown was a proof of the latter conjecture and also of the fact that the motivic MZVs from Hoffman's conjectural basis in Conjecture 6 form a basis of the corresponding $\mathcal{Z}_k$. In his proof Brown assumes certain quite specific properties of certain coefficients occurring in the relations over $\mathbb{Q}$ of some special MZVs. Specifically, he shows that the special MZVs

$$\xi(m,n) := \zeta(\{2\}^m, 3, \{2\}^n), \qquad n, m \ge 0, \tag{11.2}$$

which are part of Hoffman's basis, are $\mathbb{Q}$-linear combinations of products $\pi^{2\mu}\zeta(2\nu+1)$ with $\mu + \nu = m + n + 1$. His proof, which used motivic ideas, did not yield an explicit formula for these linear combinations, but numerical evidence suggested several properties satisfied by their coefficients (and, in particular, of the coefficient of $\zeta(2m+2n+3)$) which he could show were sufficient to imply the truth of both Hoffman's conjecture and the statement about motivic periods. The next section contains a statement and proof of an explicit formula expressing the numbers (11.2) in terms of single zeta values, as well as confirmation of the numerical properties that were required for Brown's proof.

12. Zagier's identity for $\xi(m,n)$

Before giving the formula for the numbers $\xi(m,n)$, we first recall the much easier formula from the family (5.4),

$$\xi(n) := \zeta(\{2\}^n) = \frac{\pi^{2n}}{(2n+1)!}, \qquad n \ge 0, \tag{12.1}$$

for the simplest of the Hoffman basis elements.

Theorem 12.1 (Zagier)**.** *For all integers $m,n \ge 0$, we have*

$$\xi(m,n) = 2\sum_{r=1}^{m+n+1}(-1)^{r-1}\left(\left(1-\frac{1}{2^{2r}}\right)\binom{2r}{2m+1}-\binom{2r}{2n+2}\right)\xi(m+n-r+1)\zeta(2r+1), \tag{12.2}$$

where the value of $\xi(m+n-r+1)$ is given by (12.1). Conversely, each product $\xi(\mu)\zeta(k-2\mu)$ of odd weight k is a rational combination of numbers $\xi(m,n)$ with $m+n=(k-3)/2$.

Remark. The second part of the theorem, which we only discuss as Exercise 12.6 below, gives rise to several other open questions.

The coefficients in the expressions for the products $\xi(\mu)\zeta(k-2\mu)$ as linear combinations of the numbers $\xi(m,n)$ do not seem to be given by any simple formula. For example, the inverse of the 5×5 matrix

$$\begin{pmatrix} 3 & -\frac{15}{2} & \frac{189}{16} & -\frac{255}{16} & \frac{4603}{256} \\ 0 & -\frac{15}{2} & \frac{315}{8} & -\frac{1753}{16} & \frac{9585}{64} \\ 0 & 0 & \frac{157}{16} & -\frac{889}{16} & \frac{10689}{128} \\ 0 & 2 & -30 & \frac{1985}{16} & -\frac{11535}{64} \\ -2 & 12 & -30 & 56 & -\frac{17925}{256} \end{pmatrix}$$

expressing the vector $\{\xi(m,n) : m+n=4\}$ in terms of the vector $\{\zeta(2m+3)\xi(n) : m+n=4\}$ is

$$\frac{1}{2555171}\begin{pmatrix} 11072595 & 19354609 & 23488575 & 22114173 & 15331307 \\ 59984880 & 122931470 & 160083660 & 147349978 & 89977320 \\ 246001728 & 508012288 & 669540272 & 613537008 & 369002592 \\ 494939520 & 1022542528 & 1349936640 & 1236102000 & 742409280 \\ 300405248 & 620662272 & 819546624 & 750355968 & 450607872 \end{pmatrix},$$

in which no simple pattern can be discerned and in which even the denominator (prime 2555171) cannot be recognised. This shows that the Hoffman basis, although it works over $\mathbb{Q}$, is very far from giving a basis over $\mathbb{Z}$ of $\mathbb{Z}$-linear span of MZVs, and suggests the question of finding better basis elements.

The following question is supported by numerical data for $m+n \le 30$, but remains open.

Exercise 12.1. Denote M_k the matrix from (12.2) expressing the vector $\{\xi(m,n) : m+n=k\}$ in terms of the vector $\{\zeta(2m+3)\xi(n) : m+n=k\}$, that is,

$$M_k = \left(2(-1)^\mu\left(\left(1-\frac{1}{2^{2\mu+2}}\right)\binom{2\mu+2}{2m+1}-\binom{2\mu+2}{2k-2m+2}\right)\right)_{0\le m,\mu\le k}. \tag{12.3}$$

Show that all the entries of the inverse matrix M_k^{-1} are strictly positive.

The strategy to prove Theorem 12.1 is to compare the two generating functions

$$F(x,y) = \sum_{m,n\ge 0}(-1)^{m+n+1}\xi(m,n)x^{2m+1}y^{2n+2} \tag{12.4}$$

and

$$\widehat{F}(x,y) = \sum_{m,n\ge 0}(-1)^{m+n+1}\hat{\xi}(m,n)x^{2m+1}y^{2n+2}, \tag{12.5}$$

where

$$\hat{\xi}(m,n) = 2\sum_{r=1}^{m+n+1}(-1)^{r-1}\left(\left(1-\frac{1}{2^{2r}}\right)\binom{2r}{2m+1}-\binom{2r}{2n+2}\right)\xi(m+n-r+1)\zeta(2r+1)$$

denotes the expression occurring on the right-hand side of (12.2). Of course, if the two expressions were the same, we would be done, but in fact they are *completely different*, one involving a generalized hypergeometric function

$${}_{p+1}F_p\left(\begin{matrix}a_0,\, a_1,\, \dots,\, a_p\\ b_1,\, \dots,\, b_p\end{matrix}\,\middle|\, z\right) = \sum_{n=0}^{\infty}\frac{(a_0)_n(a_1)_n\cdots(a_p)_n}{(b_1)_n\cdots(b_p)_n}\frac{z^n}{n!} \tag{12.6}$$

(cf. Section 5), and the other a complicated linear combination of the digamma functions, $\psi(x) = \Gamma'(x)/\Gamma(x)$. We therefore have to proceed indirectly, showing that both F and $\widehat{F}$ are entire functions (of order 1) in x and y and that they agree whenever $x=y$ or x or y is an integer (the details of this comparison will be however skipped). This will imply the equality $F=\widehat{F}$, and hence Theorem 12.1. There is however a belief (that is, an *open problem*!) that the use of known hypergeometric identities could lead to a direct proof of $F=\widehat{F}$; this would considerably simplify Brown's proofs mentioned above.

Lemma 12.1. *The generating function $F(x,y)$ can be expressed as the product of a sine function and a hypergeometric function:*

$$F(x,y) = \frac{\sin\pi x}{\pi}\frac{\partial}{\partial z}{}_3F_2\left(\begin{matrix}y,\, -y,\, z\\ 1+x,\, 1-x\end{matrix}\,\middle|\, 1\right)\bigg|_{z=0}. \tag{12.7}$$

Proof. The proof is similar to that for (12.1):

$$\begin{aligned}
F(x,y) &= \sum_{m,n\ge0}(-1)^{m+n+1}\zeta(\{2\}^m,3,\{2\}^n)x^{2m+1}y^{2n+2}\\
&= -xy^2\sum_{r=1}^{\infty}\prod_{k=1}^{r-1}\biggl(1-\frac{y^2}{k^2}\biggr)\cdot\frac{1}{r^3}\cdot\prod_{l=r+1}^{\infty}\biggl(1-\frac{x^2}{l^2}\biggr)\\
&= \frac{\sin\pi x}{\pi}\sum_{r=1}^{\infty}\frac{(-y)_r(y)_r}{(1-x)_r(1+x)_r}\,\frac{1}{r},
\end{aligned}$$

and this formula is seen to be equivalent to (12.7). □

Lemma 12.2. *The generating function $\widehat{F}(x,y)$ can be expressed as an integral linear combination of* fourteen *functions of the form*

$$\psi\biggl(1+\frac{u}{2}\biggr)\frac{\sin\pi v}{2\pi}\qquad\text{with}\quad u\in\{\pm x\pm y,\pm2x\pm2y,\pm2x\},\quad v\in\{x,y\}.$$

Proof. From the definition of $\widehat{F}(x,y)$ and (12.1) we find

$$\begin{aligned}
\widehat{F}(x,y) &= 2\sum_{m,n\ge0}(-1)^{m+n}x^{2m+1}y^{2n+2}\sum_{r=1}^{m+n+1}(-1)^r\Biggl((1-2^{-2r})\binom{2r}{2m+1}\\
&\qquad-\binom{2r}{2n+2}\Biggr)\frac{\pi^{2(m+n-r+1)}}{(2(m+n-r+1)+1)!}\zeta(2r+1)\\
&= \frac{2}{\pi}\sum_{s=0}^{\infty}\sum_{r=1}^{\infty}\frac{(-1)^s\pi^{2s+1}}{(2s+1)!}\zeta(2r+1)\Biggl(\sum_{n=0}^{r-1}\binom{2r}{2n+2}x^{2(s+r-n)-1}y^{2n+2}\\
&\qquad-(1-2^{-2r})\sum_{m=0}^{r-1}\binom{2r}{2m+1}x^{2m+1}y^{2(s+r-m)}\Biggr)\\
&= \frac{2\sin\pi x}{\pi}\sum_{r=1}^{\infty}\zeta(2r+1)\sum_{n=0}^{r-1}\binom{2r}{2n+2}x^{2(r-n-1)}y^{2(n+1)}\\
&\qquad-\frac{2\sin\pi y}{\pi}\sum_{r=1}^{\infty}(1-2^{-2r})\zeta(2r+1)\sum_{m=0}^{r-1}\binom{2r}{2m+1}x^{2m+1}y^{2(r-m)-1}\\
&= \frac{\sin\pi x}{\pi}\sum_{r=1}^{\infty}\zeta(2r+1)\bigl((x+y)^{2r}+(x-y)^{2r}-2x^{2r}\bigr)\\
&\qquad-\frac{\sin\pi y}{\pi}\sum_{r=1}^{\infty}(1-2^{-2r})\zeta(2r+1)\bigl((x+y)^{2r}-(x-y)^{2r}\bigr)\\
&= \frac{\sin\pi x}{\pi}\bigl(A(x+y)+A(x-y)-2A(x)\bigr)-\frac{\sin\pi y}{\pi}\bigl(B(x+y)-B(x-y)\bigr),
\end{aligned}$$

where (cf. the final part of Section 5)

$$A(t)=\sum_{r=1}^{\infty}\zeta(2r+1)t^{2r}=\sum_{r=1}^{\infty}\sum_{n=1}^{\infty}\frac{t^{2r}}{n^{2r+1}}=\sum_{n=1}^{\infty}\frac{t^2}{n(n^2-t^2)},$$

$$B(t)=\sum_{r=1}^{\infty}(1-2^{-2r})\zeta(2r+1)t^{2r}=\sum_{r=1}^{\infty}\sum_{n=1}^{\infty}(-1)^{n-1}\frac{t^{2r}}{n^{2r+1}}=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}t^2}{n(n^2-t^2)}.$$

Decomposing the summands into partial fractions allows us to represent the generating functions A and B in terms of the digamma function:

$$A(t)=\frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{1}{n-t}+\frac{1}{n+t}-\frac{2}{n}\right)=\psi(1)-\frac{1}{2}\big(\psi(1+t)+\psi(1-t)\big),$$

$$B(t)=\frac{1}{2}\sum_{n=1}^{\infty}(-1)^{n-1}\left(\frac{1}{n-t}+\frac{1}{n+t}-\frac{2}{n}\right)=A(t)-A\Big(\frac{t}{2}\Big).$$

Substituting these expressions into the previous derivation gives an expression for $\widehat{F}$ of the form stated in the lemma. □

Exercise 12.2. Show the equality $F(x,y)=\widehat{F}(x,y)$ directly by using the representations in Lemmas 12.1 and 12.2.

