MZFs19 / NWI-WM149: Multiple Zeta Functions (Notes for Spring 2018/19)

22 May 2019

Wadim Zudilin

IMAPP, Radboud University, PO Box 9010, 6500 GL Nijmegen, Netherlands

E-mail address: w.zudilin@math.ru.nl

URL: http://www.math.ru.nl/~wzudilin/

Contents

We have to set up some formal prerequisites for this course to make it successful. The course assumes basic linear algebra (vector space, subspace, quotient space, dimension, linear map, matrix), basic analysis in one variable (sequences, series and their convergence properties; continuous or differentiable functions, power series with ratio of convergence), and fair knowledge of algebra (basics of groups, rings, fields, quotient groups, quotient rings, finite fields) and complex analysis (holomorphic functions, Cauchy integral, residue sum formula). There are many textbooks that cover those topics, and it is assumed that students have successfully gone through the related courses during their undergraduate degree.

A principal aim of the course is to provide a comprehensive introduction to multiple generalizations of Riemann's zeta function and to analytic, algebraic, arithmetic, combinatorial methods used in their study.

Good sources for knowledge on principal parts in the course are the two textbooks [**1**] and [**2**]. These notes are provided to make the material more accessible to the students.

Bibliography

[1] Jianqiang Zhao: *Multiple zeta functions, multiple polylogarithms and their special values*, Series on Number Theory and its Applications **12** (World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2016).

[2] José Ignacio Burgos Gil, Javier Fresán, with contributions by Ulf Kühn: *Multiple zeta values: from numbers to motives*, Clay Mathematics Proceedings (in press).

[3] Michael Hoffman's site contains some basic information about the MZVs. Hoffman also has a comprehensive list of references on MZVs and related stuff.

CHAPTER 1

Riemann's zeta function and its multiple generalisation

1.1. Riemann's zeta function

The *Riemann zeta function* is traditionally defined in the region $\operatorname{Re} s > 1$ by the convergent series

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

this makes it an analytic function in the domain. The function is very special in number theory, because its analytic properties are ultimately linked to ones of the prime numbers; this can be seen through Euler's representation of $\zeta(s)$ as an infinite product over primes:

$$\zeta(s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1}.$$

The first thing one learns in studying the distribution of prime numbers is that $\zeta(s)$ can be analytically continued to a larger domain, and in this story Riemann's zeta function is always accompanied by Euler's gamma function $\Gamma(z)$ defined through the product expansion

$$\frac{1}{\Gamma(z)} = z e^{\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right) e^{-z/k} \tag{1.2}$$

for its reciprocal. Here

$$\begin{aligned} \gamma &= \lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \log n\right) \\ &= 0.57721566490153286060651209008240243104215933593992\dots \end{aligned}$$

is the *Euler* (or *Euler–Mascheroni*) *constant.* A theorem of Weierstrass guarantees that $1/\Gamma(z)$ is an entire function with zeros at $z = 0, -1, -2, \dots$, and many properties of the gamma function, like the difference equation

$$\Gamma(z+1) = z\Gamma(z), \tag{1.3}$$

the reflection formula

$$\Gamma(z)\Gamma(1-z) = \frac{1}{z} \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2}\right)^{-1} = \frac{\pi}{\sin \pi z} \tag{1.4}$$

and multiplication formula

$$\Gamma(z)\Gamma\Big(z+\frac{1}{n}\Big)\Gamma\Big(z+\frac{2}{n}\Big)\cdots\Gamma\Big(z+\frac{n-1}{n}\Big)=(2\pi)^{(n-1)/2}n^{-nz+1/2}\Gamma(nz), \quad (1.5)$$

follow straight from the defining product.

EXERCISE 1.1. Prove equations (1.3)–(1.5).

We also take for granted from a complex analysis course the evaluation

$$\int_0^\infty e^{-t}t^{z-1}\mathrm{d}t=\Gamma(z) \quad (1.6)$$

of the Eulerian integral (of the second kind) in the domain $\operatorname{Re} z>0$.

PROPOSITION 1.1. *The logarithmic derivative $\psi(z)=\Gamma'(z)/\Gamma(z)$ of the gamma function serves a generating function for the values of Riemann's zeta function at positive integers. More specifically,*

$$\psi(1-z)=-\gamma-\sum_{m=1}^\infty\zeta(m+1)z^m \quad \text{for } |z|<1.$$

PROOF. It follows from the logarithmic differentiation of (1.2) that

$$-\psi(z)=\frac{1}{z}+\gamma+\sum_{k=1}^\infty\Big(-\frac{1}{k}+\frac{1}{k(1+z/k)}\Big)$$

for $z\ne 0,-1,-2,\dots$. Furthermore, from (1.3) we have $\psi(1+z)=1/z+\psi(z)$. Thus,

$$\begin{aligned}-\psi(1-z)=\frac{1}{z}-\psi(-z)&=\gamma+\sum_{k=1}^\infty\frac{1}{k}\Big(-1+\frac{1}{1-z/k}\Big)\\&=\gamma+\sum_{k=1}^\infty\frac{1}{k}\sum_{m=1}^\infty\Big(\frac{z}{k}\Big)^m=\gamma+\sum_{m=1}^\infty z^m\sum_{k=1}^\infty\frac{1}{k^{m+1}},\end{aligned}$$

with all the internal series converging in the disk $|z|<1$. ☺

EXERCISE 1.2. In this exercise we compute the Eulerian integral of the first kind

$$\mathrm{B}(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,\mathrm{d}x,$$

where $\operatorname{Re}\alpha>0$ and $\operatorname{Re}\beta>0$.

(a) Verify the following properties:

$$\mathrm{B}(\alpha,\beta)=\mathrm{B}(\beta,\alpha); \qquad \mathrm{B}(\alpha,\beta+1)=\frac{\beta}{\alpha}\,\mathrm{B}(\alpha+1,\beta);$$

$$\mathrm{B}(\alpha,\beta)=\mathrm{B}(\alpha+1,\beta)+\mathrm{B}(\alpha,\beta+1); \qquad \mathrm{B}(\alpha,\beta+1)=\frac{\beta}{\alpha+\beta}\,\mathrm{B}(\alpha,\beta).$$

(b) Show that

$$\Gamma(\alpha)\Gamma(\beta) = 4 \lim_{R\to\infty} \iint_{[0,R]^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 4 \lim_{R\to\infty} \iint_{S_R} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

where $f(x,y) = e^{-(x^2+y^2)} x^{2\alpha-1} y^{2\beta-1}$ and S_R is the circular sector $x^2+y^2 \le R$, $x \ge 0$, $y \ge 0$.

(c) Pass to the polar coordinates $x = r\cos\theta$, $y = r\sin\theta$ in the integral

$$\iint_{S_R} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

and use part (b) to conclude that

$$\mathrm{B}(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

HINT. (b) Write

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1}\,\mathrm{d}t = 2\int_0^\infty e^{-x^2} x^{2\alpha-1}\,\mathrm{d}x = 2\lim_{R\to\infty}\int_0^R e^{-x^2} x^{2\alpha-1}\,\mathrm{d}x$$

and, similarly, for $\Gamma(\beta)$; then show that

$$\left| \iint_{[0,R]^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y - \iint_{S_R} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \right| \to 0 \quad \text{as } R\to\infty.$$

☺

EXERCISE 1.3. (a) Show the integral expansion

$$\psi(z) = -\gamma + \int_0^1 \frac{1-t^{z-1}}{1-t}\,\mathrm{d}t$$

in the half-plane $\operatorname{Re} z > 0$.

(b) Prove that, for $n = 1, 2, 3, \dots$,

$$\psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}.$$

1.2. Hurwitz's zeta function

In order to analyse the properties of Riemann's zeta function we turn our attention to its slightly more general version

$$\zeta(s,a) = \sum_{n=0}^\infty \frac{1}{(a+n)^s} \tag{1.7}$$

known as *Hurwitz's zeta function.* In this expression we treat a as a real constant from the interval $0 < a \le 1$ (though one can allow a to vary over the real line, and even over the complex plane); again, the series in (1.7) defines an analytic function of s in the region $\operatorname{Re} s > 1$. Observe that $\zeta(s,1) = \zeta(s)$.

Proposition 1.2. *For* $\operatorname{Re} s > 1$,

$$\zeta(s,a) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x.$$

Proof. We start with the following consequence of (1.6):

$$(a+n)^{-s}\Gamma(s) = \int_0^\infty x^{s-1}e^{-(n+a)x}\,\mathrm{d}x.$$

Taking $\delta > 0$, we have in the domain $\sigma = \operatorname{Re} s \ge 1+\delta$,

$$\begin{aligned}
\Gamma(s)\zeta(s,a) &= \lim_{N\to\infty} \sum_{n=0}^{N} \int_0^\infty x^{s-1}e^{-(n+a)x}\,\mathrm{d}x \\
&= \lim_{N\to\infty} \left(\int_0^\infty \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x - \int_0^\infty \frac{x^{s-1}e^{-(N+1+a)x}}{1-e^{-x}}\,\mathrm{d}x \right) \\
&= \lim_{N\to\infty} \left(\int_0^\infty \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x - \int_0^\infty \frac{x^{s-1}e^{-(N+a)x}}{e^{x}-1}\,\mathrm{d}x \right).
\end{aligned}$$

Since $e^x \ge 1+x$ for $x \ge 0$, the absolute value of the second integral is estimated from above by the quantity

$$\int_0^\infty x^{\sigma-2}e^{-(N+a)x}\,\mathrm{d}x = (a+N)^{1-\sigma}\Gamma(\sigma-1),$$

which clearly tends to 0 as $N \to \infty$ in view of $\sigma - 1 \ge \delta > 0$. This gives the desired formula for $\operatorname{Re} s \ge 1+\delta$, hence for $\operatorname{Re} s > 1$. ☺

For real $\rho > 0$ (possibly, $\rho = \infty$), introduce a (Hankel-type) contour $D = D(\rho)$, which starts at $z = \rho$, passes once around the origin into the positive direction (without crossing the half-line $z \ge 0$) and ends up at $z = \rho$. Our principal interest is in the integral

$$\int_{D(\infty)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z = \lim_{\rho\to\infty} \int_{D(\rho)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z$$

for a fixed s from the half-plane $\sigma = \operatorname{Re} s \ge 1+\delta$. To avoid the unwanted poles of the integrand, we further assume that the contours $D(\rho)$ do not contain the points $\pm 2\pi i n$ for $n = 1, 2, \dots$. We specify the branch of $(-z)^{s-1} = e^{(s-1)\log(-z)}$ by choosing the $\log(-z)$ to be real for negative z; then $-\pi \le \arg(-z) \le \pi$ on the contours — this makes the integrand a single-valued function on $D(\rho)$. Of course, the integrand is not analytic inside $D(\rho)$ but we can still deform it within $\mathbb{C} \setminus [0,\infty)$ to the contour going along the upper bank of the cut $[0,\infty)$ from ρ to $\varepsilon > 0$, then making a circle of radius ε around the origin and finally returning from ε to ρ along the lower bank of the cut. At the beginning we have $\arg(-z) = -\pi$, so that $(-z)^{s-1} = e^{-\pi i(s-1)}z^{s-1}$, and at the end we get $\arg(-z) = \pi$, hence $(-z)^{s-1} = e^{\pi i(s-1)}z^{s-1}$. We set $-z = \varepsilon e^{i\theta}$ on the circle.

Therefore,

$$\int_{D(\rho)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z$$

$$= e^{-\pi i(s-1)}\int_{\rho}^{\varepsilon} \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x + i\int_{-\pi}^{\pi} \frac{(\varepsilon e^{i\theta})^s e^{a\varepsilon(\cos\theta+i\sin\theta)}}{1-e^{\varepsilon(\cos\theta+i\sin\theta)}}\,\mathrm{d}\theta$$

$$+ e^{\pi i(s-1)}\int_{\varepsilon}^{\rho} \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x$$

$$= -2i\sin\pi s\int_{\varepsilon}^{\rho} \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x + i\varepsilon^{s-1}\int_{-\pi}^{\pi} \frac{\varepsilon e^{is\theta+a\varepsilon(\cos\theta+i\sin\theta)}}{1-e^{\varepsilon(\cos\theta+i\sin\theta)}}\,\mathrm{d}\theta$$

for $0<\varepsilon\le\rho$. As $\varepsilon\to 0$ we have $\varepsilon^{s-1}\to 0$ and

$$\int_{-\pi}^{\pi} \frac{\varepsilon e^{is\theta+a\varepsilon(\cos\theta+i\sin\theta)}}{1-e^{\varepsilon(\cos\theta+i\sin\theta)}}\,\mathrm{d}\theta \to \int_{-\pi}^{\pi} \frac{e^{is\theta}}{\cos\theta+i\sin\theta}\,\mathrm{d}\theta = \int_{-\pi}^{\pi} e^{i(s-1)\theta}\,\mathrm{d}\theta,$$

since the integrand uniformly converges to its limit. We conclude that

$$\int_{D(\rho)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z = -2i\sin\pi s\int_{0}^{\rho} \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x$$

implying

$$\int_{D(\infty)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z = -2i\sin\pi s\int_{0}^{\infty} \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x$$

$$= -2i\sin\pi s\,\Gamma(s)\zeta(s,a) = -2\pi i\,\frac{\zeta(s,a)}{\Gamma(1-s)}$$

on the basis of Proposition 1.2 and reflection formula (1.4). This brings us to the following result.

PROPOSITION 1.3. *For* $\operatorname{Re} s>1$,

$$\zeta(s,a) = -\frac{\Gamma(1-s)}{2\pi i}\int_{D(\infty)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z. \tag{1.8}$$

The resulting integral is a single-valued analytic function of s for *all* $s\in\mathbb{C}$. Therefore, the only potential singularities of $\zeta(s,a)$ originate from the singularities of $\Gamma(1-s)$, which are the points $s=1,2,\dots$, since the integral provide the analytic continuation of $\zeta(s,a)$ to the entire complex plane with the exception of these points. At the same time, we already now the analyticity of $\zeta(s,a)$ in the domain $\operatorname{Re} s>1$ from its defining series expansion (1.7). This leads us to the following.

COROLLARY 1.4. *The function* $\zeta(s,a)$ *is analytic in* $\mathbb{C}$ *besides* $s=1$, *where it has a simple pole with residue* 1.

When $a=1$, this implies the analytic properties of $\zeta(s)$.

PROOF. By the argument above, the point $s = 1$ is the only candidate for a singular point. Taking $s = 1$ in the integral (without the gamma prefactor) we get the expression

$$\frac{1}{2\pi i}\int_{D(\infty)} \frac{e^{-az}}{1-e^{-z}}\,\mathrm{d}z$$

which is equal to the residue of the integrand at $z = 0$: this is clearly equal to 1. Combined with (1.8) this implies

$$\lim_{s\to 1}\frac{\zeta(s,a)}{\Gamma(1-s)} = -1.$$

It remains to recall that $\Gamma(1-s)$ has a simple pole at $s = 1$ with residue -1. ☺

EXERCISE 1.4. Show for $\operatorname{Re} s > 0$,

$$(1-2^{1-s})\zeta(s) = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n^s} = \frac{1}{\Gamma(s)}\int_0^{\infty}\frac{x^{s-1}}{e^x+1}\,\mathrm{d}x.$$

EXERCISE 1.5. Show for $\operatorname{Re} s > 1$,

$$(2^s-1)\zeta(s) = \zeta\Big(s,\frac{1}{2}\Big) = \frac{2^s}{\Gamma(s)}\int_0^{\infty}\frac{x^{s-1}e^x}{e^{2x}-1}\,\mathrm{d}x.$$

EXERCISE 1.6. Show for all $s \neq 1$,

$$\zeta(s) = -\frac{2^{1-s}\Gamma(1-s)}{2\pi i\,(2^{1-s}-1)}\int_{D(\infty)}\frac{(-z)^{s-1}}{e^z+1}\,\mathrm{d}z,$$

where the contour $D(\infty)$ does not contain inside the points $\pm\pi i, \pm 3\pi i, \pm 5\pi i, \dots$.

PROPOSITION 1.5 (Hurwitz). *For $0 < a \le 1$ and $\sigma = \operatorname{Re} s < 0$,*

$$\zeta(s,a) = \frac{2\Gamma(1-s)}{(2\pi)^{1-s}}\left(\sin\frac{\pi s}{2}\sum_{n=1}^{\infty}\frac{\cos 2\pi a n}{n^{1-s}} + \cos\frac{\pi s}{2}\sum_{n=1}^{\infty}\frac{\sin 2\pi a n}{n^{1-s}}\right). \tag{1.9}$$

PROOF. Consider the integral

$$-\frac{1}{2\pi i}\int_{C_N}\frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z,$$

where N is an *odd* positive integer, the contour C_N is the circle centered at the origin of radius $N\pi$ going counter-clockwise from $N\pi$ to $N\pi$. We assume that $\arg(-z) = 0$ at $z = -N\pi$.

In the domain bounded by the contours C_N and $D(N\pi)$, the function $(-z)^{s-1}e^{-az}/(1-e^{-z})$ is analytic and single-valued, except for the poles at $\pm 2\pi i, \pm 4\pi i, \dots, \pm(N-1)\pi i$. Therefore,

$$\frac{1}{2\pi i}\int_{C_N}\frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z - \frac{1}{2\pi i}\int_{D(N\pi)}\frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z = \sum_{n=1}^{(N-1)/2}(R_n^+ + R_n^-),$$

where R_n^+ and R_n^- are the residues of the integrand at $2n\pi i$ and $-2n\pi i$, respectively. When $-z = 2n\pi e^{-\pi i/2}$, the residue is equal to $(2n\pi)^{s-1}e^{-\pi i(s-1)/2}e^{-2an\pi i}$, so that

$$R_n^+ + R_n^- = 2\,(2n\pi)^{s-1}\sin\biggl(\frac{\pi s}{2} + 2\pi an\biggr) \quad \text{for } n = 1, 2, \dots, \frac{N-1}{2}.$$

We obtain

$$-\frac{1}{2\pi i}\int_{D(N\pi)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z = \frac{2\sin\frac{\pi s}{2}}{(2\pi)^{1-s}}\sum_{n=1}^{(N-1)/2}\frac{\cos 2\pi an}{n^{1-s}} + \frac{2\cos\frac{\pi s}{2}}{(2\pi)^{1-s}}\sum_{n=1}^{(N-1)/2}\frac{\sin 2\pi an}{n^{1-s}} - \frac{1}{2\pi i}\int_{C_N}\frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z.$$

Furthermore, for $0 < a \le 1$ we can find an absolute bound $|e^{-az}/(1-e^{-z})| < M$ for $z \in C_N$, independent of N. This means that, for $\sigma = \operatorname{Re} s < 0$,

$$\left|\frac{1}{2\pi i}\int_{C_N}\frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z\right| < \frac{M}{2\pi}\int_{-\pi}^{\pi}|(N\pi)^s e^{is\theta}|\,\mathrm{d}\theta < M(N\pi)^\sigma e^{\pi|s|} \to 0 \quad \text{as } N \to \infty.$$

Thus, letting $N \to \infty$ in the above equality we arrive at the desired formula (1.9). Note the (absolute) convergence of the both series when $\operatorname{Re} s < 0$. ☺

THEOREM 1.6 (Riemann). *The following functional equation is valid for Riemann's zeta function:*

$$2^{1-s}\Gamma(s)\zeta(s)\cos\frac{\pi s}{2} = \pi^s\zeta(1-s). \tag{1.10}$$

PROOF. Take $a = 1$ in equation (1.9) and apply the reflection formula (1.4) of the gamma function. This proves (1.10) in the domain $\operatorname{Re} s < 0$. Since the both sides are analytic in the larger domain $\mathbb{C}\setminus\{0,1\}$ (besides the simple poles at $s = 0, 1$), the result remains valid there by the theory of analytic continuation. ☺

EXERCISE 1.7. Show the function $\Gamma(s/2)\pi^{-s/2}\zeta(s)$ does not change under the involution $s \leftrightarrow 1-s$.

It follows from (1.10) that $\zeta(s)$ has zeros at negative even integers; these are called trivial zeros. In his famous 1859 memoir, Riemann suggested that all other (non-trivial) zeros lie on the critical line $\operatorname{Re} s = 1/2$, which represents the symmetry of the functional equation.

1.3. Zeta values

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

$$\begin{aligned}\pi &= 4\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\\ &= 3.14159265358979323846264338327950288419716939937510\dots.\end{aligned}$$

Although we do not follow Euler's original method, the derivation is worth reproducing.

For $a \in \mathbb{R}$, the *Bernoulli polynomials* $B_s(a) \in \mathbb{Q}[a]$, where $s = 0, 1, 2, \dots$, are defined by the generating function

$$\frac{ze^{az}}{e^z - 1} = \sum_{s=0}^{\infty} B_s(a)\,\frac{z^s}{s!}\,, \tag{1.11}$$

while the *Bernoulli numbers* $B_s \in \mathbb{Q}$, where $s = 0, 1, 2, \dots$, are simply given by $B_s = B_s(0)$. The latter means that the generating function of the Bernoulli numbers is

$$\frac{z}{e^z - 1} = \sum_{s=0}^{\infty} B_s\,\frac{z^s}{s!}\,.$$

For example, $B_0 = 1$, $B_1 = -1/2$. The polynomials and numbers satisfy numerous identities, with several dedicated books devoted to them. As an example, we have the formulas $B_s'(a) = sB_{s-1}(a)$ and

$$\sum_{k=M}^{N-1} k^{s-1} = \frac{B_s(N) - B_s(M)}{s}$$

for $s = 1, 2, \dots$, and also the following ones.

Exercise 1.8. (a) Show that

$$B_s(a) = \sum_{k=0}^{s}\binom{s}{k} B_k a^{s-k} \quad \text{for } s = 0, 1, 2, \dots.$$

(b) Verify that $B_s = 0$ for odd $s \ge 3$.
(c) Verify that $B_s(1) = B_s = B_s(0)$ for even $s \ge 0$.

Lemma 1.7. *For $0 < a \le 1$ and $s = -m$ a negative integer,*

$$\zeta(-m, a) = -\frac{B_{m+1}(a)}{m+1}\,.$$

Proof. Recall the integral

$$\frac{1}{2\pi i}\int_{D(\infty)} \frac{(-z)^{s-1}e^{-az}}{1 - e^{-z}}\,\mathrm{d}z = -\frac{\zeta(s,a)}{\Gamma(1-s)}$$

from Proposition 1.3. If s is a negative integer, $s = -m$, the expression

$$\frac{(-z)^{s-1}e^{-az}}{1 - e^{-z}}$$

is a single-valued function of z, which is analytic in $|z| < 2\pi$, $z \neq 0$. By Cauchy's integral theorem, the integral over $D(\infty)$ is equal to the residue of the integrand at $z = 0$, that is, to the coefficient of $z^{-s} = z^m$ in

$$\frac{(-1)^{s-1}e^{-az}}{1-e^{-z}} = \frac{(-1)^{s-1}}{z}\,\frac{(-z)\,e^{-az}}{e^{-z}-1} = \frac{(-1)^{m-1}}{z}\sum_{k=0}^{\infty}(-1)^k B_k(a)\frac{z^k}{k!}\,.$$

It follows that

$$-\frac{\zeta(-m,a)}{m!} = -\frac{\zeta(s,a)}{\Gamma(1-s)}\bigg|_{s=-m} = \frac{B_{m+1}(a)}{(m+1)!}\,,$$

which implies the result. ☺

When $a = 1$, we get the following consequence for Riemann's zeta function (using also Exercise 1.8).

Corollary 1.8. *For $k = 1, 2, \dots,$ we have $\zeta(-2k) = 0$ and $\zeta(1-2k) = B_{2k}/(2k)$.*

Exercise 1.9. Show that $\zeta(0,a) = \frac{1}{2} - a$ and $\zeta(0) = -\frac{1}{2}$.

Corollary 1.9. *For $k = 1, 2, \dots,$ we have*

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

Proof. This follows from Corollary 1.8 and the functional equation (1.10) for $s = 2k$. ☺

In particular,

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

and so on.

