Multiple zeta values

Tasting notes

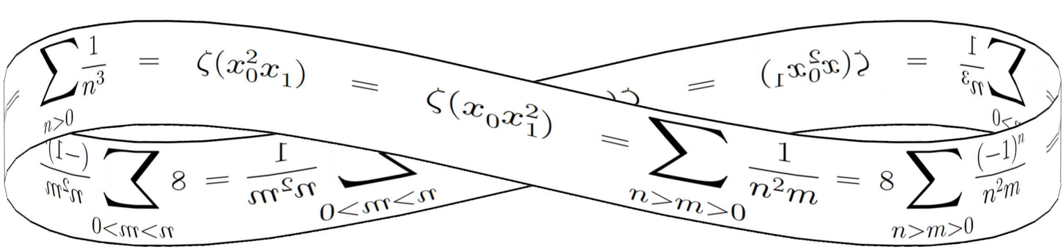
Release: 14 April 2025

Wadim Zudilin

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

Email address: w.zudilin@math.ru.nl

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



To the memory of my father Valentin

and of my grandfathers Veniamin and Stanislav

I wish I knew you better

Contents

Preface

After doing some odd mathematics choices during my university education I became a dedicated transcendental number theorist and defended my PhD thesis mysteriously entitled "On the estimates of the measure of linear independence for values of certain analytical functions" in 1995. It was a good piece of work, with potentials to develop further, but I was looking for real challenges to taste. Apéry's proof of the irrationality of $\zeta(3)$, its numerous interpretations and sequels in the literature of the 1980s and 90s lacking any progress on the irrationality of the other odds — $\zeta(5), \zeta(7), \dots$ — sounded to me like a great goal to pursue. I started to systematically dig in publications on the topic and occasionally crossed with some multiple sums generalising the values of Riemann's zeta function

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

at positive integers $s > 1$. But those multiple findings were all a nice but side story for me until I got to a lecture of Michel Waldschmidt in April 2000 on *Multiple zeta values* at the Oberwolfach meeting on *Diophantische Approximationen*. There was so much structure in those numerical quantities that I immediately fell in love with these multiple creatures and saw potentials in exploring the mechanisms behind them in irrationality and transcendence applications of ordinary zeta values. I have never managed to do this directly, and my other inspirations — the talk of Carlo Viola on his joint work [**37**] with Georges Rhin at the same Oberwolfach meeting and the later paper [**38**] of Tanguy Rivoal which appeared on the arXiv in August 2000 — offered to me more productive options. But I started a careful study of multiple zeta values and published my notes [**52**] as a review of the subject in 2003. Those who followed [**52**] can still recognise their influence on the text that follows, though the latter features many reshapings, improvements and additions through my personal involvement in this remarkable topic.

The multiple zeta values keep developing and — quite remarkably — give life to numerous generalisations and side extensions that now have stories of their own. I was truly fortunate to be part of the 17th Symposium of the Mathematical Society of Japan in February 2025 which incorporated the School on *Multiple Zeta Values and Beyond* held at Kyushu University and the Workshop on *Modular Forms and Multiple Zeta Values* in honour of Masanobu

Kaneko at Kindai University. The events gathered — quite impressively — about 200 participants in total, mainly from Japan where this mathematics topic can be called national without doubt. During the symposium I learned from Michel Waldschmidt that his personal involvement in multiple zeta values had initiated by the late Pierre Cartier, with whom he ran a successful seminar on zeta values at the Institute of Henri Poincaré from 2000 for several years. Cartier made a lasting impact on the subject through reporting at different geographic locations and educating many bright minds; I still have a good memory of his talk at the 2001 conference in Caen on the occasion of Rivoal's PhD defence. One educational fragment of his talk is Exercise 3.14 in this text.

An idea behind these notes is to provide a comprehensive but reasonably elementary introduction to multiple generalisations of Riemann's zeta function and to analytic, algebraic, arithmetic, combinatorial methods used in their study. Great sources for substantial knowledge on the topic of multiple zeta values are the two monographs [**9**] and [**51**] — in no way we compete with these advanced treatments when it comes to a systematic coverage of the themes. Further, the dedicated website [**17**] developed by Michael Hoffman includes a comprehensive list of references on multiple zeta values. For those who are already familiar with some aspects of the subject, we use the convention $n_1 > \cdots > n_l > 0$ for multiple zeta value summation (same as in [**51**] and [**9**]).

There are several personal motivations to write up these notes. First, parts of the text below grew from the material for my own lectures for national master courses in Australia (2011, taught jointly with the late Jonathan M. Borwein) and in the Netherlands (2019). Second, I have been always trying to follow methodological advances and collect the simplest proofs available — though they still have to be nice and reader-friendly. Many earlier proofs which were mysteriously tricky, are now quite elementary and can be classified as *proofs from the Book*. Third, before these notes reach a publisher I can keep them as a dynamical set of lecture notes rather than a stable book. Further, I wish to share my personal taste for multiple zeta values via particularly tasteful bites — some results can be seen as somewhat isolated but their aesthetics is truly appealing for the inclusion. I notice that the exposition is dry at quite some places and references to the original sources could be greatly improved — I plan to work on these issues in due course. I definitely welcome any constructive feedback from those who read the notes.

There are so many people who influenced me, implicitly or not, to acknowledge. As explained above my initial crash at the topic came from a lecture of Michel Waldschmidt, who further supported my education on multiple zeta values through sharing his knowledge and available materials. The papers by Jonathan Borwein, David Bradley, David Broadhurst, Herbert Gangl, Michael Hoffman, Kentaro Ihara, Masanobu Kaneko, Yasuo Ohno, Don Zagier and many others that I followed at the time formed my appreciation to the subject; I met all these mathematicians in person at later moments of my life

and some of them became my collaborators (not necessarily on multiple zeta values!) and friends. In the beginning of the 2000s we formed a little circle in Moscow of those indifferent to the arithmetic questions of zeta values and polylogarithms — in particular, their multiple generalisations — with Zhenya Ulanskii, Yuri Nesterenko and Sergey Zlobin; I owe these colleagues of mine at the time special thanks for an educational and joyful atmosphere of our seminars where we not only learned but also produced new creative results. Some related work — partly in parallel — of Stéphane Fischler, Christian Krattenthaler, Tanguy Rivoal, Vladislav Salikhov, Carlo Viola has been another personal inspiration for years. My interest in multiple zeta values was renewed in 2007 after a lecture of Yasuo Ohno and from our joint research stay at the Max-Planck Institute for Mathematics in Bonn; this was a multiply productive period of my life — thank you, Yasuo! Of course, I am grateful to the Max-Planck Institute for the unique scientific welcoming atmosphere that leads to my multiple visits there. Besides those acknowledged above, I am very thankful to Henrik Bachmann, Benjamin Brindle, Francis Brown, Steven Charlton, Kurusch Ebrahimi-Fard, Hidekazu Furusho, Tatiana Hessami Pilehrood, Jan-Willem van Ittersum, Ulf Kühn, Dominique Manchon, Kohji Matsumoto, Maki Nakasuji, Danylo Radchenko, Yoshitaka Sasaki, Nobuo Sato, Shin-ichiro Seki, Koji Tasaka, Hirofumi Tsumura, James Wan, Shuji Yamamoto, Jianqiang Zhao, for being on the multiple zeta value journey, with and without me.

The text below will not be possible without my family being next to me all the time — heartfelt thanks, Olga and Victor!

Wadim Zudilin
Nijmegen, Bonn and Newcastle
13 April 2025

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,*

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

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

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

EXERCISE 1.3. Establish the following evaluation of trigonometric integral:

$$\int_0^{\pi/2} \cos^{m-1} x\,\sin^{n-1} x\,\mathrm{d}x = \frac{1}{2}\,\mathrm{B}\big(\tfrac{m}{2},\tfrac{n}{2}\big) = \frac{1}{2}\,\frac{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{m+n}{2})}.$$

Give *closed forms* of the integral when m and n are positive integers.

EXERCISE 1.4. (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,

$$\begin{aligned}
&\int_{D(\rho)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z \\
&\quad = e^{-\pi i(s-1)} \int_\rho^\varepsilon \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x + i\int_{-\pi}^{\pi} \frac{(\varepsilon e^{i\theta})^s e^{a\varepsilon(\cos\theta+i\sin\theta)}}{1-e^{\varepsilon(\cos\theta+i\sin\theta)}}\,\mathrm{d}\theta \\
&\qquad + e^{\pi i(s-1)} \int_\varepsilon^\rho \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x \\
&\quad = -2i\sin\pi s \int_\varepsilon^\rho \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x + i\varepsilon^{s-1}\int_{-\pi}^{\pi} \frac{\varepsilon e^{is\theta+a\varepsilon(\cos\theta+i\sin\theta)}}{1-e^{\varepsilon(\cos\theta+i\sin\theta)}}\,\mathrm{d}\theta
\end{aligned}$$

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

$$\begin{aligned}
\int_{D(\infty)} \frac{(-z)^{s-1}e^{-az}}{1-e^{-z}}\,\mathrm{d}z &= -2i\sin\pi s\int_0^\infty \frac{x^{s-1}e^{-ax}}{1-e^{-x}}\,\mathrm{d}x \\
&= -2i\sin\pi s\,\Gamma(s)\zeta(s,a) = -2\pi i\,\frac{\zeta(s,a)}{\Gamma(1-s)}
\end{aligned}$$

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 1.5. 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 1.6. 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 1.7. 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* $\sigma = \operatorname{Re} s < 0$,

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

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

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

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 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 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 1.8. 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 1.9. (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 $s \geq 3$.
(c) Verify that $B_s(1) = B_s = B_s(0)$ for even $s \geq 0$.

LEMMA 1.7. *For $0 < a \leq 1$ 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 $|z| < 2\pi$, $z \ne 0$. 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 1.9).

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

EXERCISE 1.10. 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 1.11. 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* (MZF)

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

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 1.12. 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 $\sigma_1 + \dots + \sigma_j > j$, 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 $m \ge 2$,*

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

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

PROOF. Comparing Hurwitz's equation (1.9),

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

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

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

which is exactly

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

or

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

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 1.9) 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 1.13. Verify the validity of (1.15) for $m = 0, 1$.

We will also implement the (standard) notation

$$(s)_m = \frac{\Gamma(s+m)}{\Gamma(s)} = \begin{cases} s(s+1)\cdots(s+m-1) & \text{if } m = 1, 2, \dots, \\ 1 & \text{if } m = 0, \end{cases} \tag{1.16}$$

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

$$\sum_{n>M} \frac{1}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{M} \frac{1}{n^s}$$

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

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

Exercise 1.14. 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 $\mathbb{C}^l$:

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

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 $m > l + |\sigma_1| + \cdots + |\sigma_l|$, we have*

$$\zeta(\boldsymbol{s}) = \sum_{k=0}^m \frac{B_k}{k!}\,(s_1)_{k-1} \cdot \zeta(s_1 + s_2 + k - 1, s_3, \dots, s_l)$$
$$- \frac{(s_1)_m}{m!} \sum_{n_2 > \cdots > n_l \ge 1} \frac{1}{n_2^{s_2} \cdots n_l^{s_l}} \int_{n_2}^\infty \frac{\widetilde{B}_m(x)}{x^{s_1+m}}\,\mathrm{d}x. \tag{1.17}$$

Proof. The absolute convergence of the second series in the formula (1.17) 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

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

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 $n_2 > n_3 > \cdots > n_l \geq 1$ is bounded above by n_2^{l-2} (because each n_j satisfies $1 \leq n_j < n_2$), so that

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

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

Now, to get the formula (1.17) we apply Lemma 1.14 with $s = s_1$, $n = n_1$ and $M = n_2$, and then perform the summation over $n_2 > n_3 > \cdots > n_l \geq 1$.

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

Exercise 1.15. 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 (1.17) 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 multiple 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). $\square$

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

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

hence

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

and

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

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):

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

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

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

Going on in equality (2.5) we find that

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

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

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

therefore

$$\begin{aligned}&\sum_{n_0>n_1>\dots>n_l\ge 1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}m_{k-1}n_k^{s_k}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\ &\quad= \sum_{j=0}^{s_k-1}\sum_{n_0>n_1>\dots>n_l\ge 1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}n_{k-1}^{s_k-j}n_k^{j+1}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\ &\qquad+ \sum_{n_0>n_1>\dots>n_l\ge 1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}n_{k-1}^{s_k}m_{k-1}n_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\\ &\quad= \sum_{j=0}^{s_k-1}\zeta(s_1,\dots,s_{k-1},s_k-j,j+1,s_{k+1},\dots,s_l)\\ &\qquad+ \sum_{n_0>n_1>\dots>n_l\ge 1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{k-2}^{s_{k-1}}n_{k-1}^{s_k}m_kn_{k+1}^{s_{k+1}}\cdots n_l^{s_l}}\,, \end{aligned}\tag{2.8}$$

where again we swap the role of m_{k-1} and m_k. Applying consequently identities (2.8) for $k=2,\dots,l$ for the second multiple sum on the right-hand side of equality (2.7), we obtain

$$\begin{aligned}&\sum_{k=1}^{l}\bigl(\zeta(s_1,\dots,s_{k-1},s_k+1,s_{k+1},\dots,s_l)+\zeta(s_1,\dots,s_{k-1},s_k,1,s_{k+1},\dots,s_l)\bigr)\\ &\quad= \sum_{j=0}^{s_1-2}\zeta(s_1-j,j+1,s_2,\dots,s_l)\\ &\qquad+\sum_{k=2}^{l}\sum_{j=0}^{s_k-1}\zeta(s_1,\dots,s_{k-1},s_k-j,j+1,s_{k+1},\dots,s_l)\\ &\qquad+\sum_{n_0>n_1>\dots>n_l\ge 1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{l-1}^{s_k}m_l}\\ &\quad= \sum_{k=1}^{l}\sum_{j=0}^{s_k-2}\zeta(s_1,\dots,s_{k-1},s_k-j,j+1,s_{k+1},\dots,s_l)\\ &\qquad+\sum_{k=2}^{l}\zeta(s_1,\dots,s_{k-1},1,s_k,s_{k+1},\dots,s_l)+\sum_{n_0>n_1>\dots>n_l\ge 1}\frac{1}{n_0^{s_1}n_1^{s_2}\cdots n_{l-1}^{s_k}n_l}\end{aligned}$$

$$= \sum_{k=1}^{l} \sum_{j=0}^{s_k-2} \zeta(s_1, \dots, s_{k-1}, s_k - j, j+1, s_{k+1}, \dots, s_l)$$
$$+ \sum_{k=1}^{l} \zeta(s_1, \dots, s_k, 1, s_{k+1}, \dots, s_l). \tag{2.9}$$

Reducing 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*

$$\sum_{j=2}^{s-1} \zeta(j, s-j) = \zeta(s) \tag{2.10}$$

takes place.

Note also that, in the case $s = 3$, identity (2.10) 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 [**33**]). *For any $s \geq 3$,*

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

PROOF. We follow the proof from [**44**]. Write the left-hand side of (2.11) 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

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

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

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 $s \ge 4$ *even*, show that

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

EXERCISE 2.4 (Euler). For $s \ge 3$, 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 [**21**], Hoffman and Ohno proved the following result also by means of the partial-fraction method. A somewhat simpler proof was given by Ohno [**30**] (see also the later publication [**31**] 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*

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

holds.

One of the consequences of the cyclic sum theorem is independence of the sum of all multiple zeta values of fixed length and fixed weight 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 $l \geq 1$, the identity*

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

holds.

We prove this theorem in Section 4.3. In fact, Theorems 2.1 and 2.5 are particular instances of Ohno's relations (Theorem 5.1), which we discuss in Chapter 5.

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

$$\frac{1}{\Gamma(s)} \int_0^\infty x^{s-1} e^{-xn}\,\mathrm{d}x = \frac{1}{n^s\Gamma(s)} \int_0^\infty (nx)^{s-1} e^{-nx}\,\mathrm{d}(nx) = \frac{1}{n^s} \tag{2.12}$$

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

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

$$\zeta(\boldsymbol{s}) = \sum_{m_1,\dots,m_l\geq 1} \int\!\cdots\!\int_{u_l>u_{l-1}>\cdots>u_1>u_0} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)}, \tag{2.13}$$

where we set $u_0 = 0$.

PROOF. It follows from (2.12) that

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

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

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

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

in (2.13) given as follows:

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

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

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

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

Denote

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

(it differs from the one in (2.13) by reducing the integration domain to $u_l > u_{l-1} > \cdots > u_1 > u$, so that $f(\boldsymbol{s}; 0) = \zeta(\boldsymbol{s})$) and

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

(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

$$\sum_{m_1,\dots,m_l \ge 1} \int\cdots\int_{T_0(u)} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j - 1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)} = f_l(s_1, \dots, s_l; u), \tag{2.14}$$

$$\sum_{m_1,\dots,m_l \ge 1} \int\cdots\int_{T_l(u)} \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j - 1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)} = g_l(s_1, \dots, s_l; 0; u). \tag{2.15}$$

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). $\square$

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

$$\sum_{m_1,\dots,m_l\ge 1} \int\limits_{T_k(u)} \cdots \int \prod_{j=1}^{l} (u_j - u_{j-1})^{s_j-1} e^{-u_j m_j} \frac{\mathrm{d}u_j}{\Gamma(s_j)}$$
$$= \sum_{q=0}^{s_{k+1}-1} \frac{(-1)^q}{q!}\, g_k(s_1,\dots,s_k;q;u) f_{l-k}(s_{k+1}-q, s_{k+2},\dots,s_l;u). \quad (2.16)$$

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 (2.14)–(2.16) 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 $s_1,\dots,s_l$ positive integers and $u > 0$ real,*

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

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 $j = 1,\dots,l$ and $n_{l+1} = 0$, so that

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

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

$$\begin{aligned}&\int\cdots\int_{u_l>u_{l-1}>\cdots>u_1>u}\prod_{j=1}^{l} e^{-(u_j-u_{j-1})n_j}\,\frac{\mathrm{d}u_j}{\Gamma(s_j)}\\ &\quad=\int\cdots\int_{x_1>u,\,x_2>0,\,\dots,\,x_l>0}\prod_{j=1}^{l} e^{-x_jn_j}\,\frac{\mathrm{d}x_j}{\Gamma(s_j)}=\frac{e^{-un_1}}{n_1n_2\cdots n_l}\prod_{j=1}^{l}\frac{1}{(s_j-1)!}\,.\end{aligned}$$

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

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

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

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

Recall that

$$\frac{z}{e^z-1}=\sum_{k=0}^{\infty}\frac{B_k}{k!}\,z^k. \tag{2.17}$$

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

$$\begin{aligned}&\int_0^{u_2}\frac{\mathrm{d}u_1}{\Gamma(s_1)}\,\frac{(u_2-u_1)^{s_2-1}u_1^{s_1-1}}{e^{u_1}-1}\\ &\quad=\int_0^1\frac{u_2\,\mathrm{d}v_1}{\Gamma(s_1)}\,(u_2-u_2v_1)^{s_2-1}(u_2v_1)^{s_1-2}\sum_{k_1\ge 0}\frac{B_{k_1}}{k_1!}\,(u_2v_1)^{k_1}\\ &\quad=\sum_{k_1\ge 0}\frac{B_{k_1}}{k_1!}\,u_2^{k_1+s_1+s_2-2}\cdot\frac{1}{\Gamma(s_1)}\int_0^1 v_1^{k_1+s_1-2}(1-v_1)^{s_2-1}\,\mathrm{d}v_1\\ &\quad=\sum_{k_1\ge 0}\frac{B_{k_1}}{k_1!}\,u_2^{k_1+s_1+s_2-2}\cdot\frac{\Gamma(s_2)\,\Gamma(k_1+s_1-1)}{\Gamma(s_1)\,\Gamma(k_1+s_1+s_2-1)},\end{aligned}$$

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

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

and so on. The final result reads

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

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 (2.17) 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 $k_1+\dots+k_l \le K$, 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 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 $\sqcup\!\sqcup$ (the *shuffle product*) on $\mathfrak{H}$ and $*$ (the *harmonic* or *stuffle product*) on $\mathfrak{H}^1$ by the rules

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

for any word w, and

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

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

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,*

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

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

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

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, which is equivalent to Hoffman's relations in Theorem 2.1; we will discuss later its different proof.