As mentioned above, Exercise 12.2 is an open problem.

The following change of exercises sketches the remaining ingredients of the proof of Theorem 12.1.

Exercise 12.3. Show that both $F(x,y)$ and $\widehat{F}(x,y)$ are entire functions on $\mathbb{C}^2$ and are bounded by a constant multiple of $e^{\pi X}\log X$ as $X=\max\{|x|,|y|\}\to\infty$, and also by a multiple (depending on y) of $e^{\pi|\operatorname{Im}x|}$ as $|x|\to\infty$ with $y\in\mathbb{C}$ fixed.

Remark. The derivation makes use of analytic estimates of the coefficients of both $F(x,y)$ and $\widehat{F}(x,y)$ but also of certain 'standard' theorems of complex analysis, like the Phragmén–Lindelöf theorem (an extension of the maximum modulus principle to functions which are analytic in sector domains and strips).

Exercise 12.4. Show that for $x\in\mathbb{C}$ the following equality holds:

$$F(x,x)=-\frac{\sin\pi x}{\pi}A(x)=\widehat{F}(x,x),$$

where $A(x)$ is the meromorphic function defined in the proof of Lemma 12.2.

Exercise 12.5. (a) Prove that for all $n\in\mathbb{Z}_{>0}$ and $x\in\mathbb{C}$,

$$F(x,n)=\frac{\sin\pi x}{\pi}\sum_{|k|\le n}{}^{*}\,\frac{\operatorname{sgn}k}{x-k}=\widehat{F}(x,n),$$

where the asterisk means that the terms $k=\pm n$ are to be weighted with a factor 1/2.

(b) Prove that for all $m \in \mathbb{Z}_{>0}$ and $y \in \mathbb{C}$,

$$F(m,y) = (-1)^m + \frac{\sin \pi y}{\pi} \sum_{|k|\le m}^{*} \frac{(-1)^{m-k}}{k-y} = \widehat{F}(m,y),$$

with the same convention about the asterisk.

Finally, we make use of the following result.

Lemma 12.3. *An entire function $f\colon \mathbb{C} \to \mathbb{C}$ that vanishes at all integers and satisfies $f(z) = O(e^{\pi|\operatorname{Im} z|})$ as $|z| \to \infty$ is a constant multiple of $\sin \pi z$.*

Proof. Because $|\operatorname{Im} z| \le |z|$, the estimate implies $f(z) = O(e^{\pi|z|})$ as $|z| \to \infty$; in particular, $f(z)$ has order 1, and so does the function $g(z) = f(z)/\sin \pi z$ (which is indeed entire as it does not have poles). The growth hypothesis on f implies that g is bounded outside a strip of finite width around the real axis, and then it follows from the Phragmén–Lindelöf theorem that it is also bounded inside this strip (since it has finite order), so that g is constant by Liouville's theorem. ☐

Proof of Theorem 12.1. We can now complete the proof of the main equality 12.2 as follows. We have shown that $F(x,y)$ and $\widehat{F}(x,y)$ are entire functions of x and y satisfying certain (same) estimates, and that they agree whenever $x = y$ or either x or y is an integer. (The latter fact follows from Exercise 12.5 and the fact that both $F(x,y)$ and $\widehat{F}(x,y)$ are odd functions of x and even functions of y and vanish when $y = 0$.) It follows that, for fixed y, the function $f(x) = F(x,y) - \widehat{F}(x,y)$ is an entire function which vanishes at all integers and satisfies $f(x) = O(e^{\pi|\operatorname{Im} x|})$ as $|x| \to \infty$, so that by Lemma 12.3 it is a multiple of $\sin \pi x$,

$$F(x,y) - \widehat{F}(x,y) = h(y) \sin \pi x,$$

for a certain entire function $h(y)$. Substituting $y = x$ into the equality we get $h(x) = 0$ identically, so that indeed $F(x,y) - \widehat{F}(x,y) = 0$ for all x and y, implying $\xi(m,n) = \hat{\xi}(m,n)$ as required. ☐

Exercise 12.6. Prove the second statement of the theorem (that is, the invertibility of matrix M_k in (12.3)) by computing the 2-adic valuation of the entries of the matrix.

13. Double zeta values and products of single zeta values

In this section we fix an odd number $k = 2l + 1 \ge 3$ and discuss the relationship between the double zeta values $\zeta(m,n)$, the zeta products $\zeta(m)\zeta(n)$, and our latest heroes $\xi(\mu,\nu)$, all of weight $m + n = 2(\mu + \nu) + 3 = k$.

It was already found by Euler (explicitly for k up to 13) that all double zeta values of odd weight are rational linear combinations of products of single zeta values.

Theorem 13.1. *The double zeta value $\zeta(m,n)$ (with $m \ge 2$ and $n \ge 1$) of weight $m+n = k = 2l+1$ is given in terms of the products $\zeta(2s)\zeta(k-2s)$, $s = 0, 1, \dots, l-1$,* *by*

$$\zeta(m,n) = (-1)^n \sum_{s=0}^{l-1} \left(\binom{k-2s-1}{m-1} + \binom{k-2s-1}{n-1} - \delta_{m,2s} + (-1)^n \delta_{s,0} \right) \zeta(2s)\zeta(k-2s). \tag{13.1}$$

Proof. The harmonic and shuffle products in the case of single zeta values result in

$$\zeta(r)\zeta(s) = \zeta(r,s) + \zeta(s,r) + \zeta(k), \quad \text{where } r+s=k,\ r,s \geq 2, \tag{13.2}$$

$$\zeta(r)\zeta(s) = \sum_{m=2}^{k-1} \left(\binom{m-1}{r-1} + \binom{m-1}{s-1} \right) \zeta(m,k-m), \quad \text{where } r+s=k,\ r,s \geq 2, \tag{13.3}$$

In both cases we can suppose without loss of generality that $r \leq s$, since both sides of the equations are symmetric in r and s. This will give us only $2(l-1)$ equations for the $2l-1$ unknowns $\zeta(m,k-m)$, $2 \leq m \leq k-1$. However, both (13.2) and (13.3) remain true if we fix any value T (that is, any *regularization*) for the divergent zeta value $\zeta(1)$ (here 0 or Euler's constant γ would be natural choices but we can also simply take T to be an indeterminate) and use one of them to define the divergent double zeta value $\zeta(1,k-1)$, so that this gives $2l-1$ equations in $2l-1$ unknowns. To solve them, we introduce the generating functions

$$P(x,y) = \sum_{\substack{r,s\geq 1\\ r+s=k}} \zeta(r)\zeta(s) x^{r-1} y^{s-1} \quad \text{and} \quad Q(x,y) = \sum_{\substack{m,n\geq 1\\ m+n=k}} \zeta(m,n) x^{m-1} y^{n-1},$$

with the convention $\zeta(1) = T$ and $\zeta(1,k-1) = \zeta(k-1)T - \zeta(k) - \zeta(k-1,1)$. Then the (double shuffle) relations (13.2) and (13.3) translate into equations

$$\begin{aligned} P(x,y) &= Q(x,y) + Q(y,x) + \zeta(k)\frac{x^{k-1} - y^{k-1}}{x-y} \\ &= Q(x,x+y) + Q(y,x+y). \end{aligned}$$

Using $Q(-x,-y) = -Q(x,y)$ (for k odd), allows us to solve for Q:

$$\begin{aligned} Q(x,y) &= R(x,y) + R(x-y,-y) + R(x-y,x), \\ \text{where} \quad R(x,y) &= \frac{1}{2}\left(P(x,y) + P(-x,y) - \zeta(k)\frac{x^{k-1} - y^{k-1}}{x-y} \right). \end{aligned}$$

This is equivalent (because of $\zeta(0) = -\frac{1}{2}$) to (13.1). ☐

Either of the double shuffle relations (13.2) and (13.3) permits us to express the single zeta products $\zeta(2r)\zeta(k-2r)$ in terms of all double zeta values of weight k, but we would like to do this using

(a) only the 'odd-even' values $\zeta(k-2r,2r)$, where we also include $\zeta(k)$ to have the right number of quantities, or
(b) only the 'even-odd' double zeta values $\zeta(k-2r-1,2r+1)$.

This turns out to be possible only in case (a), as we now show.

Since in case (a) we have taken $\zeta(k)$ as one of the basis elements, we can omit it from the basis and work modulo $\zeta(k)$ in the right-hand side of (13.1), which simplifies to

$$\zeta(k-2r,2r) \equiv \sum_{s=1}^{l-1}\left(\binom{2l-2s}{2l-2r}+\binom{2l-2s}{2r-1}\right)\zeta(2s)\zeta(k-2s), \qquad 1 \le r \le l-1, \tag{13.4}$$

where the congruence is modulo $\mathbb{Q}\zeta(k)$.

Theorem 13.2. *For odd $k = 2l+1 \ge 3$, the products $\zeta(2s)\zeta(k-2s)$, $1 \le s \le l-1$, are expressible in terms of double zeta values $\zeta(k-2r,2r)$, $1 \le r \le l-1$.*

Proof. Let N_k be the $(l-1)\times(l-1)$ matrix whose (r,s)-entry is the sum of binomials in (13.4). It is sufficient to show that the determinant of the matrix is non-zero.