Corollary 1.9 gives us the expression for 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 proved the irrationality of the number $\zeta(3)$ and there are more recent but partial linear independence results of Rivoal and this lecturer. 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 a complete answer to 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 1.10. *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 decades 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 multiple (l-tuple) zeta function

$$\zeta(\boldsymbol{s}) = \zeta(s_1, s_2, \dots, s_l) = \sum_{n_1 > n_2 > \dots > n_l \ge 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}; \tag{1.12}$$

the corresponding multi-index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$ will be further regarded as *admissible*. The quantities (1.12) are called the *multiple zeta values* (and abbreviated MZVs), or the *multiple harmonic series*, or the *Euler sums*. The sums (1.12) for $l = 2$ were first investigated by Euler, who obtained a family of identities connecting double and ordinary zeta values (which we discuss later). In particular, Euler proved the identity

$$\zeta(2, 1) = \zeta(3), \tag{1.13}$$

which was several times rediscovered after.

Exercise 1.10. Find your own (elementary) proof of (1.13).

1.4. Analytic continuation of MZF

In this part, we discuss analytic properties of the *multiple zeta function* (MZV)

$$\zeta(\boldsymbol{s}) = \sum_{n_1 > n_2 > \dots > n_l \ge 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}} \tag{1.14}$$

as a function of complex variables $s_1, \dots, s_l$; the notation $\sigma_1, \dots, \sigma_l$ will be used for the real parts of $s_1, \dots, s_l$.

Exercise 1.11. Show that the multiple series in (1.14) converges *absolutely* in the domain

$$\sigma_1 + \dots + \sigma_j = \operatorname{Re}(s_1 + \dots + s_j) > j \quad \text{for every } j = 1, \dots, l.$$

Conclude from this that the MZV is analytic in each of its variables in the domain $\sigma_1 + \cdots + \sigma_j > j$, where $j = 1, \dots, l$.

HINT. Use mathematical induction on l and estimates

$$\sum_{n>M} \frac{1}{n^\sigma} \le \frac{1}{(\sigma - 1)M^{\sigma-1}},$$

where $M \ge 1$ is integral and $\sigma > 1$ is real, coming from the integral test (when the partial sums of a series are compared to Riemann sums). ☺

LEMMA 1.11. *For $0 < a \le 1$ and an integer $m \ge 2$,*

$$\sum_{n\in\mathbb{Z}}{}' \frac{e^{2\pi ina}}{(2\pi in)^m} = -\frac{B_m(a)}{m!},$$

where the dash in summation corresponds to omitting the (problematic) index $n = 0$.

PROOF. Comparing Hurwitz's equation (1.9),

$$\frac{\zeta(s,a)}{\Gamma(1-s)} = \frac{2}{(2\pi)^{1-s}} \sum_{n=1}^{\infty} \frac{\sin(\pi s/2 + 2\pi an)}{n^{1-s}}$$

for $s = -m + 1$, with the result of Lemma 1.7,

$$-\frac{B_m(a)}{m!} = \frac{\zeta(s,a)}{\Gamma(1-s)}\bigg|_{s=-m+1},$$

we find

$$-\frac{B_m(a)}{m!} = 2\sum_{n=1}^{\infty} \frac{\sin(-\pi(m-1)/2 + 2\pi an)}{(2\pi n)^m},$$

which is exactly

$$\begin{aligned}(-1)^k \sum_{n=1}^{\infty} \frac{2\sin 2\pi an}{(2\pi n)^{2k+1}} &= \sum_{n=1}^{\infty} \frac{e^{2\pi ina} - e^{-2\pi ina}}{(2\pi in)^{2k+1}} \\ &= \sum_{n=1}^{\infty} \frac{e^{2\pi ina}}{(2\pi in)^{2k+1}} + \sum_{n=1}^{\infty} \frac{e^{-2\pi ina}}{(-2\pi in)^{2k+1}}\end{aligned}$$

or

$$\begin{aligned}(-1)^k \sum_{n=1}^{\infty} \frac{2\cos 2\pi an}{(2\pi n)^{2k}} &= \sum_{n=1}^{\infty} \frac{e^{2\pi ina} + e^{-2\pi ina}}{(2\pi in)^{2k}} \\ &= \sum_{n=1}^{\infty} \frac{e^{2\pi ina}}{(2\pi in)^{2k}} + \sum_{n=1}^{\infty} \frac{e^{-2\pi ina}}{(-2\pi in)^{2k}}\end{aligned}$$

depending on whether $m = 2k + 1$ is odd or $m = 2k$ is even. ☺

Lemma 1.12. *For $0 < a \le 1$ and any integer $m \ge 2$,*

$$|B_m(a)| < \frac{4m!}{(2\pi)^m}.$$

Proof. It follows from Lemma 1.12 that

$$|B_m(a)| \le m! \sum_{n\in\mathbb{Z}}{}' \frac{1}{(2\pi n)^m} = \frac{2m!\,\zeta(m)}{(2\pi)^m}.$$

It remains to apply the trivial estimate $\zeta(m) \le \zeta(2) = \pi^2/6 < 2$. ☺

For the statement and application of the following classical result, it will be convenient to introduce the *periodic* Bernoulli polynomials given by $\widetilde{B}_m(a) = B_m(\{a\})$, where $\{\cdot\}$ denotes the fractional part of a real number. By Lemma 1.12 (and Exercise 1.8) we get the estimate

$$|\widetilde{B}_m(a)| < \frac{4m!}{(2\pi)^m} \quad \text{for } m = 2, 3, \dots, \tag{1.15}$$

now valid for *all* real a.

Exercise 1.12. Verify the validity of (1.15) for $m = 0, 1$.

We will also implement the (standard) notation

$$(s)_m = \frac{\Gamma(s+m)}{\Gamma(s)} = \begin{cases} s(s+1)\cdots(s+m-1) & \text{if } m = 1, 2, \dots, \\ 1 & \text{if } m = 0, \end{cases}$$

for the Pochhammer symbol, though it makes sense for *any* (not necessarily integer or nonnegative) m. For example, $(s)_{-1} = \Gamma(s-1)/\Gamma(s) = 1/(s-1)$.

Proposition 1.13 (Euler–Maclaurin summation). *Let $f(x)$ be a (complex-valued) C^∞ function on the real interval $[1, \infty)$. Then for any positive integers N and m, m even,*

$$\sum_{n=1}^{N} f(n) = \int_1^N f(x)\,\mathrm{d}x + \frac{1}{2}\big(f(1) + f(N)\big) + \sum_{k=2}^{m} \frac{B_k}{k!}\big(f^{(k-1)}(N) - f^{(k-1)}(1)\big) - \frac{1}{m!}\int_1^N \widetilde{B}_m(x) f^{(m)}(x)\,\mathrm{d}x.$$

Notice that the sum over k in the formula only involves k even, because $B_k = 0$ for odd $k \ge 2$.

Lemma 1.14. *Given $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$, for integers $M \ge 1$ and $m \ge 2$, m even, we have*

$$\sum_{n>M} \frac{1}{n^s} = \sum_{k=0}^{m} \frac{B_k}{k!} \frac{(s)_{k-1}}{M^{s+k-1}} - \frac{(s)_m}{m!} \int_M^\infty \frac{\widetilde{B}_m(x)}{x^{s+m}}\,\mathrm{d}x.$$

PROOF. Apply Proposition 1.13 with $f(x) = 1/x^s$ twice: when $N \to \infty$ and when $N = M$. Taking the difference of the results we arrive at

$$\begin{aligned}\sum_{n>M} \frac{1}{n^s} &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{M} \frac{1}{n^s} \\ &= \int_M^\infty f(x)\,\mathrm{d}x - \frac{1}{2} f(M) - \sum_{k=2}^{m} \frac{B_k}{k!} f^{(k-1)}(M) \\ &\qquad - \frac{1}{m!} \int_M^\infty \widetilde{B}_m(x) f^{(m)}(x)\,\mathrm{d}x \\ &= \frac{1}{(s-1)M^{s-1}} - \frac{1}{2M^s} - \sum_{k=2}^{m} \frac{B_k}{k!} \frac{(s)_{k-1}}{M^{s+k-1}} - \frac{(s)_m}{m!} \int_M^\infty \frac{\widetilde{B}_m(x)}{x^{s+m}}\,\mathrm{d}x,\end{aligned}$$

which can be written in the desired form because $B_0 = 1$ and $B_1 = -1/2$. ☺

EXERCISE 1.13. Use Lemma 1.14 (with $M = 1$, say) and the estimates of Lemma 1.12 to show that Riemann's zeta function can be analytically continued to the half-plane $\operatorname{Re} s > -L$ for any real $L > 0$.

Introduce the following discrete subset of $\mathbb{C}^l$:

$$\begin{aligned}\Sigma_l = \big\{ \boldsymbol{s} \in \mathbb{C}^l : s_1 \in \{1\},\ s_1 + s_2 \in \{1, 2\} \cup 2\mathbb{Z}_{\le 0}, \\ s_1 + \dots + s_j \in \mathbb{Z}_{\le j} \text{ for } j = 3, \dots, l \big\}.\end{aligned}$$

The following general result provides the analytic continuation of the MZV $\zeta(\boldsymbol{s})$ to a meromorphic function on $\mathbb{C}^l$ with (at most) simple poles given by Σ_l.

THEOREM 1.15. *Assume $l \ge 2$. Then for any $\boldsymbol{s} = (s_1, \dots, s_l) \in \mathbb{C}^l \setminus \Sigma_l$ and an even $m > l + |\sigma_1| + \dots + |\sigma_l|$, we have*

$$\begin{aligned}\zeta(\boldsymbol{s}) = \sum_{k=0}^{m} \frac{B_k}{k!} (s_1)_{k-1} \cdot \zeta(s_1 + s_2 + k - 1, s_3, \dots, s_l) \\ - \frac{(s_1)_m}{m!} \sum_{n_2 > \dots > n_l \ge 1} \frac{1}{n_2^{s_2} \cdots n_l^{s_l}} \int_{n_2}^\infty \frac{\widetilde{B}_m(x)}{x^{s_1+m}}\,\mathrm{d}x. \qquad (1.16)\end{aligned}$$

PROOF. The absolute convergence of the second series in the formula (1.16) follows from the estimate

$$\left| \int_M^\infty \frac{\widetilde{B}_m(x)}{x^{s+m}}\,\mathrm{d}x \right| \le \frac{4m!}{(2\pi)^{2m}} \int_M^\infty \frac{\mathrm{d}x}{x^{\sigma+m}} = \frac{4m!}{(2\pi)^{2m}(m-1+\sigma)M^{m-1+\sigma}},$$

where $\sigma = \operatorname{Re} s$, implying

$$\begin{aligned}\sum_{n_2 > \dots > n_l \ge 1} \left| \frac{1}{n_2^{s_2} n_3^{s_3} \cdots n_l^{s_l}} \int_{n_2}^\infty \frac{\widetilde{B}_m(x)}{x^{s_1+m}}\,\mathrm{d}x \right| \\ \le \frac{4m!}{(2\pi)^{2m}(m-1+\sigma_1)} \sum_{n_2 > \dots > n_l \ge 1} \frac{1}{n_2^{m-1+\sigma_1+\sigma_2} n_3^{\sigma_3} \cdots n_l^{\sigma_l}}.\end{aligned}$$

For the latter sum we use

$$\frac{1}{n_2^{\sigma_1+\sigma_2} n_3^{\sigma_3} \cdots n_l^{\sigma_l}} \le n_2^{|\sigma_1|+|\sigma_2|} n_3^{|\sigma_3|} \cdots n_l^{|\sigma_l|} \le n_2^{|\sigma_1|+|\sigma_2|+|\sigma_3|+\cdots+|\sigma_l|}$$

and the fact that the number of integers $n_3, \ldots, n_l$ satisfying $n_2 > n_3 > \cdots > n_l \ge 1$ is bounded above by n_2^{l-2} (because each n_j satisfies $1 \le n_j < n_2$), so that

$$\sum_{n_2 > \cdots > n_l \ge 1} \frac{1}{n_2^{m-1+\sigma_1+\sigma_2} n_3^{\sigma_3} \cdots n_l^{\sigma_l}} \le \sum_{n_2 \ge 1} \frac{n_2^{|\sigma_1|+|\sigma_2|+|\sigma_3|+\cdots+|\sigma_l|} n_2^{l-2}}{n_2^{m-1}}$$

converges when $m > l + |\sigma_1| + \cdots + |\sigma_l|$.

Now, to get the formula (1.16) we apply Lemma 1.14 with $s = s_1$, $n = n_1$ and $M = n_2$, and then perform the summation over $n_2 > n_3 > \cdots > n_l \ge 1$.

It remains to carefully control the (potential) poles by induction on l. ☺

Exercise 1.14. Show that the potential poles of $\zeta(\boldsymbol{s})$ at $\boldsymbol{s} \in \Sigma_l$ are at most simple.

Hint. Notice that the second (multiple) sum in (1.16) is analytic, so that the only source for poles comes from

$$\sum_{k=0}^{m} \frac{B_k}{k!} (s_1)_{k-1} \cdot \zeta(s_1 + s_2 + k - 1, s_3, \ldots, s_l).$$

Use mathematical induction on l and the fact that $\zeta(s)$ (when $l = 1$) has one simple pole at $s = 1$. ☺

CHAPTER 2

Multiple zeta values

2.1. First steps

The quantities

$$\zeta(\boldsymbol{s}) = \zeta(s_1, s_2, \dots, s_l) = \sum_{n_1 > n_2 > \dots > n_l \ge 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}; \tag{2.1}$$

for the admissible tuples of integers (all $s_1, \dots, s_l$ are positive and $s_1 > 1$) were introduced in the 1990s by Hoffman and, independently, by Zagier (with the opposite order of summation on the right-hand side of (2.1)). 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 since then) on the structure of algebraic relations for the family (2.1). Hoffman also introduced the alternative version

$$\zeta^\star(\boldsymbol{s}) = \zeta^\star(s_1, s_2, \dots, s_l) = \sum_{n_1 \ge n_2 \ge \dots \ge n_l \ge 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}} \tag{2.2}$$

of generalised Euler sums, with non-strict inequalities in summation; these are known by the name *multiple zeta star values.*

Exercise 2.1. For any admissible index $\boldsymbol{s} = (s_1, s_2, \dots, s_l)$, show the (dual) relations

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

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}.)

Hint. This is a purely combinatorial statement; use the inclusion-exclusion principle for part (b). ☺

Although all relations of series (2.2) may be translated, with the help of Exercise 2.1, into relations for series (2.1) and vice versa, several identities possess a more compact form by means of (2.2); for example,

$$\zeta^\star(\{2\}^k, 1) = \zeta^\star(\underbrace{2, \dots, 2}_{k \text{ times}}, 1) = 2\zeta(2k+1), \qquad k = 1, 2, \dots . \tag{2.3}$$

Observe that the particular instance $k = 1$ of (2.3) is equivalent to Euler's identity $\zeta(2, 1) = \zeta(3)$.

To each quantity (2.1) (or (2.2)), assign two characteristics: the *weight* (or *degree*) $|\boldsymbol{s}| = s_1 + s_2 + \cdots + s_l$ and the *length* (or *depth*) $\ell(\boldsymbol{s}) = l$. We shall witness in the course that all relations known so far for the MZVs (2.1) and (2.2) are weight-preserving.

2.2. 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*

$$\sum_{k=1}^{l} \zeta(s_1, \dots, s_{k-1}, s_k + 1, s_{k+1}, \dots, s_l) = \sum_{\substack{k=1 \\ s_k \ge 2}}^{l} \sum_{j=0}^{s_k - 2} \zeta(s_1, \dots, s_{k-1}, s_k - j, j + 1, s_{k+1}, \dots, s_l) \tag{2.4}$$

holds.

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

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

hence

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

and

$$\sum_{k=1}^{l}\big(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\big)$$
$$=\sum_{n_1>n_2>\dots>n_l\ge1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\sum_{m=1}^{n_1}\frac{1}{m}$$
$$=\sum_{n_1>n_2>\dots>n_l\ge1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\sum_{m=1}^{\infty}\biggl(\frac{1}{m}-\frac{1}{n_1+m}\biggr). \tag{2.5}$$

From now on, to each collection $n_1>n_2>\dots>n_l\ge1$ we will associate the set of parameters $m_1,m_2,\dots,m_l\ge1$ such that $n_k=m_k+\dots+m_l$ for $k=1,2,\dots,l$; alternatively, $m_k=n_k-n_{k+1}$ for $k=1,\dots,l-1$ and $m_l=n_l$.

Now notice the following partial-fraction decomposition (where both sides are viewed as functions of m for $n\ne0$ fixed):

$$\frac{1}{m(n+m)^s}=\frac{1}{n^sm}-\sum_{j=0}^{s-1}\frac{1}{n^{j+1}(n+m)^{s-j}}\,; \tag{2.6}$$

for the proof, it is sufficient to sum a geometric progression on the right-hand side. For $n=n_1$ and $s=s_1$ this implies

$$\frac{1}{n_1^{s_1}}\biggl(\frac{1}{m}-\frac{1}{n+m_1}\biggr)=\sum_{j=0}^{s_1-2}\frac{1}{n_1^{j+1}(n_1+m)^{s_1-j}}+\frac{1}{m(n_1+m)^{s_1}}\,.$$

Going on in equality (2.5) we find that

$$\sum_{k=1}^{l}\big(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\big)$$
$$=\sum_{j=0}^{s_1-2}\sum_{n_1>n_2>\dots>n_l\ge1}\sum_{m\ge1}\frac{1}{(n_1+m)^{s_1-j}n_1^{j+1}n_2^{s_2}\cdots n_l^{s_l}}$$
$$+\sum_{n_1>n_2>\dots>n_l\ge1}\sum_{m\ge1}\frac{1}{m(n_1+m)^{s_1}n_2^{s_2}\cdots n_l^{s_l}}$$
$$=\sum_{j=0}^{s_1-2}\zeta(s_1-j,j+1,s_2,\dots,s_l) \tag{2.7}$$
$$+\sum_{n_2>\dots>n_l\ge1}\sum_{m,m_1\ge1}\frac{1}{m(n_2+m+m_1)^{s_1}n_2^{s_2}\cdots n_l^{s_l}}$$
$$=\sum_{j=0}^{s_1-2}\zeta(s_1-j,j+1,s_2,\dots,s_l)+\sum_{n_0>n_1>n_2>\dots>n_l\ge1}\frac{1}{n_0^{s_1}m_1n_2^{s_2}\cdots n_l^{s_l}}, \tag{2.8}$$

where in the latter tuple sum we interchanged $m\leftrightarrow m_1$ and set $n_0=n_1+m$. Using now identity (2.6) with $m=m_{k-1}$, $n=n_k=n_{k-1}-m_{k-1}$ and $s=s_k$, we deduce that

$$\frac{1}{m_{k-1}n_k^{s_k}} = \frac{1}{m_{k-1}(n_k+m_{k-1})^{s_k}} + \sum_{j=0}^{s_k-1}\frac{1}{n_k^{j+1}(n_{k-1}+m_k)^{s_k-j}}$$
$$= \sum_{j=0}^{s_k-1}\frac{1}{n_{k-1}^{s_k-j}n_k^{j+1}} + \frac{1}{n_{k-1}^{s_k}m_{k-1}} \quad \text{for } k=2,\dots,l,$$

therefore

$$\sum_{n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}m_{k-1}n_k^{s_k}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}$$
$$= \sum_{j=0}^{s_k-1}\sum_{n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}n_{k-1}^{s_k-j}n_k^{j+1}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}$$
$$+ \sum_{n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}n_{k-1}^{s_k}m_{k-1}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}$$
$$= \sum_{j=0}^{s_k-1}\zeta(s_1,\dots,s_{k-1},s_k-j,j+1,s_{k+1},\dots,s_l)$$
$$+ \sum_{n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}n_{k-1}^{s_k}m_kn_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}, \tag{2.9}$$

where again we swap the role of m_{k-1} and m_k. Applying consequently identities (2.9) for $k=2,\dots,l$ for the second multiple sum on the right-hand side of equality (2.8), we obtain

$$\sum_{k=1}^{l}\big(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\big)$$
$$= \sum_{j=0}^{s_1-2}\zeta(s_1-j,j+1,s_2,\dots,s_l)$$
$$+ \sum_{k=2}^{l}\sum_{j=0}^{s_k-1}\zeta(s_1,\dots,s_{k-1},s_k-j,j+1,s_{k+1},\dots,s_l)$$
$$+ \sum_{n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{l-1}^{s_k}m_l}$$
$$= \sum_{k=1}^{l}\sum_{j=0}^{s_k-2}\zeta(s_1,\dots,s_{k-1},s_k-j,j+1,s_{k+1},\dots,s_l)$$
$$+ \sum_{k=2}^{l}\zeta(s_1,\dots,s_{k-1},1,s_k,s_{k+1},\dots,s_l) + \sum_{n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{l-1}^{s_k}n_l}$$

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

$$+ \sum_{k=1}^{l} \zeta(s_1, \dots, s_k, 1, s_{k+1}, \dots, s_l). \tag{2.10}$$

Reducing the both sides by the latter sum over k, we finally arrive at the desired identity (2.4). ☺

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

THEOREM 2.2 (Euler). *For any integer $s \geq 3$, the identity*

$$\sum_{j=2}^{s-1} \zeta(j, s-j) = \zeta(s) \tag{2.11}$$

takes place.

Note also that, in the case $s = 3$, identity (2.11) becomes nothing else but Euler's relation (1.13).

As a simple companion to Theorem 2.2 we have the following.

THEOREM 2.3 (Weighted analogue of Euler's theorem). *For any $s \geq 3$,*

$$\sum_{j=2}^{s-1} 2^j \zeta(j, s-j) = (s+1)\zeta(s). \tag{2.12}$$

PROOF (J. Wan). Write the left-hand side of (2.12) as

$$\sum_{j=2}^{s-1} 2^j \sum_{m,n=1}^{\infty} \frac{1}{(n+m)^j n^{s-j}} = \sum_{j=2}^{s-1} \sum_{\substack{m,n=1 \\ m\neq n}}^{\infty} \frac{2^j}{(n+m)^j n^{s-j}} + \sum_{j=2}^{s-1} \sum_{n=1}^{\infty} \frac{2^j}{(2n)^j n^{s-j}}$$

$$= \sum_{\substack{m,n=1 \\ m\neq n}}^{\infty} \frac{1}{n^s} \sum_{j=2}^{s-1} \frac{(2n)^j}{(n+m)^j} + (s-2)\zeta(s).$$

The geometric summation in j reduces the remaning sum to

$$\sum_{\substack{m,n=1 \\ m\neq n}}^{\infty} \left(\frac{2^s}{(n^2-m^2)(n+m)^{s-2}} - \frac{4}{(n^2-m^2)n^{s-2}} \right).$$

The first summand has antisymmetry in the variables m, n and hence vanishes when summed. For the second one we use the partial-fraction decomposition to obtain

$$\sum_{\substack{m=1 \\ m\neq n}}^{\infty} \frac{1}{m^2-n^2} = \frac{1}{2n} \sum_{\substack{m=1 \\ m\neq n}}^{\infty} \left(\frac{1}{m-n} - \frac{1}{m+n} \right)$$

$$
\begin{aligned}
&= \frac{1}{2n}\left(\sum_{m=1}^{n-1} + \sum_{m=n+1}^{\infty}\right)\left(\frac{1}{m-n} - \frac{1}{m+n}\right) \\
&= \frac{1}{2n}\left(-\sum_{k=1}^{n-1}\frac{1}{m} - \sum_{k=n+1}^{2n-1}\frac{1}{k} + \sum_{k=1}^{\infty}\left(\frac{1}{k} - \frac{1}{k+2n}\right)\right) \\
&= \frac{1}{2n}\left(-\sum_{k=1}^{n-1}\frac{1}{m} - \sum_{k=n+1}^{2n-1}\frac{1}{k} + \sum_{k=1}^{2n}\frac{1}{k}\right) \\
&= \frac{1}{2n}\cdot\left(\frac{1}{n} + \frac{1}{2n}\right) = \frac{3}{4n^2},
\end{aligned}
$$

and the result follows. ☺

Exercise 2.2. (a) Show that

$$\zeta(2)^2 = \left(\sum_{n=1}^{\infty}\frac{1}{n^2}\right)^2 = \zeta(4) + 2\zeta(2,2).$$

(b) Verify that $\zeta(2,2) = \frac{3}{4}\zeta(4)$.