Theorem 3.3. *The map ζ satisfies the relations*

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

(*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 $\zeta(s)$.

Exercise 3.4. (a) Using the stuffle product show that

$$\zeta(2m+1)\zeta(\{2\}^n) = \sum_{i=0}^{n} \zeta(\{2\}^i, 2m+1, \{2\}^{n-i}) + \sum_{i=1}^{n} \zeta(\{2\}^{i-1}, 2m+3, \{2\}^{n-i})$$

for integers $m, n \ge 1$.
(b) Deduce from part (a) that

$$\sum_{m=1}^{n} (-1)^{m-1}\zeta(2m+1)\zeta(\{2\}^{n-m}) = \sum_{i=1}^{n} \zeta(\{2\}^{i-1}, 3, \{2\}^{n-i})$$

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

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*

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

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

$$\operatorname{Li}_{\boldsymbol{s}}(1) = \zeta(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{Z}^l, \quad s_1 \ge 2, \ s_2 \ge 1, \ \dots, \ s_l \ge 1. \tag{3.10}$$

Taking, as before for multiple zeta values,

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

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, $w = x_j u$ for some word $u \in \mathfrak{H}^1$). Then*

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

where

$$\omega_j(z) = \omega_{x_j}(z) = \begin{cases} \dfrac{\mathrm{d}z}{z} & \text{if } x_j = x_0, \\ \dfrac{\mathrm{d}z}{1-z} & \text{if } x_j = x_1. \end{cases} \tag{3.13}$$

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

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

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

$$\begin{aligned}\mathrm{d}\,\mathrm{Li}_{x_0u}(z) &= \frac{1}{z}\sum_{n_1>n_2>\cdots>n_l\ge 1}\frac{z^{n_1}}{n_1^{s_1-1}n_2^{s_2}\cdots n_l^{s_l}}\,\mathrm{d}z\\ &= \frac{1}{z}\,\mathrm{Li}_{s_1-1,s_2,\dots,s_l}(z)\,\mathrm{d}z = \omega_0(z)\,\mathrm{Li}_u(z)\end{aligned}$$

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

$$\begin{aligned}\mathrm{d}\,\mathrm{Li}_{x_1u}(z) &= \sum_{n_1>n_2>\cdots>n_l\ge 1}\frac{z^{n_1-1}}{n_2^{s_2}\cdots n_l^{s_l}}\,\mathrm{d}z = \sum_{n_2>\cdots>n_l\ge 1}\frac{1}{n_2^{s_2}\cdots n_l^{s_l}}\sum_{n_1=n_2+1}^{\infty} z^{n_1-1}\,\mathrm{d}z\\ &= \frac{1}{1-z}\sum_{n_2>\cdots>n_l\ge 1}\frac{z^{n_2}}{n_2^{s_2}\cdots n_l^{s_l}}\,\mathrm{d}z = \frac{1}{1-z}\,\mathrm{Li}_{s_2,\dots,s_l}(z)\,\mathrm{d}z = \omega_1(z)\,\mathrm{Li}_u(z),\end{aligned}$$

and the result follows. $\square$

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

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

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 3.5. (a) Compute $\mathrm{Li}_{x_1x_0}(z)$.
(b) Show that

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

Hint. (a) It is standard to use

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

We get

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

(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). $\square$

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, $w_1=x_ju$ and $w_2=x_kv$, hence by Lemma 3.5 and the inductive hypothesis we have

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

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. $\square$

Proof of Theorem 3.1. Theorem 3.1 follows from Lemma 3.6 and relations (3.10). $\square$

EXERCISE 3.6. Show that

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

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 $w \mapsto \mathrm{Li}_w(z)$ of the shuffle algebra $\mathfrak{H}_{ш}$ over $\mathbb{C}$ is injective, 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 [**42**].

EXERCISE 3.7. Verify that the dilogarithm function Li_2 satisfies the identity

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

EXERCISE 3.8. (a) Demonstrate that for $n = 1, 2, \dots$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\biggl(x\sum_{j=2}^{n}(-1)^j \mathrm{Li}_n(x) + (1-x)\mathrm{Li}_1(x) - x\biggr) = (-1)^n \mathrm{Li}_n(x).$$

(b) For $n = 1, 2, \dots$, show that the function

$$f_n(x) = \sum_{j=2}^{n} \mathrm{Li}_n(x) + (1-x)\mathrm{Li}_1(x) - nx$$

satisfies

$$\frac{\partial^n}{\partial x_1 \cdots \partial x_n} f_n(x_1 \cdots x_n) = \log(1 - x_1 \cdots x_n).$$

(c) Given $k \in \mathbb{Z}_{\ge 0}$ and $n \in \mathbb{Z}_{>0}$, prove that there is a linear form $f_{n,k}(x)$ in single polylogarithms $(1-x)\mathrm{Li}_1(x), \mathrm{Li}_2(x), \dots, \mathrm{Li}_m(x), \dots$ and powers of the logarithm $\log^j x$, where $j = 1, 2, \dots$, such that

$$\frac{\partial^n}{\partial x_1 \cdots \partial x_n} f_{n,k}(x_1 \cdots x_n) = (x_1 \cdots x_n)^{k-1} \log(1 - x_1 \cdots x_n).$$

HINT. (b) Show that, more generally,

$$\frac{\partial^n}{\partial x_1 \cdots \partial x_n} f_n(tx_1 \cdots x_n) = t\log(1 - tx_1 \cdots x_n)$$

by induction on n. □

EXERCISE 3.9 ('Landen' connection formula [**35, 42**]). Prove that

$$\mathrm{Li}_{\boldsymbol{s}}(z) = \mathrm{Li}_{x_0^{s_1-1}x_1 \cdots x_0^{s_l-1}x_1}(z) = (-1)^l \sum_{\substack{w_1,\dots,w_l \\ |w_j| = s_j - 1 \text{ for } j=1,\dots,l}} \mathrm{Li}_{w_1x_1\cdots w_lx_l}\biggl(\frac{-z}{1-z}\biggr).$$

Note that $z \mapsto -z/(1-z)$ is an involution.

3.3. Duality of MZVs

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

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

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

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

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 $\omega_0(z)$ and $\omega_1(z)$.

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, $\tau(x_0^2 x_1 x_0 x_1) = x_0 x_1 x_0 x_1^2$. Clearly, τ is an involution preserving weight. It can be easily seen that τ is also the automorphism of the subalgebra $\mathfrak{H}^0$.

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). $\square$

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

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

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 3.10. 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),
$$\zeta(w) = \int\cdots\int_{z_0>z_1>\dots>z_k>z_{k+1}} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k)$$
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:
$$\begin{aligned}
\zeta(w) &= \sum_{j=0}^{k} \int\cdots\int_{z_0>\dots>z_j>z>z_{j+1}>\dots>z_{k+1}} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_k}(z_k) \\
&= \sum_{j=0}^{k} \int\cdots\int_{z_0>\dots>z_j>z} \omega_{\varepsilon_1}(z_1)\cdots\omega_{\varepsilon_j}(z_j) \\
&\qquad \times \int\cdots\int_{z>z_{j+1}>\dots>z_{k+1}} \omega_{\varepsilon_{j+1}}(z_{j+1})\cdots\omega_{\varepsilon_k}(z_k) \\
&= \sum_{j=0}^{k} \int\cdots\int_{1-z>z'_j>\dots>z'_1>0} \omega_{1-\varepsilon_j}(z'_j)\cdots\omega_{1-\varepsilon_1}(z'_1) \\
&\qquad \times \int\cdots\int_{z>z_{j+1}>\dots>z_{k+1}} \omega_{\varepsilon_{j+1}}(z_{j+1})\cdots\omega_{\varepsilon_k}(z_k) \\
&= \sum_{j=0}^{k} \operatorname{Li}_{\tau(x_{\varepsilon_1}\cdots x_{\varepsilon_j})}(1-z)\operatorname{Li}_{x_{\varepsilon_{j+1}}\cdots x_{\varepsilon_k}}(z).
\end{aligned}$$
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 $\operatorname{Li}_w(z)$ 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 3.11. (a) Give the formula for $\zeta(s)$, where $s>1$ is an integer, 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 3.6. □

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*

$$\zeta(s_1,\dots,s_l;\sigma_1,\dots,\sigma_l)=\sum_{n_1>n_2>\dots>n_l\ge1}\frac{\sigma_1^{n_1}\sigma_2^{n_2}\cdots\sigma_l^{n_l}}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}},\tag{3.22}$$

where $\sigma_j\in\{\pm1\}$ 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 $\sigma_j=-1$. For example, $\zeta(\overline{1})=\zeta(1;-1)=\mathrm{Li}_1(-1)=-\log 2$ and $\zeta(\overline{2},1)=\zeta(2,1;-1,1)$.

EXERCISE 3.12. Show that

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

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

In Section 4.4 we will see that the standard algebraic setup for the alternating 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

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

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)

$$\zeta_{<N}(\boldsymbol{s})=H(\boldsymbol{s};N)=H(s_1,\dots,s_l;N)=\sum_{N>n_1>n_2>\dots>n_l\ge1}\frac{1}{n_1^{s_1}n_2^{s_2}\cdots n_l^{s_l}},\tag{3.24}$$

where $N=1,2,\dots$; 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

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

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

Fix $N\in\mathbb{Z}_{>0}$. 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 $N\in\mathbb{Z}_{>0}$ 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 $w_1=y_ju$ and $w_2=y_kv$ for $u=y_{s_1}\cdots y_{s_l}$ and $v=y_{r_1}\cdots y_{r_i}$. Then

$$\begin{aligned}
&H(w_1;N)H(w_2;N)\\
&\quad=\sum_{N>n_0>n_1>\dots>n_l\ge1}\frac{1}{n_0^jn_1^{s_1}\cdots n_l^{s_l}}\sum_{N>m_0>m_1>\dots>m_i\ge1}\frac{1}{m_0^km_1^{r_1}\cdots m_i^{r_i}}
\end{aligned}$$

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

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

(we apply the inductive hypothesis to the internal products)

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

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

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

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

Exercise 3.13. Show that

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

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

We will have another opportunity to witness the multiple harmonic sum (3.24) as a refinement of multiple zeta value $\zeta(\boldsymbol{s})$ in Section 5.2.

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

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

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). $\square$

The approach in Exercise 3.14 (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 $\mathfrak{H}_{ш}$ 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 3.15. 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 3.16. 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 $\mathrm{Der}(\mathfrak{A})$ (naturally graded by degree).

It can be easily seen that, for defining a derivation $\delta \in \mathrm{Der}(\mathfrak{A})$, 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 $\delta_ш\colon \mathfrak{H} \to \mathfrak{H}$ and $\delta_*\colon \mathfrak{H}^1 \to \mathfrak{H}^1$, defined by the formulae

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

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:

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

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 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,

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

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

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

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

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:

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

This completes the proof. $\square$

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

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

was given by Waldschmidt. (As it is easily seen, specialising $z_2 = \cdots = z_l = 1$ 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 $z_1, \dots, z_l \in \{\pm1\}$ (see Sections 3.3 and 4.4).

EXERCISE 3.17. (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

Generating functions and periodic multi-indices

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

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

possesses the generating function

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

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

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

denotes the Pochhammer symbol (1.16).

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

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

we have

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

Lemma 4.1. *The hypergeometric function $F(a, b; c; z)$ satisfies the differential equation*

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

in equivalent form,

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

Lemma 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 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 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 $n \geq 1$, 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

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

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

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:

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

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

$$\zeta(m+2,\{1\}^n)=\zeta(n+2,\{1\}^m),\quad\text{where } m,n=0,1,2,\dots,$$

may be derived by the generating-function method (as well as by straightforward application of Theorem 3.7). The fact that both sides of this equality are expressed as polynomials in single zeta values $\zeta(s)$ with rational coefficients is the subject of Exercise 4.6.

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

A different proof of the first identity in (4.4) is discussed in Exercise 4.6 below.

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 [15].

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:

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

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. Multiple zeta values of fixed weight, length and height

In this section we discuss a different application of generating functions and the theory of hypergeometric series.

We will need the formula

$$\Gamma(1-x) = \exp\biggl(\gamma x + \sum_{k=2}^{\infty} \frac{\zeta(k)x^k}{k}\biggr). \tag{4.8}$$

This follows from Proposition 1.1.

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 $I(w, l, m)$, and set

$$\Phi(x, y, z) = \sum_{w,l,m=0}^{\infty} x^{w-l-m} y^{l-m} z^{2m-2} \sum_{\boldsymbol{s} \in I(w,l,m)} \zeta(\boldsymbol{s})$$

to be a (formal) power series with real coefficients.

THEOREM 4.6. *The generating function $\Phi(x, y, z)$ is given by*

$$\Phi(x, y, z) = \frac{1}{xy - z^2}\biggl(1 - \exp\biggl(\sum_{k=2}^{\infty} \frac{\zeta(k)}{k} S_k(x, y, z)\biggr)\biggr), \tag{4.9}$$

where the homogeneous *polynomials $S_k(x, y, z)$ of degree k are defined through the formula*

$$S_k(x, y, z) = x^k + y^k - \alpha^k - \beta^k \quad \text{with } \{\alpha, \beta\} = \frac{x + y \pm \sqrt{(x+y)^2 - 4z^2}}{2}, \tag{4.10}$$

or alternatively by the identity

$$\sum_{k=2}^{\infty} \frac{S_k(x, y, z)}{k} = \log\biggl(1 - \frac{xy - z^2}{(1-x)(1-y)}\biggr). \tag{4.11}$$

In particular, all of the coefficients $\sum_{\boldsymbol{s}\in I(w,l,m)}\zeta(\boldsymbol{s})$ of $\Phi(x,y,z)$ can be expressed as polynomials in single zeta values $\zeta(2),\zeta(3),\dots$ with rational coefficients.

PROOF. If one defines, more generally,

$$\Phi(x,y,z;t)=\sum_{w,l,m=0}^{\infty}x^{w-l-m}y^{l-m}z^{2m-2}\sum_{\boldsymbol{s}\in I(w,l,m)}\mathrm{Li}_{\boldsymbol{s}}(t),$$

$$\widetilde{\Phi}(x,y,z;t)=\sum_{w,l,m=0}^{\infty}x^{w-l-m}y^{l-m}z^{2m}\sum_{\boldsymbol{s}\in \tilde{I}(w,l,m)}\mathrm{Li}_{\boldsymbol{s}}(t),$$

where $\tilde{I}(w,l,m)$ is the set of all multi-indices $\boldsymbol{s}$ including those with $s_1=1$. Using the differential equations of the generalised polylogarithms, Lemma 3.5, we find out that

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t}=\frac{x}{t}\,\Phi+\frac{1}{yt}\,(\widetilde{\Phi}-1-z^2\Phi),\quad \frac{\mathrm{d}}{\mathrm{d}t}(\widetilde{\Phi}-z^2\Phi)=\frac{y}{1-t}\,\widetilde{\Phi}. \tag{4.12}$$

One can eliminate $\widetilde{\Phi}$ from this system and write a homogeneous linear 2nd order differential equation for $Y=1-(xy-z^2)\Phi$:

$$\frac{\mathrm{d}^2Y}{\mathrm{d}t^2}+\left(\frac{1-x}{t}-\frac{y}{1-t}\right)\frac{\mathrm{d}Y}{\mathrm{d}t}+\frac{xy-z^2}{t(1-t)}\,Y=0. \tag{4.13}$$

This equation is recognised as a hypergeometric differential equation (see Lemma 4.1) whose unqiue holomorphic solution at $t=0$ starting $Y(t)=1+O(t)$ is given by the hypergeometric function $F(\alpha-x,\beta-x;1-x;t)$, where $\alpha+\beta=x+y$ and $\alpha\beta=z^2$. Specialising to $t=1$ and using Gauss' summation formula (Lemma 4.3) arrive at

$$1-(xy-z^2)\Phi(x,y,z;1)=F(\alpha-x,\beta-x;1-x;1)=\frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-\alpha)\Gamma(1-\beta)}.$$

The rest follows from the power series expansion (4.8). □

Particular specialisations of Theorem 4.6 lead one to numerous beautiful identities of MZVs, in particular, to a simple proof of the sum formula.

PROOF OF THEOREM 2.5. Letting $z^2\to xy$ in (4.9) we obtain the generating function

$$\Phi(x,y,\sqrt{xy})=\sum_{w,l}^{\infty}x^{w-l-1}y^{l-1}\sum_{\boldsymbol{s}\in I(w,l,m)m\text{ any}}\zeta(\boldsymbol{s})$$

of multiple zeta values of weight w and length l on the left-hand side. At the same time, the right-hand side simplifies to

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{(n-x)(n-y)}&=\sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{k_1,k_2\ge 0}\left(\frac{x}{n}\right)^{k_1}\left(\frac{y}{n}\right)^{k_2}\\&=\sum_{k_1,k_2\ge 0}\zeta(k_1+k_2+2)x^{k_1}y^{k_2},\end{aligned}$$

so that the coefficient of $x^{w-l-1}y^{l-1}$ in $\Phi(x,y,\sqrt{xy})$ is equal to $\zeta(w)$ as required. $\square$

EXERCISE 4.6 ([**32**]). (a) Using Theorem 4.6 prove that

$$\zeta(\{2\}^n)=\frac{\pi^{2n}}{(2n+1)!}\quad\text{for } n=0,1,\dots.$$

(b) Show that, for each $m,n=0,1,2,\dots$, the multiple zeta values $\zeta(m+2,\{1\}^n)=\zeta(n+2,\{1\}^m)$ are polynomials in single zeta values with rational coefficients.

HINT. Use specialisation $x=y=0$ in (a) and $z=0$ in (b). $\square$

4.4. An identity for alternating Euler sums

In this part we discuss a family of relations for the alternating Euler sums (3.22) formally introduced in Section 3.3.

THEOREM 4.7 (Zhao). *The following equalities are true*:

$$\zeta(\{\overline{2},1\}^n)=\frac{1}{8^n}\zeta(\{3\}^n),\quad \textit{where } n=1,2,\dots. \tag{4.14}$$

The family of identities (4.14) was conjectured by Borwein, Bradley and Broadhurst in [**5**] (see (3.22) for the definition of alternating Euler sums); it generalises Exercises 3.10 and 3.12 (b) and looks similar to that in Theorem 4.5. It was proven more than a decade later by Zhao [**49**] using the (finite) double shuffle relations and distribution relation for the alternating Euler sums, so that it was quite from the proof given in Section 4.2. A proof by generating functions is still wanted. Here we adopt Zhao's only-known proof of (4.14).

We have essentially settled standard setup for the (alternating) Euler sums at the end of Section 3.3. The non-commutative algebra $\mathfrak{H}=\mathbb{Q}\langle x_0,x_1\rangle$ is extended to the algebra $\widehat{\mathfrak{H}}=\mathbb{Q}\langle x_0,x_1,\overline{x}_1\rangle$, and its subalgebra

$$\widehat{\mathfrak{H}}^0=\mathbb{Q}\mathbf{1}\oplus x_0\widehat{\mathfrak{H}}x_1\oplus x_0\widehat{\mathfrak{H}}\overline{x}_1\oplus \overline{x}_1\widehat{\mathfrak{H}}x_1$$

of *admissible* words is generated by words not beginning with x_1 and not ending with x_0. Furthermore, $\widehat{\mathfrak{H}}^0$ is generated by the words $y_s=x_0^{s-1}x_1$ and $\overline{y}_s=x_0^{s-1}\overline{x}_1$, where $s=1,2,\dots$, with the only restriction that the words over this newer alphabet cannot begin with y_1.