Any binomial coefficient $\binom{m}{n}$ with m even and n odd is even, because in this case

$$\binom{m}{n} = \frac{m}{n}\binom{m-1}{n-1}.$$

Thus, the matrix N_k is congruent modulo 2 to a unipotent triangular matrix and hence has odd determinant. ☐

Remark. The immediate consequence of Theorems 12.1 and 13.2 is the following result. *For each odd $k = 2l+1 \ge 3$, the l numbers $\zeta(k)$ and $\zeta(k-2r,2r)$, $1 \le r \le l-1$, span the same space over $\mathbb{Q}$ as the l numbers*

$$\{\xi(m,n) : m+n = l-1\} \quad \text{or} \quad \{\pi^{2r}\zeta(k-2r) : 0 \le r \le l-1\}.$$

Zagier made several experimental observations about the matrix N_k which we give here as open problems.

Exercise 13.1. For $k = 2l+1 \ge 3$ and the matrix $N = N_k$ defined above, show the following.

(a) $\det N = \pm(k-2)!!$, where $(k-2)!! = 1\cdot3\cdot5\cdots(k-2)$ is the ‘double factorial’ and the sign is -1 if $l \equiv 3 \pmod 4$ and $+1$ otherwise.

(b) The entries of the inverse matrix N^{-1} are explicitly given by either of the two expressions

$$\begin{aligned}(N^{-1})_{s,r} &= \frac{-2}{2s-1}\sum_{n=0}^{k-2s}\binom{k-2r-1}{k-2s-n}\binom{n+2s-2}{n}B_n \\ &= \frac{2}{2s-1}\sum_{n=0}^{k-2s}\binom{2r-1}{k-2s-n}\binom{n+2s-2}{n}B_n, \qquad 1 \le s,r \le l-1,\end{aligned}$$

where B_n denotes the nth Bernoulli number (see Section 1).

14. The two-one (conjectural) formula

In the introductory section the following alternative version of the multiple zeta values with non-strict inequalities was mentioned (see (1.7)):

$$\zeta^\star(\boldsymbol{s}) = \zeta^\star(s_1, s_2, \dots, s_l) := \sum_{n_1 \ge n_2 \ge \dots \ge n_l \ge 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}.$$

Exercise 1.2 gives a simple recipe to pass from one model to the other.

Relation (1.8) is an example of simple relations for the multiple zeta star values; its companion is

$$\zeta^\star(\{2\}^k) = 2(1 - 2^{1-2k})\zeta(2k) = 2\sum_{n=1}^\infty \frac{(-1)^{n-1}}{n^{2k}}.$$

(This expression can be compared with the one for $\zeta(\{2\}^k)$ given in (5.4) and (12.1).)

The starting goal of our joint project with Y. Ohno was finding a general form of the two families of identities for the MZSVs. On this way, we only succeeded in generalising (1.8) but conjecturally. The particular cases of our conjecture (which we dubbed as the 'two-one formula') were established by ourselves; there are some recent publications with some other particular instances proven. One of lucky accidents of our proofs is the weighted version (2.8) of Euler's original formula (2.7) (the sum formula of depth 2 in the modern terminology).

Conjecture 7 (Two-one formula)**.** *For $k = 0, 1, 2, \dots$, denote $\mu_{2k+1} = (\{2\}^k, 1)$. Then for any admissible index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$ with* odd *entries $s_1, \dots, s_l$, the following identities are valid:*

$$\zeta^\star(\mu_{s_1}, \mu_{s_2}, \dots, \mu_{s_l}) = \sum_{\boldsymbol{p}} (-1)^{\sigma(\boldsymbol{p})} 2^{l-\sigma(\boldsymbol{p})} \zeta^\star(\boldsymbol{p}) \tag{14.1}$$

$$= \sum_{\boldsymbol{p}} 2^{l-\sigma(\boldsymbol{p})} \zeta(\boldsymbol{p}), \tag{14.2}$$

where, as in Exercise 1.2, $\boldsymbol{p}$ runs through all indices of the form $(s_1 \circ s_2 \circ \dots \circ s_l)$ with '$\circ$' being either the symbol ',' or the sign '+', and the exponent $\sigma(\boldsymbol{p})$ denotes the number of signs '+' in $\boldsymbol{p}$.

Proof of the equality of the right-hand sides in (14.1) *and* (14.2)*.* By Exercise 1.2 for the right-hand side in (14.1) we have

$$\begin{aligned} &\sum_{\{\square = ,\ \text{or}\ +\}} (-1)^{\#\{\square = +\}} 2^{l - \#\{\square = +\}} \zeta^\star(s_1 \square s_2 \square \cdots \square s_l) \\ &\quad = \sum_{\{\square = ,\ \text{or}\ +\}} \sum_{\{\circ = +\ \text{or}\ \square\}} (-1)^{l - \#\{\circ = \square\} - 1} 2^{\#\{\circ = \square\} + 1} \zeta(s_1 \circ s_2 \circ \cdots \circ s_l) \end{aligned}$$

which in the notation $r = \#\{\circ = \square\} + 1$ turns out to be

$$
\begin{aligned}
&= \sum_{n=1}^{l}\biggl(\sum_{r=n}^{l}\binom{l-n}{l-r}(-1)^{l-r}2^r\biggr)\sum_{l-\#\{\circ=+\}=n}\zeta(s_1\circ s_2\circ\dots\circ s_l)\\
&= \sum_{n=1}^{l}\biggl(\sum_{m=0}^{l-n}\binom{l-n}{l-n-m}(-1)^{l-n-m}2^{n+m}\biggr)\sum_{\#\{\circ=+\}=l-n}\zeta(s_1\circ s_2\circ\dots\circ s_l)\\
&= \sum_{n=1}^{l}2^n\sum_{\#\{\circ=+\}=l-n}\zeta(s_1\circ s_2\circ\dots\circ s_l)\\
&= \sum_{\{\circ=,\ \text{or}\ +\}}2^{l-\#\{\circ=+\}}\zeta(s_1\circ s_2\circ\dots\circ s_l),
\end{aligned}
$$

and this is exactly the right-hand side of (14.2). $\square$

On the right-hand side of (14.1) and (14.2) we have MZSVs and MZVs of length at most l, while the left-hand side involves a single zeta star attached to an index with entries 2 and 1 only (and the number of 1's is equal to l); the latter circumstance is the reason of dubbing the formula as the two-one formula.

We stress that neither the two-one formula nor its special cases treated in Theorems 14.1 and 14.2 below are specializations of identities for polylogarithms (4.1).

In spite of a nicely simple (but somehow unusual) form of the two-one formula we cannot yet prove it in the full generality. Besides the case $l = 1$ given in (1.8), the following two particular cases ($l = 2$, and $s_1 = 3$, $s_2 = \dots = s_{n-2} = 1$ with $n = l + 2 \ge 3$ arbitrary) as well as our experimental results (for cases not included in the theorems below) strongly support the validity of identities (14.1), (14.2).

Theorem 14.1. *For any $n \ge 1$ and $1 \le i \le n$,*

$$\zeta^\star(\underbrace{2,\dots,2}_{i},1,\underbrace{2,\dots,2}_{n-i},1) = 4\zeta^\star(2i+1,2n-2i+1) - 2\zeta(2n+2). \tag{14.3}$$

Theorem 14.2. *For any $n \ge 3$,*

$$
\begin{aligned}
\zeta^\star(2,\underbrace{1,\dots,1}_{n-2}) &= \sum_{\{\circ=,\ \text{or}\ +\}}2^{n-2-\#\{\circ=+\}}\zeta(3\circ\underbrace{1\circ\dots\circ1}_{n-3})\\
&= \sum_{i=2}^{n-1}2^{n-i}\sum_{e_1+e_2+\dots+e_{n-i}=i-2}\zeta(3+e_1,1+e_2,1+e_3,\dots,1+e_{n-i}),
\end{aligned}
\tag{14.4}
$$

where all e_j are non-negative integers.

Before giving some details of proofs of the theorems, let us make some comments on the two-one formula.

The formula

$$\zeta^\star(\{2,\{1\}_{m-1}\}^n,1) = (m+1)\zeta((m+1)n+1)$$

for any positive integers m, n is known (two different proofs are given by Zlobin and Ohno–Wakabayashi). If $m = 1$ it is nothing but formula (1.8), while if $m \ge 2$ then its left-hand side equals $\zeta^\star(\{\mu_3, \{\mu_1\}_{m-2}\}_n, \mu_1)$. This together with the two-one formula mean that the corresponding right-hand side in (14.1) (equivalently, in (14.2)) is expected to have a closed-form evaluation by means of the single zeta value $(m+1)\zeta((m+1)n+1)$, where the integers $m \ge 2$ and $n \ge 1$ are arbitrary.

Using the integral representation of MZSVs,

$$\zeta^\star(\boldsymbol{s}) = \int \cdots \int_{[0,1]^{s_1+\dots+s_l}} \frac{\mathrm{d}t_1 \cdots \mathrm{d}t_{s_1+\dots+s_l}}{\prod_{i=1}^{l}(1 - t_1 \cdots t_{s_1+\dots+s_i})}$$

(cf. (7.10)) valid for any admissible index $\boldsymbol{s} = (s_1, \dots, s_l)$, we can write the right-hand side of (14.1) as follows:

$$2 \int \cdots \int_{[0,1]^{s_1+\dots+s_l}} \frac{\prod_{i=1}^{l-1}(1 + t_1 \cdots t_{s_1+\dots+s_i})}{\prod_{i=1}^{l}(1 - t_1 \cdots t_{s_1+\dots+s_i})} \mathrm{d}t_1 \cdots \mathrm{d}t_{s_1+\dots+s_l}. \tag{14.5}$$

The change of variable $u_j = t_1 \cdots t_j$ for $j = 1, \dots, s_1 + \dots + s_l$ gives the integral

$$\begin{aligned} 2 \int \cdots \int_{1>u_1>\dots>u_{s_1+\dots+s_l}>0} & \prod_{i=1}^{l-1} \biggl(\prod_{j=s_1+\dots+s_{i-1}+1}^{s_1+\dots+s_i-1} \frac{\mathrm{d}u_j}{u_j} \cdot \frac{(1 + u_{s_1+\dots+s_i})\,\mathrm{d}u_{s_1+\dots+s_i}}{(1 - u_{s_1+\dots+s_i})u_{s_1+\dots+s_i}} \biggr) \\ & \times \prod_{j=s_1+\dots+s_{l-1}+1}^{s_1+\dots+s_l-1} \frac{\mathrm{d}u_j}{u_j} \cdot \frac{\mathrm{d}u_{s_1+\dots+s_l}}{1 - u_{s_1+\dots+s_l}}, \end{aligned} \tag{14.6}$$

where the empty sum $s_1+\dots+s_{i-1}$ for $i = 1$ is interpreted as 0. Therefore, any of the two integrals in (14.5), (14.6) may replace the right-hand sides of (14.1) or (14.2).