(c) Conclude that $\zeta(2)^2 = \frac{5}{2}\zeta(4)$; in particular, the formula $\zeta(2) = \pi^2/6$ implies $\zeta(4) = \pi^4/90$.

Hint. (b) Use Theorems 2.2 and 2.3 for $s = 4$. ☺

Exercise 2.3. For $s \geq 4$ *even*, show that

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

Exercise 2.4 (Euler). For $s \geq 3$, show that

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

In other words, $\zeta(s-1,1)$ can be always expressed in terms of single zeta values.

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

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

holds.

Theorem 2.4 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, generalises Euler's theorem.

THEOREM 2.5 (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.

Theorems 2.1 and 2.5 are particular instances of Ohno's relations, which we discuss later in the course.

2.3. Calculation of MZVs

There are several ways of computing the MZVs efficiently based on their different representations. Notice that the original series (2.1) defining $\zeta(\boldsymbol{s})$ is somewhat inefficient, because the convergence is very slow (already for $l = 1$). The method we discuss below was designed by R. Crandall; later on in the course we shall witness a completely different strategy (see Proposition 3.8).

First observe the equality

$$\frac{1}{\Gamma(s)} \int_0^\infty x^{s-1} e^{-xn} \, \mathrm{d}x = \frac{1}{n^s \Gamma(s)} \int_0^\infty (nx)^{s-1} e^{-nx} \, \mathrm{d}(nx) = \frac{1}{n^s} \tag{2.13}$$

valid for $\operatorname{Re} s > 0$; this follows from (1.6).

LEMMA 2.6. *For admissible indices* $\boldsymbol{s} = (s_1, \dots, s_l) \in \mathbb{Z}_{>0}^l$,

$$\zeta(\boldsymbol{s}) = \sum_{m_1, \dots, m_l \ge 1} \int \cdots \int_{u_l > u_{l-1} > \dots > u_1 > u_0} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j - 1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)}, \tag{2.14}$$

where we set $u_0 = 0$.

Proof. It follows from (2.13) that

$$\zeta(\boldsymbol{s}) = \sum_{n_1>\dots>n_l\ge 1} \int\cdots\int_{x_1,\dots,x_l>0} \prod_{j=1}^{l} x_j^{s_j-1} e^{-x_j n_j} \frac{\mathrm{d}x_j}{\Gamma(s_j)}$$
$$= \sum_{m_1,\dots,m_l\ge 1} \int\cdots\int_{x_1,\dots,x_l>0} \prod_{j=1}^{l} x_j^{s_j-1} e^{-x_j (m_j+\dots+m_l)} \frac{\mathrm{d}x_j}{\Gamma(s_j)}.$$

Change the variables $u_k = \sum_{j=1}^{k} x_j$ and use the fact that the Jacobian of this transformation is the identity. ☺

For a fixed real parameter $u > 0$, consider the subdomains of the integration domain

$$T = \{(u_1,\dots,u_l) : u_l > u_{l-1} > \dots > u_1 > 0\} \in \mathbb{R}^l_{>0}$$

in (2.14) given as follows:

$$T_k = T_k(u) = \{(u_1,\dots,u_l) : u_l > \dots > u_{k+1} > u > u_k > \dots > u_0\} \in \mathbb{R}^l_{>0}$$

for $k = 0, 1, \dots, l-1$, and

$$T_l = T_l(u) = \{(u_1,\dots,u_l) : u > u_l > \dots > u_1 > u_0\} \in \mathbb{R}^l_{>0}.$$

Clearly, the domain T is the disjoint union of the $l+1$ subdomains $T_0, T_1, \dots, T_l$.

Denote

$$f(\boldsymbol{s}; u) = f_l(s_1,\dots,s_l; u)$$
$$= \sum_{m_1,\dots,m_l\ge 1} \int\cdots\int_{u_l>u_{l-1}>\dots>u_1>u} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)}$$

(it differs from the one in (2.14) by reducing the integration domain to $u_l > u_{l-1} > \dots > u_1 > u$, so that $f(\boldsymbol{s}; 0) = \zeta(\boldsymbol{s})$) and

$$g(\boldsymbol{s}; q; u) = g_l(s_1,\dots,s_l; q; u)$$
$$= \sum_{m_1,\dots,m_l\ge 1} \int\cdots\int_{u>u_l>u_{l-1}>\dots>u_1>u_0} u_l^q \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)}$$

(the integration domain is now bounded with an extra twist of the integrand by u_l^q implemented, for $q = 0, 1, 2, \dots$). These definitions immediately imply

$$\sum_{m_1,\dots,m_l\ge 1} \int\cdots\int_{T_0(u)} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)} = f_l(s_1,\dots,s_l; u), \tag{2.15}$$

$$\sum_{m_1,\dots,m_l\ge 1} \int\cdots\int_{T_l(u)} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)} = g_l(s_1,\dots,s_l; 0; u). \tag{2.16}$$

EXERCISE 2.5. It follows from the definition of $g(\boldsymbol{s};q;u)$ that

$$g(\boldsymbol{s};0;\infty)=\lim_{u\to\infty}g(\boldsymbol{s};0;u)=\zeta(\boldsymbol{s})$$

for all admissible multi-indices $\boldsymbol{s}$. Compute

$$g(\boldsymbol{s};q;\infty)=\lim_{u\to\infty}g(\boldsymbol{s};q;u)$$

for $q=1,2,\dots$.

HINT. Compute the q-th power of

$$u_l=(u_l-u_{l-1})+(u_{l-1}-u_{l-2})+\cdots+(u_2-u_1)+u_1$$

using the multinomial theorem (or apply repeatedly the binomial theorem). ☺

LEMMA 2.7. *For* $k=1,\dots,l-1$,

$$\sum_{m_1,\dots,m_l\ge1}\int\cdots\int_{T_k(u)}\prod_{j=1}^{l}(u_j-u_{j-1})^{s_j-1}e^{-u_jm_j}\frac{\mathrm{d}u_j}{\Gamma(s_j)}$$
$$=\sum_{q=0}^{s_{k+1}-1}\frac{(-1)^q}{q!}\,g_k(s_1,\dots,s_k;q;u)f_{l-k}(s_{k+1}-q,s_{k+2},\dots,s_l;u). \tag{2.17}$$

PROOF. We use the binomial expansion

$$\frac{1}{\Gamma(s_{k+1})}(u_{k+1}-u_k)^{s_{k+1}-1}=\frac{1}{\Gamma(s_{k+1})}\sum_{q=0}^{s_{k+1}-1}\frac{(s_{k+1}-1)!}{q!\,(s_{k+1}-q-1)!}\,u_{k+1}^{s_{k+1}-q-1}(-u_k)^q$$
$$=\sum_{q=0}^{s_{k+1}-1}\frac{(-1)^q}{q!}\,\frac{u_{k+1}^{s_{k+1}-q-1}}{\Gamma(s_{k+1}-q)}\,u_k^q$$

and integrate over the groups of variables $u_1,\dots,u_k$ and $u_{k+1},\dots,u_l$ separately. ☺

Combining equalities (2.15)–(2.17) we arrive at the following result.

PROPOSITION 2.8. *In the above notation,*

$$\zeta(\boldsymbol{s})=g_l(s_1,\dots,s_l;0;u)+f_l(s_1,\dots,s_l;u)$$
$$+\sum_{k=1}^{l-1}\sum_{q=0}^{s_{k+1}-1}\frac{(-1)^q}{q!}\,g_k(s_1,\dots,s_k;q;u)f_{l-k}(s_{k+1}-q,s_{k+2},\dots,s_l;u).$$

The expression found shows that, for *any* positive u, every MZV can be written as a finite sum of products of the functions $f(\boldsymbol{s};u)$ and $g(\boldsymbol{s};q;u)$. Our next step is to find efficient algorithms for computing these functions, at least for *some* $u>0$.

PROPOSITION 2.9. *For $s_1, \ldots, s_l$ positive integers and $u > 0$ real,*

$$f(\boldsymbol{s}; u) = \frac{1}{(s_1-1)!} \sum_{n_1>n_2>\cdots>n_l\geq 1} \left(-\frac{\partial}{\partial t}\right)^{s_1-1} \frac{e^{-ut}}{t}\bigg|_{t=n_1} \cdot \frac{1}{n_2^{s_2}\cdots n_l^{s_l}}.$$

The series for $f(\boldsymbol{s}; u)$ is rapidly convergent, as a geometric series with rate e^{-u}, especially for large u.

PROOF. Take $n_j = m_j + \cdots + m_l$ for $j = 1, \ldots, l$ and $n_{l+1} = 0$, so that

$$\begin{aligned} f(\boldsymbol{s}; u) &= \sum_{n_1>\cdots>n_l\geq 1} \int\cdots\int_{u_l>u_{l-1}>\cdots>u_1>u} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j(n_j-n_{j+1})} \frac{\mathrm{d}u_j}{\Gamma(s_j)} \\ &= \sum_{n_1>\cdots>n_l\geq 1} \int\cdots\int_{u_l>u_{l-1}>\cdots>u_1>u} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-(u_j-u_{j-1})n_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)}. \end{aligned}$$

Now notice that the summand is obtained from applying the differential operator

$$\left(-\frac{\partial}{\partial n_1}\right)^{s_1-1}\left(-\frac{\partial}{\partial n_2}\right)^{s_2-1}\cdots\left(-\frac{\partial}{\partial n_l}\right)^{s_l-1}$$

to

$$\begin{aligned} &\int\cdots\int_{u_l>u_{l-1}>\cdots>u_1>u} \prod_{j=1}^{l} e^{-(u_j-u_{j-1})n_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)} \\ &\quad = \int\cdots\int_{x_1>u,\, x_2>0,\, \ldots,\, x_l>0} \prod_{j=1}^{l} e^{-x_j n_j} \frac{\mathrm{d}x_j}{\Gamma(s_j)} = \frac{e^{-un_1}}{n_1 n_2\cdots n_l} \prod_{j=1}^{l} \frac{1}{(s_j-1)!}. \end{aligned}$$

This implies the desired form of $f(\boldsymbol{s}; u)$. ☺

To study $g(\boldsymbol{s}; q; u)$, we first observe that, when $u < 2\pi$, we can use

$$\sum_{m_j=1}^{\infty} e^{-u_j m_j} = \frac{e^{-u_j}}{1-e^{-u_j}} = \frac{1}{e^{u_j}-1}$$

within the range $0 < u_j < u$ for $j = 1, \ldots, l$, so that we can write

$$\begin{aligned} g(\boldsymbol{s}; q; u) &= \int\cdots\int_{u>u_l>u_{l-1}>\cdots>u_1>u_0} u_l^q \prod_{j=1}^{l} \frac{(u_j-u_{j-1})^{s_j-1}}{e^{u_j}-1} \frac{\mathrm{d}u_j}{\Gamma(s_j)} \\ &= \int_0^u \frac{\mathrm{d}u_l}{\Gamma(s_l)} \frac{u_l^q}{e^{u_l}-1} \int_0^{u_l} \frac{\mathrm{d}u_{l-1}}{\Gamma(s_{l-1})} \frac{(u_l-u_{l-1})^{s_l-1}}{e^{u_{l-1}}-1} \cdots \\ &\quad \times \int_0^{u_3} \frac{\mathrm{d}u_2}{\Gamma(s_2)} \frac{(u_3-u_2)^{s_3-1}}{e^{u_2}-1} \int_0^{u_2} \frac{\mathrm{d}u_1}{\Gamma(s_1)} \frac{(u_2-u_1)^{s_2-1} u_1^{s_1-1}}{e^{u_1}-1}. \end{aligned}$$

Recall that

$$\frac{z}{e^z - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!}\, z^k. \tag{2.18}$$

Change variables $u_j = v_j u_{j+1}$ for $j = 1, \dots, l-1$ and $u_l = v_l u$ in the resulting integral for $g(\boldsymbol{s}; q; u)$. We get

$$\begin{aligned}
&\int_0^{u_2} \frac{\mathrm{d}u_1}{\Gamma(s_1)}\, \frac{(u_2 - u_1)^{s_2-1} u_1^{s_1-1}}{e^{u_1} - 1} \\
&\quad = \int_0^1 \frac{u_2\, \mathrm{d}v_1}{\Gamma(s_1)}\, (u_2 - u_2 v_1)^{s_2-1} (u_2 v_1)^{s_1-2} \sum_{k_1 \ge 0} \frac{B_{k_1}}{k_1!}\, (u_2 v_1)^{k_1} \\
&\quad = \sum_{k_1 \ge 0} \frac{B_{k_1}}{k_1!}\, u_2^{k_1+s_1+s_2-2} \cdot \frac{1}{\Gamma(s_1)} \int_0^1 v_1^{k_1+s_1-2} (1 - v_1)^{s_2-1}\, \mathrm{d}v_1 \\
&\quad = \sum_{k_1 \ge 0} \frac{B_{k_1}}{k_1!}\, u_2^{k_1+s_1+s_2-2} \cdot \frac{\Gamma(s_2)\, \Gamma(k_1 + s_1 - 1)}{\Gamma(s_1)\, \Gamma(k_1 + s_1 + s_2 - 1)},
\end{aligned}$$

where the Euler integral of the first kind was used (see Exercise 1.2); then

$$\begin{aligned}
&\int_0^{u_3} \frac{\mathrm{d}u_2}{\Gamma(s_2)}\, \frac{(u_3 - u_2)^{s_3-1} u_2^{k_1+s_1+s_2-2}}{e^{u_2} - 1} \\
&\quad = \sum_{k_2 \ge 0} \frac{B_{k_2}}{k_2!}\, u_3^{k_1+k_2+s_1+s_2+s_3-3} \cdot \frac{\Gamma(s_3)\, \Gamma(k_1 + k_2 + s_1 + s_2 - 2)}{\Gamma(s_2)\, \Gamma(k_1 + k_2 + s_1 + s_2 + s_3 - 2)},
\end{aligned}$$

and so on. The final result reads

$$\begin{aligned}
g(\boldsymbol{s}; q; u) = \frac{1}{\Gamma(s_1)} &\sum_{k_1, \dots, k_l \ge 0} \prod_{j=1}^{l} \frac{B_{k_j}}{k_j!} \cdot u^{K_l+S_l+q-l}\, \frac{\Gamma(K_l + S_l + q - l)}{\Gamma(K_l + S_l + q + 1 - l)} \\
&\times \prod_{j=1}^{l-1} \frac{\Gamma(K_j + S_j - j)}{\Gamma(K_j + S_{j+1} - j)}\,,
\end{aligned}$$

where $K_j = k_1 + \dots + k_j$ and $S_j = s_1 + \dots + s_j$ for $j = 1, \dots, l$.

Proposition 2.10. *For* $0 < u < 2\pi$,

$$g(\boldsymbol{s}; q; u) = \frac{1}{\Gamma(s_1)} \sum_{k_1, \dots, k_l \ge 0} \prod_{j=1}^{l} \frac{B_{k_j}}{k_j!} \cdot \frac{u^{K_l+S_l+q-l}}{K_l + S_l + q - l} \cdot \prod_{j=1}^{l-1} \frac{\Gamma(K_j + S_j - j)}{\Gamma(K_j + S_{j+1} - j)}\,.$$

Because the series in (2.18) converges in the disk $|z| < 2\pi$, the sum we deduce for $g(\boldsymbol{s}; q; u)$ converges absolutely at the geometric rate $u/(2\pi)$.

Exercise 2.6. To compute the multiple sums for $f(\boldsymbol{s}; u)$ in Proposition 2.9 and for $g(\boldsymbol{s}; q; u)$ in Proposition 2.10 with a given accuracy ε, one needs (roughly) to sum the first expression over $n_1 \le N$ and the second one over

$k_1 + \cdots + k_l \le K$, where

$$N \sim \frac{-\log \varepsilon}{u} \quad \text{and} \quad K \sim \frac{\log \varepsilon}{\log(u/(2\pi))}.$$

What is an (approximate) optimal value u for the both sums? This can be used for computing the multiple zeta value $\zeta(\boldsymbol{s})$ numerically on the basis of Proposition 2.8.

EXERCISE 2.7. Implement a code for computing $\zeta(s)$ and $\zeta(s,t)$, single and double zeta values, where $s \ge 2$ and $t \ge 1$ are integers. Choose a reasonable accuracy for your numerical calculation, for example, 10^{-10}.

CHAPTER 3

Algebraic relations of multiple zeta values

In this part, we expose the standard algebraic setup of the MZVs. It is expected that all known algebraic relations (that is, numerical identities) over $\mathbb{Q}$ for the quantities (2.1) are produced by the so-called *double shuffle relations* which we describe below.

3.1. Algebra of multiple zeta values

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

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

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

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

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

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

$$x_0^j * w = w * x_0^j = wx_0^j \tag{3.5}$$

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

EXERCISE 3.1. Compute $x_0x_1 \sqcup\!\sqcup x_0x_1$ 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 $\mathfrak{H}_{\sqcup\!\sqcup} = (\mathfrak{H}, \sqcup\!\sqcup)$, $\mathfrak{H}^1_* = (\mathfrak{H}^1, *)$ (and also $\mathfrak{H}_* = (\mathfrak{H}, *)$) are examples of so-called *Hopf algebras*.

The following two statements motivate consideration of the introduced multiplications $\sqcup\!\sqcup$ and $*$.

THEOREM 3.1. *The map* ζ *is a homomorphism of the shuffle algebra* $\mathfrak{H}^0_{\sqcup\!\sqcup} = (\mathfrak{H}^0, \sqcup\!\sqcup)$ *into* $\mathbb{R}$*, that is,*

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

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

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

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

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*

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

(*in particular, the polynomials* $x_1 \sqcup\!\sqcup w - x_1 * w$ *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. This makes the following conjecture looking truthful.

CONJECTURE 3.4. *All linear relations over* $\mathbb{Q}$ *of multiple zeta values are generated by identities* (3.6)–(3.8)*; equivalently,*

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

In particular, the conjecture implies that all relations of MZVs over $\mathbb{Q}$ are homogeneous in weight.

EXERCISE 3.3. Using Theorems 3.1–3.3 show that:

(i) every MZV of weight 4 is a rational multiple of $\zeta(4)$;
(ii) every MZV of weight 5 is in the $\mathbb{Q}$-linear span of $\zeta(5)$ and $\zeta(2)\zeta(3)$;
(iii) every MZV of weight 6 is in the $\mathbb{Q}$-linear span of $\zeta(6)$ and $\zeta(3)^2$;

(iv) every MZV of weight 7 is in the $\mathbb{Q}$-linear span of $\zeta(7)$, $\zeta(2)\zeta(5)$ and $\zeta(2)^2\zeta(3)$.

In other words, any MZV of weight up to 7 can be expressed algebraically through the (single) zeta values $\zeta(s)$.

3.2. Shuffle algebra of generalised polylogarithms

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

$$\operatorname{Li}_{\boldsymbol{s}}(z)=\sum_{n_1>n_2>\cdots>n_l\ge1}\frac{z^{n_1}}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}\,,\qquad |z|<1, \tag{3.9}$$

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

$$\operatorname{Li}_{\boldsymbol{s}}(1)=\zeta(\boldsymbol{s}),\qquad \boldsymbol{s}\in\mathbb{Z}^l,\quad s_1\ge2,\ s_2\ge1,\ \dots,\ s_l\ge1. \tag{3.10}$$

Taking, as before for multiple zeta values,

$$\operatorname{Li}_{x_{\boldsymbol{s}}}(z)=\operatorname{Li}_{\boldsymbol{s}}(z),\qquad \operatorname{Li}_{\mathbf{1}}(z)=1, \tag{3.11}$$

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 3.5. *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*

$$\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_w(z)=\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_ju}(z)=\omega_j(z)\operatorname{Li}_u(z), \tag{3.12}$$

where

$$\omega_j(z)=\omega_{x_j}(z)=\begin{cases}\dfrac1z & \text{if } x_j=x_0,\\[2mm] \dfrac{1}{1-z} & \text{if } x_j=x_1.\end{cases} \tag{3.13}$$

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

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\operatorname{Li}_{x_0u}(z)&=\frac1z\sum_{n_1>n_2>\cdots>n_l\ge1}\frac{z^{n_1}}{n_1^{s_1-1}n_2^{s_2}\cdots n_l^{s_l}}\\ &=\frac1z\operatorname{Li}_{s_1-1,s_2,\dots,s_l}(z)=\frac1z\operatorname{Li}_u(z)\end{aligned}$$

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

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

and the result follows. ☺

Lemma 3.5 motivates another definition of the generalised polylogarithms, now defined for all elements of the algebra $\mathfrak{H}$. As before, it is sufficient to give it for words $w \in \mathfrak{H}$ only, distributing then over all algebra by linearity; set $\operatorname{Li}_{\mathbf{1}}(z) = 1$ and

$$\operatorname{Li}_w(z) = \begin{cases} \dfrac{\log^k z}{k!} & \text{if } w = x_0^k \text{ for some } k \ge 1, \\ \displaystyle\int_0^z \omega_j(z)\operatorname{Li}_u(z)\,\mathrm{d}z & \text{if } w = x_j u \text{ contains letter } x_1. \end{cases} \tag{3.14}$$

Evidently, Lemma 3.5 remains true for this extended version (3.14) of the polylogarithms (the fact yields coincidence of the newly-defined polylogarithms with the 'old' ones (3.11) for words w in $\mathfrak{H}^1$).

Exercise 3.4. (a) Compute $\operatorname{Li}_{x_1x_0}(z)$.
(b) Show that

$$\lim_{z\to 0^+} z^{-1/2}\operatorname{Li}_w(z) = 0 \quad \text{if the word } w \in \mathfrak{H} \text{ contains letter } x_1.$$

Hint. (a) It is standard to use

$$\log z = \frac{\mathrm{d}}{\mathrm{d}\delta}(z^\delta)\bigg|_{\delta=0}.$$

We get

$$\begin{aligned}\operatorname{Li}_{x_1x_0}(z) &= \int_0^z \omega_1(z)\operatorname{Li}_{x_0}(z)\,\mathrm{d}z = \int_0^z \frac{\log z\,\mathrm{d}z}{1-z} = \frac{\mathrm{d}}{\mathrm{d}\delta}\int_0^z \frac{z^\delta}{1-z}\,\mathrm{d}z\bigg|_{\delta=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\delta}\int_0^z \sum_{n=1}^{\infty} z^{n-1+\delta}\,\mathrm{d}z\bigg|_{\delta=0} = \frac{\mathrm{d}}{\mathrm{d}\delta}\sum_{n=1}^{\infty}\frac{z^{n+\delta}}{n+\delta}\bigg|_{\delta=0} \\ &= \sum_{n=1}^{\infty}\left(\frac{(\log z)\,z^n}{n} - \frac{z^n}{n^2}\right) = (\log z)\operatorname{Li}_1(z) - \operatorname{Li}_2(z).\end{aligned}$$

(b) Use the fact that if $f(z)$ is continuous on $(0,1)$ and $z^{-1/2}f(z) \to 0$ as $z \to 0^+$, then

$$F(z) = \int_0^z f(z)\,\mathrm{d}z$$

is also continuous on $(0,1)$ and satisfies $z^{-1/2}F(z) \to 0$ as $z \to 0^+$. Of course, this fact should be also established (by using traditional analysis techniques). ☺

LEMMA 3.6. *The map* $w \mapsto \operatorname{Li}_w(z)$ *is a homomorphism of the algebra* $\mathfrak{H}_{\sqcup\!\sqcup}$ *into* $C((0,1);\mathbb{R})$.