By assigning the three differential forms

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

(cf. (3.23)) to the three letters, for a word $w \in \widehat{\mathfrak{H}}^0$ we define the evaluation zeta map by

$$\zeta(w) = \int_0^1 w \tag{4.15}$$

(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_k after y_{s_i} (inclusive) to $\overline{y}_k$ until the next signed entry $\overline{s}_j$ occur. We then leave all the y_k 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 think of the bars as of switches between y and $\overline{y}$.

Exercise 4.7. Write the word which corresponds to the multi-index

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

Exercise 4.8. Prove the following correspondence:

$$(\{\overline{2}, 1\}^n) \mapsto (x_0\overline{x}_1^2x_0x_1^2)^{\lfloor n/2\rfloor}(x_0\overline{x}_1^2)^{2\{n/2\}} = \begin{cases} (x_0\overline{x}_1^2x_0x_1^2)^k(x_0\overline{x}_1^2) & \text{if } n = 2k+1, \\ (x_0\overline{x}_1^2x_0x_1^2)^k & \text{if } n = 2k, \end{cases}$$
$$= (\overline{y}_2\overline{y}_1y_2y_1)^{\lfloor n/2\rfloor}(\overline{y}_2\overline{y}_1)^{2\{n/2\}}.$$

The shuffle and stuffle products in (3.2)–(3.4) are extended to the algebra $\widehat{\mathfrak{H}}^0$ 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

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

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

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

Then

$$\zeta(w_1 \sqcup\!\sqcup w_2) = \zeta(w_1 * w_2) = \zeta(w_1)\zeta(w_2). \tag{4.16}$$

Exercise 4.9. (a) Prove the (finite) double shuffle relations (4.18).
(b) Verify that $x_1 \sqcup\!\sqcup w - x_1 * w \in \widehat{\mathfrak{H}}^0$ for all $w \in \widehat{\mathfrak{H}}^0$ and $\zeta(x_1 \sqcup\!\sqcup w - x_1 * w) = 0$.

Zhao's proof of (4.14) follows from the three lemmas displayed below. Though the results of Lemmas 4.8 and 4.9 represent a pure combinatorial structure, they are merely computational. In their statements we take $u = x_0(x_1\overline{x}_1 + \overline{x}_1 x_1) = y_2\overline{y}_1 + \overline{y}_2 y_1$ to be a particular word of weight 3.

LEMMA 4.8. *For* $n = 0, 1, 2, \ldots,$

$$u^n + \sum_{i=1}^{n} u^{n-i} * \overline{y}_1 u^{i-1}(y_2 + \overline{y}_2) - \sum_{i=1}^{n} \overline{y}_1 u^{n-i} * u^{i-1}(y_2 + \overline{y}_2) = (-1)^n (y_3 + \overline{y}_3)^n. \tag{4.17}$$

PROOF. Trivially, the statement is valid for $n = 0$.

Note that the words $\mathbf{1}$, $y_2 + \overline{y}_2$ and any power of u are invariant under the toggling operator $\gamma\colon w \mapsto \overline{w}$. This simplifies application of the stuffle rules when these words and their products are involved. For $v_1, v_2 \in \{\mathbf{1}, y_2 + \overline{y}_2\}$, integers $k \geq 0$ and $m \geq 1$ we have

$$\begin{aligned}
\overline{y}_1 u^k v_1 * u^m v_2 &= \overline{y}_1 u^k v_1 * (y_2\overline{y}_1 + \overline{y}_2 y_1) u^{m-1} v_2 \\
&= \overline{y}_1 (u^k v_1 * u^m v_2) \\
&\quad + y_2(\overline{y}_1 u^k v_1 * \overline{y}_1 u^{m-1} v_2) + \overline{y}_2(\overline{\overline{y}_1 u^k v_1 * \overline{y}_1 u^{m-1} v_2}) \\
&\quad + y_3(u^k v_1 * \overline{y}_1 u^{m-1} v_2) + \overline{y}_3(\overline{u^k v_1 * \overline{y}_1 u^{m-1} v_2}) \\
&= \overline{y}_1 (u^k v_1 * u^m v_2) + (1+\gamma)\big(y_2(\overline{y}_1 u^k v_1 * \overline{y}_1 u^{m-1} v_2)\big) \\
&\quad + (1+\gamma)\big(y_3(u^k v_1 * \overline{y}_1 u^{m-1} v_2)\big)
\end{aligned}$$

but also

$$\begin{aligned}
\overline{y}_1 u^k v_1 * (y_2 + \overline{y}_2) &= \overline{y}_1 (u^k v_1 * (y_2 + \overline{y}_2)) \\
&\quad + y_2 \overline{y}_1 u^k v_1 + \overline{y}_2 y_1 u^k v_1 + (y_3 + \overline{y}_3) u^k v_1 \\
&= \overline{y}_1 (u^k v_1 * (y_2 + \overline{y}_2)) + u^{k+1} v_1 + (y_3 + \overline{y}_3) u^k v_1.
\end{aligned}$$

Therefore, denoting the left-hand side of (4.17) by w_n we obtain, by telescoping,

$$\begin{aligned}
w_n = u^n &+ \overline{y}_1\Bigg(\sum_{i=1}^{n} u^{n-i} * u^{i-1}(y_2+\overline{y}_2) - \sum_{i=1}^{n} u^{n-i} * u^{i-1}(y_2+\overline{y}_2)\Bigg)\\
&+ (1+\gamma)\Bigg(y_2\Bigg(\sum_{i=1}^{n-1} \overline{y}_1 u^{n-i-1} * \overline{y}_1 u^{i-1}(y_2+\overline{y}_2)\\
&\qquad - \sum_{i=2}^{n} \overline{y}_1 u^{n-i} * \overline{y}_1 u^{i-1}(y_2+\overline{y}_2)\Bigg)\Bigg)\\
&+ (1+\gamma)\Bigg(y_3\Bigg(\sum_{i=1}^{n-1} \overline{y}_1 u^{n-i-1} * u^{i-1}(y_2+\overline{y}_2)\\
&\qquad - \sum_{i=2}^{n} u^{n-i} * \overline{y}_1 u^{i-1}(y_2+\overline{y}_2)\Bigg)\Bigg)\\
&- \big(u^n + (y_3+\overline{y}_3)u^{n-1}\big)\\
= -&(1+\gamma)(y_3 w_{n-1}),
\end{aligned}$$

and the result follows from the inductive hypothesis $w_{n-1} = (-1)^{n-1}(y_3+\overline{y}_3)^{n-1}$. $\square$

LEMMA 4.9. *For $n = 0, 1, 2, \ldots$,*

$$\begin{aligned}
u^n &+ \sum_{i=1}^{n} u^{n-i} ш \overline{x}_1 u^{i-1} x_0(x_1+\overline{x}_1) - \sum_{i=1}^{n} \overline{x}_1 u^{n-i} ш u^{i-1} x_0(x_1+\overline{x}_1)\\
&= (-2)^n (x_0\overline{x}_1^2 x_0 x_1^2)^{\lfloor n/2 \rfloor}(x_0\overline{x}_1^2)^{2\{n/2\}}.
\end{aligned} \tag{4.18}$$

PROOF. Our strategy is similar to that for the proof of Lemma 4.8 but we deal exclusively with the shuffle in this part. Again, equality (4.18) is trivially true for $n = 0$.

Let w_n denote the left-hand side of (4.18). Notice that the shuffle rules allow to swap the role of any two letters, in particular, of x_1 and $\overline{x}_1$; this means that our target equality (4.18) is equivalent to

$$\overline{w_n} = (-2)^n (x_0 x_1^2 x_0 \overline{x}_1^2)^{\lfloor n/2 \rfloor}(x_0 x_1^2)^{2\{n/2\}}.$$

Take $\{v_1, v_2\} = \{\mathbf{1}, x_0(x_1+\overline{x}_1)\}$ and recall that $u = x_0 u'$, where $u' = x_1\overline{x}_1 + \overline{x}_1 x_1$. Then for non-negative integers $k \ge 1$ and $m \ge 1$ we get

$$\begin{aligned}
\overline{x}_1 u^k v_1 ш u^m v_2 &= \overline{x}_1(u^k v_1 ш u^m v_2) + x_0(\overline{x}_1 u^k v_1 ш (x_1\overline{x}_1 + \overline{x}_1 x_1)u^{m-1} v_2)\\
&= \overline{x}_1(u^k v_1 ш u^m v_2) + x_0 x_1(\overline{x}_1 u^k v_1 ш \overline{x}_1 u^{m-1} v_2)\\
&\quad + x_0\overline{x}_1(\overline{x}_1 u^k v_1 ш x_1 u^{m-1} v_2) + x_0\overline{x}_1(u^k v_1 ш u' u^{m-1} v_2)
\end{aligned}$$

$$
\begin{aligned}
&= \overline{x}_1(u^k v_1 \sqcup\!\sqcup u^m v_2) + x_0 x_1(\overline{x}_1 u^k v_1 \sqcup\!\sqcup \overline{x}_1 u^{m-1} v_2)\\
&\quad + x_0\overline{x}_1 x_1(\overline{x}_1 u^k v_1 \sqcup\!\sqcup u^{m-1} v_2) + x_0\overline{x}_1 x_1(u^k v_1 \sqcup\!\sqcup \overline{x}_1 u^{m-1} v_2)\\
&\quad + 2x_0\overline{x}_1^2(u^k v_1 \sqcup\!\sqcup x_1 u^{m-1} v_2) + x_0\overline{x}_1 x_0(u' u^{k-1} v_1 \sqcup\!\sqcup u' u^{m-1} v_2),
\end{aligned}
$$

with the formula remaining valid for $m = 0$, $v_2 = \mathbf{1}$ (in which case all the terms containing u^{m-1} have to be dropped) and for $k = 0$, $v_1 = \mathbf{1}$ (in which case the last term containing u^{k-1} has to be dropped). In addition, with $v = x_0(x_1+\overline{x}_1)$ we have

$$
\begin{aligned}
\overline{x}_1 u^k \sqcup\!\sqcup v &= \overline{x}_1(u^k \sqcup\!\sqcup v) + x_0(x_1+\overline{x}_1)\overline{x}_1 u^k + x_0\overline{x}_1(u^k \sqcup\!\sqcup (x_1+\overline{x}_1))\\
&= \overline{x}_1(u^k \sqcup\!\sqcup v) + x_0(x_1+\overline{x}_1)\overline{x}_1 u^k + x_0\overline{x}_1(x_1+\overline{x}_1)u^k\\
&\quad + x_0\overline{x}_1 x_0(u' u^{k-1} \sqcup\!\sqcup (x_1+\overline{x}_1))\\
&= \overline{x}_1(u^k \sqcup\!\sqcup v) + u^{k+1} + 2x_0\overline{x}_1^2 u^k + x_0\overline{x}_1 x_0(u' u^{k-1} \sqcup\!\sqcup (x_1+\overline{x}_1))
\end{aligned}
$$

for $k \geq 1$ and

$$
\begin{aligned}
\overline{x}_1 v \sqcup\!\sqcup u^m &= \overline{x}_1(v \sqcup\!\sqcup u^m) + x_0 x_1(\overline{x}_1 v \sqcup\!\sqcup \overline{x}_1 u^{m-1})\\
&\quad + x_0\overline{x}_1(\overline{x}_1 v \sqcup\!\sqcup x_1 u^{m-1}) + x_0\overline{x}_1(v \sqcup\!\sqcup u' u^{m-1})\\
&= \overline{x}_1(v \sqcup\!\sqcup u^m) + x_0 x_1(\overline{x}_1 v \sqcup\!\sqcup \overline{x}_1 u^{m-1})\\
&\quad + x_0\overline{x}_1 x_1(\overline{x}_1 v \sqcup\!\sqcup u^{m-1}) + x_0\overline{x}_1 x_1(v \sqcup\!\sqcup \overline{x}_1 u^{m-1})\\
&\quad + 2x_0\overline{x}_1^2(v \sqcup\!\sqcup x_1 u^{m-1}) + x_0\overline{x}_1 x_0((x_1+\overline{x}_1) \sqcup\!\sqcup u' u^{m-1})
\end{aligned}
$$

for $m \geq 1$. Substituting these findings into the left-hand side of (4.18) we obtain, for $n \geq 1$,

$$
\begin{aligned}
w_n &= u^n + \sum_{i=1}^{n} u^{n-i} \sqcup\!\sqcup \overline{x}_1 u^{i-1} v - \sum_{i=1}^{n} \overline{x}_1 u^{n-i} \sqcup\!\sqcup u^{i-1} v\\
&= \sum_{i=1}^{n-1}\big(x_0\overline{x}_1 x_1(\overline{x}_1 u^{i-1} v \sqcup\!\sqcup u^{n-i-1}) + x_0\overline{x}_1 x_1(u^{i-1} v \sqcup\!\sqcup \overline{x}_1 u^{n-i-1})\\
&\quad + 2x_0\overline{x}_1^2(u^{i-1} v \sqcup\!\sqcup x_1 u^{n-i-1})\big)\\
&\quad - \sum_{i=2}^{n}\big(x_0\overline{x}_1 x_1(\overline{x}_1 u^{n-i} \sqcup\!\sqcup u^{i-2} v) + x_0\overline{x}_1 x_1(u^{n-i} \sqcup\!\sqcup \overline{x}_1 u^{i-2} v)\\
&\quad + 2x_0\overline{x}_1^2(u^{n-i} \sqcup\!\sqcup x_1 u^{i-2} v)\big) - 2x_0\overline{x}_1^2 u^{n-1}\\
&= -2x_0\overline{x}_1^2\overline{w_{n-1}}.
\end{aligned}
$$

Thus, the desired formula follows from the inductive hypothesis for w_{n-1} (hence for $\overline{w_{n-1}}$). $\square$

Lemma 4.10. *For $n = 0, 1, 2, \ldots,$*

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

PROOF. Observe that
$$\frac{\mathrm{d}z}{z}=\frac{1}{2}\,\frac{\mathrm{d}(z^2)}{z^2}\quad\text{and}\quad\frac{\mathrm{d}z}{1-z}+\frac{-\mathrm{d}z}{1+z}=\frac{2z\,\mathrm{d}z}{1-z^2}=\frac{\mathrm{d}(z^2)}{1-z^2}.$$
Performing the change of variables $z^2\mapsto z$ in the iterated integral (4.15) for $\zeta\big(\big(x_0^2(x_1+\overline{x}_1)\big)^n\big)$ we obtain the integral for $2^{-2n}\zeta\big((x_0^2x_1)^n\big)$. □

PROOF OF THEOREM 4.7. The statement of Lemma 4.8 can be alternatively written as
$$\begin{aligned}u^n+\sum_{i=1}^n u^{n-i}*\overline{x}_1u^{i-1}x_0(x_1+\overline{x}_1)-\sum_{i=1}^n \overline{x}_1u^{n-i}*u^{i-1}x_0(x_1+\overline{x}_1)\\ =(-1)^n\big(x_0^2(x_1+\overline{x}_1)\big)^n;\end{aligned}$$
in other words, its left-hand side coincides with that in Lemma 4.9 except that every shuffle product in the latter is replaced by the stuffle product in the former. Application of the double shuffle relations (Exercise 4.9) then implies that
$$\zeta\big(\big(x_0^2(x_1+\overline{x}_1)\big)^n\big)=2^n\zeta\big((x_0x_1^2x_0\overline{x}_1^2)^{\lfloor n/2\rfloor}(x_0x_1^2)^{2\{n/2\}}\big);$$
it remains to use Lemma 4.10 to arrive at (4.14). □

The result of Lemma 4.10 is in fact a particular instance of the distribution relation of multiple polylogarithms (3.39): for any $d\in\mathbb{Z}_{>0}$ and $\boldsymbol{s}=(s_1,\dots,s_l)$,
$$\sum_{\substack{z_j^d=a_j\\ j=1,\dots,l}}\operatorname{Li}_{\boldsymbol{s}}(z_1,\dots,z_l)=d^{l-|\boldsymbol{s}|}\operatorname{Li}_{\boldsymbol{s}}(a_1,\dots,a_l).$$
When $d=2$ and $a_1=\dots=a_l=1$, we have
$$\zeta\big(\big(x_0^2(x_1+\overline{x}_1)\big)^l\big)=\sum_{n_1>\dots>n_l>0}\frac{(1+(-1)^{n_1})\cdots(1+(-1)^{n_l})}{n_1^3\cdots n_l^3}=2^{-2l}(\{3\}^l).$$
Zhao notices that this relation does not follow from the (finite) double shuffle relations, which are involved in the other part of his proof in [**49**].

EXERCISE 4.10. Check that indeed $x_0^2(x_1+\overline{x}_1)=y_3+\overline{y}_3$ cannot be reduced to $\frac{1}{4}\,x_0^2x_1=\frac{1}{4}y_3$ in $\widehat{\mathfrak{H}}$ using the finite double shuffle relations.

HINT. You can list, for example, all linear relations in $\widehat{\mathfrak{H}}$ of weight 3. □

CHAPTER 5

Further relations of MZVs

5.1. Ohno's relations

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

THEOREM 5.1 (Ohno's relations [29]). *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 $\mathfrak{H}^1$:*

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

Then, for any integer $m \ge 0$, the identity

$$\sum_{\substack{i_1, i_2, \dots, i_l \ge 0 \\ i_1 + i_2 + \dots + i_l = m}} \zeta(y_{s_1+i_1} y_{s_2+i_2} \cdots y_{s_l+i_l}) = \sum_{\substack{i_1, i_2, \dots, i_k \ge 0 \\ i_1 + i_2 + \dots + i_k = m}} \zeta(y_{s'_1+i_1} y_{s'_2+i_2} \cdots y_{s'_k+i_k})$$

holds.

The proof of this theorem was given by Ohno in 1999. It used manipulations with multiple integral representations of (generating functions of) the sums involved in the identity. Different proofs were given later by Bradley [7] (he proved a q-version of Theorem 5.1 — see Theorem 7.2), Okuda and Ueno [35], and Ulanskii [43]. Here we follow a very simple and elementary proof given by Seki and Yamamoto [39] using the method of *connected sums*.

