Contents

Preface xvii

Introduction xix

1. Multiple zeta functions 25

1.1 Riemann zeta function 25
1.2 Multiple zeta functions 26
1.3 Analytic continuation 27
1.4 Other multiple zeta functions 30
1.5 Historical notes 33

2. Multiple polylogarithms (MPLs) 37

2.1 Chen's iterated integrals 37
2.2 Classical polylogarithms 38
2.3 Definition of multiple polylogs 40
2.3.1 Control the divergence: $\mathsf{DivLog}_{\mathbb{N}}$ and DivLog_{D} . . 41
2.3.2 Relations between $\mathsf{DivLog}_{\mathbb{N}}$ and DivLog_{D} 42
2.3.3 Application of iterated integrals 45
2.4 Analytic continuation of multiple logarithms 46
2.5 Monodromy of multiple logarithms 48
2.6 Variation matrix for MPLs 54
2.6.1 General procedure leading to pseudo row entries . 54
2.6.2 Some examples in weight 4 55
2.6.3 A general result 59
2.7 Variations of mixed Hodge structures 61
2.7.1 Definition of variations of mixed Hodge structures: a review 61

2.7.2 Mixed Hodge structures of multiple logarithms . . 64
2.8* Limit mixed Hodge structures of multiple logarithms 65
2.8.1 Limit mixed Hodge structures of double polylogs . 67
2.9 Single-valued MPLs of multi-variables 70
2.9.1 General procedure of construction 70
2.9.2 Single-valued double logarithms 71
2.9.3 Single-valued double polylogs $\mathscr{L}i_{1,2}$ and $\mathscr{L}i_{2,1}$ . . 72
2.10 Aomoto polylogs. 73
2.11 Historical notes 78

3. Multiple zeta values (MZVs) 85

3.1 Shuffle product and Euler's decomposition formula 85
3.2 Hoffman's algebraic setup 86
3.2.1 The shuffle relations 86
3.2.2 The stuffle relations 90
3.3 Double shuffle relations. 91
3.3.1 Lyndon words 91
3.3.2 Regularization in two ways 96
3.3.3 The main statement 101
3.4 Dimension conjectures over $\mathbb{Q}$. 103
3.5 Linearized double shuffle and Parity Principle 105
3.6 Historical notes 111

4. Drinfeld associator and single-valued MZVs 117

4.1 Drinfeld associator. 117
4.1.1 KZ-equations: a general setup 117
4.1.2 Some Hopf algebras and their group-like elements 120
4.1.3 Grothendieck-Teichmüller group 124
4.1.4 Pentagon and hexagon relations 127
4.2 Single-variable MPLs. 132
4.3 Single-valued single-variable MPLs 136
4.4 Single-valued MZVs and integration of SVMPLs 139
4.5 Generating series of Hoffman elements 143
4.6 Historical notes 147

5. Multiple zeta value identities 153

5.1 Duality relations. 153
5.2 Cyclic sum relation 154

5.3 Weighted sum relations . . . 158
5.4 Derivation relations . . . 169
5.5 Fixed weight, height and depth relations . . . 177
5.6 Zagier's **2**-3-**2** formula of MZVs . . . 179
5.7 An exotic shuffle relation . . . 183
5.8 Period polynomial relations of double zeta values . . . 188
5.8.1 Double zeta space . . . 188
5.8.2 Period polynomials . . . 189
5.8.3 Modular forms and cusp forms on $\mathrm{SL}_2(\mathbb{Z})$ . . . 192
5.8.4 The even weight double zeta value relations . . . 195
5.8.5 The odd weight double zeta value relations . . . 201
5.9 Conjectural multiple zeta value relations . . . 206
5.10 Historical notes . . . 207

6. Symmetrized multiple zeta values 217

6.1 Definition . . . 217
6.2 Identities involving symmetrized MZVs . . . 219
6.3 Drinfeld associator and the symmetrized MZVs . . . 224
6.4* Space generated by symmetrized MZVs . . . 228
6.4.1 Natural symmetrized MZVs . . . 228
6.4.2 An embedding $\iota_n : \mathfrak{S}_{n+1} \hookrightarrow \mathrm{GL}_n(\mathbb{Z})$ . . . 230
6.4.3 Parity results . . . 231
6.4.4 Diagonal translation invariant polynomials . . . 235
6.4.5 Symmetrized MZVs generate whole MZV space . . . 239
6.5 Historical notes . . . 241

7. (Alternating) multiple harmonic sums 243

7.1 Definitions . . . 243
7.2 Stuffle relations . . . 244
7.3 Quasi-symmetric functions with signed powers . . . 247
7.4 Symmetric sums . . . 250
7.5 Binomial identities . . . 251
7.6 Historical notes . . . 253

8. Finite multiple zeta values and finite Euler sums 257

8.1 Definitions . . . 258
8.2 Reversal and concatenation relations . . . 260
8.3 Stuffle relations . . . 264

8.4 Shuffle relations 265
8.5 Finite multiple zeta values 267
8.5.1 Homogeneous finite multiple zeta values 267
8.5.2 Duality of finite multiple zeta values 269
8.5.3 Finite multiple zeta values of lower depths 272
8.5.4 Structure of finite multiple zeta values 273
8.5.5 Finite multiple zeta values of superbity 2 275
8.5.6 Finite multiple zeta values of superbities ≤ 5 . . . 280
8.5.7 Finite multiple zeta values of arbitrary depths . . 280
8.6 Finite Euler sums 282
8.6.1 Two reduction formulas 282
8.6.2 Finite Euler sums of small depths 283
8.6.3 Finite Euler sums of superbity 1 and 2 287
8.6.4 Two applications of MHS binomial identities . . . 289
8.6.5 Finite Euler sums of arbitrary depths 292
8.7 Historical notes 293

9. q-analog of multiple harmonic (star) sums 299

9.1 Some q-combinatorial identities 301
9.2 q-MH*S identities 304
9.3 Some applications 310
9.4 Congruences of q-MHSs 311
9.5 Historical notes 320

10. Multiple zeta star values 325

10.1 An integral expression 325
10.2 Relation between MZ*V and MZV 326
10.3 Sum relations 328
10.4 Cyclic sum relations 329
10.5 Weighted sum relations. 330
10.6 Fixed weight, height and depth relations 331
10.7 Duality relations. 336
10.8 Some applications of binomial identities of q-MZ*Vs 342
10.9 Three-Two-One formulas 345
10.10 Historical notes 355

11. q-analogs of multiple zeta functions 361

11.1 Definitions 361

11.2 Analytic continuation 364
11.3 Special values at non-positive integers 372
11.4 Historical notes 383

12. q-analogs of multiple zeta (star) values 385

12.1 Various definitions of q-MZVs 385
12.2 Double shuffle relations 386
12.2.1 Rota-Baxter algebra 387
12.2.2 q-analogs of Hoffman algebras 390
12.2.3 q-stuffle relations 392
12.2.4 Iterated Jackson q-integrals 394
12.2.5 q-shuffle relations 398
12.3 Duality Relations 403
12.4 The **P**-**R** and **R**-**P** relations 405
12.5 The general type G q-MZVs 405
12.6 Application to Okounkov's conjecture 407
12.7 The q-analog of Drinfeld associator 409
12.8 q-MZVs by Bachmann and Kühn 410
12.8.1 The mono-bracket 410
12.8.2 A derivation 415
12.8.3 Multiple Eisenstein series 420
12.8.4 Bi-brackets and double shuffle relations 421
12.9 Some applications of MZ*V binomial identities 425
12.9.1 Mollified companion of q-MZVs and q-MZ*Vs . . 425
12.9.2 Applications to MZVs and MZ*Vs 427
12.10 Historical notes 430

13. Colored multiple zeta values 435

13.1 Definitions 435
13.2 Dimension upper bound by Deligne and Goncharov 436
13.3 Standard relations 437
13.3.1 Regularized double shuffle relations 437
13.3.2 Proof of regularized double shuffle relations 441
13.3.3 Weight one relations 450
13.3.4 Regularized distribution relations 451
13.3.5 Lifted relations from lower weights 453
13.3.6 The group-like element **I** 455
13.4* Generalized Drinfeld associators 457

13.5 Historical notes 459

14. Colored multiple zeta values at lower levels 465

14.1 EZ-face: an online computation resource 465
14.2 Euler sums 467
14.2.1 Dimension conjectures 467
14.2.2 A possible integral structure 468
14.2.3 A few useful integral substitutions 471
14.2.4 A family of identities $\zeta(\{3\}^n) = 8^n\zeta(\{\bar{2},1\}^n)$ . . . 471
14.2.5 Generalized doubling relations at level 2 476
14.3 Nonstandard octahedral relations at level 4 477
14.4 Multiple Clausen and Deligne values at level 6 480
14.4.1 Definitions 480
14.4.2 Central binomial sums 483
14.4.3 Dualities of multiple Clausen values 485
14.4.4 A **1**-2-**1** formula of multiple Glaishers 486
14.4.5 Dimension conjectures 490
14.5* Level p and p^2 cases ($p \geq 5$ a prime) 492
14.6 Computational results in weight ≤ 5 498
14.7 Historical notes 500

15. Application to Feynman integrals 505

15.1 Why Feynman integrals? 505
15.2 Feynman diagrams 506
15.3 Feynman integrals in momentum space 507
15.3.1 Divergence types 508
15.3.2 A short example 508
15.3.3 Massive and massless integrals 509
15.4 Graphical functions in position space 509
15.4.1 Sequential graphs and graphical functions 510
15.4.2 Properties of graphical functions 513
15.5 Feynman integrals in parametric form 523
15.5.1 Graph polynomials in edge variables 523
15.5.2 Feynman integrals in projective space 524
15.6 Historical notes 526

Appendix A Key concepts of Hopf algebras 531

Appendix B Some useful results of Lie algebras 537

Appendix C Basics of hypergeometric functions 545

C.1 Solutions of some ordinary differential equations 545
C.2 Transformation formulas 546
C.2.1 Linear transformations 546
C.2.2 Some transformations of ${}_3F_2$ 547

Appendix D Sample computer codes 549

D.1 Double shuffle relations. 549
D.2 Multiple harmonic sums 550
D.3 Finite multiple zeta values 550

Appendix E Tables of special values 553

E.1 Euler sums of lower weights. 553
E.2 Finite multiple zeta values 557

Bibliography 561

List of Abbreviations 593

List of Symbols 595

Index 599

Preface

Since my postdoc years at UPenn, I took many notes from various papers to help myself keep abreast with the current research of multiple zeta values and multiple polylogarithms. As the notes became thicker and thicker, I felt a strong urge to organize them into a book form so that not only my students and I but also other people interested in these areas would benefit from them. More recently, these objects have become more and more important due to their newly discovered close relations to numerous other branches of mathematics and even mathematical physics, so that such a textbook seems more appropriate.

Originally I planned to include the most basic topics as well as some more advanced theory such as the p-adic multiple zeta values and the motivic fundamental groups. But as more details were written out, I found it very difficulty to maintain the book to an ideal length while keeping all of them together. Consequently, I must choose to stay on the more elementary side and insert only several optional sections for the more inquisitive reader.

During the multi-year long writing process, I have benefited enormously from multiple visits to the Max Planck Institute for Mathematics, IHES, Taiwan National University, the Academia Sinica, and the Morningside Center of Mathematics at Beijing. This project was supported partially by NSF grant DMS1162116.

Many people have contributed to this work in various ways. Herbert Gangl was responsible for prompting me to prepare the first outline of this book in early 2010. He was also extremely helpful in providing with me many valuable opinions from the content of the book to the pictures on the cover. Thanks are also due to my students John Cutie and Margaret Sundberg, who suggested many grammatical improvement.

Introduction

The Riemann zeta function $\zeta(s)$ is one of the most fascinating objects in mathematics. It is a well-known historical fact that L. Euler had studied $\zeta(s)$ as a real function long before B. Riemann. In particular he was able to evaluate the function at positive even numbers up to 26 using Bernoulli numbers. He also established a few very nice relations among the double zeta values defined by the nested double sums

$$\zeta(s,t) := \sum_{m>n\geq 1} \frac{1}{m^s n^t}, \quad s,t \in \mathbb{N},$$

which are said to have weight $s+t$. As a matter of fact, Euler worked with the double zeta *star* values $\zeta^\star(s,t)$ which are defined by replacing $m > n$ by $m \geq n$ in the above. Among his many amazing discoveries one stands out as the beautiful identity $\zeta(2,1) = \zeta(3)$, simple to state but not so easy to prove. If you have not seen it before, perhaps you should try it by yourself.

Around 1990, more than 200 years after Euler, this subject was revitalized and generalized to the multiple zeta values (MZVs) by M. Hoffman and D. Zagier independently and almost simultaneously. These values provide the multiple variable version of zeta and double zeta values defined by iterated sums

$$\zeta(s_1,\ldots,s_d) := \sum_{k_1>\cdots>k_d>0} k_1^{-s_1}\cdots k_d^{-s_d}. \tag{0.1}$$

It is important to note that another common way to define these in the literature is to use the increasing order on the indices.

Since their debut MZVs have appeared unexpectedly in many diverse areas of mathematics and physics. Most prominently, they have showed up in the computation of some infinite classes of Feynman integrals and knot invariants. Further, F. Brown proved in 2012 in Ref. [104] that all

the periods of the category of mixed Tate motives unramified over $\mathbb{Z}$ are $\mathbb{Q}[\frac{1}{2\pi i}]$-linear combinations of MZVs, revealing their fundamental nature in algebraic geometry.

History often repeats itself. Just like Riemman considered $\zeta(s)$ as a complex function years after Euler's work, multiple zeta functions $\zeta(s_1, \ldots, s_d)$, where $s_1, \ldots, s_d$ are complex variables, were investigated almost a decade after Hoffman and Zagier's seminal works. We will visit this analytic theory via the Euler-Mclaurin summation formula in Chap. 1.

In the mid 1990s, mathematical physicists discovered that one needs not only the MZVs to express some Feynman integrals but also the special values of the multiple polylogarithms (MPLs) at the sixth root of unity. Here, MPLs are multiple variable generalizations of the classical polylogarithms, which will be considered in Chap. 2. Besides the analytic properties, we also look at the mixed Hodge structures associated to MPLs and, as a by-product, sketch a way to define their single-valued versions.

The next four chapters form the central part of this book. The reader will find many detailed results concerning MZVs, proved using a variety of methods, such as the theory of hypergeometric series, Hopf algebras, and Lie algebras. Since these tools are rarely found in the standard graduate curriculum, we list a few key results and properties we need in the appendices found at the back of the book.

One of the most fruitful ideas to study the rational structure of MZVs is to consider the so-called double shuffle relations. If two MZVs are multiplied, we can use either their infinite series representations or their iterated integral representations first observed by M. Kontsevich. The former leads to the quasi-shuffle structure, the latter, the shuffle. Comparing the two ways of computing the same product, one obtains the so-called finite double shuffle relations. However, even in the simplest case, these relations are insufficient. For example, Euler's identity $\zeta(2,1) = \zeta(3)$ cannot be proved this way simply because the smallest weight of a product of MZVs is 4. To prove this identity, we are naturally lead to the regularizations of MZVs in two different ways corresponding to the two different product structures. We present the beautiful work of K. Ihara, M. Kaneko and D. Zagier [298] in Chap. 3.

A surprising application of MZV theory is that they can be used to express the Drinfeld associator which plays a key role in V. Drinfeld's work on the deformation of some Hopf algebras related to the quantum Yang-Baxter equations [163,164]. The first half of Chap. 4 centers around this key connection, which provides us with an invaluable framework to study the

symmetrized MZVs in Chap. 6. The second half introduces a single variable cousin of MPLs, whose single-valued version will be crucial in evaluating several classes of Feynman integrals in Chap. 15.

Chapter 5 is a sort of compendium of various identities involving MZVs whose proofs require many different methods. Chapter 6, on the contrary, is rather short and concentrates on only one special variant of MVZs constructed by a kind of cyclic symmetry. These symmetrized MZVs will be linked to the finite version of MZVs to be considered in Chap. 8.

When the infinite sums in Eq. 0.1 are truncated to finite partial sums, one obtains the so-called multiple harmonic sums (MHSs), which will be treated in Chap. 7 together with their alternating version. As these sums are obviously rational numbers we are naturally led to examine their Van Hamme type congruence properties [279, 520, 596, 606]. Only recently do we realize that there should be close relations between the MZVs and their finite versions which are defined essentially by embedding the MHSs into a kind of adèle structure. In particular, this surprising connection motivates the study of the double shuffle relations among the finite MZVs and finite Euler sums. Notice that *a priori* we have only one way to multiply MHSs by their series definition and no integral expression is available. However, due to a beautiful conjecture from Kaneko and Zagier, we can now use some symmetrized versions of MZVs to link to finite MZVs thereby hinting at a kind of double shuffle relations among them. This will be done in Chap. 8.

For many interesting objects in number theory we often can and do consider their q-analogs. We will carry this out for MHSs in Chap. 9. Many results contained in this chapter imply the corresponding ones for q-MZVs when we take the number of terms in the finite sums to infinity. When we further let $q \to 1^-$ we arrive at statements for ordinary MZVs. And we can also extend this framework to their star version, which is the content of Chap. 10.

In the next two chapters we continue to deal with q-analogs, first for the multiple zeta functions and then for their special values (q-MZVs). It turns out there are quite a few different ways to define q-MZVs, which are often related, although sometimes implicitly at the first glance. Our attention here is mostly devoted to the double shuffle structures among these q-MZVs.

We return to MPLs in Chaps. 13 and 14 where we consider the special values of these functions at the Nth roots of unity. We call these numbers colored MZVs of level N. Unlike MZVs, the double shuffle relations only provide some of the $\mathbb{Q}$-linear relations between colored MZVs. Chaps. 13

explains how to obtain more relations from algebraic and geometric considerations. Chaps. 14 gives more detailed structural results for lower level colored MZVs.

In the last chapter we provide an attractive and dynamic link recently found between MZVs/MPLs and mathematical physics: the study of Feynman integrals. This is one of the most active research areas in mathematical physics in the past 20 years. We consider two valuable techniques in these computations. The first is the graphical functions of O. Schnetz [487] which are used to evaluate the boxed ladder graphs, the wheel graphs, and the zigzag graphs by applying the theory of single-valued MPLs and single-valued MZVs developed in Chap. 4. The second is to use the Schwinger parametric form of Feynman integrals to transform these to projective integrals, which is particularly effective in interpreting results concerning Feynman integrals in terms of algebraic geometric objects such as motives and the cosmic Galois group originally conjectured by P. Cartier and currently being developed by F. Brown.

Our primary goal in writing this book is to provide a comprehensive and accessible introduction to the theory of MZVs and MPLs for graduate students and researchers who are interested in learning this beautiful subject. Due to the rapid development of these subjects in recent years as well as space limitation, I have refrained from wading into a few advanced topics such as the motivic theory [104, 156], the p-adic theory [45, 46, 204, 206, 212, 308], the transcendence theory [32, 107, 619], the relations between MPLs to cluster algebras [231, 456], and the generalizations to elliptic MPLs [37, 55, 96, 97, 108, 189], to name just a few. The omission does not mean these are negligible. On the contrary, to have a thorough understanding of MVZs/MPLs these are often indispensable tools. For example, there is a close connection between the p-adic MZVs and the finite MZVs which provides a very promising approach to proving some of the most fundamental conjectures according to the recent works of D. Jarossay [307, 308] and S. Yasuda.

The results contained in this book are due to many people whose contributions are too numerous to name each individually. However, I have tried to compile them in the last section of each chapter as some historical notes. Further, some relevant research works are often stated without proofs in these sections. Hopefully, the interested readers can explore further and find their own next research project by themselves.

To facilitate the study, I also provide a few exercise problems at the end of each chapter. Most of these problems are straight-forward but some are

probably difficult. The serious readers are encouraged to try at least some of them, especially those that are cited in the main body of the book.

The subject of MZVs/MPLs is often very computationally intensive and has been a hot bed for experimental mathematics for a couple of decades [28, 64]. We thus provide quite a few computer programming related problems scattered throughout the book and sketch a few crucial computer pseudo-codes in an appendix. Also included in the appendices are a few tables of useful numerical data and answers to a few selected exercise problems.

Many people have read the manuscript. Although a lot of inaccuracies have been caught and corrected during the editing process, there will likely be some more errors, for which I am solely responsible. I plan to maintain an online home page for these errors at ...

Chapter 1

Multiple zeta functions

Multiple zeta functions are nested generalizations of the Riemann zeta function. In this chapter we start by summarizing briefly the most important results of the Riemann zeta function. Then we bring in one of our major players in this book — the multiple zeta function — and outline its main properties and problems. Some results will be generalized to higher levels in Chap. 13, but often with more advanced and different proofs. So, we still provide complete proofs of these results in this chapter.

1.1 Riemann zeta function

There are numerous accounts in the literature about the Riemann zeta function defined as follows:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \qquad \mathrm{Re}(s) > 1.$$

As motivation for further generalizations, we list some of its most important properties below.

- **Euler product formula.** For $\mathrm{Re}(s) > 1$

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

- **The functional equation.** The Riemann zeta function can be analytically continued to a meromorphic function on the whole complex s-plane, which is holomorphic everywhere except for a simple pole at $s = 1$ with residue 1. Moreover, it satisfies

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s)\, \zeta(1-s). \tag{1.1}$$

- **Special values.** For any positive even number $2n$,

$$\zeta(2n) = -\frac{B_{2n}}{2(2n)!}(2\pi i)^{2n}, \tag{1.2}$$

where B_{2n} is a Bernoulli number defined by the generating function

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}. \tag{1.3}$$

For negative integers $-n < 0$, one has

$$\zeta(-n) = -\frac{B_{n+1}}{n+1}. \tag{1.4}$$

- **Generating function of the special values.** Let $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ be the digamma function defined for all x unequal to a negative integer. Then for all $|x| < 1$

$$\psi(1-x) = -\sum_{k\geq 2} \zeta(k) x^{k-1} - \gamma, \tag{1.5}$$

where γ is Euler's constant. See, for example, [5, § 6.3]. Moreover, ψ satisfies the reflection, recurrence and duplication relations

$$\begin{aligned} \psi(x) &= \psi(1-x) - \pi\cot(\pi x) = \psi(x+1) - \frac{1}{x}, \\ \psi(2x) &= \frac{1}{2}\psi(x) + \frac{1}{2}\psi(x+1/2) + \log 2, \end{aligned} \tag{1.6}$$

and it has an important integral expression

$$\psi(x+1) = -\gamma + \int_0^1 \frac{1-t^x}{1-t}\mathrm{d}t, \tag{1.7}$$

- **Zeros, the critical line, and the Riemann hypothesis.** From Eq. (1.4) we see that $\zeta(s)$ has zeros at negative even integers (called trivial zeros). In his landmark 1859 paper [470] Riemann conjectured that all non-trivial zeros of $\zeta(s)$ must lie on the critical line $\mathrm{Re}(s) = 1/2$. This is the famous "Riemann hypothesis".

1.2 Multiple zeta functions

Multiple zeta functions are multiple variable generalizations of the Riemann zeta function.

Definition 1.2.1. For fixed positive integer d and d-tuple of complex variables $\mathbf{s} = (s_1, \ldots, s_d)$, the *multiple zeta function* is defined by

$$\zeta(\mathbf{s}) = \sum_{k_1 > \cdots > k_d > 0} k_1^{-s_1} \cdots k_d^{-s_d}. \tag{1.8}$$

To gurantee convergence, $\mathbf{s}$ must satisfy $\mathrm{Re}(s_1 + \cdots + s_j) > j$ for all $j = 1, \ldots, d$ where $\mathrm{Re}(s)$ is the real part of a complex number s (see Exercise 1.4). The number d is called the *depth*, denoted by $\mathrm{dp}(\mathbf{s})$. When all the variables are positive integers we call $|\mathbf{s}| := s_1 + \cdots + s_d$ the *weight*. Furthermore, the *multiple zeta star function* is defined by

$$\zeta^\star(\mathbf{s}) = \sum_{k_1 \geq \cdots \geq k_d \geq 1} k_1^{-s_1} \cdots k_d^{-s_d}. \tag{1.9}$$

Remark 1.2.2. (i) In the literature the order $0 < k_1 < \cdots < k_d$ is used sometimes, which gives rise to $\zeta(s_d, \ldots, s_1)$ in our notation.

(ii) It is not hard to see that $\zeta^\star(\mathbf{s})$ can be expressed using the multiple zeta functions of different depths:

$$\zeta^\star(s_1, \ldots, s_d) = \sum \zeta(s_1 \circ \cdots \circ s_d) \tag{1.10}$$

where $\circ$ is either "+" or ",". So we will only consider the latter in the rest of this chapter.

Similar to the Riemann zeta function one can study both the algebraic theory of the special values of the multiple zeta functions at positive integers, called the *multiple zeta values* and the analytic theory of the multiple zeta function. In this chapter we will consider mainly the analytic side and postpone the treatment of multiple zeta values to Chaps. 3–6.

1.3 Analytic continuation

Since the multiple zeta function of depth d in Eq. (1.8) is defined only for complex arguments with certain restrictions it is essential to continue it analytically to the whole d-dimensional complex space in order to study special values at arbitrary integer arguments.

Recall that the Bernoulli polynomials $B_k(x)$ are defined by the generating function

$$\frac{te^{xt}}{e^t - 1} = \sum_{k=0}^{\infty} B_k(x) \frac{t^k}{k!}$$

and the "periodic Bernoulli polynomial" defined by

$$\widetilde{B}_k(x) = B_k(\{x\}), \quad x \geq 1,$$

where $\{x\}$ is the fractional part of x. Then we have ([550, Ch. IX, Misc. Ex. 12])

$$\widetilde{B}_k(x) = -k! \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{e^{2\pi i n x}}{(2\pi i n)^k}. \tag{1.11}$$

Recall that the Bernoulli numbers satisfy $B_k = B_k(1)$ if $k \geq 2$ while $B_0 = 1$ and $B_1 = -1/2$.

Lemma 1.3.1. *For every positive integer $M \geq 2$ and $x > 1$ we have*

$$\left|\widetilde{B}_M(x)\right| \leq \frac{4M!}{(2\pi)^M}.$$

Proof. It follows from the fact that $\zeta(M) \leq \zeta(2) = \pi^2/6 < 2$ for $M \geq 2$. □

Let's recall the classical Euler-Maclaurin summation formula [550, 7.21]. Let $f(x)$ be any (complex-valued) C^∞ function on $[1,\infty)$ and let m and M be two positive integers with M even. Then we have

$$\sum_{n=1}^{m} f(n) = \int_1^m f(x)\,\mathrm{d}x + \frac{1}{2}(f(1)+f(m)) + \sum_{r=2}^{M} \frac{B_r}{r!}\left(f^{(r-1)}(m) - f^{(r-1)}(1)\right)$$
$$-\frac{1}{M!}\int_1^m \widetilde{B}_M(x) f^{(M)}(x)\,\mathrm{d}x. \tag{1.12}$$

Suppose $\mathrm{Re}(s) > 1$, $f(x) = 1/x^s$. Take $m = k$ and ∞ respectively in Eq. (1.12) and then take the difference we get:

$$\sum_{n>k} \frac{1}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{k} \frac{1}{n^s}$$
$$= \int_k^\infty f(x)\,dx - \frac{1}{2}f(k) - \sum_{r=2}^{M} \frac{B_r}{r!} f^{(r-1)}(k) + \frac{1}{M!}\int_k^\infty \widetilde{B}_M(x) f^{(M)}(x)\,\mathrm{d}x$$
$$= \frac{1}{(s-1)k^{s-1}} - \frac{1}{2k^s} + \sum_{r=2}^{M} \frac{B_r}{r!}\frac{(s)_{r-1}}{k^{s+r-1}} + \frac{(s)_M}{M!}\int_k^\infty \frac{\widetilde{B}_M(x)}{x^{s+M}}\,\mathrm{d}x,$$

where

$$(s)_M = \frac{\Gamma(s+M)}{\Gamma(s)} = \begin{cases} 1, & \text{if } M = 0; \\ s(s+1)\cdots(s+M-1), & \text{if } M > 0, \end{cases} \tag{1.13}$$

is the Pochhammer symbol. Here we have used the fact that $B_k = 0$ if k is odd ≥ 3. Setting $(s)_0 = 1$ and $(s)_{-1} = 1/(s-1)$ one has

$$\sum_{n>k} \frac{1}{n^s} = \sum_{r=0}^{M} \frac{B_r}{r!}\frac{(s)_{r-1}}{k^{s+r-1}} + \frac{(s)_M}{M!}\int_k^\infty \frac{\widetilde{B}_M(x)}{x^{s+M}}\,\mathrm{d}x. \tag{1.14}$$

Theorem 1.3.2. *Let d be a positive integer. Define*

$$\mathsf{S}_d = \left\{ (s_1, \dots, s_d) \in \mathbb{C}^d \;\middle|\; \begin{array}{l} s_1 = 1, s_1 + s_2 = 2, 1, -2m \ \forall m \in \mathbb{Z}_{\geq 0}, \\ \text{and } s_1 + \cdots + s_j \in \mathbb{Z}_{\leq j} \ \forall j \geq 3. \end{array} \right\}. \quad (1.15)$$

(i) *If $d = 1$ then we have*

$$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} + \sum_{r=1}^{M} \frac{B_{r+1}}{(r+1)!}(s)_r - \frac{(s)_{M+1}}{(M+1)!} \int_1^\infty \widetilde{B}_{M+1}(x) x^{-s-M-1} \, \mathrm{d}x. \quad (1.16)$$

(ii) *If $d \geq 2$ then for all $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{C}^d \setminus \mathsf{S}_d$ and positive integer $M > d + 1 + |\mathrm{Re}(s_1)| + \cdots + |\mathrm{Re}(s_d)|$ we have*

$$\zeta(\mathbf{s}) = \sum_{r=0}^{M} \frac{B_r}{r!}(s_1)_{r-1} \cdot \zeta(s_1 + s_2 + r - 1, s_3, \dots, s_d) + \frac{(s_1)_M}{M!} \sum_{k_2 > \cdots > k_d > 0} \frac{1}{k_2^{s_2} \cdots k_d^{s_d}} \int_{k_2}^\infty \frac{\widetilde{B}_M(x)}{x^{s_1+M}} \, \mathrm{d}x. \quad (1.17)$$

These provide the analytic continuation of multiple zeta function $\zeta(\mathbf{s})$ to a meromorphic function on $\mathbb{C}^d$ with simple poles given by S_d.

Proof. We prove the theorem by induction on the depth d. When $d = 1$, taking $k = 1$ in Eq. (1.14) we get Eq. (1.16). Since we may choose M arbitrary large, this gives the analytic continuation of $\zeta(s)$ to the whole s-plane.

Now we assume $d \geq 2$. In definition Eq. (1.8) we replace s_1, k_1, k_2 by s, n and k, respectively. Then Eq. (1.17) follows directly from Eq. (1.14) if we can show that the series in Eq. (1.17) converges. Lemma 1.3.1 implies (if $d = 2$ then take $k_3 = 1$)

$$\sum_{k_2 > \cdots > k_d > 0} \left| \frac{1}{k_2^{s_2} \cdots k_d^{s_d}} \int_{k_2}^\infty \frac{\widetilde{B}_M(x)}{x^{s_1+M}} \, \mathrm{d}x \right| \leq \frac{4M!}{(2\pi)^M (M - 1 + \mathrm{Re}(s_1))} \times \times \sum_{k_2 > \cdots > k_d > 0} \frac{1}{k_2^{M-1+\mathrm{Re}(s_1)+\mathrm{Re}(s_2)} k_3^{\mathrm{Re}(s_3)} \cdots k_d^{\mathrm{Re}(s_d)}}$$

which converges absolutely whenever $M > d + 1 + |\mathrm{Re}(s_1)| + \cdots + |\mathrm{Re}(s_d)|$.

This completes the proof of the theorem by induction. □

1.4 Other multiple zeta functions

In the Arizona Winter School 2002, Deligne designed a project for some participating graduate students to find a good candidate in the étal theory corresponding to the multiple zeta function defined by Eq. (1.8), called the p-adic multiple zeta functions. Subsequently Furusho successfully defined p-adic multiple zeta *values* by generalizing Coleman's construction of p-adic polylogarithms and explained his results with the Tannakian formalism in [204–206]. Moreover, Chatzistamatiou [123] showed that these values are p-adic integers.

Over global fields, one can generalize multiple zeta functions to multiple Dedekind zeta functions over arbitrary number field F. Let O_F be the ring of integers of F and $N(I) = |O_F : I|$ the norm for an integral ideal I in F. Let d be a positive integer and $F_1, \dots, F_d$ be d number fields. R. Masri [403] defined the multiple Dedekind zeta function over $F_1, \dots, F_d$ as (after slightly change of ordering of the indices)

$$\zeta_d(\{F_i\}_{i=1}^d; s_1, \dots, s_d) = \sum_{\substack{I_1 \in O_{F_1}, \dots, I_d \in O_{F_d} \\ N(I_1) > \cdots > N(I_d)}} N(\mathfrak{n}_1)^{-s_1} \cdots N(\mathfrak{n}_d)^{-s_d}.$$

By Riemann-Roch he proves that there are positive numbers $\varepsilon_1, \dots, \varepsilon_d$ such that the series $\zeta_d(\{F_i\}_{i=1}^d; s_1, \dots, s_d)$ is absolutely convergent and analytic in the region

$$\mathrm{Re}(s_1 + \cdots + s_j) > j + \varepsilon_1 + \cdots + \varepsilon_j, \quad \forall j = 1, 2, \dots, d.$$

In [591] we consider the case where $F_1 = \cdots = F_d = F$ and define $\zeta_F(s_1, \dots, s_d) = \zeta_d(\{F\}_{i=1}^d; s_1, \dots, s_d)$. In [291] Horozov approached this problem from a different point of view. In particular, his version of multiple Dedekind zeta functions has analytic continuations and their special values can be expressed using iterated integrals.

It would be very interesting to obtain some numerical data concerning the following generalization of Zagier's conjecture concerning Dedekind zeta functions (cf. Conjecture 2.2.1).

Problem 1.4.1. For large enough positive integers $n_1, \dots, n_d$ such that $\zeta_F(n_1, \dots, n_d)$ converges, is there an expression of $\zeta_F(n_1, \dots, n_d)$ in terms of a determinant with entries given by $\mathscr{L}i\, n_1, \dots, n_d$ evaluated at F-rational points up to some factors determined only by the number field F (such as the discriminant, the number of real and complex embeddings, etc.)? Here $\mathscr{L}i\, n_1, \dots, n_d$ is the single-valued version of multiple polylogarithms to be defined in Sec. 2.9.

In another direction, Akiyama and Ishikawa [8] defined the multiple Hurwitz zeta function by

$$\zeta(s_1,\dots,s_d;\theta_1,\dots,\theta_d)=\sum_{k_1>\cdots>k_d>0}(k_1+\theta_1)^{-s_1}\cdots(k_d+\theta_d)^{-s_d}, \quad (1.18)$$

where $0\le\theta_j<1$ for $j=1,\dots,d$ and derive its analytic continuation by using modified version of the Euler-Maclaurin summation formula. Murty and Sinha obtained the analytic continuation in [421] by using the binomial series expansion and Hartogs theorem in complex analysis. All of them consider also the cognate multiple L-series

$$L(s_1,\dots,s_d;\chi_1,\dots,\chi_d):=\sum_{k_1>\cdots>k_d\ge1}\frac{\chi_1(k_1)\cdots\chi_d(k_d)}{k_1^{s_1}\cdots k_d^{s_d}}, \quad (1.19)$$

where χ_j's are Dirichlet characters.

Kelliher and Masri [335] proved the analytic continuation of Eq. (1.18) with an extension of the generalized function approach first adopted in [585] and determine further some residues explicitly. Minh *et al.* also consider these functions in [413] and use a certain shuffle product structure to produce an algorithm for computing the special values of multiple polylogarithms (see Chap. 2 for definition).

Generalizing the ideas of Wu, Matsumoto and *et al.*, Zhao and Zhou (see [405, 408, 614]) define the *Mordell-Tornheim L-functions* by

$$L_{\mathrm{MT}}(\mathbf{s};\boldsymbol{\chi}):=\sum_{k_1=1}^{\infty}\cdots\sum_{k_d=1}^{\infty}\frac{\chi_1(k_1)\cdots\chi_d(k_d)\chi_{d+1}(k_1+\cdots+k_d)}{k_1^{s_1}\cdots k_d^{s_d}(k_1+\cdots+k_d)^{s_{d+1}}}. \quad (1.20)$$

They derive the analytic continuation by utilizing the so-called colored Mordell-Tornheim functions. When all the characters in Eq. (1.20) are principle these functions are nothing but the traditional Mordell-Tornheim zeta functions

$$\zeta_{\mathrm{MT}}(s_1,\dots,s_{d+1}):=\sum_{k_1=1}^{\infty}\cdots\sum_{k_d=1}^{\infty}\frac{1}{k_1^{s_1}\cdots k_d^{s_d}(k_1+\cdots+k_d)^{s_{d+1}}}. \quad (1.21)$$

For another type of multiple zeta functions Matsumoto (see [407]) replaces the factor $k_j+\theta_j$ in the denominator of Eq. (1.18) by linear polynomials of $k_1,\dots,k_j$. More generally, while pursuing his Ph.D. Essouabri considered the multiple zeta functions of the form

$$\sum_{\mathbf{k}=(k_1,\dots,k_d)\in\mathbb{N}^d}P_1(\mathbf{k})^{-s_1}\cdots P_d(\mathbf{k})^{-d_1}$$

where $P_j(\mathbf{k})$'s are polynomials of d variables. In his thesis [194, 195] Essouabri provided a proof of the analytic continuation in the case $d=1$

and stated that his method can be generalized to arbitrary depth for suitable P_j's including all the linear polynomials. Later, Essouabri *et al.* [196] extends this to more general types of multiple zeta functions involving vectorial functions in the denominators.

One of the most interesting multiple zeta functions originates from physics. Around 1990, related to his work on topological quantum field theory, Witten [552] found the volumes of the moduli spaces of representations of the fundamental groups of two dimensional surfaces are some special values of a new type of zeta functions attached to complex semisimple Lie algebras $\mathfrak{g}$ at positive integers. Inspired by this result, Zagier [577] defines these complex functions as follows:

$$\zeta_W(s;\mathfrak{g}) = \sum_{\rho} \frac{1}{(\dim \rho)^s},$$

where ρ runs over all finite dimensional irreducible representations of $\mathfrak{g}$. By physical considerations Witten showed that for any positive integer n

$$\zeta_W(2n;\mathfrak{g}) = c(2n;\mathfrak{g})\pi^{2nr},$$

where $c(2n;\mathfrak{g}) \in \mathbb{Q}$ and r is the number of positive roots of $\mathfrak{g}$. Such formulas are now called Witten volume formulas.

Matsumoto and his collaborators [343–345, 409] have generalized Witten's zeta functions to the multiple variable version over an arbitrary semisimple Lie algebra $\mathfrak{g}$. Let Δ_+ be the set of positive roots of $\mathfrak{g}$, $\{\alpha_1, \ldots, \alpha_d\}$ the fundamental roots with coroots $\{\alpha_1^\vee, \ldots, \alpha_d^\vee\}$, and $\{\lambda_1, \ldots, \lambda_d\}$ the fundamental weights such that $\langle \alpha_i^\vee, \lambda_j \rangle = \delta_{i,j}$ (the Kronecker symbol). Then they define the multiple Witten zeta function over $\mathfrak{g}$ by

$$\zeta_{\mathfrak{g}}(\{s_\alpha\}_{\alpha\in\Delta_+}) := \sum_{k_1,\ldots,k_d=1}^{\infty} \prod_{\alpha\in\Delta_+} \langle \alpha^\vee, k_1\lambda_1 + \cdots + k_d\lambda_d \rangle^{-s_\alpha}. \tag{1.22}$$

They showed in [343] that

$$\zeta_W(s;\mathfrak{g}) = \zeta_{\mathfrak{g}}(\{s\}^d) \prod_{\alpha\in\Delta_+} \langle \alpha^\vee, \lambda_1 + \cdots + \lambda_d \rangle,$$

where $\{s\}^d$ means the variable s repeats itself d times. One of the most exciting results is a statement on functional relations produced by the Weyl group symmetry in the underlying Lie algebra structure (see [342, 344] for details).

It is well-known that there is a huge supply of analogies between the number fields and the function fields over finite fields. Carlitz and Weil

worked on these in the early years of twentieth century, continued by many in recent years (see [246, 523] for many fascinating results along this line). In particular, Thakur devotes a section of his book [523, § 5.10] to the study of various possible multiple zeta values over the rational function field $\mathbb{F}_q[t]$ over the finite field $\mathbb{F}_q$ of q elements. For global function field K over $\mathbb{F}_q$ we now state a result of Masri [404]. For a divisor D of K, let $\deg(D)$ be its degree and $|D| = q^{\deg(D)}$ be its norm. Let $\mathcal{D}_K^+$ be the sub semi-group of effective divisors of K. Masri defines

$$Z_d\left(K; s_1, \ldots, s_d\right) = \sum_{\substack{(D_1,\ldots,D_d)\in \mathcal{D}_K^+\times\cdots\times\mathcal{D}_K^+ \\ 0\leq \deg(D_1)\leq\cdots\leq\deg(D_d)}} \prod_{k=1}^{d} |D_k|^{-s_k},$$

and shows that it has a meromorphic continuation to all of $\mathbb{C}^d$ and is a rational function in each of $q^{-s_1}, \ldots, q^{-s_d}$ with a specified denominator. Further, $Z_d\left(K; s_1, \ldots, s_d\right)$ has possible simple poles on the linear subvarieties

$$s_k + \cdots + s_d = 0, 1, \ldots, d-k+1, \quad k = 1, \ldots, d.$$

On the transcendence of the special values of multiple zeta values over the function field Yu's classical results in [566] and the new breakthrough by Chang and his collaborators [120, 121] go far beyond what we can do over number fields.

1.5 Historical notes

Despite its name the Riemann zeta function was studied first by Euler before the mid eighteenth century (see, for example, [197]) although he considered it as a function of a real variable. It is Riemann who treated it as a complex function for the first time in his famous 1859 article [470].

The genesis of the multiple zeta functions can be traced back to a series of correspondences between Leonhard Euler and Christian Goldbach (see [312]). On the Christmas Eve of 1742, with different notation Goldbach wrote down some special cases of the following infinite sum on a letter to Euler:

$$\sum_{a\geq b\geq 1} \frac{1}{a^m b^n},$$

where m and n are positive integer. In modern notation this is $\zeta(m,n) + \zeta(m+n)$ where $\zeta(m,n)$ is called a double zeta value. Later Euler discovered the so-called decomposition formula for the double zeta values:

$$2\zeta(n,1) = n\zeta(n+1) - \sum_{i=1}^{n-2} \zeta(n-i)\zeta(i+1), \tag{1.23}$$

for all positive integer $n \geq 2$.

There are two independent sources of the revival of multiple zeta functions in recent years. From late 1980s to early 1990s, motivated by some problems of Moen (see [287]) Hoffman initiated the study of multiple zeta values from the viewpoint of their underlying algebraic structures in a series of papers (see [275,276]). During the same time period, Zagier was attracted to multiple zeta functions by their mysterious connections to Witten's zeta functions, Vassiliev knot invariants and the theory of mixed Tate motives over $\mathbb{Z}$. After many discussions with Drinfeld, Kontsevich and Goncharov, Zagier included section 9 in his 1994 paper [577] which became one of the most influential papers in the modern study of multiple zeta functions.

However, the analytic properties the double zeta function was considered much earlier than the general multiple variable case, similar to the situation of Euler's study of double zeta values at positive integers. The analytic continuation of the double zeta function was first realized by Atkinson [20] using the Poisson summation formula in the 1940s, and by Apostol and Vu in [17] independently much later. For arbitrary depth there are basically two different approaches: by the generalized functions or distributions of [585], or by Euler-Maclaurin summation formula of [7, 154]. In Sec. 1.3 we gave a polished proof of the second approach outlined first in [594].

Most of the current research on multiple zeta functions is on their special values. However, there are some important results concerning their analytic properties. For example, Matsumoto showed a functional equation for the double zeta functions (see [135, 406]) and Nakamura and Pankowski discovered the existence of nontrivial zeros of general multiple zeta functions in [424,425] although their precise locations are still beyond our reach. For a good survey of the analytic theory of these functions and further generalizations Matsumoto's paper [405] is a good reference for results up to around 2001.

There are some investigations of the behavior of multiple zeta functions at non-positive integers. At depth greater than 1, strictly speaking, different ways to define these values may lead to different results. This phenomenon already occurs at depth two, see Exercise 1.6. In [7], Akiyama *et al.* provided several definitions. Using Raabe's identity, Sadaoui [476] links the multiple integral

$$Y_{\mathbf{a}}(\mathbf{n}) = \int_{[1,+\infty)^r} \frac{\mathrm{d}\mathbf{x}}{(x_1+a_1)^{n_1}\cdots(x_1+a_1+\cdots+x_r+a_r)^{n_r}}$$

to the zeta function of the form

$$Z(\mathbf{n},\mathbf{z}) = \sum_{k_1\geq 1,\ldots,k_r\geq 1} \frac{1}{(k_1+z_1)^{n_1}\cdots(k_1+z_1+\cdots+k_r+z_r)^{n_r}}$$

by

$$Y_{\mathbf{0}}(\mathbf{n}) = \int_{[0,1]^r} Z(\mathbf{n},\mathbf{z})\,\mathrm{d}\mathbf{z}.$$

He then uses a classical inversion argument to obtain an analytic continuation of the multiple zeta function defined at negative integer arguments. This process has been used to derive some closed expressions for these values by Moll *et al.* [416].

In the literature there exist a few functions named as "multiple zeta functions" but are defined completely differently from Eq. (1.8). For example, Kurokawa's multiple zeta function considered in [354] is a tensor product of the general zeta functions that are defined by Euler products and are meromorphic of finite order with functional equations over the entire space. In particular, the tensor product of two zeta functions is defined by the Hadamard product with zeros and poles given by the sum of those of the two zeta functions. So Kurokawa's multiple zeta function differs drastically from ours since we do not know any kind of Euler product associated with ours.

A caveat in reading papers on multiple zeta functions is that sometimes the word "multiple" is used to refer to multiple summation indices even though the function is a single variable function. For instances, the p-adic multiple zeta function of [516] and the Barnes' multiple zeta-function of [346] are both single variable functions.

Exercises

1.1. Show that $B_{2n+1} = 0$ for all $n > 0$ by the definition Eq. (1.3).

1.2. Prove Eq. (1.4) using the functional equation given by Eq. (1.1) and Eq. (1.2).

1.3. Prove the relation between multiple zeta star function and the multiple zeta function given by Eq. (1.10).

1.4. Show that Eq. (1.8) converges if and only if $\mathrm{Re}(s_1+\cdots+s_j) > j$ for all $j = 1,\ldots,d$.

1.5. Complete the proof of Thm. 1.3.2 by induction.

1.6. Carry out the analytic continuation of the double zeta function $\zeta(s_1, s_2)$ explicitly. Then compare the two different limits

$$\lim_{s_1 \to 0} \lim_{s_2 \to 0} \zeta(s_1, s_2) \quad \text{and} \quad \lim_{s_2 \to 0} \lim_{s_1 \to 0} \zeta(s_1, s_2)$$

at $(s_1, s_2) = (0, 0)$. Are they the same? Why?

Chapter 2

Multiple polylogarithms

In this chapter we consider single- and multi-variable generalizations of the natural logarithm, called the polylogarithms (polylogs) and the multiple polylogarithms (multiple polylogs or MPLs), respectively.

2.1 Chen's iterated integrals

We briefly recall the theory of Chen's iterated integrals in this section. It is an indispensable tool in the subsequent study of the analytic properties of MPLs. For $r > 1$ define inductively

$$\int_a^b f_1(t)\,\mathrm{d}t \cdots f_r(t)\,\mathrm{d}t = \int_a^b f_1(\tau)\,\mathrm{d}\tau \cdots f_{r-1}(\tau) \left(\int_a^\tau f_r(t)\,\mathrm{d}t \right) \mathrm{d}\tau.$$

When $r = 0$ one sets the integral to be 1 as a convention. More generally, let $w_1, \ldots, w_r$ be some 1-forms (repetition allowed) on a manifold M and let $\alpha : [0,1] \to M$ be a piecewise smooth path. Write $\alpha^* w_i = f_i(t)dt$ and define the iterated integral

$$\int_\alpha w_1 \cdots w_r := \int_0^1 f_1(t)\,\mathrm{d}t \cdots f_r(t)\,\mathrm{d}t.$$

Remark 2.1.1. Here our order of the 1-forms in the iteration is opposite to Chen's original order.

The following results are crucial in the application of the Chen's theory of iterated path integrals. However, the proofs are beyond the scope of this book so we recommend the interested reader to read Chen's original papers.

Lemma 2.1.2. *Let w_i $(i \geq 1)$ be $\mathbb{C}$-valued 1-forms on a manifold M.*

(i) *The value of $\int_\alpha w_1 \cdots w_r$ is independent of the parametrization of α.*

(ii) *If* $\alpha, \beta : [0,1] \longrightarrow M$ *are composable paths (*i.e., $\alpha(1) = \beta(0)$*), then*

$$\int_{\beta\alpha} w_1 \cdots w_r = \sum_{j=0}^{r} \int_{\beta} w_1 \cdots w_i \int_{\alpha} w_{i+1} \cdots w_r$$

where $\beta\alpha$ *denotes the composition of* α *and* β *and we set* $\int_\alpha \phi_1 \cdots \phi_m = 1$ *if* $m = 0$.

(iii) *For every path* α,

$$\int_{\alpha^{-1}} w_1 \cdots w_r = (-1)^r \int_{\alpha} w_r \cdots w_1.$$

(iv) *Let* $\mathfrak{S}_n$ *be the symmetry group of* n *letters. For every path* α,

$$\int_\alpha w_1 \cdots w_r \int_\alpha w_{r+1} \cdots w_{r+s} = \sum_{\substack{\sigma \in \mathfrak{S}_{r+s} \\ \sigma(1) < \cdots < \sigma(r) \\ \sigma(r+1) < \cdots < \sigma(r+s)}} \int_\alpha w_{\sigma^{-1}(1)} \cdots w_{\sigma^{-1}(r+s)}.$$

Proof. (i) can be derived from the theorem on p. 361 of [127]. (ii) and (iii) are formulas (1.6.1) and (1.6.2) of [127] respectively. Ree [468] discovered the shuffle relation (iv) which appeared as (1.5.1) in [127]. □

Lemma 2.1.3. *If* $w_i^{(j)}$ *are closed 1-forms for* $1 \le i \le r$ *and* $1 \le j \le n$ *such that* $\sum_j w_1^{(j)} \wedge w_2^{(j)} = \sum_j w_2^{(j)} \wedge w_3^{(j)} = \cdots = \sum_j w_{r-1}^{(j)} \wedge w_r^{(j)} = 0$ *then* $\sum_j \int_\alpha w_1^{(j)} w_2^{(j)} \cdots w_r^{(j)}$ *only depends on the homotopy class of* α.

Proof. The case $r = 1$ is proved on p. 366 of [127]. The case $r = 2$ can be found on p. 368 of [127]. The general case follows from a similar argument. □

2.2 Classical polylogarithms

The Riemann zeta values $\zeta(n)$ for positive integer $n > 1$ can be regarded as a special value of the classical polylogs.

$$\mathrm{Li}_n(x) := \sum_{k=1}^{\infty} \frac{x^k}{k^n}, \qquad |x| \le 1$$

when $x = 1$. If $n = 1$ one sees easily that

$$\mathrm{Li}_1(x) := \sum_{k=1}^{\infty} \frac{x^k}{k} = -\log(1-x), |x| < 1.$$

Notice that one can express $\mathrm{Li}_n(x)$ using the following iterated integral

$$\mathrm{Li}_n(x) = \int_0^x \left(\frac{\mathrm{d}t}{t}\right)^{n-1} \frac{\mathrm{d}t}{1-t},$$

from which one obtains an analytic continuation of $\mathrm{Li}_n(x)$ to $\mathbb{C}$ as a multi-valued function with singularities at 1. Since multi-valued functions are hard to grasp there have been several attempts to find their single-valued cousins. In [373] Lewin defined the following polylog of a real variable $x < 1$:

$$L_n(x) = \sum_{j=0}^{n} \frac{(-\log|x|)^j}{j!} \mathrm{Li}_{n-j}(x) + \frac{(-\log|x|)^{n-1}}{n!} \log|1-x|, \tag{2.1}$$

which can be extended to $x \in \mathbb{R}$ by setting $L_n(x) = (-1)^{n-1} L_n(1/x)$ for $x \geq 1$. He further conjectured that $L_n(x)$ should satisfy "clean" functional equations and proved this for those equations found by Kummer with $n < 5$. Another version of single-valued polylog is defined by Ramakrishnan (see [463, 572]) for complex variable x with $|x| \leq 1$

$$D_n(x) = \begin{cases} \mathrm{Re} \\ \mathrm{Im} \end{cases} \sum_{j=0}^{n} \frac{(-\log|x|)^j}{j!} \mathrm{Li}_{n-j}(x) - \frac{(-\log|x|)^n}{2 \cdot n!} \quad \begin{array}{l} \text{if } n \text{ is odd;} \\ \text{if } n \text{ is even,} \end{array} \tag{2.2}$$

and then extend it to $\mathbb{C}$ by $D_n(x) = (-1)^{n-1} D_n(1/x)$ for $|x| \geq 1$. Notice that this function is real analytic. But it seems that Zagier's version of single-valued polylogs $\mathcal{L}i_n$ in [575] behaves the most regularly in the sense that the functional equations of $\mathcal{L}i_n$ do not involve lower order polylogs, at least for $n \leq 3$. It is defined on the one dimensional projective complex space $\mathbb{CP}^1$ for any positive integer $n \geq 2$ by

$$\mathcal{L}i_n(x) = \begin{cases} \mathrm{Re} \\ \mathrm{Im} \end{cases} \sum_{j=0}^{n} \frac{2^j B_j}{j!} \log^j|x| \, \mathrm{Li}_{n-j}(x) \quad \begin{array}{l} \text{if } n \text{ is odd;} \\ \text{if } n \text{ is even,} \end{array} \tag{2.3}$$

where B_j is the j-th Bernoulli number and $\mathrm{Li}_0(x) = 1$. These functions are all single-valued and real-analytic by straight-forward monodromy computations. In particular $\mathcal{L}i_n(0) = \mathcal{L}i_n(\infty) = \mathcal{L}i_{2n}(1) = 0$, and $\mathcal{L}i_{2n+1}(1) = \zeta(2n+1)$.

We turn now briefly to the discussion of polylog functional equations. In high school, students already know that $\log|x| + \log|y| = \log|xy|$. For dilogarithms (dilogs) one has the following identity (see [579, p. 9]): if $1/x + 1/y + 1/z = 1$ then (see Exercise 2.1)

$$\mathrm{Li}_2(x) + \mathrm{Li}_2(y) + \mathrm{Li}_2(z) = \frac{1}{2}\left[\mathrm{Li}_2\Big(-\frac{xy}{z}\Big) + \mathrm{Li}_2\Big(-\frac{yz}{x}\Big) + \mathrm{Li}_2\Big(-\frac{zx}{y}\Big)\right]. \tag{2.4}$$

In fact, this is a consequence of the so-called five-term relation discovered by Abel [1], which says, in terms of the single-valued Bloch-Wigner dilog

function $\mathcal{L}i_2(x)$, that for all $x, y \neq 0, 1$ and $x \neq y$

$$\mathcal{L}i_2(x) - \mathcal{L}i_2(y) + \mathcal{L}i_2\left(\frac{y}{x}\right) - \mathcal{L}i_2\left(\frac{1-x^{-1}}{1-y^{-1}}\right) + \mathcal{L}i_2\left(\frac{1-x}{1-y}\right) = 0. \quad (2.5)$$

For a more symmetric form, see Exercise 2.3.

Similar result holds for the trilog $\mathcal{L}i_3(x)$: for all x, y, z, we have the 22-term relations

$$\begin{aligned}\mathcal{L}i_3(-xyz) + \bigoplus_{\mathrm{Cyc}(x,y,z)} \Big\{ & \mathcal{L}i_3(zx-x+1) + \mathcal{L}i_3\big(\frac{zx-x+1}{zx}\big) \\ & - \mathcal{L}i_3\big(\frac{zx-x+1}{z}\big) + \mathcal{L}i_3\big(\frac{x(yz-z+1)}{-(zx-x+1)}\big) + \mathcal{L}i_3\big(\frac{yz-z+1}{y(zx-x+1)}\big) \\ & + \mathcal{L}i_3(z) - \mathcal{L}i_3\big(\frac{yz-z+1}{yz(zx-x+1)}\big)\Big\} = 3\,\mathcal{L}i_3(1). \quad (2.6)\end{aligned}$$

Here $\bigoplus_{\mathrm{Cyc}(x,y,z)} f(x,y,z) := f(x,y,z) + f(y,z,x) + f(z,x,y)$. This was discovered first by Goncharov [234] when he proved the following Zagier's Conjecture for $\zeta_F(3)$ where $\zeta_F(s)$ is the Dedekind zeta function over a number field F. Zagier proved his conjecture for $\zeta_F(2)$ in [571].

Conjecture 2.2.1. (Zagier's Conjecture) *Let F be a number field. Let $\sigma_1, \ldots, \sigma_{r_1} : F \to \mathbb{R}$ be the real embeddings and $\sigma_{1+r_1} = \overline{\sigma_{r_1+1+r_2}}, \ldots, \sigma_{r_1+r_2} = \overline{\sigma_{r_1+2r_2}} : F \to \mathbb{C}$ be the complex embeddings. For integer $n \geq 2$ let $r(n) = r_2$ if n is odd and $r(n) = r_1 + r_2$ if n is even. Then there exist $y_1, \ldots, y_{r(n+1)} \in \mathbb{Q}[F \setminus \{0,1\}]$ such that*

$$\zeta_F(n) = \pi^{nr(n)} D_F^{-1/2} \det\Big[\mathcal{L}i_n\left(\sigma_i(y_j)\right)\Big]_{1 \leq i,j \leq r(n+1)},$$

where D_F is the discriminant of F.

2.3 Definition of multiple polylogs

Multiple polylogs are iterated generalizations of polylogs to the multivariable situation.

Definition 2.3.1. Let d, $s_1, \ldots, s_d$ be positive integers, and let $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{C}^d$ be a complex vector variable such that $|\prod_{j=1}^{\ell} x_j| < 1$ for all $\ell = 1, \ldots, d$. The multiple polylog (MPL) function of depth d is defined by

$$\mathrm{Li}_{s_1,\ldots,s_d}(\mathbf{x}) = \sum_{k_1 > k_2 > \cdots > k_d \geq 1} \frac{x_1^{k_1} x_2^{k_2} \cdots x_d^{k_d}}{k_1^{s_1} k_2^{s_2} \cdots k_d^{s_d}}. \quad (2.7)$$

As in the definition of multiple zeta functions, the order of indices in Eq. (2.7) is sometimes reversed in research papers, even by the author himself. Note also that occasionally the term "multiple polylog" means the single variable function $\mathrm{Li}_{s_1,s_2\ldots,s_d}(t,1,1,\ldots,1)$ in the literature, which will in fact be needed in this book later (see Sec. 4.2 and Sec. 14.1).

Now by setting $a_i = \prod_{j=1}^{i} x_j^{-1}$ it is not hard to check (see Exercise 2.4) that

$$\mathrm{Li}_{s_1,\ldots,s_d}(x_1,\ldots,x_d) = \int_0^1 \left(\frac{\mathrm{d}t}{t}\right)^{s_1-1} \frac{\mathrm{d}t}{a_1-t} \cdots \left(\frac{\mathrm{d}t}{t}\right)^{s_d-1} \frac{\mathrm{d}t}{a_d-t}. \quad (2.8)$$

This is an iterated path integral in the sense of Chen (see [127, 128]) whose path lies in $\mathbb{C}$. Thus one can easily enlarge its domain of definition to some open subset of $\mathbb{C}^d$. However, it is difficult to study the monodromy of the MPLs by this integral expression.

2.3.1 *Control the divergence:* $\mathsf{DivLog}_{\mathbb{N}}$ *and* DivLog_D

Let $D = \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disk and $\mathsf{H}(D)$ be the set of holomorphic functions on D.

Definition 2.3.2. Let $\mathbb{C}^{\mathbb{N}}$ be the set of infinite sequences in $\mathbb{C}$. Let $\mathsf{DivLog}_{\mathbb{N}} \subset \mathbb{C}^{\mathbb{N}}$ such that $\mathbf{S} = (S_1, S_2, \ldots) \in \mathsf{DivLog}_{\mathbb{N}}$ if and only if there is an asymptotic expansion

$$S_n = \mathrm{As}_{\mathbf{S}}(\log n) + O(\log^{\alpha}(n)/n) \quad \text{as } n \to \infty,$$

where $\mathrm{As}_{\mathbf{S}}(t) \in \mathbb{C}[t]$ and $\alpha \in \mathbb{N}$.

Example 2.3.3. In applications we often think $\mathbf{S}$ as coming from some partial sum sequence of an infinite series. For instance, it is well-known that the partial sums of harmonic series, namely, the harmonic numbers $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ have the asymptotic formula

$$H_n = \log n + \gamma + O(\log(n)/n) \quad \text{as } n \to \infty,$$

where $\gamma \approx 0.57721566\ldots$ is the Euler's constant. Hence $\mathrm{As}_{\mathbf{h}}(t) = t + \gamma$ for $\mathbf{h} = (H_1, H_2, \ldots)$. Similarly, for any $n \in \mathbb{N}_0$, we have $\mathrm{As}_{\mathbf{a}(n)}(t) = t^n$ where $\mathbf{a}(n) = ((H_1-\gamma)^n, (H_2-\gamma)^n, \ldots)$. Thus $\mathrm{As} : \mathsf{DivLog}_{\mathbb{N}} \to \mathbb{C}[t]$ is surjective.

Definition 2.3.4. Let DivLog_D be the set of functions $f \in \mathsf{H}(D)$ whose asymptotic expansion, as $z \to 1$ in $D \cap \mathbb{R}$, is given by

$$f(z) = \mathrm{As}_f(-\log(1-z)) + O\big(|(1-z)\log^{\alpha}(1-z)|\big),$$

where $\mathrm{As}_f(t) \in \mathbb{C}[t]$ and $\alpha \in \mathbb{N}_0$.

Example 2.3.5. For any $n \in \mathbb{N}_0$, if we take $f(z) = (\mathrm{Li}_1(z))^n = (-\log(1-z))^n$ then we see that $\mathrm{As}_{\mathrm{Li}_1^n}(t) = t^n$. Thus $\mathrm{As} : \mathsf{DivLog}_D \to \mathbb{C}[t]$ is surjective.

Remark 2.3.6. Define a ring structure on $\mathbb{C}^{\mathbb{N}}$ by the componentwise addition and multiplication. Then it is not hard to see that $\mathsf{DivLog}_{\mathbb{N}}$ becomes a sub-ring of $\mathbb{C}^{\mathbb{N}}$. It is also clear that DivLog_D is a sub-ring of $\mathsf{H}(D)$.

2.3.2 *Relations between* $\mathsf{DivLog}_{\mathbb{N}}$ *and* DivLog_D

For any sequence $\mathbf{S} = (S_1, S_2, \dots)$, we define the formal power series associated with it as follows:

$$\mathrm{Fp}(\mathbf{S}) := \sum_{n\geq 1} u_n z^n \tag{2.9}$$

where $S_n = u_1 + \cdots + u_n$.

Lemma 2.3.7. *If* $\mathbf{S} \in \mathsf{DivLog}_{\mathbb{N}}$ *and* $\mathrm{As}_{\mathbf{S}} = 0$ *then the associated power series* $f = \mathrm{Fp}(\mathbf{S}) \in \mathsf{DivLog}_D$ *and* $\mathrm{As}_f = 0$.

Proof. By the definition, there is $\alpha \in \mathbb{N}$ such that

$$S_n = u_1 + \cdots + u_n = O(\log^{\alpha}(n)/n) \quad \text{as } n \to \infty.$$

By Abel's transformation we see that

$$f(z) = \sum_{n\geq 1} u_n z^n = \sum_{M\geq 1} (z^M - z^{M+1}) \sum_{n=1}^{M} u_n = (1-z) \sum_{M\geq 1} z^M O(\log^{\alpha}(M)/M).$$

We claim that

$$\sum_{M\geq 1} z^M O(\log^{\alpha}(M)/M) = O\big(|\log^{\alpha+1}(1-z)|\big). \tag{2.10}$$

To prove this, we break the infinite sum into two parts, under the assumption that $z < 1$ but is very close to 1. First, by the Integral Test

$$\sum_{0<M<1/(1-z)} z^M \frac{\log^{\alpha}(M)}{M} \leq \int_1^{1/(1-z)} \frac{\log^{\alpha}(t)}{t} \mathrm{d}t = \frac{(-\log(1-z))^{\alpha+1}}{\alpha+1}.$$

On the other hand, if $M \geq 1/(1-z)$, then $\log^{\alpha}(M)/M \leq (-\log(1-z))^{\alpha}(1-z)$ since the function $\log^{\alpha}(t)/t$ is decreasing for all $t > e^{\alpha}$. Thus

$$\sum_{M\geq 1/(1-z)} z^M \frac{\log^{\alpha}(M)}{M} \leq (-\log(1-z))^{\alpha}(1-z) \sum_{M\geq 1/(1-z)} z^M \leq z^{1/(1-z)}(-\log(1-z))^{\alpha}.$$

Since $\lim_{z\to 1^-} z^{1/(1-z)} = \lim_{x\to 0^+} (1-x)^{1/x} = 1/e$ the claim Eq. (2.10) follows immediately. This finishes the proof of the lemma. □

By Lemma 2.3.7, we see that there exists a map cp to make the following diagram commutate:

$$\begin{array}{ccc} \mathsf{DivLog}_{\mathbb{N}} & \xrightarrow{\quad\mathrm{As}\quad} & \mathbb{C}[t] \\ \Big\downarrow{\scriptstyle \mathrm{Fp}} & & \vdots{\scriptstyle \mathrm{cp}} \\ \mathsf{DivLog}_{D} & \xrightarrow[\quad\mathrm{As}\quad]{} & \mathbb{C}[t] \end{array}$$

The comparison map cp will be constructed in the proof of Prop. 13.3.28.

Proposition 2.3.8. *Let $s \in \mathbb{N}$ and $z \in \overline{D}$, where $\overline{D}$ is the closed unit disk. If the $\mathbf{S} \in \mathsf{DivLog}_{\mathbb{N}}$ then the sequence $\mathbf{T}$ defined by*

$$T_M = \sum_{n=1}^{M-1} \frac{S_n}{n^s} z^n$$

is again in $\mathsf{DivLog}_{\mathbb{N}}$. Moreover, $\mathbf{T}$ converges if $(s,z) \neq (1,1)$. If $(s,z) = (1,1)$ then $\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{As}_{\mathbf{T}}(t) = \mathrm{As}_{\mathbf{S}}(t)$.

Proof. By linearity we may assume without loss of generality that $\mathrm{As}_{\mathbf{S}}(t) = t^k$ for some $k \in \mathbb{N}_0$. Say $S_n = \log^k(n) + O(\log^\alpha(n)/n)$.

If $s \geq 2$ then the major term of $\sum_{n\geq M} T_n$ is dominated by $\sum_{n\geq M} \log^k(n)/n^2$. Set

$$I_k = \int_M^\infty \log^k(x)/x^2 \,\mathrm{d}x.$$

Integration by parts yields easily that $I_k = k\log^{k-1}(M)/M + kI_{k-1}$ with $I_0 = 1/M$. Thus $\mathbf{T}$ converges absolutely by the Integral Test.

If $s = 1$ but $z \neq 1$, then we set $u_n = \frac{S_{n-1}}{n-1} - \frac{S_n}{n}$ for $n \geq 2$ and $u_1 = -S_1$. By Abel's transformation we get

$$\sum_{n=1}^\infty \frac{S_n}{n} z^n = -\sum_{n=1}^\infty z^n \sum_{m=1}^n u_m = -\sum_{m=1}^\infty u_m \sum_{n=m}^\infty z^m = \frac{1}{z-1}\sum_{m=1}^\infty u_m z^m. \tag{2.11}$$

Observe that $u_m > 0$ for sufficiently large m as $\log^k(t)/t$ is decreasing when $t > e^k$. Thus the series Eq. (2.11) converges absolutely since $|z| \leq 1$ and $\sum u_m$ converges (see Exercise 2.2).

If $(s,z) = (1,1)$, then the big-O term contributes to $O(\log^\alpha(n)/n^2)$ which can be handled as above. So we only need to consider the leading

term $\log^k(n)/n$. Using the decomposition $\log(n+1) = \log(n) + 1/n + O(1/n^2)$ we see that

$$\frac{\log^k(n)}{n} = \frac{\log^{k+1}(n+1) - \log^{k+1}(n)}{k+1} + O(\log^k/n^2).$$

Summing up we thus get the major term for T_M as $\log^{k+1}(M)/(k+1)$. Hence $\mathrm{As}_{\mathbf{T}}(t) = t^{k+1}/(k+1)$ whose derivative is t^k.

This finishes the proof of the proposition. □

We have a quantitative version of the the case $s = 1, z \neq 1$ of the Prop. 2.3.8.

Lemma 2.3.9. *For all $k \geq 0$ we have*

$$\sum_{m=1}^{\infty} \frac{\log^k m}{m} z^{m-1} = O\left(\log^{k+1}\left(\frac{1}{1-z}\right)\right) \text{ as } z \to 1 \text{ in } D \cap \mathbb{R}.$$

Proof. We proceed with induction on k. If $k = 0$ then the left-hand side is $z^{-1}\log\left(\frac{1}{1-z}\right) = O\Big(\log\left(\frac{1}{1-z}\right)\Big)$ as $z \to 1$ in $D \cap \mathbb{R}$. Assume the lemma holds for $k = l \geq 1$ for some l. Then we have

$$\log^{l+1} m = (l+1)\int_1^m \frac{\log^l x}{x}\,\mathrm{d}x. \leq 2(l+1)\sum_{n=1}^{m} \frac{\log^l n}{n}. \tag{2.12}$$

Here we have used the fact that the function $f(x) = \log^l x/x$ is increasing on $(1, e^l)$ and then decreasing on (e^l, ∞). Hence if $m < e^l$ then Eq. (2.12) following from a simple application of integral test technique (and the 2 is unnecessary). If $m > e^l$ then suppose $a < e^l < a+1$ for some integer $a < m$. We see that the area below $f(x)$, $1 \leq x \leq m$, can be covered by the following three kinds of rectangles:

(1). $a-1$ rectangles, each based on $[i-1, i]$ with height $f(i)$, $i = 2, \cdots, a$;

(2). $m - a - 1$ rectangles, each based on $[i, i+1]$ with height $f(i)$, $i = a+1, \ldots, m-1$;

(3). one rectangle based on $[a, a+1]$ with height $f(e^l)$.

Notice that the $m-1$ nonzero terms in $\sum_{n=1}^{m} \frac{\log^l n}{n}$ gives the first $m-2$ rectangular areas in (1) and (2) and the last term $\frac{\log^l m}{m}$ is not used. The reason is the value l/e^l is larger than all the values of the $f(x)$ at the division points $x = 2, \cdots, m$. So we need to add the one rectangle in (3) by doubling the sum. In fact we only need the two terms with $n = 2$ and 3

since $f(e^l) = l/e^l < \log^l 2/2 + \log^l 3/3$ for all $l \geq 1$. From Eq. (2.12) we get

$$\begin{aligned}\sum_{m=1}^{\infty} \frac{\log^{l+1} m}{m} z^{m-1} &\leq 2(l+1) \sum_{m=1}^{\infty} \frac{z^{m-1}}{m} \sum_{n=1}^{m} \frac{\log^l n}{n} \\ &= 2(l+1) \sum_{n=1}^{\infty} \frac{\log^l n}{n} z^{n-1} \sum_{r=1}^{\infty} \frac{z^{r-1}}{r+n-1} \\ &< 2(l+1) \left(\frac{1}{t} \log \left(\frac{1}{1-z} \right) \right) \sum_{n=1}^{\infty} \frac{\log^l n}{n} z^{n-1}.\end{aligned}$$

Thus the lemma follows from induction assumption. □

For $M \in \mathbb{N}$, we may define the Mth partial sum of of the MPL

$$\mathrm{Li}^{(M)}_{s_1,\ldots,s_d}(z_1,\ldots,z_d) := \sum_{M \geq k_1 > \cdots > k_d > 0} \frac{z_1^{k_1} \cdots z_d^{k_d}}{k_1^{s_1} \cdots k_d^{s_d}}. \tag{2.13}$$

Corollary 2.3.10. *For all* $\mathbf{s} = (s_1,\ldots,s_d) \in \mathbb{N}^d$ *and* $\mathbf{z} = (z_1,\ldots,z_d) \in \overline{D}^d$, *the sequence* $\mathrm{Li}(\mathbf{s},\mathbf{z}) := \left(\mathrm{Li}^{(M)}_{\mathbf{s}}(\mathbf{z})\right)_{M \geq 1} \in \mathsf{DivLog}_{\mathbb{N}}$. *It converges if and only if* $(s_1, z_1) \neq (1,1)$.

Proof. This can be proved by Prop. 2.3.8 and an induction on the depth d. □

2.3.3 *Application of iterated integrals*

Let $a_1,\ldots,a_k \in \mathbb{C}$ and a path $\gamma : [0,1] \to \mathbb{C}$. Define

$$\mathrm{I}_\gamma(a_1,\ldots,a_k) := \int_0^1 (\gamma^* \omega_{a_1})(t_1) \int_0^{t_1} (\gamma^* \omega_{a_2})(t_2) \int_0^{t_{k-1}} (\gamma^* \omega_{a_k})(t_k), \tag{2.14}$$

where, for all $z \in \mathbb{C}$

$$\omega_z(t) = \begin{cases} \dfrac{\mathrm{d}t}{z^{-1} - t}, & \text{if } z \neq 0; \\ \dfrac{\mathrm{d}t}{t}, & \text{if } z = 0. \end{cases}$$

By Lemma 2.1.3 we see that for all path $\gamma \in \mathbb{C} \setminus \{a_1^{-1},\ldots,a_d^{-1}\}$ the integral I_γ is well defined and depends only on the homotopy class of γ.

By convention, we set $\mathrm{I}_\gamma = 0$ if $k = 0$. Also, for any $(\alpha,\beta) \in \mathbb{C}^2$, we denote by $[\alpha,\beta]$ the path $t \mapsto \alpha + (\beta - \alpha)t$, $t \in [0,1]$.

Proposition 2.3.11. *Let* $d \in \mathbb{N}$, $\mathbf{s} = (s_1,\ldots,s_d) \in \mathbb{N}^d$ *and* $z_1,\ldots,z_d \in \overline{D} \setminus \{0\}$. *Then for all* $z \in D$, *we have*
$\mathrm{Li}_{\mathbf{s}}(zz_1, z_2, \ldots, z_d) = \mathrm{I}_{[0,z]}(\{0\}^{s_1-1}, z_1, \{0\}^{s_2-1}, z_1z_2, \ldots, \{0\}^{s_d-1}, z_1z_2\cdots z_d)$.

Proof. This is essentially Eq. (2.8). See Exercise 2.4. □

2.4 Analytic continuation of multiple logarithms

In this section we describe a method of defining the analytic continuation of multiple logarithms by Chen's iterated integrals. This can be generalized to arbitrary MPLs but the notation is too cumbersome so we leave it to the interested reader (see [595]).

For any positive integer n we set $\mathfrak{L}_n(\mathbf{x}) = \mathrm{Li}_{\underbrace{1,\dots,1}_{n \text{ times}}}(\mathbf{x})$ and define the index set

$$\mathcal{I}_n = \{0,1\}^n = \{\mathbf{i} = (i_1,\dots,i_n) :\ i_t = 0 \text{ or } 1\ \forall t = 1,\dots,n\}. \tag{2.15}$$

Put $\mathbf{0} = (0,\cdots,0)$ and $\mathbf{1} = (1,\dots,1)$. The weight function on the indices $|(i_1,\dots,i_n)| = \sum_{t=1}^n i_t$ is the number of nonzero components.

Lemma 2.4.1. *Let $\mathbf{x} = (x_1,\dots,x_n)$ be an ordered set of complex variables. Let*

$$\mathbf{x}_j = (x_1,\dots,x_{j-1},x_jx_{j+1},x_{j+2},\dots,x_n)$$

for $j = 1,\dots,n-1$ and $\mathbf{x}_n = (x_1,\dots,x_{n-1})$. Then the total differential

$$\begin{aligned} d\mathfrak{L}_n(\mathbf{x}) &= \sum_{j=1}^n d_j\mathfrak{L}_n(\mathbf{x}) \\ &= \sum_{j=1}^n \Big\{ \mathfrak{L}_{n-1}(\mathbf{x}_j)dx_j/(1-x_j) + \mathfrak{L}_{n-1}(\mathbf{x}_{j-1})dx_j/x_j(x_j-1) \Big\} \end{aligned} \tag{2.16}$$

$$= \sum_{j=1}^n \mathfrak{L}_{n-1}(\mathbf{x}_j) d\log\left(\frac{1-x_{j+1}^{-1}}{1-x_j}\right), \tag{2.17}$$

where $x_{n+1} = \infty$, $\mathfrak{L}_0(\mathbf{x}_1) = 1$ and when $t = 1$ the second term in the sum of Eq. (2.16) *does not appear.*

Proof. Suppose $j > 1$ (the case $j = 1$ is simpler). Then by the definition

$$\begin{aligned} d_j\mathfrak{L}_n(\mathbf{x}) &= \sum_{k_1>\cdots>\widehat{k_j}>\cdots>k_{j-1}} \frac{x_1^{k_1}\cdots\widehat{x_j^{k_j}}\cdots x_d^{k_d}}{k_1\cdots\widehat{k_j}\cdots k_d} \sum_{k_j=k_{j+1}+1}^{k_{j-1}-1} x_j^{k_j-1}\,dx_j \\ &= \sum_{k_1>\cdots>\widehat{k_j}>\cdots>k_{j-1}} \frac{x_1^{k_1}\cdots\widehat{x_j^{k_j}}\cdots x_d^{k_d}}{k_1\cdots\widehat{k_j}\cdots k_d} \cdot \frac{x_j^{k_{j-1}-1} - x_j^{k_{j+1}}}{x_j-1}\,dx_j \\ &= \mathfrak{L}_{n-1}(\mathbf{x}_j)dx_j/(1-x_j) + \mathfrak{L}_{n-1}(\mathbf{x}_{j-1})dx_j/x_j(x_j-1), \end{aligned}$$

which proves Eq. (2.16). Equation (2.17) now follows immediately (see Exercise 2.5). □

Suppose $\mathbf{i} = (i_1, \ldots, i_n)$ has weight k and $i_{\tau_1} = \cdots = i_{\tau_k} = 1$. We define

$$\mathbf{x}(\mathbf{i}) = \mathbf{y} = (y_1, \ldots, y_k), \text{ where } y_m = \prod_{\alpha=1+\tau_{m-1}}^{\tau_m} x_\alpha, \quad 1 \le m \le k \tag{2.18}$$

with $\tau_0 = 0$. Set $w_0(\mathbf{x}) = 0$ and

$$w_t(\mathbf{x}) := d\log\Big(\frac{1 - x_{t+1}^{-1}}{1 - x_t}\Big), \quad \text{for } \ 1 \le t \le n.$$

We now define a partial order $\prec$ on $\mathcal{I}_n$ as follows: Given $\mathbf{i} = (i_1, \ldots, i_n)$ and $\mathbf{j} = (j_1, \ldots, j_n)$, $\mathbf{j} \prec \mathbf{i}$ (or, equivalently, $\mathbf{i} \succ \mathbf{j}$) if $j_t \le i_t$ for every $1 \le t \le n$ and $j_s < i_s$ for some s.

Theorem 2.4.2. *Let $\mathbf{x} = (x_1, \ldots, x_n)$ be an ordered set of complex variables. The multiple logarithm $\mathfrak{L}_n(\mathbf{x})$ is a multi-valued holomorphic function on*

$$\mathcal{S}_n' = \mathbb{C}^n \setminus \left\{ (x_1, \ldots, x_n) : \prod_{1 \le j \le n} (1 - x_j) \prod_{1 \le j < k \le n} \big(1 - x_j \cdots x_k\big) = 0 \right\} \tag{2.19}$$

and can be expressed by

$$\mathfrak{L}_n(\mathbf{x}) = \sum_{\substack{\mathbf{0} \ne \mathbf{j}_1 \prec \cdots \prec \mathbf{j}_n, \\ \mathbf{j}_1, \ldots, \mathbf{j}_n \in \mathcal{I}_n}} \int_{\mathbf{0}}^{\mathbf{x}} w_{\mathbf{j}_n - \mathbf{j}_{n-1}}(\mathbf{x}(\mathbf{j}_n)) \cdots w_{\mathbf{j}_2 - \mathbf{j}_1}(\mathbf{x}(\mathbf{j}_2)) w_1(\mathbf{x}(\mathbf{j}_1)), \tag{2.20}$$

where the subtraction $- : \mathcal{I}_n \to \mathbb{Z}$ is a binary operation such that

$$(i_1, \ldots, i_n) - (j_1, \ldots, j_n) = \begin{cases} s, \text{ if } i_s = j_s + 1 \text{ and } i_\ell = j_\ell \quad \forall \ell \ne s; \\ 0, \text{ otherwise.} \end{cases}$$

Proof. The proof is straight-forward so we leave it as an exercise (see Exercise 2.6). □

It is clear that every nonzero term in the sum of Eq. (2.20) is provided by some vector index $(\mathbf{j}_1, \ldots, \mathbf{j}_n) \in \mathcal{I}_n^n$ with $|\mathbf{j}_\ell| = \ell$ for all $\ell = 1, \ldots, n$, which we call *admissible* for convenience.

Example 2.4.3. When $n = 1$,

$$\mathrm{Li}_1(x) = \int_0^x \mathrm{d}\log\Big(\frac{1}{1 - x}\Big) = -\log(1 - x).$$

When $n = 2$, $\mathcal{I}_2 = \{(0,0), (0,1), (1,0), (1,1)\}$ and there are two admissible elements in $\mathcal{I}_2^2$: $((1,0), (1,1))$ and $((0,1), (1,1))$. Let $\mathbf{x} = (x, y)$ then

$$\mathbf{x}(1,0) = x, \quad \mathbf{x}(0,1) = xy, \quad \mathbf{x}(1,1) = (x, y).$$

Thus

$$\begin{aligned}\mathrm{Li}_{1,1}(x,y) &= \int_{\mathbf{0}}^{\mathbf{x}} w_2(\mathbf{x})w_1(\mathbf{x}(1,0)) + w_1(\mathbf{x})w_1(\mathbf{x}(0,1)) \\ &= \int_{(0,0)}^{(x,y)} \frac{\mathrm{d}y}{1-y}\frac{\mathrm{d}x}{1-x} + \left(\frac{\mathrm{d}x}{1-x} + \frac{\mathrm{d}y}{y(y-1)}\right)\frac{\mathrm{d}(xy)}{1-xy}.\end{aligned}$$

The iterated path integrals are understood in the sense of Chen. For example, if $\alpha : [0,1] \to \mathbb{C}^2$ is the path from $(0,0)$ to (x_0, y_0) then

$$\int_{(0,0)}^{(x_0,y_0)} \frac{\mathrm{d}y}{1-y}\frac{\mathrm{d}x}{1-x} = \int_0^1 \left(\int_0^t \frac{\mathrm{d}x\circ\alpha(s)}{1-x\circ\alpha(s)}\right)\frac{\mathrm{d}y\circ\alpha(t)}{1-y\circ\alpha(t)},$$

where $x\circ\alpha(s)$ (resp. $y\circ\alpha(s)$) is the x- (resp. y-) coordinate of $\alpha(s)$. When $n=3$ let $\mathbf{x}=(x,y,z)$. Then

$$\begin{aligned}&\mathbf{x}(1,0,0) = x, \quad \mathbf{x}(0,1,0) = xy, \quad \mathbf{x}(0,0,1) = xyz, \quad \mathbf{x}(1,1,0) = (x,y),\\ &\mathbf{x}(1,0,1) = (x,yz), \quad \mathbf{x}(0,1,1) = (xy,z), \quad \mathbf{x}(1,1,1) = (x,y,z).\end{aligned}$$

Thus

$$\begin{aligned}&\mathrm{Li}_{1,1,1}(x,y,z)\\ &= \int_{(0,0,0)}^{(x,y,z)} \frac{\mathrm{d}z}{1-z}\frac{\mathrm{d}y}{1-y}\frac{\mathrm{d}x}{1-x} + \frac{\mathrm{d}z}{1-z}\left(\frac{\mathrm{d}x}{1-x} + \frac{\mathrm{d}y}{y(y-1)}\right)\frac{\mathrm{d}(xy)}{1-xy}\\ &\quad + \left(\frac{\mathrm{d}x}{1-x} + \frac{\mathrm{d}y}{y(y-1)}\right)\frac{\mathrm{d}z}{1-z}\frac{\mathrm{d}(xy)}{1-xy} + \left(\frac{\mathrm{d}y}{1-y} + \frac{\mathrm{d}z}{z(z-1)}\right)\frac{\mathrm{d}(yz)}{1-yz}\frac{\mathrm{d}x}{1-x}\\ &\quad + \left(\frac{\mathrm{d}y}{1-y} + \frac{\mathrm{d}z}{z(z-1)}\right)\left(\frac{\mathrm{d}x}{1-x} + \frac{\mathrm{d}(yz)}{yz(yz-1)}\right)\frac{\mathrm{d}(xyz)}{1-xyz}\\ &\quad + \left(\frac{\mathrm{d}x}{1-x} + \frac{\mathrm{d}y}{y(y-1)}\right)\left(\frac{\mathrm{d}(xy)}{1-xy} + \frac{\mathrm{d}z}{z(z-1)}\right)\frac{\mathrm{d}(xyz)}{1-xyz}.\end{aligned}$$

2.5 Monodromy of multiple logarithms

The main goal of this section is to define the variation matrix associated with the multiple logarithms by utilizing the vector index set $\mathfrak{I}$ as defined in Eq. (2.15), together with the partial order $\prec$.

Write $\mathbf{x}=(x_1,\ldots,x_n)$ as before. We know the n-tuple logarithm $\mathfrak{L}_n(\mathbf{x})$ is related to multiple logarithms of lower weights. This can be seen easily, for instance, when we take the derivatives as given by Lemma 2.4.1.

Definition 2.5.1. Suppose $\mathbf{i}\in\mathcal{I}_n$ has weight k. We define the $\mathbf{i}$-th *retraction* map $\rho_{\mathbf{i}} : \mathcal{I}_n \to \mathfrak{I}_k$ as follows.

(1) If $\mathbf{i} \not\succ \mathbf{j}$ then $\rho_{\mathbf{i}}(\mathbf{j}) = (0,\ldots,0)$

(2) Suppose $\mathbf{i} \succ \mathbf{j}$ and $\mathbf{j}$ has weight l. Suppose further that the 1's occur at the positions $\tau_1, \ldots, \tau_k$ in $\mathbf{i}$ and $t_1, \ldots, t_l$ in $\mathbf{j}$, respectively. Then we can put $t_r = \tau_{\alpha_r}$ for $1 \leq r \leq l$. Set the entry of $\rho_{\mathbf{i}}(\mathbf{j})$ to be 1 if it is at the α_r-th $(1 \leq r \leq l)$ component and 0 otherwise.

For instance $\rho_{(1101)}\big((1100)\big) = (110) \in \mathfrak{I}_3$ and $\rho_{(1101)}\big((0100)\big) = (010)$. In general, $\rho_{\mathbf{i}}(\mathbf{i}) = \mathbf{1}_k$ for all $\mathbf{i}$ with $|\mathbf{i}| = k$.

We are now ready to define the entries of the $2^n \times 2^n$ variation matrix $\mathcal{M}_{[n]}(\mathbf{x}) := \mathcal{M}_{\underbrace{1,\ldots,1}_{n \text{ times}}}(\mathbf{x})$ associated with the n-tuple logarithm $\mathfrak{L}_n(\mathbf{x})$.