The case $l = 2$ (Theorem 14.1) reads as

$$\zeta^\star(\{2\}^{s_1}, 1, \{2\}^{s_2}, 1) = 2\zeta(2s_1 + 2s_2 + 2) + 4\zeta(2s_1 + 1, 2s_2 + 1).$$

In particular, the latter identity implies

$$\begin{aligned} &\zeta^\star(\{2\}^{s_1}, 1, \{2\}^{s_2}, 1) + \zeta^\star(\{2\}^{s_2}, 1, \{2\}^{s_1}, 1) \\ &\qquad = 4\zeta(2s_1 + 2s_2 + 2) + 4\zeta(2s_1 + 1, 2s_2 + 1) + 4\zeta(2s_2 + 1, 2s_1 + 1) \\ &\qquad = 4\zeta(2s_1 + 1)\zeta(2s_2 + 1) = \zeta^\star(\{2\}^{s_1}, 1)\zeta^\star(\{2\}^{s_2}, 1) \end{aligned}$$

whenever $s_1 \ge 1$ and $s_2 \ge 1$. However, no further generalizations to cases $l > 2$ can be derived from Conjecture 7.

The proof of Theorem 14.1 from the joint paper with Y. Ohno is an elaborate descending inductive argument on i. The following two exercises represent the summary of this proof (given in eight lemmas).

Exercise 14.1. For $a \ge c > 0$, define the harmonic sum

$$H(a, c) = \sum_{\substack{j=1 \\ j \ne a}}^{c} \frac{1}{a - j}$$

and interpret both $H(\infty, c)$ and $H(a, 0)$ as zeroes.

(a) If $B \ge C$, we have

$$\sum_{\substack{A\ge a\ge B\\ C\ge c\ge D}} \frac{1}{a^2c} = \sum_{C\ge c\ge D} \frac{H(B,c)-H(A+1,c)}{c^2} - \sum_{A\ge a\ge B} \frac{H(a,C)-H(a,D-1)}{a^2} + \delta_{B,C}\frac{2}{B^3},$$

where $\delta_{B,C}$ stands for Kronecker's delta.

(b) For positive integers L and M satisfying $L > M$, the following identity is valid:

$$\sum_{b=1}^{M} \frac{H(L,b)}{l-b} = \sum_{a=l}^{\infty}\left(\frac{1}{a-M}-\frac{1}{a}\right)H(a+1,M+1).$$

Solution of part (a). It follows that

$$\sum_{\substack{C\ge c\ge D\\ c\ne a}} \frac{1}{a-c} = H(a,C)-H(a,D-1)$$

whenever $a \ge C$, and

$$\sum_{\substack{A\ge a\ge B\\ a\ne c}} \left(\frac{1}{a-c}-\frac{1}{a}\right) = H(B,c)-H(A+1,c)+\delta_{c,B}\frac{1}{c}$$

whenever $c \le B$. Furthermore, for $a \ne c$ the following partial fraction decomposition is valid:

$$\frac{1}{a^2c} = \left(\frac{1}{a-c}-\frac{1}{a}\right)\cdot\frac{1}{c^2} - \frac{1}{a-c}\cdot\frac{1}{a^2}.$$

Thus, under the condition $B \ge C$, we get

$$\begin{aligned}\sum_{\substack{A\ge a\ge B\\ C\ge c\ge D}} \frac{1}{a^2c} &= \sum_{\substack{A\ge a\ge B\\ C\ge c\ge D\\ a\ne c}} \left(\left(\frac{1}{a-c}-\frac{1}{a}\right)\cdot\frac{1}{c^2} - \frac{1}{a-c}\cdot\frac{1}{a^2}\right) + \delta_{B,C}\frac{1}{B^3}\\ &= \sum_{C\ge c\ge D} \frac{H(B,c)-H(A+1,c)}{c^2}\\ &\qquad - \sum_{A\ge a\ge B} \frac{H(a,C)-H(a,D-1)}{a^2} + \delta_{B,C}\frac{2}{B^3}.\end{aligned}$$

which is the desired statement. □

Remark. The proof of the cyclic sum theorem (Theorem 2.3) given by Ohno and Wakabayashi exploits the general forms of above partial-fraction identities:

$$\sum_{l=1}^{m-1} \frac{1}{a^{m+1-l}c^l} = \left(\frac{1}{a-c}-\frac{1}{a}\right)\cdot\frac{1}{c^m} - \frac{1}{a-c}\cdot\frac{1}{a^m}$$

and

$$\sum_{\substack{A\ge a\ge B\\ C\ge c\ge D}}\sum_{l=1}^{m-1}\frac{1}{a^{m+1-l}c^l}=\sum_{C\ge c\ge D}\frac{H(B,c)-H(A+1,c)}{c^m}-\sum_{A\ge a\ge B}\frac{H(a,C)-H(a,D-1)}{a^m}+\delta_{B,C}\frac{m}{B^{m+1}},$$

respectively, although the function $H(a,c)$ was not used there in an explicit form (it was introduced later by Zagier in his unpublished note on the proof of Ohno–Wakabayashi). It is an open question whether the two-one formula may be generalized further to some 'multiple cyclic' level.

Exercise 14.2. (a) For $i\ge 1$ and $j\ge 0$,

$$\zeta^\star(2i+1,\underbrace{2,\dots,2}_{j},1)=-\sum_{a_0\ge a_1\ge\dots\ge a_j\ge 1}\frac{H(a_0+1,a_j)}{a_0^{2i+1}a_1^2\cdots a_j^2}+2\zeta^\star(2i+1,2j+1).$$

(b) For $i\ge 1$ and $j\ge 0$,

$$\sum_{a_0>a_1\ge\dots\ge a_{j+1}\ge 1}\left(\frac{1}{a_0-a_{j+1}}-\frac{1}{a_0}\right)\frac{1}{a_0^{2i-1}a_1^2\cdots a_{j+1}^2}=2\zeta(2i+1,2j+1).$$

(c) Given $n\ge 1$, for any i in the range $1\le i\le n$,

$$\zeta^\star(\underbrace{2,\dots,2}_{i},1,\underbrace{2,\dots,2}_{n-i},1)=4\zeta^\star(2i+1,2n-2i+1)+\sum_{a_0\ge\dots\ge a_n\ge 1}\frac{H(a_0,a_i-1)}{a_0^2\cdots a_{n-1}^2a_n}-2\sum_{a_0\ge a_{i+1}\ge\dots\ge a_n\ge 1}\frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2}. \qquad (14.7)$$

(d) We have

$$\zeta^\star(\underbrace{2,\dots,2}_{n},1,1)=4\zeta^\star(2n+1,1)-2\zeta(2n+2).$$

(e) For $1\le i<n$,

$$\sum_{a_0\ge\dots\ge a_n\ge 1}\frac{H(a_0,a_i-1)-H(a_0,a_{i+1}-1)}{a_0^2\cdots a_{n-1}^2a_n}=2\sum_{a_0>a_{i+1}\ge\dots\ge a_n\ge 1}\frac{H(a_0,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2}-\sum_{a_0\ge\dots\ge a_n\ge 1}\left(\frac{H(a_0,a_n)}{a_0^2\cdots a_{i-1}^2a_ia_{i+1}^2\cdots a_n^2}-\frac{H(a_0,a_n)}{a_0^2\cdots a_i^2a_{i+1}a_{i+2}^2\cdots a_n^2}\right).$$

(f) For $0\le i<n$,

$$\sum_{a_0\ge\dots\ge a_n\ge 1}\left(\frac{H(a_0,a_n)}{a_0^2\cdots a_{i-1}^2a_ia_{i+1}^2\cdots a_n^2}-\frac{H(a_0,a_n)}{a_0^2\cdots a_i^2a_{i+1}a_{i+2}^2\cdots a_n^2}\right)=\zeta^\star(\underbrace{2,\dots,2}_{i+1},1,\underbrace{2,\dots,2}_{n-i-1},1)-2\zeta(2n+2).$$

Proof of Theorem 14.1. We will use the descending induction on $i = n, n-1, \dots, 1$. In the case $i = n$ (induction base) the identity of the theorem is shown in Exercise 14.2 (d). Therefore, we assume that $i < n$ and that identity (14.3) is proved with i replaced by $i+1$, that is,

$$\zeta^\star(\underbrace{2,\dots,2}_{i+1},1,\underbrace{2,\dots,2}_{n-i-1},1) = 4\zeta^\star(2i+3,2n-2i-1) - 2\zeta(2n+2). \qquad (14.8)$$

We substitute expressions

$$\begin{aligned}
&\sum_{a_0>a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} \\
&\quad= \sum_{a_0>a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} \\
&\qquad+ \sum_{a_0>a_{i+1}\ge\dots\ge a_n\ge1} \biggl(\frac{1}{a_0-a_n}-\frac{1}{a_0}\biggr)\frac{1}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} \\
&\quad= \sum_{a_0>a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} + 2\zeta(2i+3,2n-2i-1),
\end{aligned}$$

followed from Exercise 14.2 (b), and

$$\begin{aligned}
&\sum_{a_0\ge\dots\ge a_n\ge1} \biggl(\frac{H(a_0,a_n)}{a_0^2\cdots a_{i-1}^2a_ia_{i+1}^2\cdots a_n^2} - \frac{H(a_0,a_n)}{a_0^2\cdots a_i^2a_{i+1}a_{i+2}^2\cdots a_n^2}\biggr) \\
&\quad= 4\zeta^\star(2i+3,2n-2i-1) - 4\zeta(2n+2) = 4\zeta(2i+3,2n-2i-1),
\end{aligned}$$

followed from Exercise 14.2 (f) and (14.8), into the identity of Exercise 14.2 (e) to get

$$\begin{aligned}
&\sum_{a_0\ge\dots\ge a_n\ge1} \frac{H(a_0,a_i-1)-H(a_0,a_{i+1}-1)}{a_0^2\cdots a_{n-1}^2a_n} = 2\sum_{a_0>a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} \\
&\quad= 2\sum_{a_0\ge a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} - 2\sum_{a_0\ge a_{i+2}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+3}a_{i+2}^2\cdots a_n^2}.
\end{aligned}$$