We continue to use recording of multi-indices as words over the alphabets x_0, x_1 and $y_s = x_0^{s-1} x_1$, where $s = 1, 2, \dots$; all the notations are as in Section 3.1. To a (formal) variable t and two multi-indices $\boldsymbol{s} = (s_1, \dots, s_l)$ and $\boldsymbol{r} = (r_1, \dots, r_k)$, not necessarily admissible if both non-empty, we assign the 'connected' sum

$$\begin{aligned} Z\left(\begin{matrix} x_0^{s_1-1} x_1 \cdots x_0^{s_l-1} x_1 \\ x_0^{r_1-1} x_1 \cdots x_0^{r_k-1} x_1 \end{matrix} \,\middle|\, t \right) &= Z\left(\begin{matrix} \boldsymbol{s} \\ \boldsymbol{r} \end{matrix} \,\middle|\, t \right) = Z\left(\begin{matrix} s_1, \dots, s_l \\ r_1, \dots, r_k \end{matrix} \,\middle|\, t \right) \\ &= \sum_{\substack{n_1 > \dots > n_l > 0 \\ m_1 > \dots > m_k > 0}} C(n_1, m_1; t) \frac{1}{n_1^{s_1-1}(n_1 - t) \cdots n_l^{s_l-1}(n_l - t)} \\ &\qquad \times \frac{1}{m_1^{r_1-1}(m_1 - t) \cdots m_k^{r_k-1}(m_k - t)}, \end{aligned} \tag{5.1}$$

where the *connector* $C(n, m)$ is defined by

$$C(n, m; t) = \frac{(1-t)_n (1-t)_m}{(1-t)_{n+m}} = \frac{\Gamma(n+1-t)\Gamma(m+1-t)}{\Gamma(1-t)\Gamma(n+m+1-t)}$$

(we refer to the Pochhammer symbol (1.16)) and $C(n,\emptyset;t)=C(\emptyset,m;t)=1$ (this is the case when one of multi-indices $\boldsymbol{s}$ and $\boldsymbol{r}$ is not present, so that the corresponding sum in (5.1) degenerates to a sum over single group of variables). Without the connector in (5.1), the right-hand side would be simply the product of two multiple sums (the product $\zeta(\boldsymbol{s})\zeta(\boldsymbol{r})$ when $t=0$ provided both $\boldsymbol{s}$ and $\boldsymbol{r}$ are admissible). The symmetry

$$Z\left(\begin{matrix}\boldsymbol{s}\\ \boldsymbol{r}\end{matrix}\,\middle|\,t\right)=Z\left(\begin{matrix}\boldsymbol{r}\\ \boldsymbol{s}\end{matrix}\,\middle|\,t\right) \tag{5.2}$$

is clear from the definition, and we also have

$$\begin{aligned} Z\left(\begin{matrix}x_0^{s_1-1}x_1\cdots x_0^{s_l-1}x_1\\ \mathbf{1}\end{matrix}\,\middle|\,t\right)&=Z\left(\begin{matrix}\boldsymbol{s}\\ -\end{matrix}\,\middle|\,t\right)\\ &=\sum_{n_1>\dots>n_l>0}\frac{1}{n_1^{s_1-1}(n_1-t)\cdots n_l^{s_l-1}(n_l-t)} \end{aligned} \tag{5.3}$$

in the case when the second multi-index is absent.

EXERCISE 5.1. Assume that t is real, $t<1$. Prove that the multiple sum indeed converges if both $\boldsymbol{s}$ and $\boldsymbol{r}$ have length at least 1.

The main rationale behind the definition (5.1) is the following property of the connected sums.

LEMMA 5.2 (Transporting relations). *If* $s_1>1$ *then*

$$Z\left(\begin{matrix}s_1,\dots,s_l\\ r_1,\dots,r_k\end{matrix}\,\middle|\,t\right)=Z\left(\begin{matrix}s_1-1,\dots,s_l\\ 1,r_1,\dots,r_k\end{matrix}\,\middle|\,t\right); \tag{5.4}$$

if $s_1=1$ *then*

$$Z\left(\begin{matrix}1,s_2,\dots,s_l\\ r_1,\dots,r_k\end{matrix}\,\middle|\,t\right)=Z\left(\begin{matrix}s_2,\dots,s_l\\ r_1+1,\dots,r_k\end{matrix}\,\middle|\,t\right). \tag{5.5}$$

Equivalently (and uniformly),

$$Z\left(\begin{matrix}x_\varepsilon u\\ v\end{matrix}\,\middle|\,t\right)=Z\left(\begin{matrix}u\\ x_{1-\varepsilon}v\end{matrix}\,\middle|\,t\right) \tag{5.6}$$

for $x_\varepsilon\in\{x_0,x_1\}$ *and two words* u,v *over the alphabet* $\{x_0,x_1\}$.

PROOF. Clearly, property (5.4) follows from (5.5) and the symmetry (5.2). To prove (5.5), observe that

$$\begin{aligned} \frac{1}{n-t}C(n,m;t)&=\frac{(1-t)_{n-1}(1-t)_m}{(1-t)_{n+m}}\\ &=\frac{1}{m}\left(\frac{(1-t)_{n-1}(1-t)_m}{(1-t)_{n+m-1}}-\frac{(1-t)_n(1-t)_m}{(1-t)_{n+m}}\right)\\ &=\frac{1}{m}\big(C(n-1,m;t)-C(n,m;t)\big), \end{aligned}$$

hence

$$\sum_{n_1=n_2+1}^{\infty} \frac{C(n_1, m_1; t)}{n_1 - t} = \frac{C(n_2, m_1; t)}{m}$$

by telescoping, where $n_2 = 0$ if $l = 1$. □

EXERCISE 5.2. Using properties (5.2), (5.3) and (5.6) give another proof of Euler's identity (1.13).

PROOF OF THEOREM 5.1. The properties of the Seki–Yamamoto connected sums (5.1) imply

$$\sum_{n_1>\cdots>n_l>0} \frac{1}{n_1^{s_1-1}(n_1-t)\cdots n_l^{s_l-1}(n_l-t)} = Z\left(\begin{matrix} w \\ \mathbf{1} \end{matrix} \,\middle|\, t\right) = \cdots = Z\left(\begin{matrix} \mathbf{1} \\ \tau w \end{matrix} \,\middle|\, t\right)$$

$$= \sum_{n_1>\cdots>n_k>0} \frac{1}{n_1^{s_1'-1}(n_1-t)\cdots n_k^{s_k'-1}(n_k-t)}.$$

It remains to expand both sides in powers of t using

$$\frac{1}{n^{s-1}(n-t)} = \frac{1}{n^s(1-t/n)} = \sum_{i=0}^{\infty} \frac{t^i}{n^{s+i}}$$

and compare the coefficients of t^m. □

It is straightforward that case $m = 0$ in Theorem 5.1 is the duality theorem (Theorem 3.7).

EXERCISE 5.3. (a) Show that the choice $m = 1$ in Theorem 5.1 corresponds to Hoffman's relations (Theorem 2.1).
(b) Show that, if multi-index $\boldsymbol{s}$ in Theorem 5.1 is one-component (that is, $\boldsymbol{s} = (s)$), then the theorem reduces to the sum theorem (Theorem 2.5).

EXERCISE 5.4 ([**35**]). Deduce Theorem 5.1 from the Landen connection formula in Exercise 3.9.

5.2. The Maesaka–Seki–Watanabe formula

As another illustration of the power of the method of connected sums used in the previous section, we prove a refinement of the link between the multiple zeta values defined via the the sum (2.1) and their representation (3.18) as iterated integrals. For the former we use the multiple harmonic sums

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

from Section 3.4, while for the latter we introduce a new type of multiple sum

$$\zeta^\flat_{<N}(\boldsymbol{s}) = \sum_{\substack{N>n_{11}\ge\cdots\ge n_{1s_1} \\ >n_{21}\ge\cdots\ge n_{2s_2} \\ \cdots\cdots\cdots \\ >n_{l1}\ge\cdots\ge n_{ls_l}>0}} \prod_{j=1}^{l} \frac{1}{n_{j1}n_{j2}\cdots n_{j,s_j-1}(N-n_{js_j})} \tag{5.7}$$

inspired by the representation (3.18). Indeed, we have $\frac{1}{n} = \frac{1}{N} \cdot \frac{1}{n/N}$ and $\frac{1}{N-n} = \frac{1}{N} \cdot \frac{1}{1-n/N}$, and this allows us to recognise the sum in (5.7) as an N-regular Riemann sum for the integral (3.18). For example,

$$\zeta^\flat_{<N}(2,1,3) = \frac{1}{N^6} \sum_{N>n_1\ge n_2>n_3>n_4\ge n_5\ge n_6>0} \frac{1}{\frac{n_1}{N}(1-\frac{n_2}{N})\cdot(1-\frac{n_3}{N})\cdot\frac{n_4}{N}\frac{n_5}{N}(1-\frac{n_6}{N})}$$

which is a regular Riemann sum that approximates the iterated integral

$$\int_{1>z_1>z_2>z_3>z_4>z_5>z_6>0} \frac{\mathrm{d}z_1}{z_1}\frac{\mathrm{d}z_2}{1-z_2}\cdot\frac{\mathrm{d}z_3}{1-z_3}\cdot\frac{\mathrm{d}z_4}{z_4}\frac{\mathrm{d}z_5}{z_5}\frac{\mathrm{d}z_6}{1-z_6} = \zeta(2,1,3);$$

thus, we obtain

$$\lim_{N\to\infty} \zeta^\flat_{<N}(2,1,3) = \zeta(2,1,3).$$

EXERCISE 5.5. Verify that

$$\lim_{N\to\infty} \zeta^\flat_{<N}(\boldsymbol{s}) = \zeta(\boldsymbol{s})$$

for all admissible $\boldsymbol{s}$.

Quite amazingly, as shown by Maesaka, Seki and Watanabe in [**28**], the limiting equality

$$\lim_{N\to\infty} \zeta^\flat_{<N}(\boldsymbol{s}) = \zeta(\boldsymbol{s}) = \lim_{N\to\infty} \zeta_{<N}(\boldsymbol{s})$$

for admissible $\boldsymbol{s}$ has the following natural refinement.

THEOREM 5.3. *For any index* $\boldsymbol{s}$ (*not necessarily admissible*) *and any integer* $N>0$, *we have*

$$\zeta_{<N}(\boldsymbol{s}) = \zeta^\flat_{<N}(\boldsymbol{s}). \tag{5.8}$$

PROOF. It is convenient to introduce the compact notation $\Sigma_N(\boldsymbol{s})$ for the summation indices

$$\boldsymbol{n} = (n_{11},\dots,n_{1s_1},n_{21},\dots,n_{2s_2},\dots,n_{l1},\dots,n_{ls_l})$$

in (5.7). The choice of connected sums is

$$\begin{aligned} Z_N(\boldsymbol{s}\mid\boldsymbol{r}) &= Z_N(s_1,\dots,s_l\mid r_1,\dots,r_k)\\ &= \sum_{\substack{\boldsymbol{n}\in\Sigma_N(\boldsymbol{s})\\ N>m_1>\dots>m_k>0}} \prod_{j=1}^{l} \frac{1}{n_{j1}\cdots n_{j,s_j-1}(N-n_{js_j})}\cdot C_N(n_{ls_l}-1,m_1)\cdot\prod_{i=1}^{k}\frac{1}{m_i^{r_i}}, \end{aligned}$$

where the connector $C_N(n,m)$ is given by

$$C_N(n,m) = \frac{\binom{n}{m}}{\binom{N-1}{m}} = \frac{n(n-1)\cdots(n-m+1)}{(N-1)(N-2)\cdots(N-m)}.$$

The transport identity

$$Z_N(\boldsymbol{s},t\mid\boldsymbol{r}) = Z_N(\boldsymbol{s}\mid t,\boldsymbol{r}) \tag{5.9}$$

then implies that

$$\begin{aligned}\zeta^\flat_{<N}(\boldsymbol{s}) &= Z_N(s_1,\dots,s_l \mid \emptyset) = Z_N(s_1,\dots,s_{l-1} \mid s_l) = \cdots \\ &= Z_N(s_1 \mid s_2,\dots,s_l) = Z_N(\emptyset \mid s_1,s_2,\dots,s_l) = \zeta_{<N}(\boldsymbol{s}),\end{aligned}$$

which is precisely the identity in (5.8).

The remaining check of (5.9) follows from application once of the telescoping identity

$$\frac{C_N(n-1,m)}{N-n} = \sum_{a=m+1}^{N-1} \frac{C_N(n-1,a-1)-C_N(N-1,a)}{N-n} = \sum_{a=m+1}^{N-1} \frac{C_N(n,a)}{n}$$

and then t times of the identity

$$\sum_{b=1}^{n} \frac{C_N(b,m)}{b} = \sum_{b=1}^{n} \frac{C_N(b,m)-C_N(b-1,m)}{m} = \frac{C_N(n,m)}{m}. \qquad \square$$

Finally notice that Hirose, Matsusaka and Seki [**14**] generalise the formula from Theorem 5.3 to the case of multiple polylogarithms; some further variations on the theme are discussed by Yamamoto in [**46**].

5.3. Ihara–Kaneko derivations

Theorem 3.12 has a natural generalization. For any $n \ge 1$, define the anti-symmetric derivation $\partial_n \in \mathrm{Der}(\mathfrak{H})$ by the rule $\partial_n x_0 = x_0(x_0+x_1)^{n-1}x_1$; as mentioned in the proof of Theorem 3.3, we have $\partial_1 = \overline{D} - D = \delta_* - \delta_{\sqcup\!\sqcup}$. The following result is valid.

Theorem 5.4. *For any $n \ge 1$ and any word $w \in \mathfrak{H}^0$, the identity*

$$\zeta(\partial_n w) = 0 \tag{5.10}$$

holds.

In what follows, we describe a scheme of the proof of the theorem given by Kaneko and Ihara [**23**] (see also [**24**]); a different proof was provided by Hoffman and Ohno [**21**].

For each integer $n \ge 1$ 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 (Theorem 5.1) may be re-stated as follows.

THEOREM 5.5. *For any word $w \in \mathfrak{H}^0$, the inclusion*

$$(\sigma - \overline{\sigma})w \in \ker\zeta \tag{5.11}$$

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 5.1. Applying now Theorem 3.7 (with $m = n$), we arrive at the desired inclusion (5.11). $\square$

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 5.6. *The following equality holds:*

$$\exp(\partial) = \overline{\sigma}\cdot\sigma^{-1}. \tag{5.12}$$

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 (5.12). □

Proof of Theorem 5.4. Now let us show how Theorem 5.4 follows from Theorem 5.5 and Lemma 5.6. 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 5.5 yields the required identities (5.10). □

Does there exist a simpler way of proving relations (5.10)? 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 5.4 are served by induction (with Theorem 3.12 as inductive base). This circumstance motivates the following hypothesis.

Conjecture 5.7. *For any $n \geq 1$, the above-defined anti-symmetric derivation ∂_n is contained in the Lie subalgebra of* $\operatorname{Der}(\mathfrak{H})$ *generated by the derivations δ_*, $\overline{\delta}_*$, $\delta_{ш}$, and $\overline{\delta}_{ш}$.*

5.4. Open questions about MZVs

In addition to Conjectures 1.10, 3.4 and 5.7 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 5.8. *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 5.8 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 5.9. *For $k \geq 3$, the recurrence relations*

$$d_k = d_{k-2} + d_{k-3} \tag{5.13}$$

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 (5.13) (and $d_0 = d_2 = 1$, $d_1 = 0$). There are several proofs of this result, due to Terasoma [**41**], to Deligne and Goncharov [**10**], and to F. Brown [**8**]; all are algebraic and use motivic interpretations of the multiple zeta values.

Even if Conjectures 5.8 and 5.9 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 (compare, for example, with Exercise 3.4 (b)).

Conjecture 5.10 (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*

$$\bigl\{\zeta(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2, 3\},\ j = 1, \dots, \ell(\boldsymbol{s})\bigr\}. \tag{5.14}$$

A serious argument for Conjecture 5.10 to be valid, is not only experimental confirmation for $k \leq 16$ (under the hypothesis of Conjecture 3.4) but also agreement of the dimension of the $\mathbb{Q}$-space spanned by the numbers (5.14) with the dimension d_k of the spaces $\mathcal{Z}_k$ in Conjecture 5.9.

Exercise 5.6. For given $k = 0, 1, 2, \ldots$, show that the number of MZVs in (5.14) is equal to d_k. Here d_k is the same sequence defined earlier in (5.13).

Although proving Conjectures 5.8–5.10 in the form given is hopeless at the present time, the 'true' MZVs in $\mathbb{R}$ are the images under a $\mathbb{Q}$-linear map of certain *motivic* MZVs which are defined *purely algebraically*. The Terasoma [**41**] and Deligne–Goncharov [**10**] bound $\dim_{\mathbb{Q}} \mathcal{Z}_k \le d_k$, as well as Conjecture 5.8 about disjointness of the subspaces $\mathcal{Z}_k$, are shown to be true for the motivic MZVs. Terasoma and Goncharov established the bound by showing that all MZVs are periods of so-called *mixed Tate motives* that are unramified over $\mathbb{Z}$. Another well-known conjecture in the area states the converse, that is, that all periods of mixed Tate motives over $\mathbb{Z}$ can be expressed as linear combinations (over $\mathbb{Q}[(2\pi i)^{\pm 1}]$) of MZVs. Equivalently, this says that the dimension of the space of motivic MZVs of weight k is exactly d_k.

Brown [**8**] proved the latter conjecture and also the fact that the motivic MZVs from Hoffman's conjectural basis in Conjecture 5.10 form a basis of the corresponding $\mathcal{Z}_k$. Brown's proof requires quite specific properties of certain coefficients occurring in the relations over $\mathbb{Q}$ of some special MZVs; namely, that the MZVs

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

which are part of Hoffman's basis, are $\mathbb{Q}$-linear combinations of products $\pi^{2\mu}\zeta(2\nu+1)$ with $\mu + \nu = m + n + 1$. A very explicit version of such a formula was given by Zagier [**48**]; we discuss this remarkable identity of MZVs in Section 6.2.

Our final exercise in this part relates counting of the number of MZVs to partitions.

Exercise 5.7. (a) How many different MZVs of given weight k exists?

(b) Compute the limit of $d_k^{1/k}$ as $k \to \infty$ for the sequence d_k constructed in Conjecture 5.9.

(c) Any polynomial in single zeta values,

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

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

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

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

CHAPTER 6

The two-one formula and its relatives

6.1. 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 (6.11) 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 (5.14) 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)

$$\big\{\zeta(\boldsymbol{s}) : |\boldsymbol{s}| = k,\ s_j \in \{2,1\},\ j = 1, \dots, \ell(\boldsymbol{s}),\ \text{no 1s next to each other}\big\}. \tag{6.1}$$

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 in [**22**], and later confirmed by Glanois [**11**]; the equivalence of the two Hoffman's basis conjecture was also discussed by Zagier and Brown. Essentially, the original question was whether one could replace (conjecturally) the MZVs in the 'dual' Hoffman's basis (6.1) 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 (2.11) of Euler's original formula (2.10) (the sum formula of depth 2 in the modern terminology).

THEOREM 6.1 (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:*

$$\zeta^\star(\mu_{s_1},\mu_{s_2},\dots,\mu_{s_l})=\sum_{\boldsymbol{p}}(-1)^{\sigma(\boldsymbol{p})}2^{l-\sigma(\boldsymbol{p})}\zeta^\star(\boldsymbol{p}) \tag{6.2}$$

$$=\sum_{\boldsymbol{p}}2^{l-\sigma(\boldsymbol{p})}\zeta(\boldsymbol{p}), \tag{6.3}$$

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 (6.2), (6.3) 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

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

when $s_1,s_2\ge 1$, 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 $m\ge 2$ 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 (6.2) (equivalently, in (6.3)) by means of the single zeta value $(m+1)\zeta((m+1)n+1)$, where the integers $m\ge 2$ and $n\ge 1$ are arbitrary.

EXERCISE 6.1. Show the equality of the right-hand sides in (6.2) and (6.3).

HINT. Use Exercise 2.1. □

On the right-hand side of (6.2) and (6.3) 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

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

but also to a different type

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

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 6.2. 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 $k\ge 2$ and any multi-index $\boldsymbol{s}=(s_1,\dots,s_l)$,

$$\lim_{N\to\infty}\sum_{m=1}^{N}\frac{H(\boldsymbol{s};m-1)}{m^k}\biggl(1-\frac{N!^2}{(N-m)!\,(N+m)!}\biggr)=0. \tag{6.4}$$

(Notice that the expression in the parentheses is always positive.) Try first to prove (6.4) 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 $m\le\sqrt{N}$ 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 6.1 is the limiting case, as $N\to\infty$, of the following result.