PROOF. We have to verify the equalities

$$\operatorname{Li}_{w_1 \sqcup\!\sqcup w_2}(z) = \operatorname{Li}_{w_1}(z)\operatorname{Li}_{w_2}(z) \quad \text{for all } w_1, w_2 \in \mathfrak{H}; \tag{3.15}$$

it is sufficient to do this job for *words* $w_1, w_2 \in \mathfrak{H}$. We will prove equality (3.15) by induction on the quantity $|w_1| + |w_2|$. If $w_1 = \mathbf{1}$ or $w_2 = \mathbf{1}$, relation (3.15) becomes tautological by (3.2). Otherwise, $w_1 = x_j u$ and $w_2 = x_k v$, hence by Lemma 3.5 and the inductive hypothesis we have

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

Thus

$$\operatorname{Li}_{w_1}(z)\operatorname{Li}_{w_2}(z) = \operatorname{Li}_{w_1 \sqcup\!\sqcup w_2}(z) + C, \tag{3.16}$$

and letting $z \to 0^+$ 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 (3.16) becomes the required relation (3.15), and the lemma follows. ☺

PROOF OF THEOREM 3.1. Theorem 3.1 follows from Lemma 3.6 and relations (3.10). ☺

EXERCISE 3.5. Show that

$$\operatorname{Li}_{\{1\}^k}(z) = \operatorname{Li}_{x_1^k}(z) = \frac{\operatorname{Li}_1(z)^k}{k!} = \frac{(-\log(1-z))^k}{k!}$$

for $k = 1, 2, \dots$.

Explicit computation of the monodromy group for the system of differential equations (3.12) allows to Minh, Petitot, and van der Hoeven to prove that the homomorphism $w \mapsto \operatorname{Li}_w(z)$ of the shuffle algebra $\mathfrak{H}_{\sqcup\!\sqcup}$ over $\mathbb{C}$ is bijective, that is, all $\mathbb{C}$-algebraic relations for generalised polylogarithms are originated from shuffle relations (3.15) only; in particular, generalised polylogarithms are linearly independent over $\mathbb{C}$. A much simpler proof of the linear independence of functions (3.9), as a consequence of elegant identities for the functions, is due to Ulanskiĭ.

Exercise 3.6. Verify that the dilogarithm function Li_2 satisfies the identity

$$\mathrm{Li}_2(x)+\mathrm{Li}_2(y)=\mathrm{Li}_2\Big(\frac{x}{1-y}\Big)+\mathrm{Li}_2\Big(\frac{y}{1-x}\Big)-\mathrm{Li}_2\Big(\frac{xy}{(1-x)(1-y)}\Big)$$
$$-\log(1-x)\log(1-y).$$

3.3. Duality of MZVs

By Lemma 3.5, the following integral representation is valid for the word $w=x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k}\in\mathfrak{H}^1$:

$$\begin{aligned}\mathrm{Li}_w(z)&=\int_0^z\omega_{\varepsilon_1}(z_1)\,\mathrm{d}z_1\int_0^{z_1}\omega_{\varepsilon_2}(z_2)\,\mathrm{d}z_2\cdots\int_0^{z_{k-1}}\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_k\\&=\int\cdots\int_{z>z_1>z_2>\cdots>z_{k-1}>z_k>0}\omega_{\varepsilon_1}(z_1)\omega_{\varepsilon_2}(z_2)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\,\mathrm{d}z_2\cdots\mathrm{d}z_k\end{aligned}\tag{3.17}$$

if $0<z<1$. When $x_{\varepsilon_1}\ne x_1$, i.e., $w\in\mathfrak{H}^0$, the integral in (3.17) converges in the region $0<z\le 1$, hence, in accordance with (3.10), we reduce representation for the multiple zeta values

$$\zeta(w)=\int\cdots\int_{1>z_1>\cdots>z_k>0}\omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k\tag{3.18}$$

in a form of *Chen's iterated integrals.*

There is a simple mnemonic way to write down the integral representation (3.18):

$$\zeta(x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k})=\int_0^1 x_{\varepsilon_1}x_{\varepsilon_2}\cdots x_{\varepsilon_k},\tag{3.19}$$

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 $\mathfrak{H}^0$.

The following result is an immediate application of the integral representation (3.18).

Theorem 3.7 (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 $z_1'=1-z_k$, $z_2'=1-z_{k-1}$, ..., $z_k'=1-z_1$, and apply relations $\omega_0(z)=\omega_1(1-z)$ followed from (3.13). ☺

As the simplest consequence of Theorem 3.7, notice (again) identity (1.13), which follows for the word $w = x_0^2x_1$, as well as the general identity

$$\zeta(n+2) = \zeta(2, \{1\}^n) \qquad n = 1, 2, \dots, \tag{3.20}$$

for the words $w = x_0^{n+1}x_1$. Recall our convention (see (2.3)) about $\{\boldsymbol{s}\}^n$ to denote the n-repetition of multi-index $\boldsymbol{s}$.

Exercise 3.7. Show that

$$\zeta(\{2,1\}^n) = \zeta(\{3\}^n), \qquad n = 1, 2, \dots. \tag{3.21}$$

For $n = 1$, this is again Euler's (1.13).

The integral representation (3.17) leads to a recipe for computing the MZVs. For this write as in (3.18),

$$\zeta(w) = \int\cdots\int\limits_{z_0>z_1>\cdots>z_k>z_{k+1}} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k$$

where for simplicity we set $z_0 = 1$ and $z_{k+1} = 0$. Now we take an arbitrary z in the interval $0 < z < 1$ and split the integration domain into the disjoint union of $k+1$ subdomains like it was done in Section 2.3:

$$\begin{aligned}
\zeta(w) &= \sum_{j=0}^{k} \int\cdots\int\limits_{z_0>\cdots>z_j>z>z_{j+1}>\cdots>z_{k+1}} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_1\cdots\mathrm{d}z_k \\
&= \sum_{j=0}^{k} \int\cdots\int\limits_{z_0>\cdots>z_j>z} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_j}(z_j)\,\mathrm{d}z_1\cdots\mathrm{d}z_j \\
&\qquad \times \int\cdots\int\limits_{z>z_{j+1}>\cdots>z_{k+1}} \omega_{\varepsilon_{j+1}}(z_{j+1})\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_{j+1}\cdots\mathrm{d}z_k \\
&= \sum_{j=0}^{k} \int\cdots\int\limits_{1-z>z'_j>\cdots>z'_1>0} \omega_{1-\varepsilon_j}(z'_j)\cdots\omega_{1-\varepsilon_1}(z'_1)\,\mathrm{d}z'_1\cdots\mathrm{d}z'_j \\
&\qquad \times \int\cdots\int\limits_{z>z_{j+1}>\cdots>z_{k+1}} \omega_{\varepsilon_{j+1}}(z_{j+1})\cdots\omega_{\varepsilon_k}(z_k)\,\mathrm{d}z_{j+1}\cdots\mathrm{d}z_k \\
&= \sum_{j=0}^{k} \operatorname{Li}_{\tau(x_{\varepsilon_1}\cdots x_{\varepsilon_j})}(1-z)\operatorname{Li}_{x_{\varepsilon_{j+1}}\cdots x_{\varepsilon_k}}(z).
\end{aligned}$$

Finally, making the choice $z = 1/2$ leads to the following formula.

Proposition 3.8. *For the multiple zeta value $\zeta(w)$ with $w = x_{\varepsilon_1}\cdots x_{\varepsilon_k} \in \mathfrak{H}^0$, we have*

$$\zeta(w) = \sum_{w=uv} \operatorname{Li}_{\tau u}\Big(\frac{1}{2}\Big)\operatorname{Li}_v\Big(\frac{1}{2}\Big),$$

where the sum runs over all possible ways of writing the word w as uv.

The efficiency of this formula follows from the fact that the series representation of any polylogarithm $\mathrm{Li}_w(z)$ converges at the geometric rate z; the convergence of the series at $z = 1/2$ is fast. At the same time, the computational scheme implied by Proposition 3.8 is much simpler than the one coming from Proposition 2.8.

Exercise 3.8. (a) Give the formula for $\zeta(n)$ in terms of polylogarithms evaluated at $z = 1/2$.

(b) Implement it for computing the zeta values for a given accuracy.

Hint. (a) Make use of Exercise 3.5. ☺

The iterated integral representations of MZVs and generalised polylogarithms motivate considering a slightly general than (2.1) version of MZVs, namely, the (*alternating* or '*alternative*') *Euler sums*

$$\zeta(s_1,\dots,s_l;\sigma_1,\dots,\sigma_l) = \sum_{n_1>n_2>\dots>n_l\ge 1} \frac{\sigma_1^{n_1}\sigma_2^{n_2}\cdots\sigma_l^{n_l}}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}, \tag{3.22}$$

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 3.9. Show that

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

$$\text{(b) } \zeta(\overline{2},1) = \frac{\zeta(3)}{8}.$$

In Section 4.3 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 (3.19) by allowing the three differential forms

$$\begin{gathered} x_0 \mapsto a = \omega_0(z)\,\mathrm{d}z = \frac{\mathrm{d}z}{z}, \quad x_1 \mapsto b = \omega_1(z)\,\mathrm{d}z = \frac{\mathrm{d}z}{1-z}, \\ \text{and} \quad \overline{x}_1 \mapsto c = \overline{\omega}_1(z)\,\mathrm{d}z = \frac{-\mathrm{d}z}{1+z}. \end{gathered} \tag{3.23}$$

3.4. Multiple harmonic sums

Another way to cast multiple zeta values $\zeta(\boldsymbol{s})$ is through the limiting case, as $N \to \infty$, of the *multiple harmonic sums* (MHSs)

$$H(\boldsymbol{s};N) = H(s_1,\dots,s_l;N) = \sum_{N\ge n_1>n_2>\dots>n_l\ge 1} \frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}}, \tag{3.24}$$

where $N = 0,1,2,\dots$; we also set $H(\,;N) = 1$ for the empty index $\boldsymbol{s}$. Given N, notice that these are *finite* sums, therefore well defined *for any* $s_1,\dots,s_l \in \mathbb{R}$.

This allows us, in our algebraic setting, to assign the MHS $H(x_{\boldsymbol{s}};N) = H(\boldsymbol{s};N)$ to any word

$$x_{\boldsymbol{s}} = x_0^{s_1-1}x_1x_0^{s_2-1}x_1\cdots x_0^{s_l-1}x_1 = y_{s_1}y_{s_2}\cdots y_{s_l} \in \mathfrak{H}^1.$$

Then we extend the map $H(\,\cdot\,;N)\colon w \mapsto H(w;N)$ defined on words $w \in \mathfrak{H}^1$ by linearity on the graded algebra $\mathfrak{H}^1$. Notice that

$$H(s_1,\dots,s_l;N) = \sum_{n_1=1}^{N} \frac{1}{n_1^{s_1}}\,H(s_2,\dots,s_l;n_1-1); \tag{3.25}$$

this iteration shares some analogy with integrating the polylogarithms using the differential equations in (3.12).

Fix $N \in \mathbb{Z}_{\ge 0}$. An easy calculation shows that

$$H(s_1;N)H(s_2;N) = H(s_1,s_2;N) + H(s_2,s_1;N) + H(s_1+s_2;N) \tag{3.26}$$

but also motivates the fact that the product of any two MHSs (of weights k and m) can be always represented as a $\mathbb{Z}$-linear combination of MHSs (all of weight $k+m$). More specifically, the following result takes place.

Lemma 3.9. *For any $N \in \mathbb{Z}_{\ge 0}$ and words $w_1, w_2 \in \mathfrak{H}^1$, we have*

$$H(w_1;N)H(w_2;N) = H(w_1 * w_2;N).$$

Here the stuffle product is defined by the rules (3.2), (3.4).

Proof. Recall the connection between a multi-index $\boldsymbol{s} = (s_1,\dots,s_l)$ and the word $w \in \mathfrak{H}^1$: it is assigned to $w = y_{s_1}\cdots y_{s_l}$. We prove the required identity by induction on the sum of lengths of the multi-indices corresponding to w_1 and w_2. Write $w_1 = y_j u$ and $w_2 = y_k v$ for $u = y_{s_1}\cdots y_{s_l}$ and $v = y_{r_1}\cdots y_{r_i}$. Then

$$\begin{aligned} &H(w_1;N)H(w_2;N) \\ &\quad= \sum_{N\ge n_0>n_1>\cdots>n_l\ge 1} \frac{1}{n_0^j n_1^{s_1}\cdots n_l^{s_l}} \sum_{N\ge m_0>m_1>\cdots>m_i\ge 1} \frac{1}{m_0^k m_1^{r_1}\cdots m_i^{r_i}} \end{aligned}$$

(we split the sum into three, according to whether $n_0 > m_0$, $n_0 < m_0$ or $n_0 = m_0$)

$$\begin{aligned} &= \sum_{n_0\le N} \frac{1}{n_0^j}\,H(s_1,\dots,s_l;n_0-1)H(k,r_1,\dots,r_i;n_0-1) \\ &\qquad+ \sum_{m_0\le N} \frac{1}{m_0^k}\,H(j,s_1,\dots,s_l;m_0-1)H(r_1,\dots,r_i;m_0-1) \\ &\qquad+ \sum_{n_0\le N} \frac{1}{n_0^{j+k}}\,H(s_1,\dots,s_l;n_0-1)H(r_1,\dots,r_i;n_0-1) \end{aligned}$$

(we apply the inductive hypothesis to the internal products)

$$
\begin{aligned}
&= \sum_{n_0=1}^{N} \frac{1}{n_0^j} H(u * y_k v; n_0 - 1) + \sum_{m_0=1}^{N} \frac{1}{m_0^k} H(y_j u * v; m_0 - 1) \\
&\quad + \sum_{n_0=1}^{N} \frac{1}{n_0^{j+k}} H(u * v; n_0 - 1) \\
&= H(y_j(u * y_k v); N) + H(y_k(y_j u * v); N) + H(y_{j+k}(u * v); N),
\end{aligned}
$$

where the property (3.25) was implemented at the final step. The result converts into $H(w_1 * w_2; N)$ according to the definition in (3.4). ☺

Lemma 3.9 means that the map

$$w \mapsto \{H(w; N) : N = 0, 1, 2, \dots\}$$

into the $\mathbb{Q}$-linear space of (rational-valued) sequences is a homomorphism of the stuffle algebra $\mathfrak{H}^1_*$.

Exercise 3.10. Show that

$$\frac{1}{1-z} \operatorname{Li}_{\boldsymbol{s}}(z) = \sum_{N=0}^{\infty} H(\boldsymbol{s}; N) z^N.$$

In other words, the left-hand side is the generating function of the sequence $\{H(\boldsymbol{s}; N) : N = 0, 1, 2, \dots\}$.

Proof of Theorem 3.2. This follows immediately from considering the limiting case of Lemma 3.9 as $N \to \infty$. ☺

Several other proofs Theorem 3.2 are known. For example, one can invent a functional model (viewing $H(w; N)$ as functions of N, not necessarily integral!) satisfying the shuffle relations in a way similar to our treatment of generalised polylogarithms in Section 3.2. Another proof exploits 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

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

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 $n_1 > n_2 > \cdots > n_l$ and $n'_1 > n'_2 > \cdots > n'_l$. By the above means, the homomorphism in Theorem 3.2 is defined as restriction of the homomorphism ϕ on $\mathfrak{H}^0$ by setting $t_n = 1/n$ for $n = 1, 2, \dots$.

EXERCISE 3.11 (Cartier). (a) For an admissible multi-index $\boldsymbol{s}$, prove the integral representation

$$\zeta(\boldsymbol{s}) = \underset{[0,1]^{|\boldsymbol{s}|}}{\int\cdots\int} \prod_{j=1}^{l-1} \frac{t_1 t_2 \cdots t_{s_1+\cdots+s_j}}{1 - t_1 t_2 \cdots t_{s_1+\cdots+s_j}} \cdot \frac{\mathrm{d}t_1\,\mathrm{d}t_2 \cdots \mathrm{d}t_{|\boldsymbol{s}|}}{1 - t_1 t_2 \cdots t_{s_1+s_2+\cdots+s_l}}\,, \tag{3.27}$$

where $l = \ell(\boldsymbol{s})$.

(b) Using part (a) show that

$$\zeta(s_1)\zeta(s_2) = \zeta(s_1+s_2) + \zeta(s_1,s_2) + \zeta(s_2,s_1) \quad \text{for } s_1 \ge 2,\ s_2 \ge 2,$$

which corresponds to the stuffle product of words y_{s_1} and y_{s_2} (see (3.26)).

HINT. (a) Integrate termwise the series

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

(b) Substitute $u = t_1 \cdots t_{s_1}$, $v = t_{s_1+1} \cdots t_{s_1+s_2}$ into the (elementary!) identity

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

and integrate over the hypercube $[0,1]^{s_1+s_2}$ using (3.27). ☺

The approach in Exercise 3.11 (b) can be extended to demonstrate Theorem 3.2 in its generality.

3.5. Quasi-shuffle products and derivations

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.

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 (that is, 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:

$$\mathbf{1} \circ w = w \circ \mathbf{1} = w \tag{3.28}$$

for any word w, and

$$a_j u \circ a_k v = a_j(u \circ a_k v) + a_k(a_j u \circ v) + [a_j, a_k](u \circ v) \tag{3.29}$$

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

$$[\,\cdot\,,\,\cdot\,]\colon \bar{A} \times \bar{A} \to \bar{A} \tag{3.30}$$

($\bar{A} = A \cup \{\mathbf{0}\}$) 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 $\mathfrak{A}_\circ = \mathfrak{H}_{ш}$ 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

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

EXERCISE 3.12. On the algebra $\mathfrak{A}$ with the given functional (3.30), 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 (3.28) and (3.29), respectively. Then $\mathfrak{A}_{\bar{\circ}} = (\mathfrak{A}, \bar{\circ})$ is a graded $\mathcal{K}$-algebra as well (commutative, if property (S3) holds).

Show that the algebras $\mathfrak{A}_\circ$ and $\mathfrak{A}_{\bar{\circ}}$ coincide.

HINT. Use induction on $\ell(w_1) + \ell(w_2)$ to demonstrate that

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

for all words $w_1, w_2 \in \mathcal{K}\langle A \rangle$. Note that property (S3) is not required in this derivation! ☺

LEMMA 3.10. *For any letter $a \in A$ and any words $u, v \in \mathfrak{A}$, the following identity holds:*

$$a \circ uv - (a \circ u)v = u(a \circ v - av). \tag{3.31}$$

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 (3.31) 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

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

which is the desired result. ☺

By a *derivation* of the (graded non-commutative polynomial) algebra $\mathfrak{A} = \mathcal{K}\langle A\rangle$ we mean a linear map $\delta\colon \mathfrak{A} \to \mathfrak{A}$ (of the graded $\mathcal{K}$-vector spaces) that satisfies the Leibniz rule

$$\delta(uv) = \delta(u)v + u\delta(v) \quad \text{for all } \ u, v \in \mathfrak{A}. \tag{3.32}$$

Exercise 3.13. Verify that the commutator of two derivations $[\delta_1, \delta_2] = \delta_1\delta_2 - \delta_2\delta_1$ is a derivation.

Therefore, 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 (3.32).

The next assertion gives examples of derivations of $\mathfrak{A}$, when the algebra possesses an additive multiplication $\circ$ with the properties (3.28) and (3.29).

Theorem 3.11. *For any letter $a \in A$, the map*

$$\delta_a\colon w \mapsto aw - a \circ w \tag{3.33}$$

is a derivation.

Proof. Linearity of the map δ_a is clear. By Lemma 3.10, 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 (3.33) is actually a derivation. ☺

Theorem 3.11 implies that the maps $\delta_{⧢}\colon \mathfrak{H} \to \mathfrak{H}$ and $\delta_*\colon \mathfrak{H}^1 \to \mathfrak{H}^1$, defined by the formulae

$$\delta_{⧢}\colon w \mapsto x_1w - x_1 ⧢ w, \quad \delta_*\colon w \mapsto y_1w - y_1 * w = x_1w - x_1 * w, \tag{3.34}$$

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 (3.34), obtained in accordance with (3.2)–(3.5), on the generators of the algebra:

$$\delta_{⧢}x_0 = -x_0x_1, \ \delta_{⧢}x_1 = -x_1^2, \quad \delta_*x_0 = 0, \ \delta_*x_1 = -x_1^2 - x_0x_1. \tag{3.35}$$

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 3.2. 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 $Dx_0=0$, $Dx_1=x_0x_1$ (that is, $Dy_s=y_{s+1}$ 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 3.12 (Derivation theorem). *For any word $w\in\mathfrak{H}^0$, the identity*

$$\zeta(Dw)=\zeta(\overline{D}w) \tag{3.36}$$

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 (2.4) corresponds to the element

$$\begin{aligned} Dw&=D(y_{s_1}y_{s_2}\cdots y_{s_l})\\ &=y_{s_1+1}y_{s_2}\cdots y_{s_l}+y_{s_1}y_{s_2+1}y_{s_3}\cdots y_{s_l}+\cdots+y_{s_1}\cdots y_{s_{l-1}}y_{s_l+1} \end{aligned}\tag{3.37}$$

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

$$\begin{aligned} \overline{D}w&=\tau D\bigl(x_0x_1^{s_l-1}x_0x_1^{s_{l-1}-1}\cdots x_0x_1^{s_2-1}x_0x_1^{s_1-1}\bigr)\\ &=\tau\sum_{\substack{k=1\\ s_k\ge2}}^{l}\sum_{j=0}^{s_k-2}x_0x_1^{s_l-1}\cdots x_0x_1^{s_{k+1}-1}x_0x_1^{j}x_0x_1^{s_k-j-1}x_0x_1^{s_{k-1}-1}\cdots x_0x_1^{s_1-1}\\ &=\sum_{\substack{k=1\\ s_k\ge2}}^{l}\sum_{j=0}^{s_k-2}x_0^{s_1-1}x_1\cdots x_0^{s_{k-1}-1}x_1x_0^{s_k-j-1}x_1x_0^{j}x_1x_0^{s_{k+1}-1}x_1\cdots x_0^{s_l-1}x_1 \end{aligned}\tag{3.38}$$

that corresponds to the right-hand side in (2.4). Applying now the map ζ to the both sides of obtained equalities (3.37) and (3.38), by Theorem 2.1 we deduce the required identity (3.36). ☺

Note that the condition $w\in\mathfrak{H}^0$ in Theorem 3.12 cannot be weakened; equality (3.36) is false for the word $w=x_1$:

$$\zeta(Dx_1)=\zeta(x_0x_1)\ne0=\zeta(\overline{D}x_1).$$

PROOF OF THEOREM 3.3. Comparing action (3.35) of derivations (3.34) 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 3.12 to the word $w\in\mathfrak{H}^0$ leads to the required equality:

$$\zeta(x_1ш w-x_1*w)=\zeta\bigl((\delta_*-\delta_{ш})w\bigr)=\zeta\bigl((\overline{D}-D)w\bigr)=\zeta(\overline{D}w)-\zeta(Dw)=0.$$

This completes the proof. ☺

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

$$\operatorname{Li}_{\boldsymbol{s}}(\boldsymbol{z}) = \operatorname{Li}_{s_1,s_2,\dots,s_l}(z_1,z_2,\dots,z_l) = \sum_{n_1>n_2>\dots>n_l\ge1} \frac{z_1^{n_1} z_2^{n_2} \cdots z_l^{n_l}}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}}, \tag{3.39}$$

was given by Waldschmidt. (As it is easily seen, specialising $z_2 = \dots = z_l = 1$ functions (3.39) become generalised polylogarithms (3.9).) We do not discuss properties of the functional model (3.39) here, except for the cases when $z_1, \dots, z_l \in \{\pm 1\}$ (see Sections 3.3 and 4.3).

Exercise 3.14. (a) Show that

$$\operatorname{Li}_{1,1}(x,y) = \operatorname{Li}_2\biggl(-\frac{x(1-y)}{1-x}\biggr) - \operatorname{Li}_2\biggl(-\frac{x}{1-x}\biggr) - \operatorname{Li}_2(xy).$$

(b) Use part (a) to compute the integral

$$\int_0^1 \biggl(\frac{\log\frac{1+x}{2}}{1-x} - \frac{\log\frac{1-x}{2}}{1+x}\biggr)\frac{\mathrm{d}x}{x}$$

in terms of the values of logarithm and dilogarithm.