Definition 2.5.2. Fix a point $\mathbf{x} \in \mathcal{S}_n$ where

$$\mathcal{S}_n = \mathbb{C}^n \setminus \left\{ (x_1, \ldots, x_n) : \prod_{1 \leq j \leq n} x_j(1 - x_j) \prod_{1 \leq j < k \leq n} \big(1 - x_j \cdots x_k\big) = 0 \right\} \tag{2.21}$$

For $1 \leq s \leq n$ write $a_s = a_s(\mathbf{x}) =: (x_1 \cdots x_s)^{-1}$ and

$$\theta_s = \theta_s(\mathbf{x}) = \frac{\mathrm{d}t}{t - a_s} = \frac{\mathrm{d}t}{t - a_s(\mathbf{x})}.$$

(1) If $\mathbf{j} \not\succ \mathbf{i}$ then we define the $(\mathbf{i}, \mathbf{j})$-th entry of $\mathcal{M}_{[n]}(\mathbf{x})$ to be 0.

(2) Suppose $\mathbf{i} \succ \mathbf{j}$ as given in Definition 2.5.1(2). Recall from Eq. (2.18) we have

$$\mathbf{x}(\mathbf{i}) = \mathbf{y} = (y_1, \ldots, y_k), \qquad y_m = \prod_{\alpha = 1 + \tau_{m-1}}^{\tau_m} x_\alpha = \frac{a_{\tau_{m-1}}(\mathbf{x})}{a_{\tau_m}(\mathbf{x})}, \quad 1 \leq m \leq k.$$

with $\tau_0 = 0$, $a_0(\mathbf{x}) = 1$. Set $t_0 = \alpha_0 = 0$, $t_{l+1} = n + 1$, $\alpha_{l+1} = k + 1$, $a_{n+1}(\mathbf{x}) = a_{k+1}(\mathbf{y}) = 0$. Define the $(\mathbf{i}, \mathbf{j})$-th entry of $\mathcal{M}_{[n]}(\mathbf{x})$ as $(2\pi i)^l E_{\mathbf{i},\mathbf{j}}(\mathbf{x})$ where

$$\begin{aligned} E_{\mathbf{i},\mathbf{j}}(\mathbf{x}) = \gamma^k_{\rho_{\mathbf{i}}(\mathbf{j})}(\mathbf{y}) &:= (-1)^{k-l} \prod_{r=0}^{l} \int_{a_{\alpha_{r+1}}(\mathbf{y})}^{a_{\alpha_r}(\mathbf{y})} \theta_{\alpha_r + 1}(\mathbf{y}) \cdots \theta_{\alpha_{r+1} - 1}(\mathbf{y}) \\ &= (-1)^{k-l} \prod_{r=0}^{l} \int_{p_r} \theta_{\tau_{\alpha_r + 1}}(\mathbf{x}) \cdots \theta_{\tau_{\alpha_{r+1} - 1}}(\mathbf{x}) \end{aligned} \tag{2.22}$$

Here the $l + 1$ paths $p_0, \ldots, p_l$ for the $l + 1$ integrals are *independent of* $\mathbf{i}$ where p_r is any fixed contractible path from $a_{t_{r+1}}(\mathbf{x})$ to $a_{t_r}(\mathbf{x})$ in the punctured complex plane $\mathbb{C} \setminus \bigcup_{t_r < s < t_{r+1}} \{a_s(\mathbf{x})\}$, and the integral $\int_{p_r} = 1$ if $\alpha_r + 1 = \alpha_{r+1}$. We get the second equality by observing that

$$a_m(\mathbf{y}) = (y_1 \cdots y_m)^{-1} = a_{\tau_m}(\mathbf{x}) \Longrightarrow a_{\alpha_r}(\mathbf{y}) = a_{\tau_{\alpha_r}}(\mathbf{x}) = a_{t_r}(\mathbf{x}).$$

Proposition 2.5.3. *Suppose* $\mathbf{i}$ *and* $\mathbf{j}$ *are given as in Definition* 2.5.2(2). *As multi-valued functions*

$$E_{\mathbf{i},\mathbf{j}}(\mathbf{x}) \tag{2.23}$$

$$= \prod_{r=0}^{l} \mathfrak{L}_{\alpha_{r+1}-\alpha_r-1}\left(\frac{a_{t_r}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})}{a_{\tau_{\alpha_r+1}}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})},\cdots,\frac{a_{\tau_{\alpha_{r+1}-2}}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})}{a_{\tau_{\alpha_{r+1}-1}}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})}\right) \tag{2.24}$$

$$= \mathfrak{L}_{k-\alpha_l}\big(x_{1+t_l}\cdots x_{\tau_{\alpha_l+1}},\cdots,x_{1+\tau_{k-1}}\cdots x_{\tau_k}\big)\times$$
$$\times \prod_{r=0}^{l-1} \mathfrak{L}_{\alpha_{r+1}-\alpha_r-1}\left(\frac{1-x_{1+t_r}\cdots x_{t_{r+1}}}{1-x_{1+\tau_{\alpha_r+1}}\cdots x_{t_{r+1}}},\cdots,\frac{1-x_{1+\tau_{\alpha_{r+1}-2}}\cdots x_{t_{r+1}}}{1-x_{1+\tau_{\alpha_{r+1}-1}}\cdots x_{t_{r+1}}}\right). \tag{2.25}$$

Here $\mathfrak{L}_0 = 1$ *and* $a_0 = 0$.

Proof. By direct and simple calculation using substitution we get

$$(-1)^{\alpha_{r+1}-\alpha_r-1}\int_{p_r}\theta_{\tau_{\alpha_r+1}}(\mathbf{x})\cdots\theta_{\tau_{\alpha_{r+1}-1}}(\mathbf{x})$$
$$= \mathfrak{L}_{\alpha_{r+1}-\alpha_r-1}\left(\frac{a_{t_r}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})}{a_{\tau_{\alpha_r+1}}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})},\cdots,\frac{a_{\tau_{\alpha_{r+1}-2}}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})}{a_{\tau_{\alpha_{r+1}-1}}(\mathbf{x})-a_{t_{r+1}}(\mathbf{x})}\right).$$

The proposition follows immediately. □

Example 2.5.4. On the last row of $\mathcal{M}_{[n]}(\mathbf{x})$ one has

$$E_{\mathbf{1},\mathbf{j}}(\mathbf{x}) = \gamma^n_{\mathbf{j}}(\mathbf{x}) = \prod_{r=0}^{l}\mathfrak{L}_{t_{r+1}-t_r-1}\Big(\frac{a_{t_r}-a_{t_{r+1}}}{a_{t_r+1}-a_{t_{r+1}}},\cdots,\frac{a_{t_{r+1}-2}-a_{t_{r+1}}}{a_{t_{r+1}-1}-a_{t_{r+1}}}\Big)$$
$$= \prod_{r=0}^{l}\mathfrak{L}_{t_{r+1}-t_r-1}\Big(\frac{1-x_{1+t_r}\cdots x_{t_{r+1}}}{1-x_{2+t_r}\cdots x_{t_{r+1}}},\cdots,\frac{1-x_{t_{r+1}-1}x_{t_{r+1}}}{1-x_{t_{r+1}}}\Big) \tag{2.26}$$

where $\mathfrak{L}_0 = 1$ and $x_{n+1} = \infty$. In particular, $E_{\mathbf{1},\mathbf{0}} = \gamma^n_{\mathbf{0}}(\mathbf{x}) = \mathfrak{L}_n(\mathbf{x})$ and $E_{\mathbf{1},\mathbf{1}} = \gamma^n_{\mathbf{1}}(\mathbf{x}) = 1$.

For any $n \geq 2$ let $L_n = L_n(\mathbf{x}) = [C_{\mathbf{0}}\ \ldots\ C_{\mathbf{1}}]$ be the matrix with 2^n columns $C_{\mathbf{j}}$ ($\mathbf{j} \in \mathcal{I}_n$) where

$$C_{\mathbf{j}} = \sum_{\mathbf{i}\succ\mathbf{j}} \gamma^{|\mathbf{i}|}_{\rho_{\mathbf{i}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{i})\big)e_{\mathbf{i}}$$

where $e_{\mathbf{i}}$ is the $\mathbf{i}$-th standard unit column vector. In listing the columns we have used a complete order "$<$" on $\mathcal{I}_n$: If $|\mathbf{i}| < |\mathbf{j}|$ then $\mathbf{i} < \mathbf{j}$ (or,

equivalently, $\mathbf{j} > \mathbf{i}$). If $|\mathbf{i}| = |\mathbf{j}|$ then the usual lexicographic order from left to right is in force with $1 < 0$ (not $0 < 1$!).

We now fix a standard basis $\{e_{\mathbf{i}} : \mathbf{i} \in \mathcal{I}_n\}$ of $\mathbb{C}^{2^n}$ consisting of column vectors. Suppose $|\mathbf{i}| = k$. It follows from definition that the $\mathbf{i}$-th row is

$$R_{\mathbf{i}} := \sum_{\mathbf{j} \prec \mathbf{i}} (2\pi i)^{|\mathbf{j}|} \gamma^k_{\rho_{\mathbf{i}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{i})\big) e_{\mathbf{j}}^T = (2\pi i)^k e_{\mathbf{i}}^T + \sum_{\mathbf{j} \precneqq \mathbf{i}} (2\pi i)^{|\mathbf{j}|} \gamma^k_{\rho_{\mathbf{i}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{i})\big) e_{\mathbf{j}}^T \tag{2.27}$$

where $e_{\mathbf{j}}^T$ are now row vectors. Note that $\gamma^k_{\rho_{\mathbf{i}}(\mathbf{i})} = \gamma^k_{\mathbf{1}_k} = 1$ by the definition. It is clear that the first entry (*i.e.*, $\mathbf{j} = \mathbf{0}$) of this row is $\mathfrak{L}_k\big(\mathbf{x}(\mathbf{i})\big)$.

Example 2.5.5. Using the above definition we can easily get the variation matrix associated with the double logarithm:

$$\mathcal{M}_{1,1}(x,y) = \begin{bmatrix} 1 & & & \\ \mathrm{Li}_1(x) & 2\pi i & & \\ \mathrm{Li}_1(xy) & & 2\pi i & \\ \mathrm{Li}_{1,1}(x,y) & 2\pi i\,\mathrm{Li}_1(y) & 2\pi i(\mathrm{Li}_1(x) - \mathrm{Li}_1(y) - \log y) & (2\pi i)^2 \end{bmatrix}.$$

For any $k = 0, \dots, n$ let us call the minor of $\mathcal{M}_{[n]}(\mathbf{x})$ consisting of rows beginning with k-tuple logarithms (corresponding to $\mathbf{i} = (\{1\}^k, 0, \dots, 0)$) and ending just before the $(k+1)$-tuple logarithms *the k-th block*. It has $\binom{n}{k}$ rows with row indices $|\mathbf{i}| = k$.

Lemma 2.5.6. *The matrix $\mathcal{M}_{[n]}(\mathbf{x})$ is a lower triangular matrix. Moreover, the columns with $|\mathbf{j}| = k$ of the k-th block of $\mathcal{M}_{[n]}(\mathbf{x})$ is $(2\pi i)^k$ times the identity matrix of rank $\binom{n}{k}$.*

Proof. The lemma follows directly from Eq. (2.27) because if $\mathbf{j} \precneqq \mathbf{i}$ then $\mathbf{j} < \mathbf{i}$. □

Lemma 2.5.7. *The $\mathbf{j}$-th column of $\mathcal{M}_{[n]}(\mathbf{x})$ is*

$$(2\pi i)^{|\mathbf{j}|} C_{\mathbf{j}} = (2\pi i)^{|\mathbf{j}|} \sum_{\mathbf{i} \succ \mathbf{j}} \gamma^{|\mathbf{i}|}_{\rho_{\mathbf{i}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{i})\big) e_{\mathbf{i}}$$

where $\mathbf{x}(\mathbf{i})$ is defined by Eq. (2.18) depending on $\mathbf{i}$.

Proof. Use Eq. (2.27). □

Proposition 2.5.8. *The columns of $\mathcal{M}_{[n]}(\mathbf{x})$ form the set of the fundamental solutions of the following system of differential equations*

$$\begin{cases} d\,X_{\mathbf{0}} = 0, \\ d\,X_{\mathbf{i}} = \displaystyle\sum_{|\mathbf{k}| = |\mathbf{i}|-1,\, \mathbf{k} \prec \mathbf{i}} X_{\mathbf{k}}\, d\,\gamma^{|\mathbf{i}|}_{\rho_{\mathbf{i}}(\mathbf{k})}\big(\mathbf{x}(\mathbf{i})\big) \quad \textit{for all } 1 \le |\mathbf{i}| \le n \end{cases} \tag{2.28}$$

where $\mathbf{x}(\mathbf{i})$ *is determined as in Eq.* (2.18).

Proof. We prove the proposition by induction on n. It is easy to see the proposition is valid for $n = 1$ and $n = 2$. We assume that $n \geq 3$ and the proposition is true for $\leq n-1$. Let us now consider the $\mathbf{j}$-th column as expressed in Lemma 2.5.7. The cases $|\mathbf{i}| = 1$ or $\mathbf{j} > \mathbf{i}$ are obvious. Suppose

(1) $1 < |\mathbf{i}| < n$ and $\mathbf{j} \leq \mathbf{i}$. There are two cases. (i) $\mathbf{j} \not\prec \mathbf{i}$. This is trivial because each term of both sides is zero. (ii) $\mathbf{j} \prec \mathbf{i}$. Then there is a t such that $i_t = j_t = 0$. We denote $\mathbf{i}' \in \mathfrak{I}_{n-1}$ the corresponding index after deleting the i_t-th component. By induction

$$\sum_{|\mathbf{k}'|=|\mathbf{i}'|-1,\,\mathbf{j}'\prec\mathbf{k}'\prec\mathbf{i}'} \gamma^{|\mathbf{k}'|}_{\rho_{\mathbf{k}'}(\mathbf{j}')}\big(\mathbf{x}'(\mathbf{k}')\big)\, d\,\gamma^{|\mathbf{i}'|}_{\rho_{\mathbf{i}'}(\mathbf{k}')}\big(\mathbf{x}'(\mathbf{i}')\big) = d\,\gamma^{|\mathbf{i}'|}_{\rho_{\mathbf{i}'}(\mathbf{j}')}\big(\mathbf{x}'(\mathbf{i}')\big)$$

where we set $\mathbf{x}' = (x_1, \ldots, x_{i_t-1}, x_{i_t}x_{i_t+1}, x_{i_t+2}, \ldots, x_n)$. Since $|\mathbf{i}'| = |\mathbf{i}|$ and $|\mathbf{k}'| = |\mathbf{k}|$ we can get the desired equation by inserting 0 before the i_t-th components of $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$.

(2) $\mathbf{i} = \mathbf{1}$ and $|\mathbf{j}| = l$. We need to show

$$d\gamma^n_{\mathbf{j}}(\mathbf{x}) = \sum_{|\mathbf{k}|=n-1,\,\mathbf{j}\prec\mathbf{k}} \gamma^{n-1}_{\rho_{\mathbf{k}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{k})\big)\, d\gamma^n_{\mathbf{k}}(\mathbf{x}). \tag{2.29}$$

This is trivial when $l = n$. The case $l = 0$ follows from Eq. (2.17) of Lemma 2.4.1:

$$d\mathfrak{L}_n(\mathbf{x}) = \sum_{t=1}^{n} \mathfrak{L}_{n-1}(x_1, \ldots, x_{t-1}, x_t x_{t+1}, x_{t+2}, \ldots, x_n)\, d\log\frac{1-x_{t+1}^{-1}}{1-x_t}.$$

So we may assume $0 < l < n$, $j_{t_1} = \cdots = j_{t_l} = 1$ and $j_t = 0$ for all other indices t. By the definition Eq. (2.26) we have

$$d\gamma^n_{\mathbf{j}}(\mathbf{x}) = \sum_{r=0}^{l} \sum_{t_r<s<t_{r+1}} \gamma^{n-1}_{\rho_{\mathbf{v}_s}(\mathbf{j})}\big(\mathbf{x}(\mathbf{v}_s)\big)\, d\gamma^n_{\mathbf{v}_s}(\mathbf{x})$$

where $t_0 = 0$, $t_{l+1} = n+1$ and

$$\mathbf{v}_s = (1, \ldots, 1, \underset{\uparrow}{\ldots\ldots\ldots}, 1, 0, 1, \underset{\uparrow}{\ldots\ldots\ldots}, 1, \underset{\uparrow}{\ldots}, 1).$$
$$t_r\text{-th place} \qquad s\text{-th place} \qquad t_{r+1}\text{-th place}$$

Under the retraction map $\rho_{\mathbf{v}_s}$ the numbering of the indices changes as follows: $t \rightsquigarrow t$ if $t < s$ and $t \rightsquigarrow t-1$ if $t > s$. We also have

$$a_t\big(\mathbf{x}(\mathbf{v}_s)\big) = \begin{cases} a_t(\mathbf{x}) & \text{if } t < s, \\ a_{t+1}(\mathbf{x}) & \text{if } t > s. \end{cases}$$

Hence for each s such that $t_r < s < t_{r+1}$ the integral expression of $\gamma^{n-1}_{\rho_{\mathbf{k}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{k})\big)$ is unchanged under $\rho_{\mathbf{k}}$ ($\mathbf{j} \prec \mathbf{k}$) except the $\mathbf{v}_s$-term. Equation Eq. (2.29) now follows immediately from Leibniz rule and so the proposition is proved. □

Theorem 2.5.9. *Let* $\mathcal{M}_{[n]}(\mathbf{x}) = \left[E_{\mathbf{i},\mathbf{j}}(\mathbf{x})\right]_{\mathbf{i},\mathbf{j}\in\mathcal{I}_n}$ *where* $E_{\mathbf{i},\mathbf{j}}(\mathbf{x})$ *is defined by Eq.* (2.22). *Let* $1 \le j \le k \le n$ *and* $q_{jk} \in \pi_1(S_n, \mathbf{x})$ *(resp.* $1 \le j < n$ *and* q_{j0}*) enclose* $\mathcal{D}_{jk} = \{x_j \dots x_k = 1\}$, *(resp.* $\mathcal{D}_{j0} = \{x_j = 0\}$*) only once but no other irreducible component of* D_n *such that* $\int_{q_{jk}} d\log(1-x_j \dots x_k) = 2\pi i$ *(resp.* $\int_{q_{j0}} d\log x_j = 2\pi i$*). Then*

$$M(q_{j0}) = I + \left[n_{\mathbf{i},\mathbf{j}}\right]_{\mathbf{i},\mathbf{j}\in\mathcal{I}_n}, \quad M(q_{jk}) = I + \left[m_{\mathbf{i},\mathbf{j}}\right]_{\mathbf{i},\mathbf{j}\in\mathcal{I}_n}$$

where I *is the identity matrix of rank* 2^n,

$$n_{\mathbf{i},\mathbf{j}} = \begin{cases} -1 & \text{if } \mathbf{i} = \mathbf{j} + \mathbf{u}_{s+1} \text{ and } t_r \le n - j \le s \le t_{r+1} - 2,\ r \ge 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2.30}$$

and

$$m_{\mathbf{i},\mathbf{j}} = \begin{cases} 1 & \text{if } \mathbf{i} = \mathbf{j} + \mathbf{u}_{n-k+1},\ t_r = j \le n - k \le t_{r+1} - 2,\ r \ge 1, \\ -1 & \text{if } \mathbf{i} = \mathbf{j} + \mathbf{u}_i,\ t_r + 1 \le i \le n - k = t_{r+1} - 1,\ r \ge 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.31}$$

Here $\mathbf{i}$ *and* $\mathbf{j}$ *in the case of* $m_{\mathbf{i},\mathbf{j}} = \pm 1$ *and* $n_{\mathbf{i},\mathbf{j}} = -1$ *satisfy the condition in Definition* 2.5.2(2).

Proof. . If we analytically continue every integral entry of $\mathcal{M}_{[n]}(\mathbf{x})$ along a same loop $q \in \pi_1(S_n, \mathbf{x})$ we will denote this action of q by $\Theta(q)$. By the definition it is clear that if $\mathbf{i} \not\succ \mathbf{j}$ then $\Theta(q)E_{\mathbf{i},\mathbf{j}}(\mathbf{x}) = E_{\mathbf{i},\mathbf{j}}(\mathbf{x})$ which is either 0 or 1. Thus we are only concerned with $E_{\mathbf{i},\mathbf{j}}$ with $\mathbf{i} \succ \mathbf{j}$.

We now fix some $\mathbf{j}$. If $|\mathbf{j}| = n$ then clearly $(\Theta(q) - I)C_{\mathbf{1}} = [0, \dots, 0]^T$ for any loop q. This proves the proposition for $|\mathbf{j}| = n$. We now assume $|\mathbf{j}| < n$. Let $\mathbf{i}$ and $\mathbf{j}$ be given as in Definition 2.5.2(2). Then the theorem follows from the explicit expression Eq. (2.25) using the monodromy computations contained in Thm. 4.3, Props. 5.4 and 5.5 of [595]. □

Corollary 2.5.10. *The monodromy representation of* $\mathcal{M}_{[n]}(\mathbf{x})$

$$\rho_{\mathbf{x}} : \pi_1(S_n, \mathbf{x}) \longrightarrow \mathrm{GL}_{2^n}(\mathbb{Z})$$

is unipotent.

2.6 Variation matrix for MPLs

In this section we provide the procedure to construct the variation matrix $\mathcal{M}_{\mathbf{n}}(\mathbf{x})$ associated to the general MPL $\mathrm{Li}_{\mathbf{n}}(\mathbf{x})$.

Let $\mathbf{n} = (n_1, \ldots, n_d)$ and $n = |\mathbf{n}|$. Let $\mathbf{x} = (x_1, \ldots, x_d)$ and $a_j = \prod_{k=1}^{j} x_k$ for all $j = 1, \ldots, d$. We will describe precisely how to construct the entries in the last row of $\mathcal{M}_{\mathbf{n}}(\mathbf{x})$ (say, with column number c) together with the first column entry with row number also equal to c. The other rows can follow the same pattern since every row starts with some MPL by merging some of the variables in $\mathbf{x}$ (merging means multiplying some consecutive variables). The procedure always produces the correct column entries but only produces the possible candidates for row entries (called *pseudo* row entries). But the key point is that all weight $n-1$ pseudo row entries are genuine ones. And, fortunately, most of the time the pseudo row entries of lower weight are correct, too. So the most efficient way to construct the variation matrix is to write down all these pseudo entries first. Iteratively, to determine whether a certain row entry of weight k ($k < n-1$) is genuine or not one should apply the monodromy properties of the columns corresponding to those weight $k+1$ (genuine) entries to try to modify this pseudo entry. We will carry out this procedure explicitly in all the weight four double polylog cases in Sec. 2.6.2.

2.6.1 *General procedure leading to pseudo row entries*

Let $\mathbf{n}$, $\mathbf{x}$ and a_j's be the same as above. Set $a_0 = 1$ and rename $(1, \{0\}^{n_1-1}, a_1, \ldots, \{0\}^{n_d-1}, a_d, 0)$ as $(b_0, b_1, \ldots, b_n, b_{n+1})$. We now choose $i_0 < i_1 < \cdots < i_\ell < i_{\ell+1}$ subject to the following conditions:

(i) $i_0 = 0$, b_0 is the leading 1 and, $i_{\ell+1} = n+1$ and b_{n+1} is the last 0.
(ii) For all $1 \le j \le l$, either $i_j + 1 = i_{j+1}$ or one of b_{i_j} and $b_{i_{j+1}}$ is nonzero.
(iii) For fixed $1 \le j \le d$, when choosing the 0's between a_{j-1} and a_j (if there are any) always start from the first 0 (denoted by b_{i_s}) immediately after a_{j-1} and choose consecutively until, say, b_{i_t}.

Set $b'_{i_t} = b_{i_t}$ except in case (iii) if b_{i_t} is not the last 0 immediately before a_j then we define $b'_{i_t} = a_{j-1}$ (one may think of a collapse of these consecutive 0's onto a_{j-1} if not all of the 0's between a_{j-1} and a_j are chosen). For each such choice of $1 \le i_1 < \cdots < i_\ell \le n$ satisfying (i)-(iii) above we set the corresponding column entry of $\mathcal{M}_{\mathbf{n}}(\mathbf{x})$ as

$$(-1)^{\#\{j : b_{i_j} \neq 0\}} \int_0^1 \frac{\mathrm{d}t}{t - b_{i_1}} \cdots \frac{\mathrm{d}t}{t - b_{i_\ell}} \tag{2.32}$$

and the corresponding row entry as

$$I(b_0;b_1,\dots,b_{i_1-1};b_{i_1})I(b'_{i_1};b_{i_1+1},\dots,b_{i_2-1};b_{i_2})\cdots I(b'_{i_\ell};b_{i_\ell+1},\dots,b_n;b_{n+1}), \tag{2.33}$$

where

$$I(b'_k;b_{k+1},\dots,b_{m-1};b_m) = \begin{cases} 1, & \text{if } m=k+1;\\ \operatorname{sgn}(\mathbf{b}_{k,m})\displaystyle\int_{b_m}^{b'_k}\frac{\mathrm{d}t}{t-b_{k+1}}\cdots\frac{\mathrm{d}t}{t-b_{m-1}}, & \text{if } m>k+1,\end{cases}$$

where $\operatorname{sgn}(\mathbf{b}_{k,m})=1$ if $b_j=0$ for all $k<j<m$, $\operatorname{sgn}(\mathbf{b}_{k,m})=(-1)^{\#\{j:b_j\neq 0,k<j<m\}}$ if $b_k=0$, and $\operatorname{sgn}(\mathbf{b}_{k,m})=(-1)^{\#\{j:b_j\neq b_m,k<j<m\}}$ otherwise. As a check, the value in Eq. (2.33) should always be nonzero. Notice that in the above iterated integral $\mathrm{I}(b'_k;\cdots;b_m)$ we have the two boundary values reversed. Sometimes different choices of subsets of b_j's may give rise to the same column entry, then we need to combine all the corresponding pseudo row entries into one entry by adding them up (for instance, see Example 2.6.2 for $\mathcal{M}_{2,2}(x,y)$). A general choice looks like the following:

$$(1,\{0\}^{m_1-1},a_{i_1},\dots,\{0\}^{m_j-1},a_{i_j},\dots,\{0\}^{m_\ell-1},a_{i_\ell},0)$$

where $1\le i_1<\cdots<i_\ell\le d$ and the block of zeros $\{0\}^{m_j-1}$ (for $m_j>1$) must be chosen from a block of consecutive 0's between a_r and a_{r+1} (starting from the 0 immediately after a_r) for some r satisfying $0\le i_{j-1}\le r<i_j\le d$. The correspond column entry is then given by $\mathrm{Li}_{m_1,\dots,m_\ell}(1/a_{i_1},a_{i_1}/a_{i_2},\dots,a_{i_{\ell-1}}/a_{i_\ell})$. We order these entries by the following complete ordering: (i) $\mathbf{m}_1<\mathbf{m}_2$ if $|\mathbf{m}_1|<|\mathbf{m}_2|$, and (ii) if $|\mathbf{m}_1|=|\mathbf{m}_2|$ then we order by the indices $\mathbf{i}$ using the lexicographic order with $1<2<\cdots$. For example, $\mathrm{Li}_2(x)$ (with $(i_1)=(1)$) comes before $\mathrm{Li}_{1,1}(x,y)$ (with $(i_1,i_2)=(1,1)$) which in turn comes before $\mathrm{Li}_2(xy)$ (with $(i_1)=(2)$).

2.6.2 *Some examples in weight 4*

In this section we use the notation $\ell_n(x)=\log^n(x)/n!$ and set $a_0=1$, $a_1=1/x$ and $a_2=1/(xy)$. For each choice of $1\le i_1<\cdots<i_\ell\le n$ given in the last subsection we write $\alpha\otimes\beta$ to mean the corresponding column entry α and pseudo row entry β.

Example 2.6.1. . For the double polylog $\mathrm{Li}_{1,3}(x,y)$ we have

$$\mathrm{Li}_{1,3}(x,y)=I(1;a_1,0,0,a_2;0).$$

Here are the possible choices:

(1) $\emptyset$: $1 \otimes \mathrm{Li}_{1,3}(x,y)$
(2) $\{a_1\}$: $-I(1;a_1;0) \otimes -I(a_1;0,0,a_2;0) = \mathrm{Li}_1(x) \otimes \mathrm{Li}_3(y)$
(3) $\{a_2\}$: $-I(1;a_2;0) \otimes -I(1;a_1,0,0;a_2) = \mathrm{Li}_1(xy) \otimes e(x,y)$
(4) $\{a_1,a_2\}$: $I(1;a_1,a_2;0) \otimes I(a_1;0,0;a_2) = \mathrm{Li}_{1,1}(x,y) \otimes \ell_2(y)$
(5) $\{0,a_2\}$: $-I(1;0,a_2;0) \otimes -I(1;a_1;0)I(a_1;0;a_2) = \mathrm{Li}_2(xy) \otimes \beta_5 = \mathrm{Li}_2(xy) \otimes \mathrm{Li}_1(x)\ell_1(y)$ (There is a collapse of 0. And this produces the only pseudo row entry in this case.)
(6) $\{a_1,0,a_2\}$: $I(1;a_1,0,a_2;0) \otimes I(a_1;0;a_2) = \mathrm{Li}_{1,2}(x,y) \otimes \ell_1(y)$ (There is a collapse of 0.)
(7) $\{0,0,a_2\}$: $I(1;0,0,a_2;0) \otimes -I(1;a_1;0) = \mathrm{Li}_3(xy) \otimes \mathrm{Li}_1(x)$
(8) $\{0,a_1,0,a_2\}$: $\mathrm{Li}_{1,3}(x,y) \otimes 1$

Notice that variation matrix $\mathcal{M}_3(xy)$ is well-known (see [261]) and $\mathcal{M}_{1,2}(x,y)$ is given in Exercise 2.10. Thus the variation matrix $\mathcal{M}_{1,3}(x,y)$ associated with $\mathrm{Li}_{1,3}(x,y)$ over S_2 is

$$\begin{bmatrix}
1 & & & & & & & \\
\mathrm{Li}_1(x) & 1 & & & & & & \\
\mathrm{Li}_1(xy) & 0 & 1 & & & & & \\
\mathrm{Li}_{1,1}(x,y) & \mathrm{Li}_1(y) & h(x,y) & 1 & & & & \\
\mathrm{Li}_2(xy) & 0 & \ell_1(xy) & 0 & 1 & & & \\
\mathrm{Li}_{1,2}(x,y) & \mathrm{Li}_2(y) & f(x,y) & \ell_1(y) & \mathrm{Li}_1(x) & 1 & & \\
\mathrm{Li}_3(xy) & 0 & \ell_2(xy) & 0 & \ell_1(xy) & 0 & 1 & \\
\mathrm{Li}_{1,3}(x,y) & \mathrm{Li}_3(y) & e(x,y) & \ell_2(y) & \beta_5 \rightsquigarrow r_5 & \ell_1(y) & \mathrm{Li}_1(x) & 1
\end{bmatrix} \tau_{1,3}(2\pi i)$$

where

$$\begin{aligned}
h(x,y) &= -I(1;a_1;a_2) = -I\Big(1;\frac{a_1-a_2}{1-a_2};0\Big) \\
&= \mathrm{Li}_1\Big(\frac{1-xy}{1-y}\Big) = \mathrm{Li}_1(x) - \mathrm{Li}_1(y^{-1}), \\
f(x,y) &= I(1;a_1,0;a_2) = I\Big(1;\frac{a_1-a_2}{1-a_2},\frac{-a_2}{1-a_2};0\Big) \\
&= \mathrm{Li}_{1,1}\Big(\frac{1-xy}{1-y},1-y\Big) = \mathrm{Li}_2(y^{-1}) - \mathrm{Li}_2(x) + \ell_1(xy)\,\mathrm{Li}_1(x), \\
e(x,y) &= -I(1;a_1,0,0;a_2) \\
&= -I\Big(1;\frac{a_1-a_2}{1-a_2},\frac{-a_2}{1-a_2},\frac{-a_2}{1-a_2};0\Big) = \mathrm{Li}_{1,1,1}\Big(\frac{1-xy}{1-y},1-y,1\Big) \\
&= \mathrm{Li}_3(x) - \mathrm{Li}_3(y^{-1}) - \ell_1(xy)\,\mathrm{Li}_2(x) + \ell_2(xy)\,\mathrm{Li}_1(x)
\end{aligned}$$

by using Exercise 2.14, equation Eq. (2.49). Here $\beta_5 \rightsquigarrow r_5$ means that we need replace the pseudo row entry β_5 by

$$r_5 = \ell_1(xy)\,\mathrm{Li}_1(x) - \mathrm{Li}_2(x).$$

This can be found by taking monodromy of $e(x,y)$ about the four divisors $\{x=0\}$, $\{y=0\}$, $\{x=1\}$, and $\{y=1\}$. We do not need to consider the column on top of $\mathrm{Li}_3(y)$ as it will contribute nothing because of the 0 on its 5th row. Notice the diagonal entry 1 lying on top of β_5 is on the 5th row, or in other words, the column number of β_5 is 5.

Example 2.6.2. . For $\mathrm{Li}_{2,2}(x,y)$ we have

$$\mathrm{Li}_{2,2}(x,y) = I(1;0,a_1,0,a_2;0) = I(1;0_1,a_1,0_2,a_2;0).$$

Here, in order to distinguish the two interior 0's we have used the set $\{0_1,a_1,0_2,a_2\}$. Here are the possible choices:

(1) $\emptyset$: $1\otimes \mathrm{Li}_{2,2}(x,y)$
(2) $\{a_1\}$: $-I(1;a_1;0)\otimes -I(1;0_1;a_1)I(a_1;0_2,a_2;0) = \mathrm{Li}_1(x)\otimes \ell_1(x)\,\mathrm{Li}_2(y)$
(3) $\{a_2\}$: $-I(1;a_2;0)\otimes -I(1;0_1,a_1,0_2;a_2) = \mathrm{Li}_1(xy)\otimes k(x,y)$
(4) $\{0_1,a_1\}$: $-I(1;0_1,a_1;0)\otimes -I(a_1;0_2,a_2;0) = \mathrm{Li}_2(x)\otimes \mathrm{Li}_2(y)$
(5) $\{0_1,a_2\}$: $-I(1;0_1,a_2;0)\otimes -I(0_1;a_1,0_2;a_2) = \mathrm{Li}_2(xy)\otimes \mathrm{Li}_2(y^{-1})$ (Note there is no collapse of 0's! Row entry $\mathrm{Li}_2(y^{-1})$ is to be combined with $\mathrm{Li}_2(x)$ from the next case.)
(6) $\{0_2,a_2\}$: $-I(1;0_2,a_2;0)\otimes -I(1;0_1,a_1;0_2) = \mathrm{Li}_2(xy)\otimes \mathrm{Li}_2(x)$
(7) $\{a_1,a_2\}$: $I(1;a_1,a_2;0)\otimes I(1;0_1;a_1)I(a_1;0_2;a_2) = \mathrm{Li}_{1,1}(x,y)\otimes \ell_1(x)\ell_1(y)$
(8) $\{a_1,0_2,a_2\}$: $I(1;a_1,0_2,a_2;0)\otimes I(1;0_1;a_1) = \mathrm{Li}_{1,2}(x,y)\otimes \ell_1(x)$
(9) $\{0_1,a_1,a_2\}$: $I(1;0_1,a_1,a_2;0)\otimes I(a_1;0_2;a_2) = \mathrm{Li}_{2,1}(x,y)\otimes \ell_1(y)$
(10) $\{0,a_1,0,a_2\}$: $\mathrm{Li}_{2,2}(x,y)\otimes 1$

Thus the variation matrix $\mathcal{M}_{2,2}(x,y)$ associated with $\mathrm{Li}_{2,2}(x,y)$ over S_2 is

$$\begin{bmatrix}
1 & & & & & & & & \\
\mathrm{Li}_1(x) & 1 & & & & & & & \\
\mathrm{Li}_1(xy) & 0 & 1 & & & & & & \\
\mathrm{Li}_2(x) & \ell_1(x) & 0 & 1 & & & & & \\
\mathrm{Li}_{1,1}(x,y) & \mathrm{Li}_1(y) & h(x,y) & 0 & 1 & & & & \\
\mathrm{Li}_2(xy) & 0 & \ell_1(xy) & 0 & 0 & 1 & & & \\
\mathrm{Li}_{2,1}(x,y) & \ell_1(x)\,\mathrm{Li}_1(y) & g(x,y) & \mathrm{Li}_1(y) & \ell_1(x) & \alpha(y) & 1 & & \\
\mathrm{Li}_{1,2}(x,y) & \mathrm{Li}_2(y) & f(x,y) & 0 & \ell_1(y) & \mathrm{Li}_1(x) & 0 & 1 & \\
\mathrm{Li}_{2,2}(x,y) & \ell_1(x)\,\mathrm{Li}_2(y) & k(x,y) & \mathrm{Li}_2(y) & \ell_1(x)\ell_1(y) & \begin{matrix}\mathrm{Li}_2(y^{-1})\\ +\mathrm{Li}_2(x)\end{matrix} & \ell_1(y) & \ell_1(x) & 1
\end{bmatrix}\tau'$$

where $\tau' = \tau_{2,2}(2\pi i)$ and

$$\begin{aligned}\alpha(y) &= -\operatorname{Li}_1(y^{-1}) = -\operatorname{Li}_1(y) - \ell_1(y),\\ g(x,y) &= I(1;0_1,a_1;a_2) = I\Big(1;\frac{-a_2}{1-a_2},\frac{a_1-a_2}{1-a_2};0\Big)\\ &= \operatorname{Li}_{1,1}\Big(1-xy,\frac{1}{1-y}\Big) = \operatorname{Li}_2(x) - \operatorname{Li}_2(y^{-1}) - \ell_1(xy)\operatorname{Li}_1(y^{-1})\end{aligned}$$

by using Eq. (2.45), and

$$\begin{aligned}k(x,y) &= -I(1;0_1,a_1,0_2;a_2) = -I\Big(1;\frac{-a_2}{1-a_2},\frac{a_1-a_2}{1-a_2},\frac{-a_2}{1-a_2};0\Big)\\ &= \operatorname{Li}_{1,1,1}\Big(1-xy,\frac{1}{1-y},1-y\Big)\\ &= 2\operatorname{Li}_3(x) - \ell_1(x)\operatorname{Li}_2(x) - 2\operatorname{Li}_3(y^{-1}) + \ell_1(y)\operatorname{Li}_2\Big(\frac{1-xy}{x(1-y)}\Big)\\ &\quad - \ell_1(y)\operatorname{Li}_2(y^{-1}) - \ell_1(xy)\Big(\operatorname{Li}_2(y^{-1}) - \ell_1(y)\operatorname{Li}_1(y^{-1})\Big)\\ &\quad - \ell_1(y)\operatorname{Li}_2\Big(1-\frac{1}{xy}\Big) - \ell_1(y)\operatorname{Li}_2\Big(\frac{1-xy}{1-y}\Big)\end{aligned}$$

by using Exercise 2.14, equation Eq. (2.49) and the following identities:

$$\operatorname{Li}_3(y^{-1}) - \operatorname{Li}_3(y) = \ell_3(y), \quad \operatorname{Li}_2(y^{-1}) + \operatorname{Li}_2(y) = 2\operatorname{Li}_2(1) - \ell_2(y). \quad (2.34)$$

Notice in particular that all the pseudo entries are genuine in this case.

Example 2.6.3. . For $\operatorname{Li}_{3,1}(x,y)$ we have

$$\operatorname{Li}_{3,1}(x,y) = I(a_0;0,0,a_1,a_2;0) = I(1;0,0,a_1,a_2;0).$$

We have the following possible choices of the subsets of $\{0,0,a_1,a_2\} = \{0_1,0_2,a_1,a_2\}$:

(1) $\emptyset$: $1 \otimes \operatorname{Li}_{3,1}(x,y)$
(2) $\{a_1\}$: $-I(1;a_1;0) \otimes -I(1;0,0;a_1)I(a_1;a_2;0) = \operatorname{Li}_1(x) \otimes \ell_2(x)\operatorname{Li}_1(y)$
(3) $\{a_2\}$: $-I(1;a_2;0) \otimes -I(1;0,0,a_1;a_2) = \operatorname{Li}_1(xy) \otimes l(x,y)$
(4) $\{0_1,a_1\}$: $-I(1;0_1,a_1;0) \otimes -I(1;0_2;a_1)I(a_1;a_2;0) = \operatorname{Li}_2(x) \otimes \ell_1(x)\operatorname{Li}_1(y)$ (collapse of 0_1 onto $a_0 = 1$)
(5) $\{0_1,a_2\}$: $-I(1;0_1,a_2;0) \otimes I(1;0_2,a_1;a_2) = \operatorname{Li}_2(xy) \otimes g(x,y)$ (collapse of 0_1. This gives the only pseudo row entry in this case.)
(6) $\{a_1,a_2\}$: $I(1;a_1,a_2;0) \otimes I(1;0,0;a_1) = \operatorname{Li}_{1,1}(x,y) \otimes \ell_2(x)$
(7) $\{0_1,a_1,a_2\}$: $I(1;0_1,a_1,a_2;0) \otimes I(1;0_2;a_1) = \operatorname{Li}_{2,1}(x,y) \otimes \ell_1(x)$ (collapse of 0_1)
(8) $\{0,0,a_1\}$: $-I(1;0,0,a_1;0) \otimes -I(a_1;a_2;0) = \operatorname{Li}_3(xy) \otimes \operatorname{Li}_1(y)$

(9) $\{0,0,a_2\}$: $-I(1;0,0,a_2;0)\otimes -I(0;a_1;a_2)$
(10) $\{0,0,a_1,a_2\}$: $\mathrm{Li}_{3,1}(x,y)\otimes 1$

Thus the variation matrix $\mathcal{M}_{3,1}(x,y)$ associated with $\mathrm{Li}_{3,1}(x,y)$ over S_2 is

$$\begin{bmatrix}
1 & & & & & & & & & \\
\mathrm{Li}_1(x) & 1 & & & & & & & & \\
\mathrm{Li}_1(xy) & 0 & 1 & & & & & & & \\
\mathrm{Li}_2(x) & \ell_1(x) & 0 & 1 & & & & & & \\
\mathrm{Li}_{1,1}(x,y) & \mathrm{Li}_1(y) & h(x,y) & 0 & 1 & & & & & \\
\mathrm{Li}_2(xy) & 0 & \ell_1(xy) & 0 & 0 & 1 & & & & \\
\mathrm{Li}_3(x) & \ell_2(x) & 0 & \ell_1(x) & 0 & 0 & 1 & & & \\
\mathrm{Li}_{2,1}(x,y) & \ell_1(x)\mathrm{Li}_1(y) & g(x,y) & \mathrm{Li}_1(y) & \ell_1(x) & \alpha(y) & 0 & 1 & & \\
\mathrm{Li}_3(xy) & 0 & \ell_2(xy) & 0 & 0 & \ell_1(xy) & 0 & 0 & 1 & \\
\mathrm{Li}_{3,1}(x,y) & \ell_2(x)\mathrm{Li}_1(y) & l(x,y) & \ell_1(x)\mathrm{Li}_1(y) & \ell_2(x) & g\rightsquigarrow\tilde{g} & \mathrm{Li}_1(y) & \ell_1(x) & \alpha(y) & 1
\end{bmatrix}\tilde{\tau}$$

where $\tilde{\tau} = \tau_{3,1}(2\pi i)$, $\tilde{g}(x,y) = g(x,y) - \mathrm{Li}_2(x) = -\mathrm{Li}_2(y^{-1}) - \ell_1(xy)\,\mathrm{Li}_1(y^{-1})$ and

$$\begin{aligned}
l(x,y) =& -I(1;0,0,a_1;a_2) = -I\Big(1;\frac{-a_2}{1-a_2},\frac{-a_2}{1-a_2},\frac{a_1-a_2}{1-a_2};0\Big)\\
=& \,\mathrm{Li}_{1,1,1}\Big(1-xy,1,\frac{1}{1-y}\Big)\\
=& \,\mathrm{Li}_3(x)-\mathrm{Li}_3(y^{-1})-\ell_1(xy)\,\mathrm{Li}_2(y^{-1})-\ell_2(xy)\,\mathrm{Li}_1(y^{-1}).
\end{aligned}$$

by using Exercise 2.14, Eqs. (2.49), (2.34) and the following identities:

$$\begin{aligned}
\mathrm{Li}_3(1-y^{-1})+\mathrm{Li}_3(1-y)+\mathrm{Li}_3(y) =& \,\mathrm{Li}_3(1)+\ell_2(y)\,\mathrm{Li}_1(y)+\ell_3(y),\\
\mathrm{Li}_2(1-y^{-1})+\mathrm{Li}_2(1-y) =& -\ell_2(y).
\end{aligned}$$

We find $\tilde{g}(x,y)$ by considering the monodromy of $l(x,y)$. Notice that the weight 3 entry $\ell_2(x)\,\mathrm{Li}_1(y)$ does not affect $\tilde{g}(x,y)$ at all.

2.6.3 *A general result*

Usually it is tedious to write down explicitly the variation matrix associated with any given MPL $\mathrm{Li}_{m_1,\ldots,m_n}(\mathbf{x})$. However, the following general result has been proved by Deligne and Goncharov [156] (see next section for relevant definitions):

> *The MPL* $\mathrm{Li}_{n_1,\ldots,n_d}(\mathbf{x})$ *underlies a good unipotent graded-polarizable variation of mixed Hodge-Tate structures* $(V_{n_1,\ldots,n_d},W_\bullet,\mathcal{F}^\bullet)$ *over*
>
> $$\mathbb{C}^d\setminus\Big\{\prod_{i=1}^d x_i(1-x_i)\prod_{1\le i<j\le d}(1-x_i\ldots x_j)=0\Big\}$$

with the weight-graded quotients gr^W_{-2k} *being given by* c_k *copies of the Tate structure* $\mathbb{Z}(k)$ *which are nonzero only for* $0 \le k \le n := n_1 + \cdots + n_d$.

Here c_k is the number of different ways to pick ordered $(k+2)$-tuples $(b_{i_0}, \ldots, b_{i_{k+1}})$ from the ordered numbers $(b_0, \ldots, b_{n+1})$ in the following tableau

$$\Big| b_0 \Big| \cdots \Big| b_{n+1} \Big| = \Big| a_0 \Big| \underbrace{0 \Big| \cdots \Big| 0}_{n_1-1 \text{ times}} \Big| a_1 \Big| \cdots\cdots \Big| a_{d-1} \Big| \underbrace{0 \Big| \cdots \Big| 0}_{n_d-1 \text{ times}} \Big| a_d \Big| 0 \Big| \tag{2.35}$$

such that all three conditions (i)-(iii) above are satisfied. It is apparent that

$$c_k \ge d_k(n_1, \ldots, n_d) := \sum_{\substack{k_1+\cdots+k_d=k \\ 0 \le k_j < n_j \ \forall j}} 1.$$

Each term in the sum corresponds to the following choice: for every $j = 1, \ldots, n$, choose k_j 0's immediately after a_{j-1}.

Example 2.6.4. By the definition, we always have $c_0 = c_n = 1$. When $n_1 = \cdots = n_d = 1$ tableau Eq. (2.35) becomes

$$\Big| b_0 \Big| \cdots \Big| b_{n+1} \Big| = \Big| \underbrace{1 \Big| \cdots \Big| 1}_{n+1 \text{ times}} \Big| 0 \Big|$$

Hence c_k is the number of ways to choose k elements from the set $\{b_1, \ldots, b_n\}$, *i.e.*, $c_k = \binom{n}{k}$.

For ease of statement let us put a box $\boxed{\ }$ on a number whenever we choose it.

Example 2.6.5. This example will help you to do Exercise 2.9 and 2.10. First let's consider $\mathrm{Li}_{2,1}$. There are six nontrivial ways to put boxes on $\big| 1 \big| a_1 \big| a_2 \big| 0 \big| 0 \big|$:

(1) $\Big|\boxed{1}\Big| 0 \Big|\boxed{a_1}\Big| a_2 \Big|\boxed{0}\Big|$ (2) $\Big|\boxed{1}\Big| 0 \Big| a_1 \Big|\boxed{a_2}\Big|\boxed{0}\Big|$

(3) $\Big|\boxed{1}\Big|\boxed{0}\Big|\boxed{a_1}\Big| a_2 \Big|\boxed{0}\Big|$ (4) $\Big|\boxed{1}\Big| 0 \Big|\boxed{a_1}\Big|\boxed{a_2}\Big|\boxed{0}\Big|$

(5) $\Big|\boxed{1}\Big|\boxed{0}\Big| a_1 \Big|\boxed{a_2}\Big|\boxed{0}\Big|$ (6) $\Big|\boxed{1}\Big|\boxed{0}\Big|\boxed{a_1}\Big|\boxed{a_2}\Big|\boxed{0}\Big|$

Together with the trivial choice $\Big|\boxed{1}\Big| 0 \Big| a_1 \Big| a_2 \Big|\boxed{0}\Big|$ we get $c_0 = c_3 = 1$, $c_1 = 2$ and $c_2 = 3$.

However, for $\mathrm{Li}_{1,2}$ we have altogether only six ways to do this:

(1) $\left|\boxed{1}\right| a_1 \left|\right. 0 \left|\right. a_2 \left|\boxed{0}\right|$ (2) $\left|\boxed{1}\right|\left|\boxed{a_1}\right| 0 \left|\right. a_2 \left|\boxed{0}\right|$

(3) $\left|\boxed{1}\right| a_1 \left|\right. 0 \left|\boxed{a_2}\right|\boxed{0}\right|$ (4) $\left|\boxed{1}\right|\left|\boxed{a_1}\right| 0 \left|\boxed{a_2}\right|\boxed{0}\right|$

(5) $\left|\boxed{1}\right| a_1 \left|\boxed{0}\right|\boxed{a_2}\right|\boxed{0}\right|$ (6) $\left|\boxed{1}\right|\boxed{a_1}\right|\boxed{0}\right|\boxed{a_2}\right|\boxed{0}\right|$

Thus $c_0 = c_3 = 1$, $c_1 = c_2 = 2$.

2.7 Variations of mixed Hodge structures of MPLs

2.7.1 *Definition of variations of mixed Hodge structures: a review*

In this section we briefly review the theory of variations of mixed Hodge structures (for which we use the abbreviate MHS only in this chapter). See [118] for another short introduction to variations of Hodge structures.

A pure ($\mathbb{Z}$-)Hodge structure (HS) of weight k consists of a finitely generated abelian group $H(\mathbb{Z})$ and a decreasing Hodge filtration $\mathcal{F}^\bullet$ on $H(\mathbb{C}) := H(\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{C}$ such that $H(\mathbb{C}) = \mathcal{F}^p \oplus \overline{\mathcal{F}^{k-p+1}}$ for all integers p. Here the "bar" is the complex conjugation on the second factor of the tensor product. A special example is the Tate structure $\mathbb{Z}(-k)$ of weight $2k$ consists of $H(\mathbb{Z}) = \mathbb{Z}$ and the filtration $\mathcal{F}^p = 0$ for $p > k$ and $\mathcal{F}^p = H(\mathbb{C})$ for $p \leq k$. If we replace $\mathbb{Z}$ by $\mathbb{Q}$ in the above then we get a pure ($\mathbb{Q}$-)HS of weight k.

A MHS consists of a finitely generated abelian group $H(\mathbb{Z})$ and two filtrations: an increasing weight filtration $W_\bullet$ on $H(\mathbb{Q}) := H(\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Q}$ and a decreasing filtration $\mathcal{F}^\bullet$ on $H(\mathbb{C})$, which are compatible in the following sense. On each graded piece of the weight filtration $\mathrm{gr}_k^W = W_k/W_{k-1}$ the induced Hodge filtration determined by

$$\mathcal{F}^p(\mathrm{gr}_k^W)(\mathbb{C}) = {(\mathcal{F}^p \cap W_k(\mathbb{C}) + W_{k-1}(\mathbb{C}))}\big/{W_{k-1}(\mathbb{C})}$$

is a pure Hodge structure of weight k where $W_k(\mathbb{C}) := W_k \otimes_{\mathbb{Z}} \mathbb{C}$. If all the pure Hodge structures induced as above are direct sums of Tate structures then we say the MHS is a Tate structure. For a mixed Hodge-Tate structure we can put a framing as in [42, § 1.3.4, § 1.4].

Following Steenbrink and Zucker (see Definitions 3.1, 3.2 and 3.4 in [508]) we have

Definition 2.7.1. A *variation of HS of weight* k defined over $\mathbb{Q}$ and a complex manifold S is a collection of data $(\mathbb{V}_{\mathbb{Q}}, \mathcal{F}^\bullet)$ where

(a) $\mathbb{V}_{\mathbb{Q}}$ is a locally constant sheaf (local system) of $\mathbb{Q}$-vector spaces on S,

(b) $\mathcal{F}^\bullet$ is a decreasing filtration by holomorphic subbundles of the locally free sheaf $\mathcal{V} = \mathcal{O}_S \otimes_{\mathbb{Q}} \mathbb{V}_{\mathbb{Q}}$.

(c) At each $s \in S$, $\mathcal{F}^\bullet$ induces the Hodge filtration $\mathcal{F}_s^\bullet$ of a Hodge structure of weight k on the fiber $\mathcal{V}_s$ of $\mathcal{V}$ such that

(i) whenever $p + q = k$ one has $\mathcal{V}_s = \mathcal{F}_s^p \oplus \overline{\mathcal{F}_s^{q+1}}$, where the "bar" denotes the complex conjugation,

(ii) equivalently, one has $\mathcal{V}_s = \bigoplus_{p+q=k} H_s^{p,q}$ where $H_s^{p,q} = \mathcal{F}_s^p \cap \overline{\mathcal{F}_s^q}$.

(d) (Griffiths transversality) Under the connection ∇ in $\mathcal{V}$,

$$\nabla \mathcal{F}^p \subset \Omega_S^1 \otimes_{\mathcal{O}_S} \mathcal{F}^{p+1} \qquad \text{for all } p.$$

Definition 2.7.2. *A polarization over* $\mathbb{Q}$ of a variation of Hodge structure of weight k over $\mathbb{Q}$ is a non-degenerated and flat bilinear pairing:

$$\beta : \mathbb{V}_{\mathbb{Q}} \times \mathbb{V}_{\mathbb{Q}} \longrightarrow \mathbb{Q},$$

such that β is $(-1)^k$-symmetric, and the Hermitian form $\beta_s(C_s v, \overline{w})$ is positive on each fiber. Here C_s denotes the Weil operator with respect to $\mathcal{F}_s$, namely the direct sum of multiplication by i^{p-1} on $H_s^{p,q}$. A variation is called *polarizable* (over $\mathbb{Q}$) if it admits a polarization (over $\mathbb{Q}$).

Definition 2.7.3. A *variation of MHS* defined over $\mathbb{Q}$ and a complex manifold S is a collection of data $(\mathbb{V}_{\mathbb{Q}}, W_\bullet, \mathcal{F}^\bullet)$ where

(a) $\mathbb{V}_{\mathbb{Q}}$ is a local system of $\mathbb{Q}$-vector spaces on S,

(b) $W_\bullet$ is an increasing filtration of the $\mathbb{V}_{\mathbb{Q}}$ by local subsystems,

(c) $\mathcal{F}^\bullet$ is a decreasing filtration by holomorphic subbundles of $\mathcal{V} = \mathcal{O}_S \otimes_{\mathbb{Q}} \mathbb{V}_{\mathbb{Q}}$.

(d) $\nabla \mathcal{F}^p \subset \Omega_S^1 \otimes_{\mathcal{O}_S} \mathcal{F}^{p+1}$ for all p.

(e) The data

$$\left(\mathrm{gr}_k^W \mathbb{V}_{\mathbb{Q}}, \mathcal{F}^\bullet \left(\mathcal{O}_S \otimes_E W_k / \mathcal{O}_S \otimes_E W_{k-1} \right) \right) \tag{2.36}$$

is a variation of HS of weight k defined over $\mathbb{Q}$; or equivalently, on the fiber over $s \in S$, $(V_s, W_s, \mathcal{F}_s)$ is a MHS defined over $\mathbb{Q}$.

(f) If the induced collection of variations of HS Eq. (2.36) are all polarizable then the MHS is called graded-polarizable.

Remark 2.7.4. By extension of scalars in $\mathbb{V}_{\mathbb{Q}}$ one can define $\mathbb{V}_{\mathbb{F}}$ for any field $\mathbb{F}$ such that $\mathbb{Q} \subset \mathbb{F} \subset \mathbb{R}$.

Giving a local system $\mathbb{V}_{\mathbb{Q}}$ is equivalent to specifying its monodromy representation

$$\rho_x : \pi_1(S, x) \longrightarrow \mathrm{Aut}_{\mathbb{Q}}\, \mathcal{V}_x.$$

A variation is called *unipotent* if this representation is unipotent. From [265, Prop. 1.3] we know that a variation of MHS $(\mathbb{V}_{\mathbb{Q}}, W_\bullet, \mathcal{F}^\bullet)$ is unipotent if and only if each of the variations of Hodge structure $\mathrm{gr}_k^W\, \mathbb{V}_{\mathbb{Q}}$ is constant.

In general, the behavior of a variation of MHS over a non-compact base S at "infinity" is very hard to control. In [508] Steenbrink and Zucker consider the case when S is a curve and define the admissibility condition at infinity. For higher dimensional S, Kashiwara, M. Saito, and others define a variation over S to be admissible if its restriction to every curve is admissible in the sense of Steenbrink-Zucker.

However, the behavior of *unipotent* variations of MHS at infinity can be controlled rather easily. We have the classical result of Deligne (see [151, Prop. 5.2]) which defines the *canonical extension* $\tilde{\mathcal{V}}$ of $\mathcal{V}$.

Theorem 2.7.5. *Let $\tilde{S}$ be a normalization of S. Let $(\mathbb{V}_{\mathbb{Q}}, W_\bullet, \mathcal{F}^\bullet)$ be a unipotent variation of MHS over S. Then*
(a) There is a unique extension $\tilde{\mathcal{V}}$ of $\mathcal{V}$ over $\tilde{S}$ satisfying the following equivalent conditions:

> *(i) Inside every section of $\tilde{\mathcal{V}}$, every flat section of $\mathcal{V}$ increases at most at the rate of $O(\log^k \|x\|)$ (k large enough) on every compact set of $D = \tilde{S} - S$.*
>
> *(ii) Similarly, every flat section of $\mathcal{V}^\vee$ (the dual) increases at most at the rate of $O(\log^k \|x\|)$ (k large enough).*

(b) The combination of the two conditions (i) and (ii) is equivalent to the combination of the following two conditions:

> *(iii) In terms of any local basis of $\tilde{\mathcal{V}}$ the connection matrix $\boldsymbol{\omega}$ of $\mathcal{V}$ has at most logarithmic singularities along D.*
>
> *(iv) The residue of $\boldsymbol{\omega}$ along any irreducible component of D is nilpotent.*

We will verify conditions (iii) and (iv) by Prop. 2.5.8 for the multiple logarithm variations of MHS. They are unipotent variations by Thm. 2.5.9.

Definition 2.7.6. Let $\tilde{S}$ be a compactification of S. Then a unipotent variation of MHS $(\mathbb{V}_{\mathbb{Q}}, W_\bullet, \mathcal{F}^\bullet)$ over S is said to be *good* if it satisfies the following conditions at infinity

> (1) the Hodge filtration bundles $\mathcal{F}^\bullet$ extend over $\tilde{S}$ to sub-bundles $\tilde{\mathcal{F}}^\bullet$ of the canonical extension $\tilde{\mathcal{V}}$ of $\mathcal{V}$ such that they induce the corresponding thing for each pure subquotient $\mathrm{gr}_k^W\, \mathbb{V}_{\mathbb{Q}}$,

(2) for the nilpotent logarithm N_j of a local monodromy transformation about a component D_j of D, the weight filtration of N_j relative to $W_\bullet$ exists.

A slightly different definition first appeared in [264, 265] with the extra assumption that $D = \tilde{S} - S$ is a normal crossing divisor. In these papers Hain and Zucker classified good unipotent variations of MHS on algebraic manifolds. With constant pure weight subquotients these variations behave well at infinity.

2.7.2 *Mixed Hodge structures of multiple logarithms*

Fix an embedding $\mathbb{C}^n \hookrightarrow \mathbb{CP}^n$. Let $\mathcal{D}_n = D_n \cup (\mathbb{CP}^n \setminus \mathbb{C}^n)$. Let $M_{2^n}(\mathbb{C})$ be the set of $2^n \times 2^n$ matrices over $\mathbb{C}$. Put

$$\boldsymbol{\omega} = \Big(c_{\mathbf{i},\mathbf{j}}\Big)_{\mathbf{i},\mathbf{j}\in\mathcal{I}_n} \in H^0(\mathbb{CP}^n, \Omega^1_{\mathbb{CP}^n}(\log(\mathcal{D}_n))) \otimes M_{2^n}(\mathbb{C}) \tag{2.37}$$

where

$$c_{\mathbf{i},\mathbf{j}} = \begin{cases} d\gamma^{|\mathbf{i}|}_{\rho_{\mathbf{i}}(\mathbf{j})}\big(\mathbf{x}(\mathbf{i})\big) & \text{if } |\mathbf{j}| = |\mathbf{i}| - 1,\ \mathbf{j} \prec \mathbf{i}, \\ 0 & \text{otherwise.} \end{cases}$$

It is not difficult to check that all of the 1-forms in $\boldsymbol{\omega}$ have logarithmic singularity on $\mathcal{D}_n$. Moreover, because $\mathcal{M}_{[n]}(\mathbf{x})$ is invertible and $\boldsymbol{\omega}$ is closed we get

$$d\boldsymbol{\omega} = 0, \quad \boldsymbol{\omega}\wedge\boldsymbol{\omega} = 0. \tag{2.38}$$

This shows that $\boldsymbol{\omega}$ is integrable.

Now we define a meromorphic connection ∇ on the trivial bundle

$$\mathbb{CP}^n \times \mathbb{C}^{2^n} \longrightarrow \mathbb{CP}^n \tag{2.39}$$

by

$$\nabla f = df - \boldsymbol{\omega} f$$

where $f : S_n \to \mathbb{C}^{2^n}$ is a section. This connection has regular singularities along $\mathcal{D}_n$ because $\boldsymbol{\omega}$ is integrable by Eq. (2.38) and all the 1-forms in $\boldsymbol{\omega}$ are logarithmic in any compactification of S_n. By the explicit construction of $\boldsymbol{\omega}$ we see immediately that the conditions (iii) and (iv) of Thm. 2.7.5 are satisfied. Prop. 2.5.8 further implies that the columns $(2\pi i)^{|\mathbf{j}|}C_{\mathbf{j}}(\mathbf{x})$ of $\mathcal{M}_{[n]}(\mathbf{x})$ satisfy $\nabla f = 0$ and are therefore flat sections of Eq. (2.39). Even though they are multi-valued, their $\mathbb{Z}$-linear span is well defined thanks to Thm. 2.5.9. Hence $V_{[n]}(\mathbf{x})$ forms a local system over S_n.

Definition 2.7.7. The local system $V_{[n]}(\mathbf{x})$ is called the *n-tuple logarithm local system.*

To define the MHS on $V_{[n]}$ we can define the weight filtration by putting $W_{2k+1} = W_{2k}$ and

$$W_{-2k}V_{[n]}(\mathbf{x}) = \langle (2\pi i)^{|\mathbf{i}|} C_{\mathbf{i}} : \ |\mathbf{i}| \geq k \rangle_{\mathbb{Q}}$$

which is the $\mathbb{Q}$ vector space with basis $\{(2\pi i)^{|\mathbf{i}|} C_{\mathbf{i}} : \ |\mathbf{i}| \geq k\}$. In particular, $W_{-2k}V_{[n]}(\mathbf{x}) = 0$ if $k > n$ and $W_{2k}V_{[n]}(\mathbf{x}) = V_{[n]}(\mathbf{x})$ if $k \geq 0$. By regarding $e_{\mathbf{i}}$'s as column vectors one can define the Hodge filtration on $V_{[n]}(\mathbf{x}) \otimes \mathbb{C} = V_{[n],\mathbb{C}}$ as follows:

$$\mathcal{F}^{-k}V_{[n],\mathbb{C}} := \langle e_{\mathbf{i}} : \ |\mathbf{i}| \leq k \rangle_{\mathbb{C}}.$$

So in particular, $\mathcal{F}^{-k}V_{[n],\mathbb{C}} = 0$ for $k < 0$ and $\mathcal{F}^{-k}V_{[n],\mathbb{C}} = V_{[n],\mathbb{C}}$ for $k \geq n$.

By induction on n and using Lemma 2.5.6 it is easy to show that

$$\mathcal{F}^{-p} \cap W_{-2k}V_{[n],\mathbb{C}} = \begin{cases} 0 & \text{if } p \leq k-1 \\ \langle (2\pi i)^{|\mathbf{i}|} e_{\mathbf{i}} : \ k \leq |\mathbf{i}| \leq p \rangle & \text{if } k \leq p \leq n \\ \langle (2\pi i)^{|\mathbf{i}|} e_{\mathbf{i}} : \ k \leq |\mathbf{i}| \leq n \rangle & \text{if } p \geq n \end{cases}.$$

This implies that

$$\mathcal{F}^{-p} \operatorname{gr}^W_{-2k} V_{[n],\mathbb{C}} = \begin{cases} 0 & \text{if } p \leq k-1 \\ W_{-2k}V_{[n],\mathbb{C}}/W_{-2k-1}V_{[n],\mathbb{C}} & \text{if } p \geq k. \end{cases}$$

In other words, $\mathcal{F}^q \operatorname{gr}^W_{-2k} V_{[n],\mathbb{C}} = 0$ for $q \geq -k+1$ and $\mathcal{F}^q \operatorname{gr}^W_{-2k} V_{[n],\mathbb{C}} = \operatorname{gr}^W_{-2k} V_{[n],\mathbb{C}}$ for $q \leq -k$. This means that the Hodge filtration induces a pure HS of weight $-2k$ on each weight graded piece. Furthermore, it is not hard to see by checking the powers of $2\pi i$ appearing on the diagonal of $\mathcal{M}_{[n]}(\mathbf{x})$ that this induced structure on $\operatorname{gr}^W_{-2k} V_{[n],\mathbb{C}}$ is isomorphic to the direct sum of $\binom{n}{k}$ copies of the Tate structure $\mathbb{Z}(k)$ by Lemma 2.5.6.

2.8* Limit mixed Hodge structures of multiple logarithms

Let the monodromy of $\mathcal{M}_{[n]}(\mathbf{x})$ at any subvariety $\mathcal{D}$ of $\mathbb{CP}^n$ be given by the matrix $T_{\mathcal{D}}$ and the local monodromy logarithm by $N_{\mathcal{D}} = \log T_{\mathcal{D}}/(2\pi i)$. Note that $T_{\mathcal{D}}$ is unipotent so $N_{\mathcal{D}}$ is well-defined.

Now let us recall the construction of the unipotent variations of limit MHS at the "infinity" with *normal crossing*. Let S be a complex manifold of dimension d. Suppose that S is embedded in $\tilde{S}$, via the mapping j, such that $D = \tilde{S} - S$ is a divisor with normal crossings. Let $\mathbb{V}$ be any unipotent local system of complex vector spaces on S, and $\mathcal{V}$ the corresponding vector

bundle. According to Thm. 2.7.5 by Deligne there is a canonical extension $\tilde{\mathcal{V}}$ of $\mathcal{V}$ over $\tilde{S}$. Moreover, when the local monodromy is nilpotent $\tilde{\mathcal{V}}$ is a subsheaf of $j_*\mathcal{V}$. The local picture of $S \subset \tilde{S}$ is $(\Delta^*)^r \times \Delta^{d-r} \subset \Delta^d$ where Δ is the unit disk and Δ^* is the punctured one. We let $t_1, \ldots, t_r$ denote the variables on $(\Delta^*)^r$, and $N_1, \ldots, N_r$ the (commuting) local nilpotent logarithms of the associated monodromy transformations of the fibre. For $z_1, \ldots, z_r$ in the upper half-plane, the universal covering mapping for $(\Delta^*)^r$ is given by

$$t_j = \exp(2\pi i z_j), \quad j = 1, \cdots, r.$$

Let $v_1, \ldots, v_m$ be a basis of the multi-valued sections of $\mathbb{V}$ over $(\Delta^*)^r \times \Delta^{d-r}$, the formula

$$[\tilde{v}_1, \ldots, \tilde{v}_m] = [v_1, \ldots, v_m] \exp\left(-\sum_{j=1}^{r} 2\pi i z_j N_j\right) = [v_1, \ldots, v_m] \prod_{j=1}^{r} t_j^{-N_j}$$

determines a basis of the sections of $\mathcal{V}$ over Δ^d and these provide, by the definition, the generators of $\tilde{\mathcal{V}}$ over Δ^d.

In our situation, although the divisor D_n is not normal crossing Thm. 2.7.5 is still valid because the image of the global holomorphic logarithmic forms in the complex of smooth forms on S is independent of the normal crossings compactification (see [263, Prop. (3.2)]). In fact, the forms we are considering lie in the subcomplex generated by 1-forms of the type df/f where f is a rational function. Such forms are automatically logarithmic in any compactification and therefore our connection is automatically regular. Hence the admissibility and the existence of the limit MHS is an automatic consequence of the admissibility of our variations restricted to every curve in S_n. Moreover, the pullback to $\tilde{S}_n$ of our trivial bundle Eq. (2.39) restricted to S_n is exactly Deligne's canonical extension of Eq. (2.39), and the pullbacks of the subbundles $\mathcal{F}^\bullet$ and $W_\bullet$ are the correct extended Hodge and weight subbundles. Therefore we have

Theorem 2.8.1. *The n-tuple logarithm underlies a good unipotent graded-polarizable variation of mixed Hodge-Tate structures $(V_{[n]}, W_\bullet, \mathcal{F}^\bullet)$ over*

$$S_n = \mathbb{C}^n \setminus \Big\{ \prod_{1 \le j \le n} x_j(1-x_j) \prod_{1 \le i < j \le n} \big(1 - x_i \ldots x_j\big) = 0 \Big\}.$$

with the weight-graded quotients gr^W_{-2k} being given by $\binom{n}{k}$ copies of the Tate structure $\mathbb{Z}(k)$.

Proof. It is clear that all the odd graded weight quotients are zero so that we can let the polarizations on the weight graded quotients gr^W_{-2k} be the ones that give each vector $2\pi i e_{\mathbf{j}}$ ($|\mathbf{j}| = k$) length 1. Then everything is clear except the Griffiths transversality condition. But this condition is also satisfied because $dC_{\mathbf{j}} = \boldsymbol{\omega} C_{\mathbf{j}}$ for every $\mathbf{j} \in \mathcal{I}_n$ by Prop. 2.5.8 since all the entries of $\boldsymbol{\omega}$ on and above the main diagonal are zero. □

If we want to determine the limit MHS of multiple logarithms explicitly we can still apply the techniques used in the normal crossing case. We will carry this out only for the depth two and three cases. The general picture is similar but much more complicated.

2.8.1 *Limit mixed Hodge structures of double polylogs*

First we look at the double logarithm variation of MHS. We have

$$\mathcal{M}_{1,1}(x,y) = \begin{bmatrix} 1 & & & \\ \mathfrak{L}_1(x) & 2\pi i & & \\ \mathfrak{L}_1(xy) & 0 & 2\pi i & \\ \mathfrak{L}_2(x,y) & 2\pi i\mathfrak{L}_1(y) & 2\pi i h(x,y) & (2\pi i)^2 \end{bmatrix}$$

where $h(x,y) = \mathfrak{L}_1(x) - \mathfrak{L}_1(y) - \log y$.
(i) Let us first try to extend the MHS to the divisor $\mathcal{D}_{20} = \{y = 0\}$ along the tangent vector $\partial/\partial y$. We have

$$T_{\{y=0\}} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad N_{\{x=0\}} = \frac{\log T_{\{x=0\}}}{2\pi i} = \begin{bmatrix} 0 & & & \\ 0 & 0 & & \\ 0 & 0 & 0 & \\ 0 & 0 & -\frac{1}{2\pi i} & 0 \end{bmatrix}.$$

Let $\mathcal{M}_{1,1}(x,y) = [C_0(x,y) \ \cdots C_3(x,y)]$. Define

$$[s_0 \ s_1 \ s_2 \ s_3] = \lim_{t\to 0} \mathcal{M}_{1,1}(x,t) \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & \log t/(2\pi i) & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & & \\ \mathfrak{L}_1(y) & 2\pi i & & \\ 0 & 0 & 2\pi i & \\ 0 & 0 & 2\pi i\mathfrak{L}_1(x) & (2\pi i)^2 \end{bmatrix}.$$

Let $V_{\mathbb{Q},\{y=0\}}$ be the $\mathbb{Q}$-linear span of s_0, s_1, s_2, s_3, and $V_{\mathbb{C},\{y=0\}} = \mathbb{C} \otimes V_{\mathbb{Q},\{y=0\}}$. Let $\{e_j : j = 0, \cdots, 3\}$ be the standard basis of $\mathbb{C}^4$ where the

only nonzero entry of e_j is at the $(j+1)$st component. Then the limit MHS on $\{(x,y) : y = 0, x \neq 1\}$ along $\partial/\partial y$ are given by

$$((V_{\mathbb{Q},\{y=0\}}, W_\bullet), (V_{\mathbb{C},\{y=0\}}, F^\bullet))$$

where for $k = 0, \ldots, 3$

$$W_{-2k}V_{\mathbb{Q},\{x=0\}} = \langle s_k, \ldots, s_3\rangle, W_{-2k} = W_{-2k+1} \tag{2.40}$$

and

$$F^{-k}V_{\mathbb{C},\{x=0\}} = \langle e_0, \ldots, e_k\rangle. \tag{2.41}$$

(ii) A similar calculation shows that along the tangent vector $\partial/\partial y$ the limit MHS on the divisor $\mathcal{D}_{22} = \{(x,1) : x \neq 1\}$ is the $\mathbb{Q}$-linear span of $s_0, \ldots, s_3$ where

$$[s_0\ s_1\ s_2\ s_3] = \begin{bmatrix} 1 & & & \\ \mathfrak{L}_1(x) & 2\pi i & & \\ \mathfrak{L}_1(x) & 0 & 2\pi i & \\ \mathfrak{L}_2(x,1) & 0 & 2\pi i\mathfrak{L}_1(x) & (2\pi i)^2 \end{bmatrix}.$$

It is easy to see by differentiation that $\mathfrak{L}_2(x,1) = \big(\mathfrak{L}_1(x)\big)^2/2$.
(iii) The extension of MHS to $\mathcal{D}_{11} = \{(1,y) : y \neq 0, 1\}$ along the tangent vector $\partial/\partial x$ is given by the $\mathbb{Q}$-linear span of $s_0, \ldots, s_3$ where

$$[s_0\ s_1\ s_2\ s_3] = \begin{bmatrix} 1 & & & \\ 0 & 2\pi i & & \\ -\operatorname{Li}_1(\frac{y}{y-1}) & 0 & 2\pi i & \\ \operatorname{Li}_2(\frac{y}{y-1}) & 2\pi i\operatorname{Li}_1(y) & -2\pi i\log\frac{y}{y-1} & (2\pi i)^2 \end{bmatrix}.$$

(iv) Limit MHS on $\mathcal{D}_{12} = \{(x,1/x) : x \neq 0, 1\}$ along the tangent vector $\partial/\partial y$ is given by the $\mathbb{Q}$-linear span of $s_0, \ldots, s_3$ where

$$[s_0\ s_1\ s_2\ s_3] = \begin{bmatrix} 1 & & & \\ -\operatorname{Li}_1(\frac{x}{x-1}) & 2\pi i & & \\ 0 & 0 & 2\pi i & \\ -\operatorname{Li}_2(\frac{x}{x-1}) & 2\pi i\log\frac{x}{x-1} & 0 & (2\pi i)^2 \end{bmatrix}.$$

(v) $\mathcal{D}_{20} \cap \mathcal{D}_{11} = (1,0)$. From (i) we see that there are limit MHS on the open set $\mathcal{D}_{20} \setminus \{(1,0)\}$ of $\mathcal{D}_{20}$. We now can easily extend these MHS to $(1,0)$ along the vector $\partial/\partial x$ and find the limit MHS at $(1,0)$ to be the $\mathbb{Q}$-linear span of $s_0, \cdots, s_3$ where

$$[s_0\ s_1\ s_2\ s_3] = \begin{bmatrix} 1 & & & \\ 0 & 2\pi i & & \\ 0 & 0 & 2\pi i & \\ 0 & 0 & 0 & (2\pi i)^2 \end{bmatrix}.$$

If we start from (iii) and then extend the MHS to $(1,0)$ along the tangent vector $\partial/\partial x$ we will get the same limit MHS.
(vi) $\mathcal{D}_{22}\cap\mathcal{D}_{12}=\mathcal{D}_{12}\cap\mathcal{D}_{11}=\mathcal{D}_{11}\cap\mathcal{D}_{22}=(1,1)$. We can start from either case (ii) or (iii) or (iv). Extending the limit MHS of case (ii) we see immediately that the along the tangent vector $\partial/\partial y$ the limit MHS at $(1,1)$ is given by the $\mathbb{Q}$-linear span of

$$[s_0\ s_1\ s_2\ s_3]=\begin{bmatrix}1&&&\\0&2\pi i&&\\0&0&2\pi i&\\E_{4,1}&0&0&(2\pi i)^2\end{bmatrix}. \tag{2.42}$$

If we extend the limit MHS of case (iii) to $(1,1)$ along the tangent vector $\partial/\partial x$ we find that only the lower left corner entry is different from the above. Instead of 0 it is

$$E_{4,1}=\lim_{y\to1}\left(\mathrm{Li}_2(\frac{y}{y-1})+\frac{1}{2}\log^2(1-y)-\log y\log(1-y)\right)=-\mathrm{Li}_2(1)=-\frac{\pi^2}{6},$$

since for all $t\neq0$

$$\mathrm{Li}_2(1-t)+\mathrm{Li}_2(1-1/t)+\frac{1}{2}\log^2 t=0. \tag{2.43}$$

But if we take $s_0'=s_0-s_3/48$ we get the same basis as in Eq. (2.42). The same phenomenon occurs if we start from case (iv) and then use tangent vector $\partial/\partial y$.

If we extend the limit MHS of (iv) to the point $(1,1)$ along the tangent vector $\partial/\partial x$ then we find that

$$E_{4,1}=\lim_{x\to1}-\mathrm{Li}_2(\frac{x}{x-1})-\frac{1}{2}\log^2(1-x)=\mathrm{Li}_2(1)=\frac{\pi^2}{6}$$

by taking $t=1-x$ in Eq. (2.43). Now if we let $s_0'=s_0+\frac{1}{48}s_3$ then we get the same basis as in Eq. (2.42). This phenomenon happens in higher logarithm cases too.

In Exercise 2.11 you will see another example of limit MHS. From all these examples we put forward the following

Conjecture 2.8.2. *The variations of mixed Hodge-Tate structures associated to every MPL can be produced as the variations of some limit mixed Hodge-Tate structures related to some suitable choice of multiple logarithm.*

2.9 Single-valued MPLs of multi-variables

Using the monodromy of MPLs obtained from Sec. 2.5 we can construct single-valued MPLs. Although this procedure can be carried out in general the computation in each particular case will involve very complicated notations. So we will just treat several lower weight cases in detail to illustrate the main idea.

2.9.1 *General procedure of construction*

In this subsection we first describe the general procedure to define the single-valued real analytic version of $\mathrm{Li}_{m_1,\dots,m_n}(\mathbf{x})$, denoted by $\mathscr{L}i_{s_1,\dots,s_n}(\mathbf{x})$, provided the variation matrix $\mathcal{M}_{s_1,\dots,s_n}(\mathbf{x})$ is already found. We will work with the multiple logarithms below, but the idea is exactly the same for the general case.

Let $\mathcal{M}_{[n]} = \mathcal{M}_{[n]}(\mathbf{x}) = L_n(\mathbf{x})\tau_n(2\pi i)$ where

$$\tau_n(\lambda) = \mathrm{diag}\big[\lambda^{|\mathbf{j}|}\big]_{\mathbf{j}\in\mathcal{I}_n}.$$

Define the matrix

$$B_{[n]} = \tau_n(i)\mathcal{M}_{[n]}\overline{\mathcal{M}}_n^{-1}\tau_n(i)$$

where $\overline{\mathcal{M}}_n$ is the complex conjugation of $\mathcal{M}_{[n]}$. By the monodromy computation we see that $B_{[n]}$ is a single-valued matrix function defined over S_n because all the monodromy automorphisms are given by real matrices. Now both $\mathcal{M}_{[n]}$ and $\overline{\mathcal{M}}_{[n]}^{-1}$ are lower triangular matrices, so is $B_{[n]}$. Moreover its diagonal is given by $\tau_n(i)\tau_n(2\pi i)\overline{\tau_n(2\pi i)}^{-1}\tau_n(i) = I$ where I is the identity matrix. Hence we may assume $B_{[n]} = I + N$ where N a nilpotent matrix. We see that $\log B$ is well defined and since $\overline{B}_n = B_{[n]}^{-1}$ it satisfies

$$\overline{\log B_{[n]}} = -\log B_{[n]},$$

namely, $\log B_{[n]}$ is a pure imaginary matrix. Then we define $i^{2[n/2]-1}/2$ times the lower left corner entry of $\log B_{[n]}$ to be $\mathfrak{L}_n(\mathbf{x})$ which is a single-valued real analytic version of the multiple logarithm $\mathfrak{L}_n(\mathbf{x})$.

Remark 2.9.1. Our method is slightly different from that of [41]. In fact, in the polylog case the matrix $B_{[n]}$ constructed as above is the conjugate of the one in [41] by $\tau(i)$.

2.9.2 *Single-valued double logarithms*

We have seen from Example 2.5.5 that $\mathcal{M}_{1,1}(x,y) = L_{1,1}(x,y)\tau_{1,1}(2\pi i)$ where

$$L_{1,1}(x,y) = \begin{bmatrix} 1 & & & \\ \mathrm{Li}_1(x) & 1 & & \\ \mathrm{Li}_1(xy) & & 1 & \\ \mathrm{Li}_{1,1}(x,y) & \mathrm{Li}_1(y) & \log\frac{y-1}{y(1-x)} & 1 \end{bmatrix} \text{ and } \tau_{1,1}(\lambda) = \begin{bmatrix} 1 & & & \\ & \lambda & & \\ & & \lambda & \\ & & & \lambda^2 \end{bmatrix}.$$

Let $B_{1,1}(x,y) = \tau_{1,1}(i)L_{1,1}(x,y)\tau_{1,1}(-1)\overline{L_{1,1}(x,y)}^{-1}\tau_{1,1}(i)$. Then $B_{1,1}(x,y)$ is unipotent and single-valued. An easy calculation shows

$$\log B_{1,1}(x,y) = \begin{bmatrix} 0 & & & \\ -2i\log|1-x| & 0 & & \\ -2i\log|1-xy| & & 0 & \\ -2i\,\mathcal{L}i_{1,1}(x,y) & -2i\log|1-y| & 2i\log\big|\frac{y-1}{y(1-x)}\big| & 0 \end{bmatrix}$$

where

$$\mathcal{L}i_{1,1}(x,y) = \mathrm{Im}\big(\mathrm{Li}_{1,1}(x,y)\big) - \arg(1-x)\log|1-y| + \arg(1-xy)\log\Big|\frac{y-1}{y(1-x)}\Big| \tag{2.44}$$

is the single-valued real analytic version of $\mathrm{Li}_{1,1}(x,y)$.

By differentiation it is easy to check that (see Exercise 2.13)

$$\mathrm{Li}_{1,1}(x,y) = \mathrm{Li}_2\Big(\frac{xy-x}{1-x}\Big) - \mathrm{Li}_2\Big(\frac{x}{x-1}\Big) - \mathrm{Li}_2(xy). \tag{2.45}$$

So by using the single-valued Bloch-Wigner dilog function

$$\mathcal{L}i_2(z) = \mathrm{Im}\big(\mathrm{Li}_2(z)\big) + \arg(1-z)\log|z| \tag{2.46}$$

we can also recover Eq. (2.44) as

$$\mathcal{L}i_{1,1}(x,y) = \mathcal{L}i_2\Big(\frac{xy-x}{1-x}\Big) - \mathcal{L}i_2\Big(\frac{x}{x-1}\Big) - \mathcal{L}i_2(xy). \tag{2.47}$$

Further, using the five term relation we find

$$\mathcal{L}i_{1,1}(x,y) + \mathcal{L}i_{1,1}(y,x) + \mathcal{L}i_2(xy) = 0.$$

One can check quickly $\mathcal{L}i_{1,1}$ satisfies the functional equation

$$\mathcal{L}i_{1,1}(x,y) = -\mathcal{L}i_{1,1}\Big(\frac{x}{x-1}, 1-y\Big)$$

by using the functional equation $\mathcal{L}i_2(x) = -\mathcal{L}i_2(1-x) = -\mathcal{L}i_2(1/x)$.

2.9.3 *Single-valued double polylogs $\mathscr{L}i_{1,2}$ and $\mathscr{L}i_{2,1}$*

By [573] a single-valued version of $\mathrm{Li}_3(x)$ can be defined as

$$\mathscr{L}i_3(z) = \mathrm{Re}\big(\mathrm{Li}_3(z)\big) - \log|z|\,\mathrm{Re}\big(\mathrm{Li}_2(z)\big) - \frac{1}{3}(\log|z|)^2\log|1-z|. \quad (2.48)$$

We now consider $\mathrm{Li}_{2,1}(x,y)$ and $\mathrm{Li}_{1,2}(x,y)$. By the procedure outlined in Sec. 2.9.1 we find that the single-valued version of $\mathrm{Li}_{2,1}(x,y)$ is (see Exercise 2.9):

$$\begin{aligned}\mathscr{L}i_{2,1}(x,y) &= \mathrm{Re}\,\mathrm{Li}_{2,1}(x,y) - \arg(1-xy)\big[\mathscr{L}i_2(x)+\mathscr{L}i_2(y)\big] + \log|1-y|\,\mathrm{Re}\,\mathrm{Li}_2(x)\\ &\quad -\log|x|\,\mathrm{Re}\,\mathrm{Li}_{1,1}(x,y) - \log\Big|\frac{1-y}{y}\Big|\,\mathrm{Re}\,\mathrm{Li}_2(xy) - \frac{1}{3}\log|x^2y|\log|1-xy|\log\Big|\frac{1-y}{y}\Big|\\ &\quad + \frac{1}{3}\log|x|\big(2\log|1-x|\log|1-y| + \log|1-xy|\log|y(1-x)|\big).\end{aligned}$$

The single-valued version of $\mathrm{Li}_{1,2}(x,y)$ is (see Exercise 2.10):

$$\begin{aligned}\mathscr{L}i_{1,2}(x,y) &= \mathrm{Re}\,\mathrm{Li}_{1,2}(x,y) + \arg(1-xy)\big[\mathscr{L}i_2(x)+\mathscr{L}i_2(y)\big] - \arg(1-x)\,\mathscr{L}i_2(y)\\ &\quad +\log|1-x|\,\mathrm{Re}\,\mathrm{Li}_2(xy) - \log|y|\,\mathrm{Re}\,\mathrm{Li}_{1,1}(x,y) + \frac{1}{3}\log|1-x|\log|xy|\log|1-xy|\\ &\quad + \frac{1}{3}\log|y|\Big[\log|1-x|\log|1-y| + \log|1-xy|\log\Big|\frac{y(1-x)}{1-y}\Big|\Big].\end{aligned}$$

Using the single-valued versions of dilog $\mathscr{L}i_2(z)$ and trilog $\mathscr{L}i_3(z)$ we can express $\mathscr{L}i_{1,2}(x,y)$ by the trilogs

$$\begin{aligned}\mathscr{L}i_{1,2}(x,y) &= \mathscr{L}i_3\Big(\frac{y-xy}{1-xy}\Big) - \mathscr{L}i_3\Big(\frac{1-x}{1-xy}\Big)\\ &\quad + \mathscr{L}i_3(1-xy) - \mathscr{L}i_3(y) + \mathscr{L}i_3(1-x) - \mathscr{L}i_3(1),\end{aligned}$$

where $\mathscr{L}i_3$ is the single-valued trilog given by Eq. (2.48). This follows from the relation first discovered by Zagier after Goncharov's conviction that such identity should exist:

$$\begin{aligned}\mathrm{Li}_{1,2}(x,y) &= \mathrm{Li}_3\Big(\frac{y-xy}{1-xy}\Big) - \mathrm{Li}_3\Big(\frac{1-x}{1-xy}\Big) - \log\Big(\frac{1-x}{1-xy}\Big)\,\mathrm{Li}_2(y)\\ &\quad + \mathrm{Li}_3(1-xy) - \mathrm{Li}_3(y) + \mathrm{Li}_3(1-x) - \mathrm{Li}_3(1)\\ &\quad - \log(1-xy)\big(\mathrm{Li}_2(1) + \mathrm{Li}_2(1-x)\big) + \frac{1}{2}\log(y)\log^2(1-xy). \quad (2.49)\end{aligned}$$

The geometric origin of this identity can be found in [587, Remark 3.18]. By straightforward computation one can find the following interesting formula:

$$\mathscr{L}i_{1,2}(x,y) + \mathscr{L}i_{2,1}(y,x) + \mathscr{L}i_3(xy) = 2\log|y|\,\mathrm{Re}(\mathrm{Li}_1(y)\,\mathrm{Li}_1(x)).$$

One should compare this with

$$\mathrm{Li}_{1,2}(x,y)+\mathrm{Li}_{2,1}(y,x)+\mathrm{Li}_3(xy)=\mathrm{Li}_2(y)\,\mathrm{Li}_1(x).$$

Finally, one can express the triple logarithm $\mathcal{L}i_{1,1,1}(x,y,z)$ using the trilog (see Exercise 2.14). Thus in weight three we essentially only need trilog. But it may be a little surprising that similar identities in higher weight cases do not exist in general from geometric considerations (see [587, p. 146]). For example, $\mathcal{L}i_{2,2}(x,y)$ cannot be expressed by only tetralogarithms $\mathcal{L}i_4$.