THEOREM 6.2. *For any $N\in\mathbb{N}$,*

$$H^\star(\{2\}^{s_1},1,\{2\}^{s_2},1,\dots,\{2\}^{s_l},1;N)=2\sum_{\boldsymbol{p}=(2s_1+1)\circ(2s_2+1)\circ\dots\circ(2s_l+1)}2^{\bar\sigma(\boldsymbol{p})}\widehat{H}(\boldsymbol{p};N), \tag{6.5}$$

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 $H_N^\star(\boldsymbol{s})$ and $\widehat{H}_N(\boldsymbol{s})$ for $H^\star(\boldsymbol{s};N)$ and $\widehat{H}(\boldsymbol{s};N)$, respectively. The proof of (6.5) is by induction on $N+l$. As $H_1^\star(\boldsymbol{s})=1$ for any $\boldsymbol{s}$ and $\widehat{H}_1(\boldsymbol{s})=1/2$ if $l=1$ and 0 otherwise, the equality in (6.5) is trivially true when $N=1$ and $l\ge 0$ is arbitrary.

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

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

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

$$
\begin{aligned}
&H^\star_N(\{2\}^{s_1},1,\{2\}^{s_2},1,\ldots,\{2\}^{s_l},1)\\
&\quad= \frac{2}{N^{2s_1}}\sum_{k=0}^{s_1} N^{2k} \sum_{\boldsymbol{p}=(2k+1)\circ(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})}\widehat{H}_{N-1}(\boldsymbol{p})\\
&\qquad+ \frac{2}{N^{2s_1+1}} \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar{\sigma}(\boldsymbol{p})}\widehat{H}_{N}(\boldsymbol{p}).
\end{aligned}
\tag{6.6}
$$

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

$$
\begin{aligned}
&\sum_{k=0}^{s_1} N^{2k}\widehat{H}_{N-1}(p_1+2k,p_2,\ldots,p_r)\\
&\quad= N^{2s_1}\sum_{N>n_1>n_2>\cdots>n_r\geq 1}\frac{N!^2}{(N-n_1)!\,(N+n_1)!}\,\frac{1}{n_1^{p_1+2s_1}n_2^{p_2}\cdots n_r^{p_r}}\\
&\qquad-\frac{1}{N^2}\sum_{N>n_1>n_2>\cdots>n_r\geq 1}\frac{N!^2}{(N-n_1)!\,(N+n_1)!}\,\frac{1}{n_1^{p_1-2}n_2^{p_2}\cdots n_r^{p_r}}\\
&\quad= N^{2s_1}\sum_{N\geq n_1>n_2>\cdots>n_r\geq 1}\frac{N!^2}{(N-n_1)!\,(N+n_1)!}\,\frac{1}{n_1^{p_1+2s_1}n_2^{p_2}\cdots n_r^{p_r}}\\
&\qquad-\frac{1}{N^2}\sum_{N\geq n_1>n_2>\cdots>n_r\geq 1}\frac{N!^2}{(N-n_1)!\,(N+n_1)!}\,\frac{1}{n_1^{p_1-2}n_2^{p_2}\cdots n_r^{p_r}}\\
&\quad= N^{2s_1}\widehat{H}_N(p_1+2s_1,p_2,\ldots,p_r)-\frac{1}{N^2}\widehat{H}_N(p_1-2,p_2,\ldots,p_r).
\end{aligned}
$$

Therefore, the equality in (6.6) can be written as

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

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

$$
\begin{aligned}
&= \frac{2}{N^{2s_1+1}} \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar\sigma(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p}) \\
&\quad - \frac{2}{N^{2s_1+2}} \sum_{\boldsymbol{p}=(2s_2)\circ\cdots\circ(2s_l+1)} 2^{\bar\sigma(\boldsymbol{p})} \widehat{H}_N(\boldsymbol{p}) \\
&\quad - \frac{4}{N^{2s_1+2}} \sum_{m=1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m \sum_{\boldsymbol{p}=(2s_2+1)\circ\cdots\circ(2s_l+1)} 2^{\bar\sigma(\boldsymbol{p})} H_{m-1}(\boldsymbol{p}).
\end{aligned}
\tag{6.7}
$$

Finally, the other identity

$$
\begin{aligned}
&2\sum_{m=1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m\, H_{m-1}(p_1,p_2,\dots,p_r) \\
&\qquad = 2\sum_{m=1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m \sum_{n_1=1}^{m-1} \frac{H_{n_1-1}(p_2,\dots,p_r)}{n_1^{p_1}} \\
&\qquad = \sum_{n_1=1}^{N} \frac{H_{n_1-1}(p_2,\dots,p_r)}{n_1^{p_1}} \cdot 2 \sum_{m=n_1+1}^{N} \frac{N!^2}{(N-m)!\,(N+m)!}\, m
\end{aligned}
$$

(the internal sum is summed by Exercise 6.3 below)

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

simplifies the right-hand side of (6.7) to zero. $\square$

EXERCISE 6.3. 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. $\square$

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 (6.2) as follows:

$$2 \int \cdots \int_{[0,1]^{s_1+\cdots+s_l}} \frac{\prod_{i=1}^{l-1}(1 + t_1 \cdots t_{s_1+\cdots+s_i})}{\prod_{i=1}^{l}(1 - t_1 \cdots t_{s_1+\cdots+s_i})} \mathrm{d}t_1 \cdots \mathrm{d}t_{s_1+\cdots+s_l}. \tag{6.8}$$

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

$$\begin{aligned} 2 \int\cdots\int_{1>u_1>\cdots>u_{s_1+\cdots+s_l}>0} & \prod_{i=1}^{l-1}\Biggl(\prod_{j=s_1+\cdots+s_{i-1}+1}^{s_1+\cdots+s_i-1} \frac{\mathrm{d}u_j}{u_j} \cdot \frac{(1 + u_{s_1+\cdots+s_i})\,\mathrm{d}u_{s_1+\cdots+s_i}}{(1 - u_{s_1+\cdots+s_i})u_{s_1+\cdots+s_i}} \Biggr) \\ & \times \prod_{j=s_1+\cdots+s_{l-1}+1}^{s_1+\cdots+s_l-1} \frac{\mathrm{d}u_j}{u_j} \cdot \frac{\mathrm{d}u_{s_1+\cdots+s_l}}{1 - u_{s_1+\cdots+s_l}}, \end{aligned} \tag{6.9}$$

where the empty sum $s_1 + \cdots + s_{i-1}$ for $i = 1$ is interpreted as 0. Therefore, any of the two integrals in (6.8), (6.9) may replace the right-hand sides of (6.2) or (6.3).

6.2. Zagier's identity

Zagier's formula shows that the multiple zeta values

$$\xi(m, n) = \zeta(\{2\}^m, 3, \{2\}^n) \quad \text{for } m, n \geq 0, \tag{6.10}$$

which are part of Hoffman's basis in Conjecture 5.10, 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) (see Exercise 4.6(a)),

$$\xi(n) = \zeta(\{2\}^n) = \frac{\pi^{2n}}{(2n + 1)!} \quad \text{for } n \geq 0, \tag{6.11}$$

for the simplest of the Hoffman basis elements.

THEOREM 6.3 (Zagier). *For all integers $m, n \geq 0$, we have*

$$\begin{aligned} \xi(m, n) = 2 \sum_{r=1}^{m+n+1} (-1)^{r-1} & \biggl(\Bigl(1 - \frac{1}{2^{2r}}\Bigr) \binom{2r}{2m + 1} - \binom{2r}{2n + 2} \biggr) \\ & \times \xi(m + n - r + 1)\zeta(2r + 1), \end{aligned} \tag{6.12}$$

where the value of $\xi(m+n-r+1)$ is given by (6.11). Conversely, each product $\xi(\mu)\zeta(k - 2\mu)$ of odd weight k is a rational combination of numbers $\xi(m, n)$ with $m + n = (k - 3)/2$.

Zagier's original proof [**48**] is skillfully designed and worth studying on its own. Other proofs can be found in [**12, 26, 27**] (see also [**13**]). Here we follow an elementary proof given by L. Lai, C. Lupu and D. Orr in [**25**].

LEMMA 6.4. *For nonnegative integers m and k the integral*

$$I_{m,k} = \frac{2}{\pi} \int_0^{\pi/2} \frac{(2z)^{2m}}{(2m)!} \cos^{2k} z \, \mathrm{d}z$$

is evaluated as follows:

$$I_{m,k} = \frac{\binom{2k}{k}}{2^{2k}} \sum_{k_1 > \cdots > k_m > k} \frac{1}{k_1^2 \cdots k_m^2},$$

where the empty sum (when $m = 0$) is understood as 1.

PROOF. For $m > 0$ and $k > 0$, start with integration by parts

$$\begin{aligned} I_{m,k} &= \frac{2}{\pi} \int_0^{\pi/2} \frac{(2z)^{2m}}{(2m)!} \cos^{2k-1} z \, \mathrm{d}(\sin z) \\ &= \frac{2}{\pi} \int_0^{\pi/2} \left(\frac{(2z)^{2m}}{(2m)!} (2k-1) \cos^{2k-2} z \, \sin^2 z - \frac{2(2z)^{2m-1}}{(2m-1)!} \cos^{2k-1} z \, \sin z \right) \mathrm{d}z \\ &= (2k-1)(I_{m,k-1} - I_{m,k}) + \frac{4}{\pi} \int_0^{\pi/2} \frac{(2z)^{2m-1}}{(2m-1)!} \cos^{2k-1} z \, \mathrm{d}(\cos z) \end{aligned}$$

implying

$$I_{m,k} = \frac{2k-1}{2k} I_{m,k-1} + \frac{2}{\pi k} \int_0^{\pi/2} \frac{(2z)^{2m-1}}{(2m-1)!} \cos^{2k-1} z \, \mathrm{d}(\cos z).$$

Then

$$\begin{aligned} &\frac{2}{\pi} \int_0^{\pi/2} \frac{(2z)^{2m-1}}{(2m-1)!} \cos^{2k-1} z \, \mathrm{d}(\cos z) \\ &\quad = \frac{2}{\pi} \int_0^{\pi/2} \left(\frac{(2z)^{2m-1}}{(2m-1)!} (2k-1) \cos^{2k-1} z \, \sin z - \frac{2(2z)^{2m-2}}{(2m-2)!} \cos^{2k} z \right) \mathrm{d}z \\ &\quad = -(2k-1) \cdot \frac{2}{\pi} \int_0^{\pi/2} \frac{(2z)^{2m-1}}{(2m-1)!} \cos^{2k-1} z \, \mathrm{d}(\cos z) - 2I_{m-1,k} \end{aligned}$$

implying

$$\frac{2}{\pi} \int_0^{\pi/2} \frac{(2z)^{2m-1}}{(2m-1)!} \cos^{2k-1} z \, \mathrm{d}(\cos z) = -\frac{1}{k} I_{m-1,k}.$$

Thus

$$I_{m,k} = \frac{2k-1}{2k} I_{m,k-1} - \frac{1}{k^2} I_{m-1,k},$$

which can be conveniently written as

$$\frac{2^{2k}}{\binom{2k}{k}} I_{m,k} = \frac{2^{2k-2}}{\binom{2k-2}{k-1}} I_{m,k-1} - \frac{2^{2k}}{\binom{2k}{k}} \frac{I_{m-1,k}}{k^2},$$

and the same recursion $S_{m,k} = S_{m,k-1} - S_{m-1,k}/k^2$ is satisfied by the sums

$$S_{m,k} = \sum_{k_1 > \cdots > k_m > k} \frac{1}{k_1^2 \cdots k_m^2}$$

for $m, k > 0$. Therefore, the formula stated follows by induction on $m+k$ with the help of identities

$$I_{0,k} = \frac{2}{\pi}\int_0^{\pi/2} \cos^{2k} z \,\mathrm{d}z = \frac{\binom{2k}{k}}{2^{2k}}$$

(see Exercise 1.3) and

$$I_{m,0} = \frac{2^{m+1}}{\pi(2m)!}\int_0^{\pi/2} z^{2m}\,\mathrm{d}z = \frac{\pi^{2m}}{(2m+1)!} = \sum_{k_1>\dots>k_m>0} \frac{1}{k_1^2\cdots k_m^2}$$

(see (6.11)). □

We also need the following special case of the hypergeometric function (see Section 4.1).

EXERCISE 6.4 ([40, eq. (1.5.7)]). The following identity is valid:

$$F\left(\tfrac12 a, -\tfrac12 a; \tfrac12; \sin^2 z\right) = \cos az.$$

Comparing the coefficients of the expansion of the identity in powers of a (and reverting the sides) we deduce the following formula.

LEMMA 6.5. *For positive integers n,*

$$\frac{z^{2n}}{(2n)!} = \frac{1}{2^{2n}}\sum_{k=1}^{\infty}\frac{(2\sin z)^{2k}}{k^2\binom{2k}{k}}\sum_{k>l_1>\dots>l_{n-1}\ge 1}\frac{1}{l_1^2\cdots l_{n-1}^2}.$$

For the next evaluation we introduce generalised Clausen's functions

$$\mathrm{Cl}_s(\theta) = \begin{cases} \displaystyle\sum_{n=1}^{\infty}\frac{\sin n\theta}{n^s} = \operatorname{Im}\mathrm{Li}_s(e^{i\theta}) & \text{for } s \text{ even},\\ \displaystyle\sum_{n=1}^{\infty}\frac{\cos n\theta}{n^s} = \operatorname{Re}\mathrm{Li}_s(e^{i\theta}) & \text{for } s \text{ odd}. \end{cases}$$

They satisfy differential equations

$$\frac{\mathrm{d}\,\mathrm{Cl}_s(\theta)}{\mathrm{d}\theta} = (-1)^s\,\mathrm{Cl}_{s-1}(\theta) \quad \text{for } s = 2, 3, \dots .$$

EXERCISE 6.5. (a) Give a closed form expression for $\mathrm{Cl}_1(\theta)$ and show that

$$\frac{\mathrm{d}\,\mathrm{Cl}_1(\theta)}{\mathrm{d}\theta} = -\frac12\cot\frac{\theta}{2}.$$

(b) (Clausen) Prove that

$$\mathrm{Cl}_2(\theta) = -\int_0^{\theta}\log\left|2\sin\frac{t}{2}\right|\mathrm{d}t.$$

LEMMA 6.6 ([**36**, Theorem 2.1]). *For a nonnegative integer n and real x,*

$$\int_0^{\pi x} z^n \cot z \,\mathrm{d}z = (\pi x)^n \sum_{k=0}^{n} (-1)^{\lfloor (k-1)/2 \rfloor} k! \binom{n}{k} \frac{\mathrm{Cl}_{k+1}(2\pi x)}{(2\pi x)^k} + \delta_{\lfloor n/2 \rfloor, n/2} \frac{(-1)^{n/2} n!}{2^n} \zeta(n+1), \tag{6.13}$$

where $\delta_{k,m}$ stands for Kronecker's delta.

PROOF. Denote the function on the right-hand side of (6.13) by $g(x)$. Then

$$\begin{aligned} g'(x) &= (\pi x)^n \sum_{k=0}^{n-1} (-1)^{\lfloor (k-1)/2 \rfloor} k! \binom{n}{k} (n-k) \frac{\mathrm{Cl}_{k+1}(2\pi x)}{x(2\pi x)^k} \\ &\quad + (\pi x)^n \sum_{k=1}^{n} (-1)^{\lfloor (k-1)/2 \rfloor} k! \binom{n}{k} \frac{(-1)^{k+1} 2\pi \,\mathrm{Cl}_k(2\pi x)}{(2\pi x)^k} + (\pi x)^n \pi \cot(\pi x) \\ &= \pi^{n+1} x^n \cot(\pi x) \end{aligned}$$

and this clearly coincides with the derivative of the left-hand side in (6.13). In addition,

$$g(0) = (-1)^{\lfloor (n-1)/2 \rfloor} n! \frac{\mathrm{Cl}_{n+1}(0)}{2^n} + \delta_{\lfloor n/2 \rfloor, n/2} \frac{(-1)^{n/2} n!}{2^n} \zeta(n+1) = 0. \qquad \square$$

LEMMA 6.7. *For a nonnegative integer n,*

$$\int_0^{\pi/2} z^n \cot z \,\mathrm{d}z = \Big(\frac{\pi}{2}\Big)^n \bigg(\log 2 + \sum_{r=1}^{\lfloor n/2 \rfloor} (-1)^k (2r)! (2^{2r} - 1 + \delta_{r, \lfloor n/2 \rfloor}) \binom{n}{2r} \frac{\zeta(2r+1)}{(2\pi)^{2r}} \bigg).$$

In particular, for a polynomial $P(z) \in \mathbb{C}[z]$ of degree $d > 0$ with $P(0) = 0$ we have

$$\begin{aligned} \int_0^{\pi/2} P(z) \cot z \,\mathrm{d}z &= P\Big(\frac{\pi}{2}\Big) \log 2 + \sum_{r=1}^{\lfloor d/2 \rfloor} \frac{(-1)^r}{2^{2r}} P^{(2r)}\Big(\frac{\pi}{2}\Big) \Big(1 - \frac{1}{2^{2r}}\Big) \zeta(2r+1) \\ &\quad + \sum_{r=1}^{\lfloor d/2 \rfloor} \frac{(-1)^r}{2^{2r}} P^{(2r)}(0) \zeta(2r+1). \end{aligned}$$

PROOF. For the first identity, substitute $x = \pi/2$ in (6.13) and use $\mathrm{Cl}_1(\pi) = -\log 2$, $\mathrm{Cl}_{2r+1}(\pi) = -(1 - 2^{-2r})\zeta(2r+1)$ and $\mathrm{Cl}_{2r}(\pi) = 0$ for $r = 1, 2, \dots$. For the second identity, we only need to verify it for $P(z) = z^n$ where $n = 1, 2, \dots$, but this is precisely the first one. $\square$

EXERCISE 6.6 (open problem). Give a direct proof of the formulae in Lemma 6.7.

PROOF OF THEOREM 6.3. The proof compares two evaluations of the very same integral

$$\hat{\xi}(m,n) = \frac{2}{\pi} \int_0^1 \frac{(2 \arccos x)^{2m+1}}{(2m+1)!} \frac{(2 \arcsin x)^{2n+2}}{(2n+2)!} \frac{\mathrm{d}x}{x}. \tag{6.14}$$

First use the identity of Lemma 6.5 for n shifted by 1 and $z = \arcsin x$ to write

$$\hat{\xi}(m,n) = \frac{2}{\pi}\sum_{k=1}^{\infty} \frac{1}{k^2\binom{2k}{k}}\biggl(\int_0^1 \frac{(2\arccos x)^{2m+1}}{(2m+1)!}\,(2x)^{2k}\,\frac{\mathrm{d}x}{x}\biggr) \sum_{k>l_1>\dots>l_n\ge 1} \frac{1}{l_1^2\cdots l_n^2}.$$

Now observe that

$$\begin{aligned}\int_0^1 \frac{(2\arccos x)^{2m+1}}{(2m+1)!}\,x^{2k-1}\,\mathrm{d}x &= \int_0^{\pi/2} \frac{(2z)^{2m+1}}{(2m+1)!}\,\cos^{2k-1} z\,\sin z\,\mathrm{d}z \\ &= -\frac{1}{2k}\int_0^{\pi/2} \frac{(2z)^{2m+1}}{(2m+1)!}\,\mathrm{d}(\cos^{2k} z) \\ &= \frac{1}{k}\int_0^{\pi/2} \frac{(2z)^{2m}}{(2m)!}\,\cos^{2k} z\,\mathrm{d}z = \frac{\pi}{2k}\,I_{m,k}\end{aligned}$$

in the notation of Lemma 6.4, hence

$$\hat{\xi}(m,n) = \sum_{k=1}^{\infty}\frac{1}{k^3} \sum_{k_1>\dots>k_m>k} \frac{1}{k_1^2\cdots k_m^2} \sum_{k>l_1>\dots>l_n\ge 1} \frac{1}{l_1^2\cdots l_n^2} = \zeta(\{2\}^m,3,\{2\}^n).$$

On the other hand, taking $z = \arcsin x$ in the integral (6.14) and using $\arcsin x + \arccos x = \pi/2$ for $0 \le x \le 1$, we find out that

$$\begin{aligned}\hat{\xi}(m,n) &= \frac{2}{\pi}\int_0^{\pi/2} \frac{(\pi-2z)^{2m+1}}{(2m+1)!}\,\frac{(2z)^{2n+2}}{(2n+2)!}\,\cot z\,\mathrm{d}z \\ &= \frac{2^{2m+2n+4}}{\pi(2m+1)!\,(2n+2)!}\int_0^{\pi/2} \Bigl(\frac{\pi}{2}-z\Bigr)^{2m+1} z^{2n+2}\,\cot z\,\mathrm{d}z.\end{aligned}$$

For the polynomial $P(z) = (\pi/2 - z)^{2m+1} z^{2n+2}$ we get $P(\pi/2) = 0$ and

$$\begin{aligned}P^{(2r)}\Bigl(\frac{\pi}{2}\Bigr) &= -(2r)!\binom{2n+2}{2r-2m-1}\Bigl(\frac{\pi}{2}\Bigr)^{2m+2n+3-2r}, \\ P^{(2r)}(0) &= (2r)!\binom{2m+1}{2r-2n-2}\Bigl(\frac{\pi}{2}\Bigr)^{2m+2n+3-2r}\end{aligned}$$

for $r = 1,\dots,m+n+1$. Therefore, Lemma 6.7 implies

$$\begin{aligned}\hat{\xi}(m,n) = \frac{2}{(2m+1)!\,(2n+2)!} \sum_{r=1}^{m+n+1} (-1)^{r-1}(2r)!\biggl(\binom{2n+2}{2r-2m-1}\Bigl(1-\frac{1}{2^{2r}}\Bigr) \\ -\binom{2m+1}{2r-2n-2}\biggr)\pi^{2m+2n+2-2r}\zeta(2r+1),\end{aligned}$$

which after a simple manipulation with the factorials and the use of (6.11) gives precisely the right-hand side of (6.12). $\square$

EXERCISE 6.7. 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.