The last identity may be written as

$$\begin{aligned}
&\sum_{a_0\ge\dots\ge a_n\ge1} \frac{H(a_0,a_i-1)}{a_0^2\cdots a_{n-1}^2a_n} - 2\sum_{a_0\ge a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} \\
&\quad= \sum_{a_0\ge\dots\ge a_n\ge1} \frac{H(a_0,a_{i+1}-1)}{a_0^2\cdots a_{n-1}^2a_n} - 2\sum_{a_0\ge a_{i+2}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+3}a_{i+2}^2\cdots a_n^2},
\end{aligned}$$

where the right-hand side equals $-2\zeta(2n+2)$ by Exercise 14.2 (c) applied to $i+1$ instead of i and (14.8), so does the left-hand side:

$$\sum_{a_0\ge\dots\ge a_n\ge1} \frac{H(a_0,a_i-1)}{a_0^2\cdots a_{n-1}^2a_n} - 2\sum_{a_0\ge a_{i+1}\ge\dots\ge a_n\ge1} \frac{H(a_0+1,a_n)}{a_0^{2i+1}a_{i+1}^2\cdots a_n^2} = -2\zeta(2n+2). \qquad (14.9)$$

Finally, from (14.7) and (14.9) we obtain identity (14.3) for the given i, completing the proof of Theorem 14.1. ☐

Proof of Theorem 14.2. For each $i = 2, \dots, n-1$, the sum

$$\sum_{e_1+e_2+\cdots+e_{n-i}=i-2} \zeta(3+e_1, 1+e_2, 1+e_3, \dots, 1+e_{n-i})$$

can be written as

$$\sum_{e_1+e_2=i-2} \zeta(n-i+1+e_1, 1+e_2) = \sum_{l=1}^{i-1} \zeta(n-l, l)$$

by Theorem 8.4 (Ohno's relations). Therefore,

$$\begin{aligned}
&2^{n-2}\zeta(3, \underbrace{1, \dots, 1}_{n-3}) + 2^{n-3} \sum_{e_1+e_2+\cdots+e_{n-3}=1} \zeta(3+e_1, 1+e_2, \dots, 1+e_{n-3}) + \cdots \\
&\qquad + 2^{n-i} \sum_{e_1+e_2+\cdots+e_{n-i}=i-2} \zeta(3+e_1, 1+e_2, \dots, 1+e_{n-i}) + \cdots + 2\zeta(n) \\
&= \sum_{i=2}^{n-1} 2^{n-i} \sum_{l=1}^{i-1} \zeta(n-l, l) = \sum_{l=1}^{n-2} \sum_{i=2}^{n-l-1} 2^i \zeta(n-l, l) \\
&= \sum_{l=1}^{n-2} (2^{n-l} - 2)\zeta(n-l, l).
\end{aligned}$$

Applying Euler's formula (2.7) and its weighted version (2.8), the latter sum becomes

$$\sum_{l=1}^{n-2} 2^{n-l}\zeta(n-l, l) - 2\sum_{l=1}^{n-2} \zeta(n-l, l) = (n-1)\zeta(n).$$

Finally, we use the formula

$$(n-1)\zeta(n) = \zeta^\star(2, \underbrace{1, 1, \dots, 1}_{n-2})$$

which follows from Exercise 1.2 and the sum theorem (Theorem 2.4). This implies (14.4) and completes our proof of Theorem 14.2. ☐

15. Reduction of double Euler sums

Euler's original motivation to study double zeta sums was a possibility to reduce them to single zeta values. We have already discussed this problem in Section 13 for the standard multiple zeta values. Here we reproduce a result of Kentaro Ihara which addresses the alternating double Euler sums

$$\zeta(r, s; 1, \sigma) = \sum_{n > m \geq 1} \frac{\sigma^m}{n^r m^s},$$

which in fact works for any choice of σ on the unit circle $|\sigma| = 1$ (not just for $\sigma \in \{\pm 1\}$). First note the iterated integral representation

$$\zeta(r,s;1,\sigma) = \int\cdots\int_{1>t_1>\cdots>t_{r+s}>0} \frac{\mathrm{d}t_1}{t_1}\cdots\frac{\mathrm{d}t_{r-1}}{t_{r-1}}\,\frac{\mathrm{d}t_r}{1-t_r}\,\frac{\mathrm{d}t_{r+1}}{t_{r+1}}\cdots\frac{\mathrm{d}t_{r+s-1}}{t_{r+s-1}}\,\frac{\sigma\mathrm{d}t_{r+s}}{1-\sigma t_{r+s}}$$
$$= \int_0^1 \underbrace{\frac{\mathrm{d}t}{t}\cdots\frac{\mathrm{d}t}{t}}_{r-1\text{ times}}\,\frac{\mathrm{d}t}{1-t}\,\underbrace{\frac{\mathrm{d}t}{t}\cdots\frac{\mathrm{d}t}{t}}_{s-1\text{ times}}\,\frac{\sigma\mathrm{d}t}{1-\sigma t}, \tag{15.1}$$

which is given in the mnemonic form (like in (4.11)) in the second line.

Theorem 15.1. *For $k \ge 2$ and $\sigma \ne 1$,*

$$\zeta(k,1;1,\sigma) + \zeta(k,1;1,\sigma^{-1}) = \zeta(k)\big(\zeta(1;\sigma)+\zeta(1;\sigma^{-1})\big) + k\zeta(k+1) - \sum_{j=1}^{k}\zeta(j;\sigma)\zeta(k-j+1;\sigma^{-1}). \tag{15.2}$$

The particular case $\sigma = -1$ corresponds to the identity

$$\begin{aligned} 2\zeta(k,\overline{1}) = 2\sum_{n>m\ge1}\frac{(-1)^m}{n^k m} &= 2\zeta(k)\zeta(\overline{1}) + k\zeta(k+1) - \sum_{j=1}^{k}\zeta(\overline{j})\zeta(\overline{k-j+1}) \\ &= -4(1-2^{-k})\zeta(k)\log 2 + k\zeta(k+1) \\ &\quad - \sum_{j=2}^{k-1}(1-2^{1-j})(1-2^{j-k})\zeta(j)\zeta(k-j+1), \end{aligned} \tag{15.3}$$

where we use

$$\zeta(\overline{k}) = \begin{cases} -\log 2 & \text{if } k=1, \\ -(1-2^{1-k})\zeta(k) & \text{if } k>1. \end{cases}$$

Exercise 15.1. Show that the limit of the right-hand side in (15.2) as $\sigma \to 1$, $|\sigma| = 1$, exists and deduce the corresponding identity (due to Euler)

$$2\zeta(k,1) = k\zeta(k+1) - \sum_{j=2}^{k-1}\zeta(j)\zeta(k-j+1). \tag{15.4}$$

In order to prove Theorem 15.1 we use the shuffle and stuffle relations for the corresponding alternating double sums.

Lemma 15.1. *We have*

$$\zeta(k)\big(\zeta(1;\sigma)+\zeta(1;\sigma^{-1})\big) = \sum_{j=1}^{k}\big(\zeta(j,k+1-j;\sigma,\sigma^{-1}) + \zeta(k+1-j,j;\sigma^{-1},\sigma)\big) + \zeta(k,1;1,\sigma) + \zeta(k,1;1,\sigma^{-1}).$$

Proof. The shuffle product of $\zeta(1;\sigma)$ and $\zeta(k)$ reads

$$\begin{aligned}\zeta(1;\sigma)\zeta(k) &= \int_0^1 \frac{\sigma \mathrm{d}t}{1-\sigma t} \cdot \int_0^1 \underbrace{\frac{\mathrm{d}t}{t}\cdots\frac{\mathrm{d}t}{t}}_{k-1 \text{ times}} \frac{\mathrm{d}t}{1-t} \\ &= \sum_{j=1}^{k} \int_0^1 \underbrace{\frac{\mathrm{d}t}{t}\cdots\frac{\mathrm{d}t}{t}}_{j-1 \text{ times}} \frac{\sigma \mathrm{d}t}{1-\sigma t} \underbrace{\frac{\mathrm{d}t}{t}\cdots\frac{\mathrm{d}t}{t}}_{k-j \text{ times}} \frac{\mathrm{d}t}{1-t} \\ &\quad + \int_0^1 \underbrace{\frac{\mathrm{d}t}{t}\cdots\frac{\mathrm{d}t}{t}}_{k-1 \text{ times}} \frac{\mathrm{d}t}{1-t} \frac{\sigma \mathrm{d}t}{1-\sigma t} \\ &= \sum_{j=1}^{k} \zeta(j, k+1-j; \sigma, 1/\sigma) + \zeta(k,1;1,\sigma).\end{aligned}$$

It remains to add the equation obtained by replacing σ with σ^{-1}. □

Lemma 15.2. *The following identity is valid:*

$$\sum_{j=1}^{k} \zeta(j;\sigma)\zeta(k+1-j;\sigma^{-1}) = \sum_{j=1}^{k} \big(\zeta(j,k+1-j;\sigma,\sigma^{-1}) + \zeta(k+1-j,j;\sigma^{-1},\sigma)\big) + k\zeta(k+1).$$

Proof. By the shuffle product (that is, term-by-term multiplication of the corresponding series),

$$\zeta(j;\sigma)\zeta(k+1-j;\varepsilon) = \zeta(j,k+1-j;\sigma,\varepsilon) + \zeta(k+1-j,j;\varepsilon,\sigma) + \zeta(k+1;\sigma\varepsilon).$$

Putting $\varepsilon = \sigma^{-1}$ and summing for j from 1 to k, the result follows. □

Proof of Theorem 15.1*.* The identity follows by applying Lemmas 15.1 and 15.2. □

16. q-Analogues of MZVs

The classical idea of introducing an additional parameter to an expression or formula we wish to deal with, is quite fruitful in many situations. This may simplify a proof of the corresponding identity or lead to a more general identity which have several other useful specializations of the introduced parameter. We have already experienced the usefulness of the method on the example of functional models of generalised polylogarithms in Section 4 and of (no name) function in Section 7. They were used for proving the shuffle and stuffle relations of MZVs, respectively. Because they (are expected to) satisfy only 'half' of relations of MZVs, we can hardly use them as a live imitation of the latter numbers.