2.10 Aomoto polylogs

For many people it is much easier to grasp a concept when geometric figures are attached to it. This is certainly true when we study MPLs. Many deep results have been obtained for them up to weight three especially because in these lower weights we can draw pictures and visualize/discover various relations between them. Moving up to higher dimensions it is natural to consider the so-called Aomoto polylogs.

Definition 2.10.1. Let F be an arbitrary field. Fix a positive integer n. By a *simplex* in the projective spaces $\mathbb{P}^n_F$ we mean an ordered set of hyperplanes $L=(L_0,\dots,L_n)$. It is *nondegenerate* if the intersection of all the hyperplanes L_i is empty. A *face* of L is any nonempty intersection of the hyperplanes. A pair of simplices is *admissible* if they do not have common faces of the same dimension. It is a *generic pair* if all the faces of the two simplices are in general position.

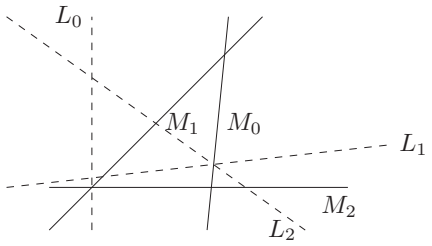


Figure 2.1: An admissible non-generic pair of triangles.

By projective invariance to be defined below, given a nondegenerate simplex L we may choose the coordinate system $[t_0,\dots,t_n]$ in $\mathbb{P}^n_F$ such that

$L_i = \{t_i = 0\}$ for $0 \leq i \leq n$. If $F = \mathbb{C}$ then we define the canonical differential form associated to L as

$$\omega_L = \frac{\mathrm{d}t_1}{t_1} \wedge \frac{\mathrm{d}t_2}{t_2} \wedge \ldots \wedge \frac{\mathrm{d}t_n}{t_n}.$$

The Aomoto n-logarithm is a multi-valued function on configurations of nondegenerate admissible pairs of simplices $(L;\, M)$ in $\mathbb{CP}^n$ defined as follows:

$$\Lambda_n(L;\, M) = \int_{\Delta_M} \omega_L,$$

where Δ_M is the n-cycle representing a generator of the relative homology $H_n(\mathbb{CP}^n, \cup_{j=0}^n M_j; \mathbb{Z})$.

We now define the double scissors congruence groups $A_n(F)$ whose defining relations should reflect (conjecturally all of) the functional equations of Aomoto polylogs. Furthermore, the associated graded object should form a Hopf algebra which provides a bridge to the study of motivic polylogic complexes.

Definition 2.10.2. Define $A_0(F) = \mathbb{Z}$. If $n > 0$ then $A_n(F)$ is the abelian group generated by admissible pairs of n-simplices $(L;\, M)$ subject to the following relations:

(R1) *Nondegeneracy.* $(L;\, M) = 0$ if and only if L or M is degenerate.

(R2) *Trivial intersection.* Suppose $L_0, \ldots, L_n$ and $M_0, \ldots, M_n$ are hyperplanes in $\mathbb{P}_F^{n+1}$. If $N = L_i$ or $N = M_i$ for some i then

$$(N|L;\, M) := \big((L_0 \cap N, \ldots, L_n \cap N);\, (M_0 \cap N \ldots, M_n \cap N)\big) = 0.$$

(R3) *Skew symmetry.* For every permutation σ of $\{0, \ldots, n\}$

$$(\sigma L;\, M) = (L; \sigma M) = \mathrm{sgn}(\sigma)(L;\, M),$$

where $\sigma(L_0, \ldots, L_n) = (L_{\sigma(0)}, \ldots, L_{\sigma(n)})$.

(R4) *Additivity in L and M.* For any hyperplanes $L_0, \ldots, L_{n+1}$ and n-simplex M in $\mathbb{P}_F^n$

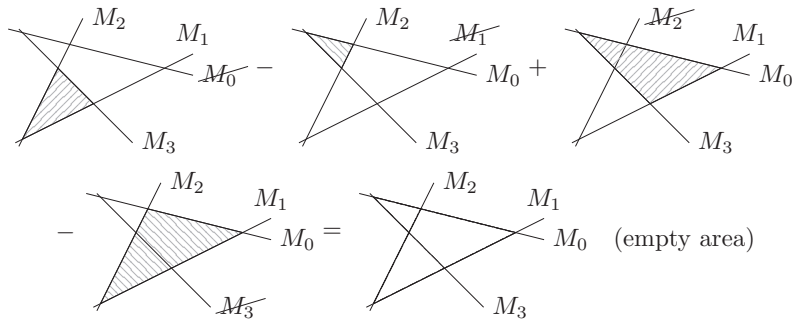
$$\sum_{j=0}^{n+1} (-1)^j ((L_0, \ldots, \widehat{L_j}, \ldots, L_{n+1});\, M) = 0$$

if every pair $((L_0, \ldots, \widehat{L_j}, \ldots, L_{n+1});\, M)$ is admissible. A similar relation is satisfied by any n-simplex L and hyperplanes $M_0, \ldots, M_{n+1}$.

(R5) *Projective invariance.* For every $g \in \mathrm{PGL}_{n+1}(F)$

$$(gL;\, gM) = (L;\, M).$$

Remark 2.10.3. The additivity relation (R4) can be visualized in dimension two by the following picture:



Example 2.10.4. We now consider the simplest case: $n = 1$. By projective invariance (R5) we may assume without loss of generality that $L = (\infty, 0)$ and $M = (1, a)$ for some $a \neq 0, 1, \infty$. Then de Rham cohomology $H^1_{\mathrm{dR}}(\mathbb{CP}^1 \setminus L) = H^1(\mathbb{C}^*)$ is generated by $[dz/z]$ where the relative singular homology $H_1(\mathbb{CP}^1, M)$ is generated by $[\Delta_M]$ where Δ_M is the line segment from 1 to a. So we get the period

$$\int_1^a \frac{\mathrm{d}z}{z} = \log a.$$

Example 2.10.5. Let $n = 2$. Take $L_1 = \{x = 0\}$, $L_2 = \{y = 0\}$ and $L_0 = \{z = 0\}$. Let $M_0 = \{x = z\}$, $M_1 = \{y = tz\}$, and $M_2 = \{x + y = z\}$. We have the following picture:

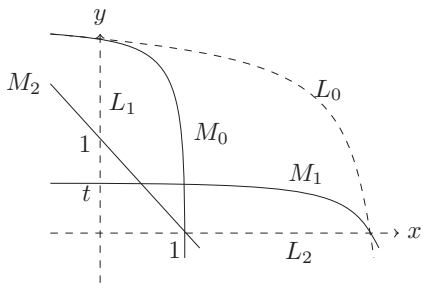


Figure 2.2: A real picture for dilog.

Hence the period is given by

$$\int_{0<1-x<y<t} \frac{\mathrm{d}x}{x}\frac{\mathrm{d}y}{y} = -\int_{0<x<y<t} \frac{\mathrm{d}y}{y}\frac{\mathrm{d}x}{1-x} = -\operatorname{Li}_2(t).$$

Definition 2.10.6. The n-simplex L whose faces are $L_i = \{t_i = 0\}$ $(0 \leq i \leq n)$ is called the *standard (coordinate) simplex*. For an arbitrary field F we denote by $\Lambda_{n_1,\ldots,n_d}$ the subgroup of A_n $(n = n_1 + \cdots + n_d)$ generated by all *MPLic pairs* $\Lambda_{n_1,\ldots,n_d}(x_1,\ldots,x_d)$ corresponding to the MPLs (all possible $x_1,\ldots,x_d$). Specifically, we can represent $\Lambda_{n_1,\ldots,n_d}(x_1,\ldots,x_d)$ by $(L;\, M)$ where L is the standard simplex in $\mathbb{P}^n_F$ and M is determined as follows: Take $a_j = 1/(x_1 \ldots x_j)$ and the vertex facing M_0 to be

$$m_0 = [z_0,\ldots,z_n] = [\underbrace{0,\ldots,0}_{n_1-1 \text{ times}}, -a_1, \underbrace{0,\ldots,0}_{n_2-1 \text{ times}}, -a_2, \ldots, \underbrace{0,\ldots,0}_{n_d-1 \text{ times}}, -a_d, 1].$$

The vertex m_j $(1 \leq j \leq n)$ facing M_j is given by

$$m_i = [z_1 + 1, \ldots, z_j + 1, z_{j+1}, \ldots, z_n, 1].$$

If $F = \mathbb{C}$ and $M = \Lambda_{n_1,\ldots,n_d}(x_1,\ldots,x_d)$ a real simplex then the integral

$$\int_{\Delta_M} \frac{\mathrm{d}t_1}{t_1} \wedge \cdots \wedge \frac{\mathrm{d}t_n}{t_n} = (-1)^d \operatorname{Li}_{n_1,\ldots,n_d}(x_1,\ldots,x_d), \tag{2.50}$$

where Δ_M is the convex real span by the faces of M. This is the origin of the name *MPLic pair*.

The key to the study of the double scissors congruence groups A_n is to equip the amalgamation $A_\bullet = \bigoplus_{n\geq 0} A_n$ with a good Hopf algebra structure. In lower weight cases (≤ 4) this can be done explicitly using some geometric combinatorial argument (see [587]) but in general this is still unclear. The product μ is easy to define but the coproduct is highly nontrivial. If this can be done for arbitrary n then we have the following surprising conjecture.

Conjecture 2.10.7. *The restricted coproduct provides a complex*

$$A_{>0} \longrightarrow A_{>0} \otimes A_{>0} \longrightarrow A_{>0} \otimes A_{>0} \otimes A_{>0} \longrightarrow \cdots$$

whose graded n-piece

$$A_n \longrightarrow \bigoplus_{k=1}^{n-1} A_k \otimes A_{n-k} \longrightarrow \cdots$$

gives

$$H^i_{(n)}(A_\bullet, \mathbb{Q}) \cong gr^\gamma_n K_{2n-i}(F)_\mathbb{Q}$$

where $gr^\gamma_n K_{2n-i}(F)_\mathbb{Q}$ is the γ-filtration of K-groups.

In fact, according to the Tannakian formalism the category $\mathsf{MTM}(F)$ of mixed Tate motives over a field F is supposed to be equivalent to the category of graded modules over a certain graded commutative Hopf algebra $\mathcal{A}_\bullet$ (see [44] and [237, Chap. 3]). Therefore the Ext groups in the category $\mathsf{MTM}(F)$ of mixed Tate motives over $\mathrm{Spec}(F)$ are isomorphic to the cohomology of the Hopf algebra $\mathcal{A}_\bullet$. Beilinson *et al.* conjecture that $A_\bullet$ is isomorphic to $\mathcal{A}_\bullet$ and therefore the groups $A_n \otimes \mathbb{Q}$ should have a Hopf algebra structure over $\mathbb{Q}$. This is the primary motivation to study the groups A_n. Further, by Beilinson's conjecture, there is a negatively graded Lie algebra $L_\bullet(F)$ over $\mathbb{Q}$ such that

$$\mathrm{Ext}^i_{\mathsf{MTM}(F)}(\mathbb{Q}(0), \mathbb{Q}(n)) \cong H^i_{(n)}(L_\bullet(F)) \cong \mathrm{gr}^n_\gamma K_{2n-i}(F) \otimes \mathbb{Q}.$$

This is why we have Conjecture 2.10.7.

One hopes to construct the motivic Lie algebra $L_\bullet(F)$ explicitly, at least when the weight is not too large. To this end, let $\mathcal{B}_n(F) = \mathbb{Z}[\mathbb{P}^1_F]/\mathcal{R}_n(F)$ where the subgroup $\mathcal{R}_n(F)$ reflects (conjecturally all of) the functional equations of the single-valued polylog $\mathcal{L}i_n(z)$ defined by Eq. (2.3). A lot of evidence in [232] shows that $\mathcal{B}_n(F) \otimes \mathbb{Q}$ (for $n = 1, 2, 3$) is dual to the motivic Lie algebra $L_{-n}(F)$ (the dual is between the ind and pro $\mathbb{Q}$-vector spaces). It follows from this line of thought that the following conjecture should be true (see [234]).

Conjecture 2.10.8. *Let* $\Pi_n = \bigoplus_{j=1}^{n-1} \mu(A_j \otimes A_{n-j})$ *be the subgroup of prisms of* A_n *then*

$$(A_n/\Pi_n) \otimes \mathbb{Q} \cong L_{-n}(F)^\vee$$

and for $n = 1, 2, 3$

$$(A_n/\Pi_n) \otimes \mathbb{Q} \cong L_{-n}(F)^\vee \cong \mathcal{B}_n(F) \otimes \mathbb{Q}.$$

The first isomorphism means that the dual to the Hopf algebra $A_\bullet \otimes \mathbb{Q}$ is isomorphic to the universal enveloping algebra of the Lie algebra $L_\bullet(F)$. This conjecture is trivial when $n = 1$ since all are isomorphic to $F^\times \otimes \mathbb{Q}$ by the cross ratio (see Exercise 2.3). For $n = 2$ the first isomorphism is proved in [43] by taking $L_{-2}(F)^\vee$ to be the Bloch group $B_2(F)$ which is the quotient group of $\mathbb{Z}[\mathbb{P}^1_F]$ modulo the five-term relations.[1] It is known that for number fields $\mathcal{B}_2(F) \cong B_2(F)$.

Similar to $B_2(F)$ the group $B_3(F)$ is defined by Goncharov [234] as the quotient of $\mathbb{Z}[\mathbb{P}^1_F]$ by the subgroup generated by the generic seven-term relations of some generalized cross ratio r_3 and the Kummer-Spence

[1] In the literature the Bloch group sometimes refers to the subgroup of $B_2(F)$ which is isomorphic to the indecomposable part of $K_3(F)$ modulo torsions.

relations (see [587, Sec. 2.4] for details). The main result of [587] is that modulo torsions one has

$$A_3(F)/\Pi_3(F) \cong B_3(F).$$

2.11 Historical notes

It is well-known that the logarithm was first invented by John Napier in 1614. Although the original intention was to simplify the multiplication and the division to the level of addition and subtraction, logarithms have found many other applications in modern sciences.

From the early days of the modern physics people knew logarithms and $\pi^2 = 6\zeta(2)$ are indispensable for computations in quantum electrodynamics (see, for example, [269, 451, 489]). The trilogs also found its way in perturbative quantum field theory in the late 1940s (see [327, 328][2]). Polylogarithms of larger weights emerged gradually throughout the later half of twentieth century, although they sometimes appeared in disguised forms (see, for example, [34,289,314]). The expressions are often very complicated even for very simple diagrams. For example, on p. 39 of [176] Eden *et al.* pointed out that there are 192 dilog terms in Wu's calculation (see [558]) on a certain fourth order single loop with all scalar particles. It was discovered later that *multiple* zeta values and Euler sums appear in the counterterms of quantum field theory in higher loop electron magnetic moments (see [93, 364]). In recent years, MPLs are related to the study of conformal field theory [201, 337], amplitudes [98, 243, 528] and Feynman graphs [62, 109, 388, 487] in physics as well as Hopf and Lie algebras, combinatorics (double shuffle relations) [413,414], algebraic geometry [218,219,232,234,587,588,592], hyperbolic volumes [429,571], modular forms [220], Mahler measures [356] and cluster algebras [200,230] in mathematics.

The first serious study of classical polylogs seemed to be carried out by Jonquière [310]. Starting from 1950s Leonard Lewin, an engineer, systematically studied polylogs by generalizing many results for the dilog in [372,373]. Apparently he chose the notation Li because of "poly**l**ogarithm **i**ntegral". Later on, with his collaborators, he found many very interesting ladder relations such as the following of order 12 [4]: let $\phi = (\sqrt{5}-1)/2$

[2]Unfortunately the paper contains an error which was corrected only around 1957 in [457, 502]

then

$$\frac{\mathrm{Li}_n(\phi^{12})}{12^{n-1}}-\frac{3}{2}\frac{\mathrm{Li}_n(\phi^6)}{6^{n-1}}-\frac{\mathrm{Li}_n(\phi^4)}{4^{n-1}}+\frac{11}{48}\frac{\mathrm{Li}_n(\phi^2)}{2^{n-1}}=\frac{13}{48}\frac{\log^n\phi}{n!}-\frac{\zeta(2)}{48}\frac{\log^{n-2}\phi}{(n-2)!}$$
$$+\frac{19\zeta(4)}{1728}\frac{\log^{n-4}\phi}{(n-4)!}+\frac{67\zeta(5)}{6912}\frac{\log^{n-5}\phi}{(n-5)!},\quad 1\le n\le 5.$$

A few higher order relations have been discovered numerically since then (see, for example, [30, 138]). For weight 6 and 7 Gangl [216] obtains the first proven ladder relations.

In a series of papers spanning over twenty years K.-T. Chen developed his far-reaching theory of iterated integrals in [126–128] which was unfortunately overlooked for several decades. Since the 1990s, especially because of its applications in multiple zeta and MPLs, its importance has been gradually recognized by more and more mathematicians. Richard Hain, one of Chen's Ph.D. students, has played a pivotal role in this process. His survey article [261] and lecture notes [262] are good sources to learn Chen's theory.

During his study of monodromy properties of $\mathrm{Li}_n(x)$ in [463] Ramakrishnan implicitly came upon the functions defined in Eq. (2.2) (written down explicitly by Zagier [572] and by Wojtkowiak [555] in a modified form). When $n=2$ this coincides with the Bloch-Wigner dilog which was brought to prominence in Bloch's famous Irvine's lectures [53]. Notice that the function defined by Eq. (2.3) also generalizes the Bloch-Wigner function.

Around year 2000 Herbert Gangl investigated various functional equations of polylogs $\mathrm{Li}_n(x)$ for $n\le 7$ (see [214, 215]). Although there is a lot of information available for these functional equations when n is small the most general functional equation of $\mathrm{Li}_n(x)$ is still a mystery whose complete solution will provide key information in the construction of some motivic complexes in [232, 234].

In [219] Gangl and Müller-Stach discovered an algebraic cycle version of the five-term relation Eq. (2.5) satisfied by the dilog in some higher Chow group. They also proved a certain Kummer relation in the cycle-version holds for the trilog. This has been generalized to Goncharov's 22-term relation for the trilog Eq. (2.6) by the author [592].

We should point out that in the literature, there is a different approach to the polylogs, treating both α and s in $\mathrm{Li}_\alpha(s)$ as complex variables, for example, see [146, 310, 460].

Turning to the MPLs, Kummer, Poincaré, Lappo-Danilevsky (see [352, 365, 366, 461], respectively) and K.-T. Chen all studied them in various disguised forms from different points of view ranging from differential

equations to iterated integrals. In the mid 1990s, Goncharov utilized these functions to solve some difficult problems in number theory and algebraic geometry. Sometimes, the term "hyperlogarithm" (for e.g., see [336]) is used instead of MPL, but this is not standard since for other people (see [461]) hyperlogarithm means a different function. In [70], a certain variant of the MPL is also called a multidimensional polylog. Further, Zlobin [617] considered some special values of MPLs, too.

Back in the early 1980s Deligne noticed that the dilog function Li_2 gives rise to a good unipotent variation of mixed Hodge-Tate structures (see [152]). In his survey paper [261] Richard Hain gave a detailed account of the polylog local systems as variations of mixed Hodge structure building on Ramakrishna's explicit computation of the monodromy of polylogs in [463]. The monodromy computation also yields the single-valued variant $\mathcal{L}i_n(z)$ of the polylogs (see [41,52]). These functions in turn have significant applications in arithmetic algebraic geometry such as Zagier's Conjecture 2.2.1. On the other hand, as pointed out by Goncharov [235], "higher cyclotomy theory" should study the MPL motives at roots of unity, not only those of the polylogs. This is one of the reasons why it is important to consider the variations of mixed Hodge structures associated with the MPLs.

The idea in Sec. 2.9 follows closely that of Beilinson and Deligne's paper [41] as given in [52]. By this approach one can easily construct the single-valued version of any MPL as long as one can determine its associated variation matrix explicitly.

In order to understand the motivic origin of MPLs it is beneficial to work with Aomoto polylogs defined first in [14] which are geometrically more intuitive. Our treatment here is quite superficial and avoids a lot of technical details but hopefully still gives the reader a glance of what can possibly be done with this powerful concept. In particular, the double scissors congruence groups are first introduced in [44] and then studied by Beilinson *et al.* in the two companion papers [42, 43]. They were able to construct the weight two motivic complex modulo torsion using the geometric combinatorial properties satisfied by the double scissors congruence group of pairs of triangles in the projective plane. This has been generalized to dimension three in [587]. In the weight four case, Gangl [214, 215] has made some important progress recently building on his previous results. Notice that in their original definition, Beilinson *et al.* imposed another type of relations called *Intersection Additivity*. However, in [586] we showed that these relations follow from the others. Finally, Dupont elucidated the coproduct

structure of double scissors congruence groups in [167].

There are several other types of MPLs. To study the so called multiple zeta star values (see Chap. 10) it is useful to consider the following variation of one variable MPLs. $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ and define

$$Le_{\mathbf{s}}(z) = \sum_{k_1 \geq \cdots \geq k_d \geq 1} \frac{z^{k_1}}{k_1^{s_1} \cdots k_d^{s_d}}, \quad |z| < 1. \tag{2.51}$$

In Chap. 4 we will consider the single-variable multiple polylogs. In the case of the classical polylogs, it is possible to generalize it to a two variable function: for $s, x \in \mathbb{C}$ with $|x| < 1$

$$\mathrm{Li}_s(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^s}.$$

Clearly when $x \to 1$ one recovers the Riemann zeta function. One can also study this function at general s, or at negative integers or even half integers (see [146, 310, 533]). There are several other analogs such as the infinitesimal polylogs, the p-adic polylogs, and Kontsevich's finite polylogs in [349] which are related to the classical polylogs via the "derivation map". See [185] for details.

Exercises

2.1. Verify Eq. (2.4) by differentiation.

2.2. Show that the infinite series

$$\sum_{n \geq 2} \left(\frac{\log^k(n-1)}{n-1} - \frac{\log^k(n)}{n} \right)$$

converges for all $k \in \mathbb{N}_0$.

2.3. Define the cross-ratio

$$r(a, b, c, d) = \frac{(a-c)(b-d)}{(a-d)(b-c)}$$

for all distinct $a, b, c, d \in \mathbb{CP}^1$. Show that for any five distinct points $x_0, \dots, x_4 \in \mathbb{CP}^1$ we have the five term relation of dilog

$$\sum_{i=0}^{4} (-1)^i \mathcal{L}i_2(r(x_0, \dots, \widehat{x_i}, \dots, x_4)) = 0,$$

where $\widehat{x_i}$ means x_i is removed.

2.4. Prove Eq. (2.8) by geometric series expansion.

2.5. Derive Eq. (2.17) from Eq. (2.16) by breaking it into two sums and making an index substitution.

2.6. Prove Thm. 2.4.2 by induction on n using Lemma 2.4.1.

2.7. Derive Eq. (2.25) from Eq. (2.23) in Prop. 2.5.3, using substitutions and the integral expression of MPLs given by Eq. (2.8).

2.8. Verify the variation matrix for the double logarithm in Example 2.5.5 by Definition 2.5.2.

2.9. Check that the variation matrix $\mathcal{M}_{2,1}(x,y)$ associated with $\mathrm{Li}_{2,1}(x,y)$ over S_2 is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathrm{Li}_1(x) & 1 & 0 & 0 & 0 & 0 & 0 \\ \mathrm{Li}_1(xy) & 0 & 1 & 0 & 0 & 0 & 0 \\ \mathrm{Li}_2(x) & \log(x) & 0 & 1 & 0 & 0 & 0 \\ \mathrm{Li}_{1,1}(x,y) & \mathrm{Li}_1(y) & \mathrm{Li}_1\left(\frac{1-xy}{1-y}\right) & 0 & 1 & 0 & 0 \\ \mathrm{Li}_2(xy) & 0 & \log(xy) & 0 & 0 & 1 & 0 \\ \mathrm{Li}_{2,1}(x,y) & \mathrm{Li}_1(y)\log(x) & g(x,y) & \mathrm{Li}_1(y) & \log x - \mathrm{Li}_1(y^{-1}) & 1 \end{bmatrix} \tau_{2,1}(2\pi i)$$

where

$$g(x,y) = -\int_{a_2}^{1} \frac{\mathrm{d}t}{t} \frac{\mathrm{d}t}{t-a_1} = \mathrm{Li}_2(x) - \mathrm{Li}_2(y^{-1}) - \log(xy)\,\mathrm{Li}_1(y^{-1}).$$

The columns of $\mathcal{M}_{2,1}(x,y)$ form the fundamental solutions of the differential equation over S_2

$$d\lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ d\,\mathrm{Li}_1(x) & 0 & 0 & 0 & 0 & 0 \\ d\,\mathrm{Li}_1(xy) & 0 & 0 & 0 & 0 & 0 \\ 0 & d\,\mathrm{Li}_1(y) & d\,\mathrm{Li}_1\left(\frac{1-xy}{1-y}\right) & 0 & 0 & 0 \\ 0 & d\log(x) & 0 & 0 & 0 & 0 \\ 0 & 0 & d\log(xy) & 0 & 0 & 0 \\ 0 & 0 & 0 & d\log(x) & d\,\mathrm{Li}_1(y) & -d\,\mathrm{Li}_1(y^{-1}) & 0 \end{bmatrix} \lambda$$

Describe the weight and Hodge filtration in detail.

2.10. Check the variation matrix $\mathcal{M}_{1,2}(x,y)$ associated with $\mathrm{Li}_{1,2}(x,y)$ over S_2 is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \mathrm{Li}_1(x) & 1 & 0 & 0 & 0 & 0 \\ \mathrm{Li}_1(xy) & 0 & 1 & 0 & 0 & 0 \\ \mathrm{Li}_{1,1}(x,y) & \mathrm{Li}_1(y) & \mathrm{Li}_1\left(\frac{1-xy}{1-y}\right) & 1 & 0 & 0 \\ \mathrm{Li}_2(xy) & 0 & \log(xy) & 0 & 1 & 0 \\ \mathrm{Li}_{1,2}(x,y) & \mathrm{Li}_2(y) & f(x,y) & \log y\,\mathrm{Li}_1(x) & 1 \end{bmatrix} \tau_{1,2}(2\pi i)$$

where

$$f(x,y) = -\int_{a_2}^{1} \frac{\mathrm{d}t}{t-a_1}\frac{\mathrm{d}t}{t} = \mathrm{Li}_2(y^{-1}) - \mathrm{Li}_2(x) + \log(xy)\,\mathrm{Li}_1(x).$$

The columns of $\mathcal{M}_{1,2}(x,y)$ form the fundamental solutions of the differential equation over S_2

$$d\lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ d\,\mathrm{Li}_1(x) & 0 & 0 & 0 & 0 & 0 \\ d\,\mathrm{Li}_1(xy) & 0 & 0 & 0 & 0 & 0 \\ 0 & d\,\mathrm{Li}_1(y) & d\,\mathrm{Li}_1\left(\frac{1-xy}{1-y}\right) & 0 & 0 & 0 \\ 0 & 0 & d\log(xy) & 0 & 0 & 0 \\ 0 & 0 & 0 & d\log y & d\,\mathrm{Li}_1(x) & 0 \end{bmatrix} \lambda$$

Describe the weight and Hodge filtration in detail.

2.11. Compute the limit MHS of the triple logarithm $\mathfrak{L}_3(x,y,z)$ on $\mathcal{D}_{11} = \{x=1\}$ over $\{(1,y,z) : yz(1-y)(1-z)(1-yz) \neq 0\}$ along the vector $\partial/\partial x$. Show that the variation matrix is essentially $\mathcal{M}_{2,1}\Big(\frac{y}{y-1}, \frac{z(y-1)}{yz-1}\Big)$. This gives some evidence for Conjecture 2.8.2.

2.12. Prove the functional equation expressed by Eq. (2.43) for the dilog by differentiation.

2.13. Verify the identity 2.45 by differentiation. Then use it to compute the integral (see (2.97) on p. 812 of [489]):

$$\int_0^1 \left(\frac{\log((1+x)/2)}{1-x} - \frac{\log((1-x)/2)}{1+x}\right)\frac{\mathrm{d}x}{x}$$

by first expanding the integrand using power series and then expressing it using double logarithms, dilogs and products of logarithms.

2.14. Prove the identity

$$\begin{aligned}\mathrm{Li}_{1,1,1}(x,y,z) &= \mathrm{Li}_{1,2}\Big(\frac{1}{1-y}, \frac{z(1-y)}{z-1}\Big) - \mathrm{Li}_{1,2}\Big(\frac{1-yx}{1-y}, \frac{z(1-y)}{z-1}\Big) \\ &\quad - \log(1-x)\,\mathrm{Li}_2\Big(\frac{z}{z-1}\Big) - \log(1-z)\,\mathrm{Li}_{1,1}(x,y)\end{aligned}$$

by differentiation. Then use Eq. (2.49) to show that

$$\begin{aligned}\mathcal{L}i_{1,1,1}(x,y,z) = {}& \mathcal{L}i_3\Big(\frac{(y-1)(1-xyz)}{y(1-x)(1-z)}\Big) + \mathcal{L}i_3\Big(\frac{y}{y-1}\Big) + \mathcal{L}i_3(yz) - \mathcal{L}i_3\Big(\frac{1-xyz}{1-z}\Big) \\ & - \mathcal{L}i_3\Big(\frac{1-xyz}{yz(1-x)}\Big) - \mathcal{L}i_3\Big(\frac{y-xy}{y-1}\Big) - \mathcal{L}i_3\Big(\frac{y-yz}{y-1}\Big) + \mathcal{L}i_3(1-z)\end{aligned}$$

modulo products of logarithms and dilogs.

2.15. Prove Eq. (2.50) for $d = 2$ and $d = 3$.

2.16. How can one decompose the pair in Figure 2.3 into two dilogic pairs? **Hint:** The dilogic pairs (L, M) are determined by the following geometric combinatorial properties: there are exactly two edges of M each of which contains a vertex of L while exactly one vertex of M facing one the above two edges must lie on an edge of L.

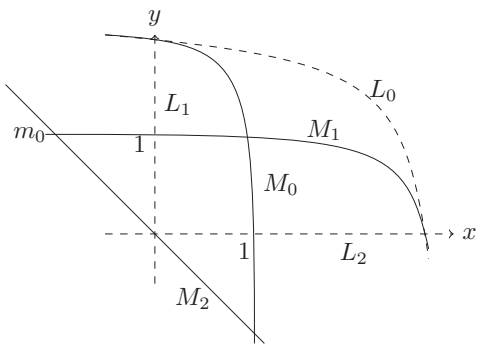


Figure 2.3: Decompose the following into two dilogic pairs.

Chapter 3

Multiple zeta values

Most of the on-going studies on multiple zeta functions concentrate on the algebraic aspect of the multiple zeta values (MZVs for abbreviation), which are the convergent special values of multiple zeta functions at positive integers. Sometimes MZVs are also called multi-zeta values, or poly-zeta values, or Euler-Zagier sums.

Throughout this chapter we will use s_j's to denote integer variables.

3.1 Shuffle product and Euler's decomposition formula

Kontsevich first observed that MZVs can be represented by the iterated integrals. Put

$$\mathtt{x}_0 = \frac{\mathrm{d}t}{t}, \qquad \mathtt{x}_1 = \frac{\mathrm{d}t}{t-1}.$$

Setting all the variables $\mathtt{x}_1 = \cdots = \mathtt{x}_d = 1$ in Eq. (2.8) we obtain an iterated integral expression of MZVs

$$\zeta(s_1, \ldots, s_d) = (-1)^d \int_0^1 \mathtt{x}_0^{s_1-1}\mathtt{x}_1 \cdots \mathtt{x}_0^{s_d-1}\mathtt{x}_1. \tag{3.1}$$

Using this we can prove the following Euler's decomposition formula quite easily.

Theorem 3.1.1. *Let $m, n \geq 2$ be two positive integers. Then*

$$\zeta(m)\zeta(n) = \sum_{\substack{i\geq 2, j\geq 1 \\ i+j=m+n}} \left[\binom{i-1}{m-1} + \binom{j-1}{n-1}\right]\zeta(i,j). \tag{3.2}$$

Proof. Set $b_1 = b_2 = \mathtt{x}_1$. By Eq. (3.1) and using Lemma 2.1.2(iv)

$$\begin{aligned}
\zeta(m)\zeta(n) &= \int_0^1 \mathtt{x}_0^{m-1} b_1 \int_0^1 \mathtt{x}_0^{n-1} b_2 \\
&= \sum_{k=0}^{n-1} \binom{m-1+k}{k} \int_0^1 \mathtt{x}_0^{m-1+k} b_1 \mathtt{x}_0^{n-1-k} b_2 \\
&+ \sum_{k=0}^{m-1} \binom{n-1+k}{k} \int_0^1 \mathtt{x}_0^{n-1+k} b_2 \mathtt{x}_0^{m-1-k} b_1 \\
&= \sum_{k=0}^{n-1} \binom{m-1+k}{k} \zeta(m+k, n-k) \\
&+ \sum_{k=0}^{m-1} \binom{n-1+k}{k} \zeta(n+k, m-k)
\end{aligned}$$

which is equivalent to Eq. (3.2). □

3.2 Hoffman's algebraic setup

One of the key ideas to study the $\mathbb{Q}$-linear relations between MZVs is to formalize them using the theory of words. In order to realize fully the product structures we actually need two algebras equipped with two different multiplications whose underlying set of words are closely related.

3.2.1 *The shuffle relations*

With the aid of Chen's theory one can define a shuffle algebra reflecting the rule of multiplication of two MZVs using these integrals. Specifically, as we have seen in the proof of Euler's decomposition formula in the proceeding section, we can apply Chen's formula (*i.e.*, Lemma 2.1.2(iv)) for the product of two iterated integrals and define some noncommutative polynomial algebra $\mathfrak{A} = \mathbb{Q}\langle \mathtt{x}_0, \mathtt{x}_1 \rangle$ so that MZVs can be viewed as values of a homomorphism (in fact, two different homomorphisms) from a sub-algebra of $\mathfrak{A}$ to the real number field $\mathbb{R}$. As an immediate consequence, the product of two MZVs can be expressed as a $\mathbb{Z}$-linear combination of MZVs, which is called a shuffle relation.

Definition 3.2.1. Set $\mathtt{e}$ to be empty word. Define $\mathfrak{A} = \mathbb{Q}\langle \mathtt{x}_0, \mathtt{x}_1 \rangle$ to be the (weight) graded noncommutative polynomial $\mathbb{Q}$-algebra of words generated by the two letters $\mathtt{x}_0$ and $\mathtt{x}_1$. The number of letters in the word $\mathbf{w}$, denoted

by $|\mathbf{w}|$, is called the *weight*[1] of $\mathbf{w}$, and the number of $\mathtt{x}_1$'s, the *depth*.

Definition 3.2.2. Let $\mathfrak{A}^0$ be the subalgebra of $\mathfrak{A}$ generated by words not beginning with $\mathtt{x}_1$ and not ending with $\mathtt{x}_0$. The words in $\mathfrak{A}^0$ are called *admissible words*. Further, we define $\mathfrak{A}^1 = \mathbb{Q}\mathtt{e} \oplus \mathfrak{A}\mathtt{x}_1$ to be the subalgebra of $\mathfrak{A}$, which is a noncommutative polynomial algebra on generators $\mathtt{y}_k = \mathtt{x}_0^{k-1}\mathtt{x}_1$ for all $k \geq 1$.

Clearly, there is a many-to-one correspondence between MZVs and admissible words and the definition of weight and depth is consistent. For example, Euler already proved $\zeta(3) = \zeta(2,1)$. So the depth of a MZV depends on its expression, however, the weight doesn't by a folklore conjecture.

Definition 3.2.3. Let $\mathfrak{A}_{ш}$ be the algebra of $\mathfrak{A}$ together with the multiplication defined by shuffle product $ш$: for any letters $x, y \in \{\mathtt{x}_0, \mathtt{x}_1\}$ and words $\mathbf{u}, \mathbf{v}$ in $\mathfrak{A}$ we define recursively

$$\mathtt{e} ш \mathbf{u} = \mathbf{u} ш \mathtt{e} = \mathbf{u}, \qquad x\mathbf{u} ш y\mathbf{v} = x(\mathbf{u} ш y\mathbf{v}) + y(x\mathbf{u} ш \mathbf{v}).$$

Here we write, for e.g., $\mathtt{x}_0\mathbf{u} ш \mathtt{x}_1\mathbf{v}$ rather than $(\mathtt{x}_0\mathbf{u}) ш (\mathtt{x}_1\mathbf{v})$ to minimize the number of parentheses. Namely, concatenation has higher operational priority than $ш$.

Proposition 3.2.4. *We have*

(i) $\mathfrak{A}_{ш}$ *is a commutative and associative* $\mathbb{Q}$*-algebra.*
(ii) *For any admissible word* $\mathbf{w}$ *of depth* d *let* $\zeta(\mathbf{w}) = (-1)^d \int_0^1 \mathbf{w}$. *Extend the map to a* $\mathbb{Q}$*-linear map* $\zeta_{ш} : \mathfrak{A}^0_{ш} \longrightarrow \mathbb{R}$. *Then* $\zeta_{ш}$ *is an algebra homomorphism.*

Proof. (i). We only prove the commutativity and leave the associativity to the interested reader. For commutativity we only need to show that for all words $\mathbf{u}$ and $\mathbf{v}$

$$\mathbf{u} ш \mathbf{v} = \mathbf{v} ш \mathbf{u}. \tag{3.3}$$

We prove this by induction on $|\mathbf{u}|+|\mathbf{v}|$. If either $\mathbf{u}$ or $\mathbf{v}$ is $\mathtt{e}$ then Eq. (3.3) is clear by the definition. Now assume Eq. (3.3) hold for any two words whose sum of weight is $< \ell$ for some positive integer ℓ. Let $\mathrm{dp}(\mathbf{u}) + \mathrm{dp}(\mathbf{v}) = \ell$

[1]In the study of words we usually use the term "length". But here we avoid using this term because in the study of MZVs this may cause confusion as some authors use it to mean depth.

and suppose without loss of generality $\mathbf{u} = x\mathbf{u}'$, and $\mathbf{v} = y\mathbf{v}'$ where $x, y \in \{\mathtt{x}_0, \mathtt{x}_1\}$. Then by the definition

$$\mathbf{u} \sqcup\!\sqcup \mathbf{v} = x(\mathbf{u}' \sqcup\!\sqcup y\mathbf{v}') + y(x\mathbf{u}' \sqcup\!\sqcup \mathbf{v}') = x(y\mathbf{v}' \sqcup\!\sqcup \mathbf{u}') + y(\mathbf{v}' \sqcup\!\sqcup x\mathbf{u}')$$

by inductive assumption. But by the definition the last expression is equal to $\mathbf{v} \sqcup\!\sqcup \mathbf{u}$. Thus Eq. (3.3) is proved and therefore $\mathfrak{A}_{\sqcup\!\sqcup}$ is a commutative $\mathbb{Q}$-algebra.

(ii). We need to show that for any $\mathbf{u}, \mathbf{v} \in \mathfrak{A}^0_{\sqcup\!\sqcup}$ we have

$$\zeta_{\sqcup\!\sqcup}(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) = \zeta_{\sqcup\!\sqcup}(\mathbf{u})\zeta_{\sqcup\!\sqcup}(\mathbf{v}).$$

To avoid ambiguity let's denote $\sqcup\!\sqcup$ the shuffle operation defined in Definition 3.2.3 and $\sqcup\!\sqcup'$ the shuffle operation defined by

$$(l_1 \cdots l_r) \sqcup\!\sqcup' (l_{r+1} \cdots l_{r+s}) := \sum_{\sigma} l_{\sigma(1)} \cdots l_{\sigma(r+s)}, \tag{3.4}$$

where the letter $l_j \in \{\mathtt{x}_0, \mathtt{x}_1\}$ for all $j = 1, \ldots, r+s$, and σ ranges over all shuffles of type (r, s), *i.e.*, permutations σ of $r+s$ letters with $\sigma^{-1}(1) < \cdots < \sigma^{-1}(r)$ and $\sigma^{-1}(r+1) < \cdots < \sigma^{-1}(r+s)$. We now prove that

$$\mathbf{u} \sqcup\!\sqcup' \mathbf{v} = \mathbf{u} \sqcup\!\sqcup \mathbf{v} \quad \forall \mathbf{u}, \mathbf{v} \in \mathfrak{A} \tag{3.5}$$

by induction on $|\mathbf{u}| + |\mathbf{v}|$. Clearly when $\mathbf{u}, \mathbf{v} \in \mathfrak{A}$ then $|\mathbf{u}| + |\mathbf{v}| \geq 0$. If $|\mathbf{u}| + |\mathbf{v}| \leq 1$ then Eq. (3.5) is trivial. Now assume Eq. (3.5) holds for $|\mathbf{u}| + |\mathbf{v}| = k \geq 1$ and let $U, V \in \mathfrak{A}$ with $|U| + |V| = k+1$. If either $U = \mathtt{e}$ or $V = \mathtt{e}$ then Eq. (3.5) is clear. Otherwise we may assume $U = x\mathbf{u}, V = y\mathbf{v} \in \mathfrak{A}$ for some $x, y \in \{\mathtt{x}_0, \mathtt{x}_1\}$. By the definition and induction assumption

$$U \sqcup\!\sqcup V = x(\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) + y(x\mathbf{u} \sqcup\!\sqcup \mathbf{v}) = x(\mathbf{u} \sqcup\!\sqcup' y\mathbf{v}) + y(x\mathbf{u} \sqcup\!\sqcup' \mathbf{v}) = U \sqcup\!\sqcup' V \tag{3.6}$$

because, according to Eq. (3.4), we can partition the terms in $U \sqcup\!\sqcup' V$ into two disjoint sets depending on whether the leading letter is $\mathtt{x}_0$ or $\mathtt{x}_1$.

The proposition now follows immediately from Eqs. (3.1), (3.5) and Lemma 2.1.2(iv). □

We now provide an interesting result which will be used later.

Definition 3.2.5. Let p, q and j be non-negative integers such that $\min(p, q) \geq j$. Let $T_{p+q,j}$ denote the sum of all those words occurring in $(\mathtt{x}_0\mathtt{x}_1)^p \sqcup\!\sqcup (\mathtt{x}_0\mathtt{x}_1)^q$ that contain the subword $\mathtt{x}_0^2$ exactly j times.

Proposition 3.2.6. *For any non-negative integers p and q we have*

$$(\mathtt{x}_0\mathtt{x}_1)^p \sqcup\!\sqcup (\mathtt{x}_0\mathtt{x}_1)^q = \sum_{j=0}^{\min(p,q)} \binom{p+q-2j}{p-j} \cdot 4^j T_{p+q,j}.$$

Proof. Let $\mathbf{w}$ be any word term in the sum $T_{p+q,j}$. We now count the multiplicity of $\mathbf{w}$ in $(\mathtt{x}_0\mathtt{x}_1)^p \sqcup\!\sqcup (\mathtt{x}_0\mathtt{x}_1)^q$, namely, the number of ways the letters of $\mathbf{w}$ can be colored in either blue (from $(\mathtt{x}_0\mathtt{x}_1)^p$) or red (from $(\mathtt{x}_0\mathtt{x}_1)^q$) following the shuffle product rule.

By the definition there are $p+q$ $\mathtt{x}_0$'s in $\mathbf{w}$, of which $2j$ $\mathtt{x}_0$'s are contained in factors $\mathtt{x}_0^2$ and $p+q-2j$ "single" $\mathtt{x}_0$'s have only $\mathtt{x}_1$'s as neighbors. This implies there are exactly j factors $\mathtt{x}_1^2$ (see Exercise 3.1). For each factor $\mathtt{x}_0^2$, one $\mathtt{x}_0$ is red and the other is blue because of the shuffle product rule. Similar pattern holds for factors $\mathtt{x}_1^2$. These contribute a factor of $2^j \cdot 2^j = 4^j$ to the multiplicity of $\mathbf{w}$. On the other hand, exactly $p-j$ of the "single" $\mathtt{x}_0$'s are colored blue, contributing a factor of $\binom{p+q-2j}{p-j}$ to the multiplicity of $\mathbf{w}$. We can now complete the proof of the proposition by observing that the coloring of the "single" $\mathtt{x}_1$'s is now determined uniquely. □

Proposition 3.2.7. *For any positive integers $n \geq m$ we have*

$$\sum_{k=1-n}^{n} (-1)^k k^{2m-1} \left[(\mathtt{x}_0\mathtt{x}_1)^{n-k} \sqcup\!\sqcup (\mathtt{x}_0\mathtt{x}_1)^{n-1+k}\right] = \sum_{j=1}^{m} 4^{n-j} T_{2n-1,n-j} \cdot a_{2m-1,j}, \tag{3.7}$$

$$\sum_{k=-n}^{n} (-1)^k k^{2m} \left[(\mathtt{x}_0\mathtt{x}_1)^{n-k} \sqcup\!\sqcup (\mathtt{x}_0\mathtt{x}_1)^{n+k}\right] = \sum_{j=1}^{m} 4^{n-j} T_{2n,n-j} \cdot a_{2m,j}, \tag{3.8}$$

where for $j = 1, \ldots, m$ we set

$$a_{2m-1,j} = \sum_{k=1-j}^{j} (-1)^k k^{2m-1} \binom{2j-1}{j-k}, \;\; a_{2m,j} = \sum_{k=-j}^{j} (-1)^k k^{2m} \binom{2j}{j-k}. \tag{3.9}$$

Proof. By Prop. 3.2.6 the left-hand side of Eq. (3.7) can be written as

$$\left(\sum_{k=1-n}^{0} \sum_{j=0}^{n-1+k} + \sum_{k=1}^{n} \sum_{j=0}^{n-k} \right) (-1)^k k^{2m-1} \binom{2n-1-2j}{n-k-j} \cdot 4^j T_{2n-1,j}$$

which after reordering is equal to

$$\sum_{j=0}^{n-1} 4^j T_{2n-1,j} \cdot \sum_{k=1-n+j}^{n-j} (-1)^k k^{2m-1} \binom{2n-1-2j}{n-k-j}.$$

Then from Eq. (3.9) (with $n-j$ in the place of j)

$$\sum_{k=1-n}^{n}(-1)^k k^{2m-1}\left[(\mathtt{x}_0\mathtt{x}_1)^{n-k}\sqcup\!\sqcup(\mathtt{x}_0\mathtt{x}_1)^{n-1+k}\right] = \sum_{j=n-m}^{n-1}4^{n-j}T_{2n-1,j}\cdot a_{2m-1,n-j}.$$

The proof of Eq. (3.8) is similar and is left to the interested reader. □

3.2.2 *The stuffle relations*

On the other hand, it is easy to see that

$$\zeta(2)\zeta(3)=\zeta(2,3)+\zeta(3,2)+\zeta(5).$$

This can be generalized to other MZVs without too much difficulty. Relations such as this are called stuffle relations. Indeed, one can even extract some formal algebra structures from MZVs.

Definition 3.2.8. Let $\mathfrak{A}_*$ be the algebra of $\mathfrak{A}$ together with the multiplication defined by the stuffle product $*$: for any nonnegative integers m,n and any words $\mathbf{u},\mathbf{v}$ in $\mathfrak{A}$ we define:

$$\mathsf{e}*\mathbf{u}=\mathbf{u}*\mathsf{e}=\mathbf{u},$$
$$\mathtt{x}_0^m*\mathbf{u}=\mathbf{u}*\mathtt{x}_0^m=\mathbf{u}\mathtt{x}_0^m,$$
$$\mathtt{x}_0^m\mathtt{x}_1\mathbf{u}*\mathtt{x}_0^n\mathtt{x}_1\mathbf{v}=\mathtt{x}_0^m\mathtt{x}_1\big(\mathbf{u}*\mathtt{x}_0^n\mathtt{x}_1\mathbf{v}\big)+\mathtt{x}_0^n\mathtt{x}_1\big(\mathtt{x}_0^m\mathtt{x}_1\mathbf{u}*\mathbf{v}\big)-\mathtt{x}_0^{m+n+1}\mathtt{x}_1(\mathbf{u}*\mathbf{v}).$$

Here, again, concatenation has higher operational priority than $*$.

Similar to Prop. 3.2.4 one can prove the following fact for the stuffle product.

Proposition 3.2.9. *We have*

(i) $\mathfrak{A}_*$ *is a commutative and associative algebra.*
(ii) *For any admissible word* $\mathbf{w}\in\mathfrak{A}^0$ *of depth* d *let* $\zeta(\mathbf{w})=(-1)^d\int_0^1\mathbf{w}$. *We then can extend the map to a* $\mathbb{Q}$*-linear map* $\zeta_*:\mathfrak{A}_*^0\longrightarrow\mathbb{R}$. *Then* ζ_* *is an algebra homomorphism.*

Proof. Similar to Prop. 3.2.4 this can be proved by induction using depth. We leave the details to the interested reader. □

3.3 Double shuffle relations

By comparing Prop. 3.2.4 and Prop. 3.2.9 we set

Definition 3.3.1. *Finite double shuffle relations* (finite DBSFs) of MZVs are those $\mathbb{Q}$-linear relations generated by

$$\zeta(\mathbf{u} * \mathbf{v} - \mathbf{u} \sqcup\!\sqcup \mathbf{v})$$

for all admissible words $\mathbf{u}$ and $\mathbf{v}$.

Example 3.3.2. The simplest finite DBSF is obtained by computing $\zeta(2)^2$ using two different ways. First,

$$\begin{aligned}\zeta(2)^2 &= \int_0^1 \mathtt{x}_0\mathtt{x}_1 \cdot \int_0^1 \mathtt{x}_0\mathtt{x}_1 = \int_0^1 \mathtt{x}_0\mathtt{x}_1 \sqcup\!\sqcup \mathtt{x}_0\mathtt{x}_1 \\ &= 2\int_0^1 \mathtt{x}_0\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1 + 4\int_0^1 \mathtt{x}_0\mathtt{x}_0\mathtt{x}_1\mathtt{x}_1 = 2\zeta(2,2) + 4\zeta(3,1).\end{aligned}$$

On the other hand, by the stuffle relation

$$\zeta(2)^2 = 2\zeta(2,2) + \zeta(4).$$

Therefore we get the finite DBSF

$$\zeta(4) = 4\zeta(3,1).$$

3.3.1 *Lyndon words*

Notice that it is impossible to prove the famous identity $\zeta(2,1) = \zeta(3)$ by finite DBSFs because the smallest weight of MZVs is two given by $\zeta(2)$ and thus $\zeta(3)$ cannot be a product of two MZVs. To solve this puzzle, we need the *regularized* DBSFs. The key idea is contained in the next theorem. Before stating it we need some results concerning the so-called Lyndon words.

Definition 3.3.3. Let A be any alphabet with an complete order $<$ on the set of words A^*. A nonempty word $\mathbf{w} \in A^*$ is called *Lyndon* if $\mathbf{w} < \mathbf{v}$ for any nontrivial factorization $\mathbf{w} = \mathbf{uv}$.

Hence, to define Lyndon words in $\mathfrak{A} = \mathbb{Q}\langle \mathtt{x}_0, \mathtt{x}_1\rangle$ we first need to put an order on it. If $\mathbf{u}, \mathbf{v}$ are words with $\mathbf{v} \neq \mathtt{e}$, set $\mathbf{u} < \mathbf{uv}$. For any words $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, set $\mathbf{w}_1\mathtt{x}_0\mathbf{w}_2 < \mathbf{w}_1\mathtt{x}_1\mathbf{w}_3$.

Example 3.3.4. It is straightforward to see that $\mathtt{x}_0$ is the smallest London word while there is no largest word. Lyndon words can have the form $\mathtt{y}_{s_1}\cdots\mathtt{y}_{s_d} := \mathtt{x}_0^{s_1-1}\mathtt{x}_1\cdots\mathtt{x}_0^{s_d-1}\mathtt{x}_1$ with $s_j < s_1$ for all $j > 1$. But the $\mathtt{x}_0^2\mathtt{x}_1\mathtt{x}_0^2\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1$ is also a Lyndon word.

Lemma-Definition 3.3.5. *Every word* $\mathbf{w} \in \mathfrak{A}$ *has a unique factorization as non-increasing Lyndon words, called the Lyndon factorization of* $\mathbf{w}$.

Proof. Write $\mathbf{w} = \mathtt{y}_{s_1} \cdots \mathtt{y}_{s_d} \mathtt{x}_0^m$ for some positive integers $s_1, \ldots, s_d$ and some nonnegative integer m. We will define a subsequence $\{i_1, \ldots, i_k\}$ of $\{1, \ldots, d\}$ such that for each $j = 1, \ldots, k$ $s_{i_j} > s_\ell$ for all $\ell \in \{i_j + 1, \ldots, i_{j+1} - 1\}$ and $s_{i_j} \leq s_{i_{j+1}}$, where $i_{k+1} = d + 1$. Clearly $\mathbf{w} = \mathbf{w}_1 \cdots \mathbf{w}_k \mathtt{x}_0^m$ where

$$\mathbf{w}_j = \mathtt{y}_{s_{i_j}} \cdots \mathtt{y}_{s_{i_{j+1}-1}}$$

is a Lyndon word for every j. However, it is not always true that $\mathbf{w}_j \geq \mathbf{w}_{j+1}$ for all j. Problem may occur if $s_{i_j} = s_{i_{j+1}}$. So we need to modify the factorization by possibly concatenating some of the w_j's as follows.

For simplicity we write $I_j(\mathtt{y}_{t_1} \cdots \mathtt{y}_{t_l}) = t_j$ for each composition $(t_1, \ldots, t_l) \in \mathbb{N}^l$. Continue with the notation in the preceding paragraph, if $I_1(\mathbf{w}_{j-1}) > I_1(\mathbf{w}_j) = I_1(\mathbf{w}_{j+1}) = \cdots = I_1(\mathbf{w}_\ell) > I_1(\mathbf{w}_{\ell+1})$ then we compare $I_j(\mathbf{w}_\alpha)$ and $I_j(\mathbf{w}_{\alpha+1})$ for all $\alpha = 1, \ldots, \ell - 1$ and for increasing $j = 2, 3, \cdots$, until we find either of the two:

(i). $I_{j-1}(\mathbf{w}_\alpha) = I_{j-1}(\mathbf{w}_{\alpha+1}) \neq 0$ but $I_j(\mathbf{w}_\alpha) \neq I_j(\mathbf{w}_{\alpha+1})$. If $I_j(\mathbf{w}_\alpha) > I_j(w_{\alpha+1}) > 0$ or $I_j(\mathbf{w}_\alpha) = 0 < I_j(\mathbf{w}_{\alpha+1})$ then we concatenate $\mathbf{w}_\alpha$ and $\mathbf{w}_{\alpha+1}$, otherwise we don't. Here $I_j(\mathbf{w}_\alpha) = 0$ means $\mathbf{w}_\alpha$ has depth $j - 1$ only.

(ii). $\mathbf{w}_\alpha = \mathbf{w}_{\alpha+1}$. Then we don't concatenate them.

Now it is not hard to see we do have non-increasing Lyndon words in the factorization. □

Example 3.3.6. We have the following Lyndon factorizations

$$(\mathtt{x}_1)(\mathtt{x}_0), \quad (\mathtt{x}_1)(\mathtt{x}_0\mathtt{x}_1)(\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1)(\mathtt{x}_0), \quad (\mathtt{x}_0^2\mathtt{x}_1)(\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0^2\mathtt{x}_1\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1),$$
$$(\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1)(\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0^2\mathtt{x}_1), \quad (\mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1)(\mathtt{x}_0^3\mathtt{x}_1).$$

Our next goal is to show that very word $\mathbf{w} \in \mathfrak{A}^1_* = (\mathfrak{A}^1, *)$ has a unique expression as a $\mathbb{Q}$-linear combination of the stuffle products of non-increasing Lyndon words.

Lemma 3.3.7. *For* $\mathtt{e} \neq \mathbf{w} \in \mathfrak{A}$ *let* $\mathrm{LW}(\mathbf{w})$ *be the largest word appearing in* $\mathbf{w}$. *Let* $\mathbf{w}, \mathbf{u}, \mathbf{v} \in \mathfrak{A}$ *such that* $\mathbf{u} < \mathbf{v}$. *Then* $\mathrm{LW}(\mathbf{u} * \mathbf{w}) < \mathrm{LW}(\mathbf{v} * \mathbf{w})$.

Proof. Use induction on the weight sum $|\mathbf{u}| + |\mathbf{v}|$. □

Lemma 3.3.8. *Let* $\mathbf{u}, \mathbf{v} \in \mathfrak{A}$ *such that* $\mathbf{u}$ *is a Lyndon word and* $\mathbf{u} \geq \mathbf{v}$. *Then* $\mathrm{LW}(\mathbf{u} * \mathbf{v}) = \mathbf{u}\mathbf{v}$.

Proof. Since $\mathbf{x}_0$ commutes with every word under stuffle we may assume that $\mathbf{u} = \mathbf{y}_{s_1} \cdots \mathbf{y}_{s_d}$ and $\mathbf{v} = \mathbf{y}_{t_1} \cdots \mathbf{y}_{t_\ell} \in \mathfrak{A}^1$ without loss of generality. Recall that in $\mathbf{u} * \mathbf{v}$ there are two kinds of words, one is obtained by pure shuffle, while the other involves stuffing. Assume $s_1 = t_1, \ldots, s_i = t_i$ and, $0 < s_{i+1} < t_{i+1}$ or $t_{i+1} = 0$ since $\mathbf{u} \geq \mathbf{v}$. Then in each term of the shuffle, except for $\mathbf{uv}$, $\mathbf{y}_{t_1}$ appears between $\mathbf{y}_{s_{j-1}}$ and $\mathbf{y}_{s_j}$ for some $1 \leq j \leq d$ ($\mathbf{y}_{s_0}$ is vacuous). If $j > i$ then clearly $s_j < s_1 \leq t_1$ since u is Lyndon word and thus this term is smaller than $\mathbf{uv}$. If $j \leq i$ then all of $\mathbf{y}_{s_j}$ up to $\mathbf{y}_{s_i}$ in this shuffle term are pushed to the right and thus producing a smaller term than $\mathbf{uv}$ since $\mathbf{uv}$'s $\mathbf{y}_{s_{i+1}}$-term corresponds to some $\mathbf{y}_{t_k}$ or some $\mathbf{y}_{s_k}$ with $k \leq i$. Finally, if some stuffing is involved, then $t_1, \ldots, t_i$ cannot become smaller resulting a stuffing term smaller than $\mathbf{uv}$. This shows that $\mathrm{LW}(\mathbf{u} * \mathbf{v}) = \mathbf{uv}$ and the lemma is proved. □

Lemma 3.3.9. *Let $\mathbf{u}, \mathbf{v} \in \mathfrak{A}^1$ such that $\mathbf{u}$ is a Lyndon word and $\mathbf{u} \geq \mathbf{v}$. Then for any nonnegative integer n we have* $\mathrm{LW}(\mathbf{u} * \mathbf{u}^n\mathbf{v}) = \mathbf{u}^{n+1}\mathbf{v}$.

Proof. We now use induction on n. Lemma 3.3.8 takes care of the case $n = 0$. By the same consideration as in Lemma 3.3.8 we see that $\mathrm{LW}(\mathbf{u} * \mathbf{u}^n\mathbf{v})$ can only come from either $\mathbf{u}(\mathbf{u}^n\mathbf{v}) = \mathbf{u}^{n+1}\mathbf{v}$ or $\mathbf{u}(\mathbf{u} * \mathbf{u}^{n-1}\mathbf{v})$. By induction we are done. □

Lemma 3.3.10. *Let $\mathbf{w}_1 \geq \mathbf{w}_2 \geq \cdots \geq \mathbf{w}_k$ be a sequence of Lyndon words. Then*

$$\mathrm{LW}(\mathbf{w}_1 * \mathbf{w}_2 * \cdots * \mathbf{w}_k) = \mathbf{w}_1\mathbf{w}_2 \cdots \mathbf{w}_k.$$

Proof. We proceed by induction on k. The case $k = 1$ is trivial. Suppose the theorem is true if the number of Lyndon words is less than k. We then have two cases.

1) If $\mathbf{w}_1 > \mathbf{w}_2$ then $\mathbf{w}_1 > \mathbf{w}_2 \cdots \mathbf{w}_k$ since $\mathbf{w}_j$'s are all Lyndon words. By Lemmas 3.3.7, Lemma 3.3.8 and the induction assumption we get

$$\mathrm{LW}(\mathbf{w}_1 * \mathbf{w}_2 * \cdots * \mathbf{w}_k) = \mathrm{LW}(\mathbf{w}_1 * \mathbf{w}_2 \cdots \mathbf{w}_k) = \mathbf{w}_1\mathbf{w}_2 \cdots \mathbf{w}_k.$$

2) If $\mathbf{w}_1 = \mathbf{w}_2$ then $\mathbf{w}_2 \cdots \mathbf{w}_k$ has the form $\mathbf{w}_1^n\mathbf{v}$ for $\mathbf{v} < \mathbf{w}_1$, and the theorem follows from Lemmas 3.3.7, Lemma 3.3.9 and the induction assumption. □

Theorem 3.3.11. *As a commutative algebra, $\mathfrak{A}_*$ is the free polynomial algebra on the Lyndon words.*

Proof. Let $\mathbf{w} \in \mathfrak{A}$ and let $\mathbf{w} = \mathbf{w}_1 \cdots \mathbf{w}_k$ be its Lyndon factorization guaranteed by Lemma 3.3.5. Then by Lemma 3.3.10 we know there is some $c \in \mathbb{Q}$ such that $\mathrm{LW}(\mathbf{w} - c\mathbf{w}_1 * \cdots * \mathbf{w}_k)$ is smaller than $\mathbf{w}$. Since there are only finitely many words having the same degree as $\mathbf{w}$, by repeating this we see that $\mathbf{w}$ can be written as a $\mathbb{Q}$-linear combination of the stuffle products of non-increasing Lyndon words. Clearly this expression is unique.

Let $\mathfrak{L}$ be the polynomial algebra with one generator for each Lyndon word. The above shows that $\mathfrak{L}$ maps onto $\mathfrak{A}$. Let $\mathfrak{A}_n$ be the $\mathbb{Q}$ vector space generated by all weight n words in $\mathfrak{A}$. We see easily that the Poincaré series of $\mathfrak{A}$ is

$$\sum_{n=0}^{\infty} (\dim_{\mathbb{Q}} \mathfrak{A}_n) t^n = \sum_{n=0}^{\infty} 2^n t^n = \frac{1}{1-2t}.$$

On the other hand, another equivalent definition of Lyndon word is the following: it is a primitive word (*i.e.*, not a true power of another word) that is minimal in its conjugacy class (*i.e.*, the set of words it produces by moving a right factor to the left). Now if a word w of weight n can be written as $\mathbf{w} = \mathbf{u}^m$ with $\mathbf{u}$ primitive of weight d then $n = dm$ and the number of conjugates of $\mathbf{w}$ is exactly d. Since every weight n word is produced in this way we get

$$2^n = \sum_{d|n} dN(d). \tag{3.10}$$

where $N(n)$ is the number of Lyndon words of weight n. Clearly the Poincaré series of $\mathfrak{L}$ is

$$f(t) := \prod_{n=1}^{\infty} (1-t^n)^{-N(n)}.$$

The coefficient of t^m in the logarithmic derivative $(\log f(t))'$ is

$$\sum_{n|(m+1)} nN(n) = 2^{m+1}$$

by Eq. (3.10). Hence $f(t) = 1/(1-2t)$ and therefore $\mathfrak{L} \cong \mathfrak{A}$. □

Theorem 3.3.12. *As a commutative algebra, $\mathfrak{A}_{ш}$ is the free polynomial algebra on the Lyndon words.*

Proof. The proof is similar to that of Thm. 3.3.11 and can be found in [469, Thm. 6.1] so we leave it to the interested readers. □

Remark 3.3.13. By Möbius inversion formula (see Exercise 3.8) the number of Lyndon words of weight n is given by

$$N(n) = \frac{1}{n}\sum_{d|n} 2^{n/d}\mu(d), \tag{3.11}$$

where the Möbius function

$$\mu(d) = \begin{cases} 1, & \text{if } d = 1; \\ 0, & \text{if } d \text{ has square factor} > 1; \\ (-1)^m, & \text{if } d \text{ is the product of } m \text{ distint primes.} \end{cases} \tag{3.12}$$

Equation (3.11) can be generalized easily to words built on a larger alphabet. One only need to replace 2 by the number of letters in the alphabet.

Theorem 3.3.14. *The evaluation maps $\zeta_{ш}$ and ζ_* can be extended to the algebra homomorphisms:*

$$\zeta_{ш} := \zeta_{ш}(-;T) : \mathfrak{A}^1_{ш} \longrightarrow \mathbb{R}[T], \quad \zeta_* := \zeta_*(-;T) : \mathfrak{A}^1_* \longrightarrow \mathbb{R}[T],$$

where $\mathfrak{A}^1_{ш} = (\mathfrak{A}^1, ш)$, $\mathfrak{A}^1_* = (\mathfrak{A}^1, *)$, and $\zeta_{ш}(\mathtt{x}_1;T) = \zeta_*(\mathtt{x}_1;T) = -T$. These provides two ways of regularizations of MZVs.

Remark 3.3.15. We will see that for any non-admissible word $\mathbf{w} \in \mathfrak{A}^1$, both $\zeta_{ш}(\mathbf{w})$ and $\zeta_*(\mathbf{w})$ are polynomials in T with all coefficients except for the leading one are given by convergent MZVs. To signify this sometimes we write $\zeta_{ш}(\mathbf{w};T)$ and $\zeta_*(\mathbf{w};T)$

Proof. By Thm. 3.3.11, as commutative algebras,

- $\mathfrak{A}^1_*$ is the free polynomial algebra on the Lyndon words (with one word $\mathtt{x}_0$ removed),
- $\mathfrak{A}^0_*$ is the free polynomial algebra on the Lyndon words (with two words $\mathtt{x}_0$ and $\mathtt{x}_1$ removed).

Thus

$$\mathfrak{A}^1_* \cong \mathfrak{A}^0_*[\mathtt{x}_1]. \tag{3.13}$$

So we can naturally extend $\zeta_* : \mathfrak{A}^0_* \longrightarrow \mathbb{R}$ to $\zeta_* : \mathfrak{A}^1_* \longrightarrow \mathbb{R}[T]$ by setting $\zeta_*(\mathtt{x}_1) = -T$. Similarly, by Thm. 3.3.12, as commutative algebras,

- $\mathfrak{A}^1_{ш}$ is the free polynomial algebra on the Lyndon words (with one word $\mathtt{x}_0$ removed),
- $\mathfrak{A}^0_{ш}$ is the free polynomial algebra on the Lyndon words (with two words $\mathtt{x}_0$ and $\mathtt{x}_1$ removed).

Thus

$$\mathfrak{A}^1_{ш} \cong \mathfrak{A}^0_{ш}[\mathtt{x}_1]. \tag{3.14}$$

So we can naturally extend $\zeta_{ш} : \mathfrak{A}^0_{ш} \longrightarrow \mathbb{R}$ to $\zeta_{ш} : \mathfrak{A}^1_{ш} \longrightarrow \mathbb{R}[T]$ by setting $\zeta_{ш}(\mathtt{x}_1) = -T$. □

Example 3.3.16. For example,

$$\begin{aligned}
\mathtt{x}_0^2\mathtt{x}_1\mathtt{x}_0^3\mathtt{x}_1 =& \mathtt{x}_0^3\mathtt{x}_1 * \mathtt{x}_0^2\mathtt{x}_1 - \mathtt{x}_0^3\mathtt{x}_1\mathtt{x}_0^2\mathtt{x}_1 + \mathtt{x}_0^6\mathtt{x}_1,\\
\mathtt{x}_0^2\mathtt{x}_1\mathtt{x}_0^2\mathtt{x}_1 =& \frac{1}{2}(\mathtt{x}_0^2\mathtt{x}_1 * \mathtt{x}_0^2\mathtt{x}_1 + \mathtt{x}_0^5\mathtt{x}_1),\\
\mathtt{x}_1^2\mathtt{x}_0 =& \frac{1}{2}\mathtt{x}_1 ш \mathtt{x}_1 ш \mathtt{x}_0 - \mathtt{x}_1 ш \mathtt{x}_0\mathtt{x}_1 + \mathtt{x}_0\mathtt{x}_1^2.
\end{aligned}$$

3.3.2 *Regularization in two ways*

Before proving the main theorem for regularized DBSFs we need a few preliminary results describing how to regularize the series definition and the iterated integral definition, and how to relate these two kinds of regularizations.

Lemma 3.3.17. *Define an* $\mathbb{R}$*-module* $\mathbb{R}$*-linear automorphism* ρ *of* $\mathbb{R}[T]$ *by*

$$\rho(\exp(Tu)) = \exp\left(\sum_{n=2}^{\infty} \frac{(-1)^n}{n}\zeta(n)u^n\right)\exp(Tu). \tag{3.15}$$

Let $P(T) \in \mathbb{R}[T]$ *and* $Q(T) = \rho(P(T))$. *Then*

$$\sum_{m=1}^{\infty} P(\log m + \gamma)t^{m-1} = \frac{1}{1-t}Q\left(\log\frac{1}{1-t}\right) + O\left(\log^{\deg P}\left(\frac{1}{1-t}\right)\right)$$

as $t \to 1^+$.

Proof. By linearity we only need to prove the identity for $P(T) = (T-\gamma)^l$ and $Q(T) = \rho((T-\gamma)^l)$. First, Weierstrass defined $\Gamma(x)$ by the following:

$$\frac{1}{\Gamma(u+1)} = \exp(\gamma u)\prod_{k=1}^{\infty}\left(1+\frac{u}{k}\right)\exp\left(-\frac{u}{k}\right).$$

Hence

$$-\log\Gamma(u+1) = \gamma u + \sum_{k=1}^{\infty}\left(\log\left(1+\frac{u}{k}\right) - \frac{u}{k}\right) = \gamma u + \sum_{k=1}^{\infty}\sum_{n=2}^{\infty}\frac{(-1)^{n-1}}{n}\frac{u^n}{k^n}. \tag{3.16}$$

Rearranging the sum we get

$$\Lambda(u) := \exp\left(\sum_{n=2}^{\infty} \frac{(-1)^n}{n}\zeta(n)u^n\right) = e^{\gamma u}\Gamma(1+u) \qquad \forall |u| < 1. \tag{3.17}$$

For a power series $g(u) = \sum_{n\geq 0} a_n u^n$ let $\mathrm{Coeff}_l(g) = a_l$. Then by linearity

$$\begin{aligned} Q(T) = \rho((T-\gamma)^l) =& l!\,\mathrm{Coeff}_l\,\rho(\exp((T-\gamma)u)) \\ =& l!\,\mathrm{Coeff}_l\left[\Lambda(u)\exp((T-\gamma)u)\right] \\ =& \frac{\mathrm{d}^l}{\mathrm{d}u^l}\left[\Lambda(u)\exp((T-\gamma)u)\right]_{u=0} \\ =& \frac{\mathrm{d}^l}{\mathrm{d}u^l}\left[\Gamma(1+u)\exp(Tu)\right]_{u=0} \end{aligned}$$

Thus

$$\begin{aligned} &\frac{1}{1-t}Q\left(\log\frac{1}{1-t}\right) = \frac{\mathrm{d}^l}{\mathrm{d}u^l}\left[\frac{\Gamma(1+u)}{(1-t)^{1+u}}\right]_{u=0} \\ =&\frac{\mathrm{d}^l}{\mathrm{d}u^l}\left[\sum_{m=1}^{\infty}\frac{\Gamma(1+u)(1+u)(2+u)\cdots(m-1+u)}{(m-1)!}t^{m-1}\right]_{u=0} \\ =&\frac{\mathrm{d}^l}{\mathrm{d}u^l}\left[\sum_{m=1}^{\infty}\frac{\Gamma(m+u)}{\Gamma(m)}t^{m-1}\right]_{u=0} \\ =&\sum_{m=1}^{\infty}\frac{\Gamma^{(l)}(m)}{\Gamma(m)}t^{m-1}. \end{aligned}$$

Using the following expression (from which one can derive the Stirling's formula)

$$\log\Gamma(x) = \left(x-\frac{1}{2}\right)\log x - x + C - \int_0^{\infty}\frac{t-[t]-1/2}{x+t}\,\mathrm{d}t$$

we get

$$\Gamma'(x) = \Gamma(x)\log x + \Gamma(x)O(x^{-1}) \quad \text{ as } x\to\infty.$$

By Leibnitz rule and complete induction on l we then have

$$\frac{\Gamma^{(l)}(m)}{\Gamma(m)} = \log^l m + O\left(\frac{\log^{l-1} m}{m}\right) \text{ as } m\to\infty,$$

for all $l \geq 0$. By Lemma 2.3.9 we get, as $t\to 1^+$,

$$\begin{aligned} \sum_{m=1}^{\infty}\frac{\Gamma^{(l)}(m)}{\Gamma(m)}t^{m-1} &= \sum_{m=1}^{\infty}(\log^l m)t^{m-1} + O\left(\log^l\frac{1}{1-t}\right) \\ &= \sum_{m=1}^{\infty}P(\log m+\gamma)t^{m-1} + O\left(\log^l\frac{1}{1-t}\right), \end{aligned}$$

as desired. □

For any composition $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ we set $\mathbf{w}(s_1, \dots, s_d) = \mathtt{x}_0^{s_1-1}\mathtt{x}_1 \cdots \mathtt{x}_0^{s_d-1}\mathtt{x}_1$ and

$$\zeta_*(\mathbf{s}; T) = (-1)^d \zeta_*(\mathbf{w}),$$
$$\zeta_{ш}(\mathbf{s}; T) = (-1)^d \zeta_{ш}(\mathbf{w}).$$

Remark 3.3.18. We will always use $\zeta_{ш}(\mathbf{s}; T)$ and $\zeta_*(\mathbf{s}; T)$ when we consider them as polynomials of T while we will write $\zeta_{ш}(\mathbf{s}) := \zeta_{ш}(\mathbf{s}; 0)$ and $\zeta_*(\mathbf{s}) := \zeta_*(\mathbf{s}; 0)$ for the constant term only.

For $M > 0$ we define the partial sum of MZVs

$$\zeta_M(\mathbf{s}) = \sum_{M > k_1 > \cdots > k_d > 0} k_1^{-s_1} \cdots k_d^{-s_d}.$$

Set $\mathtt{y}_{\mathbf{s}} = \mathtt{y}_{s_1} \cdots \mathtt{y}_{s_d}$ and $\mathtt{y}_{\mathbf{t}} = \mathtt{y}_{t_1} \cdots \mathtt{y}_{t_\ell}$. It is easy to see that

$$(-1)^{d+\ell} \zeta_M(\mathbf{s}) \zeta_M(\mathbf{t}) = \zeta_M(\mathtt{y}_{\mathbf{s}} * \mathtt{y}_{\mathbf{t}})$$

where ζ_M is the $\mathbb{Q}$-linear map $\mathfrak{A}^1_* \longrightarrow \mathbb{R}$ such that $\zeta_M(\mathtt{y}_{\mathbf{s}}) = (-1)^d \zeta_M(\mathbf{s})$.

Lemma 3.3.19. *Put $J = \mathrm{dp}(\mathbf{s})$. Then as $M \to \infty$ we have*

$$\zeta_M(\mathbf{s}) = \zeta_*(\mathbf{s}; \log M + \gamma) + O(M^{-1} \log^J M).$$

Moreover, if $\mathbf{s} = (\{1\}^\ell, \mathbf{s}')$ for admissible $\mathbf{s}'$ then the leading term of $\zeta_*(\mathbf{s}; T)$ is $(1/\ell!)\zeta(\mathbf{s}')T^\ell$, where $\zeta(\emptyset) = 1$.

Proof. We first prove that for all $d \geq 1$ and $j \leq d$ we have

$$A_{j,d}(M) = \sum_{k_1 > \cdots > k_j > M > k_{j-1} > \cdots > k_d > 0} \frac{1}{k_1^2 k_2 \cdots k_d} = O(M^{-1} \log^d M). \tag{3.18}$$

Indeed, if $d = 1$ we have $j = 1$ and

$$A_{1,1}(M) = \sum_{k > M} \frac{1}{k^2} < \frac{1}{M} \sum_{k > M} \frac{1}{k} = O(M^{-1} \log M).$$

Now by induction

$$\sum_{k_1 > \cdots > k_j > M > k_{j-1} > \cdots > k_d > 0} \frac{1}{k_1^2 k_2 \cdots k_d}$$
$$\ll \sum_{k_1 > \cdots > k_j > M} \frac{1}{k_1^2 k_2 \cdots k_j} \sum_{M > k_{j-1} > \cdots > k_{d-1} > 0} \frac{\log k_{d-1}}{k_{j+1} \cdots k_{d-1}} + O\left(\frac{\log^{d-1} M}{M}\right)$$
$$\ll \cdots$$
$$\ll (\log^{d-j} M) \sum_{k_1 > M} \frac{\log^{j-1} k_1}{k_1^2} + O\left(\frac{\log^{d-1} M}{M}\right)$$
$$\ll \frac{\log^{d-j} M}{M} \int_M^\infty \frac{\log^{j-1} x}{x} \, \mathrm{d}x + O\left(\frac{\log^{d-1} M}{M}\right) \ll \frac{\log^{d-1} M}{M}.$$

This proves Eq. (3.18).

We now proceed by induction on the number ℓ of leading 1's in $\mathbf{s}$. If $\ell = 0$ we have $s_1 > 1$.

$$\begin{aligned}\zeta(\mathbf{s}) - \zeta_M(\mathbf{s}) =& \sum_{j=1}^{d} \sum_{k_1>\cdots>k_j>M>k_{j-1}>\cdots>k_d>0} k_1^{-s_1}\cdots k_d^{-s_d} \\ \le& \sum_{j=1}^{d} \sum_{k_1>\cdots>k_j>M>k_{j-1}>\cdots>k_d>0} \frac{1}{k_1^2 k_2\cdots k_d} = O\Big(\frac{\log^{d-1} M}{M}\Big)\end{aligned}$$

by Eq. (3.18). This proves the case $\ell = 0$ with $\zeta_*(\mathbf{s}; T) = \zeta(\mathbf{s})$ a constant. Suppose the lemma is true when the number of leading 1's in $\mathbf{s}$ is less than $\ell \ge 1$. Now suppose $\mathbf{s} = (\{1\}^{\ell-1}, t_1, \dots, t_d)$ with $t_1 > 1$ for some $d \ge 0$ (when $d = 0$ no t_j's will appear). Then

$$-\zeta_M(\mathbf{x}_1) = \zeta_M(1) = \sum_{k=1}^{M} \frac{1}{k} = \log M + \gamma + O(M^{-1}) \quad \text{as } M \to \infty,$$

where γ is the Euler's constant. Using stuffle relation we have

$$\begin{aligned}\ell\zeta_M(1,\mathbf{s}) =& \zeta_M(1)\zeta_M(\mathbf{s}) - \sum_{j=1}^{d} \zeta_M(\{1\}^{\ell-1}, t_1, \dots, t_j, 1, t_{j+1}, \dots, t_d) \\ & - \sum_{j=1}^{\ell-1} \zeta_M(\{1\}^{j-1}, 2, \{1\}^{\ell-j-1}, t_1, \dots, t_d) \\ & - \sum_{j=1}^{d} \zeta_M(\{1\}^{\ell-1}, t_1, \dots, t_{j-1}, t_j + 1, t_{j+1}, \dots, t_d).\end{aligned}$$

By induction assumption

$$\begin{aligned}&\ell\zeta_M(1,\mathbf{s}) \\ =& (\log M + \gamma)\zeta_*(\mathbf{s}; \log M + \gamma) + O\Big(\frac{\log^{J+1} M}{M}\Big) \\ & - \sum_{j=1}^{d} \zeta_*(\{1\}^{\ell-1}, t_1, \dots, t_j, 1, t_{j+1}, \dots, t_d); \log M + \gamma) + O\Big(\frac{\log^{J+1} M}{M}\Big) \\ & - \sum_{j=1}^{\ell-1} \zeta_*(\{1\}^{j-1}, 2, \{1\}^{\ell-j-1}, t_1, \dots, t_d; \log M + \gamma) + O\Big(\frac{\log^{J} M}{M}\Big) \\ & - \sum_{j=1}^{d} \zeta_*(\{1\}^{\ell-1}, t_1, \dots, t_{j-1}, t_j + 1, t_{j+1}, \dots, t_d; \log M + \gamma) + O\Big(\frac{\log^{J} M}{M}\Big) \\ =& \ell\zeta_*(1, \mathbf{s}; \log M + \gamma) + O(M^{-1}\log^{J+1} M),\end{aligned}$$

since $\zeta_*(1;T)\zeta_*(\mathbf{s};T) = T\zeta_*(\mathbf{s};T)$ can be computed by the stuffle, which leads to the expression of $\zeta_*(1,\mathbf{s};T)$. In particular, the leading term of $\zeta_*(1,\mathbf{s};T)$ is equal to $\frac{T}{\ell}$ times the leading term of $\zeta_*(\mathbf{s};T)$. The lemma follows from the inductive assumption. □

Now we define the single-variable MPL by

$$\mathrm{L}_{\mathbf{s}}(z) := \int_0^z \left(\frac{\mathrm{d}x}{x}\right)^{s_1-1} \frac{\mathrm{d}x}{1-x} \cdots \left(\frac{\mathrm{d}x}{x}\right)^{s_d-1} \frac{\mathrm{d}x}{1-x} = \sum_{k_1>\cdots>k_d\geq 1} \frac{z^{k_1}}{k_1^{s_1}\cdots k_d^{s_d}}, \quad |z|<1. \tag{3.19}$$

Lemma 3.3.20. *Put $J = \mathrm{dp}(\mathbf{s})$. Then as $t \to 1^+$*

$$\mathrm{L}_{\mathbf{s}}(t) = \zeta_{ш}\left(\mathbf{s}; \log\frac{1}{1-t}\right) + O\left((1-t)\log^J\left(\frac{1}{1-t}\right)\right). \tag{3.20}$$

Proof. We prove the lemma by induction on the number ℓ of leading 1's in $\mathbf{s}$. First assume $\ell = 0$ and so $s_1 > 1$. Set paths $\alpha_t = [0,t]$ and $\beta_t = [t,1]$. By Lemma 2.1.2(iii) we have

$$\begin{aligned}
\mathrm{L}_{\mathbf{s}}(1) - \mathrm{L}_{\mathbf{s}}(t) &= \sum_{k_1>\cdots>k_d\geq 1} \frac{1-t^{k_1}}{k_1^{s_1}\cdots k_d^{s_d}} \\
&\leq \sum_{k_1>\cdots>k_d\geq 1} \frac{1-t^{k_1}}{k_1^2 k_1\cdots k_d} \\
&= \mathrm{L}_{(2,\{1\}^{d-1})}(1) - \mathrm{L}_{(2,\{1\}^{d-1})}(t) \\
&= \sum_{j=1}^{d} \int_t^1 \frac{\mathrm{d}t}{t}\left(\frac{\mathrm{d}t}{1-t}\right)^{j-1} \int_0^t \left(\frac{\mathrm{d}t}{1-t}\right)^{d-j}
\end{aligned}$$

by applying Lemma 2.1.2(ii) to $\mathrm{L}_{(2,\{1\}^{d-1})}(1) = \int_0^1 \frac{\mathrm{d}t}{t}\left(\frac{\mathrm{d}t}{1-t}\right)^{d-1}$. Now

$$\int_0^t \left(\frac{\mathrm{d}t}{1-t}\right)^{d-j} = \log^{d-j}\left(\frac{1}{1-t}\right)$$

and by the change of variables $t_j \to 1-t_j$ and Lemma 2.1.2(iv)

$$-\int_t^1 \frac{\mathrm{d}t}{t}\left(\frac{\mathrm{d}t}{t-1}\right)^{j-1} = \int_0^{1-t}\left(\frac{\mathrm{d}t}{t}\right)^{j-1}\frac{\mathrm{d}t}{1-t} = \mathrm{L}_j(1-t) = O(1-t).$$

Thus

$$\mathrm{L}_{\mathbf{s}}(t) = \zeta(\mathbf{s}) + O\left((1-t)\log^d\left(\frac{1}{1-t}\right)\right).$$

Since $\zeta_{ш}(\mathbf{s};T)=\zeta(\mathbf{s})$ this proves the lemma when $\ell=0$. The induction step can be furnished by using shuffle relation, similar to the proof of Lemma 3.3.19. We leave out the details. This finishes the proof of the lemma. □

3.3.3 *The main statement*

The following result provides the *regularized double shuffle relations* (regularized DBSFs).

Theorem 3.3.21. *For all* $\mathbf{w}\in\mathfrak{A}^1$ *one has*

$$(\zeta_{ш}-\rho\circ\zeta_*)(\mathbf{w})=0. \tag{3.21}$$

Proof. Let $J=\mathrm{dp}(\mathbf{s})$. Notice

$$\begin{aligned}\mathrm{L}_{\mathbf{s}}(t)&=\sum_{m=1}^{\infty}\sum_{m>k_2>\cdots>k_d\geq 1}\frac{t^m}{m^{s_1}k_2^{s_2}\cdots k_d^{s_d}}\\&=\sum_{m=1}^{\infty}(\zeta_{m+1}(\mathbf{s})-\zeta_m(\mathbf{s}))t^m\\&=(1-t)\sum_{m=1}^{\infty}\zeta_m(\mathbf{s})t^{m-1}\end{aligned}$$

since $\zeta_0(\mathbf{s})=0$. Thus

$$\begin{aligned}&\zeta_{ш}\left(\mathbf{s};\log\frac{1}{1-t}\right)\overset{\text{Lem. 3.3.20}}{=\!=}\mathrm{L}_{\mathbf{s}}(t)+O\left((1-t)\log^J\left(\frac{1}{1-t}\right)\right)\\&\overset{\text{Lem. 3.3.19}}{=\!=}(1-t)\sum_{m=1}^{\infty}\zeta_*(\mathbf{s};\log m+\gamma)t^{m-1}+O\left((1-t)\sum_{m=1}^{\infty}\frac{\log^J m}{m}t^{m-1}\right)\\&\overset{\text{Lem. 2.3.9}}{=\!=}(1-t)\sum_{m=1}^{\infty}\zeta_*(\mathbf{s};\log m+\gamma)t^{m-1}+O\left((1-t)\log^{J+1}\left(\frac{1}{1-t}\right)\right)\\&\overset{\text{Lem. 3.3.17}}{=\!=}Q\left(\log\frac{1}{1-t}\right)+O\left((1-t)\log^{J+1}\left(\frac{1}{1-t}\right)\right)\end{aligned}$$

with $Q(T)=\rho(\zeta_*(\mathbf{s};T))$ since $\deg_T\zeta_*(\mathbf{s};T)\leq J$. Thus $\zeta_{ш}(\mathbf{s};T)=\rho(\zeta_*(\mathbf{s};T))$. The theorem is now proved since $\mathbf{s}\in\mathbb{N}^d$ is arbitrary. □

Remark 3.3.22. One can compute ρ explicitly as follows. First expand $\Lambda(u)$ into a power series:

$$\Lambda(u)=\exp\left(\sum_{n=2}^{\infty}\frac{(-1)^n}{n}\zeta(n)u^n\right)=\sum_{k=0}^{\infty}c_ku^k,$$

where $c_0 = 1, c_1 = 0, c_2 = \zeta(2)/2, c_3 = -\zeta(3)/3, c_4 = 9\zeta(4)/16, \cdots$. Then

$$\rho(T^n) = \sum_{k=0}^{n} k! c_k \binom{n}{k} T^{n-k}, \quad \forall n \geq 0.$$

Corollary 3.3.23. *For all* $\mathbf{w}_0 \in \mathfrak{A}^0$ *and* $\mathbf{w}_1 \in \mathfrak{A}^1$ *one has*

$$\zeta_{ш}(\mathbf{w}_1 ш \mathbf{w}_0 - \mathbf{w}_1 * \mathbf{w}_0) = 0. \tag{3.22}$$

Proof. We can get the desired result multiplying $\zeta(\mathbf{w}_0)$ on both sides of $\zeta_{ш}(\mathbf{w}_1) = \rho(\zeta_*(\mathbf{w}_1))$ and using $\mathbb{R}$-linearity of ρ. □

Definition 3.3.24. The *regularized double shuffle relations* (regularized DBSFs) of MZVs are those $\mathbb{Q}$-linear relations obtained by setting $T = 0$ in Eq. (3.21) for all words $\mathbf{w} \in \mathfrak{A}^1$.