REMARK. The second part of Theorem 6.3 (Exercise 6.7 above) gives rise to several other open questions [48].

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

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

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

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

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

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

EXERCISE 6.8 (open problem). Denote M_k the matrix from (6.12) expressing the vector $\{\xi(m,n) : m+n=k\}$ in terms of the vector $\{\zeta(2m+3)\xi(n) : m+n=k\}$, that is,

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

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

6.3. 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 6.8. *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*

$$\zeta(m,n) = (-1)^n \sum_{s=0}^{l-1} \left(\binom{k-2s-1}{m-1} + \binom{k-2s-1}{n-1} - \delta_{m,2s} + (-1)^n \delta_{s,0} \right) \times \zeta(2s)\zeta(k-2s). \tag{6.16}$$

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

$$\zeta(r)\zeta(s) = \zeta(r,s) + \zeta(s,r) + \zeta(k), \quad \text{where } r+s=k,\ r,s \geq 2, \tag{6.17}$$

$$\zeta(r)\zeta(s) = \sum_{m=2}^{k-1} \left(\binom{m-1}{r-1} + \binom{m-1}{s-1} \right) \zeta(m,k-m), \quad \text{where } r+s=k,\ r,s \geq 2, \tag{6.18}$$

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 (6.17) and (6.18) 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 (6.17) and (6.18) translate into equations

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

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

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

This is equivalent (because of $\zeta(0) = -\frac{1}{2}$) to (6.16). ∎

Either of the double shuffle relations (6.17) and (6.18) 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 (6.16), which simplifies to

$$\zeta(k-2r,2r) \equiv \sum_{s=1}^{l-1}\left(\binom{2l-2s}{2l-2r}+\binom{2l-2s}{2r-1}\right)\zeta(2s)\zeta(k-2s), \quad 1 \le r \le l-1, \tag{6.19}$$

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

THEOREM 6.9. *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 (6.19). 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 6.3 and 6.9 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*

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

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

EXERCISE 6.9 (open problem). For $k = 2l+1 \ge 3$ and the matrix $N = N_k$ defined above, show the following:

(a) $\det N = (-1)^l\, 1\cdot 3\cdot 5\cdots(2l-1)$; 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).

6.4. Hirose–Sato integrals

One can cast the two-one formula (6.3) in the form

$$\zeta^\star(\{2\}^{k_1},1,\{2\}^{k_2},1,\dots,\{2\}^{k_l},1)=\sum_{n_1\ge n_2\ge\cdots\ge n_l\ge 1}\frac{2^{\#\{n_1,n_2,\dots,n_l\}}}{n_1^{2k_1+1}n_2^{2k_2+1}\cdots n_l^{2k_l+1}}, \tag{6.20}$$

where $k_1\ge 1$, $k_2,\dots,k_l\ge 0$ and $\#\{n_1,n_2,\dots,n_l\}$ counts *distinct* elements in $\{n_1,n_2,\dots,n_l\}$. In fact, Zhao gives in [**50**] a general formula expressing *any* multiple zeta star value $\zeta^\star(\boldsymbol{s})$ in terms of the zeta-star looking sums

$$\zeta^\#(\boldsymbol{s})=\zeta^\#(s_1,\dots,s_l)=\sum_{n_1\ge\cdots\ge n_l\ge 1}(-1)^{n_1(s_1-1)+\cdots+n_l(s_l-1)}\frac{2^{\#\{n_1,\dots,n_l\}}}{n_1^{s_1}\cdots n_l^{s_l}}. \tag{6.21}$$

More precisely, he show that there is a bijection σ on the set of admissible indices such that

$$\zeta^\star(\boldsymbol{s})=\zeta^\#(\sigma\boldsymbol{s})\quad\text{for all }\boldsymbol{s}. \tag{6.22}$$

For the particular shape $\boldsymbol{s}=(\{2\}^{k_1},1,\{2\}^{k_2},1,\dots,\{2\}^{k_l},1)$ of an admissible index one gets $\sigma\boldsymbol{s}=(2k_1+1,2k_2+1,\dots,2k_l+1)$ implying the two-one formula (6.20). The general description of the bijection σ in [**50**] is somewhat sophisticated and requires introducing a lot of additional notation. In order to gain a clear vision of σ we follow the generalisation of Zhao's formula (6.22) given by Hirose and Sato in [**16**].

Given holomorphic 1-forms $\phi_1,\dots,\phi_k$ on an open connected domain $V\subset\mathbb{P}^1$ and a path $\gamma\colon[0,1]\to\mathbb{P}^1$ such that $\gamma(0,1)\subset V$, define

$$I_\gamma(x;\phi_1,\dots,\phi_k;y)=\int\cdots\int_{1>z_1>\cdots>z_k>0}\phi(\gamma(z_1))\cdots\phi(\gamma(z_k)),$$

where $x=\gamma(1)$ and $y=\gamma(0)$. We omit γ if it happens to be a straight line connecting x with y. For $z\in\mathbb{C}$, set

$$e_z(t)=\frac{\mathrm{d}t}{t-z}.$$

In these settings,

$$\zeta^\star(s_1,\dots,s_l)=(-1)^{s_1+\cdots+s_l}I(1;e_{-1}^{s_1-1}e_1\cdots e_{-1}^{s_{l-1}-1}e_1e_{-1}^{s_l-1}(e_1-e_{-1});\infty).$$

To produce an integral expression for $\zeta^\#(\boldsymbol{s})$, we assign the sequence $\varepsilon_0,\varepsilon_1,\dots,\varepsilon_l\in\{\pm1\}$ to the admissible index $\boldsymbol{s}=(s_1,\dots,s_l)$ by the rule

$$\varepsilon_0=1\quad\text{and}\quad\varepsilon_j=(-1)^{s_j-1}\varepsilon_{j-1}\quad\text{for }j=1,\dots,l.$$

Then

$$\zeta^\#(\boldsymbol{s})=(-1)^{s_1+\cdots+s_l}I(1;e_0^{s_1-1}(2e_{\varepsilon_1}-e_0)\cdots e_0^{s_{l-1}-1}(2e_{\varepsilon_{l-1}}-e_0)e_0^{s_l-1}(2e_{\varepsilon_l}-2e_0);\infty).$$

Let us first see how the two-one formula looks like, on a particular example $\zeta^\star(2,2,1,2,1)=\zeta^\#(5,3)$:

$$\begin{aligned} &I(1; e_{-1}e_1\, e_{-1}e_1\, e_1\, e_{-1}e_1\, (e_1-e_{-1});\infty) \\ &\quad = I(1; e_0e_0e_0e_0(2e_1-e_0)\, e_0e_0(2e_1-2e_0);\infty). \end{aligned} \tag{6.23}$$

In order to present the correspondence between the two expressions in a more transparent form, introduce auxiliary differential 1-forms

$$f_{1,-1}=f_{-1,1}=e_0, \quad f_{1,1}=2e_1-e_0 \quad \text{and} \quad f_{-1,-1}=2e_{-1}-e_0.$$

In this notation, the right-hand side of (6.23) can be stated as

$$I(1; f_{1,-1}f_{-1,1}f_{1,-1}f_{-1,1}f_{1,1}f_{1,-1}f_{-1,1}(f_{1,1}-f_{1,-1});\infty),$$

where the rule we choose to switch between $f_{1,-1}$ and $f_{-1,1}$ (which both correspond to the same e_0) is to force *alternation* of 1 and -1 in all consecutive indices, and we always start from $f_{1,-1}$. We are now ready to witness an explicit form of Zhao's idenity (6.22).

THEOREM 6.10 (Zhao's relations). *For any $\varepsilon_1,\dots,\varepsilon_k,\varepsilon_{k+1}\in\{\pm1\}$ with $\varepsilon_1\neq1$, we have*

$$\begin{aligned} &I(1; e_{\varepsilon_1}e_{\varepsilon_2}\cdots e_{\varepsilon_{k-1}}(e_{\varepsilon_k}-e_{\varepsilon_{k+1}});\infty) \\ &\quad = I(1; f_{1,\varepsilon_1}f_{\varepsilon_1,\varepsilon_2}\cdots f_{\varepsilon_{k-2},\varepsilon_{k-1}}(f_{\varepsilon_{k-1},\varepsilon_k}-f_{\varepsilon_{k-1},\varepsilon_{k+1}});\infty). \end{aligned} \tag{6.24}$$

Though Theorem 6.10 sounds general enough, its form suggests existence of an even more general result. Suppose that we can find some differential forms $\hat f_{z,w}$ for any $z,w\in\mathbb{C}$ such that

(i) $\hat f_{z,w}=f_{z,w}$ for $z,w\in\{\pm1\}$, and
(ii) the identity

$$\begin{aligned} &I(1; e_{z_1}e_{z_2}\cdots e_{z_{k-1}}(e_{z_k}-e_{z_{k+1}});\infty) \\ &\quad = I(1; \hat f_{1,z_1}\hat f_{z_1,z_2}\cdots \hat f_{z_{k-2},z_{k-1}}(\hat f_{z_{k-1},z_k}-\hat f_{z_{k-1},z_{k+1}});\infty). \end{aligned}$$

holds for any $z_1,z_2,\dots,z_k,z_{k+1}\in\mathbb{C}$ with $z_1\neq1$ (to ensure the convergence of either side!).

The naive choice $\hat f_{z,w}=2e_{(z+w)/2}-e_0$ clearly meets condition (i) but it fails to satisfy (ii).

EXERCISE 6.10. Prove that the differential 1-form

$$\hat f_{z,w}(u)=2\mathrm{d}\log\left(\sqrt{u^2-2zu+1}-\sqrt{u^2-2wu+1}\right)-\frac{\mathrm{d}u}{u}$$

satisfy both (i) and (ii) above.

With the differential form from Exercise 6.10 in mind, after the change of variable $t=(u+u^{-1})/2$, we can now give a general identity which implies Theorem 6.10.

Theorem 6.11 (Hirose–Sato [**16**]). *For $k \ge 0$ and $z_0, z_1, \dots, z_k, z_{k+1} \in \mathbb{C}$ with $z_0 \ne z_1$ and $z_k \ne z_{k+1}$, we have*

$$\begin{aligned} &I_\gamma(z_0; e_{z_1} e_{z_2} \cdots e_{z_{k-1}} (e_{z_k} - e_{z_{k+1}}); \infty) \\ &\quad = I_\gamma(z_0; g_{z_0,z_1} g_{z_1,z_2} \cdots g_{z_{k-2},z_{k-1}} (g_{z_{k-1},z_k} - g_{z_{k-1},z_{k+1}}); \infty), \end{aligned} \tag{6.25}$$

where

$$g_{z,w}(t) = \frac{\mathrm{d}t}{\sqrt{(t-z)(t-w)}}$$

and γ is a fixed path from z_0 to ∞ avoding the singularities of the integrand.

Sketch of proof. Denoting $f_k(z_0, z_1, \dots, z_k, z_{k+1})$ either side of (6.25), one routinely verifies that

$$\frac{\partial}{\partial z_1} f_1(z_0, z_1, z_2) = \frac{1}{z_1 - z_0}$$

and

$$\frac{\partial}{\partial z_k} f_k(z_0, z_1, \dots, z_k, z_{k+1}) = \frac{1}{z_k - z_{k-1}} f_{k-1}(z_0, z_1, \dots, z_k).$$

Integrating we determine the constant by setting $z_k = z_{k+1}$, in which case both sides vanish. ☐

The argument can be generalised even further by introducing, for complex α with $\operatorname{Re}\alpha > 0$, the differential 1-forms

$$g^\alpha_{z,w}(t) = \frac{\mathrm{d}t}{(t-z)^\alpha (t-w)^{1-\alpha}}.$$

Theorem 6.12 (Hirose–Sato [**16**]). *For $k \ge 0$ and $z_0, z_1, \dots, z_k, z_{k+1} \in \mathbb{C}$ with $z_0 \ne z_1$ and $z_k \ne z_{k+1}$, the integral*

$$I_\gamma(z_0; g^\alpha_{z_0,z_1} g^\alpha_{z_1,z_2} \cdots g^\alpha_{z_{k-2},z_{k-1}} (g^\alpha_{z_{k-1},z_k} - g^\alpha_{z_{k-1},z_{k+1}}); \infty) \tag{6.26}$$

does not depend on α.

In particular, the choice $\alpha = 1$ corresponds to the left-hand side of (6.25), while $\alpha = \frac{1}{2}$ is for the right-hand side of (6.25). Thus, Theorem 6.11 follows from Theorem 6.12.

The most general form of the Hirose–Sato theorem is the following Selberg-integral-type identity.

Theorem 6.13 (Hirose–Sato [**16**]). *Take*

$$g^{\alpha,\beta}_{z,w}(t) = \frac{\mathrm{d}t}{(t-z)^\alpha (t-w)^{1-\beta}}.$$

For $k \ge 0$ and collections of complex $(k+2)$-tuples $z_0, \dots, z_{k+1}$, $\alpha_0, \dots, \alpha_{k+1}$ and $\beta_0, \dots, \beta_{k+1}$ satisfying

$$\begin{gathered} z_0 \ne z_1,\ z_k \ne z_{k+1}, \quad \operatorname{Re}\alpha_0 > 0,\ \operatorname{Re}\beta_0 > 0, \quad \operatorname{Re}\alpha_{k+1} < 1,\ \operatorname{Re}\beta_{k+1} < 1, \\ \alpha_0 - \beta_0 = \alpha_1 - \beta_1 = \cdots = \alpha_k - \beta_k = \alpha_{k+1} - \beta_{k+1}, \end{gathered}$$

the following identity is valid:

$$\frac{I_\gamma(z_0; g_{z_0,z_1}^{\alpha_0,\alpha_1} g_{z_1,z_2}^{\alpha_1,\alpha_2} \cdots g_{z_k,z_{k+1}}^{\alpha_k,\alpha_{k+1}}; z_{k+1})}{I_\gamma(z_0; g_{z_0,z_{k+1}}^{\alpha_0,\alpha_{k+1}}; z_{k+1})} = \frac{I_\gamma(z_0; g_{z_0,z_1}^{\beta_0,\beta_1} g_{z_1,z_2}^{\beta_1,\beta_2} \cdots g_{z_k,z_{k+1}}^{\beta_k,\beta_{k+1}}; z_{k+1})}{I_\gamma(z_0; g_{z_0,z_{k+1}}^{\beta_0,\beta_{k+1}}; z_{k+1})}. \tag{6.27}$$

Choosing $\alpha_0 = \cdots = \alpha_{k+1} = \alpha \to 0$ and $\beta_0 = \cdots = \beta_{k+1} = \frac{1}{2}$ in Theorem 6.13 we recover Zagier's formula from Section 6.2.

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

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

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

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

If $s\ge2$ is even, then the series $E_s(q)=1-2s\zeta_q(s)/B_s$, where the Bernoulli numbers $B_s\in\mathbb{Q}$ are defined in 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 (Bailey [4, 54]). 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 both sides of this identity by $(1-q)^3$ and letting $q \to 1$, again recover 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

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

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

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

(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 [**34**], which includes numerous implications, is a convincing argument to count the q-MZVs (7.3) appropriate enough. Recall that the *height* $m = m(\boldsymbol{s})$ of a multi-index $\boldsymbol{s} = (s_1, \dots, s_l)$ is the number of components satisfying $s_j > 1$, so that $m(\boldsymbol{s}) \ge 1$ for an admissible indices $\boldsymbol{s}$. 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 $I(w, l, m)$, and set

$$\Phi_q(x, y, z) = \sum_{w,l,m=0}^{\infty} x^{w-l-m} y^{l-m} z^{2m-2} \sum_{\boldsymbol{s} \in I(w,l,m)} \zeta_q(\boldsymbol{s}).$$

THEOREM 7.1. *The generating function Φ_q is given by*

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

where α and β are determined by

$$\alpha + \beta = x + y + (q-1)(z^2 - xy), \quad \alpha\beta = z^2.$$

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 $q \to 1$ is Theorem 4.6 of Ohno and Zagier given in Section 4.3.

COROLLARY 1. *We have the generating function identity*

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

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 $z^2 \to xy$ in (7.4) gives

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

On the other hand, it follows directly from definition that

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

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

EXERCISE 7.4. For an indeterminate t, show

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

HINT. This is equivalent to the sum theorem in Corollary 2. $\square$

The q-model (7.3) also fits the q-version of Theorem 5.1.

THEOREM 7.2 (q-analogue of Ohno's relations [**7, 39**]). *Given a non-negative integer m, for an admissible index $\boldsymbol{s} = (s_1,\dots,s_l)$ and its dual $\boldsymbol{s}' = (s_1',\dots,s_k')$ (in the sense of Section 3.3), we have the identity*

$$\sum_{\substack{i_1,\dots,i_l\ge 0\\ i_1+\dots+i_l=m}} \zeta_q(s_1+i_1,\dots,s_l+i_l) = \sum_{\substack{i_1,\dots,i_k\ge 0\\ i_1+\dots+i_k=m}} \zeta_q(s_1'+i_1,\dots,s_k'+i_k).$$

EXERCISE 7.5. Prove Theorem 7.2.