Hints. (a) Use appropriate differentiation.

(b) Expand the integrand into a power series. ☺

CHAPTER 4

The generating-function method

Another application of differential equations for generalised polylogarithms, deduced in Lemma 3.5, 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

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

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 3.5, 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 *generalised hypergeometric series*.

4.1. Hypergeometric function

In order to show any reasonable result for MZVs using generating functions, we have to familiarise ourselves 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

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

denotes the Pochhammer symbol (see Section 1.4).

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 4.1. *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,

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

Lemma 4.2 (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 $|z| < 1$,

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

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 4.3 (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 4.2 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 4.1. (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

$$F\Big(a,1-a;c;\frac{1}{2}\Big).$$

4.2. Broadhurst's MZV evaluation

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

LEMMA 4.4. *The following equality holds:*

$$L_{3,1}(z,t)=F\big(\tfrac{1}{2}(1+i)t,-\tfrac{1}{2}(1+i)t;1;z\big)\cdot F\big(\tfrac{1}{2}(1-i)t,-\tfrac{1}{2}(1-i)t;1;z\big), \quad (4.1)$$

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

PROOF. Routine verification (with a help of Lemma 3.5 for the left-hand side) shows that the both sides of the required equality are annihilated by action of the differential operator

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

in addition, the first terms in z-expansions of the both sides in (4.1) 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 4.2. 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 4.5. *For any integer $n\ge 1$, the identity*

$$\zeta(\{3,1\}^n)=\frac{2\pi^{4n}}{(4n+2)!} \quad (4.2)$$

holds.

PROOF. By Lemma 4.3 (Gauss' summation formula),

$$F(a,-a;1;1)=\frac{1}{\Gamma(1-a)\Gamma(1+a)}=\frac{\sin\pi a}{\pi a}, \tag{4.3}$$

substituting $z=1$ into equality (4.1) yields

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

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

Identity (4.2) is not the unique example of application of generating functions. We present more identities of Borwein, Bradley and Broadhurst, similar to (4.2), for which the above method is also effective:

$$\begin{gathered}\zeta(\{2\}^n)=\frac{2(2\pi)^{2n}}{(2n+1)!}\left(\frac12\right)^{2n+1},\quad \zeta(\{4\}^n)=\frac{4(2\pi)^{4n}}{(4n+2)!}\left(\frac12\right)^{2n+1},\\
\zeta(\{6\}^n)=\frac{6(2\pi)^{6n}}{(6n+3)!},\\
\zeta(\{8\}^n)=\frac{8(2\pi)^{8n}}{(8n+4)!}\left(\left(1+\frac{1}{\sqrt2}\right)^{4n+2}+\left(1-\frac{1}{\sqrt2}\right)^{4n+2}\right),\\
\zeta(\{10\}^n)=\frac{10(2\pi)^{10n}}{(10n+5)!}\left(1+\left(\frac{1+\sqrt5}{2}\right)^{10n+5}+\left(\frac{1-\sqrt5}{2}\right)^{10n+5}\right),\end{gathered}\tag{4.4}$$

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 3.7).

EXERCISE 4.3. Prove (some) identities in (4.4).

EXERCISE 4.4. Show that

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

The family of identities

$$\zeta(\{2\}^{n+3}) + 2\zeta(\{2\}^n, 3, 3) = \zeta(2, 1, \{2\}^n, 3), \qquad n = 1, 2, \dots, \tag{4.5}$$

conjectured by Hoffman, stayed a conjecture for almost 20 years. It was finally proved by M. Hirose and N. Sato in 2017.

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:

$$\begin{aligned}\sum_{n=0}^{\infty}\zeta(2n+3)t^{2n} &= \sum_{k=1}^{\infty}\frac{1}{k^3(1-t^2/k^2)}\\ &= \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}\biggl(\frac{1}{2}+\frac{2}{1-t^2/k^2}\biggr)\prod_{l=1}^{k-1}\biggl(1-\frac{t^2}{l^2}\biggr),\\ \sum_{n=0}^{\infty}\zeta(4n+3)t^{4n} &= \sum_{k=1}^{\infty}\frac{1}{k^3(1-t^4/k^4)}\\ &= \frac{5}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k^3\binom{2k}{k}}\frac{1}{1-t^4/k^4}\prod_{l=1}^{k-1}\frac{1+4t^4/l^4}{1-t^4/l^4}.\end{aligned} \tag{4.6}$$

Their proofs as well as proofs of several other identities is based on transformation and summation formulae of generalised hypergeometric functions, similar to application of formula (4.3) in deducing Theorem 4.5.

Identities (4.6) 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

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

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

$$\sum_{k=1}^{\infty}\frac{k}{k^4-s^2k^2-t^4} = \frac{1}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}r(n)}{n\binom{2n}{n}}\frac{\prod_{m=1}^{n-1}((m^2-s^2)^2+4t^4)}{\prod_{m=n}^{2n}(m^4-s^2m^2-t^4)}, \tag{4.7}$$

where

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

Formula (4.7) 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 4.5. Using (4.7), find fast converging series for $\zeta(5)$ and $\zeta(7)$.

4.3. The identity of Borwein, Bradley and Broadhurst

The family of identities

$$\zeta(\{\overline{2}, 1\}^n) = \frac{1}{8^n}\,\zeta(\{3\}^n), \qquad n = 1, 2, \dots, \tag{4.8}$$

conjectured by Borwein, Bradley and Broadhurst in 1996 (see (3.22) for the definition of alternating Euler sums), generalises Exercises 3.7 and 3.9 (b) and looks very similar to that in Theorem 4.5. It was proven only recently by Zhao using the standard (double shuffle) relations for the alternating Euler sums; a proof by generating functions is still wanted. Here we sketch Zhao's only-known proof of (4.8).

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. (3.23)) 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 (3.19)). 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 4.6. Write the word which corresponds to the multi-index

$$(2, 1, \overline{2}, 3, \overline{4}, \overline{5}).$$

Exercise 4.7. Prove the following correspondence:

$$(\{\overline{2}, 1\}^n) \mapsto (x_0\overline{x}_1^2x_0x_1^2)^{\lfloor n/2 \rfloor}(x_0\overline{x}_1^2)^{2\{n/2\}} = \begin{cases} (x_0\overline{x}_1^2x_0x_1^2)^k(x_0\overline{x}_1^2) & \text{if } n = 2k+1, \\ (x_0\overline{x}_1^2x_0x_1^2)^k & \text{if } n = 2k. \end{cases}$$

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

$$y_ju*y_kv = y_j\gamma_{y_j}(\gamma_{y_j}u*y_kv)+y_k\gamma_{y_k}(y_ju*\gamma_{y_k}v)+[y_j, y_k]\gamma_{[y_j,y_k]}(\gamma_{y_j}u*\gamma_{y_k}v), \quad (4.9)$$

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

$$\zeta(w_1 ⧢ w_2) = \zeta(w_1 * w_2) = \zeta(w_1)\zeta(w_2).$$

For a word $w = a_1a_2\cdots a_m$ over the alphabet $\{x_0, x_1, \overline{x}_1\}$, define the ith shuffle iteration by

$$⧢_i w = \begin{cases} a_1a_2\cdots a_i ⧢ a_{i+1}\cdots a_m & \text{if } i \text{ is odd}, \\ a_i\cdots a_2a_1 ⧢ a_{i+1}\cdots a_m & \text{if } i \text{ is even}, \end{cases} \qquad i = 0, 1, \ldots, m.$$

Similarly, but considering a word over the infinite alphabet $\{\overline{y}_0, y_1, \overline{y}_1, \ldots\}$, define the ith harmonic (stuffle) iteration $*_i$. Finally, define the $\star$-concatenation by settling $w_1 \star w_2 = w_1w_2$ except that

$$x_1 \star x_1 = x_1\overline{x}_1 \quad \text{and} \quad \overline{x}_1 \star \overline{x}_1 = \overline{x}_1x_1.$$

Exercise 4.8. Prove by induction that for every positive n,

$$\sum_{i=0}^{2n}(-1)^i *_i \big((\overline{x}_1z)^{\star n}\big) = (-1)^n\big(x_0^2(x_1+\overline{x}_1)\big)^n,$$

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 $z \star x_1 = z \star \overline{x}_1 = x_0(x_1\overline{x}_1 + \overline{x}_1x_1)$.

Exercise 4.9. Prove by induction that for every positive n,

$$\sum_{i=0}^{2n}(-1)^i ⧢_i \big((\overline{x}_1z)^{\star n}\big) = (-2)^n(x_0\overline{x}_1^2x_0x_1^2)^{\lfloor n/2 \rfloor}(x_0\overline{x}_1^2)^{2\{n/2\}}$$

and

$$\sum_{i=0}^{2n}(-1)^i ⧢_i \big((x_1z)^{\star n}\big) = (-2)^n(x_0x_1^2x_0\overline{x}_1^2)^{\lfloor n/2 \rfloor}(x_0x_1^2)^{2\{n/2\}}.$$

Exercise 4.10 (Distribution relation). Show that for every positive n,

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

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

Using Exercises 4.7–4.10, we deduce identity (4.8).

CHAPTER 5

Further relations of MZVs

5.1. Ihara–Kaneko derivations and Ohno's relations

Theorem 3.12 has a natural generalization. For any $n \geq 1$, 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 5.1. *For any $n \geq 1$ and any word $w \in \mathfrak{H}^0$, the identity*

$$\zeta(\partial_n w) = 0 \tag{5.1}$$

holds.

In what follows, we describe a scheme of the proof of the theorem given by Kaneko and Ihara (a different proof was provided by Hoffman and Ohno).

The following result contains as particular cases Theorems 2.1, 2.5 and 3.7 (corresponding implications are given by Ohno).

THEOREM 5.2 (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 $n \geq 0$, the identity

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

holds.

For each integer $n \geq 1$ define the derivation $D_n \in \mathrm{Der}(\mathfrak{H})$ 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 $\delta \in \mathrm{Der}(\mathfrak{H})$ 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 5.3. *For any word $w \in \mathfrak{H}^0$, the inclusion*

$$(\sigma - \overline{\sigma})w \in \ker\zeta \tag{5.2}$$

holds.

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

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

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

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

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

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

thus $\sigma w - \sigma\tau w \in \ker\zeta$ by Theorem 5.2. Applying now Theorem 3.7, we arrive at the desired inclusion (5.2). ☺

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 5.4. *The following equality holds:*

$$\exp(\partial) = \overline{\sigma}\cdot\sigma^{-1}. \tag{5.3}$$

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 that $\partial_n(x_0 + x_1)^k = 0$ for any $n \ge 1$ and $k \ge 0$, we obtain the desired property:

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

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

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

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

where $t \in \mathbb{R}$. 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 (5.3). ☺

Proof of Theorem 5.1. Now let us show how Theorem 5.1 follows from Theorem 5.3 and Lemma 5.4. 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 5.3 yields the required identities (5.1). ☺

Does there exist a simpler way of proving relations (5.1)? Explicit computations show that $\partial_1 = \delta_* - \delta_{ш}$,

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

and, in addition, $\delta_* + \overline{\delta}_* = \delta_{ш} + \overline{\delta}_{ш}$; therefore cases $n = 1, 2, 3, 4$ in Theorem 5.1 are served by induction (with Theorem 3.12 as inductive base). This circumstance motivates the following hypothesis.

Conjecture 5.5. *For any $n \geq 1$, 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* $\overline{\delta}_{ш}$.

5.2. Proof of Ohno's relations

In this section, we will discuss the original proof of Theorem 5.2 given by Ohno in 1999.

For an admissible multi-index $\boldsymbol{s} = (s_1, s_2, \ldots, s_l)$ and an integer $n \geq 0$, denote

$$Z(\boldsymbol{s}; n) = \sum_{\substack{e_1, e_2, \ldots, e_l \geq 0 \\ e_1+e_2+\cdots+e_l=n}} \zeta(s_1+e_1, s_2+e_2, \ldots, s_l+e_l)$$
$$= \sum_{\substack{e_1, e_2, \ldots, e_l \geq 0 \\ e_1+e_2+\cdots+e_l=n}} \zeta(x_0^{s_1+e_1-1} x_1 x_0^{s_2+e_2-1} x_1 \cdots x_0^{s_l+e_l-1} x_1),$$

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

$$x_0^{s_1-1} x_1 x_0^{s_2-1} x_1 \cdots x_0^{s_l-1} x_1 = x_0^{\mu_1} x_1^{\nu_1} x_0^{\mu_2} x_1^{\nu_2} \cdots x_0^{\mu_k} x_1^{\nu_k}, \tag{5.4}$$

where all the exponents are positive integers, then

$$Z(\boldsymbol{s}; n) = \sum_{\substack{q_1, \ldots, q_k \geq 0 \\ q_1+\cdots+q_k=n}} \Sigma_{\boldsymbol{s}}(\nu_1, \ldots, \nu_k; q_1, \ldots, q_k)$$

in notation

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

Furthermore, we have

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

To compute the latter expression, we use the integral representation from (3.18); performing the integration for each subword $x_0^{\mu_i-1} x_0 (x_0+T_ix_1)^{\nu_i+q_i-1} x_1$, $i =$

$1, \ldots, k$, we obtain

$$\int\!\!\cdots\!\!\int\limits_{t_{2i-2}>z_1>\cdots>z_{\mu_i-1}>t_{2i-1}>z'_1>\cdots>z'_{\nu_i+q_i-1}>t_{2i}} \frac{\mathrm{d}z_1}{z_1}\cdots\frac{\mathrm{d}z_{\mu_i-1}}{z_{\mu_i-1}}\,\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}}$$
$$\times\left(\frac{\mathrm{d}z'_1}{z'_1}+T_i\frac{\mathrm{d}z'_1}{1-z'_1}\right)\cdots\left(\frac{\mathrm{d}z'_{\nu_i+q_i-1}}{z'_{\nu_i+q_i-1}}+T_i\frac{\mathrm{d}z'_{\nu_i+q_i-1}}{1-z'_{\nu_i+q_i-1}}\right)\frac{\mathrm{d}t_{2i}}{1-t_{2i}}$$
$$=\frac{1}{(\mu_i-1)!\,(\nu_i+q_i-1)!}\iint\limits_{t_{2i-2}>t_{2i-1}>t_{2i}}\left(\log\frac{t_{2i-2}}{t_{2i-1}}\right)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}}$$
$$\times\left(\log\frac{t_{2i-1}}{t_{2i}}-T_i\log\frac{1-t_{2i-1}}{1-t_{2i}}\right)^{\nu_i+q_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}},$$

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

$$\frac{1}{(\mu_i-1)!\,q_i!\,(\nu_i-1)!}\iint\limits_{t_{2i-2}>t_{2i-1}>t_{2i}}\left(\log\frac{t_{2i-2}}{t_{2i-1}}\right)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}}\left(\log\frac{t_{2i-1}}{t_{2i}}\right)^{q_i}$$
$$\times\left(-\log\frac{1-t_{2i-1}}{1-t_{2i}}\right)^{\nu_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}}$$

and we finally arrive at

$$\Sigma_{\boldsymbol{s}}(\nu_1,\ldots,\nu_k;q_1,\ldots,q_k)$$
$$=\frac{1}{\prod_{i=1}^k(\mu_i-1)!\,q_i!\,(\nu_i-1)!}\int\!\!\cdots\!\!\int\limits_{1>t_1>\cdots>t_{2k}>0}\prod_{i=1}^k\left(\log\frac{t_{2i-2}}{t_{2i-1}}\right)^{\mu_i-1}\frac{\mathrm{d}t_{2i-1}}{t_{2i-1}}$$
$$\times\left(\log\frac{t_{2i-1}}{t_{2i}}\right)^{q_i}\left(\log\frac{1-t_{2i}}{1-t_{2i-1}}\right)^{\nu_i-1}\frac{\mathrm{d}t_{2i}}{1-t_{2i}}$$

and

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

with the convention $t_0=1$.

In the latter integral we introduce the change of variables

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

so that

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

Denote the expression

$$\begin{aligned}\biggl(\prod_{i=1}^{k}\frac{t_{2i-1}}{t_{2i}}\biggr)^{-1}&=\prod_{i=1}^{k}\frac{t_{2i-2}}{t_{2i-1}}\cdot t_{2k}\\&=\exp\biggl(\sum_{i=1}^{k}u_{2i-1}\biggr)\cdot\sum_{j=0}^{2k}(-1)^j\exp\biggl(\sum_{i=j+1}^{2k}(-1)^iu_i\biggr)\end{aligned}$$

by $f(u_1,u_2,\dots,u_{2k-1},u_{2k})$; it 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 (5.4) shows, it reduces the resulting expression to $Z(\boldsymbol{s}';n)$. This completes the proof of Theorem 5.2. ☺

It is straightforward that case $n=0$ in Theorem 5.2 is the duality theorem (Theorem 3.7).

EXERCISE 5.1. (a) Show that the choice $n=1$ in Theorem 5.2 corresponds to Hoffman's relations (Theorem 2.1).

(b) Show that, if multi-index $\boldsymbol{s}$ in Theorem 5.2 is one-component (that is, $\boldsymbol{s}=(s)$), then the theorem reduces to the sum theorem (Theorem 2.5).

CHAPTER 6

The two-one formula and its relatives

6.1. Open questions

In addition to Conjectures 1.10, 3.4 and 5.5 given earlier, 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 6.1. *As a $\mathbb{Q}$-algebra, the algebra $\mathcal{Z}$ is the direct sum of the subspaces $\mathcal{Z}_k$, where $k = 0, 1, 2, \dots$.*

It can be easily seen that relations (3.6)–(3.8) for multiple zeta values are homogeneous in weight, hence Conjecture 6.1 follows from Conjecture 3.4.

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 Exercise 3.3 (i)). For $k \geq 5$, above-deduced identities allow to compute the upper bounds; for instance, $d_5 \leq 2$, $d_6 \leq 2$, $d_7 \leq 3$ (see Exercise 3.3), and so on.

CONJECTURE 6.2. *For $k \geq 3$, the recurrence relations*

$$d_k = d_{k-2} + d_{k-3} \tag{6.1}$$

hold; equivalently,

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

It is now shown $\dim_{\mathbb{Q}} \mathcal{Z}_k \leq d_k$ for all k, where the sequence d_k is defined by the recursion (6.1) (and $d_0 = d_2 = 1$, $d_1 = 0$). There are several proofs of this result, due to Terasoma, to Deligne–Goncharov and to Brown; all are algebraic and use motivic interpretations of the multiple zeta values.

Even if Conjectures 6.1 and 6.2 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 natural.

CONJECTURE 6.3 (Hoffman's basis). *For any $k = 0, 1, 2, \dots$, a basis of the $\mathbb{Q}$-spaces $\mathcal{Z}_k$ is given by the set of numbers*

$$\bigl\{\zeta(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2, 3\},\ j = 1, \dots, \ell(\boldsymbol{s})\bigr\}. \tag{6.2}$$

Exercise 6.1. For given $k = 0, 1, 2, \dots$, show that the number of MZVs in (6.2) is equal to d_k. Here d_k is the same sequence defined earlier in (6.1).

A serious argument for Conjecture 6.3 to be valid, is not only experimental confirmation for $k \le 16$ (under the hypothesis of Conjecture 3.4) but also agreement of the dimension of the $\mathbb{Q}$-space spanned by the numbers (6.2) with the dimension d_k of the spaces $\mathcal{Z}_k$ in Conjecture 6.2 (Exercise 6.1). In his recent work, F. Brown shows that Conjecture 6.3 is true for a 'motivic' version of MZVs; in particular, that all usual MZVs can be expressed by means of the elements (6.2) of Hoffman's basis. This then implies the truth of the bounds $\dim_{\mathbb{Q}} \mathcal{Z}_k \le d_k$ for all k, discussed above in relation with Conjecture 6.3.

In the heart of Brown's 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 Section 6.3.

Exercise 6.2. (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 6.2.

(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

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

Assuming Conjecture 1.10, 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 $k \to \infty$.

Before going into details of Zagier's formula underlying Brown's proof we tackle another general relation of MZVs also motivated by Hoffman's basis.

6.2. The two-one formula

In the introductory section the following alternative version of the multiple zeta values with non-strict inequalities was mentioned (see (2.2)):

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

Exercise 2.2 gives a simple recipe to pass from one model to the other.

Relation (2.3) 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 (4.4) and reproduced in (6.13) below.)

The starting goal of our joint project with Ohno (in 2006) was not just finding a general form of the two families of identities for the MZSVs but searching for alternatives of Hoffman's basis (6.2) in terms of multiple zeta star values. Note the one can replace that basis with its dual (the order is swapped and each 3 is replaced with 2, 1)

$$\big\{\zeta(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2,1\},\ j = 1,\dots,\ell(\boldsymbol{s}),\ \text{no 1s next to each other}\big\}. \tag{6.3}$$

Another choice of the basis

$$\big\{\zeta^\star(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2,3\},\ j = 1,\dots,\ell(\boldsymbol{s})\big\}.$$

was also proposed at the time, and later shown to have similar quality as (6.2) by Zagier and Brown. Essentially, the original question was whether one could replace (conjecturally) the MZVs in the 'dual' Hoffman's basis (6.3) with MZSVs. We found that this is not the case already in weight 12 by showing that $\zeta^\star(\{2,1\}^4)$ is a rational multiple of π^{12}, hence of $\zeta^\star(\{2\}^6)$. However, on this way we succeeded in generalising (2.3), conjecturally. Some particular cases of our conjecture — dubbed as the 'two-one formula' — were established by ourselves, and it was finally proved in full generality by Zhao in 2013. One of lucky accidents of our proofs was a discovery of the weighted version (2.12) of Euler's original formula (2.11) (the sum formula of depth 2 in the modern terminology).

Theorem 6.4 (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:*

$$\zeta^\star(\mu_{s_1}, \mu_{s_2}, \dots, \mu_{s_l}) = \sum_{\boldsymbol{p}} (-1)^{\sigma(\boldsymbol{p})} 2^{l-\sigma(\boldsymbol{p})} \zeta^\star(\boldsymbol{p}) \tag{6.4}$$

$$= \sum_{\boldsymbol{p}} 2^{l-\sigma(\boldsymbol{p})} \zeta(\boldsymbol{p}), \tag{6.5}$$

where, as in Exercise 2.1, $\boldsymbol{p}$ 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 $\boldsymbol{p}$.

Surprisingly enough, the pattern in (6.4), (6.5) is similar to that in Exercise 2.1. One particular instance corresponding to $l = 2$,

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

was shown to be true in our original work with Ohno (by an elaborate descending inductive argument given in eight lemmas!). It implies the equality

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

when $s_1, s_2 \geq 1$, which does not seem to be generalisable further to cases $l > 2$. A related formula

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

for any positive integers m, n was given two different proofs are given by Zlobin and Ohno–Wakabayashi. If $m = 1$ it is nothing but formula (2.3), while if $m \geq 2$ then its left-hand side equals $\zeta^\star(\{\mu_3, \{\mu_1\}^{m-2}\}^n, \mu_1)$, so that the two-one formula implies the closed-form evaluation of the corresponding right-hand side in (6.4) (equivalently, in (6.5)) by means of the single zeta value $(m+1)\zeta((m+1)n+1)$, where the integers $m \geq 2$ and $n \geq 1$ are arbitrary.

Exercise 6.3. Show the equality of the right-hand sides in (6.4) and (6.5).