Conjecture 3.3.25. *The non-finite regularized DBSFs of MZVs are generated by the Hoffman relations* $\zeta(\mathtt{x}_1 ш \mathbf{w} - b * \mathbf{w})$ *for all admissible words* $\mathbf{w}$.

Note that if $\mathbf{w} \in \mathfrak{A}^1 \setminus \mathfrak{A}^0$ then the leading letter of $\mathbf{w}$ must be $\mathtt{x}_1$. The conjecture says that to produce all the non-finite regularized shuffle relations we only need to regularize divergent MZVs with only one leading $\mathtt{x}_1$.

Example 3.3.26. The simplest regularized DBSF is obtained by considering the word $bab \in \mathfrak{A}^1$. First, since $\zeta(2) = -\zeta_{ш}(\mathtt{x}_0\mathtt{x}_1)$ and $T = -\zeta_{ш}(\mathtt{x}_1)$

$$\zeta(2)T = \zeta_{ш}(\mathtt{x}_0\mathtt{x}_1 ш \mathtt{x}_1) = \zeta_{ш}(\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1) + 2\zeta_{ш}(\mathtt{x}_0\mathtt{x}_1^2) = \zeta_{ш}(\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1) + 2\zeta(2,1).$$

On the other hand,

$$\zeta(2)T = \zeta_*(\mathtt{x}_0\mathtt{x}_1)\zeta_*(\mathtt{x}_1) = \zeta_*(\mathtt{x}_0\mathtt{x}_1 * \mathtt{x}_1) = \zeta_*(\mathtt{x}_1\mathtt{x}_0\mathtt{x}_1) - \zeta_*(\mathtt{x}_0^2\mathtt{x}_1) + \zeta_*(\mathtt{x}_0\mathtt{x}_1^2).$$

Observe that $-\zeta_*(\mathtt{x}_0\mathtt{x}_0\mathtt{x}_1) = \zeta(3) = \zeta(2,1) = \zeta_*(\mathtt{x}_0\mathtt{x}_1\mathtt{x}_1)$ by (3.23). Applying Thm. 3.3.21 to $\mathbf{w} = \mathtt{x}_1\mathtt{x}_0\mathtt{x}_1$ we get

$$\zeta(2)T - 2\zeta(2,1) = \rho(\zeta(2)T - \zeta(3) - \zeta(2,1)) = \zeta(2)T - \zeta(3) - \zeta(2,1)$$

since $\rho(T) = T$ by Eq. (3.15). This leads to the famous identity first discovered by Euler:

$$\zeta(3) = \zeta(2,1). \tag{3.23}$$

3.4 Dimension conjectures over $\mathbb{Q}$

One of the guiding principles in the study of MZVs is the following widely believed conjecture.

Conjecture 3.4.1. *There are no nontrivial $\mathbb{Q}$-linear relations among MZVs of different weights.*

For MZVs of the same weight by Example 3.3.2 and Example 3.3.26 there may exist $\mathbb{Q}$-linear relations. Let MZ_w be the $\mathbb{Q}$-vector space spanned by all the MZVs of weight w. Zagier made the following famous conjecture.

Conjecture 3.4.2. *Define the sequence D_w ($w \geq 1$) by*

$$D_0 = 1, D_1 = 0, D_2 = 1, D_w = D_{w-2} + D_{w-3} \ \forall w \geq 3. \tag{3.24}$$

Then $\dim \mathsf{MZ}_w = D_w$ for all $w \geq 1$.

Remark 3.4.3. The numbers D_w are called Padovan numbers. According to OEIS [450] the Padovan sequence A000931 is

$$1, 0, 0, 1, 0, 1, 1, 1, 2, 2, 3, 4, 5, 7, 9, 12, 16, 21, \ldots .$$

Using the finite and regularized DBSFs one can show that each MZV of weight less than 8 is a polynomial of ordinary Riemann zeta values with $\mathbb{Q}$ coefficients. But already $\zeta(3,5)$ of weight 8 seems to fail this rule even if no proof has been found. In fact, this would be the case if the following conjecture by could be verified.

Conjecture 3.4.4. *Every linear relation among MZVs can be proved by the finite and regularized DBSFs.*

It is quite fruitful to reinterpret some of the above results in terms of a certain graded Hopf algebra (see Appendix A). Let $\mathcal{L}$ denote the free Lie algebra generated by one element e_{2n+1} in every odd degree $-2n-1$ for $n > 1$. Graded by the degree the underlying $\mathbb{Q}$-vector space is generated by,

$$e_3; \quad e_5; \quad e_7; \quad [e_3, e_5]; \quad e_9; \quad [e_3, e_7]; \quad e_{11}, [e_3, [e_5, e_3]]; \quad \cdots .$$

Let $\mathcal{UL}$ be its universal enveloping algebra, and let $\mathcal{UL}^\vee$ be its graded dual, which is a commutative Hopf algebra. One may think of $\mathcal{UL}^\vee$ as the set of all non-commutative words in letters f_{2n+1} in degree $2n+1$ dual to e_{2n+1}, equipped with the shuffle product $\sqcup\!\sqcup$. Finally, set

$$\mathcal{A} = \mathbb{Q}[f_2] \otimes \mathbb{Q}(\langle f_3, f_5, f_7, \ldots \rangle, \sqcup\!\sqcup)$$

where f_2 is a new generator of degree 2 which commutes with all the others. The generators in each weight w with $w \leq 9$ are

$$f_2; \quad f_3; \quad f_2^2; \quad f_5, f_2f_3; \quad f_2^3, f_3^2; \quad f_7, f_2f_5, f_2^2f_3;$$
$$f_2^4, f_2f_3^2, f_3 \sqcup\!\sqcup f_5, f_3f_5; \quad f_2^3f_3, f_3^3, f_2f_7, f_2^2f_5, f_9$$

which correspond to the following MZVs:

$$\zeta(2);\ \zeta(3);\ \zeta(2)^2;\ \zeta(5), \zeta(2)\zeta(3);\ \zeta(2)^3, \zeta(3,3);\ \zeta(7), \zeta(2)\zeta(5), \zeta(2)^2\zeta(3);$$
$$\zeta(2)^4, \zeta(2)\zeta(3,3), \zeta(3)\zeta(5), \zeta(3,5);\ \zeta(2)^3\zeta(3), \zeta(3)^3, \zeta(2)\zeta(7), \zeta(2)^2\zeta(5), \zeta(9)$$

The next conjecture is a more precise version of Zagier's Conjecture 3.4.2 .

Conjecture 3.4.5. *The algebra spanned by the MZVs over $\mathbb{Q}$ is isomorphic to $\mathcal{A}$.*

Actually, we only need the elements f_2 and f_3 to generate all the others by the rule

$$f_{2n+1} \to f_3f_2^{n-1} \qquad \forall n \geq 1.$$

Hence if we denote $\mathcal{F}(2,3)$ the free graded Lie algebra generated by two elements of degree -2 and -3, and $\mathcal{F}(2,3)^\vee$ its graded dual, then with the same notation we have a canonical isomorphism of graded vector spaces

$$\mathcal{A} \cong \mathcal{U}\mathcal{F}(2,3)^\vee.$$

Let d'_w be the dimension of the graded w piece of $\mathcal{A}$. Then it is clear that

$$d'_1 = 0, d'_2 = 1, d'_3 = 1, d'_w = d'_{w-2} + d'_{w-3}\ \forall w \geq 4.$$

One can even consider MZVs with both depth and the weight fixed. The next conjecture provides the hypothetical dimensions of $\mathbb{Q}$-spaces generated by such values. Let $\mathsf{MZ}_{w,\leq d}$ be the $\mathbb{Q}$-vector space spanned by all the MZVs of weight w and depth at most d. Let

$$\mathsf{MZ}_{w,+,\leq d} := \sum_{\substack{w_1,w_2\geq 1,\ w_1+w_2=w \\ 1\leq d_1,d_2<d,\ d_1+d_2\leq d}} \mathsf{MZ}_{w_1,\leq d_1} \cdot \mathsf{MZ}_{w_2,\leq d_2}$$

and

$$\mathsf{MZ}_{w,d} := \mathsf{MZ}_{w,\leq d} / \Big(\mathsf{MZ}_{w,\leq d-1} + \mathsf{MZ}_{w,+,\leq d} \Big)$$

be the $\mathbb{Q}$-vector space generated by MZVs of weight w and depth d that are not $\mathbb{Q}$-linear combination of lower depth MZVs and products of lower weight MZVs.

Conjecture 3.4.6. *We have* $\dim_{\mathbb{Q}} \mathsf{MZ}_{w,d} = D_{w,d}$ *where* $D(2,1) = 1$ *and*

$$\prod_{w\geq 3}\prod_{d\geq 1}(1-x^w y^d)^{D_{w,d}} = 1 - \frac{x^3 y}{1-x^2} + \frac{x^{12}y^2(1-y^2)}{(1-x^4)(1-x^6)}\,.$$

Remark 3.4.7. The conjectured term

$$\frac{x^3 y}{1-x^2} = \sum_{n\geq 2} x^{2n-1} y$$

is related to the Riemann zeta values $\zeta(2n)$ for $n \geq 2$ since all these values can be generated by $\zeta(2)$. For the other term, we have the expansion

$$\frac{x^{12}}{(1-x^4)(1-x^6)} = \sum_{n\geq 12, 2|n} \mathrm{sp}(n) x^n$$

where $\mathrm{sp}(n) = \lfloor (n-4)/4 \rfloor - \lfloor (n-2)/6 \rfloor$ is the dimension of the space of weight n cusp forms on the full modular group. So MZVs are expected to be intimately related to the modular forms. See Sec. 5.8.

Similar to the situation for $\dim_{\mathbb{Q}} \mathsf{MZ}_w$, it is unlikely to prove $\dim_{\mathbb{Q}} \mathsf{MZ}_{w,d} \geq D_{w,d}$ in the near future. However, Conjecture 3.4.6 implies that $D_{w,d} = 0$ unless $w \geq 3d$ and w is congruent to d modulo 2. It also implies the following conjectured formulas for depth $d \leq 3$:

$$D_{w,1} = 1, \quad D_{w,2} = \left[\frac{w-2}{6}\right], \quad D_{w,3} = \left[\frac{(w-3)^2-1}{48}\right] \tag{3.25}$$

for all w having the same parity as the depth. Notice that according to the following parity principle (see Cor. 3.5.7) $D_{w,2} = 0$ if w is odd and $D_{w,3} = 0$ if w is even.

3.5 Linearized double shuffle and Parity Principle

In this section we will study the DBSFs modulo elements of lower depths and products. Recall that $\mathsf{MZ}_{w,d}$ is the $\mathbb{Q}$-vector space generated by MZVs of weight w and depth d that cannot be expressed as a $\mathbb{Q}$-linear combination of lower depth MZVs and products of lower weight MZVs. Set

$$\mathcal{M} = \sum_{w,n\geq 1} \mathsf{MZ}_{w,n}.$$

Of course, this is a direct sum conjecturally. Write $V_n = \mathbb{Q}[x_1, \ldots, x_n]$ and denote by $V_n(d)$ its homogeneous polynomials of degree d. For $f(x_1, \ldots, x_n) \in V_n$ define the action of $M \in \mathrm{GL}_n(\mathbb{Z})$ on $f \in V_n$ by

$$(f|M)(x_1, \ldots, x_n) = f((x_1, \ldots, x_n) \cdot M^{-1}).$$

Notice that $f|(M_1M_2) = (f|M_1)|M_2$ for all $M_1, M_2 \in \mathrm{GL}_n(\mathbb{Z})$. We can then extend this action to $\mathbb{Z}[\mathrm{GL}_n(\mathbb{Z})]$ by linearity.

Let $\mathfrak{S}_n$ be the symmetry group of order n regarded as a subgroup of $\mathrm{GL}_n(\mathbb{Z})$ so that the action of $\sigma \in \mathfrak{S}_n$ on $f \in V_n$ is given by

$$(f|\sigma)(x_1, \dots, x_n) = f(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)}).$$

Explicitly, the matrix M_σ for $\sigma \in \mathfrak{S}_n$ can be obtained by applying σ to the columns of the $n \times n$ identity matrix, namely, its (i,j)th entry is $\delta_{i,\sigma(i)}$. In what follows we abuse notation and just write σ for M_σ by identifying σ with its image in $\mathrm{GL}_n(\mathbb{Z})$.

For any $j \leq n-1$ define the shuffle element

$$\mathrm{sh}_j := \mathrm{sh}_j^{(n)} = \sum_{\sigma \in \mathrm{Sh}_j^{(n)}} \sigma \tag{3.26}$$

where $Sh_j^{(n)} = \{\sigma \in \mathfrak{S}_n | \sigma(1) < \cdots < \sigma(j), \sigma(j+1) < \cdots < \sigma(w)\}$. Further, for all $f \in V_n$ we set $f^\sharp = f|P$ where

$$P = \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}, \qquad P^{-1} = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix}. \tag{3.27}$$

Clearly

$$f^\sharp(x_1, \dots, x_n) = f(x_1 + \cdots + x_n, x_2 + \cdots + x_n, \dots, x_{n-1} + x_n, x_n).$$

Definition 3.5.1. The *double shuffle space* DSh_n of V_n is defined by

$$\mathsf{DSh}_n = \left\{ f \in V_n \,\middle|\, f|\,\mathrm{sh}_j^{(n)} = f^\sharp|\,\mathrm{sh}_j^{(n)} = 0 \text{ for } j = 1, \dots, n-1 \right\}$$

for all $n \geq 2$. In particular $\mathsf{DSh}_1 = \mathbb{Q}[x_1]$. Denote by $\mathsf{DSh}_n(d)$ the degree d part of DSh_n.

Setting $S^\sharp = PSP^{-1}$ for all $S \in \mathbb{Z}[\mathrm{GL}_n(\mathbb{Z})]$ we see that $f^\sharp|\,\mathrm{sh}_j = 0$ is equivalent to $f|\,\mathrm{sh}_j^\sharp = 0$. Our goal in this section is to study DSh_n by considering the following two generating functions

$$F_{ш}^{(n)}(x_1, \dots, x_n) = \sum_{s_1, \dots, s_n \geq 1} \zeta_{ш}(s_1, \dots, s_n) x_1^{s_1 - 1} \dots x_n^{s_n - 1},$$

$$F_*^{(n)}(x_1, \dots, x_n) = \sum_{s_1, \dots, s_n \geq 1} \zeta_*(s_1, \dots, s_n) x_1^{s_1 - 1} \dots x_n^{s_n - 1}.$$

We regard these as elements in $\widetilde{\mathcal{M}}[x_1, \dots, x_n] = \widetilde{\mathcal{M}} \otimes_{\mathbb{Q}} V_n$ where $\widetilde{\mathcal{M}} = \mathcal{M} + \mathbb{Q}$ is a $\mathbb{Q}$-vector space regarded as a $\mathbb{Q}$-algebra (again, conjecturally $\widetilde{\mathcal{M}} = \mathcal{M} \oplus \mathbb{Q}$).

Lemma 3.5.2. *If $1 \le j < n$ then $(F_{ш}^{(n)})^{\sharp}|\,\mathrm{sh}_j = 0$ in $\mathcal{M} \otimes_{\mathbb{Q}} V_n$.*

Proof. For $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{N}^n$ with $|\mathbf{s}| = w$ and $\varepsilon < 1$, we define

$$\zeta_\varepsilon(\mathbf{s}) = \int\limits_{1-\varepsilon > t_1 > \cdots > t_w > 0} \cdots \int \omega(t_1) \cdots \omega(t_w)$$

where the integrand is the same as in the ordinary MZV defined by Eq. (3.1). Now applying the formula

$$\int\limits_{a > x_1 > \cdots > x_r > b} \cdots \int \frac{\mathrm{d}x_1}{x_1} \cdots \frac{\mathrm{d}x_r}{x_r} = \frac{1}{r!} \left(\log \frac{a}{b}\right)^r$$

to $(x_0, b, r) = (1-\varepsilon, u_1, s_1 - 1), (u_1, u_2, s_2 - 1), \dots, (u_{n-1}, u_n, s_n - 1)$, where $u_1 = t_{s_1}$, $u_2 = t_{s_1+s_2}, \dots,$ $u_n = t_w$, we find

$$\zeta_\varepsilon(\mathbf{s}) = \frac{1}{(s_1-1)! \cdots (s_n-1)!} \int\limits_{1-\varepsilon > u_1 > \cdots > u_n > 0} \cdots \int \left(\log \frac{1-\varepsilon}{u_1}\right)^{s_1-1} \frac{\mathrm{d}u_1}{1-u_1} \left(\log \frac{u_1}{u_2}\right)^{s_2-1} \frac{\mathrm{d}u_2}{1-u_2} \cdots \left(\log \frac{u_{n-1}}{u_n}\right)^{s_n-1} \frac{\mathrm{d}u_n}{1-u_n}.$$

Easy computation using $\sum_{i \ge 0} t^n/n! = e^t$ yields

$$F_{ш}^{(n),\varepsilon}(x_1, \dots, x_n) := \sum_{s_1, \dots, s_n \ge 1} \zeta_\varepsilon(s_1, \dots, s_n) x_1^{s_1-1} \dots x_n^{s_n-1}$$

$$= (1-\varepsilon)^{x_1} \int\limits_{1-\varepsilon > u_1 > \cdots > u_n > 0} \cdots \int \frac{u_1^{-x_1+x_2} \mathrm{d}u_1}{1-u_1} \cdots \frac{u_{n-1}^{-x_{n-1}+x_n} \mathrm{d}u_{n-1}}{1-u_{n-1}} \frac{u_n^{-x_n} \mathrm{d}u_n}{1-u_n}.$$

Thus

$$(F_{ш}^{(n),\varepsilon})^{\sharp}(x_1, \dots, x_n) = F_{ш}^{(n),\varepsilon}(x_1 + \cdots + x_n, \dots, x_{n-1} + x_n, x_n)$$

$$= (1-\varepsilon)^{x_1 + \cdots + x_n} \int\limits_{1-\varepsilon > u_1 > \cdots > u_n > 0} \cdots \int \frac{u_1^{-x_1} \mathrm{d}u_1}{1-u_1} \cdots \frac{u_{n-1}^{-x_{n-1}} \mathrm{d}u_{n-1}}{1-u_{n-1}} \frac{u_n^{-x_n} \mathrm{d}u_n}{1-u_n}.$$

Hence $(F_{ш}^{(n),\varepsilon})^{\sharp}$ satisfies the shuffle relation

$$(F_{ш}^{(j),\varepsilon})^{\sharp}(x_1, \dots, x_j)(F_{ш}^{(n-j),\varepsilon})^{\sharp}(x_{j+1}, \dots, x_n) = \big((F_{ш}^{(n),\varepsilon})^{\sharp}|\,\mathrm{sh}_j\big)(x_1, \dots, x_n)$$

for any $\varepsilon < 1$ and $1 \le j < n$. The lemma now follows from the definition of $\zeta_{ш}(\mathbf{s}; T)$. □

Theorem 3.5.3. *For all $n \geq 1$, we have*

$$F_{ш}^{(n)}(x_1, \dots, x_n) = F_{*}^{(n)}(x_1, \dots, x_n) \tag{3.28}$$

in $\mathcal{M} \otimes_{\mathbb{Q}} V_n$. Moreover, both belong to the subspace $\sum_w \mathsf{MZ}_{w,\leq n} \otimes_{\mathbb{Q}} \mathsf{DSh}_n$ modulo $\sum_w \mathsf{MZ}_{w,+,\leq n}$.

Proof. First we prove Eq. (3.28). If $n = 1$ we see by the definition that $\zeta_{ш}(s) = \zeta_*(s) = (1-\delta_{1,s})\zeta(s)$ for all positive integer s, where $\delta_{1,s}$ is the Kronecker symbol. Thus the two generating functions $F_{ш}^{(1)}(x_1)$ and $F_{*}^{(1)}(x_1)$ are exactly the same. When $n \geq 2$ by Thm. 3.3.21, we see that each coefficient of $F_{ш}^{(n)}(x_1, \dots, x_n) - F_{*}^{(n)}(x_1, \dots, x_n)$ comes from the products some of MZV factors with the coefficients of $\rho(T^i)$. Now the coefficients of $\rho(T^i)$ are all $\mathbb{Q}$-linear combinations of at most i Riemann zeta values (see Eq. (3.15)) which can be written as $\mathbb{Q}$-linear combinations of MZVs of depth at most i. Then MZV factors can be easily see to have depths at most $n - i$ (for each T we use the depth of the corresponding MZVs must be decreased by 1).

To prove the second sentence in the theorem, we have

$$\begin{aligned}
0 &= F_{*}^{(j)}(x_1, \dots, x_j) F_{*}^{(n-j)}(x_{j+1}, \dots, x_n) \\
&= \sum_{s_1, \dots, s_n \geq 1} \zeta_*(s_1, \dots, s_j)\zeta_*(s_{j+1}, \dots, s_n) x_1^{s_1-1} \dots x_n^{s_n-1} \\
&= \sum_{s_1, \dots, s_n \geq 1} \sum_{\sigma \in \mathrm{Sh}_j^{(n)}} \zeta_*(s_{\sigma^{-1}(1)}, \dots, s_{\sigma^{-1}(n)}) x_1^{s_1-1} \dots x_n^{s_n-1} \\
&= (F_{*}^{(n)} | \operatorname{sh}_j)(x_1, \dots, x_n),
\end{aligned}$$

since the "stuffing" part of the stuffle product always produces MZVs of lower depth. By Eq. (3.28) and Lemma 3.5.2 the theorem follows immediately. □

The following bound on the dimension of $\mathsf{MZ}_{w,d}$ (see Sec. 3.4) is an easy consequence of Thm. 3.5.3. The proof is left as an exercise (see Exercise 3.12).

Corollary 3.5.4. *For all $w > d > 0$, we have*

$$D_{w,d} \leq \dim_{\mathbb{Q}} \mathsf{DSh}_d(w-d).$$

Our next goal is to prove the following parity principle Cor. 3.5.7. The idea to prove this is to define a suitable space $\mathsf{ShC}_n(d)$ containing the double shuffle space $\mathsf{DSh}_n(d)$ and satisfying $\mathsf{ShC}_n(d) = \{0\}$ when d is odd and

$n \geq 1$. In order to do so we need to define an action of $\mathfrak{S}_{n+1}$ on V_n (extending that of $\mathfrak{S}_n$) by

$$(F|\sigma)(x_1,\dots,x_n) := F(x_{\sigma^{-1}(1)},\dots,x_{\sigma^{-1}(n)}), \quad \text{for all } \sigma \in \mathfrak{S}_{n+1},$$

where $x_{n+1} = -x_1 - x_2 - \cdots - x_n$. Let

$$C_{n+1} = \begin{pmatrix} 1 & \cdots & n-1 & n & n+1 \\ 2 & \cdots & n & n+1 & 1 \end{pmatrix} \in \mathfrak{S}_{n+1} \tag{3.29}$$

Then $(f|C_{n+1})(x_1,\dots,x_n) = f(x_{n+1},x_1,\dots,x_{n-1})$ so that the matrix form of C_{n+1} in $\mathrm{GL}_n(\mathbb{Z})$ is

$$C_{n+1} = \begin{pmatrix} 0 & \cdots & 0 & -1 \\ 1 & & & -1 \\ & \ddots & & \vdots \\ & & 1 & -1 \end{pmatrix}, \quad \text{and} \quad C_{n+1}^{-1} = \begin{pmatrix} -1 & 1 & & \\ -1 & 0 & \ddots & \\ \vdots & \vdots & \ddots & 1 \\ -1 & 0 & \cdots & 0 \end{pmatrix}.$$

Let $\nu(x_j) = -x_j$ for all $j \leq n$ then its matrix form in $\mathrm{GL}_n(\mathbb{Z})$ is $-I_n$ where I_n is the $n \times n$ identity matrix. Clearly $\nu(x_{n+1}) = -x_{n+1}$.

Lemma 3.5.5. *For $0 \leq j \leq n$, define the involutions*

$$T_j = \begin{pmatrix} 1 & \cdots & j & j+1 & \cdots & n \\ 1 & \cdots & j & n & \cdots & j+1 \end{pmatrix} \in \mathfrak{S}_n. \tag{3.30}$$

Then

$$\sum_{j=1}^{n-1} (-1)^{j-1} \,\mathrm{sh}_j\, T_j = T_0 + (-1)^n.$$

Proof. For $1 \leq j \leq n$, we define

$$R_j = \sum_{\sigma \in \mathfrak{S}_n, \sigma(1) < \cdots < \sigma(j) > \sigma(j+1) > \cdots > \sigma(n)} \sigma$$

Then $\mathrm{sh}_0\, T_0 = R_1$, $\mathrm{sh}_n\, T_n = R_n$, and for $1 \leq j \leq n-1$, we have

$$\mathrm{sh}_j\, T_j = \sum_{\sigma \in \mathfrak{S}_n, \sigma(1) < \cdots < \sigma(j), \sigma(j+1) > \cdots > \sigma(n)} \sigma = R_j + R_{j+1},$$

since either $\sigma(j) > \sigma(j+1)$ or $\sigma(j) < \sigma(j+1)$. The lemma follows immediately. □

Theorem 3.5.6. *For $n, d \geq 1$, we define*

$$\mathsf{ShC}_n(d) := \Big\{ f \in V_n(d) \,\Big|\, f^\sharp|\,\mathrm{sh}_j = 0 \ (1 \leq j < n), f^\sharp|C_{n+1} = \nu(f^\sharp) \Big\}. \tag{3.31}$$

Then

(i) $\mathsf{DSh}_n(d) \subset \mathsf{ShC}_n(d)$;
(ii) $\mathsf{ShC}_n(d) = 0$ *if* d *is odd.*

Proof. (i). By Lemma 3.5.5, for all $f \in \mathsf{DSh}_n(d)$ we have

$$f|T_0 = f|T_0^\sharp = (-1)^{n-1} f. \tag{3.32}$$

Define the involution

$$T' = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n & n+1 \\ n-1 & n-2 & \cdots & 1 & n+1 & n \end{pmatrix} \in \mathfrak{S}_{n+1}. \tag{3.33}$$

Then $(f|T')(x_1, \ldots, x_n) = f(x_{n-1}, \ldots, x_1, x_{n+1})$ so that the matrix form of T' in $\mathrm{GL}_n(\mathbb{Z})$ is

$$T' = T'^{-1} = \begin{pmatrix} & & 1 & -1 \\ & \iddots & & \vdots \\ 1 & & & -1 \\ 0 & \cdots & 0 & -1 \end{pmatrix}.$$

It is not hard to use the matrix forms to show that (see also Excercise 3.13)

$$\text{(A). } T_0 T' = C_{n+1}, \quad \text{and} \quad \text{(B). } PT' = \nu T_0 P. \tag{3.34}$$

Thus

$$f^\sharp|C_{n+1} \overset{(3.34)(\mathrm{A})}{=\!=} f^\sharp|T_0T' \overset{\text{defn}}{=\!=} (f|P)|T_0T' = (f|PT_0)|T' \overset{\text{defn}}{=\!=} (f|T_0^\sharp P)|T'$$
$$\overset{(3.32)}{=\!=} (-1)^{n-1}(f|P)|T' \overset{(3.34)(\mathrm{B})}{=\!=} (-1)^{n-1}\nu(f|T_0P) \overset{(3.32)}{=\!=} \nu(f|P) = \nu(f^\sharp).$$

So $f \in \mathsf{ShC}_n(d)$ and therefore (i) is proved.

(ii). Let $\tau_{1,n+1}$ be the transposition $1 \leftrightarrow n+1$. It is straight-forward to check (see Exercise 3.14) that in $\mathbb{Z}[\mathfrak{S}_{n+1}]$

$$1 + \mathrm{sh}_1^{(n)} C_{n+1} = C_{n+1}(1 + \mathrm{sh}_1^{(n)} \tau_{1,n+1}). \tag{3.35}$$

Now suppose $f \in \mathsf{ShC}_n(d)$ and let $F = f^\sharp$. Then $F|\,\mathrm{sh}_1 = 0$ and $F|C_{n+1} = \nu F$ by the definition. Thus by Eq. (3.35)

$$F = F|C_{n+1}(1 + \mathrm{sh}_1^{(n)} \tau_{1,n+1}) = \nu\big(F|(1 + \mathrm{sh}_1^{(n)} \tau_{1,n+1})\big) = \nu F = -F$$

if the degree d is odd. Hence $F = 0$ and thus $f = F|P^{-1} = 0$. This completes the proof of the theorem. □

Corollary 3.5.7. (Parity Principle) *If the weight* w *and depth* n *of* $\zeta(\mathbf{s})$ *have different parities then* $\zeta(\mathbf{s})$ *is a* $\mathbb{Q}$*-linear combination of lower depth MZVs and products of lower weight MZVs whose sum of depths is bounded by that of* $\mathbf{s}$.

Proof. If $|\mathbf{s}|$ and n has different parity then the degree of $x_1^{s_1-1}\dots x_n^{s_n-1}$ is odd. So the corollary follows from part (ii) of the above theorem. □

The next result is another kind of Parity Principle.

Corollary 3.5.8. *If the $w = |\mathbf{s}|$ and n have the same parity then both $\zeta_*(\mathbf{s})+(-1)^w\zeta_*(\tilde{\mathbf{s}})$ and $\zeta_{ш}(\mathbf{s})+(-1)^w\zeta_s a(\tilde{\mathbf{s}})$ $\mathbb{Q}$-linear combinations of lower depth MZVs and products of lower weight MZVs.*

Proof. By Eq. (3.32) we see that

$$\begin{aligned}&\sum_{\mathbf{s}\in\mathbb{N}^n}\zeta_*(\overleftarrow{\mathbf{s}})\mathbf{x}^{|\mathbf{s}|-\mathbf{1}}\\ =&\sum_{\mathbf{s}\in\mathbb{N}^n}\zeta_*(\mathbf{s})\mathbf{x}^{|\overleftarrow{\mathbf{s}}|-\mathbf{1}}\\ =&(F_*^{(n)}|T_0)(\mathbf{x}) = (-1)^{n-1}F_*^{(n)}(\mathbf{x}).\end{aligned}$$

So we get

$$\zeta(s_1,\dots,s_n)+(-1)^w\zeta(\tilde{\mathbf{s}})\equiv\zeta(s_1,\dots,s_n)+(-1)^{w+n-1}\zeta(\mathbf{s})\equiv 0$$

modulo $\mathsf{MZ}_{w,+,\le n}$. Similar argument is valid for the ш-version, too. This completes the proof of the corollary. □

3.6 Historical notes

Even though some special cases of the theory of multiple zeta values were discovered by Goldbach and Euler, their formal definition and study was initiated only recently by Hoffman [275] and Zagier [577] independently. Incidentally, in early 1980s, while developing a theory of resurgence in complex analysis Ecalle studied some quite general mathematical objects called "Moulds" (functions with variable number of variables) in which MZVs are one of the examples [174, p. 429]. He even mentioned their iterated integral representation [174, p. 431, Remark 4] without explicit producing it. For an introduction to this theory, see Schneps' exposition [486].

In recent years, the study of MZVs has opened up quite a lot fascinating connections to other branches of mathematics such as knot theory, Mahler measures and the theory of motives (see [87,104,106,156,355]). MZVs have also emerged unexpectedly in high energy physics such as Feyman diagrams, amplitude, mirror symmetry and quantum field theory (see Chap. 15 and Refs. [87, 93, 109, 278, 482]).

One of the most important algebraic approaches to the study of MZVs is through the connection between the two kinds of products: shuffle and

stuffle, namely, the regularized DBSFs. The name "stuffle" first appeared in [69, 70]. It is also called quasi-shuffle in [277], harmonic shuffle in [526] or sticky shuffle in [293]. In fact, it is a special case of the algebra structure called mixable shuffle treated in [257]. In a series of papers [276–278], Hoffman first realized that one may equip MZVs with two different algebra structures reflecting the two different ways to multiply MZVs. This has become one of the cornerstones in various studies of MZVs and their generalizations.

It turns out the MZV relations obtained by the shuffle product can often be proved using other methods. In fact, Euler proved his decomposition formula, Thm. 3.1.1, using partial fractions. This is important, because in some generalizations of MZVs the integral representation is not easily available, but the partial fraction technique can still be applied. For example, Thm. 3.1.1 were extended to q-analogs by the author [594] by using Jackson's q-iterated integrals in a very complicated manner. But by using partial fractions, Bradley [83] discovered a much cleaner form with a shorter proof. Another q-analog of Eulers decomposition formula can be found in [115]. This formula is later generalized to the product of any two MZVs by the algebraic method of double shuffle product by Guo and Xie [256]. With different approaches, it is also generalized to the cases involving MZVs of the form $\zeta(m, \{1\}^n)$ in [183, 252].

Lemma-Definition 3.3.5 is a special case of Chen-Fox-Lyndon Theorem since it can be derived from [129], but the first explicit statement for the algebra of words representing MZVs did not appear until the appearance of [490], although Lemma 3.3.7 and many other results in Sec. 3.3 can be found in [276, 469]. For Remark 3.3.13 see [391, (1.3.7)].

Thm. 3.3.21 giving regularized DBSFs was stated by Ihara, Kaneko and Zagier in [298] parts of which had circulated as preprints long before its formal publication. Incidentally, regularized DBSFs are called extended DBSFs in [298]. In the same paper they also considered derivation relations which we will discuss in Chap. 5. In [322] Kaneko *et al.* further conjectured that a certain concrete subset of the RDS can generate all regularized DBSFs which was verified by Combariza [139] with some numerical data. Notice that the regularized DBSFs themselves are not explicit formulas of MZVs, neither are the finite ones. In [256] Guo and Xie described an idea to write down these explicitly and carried this out for products of two double zeta values.

The folklore Conjecture 3.4.1 is a consequence of some deep conjectures in the theory of algebraic geometry and motives (see [233, Conjecture B]).

Zagier further made Conjecture 3.4.4 while Conjecture 3.3.25 is a much stronger form of it. Clearly, Conjecture 3.4.5 also implies the next conjecture.

Conjecture 3.6.1. (Hoffman's Basis Conjecture) *The $\mathbb{Q}$-space generated by MZVs has a basis consisting of MZVs whose arguments are equal to either* 2 *or* 3 *(such MZVs are called Hoffman elements).*

A deep result of Brown [104] implies that every MZV is a $\mathbb{Q}$-linear combination of Hoffman elements.

Prop. 3.2.6 is due to Borwein, Bradley and Broadhurst (see [69, Prop. 1]). The idea to use relations between the words in Hoffman algebra is one of the most powerful tools to study the $\mathbb{Q}$-linear relations among MZVs.

Zagier's Dimension Conjecture 3.4.2 has been one of the driving forces in the study of MZVs. Its motivic version, Conjecture 3.4.5, was stated in [239]. Since Conjecture 3.4.2 is an equality one may approach it by proving two inequalities. Goncharov and Terasoma (see [239, 524]) independently proved that $\dim_{\mathbb{Q}} \mathsf{MZ}_w \leq D_w$. This inequality also follows from Brown's recent result mentioned above, which confirms only half of Hoffman's Basis Conjecture.

To prove the other half of Conjecture 3.4.2 one now needs to show $\dim_{\mathbb{Q}} \mathsf{MZ}_w \geq D_w$ which can be deduced from some notoriously difficult conjectures in transcendental number theory concerning MZVs, which is also related to Grothendieck's period conjecture (see [156, § 5.27] for details). For example, we know π is irrational by Lampert's work in [358] of 1768 (and transcendental by Lindemann [387] more than a century later). Almost another century passed before Apéry [16] proved $\zeta(3)$ is irrational in 1978. More than twenty years later, Ball and Rivoal [32, 467] showed that infinitely many values of $\zeta(2n+1)$ are irrational. More recently Zudilin [619] and others made some remarkable refinements in this direction. The progress seems to be quite slow, and in fact, until today we don't even know if $\zeta(5)$ is irrational or not.

To study the fine structure of MZVs, Broadhurst and his collaborators computed extensively with the aid of computers and observed a lot of subtle properties. In [87] he gave the precise Conjecture 3.4.6 for the number of irreducible MZVs of given weight and depth just by observing the patterns of the known dimensions. This formula has since been explained partly by a hidden relation between MZVs and modular forms (see [220]). Notice that each of the three conjectural formulas in Eq. (3.25) is known to give

an upper bound for the corresponding true dimension (see [576] for $d = 2$ and [236] for $d = 3$ or [300] by a different method). The ideas from [300] may be useful to treat the $d = 4$ case. In the general case, Carr, Gangl and Schneps have found some convincing evidence in [113].

Tsumura's original proof of the Parity Principle contained in [531] (Cor. 3.5.7) uses heavy analysis. Another different proof was later given by Ihara, Kaneko and Zagier (see [298, Thm. 7]) which we have modified and adopted in Sec. 3.5. Note that the case $d = 2$ was proved by Tornheim [529] and the case $d = 3$ was proved by Borwein and Girgensohn [74]. Similar parity result for other kind of multiple zeta functions should exist. For example, Okamoto [440] showed such a result for the Witten zeta function $\zeta_2(\mathbf{s}; G_2)$ for Lie algebra G_2 in some extended sense.

For general surveys of the results on MZVs up to the beginning of twenty-first century, the interested reader may peruse Refs. [114, 543, 620] by Cartier, Waldschmidt and Zudilin, respectively.

Exercises

3.1. Let $T_{p+q,j}$ be defined as in 3.2.5. Show that the factor x_1^2 appears exactly j times.

3.2. Prove Prop. 3.2.9 by induction.

3.3. Complete the proof of Lemma 3.3.7.

3.4. Complete the proof of Thm. 3.3.12.

3.5. Find the finite DBSF by computing $\zeta(2)\zeta(3)$ in two different ways.

3.6. Using Definition 3.3.3 how many Lyndon words of weight 6 can you find in $\mathbb{Q}\langle \mathtt{x}, \mathtt{y}\rangle$?

3.7. (1). Write down the Lyndon factorization of $\mathtt{x}_1^2\mathtt{x}_0^2\mathtt{x}_1$ and $\mathtt{x}_1\mathtt{x}_0^3\mathtt{x}_1^3\mathtt{x}_0^5\mathtt{x}_1^2\mathtt{x}_0^2\mathtt{x}_1\mathtt{x}_0^5\mathtt{x}_1\mathtt{x}_0^3$.

(2). Express the words in (1) as $\mathbb{Q}$-linear combinations of $*$-product of Lyndon words.

(3). Express the words in (1) $\mathbf{w}$ as $\mathbb{Q}$-linear combinations of $\sqcup\!\sqcup$-product of Lyndon words.

3.8. Let f and g be two real functions defined on $\mathbb{N}$. If $g(n) = \sum_{d|n} f(d)$ for all $n \in \mathbb{N}$, prove that $f(n) = \sum_{d|n} f(d)\mu(n/d)$ for all $n \in \mathbb{N}$ where μ is the Möbius function defined by Eq. (3.12). This is called the Möbius inversion formula.

3.9. Find the regularized DBSF by computing $\zeta(3)T$ in two different ways using Thm. 3.3.21 and Remark 3.3.22.

3.10. By hand computation verify Zagier's Dimension Conjectuer 3.4.2 provides the upper bound for $\dim_{\mathbb{Q}} \mathsf{MZ}_w$ for $w \leq 5$.

3.11. This exercise is for Maple users. First, download the software mzv16.m (or mzv12.txt if your computer has moderate RAM) from Petitot's website http://www.lifl.fr/ petitot/recherche/MZV/. Then

- Verify that Broadhurst Conjecture 3.4.6 provides the upper bound for $\dim_{\mathbb{Q}} \mathsf{MZ}_{w,d}$ for $d < w \leq 12$.
- Verify the Parity Principle (Cor. 3.5.7) for all MZVs of weight 10.

3.12. Prove Cor. 3.5.4.

3.13. Recall that P, T' and T_0 are defined by Eqs. (3.27), (3.33) and (3.30), respectively. Show that for all $f \in V_n$

$$\begin{aligned}((f|P)T')(x_1,\ldots,x_n) &= (f|T_0P)(x_1,\ldots,x_n)\\ &= f(-x_n,-x_n-x_{n-1},\ldots,-x_n-\cdots-x_1).\end{aligned}$$

3.14. Recall that $\tau_{1,n+1}$ is the transposition $1 \leftrightarrow n+1$, and $\mathrm{sh}_1^{(n)}$ and C_{n+1} are defined by Eqs. (3.31) and (3.29), respectively. Show that by regarding $a_1 \ldots a_{n+1}$ as a word

$$\begin{aligned}&(1+\mathrm{sh}_1^{(n)} C_{n+1})(a_1\ldots a_{n+1}) = C_{n+1}(1+\mathrm{sh}_1^{(n)}\tau_{1n+1})(a_1\ldots a_{n+1})\\ &= a_1\ldots a_{n+1}+a_{n+1}a_1\ldots a_n+a_1[a_{n+1}\sqcup\!\sqcup(a_2\ldots a_{n-1})]a_n = a_{n+1}\sqcup\!\sqcup(a_1\ldots a_n).\end{aligned}$$

3.15. For any $w, d \in \mathbb{N}$, show that

$$\zeta^\star(w,\{1\}^{d-1}) = \sum_{t=1}^{d} \sum_{\substack{c_1+\cdots+c_{t-1}=d\\ c_j\geq 0 \forall j=1,\ldots,t-1}} \zeta(c_{t-1}+w, c_{t-2}+1,\ldots,c_1+1).$$

Chapter 4

Drinfeld associator and single-valued MZVs

A crucial observation by Deligne is that one can attach tangential base point on $\mathbb{P}\setminus\{0,1,\infty\}$ to study the torsors of paths and the motivic fundamental group. Let 0_1 and 1_{-1} be the tangential base point at 0 (resp. at 1) pointing to the positive (resp. negative) direction. Then the de Rham version of the Betty path from 0_1 going straight to 1_{-1} gives rise to the Drinfeld associator whose coefficients are all expressible by MZVs.

4.1 Drinfeld associator

4.1.1 *KZ-equations: a general setup*

For future reference, we consider the general case by setting $\mathsf{X} = \{\mathtt{x}_1,\dots,\mathtt{x}_N\}$ to be an alphabet of $N+1$ letters and X^* the set of words on X. Let $\mathbb{C}\langle\langle\mathsf{X}\rangle\rangle$ be the algebra of non-commutative formal power series in X^*. Fix an injection $j:\mathsf{X}\hookrightarrow\mathbb{C}$ and set $p_i = j(\mathtt{x}_i)$ for all $i=0,\dots,N$. Consider the differential equation

$$F'(z) = \sum_{i=0}^{N}\frac{\mathtt{x}_i}{z-p_i}F(z). \tag{4.1}$$

For each i fix a half line $\ell(p_i)$ starting from p_i such that $\bigcap_{i=0}^N \ell(p_i)=\emptyset$. Let $U=\mathbb{C}\setminus\bigcup_{i=0}^N\ell(p_i)$ and fix a branch of logarithm $\log(z-p_0)$ on $\mathbb{C}\setminus\ell(p_0)$.

The following general theorem guarantees the existence of the Drinfeld associator.

Theorem 4.1.1. *The differential equation* (4.1) *has a unique solution $L(z)$ on U such that*

$$L(z) = f_0(z)\exp\big(\mathtt{x}_0\log(z-p_0)\big),$$

where $f_0(z)$ is a holomorphic function on $\mathbb{C}\setminus\bigcup_{i=1}^N\ell(p_i)$ with $f_0(p_0)=1$. Moreover, every holomorphic solution of (4.1) *on $\mathbb{C}\setminus\bigcup_{i=0}^N\{p_i\}$ is equal to $L(z)C$ for some $C\in\mathbb{C}\langle\langle\mathsf{X}\rangle\rangle$ depending only on p_i's but not on z.*

Proof. Set $L(z) = \sum_{\mathbf{w}\in\mathsf{X}^*} L_{\mathbf{w}}(z)\mathbf{w}$. Note that $L(z)$ satisfying (4.1) is equivalent to

$$\frac{\partial}{\partial z} L_{\mathtt{x}_i\mathbf{w}}(z) = \frac{L_{\mathbf{w}}(z)}{z-p_i}$$

for all $i = 0, \ldots, N$ and all $\mathbf{w} \in \mathsf{X}^*$. We now recursively define functions $L_{\mathbf{w}}(z)$ satisfying (4.1) for all words $\mathbf{w} \in \mathsf{X}^*$ not ending in the letter $\mathtt{x}_0$. For such words, the limiting condition is just $\lim_{z\to p_0} L_{\mathbf{w}}(z) = 0$. If we write $\mathbf{w} = \mathtt{x}_0^{s_1-1}\mathtt{x}_{i_1}\ldots\mathtt{x}_0^{s_d-1}\mathtt{x}_{i_d}$, where $1 \le i_1, \ldots, i_d \le N$, then in a neighborhood of p_0 the following function satisfies (4.1):

$$L_{\mathbf{w}}(z) = \sum_{k_1>\cdots>k_d} \frac{(-1)^r}{k_1^{s_1}\cdots k_d^{s_d}} \prod_{m=1}^{d} \Big(\frac{z-p_0}{p_{i_m}-p_0}\Big)^{k_m-k_{m+1}} \quad (k_{d+1}=0),$$

which converges absolutely for $|z-p_0| < \min\{|p_{i_1}-p_0|, \ldots, |p_{i_d}-p_0|\}$. The functions $L_{\mathbf{w}}(z)$ can be extended analytically to the whole of U by the recursive formula

$$L_{\mathtt{x}_i\mathbf{w}}(z) = \int_0^z \frac{L_{\mathbf{w}}(t)}{t-p_i}\mathrm{d}t, \quad \forall i = 0, \ldots, N.$$

Now we write $f_0(z) = \sum_{\mathbf{w}\in\mathsf{X}^*} \Lambda_{\mathbf{w}}(z)\mathbf{w}$ where

$$\Lambda_{\mathbf{w}}(z) = \begin{cases} 1, \text{ if } \mathbf{w} = \mathtt{e}; \\ 0, \text{ if } \mathbf{w} = \mathtt{x}_0^n,\ n \in \mathbb{N}; \end{cases} \tag{4.2}$$

and if $\mathbf{w} = \mathtt{x}_0^{s_1}\mathtt{x}_{i_1}\ldots\mathtt{x}_0^{s_d}\mathtt{x}_{i_d}\mathtt{x}_0^n$, where $s_1, \ldots, s_d \ge 0$ and $1 \le i_1, \ldots, i_d \le N$, then

$$\Lambda_{\mathbf{w}}(z) = (-1)^n \sum_{s_j\le k_j, |\mathbf{k}|=|\mathbf{s}|+n} \prod_{j=1}^{d} \binom{k_j}{s_j} L_{\mathtt{x}_0^{k_1}\mathtt{x}_{i_1}\ldots\mathtt{x}_0^{k_d}\mathtt{x}_{i_d}}(z). \tag{4.3}$$

Then it is not hard to see that $L(z) = f_0(z)\exp\big(\mathtt{x}_0\log(z-p_0)\big)$ satisfies equation (4.1). Note that $f_0(z)$ is holomorphic in a neighborhood of p_0 since its coefficients are just linear combinations of the functions $L_{\mathbf{w}}(z)$.

Finally, the uniqueness can be proved in a straightforward manner. This completes the proof of the theorem. □

Definition 4.1.2. Let $\mathtt{x}_0$ and $\mathtt{x}_1$ be two non-commuting symbols. Let G_0 and G_1 be the analytic solutions of the *Knizhnik-Zamolodchikov (KZ) differential equation*

$$G'(z) = \left(\frac{\mathtt{x}_0}{z} + \frac{\mathtt{x}_1}{z-1}\right) G(z), \tag{4.4}$$

that are defined for $0 < z < 1$, such that

$$G_0(z) \sim z^{x_0} \text{ near } z = 0 \quad \text{and} \quad G_1(z) \sim (1-z)^{x_1} \text{ near } z = 1.$$

This means both $G_0(z)$ and $G_1(z)$ are formal series in $\mathtt{x}_0$ and $\mathtt{x}_1$ with coefficients given by analytic functions of z in $(0,1)$. Moreover, $G_0(z)e^{-\mathtt{x}_0 \log z}$ (resp. $G_1(z)e^{-\mathtt{x}_1 \log(1-z)}$) has an analytic continuation into a neighborhood of $z = 0$ (reps. $z = 1$) and becomes 1 at this point. The *Drinfeld associator* is define by

$$\Phi(\mathtt{x}_0, \mathtt{x}_1) := 1 + \sum_{\mathbf{w} \in \mathsf{X}^*} \Phi[\mathbf{w}]\mathbf{w} := G_1^{-1} G_0. \tag{4.5}$$

Here and in what follows $f[\mathbf{w}]$ always means the coefficient of $\mathbf{w}$ in f.

Remark 4.1.3. The existence of $G_0(z)$ and $G_1(z)$ follows from Thm. 4.1.1 by taking $p_0 = 0$ and 1, respectively.

Notice that $\Phi(\mathtt{x}_0, \mathtt{x}_1)$ should only depend on $\mathtt{x}_0$ and $\mathtt{x}_1$. In order to understand this we may consider the analogous but much easier differential equation $f'(z) = f(z)$ with two solutions $f_0(z)$ and $f_1(z)$ determined by the condition that $f_0(0) = 1$ and $f_1(1) = 1$. Then $f_0(z) = e^z$ and $f_1(z) = e^{z-1}$. So $f_1(z)^{-1} f_0(z) = e$ which is a constant.

Theorem 4.1.4. *The Drinfeld associator satisfy the following duality relation:*

$$\Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1} = \Phi(\mathtt{x}_1, \mathtt{x}_0). \tag{4.6}$$

Proof. First observe that by $z \leftrightarrow 1 - z$ we can rewrite Eq. (4.4) as

$$H'(z) = \left(\frac{\mathtt{x}_0}{z-1} + \frac{\mathtt{x}_1}{z} \right) H(z), \quad H(z) = G(1-z).$$

Thus $\Phi^{-1} = G_0^{-1} G_1 = H_1^{-1} H_0$ which implies that Φ^{-1} can be obtained from Φ by the substitution $\mathtt{x}_0 \leftrightarrow \mathtt{x}_1$. □

Remark 4.1.5. Here is another way to understand $\Phi(\mathtt{x}_0, \mathtt{x}_1)$. Since $1/z$ and $1/(1-z)$ are both analytic in $(0,1)$ for any $0 < u < 1$ there is a unique solution to Eq. (4.4) with $G(u) = 1$. Let $Z_u^v(\mathtt{x}_0, \mathtt{x}_1) = G(v)$ for $v \in (0,1)$. Since $Z_u^u(\mathtt{x}_0, \mathtt{x}_1) = 1$ we can write

$$Z_u^v(\mathtt{x}_0, \mathtt{x}_1) = 1 + \sum_{\mathbf{w} \in \mathsf{X}^*} f_{\mathbf{w}}(u, v)\mathbf{w},$$

where $\mathfrak{A}$ is defined in Definition 3.2.1 and $f_{\mathbf{w}}(u, v)$ is an analytic function in u and v. Define the iterated integral $I_{\mathbf{w}}(u, v) = \int_u^v w(\mathrm{d}t/t, \mathrm{d}t/(t-1))$ for any

word $\mathbf{w} = \mathbf{w}(\mathtt{x}_0, \mathtt{x}_1)$. By induction on the weight of the word we see easily that $f_{\mathbf{w}}(u, v) = I_{\mathbf{w}}(u, v)$. If $\mathbf{w} = \mathtt{x}_0^{s_1-1}\mathtt{x}_1 \cdots \mathtt{x}_0^{s_d-1}\mathtt{x}_1 \in \mathfrak{A}^0$ (*i.e.*, $s_1 \geq 2$, see Definition 3.2.2) then we have $\lim_{\varepsilon\to 0^+} I_{\mathbf{w}}(\varepsilon, 1-\varepsilon) = (-1)^d\zeta(s_1, \ldots, s_d)$ by Eq. (3.1). But if $\mathbf{w}$ is not admissible we have to regularize $I_{\mathbf{w}}(\varepsilon, 1-\varepsilon)$. It is possible to compute this explicitly (see [368, Appendix]) but this would take us too far afield. Another high powered machinery to do this is to use some Hopf algebra fact whose proof will be omitted.

4.1.2 *Some Hopf algebras and their group-like elements*

Let $\mathsf{X} = \{\mathtt{x}_0, \mathtt{x}_1\}$ and $\mathscr{R} := \mathbb{C}\langle\langle \mathsf{X} \rangle\rangle$ be the Hopf algebra as defined in Example A.12 with concatenation product and with coproduct $\Delta_{ш}$ such that $\Delta_{ш}(\mathtt{x}_i) = 1 \otimes \mathtt{x}_i + \mathtt{x}_i \otimes 1$ for $i = 0, 1$. The counit $\epsilon_{ш}$ is defined by $\epsilon_{ш}(\mathbf{w}) = 0$ for all nontrivial word $\mathbf{w}$ and $\epsilon_{ш}(\mathtt{e}) = 1$. Notice that $\mathscr{R}$ can be obtained from $\mathfrak{A}$ by an extension of scalars to $\mathbb{C}$ and then by completion with respect to the weight grading.

We now consider $\mathscr{R}_{ш} := \mathscr{R}/\mathscr{R}\mathtt{x}_0$ which can be identified with the subalgebra of $\mathscr{R}$ generated by words ending with $\mathtt{x}_1$. It can be thought as the completion of $\mathfrak{A}^1_{ш} \otimes \mathbb{C}$ with respect to the weight grading. By the above we see that the Drinfeld associator is closely related to the shuffle regularization of the MVZs. Similarly we can treat the stuffle version using a $*$-version of Φ.

Let $\mathscr{R}_* := \mathbb{C}\langle\langle \mathtt{y}_n,\ n \in \mathbb{N} \rangle\rangle$ be the Hopf algebra with $\Delta_*(\mathtt{y}_n) = \sum_{k=0}^n \mathtt{y}_k \otimes \mathtt{y}_{n-k}$ for all $n \in \mathbb{N}$. Clearly, this is the completion of $\mathfrak{A}^1_* \otimes \mathbb{C}$ with respect to the weight grading. Define the projection $\pi : \mathscr{R} \longrightarrow \mathscr{R}_{ш}$ by forgetting the terms corresponding to the words ending with $\mathtt{x}_0$. Notice that if we set $\mathtt{y}_n = \mathtt{x}_0^{n-1}\mathtt{x}_0$ for all $n \in \mathbb{N}$ then as concatenation algebras $\mathscr{R}_{ш} = \mathscr{R}_*$. Define

$$\Phi_{ш} = \pi(\Phi).$$

Theorem 4.1.6. *Let $f_* = \sum_{\mathbf{w}\in\mathsf{X}^*\mathtt{x}_1} f_*[\mathbf{w}]\mathbf{w} \in \mathscr{R}_*$. Then f_* satisfies stuffle relation if and only if $\Delta_*(f_*) = f_* \otimes f_*$,* i.e.*, f_* is group-like for Δ_*.*

Proof. Throughout this proof, by abuse of notation, for any composition $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ we also identify it with $(-1)^d\mathtt{x}_0^{s_1-1}\mathtt{x}_1 \ldots \mathtt{x}_0^{s_d-1}\mathtt{x}_1 \in \mathfrak{A}^1_*$ when we write $\mathbf{s} \in \mathfrak{A}^1_*$.

By the definition $\Delta_*(f_*) = f_* \otimes f_*$ means for all $\mathbf{s}, \mathbf{t} \in \mathfrak{A}^1_*$ we have

$$f_*[\mathbf{s}]f_*[\mathbf{t}] = \sum_{\substack{(r'_1,\ldots,r'_\ell)=\mathbf{s}\\(r''_1,\ldots,r''_\ell)=\mathbf{t}\\r'_j+r''_j>0}} f_*[r'_1 + r''_1, \ldots, r'_\ell + r''_\ell]. \tag{4.7}$$

Extending f_* linearly to $\mathfrak{A}^1_*$ we need to show Eq. (4.7) is equivalent to

$$f_*[\mathbf{s}]f_*[\mathbf{t}] = f_*[\mathbf{s} * \mathbf{t}] \text{ for all } \mathbf{s}, \mathbf{t} \in \mathfrak{A}^1_*. \tag{4.8}$$

Assuming Eq. (4.7) we now prove Eq. (4.8) by induction on $|\mathbf{s}| + |\mathbf{t}|$. If either $\mathbf{s}$ or $\mathbf{t}$ is $\mathtt{e}$ then this is obvious. In general, we have

$$f_*[m, \mathbf{s}]f_*[n, \mathbf{t}] = \sum_{\substack{(r'_1,\dots,r'_\ell)=(m,\mathbf{s}) \\ (r''_1,\dots,r''_\ell)=(n,\mathbf{t}) \\ r'_j+r''_j>0 \ \forall j=1,\dots,\ell}} f_*[r'_1 + r''_1, \dots, r'_\ell + r''_\ell].$$

On the right-hand side we have three cases to consider: $(r'_1, r''_1) = (m, 0)$, $(r'_1, r''_1) = (0, n)$, and $(r'_1, r''_1) = (m, n)$. By induction assumption we thus get

$$\begin{aligned} f_*[m, \mathbf{s}]f_*[n, \mathbf{t}] &= f_*\big[m, \mathbf{s} * (n, \mathbf{t})\big] + f_*\big[n, (m, \mathbf{s}) * \mathbf{t}\big] + f_*\big[m + n, \mathbf{s} * \mathbf{t}\big] \\ &= f_*\big[(m, \mathbf{s}) * (n, \mathbf{t})\big]. \end{aligned}$$

It is not too hard to check that Eq. (4.8) implies Eq. (4.7) (see Exercise 4.2). This finishes the proof of the theorem. □

Since the $*$-regularized MZVs $\zeta_*(s_1, \dots, s_d; T)$ satisfy stuffle relations the next result follows immediately.

Corollary 4.1.7. *Let*

$$\Phi_* = \sum_{s_1,\dots,s_d \in \mathbb{N}} (-1)^d \zeta_*(s_1, \dots, s_d) \mathtt{y}_{s_1} \cdots \mathtt{y}_{s_d}.$$

Then $\Delta_*(\Phi_*) = \Phi_* \otimes \Phi_*$.

Theorem 4.1.8. *Let* $f = \sum_{\mathbf{w} \in \mathtt{x}_0 \mathtt{X}^* \mathtt{x}_1} f[\mathbf{w}]\mathbf{w} \in (\mathcal{R}_{\mathrm{cv}}, \Delta_{ш})$. *Then* $f[\mathbf{w}]$ *satisfies shuffle relation if and only if* $\Delta_{ш}(f) = f \otimes f$, i.e., *$f$ is group-like. Similar result for* $(\mathcal{R}_{ш}, \Delta_{ш})$ *holds, too.*

Proof. The computation is straightforward so we leave it as an exercise (see Exercise 4.3). □

We leave the proof of next two technical lemmas to the interested reader (see Exercise 4.4).

Lemma 4.1.9. *Let* $\mathcal{R}^T_* = \mathcal{R}_* \otimes_{\mathbb{C}} \mathbb{C}[T]$. *Any group-like element* f_{cv} *of* $(\mathcal{R}_{\mathrm{cv}}, \Delta_*)$ *is the image of a group-like element of* $(\mathcal{R}^T_*, \Delta_*)$. *Moreover, two such elements* f_1 *and* f_2 *are related by*

$$f_2 = \exp((\lambda_2 - \lambda_1)\mathtt{y}_1) f_1,$$

where λ_1 *and* λ_2 *are the coefficients of* $\mathtt{y}_1$ *in* f_1 *and* f_2 *respectively.*

Similarly one can prove

Lemma 4.1.10. *Let $\mathscr{R}_{ш}^{T} = \mathscr{R}_{ш} \otimes_{\mathbb{C}} \mathbb{C}[T]$. Any group-like element f_{cv} of $(\mathscr{R}_{\mathrm{cv}}, \Delta_{ш})$ is the image of a group-like element of $(\mathscr{R}_{ш}^{T}, \Delta_{ш})$. Moreover, two such elements f_1 and f_2 are related by*

$$f_2 = \exp((\lambda_2 - \lambda_1)\mathrm{x}_1) f_1,$$

where λ_1 and λ_2 are the coefficients of x_1 in f_1 and f_2 respectively.

Theorem 4.1.11. *The Drinfeld associator $\Phi(\mathrm{x}_0, \mathrm{x}_1)$ is the only element in Hopf algebra $(\mathscr{R}, \Delta_{ш}, \epsilon_{ш})$ such that*

(i) $\Delta_{ш}(\Phi) = \Phi \otimes \Phi$ *(*i.e. Φ *is group-like, see Definition* A.6*)*
(ii) $\Phi[\mathrm{x}_0] = \Phi[\mathrm{x}_1] = 0$
(iii) $\Phi[\mathbf{w}] = (-1)^d \zeta(s_1, \ldots, s_d) \quad \forall \mathbf{w} = \mathrm{x}_0^{s_1-1}\mathrm{x}_1 \cdots \mathrm{x}_0^{s_d-1}\mathrm{x}_1 \in \mathfrak{A}^0$.

Proof. This will follow from Lemma 4.1.10 easily. □

Remark 4.1.12. In Chap. 13 we will prove a more general statement in (see Thm. 13.4.1).

It is easy to see that the shuffle relations between MZVs are direct consequences of the fact that $\Delta_{ш}(\Phi) = \Phi \otimes \Phi$. See Thm. 4.1.8 and Exercise 4.3.

Using Thm. 4.1.11 we can prove the following result which provides the recursive relations one may use to compute all the coefficients of Φ.

Proposition 4.1.13. *We have*

(i) *For $\mathbf{w} = \mathrm{x}_0^{s_1-1}\mathrm{x}_1 \cdots \mathrm{x}_0^{s_d-1}\mathrm{x}_1 \in \mathfrak{A}^0$, we have $\Phi[\mathbf{w}] = (-1)^d \zeta(s_1, \ldots, s_d)$.*
(ii) *For $\mathbf{w} = \mathrm{x}_1^m \mathbf{v} \mathrm{x}_0^n \notin \mathfrak{A}^1$ with $m + n > 0$ we may set $\mathbf{w} = \mathrm{x}_1^m \mathbf{v} \mathrm{x}_0^n = \mathrm{x}_1^m l_1 \cdots l_p \mathrm{x}_0^n$ where $l_j \in \{\mathrm{x}_0, \mathrm{x}_1\}$ and if $p > 0$ then $l_1 \neq \mathrm{x}_1, l_p \neq \mathrm{x}_0$. Further set $l_1 \cdots l_p \mathrm{x}_0^s = l_1 \cdots l_q$. Then*

$$\Phi[\mathbf{w}] = \begin{cases} 0, & \text{if } mn = p = 0; \\ -\dfrac{1}{m} \displaystyle\sum_{i=1}^{q} \Phi[\mathrm{x}_1^{m-1} l_1 \cdots l_i \mathrm{x}_1 l_{i+1} \ldots l_q], & \text{if } m > 0; \\ -\dfrac{1}{n} \displaystyle\sum_{i=1}^{p} \Phi[l_1 \cdots l_{i-1} \mathrm{x}_0 l_i \cdots l_p \mathrm{x}_0^{n-1}], & \text{if } m = 0, n > 0. \end{cases} \tag{4.9}$$

Proof. We only need to prove (ii). Since Φ is group-like one has

$$\Delta_{ш}\Phi = \Phi \otimes \Phi. \tag{4.10}$$

First we assume $mn = p = 0$. We may further assume $m = 0$ ($n = 0$ is completely similar). We need to show that $\Phi[\mathtt{x}_0^n] = 0$ for all $n \in \mathbb{N}$. We prove this by induction on n. Suppose $\Phi[\mathtt{x}_0^{n-1}] = 0$ for some $n \geq 1$. By Thm. 4.1.11 we have $\Phi[\mathtt{x}_0] = 0$ and by Eq. (4.10)

$$\begin{aligned}\Delta_{ш}(1) + \Phi[\mathtt{x}_0^2](\Delta_{ш}(\mathtt{x}_0))^2 + \cdots &= 1 \otimes 1 + \Phi[\mathtt{x}_0^2](\mathtt{x}_0 \otimes 1 + 1 \otimes \mathtt{x}_0)^2 + \cdots \\ &= (1 + \Phi[\mathtt{x}_0]\mathtt{x}_0 + \cdots) \otimes (1 + \Phi[\mathtt{x}_0]\mathtt{x}_0 + \cdots).\end{aligned}$$

We now compare the coefficient of $\mathtt{x}_0^{n-1} \otimes \mathtt{x}_0$ on the two sides. The only way to produce this on the left-hand side is by $\Phi[\mathtt{x}_0^n]\Delta_{ш}(\mathtt{x}_0^n)$ which in fact produces n copies of it. On the right-hand side the only way is to use $\Phi[\mathtt{x}_0^{n-1}]\mathtt{x}_0^{n-1} \otimes \Phi[\mathtt{x}_0]\mathtt{x}_0$. Hence $n\Phi[\mathtt{x}_0^n] = \Phi[\mathtt{x}_0^{n-1}]\Phi[\mathtt{x}_0] = 0$ and therefore $\Phi[\mathtt{x}_0^n] = 0$.

By similar argument if $m > 0$ then one can compare the coefficient of $\mathtt{x}_1 \otimes (\mathtt{x}_1^{m-1} l_1 \cdots l_q)$ of the two sides of Eq. (4.10) (which is 0) and find the relation (4.9). Finally, if $m = 0$ and $n > 0$ then one may similarly consider the coefficient of $(l_1 \cdots l_p \mathtt{x}_0^{n-1}) \otimes \mathtt{x}_0$ in Eq. (4.10). This finishes the proof of the proposition. □

Theorem 4.1.14. *If* $\mathbf{w} = \{0\}^{s_1-1}1\cdots\{0\}^{s_d-1}1\{0\}^{n-1}$ *then*

$$\Phi[\mathbf{w}] = (-1)^n \sum_{k_1+\cdots+k_d=n} \prod_{j=1}^{d} \binom{k_j + s_j - 1}{k_j} \zeta(s_1 + k_1, \ldots, s_d + k_d).$$

Proof. This can be proved easily by induction on n using Prop. 4.1.13 so is left to the interested reader. □

Theorem 4.1.15. *We have* $\Phi_{ш} = \Lambda(\mathtt{x}_1)\Phi_*$ *where* Λ *is defined by Eq.* (3.17)*:*

$$\Lambda(u) = \exp\left(\sum_{n=2}^{\infty} \frac{(-1)^n \zeta(n)}{n} u^n\right).$$

Proof. Let $\mathscr{R}_{\mathrm{cv}} := \mathscr{R}_*/\mathtt{y}_1\mathscr{R}_* \cong \mathbb{C}\langle\langle \mathtt{y}_n, n \geq 2\rangle\rangle$ which is also equal to $\mathscr{R}_*/\mathtt{x}_1\mathscr{R}_*$. Let $\pi_{\mathrm{cv}} : \mathscr{R}_* \to \mathscr{R}_{\mathrm{cv}}$ be the projection map by forgetting all the terms corresponding to the words beginning with $\mathtt{x}_1$. The generating series of $\zeta(s_1, \ldots, s_d)$ is defined as follows:

$$\begin{aligned}\mathbf{L}_{\mathrm{cv}} &:= \sum_{s_1,\cdots,s_d \in \mathbb{N}, s_1 \geq 2} (-1)^d \zeta(s_1, \ldots, s_d) \mathtt{y}_{s_1} \cdots \mathtt{y}_{s_d} (\in \mathscr{R}_{\mathrm{cv}}, \Delta_*), \\ &:= \sum_{s_1,\cdots,s_d \in \mathbb{N}, s_1 \geq 2} (-1)^d \zeta(s_1, \ldots, s_d) \mathtt{x}_0^{s_1-1}\mathtt{x}_1 \cdots \mathtt{x}_0^{s_d-1}\mathtt{x}_1 (\in \mathscr{R}_{\mathrm{cv}}, \Delta_{ш}).\end{aligned}$$

Since MZVs satisfy the stuffle relations we see that $\mathbf{L}_{\mathrm{cv}}$ is a group-like elements in $(\mathscr{R}_{\mathrm{cv}}, \Delta_*)$ by a similar argument as in the proof of Thm. 4.1.6. On the other hand, since MZVs also satisfy the shuffle relations $\mathbf{L}_{\mathrm{cv}}$ is also a group-like elements in $(\mathscr{R}_{\mathrm{cv}}, \Delta_{ш})$ by Thm. 4.1.8.

Let $\mathbf{L} := \exp(T\mathtt{y}_1) \cdot \Phi_* \in \mathscr{R}_*^T$ and $\mathbf{I} := \exp(Tb) \cdot \Phi_{ш} \in \mathscr{R}_{ш}^T$. Clearly $\mathbf{L}$ is group-like for Δ_* and $\mathbf{I}$ is group-like for $\Delta_{ш}$. Moreover, it is straightforward to check the coefficients of $\mathtt{y}_1 = \mathtt{x}_0$ in $\mathbf{L}$ and $\mathbf{I}$ are both equal to T. By the uniqueness guaranteed by Lemma 4.1.9 and Lemma 4.1.10

$$\mathbf{L} = \sum_{\mathbf{w} \in \mathsf{X}^* \mathtt{x}_1} \zeta_*(\mathbf{w}; T)\mathbf{w}, \qquad \mathbf{I} = \sum_{\mathbf{w} \in \mathsf{X}^* \mathtt{x}_1} \zeta_{ш}(\mathbf{w}; T)\mathbf{w}$$

since both of the right-hand sides are group-like (with respect to Δ_* by Thm. 4.1.6 and $\Delta_{ш}$ by Thm. 4.1.8, respectively). By Thm. 3.3.21 we see that

$$\rho(\mathbf{L}) = \mathbf{I} = \exp(Tb) \cdot \Phi_{ш}.$$

But by the definition of ρ in Lemma 3.3.17

$$\rho(\mathbf{L}) = \rho(\exp(T\mathtt{y}_1) \cdot \Phi_*) = \Lambda(\mathtt{x}_1) \exp(T\mathtt{x}_1)\Phi_* = \exp(T\mathtt{x}_1) \cdot \Lambda(\mathtt{x}_1)\Phi_*.$$

The theorem follows immediately. □

4.1.3 *Grothendieck-Teichmüller group*

To have a better understanding of the Drinfeld associator Φ and also for future reference we need the following definitions.

Definition 4.1.16. Let $\mathbb{K}$ be a field of characteristic 0. For $n \geq 2$, denote by $\mathfrak{t}_n$ the completed $\mathbb{K}$-Lie algebra with generators t^{ij} $(i \neq j,\ 1 \leq i, j \leq n)$ subject to the following *(infinitesimal) pure braids relations*[1]:

$$t^{ij} = t^{ji},\ [t^{ij}, t^{ik} + t^{jk}] = 0 \text{ and } [t^{ij}, t^{kl}] = 0 \quad \text{for all distinct } i,\ j,\ k,\ l. \tag{4.11}$$

Let $\mathfrak{t}_n^0$ be the Lie subalgebra of $\mathfrak{t}_n$ with the generators t^{ij} and the same relations as $\mathfrak{t}_n$ plus one more relation $\sum_{1 \leq i < j \leq n} t^{ij} = 0$.

Example 4.1.17. We note that $\mathfrak{t}_2$ is the 1-dimensional abelian Lie algebra generated by t^{12}. Further, $\mathfrak{t}_3^0$ is the free Lie algebra $\mathbb{K}\langle \mathtt{x}_0, \mathtt{x}_1 \rangle$ of rank 2 with generators $\mathtt{x}_0 := t^{12}$ and $\mathtt{x}_1 := t^{23}$. See Exercise 4.5.

[1] Due to its relation to the braid groups this Lie algebra is called the *Lie algebra of pure braids*. But we will not go into the details of these relations in this book.

Definition 4.1.18. The *Grothendieck-Teichmüller Lie algebra* $\mathfrak{grt}_1(\mathbb{K})$ is defined to be the set of $\varphi = \varphi(\mathtt{x}_0, \mathtt{x}_1) \in \mathfrak{t}_3^0$ satisfying the *duality and hexagon relations* in $\mathfrak{t}_3^0$

$$\varphi(\mathtt{x}_0, \mathtt{x}_1) + \varphi(\mathtt{x}_1, \mathtt{x}_0) = 0, \qquad \varphi(\mathtt{x}_0, \mathtt{x}_1) + \varphi(\mathtt{x}_1, \mathtt{x}_\infty) + \varphi(\mathtt{x}_\infty, \mathtt{x}_0) = 0 \tag{4.12}$$

with $\mathtt{x}_0 + \mathtt{x}_1 + \mathtt{x}_\infty = 0$, and the *pentagon relation* in $\mathfrak{t}_4^0$

$$\begin{aligned}\varphi(t^{12}, t^{23} + t^{24}) + \varphi(t^{13} + t^{23}, t^{34}) \\ = \varphi(t^{23}, t^{34}) + \varphi(t^{12} + t^{13}, t^{24} + t^{34}) + \varphi(t^{12}, t^{23}).\end{aligned} \tag{4.13}$$

See Exercise 4.6 for the definition of the corresponding Lie bracket.

Proposition 4.1.19. *If* $\varphi \in \mathfrak{grt}_1(\mathbb{K})$ *then it satisfies the the* special derivation condition *in* $\mathfrak{t}_3^0$

$$[\mathtt{x}_1, \varphi(\mathtt{x}_0, \mathtt{x}_1)] + [\mathtt{x}_\infty, \varphi(\mathtt{x}_0, \mathtt{x}_\infty)] = 0 \tag{4.14}$$

for $\mathtt{x}_0 + \mathtt{x}_1 + \mathtt{x}_\infty = 0$.

Proof. Let $s(\mathtt{x}_1, \mathtt{x}_\infty) = [\mathtt{x}_1, \varphi(\mathtt{x}_0, \mathtt{x}_1)] + [\mathtt{x}_\infty, \varphi(\mathtt{x}_0, \mathtt{x}_\infty)]$. Then $s(\mathtt{x}_1, \mathtt{x}_\infty) = s(\mathtt{x}_\infty, \mathtt{x}_1)$. Now let Y_1, Y_2 and Y_3 be any generators of $\mathfrak{t}_4$ and set

$$u(Y_1, Y_2, Y_3) := s(Y_1, Y_2) - s(Y_1, Y_2 + Y_3) + s(Y_1 + Y_2, Y_3) - s(Y_2, Y_3).$$

By Eq. (4.12) we have

$$u(t^{14}, t^{24}, t^{34}) = [t^{14} + t^{24} + t^{34}, m^{1234}] - [t^{14} + t^{24}, m^{1243}] + [t^{14}, m^{1423}], \tag{4.15}$$

where $m^{1234} = \{\text{left-hand side of (4.13)}\} - \{\text{right-hand side of (4.13)}\} = 0$ (see Exercise 4.12). Hence $u(Y_1, Y_2, Y_3) = 0$ and by Thm. B.27 we see that $s = 0$ since s is a Lie polynomial by Thm. B.16. This finishes the proof of the proposition. □

Definition 4.1.20. The *Grothendieck-Teichmüller group* $\mathrm{GRT}_1(\mathbb{K})$ is defined to be the set of series $g \in \exp \mathfrak{t}_3^0$ satisfying the *duality and hexagon relations* in $\exp \mathfrak{t}_3^0$

$$g(\mathtt{x}_0, \mathtt{x}_1) g(\mathtt{x}_1, \mathtt{x}_0) = 1, \qquad g(\mathtt{x}_\infty, \mathtt{x}_0) g(\mathtt{x}_1, \mathtt{x}_\infty) g(\mathtt{x}_0, \mathtt{x}_1) = 1 \tag{4.16}$$

with $\mathtt{x}_0 + \mathtt{x}_1 + \mathtt{x}_\infty = 0$, and the *pentagon equation* in $\exp \mathfrak{t}_4^0$

$$g(t^{12}, t^{23} + t^{24}) g(t^{13} + t^{23}, t^{34}) = g(t^{23}, t^{34}) g(t^{12} + t^{13}, t^{24} + t^{34}) g(t^{12}, t^{23}). \tag{4.17}$$

See Exercise 4.7 for the definition of the corresponding group multiplication. We note that the associated Lie algebra is $\mathfrak{grt}_1(\mathbb{K})$.

Definition 4.1.21. The *associator set* $\mathrm{M}_\lambda(\mathbb{K})$ is defined to be the set of series $g \in \exp \mathfrak{t}_3^0$ satisfying the pentagon relation (4.17) and the following duality and λ-hexagon relations

$$g(\mathtt{x}_0,\mathtt{x}_1)g(\mathtt{x}_1,\mathtt{x}_0)=1, \quad e^{\lambda\mathtt{x}_0/2}g(\mathtt{x}_\infty,\mathtt{x}_0)e^{\lambda\mathtt{x}_\infty/2}g(\mathtt{x}_1,\mathtt{x}_\infty)e^{\lambda\mathtt{x}_1/2}g(\mathtt{x}_0,\mathtt{x}_1)=1 \tag{4.18}$$

with $\mathtt{x}_0+\mathtt{x}_1+\mathtt{x}_\infty=0$ and $\lambda \in \mathbb{K}^\times$.

By the above definitions it is clear that $\mathrm{GRT}_1(\mathbb{K}) = \mathrm{M}_0(\mathbb{K})$. This is different from the original definition of $\mathrm{GRT}_1(\mathbb{K})$ where *special action condition* is also required. But by the next proposition proved by Drinfeld in [164] we see this condition is not needed.

Proposition 4.1.22. *If* $g(\mathtt{x}_0,\mathtt{x}_1) \in \mathrm{GRT}_1(\mathbb{K})$ *then it satisfies the the* special action condition *in* $\exp \mathfrak{t}_3^0$

$$\mathtt{x}_0+g(\mathtt{x}_0,\mathtt{x}_1)^{-1}\mathtt{x}_1 g(\mathtt{x}_0,\mathtt{x}_1)+g(\mathtt{x}_0,\mathtt{x}_\infty))^{-1}\mathtt{x}_\infty g(\mathtt{x}_0,\mathtt{x}_\infty)=0, \tag{4.19}$$

for $\mathtt{x}_0+\mathtt{x}_1+\mathtt{x}_\infty=0$.

Proof. Let $g(\mathtt{x}_0,\mathtt{x}_1) \in \mathrm{GRT}_1(\mathbb{K})$, then $\log(g(\mathtt{x}_0,\mathtt{x}_1)) \in \mathfrak{t}_3^0$. For each positive integer n let φ_n be the degree $\le n$ part of $\log(g(\mathtt{x}_0,\mathtt{x}_1))$. Then $\varphi_n \, \mathfrak{grt}_1(\mathbb{K})$ so it satisfies the special derivation condition Eq. (4.14) by Prop. 4.1.19. Since $\exp(-\varphi_n)g(\mathtt{x}_0,\mathtt{x}_1) \equiv 1$ (see Exercise 4.8 for the multiplication definition) modulo degree $n+1$ we see that Eq. (4.19) is true modulo degree $n+1$. This proves the proposition since n is arbitrary. □

The next result of Furusho shows that the pentagon relation is in some sense more fundamental than the other relations. We refer the interested reader to his original proof in [207].

Theorem 4.1.23. (i) *Let* $\varphi \in \mathfrak{t}_3^0$ *with* $\varphi[\mathtt{x}_1]=\varphi[\mathtt{x}_0\mathtt{x}_1]=0$ *(*i.e.*, both coefficients of* $\mathtt{x}_1$ *and* $\mathtt{x}_0\mathtt{x}_1$ *vanish). If* φ *satisfies the pentagon equation* (4.13)*, then it also satisfies duality and hexagon equations* (4.12)*.*

(ii) *Let* $g \in \exp \mathfrak{t}_3^0$ *with* $g[\mathtt{x}_1]=g[\mathtt{x}_0\mathtt{x}_1]=0$*. If* g *satisfies the pentagon equation* (4.17)*, then it also satisfies duality and hexagon equations* (4.16)*.*

(iii) *Let* $g \in \exp \mathfrak{t}_3^0$ *with* $g[\mathtt{x}_1]=0$ *and* $g[\mathtt{x}_0\mathtt{x}_1] \in \mathbb{K}^\times$*. If* g *satisfies the pentagon equation* (4.17)*, then the duality and hexagon equations* (4.18) *hold for* $g(\mathtt{x}_0/\alpha,\mathtt{x}_1/\alpha)$ *with* $\alpha=\pm\sqrt{24g[\mathtt{x}_0\mathtt{x}_1]} \in \bar{\mathbb{K}}$*.*

4.1.4 *Pentagon and hexagon relations*

We shall prove that the Drinfeld associator Φ satisfies both the pentagon and hexagon relations. It is easier to study these relations in the context of quasi-Hopf (for pentagon) and quasi-triangular (for hexagon) algebras. We will not define these formally although we will use some of the key ideas related to these structures in this subsection.

Let $\mathbb{K}$ be a field containing $\mathbb{Q}$. Let h be a formal variable (which is related to the Planck constant in theoretical physics) and $\hbar = h/2\pi i$. Let $\mathfrak{g}$ be a Lie algebra (see Appendix B) over $\mathbb{K}[\![h]\!]$ which as a $\mathbb{K}[\![h]\!]$-module is isomorphic to $\mathcal{V}[\![h]\!]$ for some vector space $\mathcal{V}$ over $\mathbb{K}$. Let $t \in \mathfrak{g} \otimes \mathfrak{g}$ be a $\mathfrak{g}$-invariant symmetric tensor, *i.e.*, $t = \tau(t)$ with $\tau(x \otimes y) = y \otimes x$. For any positive integer $n \geq 3$ we denote by t^{ij} the image of t under the (i,j)-th embedding $\mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}^{\otimes n}$. For example, if $n = 3$ then $t^{12} = t \otimes 1, t^{23} = 1 \otimes t \in \mathfrak{g}^{\otimes 3}$.

Lemma 4.1.24. *The t^{ij} satisfy all the pure braids relations in Eq.* (4.11).

Proof. This is straight-forward by the symmetry of t. □

Definition 4.1.25. Let $\Phi_h = \Phi_h(t^{12}, t^{23}) := G_1^{-1} G_0$ where G_0 and G_1 are the analytic solutions of the KZ differential equation

$$G'(z) = \hbar \left(\frac{t^{12}}{z} + \frac{t^{23}}{z-1} \right) G(z), \tag{4.20}$$

that are defined for $0 < z < 1$, such that

$$G_0(z) \sim z^{\hbar t^{12}} \text{ near } z = 0 \quad \text{and} \quad G_1(z) \sim (1-z)^{\hbar t^{23}} \text{ near } z = 1.$$

Here, $f(z) \sim 1$ near $z = 0$ means that there is an expansion of $f(z) = 1 + f_1(z)h + f_2(z)h^2 + \cdots$ such that $f_j(z)$ are all analytic near $z = 0$.

Remark 4.1.26. By the definition Φ_h is an invertible element of $(U\mathfrak{g})^{\otimes 3}$. From Definition 4.1.2 we see that $\Phi_{2\pi i}(t^{12}, t^{23}) = \Phi(t^{12}, t^{23})$ (or, equivalently, when $\hbar = 1$).

The KZ system (4.20) has the form

$$\frac{\partial W}{\partial z_i} = \hbar \sum_{j \neq i} \frac{t^{ij}}{z_i - z_j} W, \quad i = 1, \ldots, n, \tag{4.21}$$

where $W(z_1, \ldots, z_n) \in (U\mathfrak{g})^{\otimes n}$. Since $\partial W/\partial z_1 + \cdots + \partial W/\partial z_n = 0$ the function W depends only on the difference $z_i - z_j$. Another restriction $\sum_{i=1}^n z_i \partial W/\partial z_i = \hbar \sum_{1=i<j\leq n} t^{ij} W$. further reduces (4.21) to a system of

$n-2$ variables. In particular, if $n=3$ the solution of (4.20) has the form $(z_3-z_1)^{\hbar(t^{12}+t^{13}+t^{23})}G((z_2-z_1)/(z_3-z_1))$ where G satisfies (4.20). Hence

$$W_1 = W_2 \cdot \Phi_h \tag{4.22}$$

where W_1 and W_2 are the solutions of (4.21) for $n=3$ in the region $z_1<z_2<z_3$ of $\mathbb{R}^3$ with the asymptotic behaviors

$$\begin{aligned} W_1 &\sim (z_2-z_1)^{\hbar t^{12}}(z_3-z_1)^{\hbar(t^{13}+t^{23})} \quad \text{for} \quad z_2-z_1 \ll z_3-z_1, \\ W_2 &\sim (z_3-z_2)^{\hbar t^{23}}(z_3-z_1)^{\hbar(t^{13}+t^{23})} \quad \text{for} \quad z_3-z_2 \ll z_3-z_1. \end{aligned} \tag{4.23}$$

Lemma 4.1.27. *For $n=4$, there exist unique solutions $W_1,\dots,W_5$ of the system* (4.21) *with the following given asymptotic behaviors in the given zones:*

$$\begin{aligned} W_1 &\sim (z_2-z_1)^{\hbar t^{12}}(z_3-z_1)^{\hbar(t^{13}+t^{23})}(z_4-z_1)^{\hbar(t^{14}+t^{24}+t^{34})} \\ &\quad \textit{for} \quad z_2-z_1 \ll z_3-z_1 \ll z_4-z_1, \\ W_2 &\sim (z_3-z_2)^{\hbar t^{23}}(z_3-z_1)^{\hbar(t^{12}+t^{13})}(z_4-z_1)^{\hbar(t^{14}+t^{24}+t^{34})} \\ &\quad \textit{for} \quad z_3-z_2 \ll z_3-z_1 \ll z_4-z_1, \\ W_3 &\sim (z_3-z_2)^{\hbar t^{23}}(z_4-z_2)^{\hbar(t^{24}+t^{34})}(z_4-z_1)^{\hbar(t^{12}+t^{13}+t^{14})} \\ &\quad \textit{for} \quad z_3-z_2 \ll z_4-z_2 \ll z_4-z_1, \\ W_4 &\sim (z_4-z_3)^{\hbar t^{34}}(z_4-z_2)^{\hbar(t^{23}+t^{24})}(z_4-z_1)^{\hbar(t^{12}+t^{13}+t^{14})} \\ &\quad \textit{for} \quad z_4-z_3 \ll z_4-z_2 \ll z_4-z_1, \\ W_5 &\sim (z_2-z_1)^{\hbar t^{12}}(z_4-z_3)^{\hbar t^{34}}(z_4-z_1)^{\hbar(t^{13}+t^{14}+t^{23}+t^{24})} \\ &\quad \textit{for} \quad z_2-z_1 \ll z_4-z_1, z_4-z_3 \ll z_4-z_1. \end{aligned}$$

Proof. For W_1 we set $u=(z_2-z_1)/(z_3-z_1)$ and $v=(z_3-z_1)/(z_4-z_1)$. Then we have

$$uv=\frac{z_2-z_1}{z_4-z_1}, \quad 1-uv=\frac{z_4-z_2}{z_4-z_1} \quad 1-u=\frac{z_3-z_2}{z_3-z_1}, \quad 1-v=\frac{z_4-z_3}{z_4-z_1}.$$

Define $W=g(u,v)(z_4-z_1)^{\hbar T}$ where $T=\sum_{1=i<j\le n}t^{ij}$. Then

$$\begin{aligned} \frac{\partial g}{\partial z_2} &= \frac{1}{z_3-z_1}\frac{\partial g}{\partial u}, \qquad \frac{\partial g}{\partial z_3} = \frac{z_1-z_2}{(z_3-z_1)^2}\frac{\partial g}{\partial u}+\frac{1}{z_4-z_1}\frac{\partial g}{\partial v} \\ \Longrightarrow \frac{\partial g}{\partial u} &= (z_3-z_1)\frac{\partial g}{\partial z_2}, \qquad \frac{\partial g}{\partial v} = u(z_4-z_1)\frac{\partial g}{\partial z_2}+(z_4-z_1)\frac{\partial g}{\partial z_3} \end{aligned}$$

Hence g satisfies

$$\frac{\partial g}{\partial u} = \hbar\left(\frac{t^{12}}{u} - \frac{vt^{24}}{1-uv} - \frac{t^{23}}{1-u}\right)g = \hbar\left(\frac{A}{u} + R(u,v)\right)g, \tag{4.24}$$

$$\begin{aligned}\frac{\partial g}{\partial v} &= \hbar\left(\frac{t^{12}}{v} - \frac{ut^{23}}{v(1-u)} - \frac{ut^{24}}{1-uv} + \frac{t^{13}}{v} + \frac{t^{23}}{v(1-u)} - \frac{t^{34}}{1-v}\right)g(u,v)\\ &= \hbar\left(\frac{B}{v} - \frac{ut^{24}}{1-uv} - \frac{t^{34}}{1-v}\right)g = \hbar\left(\frac{B}{v} + S(u,v)\right)g,\end{aligned} \tag{4.25}$$

where $A = t^{12}$ and $B = t^{12} + t^{13} + t^{23}$. Let $g(u,v) = f(u,v)u^A v^B$ then we need to show the following system has a solution:

$$f^{-1}\nabla_u f = hAu^{-1}, \quad f^{-1}\nabla_v f = hBv^{-1},$$

where the linear differential operators $\nabla_u = h(Au^{-1} + R(u,v)) - \partial/\partial u$ and $\nabla_v = h(Bv^{-1} + S(u,v)) - \partial/\partial u$. Notice that $[A,B] = 0$ by Lemma 4.1.24 so the above system has a solution if we notice that

$$R[u^i v^{j+1}] = S[u^{i+1}v^j] = -\delta_{i,j}t^{24} \quad \forall i,j \geq 0, \tag{4.26}$$

where $\delta_{i,j}$ is the Kronecker symbol. To derive the uniqueness, we only need to find the unique power series expansion $f(u,v) = \sum_{i,j\geq 0} c_{i,j}u^i v^j$ so that

$$g(u,v) = f(u,v)u^A v^B = f(u,v)\left(\frac{z_2 - z_1}{z_4 - z_1}\right)^{\hbar t^{12}}\left(\frac{z_3 - z_1}{z_4 - z_1}\right)^{\hbar(t^{13}+t^{23})}$$

satisfies Eqs. (4.24) and (4.25). Starting from $c_{0,0} = 1$ all the coefficients of f can be determined uniquely by iterations. Thus W_1 exists and is unique. See Exercise 4.9 for details.