The story of introducing the parameter q (or, the 'quantum' parameter) often has a different flavor. Note that the basic idea is simply to replace a number n (not necessarily an integer!) by the function $[n] = [n]_q := (1 - q^n)/(1 - q)$; this is, of course, nothing else but a polynomial for positive $n \in \mathbb{Z}$. The actual motivation of the replacement has strong analytic grounds:

$$\lim_{\substack{q\to 1\\ 0<q<1}} [n]_q = n,$$

so that the (sometimes formal) limit as $q \to 1$ produces back the original limits. Note however that this is only a part of the recipe, as multiplying the 'q-number' $[n]_q$ by any power of q makes exactly the same job as $q \to 1$. *Getting the right exponents of q is an art.*

The main requirement from a q-model of MZVs (or MZSVs) is a better understanding of the structure of linear and algebraic relations between the corresponding numbers. An important advantage of the q-model is that proving the absence of such relations and guessing their existence are usually a much easier task: for example, the linear independence of any version of q-MZVs (and much more) is known, while just the irrationality of odd single zeta values seems to be hard. On the other hand, showing that some relations hold is normally easier for numbers than for functions. The main problem here is finding an appropriate q-analogue which is often dictated by already existing proofs of the corresponding original identities.

An unfortunate thing about MZVs is that there is no uniform q-generalization of the multiple zeta (star) values. Having however several q-analogues in mind and a simple way to pass from one q-model to another gives one a very natural parallel between the numbers and their q-analogues.

There are very good reasons to believe that the most perfect q-extension of MZVs is given by

$$\zeta_q(s_1, s_2, \dots, s_l) = \sum_{n_1>n_2>\dots>n_l\ge 1} \frac{q^{n_1(s_1-1)+n_2(s_2-1)+\dots+n_l(s_l-1)}}{[n_1]^{s_1}[n_2]^{s_2}\cdots[n_l]^{s_l}}, \tag{16.1}$$

where conditions on the multi-index $\boldsymbol{s} = (s_1, \dots, s_l)$ are exactly the same as for the MZVs (1.5) (that is, the multi-index is admissible). The corresponding q-analogues of the values of Riemann's zeta function are in this case as follows:

$$\zeta_q(s) = \sum_{n\ge1} \frac{q^{n(s-1)}}{[n]^s}.$$

The q-model (16.1) inherits many relations available for MZVs $\zeta(\boldsymbol{s})$. There is a version of stuffle relations, which is based on the identity

$$\frac{q^{n(s-1)}}{[n]_q^s}\,\frac{q^{m(r-1)}}{[m]_q^r}\bigg|_{m=n} = (1-q)\frac{q^{n(s+r-2)}}{[n]_q^{s+r-1}} + \frac{q^{n(s+r-1)}}{[n]_q^{s+r}};$$

there is however no reasonably nice version of shuffle relations. The following result of Okuda and Takeyama, which includes numerous implications, is a convincing argument to count the q-MZVs (16.1) appropriate enough. In order to state it, we define the *height* $m = m(\boldsymbol{s})$ of a multi-index $\boldsymbol{s} = (s_1,\dots,s_l)$ to be the number of components satisfying $s_j > 1$; for an admissible $\boldsymbol{s}$ we have $s_1 > 1$, so that $m(\boldsymbol{s}) \ge 1$. Denote the set of admissible multi-indices of fixed weight $w = |\boldsymbol{s}|$, length $l = \ell(\boldsymbol{s})$ and height $m = m(\boldsymbol{s})$ by $I_0(w,l,m)$, and set

$$\Phi_q(x,y,z) := \sum_{w,l,m=0}^{\infty} x^{w-l-m}y^{l-m}z^{m-1} \sum_{\boldsymbol{s}\in I_0(w,l,m)} \zeta_q(\boldsymbol{s}).$$

Theorem 16.1. *The generating function Φ_q is given by*

$$\begin{aligned} 1 + (z-xy)\Phi_q(x,y,z) &= \prod_{n=1}^{\infty} \frac{([n]_q - \alpha q^n)([n]_q - \beta q^n)}{([n]_q - xq^n)([n]_q - yq^n)} \\ &= \exp\biggl(\sum_{k=2}^{\infty} \frac{x^k + y^k - \alpha^k - \beta^k}{k} \sum_{j=2}^{k} (q-1)^{k-j}\zeta_q(j)\biggr), \qquad (16.2) \end{aligned}$$

where α and β are determined by

$$\alpha + \beta = x + y + (q-1)(z-xy), \quad \alpha\beta = z.$$

In particular, the sum of the multiple q-zeta values of fixed weight, length and height is a polynomial in q and single q-zeta values.

The limiting case $q \to 1$ was established earlier by Ohno and Zagier.

Corollary 1. *We have the generating function identity*

$$\begin{aligned} &\sum_{s,r=0}^{\infty} x^{s+1}y^{r+1}\zeta_q(s+2,\{1\}^r) \\ &\qquad = \exp\biggl(\sum_{k=2}^{\infty} \frac{x^k + y^k - (x+y+(1-q)xy)^k}{k} \sum_{j=2}^{k} (q-1)^{k-j}\zeta_q(j)\biggr). \end{aligned}$$

In particular, because of the symmetry in x and y,

$$\zeta_q(s+2,\{1\}^r) = \zeta_q(r+2,\{1\}^s).$$

Proof. The identity follows by taking $z = 0$ in (16.2). ∎

Corollary 2 (Sum theorem)**.** *The sum of all admissible multiple q-zeta values of fixed weight w and fixed length is equal to $\zeta_q(w)$,*

$$\sum_{\boldsymbol{s}:|\boldsymbol{s}|=w,\ \ell(\boldsymbol{s})=l} \zeta_q(\boldsymbol{s}) = \zeta_q(w).$$

Proof. This derivation is more subtle. Taking the limit as $z \to xy$ in (16.2) gives

$$\begin{aligned}\Phi_q(x,y,xy) &= \sum_{r=1}^{\infty} \frac{q^r}{([r]_q - xq^r)([r]_q - yq^r)} \\ &= \sum_{r=1}^{\infty} \frac{q^r}{[r]_q^2}\left(1 - \frac{xq^r}{[r]_q}\right)^{-1}\left(1 - \frac{yq^r}{[r]_q}\right)^{-1} \\ &= \sum_{m,n=0}^{\infty} x^m y^n \zeta_q(m+n+2) = \sum_{w>l\ge1} x^{w-l-1}y^{l-1}\zeta_q(w).\end{aligned}$$

On the other hand, it follows directly from definition that

$$\Phi_q(x,y,xy) = \sum_{w,l=0}^{\infty} x^{w-l-1}y^{l-1} \sum_{\boldsymbol{s}:|\boldsymbol{s}|=w,\ \ell(\boldsymbol{s})=l} \zeta_q(\boldsymbol{s}).$$

It remains to compare the coefficients in the two representations of $\Phi_q(x,y,xy)$. □

Exercise 16.1. For an indeterminate z, show

$$\sum_{n_1>\dots>n_l\ge1} \frac{q^{n_1}}{[n_1]_q} \prod_{j=1}^{l} \frac{1}{[n_j]_q - zq^{n_j}} = \sum_{n=1}^{\infty} \frac{q^{ln}}{[n]_q^l([n]_q - zq^n)}.$$

Hint. This is equivalent to the sum theorem in Corollary 2. □

In spite of the above 'naturalness' of the q-MZVs (16.1), there are other variations, and we indicate more in what follows. The main difficulty of all these q-models occurs when we look for a reasonable q-generalization of the shuffle product from Theorem 3.1, the product originated from the differential equations for the multiple polylogarithms (4.1). Lemma 4.1 tells us that

$$\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{s_1,s_2,\dots,s_l}(z) = \begin{cases} \dfrac{1}{z}\operatorname{Li}_{s_1-1,s_2,\dots,s_l}(z) & \text{if } s_1 \ge 2, \\ \dfrac{1}{1-z}\operatorname{Li}_{s_2,\dots,s_l}(z) & \text{if } s_1 = 1, \end{cases} \tag{16.3}$$

and this comes from the *fundamental theorem of calculus*,

$$\frac{\mathrm{d}}{\mathrm{d}z}\big(f(z)g(z)\big) = \frac{\mathrm{d}}{\mathrm{d}z}f(z)\cdot g(z) + f(z)\cdot\frac{\mathrm{d}}{\mathrm{d}z}g(z). \tag{16.4}$$

The differential equations (16.3) give rise to an integral representation of the polylogarithms (4.1) (hence, of the multiple zeta values), where the participating differential forms $\mathrm{d}z/z$ and $\mathrm{d}z/(1-z)$ are assigned as two non-commutative letters, so that the integrals themselves are interpreted as words on these letters.

The q-analogue of (16.4) reads as

$$D_q\big(f(z)g(z)\big) = D_q f(z)\cdot g(z) + f(z)\cdot D_q g(z) - (1-q)z\cdot D_q f(z)\cdot D_q g(z), \quad (16.5)$$

where

$$D_q f(z) = \frac{f(z)-f(qz)}{(1-q)z}.$$

Defining a q-analogue of the multiple polylogarithms (4.1) as

$$\mathrm{Li}_{s_1,\dots,s_l}(z;q) = \sum_{n_1>\dots>n_l\ge1} \frac{z^{n_1}}{[n_1]^{s_1}\cdots[n_l]^{s_l}}, \quad (16.6)$$

from (16.5) we deduce the following analogue of (16.3):

$$D_q \mathrm{Li}_{s_1,s_2,\dots,s_l}(z;q) = \begin{cases} \dfrac{1}{z}\mathrm{Li}_{s_1-1,s_2,\dots,s_l}(z;q) & \text{if } s_1\ge 2,\\ \dfrac{1}{1-z}\mathrm{Li}_{s_2,\dots,s_l}(z;q) & \text{if } s_1 = 1.\end{cases}$$

This q-model of the multiple polylogarithms, together with classical formulae in the theory of basic hypergeometric series (which we 'touch' below), were used in the derivation of Theorem 16.1 by Okuda and Takeyama. This is a reason to believe that the q-multiple polylogarithms (16.6) are 'motivated' q-analogues of (4.1), and that their values at $z=q$,

$$\begin{aligned}\mathfrak{z}_q(s_1,s_2,\dots,s_l) &= (1-q)^{-|\boldsymbol{s}|}\,\mathrm{Li}_{s_1,s_2,\dots,s_l}(q;q)\\ &= \sum_{n_1>n_2>\dots>n_l\ge1} \frac{q^{n_1}}{(1-q^{n_1})^{s_1}(1-q^{n_2})^{s_2}\cdots(1-q^{n_l})^{s_l}}, \end{aligned} \quad (16.7)$$

are reasonable q-analogues of multiple zeta values. Note the normalization factor $(1-q)^{-|\boldsymbol{s}|}$ in the latter specialization; it makes many formulae for q-MZVs 'cleaner' and could be also used for the q-model (16.1).