Hint. Use Exercise 2.1. ☺

On the right-hand side of (6.4) and (6.5) 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 was the reason of dubbing the formula as the two-one formula. The formula does not seem to be a specialization of identities for polylogarithms (3.9) but, after Zhao's proof, is linked to the multiple harmonic sums (3.24), their star counterparts

$$H^\star(\boldsymbol{s}; N) = H^\star(s_1, \dots, s_l; N) = \sum_{N \geq n_1 \geq n_2 \geq \dots \geq n_l \geq 1} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}},$$

but also to a different type

$$\widehat{H}(\boldsymbol{s}; N) = \sum_{N \geq n_1 > n_2 > \dots > n_l \geq 1} \frac{N!^2}{(N-n_1)!\,(N+n_1)!} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_l^{s_l}},$$

where $N = 0, 1, 2, \dots$ and $\widetilde{H}(\,;N) = \widehat{H}(\,;N) = 1$ for the empty index $\boldsymbol{s}$. As seen earlier

$$\lim_{n\to\infty} H(\boldsymbol{s}; N) = \zeta(\boldsymbol{s}) \quad \text{and} \quad \lim_{n\to\infty} H^\star(\boldsymbol{s}; N) = \zeta^\star(\boldsymbol{s})$$

when $s_1 > 1$.

Exercise 6.4. Show that for admissible multi-indices $\boldsymbol{s}$, we have

$$\lim_{n\to\infty} \widehat{H}(\boldsymbol{s}; N) = \zeta(\boldsymbol{s}).$$

Hint. The limit relation is equivalent to showing that, for $k \geq 2$ and any multi-index $\boldsymbol{s} = (s_1, \dots, s_l)$,

$$\lim_{N\to\infty} \sum_{m=1}^{N} \frac{H(\boldsymbol{s}; m-1)}{m^k} \left(1 - \frac{N!^2}{(N-m)!\,(N+m)!}\right) = 0. \tag{6.6}$$

(Notice that the expression in the parentheses is always positive.) Try first to prove (6.6) in the toughest possible case $k = 2$, $s_1 = \dots = s_l = 1$. One possible strategy is to split the sum into two, according to $m \leq \sqrt{N}$ and $m > \sqrt{N}$; use an estimate for the expression in the parentheses for the first sum and some trivial estimates for the second one. ☺

With the above notation in mind Theorem 6.4 is the limiting case, as $N \to \infty$, of the following result.

THEOREM 6.5. *For any* $N \in \mathbb{N}$,

$$H^\star(\{2\}^{s_1}, 1, \{2\}^{s_2}, 1, \dots, \{2\}^{s_l}, 1; N) = 2 \sum_{\boldsymbol{p}=(2s_1+1)\circ(2s_2+1)\circ\dots(2s_l+1)} 2^{\bar\sigma(\boldsymbol{p})} \widehat{H}(\boldsymbol{p}; N), \tag{6.7}$$

where $\circ$ *is either comma or plus and* $\bar\sigma(\boldsymbol{p})$ *denotes the exact number of commas.*

PROOF. For aesthetic reasons we will write $H^\star_N(\boldsymbol{s})$ and $\widehat{H}_N(\boldsymbol{s})$ for $H^\star(\boldsymbol{s}; N)$ and $\widehat{H}(\boldsymbol{s}; N)$, respectively. The proof of (6.7) is by induction on $N + l$. As $H^\star_1(\boldsymbol{s}) = 1$ for any $\boldsymbol{s}$ and $\widehat{H}_1(\boldsymbol{s}) = 1/2$ if $l = 1$ and 0 otherwise, the equality in (6.7) is trivially true when $N = 1$ and $l \geq 0$ is arbitrary.

Furthermore, assume that $N > 1$ and use the definition to write

$$\begin{aligned} &H^\star_N(\{2\}^{s_1}, 1, \{2\}^{s_2}, 1, \dots, \{2\}^{s_l}, 1) \\ &\quad= \sum_{k=0}^{s_1} \frac{1}{n^{2s_1-2k}} H^\star_{N-1}(\{2\}^k, 1, \{2\}^{s_2}, 1, \dots, \{2\}^{s_l}, 1) \\ &\qquad+ \frac{1}{N^{2s_1+1}} H^\star_N(\{2\}^{s_2}, 1, \dots, \{2\}^{s_l}, 1). \end{aligned}$$

Applying the induction statement to the newer multiple harmonic sums we obtain

$$\begin{aligned} &H^\star_N(\{2\}^{s_1}, 1, \{2\}^{s_2}, 1, \dots, \{2\}^{s_l}, 1) \\ &\quad= \frac{2}{N^{2s_1}} \sum_{k=0}^{s_1} N^{2k} \sum_{\boldsymbol{p}=(2k+1)\circ(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar\sigma(\boldsymbol{p})} \widehat{H}_{N-1}(\boldsymbol{p}) \\ &\qquad+ \frac{2}{N^{2s_1+1}} \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar\sigma(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p}). \end{aligned} \tag{6.8}$$

Using then the geometric sum

$$\sum_{k=0}^{s_1} \left(\frac{N}{n_1}\right)^{2k} = \frac{1}{n_1^{2s_1}} \frac{N^{2s_1+2} - n_1^{2s_1+2}}{(N - n_1)(N + n_1)}$$

we deduce that

$$\sum_{k=0}^{s_1} N^{2k} \widehat{H}_{N-1}(p_1 + 2k, p_2, \dots, p_r)$$

$$= N^{2s_1} \sum_{N>n_1>n_2>\cdots>n_r\geq 1} \frac{N!^2}{(N-n_1)!\,(N+n_1)!} \frac{1}{n_1^{p_1+2s_1} n_2^{p_2} \cdots n_r^{p_r}}$$

$$- \frac{1}{N^2} \sum_{N>n_1>n_2>\cdots>n_r\geq 1} \frac{N!^2}{(N-n_1)!\,(N+n_1)!} \frac{1}{n_1^{p_1-2} n_2^{p_2} \cdots n_r^{p_r}}$$

$$= N^{2s_1} \sum_{N\geq n_1>n_2>\cdots>n_r\geq 1} \frac{N!^2}{(N-n_1)!\,(N+n_1)!} \frac{1}{n_1^{p_1+2s_1} n_2^{p_2} \cdots n_r^{p_r}}$$

$$- \frac{1}{N^2} \sum_{N\geq n_1>n_2>\cdots>n_r\geq 1} \frac{N!^2}{(N-n_1)!\,(N+n_1)!} \frac{1}{n_1^{p_1-2} n_2^{p_2} \cdots n_r^{p_r}}$$

$$= N^{2s_1} \widehat{H}_N(p_1+2s_1, p_2, \ldots, p_r) - \frac{1}{N^2} \widehat{H}_N(p_1-2, p_2, \ldots, p_r).$$

Therefore, the equality in (6.8) can be written as

$$H_N^\star(\{2\}^{s_1}, 1, \{2\}^{s_2}, 1, \ldots, \{2\}^{s_l}, 1) - 2 \sum_{\boldsymbol{p}=(2s_1+1)\circ(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p})$$

$$= \frac{2}{N^{2s_1+1}} \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p})$$

$$- \frac{2}{N^{2s_1+2}} \sum_{\boldsymbol{p}=(-1)\circ(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p})$$

(we expand the first $\circ$ in $\boldsymbol{p} = (-1) \circ (2s_2+1) \circ \cdots \circ (2s_l+1)$)

$$= \frac{2}{N^{2s_1+1}} \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p})$$

$$- \frac{2}{N^{2s_1+2}} \sum_{\boldsymbol{p}=(2s_2)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p})$$

$$- \frac{4}{N^{2s_1+2}} \sum_{m=1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})} H_{m-1}(\boldsymbol{p}). \tag{6.9}$$

Finally, the other identity

$$2 \sum_{m=1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m\, H_{m-1}(p_1, p_2, \ldots, p_r)$$

$$= 2 \sum_{m=1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m \sum_{n_1=1}^{m-1} \frac{H_{n_1-1}(p_2, \ldots, p_r)}{n_1^{p_1}}$$

$$= \sum_{n_1=1}^{N} \frac{H_{n_1-1}(p_2, \ldots, p_r)}{n_1^{p_1}} \cdot 2 \sum_{m=n_1+1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m$$

(the internal sum is summed by Exercise 6.5 below)

$$= \sum_{n_1=1}^{N} \frac{H_{n_1-1}(p_2,\dots,p_r)}{n_1^{p_1}} \cdot \frac{(N-n_1)\,N!^2}{(N-n_1)!\,(N+n_1)!}$$
$$= N\widehat{H}_N(p_1,p_2,\dots,p_r) - \widehat{H}_N(p_1-1,p_2,\dots,p_r)$$

simplifies the right-hand side of (6.9) to zero. ☺

EXERCISE 6.5. For integers $N>0$ and $n\ge 0$, show

$$2\sum_{m=n+1}^{N} \frac{m\binom{N}{m}}{\binom{N+m}{m}} = \frac{N\binom{N-1}{n}}{\binom{N+n}{n}}.$$

HINT. Use a telescoping argument: verify that

$$\frac{2m\binom{N}{m}}{\binom{N+m}{m}} = G(N,m+1) - G(N,m), \quad \text{where } G(N,m) = -\frac{(N+m)\binom{N}{m}}{\binom{N+m}{m}},$$

and sum both sides of the identity over m from $n+1$ to N. ☺

Finally, we point out that using the integral representation of MZSVs,

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

(compare with (3.27)) valid for any admissible multi-index $\boldsymbol{s}=(s_1,\dots,s_l)$, we can write the right-hand side of (6.4) as follows:

$$2\int\cdots\int_{[0,1]^{s_1+\cdots+s_l}} \frac{\prod_{i=1}^{l-1}(1+t_1\cdots t_{s_1+\cdots+s_i})}{\prod_{i=1}^{l}(1-t_1\cdots t_{s_1+\cdots+s_i})}\mathrm{d}t_1\cdots\mathrm{d}t_{s_1+\cdots+s_l}. \tag{6.10}$$

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

$$2\int\cdots\int_{1>u_1>\cdots>u_{s_1+\cdots+s_l}>0} \prod_{i=1}^{l-1}\Bigg(\prod_{j=s_1+\cdots+s_{i-1}+1}^{s_1+\cdots+s_i-1}\frac{\mathrm{d}u_j}{u_j}\cdot\frac{(1+u_{s_1+\cdots+s_i})\,\mathrm{d}u_{s_1+\cdots+s_i}}{(1-u_{s_1+\cdots+s_i})u_{s_1+\cdots+s_i}}\Bigg)$$
$$\times \prod_{j=s_1+\cdots+s_{l-1}+1}^{s_1+\cdots+s_l-1}\frac{\mathrm{d}u_j}{u_j}\cdot\frac{\mathrm{d}u_{s_1+\cdots+s_l}}{1-u_{s_1+\cdots+s_l}}, \tag{6.11}$$

where the empty sum $s_1+\cdots+s_{i-1}$ for $i=1$ is interpreted as 0. Therefore, any of the two integrals in (6.10), (6.11) may replace the right-hand sides of (6.4) or (6.5).

6.3. Zagier's identity

Zagier's formula shows that the multiple zeta values

$$\xi(m,n) = \zeta(\{2\}^m,3,\{2\}^n) \quad \text{for } n,m\ge 0, \tag{6.12}$$

which are part of Hoffman's basis in Conjecture 6.3, are $\mathbb{Q}$-linear combinations of products $\pi^{2\mu}\zeta(2\nu+1)$ with $\mu+\nu=m+n+1$.

Before giving the formula for the numbers $\xi(m,n)$, we first recall the much easier formula from the family (4.4),

$$\xi(n) = \zeta(\{2\}^n) = \frac{\pi^{2n}}{(2n+1)!} \quad \text{for } n \geq 0, \tag{6.13}$$

for the simplest of the Hoffman basis elements.

Theorem 6.6 (Zagier). *For all integers $m, n \geq 0$, we have*

$$\xi(m,n) = 2\sum_{r=1}^{m+n+1}(-1)^{r-1}\left(\left(1-\frac{1}{2^{2r}}\right)\binom{2r}{2m+1} - \binom{2r}{2n+2}\right) \times \xi(m+n-r+1)\zeta(2r+1), \tag{6.14}$$

where the value of $\xi(m+n-r+1)$ is given by (6.13). 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 6.11 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 $m+n \leq 30$, but remains open.

Exercise 6.6 (open problem). Denote M_k the matrix from (6.14) expressing the vector $\{\xi(m,n) : m+n=k\}$ in terms of the vector $\{\zeta(2m+3)\xi(n) : m+n=k\}$, that is,

$$M_k = \left(2(-1)^\mu\left(\left(1-\frac{1}{2^{2\mu+2}}\right)\binom{2\mu+2}{2m+1} - \binom{2\mu+2}{2k-2m+2}\right)\right)_{0\le m,\mu\le k}. \quad (6.15)$$

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

The strategy to prove Theorem 6.6 is to compare the two generating functions

$$F(x,y) = \sum_{m,n\ge 0} (-1)^{m+n+1}\xi(m,n)x^{2m+1}y^{2n+2} \quad (6.16)$$

and

$$\widehat{F}(x,y) = \sum_{m,n\ge 0} (-1)^{m+n+1}\hat{\xi}(m,n)x^{2m+1}y^{2n+2}, \quad (6.17)$$

where

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

denotes the expression occurring on the right-hand side of (6.14). Of course, if the two expressions were the same, we would be done, but in fact they are *completely different*, one involving a generalised hypergeometric function

$${}_{p+1}F_p\left(\begin{matrix} a_0,\ a_1,\ \ldots,\ a_p \\ b_1,\ \ldots,\ b_p \end{matrix}\,\middle|\, z\right) = \sum_{n=0}^{\infty} \frac{(a_0)_n(a_1)_n\cdots(a_p)_n}{(b_1)_n\cdots(b_p)_n}\,\frac{z^n}{n!} \quad (6.18)$$

(cf. Section 4.1), 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 6.6. 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 6.7. *The generating function $F(x,y)$ can be expressed as the product of a sine function and a hypergeometric function:*

$$F(x,y) = \frac{\sin\pi x}{\pi}\,\frac{\partial}{\partial z}\,{}_3F_2\left(\begin{matrix} y,\ -y,\ z \\ 1+x,\ 1-x \end{matrix}\,\middle|\, 1\right)\Bigg|_{z=0}. \quad (6.19)$$

Proof. The proof is similar to that for (6.13):

$$\begin{aligned}F(x,y) &= \sum_{m,n\geq 0}(-1)^{m+n+1}\zeta(\{2\}^m,3,\{2\}^n)x^{2m+1}y^{2n+2}\\ &= -xy^2\sum_{r=1}^{\infty}\prod_{k=1}^{r-1}\left(1-\frac{y^2}{k^2}\right)\cdot\frac{1}{r^3}\cdot\prod_{l=r+1}^{\infty}\left(1-\frac{x^2}{l^2}\right)\\ &= \frac{\sin\pi x}{\pi}\sum_{r=1}^{\infty}\frac{(-y)_r(y)_r}{(1-x)_r(1+x)_r}\,\frac{1}{r},\end{aligned}$$

and this formula is seen to be equivalent to (6.19). ☺

Lemma 6.8. *The generating function $\widehat{F}(x,y)$ can be expressed as an integral linear combination of* fourteen *functions of the form*

$$\psi\left(1+\frac{u}{2}\right)\frac{\sin\pi v}{2\pi}\quad\text{with } u\in\{\pm x\pm y,\pm 2x\pm 2y,\pm 2x\},\quad v\in\{x,y\}.$$

Proof. From the definition of $\widehat{F}(x,y)$ and (6.13) we find

$$\begin{aligned}\widehat{F}(x,y) &= 2\sum_{m,n\geq 0}(-1)^{m+n}x^{2m+1}y^{2n+2}\sum_{r=1}^{m+n+1}(-1)^r\left((1-2^{-2r})\binom{2r}{2m+1}\right.\\ &\qquad\left.-\binom{2r}{2n+2}\right)\frac{\pi^{2(m+n-r+1)}}{(2(m+n-r+1)+1)!}\zeta(2r+1)\\ &= \frac{2}{\pi}\sum_{s=0}^{\infty}\sum_{r=1}^{\infty}\frac{(-1)^s\pi^{2s+1}}{(2s+1)!}\zeta(2r+1)\left(\sum_{n=0}^{r-1}\binom{2r}{2n+2}x^{2(s+r-n)-1}y^{2n+2}\right.\\ &\qquad\left.-(1-2^{-2r})\sum_{m=0}^{r-1}\binom{2r}{2m+1}x^{2m+1}y^{2(s+r-m)}\right)\\ &= \frac{2\sin\pi x}{\pi}\sum_{r=1}^{\infty}\zeta(2r+1)\sum_{n=0}^{r-1}\binom{2r}{2n+2}x^{2(r-n-1)}y^{2(n+1)}\\ &\qquad-\frac{2\sin\pi y}{\pi}\sum_{r=1}^{\infty}(1-2^{-2r})\zeta(2r+1)\sum_{m=0}^{r-1}\binom{2r}{2m+1}x^{2m+1}y^{2(r-m)-1}\\ &= \frac{\sin\pi x}{\pi}\sum_{r=1}^{\infty}\zeta(2r+1)\big((x+y)^{2r}+(x-y)^{2r}-2x^{2r}\big)\\ &\qquad-\frac{\sin\pi y}{\pi}\sum_{r=1}^{\infty}(1-2^{-2r})\zeta(2r+1)\big((x+y)^{2r}-(x-y)^{2r}\big)\\ &= \frac{\sin\pi x}{\pi}\big(A(x+y)+A(x-y)-2A(x)\big)\\ &\qquad-\frac{\sin\pi y}{\pi}\big(B(x+y)-B(x-y)\big),\end{aligned}$$

where (cf. the final part of Section 4.1)

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

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

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 6.7 (open problem). Show the equality $F(x,y)=\widehat{F}(x,y)$ directly by using the representations in Lemmas 6.7 and 6.8.

The following chain of exercises sketches the remaining ingredients of the proof of Theorem 6.6.

Exercise 6.8. 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 6.9. 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 6.8.

Exercise 6.10. (a) Prove that for all $n\in\mathbb{Z}_{>0}$ and $x\in\mathbb{C}$,

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

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 6.9. *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 6.6. We can now complete the proof of the main equality 6.14 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 6.10 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 6.9 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 6.11. Prove the second statement of the theorem (that is, the invertibility of matrix M_k in (6.15)) by computing the 2-adic valuation of the entries of the matrix.

6.4. 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 6.10. *The double zeta value* $\zeta(m,n)$ (*with* $m \geq 2$ *and* $n \geq 1$) *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*

$$\zeta(m,n) = (-1)^n \sum_{s=0}^{l-1} \left(\binom{k-2s-1}{m-1} + \binom{k-2s-1}{n-1} - \delta_{m,2s} + (-1)^n \delta_{s,0} \right) \times \zeta(2s)\zeta(k-2s). \tag{6.20}$$

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

$$\zeta(r)\zeta(s) = \zeta(r,s) + \zeta(s,r) + \zeta(k), \quad \text{where } r+s=k,\ r,s \geq 2, \tag{6.21}$$

$$\zeta(r)\zeta(s) = \sum_{m=2}^{k-1} \left(\binom{m-1}{r-1} + \binom{m-1}{s-1} \right) \zeta(m,k-m), \quad \text{where } r+s=k,\ r,s \geq 2, \tag{6.22}$$

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 (6.21) and (6.22) 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 (6.21) and (6.22) 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:

$$Q(x,y) = R(x,y) + R(x-y,-y) + R(x-y,x),$$

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

This is equivalent (because of $\zeta(0) = -\frac{1}{2}$) to (6.20). ☺

Either of the double shuffle relations (6.21) and (6.22) 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 (6.20), which simplifies to

$$\zeta(k-2r,2r) \equiv \sum_{s=1}^{l-1}\left(\binom{2l-2s}{2l-2r}+\binom{2l-2s}{2r-1}\right)\zeta(2s)\zeta(k-2s), \quad 1 \le r \le l-1, \tag{6.23}$$

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

THEOREM 6.11. *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 (6.23). 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 6.6 and 6.11 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*

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

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

EXERCISE 6.12 (open problem). For $k = 2l+1 \ge 3$ and the matrix $N = N_k$ defined above, show the following:

(a) $\det N = (-1)^l\, 1\cdot 3\cdot 5\cdots(k-2)$; and
(b) the entries of the inverse matrix N^{-1} are explicitly given by either of the two expressions

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

where B_n denotes the nth Bernoulli number (see Section 1.3).

CHAPTER 7

q-Analogues of multiple zeta values

7.1. q-Zeta values

The classical idea of introducing an additional parameter in an expression or formula we wish to deal with, is quite fruitful in many situations. This may significantly simplify a proof of the corresponding identity or lead to a more general identity which have several other useful specializations of the parameter introduced. We have already witnessed a usefulness of this approach on examples of functional models of generalised polylogarithms in Section 3.2 and of multiple harmonic sums in Section 3.4. These were used for proving the shuffle and stuffle relations of MZVs, respectively. Because the functional versions satisfy only 'half' of relations of MZVs, we can hardly use either of them as a parametric version of the latter numbers.

There is a different way of introducing a parameter. The story usually refers to the parameter q (from 'quantum', whatever it means) and often has a different flavour. The basic idea is simply replacing 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 clear analytical 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 quantities. 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.

EXERCISE 7.1. For integers $m \ge n \ge 0$, define the q-binomial coefficients

$$\begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} m \\ n \end{bmatrix}_q = \frac{[m]!}{[n]!\,[m-n]!}, \quad \text{where } [n]! = [1]\,[2]\cdots[n].$$

Show that all are *polynomials* in q with integer *nonnegative* coeffcients. These are also known by the name Gaussian (binomial) polynomials in the literature.

HINT. Show first the q-analogue

$$\begin{bmatrix} m \\ n \end{bmatrix} = q^n \begin{bmatrix} m-1 \\ n \end{bmatrix} + \begin{bmatrix} m-1 \\ n-1 \end{bmatrix}$$

of Pascal identities, or the q-binomial theorem

$$\prod_{k=1}^{m}(1+q^k z) = \sum_{n=0}^{m} q^{n(n+1)/2} \begin{bmatrix} m \\ n \end{bmatrix} z^n. \qquad ☺$$

Before going into details of q-generalization of multiple zeta values, let us examine the zeta values — MZVs of length 1. For this purpose, we introduce an 'arithmetically motivated' q-model of $\zeta(s)$, namely,

$$\tilde{\zeta}_q(s) = \sum_{n=1}^{\infty} \sigma_{s-1}(n) q^n = \sum_{n=1}^{\infty} \frac{n^{s-1} q^n}{1-q^n}, \quad s = 1, 2, \dots, \tag{7.1}$$

where $\sigma_{s-1}(n) = \sum_{d|n} d^{s-1}$ denotes the sum of powers of the divisors. Here are the first few instances:

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

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

and, in general,

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

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

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

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

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

If $s \ge 2$ is even, then the series $E_s(q) = 1 - 2s\zeta_q(s)/B_s$, where the Bernoulli numbers $B_s \in \mathbb{Q}$ are defined in Section 1.3, are known as the *Eisenstein series*. In particular, they are examples of (quasi-)modular forms whose structural properties are well studied. This circumstance allows one 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 established in Corollary 1.9. 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

$$\delta E_2 = \frac{1}{12}(E_2^2 - E_4), \quad \delta E_4 = \frac{1}{3}(E_2E_4 - E_6), \quad \delta E_6 = \frac{1}{2}(E_2E_6 - E_4^2), \tag{7.2}$$

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

EXERCISE 7.2 (W. N. Bailey). Show that

$$\tilde{\zeta}_q(3) = \sum_{n_1 \ge n_2 \ge 1} \frac{q^{n_1}}{(1-q^{n_1})^2(1-q^{n_2})}.$$

Multiplying the both sides of this identity by $(1-q)^3$ and letting $q \to 1$, again recovers Euler's identity (1.13).

7.2. q-Models of MZVs

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

$$\zeta_q(s_1, s_2, \dots, s_l) = \sum_{n_1 > n_2 > \cdots > n_l \ge 1} \frac{q^{n_1(s_1-1)+n_2(s_2-1)+\cdots+n_l(s_l-1)}}{[n_1]^{s_1}[n_2]^{s_2}\cdots[n_l]^{s_l}}, \tag{7.3}$$

where conditions on the multi-index $\boldsymbol{s} = (s_1, \dots, s_l)$ are exactly the same as for the MZVs (2.1) (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:

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

The q-model (7.3) inherits many relations available for MZVs $\zeta(\boldsymbol{s})$. There is a version of stuffle relations, which is based on the identity from the following exercise.