We also leave the similar proofs of the other solutions to the interested reader. □

Lemma 4.1.28. *Let $W_1, \ldots, W_5$ be the analytic continuation of the functions in Lemma* 4.1.27 *to the whole region $z_1 < z_2 < z_3 < z_4$ in $\mathbb{R}^4$. Then*

$$W_1 = W_2(\Phi_h \otimes 1), \quad W_2 = W_3(\mathrm{id}\otimes\Delta\otimes\mathrm{id})(\Phi_h), \quad W_3 = W_4(1\otimes\Phi_h),$$
$$W_1 = W_5(\Delta\otimes\mathrm{id}\otimes\mathrm{id})(\Phi_h), \quad \textit{and} \quad W_5 = W_4(\mathrm{id}\otimes\mathrm{id}\otimes\Delta)(\Phi_h).$$

Proof. Define two analytic functions on $z_1 < z_2 < z_3 < z_4$:

$$V_1 = W_1(z_4 - z_1)^{-\hbar(t^{14}+t^{24}+t^{34})},$$
$$V_2 = W_2(z_4 - z_1)^{-\hbar(t^{14}+t^{24}+t^{34})}(\Phi_h \otimes 1).$$

In fact, since the factor $(z_4 - z_1)$ does not appear in the asymptotics anymore, the two functions are still well-defined for $z_4 < z_1$. We now prove $V_1 = V_2$. First, both V_1 and V_2 satisfy the equations

$$\frac{\partial V}{\partial z_2} = \hbar \sum_{j \neq i} \frac{t^{ij}}{z_i - z_j} V, \quad i \neq 2, 3, \tag{4.27}$$

$$\frac{\partial V}{\partial z_1} = \hbar \sum_{j \neq 1} \frac{t^{1j}}{z_1 - z_j} V - \hbar V \cdot \frac{t^{14} + t^{24} + t^{34}}{z_1 - z_4}, \tag{4.28}$$

$$\frac{\partial V}{\partial z_4} = \hbar \sum_{j \neq 4} \frac{t^{4j}}{z_4 - z_j} V - \hbar V \cdot \frac{t^{14} + t^{24} + t^{34}}{z_4 - z_1}. \tag{4.29}$$

Now as $z_4 \to \infty$ we see that the differential equations in Eqs. (4.27) and (4.28) are exactly the same as those of Eq. (4.21) when $n = 3$. Moreover,

$$V_1 \sim (z_2 - z_1)^{\hbar t^{12}} (z_3 - z_1)^{\hbar (t^{13} + t^{23})},$$
$$V_2 \cdot \Phi_h^{-1} \sim (z_3 - z_2)^{\hbar t^{23}} (z_3 - z_1)^{\hbar (t^{12} + t^{13})},$$

which are exactly the same as those of Eq. (4.23). By Eq. (4.22) we see that $V_1 = V_2$ as $z_4 \to \infty$. Together with Eq. (4.29) this implies $V_1 = V_2$.

We leave the proofs of the other cases to the interested reader since they are almost identical to the above (see Exercise 4.10). □

Lemma 4.1.28 quickly yields the pentagon relations for Φ_h.

Corollary 4.1.29. *We have*

$$(\mathrm{id} \otimes \mathrm{id} \otimes \Delta)(\Phi_h)(\Delta \otimes \mathrm{id} \otimes \mathrm{id})(\Phi_h) = (1 \otimes \Phi_h)(\mathrm{id} \otimes \Delta \otimes \mathrm{id})(\Phi_h)(\Phi_h \otimes 1).$$

Or, equivalently,

$$\Phi_h(t^{12}, t^{23} + t^{24}) \Phi_h(t^{13} + t^{23}, t^{34}) = \Phi_h(t^{23}, t^{34}) \Phi_h(t^{12} + t^{13}, t^{24} + t^{34}) \Phi_h(t^{12}, t^{23}).$$

We now turn to the hexagon relations.

Lemma 4.1.30. *For $n = 3$, define the unique solutions $W_1, \ldots, W_6$ of the system* (4.21) *with the following given asymptotic behaviors in the given zones in $\{(z_1, z_2, z_3) \in \mathbb{C}^3 : \mathrm{Im}(z_1) \leq \mathrm{Im}(z_2) \leq \mathrm{Im}(z_3), z_1 \neq z_2 \neq z_3 \neq z_1\}$:*

$$W_1 \sim (z_2 - z_1)^{\hbar t^{12}} (z_3 - z_1)^{\hbar (t^{13} + t^{23})} \quad \textit{for} \quad |z_2 - z_1| \ll |z_3 - z_1|,$$
$$W_2 \sim (z_3 - z_2)^{\hbar t^{23}} (z_3 - z_1)^{\hbar (t^{12} + t^{23})} \quad \textit{for} \quad |z_3 - z_2| \ll |z_3 - z_1|,$$
$$W_3 \sim (z_2 - z_3)^{\hbar t^{23}} (z_2 - z_1)^{\hbar (t^{12} + t^{13})} \quad \textit{for} \quad |z_3 - z_2| \ll |z_2 - z_1|,$$

$$
\begin{aligned}
W_4 &\sim (z_3-z_1)^{\hbar t^{13}}(z_2-z_1)^{\hbar(t^{12}+t^{23})} \quad \textit{for} \quad |z_1-z_3| \ll |z_2-z_1|,\\
W_5 &\sim (z_1-z_3)^{\hbar t^{13}}(z_2-z_3)^{\hbar(t^{12}+t^{23})} \quad \textit{for} \quad |z_1-z_3| \ll |z_2-z_3|,\\
W_6 &\sim (z_2-z_1)^{\hbar t^{12}}(z_2-z_3)^{\hbar(t^{13}+t^{23})} \quad \textit{for} \quad |z_2-z_1| \ll |z_3-z_2|.
\end{aligned}
$$

Then we have

$$W_1 = W_2\Phi_h, \qquad W_4 = W_3\Phi_h^{132}, \qquad W_5 = W_6\Phi_h^{312}, \tag{4.30}$$

$$W_2 = W_3 e^{ht^{23}/2}, \quad W_4 = W_5 e^{ht^{13}/2}, \quad W_1 = W_6 e^{h(t^{13}+t^{23})/2}. \tag{4.31}$$

Proof. By Eqs. (4.23) and (4.22) we see that $W_1 = W_2\Phi_h$ when $z_1 < z_2 < z_3$. Similarly, by transposition of indices $2 \leftrightarrow 3$ we see that $W_4 = W_3\Phi_h^{132}$ when $z_1 < z_3 < z_2$, and by a further transposition of indices $1 \leftrightarrow 3$ we see that $W_5 = W_6\Phi_h^{312}$ when $z_3 < z_1 < z_2$.

Now define $V_2 = W_2(z_3-z_2)^{-\hbar t^{23}}$ and $V_3 = W_3(z_3-z_2)^{-\hbar t^{23}}e^{ht^{23}/2} = W_3(z_2-z_3)^{-\hbar t^{23}}$. We now show that $V_2 = V_3$. We see that both functions satisfy

$$
\begin{aligned}
\frac{\partial V}{\partial z_1} &= \hbar\left(\frac{t^{12}}{z_1-z_2}\frac{t^{13}}{z_1-z_3}\right) + V,\\
\frac{\partial V}{\partial z_2} &= \hbar\sum_{j\neq 2}\frac{t^{2j}}{z_2-z_j}V - \hbar V\cdot\frac{t^{23}}{z_2-z_3} = \frac{t^{12}V}{z_2-z_1} + \frac{[t^{23},V]}{z_2-z_3},\\
\frac{\partial V}{\partial z_3} &= \hbar\sum_{j\neq 3}\frac{t^{3j}}{z_3-z_j}V - \hbar V\cdot\frac{t^{23}}{z_3-z_2} = \frac{t^{13}V}{z_3-z_1} + \frac{[t^{23},V]}{z_3-z_2}.
\end{aligned}
$$

Thus $V_2 - V_3$ is an absolute constant. Moreover, both functions have the same asymptotics as $z_2 \to z_3$ which implies $V_2 = V_3$

The other two identities are left as an exercise for the interested reader (see Exercise 4.11). □

Lemma 4.1.30 immediately yields the h-hexagon relations for Φ_h given by Eq. (4.18).

Corollary 4.1.31. *We have*

$$e^{ht^{13}/2}\Phi_h(t^{23},t^{13})e^{ht^{23}/2}\Phi_h(t^{12},t^{23})e^{-h(t^{13}+t^{23})/2}\Phi_h(t^{13},t^{12}) = 1.$$

Proof. In $\exp(\mathfrak{t}_3^0)$ we have $t^{12} = -t^{13} - t^{23}$. □

The following theorem due to Drinfeld shows in particular that the associator Φ satisfies the pentagon relation (4.13). Together with Thm. 4.1.14 one may derive a lot relations among MZVs.

Theorem 4.1.32. *The Drinfeld associator* $\Phi(a,b) \in \mathrm{M}_{2\pi i}(\mathbb{C})$. *Moreover, it satisfies the special action condition*

$$\mathtt{x}_0 + \Phi(\mathtt{x}_0,\mathtt{x}_1)^{-1}\mathtt{x}_1\Phi(\mathtt{x}_0,\mathtt{x}_1) + \Phi(\mathtt{x}_0,\mathtt{x}_\infty)^{-1}\mathtt{x}_\infty\Phi(\mathtt{x}_0,\mathtt{x}_\infty) = 0 \qquad (4.32)$$

if $\mathtt{x}_0 + \mathtt{x}_1 + \mathtt{x}_\infty = 0$.

Proof. By Remark 4.1.26 this follows from Cor. 4.1.29 and 4.1.31. □

In particular, the Drinfeld associator Φ does not lie in the Grothendieck-Teichmüller group $\mathrm{GRT}_1(\mathbb{C})$ but instead it is an element in the associator set $\mathrm{M}_{2\pi i}(\mathbb{C})$ by its duality relation Eq. (4.6) and Thm. 4.1.32.

4.2 Single-variable MPLs

Recall that in the study of MZVs in the last chapter, the single-variable MPLs defined by Eq. (3.19) plays an important role. By definition, it is clear that $\mathrm{L}_{\mathbf{s}}(z) = \mathrm{Li}_{\mathbf{s}}(z,1,\ldots,1)$.

Proposition 4.2.1. *We have*

$$\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{L}_{\mathbf{s}}(z) = \begin{cases} \mathrm{L}_{s_1-1,s_2,\ldots,s_d}(z)/z, & \text{if } s_1 > 1; \\ \mathrm{L}_{s_2,s_3,\ldots,s_d}(z)/(1-z), & \text{if } s_1 = 1. \end{cases}$$

Proof. The case $s_1 > 1$ is obvious. An easy application of geometric series summation formula proves the case $s_1 = 1$ immediately. □

Let $\omega_0(t) = dt/t$ and $\omega_1(t) = dt/(t-1)$. For any path $\gamma : [0,1] \to \mathbb{C}\backslash\{0,1\}$ from $y = \gamma(0)$ to $z = \gamma(1)$ and $a_1,\ldots,a_n \in \{0,1\}$ we define

$$I_\gamma(y; a_1 \ldots a_n; z) = \int_{0<t_1<\ldots<t_n<1} \gamma^*\omega_{a_1}(t_1) \wedge \ldots \wedge \gamma^*\omega_{a_n}(t_n)$$

where $\gamma^*\omega$ is the pullback of ω by γ (see Sec. 2.1). Put $I_\gamma(y;\emptyset;z) = 1$. Let ϕ be one of the six Möbius transformations

$$\phi(z) = z, 1-z, 1-z^{-1}, (1-z^{-1})^{-1}, (1-z)^{-1}, z^{-1}. \qquad (4.33)$$

Then we can extend ϕ linearly to words and to iterated integrals:

$$\begin{aligned} 1-z &: \omega_0 \mapsto \omega_1, & \omega_1 &\mapsto \omega_0, \\ 1-z^{-1} &: \omega_0 \mapsto -\omega_0+\omega_1, & \omega_1 &\mapsto -\omega_0, \\ (1-z^{-1})^{-1} &: \omega_0 \mapsto \omega_0-\omega_1, & \omega_1 &\mapsto -\omega_1, \\ (1-z)^{-1} &: \omega_0 \mapsto -\omega_1, & \omega_1 &\mapsto \omega_0-\omega_1, \\ z^{-1} &: \omega_0 \mapsto -\omega_0, & \omega_1 &\mapsto -\omega_0+\omega_1. \end{aligned}$$

Lemma 4.2.2. *For any path γ from y to z all the iterated integrals $\{I_\gamma(y;\mathbf{w};z) : \mathbf{w} \in \{0,1\}^*\}$ form a shuffle algebra. Moreover, for every ϕ in* (4.33)*, we have*

$$I_{\phi(\gamma)}(\phi(y);\mathbf{w};\phi(z)) = I_\gamma(y;\phi^*\mathbf{w};z) \tag{4.34}$$

and for any word $\mathbf{w} \in \{0,1\}^*$ and $a \in \{0,1\}$

$$\partial_z I(y;a\mathbf{w};z) = \frac{1}{z-a} I(y;\mathbf{w};z), \quad \partial_y I(y;a\mathbf{w};z) = -\frac{1}{y-a} I(y;\mathbf{w};z).$$

Proof. These follow easily from the properties of iterated integrals. □

In the rest of this chapter let $\mathsf{X} = \{\mathtt{x}_0, \mathtt{x}_1\}$ and let X^* be the set of words over X. Denote by $\mathtt{e}$ the empty word and set $I_\gamma(y;\mathtt{e};z) = 1$. Write $I_\gamma(y;\mathtt{x}_{a_1}\dots\mathtt{x}_{a_n};z) = I_\gamma(y;a_1\dots a_n;z)$ for any $\mathtt{x}_{a_1}\dots\mathtt{x}_{a_n} \in \mathsf{X}^*$. We now define the generating series

$$S_{\mathsf{X}}(\gamma) = \sum_{\mathbf{w}\in\mathsf{X}^*} I_\gamma(y;\mathbf{w};z)\,\mathbf{w} \in \mathbb{C}\langle\langle \mathsf{X} \rangle\rangle. \tag{4.35}$$

By the properties of the iterated integrals one can prove

Lemma 4.2.3. *If $\gamma(0) = y$ and $\gamma(1) = z$ then*

$$\partial_z S_{\mathsf{X}}(\gamma) = \left(\frac{\mathtt{x}_0}{z} + \frac{\mathtt{x}_1}{z-1}\right) S_\gamma, \quad \partial_y S_{\mathsf{X}}(\gamma) = -S_\gamma \left(\frac{\mathtt{x}_0}{y} + \frac{\mathtt{x}_1}{y-1}\right),$$

Write $S(y,z)$ when the homotopy class of γ is clear from the context.

Proposition 4.2.4. *The following limits are well-defined*

$$F_0(z) := \lim_{\varepsilon\to 0^+} S_{\mathsf{X}}(\varepsilon,z)e^{\mathtt{x}_0\log\varepsilon}, \qquad F_0(y)^{-1} := \lim_{\varepsilon\to 0^+} e^{-\mathtt{x}_0\log\varepsilon} S_{\mathsf{X}}(y,\varepsilon),$$
$$F_1(z) := \lim_{\varepsilon\to 0} S_{\mathsf{X}}(1-\varepsilon,z)e^{\mathtt{x}_1\log\varepsilon}, \quad F_1(y)^{-1} := \lim_{\varepsilon\to 0^+} e^{-\mathtt{x}_1\log\varepsilon} S_{\mathsf{X}}(y,1-\varepsilon).$$

Proof. This follows immediately from Thm 4.1.1. □

For simplicity henceforth we will simply write $S_{\mathsf{X}}(0,z) = S_{\mathsf{X}}(z,0)^{-1} = F_0(z)$ and $S_{\mathsf{X}}(1,z) = S_{\mathsf{X}}(z,1)^{-1} = F_1(z)$. In particular we have

$$S_{\mathsf{X}}(0,1) = \Phi(\mathtt{x}_0,\mathtt{x}_1).$$

Remark 4.2.5. One can also prove the proposition directly the regularization procedure where all powers of $\log\varepsilon$ are assigned to be 0 in the limit $\varepsilon \to 0$. This is also equivalent to the regularization by tangential base points used by Deligne [153].

For simplicity, in what follows we will sometimes identify words in X^* with words in $\{0,1\}^*$ by identify $\mathtt{x}_a$ with $a \in \{0,1\}$. It should be clear which type we mean from the context.

Fix a path γ from 0 to z. If z is in the unit open disk $|z| < 1$ and not a negative real number we assume γ to be a straight line from 0 to z. For any $\mathbf{w} \in \{0,1\}^*$ (or $\mathbf{w} \in \mathsf{X}^*$) we define recursively

$$\mathrm{L}_{a\mathbf{w}}(z)(=\mathrm{L}_{\mathtt{x}_a\mathbf{w}}(z)) = \int_\gamma \frac{\mathrm{L}_{\mathbf{w}}(t)}{t-a}\mathrm{d}t, \quad a \in \{0,1\}, \tag{4.36}$$

with the initial value $\mathrm{L}_{\mathbf{w}}(0) = 0$ unless $\mathbf{w} = \{0\}^n (= \mathtt{x}_0^n)$ when $\mathrm{L}_{\mathbf{w}}(0) = \frac{\log^n z}{n!}$. This leads to the multi-valued functions on $\mathbb{C}^* \setminus \{0,1\}$

$$\mathrm{L}_{\mathbf{w}}(z) = I_\gamma(0;\mathbf{w};z).$$

So the generating series of the one-varable MPLs (in word form) is given by

$$\mathrm{L}_{\mathsf{X}}(z) := \sum_{\mathbf{w}\in\mathsf{X}^*} \mathrm{L}_{\mathbf{w}}(z)\mathbf{w}. \tag{4.37}$$

By Lemma 4.2.3 and Thm. 4.1.1 $\mathrm{L}_{\mathsf{X}}(z)$ is the unique solution to the KZ equation

$$\partial_z \mathrm{L}_{\mathsf{X}}(z) = \left(\frac{\mathtt{x}_0}{z} + \frac{\mathtt{x}_1}{z-1}\right)\mathrm{L}_{\mathsf{X}}(z) \tag{4.38}$$

which satisfies the asymptotic condition

$$\mathrm{L}_{\mathsf{X}}(z) = h_0(z)e^{\mathtt{x}_0 \log z} \quad \text{if } z \to 0, \tag{4.39}$$

where $h_0(z)$ is a holomorphic function at 0 taking values in $\mathbb{C}\langle\langle\mathsf{X}\rangle\rangle$ with $h(0) = 1$. Hence $\mathrm{L}_{\mathsf{X}}(z) = G_0(z)$ appearing in the Definition 4.1.2 for the Drinfeld associator.

The series $\mathrm{L}_{\mathsf{X}}(z)$ is a group-like (see [469, Cor. 3.5]) so that we have the shuffle product formula

$$\mathrm{L}_{\mathbf{u}\sqcup\!\sqcup\mathbf{v}}(z) = \mathrm{L}_{\mathbf{u}}(z)\,\mathrm{L}_{\mathbf{v}}(z) \quad \forall \mathbf{u},\mathbf{v} \in \mathsf{X}^*. \tag{4.40}$$

Lemma 4.2.6. *There exists a function $h_1(z)$ taking values in the series $\mathbb{C}\langle\langle\mathsf{X}\rangle\rangle$, which is holomorphic at $z = 1$ with $h_1(1) = 1$, such that*

$$\mathrm{L}_{\mathsf{X}}(z) = h_1(z)e^{\mathtt{x}_1 \log(1-z)}\Phi(\mathtt{x}_0,\mathtt{x}_1). \tag{4.41}$$

Proof. By definition of Drinfeld associator in Definition 4.1.2, we know $\Phi(\mathtt{x}_0,\mathtt{x}_1) = G_1(z)^{-1}G_0(z)$ where $G_0(z) = h_0 e^{\mathtt{x}_0 \log z}$ and $G_1(z) = h_1(z)e^{\mathtt{x}_1 \log(1-z)}$ for some $h_1(z)$ satisfying the condition for $h_1(z)$ stated in the lemma. Hence the lemma follows from Eq. (4.39) immediately. □

For $a = 0, 1$ let σ_a be the loop centered at a oriented clockwise. Let $M(\sigma_a)$ denote analytic continuation around σ_a. Then $M(\sigma_a)$ acts on the series $\mathrm{L}_{\mathsf{X}}(z)$ and its complex conjugate $\mathrm{L}(\bar{z})$, and commutes with multiplication, ∂_z, and $\partial_{\bar{z}}$.

Lemma 4.2.7. *The monodromy operators $M(\sigma_0)$, $M(\sigma_1)$ act as follows:*

$$\begin{aligned} M(\sigma_0)\,\mathrm{L}_{\mathsf{X}}(z) &= \mathrm{L}_{\mathsf{X}}(z)e^{2\pi i\mathsf{x}_0}, \\ M(\sigma_1)\,\mathrm{L}_{\mathsf{X}}(z) &= \mathrm{L}_{\mathsf{X}}(z)\Phi(\mathsf{x}_0,\mathsf{x}_1)^{-1}e^{2\pi i\mathsf{x}_1}\Phi(\mathsf{x}_0,\mathsf{x}_1). \end{aligned} \tag{4.42}$$

Proof. These follow from (4.39), (4.41) and the equations $M(\sigma_0)\log z = \log z + 2\pi i$, $M(\sigma_1)\log(1-z) = \log(1-z) + 2\pi i$. □

A word form single-variable MPLs is related to those defined by Eq. (3.19):

$$(-1)^r\,\mathrm{L}_{\{0\}^{n_1-1}1\cdots\{0\}^{n_r-1}1}(z) = \mathrm{L}_{n_1,\ldots,n_r}(z) = \sum_{k_1>\cdots>k_r\geq 1}\frac{z^{k_1}}{k_1^{n_1}\ldots k_r^{n_r}}. \tag{4.43}$$

Remark 4.2.8. It is unfortunate that $\mathrm{L}_1(z)$ could have two different meanings: the word form $\mathrm{L}_{\mathsf{x}_1}(z) = -\mathrm{Li}_1(z)$ or the single-variable MPL $L_s(z)$ when $s = 1$, *i.e.*, $\mathrm{Li}_1(z)$. So in this book we will always treat the subscript 1 in $\mathrm{L}_1(z)$ as a letter in $\{0, 1\}$, meaning $\mathrm{L}_1(z) = -\mathrm{Li}_1(z) = \log(1-z)$.

Lemma 4.2.9. *If* $\mathbf{w} = \{0\}^{s_1-1}1\cdots\{0\}^{s_d-1}1\{0\}^n$ *then*

$$\mathrm{L}_{\mathbf{w}}(z) = \sum_{\substack{\ell\geq 0, k_i\geq s_i\ (i\geq 1)\\ |\mathbf{k}|+\ell=|\mathbf{s}|+n}}(-1)^{\ell+n+d}\prod_{i=1}^{d}\binom{k_i-1}{s_i-1}\frac{\log^\ell z}{\ell!}\,\mathrm{L}_{k_1,\ldots,k_r}(z). \tag{4.44}$$

Proof. This directly follows from Eq. (4.43) and Thm. 4.1.1 in which $f_0(z)$ is explicitly given by Eqs. (4.2) and (4.3). □

Example 4.2.10. By the lemma we get (also see Exercise 4.19) in the word form

$$\begin{aligned} &\mathrm{L}_0(z) = \log z, && \mathrm{L}_1(z) = -\mathrm{Li}_1(z) = \log(1-z)\ \text{ (see Remark 4.2.8)},\\ &\mathrm{L}_{01}(z) = -\mathrm{Li}_2(z), && \mathrm{L}_{10}(z) = \mathrm{Li}_2(z) + \log z\log(1-z). \end{aligned}$$

4.3 Single-valued single-variable MPLs

In this section we construct the single-valued version of the single-variable MPLs defined by Eq. (3.19). This will be crucial in the study of several classes of Feynman integrals in Chap. 15.

We begin with a technical lemma. Let $\mathsf{X} = \{\mathtt{x}_0, \mathtt{x}_1\}$ be the same as in the proceeding section. Let $\mathsf{MZ}_{\mathbb{Z}}$ be the ring over $\mathbb{Z}$ generated by MZVs and let $\mathsf{MZ}_{\mathbb{Z}}\langle\langle\mathsf{X}\rangle\rangle$ be the Hopf algebra with coefficients in $\mathsf{MZ}_{\mathbb{Z}}$ so that both $\mathtt{x}_0$ and $\mathtt{x}_1$ are primitive elements. Further, we define $\leftarrow\colon \mathsf{X}^* \to \mathsf{X}^*$ to be the anti-homomorphism such that $\overleftarrow{\mathbf{w}}$ is the order reversal of $\mathbf{w}$. Note that $\overleftarrow{\mathtt{x}_a} = \mathtt{x}_a$ for $a = 0, 1$.

Lemma 4.3.1. *There exists* $\mathtt{y}_1 \in \mathsf{MZ}_{\mathbb{Z}}\langle\langle\mathsf{X}\rangle\rangle$ *such that*

$$\overleftarrow{\Phi}(\mathtt{x}_0, \mathtt{y}_1)\mathtt{y}_1\overleftarrow{\Phi}(\mathtt{x}_0, \mathtt{y}_1)^{-1} = \Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1}\mathtt{x}_1\Phi(\mathtt{x}_0, \mathtt{x}_1) \tag{4.45}$$

Moreover, $\mathtt{y}_1 \equiv \mathtt{x}_1$ *modulo words with at least two* $\mathtt{x}_1$*'s (depth two) or modulo words of weight* ≥ 4*.*

Proof. Let $\Phi(\mathtt{x}_0, \mathtt{x}_1)$ be the Drinfeld associator. Define

$$H_{\mathsf{X}}(t) := \Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1}e^{t\mathtt{x}_1}\Phi(\mathtt{x}_0, \mathtt{x}_1) \in \mathsf{MZ}_{\mathbb{Z}}[t]\langle\langle\mathsf{X}\rangle\rangle$$

because $\Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1} = \Phi(\mathtt{x}_1, \mathtt{x}_0)$. Moreover, by Thm. B.22(i) and the famous Baker-Campbell-Hausdorff formula $H_{\mathsf{X}}(t) = e^{h(t)}$ is a Lie exponential. Since $h(0) = 0$ we have $H'_{\mathsf{X}}(0) = h'(0) \in \mathsf{MZ}_{\mathbb{Z}}\langle\langle\mathsf{X}\rangle\rangle$. This implies that

$$F(\mathtt{x}_0, \mathtt{x}_1) := \Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1}\mathtt{x}_1\Phi(\mathtt{x}_0, \mathtt{x}_1) - \mathtt{x}_1 \in \mathsf{MZ}_{\mathbb{Z}}\langle\langle\mathsf{X}\rangle\rangle. \tag{4.46}$$

By straight-forward computation (see Exercise 4.17) we have $F \equiv 0$ modulo words of depth two and $F \equiv \zeta(2)[[\mathtt{x}_0, \mathtt{x}_1], \mathtt{x}_1]$ modulo words of weight four. Now we can solve Eq. (4.45) recursively by setting

$$\mathtt{y}_1^{(0)} = \mathtt{x}_1, \quad \mathtt{y}_1^{(k+1)} = \mathtt{x}_1 + F(\mathtt{x}_0, \mathtt{x}_1) + F(-\mathtt{x}_0, -\mathtt{y}_1^{(k)}).$$

Because F has lowest word weight three, the recursion converges by weight filtration. Let $\mathtt{y}_1 = \lim_{k\to\infty} \mathtt{y}_1^{(k)}$ then

$$\mathtt{y}_1 = \mathtt{x}_1 + F(\mathtt{x}_0, \mathtt{x}_1) + F(-\mathtt{x}_0, -\mathtt{y}_1) \tag{4.47}$$

which yields Eq. (4.45). Here we need the fact that $\overleftarrow{\Phi}(-\mathtt{x}_0, -\mathtt{x}_1) = \Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1}$ by Thm. B.22(ii). By induction every $\mathtt{y}_1^{(k)}$ is in $\mathsf{MZ}_{\mathbb{Z}}\langle\langle\mathsf{X}\rangle\rangle$ so must be $\mathtt{y}_1$. This completes the proof of the lemma. □

Let ∂_z and $\partial_{\bar{z}}$ be the holomorphic and antiholomorphic differentiation and set

$$\mathsf{O} := \mathbb{Q}\left[z, \frac{1}{z}, \frac{1}{z-1}\right], \quad \overline{\mathsf{O}} := \mathbb{Q}\left[\bar{z}, \frac{1}{\bar{z}}, \frac{1}{\bar{z}-1}\right].$$

Theorem 4.3.2. *For each* $\mathbf{w} \in \mathsf{X}^*$ *there is a single-valued function* $\mathscr{L}_{\mathbf{w}}(z)$ *on* $\mathbb{C}^*\backslash\{0,1\}$*, called the single-valued single-variable multiple polylog (SVMPL for abbreviation), such that*

(i) $\mathscr{L}_{\mathbf{w}}(z)$ *is a linear combination of* $\mathrm{L}_{\mathbf{u}}(\bar{z})\,\mathrm{L}_{\mathbf{v}}(z)$*, where* $\mathbf{u},\mathbf{v} \in \mathsf{X}^*$*;*

(ii) $\mathscr{L}_{\mathbf{e}}(z) = 1$, $\mathscr{L}_{\{0\}^n}(z) = \dfrac{\log^n(z\bar{z})}{n!}$ *for all* $n \geq 1$*, and* $\lim\limits_{z\to 0}\mathscr{L}_{\mathbf{w}}(z) = 0$ *for all* $\mathbf{w} \neq \{0\}^*$*;*

(iii) *they satisfy the recursive relations*

$$\partial_z\,\mathscr{L}_{a\mathbf{w}}(z) = \frac{\mathscr{L}_{\mathbf{w}}(z)}{z-a}, \qquad \text{for } a \in \{0,1\}, \tag{4.48}$$

(iv) *they satisfy the shuffle relations, and are linearly independent over* $\mathsf{O}\overline{\mathsf{O}}$*;*

(v) *every linear combination of the functions* $\mathrm{L}_{\mathbf{u}}(z)\,\mathrm{L}_{\mathbf{v}}(\bar{z})$*, where* $\mathbf{u},\mathbf{v} \in \mathsf{X}^*$*, which is single-valued, can be written as a unique linear combination of the functions* $\mathscr{L}_{\mathbf{w}}(z)$*.*

Proof. Let $\mathsf{Y} = \{\mathtt{y}_0, \mathtt{y}_1\}$ and Y^* the set of words on Y. Denote by $\overleftarrow{\mathbf{w}}$ the reversal of $\mathbf{w} \in \mathsf{Y}^*$. By Lemma 4.3.1 we see that

$$\mathtt{y}_0 = \mathtt{x}_0, \quad \overleftarrow{\Phi}(\mathtt{y}_0,\mathtt{y}_1)e^{2\pi i\mathtt{y}_1}\overleftarrow{\Phi}(\mathtt{y}_0,\mathtt{y}_1)^{-1} = \Phi(\mathtt{x}_0,\mathtt{x}_1)^{-1}e^{2\pi i\mathtt{x}_1}\Phi(\mathtt{x}_0,\mathtt{x}_1) \tag{4.49}$$

has a solution $\mathtt{y}_1 = \mathtt{x}_1 + \epsilon$ where $\epsilon \neq 0$ is a series of words with length ≥ 4. Define $\psi : \mathsf{Y}^* \to \mathsf{X}^*$, $\mathtt{y}_0 \mapsto \mathtt{x}_0$ and $\mathtt{y}_1 \mapsto \mathtt{x}_1 + \epsilon$. Set

$$\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z}) = \sum_{\mathbf{w}\in\mathsf{Y}^*} \mathrm{L}_{\psi(\mathbf{w})}(\bar{z})\overleftarrow{\mathbf{w}} = 1 + \mathrm{L}_{\mathtt{x}_0}(\bar{z})\mathtt{x}_0 + \mathrm{L}_{\mathtt{x}_1}(\bar{z})(\mathtt{x}_1+\epsilon) + \cdots.$$

We now prove that

$$\mathscr{L}_{\mathsf{X}}(z) = \sum_{\mathbf{w}\in\mathsf{X}^*}\mathscr{L}_{\mathbf{w}}(z) := \mathrm{L}_{\mathsf{X}}(z)\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z}) \tag{4.50}$$

is single-valued. In fact, by monodromy computation

$$\begin{aligned} M(\sigma_0)\,\mathscr{L}_{\mathsf{X}}(z) &= \mathrm{L}_{\mathsf{X}}(z)e^{2\pi i\mathtt{x}_0}e^{-2\pi i\mathtt{x}_0}\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z}) = \mathscr{L}_{\mathsf{X}}(z),\\ M(\sigma_1)\,\mathscr{L}_{\mathsf{X}}(z) &= \mathrm{L}_{\mathsf{X}}(z)\Phi(\mathtt{x}_0,\mathtt{x}_1)^{-1}e^{2\pi i\mathtt{x}_1}\Phi(\mathtt{x}_0,\mathtt{x}_1)\times\\ &\qquad\times\overleftarrow{\Phi}(\mathtt{y}_0,\mathtt{y}_1)^{-1}e^{-2\pi i\mathtt{y}_1}\overleftarrow{\Phi}(\mathtt{y}_0,\mathtt{y}_1)\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z}) = \mathscr{L}_{\mathsf{X}}(z) \end{aligned}$$

by Eq. (4.49) and the fact that the word reversal map $\leftarrow$ is an anti-homomorphism.

(i) is clear from the definition (4.50) of $\mathscr{L}_{\mathbf{w}}(z)$.

(ii) By setting $\mathtt{x}_1 = 0$ we have

$$\begin{aligned} \mathscr{L}_{\mathtt{x}_0,0}(z) = \mathrm{L}_{\mathsf{X}}(0)\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(0) &= \mathrm{L}_{\mathtt{x}_0,0}(z)\overleftarrow{\mathrm{L}}_{\mathtt{x}_0,0}(\bar{z})\\ &= \mathrm{L}_{\mathtt{x}_0,0}(z)\,\mathrm{L}_{\mathtt{x}_0,0}(\bar{z}) = e^{\mathtt{x}_0\log z}e^{\mathtt{x}_0\log\bar{z}} = e^{\mathtt{x}_0\log z\bar{z}}. \end{aligned}$$

(iii) is equivalent to showing that

$$\partial_z \mathscr{L}_{\mathsf{X}}(z) = \left(\frac{\mathtt{x}_0}{z} + \frac{\mathtt{x}_1}{z-1}\right) \mathscr{L}_{\mathsf{X}}(z)$$

which follows from Eq. (4.38) immediately since $\partial_z \overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z}) = 0$.

(iv) Both $\mathrm{L}_{\mathsf{X}}(z)$ and $\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z})$ are group-like so is their product $\mathscr{L}_{\mathsf{X}}(z)$. Hence the functions $\mathscr{L}_{\mathbf{w}}(z)$ satisfy the shuffle relations. The second half of (iv) and (v) are not needed in the rest of the book so we refer the interested reader to Brown's original manuscripts [99, 102]. □

Example 4.3.3. At weight one and two we obtain the following SVMPLs by using Example 4.2.10:

$$\begin{aligned}
\mathscr{L}_0(z) &= \mathrm{L}_0(z) + \mathrm{L}_0(\bar{z}) = \log z + \log \bar{z} = 2\log|z|,\\
\mathscr{L}_1(z) &= \mathrm{L}_1(z) + \mathrm{L}_1(\bar{z}) = \log(1-z) + \log(1-\bar{z}) = 2\log|1-z|,\\
\mathscr{L}_{00}(z) &= \frac{1}{2}\mathscr{L}_0(z)^2 = \frac{1}{2}\log^2(z\bar{z}) = 2\log^2|z|,\\
\mathscr{L}_{01}(z) &= \mathrm{L}_{01}(z) + \mathrm{L}_0(z)\,\mathrm{L}_1(\bar{z}) + \mathrm{L}_{10}(\bar{z}) = 2\log|z|\log|1-z| - 2i\mathscr{L}i_2(z),\\
\mathscr{L}_{10}(z) &= \mathrm{L}_{10}(z) + \mathrm{L}_1(z)\,\mathrm{L}_0(\bar{z}) + \mathrm{L}_{01}(\bar{z}) = 2\log|z|\log|1-z| + 2i\mathscr{L}i_2(z),\\
\mathscr{L}_{11}(z) &= \frac{1}{2}\mathscr{L}_1(z)^2 \;=\; \frac{1}{2}\log(1-z)(1-\bar{z}) = 2\log^2|1-z|,
\end{aligned}$$

where $\mathscr{L}i_2(z)$ is the Bloch-Wigner dilog (see Eq. (2.46) and also Exercise 4.20). So there is only one algebraically independent SVMPL of weight two, which can also be seen by the shuffle relation $\mathscr{L}_{01} + \mathscr{L}_{10} = \mathscr{L}_0\,\mathscr{L}_1$.

In higher weights, due to the appearance of $\mathtt{y}_1$, the expression of $\mathscr{L}_{\mathbf{w}}$ often involve MZVs. Let the total weight of a product of MZVs with holomorphic and antiholomorphic MPLs be the sum of the individual weights.

Proposition 4.3.4. *$\mathscr{L}_{\mathbf{w}}$ is homogeneous of total weight $|\mathbf{w}|$.*

Proof. Assign weight -1 to $\mathtt{x}_0$ and $\mathtt{x}_1$. By its construction in the proof of Lemma 4.3.1 $\mathtt{y}_1$ has weight -1, the same as $\mathtt{x}_1$. Then $\mathrm{L}_{\mathsf{X}}(z)$ and $\overleftarrow{\mathrm{L}}_{\mathsf{Y}}(\bar{z})$ both have weight 0, so does $\mathscr{L}_{\mathsf{X}}(z)$, which yields the proposition. □

The shuffle-algebra over $\mathbb{C}$ spanned by SVMPLs graded by the total weight is denoted by

$$\mathsf{Ls} = \bigoplus_{n\geq 0} \mathsf{Ls}_n := \mathbb{C}\langle \mathscr{L}_{\mathbf{w}} : \mathbf{w} \in \mathsf{X}^* \rangle.$$

For example, $\mathsf{Ls}_1 = \mathbb{C}\langle \log|z|, \log|1-z|\rangle$ and $\mathsf{Ls}_2 = \mathbb{C}\langle \log|z|\log|1-z|, \log^2|z|, \log^2|1-z|, \mathscr{L}i_2(z)\rangle$. It turns out, however, the most natural coefficient ring for SVMPLs is much smaller than $\mathbb{C}$. We will consider this in the next section.

4.4 Single-valued MZVs and integration of SVMPLs

The SVMPLs are not defined at 0, 1 and ∞. We now extend SVMPLs to these points by applying regularizations.

Definition 4.4.1. For a single-valued function on $\mathbb{C}\setminus\{0,1\}$ written in the form $f(z) = \sum_{j=0}^{n} f_j(z)\log^j|z-a|$ ($a = 0$ or 1) where $f_j(z)$ are all single-valued function continuous and finite at a we define the regularized value of $f(z)$ at a to be $\mathrm{reg}_{z=a} f(z) = f_0(a)$. At ∞, if there is $M > 0$ such that $f(z) = \sum_{j=0}^{n} f_j(z)\log^j|z|$ where $f_j(z)$ are all single-valued function continuous for all $|z| > M$ and remain finite when $z \to \infty$ then we define the regularized value of $f(z)$ at ∞ to be $\mathrm{reg}_{z=\infty} f(z) = \lim_{z\to\infty} f_0(z)$.

Lemma 4.4.2. *For any single-valued functions f and g defined near $a \in \mathbb{C}\cup\infty$, we have*

$$\mathrm{reg}_{z=a}\big(f(z)g(z)\big) = \mathrm{reg}_{z=a} f(z)\,\mathrm{reg}_{z=a} g(z),$$

provided all the regularized values are defined.

Proof. This is clear from the definition. □

By the definition we see that $\mathscr{L}_{\mathbf{w}}(0) := \mathrm{reg}_{z=0}\mathscr{L}_{\mathbf{w}}(z) = 0$ for all non-empty word $\mathbf{w} \in \mathsf{X}^*$.

Definition 4.4.3. We set $\mathscr{L}_{\mathbf{w}}(1) := \mathrm{reg}_{z=1}\mathscr{L}_{\mathbf{w}}(z)$ and call it a *single-valued MZV* (SVMZV).

Lemma 4.4.4. *Let $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ be the set of all $\mathbb{Z}$-linear combinations of SVMZVs. Then $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ is a sub-ring of $\mathsf{MZ}_{\mathbb{Z}}$, namely, $\mathscr{L}_{\mathbf{u}}(1)\mathscr{L}_{\mathbf{v}}(1) \in \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ for all $\mathbf{u},\mathbf{v} \in \mathsf{X}^*$.*

Proof. Without loss of generality, let $\mathbf{u},\mathbf{v} \in \mathsf{X}^*$ be two non-empty words. Then by Lemma 4.4.2

$$\begin{aligned}\mathscr{L}_{\mathbf{u}}(1)\mathscr{L}_{\mathbf{v}}(1) &= \mathrm{reg}_{z=1}\mathscr{L}_{\mathbf{u}}(z)\cdot\mathrm{reg}_{z=1}\mathscr{L}_{\mathbf{v}}(z)\\ &= \mathrm{reg}_{z=1}\big(\mathscr{L}_{\mathbf{u}}(z)\mathscr{L}_{\mathbf{v}}(z)\big) = \mathrm{reg}_{z=1}\mathscr{L}_{\mathbf{u}\sqcup\!\sqcup\mathbf{v}}(z) = \mathscr{L}_{\mathbf{u}\sqcup\!\sqcup\mathbf{v}}(1)\end{aligned}$$

where we have extended $\mathscr{L}$ $\mathbb{Z}$-linearly over $\mathbb{Z}\langle X\rangle$. So the set $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ is closed under multiplication, as desired. □

Example 4.4.5. Maple computation shows that up to weight ten $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Q}} := \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}} \otimes \mathbb{Q}$ is generated by $\zeta(3)$, $\zeta(5)$, $\zeta(7)$, $\zeta(9)$. At weight 11 we have two generators: $\zeta(11)$ and

$$g_{533} = \zeta(5,3,3) - \frac{4}{7}\zeta(5)\zeta(2)^3 + \frac{6}{5}\zeta(7)\zeta(2)^2 + 45\zeta(9)\zeta(2).$$

More precisely, up to weight 11 the ring $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ is spanned by $2\zeta(3)$, $\zeta(5)$, $2\zeta(3)^2$, $\frac{1}{8}\zeta(7)$, $\zeta(3)\zeta(5)$, $\frac{1}{72}\zeta(9)$, $\frac{4}{3}\zeta(3)^3$, $\frac{1}{8}\zeta(3)\zeta(7)$, $\frac{1}{2}\zeta(5)^2$, $\frac{1}{384}\zeta(11)$, $\frac{1}{5}(g_{533}+\zeta(11)) + \frac{3}{2}\zeta(3)^2\zeta(5)$, and $2\zeta(3)^2\zeta(5)$.

Let $\mathsf{Ls}^{\mathrm{sv}}$ be the shuffle-algebra over $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ spanned by SVMPLs. Recall that $\mathsf{O} = \mathbb{Q}\left[z, \frac{1}{z}, \frac{1}{z-1}\right]$ and $\overline{\mathsf{O}} = \mathbb{Q}\left[\bar{z}, \frac{1}{\bar{z}}, \frac{1}{\bar{z}-1}\right]$. Define two bi-differential algebras

$$\mathsf{A} := \mathsf{O}\overline{\mathsf{O}}\mathsf{Ls}, \quad \mathsf{A}^{\mathrm{sv}} := \mathsf{O}\overline{\mathsf{O}}\mathsf{Ls}^{\mathrm{sv}}. \tag{4.51}$$

Let $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z},n}(\mathbb{Z})$, $\mathsf{Ls}^{\mathrm{sv}}_n$ and $\mathsf{A}^{\mathrm{sv}}_n$ denote the subspaces of total weight n. Due to difficulty in transcendental number theory we do not know if these subspaces are $\mathbb{Q}$-linearly independent.

Remark 4.4.6. The only difference between A (resp. Ls) and A^{sv} (resp. $\mathsf{Ls}^{\mathrm{sv}}$) is the coefficients for the former are in $\mathbb{C}$ while those for the latter are in $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$.

By the definition, A^{sv} is closed under ∂_z by Thm. 4.3.2(iii). For $\partial_{\bar{z}}$, by the definition (4.50) we see that

$$\partial_{\bar{z}}\, \mathscr{L}_{\mathsf{X}}(z) = \mathscr{L}_{\mathsf{X}}\left(\frac{\mathtt{x}_0}{\bar{z}} + \frac{\mathtt{y}_1}{\bar{z}-1}\right)$$

which yields

$$\begin{aligned}\partial_{\bar{z}}\, \mathscr{L}_{\mathbf{w}a}(z) &= \frac{\delta_{a,0}}{\bar{z}}\, \mathscr{L}_{\mathbf{w}}(z) + \frac{1}{\bar{z}-1} \sum_{\mathbf{w}a=\mathbf{uv}} \mathscr{L}_{\mathbf{u}}(z)\mathtt{y}_1[\mathbf{v}] \\ &= \frac{\mathscr{L}_{\mathbf{w}}(z)}{\bar{z}-a} + \frac{1}{\bar{z}-1} \sum_{\mathbf{w}a=\mathbf{uv}} \mathscr{L}_{\mathbf{u}}(z)(\mathtt{y}_1[\mathbf{v}] - \delta_{\mathtt{x}_1,\mathbf{v}}), \end{aligned} \tag{4.52}$$

where $\mathtt{y}_1[\mathbf{v}]$ is the coefficient of the word $\mathbf{v}$ in the series $\mathtt{y}_1$.

Definition 4.4.7. For $P \in \mathsf{A}$ let $\int_0 P(z)\mathrm{d}z$ be the unique function $F \in \mathsf{A}$, if it exists, such that $\partial_z F(z) = P(z)$ with $\int_0$ denoting the regularized value $\mathrm{reg}_{z=0} F(z) = 0$. Similarly, $\int_0 P(z)\mathrm{d}\bar{z}$ is the unique function $G(z) \in \mathsf{A}$, if it exists, such that $\partial_{\bar{z}} G(z) = P(z)$. For $b \in \{1, \infty\}$ we define $\int_b P(z)\mathrm{d}z = F(z) - F(b)$ and $\int_b P(z)\mathrm{d}\bar{z} = G(z) - G(b)$, where $F(b)$ and $G(b)$ are regularized values.

Lemma 4.4.8. *We have a well-defined integral operator on* A*,*

$$\int_0 \mathrm{d}z : \mathsf{A} \longrightarrow \mathsf{A}, \quad F(z) \longmapsto \int_0 F(z)\mathrm{d}z.$$

Moreover, $\int_0 \mathrm{d}z$ *maps* A^{sv} *into* A^{sv}*.*

Proof. For the first sentence, we only need to show that the images of the operators lie in A. First note that for $a \in \{0,1\}$ and $n \neq -1$

$$\int_0 \frac{\mathscr{L}_{\mathsf{e}}(z)}{z-a}\mathrm{d}z = \mathscr{L}_1(z-a), \quad \int_0 \mathscr{L}_{\mathsf{e}}(z)(z-a)^n \mathrm{d}z = \frac{(z-a)^{n+1}}{n+1}.$$

both of which are in A^{sv}. For general $\mathbf{w}$, by Thm. 4.3.2(iii) we have

$$\int_0 \frac{\mathscr{L}_{\mathbf{w}}(z)}{z-a}\mathrm{d}z = \mathscr{L}_{a\mathbf{w}}(z) \in \mathsf{A}^{\mathrm{sv}}.$$

For any $n \in \mathbb{Z}$, $n \neq -1$, we have

$$\int_0 \mathscr{L}_{\mathbf{w}}(z)(z-a)^n \mathrm{d}z = \mathscr{L}_{\mathbf{w}}(z)\frac{(z-a)^{n+1}}{n+1} - \int_0 \mathscr{L}_{\mathbf{w}'}(z)\frac{(z-a)^{n+1}}{n+1}\mathrm{d}z \in \mathsf{A}$$

by induction on $|\mathbf{w}|$, where $\mathbf{w} = a\mathbf{w}'$. The claim for stability of A^{sv} under $\int_0 \mathrm{d}z$ follows from the same argument. □

Proposition 4.4.9. *Let* $\mathfrak{S}_3$ *be the group of Möbius transformations permuting* $\{0,1,\infty\}$*. Then* $\mathsf{Ls}_n^{\mathrm{sv}}$ *is stable under the* $\mathfrak{S}_3$*. The integral operator* $\int_b \mathrm{d}z/(z-a)$ *(*$a \in \{0,1\}, b \in \{0,1,\infty\}$*) satisfies*

$$\int_b \mathrm{d}z/(z-a) : \mathsf{Ls}_n^{\mathrm{sv}} \hookrightarrow \mathsf{Ls}_{n+1}^{\mathrm{sv}}. \tag{4.53}$$

Furthermore,

$$\partial_z \mathsf{Ls}_n^{\mathrm{sv}} = \frac{\mathsf{Ls}_{n-1}^{\mathrm{sv}}}{z} + \frac{\mathsf{Ls}_{n-1}^{\mathrm{sv}}}{z-1}. \tag{4.54}$$

Proof. First, Eq. (4.54) follows easily from Thm. 4.3.2(iii). The stability of $\mathsf{Ls}^{\mathrm{sv}}$ under $z \to \frac{z}{z-1}$ and under $\int_0 \mathrm{d}z/z$, $\int_0 \mathrm{d}z/(z-1)$ follows from straightforward computations (see Exercises 4.25 and Lemma 4.4.8). If $|\mathbf{w}| = n$ and $b = 1$ or ∞ we have

$$\int_b \mathscr{L}_{\mathbf{w}}(z)\mathrm{d}z/(z-a) = \int_0 \mathscr{L}_{\mathbf{w}}(z)\mathrm{d}z/(z-a) + \mathscr{L}_{a\mathbf{w}}(b) \in \mathsf{Ls}_{n+1}^{\mathrm{sv}}.$$

This proves Eq. (4.53).

Since the group $\mathfrak{S}_3$ is generated by $z \to \frac{z}{z-1}$ and $z \to 1-z$ we now only need to prove the stability of $\mathsf{Ls}^{\mathrm{sv}}$ under $z \to 1-z$. We proceed by

an induction on the total weight. In weight zero this is trivial. Assume $P \in \mathsf{Ls}_n^{\mathrm{sv}}$. Then by Eq. (4.54),

$$P(z) = \int_1 \left(\frac{Q_0(z)}{z} + \frac{Q_1(z)}{z-1} \right) \mathrm{d}z + P(1)$$

where $Q_0, Q_1 \in \mathsf{Ls}_{n-1}^{\mathrm{sv}}$. Since $Q_0(1-z), Q_1(1-z) \in \mathsf{Ls}_{n-1}^{\mathrm{sv}}$ by induction we finally get

$$P(1-z) = \int_0 \left(\frac{Q_0(1-z)}{z-1} + \frac{Q_1(1-z)}{z} \right) \mathrm{d}z + P(1) \in \mathsf{Ls}_n^{\mathrm{sv}}$$

by Eq. (4.53). □

Proposition 4.4.10. *The series* $\mathtt{y}_1 \in \mathsf{MZ}_{\mathbb{Z}}^{\mathrm{sv}} \langle\langle \mathsf{X} \rangle\rangle$.

Proof. We show that $\mathtt{y}_1 = \sum_{\mathbf{w} \in \mathsf{X}^*} \mathtt{y}_1[\mathbf{w}] \in \mathsf{MZ}_{\mathbb{Z}}^{\mathrm{sv}} \langle\langle \mathsf{X} \rangle\rangle$ by induction on the weight of $\mathbf{w}$. Suppose $\mathbf{w}$ has weight n. By (4.52), for any word $\mathbf{u}$, $\lim_{z \to 0} \bar{z} \partial_{\bar{z}} \mathscr{L}_{\mathbf{u}a}(z) = \delta_{a,0} \mathscr{L}_{\mathbf{u}}(0) \in \mathsf{MZ}_{\mathbb{Z}}^{\mathrm{sv}}$. By stability of $\mathsf{Ls}^{\mathrm{sv}}$ under $z \to 1-z$ we obtain

$$\lim_{z \to 1} (\bar{z} - 1) \partial_{\bar{z}} \mathscr{L}_{\mathbf{w}}(z) = \lim_{z \to 0} \bar{z} \partial_{\bar{z}} \mathscr{L}_{\mathbf{w}}(1-z) \in \mathsf{MZ}_{\mathbb{Z}}^{\mathrm{sv}}.$$

On the other hand from (4.52) we obtain

$$\lim_{z \to 1} (\bar{z} - 1) \partial_{\bar{z}} \mathscr{L}_{\mathbf{w}}(z) = \sum_{\mathbf{w} = \mathbf{uv}} \mathscr{L}_{\mathbf{u}}(1) \mathtt{y}_1[\mathbf{v}] = \mathtt{y}_1[\mathbf{w}] + \sum_{\substack{\mathbf{w} = \mathbf{uv} \\ |\mathbf{u}| < |\mathbf{w}|}} \mathscr{L}_{\mathbf{u}}(1) \mathtt{y}_1[\mathbf{v}].$$

This proves $\mathtt{y}_1 \in \mathsf{MZ}_{\mathbb{Z}}^{\mathrm{sv}} \langle\langle \mathsf{X} \rangle\rangle$ by induction. □

Lemma 4.4.11. *We have a well-defined integral operator on* A*,*

$$\int_0 \mathrm{d}\bar{z} : \mathsf{A} \longrightarrow \mathsf{A}, \quad F(z) \longmapsto \int_0 F(z) \mathrm{d}\bar{z}.$$

Moreover, $\int_0 \mathrm{d}\bar{z}$ *maps* A^{sv} *into* A^{sv}*.*

Proof. As in the proof of Lemma 4.4.8 we first show that the images of $\int_0 \mathrm{d}\bar{z}$ lie in A. Note that for $a \in \{0, 1\}$ and $n \neq -1$

$$\int_0 \frac{\mathscr{L}_{\mathtt{e}}(z)}{\bar{z} - a} \mathrm{d}\bar{z} = \mathscr{L}_1(z - a), \quad \int_0 \mathscr{L}_{\mathtt{e}}(z) (\bar{z} - a)^n \mathrm{d}\bar{z} = \frac{(\bar{z} - a)^{n+1}}{n+1}.$$

both of which are in A^{sv}. For general $\mathbf{w}$, note that all the terms on the right-hand side of Eq. (4.52) involving $\mathbf{u}$ vanish unless $|\mathbf{u}| \leq |\mathbf{w}a| - 4$ according to the property that $\mathtt{y}_1 \equiv \mathtt{x}_1$ modulo weight 4 words. Thus $\int_0 \mathscr{L}_{\mathbf{w}}(z) (\bar{z} - a)^{-1} \mathrm{d}\bar{z} \in \mathsf{A}^{\mathrm{sv}}$ by Eq. (4.52) and Prop. 4.4.10. If $n \neq -1 \in \mathbb{Z}$ then by an induction on $|\mathbf{w}|$ we can show that $\int_0 \mathscr{L}_{\mathbf{w}}(z) (\bar{z} - a)^n \mathrm{d}\bar{z} \in \mathsf{A}$ for any by integration by parts using Eq. (4.52).

The claim for stability of A^{sv} under $\int_0 \mathrm{d}z$ follows from the same argument by using Prop. 4.4.10. □

Proposition 4.4.12. *The integral operator* $\int_b \mathrm{d}\bar{z}/(\bar{z}-a)$ *(*$a \in \{0,1\}, b \in \{0,1,\infty\}$*) satisfies*

$$\int_b \mathrm{d}\bar{z}/(\bar{z}-a) : \mathsf{Ls}^{\mathrm{sv}}_n \hookrightarrow \mathsf{Ls}^{\mathrm{sv}}_{n+1}. \tag{4.55}$$

Furthermore,

$$\partial_{\bar{z}} \mathsf{Ls}^{\mathrm{sv}}_n = \frac{\mathsf{Ls}^{\mathrm{sv}}_{n-1}}{\bar{z}} + \frac{\mathsf{Ls}^{\mathrm{sv}}_{n-1}}{\bar{z}-1}. \tag{4.56}$$

Proof. Eq. (4.56) follows readily from Eq. (4.52) by using Prop. 4.4.10. Eq. (4.52) also yields the inductively defined antiholomorphic integral

$$\int_0 \frac{\mathscr{L}_{\mathbf{w}}(z)}{\bar{z}-a}\mathrm{d}\bar{z} = \mathscr{L}_{\mathbf{w}a}(z) - \sum_{\mathbf{w}a=\mathbf{uv}} (\mathsf{y}_1[\mathbf{u}] - \mathsf{x}_1[\mathbf{u}]) \int_0 \frac{\mathscr{L}_{\mathbf{v}}(z)}{\bar{z}-1}\mathrm{d}\bar{z}, \tag{4.57}$$

where $a \in \{0,1\}$. This implies Eq. (4.55) by an induction on the weight. We leave the details to the interested reader. □

Proposition 4.4.13. *For all* $n \in \mathbb{N}$*, we have* $\mathsf{Ls}^{\mathrm{sv}}_n(0) = \mathsf{Ls}^{\mathrm{sv}}_n(1) = \mathsf{Ls}^{\mathrm{sv}}_n(\infty) = \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z},n}(\mathbb{Z}) \subset \mathsf{MZ}_{\mathbb{Z},n}$.

Proof. By definition $\mathsf{Ls}^{\mathrm{sv}}(0) = \mathscr{L}_{\mathbf{e}}(0)\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}} = \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$. The second equality $\mathsf{Ls}^{\mathrm{sv}}(1) = \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ follows from Lemma 4.4.4. Finally, $\mathsf{Ls}^{\mathrm{sv}}(\infty) = \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ since $\mathsf{Ls}^{\mathrm{sv}}(\infty) = \mathsf{Ls}^{\mathrm{sv}}(1)$ (see Exercise 4.25). By Lemma 4.3.1 $\mathsf{y}_1 \in \mathsf{MZ}_{\mathbb{Z}}\langle\langle \mathsf{X} \rangle\rangle$ hence by the definition $\sum_{\mathbf{w}\in\mathsf{X}^*} \mathscr{L}_{\mathbf{w}}(z) = \mathrm{L}_{\mathbf{x}_0,\mathbf{x}_1}(z)\overleftarrow{\mathrm{L}}_{\mathbf{x}_0,\mathsf{y}_1}(\bar{z})$ we finally arrived at $\mathsf{Ls}^{\mathrm{sv}}(1) \subset \mathsf{MZ}_{\mathbb{Z}}$. □

4.5 Generating series of Hoffman elements

According to Hoffman's Basis Conjecture 3.6.1 partly proved by Brown [104] all the MZVs can be expressed as a $\mathbb{Q}$-linear combination of MZVs with arguments equal to 2 or 3. If exactly k of 3's appear in such an MZV then it is called a *k-Hoffman element*. It is clear that the word representation of a convergent 0-Hoffman element is an alternating word containing neither x_0^2 nor x_1^2. These reduce to even powers of π according to Cor. 5.5.4.

The word of a 1-Hoffman element contains exactly one subword x_0^2 but no subword x_1^2. The exact evaluation of such MZVs is given by Thm. 5.6.4.

Definition 4.5.1. Let $\mathsf{X} = \{\mathsf{x}_0, \mathsf{x}_1\}$. For $a = 0$ or 1, denote by H^a the set of words $\mathbf{w} \in \mathsf{X}^*$ containing the subword x_0^2 exactly a times but no subword x_1^2. Let H be the subspace of $\mathbb{C}\langle\langle \mathsf{X} \rangle\rangle$ generated by the set $\mathsf{H}^0 \cup \mathsf{H}^1$ and $I_H \subset \mathbb{C}\langle\langle \mathsf{X} \rangle\rangle$ denote the ideal generated by $\mathsf{X}^* \setminus (\mathsf{H}^0 \cup \mathsf{H}^1)$.

Remark 4.5.2. To prove any identity in $\mathbb{C}\langle\langle \mathsf{X}\rangle\rangle$ it suffices to prove it under both the projection to H and modulo I_H, since H is stable under the deconcatenation coproduct and the coalgebra H is dual to $\mathbb{C}\langle\langle \mathsf{X}\rangle\rangle/I_H$.

Definition 4.5.3. Let $\mathsf{S}^1_{1,0}(m,n) = 4\binom{2m+2n}{2m}\zeta(2m+2n+1)$ for all $m \geq 1, n \geq 0$. Define a formal power series $\mathsf{S}_{\mathsf{X}} \in \mathbb{R}\langle\langle \mathsf{X}\rangle\rangle$ by

$$\mathsf{S}_{\mathsf{X}} = 1 + \mathsf{S}^0_{0,0} + \mathsf{S}^1_{0,1} + \mathsf{S}^1_{1,0} + \mathsf{S}^1_{0,0},$$

where $\mathsf{S}^1_{0,1} = \overleftarrow{\mathsf{S}^1_{1,0}}$, and

$$\begin{aligned}
\mathsf{S}^0_{0,0} &= 4\sum_{n\geq 1}\zeta(2n+1)\,(\mathtt{x}_0\mathtt{x}_1)^n\mathtt{x}_0,\\
\mathsf{S}^1_{1,0} &= \sum_{m\geq 1, n\geq 0}\mathsf{S}^1_{1,0}(m,n)(\mathtt{x}_1\mathtt{x}_0)^m\mathtt{x}_0(\mathtt{x}_1\mathtt{x}_0)^n \qquad (4.58)\\
\mathsf{S}^1_{0,0} &= \frac{1}{2}\big(\mathsf{S}^0_{0,0}\,\mathsf{S}^0_{0,0} + \mathsf{S}^0_{0,0}\,\mathsf{S}^1_{1,0} + \mathsf{S}^1_{0,1}\,\mathsf{S}^0_{0,0}\big).
\end{aligned}$$

In order to compute the the Feynman amplitudes of the so called zig-zag graphs in Chap. 15 we also need to study another family of series in $\mathbb{R}\langle\langle \mathsf{X}\rangle\rangle$.

Definition 4.5.4. Let $\hat{\mathsf{S}}^1_{1,0}(m,n) = 4\,(1-4^{-m-n})\binom{2m+2n}{2m+1}\zeta(2m+2n+1)$ for all $m \geq 0, n \geq 1$. Define a formal power series $\hat{\mathsf{S}}_{\mathsf{X}} \in \mathbb{R}\langle\langle \mathsf{X}\rangle\rangle$ by

$$\hat{\mathsf{S}}_{\mathsf{X}} = 1 + \hat{\mathsf{S}}^0_{0,0} + \hat{\mathsf{S}}^1_{0,1} + \hat{\mathsf{S}}^1_{1,0} + \hat{\mathsf{S}}^1_{0,0},$$

where $\hat{\mathsf{S}}^0_{0,0} = \mathsf{S}^0_{0,0}$, $\hat{\mathsf{S}}^1_{0,1} = \overleftarrow{\hat{\mathsf{S}}^1_{1,0}}$, and

$$\begin{aligned}
\hat{\mathsf{S}}^1_{1,0} &= \sum_{m\geq 0, n\geq 1}\hat{\mathsf{S}}^1_{1,0}(m,n)(\mathtt{x}_1\mathtt{x}_0)^m\mathtt{x}_1(\mathtt{x}_1\mathtt{x}_0)^n, \qquad (4.59)\\
\hat{\mathsf{S}}^1_{0,0} &= \frac{1}{2}\big(\hat{\mathsf{S}}^0_{0,0}\,\hat{\mathsf{S}}^1_{1,0} + \hat{\mathsf{S}}^1_{0,1}\,\hat{\mathsf{S}}^0_{0,0}\big).
\end{aligned}$$

Let $\hat{\mathsf{H}}^a$ (resp. $\hat{H}$, resp. $I_{\hat{H}}$) denote the set of elements obtained from H^a (resp. H, resp. I_H) by interchanging $\mathtt{x}_0$ and $\mathtt{x}_1$.

For convenience, if $T = \sum_{\mathbf{w}\in\mathsf{X}^*} T_{\mathbf{w}}\mathbf{w} \in \mathbb{C}\langle\langle \mathsf{X}\rangle\rangle$ then for $a, b = 0$ or 1, we set

$$T_{a,b} := \delta_{a,b}T_{\mathtt{x}_a}\mathtt{x}_a + \sum_{\mathbf{w}\in\mathtt{x}_a\mathsf{X}^*\mathtt{x}_b} T_{\mathbf{w}}\mathbf{w},$$

and for $a = 0$ or 1, we set $T^a := \sum_{\mathbf{w}\in\mathsf{H}^a} T_{\mathbf{w}}\mathbf{w}$. So there are eight possible subseries $T^c_{a,b}$ for $a, b, c = 0$ or 1. For example, we clearly have $\mathsf{S}^0_{1,1} = \mathsf{S}^1_{1,1} = \mathsf{S}^0_{1,0} = \mathsf{S}^0_{0,1} = 0$.

Let $V = \Phi^{-1}\mathtt{x}_1\Phi \in \mathbb{C}\langle\langle \mathsf{X}\rangle\rangle$ where $\Phi = \Phi(\mathtt{x}_0, \mathtt{x}_1)$ is the Drinfeld associator defined by Eq. (4.5).

Lemma 4.5.5. *We have*

(i) $[\mathsf{S}_{\mathsf{X}}, \mathtt{x}_0] \equiv 0 \pmod{I_H}$ *and* $[\hat{\mathsf{S}}_{\mathsf{X}}, \mathtt{x}_0] \equiv 0 \pmod{I_{\hat{H}}}$*;*

(ii) $\mathsf{S}_{\mathsf{X}} V_- + V \mathsf{S}_{\mathsf{X}} \equiv 0 \pmod{I_H}$ *and* $\hat{\mathsf{S}}_{\mathsf{X}} V_- + V \hat{\mathsf{S}}_{\mathsf{X}} \equiv 0 \pmod{I_{\hat{H}}}$*, where* $V_- := V(-x_0, -x_1)$*;*

(iii) *both* S_{X} *and and* $\hat{\mathsf{S}}_{\mathsf{X}}$ *are real and they satisfy* $\mathsf{S}_{\mathsf{X}} \equiv \overleftarrow{\mathsf{S}}_{\mathsf{X}} \pmod{I_H}$ *and* $\hat{\mathsf{S}}_{\mathsf{X}} \equiv \overleftarrow{\hat{\mathsf{S}}}_{\mathsf{X}} \pmod{I_{\hat{H}}}$.

Proof. (i) To prove $[\mathsf{S}_{\mathsf{X}}, \mathtt{x}_0] \equiv 0 \pmod{I_H}$, by Definition 4.5.3 we only need to show

$$\mathsf{S}^0_{0,0}\, \mathtt{x}_0 + \mathsf{S}^1_{0,1}\, \mathtt{x}_0 = \mathtt{x}_0 \mathsf{S}^0_{0,0} + \mathtt{x}_0 \mathsf{S}^1_{1,0}\,. \tag{4.60}$$

This can be prove easily by comparing the coefficients of the words of the form $(\mathtt{x}_0\mathtt{x}_1)^a \mathtt{x}_0^2 (\mathtt{x}_1\mathtt{x}_0)^b$. Similarly we can show that $[\hat{\mathsf{S}}_{\mathsf{X}}, \mathtt{x}_0] \equiv 0 \pmod{I_{\hat{H}}}$. We leave the details to the interested reader.

(ii) For S_{X}, it suffices to prove $(V_-\mathsf{S}_{\mathsf{X}} + \mathsf{S}_{\mathsf{X}} V)_{a,b} \equiv 0 \pmod{I_H}$ for $a, b = 0$ or 1. Breaking these further into the odd and even weight parts we get seven nontrivial equations all of which follow from Exercise 4.27. Similar results can be prove for $\hat{\mathsf{S}}_{\mathsf{X}}$ using Exercise 4.28.

(iii) is obvious by the definitions. □

Define the generating series

$$F(z) = \sum_{\mathbf{w} \in \mathsf{X}^*} F_{\mathbf{w}}(z) \mathbf{w} = \mathrm{L}_{\mathsf{X}}(z) \mathsf{S}_{\mathsf{X}} \overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}). \tag{4.61}$$

Theorem 4.5.6. *The functions* $F_{\mathbf{w}}(z)$ *are single-valued. Moreover, for all* $\mathbf{w} \in H$

$$\partial_z \partial_{\bar{z}} F_{\mathtt{x}_a \mathbf{w} \mathtt{x}_b}(z) = \frac{F_{\mathbf{w}}(z)}{(z-a)(\bar{z}-b)} \tag{4.62}$$

for all $a, b \in \{0, 1\}$*, and they satisfy* $F_{\mathbf{w}}(z) = F_{\overleftarrow{\mathbf{w}}}(\bar{z})$.

Proof. After projecting to the subspace H we clearly have $F(z) = \mathrm{L}_{\mathsf{X}}(z) \overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z})$ which has no monodromy at 0 and 1 by Lemma 4.2.7. By Remark 4.5.2 it suffices to prove

$$\begin{aligned} M(\sigma_0) F(z) &\equiv F(z) \pmod{I_H}\,, \\ M(\sigma_1) F(z) &\equiv F(z) \pmod{I_H}\,. \end{aligned} \tag{4.63}$$

For the monodromy at 0, by Lemma 4.5.5(i)

$$e^{2i\pi \mathtt{x}_0} \mathsf{S}_{\mathsf{X}} \equiv \mathsf{S}_{\mathsf{X}} e^{2i\pi \mathtt{x}_0} \pmod{I_H}.$$

Thus

$$\begin{aligned} M(\sigma_0)F(z) &= \mathrm{L}_{\mathsf{X}}(z)e^{2i\pi \mathtt{x}_0}\mathsf{S}_{\mathsf{X}}e^{-2i\pi \mathtt{x}_0}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}) \\ &\equiv \mathrm{L}_{\mathsf{X}}(z)\mathsf{S}_{\mathsf{X}}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}) \equiv F(z) \pmod{I_H}. \end{aligned} \tag{4.64}$$

For the monodromy at 1, first we notice that $W := \Phi^{-1}e^{2i\pi \mathtt{x}_1}\Phi \equiv 1 + 2i\pi V \pmod{I_H}$ since $\mathtt{x}_1^2 \in I_H$. Moreover, from $\overleftarrow{\Phi}(-\mathtt{x}_0, -\mathtt{x}_1) = \Phi(\mathtt{x}_0, \mathtt{x}_1)^{-1}$ by Thm. B.22(ii)

$$\overleftarrow{V}(\mathtt{x}_0, \mathtt{x}_1) = \Phi(-\mathtt{x}_0, -\mathtt{x}_1)^{-1}\mathtt{x}_1\Phi(-\mathtt{x}_0, -\mathtt{x}_1) = -V_-. \tag{4.65}$$

Since $V^2 \equiv \Phi^{-2}\mathtt{x}_1^2\Phi^2 \equiv 0 \pmod{I_H}$, by multiplying Lemma 4.5.5(ii) by V from the left and using Eq. (4.65) we get $V\mathsf{S}_{\mathsf{X}}\overleftarrow{V} \equiv 0 \pmod{I_H}$. Therefore

$$W\mathsf{S}_{\mathsf{X}}\overleftarrow{W} - \mathsf{S}_{\mathsf{X}} \equiv 2i\pi(-\mathsf{S}_{\mathsf{X}}\overleftarrow{V} + V\mathsf{S}_{\mathsf{X}}) + (2i\pi)^2 V\mathsf{S}_{\mathsf{X}}\overleftarrow{V} \equiv 0 \pmod{I_H}$$

by Lemma 4.5.5(ii) and Eq. (4.65) again. By Lemma 4.2.7 we obtain

$$M(\sigma_1)F(z) = \mathrm{L}_{\mathsf{X}}(z)W\mathsf{S}_{\mathsf{X}}\overleftarrow{W}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}) \equiv \mathrm{L}_{\mathsf{X}}(z)\mathsf{S}_{\mathsf{X}}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}) \equiv F(z) \pmod{I_H}.$$

From this and (4.64) we see that $F_{\mathbf{w}}(z)$ are single-valued for all $\mathbf{w} \in \mathsf{X}^*$.

Equation (4.62) follows easily from Prop. 4.2.1 and the last sentence of the theorem follows easily Lemma 4.5.5(iii). □

Define the generating series

$$\hat{F}(z) = \sum_{\mathbf{w}\in\mathsf{X}^*} \hat{F}_{\mathbf{w}}(z)\mathbf{w} = \mathrm{L}_{\mathsf{X}}(z)\hat{\mathsf{S}}_{\mathsf{X}}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}). \tag{4.66}$$

Parallel to Thm. 4.5.6 we have

Theorem 4.5.7. *The functions $\hat{F}_{\mathbf{w}}(z)$ are single-valued. Moreover, for all $\mathbf{w} \in \hat{H}$*

$$\partial_z\partial_{\bar{z}}\hat{F}_{\mathtt{x}_a\mathbf{w}\mathtt{x}_b}(z) = \frac{F_w(z)}{(z-a)(\bar{z}-b)} \tag{4.67}$$

for all $a, b \in \{0, 1\}$, and they satisfy $\hat{F}_{\mathbf{w}}(z) = \hat{F}_{\overleftarrow{\mathbf{w}}}(\bar{z})$.

Proof. We only need to show the triviality of the monodromy at 1. The proofs of the triviality of the monodromy at 0 and all the other statements in the theorem are exactly the same as those of Thm. 4.5.6.

For the monodromy at 1, we have from Lemma 4.5.5(ii)

$$V\hat{\mathsf{S}}_{\mathsf{X}} + \hat{\mathsf{S}}_{\mathsf{X}}V_- \equiv 0 \pmod{I_{\hat{H}}}. \tag{4.68}$$

Multiplying this on the left (resp. right) by V^2 (resp. V_-^2), we get

$$V^2\hat{\mathsf{S}}_{\mathsf{X}}V_- \equiv V\hat{\mathsf{S}}_{\mathsf{X}}V_-^2 \equiv 0 \pmod{I_{\hat{H}}}. \tag{4.69}$$

Now let $W = \Phi^{-1}e^{2i\pi\mathrm{x}_1}\Phi$ as before. Then $W \equiv 1 + 2i\pi V + \frac{1}{2}(2i\pi)^2V^2$ and $\overleftarrow{W} \equiv 1 + 2i\pi V_- + \frac{1}{2}(2i\pi)^2V_-^2 \pmod{I_{\hat{H}}}$ by Eq. (4.65). Thus

$$W\hat{\mathsf{S}}_{\mathsf{X}}\overleftarrow{W} - \hat{\mathsf{S}}_{\mathsf{X}} \equiv 2i\pi(V\hat{\mathsf{S}}_{\mathsf{X}} + \hat{\mathsf{S}}_{\mathsf{X}}V_-) - 4i\pi^3(V\hat{\mathsf{S}}_{\mathsf{X}}V_-^2 + V^2\hat{\mathsf{S}}_{\mathsf{X}}V_-)$$
$$-2\pi^2\big(V(V\hat{\mathsf{S}}_{\mathsf{X}}+\hat{\mathsf{S}}_{\mathsf{X}}V_-)+(V\hat{\mathsf{S}}_{\mathsf{X}}+\hat{\mathsf{S}}_{\mathsf{X}}V_-)V_-\big)+4\pi^4V^2\hat{\mathsf{S}}_{\mathsf{X}}V_-^2) \equiv 0 \pmod{I_{\hat{H}}},$$

by Eqs. (4.68) and (4.69). Hence

$$M(\sigma_1)\hat{F}(z) = \mathrm{L}_{\mathsf{X}}(z)W\hat{\mathsf{S}}_{\mathsf{X}}\overleftarrow{W}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}) \equiv \mathrm{L}_{\mathsf{X}}(z)\hat{\mathsf{S}}_{\mathsf{X}}\overleftarrow{\mathrm{L}}_{\mathsf{X}}(\bar{z}) \equiv \hat{F}(z) \pmod{I_H}.$$

This completes the proof of the theorem. □

4.6 Historical notes

By studying the twisting (a certain gauge transformation) of quasi-Hopf algebras Drinfeld (see [163,164]) was able to not only give a beautiful explanation of Kohno's Monodromy Theorem in [341] but also describe the absolute Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ in terms of the Grothendieck-Teichmüller group GT which acts transitively on a formal Lie series P with rational coefficients such that his associator $\Phi(t^{12}, t^{23}) = \exp P(t^{12}, t^{23})$. He proved that Φ satisfies both the pentagon relation (4.13) and the hexagon relation (4.12). Moreover, he showed that the special derivation condition (4.14) for $\varphi \in \mathfrak{t}_3^0$ (resp.(4.19) for $g \in \exp\mathfrak{t}_3^0$) follows from duality and hexagon equations (4.12) and the pentagon equation (4.13) in [164, Prop. 5.7]. (resp.(4.16) and (4.17) in [164, Prop. 5.9].) Our proof of the special derivation condition in Props. 4.1.19 and 4.1.22 follows Drinfeld's original proof closely.

In [207] Furusho proved that in the MZV setting, the hexagon relation is a consequence of the pentagon relation. Later, in [208] he further proved that the pentagon relation implies all the regularized DBSFs for MZVs. The computation of all the coefficients of Φ including its regularized terms, which generalizes Thm. 4.1.14, was stated first by Le and Murakami [368] but revised later by Furusho as [203, Prop. 3.2.3] using Le-Murakami's method. Our Lie algebra approach is influenced by [156,465]. For example, Lemma 4.1.9 is proved by Racinet as Cor. 2.4.4 of [465] by using Poincaré-Birkhoff-Witt Theorem.

Using higher order of roots of unity Enriquez [188] generalized Drinfeld associator to the cyclotomic setting by defining the notion of a quasi-reflection algebra. These associators satisfy some relations called the octagon and the mixed pentagon relations. We will see in Chap. 13 that these

generalizations provide relations among special values of MPLs at roots of unity. In [95] Brochier computed explicitly the monodromy representations of the cyclotomic KZ differential equation system and obtained the corresponding cyclotomic Kohno-Drinfeld Theorem. Enriquez and Furusho [191] further consider various implications among these and some more relations, one of which is the special action condition we refer to in Eq. (4.32).

As one of the most beautiful applications of the theory of MZVs and MPLs, the SVMPLs and the SVMZVs have appeared in a few computations in mathematical physics in the past few years, such as Brown and Schnetz's solution of the zig-zag conjecture [109] (see Chap. 15), the expression of the coefficients of the closed super-string tree-level amplitude [482], and the study of wrapping functions1 in $N = 4$ super Yang-Mills [371] and some other related works [124, 162]. Our proof of Thm. 4.1.1 follows that of [99, Thm. 2.1] and [245, Prop. 1.2]. For the proof of the zig-zag conjecture the generating series in Sec. 4.5 will play crucial roles. In both Thm. 4.5.7 and 4.5.7 the single-valuedness is the most techinical part. Note that the coefficients of both S and $\hat{\mathsf{S}}$ series here have opposite signs to those in the original paper [109] due to our different convention in setting up the KZ-equations.

Exercises

4.1. Use the method of undetermined coefficients to find the first few terms of $G_0(x)$ and $G_1(x)$ satisfying Eq. (4.4). Then compute the first few term of $G_1^{-1}G_0$ and show that the Drinfeld associator has the following expansion:

$$\begin{aligned}\Phi(\mathtt{x}_0,\mathtt{x}_1) = 1 &- \zeta(2)[\mathtt{x}_0,\mathtt{x}_1] - \zeta(3)\Big([\mathtt{x}_0,[\mathtt{x}_0,\mathtt{x}_1]] - [[\mathtt{x}_0,\mathtt{x}_1],\mathtt{x}_1]\Big)\\ &+ \zeta(4)[\mathtt{x}_0,[\mathtt{x}_0,[\mathtt{x}_0,\mathtt{x}_1]]] + \zeta(3,1)[\mathtt{x}_0,[[\mathtt{x}_0,\mathtt{x}_1],\mathtt{x}_1]]\\ &- \zeta(2,1,1)[[[\mathtt{x}_0,\mathtt{x}_1],\mathtt{x}_1],\mathtt{x}_1] + \frac{1}{2}\zeta(2)^2[\mathtt{x}_0,\mathtt{x}_1]^2 + \cdots .\end{aligned}$$

4.2. Show that if $f_*[\ ] : \mathfrak{A}^1_* \to \mathbb{R}$ satisfies the stuffle relations Eq. (4.7) then it satisfies Eq. (4.7) which implies that $f_* = \sum_{\mathbf{w}\in\mathsf{X}^*\mathtt{x}_1} f_*[\mathbf{w}]\mathbf{w}$ is a group-like for Δ_*.

4.3. Let $f = \sum_{\mathbf{w}\in\mathtt{x}_0\mathsf{X}^*\mathtt{x}_1} f[\mathbf{w}]\mathbf{w} \in (\mathscr{R}_{\mathrm{cv}}, \Delta_{\sqcup\!\sqcup})$. Prove that f satisfies shuffle relation if and only if $\Delta_{\sqcup\!\sqcup}(f) = f \otimes f$ (*i.e.*, f is group-like for $\Delta_{\sqcup\!\sqcup}$) by expanding both $\Delta_{\sqcup\!\sqcup}(f)$ and $f \otimes f$ and then compare the corresponding terms. Prove the above result is still valid if we consider $(\mathscr{R}_{\sqcup\!\sqcup}, \Delta_{\sqcup\!\sqcup})$. This proves Thm. 4.1.8.

4.4. Prove that Lemma 4.1.9 and Lemma 4.1.9 both follow from the Cor. B.26 of the Poincaré-Birkhoff-Witt Theorem B.25.

4.5. Prove that $z_n = \sum_{1\leq i<j\leq n} t^{ij}$ is central in $\mathfrak{t}_n$ and that $\mathfrak{t}_n \cong \mathfrak{t}_n^0 \oplus \mathbb{K}\cdot z_n$.

4.6. Show that the Grothendieck-Teichmüller Lie algebra $\mathfrak{grt}_1(\mathbb{K})$ indeed forms a Lie algebra with the Lie bracket given by

$$\langle \varphi_1, \varphi_2 \rangle = [\varphi_1, \varphi_2] + D_{\varphi_2}(\varphi_1) - D_{\varphi_1}(\varphi_2), \tag{4.70}$$

where D_φ is the derivation of $\mathfrak{t}_3^0$ defined by

$$D_\varphi(a) = [\varphi, a] \quad \text{and } D_\varphi(b) = 0.$$

The Lie algebra structure can be realized as an embedding

$$\mathfrak{grt}_1(\mathbb{K}) \hookrightarrow \mathrm{Der}(\mathfrak{t}_3^0)$$

by sending $\varphi \mapsto D_\varphi$.

4.7. Show that the Grothendieck-Teichmüller group $\mathrm{GRT}_1(\mathbb{K})$ indeed forms a group by the multiplication

$$g_2 \circ g_1 = g_1(g_2^{-1}ag_2, b)\cdot g_2(a,b) = g_2(a,b)\cdot g_1(a, g_2^{-1}bg_2). \tag{4.71}$$

The group structure can be realized as an embedding (with anti-homomorphism)

$$\mathrm{GRT}_1(\mathbb{K}) \hookrightarrow \mathrm{Aut}\exp \mathfrak{t}_3^0$$

sending by sending g to the automorphism A_g defined by

$$a \mapsto a \quad \text{and } b \mapsto g^{-1}bg.$$

4.8. Recall that for a group G a G-torsor is a non-empty set X on which G acts freely and transitively, meaning that for any $x, y \in X$ there exists a unique $g \in G$ such that $x * g = y$ where $*$ denotes the right action of G on X. Show that the associator set $\mathrm{M}_\lambda(\mathbb{K})$ is a left $\mathrm{GRT}_1(\mathbb{K})$-torsor by (4.71) defined in Exercise 4.7 with $g_1 \in \mathrm{M}_\lambda(\mathbb{K})$ and $g_2 \in \mathrm{GRT}_1(\mathbb{K})$.

4.9. Complete the proof of Lemma 4.1.27. For the uniqueness of W_1, let $\mathbf{w}$ be any word neither beginning nor ending with A or B. First show that the coefficient $c^{i,j}_{k,l,m,n}$ of $A^kB^l\mathbf{w}A^mB^nu^iv^j$ for $f(u,v)$ satisfies

$$ic^{i,j}_{k,l,m,n} = c^{i,j}_{k-1,l,m,n} - c^{i,j}_{k,l,m-1,n} + (Rf)[A^kB^l\mathbf{w}A^mB^nu^{i-1}v^j] \tag{4.72}$$

and

$$jc^{i,j}_{k,l,m,n} = c^{i,j}_{k,l-1,m,n} - c^{i,j}_{k,l,m,n-1} + (Sf)[A^kB^l\mathbf{w}A^mB^nu^iv^{j-1}] \tag{4.73}$$

for all $i+j \geq 0$ and $k+l+m+n \geq 0$. Here if a word contains either A^{-1} or B^{-1} the corresponding term is vacuous. Then use double induction starting with $k+l+m+n=0$ to show that $c^{i,j}_{0,0,0,0}$ can be determined uniquely, with the inner induction loop on $i+j$. Then treat the general $k+l+m+n>0$ case by induction. In fact, if $i=j$ then Eq. (4.72) coincides with Eq. (4.73) by Eq. (4.26). This determines recursively the coefficient $c^{i,i}_{k,l,m,n}$. Furthermore, for $i,j>0$ the constraint $(1/i)\times$Eq. (4.72)$=(1/j)\times$Eq. (4.73) determines recursively the coefficient $c^{i-1,j}_{k,l,m,n}$ (resp. $c^{i,j-1}_{k,l,m,n}$) if $i<j$ (resp. $i>j$).

4.10. Complete the proof of Lemma 4.1.28.

4.11. Complete the proof of Eq. (4.31) by mimicking the proof of the first identity $W_2 = W_3 e^{ht^{23}/2}$.

4.12. Prove Eq. (4.15) by using Eq. (4.12).

4.13. Prove that $f_0(z)$ defined by (4.2) and (4.3) indeed provides a solution in Thm. 4.1.1.

4.14. Prove the uniqueness of the solution in Thm. 4.1.1.

4.15. Show that by the change of variables $z-p_0 \to 1/(z-p_0)$ one obtains a solution $L^{\infty}(z)$ to Eq. (4.1) on U which corresponds to the point at infinity. Fix any branch of $\log z$ on U. Show that $L^{\infty}(z) \sim \exp(\mathbf{x}_{\infty} \log z)$ where $\mathbf{x}_{\infty} = \sum_{i=0}^{N} \mathbf{x}_i$.