HINT. Replace the Seki–Yamamoto 'connected' sums (5.1) by their q-analogues using the rules

$$\frac{1}{n^{s-1}(n-t)} \mapsto \frac{q^{n(s-1)}}{[n]_q^{s-1}([n]_q - tq^n)}$$

and

$$C(n,m;t) \mapsto C_q(n,m;t) = q^{nm} \frac{\prod_{j=1}^{n}([j]_q - tq^j)\cdot\prod_{j=1}^{m}([j]_q - tq^j)}{\prod_{j=1}^{n+m}([j]_q - tq^j)}. \qquad \square$$

In spite of the above naturalness of the q-MZVs (7.3), there exist 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

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

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

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

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

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

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

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

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 [**34**]. 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 $z=q$,

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

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

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

(cf. Exercise 3.12 (a)) in terms of (7.8) involve certain undesired 'parasites': if $r=2$, from

$$D_q\bigl(\operatorname{Li}_1(z;q)\operatorname{Li}_1(z;q)\bigr) = \frac{1}{1-z}\operatorname{Li}_1(z;q) + \operatorname{Li}_1(z;q)\frac{1}{1-z} - (1-q)\frac{z}{(1-z)^2}$$

we have

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

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 (6.17), (6.18)), 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

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

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

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

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):

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

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

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

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

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

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

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

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

$$\begin{aligned}\sum_{k=0}^\infty F_k(\varepsilon)x^k &= \frac{1}{1-(1-e^{-\varepsilon x})/\varepsilon}\\ &= 1+x+\left(-\frac12\varepsilon+1\right)x^2+\left(\frac16\varepsilon^2-\varepsilon+1\right)x^3\\ &\qquad+\left(-\frac{1}{24}\varepsilon^3+\frac{7}{12}\varepsilon^2-\frac32\varepsilon+1\right)x^4\\ &\qquad+\left(\frac{1}{120}\varepsilon^4-\frac14\varepsilon^3+\frac54\varepsilon^2-2\varepsilon+1\right)x^5+\cdots\end{aligned}$$

and

$$\sum_{s=0}^\infty \hat\lambda_s x^s = -\frac{xe^x}{1-e^x} = 1+\frac12 x+\sum_{k=1}^\infty \frac{B_{2k}}{(2k)!}x^{2k}$$

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

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

$$\begin{aligned}\frac12+\frac{\hat\rho_1(q^n)}{1-q^n} &= \frac1n\cdot\left(\varepsilon^{-1}-\frac12\right)+O(\varepsilon),\\ \frac{1}{12}+\frac{\hat\rho_2(q^n)}{(1-q^n)^2} &= \frac{1}{n^2}\cdot\left(\varepsilon^{-2}-\varepsilon^{-1}+\frac{1}{12}\right)+O(\varepsilon),\\ \frac{\hat\rho_3(q^n)}{(1-q^n)^3} &= \frac{1}{n^3}\cdot\left(\varepsilon^{-3}-\frac32\varepsilon^{-2}+\frac12\varepsilon^{-1}\right)+O(\varepsilon),\\ -\frac{1}{720}+\frac{\hat\rho_4(q^n)}{(1-q^n)^4} &= \frac{1}{n^4}\cdot\left(\varepsilon^{-4}-2\varepsilon^{-3}+\frac76\varepsilon^{-2}-\frac16\varepsilon^{-1}-\frac{1}{720}\right)+O(\varepsilon),\end{aligned}$$

and so on.

PROPOSITION 7.3. *Assume that $s_1>r_1+1$ and $s_j\ge r_j+1$ for $j=2,\dots,l$. Then*

$$\begin{bmatrix}s_1,\dots,s_l\\ r_1,\dots,r_l\end{bmatrix}\sim\frac{\zeta(s_1-r_1,s_2-r_2,\dots,s_l-r_l)}{r_1!\,r_2!\cdots r_l!}\,\frac{1}{(1-q)^{s_1+s_2+\cdots+s_l}}\qquad \textit{as } q\to1^-,$$

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 7.4. *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

$$\begin{aligned} z_{s_1,r_1}\diamond z_{s_2,r_2}=\binom{r_1+r_2-2}{r_1-1}\Bigg(&z_{s_1+s_2,r_1+r_2-1}\\ &+\sum_{j=1}^{s_1}(-1)^{s_2-1}\binom{s_1+s_2-j-1}{s_1-j}\lambda_{s_1+s_2-j}z_{j,r_1+r_2-1}\\ &+\sum_{j=1}^{s_2}(-1)^{s_1-1}\binom{s_1+s_2-j-1}{s_2-j}\lambda_{s_1+s_2-j}z_{j,r_1+r_2-1}\Bigg), \end{aligned} \tag{7.15}$$

where

$$\sum_{s=0}^{\infty}\lambda_sx^s=-\frac{x}{1-e^x}=1+\sum_{s=1}^{\infty}\frac{B_s}{s!}x^s$$

is the generating function of Bernoulli numbers. Note that $\hat\lambda_s=\lambda_s$ for $s\ge2$, while $\hat\lambda_1=\frac12=-\lambda_1$ in the notation of Exercise 7.6.

EXERCISE 7.7. 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

$$aw \sqcap\!\sqcap\, bv = a(w \sqcap\!\sqcap\, bv) + b(aw \sqcap\!\sqcap\, v) + (a \diamond b)(w \sqcap\!\sqcap\, v), \tag{7.16}$$

for arbitrary $w, v \in \mathbb{Q}\langle Z\rangle$ and $a, b \in Z$.

PROPOSITION 7.5. *The evaluation map*

$$[\,\cdot\,]\colon z_{s_1,r_1}\cdots z_{s_l,r_l} \mapsto \begin{bmatrix} s_1,\dots,s_l \\ r_1-1,\dots,r_l-1 \end{bmatrix} \tag{7.17}$$

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

$$\begin{aligned}
&\frac{n^{r_1-1}\hat\rho_{s_1}(q^n)}{(s_1-1)!\,(r_1-1)!\,(1-q^n)^{s_1}}\cdot\frac{n^{r_2-1}\hat\rho_{s_2}(q^n)}{(s_2-1)!\,(r_2-1)!\,(1-q^n)^{s_2}} \\
&= \binom{r_1+r_2-2}{r_1-1}\frac{n^{r_1+r_2-2}}{(r_1+r_2-2)!}\biggl(\frac{\hat\rho_{s_1+s_2}(q^n)}{(s_1+s_2-1)!\,(1-q^n)^{s_1+s_2}} \\
&\qquad + \sum_{j=1}^{s_1}(-1)^{s_2-1}\binom{s_1+s_2-j-1}{s_1-j}\lambda_{s_1+s_2-j}\frac{\hat\rho_j(q^n)}{(j-1)!\,(1-q^n)^j} \\
&\qquad + \sum_{j=1}^{s_2}(-1)^{s_1-1}\binom{s_1+s_2-j-1}{s_2-j}\lambda_{s_1+s_2-j}\frac{\hat\rho_j(q^n)}{(j-1)!\,(1-q^n)^j}\biggr). \qquad\square
\end{aligned}$$

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 7.3 results in the following statement.

THEOREM 7.6. *For admissible words $w = z_{s_1}\cdots z_{s_l}$ and $v = z_{s'_1}\cdots z_{s'_m}$ of weight $|w| = s_1+\dots+s_l$ and $|v| = s'_1+\dots+s'_m$, respectively,*

$$[w \sqcap\!\sqcap\, v] \sim (1-q)^{-|w|-|v|}\zeta(w * v) \quad \textit{as } q \to 1^-,$$

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 $[w \sqcap v] = [w] \cdot [v]$, Theorem 7.6 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 7.8. Show that

$$\begin{aligned}
&\begin{bmatrix} s_1, & s_2, & \dots, & s_l \\ r_1-1, & r_2-1, & \dots, & r_l-1 \end{bmatrix} \\
&= \sum_{j_2=1}^{r_1+r_2-1} \binom{j_2-1}{r_2-1} \sum_{j_3=1}^{j_2+r_3-1} \binom{j_3-1}{r_3-1} \cdots \sum_{j_l=1}^{j_{l-1}+r_l-1} \binom{j_l-1}{r_l-1} \\
&\quad \times \mathfrak{z}\begin{bmatrix} s_1, & s_2, & \dots, & s_{l-1}, & s_l \\ r_1+r_2-j_2, & j_2+r_3-j_3, & \dots, & j_{l-1}+r_l-j_l, & j_l \end{bmatrix}
\end{aligned}$$

and

$$\mathfrak{Z}\begin{bmatrix} s_1,\dots,s_l \\ r_1,\dots,r_l \end{bmatrix} = \sum_{i_1=1}^{r_1}\sum_{i_2=1}^{r_2}\cdots\sum_{i_{l-1}=1}^{r_{l-1}}(-1)^{r_1+\dots+r_{l-1}-i_1-\dots-i_{l-1}}$$
$$\times\binom{r_1-i_1+i_2-1}{r_1-i_1}\cdots\binom{r_{l-2}-i_{l-2}+i_{l-1}-1}{r_{l-2}-i_{l-2}}\binom{r_{l-1}-i_{l-1}+r_l-1}{r_{l-1}-i_{l-1}}$$
$$\times\begin{bmatrix} s_1, & s_2, & \dots, & s_{l-1}, & s_l \\ i_1-1, & r_1-i_1+i_2-1, & \dots, & r_{l-2}-i_{l-2}+i_{l-1}-1, & r_{l-1}-i_{l-1}+r_l-1 \end{bmatrix}.$$

Exercise 7.8 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 7.7 (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 $m_j' = d_{l+1-j}$. $\square$

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 7.7 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 7.5 and 7.7 that

PROPOSITION 7.8. *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

$$\begin{aligned}
[w \,\overline{\sqcap}\, v] &= [\varphi^{-1}\tau\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v)] \\
&= \mathfrak{Z}[\tau\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v)] = \mathfrak{Z}[\varphi(\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v)] \\
&= [\varphi^{-1}\tau\varphi w \sqcap \varphi^{-1}\tau\varphi v] = [\varphi^{-1}\tau\varphi w]\cdot[\varphi^{-1}\tau\varphi v] \\
&= \mathfrak{Z}[\tau\varphi w]\cdot\mathfrak{Z}[\tau\varphi v] = \mathfrak{Z}[\varphi w]\cdot\mathfrak{Z}[\varphi v] = [w]\cdot[v]. \qquad \square
\end{aligned}$$

Note that (7.18) is also equivalent to the expansion from the right (this is established in Exercise 3.15):

$$\begin{aligned}
wz_{s_1,r_1} \sqcap vz_{s_2,r_2} \equiv (w \sqcap vz_{s_2,r_2})z_{s_1,r_1} &+ (wz_{s_1,r_1} \sqcap v)z_{s_2,r_2} \\
&+ \binom{r_1+r_2-2}{r_1-1}(w \sqcap v)z_{s_1+s_2,r_1+r_2-1}.
\end{aligned} \tag{7.22}$$

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 7.9. *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

$$wz_{s_1,r_1} \sqcap vz_{s_2,r_2} \equiv (w \sqcap vz_{s_2,r_2})z_{s_1,r_1} + (wz_{s_1,r_1} \sqcap v)z_{s_2,r_2},$$

and that we also have

$$\begin{aligned}
wz_{s_1,r_1+1} \sqcap vz_{s_2,r_2} &\equiv (wz_{s_1,r_1} \sqcap vz_{s_2,r_2})x_1 + (wz_{s_1,r_1+1} \sqcap v)z_{s_2,r_2}, \\
wz_{s_1,r_1+1} \sqcap vz_{s_2,r_2+1} &\equiv (wz_{s_1,r_1} \sqcap vz_{s_2,r_2+1})x_1 + (wz_{s_1,r_1+1} \sqcap vz_{s_2,r_2})x_1.
\end{aligned}$$

The relations already show that

$$wa' \sqcap vb' \equiv (w \sqcap vb')a' + (wa' \sqcap v)b' \tag{7.24}$$

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 7.8) acts trivially on the words from $\mathbb{Q}\langle Z'\rangle$. Therefore, applying $\tau\varphi$ to both sides of (7.21) and extracting the homogeneous part of the result corresponding to the highest weight and length we arrive at

$$\tau(w \,\overline{\sqcap}\, v) \equiv \tau w \sqcap \tau v \quad \text{for all } w, v \in \mathbb{Q}\langle Z'\rangle.$$

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

$$\begin{aligned} \tau(aw \,\overline{\sqcap}\, bv) &\equiv \tau(aw) \sqcap \tau(bv) \equiv (\tau w)\overline{a} \sqcap (\tau v)\overline{b} \\ &\equiv (\tau w \sqcap (\tau v)\overline{b})\overline{a} + ((\tau w)\overline{a} \sqcap \tau v)\overline{b} \\ &\equiv (\tau w \sqcap \tau(bv))\overline{a} + (\tau(aw) \sqcap \tau v)\overline{b} \equiv (\tau(w \,\overline{\sqcap}\, bv))\overline{a} + (\tau(aw \,\overline{\sqcap}\, v))\overline{b} \\ &\equiv \tau(a(w \,\overline{\sqcap}\, bv) + b(aw \,\overline{\sqcap}\, v)), \end{aligned}$$

which implies the desired result. □

Theorem 7.10. *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,*

$$[w \,\overline{\sqcap}\, v] \sim (1-q)^{-|w|-|v|}\zeta(w \sqcup\!\sqcup v) \qquad \textit{as} \quad q \to 1^-,$$

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 7.9 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 $r_j \ge 2$ for at least one j. The latter circumstance and Proposition 7.3 then imply

$$\lim_{q \to 1^-} (1-q)^{|w|+|v|}[z_{q_1,r_1} \cdots z_{q_n,r_n}] = 0. \qquad \square$$

Theorem 7.10 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 7.9 we obtain

Theorem 7.11. *For two words $w = z_{s_1} \cdots z_{s_l}$ and $v = z_{s'_1} \cdots z_{s'_m}$, not necessarily admissible,*

$$[w \sqcap v - w \,\overline{\sqcap}\, v] \sim (1-q)^{-|w|-|v|}\zeta(w * v - w \sqcup\!\sqcup v) \qquad \textit{as} \quad q \to 1^-,$$

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 7.12 (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 $x_0^{s_1}x_1x_0^{s_2}x_1\cdots x_0^{s_l}x_1$, so that they become $x_0^{s_1-1}x_1x_0^{s_2-1}x_1\cdots x_0^{s_l-1}x_1$. 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 7.13 (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

$$\zeta(s_1, s_2, \dots, s_l) = \zeta(x_0^{s_1-1}x_1x_0^{s_2-1}x_1\cdots x_0^{s_l-1}x_1)$$

with all s_j to be at least 2 (so that there is no appearance of x_1^r with $r \ge 2$). 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 $\mathfrak{Z}\left[\begin{smallmatrix}2\\2\end{smallmatrix}\right]$ and we find out that

$$\mathfrak{Z}\left[\begin{smallmatrix}2\\2\end{smallmatrix}\right] = \tfrac{1}{2}\mathfrak{Z}\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right] + \mathfrak{Z}\left[\begin{smallmatrix}3\\1\end{smallmatrix}\right] - \mathfrak{Z}\left[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\right].$$

There are 34 totally q-zeta brackets of weight up to 4,

$$\mathfrak{Z}[\ ]^*,\ \mathfrak{Z}\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]^*,\ \mathfrak{Z}\left[\begin{smallmatrix}2\\1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}2\\2\end{smallmatrix}\right]^*,\ \mathfrak{Z}\left[\begin{smallmatrix}3\\1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1\\3\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}3\\2\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}2\\3\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}4\\1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1\\4\end{smallmatrix}\right],$$

$$\mathfrak{Z}\left[\begin{smallmatrix}1,1\\1,1\end{smallmatrix}\right]^*,\ \mathfrak{Z}\left[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1\\1,2\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}1,2\\1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1\\2,1\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}2,1\\2,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,2\\1,2\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}2,1\\1,2\end{smallmatrix}\right]^*,\ \mathfrak{Z}\left[\begin{smallmatrix}1,2\\2,1\end{smallmatrix}\right]^*,$$

$$\mathfrak{Z}\left[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1\\2,2\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}3,1\\1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1\\1,3\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}1,3\\1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1\\3,1\end{smallmatrix}\right],$$

$$\mathfrak{Z}\left[\begin{smallmatrix}1,1,1\\1,1,1\end{smallmatrix}\right]^*,\ \mathfrak{Z}\left[\begin{smallmatrix}2,1,1\\1,1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1,1\\1,1,2\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}1,2,1\\1,1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1,1\\1,2,1\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}1,1,2\\1,1,1\end{smallmatrix}\right] = \mathfrak{Z}\left[\begin{smallmatrix}1,1,1\\2,1,1\end{smallmatrix}\right],\ \mathfrak{Z}\left[\begin{smallmatrix}1,1,1,1\\1,1,1,1\end{smallmatrix}\right]^*,$$

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 $\mathfrak{Z}\bigl[\begin{smallmatrix}2\\2\end{smallmatrix}\bigr]$ these are $\mathfrak{Z}\bigl[\begin{smallmatrix}3\\2\end{smallmatrix}\bigr]$, $\mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\2,1\end{smallmatrix}\bigr]$, $\mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\1,2\end{smallmatrix}\bigr]$ and $\mathfrak{Z}\bigl[\begin{smallmatrix}1,2\\2,1\end{smallmatrix}\bigr]$. We find out that

$$
\begin{aligned}
\mathfrak{Z}\bigl[\begin{smallmatrix}3\\2\end{smallmatrix}\bigr] &= \tfrac14\mathfrak{Z}\bigl[\begin{smallmatrix}2\\1\end{smallmatrix}\bigr] + \tfrac32\mathfrak{Z}\bigl[\begin{smallmatrix}4\\1\end{smallmatrix}\bigr] - 2\mathfrak{Z}\bigl[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\bigr],\\
\mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\2,1\end{smallmatrix}\bigr] &= \mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\bigr] + \tfrac12\mathfrak{Z}\bigl[\begin{smallmatrix}1,2\\1,1\end{smallmatrix}\bigr] - \mathfrak{Z}\bigl[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\bigr] + \mathfrak{Z}\bigl[\begin{smallmatrix}1,3\\1,1\end{smallmatrix}\bigr] - \mathfrak{Z}\bigl[\begin{smallmatrix}2,1,1\\1,1,1\end{smallmatrix}\bigr] - \mathfrak{Z}\bigl[\begin{smallmatrix}1,2,1\\1,1,1\end{smallmatrix}\bigr],\\
\mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\1,2\end{smallmatrix}\bigr] &= -\tfrac12\mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\bigr] - \tfrac12\mathfrak{Z}\bigl[\begin{smallmatrix}1,2\\1,1\end{smallmatrix}\bigr] + 2\mathfrak{Z}\bigl[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\bigr] + \mathfrak{Z}\bigl[\begin{smallmatrix}3,1\\1,1\end{smallmatrix}\bigr] - \mathfrak{Z}\bigl[\begin{smallmatrix}1,3\\1,1\end{smallmatrix}\bigr] + \mathfrak{Z}\bigl[\begin{smallmatrix}1,2,1\\1,1,1\end{smallmatrix}\bigr],\\
\mathfrak{Z}\bigl[\begin{smallmatrix}1,2\\2,1\end{smallmatrix}\bigr] &= -\mathfrak{Z}\bigl[\begin{smallmatrix}2,1\\1,1\end{smallmatrix}\bigr] + 2\mathfrak{Z}\bigl[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\bigr] + \mathfrak{Z}\bigl[\begin{smallmatrix}2,1,1\\1,1,1\end{smallmatrix}\bigr],
\end{aligned}
$$

and there is one more relation in this weight between the q-zeta brackets from $\mathcal{MD}$:

$$\tfrac13\mathfrak{Z}\bigl[\begin{smallmatrix}2\\1\end{smallmatrix}\bigr] - \mathfrak{Z}\bigl[\begin{smallmatrix}3\\1\end{smallmatrix}\bigr] + \mathfrak{Z}\bigl[\begin{smallmatrix}4\\1\end{smallmatrix}\bigr] - 2\mathfrak{Z}\bigl[\begin{smallmatrix}2,2\\1,1\end{smallmatrix}\bigr] + 2\mathfrak{Z}\bigl[\begin{smallmatrix}3,1\\1,1\end{smallmatrix}\bigr] = 0.$$

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 7.13 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 over $\mathbb{C}(q)$ seem to be always liftable to relations over $\mathbb{Q}$.