EXERCISE 7.3. (a) Show that

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

(b) One may also interpret ζ_q as a linear evaluation map on the $\mathbb{Q}$-algebra $\mathfrak{H}^0$ generated by admissible words over the alphabet $\{y_1, y_2, \dots\}$ (as in Section 3.1). Use part (a) to define the harmonic (stuffle) product $*_q$ on the algebra $\mathfrak{H}^1$ in such a way that

$$\zeta_q(w_1 *_q w_2) = \zeta_q(w_1)\zeta_q(w_2) \quad \text{for words } w_1, w_2 \in \mathfrak{H}^0.$$

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 (7.3) 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 7.1. *The generating function Φ_q is given by*

$$\begin{aligned} 1 + (z - xy)\Phi_q(x, y, z) &= \prod_{n=1}^{\infty} \frac{([n]_q - \alpha q^n)([n]_q - \beta q^n)}{([n]_q - x q^n)([n]_q - y q^n)} \\ &= \exp\biggl(\sum_{k=2}^{\infty} \frac{x^k + y^k - \alpha^k - \beta^k}{k} \sum_{j=2}^{k} (q-1)^{k-j} \zeta_q(j)\biggr), \end{aligned} \tag{7.4}$$

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*

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

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 (7.4). ☺

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 (7.4) gives

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

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 7.4. For an indeterminate z, show

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

HINT. This is equivalent to the sum theorem in Corollary 2. ☺

In spite of the above 'naturalness' of the q-MZVs (7.3), 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 (3.9). Lemma 3.5 tells us that

$$\frac{\mathrm{d}}{\mathrm{d}z}\,\mathrm{Li}_{s_1,s_2,\dots,s_l}(z) = \begin{cases} \dfrac{1}{z}\,\mathrm{Li}_{s_1-1,s_2,\dots,s_l}(z) & \text{if } s_1 \geq 2, \\ \dfrac{1}{1-z}\,\mathrm{Li}_{s_2,\dots,s_l}(z) & \text{if } s_1 = 1, \end{cases} \tag{7.5}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}z}\big(f(z)g(z)\big) = \frac{\mathrm{d}}{\mathrm{d}z}f(z)\cdot g(z) + f(z)\cdot\frac{\mathrm{d}}{\mathrm{d}z}g(z). \tag{7.6}$$

The differential equations (7.5) give rise to an integral representation of the polylogarithms (3.9) (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 (7.6) reads as

$$D_q\big(f(z)g(z)\big) = D_qf(z)\cdot g(z) + f(z)\cdot D_qg(z) - (1-q)z\cdot D_qf(z)\cdot D_qg(z), \tag{7.7}$$

where

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

Defining a q-analogue of the multiple polylogarithms (3.9) as

$$\mathrm{Li}_{s_1,\dots,s_l}(z;q) = \sum_{n_1>\dots>n_l\ge 1} \frac{z^{n_1}}{[n_1]^{s_1}\cdots[n_l]^{s_l}}, \tag{7.8}$$

from (7.7) we deduce the following analogue of (7.5):

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

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 7.1 by Okuda and Takeyama. This is a reason to believe that the q-multiple polylogarithms (7.8) are 'motivated' q-analogues of (3.9), and that their values at $z = q$,

$$\begin{aligned} \mathfrak{z}_q(s_1, s_2, \dots, s_l) &= (1-q)^{-|\boldsymbol{s}|}\,\mathrm{Li}_{s_1,s_2,\dots,s_l}(q;q) \\ &= \sum_{n_1>n_2>\dots>n_l\ge 1} \frac{q^{n_1}}{(1-q^{n_1})^{s_1}(1-q^{n_2})^{s_2}\cdots(1-q^{n_l})^{s_l}}, \end{aligned} \tag{7.9}$$

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 (7.3).

Although the rule (7.7) 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 3.9 (a)) in terms of (7.8) 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 (7.8).

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

$$\zeta(r)\zeta(s) = \sum_{i=0}^{r-1}\binom{s-1+i}{i}\zeta(s+i,r-i) + \sum_{i=0}^{s-1}\binom{r-1+i}{i}\zeta(r+i,s-i) \tag{7.10}$$

(which follows from the double shuffle relations (6.21), (6.22)), 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 (7.10) when $r = s = 2$,

$$\zeta(2)^2 = 2\zeta(2,2) + 4\zeta(3,1), \tag{7.11}$$

by means of (7.9). As Bradley shows, even this particular case involves something, which is not expressible by means of q-MZVs (7.3).

We start with the partial-fraction identity

$$\frac{1}{(1-x)(1-y)} = \frac{1}{2}\big(f(x,y) + f(y,x)\big), \quad \text{where } f(x,y) = \frac{1+x}{(1-x)(1-xy)},$$

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 (7.11).

One can also use Ramanujan's system of differential equations (7.2) 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 (7.11). But for a general q-analogue of (7.10) 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 (7.9). Is there a nice form of double shuffle relations in this differential algebra?

7.3. Multiple q-zeta brackets

Apart from standard q-model of the multiple zeta values (7.3) and (7.9) discussed above, there is a somewhat different version introduced recently by Bachmann (partly in collaboration with Kühn):

$$
\begin{aligned}
[s_1,\dots,s_l] &= \frac{1}{(s_1-1)!\cdots(s_l-1)!}\sum_{\substack{n_1>\cdots>n_l>0\\ d_1,\dots,d_l>0}} d_1^{s_1-1}\cdots d_l^{s_l-1}q^{n_1d_1+\cdots+n_ld_l}\\
&= \frac{1}{(s_1-1)!\cdots(s_l-1)!}\\
&\quad\times\sum_{\substack{m_1,\dots,m_l>0\\ d_1,\dots,d_l>0}} d_1^{s_1-1}\cdots d_l^{s_l-1}q^{(m_1+\cdots+m_l)d_1+(m_2+\cdots+m_l)d_2+\cdots+m_ld_l}.
\end{aligned}
\tag{7.12}
$$

The series are generating functions of multiple divisor sums, called (*mono-*) *brackets*, with the $\mathbb{Q}$-algebra spanned by them denoted by $\mathcal{MD}$. These clearly include the single q-zeta values (7.1) from Section 7.1. Note that the q-series (7.12) can be alternatively written

$$
[s_1,\dots,s_l] = \frac{1}{(s_1-1)!\cdots(s_l-1)!}\sum_{n_1>\cdots>n_l>0}\frac{\hat\rho_{s_1}(q^{n_1})\cdots\hat\rho_{s_l}(q^{n_l})}{(1-q^{n_1})^{s_1}\cdots(1-q^{n_l})^{s_l}},
$$

where $\hat\rho_s(x)=x\rho_s(x)$ are (essentially) the polynomials from Section 7.1:

$$
\frac{\hat\rho_s(x)}{(1-x)^s} = \left(x\,\frac{\mathrm{d}}{\mathrm{d}x}\right)^{s-1}\frac{x}{1-x} = \sum_{d=1}^{\infty} d^{s-1}x^d.
$$

Since $\hat\rho_s(1)=\rho_s(1)=(s-1)!$, we have

$$
\lim_{q\to1^-}(1-q)^{s_1+\cdots+s_l}[s_1,\dots,s_l] = \zeta(s_1,\dots,s_l). \tag{7.13}
$$

In addition to (7.12), Bachmann introduced a more general model of the brackets

$$
\begin{aligned}
\begin{bmatrix} s_1,\dots,s_l\\ r_1,\dots,r_l\end{bmatrix} &= \frac{1}{r_1!\,(s_1-1)!\cdots r_l!\,(s_l-1)!}\\
&\quad\times\sum_{\substack{n_1>\cdots>n_l>0\\ d_1,\dots,d_l>0}} n_1^{r_1}d_1^{s_1-1}\cdots n_l^{r_l}d_l^{s_l-1}q^{n_1d_1+\cdots+n_ld_l}\\
&= \frac{1}{r_1!\,(s_1-1)!\cdots r_l!\,(s_l-1)!}\\
&\quad\times\sum_{n_1>\cdots>n_l>0}\frac{n_1^{r_1}\hat\rho_{s_1}(q^{n_1})\cdots n_l^{r_l}\hat\rho_{s_l}(q^{n_l})}{(1-q^{n_1})^{s_1}\cdots(1-q^{n_l})^{s_l}},
\end{aligned}
\tag{7.14}
$$

which he called *bi-brackets*, in order to describe, in a natural way, the double shuffle relations of these q-analogues of MZVs. Note that the stuffle (or harmonic) product for the both models (7.12) and (7.14) in Bachmann's work comes from the standard rearrangement of the multiple sums obtained from the term-by-term multiplication of two series. The other shuffle product is then interpreted for the model (7.14) only, as a dual product to the stuffle one via a *partition duality*. Bachmann further conjectures that the $\mathbb{Q}$-algebra $\mathcal{BD}$ spanned by the bi-brackets (7.14) coincides with the $\mathbb{Q}$-algebra $\mathcal{MD}$.

The goal of this section is to make an algebraic setup for Bachmann's double stuffle relations as well as to demonstrate that those relations indeed reduce to the corresponding stuffle and shuffle relations in the limit as $q \to 1^-$. We also briefly address the reduction of the bi-brackets to the mono-brackets.

The following result allows one to control the asymptotic behaviour of the bi-brackets not only as $q \to 1^-$ but also as q approaches radially a root of unity.

EXERCISE 7.5. As $q = 1 - \varepsilon \to 1^-$,

$$\frac{1}{(s-1)!}\,\frac{\hat\rho_s(q^n)}{(1-q^n)^s} = \frac{1}{n^s\varepsilon^s}\big((1-\varepsilon)F_{s-1}(\varepsilon) + \hat\lambda_s\cdot\varepsilon^s\big) - \hat\lambda_s + O(\varepsilon)$$

where the polynomials $F_k(\varepsilon) \in \mathbb{Q}[\varepsilon]$ of degree $\max\{0, k-1\}$ are generated by

$$\begin{aligned}\sum_{k=0}^{\infty} F_k(\varepsilon)x^k &= \frac{1}{1-(1-e^{-\varepsilon x})/\varepsilon} \\ &= 1 + x + \left(-\frac{1}{2}\varepsilon + 1\right)x^2 + \left(\frac{1}{6}\varepsilon^2 - \varepsilon + 1\right)x^3 \\ &\quad + \left(-\frac{1}{24}\varepsilon^3 + \frac{7}{12}\varepsilon^2 - \frac{3}{2}\varepsilon + 1\right)x^4 \\ &\quad + \left(\frac{1}{120}\varepsilon^4 - \frac{1}{4}\varepsilon^3 + \frac{5}{4}\varepsilon^2 - 2\varepsilon + 1\right)x^5 + \cdots\end{aligned}$$

and

$$\sum_{s=0}^{\infty}\hat\lambda_s x^s = -\frac{xe^x}{1-e^x} = 1 + \frac{1}{2}x + \sum_{k=1}^{\infty}\frac{B_{2k}}{(2k)!}x^{2k}$$

is the (modified) generating function of the Bernoulli numbers.

By moving the constant term $\hat\lambda_s$ to the right-hand side, we get

$$\begin{aligned}\frac{1}{2} + \frac{\hat\rho_1(q^n)}{1-q^n} &= \frac{1}{n}\cdot\left(\varepsilon^{-1} - \frac{1}{2}\right) + O(\varepsilon), \\ \frac{1}{12} + \frac{\hat\rho_2(q^n)}{(1-q^n)^2} &= \frac{1}{n^2}\cdot\left(\varepsilon^{-2} - \varepsilon^{-1} + \frac{1}{12}\right) + O(\varepsilon), \\ \frac{\hat\rho_3(q^n)}{(1-q^n)^3} &= \frac{1}{n^3}\cdot\left(\varepsilon^{-3} - \frac{3}{2}\varepsilon^{-2} + \frac{1}{2}\varepsilon^{-1}\right) + O(\varepsilon), \\ -\frac{1}{720} + \frac{\hat\rho_4(q^n)}{(1-q^n)^4} &= \frac{1}{n^4}\cdot\left(\varepsilon^{-4} - 2\varepsilon^{-3} + \frac{7}{6}\varepsilon^{-2} - \frac{1}{6}\varepsilon^{-1} - \frac{1}{720}\right) + O(\varepsilon),\end{aligned}$$

and so on.

PROPOSITION 7.2. *Assume that $s_1 > r_1+1$ and $s_j \ge r_j+1$ for $j = 2,\dots,l$. Then*

$$\begin{bmatrix} s_1,\dots,s_l \\ r_1,\dots,r_l \end{bmatrix} \sim \frac{\zeta(s_1-r_1, s_2-r_2, \dots, s_l-r_l)}{r_1!\,r_2!\cdots r_l!} \frac{1}{(1-q)^{s_1+s_2+\cdots+s_l}} \quad \text{as } q \to 1^-,$$

where $\zeta(s_1,\dots,s_l)$ denotes the standard MZV.

Another way to tackle the asymptotic behaviour of the (bi-)brackets is based on the Mellin transform

$$\varphi(t) \mapsto \widetilde{\varphi}(s) = \int_0^\infty \varphi(t) t^{s-1} \mathrm{d}t$$

which maps

$$q^{n_1d_1+\cdots+n_ld_l}\Big|_{q=e^{-t}} \mapsto \frac{\Gamma(s)}{(n_1d_1+\cdots+n_ld_l)^s}.$$

Note that the bijective correspondence between the bi-brackets and the zeta functions

$$\frac{\Gamma(s)}{r_1!\,(s_1-1)!\cdots r_l!\,(s_l-1)!} \sum_{\substack{n_1>\cdots>n_l>0 \\ d_1,\dots,d_l>0}} \frac{n_1^{r_1} d_1^{s_1-1} \cdots n_l^{r_l} d_l^{s_l-1}}{(n_1d_1+\cdots+n_ld_l)^s}$$

can be potentially used for determining the linear relations of the former. A simple illustration is the linear independence of the length 1 bi-brackets.

THEOREM 7.3. *The bi-brackets $\left[\begin{smallmatrix} s_1 \\ r_1 \end{smallmatrix}\right]$, where $0 \le r_1 < s_1 \le n$, $s_1 + r_1 \le n$, are linearly independent over $\mathbb{Q}$. Therefore, the dimension $d_n^{\mathcal{BD}}$ of the $\mathbb{Q}$-space spanned by all bi-brackets of weight at most n is bounded from below by $\lfloor (n+1)^2/4 \rfloor \ge n(n+2)/4$.*

PROOF. Indeed, the functions

$$\frac{\Gamma(s)}{r_1!\,(s_1-1)!} \sum_{n_1,d_1>0} \frac{n_1^{r_1} d_1^{s_1-1}}{(n_1d_1)^s} = \Gamma(s) \frac{\zeta(s-s_1+1)\zeta(s-r_1)}{(s_1-1)!\,r_1!},$$
$$\text{where } 0 \le r_1 < s_1 \le n, \; s_1+r_1 \le n,$$

are linearly independent over $\mathbb{Q}$ (because of their disjoint sets of poles at $s = s_1$ and $s = r_1 + 1$, respectively); thus the corresponding bi-brackets $\left[\begin{smallmatrix} s_1 \\ r_1 \end{smallmatrix}\right]$ are $\mathbb{Q}$-linearly independent as well. ☺

A similar (though more involved) analysis can be applied to describe the Mellin transform of the length 2 bi-brackets; note that it is more easily done for another q-model we introduce below.

Consider now the alphabet $Z = \{z_{s,r} : s, r = 1, 2, \dots\}$ on the double-indexed letters $z_{s,r}$ of the pre-defined weight $s + r - 1$. On $\mathbb{Q}Z$ define the product

$$z_{s_1,r_1} \diamond z_{s_2,r_2} = \binom{r_1+r_2-2}{r_1-1}\bigg(z_{s_1+s_2,r_1+r_2-1} + \sum_{j=1}^{s_1}(-1)^{s_2-1}\binom{s_1+s_2-j-1}{s_1-j}\lambda_{s_1+s_2-j}z_{j,r_1+r_2-1} + \sum_{j=1}^{s_2}(-1)^{s_1-1}\binom{s_1+s_2-j-1}{s_2-j}\lambda_{s_1+s_2-j}z_{j,r_1+r_2-1}\bigg), \quad (7.15)$$

where

$$\sum_{s=0}^{\infty}\lambda_s x^s = -\frac{x}{1-e^x} = 1 + \sum_{s=1}^{\infty}\frac{B_s}{s!}x^s$$

is the generating function of Bernoulli numbers. Note that $\hat\lambda_s = \lambda_s$ for $s \ge 2$, while $\hat\lambda_1 = \frac{1}{2} = -\lambda_1$ in the notation of Exercise 7.5.

Exercise 7.6. Show that the the product $\diamond$ is (associative and) commutative.

With the help of (7.15) define the stuffle product on the $\mathbb{Q}$-algebra $\mathbb{Q}\langle Z\rangle$ recursively by $\mathbf{1} \sqcap\!\!\sqcap w = w \sqcap\!\!\sqcap \mathbf{1} = w$ and

$$aw \sqcap\!\!\sqcap bv = a(w \sqcap\!\!\sqcap bv) + b(aw \sqcap\!\!\sqcap v) + (a\diamond b)(w \sqcap\!\!\sqcap v), \quad (7.16)$$

for arbitrary $w, v \in \mathbb{Q}\langle Z\rangle$ and $a, b \in Z$.

Proposition 7.4. *The evaluation map*

$$[\,\cdot\,]\colon z_{s_1,r_1}\cdots z_{s_l,r_l} \mapsto \begin{bmatrix} s_1,\dots,s_l \\ r_1-1,\dots,r_l-1\end{bmatrix} \quad (7.17)$$

extended to $\mathbb{Q}\langle Z\rangle$ *by linearity satisfies* $[w \sqcap\!\!\sqcap v] = [w]\cdot[v]$, *so that it is a homomorphism of the* $\mathbb{Q}$-*algebra* $(\mathbb{Q}\langle Z\rangle, \sqcap\!\!\sqcap)$ *onto* $(\mathcal{BD}, \cdot)$, *the latter hence being a* $\mathbb{Q}$-*algebra as well.*

Proof. The proof is based on the identity

$$\frac{n^{r_1-1}\hat\rho_{s_1}(q^n)}{(s_1-1)!\,(r_1-1)!\,(1-q^n)^{s_1}}\cdot\frac{n^{r_2-1}\hat\rho_{s_2}(q^n)}{(s_2-1)!\,(r_2-1)!\,(1-q^n)^{s_2}} = \binom{r_1+r_2-2}{r_1-1}\frac{n^{r_1+r_2-2}}{(r_1+r_2-2)!}\bigg(\frac{\hat\rho_{s_1+s_2}(q^n)}{(s_1+s_2-1)!\,(1-q^n)^{s_1+s_2}} + \sum_{j=1}^{s_1}(-1)^{s_2-1}\binom{s_1+s_2-j-1}{s_1-j}\lambda_{s_1+s_2-j}\frac{\hat\rho_j(q^n)}{(j-1)!\,(1-q^n)^j} + \sum_{j=1}^{s_2}(-1)^{s_1-1}\binom{s_1+s_2-j-1}{s_2-j}\lambda_{s_1+s_2-j}\frac{\hat\rho_j(q^n)}{(j-1)!\,(1-q^n)^j}\bigg). \quad ☺$$

Modulo the highest weight, the commutative product (7.15) on Z assumes the form

$$z_{s_1,r_1} \diamond z_{s_2,r_2} \equiv \binom{r_1+r_2-2}{r_1-1} z_{s_1+s_2,r_1+r_2-1},$$

so that the stuffle product (7.16) reads

$$\begin{aligned} z_{s_1,r_1} w \sqcap\!\!\sqcap\, z_{s_2,r_2} v \equiv\ & z_{s_1,r_1}(w \sqcap\!\!\sqcap\, z_{s_2,r_2} v) + z_{s_2,r_2}(z_{s_1,r_1} w \sqcap\!\!\sqcap\, v) \\ & + \binom{r_1+r_2-2}{r_1-1} z_{s_1+s_2,r_1+r_2-1}(w \sqcap\!\!\sqcap\, v) \end{aligned} \tag{7.18}$$

for arbitrary $w, v \in \mathbb{Q}\langle Z\rangle$ and $z_{s_1,r_1}, z_{s_2,r_2} \in Z$. If we set $z_s = z_{s,1}$ and further restrict the product to the subalgebra $\mathbb{Q}\langle Z'\rangle$, where $Z' = \{z_s : s = 1, 2, \dots\}$, then Proposition 7.2 results in the following statement.

Theorem 7.5. *For admissible words $w = z_{s_1} \cdots z_{s_l}$ and $v = z_{s'_1} \cdots z_{s'_m}$ of weight $|w| = s_1 + \cdots + s_l$ and $|v| = s'_1 + \cdots + s'_m$, respectively,*

$$[w \sqcap\!\!\sqcap\, v] \sim (1-q)^{-|w|-|v|} \zeta(w * v) \quad \text{as } q \to 1^-,$$

where $*$ denotes the standard stuffle (harmonic) product of MZVs on $\mathbb{Q}\langle Z'\rangle$.

Since $[w] \sim (1-q)^{-|w|}\zeta(w)$, $[v] \sim (1-q)^{-|v|}\zeta(v)$ as $q \to 1^-$ and $[w \sqcap\!\!\sqcap\, v] = [w]\cdot[v]$, Theorem 7.5 asserts that the stuffle product (7.16) of the algebra $\mathcal{MD}$ reduces to the stuffle product of the algebra of MZVs in the limit as $q \to 1^-$.

To analyse the duality of bi-brackets, we introduce the following alternative extension of the mono-brackets (7.12), called *multiple q-zeta brackets*:

$$\begin{aligned} \mathfrak{z}\begin{bmatrix} s_1, \dots, s_l \\ r_1, \dots, r_l \end{bmatrix} &= \mathfrak{z}_q\begin{bmatrix} s_1, \dots, s_l \\ r_1, \dots, r_l \end{bmatrix} \\ &= c \sum_{\substack{m_1,\dots,m_l>0 \\ d_1,\dots,d_l>0}} m_1^{r_1-1} d_1^{s_1-1} \cdots m_l^{r_l-1} d_l^{s_l-1} q^{(m_1+\dots+m_l)d_1+(m_2+\dots+m_l)d_2+\dots+m_l d_l} \\ &= c \sum_{m_1,\dots,m_l>0} \frac{m_1^{r_1-1}\hat\rho_{s_1}(q^{m_1+\dots+m_l}) m_2^{r_2-1}\hat\rho_{s_2}(q^{m_2+\dots+m_l}) \cdots m_l^{r_l-1}\hat\rho_{s_l}(q^{m_l})}{(1-q^{m_1+\dots+m_l})^{s_1}(1-q^{m_2+\dots+m_l})^{s_2}\cdots(1-q^{m_l})^{s_l}} \end{aligned} \tag{7.19}$$

where

$$c = \frac{1}{(r_1-1)!\,(s_1-1)!\cdots(r_l-1)!\,(s_l-1)!}.$$

Then

$$\begin{bmatrix} s_1 \\ r_1 - 1 \end{bmatrix} = \mathfrak{z}\begin{bmatrix} s_1 \\ r_1 \end{bmatrix} \qquad \text{and} \qquad [s_1, \dots, s_l] = \begin{bmatrix} s_1, \dots, s_l \\ 0, \dots, 0 \end{bmatrix} = \mathfrak{z}\begin{bmatrix} s_1, \dots, s_l \\ 1, \dots, 1 \end{bmatrix}.$$

By applying iteratively the binomial theorem in the forms

$$\frac{(m+n)^{r_1-1}}{(r_1-1)!}\,\frac{n^{r_2-1}}{(r_2-1)!} = \sum_{j=1}^{r_1+r_2-1} \binom{j-1}{r_2-1} \frac{m^{r_1+r_2-j-1}}{(r_1+r_2-j-1)!}\,\frac{n^{j-1}}{(j-1)!}$$

and

$$\frac{(n-m)^{r-1}}{(r-1)!}=\sum_{i=1}^{r}(-1)^{r+i}\frac{n^{i-1}}{(i-1)!}\,\frac{m^{r-i}}{(r-i)!}$$

we see that the $\mathbb{Q}$-algebras spanned by either (7.14) or (7.19) coincide. More precisely, the following formulae link the two versions of brackets.