4.16. Show that for any $\mathbf{w} \in \mathsf{X}^*$ and $a \in \{0,1\}$

$$\partial_{\bar{z}} \mathscr{L}_{\mathbf{w}\mathbf{x}_a}(\bar{z}) = \frac{\mathscr{L}_{\mathbf{w}}(\bar{z})}{\bar{z}-a},$$

4.17. Show that $F(\mathbf{x}_0,\mathbf{x}_1)$ defined by Eq. (4.46) satisfies $F \equiv 0$ modulo words of depth two. Moreover, it has the following explicit expansion

$$F(\mathbf{x}_0,\mathbf{x}_1) = \mathbf{x}_1 + \zeta(2)[[\mathbf{x}_0,\mathbf{x}_1],\mathbf{x}_1] + \zeta(3)[\mathbf{x}_1,\mathbf{u}_3] + \cdots$$

where $\mathbf{u}_3$ is the coefficient of $\zeta(3)$ in $\Phi(\mathbf{x}_0,\mathbf{x}_1)$ given by Exercise 4.1. Then show that by Eq. (4.47)

$$\mathrm{y}_1 = \mathbf{x}_1 + 2\zeta(3)[\mathbf{x}_1,\mathbf{u}_3] + \cdots$$

4.18. Show that if y_1 satisfies Eq. (4.47) then it satisfies Eq. (4.45).

4.19. Show that

$$\mathrm{L}_{\{0\}^m 1\{0\}^n}(z) = \sum_{k=0}^{n} (-1)^{k+1} \binom{m+k}{m} \frac{\log^{n-k} z}{(n-k)!} \mathrm{Li}_{m+k+1}(z).$$

4.20. Use Exercise 4.19 to show that

$$\begin{aligned}\mathscr{L}_{\{0\}^m1\{0\}^n}(z) = &\sum_{k=0}^{n}(-1)^{k+1}\binom{m+k}{m}\frac{\log^{n-k}(z\bar{z})}{(n-k)!}\operatorname{Li}_{m+k+1}(z)\\ &+\sum_{k=0}^{m}(-1)^{k+1}\binom{n+k}{n}\frac{\log^{m-k}(z\bar{z})}{(m-k)!}\operatorname{Li}_{n+k+1}(\bar{z}).\end{aligned}$$

4.21. Show that the regularizations in Definition 4.4.1 are well-defined (*i.e.*, unique).

4.22. Show that if $f(z)\in\mathsf{Ls}$ then $f(z)=o(|z|^{-\epsilon})$ as $z\to 0$, i.e.,

$$\lim_{z\to 0}|z|^{\epsilon}f(z)=0.$$

4.23. Show that if $f(z)\in\mathsf{Ls}$ then $f(z)=o(|z|^{\epsilon})$ as $z\to\infty$, i.e.,

$$\lim_{z\to\infty}|z|^{-\epsilon}f(z)=0.$$

4.24. Let $P\in\mathsf{Ls}$ and $a,b\in\{0,1\}$. Show that

$$\begin{aligned}&\left(\int_0\frac{\mathrm{d}z}{z-a}\int_0\frac{\mathrm{d}\bar{z}}{\bar{z}-b}-\int_0\frac{\mathrm{d}\bar{z}}{\bar{z}-b}\int_0\frac{\mathrm{d}z}{z-a}\right)P(z)\\ &\quad=\left(\int_0\frac{\mathrm{d}\bar{z}}{\bar{z}-b}P(z)\Big|_{z=a}-\int_0\frac{\mathrm{d}z}{z-a}P(z)\Big|_{z=b}\right)\cdot\mathscr{L}_1(z).\end{aligned}$$

4.25. Prove the stability of $\mathsf{Ls}^{\mathrm{sv}}$ under the transformations $z\to\frac{z}{z-1}$ by using the following identities:

$$\begin{aligned}\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}(1-z)&=\mathrm{L}_{\mathrm{x}_1,\mathrm{x}_0}(z)\Phi_{\mathrm{x}_0,\mathrm{x}_1}\overleftarrow{\Phi}_{\mathrm{x}_0,\mathrm{y}_1}\overleftarrow{\mathrm{L}}_{\mathrm{y}_1,\mathrm{x}_0}(\bar{z})\\&=\mathscr{L}_{\mathrm{x}_1,\mathrm{x}_0}(z)\,\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}(1)\;=\;\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}(1)\,\mathscr{L}_{\mathrm{y}_1,\overleftarrow{\mathrm{y}_1}(\mathrm{y}_1,\mathrm{x}_0)}(z),\\
\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}\left(\frac{z-1}{z}\right)&=\mathrm{L}_{-\mathrm{x}_0-\mathrm{x}_1,\mathrm{x}_0}(z)\Phi_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{x}_1}\overleftarrow{\Phi}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{y}_1}\overleftarrow{\mathrm{L}}_{-\mathrm{x}_0-\mathrm{y}_1,\mathrm{x}_0}(\bar{z})\\&=\mathscr{L}_{-\mathrm{x}_0-\mathrm{x}_1,\mathrm{x}_0}(z)\,\mathscr{L}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{x}_1}(1)\\&=\mathscr{L}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{x}_1}(1)\,\mathscr{L}_{-\mathrm{x}_0-\mathrm{y}_1,\overleftarrow{\mathrm{y}_1}(-\mathrm{x}_0-\mathrm{y}_1,\mathrm{x}_0)}(z),\\
\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}\left(\frac{z}{z-1}\right)&=\mathrm{L}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{x}_1}(z)\overleftarrow{\mathrm{L}}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{y}_1}(\bar{z})\;=\;\mathscr{L}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{x}_1}(z),\\
\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}\left(\frac{1}{1-z}\right)&=\mathrm{L}_{\mathrm{x}_1,-\mathrm{x}_0-\mathrm{x}_1}(z)\Phi_{\mathrm{x}_0,\mathrm{x}_1}\overleftarrow{\Phi}_{\mathrm{x}_0,\mathrm{y}_1}\overleftarrow{\mathrm{L}}_{\mathrm{y}_1,-\mathrm{x}_0-\mathrm{y}_1}(\bar{z})\\&=\mathscr{L}_{\mathrm{x}_1,-\mathrm{x}_0-\mathrm{x}_1}(z)\,\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}(1)\\&=\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}(1)\,\mathscr{L}_{\mathrm{y}_1,\overleftarrow{\mathrm{y}_1}(\mathrm{y}_1,-\mathrm{x}_0-\mathrm{y}_1)}(z),\\
\mathscr{L}_{\mathrm{x}_0,\mathrm{x}_1}\left(\frac{1}{z}\right)&=\mathrm{L}_{-\mathrm{x}_0-\mathrm{x}_1,\mathrm{x}_1}(z)\Phi_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{x}_1}\overleftarrow{\Phi}_{\mathrm{x}_0,-\mathrm{x}_0-\mathrm{y}_1}\overleftarrow{\mathrm{L}}_{-\mathrm{x}_0-\mathrm{y}_1,\mathrm{y}_1}(\bar{z})\end{aligned}$$

$$= \mathscr{L}_{-\mathtt{x}_0-\mathtt{x}_1,\mathtt{x}_1}(z)\,\mathscr{L}_{\mathtt{x}_0,-\mathtt{x}_0-\mathtt{x}_1}(1)$$
$$= \mathscr{L}_{\mathtt{x}_0,-\mathtt{x}_0-\mathtt{x}_1}(1)\,\mathscr{L}_{-\mathtt{x}_0-\mathtt{y}_1,\overleftarrow{\mathtt{y}_1}(-\mathtt{x}_0-\mathtt{y}_1,\mathtt{y}_1)}(z).$$

4.26. Show that modulo I_H the series $V = \Phi^{-1}\mathtt{x}_1\Phi$ satisfies

$$V \equiv (1 - \Phi^0_{0,0} + \Phi^0_{0,1} + \Phi^1_{0,0} - \Phi^1_{0,1})\mathtt{x}_1(1 + \Phi^0_{0,0} + \Phi^0_{1,0} + \Phi^1_{0,0} + \Phi^1_{1,0}).$$

4.27. Show that modulo I_H

$$\begin{aligned}
2V^0_{0,1} + \mathsf{S}^0_{0,0}\,V^0_{1,1} \equiv 2V^0_{1,0} - V^0_{1,1}\,\mathsf{S}^0_{0,0} \equiv V^0_{0,1}\,\mathsf{S}^0_{0,0} + \mathsf{S}^0_{0,0}\,V^0_{1,0} &\equiv 0\,,\\
2V^1_{1,1} - V^0_{1,1}\,\mathsf{S}^1_{0,1} + \mathsf{S}^1_{1,0}\,V^0_{1,1} &\equiv 0,\\
V^0_{0,1}\,\mathsf{S}^1_{0,1} + \mathsf{S}^1_{0,1}\,V^0_{0,1} + \mathsf{S}^0_{0,0}\,(V^0_{0,1} + V^1_{1,1}) + \mathsf{S}^1_{0,0}\,V^0_{1,1} &\equiv 0,\\
V^0_{1,0}\,\mathsf{S}^1_{1,0} + \mathsf{S}^1_{1,0}\,V^0_{1,0} + (V^0_{1,0} + V^1_{1,1})\mathsf{S}^0_{0,0} - V^0_{1,1}\,\mathsf{S}^1_{0,0} &\equiv 0,\\
\mathbf{W} + (\mathsf{S}^0_{0,0} + \mathsf{S}^1_{0,1})V^0_{0,0} - V^0_{0,0}\,(\mathsf{S}^0_{0,0} + \mathsf{S}^1_{1,0}) &\equiv 0,
\end{aligned}$$

where $\mathbf{W} = 2V^1_{0,0} + V^0_{0,1}\,\mathsf{S}^1_{0,0} - V^1_{0,1}\,\mathsf{S}^0_{0,0} + \mathsf{S}^0_{0,0}\,V^1_{1,0} + \mathsf{S}^1_{0,0}\,V^0_{1,0}$.

4.28. Decompose V as $V = \sum_{a,b,c\in\{0,1\}} \hat{V}^a_{b,c}$ in $\hat{H}$ where $\hat{V}^a$ is the sum of the subset of words of V that begins with $\mathtt{x}_b$, ends with $\mathtt{x}_c$, and contains the subword $\mathtt{x}_1^2$ exactly a times. Show that modulo $I_{\hat{H}}$

$$\begin{aligned}
2\hat{V}^0_{0,1} + \hat{\mathsf{S}}^0_{0,0}\,\hat{V}^0_{1,1} \equiv 2\hat{V}^1_{1,1} - \hat{V}^0_{1,1}\,\hat{\mathsf{S}}^1_{0,1} + \hat{\mathsf{S}}^1_{1,0}\,\hat{V}^0_{1,1} &\equiv 0,\\
\hat{V}^0_{0,1}\,\hat{\mathsf{S}}^1_{0,1} + \hat{\mathsf{S}}^1_{0,1}\,\hat{V}^0_{0,1} + \hat{\mathsf{S}}^0_{0,0}\,\hat{V}^1_{1,1} + \hat{\mathsf{S}}^1_{0,0}\,\hat{V}^0_{1,1} &\equiv 0,\\
\mathbf{W} - \hat{V}^0_{0,0}\,\hat{\mathsf{S}}^1_{1,0} + \hat{\mathsf{S}}^1_{0,1}\,\hat{V}^0_{0,0} &\equiv 0,
\end{aligned}$$

where $\mathbf{W} = 2\hat{V}^1_{0,0} + \hat{V}^0_{0,1}\,\hat{\mathsf{S}}^1_{0,0} + \hat{\mathsf{S}}^1_{0,0}\,\hat{V}^0_{1,0} - \hat{V}^1_{0,1}\,\hat{\mathsf{S}}^0_{0,0} + \hat{\mathsf{S}}^0_{0,0}\,\hat{V}^1_{1,0}$.

Chapter 5

Multiple zeta value identities

There are many known relations among MZVs most of which have been proved rigorously in recent years with a variety of techniques. In Sec. 3.3 we have seen that (regularized) DBSFs give rise to many linear relations over $\mathbb{Q}$ and conjecturally these imply all the other linear relations. In this chapter we collect some special families of these relations and also present some conjectural ones at the end. We will often resort to the properties of Bernoulli numbers thanks to Euler's famous evaluation in Eq. (1.2).

5.1 Duality relations

Let k be a positive integer and $\mathbf{s} = (s_1, \ldots, s_d)$ of weight $w = |\mathbf{s}|$. Write

$$\mathbf{s} = (a_1 + 1, \{1\}^{b_1 - 1}, \ldots, a_\ell + 1, \{1\}^{b_\ell - 1})$$

for $a_1, b_1, \ldots, a_\ell, b_\ell \geq 1$. Then its dual is defined by

$$\mathbf{s}^* = (b_\ell + 1, \{1\}^{a_\ell - 1}, \ldots, b_1 + 1, \{1\}^{a_1 - 1}). \tag{5.1}$$

An important relation between $\mathbf{s}$ and its dual is the following:

$$\mathrm{dp}(\mathbf{s}) + \mathrm{dp}(\mathbf{s}^*) = |\mathbf{s}|. \tag{5.2}$$

Example 5.1.1. One can check easily that

$$(n + 1, \{1\}^{k-1})^* = (k + 1, \{1\}^{n-1})$$

Recall from Definition 3.2.2 that $\mathfrak{A}$ is the algebra of words generated over $\mathtt{x}_0$ and $\mathtt{x}_1$. In this chapter, to save space, we will write $\mathtt{x}_0 = \mathtt{a}$ and $\mathtt{x}_1 = \mathtt{b}$. The following result gives the so-called *duality relation* of MZVs.

Theorem 5.1.2. *For every admissible composition* $\mathbf{s}$*, we have*

$$\zeta(\mathbf{s}) = \zeta(\mathbf{s}^*). \tag{5.3}$$

Using admissible words (see Definition 3.2.2*), this can be expressed as*

$$\zeta(\mathbf{w}) = \zeta(\tau\mathbf{w}) \quad \forall \mathbf{w} \in \mathfrak{A}^0 \tag{5.4}$$

where τ *is the anti-automorphism of* $\mathfrak{A}^0$ *by exchanging* $\mathtt{a}$ *and* $\mathtt{b}$.

Proof. By Eq. (5.1), given an admissible word $\mathbf{w} = \mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}$ (see Definition 3.2.1) and $\mathbf{s}^* = \mathbf{t} = (t_1, \ldots, t_\ell)$ one checks that $\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_\ell-1}\mathtt{b} = \mathtt{ab}^{s_d-1}\cdots\mathtt{ab}^{s_1-1} = \tau(\mathbf{w})$ (see Exercise 5.2). Set the 1-forms $\mathtt{a} = \mathrm{d}t/t$ and $\mathtt{b} = \mathrm{d}t/(1-t)$ as before. By Chen's formula on the reversal of iterated integrals Lemma 2.1.2(iii) followed by the substitution $t \to 1-t$ one gets (notice $\mathrm{d}(1-t) = -\mathrm{d}t$)

$$\zeta(\mathbf{s}) = \int_0^1 w = (-1)^{|\mathbf{s}|}\int_1^0 \mathtt{ba}^{s_d-1}\cdots\mathtt{ba}^{s_1-1} = \int_0^1 \mathtt{ab}^{s_d-1}\mathtt{b}\cdots\mathtt{ab}^{s_1-1} = \zeta(\mathbf{s}^*).$$

This proves the theorem. □

5.2 Cyclic sum relation

First we define a cyclic equivalence classes of multiple indices on the set

$$I(w,d) = \{(k_1,\ldots,k_d) | k_1 + \cdots + k_d = w,\ k_1,\ldots,k_d \geq 1\}.$$

Definition 5.2.1. We say two elements of $I(w,d)$ are *cyclically equivalent* if they are cyclic permutations of each other, *i.e.*, for the permutation $\sigma = (1,\ldots,d)$ and $j = 1,\ldots,d$, we define $(k_1,\ldots,k_d) \equiv (\sigma^j(k_1),\ldots,\sigma^j(k_d))$ whose equivalence class is denoted by $[(k_1,\ldots,k_d)]$. Let $\mathcal{C}(w,d)$ be the set of cyclic equivalence classes of $I(w,d)$.

The main idea of the proof of the cyclic sum relation is to apply partial fractions. Already used by Euler this technique is ubiquitous in many proofs of MZV identities.

We first need a lemma. Let $\mathbb{N}_0 = \mathbb{N}\cup\{0\}$ be the set of nonnegative integers For $k_1, k_2, \ldots, k_d \in \mathbb{N}$ and $k_{d+1} \in \mathbb{N}_0$, let

$$A(k_1,\ldots,k_d) = \sum_{n_1>n_2>\cdots>n_d>n_{d+1}\geq 0} \frac{1}{(n_1-n_{d+1})n_1^{k_1}\ldots n_d^{k_d}}$$

and

$$B(k_1,\ldots,k_d,k_{d+1}) = \sum_{n_1>n_2>\cdots>n_d>n_{d+1}>0} \frac{1}{(n_1-n_{d+1})n_1^{k_1}\ldots n_d^{k_d}n_{d+1}^{k_{d+1}}}.$$

Lemma 5.2.2. *The infinite sum* $A(k_1,\ldots,k_d)$ *converges when one of* $k_1,\ldots,k_d$ *is* > 1, *and* $B(k_1,\ldots,k_d,k_{d+1})$ *converges when one of* $k_1,\ldots,k_d,k_{d+1}+1$ *is* > 1.

Proof. By the definition

$$B(k_1,\dots,k_d,0)=A(k_1,\dots,k_d)-\zeta(k_1+1,k_2,\dots,k_d). \tag{5.5}$$

Applying the partial fraction identity

$$\frac{1}{n_1(n_1-n_{d+1})}=\frac{1}{n_{d+1}}\left(\frac{1}{n_1-n_{d+1}}-\frac{1}{n_1}\right) \tag{5.6}$$

to $B(k_1,\dots,k_d,k_{d+1})$ we get

$$B(k_1,\dots,k_d,k_{d+1})=B(k_1-1,k_2,\dots,k_d,k_{d+1}+1)-\zeta(k_1,\dots,k_d,k_{d+1}+1). \tag{5.7}$$

Applying Eq. (5.6) to $B(1,k_2,\dots,k_d,k_{d+1})$ we get

$$\begin{aligned}
&\sum_{n_1>n_2>\cdots>n_d>n_{d+1}>0}\frac{1}{n_2^{k_2}\dots n_d^{k_d}n_{d+1}^{k_{d+1}+1}}\left(\frac{1}{n_1-n_{d+1}}-\frac{1}{n_1}\right)\\
&=\sum_{n_2>\cdots>n_d>n_{d+1}>0}\frac{1}{n_2^{k_2}\dots n_d^{k_d}n_{d+1}^{k_{d+1}+1}}\sum_{n_1=n_2+1}^{\infty}\left(\frac{1}{n_1-n_{d+1}}-\frac{1}{n_1}\right)\\
&=\sum_{n_2>\cdots>n_d>n_{d+1}>0}\frac{1}{n_2^{k_2}\dots n_d^{k_d}n_{d+1}^{k_{d+1}+1}}\sum_{j=0}^{n_{d+1}-1}\frac{1}{n_2-j}\\
&=\sum_{n_2>\cdots>n_d>n_{d+1}>j\ge 0}\frac{1}{(n_2-j)n_2^{k_2}\dots n_d^{k_d}n_{d+1}^{k_{d+1}+1}}
\end{aligned}$$

and so

$$B(1,k_2,\dots,k_d,k_{d+1})=A(k_2,\dots,k_d,k_{d+1}+1). \tag{5.8}$$

By Eq. (5.5)

$$B(k_1,\dots,k_d,k_{d+1})\le B(k_1,\dots,k_d,0)\le A(k_1,\dots,k_d).$$

If $k_1\ge 2$ then $B(k_1,\dots,k_{d+1})$ converges whenever $A(k_1,\dots,k_d)$ does; otherwise, if $k_1=1$ then Eq. (5.8) yields $B(k_1,\dots,k_{d+1})=A(k_2,\dots,k_d,k_{d+1}+1)$. So the statement about the B's follows from the one about the A's.

Turning to the first assertion of lemma we clearly only need to consider the case $k_1+\cdots+k_d=d+1$ when $A(k_1,\dots,k_d)$ are among the largest. Now

$$\begin{aligned}
A(2,1,\dots,1)&=\sum_{n_1>n_2>\cdots>n_{d+1}\ge 0}\frac{1}{n_1^2(n_1-n_{d+1})n_2\dots n_d}\\
&\le\sum_{\substack{n_1>n_2>\cdots>n_d>0\\ n_1\ge j>0}}\frac{1}{n_1^2 j n_2\dots n_d}\\
&=\zeta(3,\{1\}^{d-1})+d\zeta(2,\{1\}^d)+\sum_{i=1}^{d-1}\zeta(2,\{1\}^{i-1},2,\{1\}^{d-i-1})
\end{aligned}$$

by inserting j into all possible places inside $(0, n_1)$. Hence $A(2,1,\ldots,1)$ converges and by equations Eq. (5.5) and Eq. (5.8) we have

$$\begin{aligned} A(1,2,1,\ldots,1) &= B(1,2,1,\ldots,1,0) + \zeta(2,2,1,\ldots,1) \\ &= A(2,1,\ldots,1,1) + \zeta(2,2,1,\ldots,1). \end{aligned}$$

We can repeat this idea and show that all the sums $A(1,\ldots,1,2,1,\ldots,1)$ converge. □

The next result provides the *cyclic sum relations* of MZVs.

Theorem 5.2.3. *For any* $\mathbf{k} = (k_1, k_2, \ldots, k_d) \in \mathbb{N}^d$ *with some* $k_i \geq 2$,

$$\sum_{j=1}^{d} \zeta(k_j + 1, k_{j+1}, \ldots, k_d, k_1, \ldots, k_{j-1}) = \sum_{\{j:k_j \geq 2\}} \sum_{q=0}^{k_j - 2} \zeta(k_j - q, k_{j+1}, \ldots, k_d, k_1, \ldots, k_{j-1}, q+1). \quad (5.9)$$

Equivalently,

$$\sum_{(p_1,\ldots,p_d)\in[\mathbf{k}]} \zeta(p_1+1, p_2, \ldots, p_d) = \sum_{(q_1,\ldots,q_{|\mathbf{k}|-d})\in[\mathbf{k}^*]} \zeta(q_1+1, q_2, \ldots, q_{|\mathbf{k}|-d}). \quad (5.10)$$

Proof. We will prove Eq. (5.9) and leave the interested reader to check that Eq. (5.9) and Eq. (5.10) are equivalent (see Exercise 5.3). Let us first prove

$$\begin{aligned} &A(k_1,\ldots,k_d) - A(k_2,\ldots,k_d,k_1) \\ &\quad = \zeta(k_1+1, k_2, \ldots, k_d) - \sum_{j=0}^{k_1-2} \zeta(k_1 - j, k_2, \ldots, k_d, j+1) \end{aligned} \quad (5.11)$$

where the sum on the right is understood as 0 if $k_1 = 1$. We have

$$\begin{aligned} &A(k_1,\ldots,k_d) - \zeta(k_1+1, k_2, \ldots, k_d) \\ =&B(k_1,\ldots,k_d,0) && \text{(by Eq. (5.5))} \\ =&B(k_1-1,k_2,\ldots,k_d,1) - \zeta(k_1,\ldots,k_d,1) = \cdots && \text{(by Eq. (5.7))} \\ =&B(1,k_2,\ldots,k_d,k_1-1) - \sum_{j=0}^{k_1-2} \zeta(k_1-j,k_2,\ldots,k_d,j+1) && \text{(by Eq. (5.7))} \\ =&A(k_2,\ldots,k_d,k_1) - \sum_{j=0}^{k_1-2} \zeta(k_1-j,k_2,\ldots,k_d,j+1) && \text{(by Eq. (5.8)).} \end{aligned}$$

This proves Eq. (5.11). Equation Eq. (5.9) now follows from adding Eq. (5.11) over all cyclic permutations of the sequence $(k_1, \ldots, k_d)$. This completes the proof of the theorem. □

An interesting corollary of the cyclic sum relations provides the so-called *sum relations*:

Corollary 5.2.4. *For any positive integers w and d we have*

$$\sum_{k_1+\cdots+k_d=w} \zeta(k_1+1, k_2, \ldots, k_d) = \zeta(w+1). \tag{5.12}$$

Proof. Indeed, if $d = 1$ then Eq. (5.12) is trivial. Suppose it is true when the depth on the LHS is d. Applying cyclic permutation of the indices $(k_1, \ldots, k_d)$ to the LHS of Eq. (5.12) we get by Theorm 5.2.3

$$\begin{aligned} d\zeta(w+1) &= \sum_{|\mathbf{k}|=w} \sum_{j=1}^{d} \zeta(k_j+1, k_{j+1}, \ldots, k_d, k_1, \ldots, k_{j-1}) \\ &= \sum_{|\mathbf{k}|=w} \sum_{\{j|k_j\geq 2\}} \sum_{q=0}^{k_j-2} \zeta(k_j-q, k_{j+1}, \ldots, k_d, k_1, \ldots, k_{j-1}, q+1) \\ &= d \sum_{|\mathbf{k}|=w} \sum_{q=1}^{k_1-1} \zeta(k_1-q+1, k_2, \ldots, k_d, q) \\ &= d \sum_{|\mathbf{k}|+q=w} \zeta(k_1+1, k_2, \ldots, k_d, q). \end{aligned}$$

This proves the sum relation Eq. (5.12) by induction. □

With the notation of Hoffman algebras as in Sec. 3.2 we define for $k \geq 1$

$$\mathfrak{A}^1(k) = \langle \mathbf{w} : \text{ word } \mathbf{w} \in \{\mathtt{a}, \mathtt{b}\}^* \mathtt{ab}, |\mathbf{w}| = k \rangle_{\mathbb{Q}}.$$

So an element in $\mathfrak{A}^1(k)$ is a homogeneous polynomial of degree k with each monomial given by a word ending with $\mathtt{b}$ but not a $\mathtt{b}$-power. Define a $\mathbb{Q}$-linear map $\rho_0 : \bigcup_{k\geq 1} \mathfrak{A}^1(k) \to \mathfrak{A}^0$ by setting

$$\rho_0\big(\mathtt{y}_{k_1} \cdots \mathtt{y}_{k_\ell}\big) = \sum_{j=1}^{\ell} \sum_{i=1}^{k_j-1} \mathtt{y}_{k_j-i+1} \mathtt{y}_{k_{j+1}} \cdots \mathtt{y}_{k_{j+\ell-1}} \mathtt{y}_i - \sum_{j=1}^{\ell} \mathtt{y}_{k_j+1} \mathtt{y}_{k_{j+1}} \cdots \mathtt{y}_{k_{j+\ell-1}},$$

where $\mathtt{y}_m = \mathtt{y}_n$ if $m \equiv n \bmod \ell$. Define recursively

$$\rho_{n+1}(\mathbf{w}) = \rho_n\big((\mathtt{a}+\mathtt{b})\mathbf{w}\big) \quad \forall n \geq 0.$$

Recall we have the map $Z : \mathfrak{A}^0 \to \mathbb{R}$ defined by $\zeta(\mathbf{w}) = \int_0^1 \mathbf{w}$ (see Eq. (3.1)).

Theorem 5.2.5. *If $n \in \mathbb{N}_0$ and $k \in \mathbb{N}$, then*

$$\{0\} = \rho_{k-1}(\mathfrak{A}_1^1) \subset \rho_{k-2}(\mathfrak{A}_2^1) \subset \cdots \subset \rho_0(\mathfrak{A}_k^1) \subset \ker Z \cap \mathfrak{A}_{k+1}^0,$$

where $\mathfrak{A}_{k+1}^0$ is the homogeneous weight $k+1$ piece of $\mathfrak{A}^0$. Moreover, if $k > n$ then

$$\dim \rho_n(\mathfrak{A}_{k-n}^1) = \frac{1}{k}\sum_{m|k} \varphi\left(\frac{k}{m}\right)(2^m - L_{m,n}) - 2,$$

where φ is Euler's totient function and

$$L_{m,n} = \begin{cases} 2^m - 1, & \text{if } m = 1, \ldots, n; \\ L_{m-1,n} + \cdots + L_{m-n,n}, & \text{if } m \geq n+1. \end{cases}$$

5.3 Weighted sum relations

The following is a type of shuffle formula whose proof uses partial fractions. Another proof is outlined in Exercise 5.4.

Theorem 5.3.1. *For all positive integers $2 \leq s \leq w-2$ we have*

$$\sum_{k=2}^{w-1} \left\{ \binom{k-1}{s-1} + \binom{k-1}{w-s-1} \right\} \zeta(k, w-k) = \zeta(s)\zeta(w-s).$$

Proof. We start with an easy but crucial observation: for any positive integers m and n

$$\frac{1}{mn} = \frac{m+n}{(m+n)mn} = \frac{1}{(m+n)m} + \frac{1}{(m+n)n}. \tag{5.13}$$

Repeating this procedure one quickly find that for any positive integers s and t

$$\frac{1}{m^s n^t} = \sum_{j,k>0,\, j+k=s+t} \binom{k-1}{s-1} \frac{1}{(m+n)^k n^j} + \binom{k-1}{t-1} \frac{1}{(m+n)^k m^j}. \tag{5.14}$$

This follows from induction on w since

$$\begin{aligned} \frac{1}{m^{s+1}n^t} &= \frac{1}{m} \sum_{j,k>0,\, j+k=s+t} \binom{k-1}{s-1} \frac{1}{(m+n)^k n^j} + \binom{k-1}{t-1} \frac{1}{(m+n)^k m^j} \\ &= \sum_{j,k>0,\, j+k=s+t} \binom{k-1}{s-1} \frac{1}{(m+n)^k n^j m} + \binom{k-1}{t-1} \frac{1}{(m+n)^k m^{j+1}}. \end{aligned}$$

Thus we can finish the induction by the easily checked identity

$$\frac{1}{mn^j} = \frac{1}{(m+n)^j m} + \sum_{i=1}^{j} \frac{1}{(m+n)^{j-i+1} n^i}$$

using induction on j and Eq. (5.13). Now the theorem follows from taking the sum $\sum_{m,n=1}^{\infty}$ on both sides of Eq. (5.14). □

The next result is a weighted form of Euler's sum formula.

Theorem 5.3.2. *For all integer $w \geq 3$*

$$\sum_{k=2}^{w-1} 2^k \zeta(k, w-k) = (w+1)\zeta(w). \tag{5.15}$$

Proof. By stuffle relation and Thm. 5.3.1 the product $\zeta(s)\zeta(r)$ can be expressed in two different ways:

$$\zeta(r,s) + \zeta(s,r) + \zeta(r+s) = \sum_{k=2}^{r+s-1} \left\{ \binom{k-1}{r-1} + \binom{k-1}{s-1} \right\} \zeta(k, r+s-k). \tag{5.16}$$

For any positive integer d we define the generating function of MZVs by

$$G_d(x_1, \ldots, x_d) = \sum_{s_1, \ldots, s_d \in \mathbb{N}, s_1 > 1} x_1^{s_1 - 1} \cdots x_d^{s_d - 1} \zeta(s_1, \ldots, s_d). \tag{5.17}$$

Then by multiplying $x^{a-1}y^{b-1}$ on Eq. (5.16) and then summing over a and b we get

$$G_2(x+y, x) + G_2(x+y, y) = G_2(x, y) + G_2(y, x) + \frac{G_1(x) - G_1(y)}{x - y}, \tag{5.18}$$

Replacing (x, y) by (tx, ty) and considering the coefficients of t^{w-2} we get

$$\begin{aligned} &\sum_{k=2}^{w-1} (x+y)^{k-1} (x^{w-k-1} + y^{w-k-1}) \zeta(k, w-k) \\ &= \sum_{k=2}^{w-1} (x^{k-1} y^{w-k-1} + y^{k-1} x^{w-k-1}) \zeta(k, w-k) + \zeta(w) \sum_{i=0}^{w-2} x^i y^{w-2-i}. \end{aligned}$$

Taking $x = y = 1$ and using the sum relation Cor. 5.2.4 we arrive at Eq. (5.15). □

There are some other families of weighted sum relations implicitly related to the above. The results in the next theorem are called *restricted sum formulas*. But they can be regarded as weighted sum formulas with weight factors given by 0 or 1.

Theorem 5.3.3. *For all $n \geq 2$ we have*

$$\sum_{k=1}^{n-1} \zeta(2k, 2n-2k) = \frac{3}{4} \zeta(2n), \tag{5.19}$$

$$\sum_{k=1}^{n-1} \zeta(2k+1, 2n-2k-1) = \frac{1}{4} \zeta(2n). \tag{5.20}$$

Proof. Clearly Eq. (5.20) follows from Eq. (5.19) and the sum relation Eq. (5.12) which says in our case

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

Now we prove Eq. (5.19) using the following well-known formula for Bernoulli numbers (see for e.g., [137, p. 119]):

$$(2n+1)B_{2n} = -\sum_{k=1}^{n-1} \binom{2n}{2k} B_{2k} B_{2n-2k}, \qquad \forall n \geq 2. \tag{5.21}$$

By Euler famous evaluation Eq. (1.2) we get

$$\begin{aligned}\sum_{k=1}^{n-1} \zeta(2k)\zeta(2n-2k) &= \frac{(2\pi i)^{2n}}{4(2n)!} \sum_{k=1}^{n-1} \binom{2n}{2k} B_{2k} B_{2n-2k} \\ &= \frac{(2\pi i)^2 n}{4(2n)!}(2n+1)B_{2n} = \frac{2n+1}{2}\zeta(2n).\end{aligned}$$

Now adding the following stuffle relation

$$\zeta(2k)\zeta(2n-2k) = \zeta(2k, 2n-2k) + \zeta(2n-2k, 2k) + \zeta(2n) \tag{5.22}$$

for $k = 1$ up to $n-1$ we get

$$2\sum_{k=1}^{n-1} \zeta(2k, 2n-2k) + (n-1)\zeta(2n) = \frac{2n+1}{2}\zeta(2n),$$

which yields Eq. (5.19) immediately. This proves the theorem. □

Theorem 5.3.4. *For all positive integer $n \geq 2$ we have*

$$\sum_{k=1}^{n-1} (4^k + 4^{n-k})\zeta(2k, 2n-2k) = \left(n + \frac{4}{3} + \frac{4^n}{6}\right)\zeta(2n). \tag{5.23}$$

For all positive integer $n \geq 2$ we have

$$\sum_{k=2}^{n-2} (2k-1)(2n-2k-1)\zeta(2k, 2n-2k) = \frac{3}{4}(n-3)\zeta(2n). \tag{5.24}$$

Proof. We have the key identity of Bernoulli polynomials (see [137, p. 119]):

$$\sum_{i=0}^{n} \binom{n}{i} B_i(x) B_{n-i}(y) = n(x+y-1)B_{n-1}(x+y) - (n-1)B_n(x+y) \tag{5.25}$$

which can be proved by observing that (setting $z = x + y$)

$$\frac{te^{xt}}{e^t-1}\frac{te^{yt}}{e^t-1} = \frac{t^2e^{zt}}{(e^t-1)^2} = \frac{(z-1)t^2e^{zt}}{e^t-1} - t^2\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{e^{zt}}{e^t-1}\right).$$

Taking $x = 1/2$ and $y = 0$ and using the fact that $B_n(1/2) = -(1-2^{1-n})B_n$ we get

$$\sum_{k=1}^{n-1}\binom{2n}{2k}(2^{1-2k}-1)B_{2k}B_{2n-2k} = 2n(1-2^{1-2n})B_{2n} - B_{2n}.$$

By Eq. (5.21) this reduces to

$$\sum_{k=1}^{n-1}\binom{2n}{2k}2^{1-2k}B_{2k}B_{2n-2k} = -(2+n4^{1-n})B_{2n}.$$

Rewriting this using zeta values according to Eq. (1.2) we get

$$\sum_{k=1}^{n-1}4^{1-k}\zeta(2k)\zeta(2n-2k) = (2+n4^{1-n})\zeta(2n).$$

Using stuffle relation Eq. (5.22) this quick leads to Eq. (5.23).

Let $f(t) = 1/(e^t-1) = \sum_{j=0}^{\infty}B_jt^{j-1}/j!$ then

$$f'(t) = \frac{-e^t}{(e^t-1)^2} = \sum_{j=0}^{\infty}(j-1)B_j\frac{t^{j-2}}{j!}$$

has only even powers of t. We have

$$f''(t) = -f'(t) + \frac{2e^t}{(e^t-1)^3} = \sum_{j=0}^{\infty}(j-1)(j-2)B_j\frac{t^{j-3}}{j!},$$

$$f'''(t) = f'(t) - \frac{6e^{2t}}{(e^t-1)^4} = \sum_{j=0}^{\infty}(j-1)(j-2)(j-3)B_j\frac{t^{j-4}}{j!}.$$

Therefore

$$t^4\big(f'(t)\big)^2 = \frac{t^4e^{2t}}{(e^t-1)^4} = \frac{t^4}{6}\big(f'(t) - f'''(t)\big).$$

Thus

$$\left(\sum_{j=0}^{\infty}(j-1)B_j\frac{t^j}{j!}\right)^2 = \frac{1}{6}\sum_{j=0}^{\infty}\left\{(j-1)B_j\frac{t^{j+2}}{j!} - (j-1)(j-2)(j-3)B_j\frac{t^j}{j!}\right\}.$$

Comparing the coefficients of t^k for $k \geq 4$ even we get

$$\sum_{i+j=k}(i-1)(j-1)B_iB_j\binom{k}{i} = \frac{1}{6}\Big((k-3)(k-1)kB_{k-2} - (k-1)(k-2)(k-3)B_k\Big).$$

Magically B_{k-2}-terms cancel each other on the two side since $B_2 = 1/6$ so we have

$$6 \sum_{i,j\geq 4, i+j=k} (i-1)(j-1)B_i B_j \binom{k}{i} = -(k-1)(k^2-5k-6)B_k. \quad (5.26)$$

Setting $k = 2n$, multiplying both sides by $(2\pi i)^{2n}/4$ and using Eq. (1.2) we get

$$6\sum_{k=2}^{n-2}(2k-1)(2n-2k-1)\zeta(2k)\zeta(2n-2k) = (n-3)(4n^2-1)\zeta(2n). \quad (5.27)$$

By the stuffle relation Eq. (5.22) and the fact

$$\sum_{k=2}^{n-2}(2k-1)(2n-2k-1) = \frac{(n-3)(2n^2-5)}{3}$$

we can finally deduce Eq. (5.24). This finishes the proof of the theorem. □

Theorem 5.3.5. *For all positive integer r and $n \geq 2r$ we have*

$$\begin{aligned}
&\sum_{k=r+1}^{n-r-1} \frac{(2n-2r-2)!}{(2k-r-1)!(2n-2k-r-1)!}\frac{B_{2k}}{2k}\frac{B_{2n-2k}}{2n-2k} \\
&= -\left(\frac{2(-1)^r r!(2n-2r-2)!}{(2n-r-1)!} + \frac{(r!)^2}{(2r+1)!}\right)\frac{B_{2n}}{2n} \qquad (5.28)\\
&- \sum_{k=\lceil r/2\rceil+1}^{r} 2\left\{(-1)^r\binom{r}{2k-r-1} + \binom{2n-2r-2}{2k-r-1}\right\}\frac{B_{2k}}{2k}\frac{B_{2n-2k}}{2n-2k}.
\end{aligned}$$

Proof. Define a zeta function

$$Z_r(s) = \sum_{\alpha,\beta=1}^{\infty} \alpha^r\beta^r(\alpha+\beta)^{-s} = \sum_{n=1}^{\infty}\sum_{\beta=1}^{n-1}\beta^r(n-\beta)^r n^{-s}$$

by substitution $n = \alpha + \beta$. Let $B_j(x)$ be the jth Bernoulli polynomial. Then we have

$$\begin{aligned}
Z_r(s) &= \sum_{j=0}^{r}\binom{r}{j}\sum_{n=1}^{\infty}(-1)^j n^{r-j-s}\sum_{\beta=1}^{n-1}\beta^{r+j} \\
&= \sum_{j=0}^{r}\binom{r}{j}\sum_{n=1}^{\infty}(-1)^j n^{r-j-s}\frac{B_{r+j+1}(n)-B_{r+j+1}}{r+j+1} \\
&= \sum_{j=0}^{r}\binom{r}{j}\sum_{n=1}^{\infty}(-1)^j n^{r-j-s}\sum_{k=0}^{r+j}\binom{r+j+1}{k}\frac{B_k}{r+j+1}n^{r+j+1-k} \\
&= \sum_{j=0}^{r}\binom{r}{j}\sum_{k=0}^{r+j}\binom{r+j+1}{k}(-1)^j\frac{B_k}{r+j+1}\zeta(s+k-2r-1).
\end{aligned}$$

Now exchange the order of summation and break it into three parts: (i) $\sum_{k=0}^{0}\sum_{j=0}^{r}$, (ii) $\sum_{k=1}^{r}\sum_{j=0}^{r}$, and (iii) $\sum_{j=1}^{r}\sum_{k=r+1}^{r+j}=\sum_{k=r+1}^{2r}\sum_{j=k-r}^{r}$.

(i) Define

$$f(x)=\sum_{j=0}^{r}\binom{r}{j}(-1)^j\frac{x^{r+j+1}}{r+j+1}.$$

Then

$$f'(x)=\sum_{j=0}^{r}\binom{r}{j}(-1)^j x^{r+j}=x^r(1-x)^r.$$

Thus using beta function $B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\,dx=\Gamma(\mathtt{a})\Gamma(\mathtt{b})/\Gamma(\mathtt{a}+\mathtt{b})$ we have

$$\sum_{j=0}^{r}\binom{r}{j}(-1)^j\frac{x^{r+j+1}}{r+j+1}=\int_0^1 x^r(1-x)^r\,dx=B(r+1,r+1)=\frac{(r!)^2}{(2r+1)!}.$$

(ii) We want to show this sum is 0. Since $\binom{r+j+1}{k}/(r+j+1)$ is a polynomial of j of degree $k-1$ for all $k\geq 1$ it suffices to prove that for all nonnegative integer $d<r$ we have $g(1)=0$ where

$$g(x)=\sum_{j=0}^{r}j^d\binom{r}{j}(-1)^j x^j=\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^d\sum_{j=0}^{r}\binom{r}{j}(-1)^j x^j=\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^d(1-x)^r.$$

Since $r>d$ by Leibnitz rule it is now obvious that $g(1)=0$.

(iii) For each fixed integer $k\in[r+1,2r]$ we claim that

$$kS_{iii}=\sum_{j=k-r}^{r}(-1)^j\binom{r}{j}\binom{r+j}{k-1}=(-1)^r(1+(-1)^k)\binom{r}{k-r-1}. \quad (5.29)$$

To show this we set

$$h(x):=\sum_{k=r+1}^{2r}\sum_{j=k-r}^{r}(-1)^j\binom{r}{j}\binom{r+j}{k-1}x^k=\sum_{j=1}^{r}\sum_{k=r+1}^{r+j}(-1)^j\binom{r}{j}\binom{r+j}{k-1}x^k.$$

For a polynomial $f(x)$ we denote by $\mathrm{Trunc}_d f(x)$ the truncation of $f(x)$ by

throwing away all the terms of degree no greater than d. Then we see that

$$\begin{aligned}
h(x) =&\mathrm{Trunc}_r \sum_{j=1}^{r}(-1)^j\binom{r}{j}\sum_{k=1}^{r+j}\binom{r+j}{k-1}x^k\\
=&\mathrm{Trunc}_r \sum_{j=0}^{r}(-1)^j\binom{r}{j}\sum_{k=0}^{r+j-1}\binom{r+j}{k}x^{k+1}\\
=&\mathrm{Trunc}_r \sum_{j=0}^{r}(-1)^j\binom{r}{j}\Big(x(x+1)^{r+j}-x^{r+j+1}\Big)\\
=&\mathrm{Trunc}_r\Big\{x(x+1)^r(-x)^r-x^{r+1}(1-x)^r\Big\}\\
=&(-1)^r x^{r+1}\Big\{(x+1)^r-(x-1)^r\Big)\Big\}.
\end{aligned}$$

Hence Eq. (5.29) follows immediately. Putting (i), (ii) and (iii) together we get

$$\begin{aligned}
Z_r(s) =& \frac{(r!)^2}{(2r+1)!}\zeta(s-2r-1)\\
&+(-1)^r\sum_{k=r+1}^{2r}(1+(-1)^k)\binom{r}{k-r-1}\frac{B_k}{k}\zeta(s+k-2r-1). \quad (5.30)
\end{aligned}$$

On the other hand Eie showed that functions like $Z_r(s)$ have analytic continuations over the whole complex plan and further they are defined at negative integers (see [179, 180]). In fact the expressions in (i) to (iii) shows clearly that this can be done using Riemann zeta functions. More importantly, Eie showed that these special values at negative integers can also be computed using some integrals over clearly specified simplices. In our situation his theory implies that

$$Z(2r+2-2n) = J^2(x^r y^r (x+y)^{2n-2r-2})+2J^1\left(\int_0^{-x} x^r y^r (x+y)^{2n-2r-2}\,\mathrm{d}y\right),$$

by symmetry between x and y, where for positive integers a_j $(j=1,\dots,m)$

$$J^m(x_1^{a_1}\dots x_m^{a_m}) = \prod_{j=1}^{m}(-1)^{a_j}\frac{B_{a_j+1}}{a_j+1}. \quad (5.31)$$

For the integral we may use substitution $y=-xt$ to get

$$\begin{aligned}
\int_0^{-x} x^r y^r (x+y)^{2n-2r-2} =&(-1)^{r+1}x^{2n-1}\int_0^1 t^r(1-t)^{2n-2r-2}\,\mathrm{d}t\\
=&(-1)^{r+1}x^{2n-1}\frac{r!(2n-2r-2)!}{(2n-r-1)!}.
\end{aligned}$$

Applying operator J we get

$$Z(2r+2-2n) = \sum_{a+b=2n-2r-2} \frac{(2n-2r-2)!}{a!b!} \frac{B_{a+r+1}}{a+r+1} \frac{B_{b+r+1}}{b+r+1} + \frac{2(-1)^r r!(2n-2r-2)!}{(2n-r-1)!} \frac{B_{2n}}{2n}.$$

Comparing with evaluation of Eq. (5.30) at $s = 2r+2-2n$ we can now finish the proof of the theorem by noticing finally that in Eq. (5.28) if r is odd and $k = (r+1)/2$ then the term happens to be zero. Thus we can improve the lower limit of k from $\lceil (r+1)/2 \rceil$ to $\lceil r/2 \rceil + 1$. □

Remark 5.3.6. In [6] Agoh and Dilcher gave a formula for $(B_k + B_m)^n$ (with classical umbral calculus notation) which leads to very similar results when $k = m$.

Example 5.3.7. Taking $r = 3$ in Thm. 5.3.5 we get for all $n \geq 2r$

$$\sum_{k=4}^{n-4} \frac{(2n-8)!}{(2k-4)!(2n-2k-4)!} \frac{B_{2k}}{2k} \frac{B_{2n-2k}}{2n-2k} = -\frac{(n-6)(2n+1)(2n^2-11n+35)}{140(n-2)(n-3)(2n-5)(2n-7)} \frac{B_{2n}}{2n} - \frac{2n-11}{6} B_6 B_{2n-6}.$$

Theorem 5.3.8. *For all positive integer r and $n \geq 2r$ we have*

$$\sum_{k=r+1}^{n-r-1} \zeta(2k)\zeta(2n-2k) \prod_{\alpha+1}^{r} \Big\{ (2k-\alpha)(2n-2k-\alpha) \Big\}$$

$$= \left((-1)^r r! + \frac{(r!)^2}{2(2r+1)!} \prod_{\beta=r+1}^{2r+1} (2n-\beta) \right) \zeta(2n) \prod_{\alpha=1}^{r} (2n-\alpha) \tag{5.32}$$

$$- \sum_{k=\lceil r/2 \rceil+1}^{r} \prod_{\alpha=1}^{r} (2k-\alpha) \prod_{\beta=2k+1}^{2r+1} (2n-\beta) \times \times \left\{ \frac{2(-1)^r r!}{(2r-2k+1)!} + \frac{2(2n-2r-2)!}{(2n-2k-r-1)!} \right\} \zeta(2k)\zeta(2n-2k).$$

Proof. This follows from Thm. 5.3.5 and Eq. (1.2) easily. □

Example 5.3.9. When $r = 1$ we recover the formula Eq. (5.27). When $r = 2$ we find for all $n \geq 4$

$$\sum_{j,k \geq 3, j+k=n} (2j-1)(2j-2)(2k-1)(2k-2)\zeta(2j)\zeta(2k) = \frac{1}{15}(n-1)(4n^2-1)(2n^2-13n+30)\zeta(2n) - 24(n-2)(2n-5)\zeta(4)\zeta(2n-4).$$

When $r=3$ for all $n \geq 6$

$$\begin{aligned}
&\sum_{j,k\geq 4, j+k=n} (2j-1)(2j-2)(2j-3)(2k-1)(2k-2)(2k-3)\zeta(2j)\zeta(2k)\\
=&\frac{1}{35}(n-6)(2n-3)(n-1)(4n^2-1)(2n^2-11n+35)\zeta(2n)\\
&-240(2n-11)(n-3)(2n-7)\zeta(6)\zeta(2n-6).
\end{aligned}$$

When $r=4$ for all $n \geq 8$

$$\begin{aligned}
&\sum_{j,k\geq 5, j+k=n} \zeta(2j)\zeta(2k)\prod_{\alpha=1}^{4}\Big\{(2j-\alpha)(2k-\alpha)\Big\}\\
=&\frac{8}{315}\binom{n-1}{2}(2n-3)(4n^2-1)(4n^4-72n^3+521n^2-1923n+3780)\zeta(2n)\\
&-960(n-3)(n-4)(2n-7)(2n-9)\zeta(6)\zeta(2n-6)\\
&-6720(n-4)(2n-9)(2n^2-25n+81)\zeta(8)\zeta(2n-8).
\end{aligned}$$

Corollary 5.3.10. *We have for all* $n \geq 4$

$$\begin{aligned}
&\sum_{k=3}^{n-3}\zeta(2k,2n-2k)\prod_{\alpha=1}^{2}\Big\{(2n-2k-\alpha)(2k-\alpha)\Big\}\\
=&\frac{1}{2}(57n^2-279n+366)\zeta(2n)-12(n-2)(2n-5)\zeta(4)\zeta(2n-4). \quad (5.33)
\end{aligned}$$

For all $n \geq 6$

$$\begin{aligned}
&\sum_{k=4}^{n-4}\zeta(2k,2n-2k)\prod_{\alpha=1}^{3}\Big\{(2n-2k-\alpha)(2k-\alpha)\Big\}\\
=&\frac{1}{2}(1005n^3-12222n^2+48243n-62946)\zeta(2n)\\
&-120(2n-11)(n-3)(2n-7)\zeta(6)\zeta(2n-6). \quad (5.34)
\end{aligned}$$

For all $n \geq 8$

$$\begin{aligned}
&\sum_{k=5}^{n-5}\zeta(2k,2n-2k)\prod_{\alpha=1}^{4}\Big\{(2n-2k-\alpha)(2k-\alpha)\Big\}\\
=&\frac{1}{2}(31116n^4-631800n^3+4846020n^2-16543800n+21168864)\zeta(2n)\\
&-480(n-3)(n-4)(2n-7)(2n-9)\zeta(6)\zeta(2n-6) \quad (5.35)\\
&-3360(n-4)(2n-9)(2n^2-25n+81)\zeta(8)\zeta(2n-8).
\end{aligned}$$

Corollary 5.3.11. *We have for all* $n \geq 2$

$$\sum_{k=1}^{n-1} k(n-k)\zeta(2k, 2n-2k) = \frac{n}{16}\zeta(2n) + \frac{2n-3}{4}\zeta(2)\zeta(2n-2). \quad (5.36)$$

Proof. Let S be the left-hand side of Eq. (5.36). By expanding $(2k-1)(2n-2k-1)$ in Eq. (5.24) we see that

$$\begin{aligned}&\sum_{k=1}^{n-1}(4k(n-k)-2n+1)\zeta(2k,2n-2k)\\ =&\frac{3}{4}(n-3)\zeta(2n)+(2n-3)(\zeta(2,2n-2)+\zeta(2n-2,2))\\ =&\frac{3-5n}{4}\zeta(2n)+(2n-3)\zeta(2)\zeta(2n-2)\end{aligned}$$

By Eq. (5.19) we have

$$\begin{aligned}4S =&(2n-1)\sum_{k=1}^{n-1}\zeta(2k,2n-2k)+\frac{3-5n}{4}\zeta(2n)+(2n-3)\zeta(2)\zeta(2n-2)\\ =&\frac{n}{4}\zeta(2n)+(2n-3)\zeta(2)\zeta(2n-2)\end{aligned}$$

which yields the corollary at once. □

Similarly we can replace the factor $k(n-k)$ in Cor. 5.3.11 by $k^r(n-k)^r$ for any positive integer r. For example, we have

Corollary 5.3.12. *We have for all* $n \geq 2$

$$\begin{aligned}\sum_{j+k=n,j,k>0} k^2j^2\zeta(2k,2j) =& \frac{n}{64}\zeta(2n)+\frac{(2n-3)(6n-5)}{16}\zeta(2)\zeta(2n-2)\\ &-\frac{3(2n-5)}{4}\zeta(4)\zeta(2n-4),\\ \sum_{j+k=n,j,k>0} k^3j^3\zeta(2k,2j) =& +\frac{45}{8}(2n-7)\zeta(6)\zeta(2n-6)\\ -\frac{1}{256}n(2n^2-3)\zeta(2n)+&\frac{1}{64}(2n-3)(28n^2-48n+21)\zeta(2)\zeta(2n-2)\\ +\frac{3}{16}(2n-5)(2n^2-25n+35)&\zeta(4)\zeta(2n-4),\\ \sum_{j+k=n,j,k>0} k^4j^4\zeta(2k,2j) =& -\frac{315}{4}(2n-9))\zeta(8)\zeta(2n-8)\\ -\frac{n(16n^2-17)}{1024}\zeta(2n)+&\frac{1}{256}(2n-3)(6n-5)(20n^2-36n+17)\zeta(2)\zeta(2n-2)\end{aligned}$$

$$+\frac{3}{64}(2n-5)(40n^3-420n^2+1050n-777)\zeta(4)\zeta(2n-4)$$
$$-\frac{15}{16}(2n-7)(8n^2-98n+189)\zeta(6)\zeta(2n-6).$$

Proof. Suppose $j+k=n$. Then we easily see that

$$\prod_{\alpha=1}^{2}\Big\{(2j-\alpha)(2k-\alpha)\Big\}=16j^2k^2+4(5-6n)jk+4(2n-1)(n-1).$$

The corollary follows from Eq. (5.19), Eq. (5.33), and Eq. (5.36) immediately. □

Equation Eq. (5.21) can be generalized as follows.

Theorem 5.3.13. *Let $N\in\mathbb{N}$ and define the* unsigned Stirling numbers of the first kind *by*

$$\sum_{k=1}^{N}c(N,k)x^k=x(x+1)\cdots(x+N-1). \tag{5.37}$$

Then for any positive integer $m\geq N$ we have

$$\sum_{\substack{j_1+\cdots+j_N=m\\ j_1,\ldots,j_N\geq 0}}\binom{m}{j_1,\cdots,j_N}B_{j_1}\cdots B_{j_N}$$
$$=(-1)^{N-1}N\binom{m}{N}\sum_{k=1}^{N}c(N,k)\frac{B_{m+k-N}}{m+k-N}.$$

Proof. For any function $\sum_{n\geq 0}a_nt^n$ of t we define its associated zeta function by $\sum_{n\geq 0}a_n/n^s$. Then for $F(t)=(1-t)^{-N}$ we know its associated zeta function is

$$Z_F(s)=\sum_{n_1=0}^{\infty}\cdots\sum_{n_N=0}^{\infty}(n_1+\cdots+n_N)^{-s}.$$

By Eie and Lai's result [181, Prop. 5] we see that

$$Z_F(-m)=(-1)^N\sum_{\substack{j_1+\cdots+j_N=m+N\\ j_1,\ldots,j_N\geq 0}}\binom{m}{j_1,\cdots,j_N}B_{j_1}\cdots B_{j_N}. \tag{5.38}$$

On the other hand, by the definition

$$F(t)=\frac{1}{(1-t)^N}=\sum_{n=0}^{\infty}\binom{-N}{n}(-t)^n=\sum_{n=0}^{\infty}\frac{t^n}{(N-1)!}\sum_{k=1}^{N}c(N,k)n^{k-1}.$$

The associated zeta function is

$$Z_F(s) = \sum_{n=0}^{\infty} \frac{1}{(N-1)!} \sum_{k=1}^{N} c(N,k) n^{k-1-s}.$$

Thus

$$Z_F(-m) = \sum_{k=1}^{N} \frac{c(N,k)}{(N-1)!} \frac{-B_{m+k}}{m+k}. \tag{5.39}$$

Changing the index m to $m-N$ we see that the theorem follows from Eq. (5.38) and Eq. (5.39) immediately. □

For example, putting $b_{a_1,\dots,a_k} = \prod_{j=1}^{k} \left(B_{2a_j}/(2a_j)!\right)$, we have

$$\sum_{\substack{a,b,c\geq 1,\\ a+b+c=2n}} (2n)! b_{a,b,c} = \binom{2n+2}{2} B_{2n} + n\left(n-\frac{1}{2}\right) B_{2n-2}, \tag{5.40}$$

$$\sum_{\substack{a,b,c,d\geq 1,\\ a+b+c+d=2n}} (2n)! b_{a,b,c,d} = -\binom{2n+3}{3} B_{2n} + \frac{4}{3} n^2 (2n-1) B_{2n-2}. \tag{5.41}$$

These clearly yield identities between Riemann zeta values at positive even integers. It is also related to the next general restricted sum relations.

Theorem 5.3.14. *Set $\zeta(0) = -1/2$. Then for all positive integer $d \leq n$ we have*

$$\sum_{\substack{j_1,\dots,j_d>0\\ j_1+\cdots+j_d=n}} \zeta(2j_1,\dots,2j_d)$$
$$= -\sum_{j=0}^{\lfloor \frac{d-1}{2} \rfloor} \frac{1}{2^{2d-3}(2j+1)B_{2j}} \binom{2d-2j-1}{d} \zeta(2j)\zeta(2n-2j).$$

Proof. See [284] for details. □

5.4 Derivation relations

Let $\mathfrak{A} = \mathbb{Q}\langle \mathtt{x}, \mathtt{y} \rangle$ be the noncommutative polynomial algebra as in Definition 3.2.1, regarded as a graded $\mathbb{Q}$-algebra with $\mathtt{a}$ and $\mathtt{b}$ both of degree one. For any word $\mathbf{w}$ of $\mathfrak{A}$, denote by $|\mathbf{w}|$ its total degree and by $\mathrm{dp}(\mathbf{w})$ the *depth* of $\mathbf{w}$ which is the number of occurrences of $\mathtt{b}$ in $\mathbf{w}$. Some authors use length here, but it has a established meaning in the combinatorics of words (which is called the *weight* in this book). We call $|\mathbf{w}| - \mathrm{dp}(\mathbf{w})$ the codepth of $\mathbf{w}$, which gives the number of occurrences of $\mathtt{a}$ in $\mathbf{w}$.

As in Definition 3.2.2 we let $\mathfrak{A}^1 = \mathbb{Q}\mathtt{e} \oplus \mathfrak{A}b$ and $\mathfrak{A}^0 = \mathbb{Q}\mathtt{e} \oplus \mathtt{a}\mathfrak{A}\mathtt{b}$. Denote by $\widehat{\mathfrak{A}}$, $\widehat{\mathfrak{A}}^1$ and $\widehat{\mathfrak{A}}^0$ the completions of the corresponding objects. Let τ be the anti-automorphism of $\widehat{\mathfrak{A}}$ exchanging $\mathtt{a}$ and $\mathtt{b}$. Applied to words, τ preserves weight and exchanges depth and codepth: note that $\widehat{\mathfrak{A}}^0$ (but not $\widehat{\mathfrak{A}}^1$) is closed under τ. It is easy to check that $\tau(z(\mathbf{s})) = z(\mathbf{s}^*)$ where $z(\mathbf{s}) = \mathtt{y}_{s_1} \cdots, \mathtt{y}_{s_d}$ for $\mathbf{s} = (s_1, s_2, \ldots, s_d)$.

As usual, by a derivation of $\widehat{\mathfrak{A}}$ we mean a map $\delta : \widehat{\mathfrak{A}} \to \widehat{\mathfrak{A}}$ (of graded rational vector spaces) such that $\delta(\mathbf{uv}) = \delta(\mathbf{u})\mathbf{v} + \mathbf{u}\delta(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in \widehat{\mathfrak{A}}$. It is straight-forward to see that the commutator of two derivations is a derivation, so the set of derivations of $\widehat{\mathfrak{A}}$ is a Lie algebra graded by the degree, denoted by $\mathrm{Der}(\widehat{\mathfrak{A}})$.

Lemma 5.4.1. *If δ is a derivation, then $\bar{\delta} := \tau\delta\tau$ is also a derivation (of the same degree).*

Proof. Let $\mathbf{s}, \mathbf{t} \in \mathbb{N}^d$, and $\mathbf{u} = z(\mathbf{s})\mathtt{a}^i, \mathbf{v} = z(\mathbf{t})\mathtt{a}^j \in \mathfrak{A}$. Then

$$\begin{aligned}
\tau\delta\tau(\mathbf{uv}) =& \tau\delta\Big(\tau\big(z(\mathbf{s})\mathtt{a}^i z(\mathbf{t})\mathtt{a}^j\big)\Big)\\
=& \tau\delta\Big(\tau\big(\mathtt{a}^j\big)\tau\big(z(\mathbf{t})\big)\tau\big(\mathtt{a}^i\big)\tau\big(z(\mathbf{s})\big)\Big)\\
=& \tau\delta\Big(\mathtt{b}^j z(\mathbf{t}^*)\mathtt{b}^i z(\mathbf{s}^*)\Big)\\
=& \tau\Big(\delta\big(\mathtt{b}^j z(\mathbf{t}^*)\big)\mathtt{b}^i z(\mathbf{s}^*)\Big) + \tau\Big(\mathtt{b}^j z(\mathbf{t}^*)\delta\big(\mathtt{b}^i z(\mathbf{s}^*)\big)\Big)\\
=& z(\mathbf{s})\mathtt{a}^i\tau\Big(\delta\tau\big(z(\mathbf{t})\mathtt{a}^j\big)\Big) + \tau\Big(\delta\tau\big(z(\mathbf{s})\mathtt{a}^i\big)\Big)z(\mathbf{t})\mathtt{a}^j\\
=& \mathbf{u}\tau\delta\tau(\mathbf{v}) + \tau\delta\tau(\mathbf{u})\mathbf{v},
\end{aligned}$$

as desired. □

Lemma 5.4.2. *There is an isomorphism*

$$\begin{aligned}
\exp : \mathrm{Der}(\widehat{\mathfrak{A}}) &\longrightarrow \mathrm{Aut}(\widehat{\mathfrak{A}})\\
\delta &\longmapsto \sum_{n=0}^{\infty} \frac{\delta^n}{n!},
\end{aligned}$$

whose inverse is given by

$$\begin{aligned}
\log : \mathrm{Aut}(\widehat{\mathfrak{A}}) &\longrightarrow \mathrm{Der}(\widehat{\mathfrak{A}})\\
\sigma &\longmapsto -\sum_{n=1}^{\infty} \frac{(1-\sigma)^n}{n},
\end{aligned}$$

To state the next theorem we now define $\mathfrak{d}$ to be the derivation on $\widehat{\mathfrak{A}}$ such that

$$\mathfrak{d}(\mathbf{w}) = \mathtt{b} \sqcup\!\sqcup \mathbf{w} - \mathtt{b}\mathbf{w} \quad \forall \mathbf{w} \in \widehat{\mathfrak{A}}. \tag{5.42}$$

It can be verified that (see Exercise 5.9) for all $n \geq 1$ and $\mathbf{w} \in \widehat{\mathfrak{A}}$

$$\mathfrak{d}(\mathtt{a}) = \mathtt{ab},\ \mathfrak{d}(\mathtt{b}) = \mathtt{b}^2,\ \text{and}\ \frac{1}{n!}\mathfrak{d}^n(\mathbf{w}) = \mathtt{b}^n \sqcup\!\sqcup \mathbf{w} - \mathtt{b}(\mathtt{b}^{n-1} \sqcup\!\sqcup \mathbf{w}). \tag{5.43}$$

Let $\mathfrak{z}$ be the (possibly infinite) $\mathbb{Q}$-linear span of $\mathtt{y}_k$ for $k \geq 1$. For all $z \in \mathfrak{z}$ we define the derivation δ_z on $\widehat{\mathfrak{A}}$ (see Exercise 5.10) such that

$$\delta_z(\mathtt{a}) = 0 \quad \text{and} \quad \delta_z(\mathtt{b}) = (\mathtt{a}+\mathtt{b})z, \tag{5.44}$$

and a map $\chi_z : \widehat{\mathfrak{A}}^1 \to \widehat{\mathfrak{A}}^1$ by

$$\chi_z(\mathbf{w}) = (1-z)\Big((1-z)^{-1} * \mathbf{w}\Big) \quad \forall \mathbf{w} \in \widehat{\mathfrak{A}}^1. \tag{5.45}$$

Lemma 5.4.3. (i). *The vector space $\mathfrak{z}$ becomes a commutative and associative algebra with respect to the multiplication $\circ$ defined by*

$$z \circ z' = z * z' - zz' - z'z, \quad \forall z, z' \in \mathfrak{z}.$$

(ii). *The $\mathbb{Q}$-linear map*

$$\gamma : X\mathbb{Q}\langle\!\langle X \rangle\!\rangle \longrightarrow \mathfrak{z}, \quad \gamma(X^k) = \mathtt{y}_k\ \forall k \geq 1, \tag{5.46}$$

is an algebra isomorphism for $\circ$. Moreover, for every $f \in X\mathbb{Q}\langle\!\langle X \rangle\!\rangle$ we have

$$\delta_{\gamma(f)}(\mathtt{a}) = 0, \quad \delta_{\gamma(f)}(\mathtt{b}) = (\mathtt{a}+\mathtt{b})\frac{f(\mathtt{a})}{\mathtt{a}}\mathtt{b}. \tag{5.47}$$

Proof. (i) By the definition of $*$ we see that

$$\mathtt{y}_k \circ \mathtt{y}_l = \mathtt{y}_k * \mathtt{y}_l - \mathtt{y}_k\mathtt{y}_l - \mathtt{y}_l\mathtt{y}_k = \mathtt{y}_{k+l}.$$

Thus (i) follows immediately.

(ii) The first sentence is straight-forward (see Exercise 5.11). To prove Eq. (5.47) we only need to assume $f(X) = X^k$ for some $k \geq 1$. Then $\gamma(X^k) = \mathtt{y}_k$ and by the definition Eq. (5.44)

$$\delta_{\mathtt{y}_k}(\mathtt{a}) = 0, \quad \delta_{\mathtt{y}_k}(\mathtt{b}) = (\mathtt{a}+\mathtt{b})\mathtt{y}_k = (\mathtt{a}+\mathtt{b})\mathtt{a}^{k-1}\mathtt{b} = (\mathtt{a}+\mathtt{b})\frac{f(\mathtt{a})}{\mathtt{a}}\mathtt{b}.$$

This completes the proof of the lemma. □

Proposition 5.4.4. *Let $z \in \mathfrak{z}$. Then we have*

(i). $\chi_z \in \mathrm{Aut}(\widehat{\mathfrak{A}}^1)$ *and*

$$\chi_z(\mathbf{w}) = \exp(\delta_t)(\mathbf{w}),\ t = \log_\circ(1+z) = -\sum_{n=1}^{\infty}(-1)^n \frac{z^{\circ n}}{n} \in \mathfrak{z}. \quad (5.48)$$

(ii). χ_z *can be extended to* $\mathrm{Aut}(\widehat{\mathfrak{A}})$ *with* $\chi_z(\mathtt{a}) = \mathtt{a}$ *and* $\chi_z(\mathtt{a}+\mathtt{b}) = (\mathtt{a}+\mathtt{b})(1-z)^{-1}$. *Further,* χ_z *induces an automorphism of* $\widehat{\mathfrak{A}}^0$.

Proof. (i). For any $t \in \mathfrak{z}$ we have (see Exercise 5.7)

$$\exp_*(t) = (2 - \exp_\circ(t))^{-1}. \quad (5.49)$$

Taking $t = \log_\circ(1+z)$ we thus get $\exp_*(t) = (1-z)^{-1}$. Also, for $t \in \mathfrak{z}$ and $\mathbf{w} \in \widehat{\mathfrak{A}}^1$ we have (see Exercise 5.8)

$$\exp(\delta_t)(\mathbf{w}) = (\exp_*(t))^{-1}(\exp_*(t) * \mathbf{w}) = (1-z)((1-z)^{-1} * \mathbf{w}) = \chi_z(\mathbf{w}). \quad (5.50)$$

(ii). By the definition $\delta_t(\mathtt{a}) = 0$ and $\delta_t(\mathtt{b}) = (\mathtt{a}+\mathtt{b})t$. So

$$\delta_t\big((\mathtt{a}+\mathtt{b})\mathbf{w}\big) = (\mathtt{a}+\mathtt{b})\big(t\mathbf{w} + \delta_t(\mathbf{w})\big).$$

By induction it is straight-forward to show that $\delta_t^n(\mathtt{b}) = (\mathtt{a}+\mathtt{b})t^{*n}$ for all $n \in \mathbb{N}$. Hence by Eq. (5.48) $\chi_z(\mathtt{a}) = \exp(\delta_t)(\mathtt{a}) = \mathtt{a}$ and $\chi_z(\mathtt{a}+\mathtt{b}) = \exp(\delta_t)(\mathtt{a}+\mathtt{b}) = (\mathtt{a}+\mathtt{b})\exp_*(t) = (\mathtt{a}+\mathtt{b})(1-z)^{-1}$ by Eq. (5.49). To show χ_z induces an automorphism of $\widehat{\mathfrak{A}}^0$ it suffices to show that $\chi_z(\mathbf{w}) \in \widehat{\mathfrak{A}}^0$ for all $\mathbf{w} \in \widehat{\mathfrak{A}}^0$ since clearly $\exp(-\delta_t)$ is the inverse. Now suppose $\mathbf{w} = \mathtt{a}\mathbf{u}\mathtt{b}$. Then $\chi_z(\mathtt{a}\mathbf{u}\mathtt{b}) = \mathtt{a}\chi_z(\mathbf{u})\chi_z(\mathtt{b})$. But $\chi_z(\mathtt{b}) = \mathtt{b} + \sum_{n=1}^{\infty}(\mathtt{a}+\mathtt{b})z^n$ so that each word ends with b. Hence $\chi_z(\mathbf{w}) \in \widehat{\mathfrak{A}}^0$.

This completes the proof of the proposition. □

For $n \geq 1$ let ∂_n be the derivation with

$$\partial_n(\mathtt{a}) = \mathtt{a}(\mathtt{a}+\mathtt{b})^{n-1}\mathtt{b}, \quad \text{and} \quad \partial_n(\mathtt{b}) = -\mathtt{a}(\mathtt{a}+\mathtt{b})^{n-1}\mathtt{b}. \quad (5.51)$$

Theorem 5.4.5. (Drivation Relations) *For all* $n \in \mathbb{N}$ *and word* $\mathbf{w} \in \mathfrak{A}^0$, $\zeta(\partial_n(\mathbf{w})) = 0$.

Proof. By Thm. 3.3.21 for all positive integer n one has

$$Z^{\sqcup\!\sqcup}(\mathtt{b}^n) = \rho \circ Z^*(\mathtt{b}^n).$$

Multiplying $\zeta(\mathbf{w})$ on both sides and using the fact that both $Z^{\sqcup\!\sqcup}$ and Z^* are homomorphisms extending ζ we get

$$Z^{\sqcup\!\sqcup}(\mathtt{b}^n \sqcup\!\sqcup \mathbf{w}) = \rho \circ Z^*(\mathtt{b}^n * \sqcup\!\sqcup\mathbf{w}) = Z^{\sqcup\!\sqcup}(\mathtt{b}^n * \mathbf{w}),$$

where the second equality follows from Thm. 3.3.21 again. Recall that we have the isomorphism Eq. (3.14): $\mathfrak{A}^1_{\sqcup\!\sqcup} \cong \mathfrak{A}^0_{\sqcup\!\sqcup}[T]$. Let's denote by $\mathrm{reg}_{\sqcup\!\sqcup}$ the

regularization map $\mathfrak{A}^1 \cong \mathfrak{A}^0_{ш}[T] \to \mathfrak{A}^0$ by specializing the map to $T = 0$. We see that

$$\mathrm{reg}_{ш}(\mathfrak{b}^n ш \mathbf{w}) = \frac{1}{n!}\mathrm{reg}_{ш}(\mathfrak{b}^{ш n} ш \mathbf{w}) = \frac{1}{n!}\mathrm{reg}_{ш}(\mathfrak{b})^n\mathbf{w} = 0.$$

Hence $\zeta \circ \mathrm{reg}_{ш}(\mathfrak{b}^n * \mathbf{w}) = 0$ for all positive integer n. So

$$\zeta\Big(\sum_{n=1}^{\infty}\mathrm{reg}_{ш}(\mathfrak{b}^n * \mathbf{w})u^n\Big) = 0 \tag{5.52}$$

where u is a formal variable. Define $\Delta_u \in \mathrm{Aut}(\widehat{\mathfrak{A}})$ by

$$\Delta_u = \exp(-\mathfrak{d}u) \circ \chi_{\mathfrak{b}u}. \tag{5.53}$$

Now multiplying Eq. (5.43) by u^n and summing over n yields (after replacing $\mathbf{w}$ by $\mathbf{v}$ there)

$$\exp(\mathfrak{d}u)(\mathbf{v}) = (1-\mathfrak{b}u)\Big((1-\mathfrak{b}u)^{-1} ш \mathbf{v}\Big), \quad \forall \mathbf{v} \in \widehat{\mathfrak{A}}. \tag{5.54}$$

Substituting $\Delta_u(\mathbf{w})$ for $\mathbf{v}$ we see that

$$(1-\mathfrak{b}u)\Big((1-\mathfrak{b}u)^{-1} ш \Delta_u(\mathbf{w})\Big) = \chi_{\mathfrak{b}u}(\mathbf{w}) = (1-\mathfrak{b}u)\Big((1-\mathfrak{b}u)^{-1} * \mathbf{w}\Big).$$

Dividing by $1-\mathfrak{b}u$ and then applying $\mathrm{reg}_{ш}$ we get

$$\mathrm{reg}_{ш}\left(\Delta_u(\mathbf{w})\right) = \sum_{n=0}^{\infty}\mathrm{reg}_{ш}(\mathfrak{b}^n * \mathbf{w})u^n.$$

Here we use the fact that $\mathrm{reg}_{ш}\left((1-\mathfrak{b}u)^{-1}\right) = 1$. Since $\mathbf{w} \in \widehat{\mathfrak{A}}^0$ and both $\exp(\mathfrak{d}u)$ and $\chi_{\mathfrak{b}u}$ induce automorphisms on $\widehat{\mathfrak{A}}^0$ (see Excercise 5.9 and Prop. 5.4.4(ii)) we see that $\Delta_u(\mathbf{w}) \in \widehat{\mathfrak{A}}^0$. Hence $\mathrm{reg}_{ш}\left(\Delta_u(\mathbf{w})\right) = \Delta_u(\mathbf{w})$ and Eq. (5.52) gives $\zeta(\Delta_u(\mathbf{w})) = \zeta(\mathbf{w})$ for all $\mathbf{w} \in \widehat{\mathfrak{A}}^0$. The theorem now follows from the following lemma. □

Lemma 5.4.6. *As an automorphism of* $\widehat{\mathfrak{A}}$

$$\Delta_u = \exp\left(\sum_{n=1}^{\infty}(-1)^n\frac{\partial_n}{n}u^n\right). \tag{5.55}$$

Proof. We only need to check their images on the generators $\mathfrak{a}$ and $\mathfrak{b}$ are the same. First, by Prop. 5.4.4(ii)

$$\begin{aligned}\Delta_u(\mathfrak{a}) &= \exp(-\mathfrak{d}u) \circ \chi_{\mathfrak{b}u}(\mathfrak{a}) = \exp(-\mathfrak{d}u)(\mathfrak{a}) \\ &= \sum_{n=0}^{n}(-1)^n\frac{\mathfrak{d}^n(\mathfrak{a})}{n!}u^n = \sum_{n=0}^{n}(-1)^n\mathfrak{a}\mathfrak{b}^n u^n = \mathfrak{a}(1+\mathfrak{b}u)^{-1}.\end{aligned} \tag{5.56}$$

Now by Eq. (5.43) we have $\exp(\mathfrak{d}u)(\mathtt{a}+\mathtt{b}) = (\mathtt{a}+\mathtt{b})(1-\mathtt{b}u)^{-1}$ (see Exercise 5.10) so that

$$\begin{aligned}\Delta_u(\mathtt{a}+\mathtt{b}) &= \exp(-\mathfrak{d}u)\circ\chi_{\mathtt{b}u}(\mathtt{a}+\mathtt{b})\\ &= \exp(-\mathfrak{d}u)\big((\mathtt{a}+\mathtt{b})(1-\mathtt{b}u)^{-1}\big) = \mathtt{a}+\mathtt{b}. \quad (5.57)\end{aligned}$$

Comparing to Prop. 5.4.4(ii) we see that

$$\Delta_u = \varepsilon\chi_{\mathtt{b}u}\varepsilon \tag{5.58}$$

where ε is the involution on $\widehat{\mathfrak{A}}$ such that $\varepsilon(\mathtt{a}) = \mathtt{a}+\mathtt{b}, \varepsilon(\mathtt{b}) = -\mathtt{b}$. Therefore, Prop. 5.4.4(i) yields

$$\Delta_u = \varepsilon\exp\big(\delta_{\log_\circ(1+\mathtt{b}u)}\big)\varepsilon.$$

Hence

$$\begin{aligned}\Delta_u(\mathtt{a}) =& \varepsilon\exp\big(\delta_{\log_\circ(1+\mathtt{b}u)}\big)(\mathtt{a}+\mathtt{b})\\ =& \varepsilon\exp\big(\delta_{\gamma(\log(1+uX))}\big)(\mathtt{a}+\mathtt{b}) && \text{(by Lemma 5.4.3(ii))}\\ =& \varepsilon\exp\left((\mathtt{a}+\mathtt{b})\frac{\log(1+\mathtt{a}u)}{\mathtt{a}}\mathtt{b}\right) && \text{(by Eq. (5.47))}\\ =& \exp\left(-a\frac{\log(1+(\mathtt{a}+\mathtt{b})u)}{\mathtt{a}+\mathtt{b}}\mathtt{b}\right)\\ =& \exp\left(\sum_{n=1}^{\infty}(-1)^n\frac{\mathtt{a}(\mathtt{a}+\mathtt{b})^{n-1}\mathtt{b}}{n}u^n\right)\\ =& \exp\left(\sum_{n=1}^{\infty}(-1)^n\frac{\partial_n}{n}u^n\right)(\mathtt{a}) && \text{(by Eq. (5.51))},\end{aligned}$$

and

$$\begin{aligned}\Delta_u(\mathtt{a}+\mathtt{b}) &= \varepsilon\exp\big(\delta_{\log_\circ(1+\mathtt{b}u)}\big)(\mathtt{a})\\ &= \varepsilon(\mathtt{a}) = \mathtt{a}+\mathtt{b} = \exp\left(\sum_{n=1}^{\infty}(-1)^n\frac{\partial_n}{n}u^n\right)(\mathtt{a}+\mathtt{b}),\end{aligned}$$

since $\partial_n(\mathtt{a}+\mathtt{b}) = 0$ for all $n \geq 1$. The theorem is now proved. □

Corollary 5.4.7. *For all $m \geq 0$ and $\mathbf{w} = \mathtt{a}\mathbf{w}_0 \in \mathfrak{A}^0$, the regularization map satisfies*

$$\mathrm{reg}_{ш}(\mathtt{b}^m\mathbf{w}) = \frac{(-1)^m}{m!}\mathfrak{d}^m(\mathbf{w}) = (-1)^m\mathtt{a}(\mathtt{b}^m ш \mathbf{w}_0). \tag{5.59}$$

Proof. First notice that $(1-\mathtt{b}u)^{-1} = \exp_{ш}(\mathtt{b}u)$. Taking $\mathbf{v} = \exp(-\mathfrak{d}u)(\mathbf{w})$ in Eq. (5.54) and multiplying $(1-\mathtt{b}u)^{-1}$ from the left, we have

$$(1-\mathtt{b}u)^{-1}\mathbf{w} = (1-\mathtt{b}u)^{-1} ш \exp(-\mathfrak{d}u)(\mathbf{w}).$$

Applying $\mathrm{reg}_{ш}$ we get

$$\begin{aligned}\mathrm{reg}_{ш}\Big((1-\mathtt{b}u)^{-1}\mathbf{w}\Big) &= \mathrm{reg}_{ш}\Big((1-\mathtt{b}u)^{-1}\Big)\,\mathrm{reg}_{ш}\Big(\exp(-\mathfrak{d}u)(\mathbf{w})\Big)\\ &= \mathrm{reg}_{ш}\Big(\exp(-\mathfrak{d}u)(\mathbf{w})\Big).\end{aligned}$$

Comparing the coefficients of u^m on both sides we prove the first equality in Eq. (5.59) easily.

Now observe that $1/m!\mathfrak{d}^m(\mathtt{a}) = \mathtt{a}\mathtt{b}^m$ and $1/m!\mathfrak{d}^m(\mathtt{b}) = \mathtt{b}^{m+1}$ (see Exercise 5.9) we get

$$\exp(\mathfrak{d}u)(\mathtt{a}) = \mathtt{a}(1-\mathtt{b}u)^{-1}, \qquad \exp(\mathfrak{d}u)(\mathtt{b}) = \mathtt{b}(1-\mathtt{b}u)^{-1}. \tag{5.60}$$

Taking $\mathbf{v} = \mathtt{a}\mathbf{w}_0$ in Eq. (5.54) and using the first identity of Eq. (5.60) (with $u \to -u$) we have

$$\exp(-\mathfrak{d}u)(\mathbf{w}) = \exp(-\mathfrak{d}u)(x)\exp(-\mathfrak{d}u)(\mathbf{w}_0) = \mathtt{a}\Big((1+\mathtt{b}u)^{-1} ш \mathbf{w}_0\Big).$$

The second equality in Eq. (5.59) follows by comparing the coefficients of u^m on both sides. We have finished the proof of the corollary. □

Define $D_n(\mathtt{a}) = 0$ and $D_n(\mathtt{b}) = \mathtt{a}^n\mathtt{b}$ for all $n \geq 1$. For each $n \geq 1$, define the linear maps σ_n and $\overline{\sigma_n} : \widehat{\mathfrak{A}} \to \widehat{\mathfrak{A}}$ as homogeneous components of degree n of the homomorphisms

$$\sigma = \exp\left(\sum_{n\geq 1}\frac{D_n}{n}\right) = \sum_{n=1}^{\infty}\sigma_n, \quad \overline{\sigma} = \tau\sigma\tau = \exp\left(\sum_{n\geq 1}\frac{\overline{D_n}}{n}\right) = \sum_{n=1}^{\infty}\overline{\sigma_n}.$$

Lemma 5.4.8. *Set* $\partial = \sum_{n=1}^{\infty}\frac{\partial_n}{n}$. *Then*

$$\exp(\partial) = \overline{\sigma}\sigma^{-1}. \tag{5.61}$$

Proof. We check the two automorphisms' images on the generators $\mathtt{a}$ and $\mathtt{a}+\mathtt{b}$ are the same. First note that $\exp(\partial) = \Delta_{-1}$. Thus by Eq. (5.56) and Eq. (5.57)

$$\exp(\partial)(\mathtt{a}) = \Delta_{-1}(\mathtt{a}) = \mathtt{a}(1-\mathtt{b})^{-1} \quad \text{and} \quad \exp(\partial)(\mathtt{a}+\mathtt{b}) = \mathtt{a}+\mathtt{b}. \tag{5.62}$$

Now we set

$$D = \sum_{n\geq 1}\frac{D_n}{n}.$$

Then $D(\mathtt{a}) = \sum_{n\geq 1} \frac{D_n(\mathtt{a})}{n} = 0$ and

$$D(\mathtt{b}) = \sum_{n\geq 1} \frac{D_n(\mathtt{b})}{n} = \sum_{n\geq 1} \frac{\mathtt{a}^n b}{n} = (-\log(1-\mathtt{a}))\mathtt{b}.$$

Hence $D^n(\mathtt{a}) = 0$ and $D^n(\mathtt{b}) = (-\log(1-\mathtt{a}))^n\mathtt{b}$ and therefore

$$\sigma(\mathtt{a}) = \exp(D)(\mathtt{a}) = \sum_{n\geq 0} \frac{D^n(\mathtt{a})}{n!} = \mathtt{a}, \tag{5.63}$$

so $\sigma^{-1}(\mathtt{a}) = \mathtt{a}$. Further,

$$\begin{aligned}\sigma(\mathtt{b}) = \exp(D)(\mathtt{b}) &= \sum_{n\geq 0} \frac{D^n(\mathtt{a})}{n!} \\ &= \sum_{n\geq 0} \frac{(-\log(1-\mathtt{a}))^n}{n!}\mathtt{b} = (1-\mathtt{a})^{-1}\mathtt{b} = \exp(\partial)(\mathtt{a}). \end{aligned} \tag{5.64}$$

Similarly $\sigma^{-1}(\mathtt{b}) = (1-a)b$. Thus

$$\overline{\sigma}\sigma^{-1}(\mathtt{a}) = \overline{\sigma}(\mathtt{a}) = \tau(\sigma(\mathtt{b})) = \mathtt{a}(1-\mathtt{b})^{-1}$$

and

$$\begin{aligned}\overline{\sigma}\sigma^{-1}(\mathtt{b}) = \overline{\sigma}((1-\mathtt{a})\mathtt{b}) &= \tau\sigma(\mathtt{a}(1-\mathtt{b})) \\ &= \tau\big(\mathtt{a}(1-(1-\mathtt{a})^{-1}\mathtt{b})\big) = (1-\mathtt{a}(1-\mathtt{b})^{-1})\mathtt{b}.\end{aligned}$$

Adding the above two equations we get $\overline{\sigma}\sigma^{-1}(\mathtt{a}+\mathtt{b}) = \mathtt{a}+\mathtt{b}$. By comparing with Eq. (5.62) this proves the lemma. □

We now prove the next theorem using derivations.

Theorem 5.4.9. (Ohno's relations) *Suppose* $\mathbf{s} = (s_1, \ldots, s_d)$ *with* $s_1 > 1$ *and its dual* $\mathbf{s}^*$ *has depth* d'. *Then for any integer* $n \geq 0$

$$\sum_{\mathbf{e}\in(\mathbb{N}_0)^d, |\mathbf{e}|=n} \zeta(\mathbf{s}+\mathbf{e}) = \sum_{\mathbf{e}'\in(\mathbb{N}_0)^{d'}, |\mathbf{e}'|=n} \zeta(\mathbf{s}^*+\mathbf{e}'). \tag{5.65}$$

Proof. By Eq. (5.63) andEq. (5.64) we see that

$$\begin{aligned}\sigma(\mathtt{y}_{s_1}\cdots\mathtt{y}_{s_d}) =& \mathtt{a}^{s_1-1}(1-\mathtt{a})^{-1}\mathtt{b}\mathtt{a}^{s_2-1}(1-\mathtt{a})^{-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}(1-\mathtt{a})^{-1}\mathtt{b} \\ =& \sum_{n=0}^{\infty} \sum_{\substack{e_1+\cdots+e_d=n \\ e_j\in\mathbb{N}_0}} \mathtt{y}_{s_1+e_1}\cdots\mathtt{y}_{s_d+e_d}.\end{aligned}$$

Thus

$$\sigma_n(\mathtt{y}_{s_1}\cdots\mathtt{y}_{s_d}) = \sum_{\substack{e_1+\cdots+e_d=n\\ e_j\in\mathbb{N}_0}} \mathtt{y}_{s_1+e_1}\cdots\mathtt{y}_{s_d+e_d}.$$

Hence, in terms of derivation, Ohno's relation says that

$$\zeta\big((\sigma_n - \sigma_n\tau)\mathbf{w}\big) = 0 \quad \forall \mathbf{w}\in\mathfrak{A}^0, n\geq 1.$$

By duality Eq. (5.4) we only need to show that

$$\zeta\big((\sigma_n - \overline{\sigma_n})\mathbf{w}\big) = 0 \quad \forall \mathbf{w}\in\mathfrak{A}^0, n\geq 1. \tag{5.66}$$

But by Lemma 5.4.8 we see that $\sigma - \overline{\sigma} = (1-\exp(\partial))\sigma$. Hence Eq. (5.66) follows from Thm. 5.4.5 componentwise. This finishes the proof of the theorem. □

5.5 Fixed weight, height and depth relations

Another important technique in the study of MZVs is to use the generating functions, which frequently lead to the generalized hypergeometric functions (see Appendix C for a brief overview and [500] for a comprehensive treatment). For example, we will use Lemma C.28 to prove the main theorem of this section.