Although the rule (16.5) might be interpreted as a shuffle product of a suitable functional q-model of the multiple polylogarithms and the corresponding q-MZVs, these models are different from and even 'incompatible' with already given models. For example, the q-analogue of the formula

$$\mathrm{Li}_1(z)^r = r!\,\mathrm{Li}_{\{1\}_r}(z)$$

(cf. Exercise 4.2 (a)) in terms of (16.6) involve certain undesired 'parasites': if $r=2$, from

$$D_q\big(\mathrm{Li}_1(z;q)\,\mathrm{Li}_1(z;q)\big) = \frac{1}{1-z}\mathrm{Li}_1(z;q) + \mathrm{Li}_1(z;q)\frac{1}{1-z} - (1-q)\frac{z}{(1-z)^2}$$

we have

$$\mathrm{Li}_1(z;q)^2 = 2\,\mathrm{Li}_{1,1}(z;q) - (1-q)\sum_{n=1}^{\infty}\frac{(n-1)z^n}{[n]},$$

where the latter series cannot be expressed by means of (16.6).

A related problem is a q-generalization of Euler's decomposition formula

$$\zeta(r)\zeta(s)=\sum_{i=0}^{r-1}\binom{s-1+i}{i}\zeta(s+i,r-i)+\sum_{i=0}^{s-1}\binom{r-1+i}{i}\zeta(r+i,s-i) \quad (16.8)$$

(which follows from the double shuffle relations (13.2), (13.3)), since the known proofs make use (explicitly or not) of the shuffle relations. It seems that a way to overcome this difficulty is to extend the algebra of q-MZVs *differentially*, that is, to consider a differential algebra of q-MZVs and all their δ-derivatives of arbitrary order, where $\delta=q\frac{\mathrm{d}}{\mathrm{d}q}$. Although it is hard to justify this claim, let us see how the problem may be fixed on the example of a q-analogue of (16.8) when $r=s=2$,

$$\zeta(2)^2=2\zeta(2,2)+4\zeta(3,1), \quad (16.9)$$

by means of (16.7). As Bradley shows, even this particular case involves something, which is not expressible by means of q-MZVs (16.1).

We start with the partial-fraction identity

$$\frac{1}{(1-x)(1-y)}=\frac12\big(f(x,y)+f(y,x)\big), \qquad \text{where} \quad f(x,y)=\frac{1+x}{(1-x)(1-xy)},$$

and differentiate both sides with respect to x and y,

$$\frac{\partial f(x,y)}{\partial x\,\partial y}=\frac{2}{(1-x)^2(1-xy)^2}+\frac{4}{(1-x)(1-xy)^3}-\frac{4}{(1-x)(1-xy)^2}-\frac{1+xy}{(1-xy)^3}.$$

Multiplying the result by xy, substituting $x=q^n$ and $y=q^m$, and using

$$\begin{aligned}\sum_{n,m=1}^{\infty}\frac{xy(1+xy)}{(1-xy)^3}\bigg|_{x=q^n,\,y=q^m}&=\sum_{l=1}^{\infty}(l-1)\frac{q^l(1+q^l)}{(1-q^l)^3}\\ &=\delta\sum_{l=1}^{\infty}\frac{q^l}{(1-q^l)^2}-\sum_{l=1}^{\infty}\frac{q^l(1+q^l)}{(1-q^l)^3}=\delta\mathfrak{z}_q(2)-2\mathfrak{z}_q(3)+\mathfrak{z}_q(2),\end{aligned}$$

we finally arrive at

$$\mathfrak{z}_q(2)^2+\delta\mathfrak{z}_q(2)=2\mathfrak{z}_q(2,2)+4\mathfrak{z}_q(3,1)-4\mathfrak{z}_q(2,1)+2\mathfrak{z}_q(3)-\mathfrak{z}_q(2),$$

which is the desired q-analogue of (16.9).

One can also use Ramanujan's system of differential equations (16.11) to get rid of the term $\delta\mathfrak{z}_q(2)$. Namely, using

$$\delta\mathfrak{z}_q(2)=\mathfrak{z}_q(2)-5\mathfrak{z}_q(3)+5\mathfrak{z}_q(4)-2\mathfrak{z}_q(2)^2$$

we obtain

$$\mathfrak{z}_q(2)^2=-2\mathfrak{z}_q(2,2)-4\mathfrak{z}_q(3,1)+4\mathfrak{z}_q(2,1)+5\mathfrak{z}_q(4)-7\mathfrak{z}_q(3)+2\mathfrak{z}_q(2),$$

which is also a q-analogue of (16.9). But for a general q-analogue of (16.8) we do expect terms involving $\delta\mathfrak{z}_q(s)$ and $\delta\mathfrak{z}_q(t)$, hence working in the δ-*differential* algebra generated by the multiple q-zeta values (16.7). Is there a nice form of double shuffle relations in this differential algebra?

There is also an arithmetically motivated q-model, but for single (non-multiple) zeta values:

$$\tilde{\zeta}_q(s)=\sum_{n=1}^{\infty}\sigma_{s-1}(n)q^n=\sum_{n=1}^{\infty}\frac{n^{s-1}q^n}{1-q^n}\,,\qquad s=1,2,\dots, \tag{16.10}$$

where $\sigma_{s-1}(n)=\sum_{d|n}d^{s-1}$ denotes the sum of powers of the divisors. These can be readily recalculated in terms of the q-zeta values (16.1) and (16.7) with $l=1$, because

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

and, in general,

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

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

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

The latter imply $\rho_{k+1}(1)=k!$ that results in the limiting relations

$$\lim_{\substack{q\to1\\0<q<1}}(1-q)^s\tilde{\zeta}_q(s)=(s-1)!\cdot\zeta(s),\qquad s=2,3,\dots.$$

If $s\ge2$ 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 defined in (1.2), are known as the *Eisenstein series*. This circumstance allows to prove the coincidence of the rings

$$\mathbb{Q}[q,\tilde{\zeta}_q(2),\tilde{\zeta}_q(4),\tilde{\zeta}_q(6),\tilde{\zeta}_q(8),\tilde{\zeta}_q(10),\dots]\quad\text{and}\quad\mathbb{Q}[q,\tilde{\zeta}_q(2),\tilde{\zeta}_q(4),\tilde{\zeta}_q(6)];$$

the fact can be viewed as a q-analogue of the coincidence of the numerical rings

$$\mathbb{Q}[\zeta(2),\zeta(4),\zeta(6),\zeta(8),\zeta(10),\dots]\quad\text{and}\quad\mathbb{Q}[\zeta(2)]=\mathbb{Q}[\pi^2]$$

which we proved in Lemma 1.2. Even more, the ring $\mathbb{Q}[q,\tilde{\zeta}_q(2),\tilde{\zeta}_q(4),\tilde{\zeta}_q(6)]$ is *differentially stable* because of Ramanujan's system of differential equations

$$\delta E_2=\frac{1}{12}(E_2^2-E_4),\quad \delta E_4=\frac{1}{3}(E_2E_4-E_6),\quad \delta E_6=\frac{1}{2}(E_2E_6-E_4^2), \tag{16.11}$$

where, as before, $\delta=q\frac{\mathrm{d}}{\mathrm{d}q}$.

There are other examples of q-generalizations of both MZVs and generalised polylogarithms, motivated by the theory of modular forms, basic (q-) hypergeometric series and mathematical physics. They are not yet systematically investigated. A basic example here is related to the *q-exponential function*

$$e(z)=e_q(z)=\frac{1}{\prod_{m=0}^{\infty}(1-zq^m)}=\frac{1}{(z;q)_\infty}, \tag{16.12}$$

where $z \in \mathbb{C}$, $|z| < 1$. Here we use the standard q-Pochhammer notation (cf. Section 5)

$$(a;q)_n = \begin{cases} 1 & \text{if } n = 0, \\ \prod_{m=0}^{n-1}(1-aq^m) & \text{if } n \ge 1, \end{cases}$$

which, of course, has perfect sense for $n = \infty$ as well, because $|q| < 1$. The similarity with the classical exponential function comes from the expansion

$$e_q(z) = \sum_{n=0}^{\infty} \frac{z^n}{(q;q)_n},$$

which is the special case $x = 0$, $y = z$ of the q-binomial theorem

$$\sum_{n=0}^{\infty} \frac{(x;q)_n}{(q;q)_n} y^n = \frac{(xy;q)_\infty}{(y;q)_\infty}. \tag{16.13}$$

The q-polynomials

$$[n]_q! = \prod_{k=1}^{n} [k]_q = \prod_{k=1}^{n} \frac{1-q^k}{1-q} = \frac{(q;q)_n}{(1-q)^n},$$

and so the product $(q;q)_n$, are regarded as natural q-extensions of $n!$. Moreover, the function $e(z)$ satisfies the 'standard' exponential functional identity

$$e(X+Y) = e(X)e(Y),$$

if $e(X) = e_q(X)$, $e(Y) = e_q(Y)$ and $e(X+Y) = e_q(X+Y)$ are viewed as elements in the algebra $\mathbb{C}_q[[X,Y]]$ of formal power series in two elements X, Y linked by the commutation relation $XY = qYX$. This noncommutative combinatorial interpretation was given by Schutzenberger in the 1950s.

On the other hand, from (16.12) we have the asymptotic behaviour

$$\begin{aligned} \log e(z) &= \sum_{n=0}^{\infty} \bigl(-\log(1-q^n z)\bigr) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{q^{mn} z^m}{m} = \sum_{m=1}^{\infty} \frac{z^m}{m(1-q^m)} \\ &= \frac{1}{1-q} \sum_{m=1}^{\infty} \frac{z^m}{m[m]_q} \sim \frac{-1}{\log q} \sum_{m=1}^{\infty} \frac{z^m}{m^2} \qquad \text{as} \quad q \to 1 \end{aligned} \tag{16.14}$$

(known already to Ramanujan), since $-\log q \sim 1-q$ as $q \to 1$. This allows to think of $\log e(z)$ as of a q-analogue of the dilogarithm function

$$\mathrm{Li}_2(z) = \sum_{m=1}^{\infty} \frac{z^m}{m^2},$$

the *quantum dilogarithm*. This analogy is much deeper than just the asymptotics above because it is not hard to check that the q-binomial theorem (16.13) is equivalent to the so-called quantum pentagonal identity

$$e(X)e(Y) = e(Y)e(-YX)e(X), \tag{16.15}$$

where as before $e(X) = e_q(X)$, $e(Y) = e_q(Y)$ and $e(-YX) = e_q(-YX)$ are elements in the algebra $\mathbb{C}_q[[X,Y]]$ of formal power series in two elements X, Y linked by the commutation relation $XY = qYX$. It seems that Richmond and Szekeres were the first to realise that the limiting case $q \to 1$ of certain q-hypergeometric identities (actually, they considered the Andrews–Gordon generalisation of the Rogers–Ramanujan identities) produces non-trivial identities for the dilogarithm values; the argument was later exploited by Loxton and rediscovered in the context of (16.13), (16.15) by Faddeev and Kashaev.