Exercise 7.7. Show that

$$\begin{aligned}
&\begin{bmatrix} s_1, & s_2, & \dots, & s_l \\ r_1-1, & r_2-1, & \dots, & r_l-1 \end{bmatrix}\\
&=\sum_{j_2=1}^{r_1+r_2-1}\binom{j_2-1}{r_2-1}\sum_{j_3=1}^{j_2+r_3-1}\binom{j_3-1}{r_3-1}\cdots\sum_{j_l=1}^{j_{l-1}+r_l-1}\binom{j_l-1}{r_l-1}\\
&\qquad\times\mathfrak{Z}\begin{bmatrix} s_1, & s_2, & \dots, & s_{l-1}, & s_l \\ r_1+r_2-j_2, & j_2+r_3-j_3, & \dots, & j_{l-1}+r_l-j_l, & j_l \end{bmatrix}
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{Z}\begin{bmatrix} s_1, \dots, s_l \\ r_1, \dots, r_l \end{bmatrix} &= \sum_{i_1=1}^{r_1}\sum_{i_2=1}^{r_2}\cdots\sum_{i_{l-1}=1}^{r_{l-1}}(-1)^{r_1+\dots+r_{l-1}-i_1-\dots-i_{l-1}}\\
&\times\binom{r_1-i_1+i_2-1}{r_1-i_1}\cdots\binom{r_{l-2}-i_{l-2}+i_{l-1}-1}{r_{l-2}-i_{l-2}}\binom{r_{l-1}-i_{l-1}+r_l-1}{r_{l-1}-i_{l-1}}\\
&\times\begin{bmatrix} s_1, & s_2, & \dots, & s_{l-1}, & s_l \\ i_1-1, & r_1-i_1+i_2-1, & \dots, & r_{l-2}-i_{l-2}+i_{l-1}-1, & r_{l-1}-i_{l-1}+r_l-1 \end{bmatrix}.
\end{aligned}$$

Exercise 7.7 allows us to construct an isomorphism φ of the two $\mathbb{Q}$-algebras $\mathbb{Q}\langle Z\rangle$ with two evaluation maps $[\,\cdot\,]$ and $\mathfrak{Z}[\,\cdot\,]$,

$$\mathfrak{Z}[z_{s_1,r_1}\cdots z_{s_l,r_l}]=\mathfrak{Z}\begin{bmatrix} s_1, \dots, s_l \\ r_1, \dots, r_l \end{bmatrix},$$

such that

$$[w]=\mathfrak{Z}[\varphi w] \qquad \text{and} \qquad \mathfrak{Z}[w]=[\varphi^{-1}w].$$

Note however that the isomorphism breaks the simplicity of defining the stuffle product $\sqcap\!\sqcap$.

Another algebraic setup can be used for the $\mathbb{Q}$-algebra $\mathbb{Q}\langle Z\rangle$ with evaluation $\mathfrak{Z}$. We can recast it as the familiar $\mathbb{Q}$-subalgebra $\mathfrak{H}^0=\mathbb{Q}\mathbf{1}\oplus x_0\mathfrak{H}x_1$ of the $\mathbb{Q}$-algebra $\mathfrak{H}=\mathbb{Q}\langle x_0,x_1\rangle$ by setting $\mathfrak{Z}[\mathbf{1}]=1$ and

$$\mathfrak{Z}[x_0^{s_1}x_1^{r_1}\cdots x_0^{s_l}x_1^{r_l}]=\mathfrak{Z}\begin{bmatrix} s_1, \dots, s_l \\ r_1, \dots, r_l \end{bmatrix}.$$

The *length* (or *depth*) is defined as the number of appearances of the subword x_0x_1, while the *weight* is the number of letters x_0 or x_1 minus the length.

Proposition 7.6 (Duality). *We have*

$$\mathfrak{Z}\begin{bmatrix} s_1, s_2, \dots, s_l \\ r_1, r_2, \dots, r_l \end{bmatrix}=\mathfrak{Z}\begin{bmatrix} r_l, r_{l-1}, \dots, r_1 \\ s_l, s_{l-1}, \dots, s_1 \end{bmatrix}.$$

Proof. This follows from the rearrangement of the summation indices:

$$\sum_{i=1}^{l} d_i \sum_{j=i}^{l} m_j = \sum_{i=1}^{l} d'_i \sum_{j=i}^{l} m'_j$$

where $d'_i = m_{l+1-i}$ and $m'_j = d_{l+1-j}$. ☺

If τ denotes the familiar anti-automorphism of the algebra $\mathfrak{H}$ (and of its subalgebra $\mathfrak{H}^0$), interchanging x_0 and x_1, then, clearly, τ is an involution preserving both the weight and length. The duality can be then stated as

$$\mathfrak{Z}[\tau w] = \mathfrak{Z}[w] \quad \text{for any } w \in \mathfrak{H}^0. \tag{7.20}$$

We also extend τ to $\mathbb{Q}\langle Z\rangle$ by linearity.

The duality in Proposition 7.6 transfered to the bi-bracket setting (7.14), namely $\varphi^{-1}\tau\varphi$, is exactly the partition duality given by Bachmann.

We can now introduce the product which is dual to the stuffle one. Namely, it is the duality composed with the stuffle product and, again, with the duality:

$$w \,\overline{\sqcap}\, v = \varphi^{-1}\tau\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v) \quad \text{for } w, v \in \mathbb{Q}\langle Z\rangle. \tag{7.21}$$

It follows then from Propositions 7.4 and 7.6 that

Proposition 7.7. *The evaluation map* (7.17) *on* $\mathbb{Q}\langle Z\rangle$ *satisfies* $[w \,\overline{\sqcap}\, v] = [w]\cdot[v]$*, so that it is also a homomorphism of the* $\mathbb{Q}$*-algebra* $(\mathbb{Q}\langle Z\rangle, \overline{\sqcap})$ *onto* $(\mathcal{BD}, \cdot)$*.*

Proof. We have

$$\begin{aligned}
[w \,\overline{\sqcap}\, v] &= [\varphi^{-1}\tau\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v)] \\
&= \mathfrak{Z}[\tau\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v)] = \mathfrak{Z}[\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v)] \\
&= [\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v] = [\varphi^{-1}\tau\varphi w]\cdot[\varphi^{-1}\tau\varphi v] \\
&= \mathfrak{Z}[\tau\varphi w]\cdot\mathfrak{Z}[\tau\varphi v] = \mathfrak{Z}[\varphi w]\cdot\mathfrak{Z}[\varphi v] = [w]\cdot[v].
\end{aligned}$$

☺

Note that (7.18) is also equivalent to the expansion from the right (this is established in Exercise 3.12):

$$\begin{aligned}
wz_{s_1,r_1} \sqcap vz_{s_2,r_2} \equiv (w \sqcap vz_{s_2,r_2})z_{s_1,r_1} + (wz_{s_1,r_1} \sqcap v)z_{s_2,r_2} \\
+ \binom{r_1+r_2-2}{r_1-1}(w \sqcap v)z_{s_1+s_2,r_1+r_2-1}.
\end{aligned} \tag{7.22}$$

The next statement addresses the structure of the dual stuffle product (7.21) for the words over the sub-alphabet $Z' = \{z_s = z_{s,1} : s = 1, 2, \dots\} \subset Z$. Note that the words from $\mathbb{Q}\langle Z'\rangle$ can be also presented as the words from $\mathbb{Q}\langle x_0, x_0x_1\rangle$ necessarily ending with x_0x_1.

Proposition 7.8. *Modulo the highest weight and length,*

$$aw \,\overline{\sqcap}\, bv \equiv a(w \,\overline{\sqcap}\, bv) + b(aw \,\overline{\sqcap}\, v) \tag{7.23}$$

for arbitrary words $w, v \in \mathbb{Q}\mathbf{1} \oplus \mathbb{Q}\langle x_0, x_0x_1\rangle x_0x_1$ *and* $a, b \in \{x_0, x_0x_1\}$*.*

Proof. First note that restricting (7.22) further modulo the highest length implies

$$wz_{s_1,r_1} \mathbin{\overline{\sqcap}} vz_{s_2,r_2} \equiv (w \mathbin{\overline{\sqcap}} vz_{s_2,r_2})z_{s_1,r_1} + (wz_{s_1,r_1} \mathbin{\overline{\sqcap}} v)z_{s_2,r_2},$$

and that we also have

$$\begin{aligned} wz_{s_1,r_1+1} \mathbin{\sqcap} vz_{s_2,r_2} &\equiv (wz_{s_1,r_1} \mathbin{\sqcap} vz_{s_2,r_2})x_1 + (wz_{s_1,r_1+1} \mathbin{\sqcap} v)z_{s_2,r_2}, \\ wz_{s_1,r_1+1} \mathbin{\sqcap} vz_{s_2,r_2+1} &\equiv (wz_{s_1,r_1} \mathbin{\sqcap} vz_{s_2,r_2+1})x_1 + (wz_{s_1,r_1+1} \mathbin{\sqcap} vz_{s_2,r_2})x_1. \end{aligned}$$

The relations already show that

$$wa' \mathbin{\sqcap} vb' \equiv (w \mathbin{\sqcap} vb')a' + (wa' \mathbin{\sqcap} v)b' \tag{7.24}$$

for arbitrary words $w, v \in \mathbb{Q} + \mathbb{Q}\langle Z\rangle$ and $a', b' \in Z \cup \{x_1\}$, where

$$z_{s_1,r_1} \cdots z_{s_{l-1},r_{l-1}} z_{s_l,r_l} x_1 = z_{s_1,r_1} \cdots z_{s_{l-1},r_{l-1}} z_{s_l,r_l+1}.$$

Secondly note that the isomorphism φ (based on Exercise 7.7) acts trivially on the words from $\mathbb{Q}\langle Z'\rangle$. Therefore, applying $\tau\varphi$ to the both sides of (7.21) and extracting the homogeneous part of the result corresponding to the highest weight and length we arrive at

$$\tau(w \mathbin{\overline{\sqcap}} v) \equiv \tau w \mathbin{\sqcap} \tau v \quad \text{for all } w, v \in \mathbb{Q}\langle Z'\rangle.$$

Denoting

$$\overline{a} = \tau a = \begin{cases} x_1 & \text{if } a = x_0, \\ x_0 x_1 & \text{if } a = x_0 x_1, \end{cases}$$

and using (7.24) we find out that

$$\begin{aligned} \tau(aw \mathbin{\overline{\sqcap}} bv) &\equiv \tau(aw) \mathbin{\sqcap} \tau(bv) \equiv (\tau w)\overline{a} \mathbin{\sqcap} (\tau v)\overline{b} \\ &\equiv (\tau w \mathbin{\sqcap} (\tau v)\overline{b})\overline{a} + ((\tau w)\overline{a} \mathbin{\sqcap} \tau v)\overline{b} \\ &\equiv (\tau w \mathbin{\sqcap} \tau(bv))\overline{a} + (\tau(aw) \mathbin{\sqcap} \tau v)\overline{b} \equiv (\tau(w \mathbin{\overline{\sqcap}} bv))\overline{a} + (\tau(aw \mathbin{\overline{\sqcap}} v))\overline{b} \\ &\equiv \tau(a(w \mathbin{\overline{\sqcap}} bv) + b(aw \mathbin{\overline{\sqcap}} v)), \end{aligned}$$

which implies the desired result. ☺

Theorem 7.9. *For admissible words $w = z_{s_1} \cdots z_{s_l}$ and $v = z_{s'_1} \cdots z_{s'_m}$ of weight $|w| = s_1 + \cdots + s_l$ and $|v| = s'_1 + \cdots + s'_m$, respectively,*

$$[w \mathbin{\overline{\sqcap}} v] \sim (1-q)^{-|w|-|v|}\zeta(w \sqcup\!\sqcup v) \qquad \text{as} \quad q \to 1^-,$$

where $\sqcup\!\sqcup$ denotes the standard shuffle product of MZVs on $\mathbb{Q}\langle Z'\rangle$.

Proof. Because both φ and τ respect the weight, Proposition 7.8 shows that the only terms that can potentially interfere with the asymptotic behaviour as $q \to 1^-$ correspond to the same weight but lower length. However, according to (7.21) and (7.22), the 'shorter' terms do not belong to $\mathbb{Q}\langle Z'\rangle$, that is, they are linear combinations of the monomials $z_{q_1,r_1} \cdots z_{q_n,r_n}$ with

$r_1 + \cdots + r_n = l + m > n$, hence $r_j \geq 2$ for at least one j. The latter circumstance and Proposition 7.2 then imply

$$\lim_{q \to 1^-} (1-q)^{|w|+|v|} [z_{q_1,r_1} \cdots z_{q_n,r_n}] = 0. \qquad ☺$$

Theorem 7.9 asserts that the dual stuffle product (7.21) restricted from $\mathcal{BD}$ to the subalgebra $\mathcal{MD}$ reduces to the shuffle product of the algebra of MZVs in the limit as $q \to 1^-$. More is true: using (7.18) and Proposition 7.8 we obtain

Theorem 7.10. *For two words $w = z_{s_1} \cdots z_{s_l}$ and $v = z_{s'_1} \cdots z_{s'_m}$, not necessarily admissible,*

$$[w \sqcap v - w \,\overline{\sqcap}\, v] \sim (1-q)^{-|w|-|v|} \zeta(w * v - w \sqcup\!\sqcup\, v) \qquad \text{as} \quad q \to 1^-,$$

whenever the MZV on the right-hand side makes sense.

In other words, the q-zeta model of bi-brackets provides us with a (far reaching) regularisation of the MZVs: the former includes the extended double shuffle relations as the limiting $q \to 1^-$ case.

Conjecture 7.11 (Bachmann). *The resulting double stuffle (that is, stuffle and dual stuffle) relations exhaust all the relations between the bi-brackets. Equivalently (and simpler), the stuffle relations and the duality exhaust all the relations between the bi-brackets.*

We would like to point out that the duality τ we introduced in this section is similar to the duality of MZVs from Section 3.3. However the two dualities are not related: the limiting $q \to 1^-$ process squeezes the appearances of x_0 preceding x_1 in the words $x_0^{s_1} x_1 x_0^{s_2} x_1 \cdots x_0^{s_l} x_1$, so that they become $x_0^{s_1-1} x_1 x_0^{s_2-1} x_1 \cdots x_0^{s_l-1} x_1$. Furthermore, the duality of MZVs respects the shuffle product: the dual shuffle product coincides with the shuffle product itself. On the other hand, the dual stuffle product of MZVs is very different from the stuffle (and shuffle) products. It may be an interesting problem to understand the double stuffle relations of the algebra of MZVs.

Finally, we present some observations towards another conjecture of Bachmann about the coincidence of the $\mathbb{Q}$-algebras of bi- and mono-brackets.

Conjecture 7.12 (Bachmann). *$\mathcal{MD} = \mathcal{BD}$.*

Based on the representation of the elements from $\mathcal{BD}$ as the polynomials from $\mathbb{Q}\langle x_0, x_1 \rangle$ (see also the above comment about duality τ), we can loosely interpret this conjecture for the algebra of MZVs as follows: all MZVs lie in the $\mathbb{Q}$-span of

$$\zeta(s_1, s_2, \ldots, s_l) = \zeta(x_0^{s_1-1} x_1 x_0^{s_2-1} x_1 \cdots x_0^{s_l-1} x_1)$$

with all s_j to be at least 2 (so that there is no appearance of x_1^r with $r \geq 2$). The latter statement is already known to be true: Brown proves that one can span the $\mathbb{Q}$-algebra of MZVs by the set with all $s_j \in \{2, 3\}$.

In what follows we analyse the relations for the model (7.19), because it makes simpler keeping track of the duality relation. We point out from the very beginning that the linear relations given below are all experimentally found (with the check of 500 terms in the corresponding q-expansions) but we believe that it is possible to establish them rigorously using the double stuffle relations given above.

The first presence of the q-zeta brackets that are not reduced to ones from $\mathcal{MD}$ by the duality relation happens in weight 3. It is $\mathfrak{z}\big[\begin{smallmatrix}2\\2\end{smallmatrix}\big]$ and we find out that

$$\mathfrak{z}\big[\begin{smallmatrix}2\\2\end{smallmatrix}\big] = \tfrac{1}{2}\mathfrak{z}\big[\begin{smallmatrix}2\\1\end{smallmatrix}\big] + \mathfrak{z}\big[\begin{smallmatrix}3\\1\end{smallmatrix}\big] - \mathfrak{z}\big[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\big].$$

There are 34 totally q-zeta brackets of weight up to 4,

$$\begin{gathered}
\mathfrak{z}[\ ]^*,\ \mathfrak{z}\big[\begin{smallmatrix}1\\1\end{smallmatrix}\big]^*,\ \mathfrak{z}\big[\begin{smallmatrix}2\\1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1\\2\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}2\\2\end{smallmatrix}\big]^*,\ \mathfrak{z}\big[\begin{smallmatrix}3\\1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1\\3\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}3\\2\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}2\\3\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}4\\1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1\\4\end{smallmatrix}\big],\\
\mathfrak{z}\big[\begin{smallmatrix}1,1\\1,1\end{smallmatrix}\big]^*,\ \mathfrak{z}\big[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1\\1,2\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}1,2\\1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1\\2,1\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}2,1\\2,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,2\\1,2\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}2,1\\1,2\end{smallmatrix}\big]^*,\ \mathfrak{z}\big[\begin{smallmatrix}1,2\\2,1\end{smallmatrix}\big]^*,\\
\mathfrak{z}\big[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1\\2,2\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}3,1\\1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1\\1,3\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}1,3\\1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1\\3,1\end{smallmatrix}\big],\\
\mathfrak{z}\big[\begin{smallmatrix}1,1,1\\1,1,1\end{smallmatrix}\big]^*,\ \mathfrak{z}\big[\begin{smallmatrix}2,1,1\\1,1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1,1\\1,1,2\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}1,2,1\\1,1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1,1\\1,2,1\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}1,1,2\\1,1,1\end{smallmatrix}\big] = \mathfrak{z}\big[\begin{smallmatrix}1,1,1\\2,1,1\end{smallmatrix}\big],\ \mathfrak{z}\big[\begin{smallmatrix}1,1,1,1\\1,1,1,1\end{smallmatrix}\big]^*,
\end{gathered}$$

where the asterisk marks the self-dual ones. Only 21 of those listed are not dual-equivalent, and only five of the latter are not reduced to the q-zeta brackets from $\mathcal{MD}$; besides the already mentioned $\mathfrak{z}\big[\begin{smallmatrix}2\\2\end{smallmatrix}\big]$ these are $\mathfrak{z}\big[\begin{smallmatrix}3\\2\end{smallmatrix}\big]$, $\mathfrak{z}\big[\begin{smallmatrix}2,1\\2,1\end{smallmatrix}\big]$, $\mathfrak{z}\big[\begin{smallmatrix}2,1\\1,2\end{smallmatrix}\big]$ and $\mathfrak{z}\big[\begin{smallmatrix}1,2\\2,1\end{smallmatrix}\big]$. We find out that

$$\begin{aligned}
\mathfrak{z}\big[\begin{smallmatrix}3\\2\end{smallmatrix}\big] &= \tfrac{1}{4}\mathfrak{z}\big[\begin{smallmatrix}2\\1\end{smallmatrix}\big] + \tfrac{3}{2}\mathfrak{z}\big[\begin{smallmatrix}4\\1\end{smallmatrix}\big] - 2\mathfrak{z}\big[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\big],\\
\mathfrak{z}\big[\begin{smallmatrix}2,1\\2,1\end{smallmatrix}\big] &= \mathfrak{z}\big[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\big] + \tfrac{1}{2}\mathfrak{z}\big[\begin{smallmatrix}1,2\\1,1\end{smallmatrix}\big] - \mathfrak{z}\big[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\big] + \mathfrak{z}\big[\begin{smallmatrix}1,3\\1,1\end{smallmatrix}\big] - \mathfrak{z}\big[\begin{smallmatrix}2,1,1\\1,1,1\end{smallmatrix}\big] - \mathfrak{z}\big[\begin{smallmatrix}1,2,1\\1,1,1\end{smallmatrix}\big],\\
\mathfrak{z}\big[\begin{smallmatrix}2,1\\1,2\end{smallmatrix}\big] &= -\tfrac{1}{2}\mathfrak{z}\big[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\big] - \tfrac{1}{2}\mathfrak{z}\big[\begin{smallmatrix}1,2\\1,1\end{smallmatrix}\big] + 2\mathfrak{z}\big[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\big] + \mathfrak{z}\big[\begin{smallmatrix}3,1\\1,1\end{smallmatrix}\big] - \mathfrak{z}\big[\begin{smallmatrix}1,3\\1,1\end{smallmatrix}\big] + \mathfrak{z}\big[\begin{smallmatrix}1,2,1\\1,1,1\end{smallmatrix}\big],\\
\mathfrak{z}\big[\begin{smallmatrix}1,2\\2,1\end{smallmatrix}\big] &= -\mathfrak{z}\big[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\big] + 2\mathfrak{z}\big[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\big] + \mathfrak{z}\big[\begin{smallmatrix}2,1,1\\1,1,1\end{smallmatrix}\big],
\end{aligned}$$

and there is one more relation in this weight between the q-zeta brackets from $\mathcal{MD}$:

$$\tfrac{1}{3}\mathfrak{z}\big[\begin{smallmatrix}2\\1\end{smallmatrix}\big] - \mathfrak{z}\big[\begin{smallmatrix}3\\1\end{smallmatrix}\big] + \mathfrak{z}\big[\begin{smallmatrix}4\\1\end{smallmatrix}\big] - 2\mathfrak{z}\big[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\big] + 2\mathfrak{z}\big[\begin{smallmatrix}3,1\\1,1\end{smallmatrix}\big] = 0.$$

The computation implies that the dimension $d_4^{\mathcal{BD}}$ of the $\mathbb{Q}$-space spanned by all multiple q-zeta brackets of weight not more than 4 is equal to the dimension $d_4^{\mathcal{MD}}$ of the $\mathbb{Q}$-space spanned by all such brackets from $\mathcal{MD}$ and that both are equal to 15. A similar analysis demonstrates that

$$d_5^{\mathcal{BD}} = d_5^{\mathcal{MD}} = 28 \quad \text{and} \quad d_6^{\mathcal{BD}} = d_6^{\mathcal{MD}} = 51,$$

and it seems less realistic to compute and verify that $d_n^{\mathcal{BD}} = d_n^{\mathcal{MD}}$ for $n \ge 7$ though Conjecture 7.12 supports

$$\sum_{n=0}^{\infty} d_n^{\mathcal{MD}} x^n \overset{?}{=} \frac{1 - x^2 + x^4}{(1-x)^2(1-2x^2-2x^3)}.$$

We can compare this with the count $c_n^{\mathcal{MD}}$ and $c_n^{\mathcal{BD}}$ of total number of mono- and bi-brackets of weight $\le n$, respectively:

$$\sum_{n=0}^{\infty} c_n^{\mathcal{MD}} x^n = \frac{1}{1-2x} \quad \text{and} \quad \sum_{n=0}^{\infty} c_n^{\mathcal{BD}} x^n = \frac{1-x}{1-3x+x^2} = \sum_{n=0}^{\infty} F_{2n} x^n,$$

where F_n denotes the Fibonacci sequence.

In addition, we would like to point out one more expectation for the algebra of (both mono- and bi-) brackets, which is not shared by other q-models of MZVs: all linear (hence algebraic) relations between them seem over $\mathbb{C}(q)$ seem to be always liftable to relations over $\mathbb{Q}$.

Exercise 7.8 (Open problem). Show that a collection of (bi-)brackets is linearly dependent over $\mathbb{C}(q)$ if and only if it is linearly dependent over $\mathbb{Q}$.