First we need to introduce some new notation.

Definition 5.5.1. Let $\mathbf{s} = (s_1,\ldots,s_d)\in\mathbb{N}^d$. We define the *height* of $\mathbf{s}$ to be the number of components of $\mathbf{s}$ which are greater than 1, denoted by $\mathrm{ht}(\mathbf{s})$. For non-negative integers w, d and h set

$$\begin{aligned} I(w,d,h) =& \big\{\mathbf{s}=(s_1,\ldots,s_d)\in\mathbb{N}^d : |\mathbf{s}| = w, \mathrm{ht}(\mathbf{s}) = h\big\},\\ I_0(w,d,h) =& \big\{\mathbf{s}=(s_1,\ldots,s_d)\in\mathbb{N}^d : |\mathbf{s}| = w, \mathrm{ht}(\mathbf{s}) = h, s_1\geq 2\big\}. \end{aligned}$$

Theorem 5.5.2. *For any non-negative integers w, d and h let*

$$G_0(w,d,h) = \sum_{\mathbf{s}\in I_0(w,d,h)} \zeta(\mathbf{s}).$$

Then

$$\sum_{w,d,h\geq 1} G_0(w,d,h)x^{w-d-h}y^{d-h}z^{h-1} = \frac{1-\exp\left(\sum_{n=2}^{\infty}\frac{\zeta(n)}{n}S_n(x,y,z)\right)}{xy-z},$$

where

$$S_n(x,y,z) = x^n+y^n-\alpha^n-\beta^n, \quad \alpha,\beta = \frac{x+y\pm\sqrt{(x+y)^2-4z}}{2}.$$

Proof. The basic idea is to treat $G_0(w,d,h)$ as a special point in a one parameter family $G_0(w,d,h;t)$ and study the generating function of the whole family. The advantage is that we can now take derivatives and solve certain related differential equations.

Recall the one variable multiple polylog $\widetilde{\mathrm{Li}}_{\mathbf{s}}(t)$ defined by Eq. (3.19) and further set $\widetilde{\mathrm{Li}}_{\mathbf{s}}(t)=0$ if $\mathbf{s}$ is empty. Define

$$G(w,d,h;t)=\sum_{\mathbf{s}\in I(w,d,h)}\widetilde{\mathrm{Li}}_{\mathbf{s}}(t),\quad G_0(w,d,h;t)=\sum_{\mathbf{s}\in I_0(w,d,h)}\widetilde{\mathrm{Li}}_{\mathbf{s}}(t).$$

By Prop. 4.2.1 we get

$$t\frac{\mathrm{d}}{\mathrm{d}t}G_0(w,d,h;t)=G(w-1,d,h-1;t)-G_0(w-1,d,h-1;t)\\ +G_0(w-1,d,h;t), \tag{5.67}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(G(w,d,h;t)-G_0(w,d,h;t)\Big)=\frac{1}{1-t}G_0(w-1,d-1,h;t). \tag{5.68}$$

Consider the generating functions

$$\Psi(x,y,z;t)=1+\sum_{w,d,h\geq 0}G(w,d,h;t)x^{w-d-h}y^{d-h}z^h,$$

$$\Psi_0(x,y,z;t)=\sum_{w,d,h\geq 0}G_0(w,d,h;t)x^{w-d-h}y^{d-h}z^{h-1}.$$

Then Eq. (5.67) and Eq. (5.68) yield

$$t\frac{\mathrm{d}\Psi}{\mathrm{d}t}=\frac{\Psi-1-z\Psi_0}{y}+x\Psi_0,\quad (1-t)\frac{\mathrm{d}}{\mathrm{d}t}(\Psi-z\Psi_0)=y\Psi.$$

Thus we get by elimination of Ψ

$$t(1-t)\frac{\mathrm{d}^2\Psi_0}{\mathrm{d}t^2}+\Big(1-x-(1-x+y)t\Big)\frac{\mathrm{d}\Psi_0}{\mathrm{d}t}+(xy-z)\Psi_0=1.$$

Clearly $\Psi_0(0)=0$. To compute $\Psi_0'(0)$ first notice that $\lim_{t\to 0}\widetilde{\mathrm{Li}}_{\mathbf{k}}(t)/t=0$ unless $\mathrm{dp}(\mathbf{k})=\mathrm{ht}(\mathbf{k})=1$ in which case $\lim_{t\to 0}\widetilde{\mathrm{Li}}_{\mathbf{k}}(t)/t=1$. Thus $\Psi_0'(0)=\sum_{w=2}^{\infty}x^{w-2}=1/(1-x)$. Setting $\alpha+\beta=x+y$ and $\alpha\beta=z$ we see that

$$\Psi_0(x,y,z;t)=\frac{1}{xy-z}\Big(1-{}_2F_1(\alpha-x,\beta-x;1-x;t)\Big),$$

since they are both solutions of Eq. (C.3) satisfying the same initial conditions $y(0)=0, y'(0)=1/(1-x)$. Thus by Gauss summation theorem Eq. (C.2)

$$\Psi_0(x,y,z;1)=\frac{1}{xy-z}\left(1-\frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-\alpha)\Gamma(1-\beta)}\right).$$

Finally, the theorem follows from the expansion (see Eq. (3.17))

$$\Gamma(1-x)=\exp\left(\gamma x+\sum_{n=2}^{\infty}\frac{\zeta(n)}{n}x^n\right).$$

This concludes the proof of the theorem. □

Corollary 5.5.3. *For any non-negative integers w, d and h the quantity*

$$G_0(w,d,h) = \sum_{\mathbf{s}\in I_0(w,d,h)} \zeta(\mathbf{s})$$

can be expressed as a polynomial of $\zeta(2), \zeta(3), \ldots$ with rational coefficients.

Proof. This follows from the theorem immediately. □

Corollary 5.5.4. *For any positive integer m*

$$\zeta\big(\{2\}^m\big) = \frac{\pi^{2m}}{(2m+1)!}.$$

Proof. In the generating function $\Psi_0(x,y,z;1)$ if we set $x = y = 0$ then the only nonzero terms are given by $d = h$ and $w = 2d$ meaning that

$$\Psi_0(x,y,z;1) = \sum_{s=1}^{\infty} \zeta\big(\{2\}^s\big) z^{s-1}.$$

By the theorem we get

$$\begin{aligned}
\sum_{m=1}^{\infty} \zeta\big(\{2\}^m\big) z^{m-1} =& \frac{1}{z}\left(\exp\left(-\sum_{n=1}^{\infty}\frac{\zeta(2n)}{n}(-z)^n\right) - 1\right)\\
=& \frac{1}{z}\left(\prod_{k=1}^{\infty}\exp\left(-\sum_{n=1}^{\infty}\frac{(-z/k^2)^n}{n}\right) - 1\right)\\
=& \frac{1}{z}\left(\prod_{k=1}^{\infty}\left(1+\frac{z}{k^2}\right) - 1\right)\\
=& \frac{1}{z}\left(\frac{\sinh\pi\sqrt{z}}{\pi\sqrt{z}} - 1\right) = \sum_{m=1}^{\infty}\frac{\pi^{2m}}{(2m+1)!} z^{m-1}.
\end{aligned}$$

The corollary follows from this immediately. □

5.6 Zagier's 2-3-2 formula of MZVs

In this section we will prove a **2-3-2** formula of MZVs, *i.e.*, a formula which involves the MZVs of the form

$$L(a,b) := \zeta(\{2\}^a, 3, \{2\}^b), \qquad a, b \in \mathbb{N}_0. \tag{5.69}$$

Zagier first proved this formula which has an important application in Brown's proof of the Deligne-Ihara conjecture and part of Hoffman's Basis Conjecture (see Conjection 3.6.1) in [104]. The key step is to express the quantities in Thm. 5.6.4 using the hypergeometric series. The formula in the theorem then follows from some transformation relations of the hypergeometric series.

Lemma 5.6.1. *The generating function*

$$\begin{aligned} F_L(x,y) &:= \sum_{a,b=0}^{\infty} (-1)^{a+b+1} L(a,b) x^{2a+1} y^{2b+2} \\ &= \frac{\sin \pi x}{\pi} \frac{\mathrm{d}}{\mathrm{d}z}\bigg|_{z=0} {}_3F_2\left(\begin{matrix} y, -y, z \\ 1+x, 1-x \end{matrix}\right), \end{aligned}$$

where the hypergeometric series ${}_3F_2\left(\begin{smallmatrix}\alpha_1,\alpha_2,\alpha_3\\ \beta_1,\beta_2\end{smallmatrix}\right) = {}_3F_2\left(\begin{smallmatrix}\alpha_1,\alpha_2,\alpha_3\\ \beta_1,\beta_2\end{smallmatrix};1\right)$ *is defined in Appendix* C.

Proof. Using the definition of MZV and the Pochhammer symbol Eq. (1.13), we have

$$\begin{aligned} F_L(x,y) &= -xy^2 \sum_{n=1}^{\infty} \prod_{l>n} \left(1 - \frac{x^2}{l^2}\right) \frac{1}{n^3} \prod_{n>k>0} \left(1 - \frac{y^2}{k^2}\right) \\ &= \frac{\sin \pi x}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \frac{(-y)_n (y)_n}{(1-x)_n (1+x)_n} \end{aligned}$$

which can be easily transformed to the formula in Lemma 5.6.1. □

For any nonnegative integers a, b, define

$$\begin{aligned} R(a,b) := 2 \sum_{n=1}^{a+b+1} (-1)^n \left[\binom{2n}{2b+2} - \left(1 - \frac{1}{4^n}\right) \binom{2n}{2a+1} \right] \times \\ \times \zeta(2n+1) \zeta\big(\{2\}^{a+b+1-n}\big). \end{aligned}$$

Lemma 5.6.2. *The generating function*

$$F_R(x,y) = \sum_{a,b=0}^{\infty} (-1)^{a+b+1} R(a,b) x^{2a+1} y^{2b+2}$$

satisfies the relation

$$\begin{aligned} \frac{\pi \cdot F_R(x,y)}{\sin \pi x} &= \psi(1+x) + \psi(1-x) - \frac{1}{2} \sum_{\kappa,\lambda=\pm 1} \psi(1 + \kappa x + \lambda y) \\ &\quad - \frac{\sin \pi y}{2 \sin \pi x} \sum_{\kappa,\lambda=\pm 1} \kappa\lambda \left(\psi\Big(1 + \frac{\kappa x + \lambda y}{2}\Big) - \psi(1 + \kappa x + \lambda y) \right). \quad (5.70) \end{aligned}$$

Proof. By the definition and Cor. 5.5.4, we have

$$\begin{aligned}F_R(x,y) &= 2\sum_{k,n=1}^{\infty}\frac{(-1)^k\pi^{2k}}{(2k+1)!}\zeta(2n+1)\Big[\sum_{b=0}^{n-1}\binom{2n}{2b+2}x^{2k+2n-2b-1}y^{2b+2}\\ &\qquad -\Big(1-\frac{1}{4^n}\Big)\sum_{a=0}^{n-1}\binom{2n}{2a+1}y^{2k+2n-2a}x^{2a+1}\Big]\\ &= \frac{\sin\pi x}{\pi}\sum_{n=1}^{\infty}\zeta(2n+1)\big[(x+y)^{2n}+(x-y)^{2n}-2x^{2n}\big]\\ &\quad -\frac{\sin\pi y}{\pi}\sum_{n=1}^{\infty}\Big(1-\frac{1}{4^n}\Big)\zeta(2n+1)\big[(x+y)^{2n}-(x-y)^{2n}\big]\\ &= \frac{\sin\pi x}{\pi}\big[A(x+y)+A(x-y)-2A(x)\big]-\frac{\sin\pi y}{\pi}\big[B(x+y)-B(x-y)\big],\end{aligned}$$

where, for all $|z|<1$

$$A(z)=\sum_{n=1}^{\infty}\zeta(2n+1)z^{2n},\quad B(z)=\sum_{n=1}^{\infty}\Big(1-\frac{1}{4^n}\Big)\zeta(2n+1)z^{2n}=A(z)-A(z/2).$$

Moreover, using the expansion Eq. (1.5) of the digamma function $\psi(x)$ we can see easily that

$$A(z)=\psi(1)-\frac{1}{2}\psi(1+z)-\frac{1}{2}\psi(1-z).$$

This finishes the proof of the lemma. □

We also need a result concerning the digamma function.

Lemma 5.6.3. *We have*

$$\frac{\Gamma(x)}{\Gamma(x-z)}=1+\psi(x)z+o(z).$$

Proof. By L'Hôpital's Rule, we have

$$\lim_{z\to 0}\frac{\Gamma(x)-\Gamma(x-z)}{\Gamma(x-z)z}=\lim_{z\to 0}\frac{\Gamma'(x-z)}{\Gamma'(x-z)z+\Gamma(x-z)}=\frac{\Gamma'(x)}{\Gamma(x)}=\psi(x).$$

□

We are now ready to prove Zagier's **2**-3-**2** formula of MZVs.

Theorem 5.6.4. *For any $a,b\in\mathbb{N}_0$, we have $L(a,b)=R(a,b)$. Namely,*

$$\zeta(\{2\}^a,3,\{2\}^b)=2\sum_{n=1}^{a+b+1}(-1)^n\left[\binom{2n}{2b+2}-\left(1-\frac{1}{4^n}\right)\binom{2n}{2a+1}\right]\times\times\zeta(2n+1)\zeta\big(\{2\}^{a+b+1-n}\big). \tag{5.71}$$

Proof. First we observe that

$$(y)_n(-y)_n = \frac{1}{2}(y)_n(1-y)_n + \frac{1}{2}(1+y)_n(-y)_n.$$

Thus

$${}_3F_2\left(\begin{matrix} y, -y, z \\ 1+x, 1-x \end{matrix}\right) = \frac{1}{2}\,{}_3F_2\left(\begin{matrix} y, 1-y, z \\ 1+x, 1-x \end{matrix}\right) + \frac{1}{2}\,{}_3F_2\left(\begin{matrix} 1+y, -y, z \\ 1+x, 1-x \end{matrix}\right). \tag{5.72}$$

Because of the symmetry $y \leftrightarrow -y$ on the right-hand side of Eq. (5.72) we only need to consider the first ${}_3F_2$-term. Applying the transformation formula Eq. (C.6) with $\alpha_1 = y, \alpha_2 = z, \alpha_3 = 1-y, \beta_1 = 1+x, \beta_2 = 1-x$, we get

$$\begin{aligned} {}_3F_2\left(\begin{matrix} y, 1-y, z \\ 1+x, 1-x \end{matrix}\right) &= \frac{\Gamma(1+x)\Gamma(1+x-y-z)}{\Gamma(1+x-y)\Gamma(1+x-z)}\,{}_3F_2\left(\begin{matrix} y, y-x, z \\ y-x+z, 1-x \end{matrix}\right) \\ &\quad + \frac{\Gamma(1+x)\Gamma(1-x)\Gamma(y-x+z-1)\Gamma(1-z)}{\Gamma(y)\Gamma(z)\Gamma(y-x)\Gamma(2-y-z)}\times \\ &\qquad \times\,{}_3F_2\left(\begin{matrix} 1+x-y, 1+x-z, 1-z \\ 2+x-y-z, 2-y-z \end{matrix}\right). \end{aligned} \tag{5.73}$$

Applying the transformation Eq. (C.7) to the first ${}_3F_2$-term on the right-hand side of Eq. (5.73) with $\alpha_1 = y, \alpha_2 = y-x, \alpha_3 = z, \beta_1 = y-x+z, \beta_2 = 1-x$, we get

$$\begin{aligned} {}_3F_2\left(\begin{matrix} y, 1-y, z \\ 1+x, 1-x \end{matrix}\right) &= \frac{\Gamma(1+x)}{\Gamma(1+x-z)}\frac{\Gamma(1-x)}{\Gamma(1-x-z)}\frac{\Gamma(1-x-y)}{\Gamma(1-x-y+z)}\times \\ &\qquad \times \frac{\Gamma(1+x-y-z)}{\Gamma(1+x-y)}\,{}_3F_2\left(\begin{matrix} -x+z, z, z \\ y-x+z, 1-x-y+z \end{matrix}\right) \\ &\quad + \frac{x\Gamma(x)\Gamma(1-x)}{\Gamma(y)\Gamma(2-y-z)}\frac{\Gamma(y-x-1+z)}{\Gamma(y-x)}\times \\ &\qquad \times \frac{\Gamma(1-z)}{\Gamma(z)}\,{}_3F_2\left(\begin{matrix} 1+x-y, 1+x-z, 1-z \\ 2+x-y-z, 2-y-z \end{matrix}\right) \\ &= [1+\psi(1+x)z][1+\psi(1-x)z][1-\psi(1-x-y)z][1-\psi(1+x-y)z] \\ &\quad + \frac{x}{1-y}\frac{\sin\pi y}{\sin\pi x}\frac{z}{y-x-1}\,{}_3F_2\left(\begin{matrix} 1+x-y, 1+x, 1 \\ 2+x-y, 2-y \end{matrix}\right) + o(z) \end{aligned}$$

as $z \to 0$ by Lemma 5.6.3. Here we have also used the reflection formula

$$\Gamma(t)\Gamma(1-t) = \frac{\pi}{\sin\pi t}$$

for the Γ-product coefficient of the second ${}_3F_2$-term.

Now taking $\alpha = 1 + x - y, \beta = 1 + x$ in the summation formula of Prop. C.30, we get

$$\left.\frac{\mathrm{d}}{\mathrm{d}z}\right|_{z=0} {}_3F_2\left(\begin{matrix} y, 1-y, z \\ 1+x, 1-x \end{matrix}\right) = \psi(1+x)+\psi(1-x)-\psi(1+x-y)-\psi(1-x-y)$$
$$-\frac{\sin \pi y}{\sin \pi x} \cdot \Big[\psi(1+x-y)-\psi(1-x-y)-\psi\Big(1+\frac{x-y}{2}\Big)+\psi\Big(1-\frac{x+y}{2}\Big)\Big]. \tag{5.74}$$

Combining Eq. (5.72) and Eq. (5.74) we see that $F_L(x, y) = F_R(x, y)$ which yields the theorem immediately. □

If $n_1, \ldots, n_d \in \mathbb{N}$, we have defined the following shuffle-regularized MZV (see Prop. 3.2.4):

$$\zeta_k(n_1, \ldots, n_d) := (-1)^d \zeta_{ш}(\mathtt{a}^{n_1-1} \cdots \mathtt{ba}^{n_d-1}\mathtt{ba}^k).$$

Corollary 5.6.5. *For any $n \in \mathbb{N}$, we have*

$$\zeta_1\big(\{2\}^n\big) = 2\sum_{i=1}^{n}(-1)^i \zeta(2i+1)\zeta\big(\{2\}^{n-i}\big).$$

Proof. By Thm. 4.1.14 we have

$$\zeta_1\big(\{2\}^d\big) = 2\sum_{a=0}^{d-1}\zeta\big(\{2\}^a, 3, \{2\}^{d-a-1}\big)$$
$$= 4\sum_{i=1}^{d}(-1)^i\zeta(2i+1)\zeta\big(\{2\}^{d-i}\big)\sum_{a=0}^{d-1}\left[\binom{2i}{2d-2a} - \left(1-\frac{1}{4^i}\right)\binom{2i}{2a+1}\right]$$

by Thm. 5.6.4. So the corollary follows from the well-known combinatorial identities:

$$\sum_{b=0}^{i-1}\binom{2i}{2b+2} = 2^{2n-1}-1, \quad \sum_{a=0}^{i-1}\binom{2i}{2a+1} = 2^{2n-1},$$

which can be proved, for e.g., by using generating functions. □

5.7 An exotic shuffle relation

Using shuffle relations of MZVs we can prove the following identity. Usually, the shuffle product is defined on the algebra of words. But in the theorem (but not in the proof) we treat the numbers 1,2 and 3 as letters and then extend the zeta function ζ linearly over $\mathbb{Z}$.

Theorem 5.7.1. *For any nonnegative integers m and n,*

$$\zeta\big(\{2\}^m \sqcup\!\sqcup \{3,1\}^n\big) = \binom{2n+m}{m} \frac{\pi^{4n+2m}}{(2n+1)\cdot(4n+2m+1)!}. \tag{5.75}$$

Proof. The case $n = 0$ is just Cor. 5.5.4: for every positive integer r

$$\zeta(\{2\}^r) = \frac{\pi^{2r}}{(2r+1)!}. \tag{5.76}$$

Applying the evaluation map ζ defined in Prop. 3.2.4 to the two equations in Prop. 3.2.7 and using Eq. (5.76) we see that

$$\sum_{k=1-n}^{n} \frac{(-1)^k k^{2m-1}\pi^{4n-2}}{(2n+2k-1)!(2n-2k+1)!} = 4^n \sum_{j=1}^{m} x_{2n-1,n-j}\cdot a_{2m-1,j}, \tag{5.77}$$

$$\sum_{k=-n}^{n} \frac{(-1)^k k^{2m}\pi^{4n}}{(2n+2k+1)!(2n-2k+1)!} = 4^n \sum_{j=1}^{m} x_{2n,n-j}\cdot a_{2m,j}, \tag{5.78}$$

where by setting $a = \frac{\mathrm{d}t}{t}$, $\quad b = \frac{\mathrm{d}t}{1-t}$

$$x_{2n-1,n-j} = \int_0^1 T_{2n-1,n-j}, \qquad x_{2n,n-j} = \int_0^1 T_{2n,n-j}.$$

Observe that $\zeta(\{2\}^r) = \zeta((\mathtt{ab})^r)$ and $\zeta(\{3,1\}^r) = \zeta((\mathtt{a}^2\mathtt{b}^2)^r)$ so it is not too hard to see that

$$x_{2n-1,n-j} = \zeta(\{2\}^{2j-1} \sqcup\!\sqcup \{3,1\}^{n-j}), \qquad x_{2n,n-j} = \zeta(\{2\}^{2j} \sqcup\!\sqcup \{3,1\}^{n-j}).$$

Note that for arbitrary fixed n, $x_{2n-1,n-j}$ (resp. $x_{2n,n-j}$) are recursively defined by Eq. (5.77) (resp. Eq. (5.78)) when we take $m = 1,\ldots,n-1$. Thus Thm. 5.7.1 is quickly reduced to the following combinatorial identities: For all $m, n \geq 1$ we have

$$\sum_{k=1-n}^{n} \frac{(-1)^k k^{2m-1}}{2n-2k+1}\binom{4n-1}{2n+2k-1} = \sum_{j=1}^{m}\sum_{k=1-j}^{j} \frac{4^{n-j}(-1)^k k^{2m-1}}{2n-2j+1}\binom{2n-1}{2j-1}\binom{2j-1}{j-k}, \tag{5.79}$$

$$\sum_{k=-n}^{n} \frac{(-1)^k k^{2m}}{2n-2k+1}\binom{4n+1}{2n+2k+1} = \sum_{j=1}^{m}\sum_{k=-j}^{j} \frac{4^{n-j}(-1)^k k^{2m}}{2n-2j+1}\binom{2n}{2j}\binom{2j}{j-k}. \tag{5.80}$$

Now observe that we may replace the outer sum of the right-hand side in both Eq. (5.79) and Eq. (5.80) by $\sum_{j=1}^{\infty}$ by the following computation: if $j > m$ then

$$\sum_{k=1-j}^{j} (-1)^k k^{2m-1} \binom{2j-1}{j-k} = \left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{2m-1} \sum_{k=1-j}^{j} (-1)^k x^k \binom{2j-1}{j-k}\Big|_{x=1}$$
$$= (-1)^{1-j} \left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{2m-1} x^{1-j}(1-x)^{2j-1}\Big|_{x=1} = 0.$$

Similarly, if $j > m$ then

$$\sum_{k=-j}^{j} (-1)^k k^{2m} \binom{2j}{j-k} = 0.$$

We therefore only need to prove that for all $m, n \geq 1$

$$\sum_{k=1-n}^{n} \frac{(-1)^k k^m}{2n-2k+1} \binom{4n-1}{2n+2k-1}$$
$$= \sum_{j=1}^{\infty} \sum_{k=1-j}^{j} \frac{4^{n-j}(-1)^k k^m}{2n-2j+1} \binom{2n-1}{2j-1} \binom{2j-1}{j-k}, \tag{5.81}$$
$$\sum_{k=-n}^{n} \frac{(-1)^k k^m}{2n-2k+1} \binom{4n+1}{2n+2k+1}$$
$$= \sum_{j=1}^{\infty} \sum_{k=-j}^{j} \frac{4^{n-j}(-1)^k k^m}{2n-2j+1} \binom{2n}{2j} \binom{2j}{j-k}. \tag{5.82}$$

In fact, we only need the case m is odd in Eq. (5.81) and m is even in Eq. (5.82).

It is a well-known fact that $x\mathbb{Q}[x]$ have two bases over $\mathbb{Q}$:$\{x^m : m \geq 1\}$ and $\{\binom{x}{m} : m \geq 1\}$. So Thm. 5.7.1 follows immediately from the following Prop. 5.7.2. Note that we will exchange the index j and k on the right-hand side. $\square$

Proposition 5.7.2. *For all $m, n \geq 1$ we have*

$$\sum_{k=1-n}^{n} \frac{(-1)^k}{2n-2k+1}\binom{k}{m}\binom{4n-1}{2n+2k-1} = \sum_{k=1}^{\infty}\sum_{j=1-k}^{k} \frac{4^{n-k}(-1)^j}{2n-2k+1}\binom{j}{m}\binom{2n-1}{2k-1}\binom{2k-1}{k-j}, \quad (5.83)$$

$$\sum_{k=-n}^{n} \frac{(-1)^k}{2n-2k+1}\binom{k}{m}\binom{4n+1}{2n+2k+1} = \sum_{k=1}^{\infty}\sum_{j=-k}^{k} \frac{4^{n-k}(-1)^j}{2n-2k+1}\binom{j}{m}\binom{2n}{2k}\binom{2k}{k-j}. \quad (5.84)$$

Proof. We first break the inner sum on the right-hand side of Eq. (5.83) as follows:

$$\sum_{j=1-k}^{k} (-1)^j\binom{j}{m}\binom{2k-1}{k-j} = \sum_{j=1}^{k}(-1)^j\binom{j}{m}\binom{2k-1}{k-j} + \sum_{j=1}^{k-1}(-1)^{m-j}\binom{m+j-1}{m}\binom{2k-1}{j+k}.$$

Re-indexing and using the Chu-Vandermonde identity (see [459, p. 182]) we change the above to

$$\sum_{j=0}^{k-m}(-1)^m\binom{-m-1}{j}\binom{2k-1}{k-m-j} + \sum_{j=0}^{k-2}(-1)^{m-1}\binom{-m-1}{j}\binom{2k-1}{k-2-j} = (-1)^m\binom{2k-m-2}{k-m} + (-1)^{m-1}\binom{2k-m-2}{k-2}.$$

Notice that the sum is 0 when $k \geq m$, however, when $k < m$ only the first term in the sum is always 0. By denoting the left-hand (resp. right-hand) side of Eq. (5.83) by $L(m,n)$ (resp. $R(m,n)$) we get

$$R(m,n) = \sum_{k=1}^{n} \frac{(-1)^m 4^{n-k}}{(2n-2k+1)}\binom{2n-1}{2k-1}\left(\binom{2k-m-2}{k-m} - \binom{2k-m-2}{k-2}\right). \quad (5.85)$$

Set

$$F_0(n,k) := \frac{(-1)^k}{2n-2k+1}\binom{k}{m}\binom{4n-1}{2n+2k-1},$$
$$F_1(n,k) := \frac{(-1)^{m+1}4^{n-k}}{2n-2k+1}\binom{2n-1}{2k-1}\binom{2k-m-2}{k-2},$$
$$F_2(n,k) := \frac{(-1)^m 4^{n-k}}{2n-2k+1}\binom{2n-1}{2k-1}\binom{2k-m-2}{k-m}.$$

Then clearly $\lim_{k\to\pm\infty} F_i(n,k) = 0$ for each n and

$$L(m,n) = \sum_{k\in\mathbb{Z}} F_0(n,k), \quad R(m,n) = \sum_{k\in\mathbb{Z}} F_1(n,k) + F_2(n,k).$$

Now by WZ-method (see [459]) we can find functions $G_i(n,k)$ such that all the three functions F_i above satisfy the following: for $i = 0, 1, 2$ and all k

$$c_3F_i(n+3,k) - c_2F_i(n+2,k) + c_1F_i(n+1,k) - c_0F_i(n,k) = G_i(n,k+1) - G_i(n,k), \quad (5.86)$$

where $\lim_{k\to\pm\infty} G_i(n,k) = 0$ for each fixed n and

$$\begin{aligned} c_0 &= 32n(4n+1)(2n+1)(4n+3)(n+1)[144n^4 - 168n^3m \\ &\quad + 40n^2m^2 + 6nm^3 + 1128n^3 - 876n^2m + 97nm^2 + 10m^3 \\ &\quad + 3188n^2 - 1457nm + 51m^2 + 3868n - 775m + 1704], \\ c_3 &= (2n+3)(n+3)(n+1)(2n-2m+7)(n-m+3)[144n^4 \\ &\quad - 168n^3m + 40n^2m^2 + 6nm^3 + 552n^3 - 372n^2m + 17nm^2 \\ &\quad + 4m^3 + 668n^2 - 209nm - 6m^2 + 300n - 26m + 40]. \end{aligned}$$

Similarly we can find explicit expression of c_1 and c_2 (see Exercise 5.12). Now summing up Eq. (5.86) for $-\infty < k < \infty$ we find that both $L(m,n)$ and $R(m,n)$ satisfy the same recurrence relation

$$c_3A(n+3) = c_2A(n+2) - c_1A(n+1) + c_0A(n).$$

Furthermore it is easy to check by using Eq. (5.83) that for all $m \geq 1$ ($\delta_{m,1}$ is the Kronecker symbol)

$$L(m,1) = R(m,1) = -\delta_{m,1}, \quad L(m,2) = R(m,2) = -5\delta_{m,1} - (-1)^m.$$

When $n = 3$ direct computation using Eq. (5.85) shows that

$$L(m,3) = R(m,3) = \begin{cases} -16 & \text{if } m = 1, \\ 0 & \text{if } m = 2, \\ 40/3 & \text{if } m = 3, \\ (-1)^m(m - 52/3) & \text{if } m \geq 4. \end{cases}$$

This implies that Eq. (5.83) is true.

The proof of Eq. (5.84) can be carried out by exactly the same method as above so we leave the details to the interested readers. This completes the proof of the proposition and therefore Thm. 5.7.1. □

5.8 Period polynomial relations of double zeta values

In Remark 3.4.7 concerning Conjecture 3.4.6 about the dimensions of MZV space we mentioned that MZVs are expected to be intimately related to the modular forms. In this section we will explore this important relation in some details.

We will always assume k to be an even positive integer throughout this section, except when we point out otherwise.

5.8.1 *Double zeta space*

Recall that the double zeta values satisfy the following stuffle and shuffle relations (see Eq. (5.16)):

$$\zeta(j)\zeta(w-j) = \zeta(j,w-j) + \zeta(w-j,j) + \zeta(w), \tag{5.87}$$

$$\zeta(j)\zeta(w-j) = \sum_{r=2}^{w-1} \left[\binom{r-1}{j-1} + \binom{r-1}{w-j-1} \right] \zeta(r,w-r). \tag{5.88}$$

for $2 \leq j \leq w/2$. Equating the right-hand side of these two relations we get

$$\zeta(j,w-j) + \zeta(w-j,j) + \zeta(w) = \sum_{r=2}^{w-1} \left[\binom{r-1}{j-1} + \binom{r-1}{w-j-1} \right] \zeta(r,w-r). \tag{5.89}$$

We define the *formal double zeta space* $\mathcal{D}_w$ to be the $\mathbb{Q}$-vector space generated by formal symbols $Z_{r,s}$, $P_{r,s}$ and Z_w subject to the above two sets of relations, with $\zeta(r,s)$, $\zeta(r)\zeta(s)$ and $\zeta(w)$ replaced by $Z_{r,s}$, $P_{r,s}$ and Z_w, respectively, by formally allowing r and s to be 1. Clearly we can also write

$$\mathcal{D}_w = \frac{\mathbb{Q}\langle Z_{r,w-r}, Z_w : 1 \leq r < w \rangle}{\mathbb{Q}\langle \text{relations (5.89)} \rangle},$$

with w generators and $w/2$ relations, which shows that

$$\dim \mathcal{D}_w \geq \frac{w}{2}. \tag{5.90}$$

We will show in fact the equality holds in Thm. 5.8.11.

Theorem 5.8.1. *Let w be a positive even integer. Then*

$$\sum_{1<r<w,\ r\ odd} Z_{r,w-r} = \frac{1}{4}Z_w, \qquad \sum_{1<r<w,\ r\ even} Z_{r,w-r} = \frac{3}{4}Z_w. \tag{5.91}$$

Remark 5.8.2. This is the formal version of Thm. 5.3.3.

Proof. Consider the generating functions

$$\mathfrak{Z}_w(X,Y) = \sum_{r+s=w} Z_{r,s}X^{r-1}Y^{s-1}.$$

By (5.89) we see that

$$\mathfrak{Z}_w(X,Y)+\mathfrak{Z}_w(Y,X)+Z_w\frac{X^{w-1}-Y^{w-1}}{X-Y} = \mathfrak{Z}_w(X+Y,Y)+\mathfrak{Z}_w(X+Y,X).$$

Set $(X,Y)=(1,0)$ and then $(X,Y)=(1,-1)$ we get, respectively,

$$\sum_{r=2}^{w-1} Z_{r,w-r} = Z_w, \tag{5.92}$$

$$2\sum_{r=2}^{w-1}(-1)^{r-1}Z_{r,w-r} + Z_w = 0. \tag{5.93}$$

By adding (resp. subtracting) twice of (5.92) to (resp. from) (5.93) we obtain (5.91). □

5.8.2 *Period polynomials*

Let w be an even positive integer. V_w be the $\mathbb{Q}$-vector space of homogeneous polynomials of degree $w-2$ inside $\mathbb{Q}[X,Y]$.

To study period polynomials, we need to use the following special elements in $\Gamma := \mathrm{PGL}_2(\mathbb{Z})$ (notation only valid in Sec. 5.8):

$$U := \begin{pmatrix}1 & -1\\ 1 & 0\end{pmatrix}, \quad S := \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \quad T := US = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix},$$
$$\varepsilon := \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad \delta := \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}, \quad T' := U^2S = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}.$$

Let $I = \pm\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$. The above matrices satisfy the following relations:

$$\begin{aligned} &S^2 = U^3 = I, \quad \varepsilon T\varepsilon = T', \quad \delta T\delta = T^{-1},\\ &S = \varepsilon\delta, \qquad\quad \varepsilon U\varepsilon = U^2, \quad STS = T'^{-1}. \end{aligned} \tag{5.94}$$

For any $F(X,Y)\in V_w$ and $\gamma = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)\in\Gamma$ we define

$$F|_\gamma(X,Y) = F(\mathtt{a}X+bY, cX+dY).$$

Denote by $V_w^{\pm}$ the $(\pm)$-eigenspaces of ε, V_w^{ev} and V_w^{od} the $(\pm)$-eigenspaces of δ, respectively.

Proposition 5.8.3. *We have a surjective map:*

$$\pi^{\mathrm{od}} := \frac{1}{2}(1-\delta) : V_w^+|_{T-I} \longrightarrow V_w^{\mathrm{od}}.$$

Proof. We prove that the following diagram is commutative:

$$\begin{array}{ccccc} & & V_w^{T\delta} & \xrightarrow{T-I} & V_w^{\mathrm{od}} \longrightarrow 0 \\ & & \Big\downarrow{\scriptstyle 1+\varepsilon-\varepsilon U} & & \Big\uparrow{\scriptstyle \pi^{\mathrm{od}}} \\ & & V_w^+ & \xrightarrow{T-I} & V_w^+|_{T-I} \end{array}$$

and the top row is exact, namely, the top horizontal map is surjective.

(1). The two maps from $V_w^{T\delta}$ are well-defined. Let $g \in V_w^{T\delta}$.

(i) Horizontal: Since $g|_{T\delta} = g$ and $T\delta T = \delta$ we have $g|_\delta = g|_{T\delta T} = g|_T$ and

$$(g|_{T-I})|_\delta = g|_{T\delta-\delta} = g|_{I-T\delta T} = g|_{I-T} = -g|_{T-I}. \tag{5.95}$$

Thus $g|_{T-I} \in V_w^{\mathrm{od}}$.

(ii) Vertical: By Eq. (5.94) we have $\varepsilon U\varepsilon = U^2 = TSU$. Thus

$$(g|_{1+\varepsilon-\varepsilon U})|_\varepsilon = g|_{1+\varepsilon-TSU} = g|_{1+\varepsilon-\delta SU} = g|_{1+\varepsilon-\varepsilon U}.$$

Thus $g|_{1+\varepsilon-\varepsilon U} \in V_w^+$.

(2). The top horizontal map $T-I$ is surjective. Indeed, if $f \in V_w^{\mathrm{od}}$ then there exists $g \in V_w$ such that $f = g|_{T-I}$ (see Exercise 5.17). But then

$$(g|_{T\delta-I})|_{I-T} = g|_{T\delta-I-T\delta T+T} = g|_{T\delta-I-\delta+T} = f|_{\delta+I} = 0.$$

Hence $g|_{T\delta-I}(X,Y) = g(Y-X,Y) - g(X,Y)$ must be a constant. Taking $(X,Y) = (1,2)$ we see that $g|_{T\delta-I} = 0$ which implies $g \in V_w^{T\delta}$.

(3). The square commutes. Direct computation shows that

$$(\varepsilon-\varepsilon U)(T-I)(I-\delta) = (I-T\delta)(\varepsilon-\varepsilon U)T(I-\delta).$$

Therefore,

$$(g|_{1+\varepsilon-\varepsilon U} - g)|_{T-I}|_{I-\delta} = g|_{(I-T\delta)(\varepsilon-\varepsilon U)T(I-\delta)} = 0.$$

But by Eq. (5.95) $(g|_{T-I})|_{I-\delta} = 2g|_{T-I}$. Hence

$$\pi^{\mathrm{od}}\big((g|_{1+\varepsilon-\varepsilon U})|_{T-I}\big) = g|_{T-I}.$$

Combining (2) and (3) we can finish the proof of the proposition easily. □

We now define the subspace $W_w \subset V_w$ by

$$W_w = \{P(X,Y) \in V_w : P|_{I+S} = P|_{I+U+U^2} = 0\}.$$

This is equivalent to saying that W_w consists of all the polynomials $P(X,Y)$ in V_w satisfying

$$P(X,Y) + P(-Y,X) = P(X,Y) + P(X-Y,X) + P(Y,Y-X) = 0.$$

Since $\varepsilon S \varepsilon = S$ and $\varepsilon U \varepsilon = U^2$ we see that W_w is invariant under ε, implying that $W^{\pm} = W \cap V_w^{\pm}$ are the $(\pm)$-eigenspaces of ε in W, $W^{\mathrm{ev}} = W \cap V_w^{\mathrm{ev}}$ and $W_w^{\mathrm{od}} = W \cap V_w^{\mathrm{od}}$ the $(\pm)$-eigenspaces of δ, respectively. As every polynomial in W is a sum of its symmetric part $P^+(X,Y) := (P(X,Y) + P(Y,X))/2$ and anti-symmetric part $P^-(X,Y) := (P(X,Y) - P(Y,X))/2$, we see that $W_w = W_w^+ \oplus W_w^-$ where

$$W_w^{\pm} = \{P \in W_w : P|_{\varepsilon} = \pm P\}.$$

Noticing that $P|_{\delta} = P|_{S\varepsilon} = -P|_{\varepsilon}$ for all $P \in W_w$ we can also write $W_w^+ = W_w^{\mathrm{od}}$ and $W_w^- = W_w^{\mathrm{ev}}$ where

$$W_w^{\mathrm{ev}} = \{P \in W_w : P|_{\delta} = P\} \quad \text{and} \quad W_w^{\mathrm{od}} = \{P \in W_w : P|_{\delta} = -P\}.$$

Consequently, we have

Lemma 5.8.4. *If* $p(X,Y) = \sum \binom{w-2}{r-1} p_{r,s} X^{r-1} Y^{s-1} \in W_w^-$ *then* $p_{r,s} = 0$ *for all* r, s *even.*

Proposition 5.8.5. *Suppose* $p(X,Y) \in W_w^+$. *Then* $F(x) := p(x,1)$ *enjoys the following properties.*

(i) $F(x)$ *is a palindrome polynomial of degree* $\leq w-3$.
(ii) $F(x) + x^{w-2} F(1-1/x) + F(1-x) = 0$.
(iii) $F(0) = F(1) = F'(1) = 0$ *so* $F(x+1)$ *has neither constant nor linear terms.*

Proof. (i) This follows from the symmetry of $p(X,Y)$. Since the degree of F is bounded by $w-2$ and is odd, it must be bounded by $w-3$.

(ii) This follows from $p|_{I+U+U^2} = p(X,Y) + p(X-Y,X) + p(-Y,X-Y) = p(X,Y) + p(X-Y,X) - p(X-Y,Y) = 0$.

(iii) Clearly $F(0) = 0$ since $F(x)$ has only odd degree terms. Setting $x = 1$ in (ii) we get $F(1) = 0$. By (i) we have $F(x) = x^d F(1/x)$ where $d = \deg F$. Differentiating both sides yields $F'(x) = dx^{d-1} F(1/x) - x^{d-2} F'(1/x)$. Taking $x = 1$ we see that $F'(1) = 0$. □

Proposition 5.8.6. *Suppose $p(X,Y) \in W_w^-$. Then $F(x) := p(x,1)$ enjoys the following properties. Let $p_1(X,Y)$ and $p_2(X,Y)$ be the partial derivatives with respect to X and Y, respectively.*

(i) $F(x) = -x^{w-2}F(1/x)$.
(ii) $F(x) + x^{w-2}F(1-1/x) - F(x-1) = 0$.
(iii) $F(1) = F'(0) = 0$.
(iv) $p_1(1,1) - p_1(1,0) = 0$.
(v) $p_{1,1}(1,1) - p_{1,1}(1,0) = 0$

Proof. (i) We have $p(X,Y) = -p(Y,X)$ since $p(X,Y) \in W_w^-$. Since p has homogeneous degree $w-2$ (i) follows immediately.

(ii) This follows from $p|_{I+U+U^2} = p(X,Y) + p(X-Y,X) + p(-Y,X-Y) = p(X,Y) + p(X-Y,X) - p(X-Y,Y) = 0$.

(iii) Clearly $F'(0) = 0$ since $F'(x)$ has only odd degree terms. Setting $x = 1$ in (ii) we get $F(1) = 0$.

(iv) Differentiating $p(Y,X) = -p(X,Y)$ and $p|_{I+U+U^2} = 0$ with respect to X we get $p_2(Y,X) = -p_1(X,Y)$ and $p_1(X,Y) + p_1(X-Y,X) + p_2(X-Y,X) - p_1(X-Y,Y) = 0$. Thus

$$p_1(X,Y) + p_1(X-Y,X) - p_1(X,X-Y) - p_1(X-Y,Y) = 0. \quad (5.96)$$

We can prove (iv) by setting $X = Y = 1$.

(v) Differentiating Eq. (5.96) with respect to X we get

$$p_{1,1}(X,Y) + p_{1,1}(X-Y,X) + p_{1,2}(X-Y,X) - p_{1,1}(X,X-Y) - p_{1,2}(X,X-Y) - p_{1,1}(X-Y,Y) = 0.$$

Taking $X = Y = 1$ we get

$$p_{1,1}(1,1) + p_{1,2}(0,1) - p_{1,1}(1,0) - p_{1,2}(1,0) = 0.$$

Notice that $p_{1,2}(X,Y)$ all the powers of X and Y are odd. So $p_{1,2}(0,1) = p_{1,2}(1,0) = 0$ which yields (v). □

5.8.3 *Modular forms and cusp forms on* $\mathbf{SL_2(\mathbb{Z})}$

Recall that a modular form on $\mathrm{PSL}_2(\mathbb{Z})$ of weight w is a holomorphic function on the upper half plane $\mathbb{H}$ satisfying the following two conditions:

(i) $f\Big(\dfrac{az+b}{cz+d}\Big) = (cz+d)^w f(z)$ for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{Z})$;

(ii) The Fourier expansion of f at $i\infty$ has the form $\displaystyle\sum_{n=0}^{i\infty} a_n e^{2\pi i n z}$.

A modular form f is a *cusp form* if it further satisfies $a_0 = 0$. We denote by S_w (resp. M_w) the $\mathbb{Q}$-vector space generated by all the cusp (resp. modular) forms of weight w. For example, for every even integer $w \geq 4$, we have the Eisenstein series defined by

$$G_w(z) := \frac{1}{2} \sum_{(m,n)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m+nz)^w}.$$

This can be shown to be a modular form of weight w (see [359]). One can also show that for each w all the modular forms of weight w form a finite dimension vector space over $\mathbb{Q}$ and

$$\dim S_w = \dim M_w - 1, \quad \dim M_w = \begin{cases} \left[\frac{w}{12}\right] + 1, & \text{if } 12 \nmid w-2; \\ \left[\frac{w}{12}\right], & \text{if } 12 \mid w-2. \end{cases}$$

Definition 5.8.7. The period polynomial associated with a cusp form $f \in S_w$ is defined by

$$P_f(X,Y) = \int_0^{i\infty} (X - Y\tau)^{w-2} f(\tau)\,\mathrm{d}\tau. \tag{5.97}$$

By integration by parts one can show that $P_f(X,Y) \in V_w$ (see [359, Chap. VI]). The following result and Thm. 5.8.18 to be proved later give the reason why W_w is called the *space of period polynomials*.

Theorem 5.8.8. *For any cusp form $f \in S_w$ its associated period polynomial $P_f(X,Y) \in W_w$.*

Proof. For any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{Z})$, we have

$$P_f|_\gamma(X,Y) = \int_{\gamma^{-1}(0)}^{\gamma^{-1}(i\infty)} (X - Y\tau)^{w-2} f(\tau)\,\mathrm{d}\tau. \tag{5.98}$$

(See Exercise 5.15). Since $S^{-1} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$ and $U^{-1} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right)$, we get

$$P_f|_{I+S} = \left(\int_0^{i\infty} + \int_{i\infty}^0\right) (X - Y\tau)^{w-2} f(\tau)\,\mathrm{d}\tau = 0$$

and

$$P_f|_{I+U+U^2} = \left(\int_0^{i\infty} + \int_1^0 + \int_{i\infty}^1\right) (X - Y\tau)^{w-2} f(\tau)\,\mathrm{d}\tau = 0.$$

Thus $P_f(X,Y) \in W_w$ by the definition. □

In order to study double zeta values using period polynomial we still need two auxiliary results.

Proposition 5.8.9. *Let w be any positive integer, **either even or odd**. Let $a_{r,s}$ and λ be rational numbers. Then the following two statements are equivalent:*

(i) *The relation*

$$\sum_{r+s=w} a_{r,s} Z_{r,s} = \lambda Z_w \tag{5.99}$$

holds in $\mathcal{D}_w$.

(ii) *The generating function*

$$A(X,Y) = \sum_{r+s=w} \binom{w-2}{r-1} a_{r,s} X^{r-1} Y^{s-1} \in V_w \tag{5.100}$$

can be written as $H|_{T'-I}(X,Y) = H(X,X+Y) - H(X,Y)$ for some $H \in V_w^+$, and

$$\lambda = \frac{w-1}{2} \int_0^1 H(t,1-t)\mathrm{d}t. \tag{5.101}$$

Proof. Since V_w^+ is spanned by symmetric polynomials of the form $X^{i-1}Y^{j-1} + X^{j-1}Y^{i-1}$ $(i = w-j)$ we may assume just $H(X,Y) = X^{i-1}Y^{j-1} + X^{j-1}Y^{i-1}$. Then

$$\lambda = (w-1)\int_0^1 t^{i-1}(1-t)^{j-1}\mathrm{d}t = \frac{(i-1)!(j-1)!}{(w-2)!}$$

by beta integral. An easy binomial expansion shows that the $a_{r,s}$'s in (5.100) reduces Eq. (5.99) to Eq. (5.89). □

We now define a $\mathbb{Q}$-bilinear pairing $\langle -,- \rangle : V_w \times V_w \to \mathbb{Q}$ such that

$$\langle X^{r-1}Y^{s-1}, X^{m-1}Y^{n-1} \rangle := \frac{(-1)^r}{\binom{w-2}{s-1}} \delta_{(r,s),(n,m)}. \tag{5.102}$$

Watch out for the change of order in (n,m).

Proposition 5.8.10. *Let $a_{r,s}$ and λ be rational numbers with $a_{r,s} = a_{s,r}$. Then the following three statements are equivalent:*

(i) *The relation*

$$\sum_{r+s=w} a_{r,s} P_{r,s} = \lambda' Z_w$$

holds in $\mathcal{D}_w$ for some λ'.

(ii) *The generating function*

$$A(X,Y) = \sum_{r+s=w} \binom{w-2}{r-1} a_{r,s} X^{r-1} Y^{s-1} \in V_w$$

can be written as $H|_{I-S}(X,Y) = H(X,Y) - H(-Y,X)$ *for some* $H \in V_w^{U,+} := V_w^U \cap V_w^+$.

(iii) *The generating function* $A(X,Y) \in V_w^{\mathrm{od}}$ *and* $A \perp W_w^+$ *with respect to the scalar product Eq.* (5.102).

If these statements hold, then $a_{r,s} = 0$ *for all odd* r *and* s.

Proof. The equivalence of (i) and (ii) follows almost immediately from Prop. 5.8.9 except for the extra condition $a_{r,s} = a_{s,r}$. Then (i) is obvious. For (ii), we notice that if $H \in V_w^+$ then $H|_{(T'-I)(\varepsilon-I)} = H|_{T-T\varepsilon} = H|_{(I-U)T}$. Thus $H|_{T'-I}$ is symmetric $\Leftrightarrow H = H|_U \Leftrightarrow H|_{T'-I} = H|_{S-I}$ since $T' = U^2 S = U^{-1} S$.

We now prove the equivalence of (ii) and (iii). We have

$$\begin{aligned}
&v \text{ is orthogonal to } V_w^{U,+}|_{I-S} = V_w|_{(I+U+U^2)(I-S)} \\
&\Leftrightarrow v|_{(I-S)(I+\varepsilon)(I+U+U^2)} = 0 \\
&\Leftrightarrow v|_{(I-S)(I+\varepsilon)} \in \ker(I+U+U^2) \cap \ker(I+S) \cap \ker(I-\varepsilon) = W_w^+ \\
&\Leftrightarrow v^{\mathrm{od},+} \in W_w^+ \Leftrightarrow v \in W_w^+ + V_w^- + V_w^{\mathrm{ev}}
\end{aligned}$$

by writing $v = v^{\mathrm{ev},+} + v^{\mathrm{ev},-} + v^{\mathrm{od},+} + v^{\mathrm{od},-}$. Therefore,

$$A \in V_w^{U,+}|_{I-S} \Leftrightarrow A \perp W_w^+ + V_w^- + V_w^{\mathrm{ev}} \Leftrightarrow A \perp W_w^+ \text{ and } A \in V_w^{\mathrm{od}}.$$

The last sentence of the proposition follows from $A \in V_w^{\mathrm{od}}$ because this means all the even X-power terms vanish. □

5.8.4 *The even weight double zeta value relations*

Theorem 5.8.11. *Let* $w \geq 4$ *be an even integer. Then the set*

$$\left\{ Z_{r,w-r} : 0 < r < w, r \text{ is odd} \right\}$$

form a basis of $\mathcal{D}_w$.

Proof. We already know $\dim \mathcal{D}_w \geq w/2$ and Z_w is a linear combination of $Z_{\mathrm{od,od}}$'s by Thm. 5.8.1. So it suffices to prove that if r is even then $Z_{r,w-r}$ can be expressed as a $\mathbb{Q}$-linear combination of $Z_{\mathrm{od,od}}$'s. In other words, for the choice $a_{r,w-r} = 1$ and $a_{j,w-j} = 0$ for all other even j, there exist suitable choices of $a_{\mathrm{od,od}}$'s and λ such that Eq. (5.99) holds. But by Prop. 5.8.9 this

means we need to show that the odd polynomial $X^{r-1}Y^{w-r-1}$ is the odd part of $H|_{T'-I}$ for some $H \in V_w^+$ or, equivalently, $X^{w-r-1}Y^{r-1}$ is the odd part of $H|_{T-I}$. Indeed, in general $F = 2\pi^{\mathrm{od}}(H|_{T-I}) = H|_{(T-I)(I-\delta)}$ for $H \in V_w^+$ if and only if

$$F|\varepsilon = H|_{(T-I)(\varepsilon-\delta\varepsilon)} = H|_{(T\varepsilon-\varepsilon)(I-\delta)} = H|_{(T'-I)(I-\delta)} = 2\pi^{\mathrm{od}}(H|_{T'-I})$$

since $\delta\varepsilon = \varepsilon\delta$ in $\mathrm{PGL}_2(\mathbb{Z})$ and $H|_\varepsilon = H$. Finally, the existence of such an $H \in V_w^+$ is guaranteed by Prop. 5.8.3.

We have finished the proof of the theorem. □

Remark 5.8.12. We caution the reader that the theorem is true only in the context of formal double zeta space. There will be non-trivial $\mathbb{Q}$-linear relations among double zeta values $\zeta(\mathrm{od},\mathrm{od})$'s. In fact, this will be the content of Cor. 5.8.19 below.

Let $\mathcal{P}_w^{\mathrm{ev}}$ be the space spanned by $P_{\mathrm{ev,ev}}$ over $\mathbb{Q}$. We define a map $\phi : W_w^- \to \mathcal{P}_w^{\mathrm{ev}}$ as follows. For each $p(X,Y) \in W_w^-$ we write

$$\begin{aligned} p(X,Y) &= \sum_{r+s=w} \binom{w-2}{r-1} p_{r,s} X^{r-1}Y^{s-1}, \\ p|_T(X,Y) = p(X+Y,Y) &= \sum_{r+s=w} \binom{w-2}{r-1} q_{r,s} X^{r-1}Y^{s-1}. \end{aligned}$$

Then we define

$$\phi\big(p(X,Y)\big) = \sum_{r+s=w, r,s \text{ even}} q_{r,s} Z_{r,s} \in \mathcal{P}_w^{\mathrm{ev}}. \tag{5.103}$$

Lemma 5.8.13. *For all r,s, we have $p_{r,s} = q_{r,s} - q_{s,r}$. Moreover, $q_{r,s} = q_{s,r}$ for all r,s even so the map ϕ is well-defined.*

Proof. Let $q = p|_T$. Since $T + \varepsilon T\varepsilon = I + (I+U+U^2)S - (I+S)S$ and $p|_\varepsilon = -p$ we have

$$q|_{I-\varepsilon} = p|_{T-T\varepsilon} = p|_{T+\varepsilon T\varepsilon} = p.$$

Thus $p_{r,s} = q_{r,s} - q_{s,r}$. Then $p_{r,s} = 0$ for all r,s even by Lemma 5.8.4. Now by the second equation of Eq. (5.91) we see that

$$Z_w = \frac{2}{w+1} \sum_{1<r<w,\ r \text{ even}} P_{r,w-r} \in \mathcal{P}_w^{\mathrm{ev}}.$$

So for all r,s we have $Z_{r,s} + Z_{s,r} = P_{r,s} - Z_w \in \mathcal{P}_w^{\mathrm{ev}}$ and therefore the map ϕ is well-define. □

Theorem 5.8.14. *The map $\phi : W_w^- \to \mathcal{P}_w^{\mathrm{ev}}$ is an isomorphism. Moreover,*

$$\phi\big(p(X,Y)\big) = \sum_{\substack{r+s=w \\ r,s \ even}} q_{r,s} Z_{r,s} \equiv 3 \sum_{\substack{r+s=w \\ r,s \ odd}} q_{r,s} Z_{r,s} \pmod{Z_w}. \tag{5.104}$$

Conversely, an element $\sum_{r+s=w,r,s \ odd} c_{r,s} Z_{r,s} \in \mathcal{D}_w$ belongs to $\mathcal{P}_w^{\mathrm{ev}}$ only if $c_{r,s} = q_{r,s}$ arising in this way.

Proof. The injectivity of ϕ follows from Eq. (5.104) since the coefficients on the right-hand side of Eq. (5.104) is anti-symmetric. The surjectivity, which is equivalence to the last sentence in the statement of the theorem, follows from Cor. 5.8.22 since injectivity and $\dim W_w^- = \dim \mathcal{P}_w^{\mathrm{ev}}$ implies bijectivity.

We now prove Eq. (5.104). We decompose q as $q = q^{\mathrm{ev},+} + q^{\mathrm{ev},-} + q^{\mathrm{od},+} + q^{\mathrm{od},-}$. Then $q^{\mathrm{ev},-} = p/2$ and $q^{\mathrm{od},-} = 0$ by Lemma 5.8.13. Set $[a,b,c] = \mathsf{a} q^{\mathrm{ev},+} + b q^{\mathrm{ev},-} + c q^{\mathrm{od},+}$. Then we have

$$\begin{aligned} [0,2,0]|_{T'} &= p|_{T'} = p|_{\varepsilon T \varepsilon} = -p|_{T\varepsilon} = -q|_{\varepsilon} = [-1,1,-1], \\ [1,-1,-1]|_{T'} &= q|_{ST'} = p|_{TST'} = p|_S = -p = [0,-2,0]. \end{aligned}$$

So

$$[2,0,-2]|_{T'-I} = [-1,-3,-1] - [2,0,-2] = [-3,-3,1].$$

Hence setting $H = 2(q^{\mathrm{ev},+} - q^{\mathrm{od},+}) \in V_w^+$, we get

$$q^{\mathrm{od}} - 3q^{\mathrm{ev}} = H|_{T'-I}.$$

So Eq. (5.104) follows from Prop. 5.8.9. This concludes the proof of the theorem. □

The proofs above provides us with an explicit way to write any $\zeta(\mathrm{ev},\mathrm{ev})$ or $\zeta(\mathrm{ev})\zeta(\mathrm{ev})$ as a $\mathbb{Q}$-linear combination of $\zeta(\mathrm{od},\mathrm{od})$'s. In particular, every asymmetric polynomial in W_w^- leads to some nontrivial relations among double zeta values.

Example 5.8.15. The polynomials $X^2Y^2(X^2 - Y^2)^3$ in W_{12}^- and $X^2Y^2(X^2 - Y^2)^3(2X^4 - X^2Y^2 + 2Y^4)$ in W_{16}^- yield

$$\begin{aligned} &28P_{8,4} + \frac{95}{3}P_{6,6} - \frac{167}{3}Z_{12} = 28Z_{9,3} + 150Z_{7,5} + 168Z_{5,7}, \\ &66P_{4,12} + 185P_{6,10} + \frac{364}{3}P_{8,8} - \frac{1081}{3}Z_{16} \\ &\qquad = 66Z_{13,3} + 375Z_{11,5} + 686Z_{9,7} + 675Z_{7,9} + 396Z_{5,11}, \end{aligned}$$

see Exercise 5.20. These lead to the identities:

$$\tfrac{5197}{691}\zeta(12) = 28\zeta(9,3)+150\zeta(7,5)+168\zeta(5,7),$$
$$\tfrac{78967}{3617}\zeta(16) = 66\zeta(13,3)+375\zeta(11,5)+686\zeta(9,7)+675\zeta(7,9)+396\zeta(5,11).$$

It is by no means some accidents that 691 and 3617 are exactly the numerators of B_{12} and B_{16}, respectively.

In order to prove the surjectivity of ϕ in Thm. 5.8.14 we need to study not only the space W_w but also its 1-dimensional extension $\widehat{W}_w$, called the *space of extended period polynomials*, which is defined by

$$\widehat{W}_w = \ker(I+S, \widehat{V}_w) \cap \ker(I+U+U^2, \widehat{V}_w)$$

where we have extended the maps $S, U,$ *etc.* to

$$S, U, \textit{etc.} : \widehat{V}_w := \bigoplus_{\substack{r+s=w \\ r,s\geq 0}} \mathbb{C}X^{r-1}Y^{s-1} \longrightarrow \mathbb{C}(X,Y).$$

We define $\widehat{W}_w^{\pm}$ to be the $(\pm)$-eigenspace of ε.

Lemma 5.8.16. *Let $f \in M_w$ with $f(i\infty) = a_0 \neq 0$. Let $\tau_0 \in \mathbb{H}$. Then*

$$\widehat{P}_f(X,Y) = \int_{\tau_0}^{i\infty} (X-Y\tau)^{w-2}(f(\tau)-a_0)\mathrm{d}\tau$$
$$+\int_0^{\tau_0} (X-Y\tau)^{w-2}\Big(f(\tau)-\frac{a_0}{\tau^w}\Big)\mathrm{d}\tau + \frac{a_0}{w-1}\left(\frac{1}{Y}-\frac{\tau^{1-w}}{X}\right)(X-Y\tau)^{w-1}$$

is independent of τ_0 and lies in $\widehat{W}_w$.

Proof. It is straight-forward to check that $\widehat{P}_f(X,Y)$ is independent of τ_0 by showing the derivative with respect to τ_0 vanishes. The first integral converges because the fast decay of e^{-y} as $y \to \infty$. The second integral also converges because $f(-1/it) = (it)^w f(it)$ as $f \in M_w$. This finishes the proof of the lemma. □

One may wonder how the expression in the lemma is found in the first place. For this, we need to use the L-function $L(s,f) = \sum_{n\geq 1} a_n/n^s$ associated to f. For $\mathrm{Re}(s) \gg 0$ it is easy to see that

$$L^*(s,f) := \int_0^{\infty} (f(it)-a_0)t^{s-1}\mathrm{d}t = (2\pi)\Gamma(s)L(s,f)$$

which can be rewritten as (for any $t_0 > 0$)

$$L^*(s,f) := \int_{t_0}^{\infty} \Big(f(it) - a_0\Big)t^{s-1}\mathrm{d}t + \int_0^{t_0} \Big(f(it) - \frac{a_0}{(it)^w}\Big)t^{s-1}\mathrm{d}t - a_0\Big(\frac{t_0^s}{s} + \frac{(-1)^{w/2}t_0^{w-s}}{w-s}\Big). \quad (5.105)$$

From Eq. (5.105) we see immediately that $L^*(s,f)$ can be analytically continued as a meromorphic function on $\mathbb{C}$ satisfying the functional equation $L^*(s,f) = (-1)^{w/2}L^*(w-s,f)$. It has simple poles at $s=0$ with residue $-a_0$ and at $s=w$ with residue $(-1)^{w/2}a_0$.

Now if $f \in S_w$ we have

$$\begin{aligned} P_f(X,Y) &= \int_0^{i\infty} (X - Y\tau)^{w-2} f(\tau)\mathrm{d}\tau \\ &= \int_0^{\infty} \sum_{n=0}^{w-2} i^{1-n}\binom{w-2}{n} X^{w-n-2}Y^n f(it)t^n \mathrm{d}t \\ &= \sum_{n=0}^{w-2} i^{1-n}\binom{w-2}{n} X^{w-n-2}Y^n L^*(n+1,f) \end{aligned}$$

For $f \in M_w$ by adding two terms in the above sum when $n \to (-1)^+$ and $n \to (w-1)^-$ we would have added

$$\lim_{n\to c} \frac{\Gamma(w-1)}{\Gamma(n+1)\Gamma(w-n-1)}(-i)^n X^{w-n-2}Y^n L^*(n+1,f) = \begin{cases} -a_0/(w-1)\cdot X^{w-1}/Y, \text{ if } c = (-1)^+; \\ -a_0/(w-1)\cdot Y^{w-1}/X, \text{ if } c = (w-1)^-. \end{cases}$$

Thus the the expression

$$\frac{a_0}{w-1}\Big(\frac{X^{w-1}}{Y} + \frac{X^{w-1}}{Y}\Big) + \sum_{n=0}^{w-2} i^{1-n}\binom{w-2}{n} X^{w-n-2}Y^n L^*(n+1,f) \quad (5.106)$$

is well-defined in $\widehat{V}_w$. Now we can easily obtain the expression in the lemma by applying Eq. (5.105) to Eq. (5.106).

Lemma 5.8.17. *The element*

$$\widehat{\mathcal{E}}_w = \sum_{r+s=w, r,s\geq 0} \binom{w}{r} B_r B_s X^{r-1}Y^{s-1} \in \widehat{W}_w^+.$$

Proof. The function is clearly symmetric. We leave the interested reader to check it is annihilated by both $I+S$ and $I+U+U^2$. □

The Eichler-Shimura-Manin theory completed by Kohnen-Zagier (see [340, 359] or [578]) provides the following fundamental result whose proof is far beyond the scope of this book.

Theorem 5.8.18. *Let w be an even positive integer. Let $P_f^{\pm}$ be the projection of P_f into $W_w^{\pm}$. Then we have*

(i) *The map $f \mapsto P_f^+$ gives an isomorphism $S_w \cong W_w^+$.*
(ii) *The map $f \mapsto P_f^-$ gives an isomorphism of S_w onto a codimension 1 subspace of W_w^-.*
(iii) *The pairing (5.102) induces a pairing $\widehat{W}_w^+ \times W_w^- \to \mathbb{C}, (P_f^+, P_g^-) \mapsto \langle P_f^+|_{T-T^{-1}}, P_g^- \rangle$ such that the a codimension 1 subspace of W_w^- in (ii) is precisely the space orthogonal to $\widehat{\mathcal{E}}_w$.*

Corollary 5.8.19. *The values $\zeta(\mathrm{od}, \mathrm{od})$ of weight w satisfy at least $\dim S_w$ linearly independent relations.*

Proof. By Thm. 5.8.14 the values $\zeta(\mathrm{od}, \mathrm{od})$ of weight w and $\zeta(w)$ satisfy at least $\dim W_w^-$ linearly independent relations. Noticing that $\zeta(\mathrm{ev})\zeta(w - \mathrm{ev})$ is always a rational multiple of $\zeta(w)$, after eliminating all such terms we arrive at the conclusion of the corollary in view of Thm. 5.8.18 (ii). □

Corollary 5.8.20. *We have $1 + \dim W_w^+ = \dim W_w^- = \dim M_w$.*

Proof. It is well-known that $M_w = S_w \oplus \mathbb{Q} \cdot G_w$ where G_w is the Eisenstein series of weight w, see [158, Chap. 4]. □

Theorem 5.8.21. *If $\{Z_{r,s}, P_{r,s}, Z_w\}$ satisfy Eqs. (5.87)-(5.89), then*

$$\sum_{\substack{r+s=w \\ r,s \text{ even}}} P_{r,s} X^{r-1} Y^{s-1} - \frac{Z_w}{2} \left(\frac{X^{w-1}}{Y} + \frac{Y^{w-1}}{X} \right) \in \widehat{W}_w^+.$$

Moreover, every element of $\widehat{W}_w^+$ arises in this way.

Proof. Prop. 5.8.10 (i) is equivalent to the statement that $\sum a_{r,s} p_{r,s} = 0$ for any realization $\omega : \mathcal{D}_w \to \mathbb{C}$ with $\omega(P_{r,s}) = p_{r,s}$ and $\omega(Z_w) = 0$, while (iii) means that all $a_{\mathrm{od},\mathrm{od}} = 0$ and $\sum a_{r,s} p_{r,s} = 0$ for all $\sum p_{r,s} X^{r-1} Y^{s-1} \in W_w^+$. The equivalence of (i) and (iii) implies that every collection of numbers $\{p_{r,s} = p_{s,r} : r + s = w, r, s \text{ odd}\}$ can be extended to of realization $\omega : \mathcal{D}_w \to \mathbb{C}$ with $\omega(Z_w) = 0$, while every collection of numbers $\{p_{r,s} = p_{s,r} : r + s = w, r, s \text{ even}\}$ can be extended to of realization $\omega : \mathcal{D}_w \to \mathbb{C}$ with

$\omega(Z_w) = 0$ if and only $\sum \omega(P_{r,s}) X^{r-1}Y^{s-1} \in W_w^+$. Thus the case when $\omega(Z_w) = 0$ of the theorem is proved.

In general, we need to show that given any extended symmetric function

$$\widehat{P} = \sum_{r+s=w; r,s \geq 0 \text{ and even}} p_{r,s} X^{r-1}Y^{s-1} \in \widehat{W}_w^+$$

there is a realization $\omega : \mathcal{D}_w \to \mathbb{C}$ with

$$\omega(P_{r,s}) = \begin{cases} p_{s,r}, & \text{if } r,s \text{ even;} \\ 0, & \text{if } r,s \text{ odd,} \end{cases} \qquad \omega(Z_w) = -2p_{w,0}.$$

To see this, it suffices to check it for any ONE extended element $\widehat{P} \in \widehat{W}_w^+ \setminus W_w^+$ because the codimension of W_w^+ in $\widehat{W}_w^+$ is 1. But by taking $\widehat{\mathcal{E}}_w$ we see that $p_{r,s} = \binom{w}{r} B_r B_s = 4w!\beta_r\beta_s$, where $\beta_r = \zeta(r)/(2\pi i)^r = -B_r/(2 \cdot r!)$ for all even $r \geq 0$. This is exactly $4w!/(2\pi i)^w$ times the original double zeta realization (denoted by ω_0) of $\mathcal{D}_w$ since

$$\frac{4w!}{(2\pi i)^w}\omega_0(P_{r,s}) = \frac{4w!\zeta(r)\zeta(s)}{(2\pi i)^w} = 4w!\beta_r\beta_s,$$
$$\frac{4w!}{(2\pi i)^w}\omega_0(Z_w) = \frac{4w!\zeta(w)}{(2\pi i)^w} = 4w!\beta_w = -2(4w!\beta_w\beta_0).$$

This completes the proof of the theorem. □

Corollary 5.8.22. *We have* $\dim \mathcal{P}_w^{\text{ev}} = \dim \widehat{W}_w^+ = \dim W_w^-$.

Proof. This follows from Thm. 5.8.21 and Cor. 5.8.20. □

5.8.5 *The odd weight double zeta value relations*

First of all, we want to remind the reader that by the Parity Principle of Sec. 3.5, all double zeta values of odd weight can be expressed as a $\mathbb{Q}$-linear combination of products of Riemann zeta values. The following theorem provides the explicit relations.

Theorem 5.8.23. *Let $m \geq 2$, $n \geq 1$ and the weight $w = m + n = 2k + 1$ be an odd number. Then*

$$\zeta(m,n) = (-1)^n \sum_{j=0}^{k-1} \left[\binom{w-2j-1}{m-1} + \binom{w-2j-1}{n-1} - \delta_{m,2j} + (-1)^n \delta_{j,0} \right] \times \\ \times \zeta(2j)\zeta(w-2j),$$

where $\zeta(0) = -1/2$.

Proof. Recall that in the formal double zeta space we have

$$P_{j,w-j} = Z_{j,w-j} + Z_{w-j,j} + Z_w, \tag{5.107}$$

$$P_{j,w-j} = \sum_{r=1}^{w-1} \left[\binom{r-1}{j-1} + \binom{r-1}{w-j-1} \right] Z_{r,w-r}, \tag{5.108}$$

for all $j = 1, \dots, w-1$. Define the generating function

$$\mathfrak{P}(X,Y) = \sum_{j=1}^{w-1} P_{j,w-j} X^{w-j-1} Y^{j-1}, \quad \mathfrak{D}(X,Y) = \sum_{j=1}^{w-1} Z_{j,w-j} X^{w-j-1} Y^{j-1},$$

and $\mathfrak{Z}(X,Y) = Z_w (X^{w-1} - Y^{w-1})/(X-Y)$. Then we have

$$\mathfrak{P}(X,Y) = \mathfrak{D}(X,Y) + \mathfrak{D}(Y,X) + \mathfrak{Z}(X,Y) = \mathfrak{D}(X,X+Y) + \mathfrak{D}(Y,X+Y).$$

Repeatedly using this we get

$$\mathfrak{D}(X,Y) = \mathfrak{D}(-X,-Y) + \mathfrak{Q}(X,Y) + \mathfrak{Q}(X-Y,-Y) + \mathfrak{Q}(X-Y,X),$$

where $\mathfrak{Q}(X,Y) = \mathfrak{P}(X,Y) - \mathfrak{P}(X,-Y) - \mathfrak{Z}(X,Y)$. Since the weight w is odd we get $\mathfrak{D}(-X,-Y) = -\mathfrak{D}(X,Y)$. Easy computation shows that

$$\mathfrak{Q}(X,Y) = 2 \sum_{j=1}^{k} P_{2j,w-2j} X^{w-2j-1} Y^{2j-1} - Z_w \sum_{m=1}^{w-1} X^{w-m-1} Y^{m-1},$$

$$\begin{aligned}
\mathfrak{Q}(X-Y,-Y) &= 2 \sum_{j=1}^{k} P_{2j,w-2j} (X-Y)^{w-2j-1} (-Y)^{2j-1} \\
&\quad - Z_w \sum_{j=1}^{w-1} (X-Y)^{w-j-1} (-Y)^{j-1} \\
&= 2 \sum_{j=1}^{k} P_{2j,w-2j} \sum_{n=1}^{w-2j} (-1)^n \binom{w-2j-1}{n-1} X^{n-1} Y^{w-n-1} \\
&\quad - Z_w \sum_{j=1}^{w-1} \sum_{n=1}^{w-j} (-1)^n \binom{w-j-1}{n-1} X^{n-1} Y^{w-n-1} \\
&= 2 \sum_{n=1}^{w-1} \sum_{j=1}^{k} P_{2j,w-2j} (-1)^n \binom{w-2j-1}{n-1} X^{n-1} Y^{m-1} \\
&\quad - Z_w \sum_{n=1}^{w-1} (-1)^n \sum_{j=1}^{w-1} \binom{w-j-1}{n-1} X^{n-1} Y^{m-1}
\end{aligned}$$

$$\mathfrak{Q}(X-Y,X) = 2 \sum_{j=1}^{k} P_{2j,w-2j} (X-Y)^{w-2j-1} X^{2j-1}$$

$$
\begin{aligned}
&- Z_w \sum_{j=1}^{w-1} (X-Y)^{w-j-1} X^{j-1} \\
&= 2 \sum_{j=1}^{k} P_{2j,w-2j} \sum_{n=2j}^{w-1} (-1)^n \binom{w-2j-1}{n-2j} X^{n-1} Y^{w-n-1} \\
&\quad - Z_w \sum_{j=1}^{w-1} \sum_{n=j}^{w-1} (-1)^n \binom{w-j-1}{n-j} X^{n-1} Y^{w-n-1} \\
&= 2 \sum_{n=1}^{w-1} \sum_{j=1}^{k} P_{2j,w-2j} (-1)^n \binom{w-2j-1}{m-1} X^{n-1} Y^{m-1} \\
&\quad - Z_w \sum_{n=1}^{w-1} (-1)^n \sum_{j=1}^{w-1} \binom{w-j-1}{m-1} X^{n-1} Y^{m-1}.
\end{aligned}
$$

The theorem now follows from the identity

$$
\sum_{j=1}^{w-1} \binom{w-j-1}{n-1} = \sum_{j=n}^{w-2} \binom{j-1}{n-1} = \binom{w-1}{n} = \binom{w-1}{m-1}
$$

and a similar identity with $n \leftrightarrow m$. □

The next result is an analog of Thm. 5.8.14 in the odd weight case.

Theorem 5.8.24. *Let $w \geq 12$ be an even integer. For $p \in W_w^+$ we define $q_{r,s}$ $(r+s = w+1, r, s \geq 1)$ by*

$$
p(X+Y, Y) = \sum_{r+s=w+1} \binom{w-1}{r-1} q_{r,s} X^{r-1} Y^{s-2}.
$$

Then

$$
\sum_{\substack{r+s=w+1 \\ 4 \leq r \leq w-2 \ even}} (q_{r,s} - q_{s,r}) Z_{r,s} \equiv 0 \pmod{Z_{w+1}}. \tag{5.109}
$$

Proof. Let $g(X,Y) := p|_T = p(X+Y, Y)$. Since $\varepsilon S = S\varepsilon$ in $\mathrm{PGL}_2(Z)$ we have $\varepsilon T \varepsilon S = \varepsilon T S \varepsilon = \varepsilon U \varepsilon = U^2$. Thus

$$
\begin{aligned}
\varepsilon T \varepsilon + TST' + \varepsilon T \varepsilon ST' &= (I + U + U^2) T' \\
\varepsilon T \varepsilon ST' - T &= U^2 T' - T = US - T = 0.
\end{aligned}
$$

Hence

$$
\begin{aligned}
g|_{\varepsilon + ST' + \varepsilon ST'} &= p|_{T\varepsilon + TST' + T\varepsilon ST'} = p|_{\varepsilon T \varepsilon + TST' + \varepsilon T \varepsilon ST'} = 0, \\
g|_{\varepsilon ST' - I} &= p|_{T \varepsilon ST' - T} = p|_{\varepsilon T \varepsilon ST' - T} = 0.
\end{aligned} \tag{5.110}
$$

We now define

$$f(X,Y) := g\cdot Y - g|_\varepsilon \cdot X = \sum_{r+s=w+1} \binom{w-1}{r-1}(q_{r,s}-q_{s,r})X^{r-1}Y^{s-1}.$$

Then

$$\begin{aligned} f|_{ST'} - f &= (g\cdot Y - g|_\varepsilon\cdot X)|_{ST'} - (g\cdot Y - g|_\varepsilon\cdot X)\\ &= g|_{ST'}\cdot X - g|_{\varepsilon ST'}\cdot(-(X+Y)) - g\cdot Y + g|_\varepsilon\cdot X\\ &= g|_{\varepsilon+ST'+\varepsilon ST'}\cdot X + g|_{\varepsilon ST'-I}\cdot Y = 0 \end{aligned}$$

by Eq. (5.110). So we get $f = f|_{ST'}$. Moreover,

$$\begin{aligned} (f|_S)|_\varepsilon - f|_S &= (g\cdot Y - g|_\varepsilon\cdot X)|_{S\varepsilon} - (g\cdot Y - g|_\varepsilon\cdot X)|_S\\ &= g|_{\varepsilon S\varepsilon - S}\cdot X + g|_{S\varepsilon-\varepsilon S}\cdot Y = 0. \end{aligned}$$

Thus $f|_S$ is a symmetric homogeneous polynomial of degree $w-1$ which is odd. Therefore,

$$f|_{I-S} = \sum_{r+s=w+1,\, r \text{ even}, 4\le r\le w-2} 2\binom{w-1}{r-1}(q_{r,s}-q_{s,r})X^{r-1}Y^{s-1}.$$

Notice that g does not have terms with X-degree equal to 1 or $w-1$ by Prop. 5.8.5 so neither does f. Finally, since

$$f|_{I-S} = f|_{ST'-S} = (f|_S)|_{T'-I}$$

the theorem follows from Prop. 5.8.9 immediately. □

Example 5.8.25. The space W_{12}^+ is 1-dimensional, spanned by the odd period polynomial $p(x) = x(x^2-4)(4x^2-1)(x^2-1)^2$. We have $p(x+1) = x^2(x^2-1)(2x+3)(2x+1)(x+3)(x+2)^2$, so the $q_{r,s}$ of the theorem can be found using the table

r	10	9	8	7	6	5	4	3
s	3	4	5	6	7	8	9	10
$330q_{r,s}$	24	72	119	115	15	-161	-288	-216

The relations in the theorem, divided by 10, become

$$24Z_{10,3} + 28Z_{8,5} - 10Z_{6,7} - 36Z_{4,9} \equiv 0 \pmod{Z_{13}}.$$

In the double zeta value realization, this can be written as

$$24\zeta(10,3) + 28\zeta(8,5) - 10\zeta(6,7) - 36\zeta(4,9) = -3\zeta(13).$$

Theorem 5.8.26. *Let $w \geq 12$ be an even integer. To each $p \in W_w^-$ we associate the coefficients $q_{r,s}$ $(r+s=w-1)$ which are defined by*

$$\frac{\partial}{\partial X}p(X+Y,Y) = \sum_{r+s=w-1} \binom{w-3}{r-1} q_{r,s} X^{r-1} Y^{s-1}.$$

Then

$$\sum_{\substack{r+s=w-1 \\ 4\leq r\leq w-4 \ even}} (q_{r,s} - q_{s,r}) Z_{r,s} \equiv 0 \pmod{Z_{w-1}}. \tag{5.111}$$

Proof. Let $g = \frac{\partial}{\partial X}p|_T = \frac{\partial}{\partial X}p(X+Y,Y)$. We now define

$$f(X,Y) := g|_{I-\varepsilon} = \sum_{r+s=w-1} \binom{w-3}{r-1} (q_{r,s} - q_{s,r}) X^{r-1} Y^{s-1}.$$

Easy computation shows that $T(I-\varepsilon) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & -1 \end{smallmatrix}\right)$. Thus

$$f|_{ST'-I} = g|_{(I-\varepsilon)(ST'-I)} = \frac{\partial}{\partial X}p|_{T(I-\varepsilon)(ST'-I)} = \frac{\partial}{\partial X}p(0,-3X) = 0$$

by Prop. 5.8.6 (iii). So we get $f = f|_{ST'}$. Moreover, $(I-\varepsilon)S(\varepsilon - I) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ so we get

$$(f|_S)|_\varepsilon - f|_S = g|_{(I-\varepsilon)S(\varepsilon-I)} = \frac{\partial}{\partial X}p(0,0) = 0$$

by Prop. 5.8.6 (iii). Thus $f|_S$ is a symmetric homogeneous polynomial of degree $w-3$. Furthermore, easy computation shows that the coefficient of $X^{w-3}Y$ in f is $p_1(1,1) - p_1(1,0) = 0$ by Prop. 5.8.6 (iv). Similarly, the coefficient of XY^{w-3} in f is $p_{1,1}(1,1) - p_{1,1}(1,0) = 0$ by Prop. 5.8.6 (v). Hence

$$f|_{I-S} = \sum_{r+s=w-1, r \text{ even}, 4\leq r\leq w-4} 2\binom{w-3}{r-1}(q_{r,s} - q_{s,r}) X^{r-1} Y^{s-1}.$$

Finally, since

$$f|_{I-S} = f|_{ST'-S} = (f|_S)|_{T'-I}$$

the theorem follows from Prop. 5.8.9 immediately. □

Example 5.8.27. The space W_{12}^- is 2-dimensional, spanned by the even period polynomial $p(x) = x^{10} - 1$ and $x^2(x^2-1)^3$. For the latter, we have $p'(x+1) = 8x^7 + 56x^6 + 150x^5 + 190x^4 + 112x^3 + 24x^2$, so the $q_{r,s}$ of the theorem are given (after multiplication by 63) by the table

r	8	7	6	5	4	3
s	3	4	5	6	7	8
$63q_{r,s}$	14	42	75	95	84	42

The relations in the theorem, divided by -1, become

$$28Z_{8,3} + 20Z_{6,5} - 42Z_{4,7} \equiv 0 \pmod{Z_{11}}.$$

In the double zeta value realization, this can be written as

$$28\zeta(8,3) + 20\zeta(6,5) - 42\zeta(4,7) = -3\zeta(11). \tag{5.112}$$

5.9 Conjectural multiple zeta value relations

Hoffman posted on the internet the following conjecture and mentioned that J. Vermaseren had used the MZV data mine presented in [59] to check it through weight 22.

Conjecture 5.9.1. *For every nonnegative integer n*

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

During mid 1990s Borwein *et al.* started the development of EZ-Face (an abbreviation for Euler Zetas interFace, see Sec. 14.1), an on-line calculator for alternating MZVs called Euler sums. They detected many relations and the following is one such family (see [69]). First define

$$Z(m_0,\ldots,m_{2n}) := \zeta(\{2\}^{m_0},3,\{2\}^{m_1},1,\{2\}^{m_2},\ldots,3,\{2\}^{m_{2n-1}},1,\{2\}^{m_{2n}}), \tag{5.113}$$

with $\{2\}^{m_j}$ inserted after the j-th element of the string $\{3,1\}^n$. For example, $Z(1,3,2) = \zeta(2,3,2,2,2,1,2,2)$.

Conjecture 5.9.2. *For any sequence $S = (m_0,\ldots,m_{2n})$ of $2n+1$ nonnegative integers, we have*

$$\sum_{j=0}^{2n} Z(\mathcal{C}^j S) = \frac{\pi^{4n+2M}}{(4n+2M+1)!}, \tag{5.114}$$

where $M := \sum_{i=0}^{2n} m_i$ and $\mathcal{C}$ is the cyclic permutation operator, that is,

$$\mathcal{C}^j(m_0,\ldots,m_{2n}) = (m_{2n-j+1},\ldots,m_{2n},m_0,\ldots,m_{2n-j}).$$

Note that by Eq. (5.75) the right-hand side of Eq. (5.114) is equal to $\zeta\left(\{2\}^{2n+M}\right)$. Borwein *et al.* proved Eq. (5.114) in [69] for the cases $M = 0$ (which is Eq. (5.75) when $m = 0$) and $M = 1$. For $n = 0$, Eq. (5.114) is exactly Eq. (5.75) when $n = 0$. If all m_i's are equal, Eq. (5.114) reduces to [68, Conjecture (18)].

There are many conjectured identities between MZVs with components consisting of only 1, 2 or 3's. Here is a typical one given near the end of [69]. One can verify it with the aid of Petitot's table in [458] up to weight 16 and by EZ-face numerically for larger weight.

Conjecture 5.9.3. *For any non-negative integers a_1, a_2, a_3, b_1, b_2, we have*

$$\begin{aligned} &Z(a_1, b_1, a_2, b_2, a_3) + Z(a_2, b_1, a_3, b_2, a_1) + Z(a_3, b_1, a_1, b_2, a_2) \\ &\quad = Z(a_1, b_2, a_2, b_1, a_3) + Z(a_2, b_2, a_3, b_1, a_1) + Z(a_3, b_2, a_1, b_1, a_2). \end{aligned}$$

The following conjecture, which is supported by many examples in depth 2 and 3, provides a general form of the weighted sum relations.

Conjecture 5.9.4. *Let $F(x_1, \dots, x_m) \in \mathbb{Q}[x_1, \dots, x_m]$ be a symmetric polynomial of total degree r. Suppose $d = \deg_{x_1} F(x_1, \dots, x_m)$. Then for every positive integer $n \geq m$ we have*

$$\sum_{\substack{k_1+\cdots+k_m=n \\ k_1,\dots,k_m\geq 1}} F(k_1, \dots, k_m)\zeta(2k_1)\dots\zeta(2k_m) = \sum_{k=0}^{T} e_{F,k}(n)\zeta(2k)\zeta(2n-2k),$$

$$\sum_{\substack{k_1+\cdots+k_m=n \\ k_1,\dots,k_m\geq 1}} F(k_1, \dots, k_m)\zeta(2k_1, \dots, 2k_m) = \sum_{k=0}^{T} c_{F,k}(n)\zeta(2k)\zeta(2n-2k),$$

where $T = \max\{\lfloor (r+m-2)/2 \rfloor, \lfloor (m-1)/2 \rfloor\}$, $e_{F,k}(x), c_{F,k}(x) \in \mathbb{Q}[x]$ depend only on k and F, $\deg e_{F,k}(x) \leq r-1$ and $\deg c_{F,k}(x) \leq d$.

5.10 Historical notes

Hoffman proved Thm. 5.1.2 for $\mathbf{s}$ of some special forms such as in Example 5.1.1. The general case was first proved by the above approach discovered by Drinfeld, Kontsevich and Zagier (see also the Appendix of [368]). Ohno [431] generalized this to Thm. 5.4.9 by using iterated integrals. Our proof here is adopted from the paper [298] by Ihara *et al.* Another approach is presented in [446]. Using generating functions, Igarashi [294] extended Ohno's relation to the following parametric form:

Theorem 5.10.1. *Let $\alpha \in \mathbb{C}$ such that $\mathrm{Re}(\alpha) > 0$. Denote the Pochhammer symbol $(\alpha)_0 = 1$ and $(\alpha)_k = \prod_{j=0}^{k-1}(\alpha + j)$ for all $k \geq 1$. Define*

$$Z(\mathbf{s}; \alpha) = \sum_{k_1>\cdots>k_d>0} \frac{(\alpha)_{k_1}}{k_1!} \frac{k_d!}{(\alpha)_{k_d+1}} \frac{1}{(k_1+\alpha)^{s_1}\cdots(k_d+\alpha)^{s_d}}.$$

Then for any integer $k \geq 0$ admissible $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ we have

$$\sum_{\substack{\varepsilon_1+\cdots+\varepsilon_d=k \\ \varepsilon_j \in \mathbb{N}_0}} Z(s_1+\varepsilon_1, \ldots, s_d+\varepsilon_d; \alpha) = \sum_{\substack{\varepsilon'_1+\cdots+\varepsilon'_\ell=k \\ \varepsilon'_j \in \mathbb{N}_0}} Z(s'_1+\varepsilon'_1, \ldots, s'_\ell+\varepsilon'_\ell; \alpha). \tag{5.115}$$

Kawashima [331] showed another class of relations among MZVs containing Ohno's relations. He conjectured that his relations contain all cyclic sum relations given by Thm. 5.2.3 which was conjectured by Hoffman and proved by Ohno [288] whose proof using partial fractions was reproduced here.