Theorem 16.2. *The limiting case $q \to 1$ of the q-binomial theorem* (16.13) *is the equality*

$$\operatorname{Li}_2(x) + \operatorname{Li}_2(y) = \operatorname{Li}_2\biggl(\frac{x}{1-y}\biggr) + \operatorname{Li}_2\biggl(\frac{y}{1-x}\biggr) - \operatorname{Li}_2\biggl(\frac{xy}{(1-x)(1-y)}\biggr) - \log(1-x)\log(1-y), \qquad 0<x<1, \quad 0<y<1. \tag{16.16}$$

Remark. Although we prove relation (16.16) for x and y restricted to the interval $(0,1)$, and this positivity is always crucial in application of the allied asymptotical formulae, the identity remains valid for $x, y \in \mathbb{C} \setminus (1, +\infty)$ by analytic continuation.

Formula (16.16) is due to Abel but an equivalent formula was published by Spence nearly twenty years earlier. Another equivalent form of (16.16) (see (16.27) below) was given by Rogers.

Proof. Without loss of generality assume that q is sufficiently close to 1, namely, that

$$\max\{x, y, 1-y(1-x)\} < q < 1.$$

The easy part of the theorem is the asymptotics of the right-hand side in (16.13):

$$\log\frac{(xy;q)_\infty}{(y;q)_\infty} = \log\frac{e(y)}{e(xy)} \sim \frac{1}{\log q}\bigl(\operatorname{Li}_2(xy) - \operatorname{Li}_2(y)\bigr) \qquad \text{as} \quad q \to 1, \tag{16.17}$$

which is obtained on the basis of (16.14).

For the left-hand side of (16.13), write

$$\sum_{n=0}^\infty \frac{(x;q)_n}{(q;q)_n} y^n = \sum_{n=0}^\infty c_n, \qquad \text{where} \quad c_n = \frac{(x;q)_n}{(q;q)_n} y^n > 0. \tag{16.18}$$

Then the sequence

$$d_n = \frac{c_{n+1}}{c_n} = \frac{1-xq^n}{1-q^{n+1}} y > 0, \qquad n = 0, 1, 2, \dots, \tag{16.19}$$

satisfies

$$\frac{d_{n+1}}{d_n} = \frac{(1-xq^{n+1})(1-q^{n+1})}{(1-xq^n)(1-q^{n+2})} = 1 - \frac{q^n(1-q)(q-x)}{(1-xq^n)(1-q^{n+2})} < 1 - q^n(1-q)(q-x), \qquad n = 0, 1, 2, \dots \tag{16.20}$$

(we use $0 < x < q < 1$), hence it is strictly decreasing. On the other hand, $1 - y(1-x) < q$ implies

$$d_0 = \frac{c_1}{c_0} = \frac{1-x}{1-q} y > 1,$$

while

$$\lim_{n\to\infty} d_n = \lim_{n\to\infty} \frac{1-xq^n}{1-q^{n+1}} y = y < 1;$$

thus, there exists the unique index $N \ge 1$ such that

$$d_{N-1} = \frac{c_N}{c_{N-1}} \ge 1 \qquad \text{and} \qquad d_N = \frac{c_{N+1}}{c_N} < 1. \tag{16.21}$$

Solving the inequality $c_{n+1}/c_n < 1$ or, equivalently, $(1-xq^n)y < 1-q^{n+1}$ we obtain $n > T$, where

$$T = \frac{1}{\log q} \cdot \log \frac{1-y}{q-xy}, \tag{16.22}$$

hence $N = \lfloor T \rfloor$, the integral part of T. From (16.19)–(16.21) we conclude that c_N is the main term contributing the sum in (16.18), namely,

$$1 < \frac{\sum_{n=0}^{\infty} c_n}{c_N} < \text{const}\,.$$

This implies

$$\begin{aligned} \log \sum_{n=0}^{\infty} c_n \sim \log c_N &= \log\biggl(\frac{e(q)e(xq^N)}{e(x)e(q^{N+1})} y^N \biggr) \\ &\sim \log\biggl(\frac{e(q)e(xq^T)}{e(x)e(q^{T+1})} y^T \biggr) \qquad \text{as} \quad q \to 1. \end{aligned} \tag{16.23}$$

Note now that from (16.22)

$$q^T = \frac{1-y}{q-xy},$$

whence the asymptotics in (16.23) may be continued as follows:

$$\begin{aligned} \log \sum_{n=0}^{\infty} c_n &\sim \log e(q) + \log e\biggl(x\frac{1-y}{q-xy} \biggr) - \log e(x) - \log e\biggl(q\frac{1-y}{q-xy} \biggr) \\ &\qquad + \frac{\log y}{\log q} \cdot \log \frac{1-y}{q-xy} \\ &\sim \frac{1}{\log q} \biggl(\mathrm{Li}_2(x) + \mathrm{Li}_2\biggl(\frac{1-y}{1-xy} \biggr) - \mathrm{Li}_2(1) - \mathrm{Li}_2\biggl(x\frac{1-y}{1-xy} \biggr) \\ &\qquad + \log y \cdot \log \frac{1-y}{1-xy} \biggr) \qquad \text{as} \quad q \to 1, \end{aligned} \tag{16.24}$$

where (16.14) is used.

Comparing the asymptotics (16.17) and (16.24) of the both sides of (16.13) we arrive at the identity

$$\begin{aligned} \mathrm{Li}_2(x) + \mathrm{Li}_2\biggl(\frac{1-y}{1-xy} \biggr) - \mathrm{Li}_2(1) - \mathrm{Li}_2\biggl(x\frac{1-y}{1-xy} \biggr) + \log y \cdot \log \frac{1-y}{1-xy} \\ = \mathrm{Li}_2(xy) - \mathrm{Li}_2(y). \end{aligned} \tag{16.25}$$

Take $x = 0$ in (16.25) to get

$$\mathrm{Li}_2(y) + \mathrm{Li}_2(1-y) - \mathrm{Li}_2(1) + \log y \cdot \log(1-y) = 0.$$

This identity, in particular, implies

$$\operatorname{Li}_2\biggl(\frac{1-y}{1-xy}\biggr)-\operatorname{Li}_2(1)=-\operatorname{Li}_2\biggl(1-\frac{1-y}{1-xy}\biggr)-\log\frac{1-y}{1-xy}\cdot\log\biggl(1-\frac{1-y}{1-xy}\biggr). \quad (16.26)$$

Substituting (16.26) into (16.25) results in

$$\operatorname{Li}_2(xy)+\operatorname{Li}_2\biggl(\frac{x(1-y)}{1-xy}\biggr)+\operatorname{Li}_2\biggl(\frac{y(1-x)}{1-xy}\biggr)+\log\frac{1-y}{1-xy}\cdot\log\frac{1-x}{1-xy}$$
$$=\operatorname{Li}_2(x)+\operatorname{Li}_2(y). \quad (16.27)$$

Finally, changing variable $\tilde x=x(1-y)/(1-xy)$, $\tilde y=y(1-x)/(1-xy)$, hence

$$1-\tilde x=\frac{1-x}{1-xy}, \quad 1-\tilde y=\frac{1-y}{1-xy}, \qquad x=\frac{\tilde x}{1-\tilde y}, \quad y=\frac{\tilde y}{1-\tilde x},$$

reduces identity (16.27) to the required form (16.16). □

A similar 'mixed' q-extension of the multiple zeta values might be possible. The following example is due to Zagier.

Theorem 16.3. *The following identity is valid:*

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{mn(m+n)}\,\frac{q^{m+n}}{(1-q^m)(1-q^n)(1-q^{m+n})}=\sum_{m=1}^{\infty}\frac{1}{6m^3}\,\frac{q^{2m}(3-q^m)}{(1-q^m)^3}\,. \quad (16.28)$$

Remark. The limiting case as $q\to 1$ of the identity reads

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{m^2n^2(m+n)^2}=\frac{1}{3}\sum_{m=1}^{\infty}\frac{1}{m^6}=\frac{\pi^6}{2835}\,.$$

Although the left-hand side here is not a standard MZV, the identity reduces it to a single zeta value.

By comparing the coefficients of q^N, we see that (16.28) is equivalent to the number-theoretic identity

$$\sum_{\substack{m,n,r,s>0\\ mr+ns=N}}\frac{\min(r,s)}{mn(m+n)}=\frac{\sigma_5(N)-\sigma_3(N)}{6N^3}, \qquad N\in\mathbb{N}.$$

Lemma 16.1. *Let $\{\alpha(m,n)\}_{m,n\in\mathbb{N}}$ be a collection of complex numbers which can be written in the form*

$$\alpha(m,n)=\beta(m,n)-\beta(m+n,n)-\beta(m,m+n), \qquad m,n\in\mathbb{N}, \quad (16.29)$$

where $\sum_{m,n>0}\beta(m,n)$ is absolutely convergent. Then

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\alpha(m,n)=\sum_{m=1}^{\infty}\beta(m,m).$$

Proof. It follows from (16.29) that

$$\sum_{m,n>0}\alpha(m,n)=\biggl(\sum_{m,n>0}-\sum_{m>n>0}-\sum_{n>m>0}\biggr)\beta(m,n)=\sum_{m=n>0}\beta(m,n),$$

which is the wanted identity. □

Proof of Theorem 16.3. It is interesting that the summand on the left-hand side of (16.28) cannot be given in the form (16.29). However, there is an identity of this type at the level of *derivatives*, namely

$$q\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{1}{mn(m+n)}\,\frac{q^{m+n}}{(1-q^m)(1-q^n)(1-q^{m+n})}\right)=\beta(m,n)-\beta(m+n,n)-\beta(m,m+n)$$

with

$$\beta(m,n)=\frac{1}{m}\,\frac{q^m}{(1-q^m)^2}\cdot\frac{1}{n}\,\frac{q^n}{(1-q^n)^2},$$

and now the required identity follows from Lemma 16.1 and

$$\beta(m,m)=q\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{q^{2m}(3-q^m)}{6m^3(1-q^m)^3}\right)$$

after integration. □