EXERCISE 7.9 (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 $\mathbb{Q}$.

7.4. q-Differentiation of brackets

LEMMA 7.14 ([2, Theorem 3.1]). *The generating series*

$$T(x_1, \dots, x_l) = \sum_{s_1,\dots,s_l>0} [s_1, \dots, s_l] x_1^{s_1-1} \cdots x_l^{s_l-1}$$

of brackets of length l can be written as

$$T(x_1,\dots,x_l)=\sum_{d_1,\dots,d_l>0}\prod_{j=1}^{l}\frac{e^{d_jx_j}q^{d_1+\cdots+d_j}}{1-q^{d_1+\cdots+d_j}}.$$

PROOF. Indeed,

$$\begin{aligned}T(x_1,\dots,x_l)&=\sum_{s_1,\dots,s_l>0}\sum_{\substack{n_1>\cdots>n_l>0\\ d_1,\dots,d_l>0}}\frac{(d_1x_1)^{s_1-1}\cdots(d_lx_l)^{s_l-1}}{(s_1-1)!\cdots(s_l-1)!}q^{n_1d_1+\cdots+n_ld_l}\\ &=\sum_{\substack{n_1>\cdots>n_l>0\\ d_1,\dots,d_l>0}}e^{d_1x_1+\cdots+d_lx_l}q^{n_1d_1+\cdots+n_ld_l}\\ &=\sum_{\substack{d_1,\dots,d_l>0\\ m_1,\dots,m_l>0}}e^{d_1x_1+\cdots+d_lx_l}q^{(m_1+\cdots+m_l)d_1+(m_2+\cdots+m_l)d_2+\cdots+m_ld_l}\\ &=\sum_{d_1,\dots,d_l>0}\prod_{j=1}^{l}e^{d_jx_j}\sum_{m_j>0}q^{m_j(d_1+\cdots+d_j)}.\end{aligned}$$

□

THEOREM 7.15 ([2, Theorem 1.7]). *The $\mathbb{Q}$-algebra $\mathcal{MD}$ is closed under the derivative $\delta=q\frac{\mathrm{d}}{\mathrm{d}q}$.*

PROOF. We start from observing that

$$\delta\left(\frac{q^d}{1-q^d}\right)=\frac{dq^d}{(1-q^d)^2}$$

implying

$$\begin{aligned}\delta T(x_1,\dots,x_l)&=\sum_{d_1,\dots,d_l>0}e^{d_1x_1+\cdots+d_lx_l}\frac{q^{d_1}}{1-q^{d_1}}\cdots\frac{q^{d_1+\cdots+d_l}}{1-q^{d_1+\cdots+d_l}}\sum_{j=1}^{l}\frac{d_1+\cdots+d_j}{1-q^{d_1+\cdots+d_l}}\\ &=\frac{\partial}{\partial x}\sum_{d_1,\dots,d_l>0}e^{d_1x_1+\cdots+d_lx_l}\frac{q^{d_1}}{1-q^{d_1}}\cdots\frac{q^{d_1+\cdots+d_l}}{1-q^{d_1+\cdots+d_l}}\sum_{j=1}^{l}\frac{e^{(d_1+\cdots+d_j)x}}{1-q^{d_1+\cdots+d_l}}\bigg|_{x=0}.\end{aligned}$$

The product of two generating series

$$T(x)T(x_1,\dots,x_l)=\sum_{d,d_1,\dots,d_l>0}e^{dx+d_1x_1+\cdots+d_lx_l}\frac{q^d}{1-q^d}\frac{q^{d_1}}{1-q^{d_1}}\cdots\frac{q^{d_1+\cdots+d_l}}{1-q^{d_1+\cdots+d_l}}$$

can be viewed as shuffling of single and l-tuple sums. We split the sum over d into the sums

$$\Sigma_j=\sum_{d_1,\dots,d_l>0}\sum_{d_1+\cdots+d_j<d<d_1+\cdots+d_{j+1}}e^{dx+d_1x_1+\cdots+d_lx_l}\frac{q^d}{1-q^d}\frac{q^{d_1}}{1-q^{d_1}}\cdots\frac{q^{d_1+\cdots+d_l}}{1-q^{d_1+\cdots+d_l}}$$

for $j = 0, 1, \dots, l$, where the internal sums are $\sum_{0<d<d_1}$ and $\sum_{d>d_1+\dots+d_l}$ for $j = 0$ and $j = l$, respectively, and the sums

$$\begin{aligned}\Sigma_j' &= \sum_{d_1,\dots,d_l>0} \sum_{d=d_1+\dots+d_j} e^{dx+d_1x_1+\dots+d_lx_l} \frac{q^d}{1-q^d} \frac{q^{d_1}}{1-q^{d_1}} \cdots \frac{q^{d_1+\dots+d_l}}{1-q^{d_1+\dots+d_l}} \\ &= \sum_{d_1,\dots,d_l>0} e^{d_1x_1+\dots+d_lx_l} \frac{q^{d_1}}{1-q^{d_1}} \cdots \frac{q^{d_1+\dots+d_l}}{1-q^{d_1+\dots+d_l}} \frac{e^{(d_1+\dots+d_j)x} q^{d_1+\dots+d_l}}{1-q^{d_1+\dots+d_l}}\end{aligned}$$

for $j = 1, \dots, l$. Writing $d = d_1 + \dots + d_j + \hat{d}$ and $d_{j+1} = \hat{d} + \hat{d}_{j+1}$, we see that

$$\Sigma_j = T(x + x_1, \dots, x + x_{j+1}, x_{j+1}, \dots, x_l) \quad \text{for } j = 0, 1, \dots, d-1$$

and $\Sigma_l = T(x + x_1, \dots, x + x_l, x)$. For the second group of sums we use $q^d = 1 - (1 - q^d)$ to write them as

$$\begin{aligned}\Sigma_j' = \sum_{d_1,\dots,d_l>0} e^{d_1x_1+\dots+d_lx_l} \frac{q^{d_1}}{1-q^{d_1}} \cdots \frac{q^{d_1+\dots+d_l}}{1-q^{d_1+\dots+d_l}} \frac{e^{(d_1+\dots+d_j)x}}{1-q^{d_1+\dots+d_l}} \\ - T(x + x_1, \dots, x + x_j, x_{j+1}, \dots, x_l)\end{aligned}$$

for $j = 1, \dots, l$. It follows then that

$$\begin{aligned}&\sum_{d_1,\dots,d_l>0} e^{d_1x_1+\dots+d_lx_l} \frac{q^{d_1}}{1-q^{d_1}} \cdots \frac{q^{d_1+\dots+d_l}}{1-q^{d_1+\dots+d_l}} \sum_{j=1}^{l} \frac{e^{(d_1+\dots+d_j)x}}{1-q^{d_1+\dots+d_l}} \\ &= T(x)T(x_1, \dots, x_l) + \sum_{j=1}^{l} T(x + x_1, \dots, x + x_j, x_{j+1}, \dots, x_l) \\ &\quad - \sum_{j=1}^{l} T(x + x_1, \dots, x + x_j, x_j, x_{j+1}, \dots, x_l) - T(x + x_1, \dots, x + x_l, x).\end{aligned}$$

Using $T(x) = [1] + [2]x + \cdots$, differentiating both side of the identity and substituting $x = 0$ we obtain

$$\begin{aligned}\delta T(x_1, \dots, x_l) = [2]T(x_1, \dots, x_l) + \sum_{j=1}^{l} \sum_{k=1}^{j} \frac{\partial T(x_1, \dots, x_l)}{\partial x_k} \\ - \sum_{j=1}^{l} \sum_{k=1}^{j} \frac{\partial T(z_1, \dots, z_{l+1})}{\partial z_k} \bigg|_{(z_1,\dots,z_{l+1})=(x_1,\dots,x_j,x_j,\dots,x_l)} \\ - \sum_{k=1}^{l+1} \frac{\partial T(z_1, \dots, z_{l+1})}{\partial z_k} \bigg|_{(z_1,\dots,z_{l+1})=(x_1,\dots,x_l,0)}.\end{aligned}$$

It remains to take into account that the coefficients in the (multi-variable) Taylor expansion of *any* order partial derivative of $T(x_1, \dots, x_l)$ are integer multiples of mono-brackets. $\square$

Now we turn our attention to bi-brackets (7.14). A straightforward computation implies the following result.

LEMMA 7.16. *We have*

$$\delta\begin{bmatrix}s_1,\dots,s_l\\ r_1,\dots,r_l\end{bmatrix}=\sum_{j=1}^{l}s_j(r_j+1)\begin{bmatrix}s_1,\dots,s_{j-1},s_j+1,s_{j+1},\dots,s_l\\ r_1,\dots,r_{j-1},r_j+1,r_{j+1},\dots,r_l\end{bmatrix}.$$

As a corollary of Lemma 7.16 and Theorem 7.15 we deduce the following.

THEOREM 7.17. *Any bi-bracket* $\left[\begin{smallmatrix}s\\ r\end{smallmatrix}\right]$ *of length* 1 *is an element of* $\mathbb{Q}$*-algebra* $\mathcal{MD}$.

PROOF. Since

$$\begin{bmatrix}s+1\\ r\end{bmatrix}=\sum_{n,d>0}n^r d^s q^{nd}=\begin{bmatrix}r+1\\ s\end{bmatrix},$$

we assume without loss of generality that the entries of our bi-bracket $\left[\begin{smallmatrix}s\\ r\end{smallmatrix}\right]$ under consideration satisfy $s>r$. Then a repeated application of Lemma 7.16 implies

$$\begin{bmatrix}s\\ r\end{bmatrix}=\frac{1}{(s-1)r}\,\delta\begin{bmatrix}s-1\\ r-1\end{bmatrix}=\dots=\frac{1}{(s-1)(s-2)\cdots(s-r)\,r!}\,\delta^r\begin{bmatrix}s-r\\ 0\end{bmatrix},$$

and the result follows from noticing that the mono-bracket $\left[\begin{smallmatrix}s-r\\ 0\end{smallmatrix}\right]=[s-r]$ and all its δ-derivatives are in $\mathcal{MD}$ by Theorem 7.15. $\square$

Bibliography

[1] H. Bachmann, Multiple Eisenstein series and q-analogues of multiple zeta values, in *Periods in Quantum Field Theory and Arithmetic* (RTMZV 2014, Madrid, Spain), J. I. Burgos Gil, K. Ebrahimi-Fard and H. Gangl (eds.), Springer Proc. Math. Stat. **314** (Springer, 2020), 173–235.

[2] H. Bachmann and U. Kühn, The algebra of generating functions for multiple divisor sums and applications to multiple zeta values, *Ramanujan J.* **40** (2016), no. 3, 605–648.

[3] H. Bachmann and U. Kühn, A dimension conjecture for q-analogues of multiple zeta values, in *Periods in Quantum Field Theory and Arithmetic* (RTMZV 2014, Madrid, Spain), J. I. Burgos Gil, K. Ebrahimi-Fard and H. Gangl (eds.), Springer Proc. Math. Stat. **314** (Springer, 2020), 237–258.

[4] W. N. Bailey, An algebraic identity, *J. London Math. Soc.* **11** (1936), 156–160.

[5] J. M. Borwein, D. M. Bradley and D. J. Broadhurst, Evaluations of k-fold Euler/Zagier sums: a compendium of results for arbitrary k, *Electron. J. Combin.* **4** (1997), no. 2, #R5.

[6] J. M. Borwein, D. M. Bradley, D. J. Broadhurst and P. Lisoněk, Special values of multiple polylogarithms, *Trans. Amer. Math. Soc.* **353** (2001), no. 3, 907–941.

[7] D. M. Bradley, Multiple q-zeta values, *J. Algebra* **283** (2005), no. 2, 752–798.

[8] F. Brown, Mixed Tate motives over $\mathbb{Z}$, *Ann. of Math.* (2) **175** (2012), 949–976.

[9] J. I. Burgos Gil, J. Fresán, with contributions by U. Kühn, *Multiple zeta values: from numbers to motives*, Clay Mathematics Proceedings (in press).

[10] P. Deligne and A. B. Goncharov, Groupes fondamentaux motiviques de Tate mixte, *Ann. Sci. École Norm. Sup.* (4) **38** (2005), no. 1, 1–56.

[11] C. Glanois, Motivic unipotent fundamental groupoid of $\mathbb{G}_m \setminus \mu_N$ for $N = 2, 3, 4, 6, 8$ and Galois descents, *J. Number Theory* **160** (2016), 334–384.

[12] Kh. Hessami Pilehrood and T. Hessami Pilehrood, An alternative proof of a theorem of Zagier, *J. Math. Anal. Appl.* **449** (2017), no. 1, 168–175.

[13] Kh. Hessami Pilehrood and T. Hessami Pilehrood, Generating functions for multiple zeta star values, *J. Théorie Nombres Bordeaux* **31** (2019), no. 2, 343–360.

[14] M. Hirose, T. Matsusaka and S. Seki, A discretization of the iterated integral expression of the multiple polylogarithm, *Preprint* `arXiv:2404.15210 [math.NT]`.

[15] M. Hirose and N. Sato, Hoffman's conjectural identity, *Intern. J. Number Theory* **15** (2019), no. 1, 167–171.

[16] M. Hirose and N. Sato, A new proof of the two-one formula, *in progress* (2020).

[17] M. E. Hoffman, Multiple zeta values, Michael Hoffman's homepage; also a comprehensive list of references on MZVs and related stuff.

[18] M. E. Hoffman, Multiple harmonic series, *Pacific J. Math.* **152** (1992), no. 2, 275–290.

[19] M. E. Hoffman, The algebra of multiple harmonic series, *J. Algebra* **194** (1997), no. 2, 477–495.

[20] M. E. Hoffman, Quasi-shuffle products, *J. Algebraic Combin.* **11** (2000), no. 1, 49–68.

[21] M. E. Hoffman and Y. Ohno, Relations of multiple zeta values and their algebraic expression, *J. Algebra* **262** (2003), no. 2, 332–347.

[22] K. Ihara, J. Kajikawa, Y. Ohno and J. Okuda, Multiple zeta values vs. multiple zeta-star values, *J. Algebra* **332** (2011), 187–208.

[23] K. Ihara and M. Kaneko, Derivation relations and regularized double shuffle relations of multiple zeta values, *Preprint* (2000).

[24] K. Ihara, M. Kaneko and D. Zagier, Derivation and double shuffle relations for multiple zeta values, *Compos. Math.* **142** (2006), no. 2, 307–338.

[25] L. Lai, C. Lupu and D. Orr, Elementary proofs of Zagier's formula for multiple zeta values and its odd variant, *Preprint* arXiv:2201.09262 [math.NT].

[26] T. Lee-Peng, Alternating double Euler sums, hypergeometric identities and a theorem of Zagier, *J. Math. Anal. Appl.* **462** (2018), 777–800.

[27] Z-H. Li, Another proof of Zagier's evaluation formula of the multiple zeta values $\zeta(2,\dots,2,3,2,\dots,2)$, *Math. Res. Lett.* **20** (2013), 947–950.

[28] T. Maesaka, S. Seki and T. Watanabe, Deriving two dualities simultaneously from a family of identities for multiple harmonic sums. arXiv:2402.05730 *Preprint* arXiv:2402.05730 [math.NT].

[29] Y. Ohno, A generalization of the duality and sum formulas on the multiple zeta values, *J. Number Theory* **74** (1999), no. 1, 39–43.

[30] Y. Ohno, A proof of the cyclic sum conjecture for multiple zeta values, *Preprint* MPIM 2000 (21).

[31] Y. Ohno and N. Wakabayashi, Cyclic sum of multiple zeta values, *Acta Arith.* **123** (2006), no. 3, 289–295.

[32] Y. Ohno and D. Zagier, Multiple zeta values of fixed weight, depth, and height, *Indag. Math. (N.S.)* **12** (2001), no. 4, 483–487.

[33] Y. Ohno and W. Zudilin, Zeta stars, *Commun. Number Theory Phys.* **2** (2008), no. 2, 325–347.

[34] J. Okuda and Y. Takeyama, On relations for the multiple q-zeta values, *Ramanujan J.* **14** (2007), no. 3, 379–387.

[35] J. Okuda and K. Ueno, Relations for multiple zeta values and Mellin transforms of multiple polylogarithms, *Publ. Res. Inst. Math. Sci.* **40** (2004), no. 2, 537–564.

[36] D. Orr, Generalized rational zeta series for $\zeta(2n)$ and $\zeta(2n+1)$, *Integral Transforms Spec. Funct.* **28** (2017), no. 12, 966–987.

[37] G. Rhin and C. Viola, The group structure for $\zeta(3)$, *Acta Arith.* **97** (2001), no. 3, 269–293.

[38] T. Rivoal, La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs, *C. R. Acad. Sci. Paris Sér. I Math.* **331** (2000), no. 4, 267–270.

[39] S. Seki and S. Yamamoto, A new proof of the duality of multiple zeta values and its generalizations, *Intern. J. Number Theory* **15** (2019), no. 6, 1261–1265.

[40] L. J. Slater, *Generalized hypergeometric functions* (Cambridge University Press, Cambridge, 1966).

[41] T. Terasoma, Mixed Tate motives and multiple zeta values, *Invent. Math.* **149** (2002), no. 2, 339–369.

[42] E. A. Ulanskii, Identities for generalized polylogarithms, *Mat. Zametki* **73** (2003), no. 4, 613–624; English transl., *Math. Notes* **73** (2003), 571–581.

[43] E. Ulanskii, A new proof of the theorem on Ohno relations for MZVs, *Moscow J. Combin. Number Theory* **7** (2017), no. 1, 79–88.

[44] J. Wan, Some notes on weighted sum formulae for double zeta values, in *Number theory and related fields*, Springer Proc. Math. Stat. **43** (Springer, New York, 2013), 361–379.

[45] S. Yamamoto, Interpolation of multiple zeta and zeta-star values, *J. Algebra* **385** (2013), 102–114.

[46] S. Yamamoto, Some remarks on Maesaka–Seki–Watanabe's formula for the multiple harmonic sums, *Preprint* arXiv:2403.03498 [math.NT].

[47] D. ZAGIER, Values of zeta functions and their applications, in *First European Congress of Mathematics*, Vol. II (Paris, 1992), Progress in Math. **120** (Birkhäuser, Basel, 1994), 497–512.

[48] D. ZAGIER, Evaluation of the multiple zeta values $\zeta(2,\dots,2,3,2,\dots,2)$, *Ann. of Math.* (2) **175** (2012), 977–1000.

[49] J. ZHAO, On a conjecture of Borwein, Bradley and Broadhurst, *J. Reine Angew. Math.* **639** (2010), 223–233.

[50] J. ZHAO, Identity families of multiple harmonic sums and multiple zeta star values, *J. Math. Soc. Japan* **68** (2016), no. 4, 1669–1694.

[51] J. ZHAO, *Multiple zeta functions, multiple polylogarithms and their special values*, Series on Number Theory and its Applications **12** (World Sci. Publ., Hackensack, NJ, 2016).

[52] W. ZUDILIN, Algebraic relations for multiple zeta values, *Russian Math. Surveys* **58** (2003), no. 1, 1–29.

[53] W. ZUDILIN, Multiple q-zeta brackets, *Mathematics* **3** (2015), no. 1, 119–130.

[54] W. ZUDILIN, Hypergeometric heritage of W. N. Bailey, *Not. Intern. Congress Chinese Mathematicians* **7** (2019), no. 2, 32–46.

[55] W. ZUDILIN, On a family of polynomials related to $\zeta(2,1)=\zeta(3)$, in *Periods in Quantum Field Theory and Arithmetic* (RTMZV 2014, Madrid, Spain), J.I. Burgos Gil, K. Ebrahimi-Fard and H. Gangl (eds.), Springer Proc. Math. Stat. **314** (Springer, 2020), 621–630.

Index