The sum relations in Cor. 5.2.4 were conjectured by Moen [275] and by Schmidt and Markett [401] independently. The depth two case was know for quite a long time, see [430, III § 19 (1)]. Hoffman proved the general case under the assumption of the duality relation. The general case of Eq. (5.12) was proved independently by A. Granville [248] and D. Zagier using generating functions.

Eie *et al.* [182] improved the sum relations Eq. (5.12) in the following way.

$$\sum_{k_1+\cdots+k_d=w} \zeta(k_1+1, k_2, \ldots, k_d, \{1\}^e) = \sum_{l_1+\cdots+l_{e+1}=d+e} \zeta(l_1+(w-d)+1, l_2, \ldots, l_{e+1}). \tag{5.116}$$

This reduces to Eq. (5.12) if one sets $e = 0$. Tanaka [514] showed that this follows from the derivation relations Thm. 5.4.5. In [477] S. Saito *et al.* gave a detailed analysis of the cyclic sum relations by stratifying them using Hoffman's algebra setup and then proved Thm. 5.2.5.

During their study of multiple zeta star values (see Chap. 10) Ohno and Zudilin [437] proved the weighted sum relations Eq. (5.15) using the DBSFs. Our proof follows the idea in [298]. Equation Eq. (5.15) was generalized by Gue and Xie [254] as follows.

Theorem 5.10.2. *For positive integers $d \geq 2$ and $w \geq d+1$, we have*

$$\sum_{\substack{s_i \geq 1, s_1 \geq 2 \\ s_1+\cdots+s_d=w}} \Big[2^{s_1-1} + (2^{s_1-1}-1)\Big((\sum_{i=2}^{d-1} 2^{S_i-s_1-(i-1)}) + 2^{S_{d-1}-s_1-(d-2)}\Big)\Big] \times \zeta(s_1, \ldots, s_d) = w\zeta(w),$$

where $S_i = s_1 + \cdots + s_i$ for $i = 1, \ldots, d-1$.

Using some key ideas from [220] Machide in [393, 394, 397, 398] further studied more general weighted sum formulas in depth two to four in which he used some parameters whose specializations would give the proceeding formulas in these depths. See Exercises 5.13 and 5.14.

Many identities of MZVs (and related sums) have been derived directly by the generating function method. One may further consult the papers [11, 13, 220, 284, 567, 567, 610, 611, 616].

Thm. 5.3.3 was discovered by Gangl *et al.* [220] during their study of double zeta values and modular forms. In fact, it is contained implicitly already in Euler's work [198]. Our proof here is due to Nakamura [423] which is much shorter than the original proof by Gangl *et al.* However, using their approach Gangl *et al.* can further show that every double zeta value of even weight w is a $\mathbb{Q}$-linear combination of double zeta values of the form $\zeta(\text{odd}, \text{odd})$ while these values themselves satisfy at least $\dim S_w$ linearly independent relations, where S_w denotes the space of cusp forms of weight w on $\mathrm{PSL}_2(\mathbb{Z})$. Moreover, they conjectured that these relations are the only ones between these values. They also gave a more complicated identity.

Theorem 5.10.3. *For all odd integers $m, n \geq 2$ let $w = m + n$. Then we have*

$$2\sum_{i=0}^{n-1}\binom{-m}{i}B_i\zeta(n-i,m+i) = \sum_{j+k=w}(-1)^{k-1}\lambda_{m,n}(j,k)\zeta(j)\zeta(k).$$

where B_i is the ith Bernoulli number and

$$\lambda_{m,n}(j,k) = \sum_{i=0}^{n-1}\binom{m+i-1}{i}\binom{j-1}{n-i-1}B_i.$$

Thm. 5.3.4 is due to Nakamura [423]. The identity Eq. (5.26) between Bernoulli numbers have been obtained by a few diverse methods. Rademacher [466, p. 121] derived it as a consequence of an identity among Eisenstein series of different weights. Shimura essential did the same in his book [495, (11.10)]. Eie [179] proved it using the zeta function associated with some polynomials. His method can be generalized to obtain infinitely many families of Bernoulli number identities similar to Eq. (5.26) which lead to infinitely many different weighted sum formulas.

By studying higher derivatives of the function $f(t) = 1/(e^t - 1)$ Dilcher [159] generalized Eq. (5.21) in Thm. 5.3.13 involving arbitrary N factors in the product, which was also discovered in [498] for $N = 3$ and $N = 4$,

Sankarayanan [480] for $N = 5$ and by Zhang [584] for $N \leq 7$. This was proved by Eie [179] in special cases and Chen [130, 179] in general using completely different method. Our proof here follows that in [130].

With the aid of formulas Eqs. (5.40) and (5.41) Shen and Cai [492] generalized Thm. 5.3.3 to depth 3 and 4 cases. As a kind of corollary Shen and Cai [492] also found the following whose generalization to larger depth is unknown.

Theorem 5.10.4. *For any positive integer $n > 2$, we have*

$$\sum_{\substack{a+b+c=2n \\ a\geq 2, b, c\geq 1}} \left[(-1)^a + (-1)^b + (-1)^c\right]\zeta(a, b, c) = \frac{3}{2}\zeta(2n) - \zeta(2)\zeta(2n-2).$$

The case $r = 1$ of Example 5.3.7 gives [179, Prop. 1] (note that there is typo there where $(2n-2)!$ on the left numerator should be $(2n-4)!$). Taking $r = 2$ we recover [180, Prop. 4.2.2]. Cor. 5.3.11 is a result used by Shen and Cai [492, Lemma 5] which they derived by some complicated method using Bernoulli number. More formulas along this line and some involving alternating sums can be found in [545].

Thm. 5.4.5 was proved by K. Ihara and M. Kaneko [298] which generalized the case $n = 1$ proved by Hoffman [275]. In fact, they showed that the claim in Thm. 5.4.5 is equivalent to the validity of the DBSFs. Our proof only goes in one direction and therefore is much shorter. Kaneko [316] defined a generalization $\partial_n^{(c)}$ of the derivation ∂_n so that $\partial_n^{(0)} = \partial_n$ and he conjectured $\zeta(\partial_n^{(c)}(\mathbf{w})) = 0$ for all $c \in \mathbb{Q}$ and $n \geq 1$. This conjecture was proved later by Tanaka [513].

The special case of Thm. 5.5.2 with $z = 0$ (*i.e.*, height 1 case) can be found in [15,68,164]. The general case is proved by Ohno and Zagier in [436] and has been generalized to i-height case by Z. Li [376], where i-height of a composition is the number of components greater than i. So the classical height is just 1-height. As a related result of Cor. 5.5.4 Kajikawa [313] showed that the relation

$$G_0(w, d, h) = \sum \text{their duals}$$

can be derived from the finite and regularized DBSFs. But in general we don't know how to prove duality using only finite and regularized DBSFs.

Zagier's evaluation Thm. 5.6.4 was proved originally using some deep complex analysis results. Here, we choose to present Z.-H. Li's shorter hypergeometric series proof contained in [383]. Thm. 5.6.4 plays an important role in the work of F. Brown [104] on the mixed Tate motives over $\mathbb{Z}$. As

a corollary, Brown proved that every MZV can be expressed as a $\mathbb{Q}$-linear combination of MZVs whose arguments are either 2 or 3. Thm. 5.6.4 is also used in F. Brown and O. Schnetz's proof of the ZigZag Conjecture in [109].

Thm. 5.7.1 was a special case of Conjecture and was proved first by Bowman and Bradley [78, Cor. 5.1]. Two simplified proof can be found in [419] and [598]. The proof here is adopted from [598]. The case when $n = 0$, *i.e.*, Cor. 5.5.4, was known to Euler and re-proved by Hoffman [275]. The case $m = 0$ was conjectured by Zagier [577] and proved first by Broadhurst [69]. With the aid of Brown's theory of motivic MZVs [104, 105] Charlton [122] improved Eq. (5.75) to show that the sum of inserting all possible permutations of some fixed blocks of 2's into $\zeta(3, 1, \ldots, 3, 1)$ produces a rational multiple of a power of π.

One of the most fascinating aspects of MZVs is the intimate relations between the double zeta values, modular forms and the period polynomials. For the even weight case, we adopt the framework of [220] in Sec. 5.8. Schneps gives a different approach to these results in [485]. For the odd case, Thms. 5.8.24 and 5.8.26 are due to Ma (see [392]).

The input from the knot theory stimulated quite a lot of the study of MZVs at the initial stage. For example, the knot invariant of the 3-braid torus knot 8_{19} at 8 crossings (left of Fig. 5.1, according to Rolfsen's well-known table [475], is related to the (conjecturally) irreducible MZV of depth 2 occurring first at weight 8 (namely, $\zeta(5, 3)$), while knot 10_{124} (right of Fig. 5.1) involves the double zeta $\zeta(7, 3)$ (see [93]). Likewise the

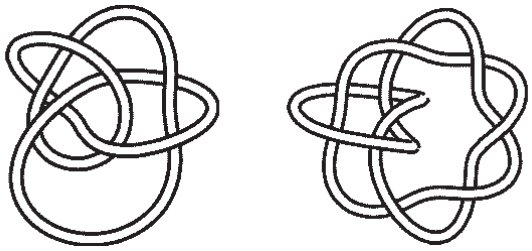


Figure 5.1: Knots 8_{19} and 10_{124} are related to irreducible depth 2 MZVs.

first (conjecturally) irreducible MZVs of depth 3 occurs at weight 11 (for e.g., $\zeta(3, 3, 5)$) [74], matching the appearance of the first positive 4-braid at 11 crossings, obtained from skeining link diagrams that encode momentum flow in 7-loop counterterms [92]. Later, they also found the first weight

12 MZV reduction that requires alternating Euler sums in [91]. A more systematic study of the relation between knots and MZVs can be found in [93].

In [367] Le and Murakami studied Kontsevich integrals of Homfly polynomials using representations of the chord diagram algebra. In the process they found many interesting relations of MZVs. For example,

$$\big(1-2^{2-2n}\big)\zeta(2n)=\sum_{k=2}^{2n-1}(-1)^k\zeta\big(k,\{1\}^{2n-k}\big)\quad\forall n\geq 2.$$

They also discovered the following

$$\sum_{\mathbf{s}:|_{\mathbf{s}}|_=2n,\ \mathrm{ht}(\mathbf{s})=k}(-1)^{\mathrm{dp}(\mathbf{s})}\zeta(\mathbf{s})=(-1)^n\zeta(\{2\}^n)\sum_{j=0}^{n-k}\binom{2n+1}{2j}(2-2^{2j})B_{2j}$$

by examining the Kontsevich integral of the unknot. This can also be proved by setting $y=-x$ in Thm. 5.5.2. More identities produced from knot theory can be found in [210].

There is a proof of a weak form of Conjecture 5.9.2 by Charlton [122] which partially generalized Eq. (5.75). The novelty of this proof is the use of the motivic MZV theory developed by Deligne, Goncharov and Brown. However, it does not provide the explicit multiple of the π-power specified in the conjecture.

Exercises

5.1. Find the dual $\mathbf{s}^*$ if $\mathbf{s}$ is given by
(i) $(4,2,1)$ (ii) $(3,\{2\}^n)$

5.2. Suppose $\mathbf{s}=(s_1,\ldots,s_d)$ and $\mathbf{w}=\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}\in\mathfrak{A}^0$. If $\mathbf{s}^*=(t_1,\ldots,t_\ell)$ then prove that $\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_\ell-1}\mathtt{b}=\tau(\mathbf{w})=\mathtt{ab}^{s_d-1}\cdots\mathtt{ab}^{s_1-1}$.

5.3. Show that Eq. (5.9) and Eq. (5.10) are equivalent.

5.4. Prove Thm. 5.3.1 by using integral expressions of MZVs and Chen's product formula Lemma 2.1.2(iv).

5.5. One may use the following idea to give an alternative elementary proof of Thm. 5.3.2. First, rewrite the double sum as

$$\sum_{k=2}^{w-1}\sum_{m,n=1}^{\infty}\frac{2^k}{n^{w-k}(n+m)^k}.$$

Then break the sum into two cases $m=n$ and $m\neq n$. Show that the first case yields $(w-2)\zeta(w)$ and the second yields $3\zeta(w)$ by summing over k first and then using telescoping series.

5.6. Show that the commutator of two derivations is a derivation.

5.7. For any $t \in \mathfrak{z}$ define a power series $f(u) \in \mathfrak{z}[\![u]\!]$ by

$$f(u) = \exp_\circ(tu) - 1 = \sum_{n\geq 1} \frac{t^{\circ n}}{n!} u^n.$$

(i). Show that

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{1}{1-f(u)}\right) = t * \left(\frac{1}{1-f(u)}\right)$$

by using the fact that $f'(u) = t \circ (1 + f(u))$.

(ii). Show that $F(u) = (1-f(u))^{-1}$ satisfies $F'(u) = z * F(u)$ and $F(0) = 1$. Consequently, one has

$$\exp_*(tu) = (2 - \exp_\circ(tu))^{-1}.$$

5.8. Let $t \in \mathfrak{z}$ and $\mathbf{w} \in \widehat{\mathfrak{A}}^1$.

(i). Show that for any $z \in \mathfrak{z}$

$$\delta_t(z\mathbf{w}) = (t \circ z)\mathbf{w} + z(t * \mathbf{w}).$$

(ii). Show that for any positive integer

$$\delta_t^n(z\mathbf{w}) = \sum_{\alpha+\beta=n,\alpha,\beta\geq 0} \binom{n}{\alpha} (t^{\circ\alpha} \circ z)(t^{*\beta} * \mathbf{w}).$$

(iii). Show that

$$\exp(\delta_t)(\mathbf{w}) = (\exp_*(t))^{-1}(\exp_*(t) * \mathbf{w}).$$

5.9. Show that $\mathfrak{d} : \mathfrak{A} \to \mathfrak{A}$ defined by $\mathfrak{d}(w) = y \sqcup\!\sqcup w - yw$ is a derivation on $\mathfrak{A}$ such that $1/m!\mathfrak{d}^m(\mathtt{a}) = \mathtt{a}\mathtt{b}^m$ and $1/m!\mathfrak{d}^m(\mathtt{b}) = \mathtt{b}^{m+1}$ for all $m \geq 1$. Further, show that $\mathfrak{d}$ induces a derivation on $\widehat{\mathfrak{A}}^0$ and then verify Eq. (5.43) by induction. In particular,

$$\exp(\mathfrak{d}u)(\mathtt{a}) = \mathtt{a}(1-\mathtt{b}u)^{-1},\ \exp(\mathfrak{d}u)(\mathtt{b}) = \mathtt{b}(1-\mathtt{b}u)^{-1},$$

and

$$\exp(\mathfrak{d}u)(\mathtt{a}+\mathtt{b}) = (\mathtt{a}+\mathtt{b})(1-\mathtt{b}u)^{-1}.$$

5.10. Show that $\delta_z : \widehat{\mathfrak{A}}^1 \to \widehat{\mathfrak{A}}^1$ defined by $\delta(w) = z * w - zw$ is a derivation on $\widehat{\mathfrak{A}}^1$ for every $z \in \mathfrak{z}$, and these derivations all commute. Furthermore, δ_z extends to a derivation on all of $\widehat{\mathfrak{A}}$, with

$$\delta_z(\mathtt{a}) = 0 \quad \text{and} \quad \delta_z(\mathtt{b}) = (\mathtt{a}+\mathtt{b})z.$$

5.11. Show that γ defined by Eq. (5.46) provides an algebra isomorphism

$$\gamma : X\mathbb{Q}[\![X]\!] \xrightarrow{\simeq} (\mathfrak{z}, \circ).$$

5.12. Using computer algebra systems Mathematica or Maple to find c_1 and c_2 in the proof of Prop. 5.7.2. One can download the free program of the WZ-method from the internet after searching "WZ-method".

5.13. Using appropriate specializations of the generating function for double and triple zeta values to show that for all $w \geq 1$

$$\sum_{a\geq 2, b\geq 1, a+b<w} \Big(3^{a-1}2^b - 2^{a+b-1} - 2^{a-1}\Big)\zeta(a,b,w-a-b)$$
$$= \frac{(w-4)(w+1)}{6}\zeta(w), \tag{5.117}$$

$$\sum_{a\geq 2, b\geq 1, a+b<w} \Big(2^{a+b-1} + 2^{a-1} - 2^b\Big)\zeta(a,b,w-a-b) = w\zeta(w), \tag{5.118}$$

$$\sum_{a\geq 2, b\geq 1, a+b<w} 2^b\Big(3^{a-1} - 1\Big)\zeta(a,b,w-a-b) = \frac{(w-1)(w+4)}{6}\zeta(w).$$

Notice that Eq. (5.118) is the depth $d = 3$ case of Thm. 5.10.2 by Guo and Xie.

5.14. Using the generating function for depth 4 MZVs to show that

$$\sum_{s_1+\cdots+s_4=w, s_1\geq 2, s_j, \geq 1} \big(3^{s_2}2^{s_1-1} - 3^{s_2} - 1\big)2^{s_1+s_3}\zeta(s_1,\ldots,s_4)$$
$$= \left(\frac{(w+7)(w+2)(w-3)}{12} + 2\right)\zeta(w).$$

5.15. Prove Eq. (5.98) using the identity $(\mathtt{a}X + bY) - (cX + dY)\tau = (\mathtt{a} - c\tau)(X - \gamma^{-1}(\tau)Y)$ and the modularity of f.

5.16. Prove that in Part (ii) of Prop. 5.8.10 we can take

$$\lambda' = \Big\langle H, \frac{X^{w-1} - Y^{w-1}}{X - Y}\Big\rangle.$$

5.17. Show that if $f(x)$ is an odd polynomial in $\mathbb{Q}[x]$ of degree $2d-1$ then there is $g(x) \in \mathbb{Q}[x]$ of degree $\leq 2d$ such that $f(x) = g(x+1) - g(x)$.

5.18. By direct computation show that

$$(\varepsilon - \varepsilon U)(T - I)(I - \delta) = (I - T\delta)(\varepsilon - \varepsilon U)T(I - \delta),$$
$$T + \varepsilon T\varepsilon = I + (I + U + U^2)S - (I + S)S.$$

5.19. Instead of using generating function of double zeta values, prove Thm. 5.8.1 by two suitable choices of H in Prop. 5.8.9.

5.20. Show that the Eq. (5.104) in Thm. 5.8.14 can be strengthened as

$$\sum_{\substack{r+s=w\\ r,s \text{ even}}} q_{r,s} Z_{r,s} + \sum_{r+s=w} (-1)^r q_{r,s} Z_w = 3 \sum_{\substack{r+s=w\\ r,s \text{ odd}}} q_{r,s} Z_{r,s}.$$

5.21. Complete the proof of Lemma 5.8.17 by using the method of generating functions and the fact that

$$\left(\frac{X^{w-1}}{Y} + \frac{X^{w-1}}{Y} \right) \Big|_{I+U+U^2}$$
$$= \frac{X^{w-1} - (X-Y)^{w-1}}{Y} + \frac{Y^{w-1} + (X-Y)^{w-1}}{X} + \frac{X^{w-1} - Y^{w-1}}{X-Y}.$$

.

Chapter 6

Symmetrized multiple zeta values

In the proceeding two chapters, we have considered many interesting properties of the MZVs. However, one unpleasant feature causes a lot of trouble, namely, admissible MZVs must start with an argument at least two. One effective way to overcome this difficulty is to define some *symmetrized* versions of MZVs, which will be carried in this chapter.

In Sec. 8.5.4 we will see a mysterious relation between MZVs and the so called finite version of MZVs. This relation will be established by using the symmetrized MZVs.

6.1 Definition

There are in fact two versions of the symmetrized MZVs.

Definition 6.1.1. For any $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$, we define

$$\zeta^{\mathcal{S}}_{ш}(\mathbf{s}) = \sum_{i=0}^{d} (-1)^{s_1+\cdots+s_i} \zeta_{ш}(s_i, \dots, s_1; T)\zeta_{ш}(s_{i+1}, \dots, s_d; T), \tag{6.1}$$

$$\zeta^{\mathcal{S}}_{*}(\mathbf{s}) = \sum_{i=0}^{d} (-1)^{s_1+\cdots+s_i} \zeta_{*}(s_i, \dots, s_1; T)\zeta_{*}(s_{i+1}, \dots, s_d; T). \tag{6.2}$$

These are called the *ш-symmetrized MZVs* and *$*$-symmetrized MZVs*, respectively. Both are called symmetrized MZVs (SMZVs) if we don't need to distinguish the two versions.

Remark 6.1.2. Although perhaps looking strange at the first glance, the $*$-symmetrized MZVs have in fact a natural series definition (see Prop. 6.1.4).

First, we show that neither $\zeta^{\mathcal{S}}_{ш}(\mathbf{s})$ nor $\zeta^{\mathcal{S}}_{*}(\mathbf{s})$ depends on the variable T, namely, they are actually real numbers.

Proposition 6.1.3. *For all* $s_1, \dots, s_d \in \mathbb{N}$ *both* $\zeta^{\mathcal{S}}_{ш}(s_1, \dots, s_d)$ *and* $\zeta^{\mathcal{S}}_{*}(s_1, \dots, s_d)$ *are finite real numbers. In particular,*

$$\zeta^{\mathcal{S}}_{ш}(\mathbf{s}) = \sum_{i=0}^{d} (-1)^{s_1+\cdots+s_i} \zeta_{ш}(s_i, \dots, s_1)\zeta_{ш}(s_{i+1}, \dots, s_d),$$

$$\zeta^{\mathcal{S}}_{*}(\mathbf{s}) = \sum_{i=0}^{d} (-1)^{s_1+\cdots+s_i} \zeta_{*}(s_i, \dots, s_1)\zeta_{*}(s_{i+1}, \dots, s_d).$$

Proof. We start with the shuffle version $\zeta^{\mathcal{S}}_{ш}(s_1, \dots, s_d)$. If all $s_i \geq 2$ then clearly it is finite. If all $s_i = 1$ then

$$\zeta^{\mathcal{S}}_{ш}(\mathbf{s}) = \sum_{i=0}^{d} (-1)^{d-i} \frac{T^i}{i!} \frac{T^{d-i}}{(d-i)!} = 0.$$

Otherwise, we may assume for some i and $l \geq 1$ we have $s_i \geq 2$, $s_{i+1} = \cdots = s_{i+l} = 1$, and $s_{i+l+1} \geq 2$. We only need to show that

$$\sum_{j=0}^{l} (-1)^j \zeta_{ш}(s_{i+j}, \dots, s_1)\zeta_{ш}(s_{i+j+1}, \dots, s_d)$$

is finite. Consider the words in $\mathfrak{A}^1_{ш}$ (see Definition 3.2.2 corresponding to the above values. We may assume $\mathbf{x}_0\mathbf{u} = \mathbf{x}_0^{s_i-1}\mathbf{x}_1 \cdots \mathbf{x}_0^{s_1-1}\mathbf{x}_1$ and $\mathbf{x}_0\mathbf{v} = \mathbf{x}_0^{s_{i+l+1}-1}\mathbf{x}_1 \cdots \mathbf{x}_0^{s_d-1}\mathbf{x}_1$. So we only need to show that

$$\sum_{j=0}^{l} (-1)^j (\mathbf{x}_1^j \mathbf{x}_0\mathbf{u}) ш (\mathbf{x}_1^{l-j}\mathbf{x}_0\mathbf{v}) \in \mathfrak{A}^0. \tag{6.3}$$

Indeed,

$$\begin{aligned}
&\sum_{j=0}^{l} (-1)^j (\mathbf{x}_1^j \mathbf{x}_0\mathbf{u}) ш (\mathbf{x}_1^{l-j}\mathbf{x}_0\mathbf{v}) \\
&= \mathbf{x}_0(\mathbf{u} ш \mathbf{x}_1^l\mathbf{x}_0\mathbf{v}) + \sum_{j=1}^{l} (-1)^j \mathbf{x}_1(\mathbf{x}_1^{j-1}\mathbf{x}_0\mathbf{u} ш \mathbf{x}_1^{l-j}\mathbf{x}_0\mathbf{v}) \\
&+ (-1)^l \mathbf{x}_0(\mathbf{x}_1^l\mathbf{x}_0\mathbf{u} ш \mathbf{v}) + \sum_{j=0}^{l-1} (-1)^j \mathbf{x}_1(\mathbf{x}_1^j\mathbf{x}_0\mathbf{u} ш \mathbf{x}_1^{l-j-1}\mathbf{x}_0\mathbf{v}) \\
&= \mathbf{x}_0(\mathbf{u} ш \mathbf{x}_1^l\mathbf{x}_0\mathbf{v}) + (-1)^l \mathbf{x}_0(\mathbf{x}_1^l\mathbf{x}_0\mathbf{u} ш \mathbf{v}) \in \mathfrak{A}^0,
\end{aligned}$$

where we have used the substitution $j \to j+1$ in the first sigma summation which is canceled by the second sigma summation.

For the stuffle version $\zeta^{\mathcal{S}}_{*}(s_1, \dots, s_d)$ we can use the same idea and show it's finite by noticing that the "stuffing" part of the stuffle product produces admissible words in $\mathfrak{A}^0$. We leave the details to the interested reader (see Exercise 6.2). □

Another useful way to define the $*$-symmetrized MZVs is to use some infinite series.

Proposition 6.1.4. *For any* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ *the limit*

$$\lim_{M\to\infty} \sum_{\substack{\frac{1}{k_1}>\cdots>\frac{1}{k_d} \\ 0\neq|k_j|\in\mathbb{Z}_{\leq M} \ \forall j=1,\dots,d}} \frac{1}{k_1^{s_1}\cdots k_d^{s_d}}$$

is equal to $(-1)^{|\mathbf{s}|}\zeta_*^{\mathcal{S}}(\mathbf{s})$.

Proof. We see that for all for $\frac{1}{k_1} > \cdots > \frac{1}{k_d}$ there is some i $(i = 0, \dots, d)$ such that $\frac{1}{k_1} > \cdots \frac{1}{k_i} > 0 > \frac{1}{k_{i+1}} > \cdots > \frac{1}{k_d}$. By the regularization procedure presented in Sec. 3.3.2 we see that

$$\begin{aligned}
&(-1)^{|\mathbf{s}|} \sum_{\substack{\frac{1}{k_1}>\cdots>\frac{1}{k_d} \\ 0\neq|k_j|\in\mathbb{Z}_{\leq M} \ \forall j=1,\dots,d}} \frac{1}{k_1^{s_1}\cdots k_d^{s_d}} \\
&= (-1)^{|\mathbf{s}|} \sum_{i=0}^{d} \sum_{\substack{\frac{1}{k_1}>\cdots>\frac{1}{k_i}>0>\frac{1}{k_{i+1}}>\cdots>\frac{1}{k_d} \\ 0\neq|k_j|\in\mathbb{Z}_{\leq M} \ \forall j=1,\dots,d}} \frac{1}{k_1^{s_1}\cdots k_d^{s_d}} \\
&= \sum_{i=0}^{d} (-1)^{s_1+\cdots+s_i} \zeta_*(s_i, \dots, s_1; T)\zeta_*(s_{i+1}, \dots, s_d; T) + O(T^{-1}\log^A(T)),
\end{aligned}$$

where $T = \log M + \gamma$ as $M \to \infty$, for some constant A. By Prop. 6.1.3 we see that the sum above is independent of T so that when $M \to \infty$ we can recover the definition of $\zeta_*^{\mathcal{S}}(\mathbf{s})$ given by Eq. (6.2). □

This proposition clearly implies the following corollary.

Corollary 6.1.5. *The* $*$*-symmetrized MZVs satisfy the stuffle relations.*

6.2 Identities involving symmetrized MZVs

In this subsection, we will study some relations among the SMZVs that are analogs of those for ordinary MZVs. First, we want to prove a kind of sum formula. Before proving the result we need a combinatorial lemma. For any $w, d, i \in \mathbb{N}$ with $1 \leq i \leq d \leq w-1$, we set

$$J_{w,d,i} = \{(s_1, \dots, s_d) \in \mathbb{N}^n \mid s_1 + \cdots + s_d = w, \ s_i \geq 2\}.$$

Lemma 6.2.1. *Let f be a multiple integer variable function satisfying stuffle relations, namely, $f(\mathbf{t}_1)f(\mathbf{t}_2) = f(\mathbf{t}_1 * \mathbf{t}_2)$ for any $\mathbf{t}_j \in \mathbb{N}^{d_j}$ $(j = 1, 2)$. Set*

$$S_{w,d,i} := \sum_{(s_1,\ldots,s_d)\in J_{w,d,i}} f(s_1, \ldots, s_d).$$

If $f(s) = 0$ for all $s \in \mathbb{N}$ then for any $w, d, i \in \mathbb{N}$ with $2 \le i+1 \le d \le w-1$, we have

$$(d-i)S_{w,d,i} + iS_{w,d,i+1} + (w-d)S_{w,d-1,i} = 0.$$

Proof. By stuffle relation

$$\begin{aligned} 0 &= f(l)f(s_1, \ldots, s_{d-1}) \\ &= \sum_{j=1}^{d} f(s_1, \ldots, s_{j-1}, l, s_j, \ldots, s_{d-1}) + \sum_{j=1}^{d-1} f(s_1, \ldots, s_{j-1}, s_j + l, s_{j+1}, \ldots, s_{d-1}). \end{aligned}$$

Summing this over all $(s_1, \ldots, s_{d-1}, l) \in J_{w,d,i}$ gives the desired recurrence relations. Indeed, the map

$$(s_1, \ldots, s_{d-1}, l) \mapsto (s_1, \ldots, s_{j-1}, l, s_j, \ldots, s_{d-1})$$

defined on $J_{w,d,i}$ is a bijection to $J_{w,d,i+1}$ for $1 \le j \le i$ and is surjection onto $J_{w,d,i}$ for $i < j \le d$. Moreover, under the map

$$\begin{aligned} J_{w,d,i} \times \{1, \ldots, d-1\} &\longrightarrow J_{w,d-1,i} \\ \big((s_1, \ldots, s_{d-1}, l), j\big) &\longmapsto (s_1, \ldots, s_{j-1}, s_j + l, s_{j+1}, \ldots, s_{d-1}), \end{aligned}$$

the pre-images of each $(s'_1, \ldots, s'_{d-1}) \in J_{w,d-1,i}$ form a set of cardinality

$$\sum_{\substack{1\le j\le d-1 \\ j\ne i}} (s'_j - 1) + (s'_i - 2) = w - d.$$

This completes the proof of the lemma. □

The next theorem is often called a sum formula.

Theorem 6.2.2. *For any $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ and $w, i \in \mathbb{N}$ with $1 \le i \le d \le w-1$, we have, modulo $\zeta(2)$,*

$$\sum_{|\mathbf{s}|=w,\ s_i\ge 2} \zeta^{\mathcal{S}}_*(\mathbf{s}) \equiv (-1)^{i-1}\left(\binom{w-1}{i-1} + (-1)^d\binom{w-1}{d-i}\right)\zeta(w). \qquad (6.4)$$

Proof. We will prove the theorem by backward induction on the depth d. When $d = w-1$ the right-hand side of Eq. (6.4) is 0 when w is even and is

$$(-1)^{i-1}\left(\binom{w-1}{i-1}+\binom{w-1}{i}\right)\zeta(w) = (-1)^{i-1}\binom{w}{i}\zeta(w)$$

when w is odd. Now consider the left-hand side of Eq. (6.4). Since $\zeta_*^{\mathcal{S}}$ satisfies the stuffle relation by Cor. 6.1.5 we may apply Lemma 6.2.1 with $f = \zeta_*^{\mathcal{S}}$. By Cor. 5.4.7 we see that

$$\mathrm{reg}_{\sqcup\!\sqcup}(\mathtt{x}_1^m\mathtt{x}_0\mathtt{x}_1^j) = (-1)^m\mathtt{x}_0(\mathtt{x}_1^m \sqcup\!\sqcup \mathtt{x}_1^j) = (-1)^m\binom{j+m}{m}\mathtt{x}_0\mathtt{x}_1^{m+j}.$$

Therefore

$$\begin{aligned}\zeta_{\sqcup\!\sqcup}(\{1\}^{i-1},2,\{1\}^{w-i-1}) &= (-1)^{i-1}\binom{w-1}{i-1}\zeta(2,\{1\}^{w-2})\\ &= (-1)^{i-1}\binom{w-1}{i-1}\zeta(w)\end{aligned}$$

by duality relation (5.3). Similarly

$$\zeta_{\sqcup\!\sqcup}(\{1\}^{w-i-1},2,\{1\}^{i-1}) = (-1)^{w-i-1}\binom{w-1}{i}\zeta(w).$$

Thus by the definition and Thm. 6.3.2

$$\begin{aligned}S_{w,w-1,i} &= \zeta_*^{\mathcal{S}}(\{1\}^{i-1},2,\{1\}^{w-i-1} \equiv \zeta_{\sqcup\!\sqcup}^{\mathcal{S}}(\{1\}^{i-1},2,\{1\}^{w-i-1})\\ &\equiv \zeta_{\sqcup\!\sqcup}(\{1\}^{i-1},2,\{1\}^{w-i-1}) + (-1)^w\zeta_{\sqcup\!\sqcup}(\{1\}^{w-i-1},2,\{1\}^{i-1})\\ &\equiv (-1)^{i-1}\binom{w}{i}\zeta(w) \qquad (\mathrm{mod}\ \zeta(2)).\end{aligned}$$

So the left-hand side of Eq. (6.4) agrees with the right-hand side.

Now we assume that Eq. (6.4) holds for some $d \le w-1$. By Lemma 6.2.1

$$\begin{aligned}&(d-w)S_{w,d-1,i} = (d-i)S_{w,d,i} + iS_{w,d,i+1}\\ &= (d-i)(-1)^{i-1}\left(\binom{w-1}{i-1}+(-1)^d\binom{w-1}{d-i}\right)\zeta(w)\\ &\quad + i(-1)^i\left(\binom{w-1}{i}+(-1)^d\binom{w-1}{d-i-1}\right)\zeta(w)\\ &= (-1)^{i-1}\left((d-i)\binom{w-1}{i-1}+(w-d+i)(-1)^d\binom{w-1}{d-i-1}\right)\zeta(w)\\ &\quad + (-1)^i\left((w-i)\binom{w-1}{i-1}+i(-1)^d\binom{w-1}{d-i-1}\right)\zeta(w)\\ &= (d-w)(-1)^{i-1}\left(\binom{w-1}{i-1}+(-1)^{d-1}\binom{w-1}{d-i-1}\right)\zeta(w).\end{aligned}$$

Thus Eq. (6.4) holds for $d-1$. This finishes the proof of the lemma. □

Lemma 6.2.3. *For positive integers k and n with $k \geq 2$, we have the duality in height* 1*:*

$$\zeta_{ш}(\{1\}^{d-1}, k) = (-1)^{d-1}\zeta^{\star}(k, \{1\}^{d-1}).$$

Proof. Set $(\ldots, i, \{1\}^{-1}, j, \ldots) = (\ldots, i+j-1, \ldots)$. By Cor. 5.4.7, we have

$$\begin{aligned}
&\zeta_{ш}(\{1\}^{d-1}, k) = (-1)^{d+k}\zeta_{ш}(\mathtt{x}_0(\mathtt{x}_1^{d-1} ш \mathtt{x}_0^{k-2}\mathtt{x}_1)) \quad (\text{word depth}=k-1)\\
&= (-1)^{d+k} \sum_{\substack{e_1+\cdots+e_{k-1}=d-1\\ e_i\geq 0\,(i=1,2,\ldots,k-1)}} (e_1+1)\zeta(\mathtt{x}_0\mathtt{x}_1^{e_{k-1}}\cdots\mathtt{x}_0\mathtt{x}_1^{e_1+1})\\
&= (-1)^{d+k} \sum_{\substack{e_1+\cdots+e_{k-1}=d-1\\ e_i\geq 0\,(i=1,2,\ldots,k-1)}} (e_1+1)\zeta(\mathtt{x}_0^{e_1+1}\mathtt{x}_1\cdots\mathtt{x}_0^{e_{k-1}}\mathtt{x}_1) \quad (\text{by Eq. (5.3)})\\
&= (-1)^{d-1} \sum_{\substack{e_1+\cdots+e_{k-1}=d-1\\ e_i\geq 0\,(i=1,2,\ldots,k-1)}} (e_1+1)\zeta(2+e_1, 1+e_2, \ldots, 1+e_{k-1})\\
&= (-1)^{d-1} \sum_{t=1}^{d} \sum_{\substack{e_1+\cdots+e_{k-1}=d-t\\ e_i\geq 0\,(i=1,2,\ldots,k-1)}} \zeta(t+1+e_1, 1+e_2, \ldots, 1+e_{k-1})\\
&= (-1)^{d-1} \sum_{t=1}^{d} \sum_{\substack{e'_1+\cdots+e'_t=d-t\\ e'_j\geq 0\,(j=1,2,\ldots,t)}} \zeta(k+e'_1, 1+e'_2, \ldots, 1+e'_t).
\end{aligned}$$

by Thm. 5.4.9 of Ohno's Relations. It is not hard to see that the last sum can be expressed as (see Eq. (1.10) and Exercise 3.15)

$$(-1)^{d-1}\sum \zeta(k \circ \underbrace{1 \circ \cdots \circ 1}_{(d-1)\text{-times}}) = (-1)^{d-1}\zeta^{\star}(k, \{1\}^{d-1}),$$

where $\circ$ is either "+" or ",". The lemma is now proved. □

For SMZVs we have the following height 1 duality.

Theorem 6.2.4. *For any $k, n \in \mathbb{N}$, we have*

$$\zeta^{\mathcal{S}}_{ш}(k, \{1\}^{d-1}) - \zeta^{\mathcal{S}}_{ш}(d, \{1\}^{k-1})$$

is equal to $\zeta(2)$ times a polynomial of Riemann zeta values with rational coefficients.

Proof. The claim is trivial if $k = 1$ or $d = 1$ so we may assume $k, d \geq 2$. By Lemma 6.2.3, we have

$$\begin{aligned}&\zeta^{\mathcal{S}}_{ш}(k, \{1\}^{d-1}) - \zeta^{\mathcal{S}}_{ш}(d, \{1\}^{k-1})\\ &= \zeta(k, \{1\}^{d-1}) + (-1)^k \zeta^\star(k, \{1\}^{d-1}) - \zeta(d, \{1\}^{k-1}) - (-1)^d \zeta^\star(d, \{1\}^{k-1}).\end{aligned}$$

Taking $z = 0$ in Thm. 5.5.2, we get

$$\begin{aligned}&1 - \sum_{k,d\geq 1} \zeta(k+1, \{1\}^{d-1}) x^k y^d\\ &\qquad = \exp\biggl(\sum_{d\geq 2} \zeta(d) \frac{x^d + y^d - (x+y)^d}{d}\biggr) = \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}. \quad (6.5)\end{aligned}$$

Hence by Eq. (1.5) we have

$$\begin{aligned}&\sum_{k,d\geq 2} \Bigl(\zeta(k, \{1\}^{d-1}) - \zeta(d, \{1\}^{k-1})\Bigr) x^{k-1} y^{d-1}\\ &= \biggl(\frac{1}{y} - \frac{1}{x}\biggr)\biggl(1 - \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}\biggr) + \psi(1-x) - \psi(1-y).\end{aligned}$$

On the other hand, by Cor. 10.7.6

$$\begin{aligned}&\sum_{k,d\geq 2} \Bigl((-1)^k \zeta^\star(k, \{1\}^{d-1}) - (-1)^d \zeta^\star(d, \{1\}^{k-1})\Bigr) x^{k-1} y^{d-1}\\ &= -\psi(x) + \psi(y) - \pi(\cot(\pi x) - \cot(\pi y)) \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}.\end{aligned}$$

Combining with the well-known reflection formula (1.6) of the digamma function, we have

$$\begin{aligned}&\sum_{k,d\geq 2} \Bigl(\zeta^{\mathcal{S}}_{ш}(k, \{1\}^{d-1}) - \zeta^{\mathcal{S}}_{ш}(n, \{1\}^{k-1})\Bigr) x^{k-1} y^{d-1}\\ &= \biggl(\frac{1}{y} - \frac{1}{x}\biggr)\biggl(1 - \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}\biggr) + \psi(1-x) - \psi(1-y)\\ &\quad - \psi(x) + \psi(y) - \pi(\cot(\pi x) - \cot(\pi y)) \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}\\ &= \biggl(1 - \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}\biggr)(\psi(1-x) - \psi(1+x) - \psi(1-y) + \psi(1+y))\\ &= -2\biggl(1 - \frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}\biggr) \sum_{l\geq 1} \zeta(2l)(x^{2l-1} - y^{2l-1}),\end{aligned}$$

by Eq. (1.5). Since all the coefficients of $\Gamma(1-x)\Gamma(1-y)/\Gamma(1-x-y)$ can be expressed by as a polynomial of Riemann zeta values with rational coefficients according to Eq. (6.5), Thm. 6.2.4 follows immediately. □

6.3 Drinfeld associator and the symmetrized MZVs

To investigate other useful properties of the SMZVs it is more enlightening to consider their relations to the Drinfeld associator discussed in Sec. 4.1. Recall that $\Phi_{ш} = \pi(\Phi(\mathtt{a},\mathtt{b}))$ where $\Phi(\mathtt{a},\mathtt{b})$ is the Drinfeld associator and $\pi : \mathscr{R} \to \mathscr{R}/\mathscr{R}\mathtt{a}$. Here for simplicity, we rewrite $\mathtt{x}_0 = \mathtt{a}$ and $\mathtt{x}_1 = \mathtt{b}$. Then it follows from Thm. 4.1.11 that

$$\Phi_{ш}[\mathbf{w}] = (-1)^d \zeta_{ш}(\mathbf{w};0) \quad \text{for all } \mathbf{w} \in \mathfrak{A}^1_{ш}.$$

Therefore we get

$$\Phi_{ш}(\mathtt{a},\mathtt{b}) = \sum_{\mathbf{w}\in \mathtt{X}^*\mathtt{b}} (-1)^{\mathrm{dp}(\mathbf{w})} \zeta_{ш}(\mathbf{w};0)\mathbf{w}.$$

Motivated by this, we define

$$\Phi_{*}(\mathtt{a},\mathtt{b}) = \sum_{\mathbf{w}\in \mathtt{X}^*\mathtt{b}} (-1)^{\mathrm{dp}(\mathbf{w})} \zeta_{*}(\mathbf{w};0)\mathbf{w}.$$

As before we put $\zeta_{ш}(s_1,\dots,s_d) := (-1)^d \zeta_{ш}(\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b};0)$ and $\zeta_{*}(s_1,\dots,s_d) := (-1)^d \zeta_{*}(\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b};0)$. Then

Theorem 6.3.1. *For every word* $\mathbf{w} = \mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b} \in \mathfrak{A}^1_{ш}$, *the SMZV*

$$\zeta^{\mathcal{S}}_{ш}(s_1,\dots,s_d) = (-1)^d (\Phi^{-1}_{ш}\mathtt{b}\Phi_{ш})[\mathtt{b}\mathbf{w}].$$

Proof. By the duality relation Eq. (4.6), for all $i = 0,\dots,d$

$$\begin{aligned}\Phi^{-1}_{ш}[\mathtt{b}\mathtt{a}^{s_1-1}\cdots\mathtt{b}\mathtt{a}^{s_i-1}] &= \Phi_{ш}[\mathtt{a}\mathtt{b}^{s_1-1}\cdots\mathtt{a}\mathtt{b}^{s_i-1}]\\ &= (-1)^{s_1+\cdots+s_i-i}\zeta_{ш}(\mathtt{a}\mathtt{b}^{s_1-1}\cdots\mathtt{a}\mathtt{b}^{s_i-1};0)\\ &= (-1)^{s_1+\cdots+s_i-i}\zeta_{ш}(\mathtt{a}^{s_i-1}\mathtt{b}\cdots\mathtt{a}^{s_1-1}\mathtt{b};0)\end{aligned}$$

by duality relation (see Thm. 5.1.2). Therefore

$$\begin{aligned}&(\Phi^{-1}_{ш}\mathtt{b}\Phi_{ш})[\mathtt{b}\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}]\\ &= \sum_{i=0}^{d} \Phi^{-1}_{ш}[\mathtt{b}\mathtt{a}^{s_1-1}\cdots\mathtt{b}\mathtt{a}^{s_i-1}]\Phi_{ш}[\mathtt{a}^{s_{i+1}-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}]\\ &= \sum_{i=0}^{d} (-1)^{s_1+\cdots+s_i+d}\zeta_{ш}(s_i,\dots,s_1)\zeta_{ш}(s_{i+1},\dots,s_d) = (-1)^d\zeta^{\mathcal{S}}_{ш}(\mathbf{s})\end{aligned}$$

by Prop. 6.1.3, as desired. □

We can also use the Drinfeld associator to give another proof of Cor. 6.1.5. In fact, setting $\mathrm{inv}(\mathtt{e}) = 1$ and

$$\mathrm{inv}(\mathtt{x}_{i_1} \dots \mathtt{x}_{i_r}) = (-1)^r \mathtt{x}_{i_r} \dots \mathtt{x}_{i_1} \quad \text{for all } i_1, \dots, i_r = 0 \text{ or } 1. \tag{6.6}$$

one can check (see Exercise 6.3) that Cor. 4.1.7 implies that Φ_*^{inv} is group-like, namely, $\Delta_*(\Phi_*^{\mathrm{inv}}) = \Phi_*^{\mathrm{inv}} \otimes \Phi_*^{\mathrm{inv}}$. Hence

$$\Delta_*(\Phi_*^{\mathrm{inv}}\Phi_*) = \Delta_*(\Phi_*^{\mathrm{inv}})\Delta_*(\Phi_*) = \Phi_*^{\mathrm{inv}}\Phi_* \otimes \Phi_*^{\mathrm{inv}}\Phi_*.$$

But clearly

$$\Phi_*^{\mathrm{inv}}\Phi_* = \sum_{\mathbf{w} \in \mathsf{X}^*\mathtt{x}_1} (-1)^{\mathrm{dp}(\mathbf{w})} \zeta_*^{\mathcal{S}}(\mathbf{w}; 0)\mathbf{w}. \tag{6.7}$$

So Cor. 6.1.5 follows from Thm. 4.1.6 directly.

Similar to Eq. (6.7) we also have

$$\Phi_{ш}^{\mathrm{inv}}\Phi_{ш} = \sum_{\mathbf{w} \in \mathsf{X}^*\mathtt{x}_1} (-1)^{\mathrm{dp}(\mathbf{w})} \zeta_{ш}^{\mathcal{S}}(\mathbf{w}; 0)\mathbf{w}. \tag{6.8}$$

Let MZ_w, $\mathsf{MZ}_{*,w}^{\mathcal{S}}$ and $\mathsf{MZ}_{ш,w}^{\mathcal{S}}$ be the $\mathbb{Q}$-vector space generated by all weight w MZVs, $*$-symmetrized MZVs and $ш$-symmetrized MZVs, respectively.

Theorem 6.3.2. *For all* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ *the SMZVs*

$$\zeta_{ш}^{\mathcal{S}}(\mathbf{s}) \equiv \zeta_*^{\mathcal{S}}(\mathbf{s}) \pmod{\zeta(2)},$$

where $\zeta(2)$ *means either* $\zeta(2)\mathsf{MZ}_{|\mathbf{s}|-2}$, *or* $\zeta(2)\mathsf{MZ}_{*,|\mathbf{s}|-2}^{\mathcal{S}}$, *or* $\zeta(2)\mathsf{MZ}_{ш,|\mathbf{s}|-2}^{\mathcal{S}}$. *Consequently, in any statement modulo products of MZVs we may write* $\zeta^{\mathcal{S}}(s_1, \dots, s_d)$ *for either one of the two symmetrized versions.*

Proof. According to Thm. 4.1.15 we have $\Phi_{ш}(\mathtt{a}, \mathtt{b}) = \Lambda\Phi_*(\mathtt{a}, \mathtt{b})$ where

$$\Lambda = \exp\Big(\sum_{n=2}^{\infty} (-1)^n \frac{\zeta(n)}{n} \mathtt{b}^n\Big) \text{ and } \Lambda^{\mathrm{inv}} = \exp\Big(\sum_{n=2}^{\infty} \frac{\zeta(n)}{n} \mathtt{b}^n\Big).$$

Since inv is an anti-automorphism

$$\Phi_{ш}^{\mathrm{inv}}\Phi_{ш} = \Phi_*^{\mathrm{inv}}\Lambda^{\mathrm{inv}}\Lambda\Phi_* = \Phi_*^{\mathrm{inv}} \exp\Big(\sum_{n=2}^{\infty} \frac{\zeta(2n)}{n} \mathtt{b}^{2n}\Big)\Phi_* \equiv \Phi_*^{\mathrm{inv}}\Phi_*$$

modulo $\zeta(2)\mathsf{MZ}_{*,|\mathbf{s}|-2}^{\mathcal{S}}$. Similarly, we have

$$\Phi_*^{\mathrm{inv}}\Phi_* = \Phi_{ш}^{\mathrm{inv}}(\Lambda^{\mathrm{inv}})^{-1}\Lambda^{-1}\Phi_{ш} = \Phi_{ш}^{\mathrm{inv}} \exp\Big(-\sum_{n=2}^{\infty} \frac{\zeta(2n)}{n} b^{2n}\Big)\Phi_{ш} \equiv \Phi_{ш}^{\mathrm{inv}}\Phi_{ш}$$

modulo $\zeta(2)\mathsf{MZ}_{ш,|\mathbf{s}|-2}^{\mathcal{S}}$ (or $\zeta(2)\mathsf{MZ}_{|\mathbf{s}|-2}$ by Prop. 6.1.3). Together with Eqs. (6.7) and (6.8) this completes the proof of the theorem. □

Remark 6.3.3. More detailed analysis reveals that $\zeta(2)\mathsf{MZ}_{|\mathbf{s}|-2}$ can be replaced by $\zeta(2)\mathsf{MZ}_{|\mathbf{s}|-2,\leq \mathrm{dp}(\mathbf{s})-1}$ where $\mathsf{MZ}_{w,\leq n}$ is the subspace of MZ_w generated by MZVs of depth at most n. Similar claims are true for the $*$- and $\sqcup\!\sqcup$-versions of MZVs, too.

The next theorem shows that the SMZVs satisfy not only stuffle relations but also some kind of the shuffle relations. Define

$$\tau(\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}) = (-1)^{s_1+\cdots+s_d}\mathtt{a}^{s_d-1}\mathtt{b}\cdots\mathtt{a}^{s_1-1}\mathtt{b}. \tag{6.9}$$

Theorem 6.3.4. *For all words* $\mathbf{u},\mathbf{w},\mathbf{w}' \in \mathfrak{A}^1$*, we have*

(i) $\zeta^{\mathcal{S}}_{\sqcup\!\sqcup}(\mathbf{u}\sqcup\!\sqcup\mathbf{v}) = \zeta^{\mathcal{S}}_{\sqcup\!\sqcup}(\tau(\mathbf{u})\mathbf{v})$,
(ii) $\zeta^{\mathcal{S}}_{\sqcup\!\sqcup}((\mathbf{w}\mathbf{u})\sqcup\!\sqcup\mathbf{v}) = \zeta^{\mathcal{S}}_{\sqcup\!\sqcup}(\mathbf{u}\sqcup\!\sqcup\tau(\mathbf{w})\mathbf{v})$,
(iii) $\zeta^{\mathcal{S}}_{\sqcup\!\sqcup}((\mathtt{a}^{s-1}\mathtt{b}\mathbf{u})\sqcup\!\sqcup\mathbf{v}) = (-1)^s\zeta^{\mathcal{S}}_{\sqcup\!\sqcup}(\mathbf{u}\sqcup\!\sqcup(\mathtt{a}^{s-1}\mathtt{b}\mathbf{v}))$.

Proof. Taking $\mathbf{u}=\emptyset$ and then $\mathbf{w}=\mathbf{u}$ we see that (ii) implies (i). Decomposing $\mathbf{w}$ into strings of the type $\mathtt{a}^{s-1}\mathtt{b}$ we see that (iii) implies (ii). To prove (iii), we notice that

$$\mathtt{b}\Big((\mathtt{a}^{s-1}\mathtt{b}\mathbf{u})\sqcup\!\sqcup\mathbf{v} - (-1)^s\mathbf{u}\sqcup\!\sqcup(\mathtt{a}^{s-1}\mathtt{b}\mathbf{v})\Big) = \sum_{i=0}^{s-1}(-1)^i(\mathtt{a}^{s-1-i}\mathtt{b}\mathbf{u})\sqcup\!\sqcup(\mathtt{a}^i\mathtt{b}\mathbf{v}). \tag{6.10}$$

By Thm. 6.3.1 it suffices to show that the image of the above under $E = \Phi_{\sqcup\!\sqcup}^{-1}\mathtt{b}\Phi_{\sqcup\!\sqcup}$ vanishes. Now

$$\Delta_{\sqcup\!\sqcup}(E) = (\Phi_{\sqcup\!\sqcup}^{-1}\otimes\Phi_{\sqcup\!\sqcup}^{-1})(\mathtt{b}\otimes 1+1\otimes\mathtt{b})(\Phi_{\sqcup\!\sqcup}\otimes\Phi_{\sqcup\!\sqcup}) = E\otimes 1+1\otimes E.$$

Thus E is a primitive element for $\Delta_{\sqcup\!\sqcup}$ so that we can regard it as a Lie element, namely, it acts on shuffle products like a "derivation". By the definition the counit $\epsilon_{\sqcup\!\sqcup}$ satisfies $\epsilon_{\sqcup\!\sqcup}[\mathbf{w}]=0$ for all $\mathbf{w}\neq\mathtt{e}$. Hence for any nonempty words $\mathbf{u},\mathbf{v}\in\mathfrak{A}^1$,

$$E[\mathbf{u}\sqcup\!\sqcup\mathbf{v}] = E[\mathbf{u}]\epsilon_{\sqcup\!\sqcup}[\mathbf{v}] + \epsilon_{\sqcup\!\sqcup}[\mathbf{u}]E[\mathbf{v}] = 0.$$

This completes the proof by Thm. 6.3.1 since none of the factors in the shuffle products of the right-hand side of Eq. (6.10) is empty word ($\mathtt{b}$ is always contained). □

Besides the stuffle and the shuffle relations MZVs satisfy quite a few other well-known relations such as the duality relation and the sum formula which will be considered in the next chapter. For the SMZVs, there is in fact another kind of duality which we call *φ-dual*. On the level of words it corresponds to the involution

$$\mathtt{a}^{\varphi} = \mathtt{a}+\mathtt{b}, \mathtt{b}^{\varphi} = -\mathtt{b}. \tag{6.11}$$

Theorem 6.3.5. *For all* $\mathbf{w} \in \mathfrak{A}^1_{ш}$ *we have*

$$\zeta^{\mathcal{S}}(\mathbf{w}) \equiv \zeta^{\mathcal{S}}(\mathbf{w}^{\varphi}) \pmod{\zeta(2)},$$

where $\zeta^{\mathcal{S}}(\mathbf{w})$ *means either* $\zeta^{\mathcal{S}}_{ш}(\mathbf{w})$ *or* $\zeta^{\mathcal{S}}_{*}(\mathbf{w})$.

Proof. Throughout this proof we write $\Phi_{\mathtt{a},\mathtt{b}} = \Phi(\mathtt{a},\mathtt{b})$. Set $\mathtt{c} = -(\mathtt{a}+\mathtt{b})$. For any $\mathbf{t} \in \mathbb{N}^d$, we first decipher the coefficient

$$\Phi_{\mathtt{a},\mathtt{c}}[\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{b}\mathtt{a}^{t_d-1}\mathtt{b}].$$

Observe that the expansion of $\mathtt{a}^{r_1-1}(\mathtt{a}+\mathtt{b})\cdots\mathtt{a}^{r_\ell-1}(\mathtt{a}+\mathtt{b})$ contains $\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_d-1}\mathtt{b}$ if and only if it has the form

$$\alpha_{1,1}\cdots\alpha_{1,t_1-1}(\mathtt{a}+\mathtt{b})\cdots\alpha_{d,1}\cdots\alpha_{d,t_d-1}(\mathtt{a}+\mathtt{b})$$

where $\alpha_{\mu,\nu} = \mathtt{a}$ or $\alpha_{\mu,\nu} = (\mathtt{a}+\mathtt{b})$. Since the coefficient of $\mathtt{a}^{r_1-1}(\mathtt{a}+\mathtt{b})\cdots\mathtt{a}^{r_\ell-1}(\mathtt{a}+\mathtt{b})$ in $\Phi_{\mathtt{a},\mathtt{c}}$ is $(-1)^\ell\Phi_{a,b}[\mathtt{a}^{r_1-1}\mathtt{b}\cdots\mathtt{a}^{r_\ell-1}\mathtt{b}]$ we get

$$\begin{aligned}&\Phi_{\mathtt{a},\mathtt{c}}[\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_d-1}\mathtt{b}]\\ &= \sum_{\substack{\beta_{\mu,\nu}=\mathtt{a},\mathtt{b}\\ 1\le\mu\le d, 1\le\nu<t_\mu}} (-1)^{d+\sharp\{(\mu,\nu):\beta_{\mu,\nu}=\mathtt{b}\}}\Phi_{\mathtt{a},\mathtt{b}}[\beta_{1,1}\cdots\beta_{1,t_1-1}\mathtt{b}\cdots\beta_{\ell,1}\cdots\beta_{d,t_d-1}\mathtt{b}].\end{aligned}$$

Putting $\mathtt{a}^\natural = \mathtt{a}-\mathtt{b}, \mathtt{b}^\natural = -b$ we can rewrite this more compactly as

$$\Phi_{\mathtt{a},\mathtt{c}}[\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_d-1}\mathtt{b}] = \Phi_{\mathtt{a},\mathtt{b}}[(\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_d-1}\mathtt{b})^\natural]. \tag{6.12}$$

For any $\mathbf{t} = (t_1,\dots,t_d) \in \mathbb{N}^d$ with $t_d \ge 2$, we turn to the coefficient

$$\Phi^{-1}_{\mathtt{a},\mathtt{c}}[\mathtt{b}\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{b}\mathtt{a}^{t_d-1}].$$

Observe that the expansion of $(\mathtt{a}+\mathtt{b})^{r_1-1}\mathtt{a}\cdots(\mathtt{a}+\mathtt{b})^{r_\ell-1}\mathtt{a}$ contains $b\mathtt{a}^{t_1-1}\cdots\mathtt{b}\mathtt{a}^{t_d-1}$ if and only if it has the form

$$(\mathtt{a}+\mathtt{b})\alpha_{1,1}\cdots\alpha_{1,t_1-1}\cdots(\mathtt{a}+\mathtt{b})\alpha_{d,1}\cdots\alpha_{d,t_d-1}$$

where $\alpha_{\mu,\nu} = \mathtt{a}$ or $\alpha_{\mu,\nu} = (\mathtt{a}+\mathtt{b})$. Since the coefficient of $(\mathtt{a}+\mathtt{b})^{r_1-1}\mathtt{a}\cdots(\mathtt{a}+\mathtt{b})^{r_\ell-1}\mathtt{a}$ in $\Phi^{-1}_{\mathtt{a},\mathtt{c}}$ is $(-1)^{|\mathbf{r}|-\ell}\Phi_{a,b}[\mathtt{a}^{r_1-1}\mathtt{b}\cdots\mathtt{a}^{r_\ell-1}\mathtt{b}]$, by duality (essentially by Lemma 2.1.2(iii)) we get

$$\begin{aligned}&\Phi^{-1}_{\mathtt{a},\mathtt{c}}[\mathtt{b}\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_d-1}]\\ &= \sum_{\substack{\beta_{\mu,\nu}=\mathtt{a},\mathtt{b}\\ 1\le\mu\le d, 1\le\nu<t_\mu}} (-1)^{d+\boldsymbol{\beta}(\mathtt{a})}\Phi_{\mathtt{a},\mathtt{b}}[\mathtt{a}\beta_{1,1}\cdots\beta_{1,t_1-1}\cdots\mathtt{a}\beta_{d,1}\cdots\beta_{d,t_d-1}]\\ &= \sum_{\substack{\beta_{\mu,\nu}=\mathtt{a},\mathtt{b}\\ 1\le\mu\le d, 1\le\nu<t_\mu}} (-1)^{d+\boldsymbol{\beta}(\mathtt{b})+|\mathbf{t}|}\Phi_{\mathtt{a},\mathtt{b}}[\beta_{d,t_d-1}\cdots\beta_{d,1}\mathtt{b}\cdots\beta_{1,t_1-1}\cdots\beta_{1,1}\mathtt{b}],\end{aligned}$$

where $\boldsymbol{\beta}(x) = \sharp\{(\mu,\nu) : \beta_{\mu,\nu} = x\}$ for $x = \mathtt{a}$ or $\mathtt{b}$. Here we used the fact that for any admissible word $\mathbf{w} = l_1 \dots l_d$, $\Phi_{\mathtt{a},\mathtt{b}}[\mathbf{w}] = (-1)^{\sharp\{i:l_i=\mathtt{b}\}} \int_0^1 \mathbf{w}$. Hence

$$\Phi_{\mathtt{a},\mathtt{c}}^{-1}[\mathtt{b}\mathtt{a}^{t_1-1}\mathtt{b}\cdots\mathtt{a}^{t_d-1}] = (-1)^{|\mathbf{t}|}\Phi_{\mathtt{a},\mathtt{b}}[(\mathtt{a}^{t_d-1}\mathtt{b}\cdots\mathtt{a}^{t_1-1}\mathtt{b})^\natural]. \tag{6.13}$$

Combining the two Eqs. (6.12) and (6.13) we see that for any $\mathbf{s} \in \mathbb{N}^d$,

$$\begin{aligned}
&\Phi_{\mathtt{a},\mathtt{c}}^{-1}\mathtt{b}\Phi_{\mathtt{a},\mathtt{c}}[\mathtt{b}\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}]\\
&= \sum_{\substack{0\le i\le d\\ s_i\ge 2}} \Phi_{\mathtt{a},\mathtt{c}}^{-1}[\mathtt{b}\mathtt{a}^{s_1-1}\mathtt{b}\cdots\mathtt{a}^{s_i-1}]\Phi_{\mathtt{a},\mathtt{c}}[\mathtt{a}^{s_{i+1}-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b}]\\
&= \sum_{\substack{0\le i\le d\\ s_i\ge 2}} (-1)^{s_1+\cdots+s_i}\Phi_{\mathtt{a},\mathtt{b}}[(\mathtt{a}^{s_i-1}\mathtt{b}\cdots\mathtt{a}^{s_1-1}\mathtt{b})^\natural]\Phi_{\mathtt{a},\mathtt{b}}[(\mathtt{a}^{s_{i+1}-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b})^\natural]\\
&= \sum_{\substack{0\le i\le d\\ s_i\ge 2}} (-1)^{d+s_1+\cdots+s_i}\zeta_{ш}\big((\mathtt{a}^{s_i-1}\mathtt{b}\cdots\mathtt{a}^{s_1-1}\mathtt{b})^\varphi\big)\zeta_{ш}\big((\mathtt{a}^{s_{i+1}-1}\mathtt{b}\cdots\mathtt{a}^{s_d-1}\mathtt{b})^\varphi\big)\\
&\equiv (-1)^d\zeta_{ш}^{\mathcal{S}}(\mathbf{w}^\varphi)
\end{aligned} \tag{6.14}$$

modulo $\zeta(2)$ by Prop. 6.1.3, where $\mathtt{a}^\varphi = \mathtt{a}+\mathtt{b}, b^\varphi = -b$.

Mimicking the proof of Eq. (6.14) and using an argument similar to the proof of Eq. (6.3) one can show that (see Exercise 6.4)

$$\Phi_{\mathtt{a},\mathtt{c}}^{-1}\mathtt{a}\Phi_{\mathtt{a},\mathtt{c}}[\mathtt{b}\mathbf{w}] = \mathtt{a} \quad \forall \mathbf{w} \in \mathfrak{A}^1_{ш}. \tag{6.15}$$

So the theorem follows from Thm. 6.3.1 and Eq. (4.32). □

6.4* Space generated by symmetrized MZVs

In this section we show that no information is lost if we consider SMZVs instead of MZVs (see Thm. 6.4.19). In Chap. 8 this result will play a key role in a conjecture relating the MZVs to the finite version of MZVs. We break the proof of this theorem into a series of lemmas and propositions which will be used again in proving a similar result for the alternating Euler sums in Chap. 14.

6.4.1 *Natural symmetrized MZVs*

We will be interested in the behavior of depth reduction of MZVs so we let $\mathrm{Surj}(n,m)$ be the set of surjective maps from $\{1,\dots,n\}$ to $\{1,\dots,m\}$ for all positive integers $n \ge m$. For any $\phi \in \mathrm{Surj}(n,m)$ and $\mathbf{s} = (s_1,\dots,s_n) \in \mathbb{N}^n$ we set

$$\phi_*\mathbf{s} = \left(\sum_{i\in\phi^{-1}(1)} s_i, \sum_{i\in\phi^{-1}(2)} s_i, \dots, \sum_{i\in\phi^{-1}(m)} s_i\right).$$

Then we define

$$\zeta^{\natural,\mathcal{S}}(\mathbf{s}) := \sum_{1\le m\le n}\sum_{\phi\in \mathrm{Surj}(n,m)} \frac{1}{\prod_{j=1}^{m}(\sharp\phi^{-1}(j))!}\zeta_*^{\mathcal{S}}(\phi_*\mathbf{s}).$$

Remark 6.4.1. If we use $\zeta_{\sqcup\!\sqcup}^{\mathcal{S}}(\mathbf{s})$ instead of $\zeta_*^{\mathcal{S}}(\mathbf{s})$ to define $\zeta^{\natural,\mathcal{S}}(\mathbf{s})$ then the difference lies in $\zeta(2)\mathsf{MZ}_{|\mathbf{s}|-2,\le \mathrm{dp}(\mathbf{s})-1}$ by Remark 6.3.3.

Example 6.4.2. We have

$$\begin{aligned}\zeta^{\natural,\mathcal{S}}(s_1,s_2) &= \zeta_*^{\mathcal{S}}(s_1,s_2)+\frac{1}{2}\zeta_*^{\mathcal{S}}(s_1+s_2),\\ \zeta^{\natural,\mathcal{S}}(s_1,s_2,s_3) &= \zeta_*^{\mathcal{S}}(s_1,s_2,s_3)+\frac{1}{2}\zeta_*^{\mathcal{S}}(s_1+s_2,s_3)\\ &\quad+\frac{1}{2}\zeta_*^{\mathcal{S}}(s_1,s_2+s_3)+\frac{1}{6}\zeta_*^{\mathcal{S}}(s_1+s_2+s_3).\end{aligned}$$

Recall that MZ_w (resp. $\mathsf{MZ}_{w,\le d}$) is the $\mathbb{Q}$-vector space spanned by all the MZVs of weight w (resp. weight w and depth at most d). Now we further set

$$\begin{aligned}\mathsf{MZ}_{w,+} &= \sum_{w',w''\ge 1,\ w'+w''=w} \mathsf{MZ}_{w'}\cdot\mathsf{MZ}_{w''}\subseteq \mathsf{MZ}_w,\\ \mathsf{MZ}_{w,+,\le d} &= \sum_{\substack{w',w''\ge 1,\ w'+w''=w,\\ 1\le k,l\le d-1,\ k+l\le d}} \mathsf{MZ}_{w',\le k}\cdot\mathsf{MZ}_{w'',\le l}\subseteq \mathsf{MZ}_{w,\le d}.\end{aligned}$$

For each $d\in\mathbb{N}$ define the $\mathbb{Q}$-vector space of formal generating series

$$M_d := \left\{ \sum_{\mathbf{s}=(s_1,\dots,s_d)\in\mathbb{N}^d} a_{\mathbf{s}}\mathbf{x}^{\mathbf{s-1}} \;\middle|\; a_{\mathbf{s}}\in \mathsf{MZ}_{|\mathbf{s}|}\ \forall \mathbf{s}\in\mathbb{N}^d \right\}$$

where $\mathbf{x}^{\mathbf{s-1}} = x_1^{s_1-1}\cdots x_d^{s_d-1}$, and its subspace

$$M_{d,+} := \left\{ \sum_{\mathbf{s}=(s_1,\dots,s_d)\in\mathbb{N}^d} a_{\mathbf{s}}\mathbf{x}^{\mathbf{s-1}} \;\middle|\; a_{\mathbf{s}}\in \mathsf{MZ}_{|\mathbf{s}|,+}\cup \mathsf{MZ}_{|\mathbf{s}|,\le 1}\ \forall \mathbf{s}\in\mathbb{N}^d \right\}.$$

One can similarly define subspaces $M_{d,\le m}$ and $M_{d,+,\le m}$ Similar as in Chap. 3 we define the generating functions

$$\begin{aligned}F_{\sqcup\!\sqcup}(x_1,\dots,x_n) = F_{\sqcup\!\sqcup}^{(n)}(x_1,\dots,x_n) &:= \sum_{\mathbf{s}\in\mathbb{N}^n}\zeta_{\sqcup\!\sqcup}(\mathbf{s})\mathbf{x}^{\mathbf{s-1}},\\ F_*(x_1,\dots,x_n) = F_*^{(n)}(x_1,\dots,x_n) &:= \sum_{\mathbf{s}\in\mathbb{N}^n}\zeta_*(\mathbf{s})\mathbf{x}^{\mathbf{s-1}},\\ F_\natural(x_1,\dots,x_n) = F_\natural^{(n)}(x_1,\dots,x_n) &:= \sum_{\mathbf{s}\in\mathbb{N}^n}\zeta^\natural(\mathbf{s})\mathbf{x}^{\mathbf{s-1}}.\end{aligned}$$

where

$$\zeta^{\natural}(\mathbf{s}) = \sum_{1\le m\le n} \sum_{\phi\in \mathrm{Surj}(n,m)} \frac{1}{\prod_{j=1}^{m} (\sharp\phi^{-1}(j))!} \zeta_*(\phi_*\mathbf{s}).$$

By Remark 6.4.1 we could as well use $\zeta_{ш}(\phi_*\mathbf{s})$ to define $\zeta^{\natural}(\mathbf{s})$ when we consider identities modulo $\mathsf{MZ}_{|\mathbf{s}|,+}$. By convention, $F_{ш}^{(0)} = F_{*}^{(0)} = F_{\natural}^{(0)} = 1$.

6.4.2 *An embedding* $\iota_n : \mathfrak{S}_{n+1} \hookrightarrow \mathrm{GL}_n(\mathbb{Z})$

In the proof of the Parity Principle of MZVs, we defined an action of $\mathfrak{S}_{n+1}$ on V_n extending that of $\mathfrak{S}_n$. We now need to use a new construction. Recall that $-\nu_n$ is the $n\times n$ identity matrix and

$$P = P_n = \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}, \qquad P^{-1} = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

As before, by abuse of notation we let $T_0 = T_{n,0}$ denote the matrix corresponding to the reversal $T_{n,0} = \left(\begin{smallmatrix} 1 & 2 & \cdots & n \\ n & n-1 & \cdots & 1 \end{smallmatrix}\right) \in \mathfrak{S}_n$.

Lemma 6.4.3. *There exists a unique injective group homomorphism* $\iota_n : \mathfrak{S}_{n+1} \hookrightarrow \mathrm{GL}_n(\mathbb{Z})$ *which satisfies the following properties:*

- *If* $\sigma \in \mathfrak{S}_n$, *then* $\iota_n(\sigma) = M_\sigma$ *so that*

$$(x_{\sigma^{-1}(1)} \;\; x_{\sigma^{-1}(2)} \;\; \cdots \;\; x_{\sigma^{-1}(n)}) = (x_1 \;\; x_2 \;\; \cdots \;\; x_n) M_\sigma^{-1}.$$

- *The image under* ι_n *of the permutation*

$$T' = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n & n+1 \\ n-1 & n-2 & \cdots & 1 & n+1 & n \end{pmatrix}$$

  is equal to $\nu T_0 T_0^{\sharp} T_0$.

Proof. Let $\mathbb{Z}^n$ denote the abelian group of n-dimensional row vectors with integer components. By the multiplication from the right, the group $\mathrm{GL}_n(\mathbb{Z})$ acts on $\mathbb{Z}^n$ from the right. This yields the embedding $\mathfrak{S}_n \hookrightarrow \mathrm{GL}_n(\mathbb{Z})$, $\sigma \to M_\sigma^{-1}$.

But the embedding $\mathbb{Z}^n \hookrightarrow \mathbb{Z}^{(n+1)}$, $(a_1,\ldots,a_n) \to (a_1,\ldots,a_n,-(a_1+\cdots+a_n))$ has a stable image under the action of $\mathfrak{S}_{n+1}$, leading to a homomorphism $\iota_n : \mathfrak{S}_{n+1} \hookrightarrow \mathrm{GL}_n(\mathbb{Z})$. It is not hard to show that the image

$\iota_n(T')$ is given by

$$\iota_n(T') = \gamma_2 := \begin{pmatrix} & & 1 & 0 \\ & \iddots & & \vdots \\ 1 & & & 0 \\ -1 & \cdots & -1 & -1 \end{pmatrix} = \nu T_0 T_0^\sharp T_0 \quad \text{with} \quad \gamma_2^{-1} = \gamma_2. \tag{6.16}$$

Finally, the map ι_n is unique since $\mathfrak{S}_{n+1}$ is generated by $\mathfrak{S}_n$ and T'. □

6.4.3 *Parity results*

We have seen in Sec. 3.5 that if the depth and the weight of a MZV have different parity then they can be expressed as products of MZVs of lower depths. For SMZVs, we can say a lot more. Even in the same parity case, they can be reduced.

Proposition 6.4.4. *For all* $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{N}^n$*, if* $w := |\mathbf{s}|$ *and* n *have the same parity, then* $\zeta^{\natural,\mathcal{S}}(\mathbf{s}) \in \mathsf{MZ}_{w,+,\leq n}$.

Proof. First, Lemma 3.5.5 implies that $\zeta_*(\mathbf{s}) + (-1)^n \zeta_*(\overleftarrow{\mathbf{s}}) \equiv 0$ modulo $\mathsf{MZ}_{w,+,\leq n}$. Thus $\zeta_*(\mathbf{s}) + (-1)^w \zeta_*(\overleftarrow{\mathbf{s}}) \equiv 0$. So for all $\phi \in \mathrm{Surj}(n, m)$ we have $\zeta_*(\phi_* \mathbf{s}) + (-1)^w \zeta_*(\phi_* \overleftarrow{\mathbf{s}}) \equiv 0$. Hence by the definition

$$\zeta^{\natural,\mathcal{S}}(\mathbf{s}) \equiv \zeta^{\natural}(\mathbf{s}) + (-1)^w \zeta^{\natural}(\overleftarrow{\mathbf{s}}) \equiv 0$$

modulo $\mathsf{MZ}_{w,+,\leq n}$, as desired. □

Corollary 6.4.5. *For all* $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{N}^n$*, if* $w := |\mathbf{s}|$ *and* n *have the same parity, then*

$$\zeta^{\mathcal{S}}_{ш}(\mathbf{s}) \equiv \zeta^{\mathcal{S}}_*(\mathbf{s}) \equiv -\sum_{i=1}^{n-1} \zeta_*(s_1, \dots, s_{i-1}, s_i + s_{i+1}, s_{i+1}, \dots, s_n)$$

modulo $\mathsf{MZ}_{w,+,\leq n} + \mathsf{MZ}_{w,\leq n-2}$.

Proof. By the definition and Prop. 6.4.4

$$\zeta^{\natural,\mathcal{S}}(\mathbf{s}) \equiv \zeta^{\mathcal{S}}_*(\mathbf{s}) + \frac{1}{2} \sum_{i=1}^{n-1} \zeta^{\mathcal{S}}_*(s_1, \dots, s_{i-1}, s_i + s_{i+1}, s_{i+1}, \dots, s_n)$$

modulo $\mathsf{MZ}_{w,\leq n-2}$. But for all $\mathbf{t}$ of depth $n-1$ we have

$$\zeta^{\mathcal{S}}_*(\mathbf{t}) \equiv \zeta^{\mathcal{S}}_*(\mathbf{t}) + (-1)^w \zeta_*(\overleftarrow{\mathbf{t}}) \equiv \zeta^{\mathcal{S}}_*(\mathbf{t}) + (-1)^{w+n} \zeta^{\mathcal{S}}_*(\mathbf{t}) \equiv 2\zeta^{\mathcal{S}}_*(\mathbf{t})$$

modulo $\mathsf{MZ}_{w,+,\leq n}$. So the corollary follows. □

The case of different parities is much more involved. In the rest of this subsection, we assume this is the case. By Theorem 3.5.3 and Eq. (3.32) we see that $F_*|T_0 \equiv F_*|T_0^\sharp \equiv (-1)^{n-1}F_* \pmod{\mathsf{MZ}_{w,+,\le n}}$. Similar congruences hold for $F_\natural$ and $F_{\sqcup\!\sqcup}$, too. The same remark can be applied throughout the rest of this subsection.

To save space, we put $M'_n = M_{n,+,\le n} + M_{n,\le n-2}$. By Eq. (3.34) (A) we have $T_0T' = C_{n+1}$ which yields $\iota(C_{n+1}) = \nu_n T_0^\sharp T_0$. Hence, by Eq. (3.35) we have

$$F_\natural \equiv F_\natural\big|\big[1 + \mathrm{sh}_1^{(n)}\,\nu_n T_0^\sharp T_0\big] = F_\natural\big|\big[\nu_n T_0^\sharp T_0\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\big] \pmod{M'_n}$$

where τ is the transposition $(1, n+1)$. Further, since ν commutes with everything else we get

$$\begin{aligned}
F_\natural &\equiv (-1)^{n-1}(F_\natural - F_*)\big|\big[T_0T_0^\sharp T_0\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\nu\big] + F_*\big|\big[\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\nu\big] \\
&\equiv (-1)^{n-1}(F_\natural - F_*)\big|\big[\tilde{T}\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\nu\big] \\
&\quad + (F_* - F_\natural)\big|\big[\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\nu\big] + F_\natural\big|\big[\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\nu\big] \\
&\equiv F_\natural|\nu - (F_\natural - F_*)\big|\big[\big((-1)^n\tilde{T} + 1\big)\big(1 + \mathrm{sh}_1^{(n)}\,\iota(\tau)\big)\nu\big] \qquad (6.17)
\end{aligned}$$

modulo M'_n, where $\tilde{T} = \tilde{T}_n := T_{n,0}T_{n,0}^\sharp T_{n,0}$. Now

$$F_\natural - F_* \equiv \frac{1}{2}\sum_{i=1}^{n-1} \frac{F_\natural(x_1, \dots, \widehat{x_{i+1}}, \dots, x_n) - F_\natural(x_1, \dots, \widehat{x_i}, \dots, x_n)}{x_i - x_{i+1}} \qquad (6.18)$$

modulo M'_n, where $F_\natural = F_\natural^{(n-1)}$ on the right-hand side. Note that

$$\begin{aligned}
&\frac{F_\natural(x_1, \dots, x_i, \widehat{x_{i+1}}, \dots, x_n) - F_\natural(x_1, \dots, \widehat{x_i}, x_{i+1}, \dots, x_n)}{x_i - x_{i+1}}\Big|\,\tilde{T} \\
&= \begin{cases} \dfrac{F_\natural(y_1, \dots, \widehat{y_{i+1}}, \dots, y_n) - F_\natural(y_1, \dots, \widehat{y_i}, \dots, y_n)}{x_{n-i-1} - x_{n-i}} & \text{if } i \le n-2; \\ \dfrac{F_\natural(y_1, \dots, y_{n-1}) - F_\natural(y_1, \dots, y_{n-2}, x_n)}{-x_1} & \text{if } i = n-1, \end{cases}
\end{aligned}$$

where $y_i = x_n - x_{n-i}$ for $i = 1, \dots, n$ and $x_0 = 0$. Since $F_\natural\big|T_{n-1,0} = F_\natural\big|T_{n-1,0}^\sharp \equiv (-1)^n F_\natural \pmod{M_{n-1,+,\le n-1}}$ we get

$$\begin{aligned}
F_\natural(y_1, \dots, \widehat{y_i}, \dots, y_n) &= F_\natural\big|\tilde{T}_{n-1}(x_1, \dots, \widehat{x_{n-i}}, \dots, x_n) \\
&\equiv (-1)^n F_\natural(x_1, \dots, \widehat{x_{n-i}}, \dots, x_n)
\end{aligned}$$

and similarly

$$\begin{aligned}
F_\natural(y_1, \dots, y_{n-1}) &= F_\natural\big|\tilde{T}_{n-1}(x_2 - x_1, \dots, x_n - x_1) \\
&\equiv (-1)^n F_\natural(x_2 - x_1, \dots, x_n - x_1)
\end{aligned}$$

modulo M'_n. Hence

$$\begin{aligned}
&\left.\frac{F_\natural(x_1,\dots,\widehat{x_{i+1}},\dots,x_n)-F_\natural(x_1,\dots,\widehat{x_i},\dots,x_n)}{x_i-x_{i+1}}\right|(-1)^n\tilde{T}\\
&\equiv\begin{cases}\dfrac{F_\natural(x_1,\dots,\widehat{x_{n-i-1}},\dots,x_n)-F_\natural(x_1,\dots,\widehat{x_{n-i}},\dots,x_n)}{x_{n-i-1}-x_{n-i}} & \text{if } i\le n-2;\\ \dfrac{F_\natural(x_2-x_1,\dots,x_n-x_1)-F_\natural(x_2,\dots,x_n)}{-x_1} & \text{if } i=n-1,\end{cases}
\end{aligned}$$

modulo M'_n. Therefore, by cancelation Eq. (6.18) yields

$$\begin{aligned}
2(F_\natural^{(n)}-F_*^{(n)})\big|((-1)^n\tilde{T}+1)\equiv{}&\frac{F_\natural(x_2,\dots,x_n)-F_\natural(x_2-x_1,\dots,x_n-x_1)}{x_1}\\
&+\frac{F_\natural(x_1,\dots,x_{n-1})-F_\natural(x_1,\dots,\widehat{x_{n-1}},x_n)}{x_{n-1}-x_n}.
\end{aligned}\tag{6.19}$$

modulo M'_n. Applying $\mathrm{sh}_1^{(n)}$ to both sides of Eq. (6.19) we get

$$\begin{aligned}
2(F_\natural-F_*)\big|((-1)^n\tilde{T}+1)\,\mathrm{sh}_1^{(n)}\equiv{}&\frac{F_\natural(x_2,\dots,x_n)-F_\natural(x_2-x_1,\dots,x_n-x_1)}{x_1}\\
&+\frac{F_\natural(x_2,\dots,\widehat{x_{n-1}},x_n,x_1)-F_\natural(x_2,\dots,x_{n-1},x_1)}{x_{n-1}-x_n}\\
&+2\cdot\frac{F_\natural(x_2,\dots,x_{n-1},x_1)-F_\natural(x_2,\dots,x_n)}{x_1-x_n}
\end{aligned}\tag{6.20}$$

modulo M'_n. Define $\mathbf{1}=(1,\dots,1)$, $\mathbf{0}=(0,\dots,0)$, and

$$Q:=Q_{n-1}=\nu_{n-1}T^\sharp_{n-1,0}T_{n-1,0}=\begin{pmatrix}-\mathbf{1} & -1\\ I_{n-2} & \mathbf{0}^t\end{pmatrix}\in\mathrm{GL}_{n-1}(\mathbb{Z}),$$

where $\mathbf{0}^t$ is the transpose of $\mathbf{0}$. Then by applying

$$\iota(\tau)=\begin{pmatrix}-1 & -\mathbf{1}\\ \mathbf{0}^t & I_{n-1}\end{pmatrix}\in\mathrm{GL}_n(\mathbb{Z})$$

to Eq. (6.20) and using $F_\natural|Q^{-1}\equiv F_\natural|\nu$ we get

$$\begin{aligned}
&2(F_\natural-F_*)\big|((-1)^n\tilde{T}+1)\,\mathrm{sh}_1^{(n)}\,\iota(\tau)\\
\equiv{}&\frac{F_\natural(x_2,\dots,x_n)-F_\natural(x_2-x_1,\dots,x_n-x_1)}{x_1}\\
&+\frac{F_\natural|Q^{-1}(x_1,\dots,\widehat{x_{n-1}},x_n)-F_\natural|Q^{-1}(x_1,\dots,x_{n-1})}{x_{n-1}-x_n}\\
&+2\cdot\frac{F_\natural|Q^{-1}(x_1,\dots,x_{n-1})-F_\natural|Q^{-1}(x_1-x_n,\dots,x_{n-1}-x_n)}{-x_n}
\end{aligned}$$

$$\begin{aligned}
&\equiv \frac{F_\natural(x_2,\dots,x_n) - F_\natural(x_2-x_1,\dots,x_n-x_1)}{x_1} \\
&+ \frac{F_\natural|\nu(x_1,\dots,\widehat{x_{n-1}},x_n) - F_\natural|\nu(x_1,\dots,x_{n-1})}{x_{n-1}-x_n} \\
&+ 2\cdot \frac{F_\natural|\nu(x_1,\dots,x_{n-1}) - F_\natural|\nu(x_1-x_n,\dots,x_{n-1}-x_n)}{-x_n}
\end{aligned}$$

modulo M_n'. Combining with Eq. (6.19) we have

$$\begin{aligned}
&2(F_\natural - F_*)\big|((-1)^n\tilde{T}+1)(1+\mathrm{sh}_1^{(n)}\,\iota(\tau)) \\
\equiv& 2\cdot \frac{F_\natural(x_2,\dots,x_n) - F_\natural(x_2-x_1,\dots,x_n-x_1)}{x_1} \\
&- \frac{F_\natural\big|(1-\nu)(x_1,\dots,\widehat{x_{n-1}},x_n) - F_\natural\big|(1-\nu)(x_1,\dots,x_{n-1})}{x_{n-1}-x_n} \\
&+ 2\cdot \frac{F_\natural|\nu(x_1-x_n,\dots,x_{n-1}-x_n) - F_\natural|\nu(x_1,\dots,x_{n-1})}{x_n}
\end{aligned} \tag{6.21}$$

modulo M_n'. Hence by Eq. (6.17), we have

$$F_\natural\big|(\nu-1) \equiv \frac{1}{2}\Big(\text{right-hand side of Eq.(6.21)}\Big) \tag{6.22}$$

modulo M_n'. Let $F_\natural^{\mathrm{ev}}$ and $F_\natural^{\mathrm{od}}$ denote the even degree part and the odd degree part of $F_\natural$, respectively. Since $F_\natural^{\mathrm{od}}|\nu = -F_\natural^{\mathrm{od}}$ and $F_\natural^{\mathrm{ev}}|\nu = F_\natural^{\mathrm{ev}}$, by Eq. (6.22), we have

$$\begin{aligned}
2F_\natural^{\mathrm{od}}(x_1,\dots,x_n) \equiv& \frac{F_\natural^{\mathrm{ev}}(x_2-x_1,\dots,x_n-x_1) - F_\natural^{\mathrm{ev}}(x_2,\dots,x_n)}{x_1} \\
&- \frac{F_\natural^{\mathrm{ev}}(x_1-x_n,\dots,x_{n-1}-x_n) - F_\natural^{\mathrm{ev}}(x_1,\dots,x_{n-1})}{x_n}
\end{aligned} \tag{6.23}$$

modulo M_n'.

Remark 6.4.6. (1) If we define $\zeta^{\natural,\mathcal{S}}$ using $\zeta_{ш}$ as mentioned in Remark 6.4.1 then all the arguements in this subsection are still valid modulo M_n'.

(2) We may also replace F_* in the above.

Proposition 6.4.7. *For all* $\mathbf{s} = (s_1,\dots,s_n) \in \mathbb{N}^n$*, if* $w := |\mathbf{s}|$ *and* n *have*

different parities and $|\mathbf{s}| \neq 2$*, then modulo* $\mathsf{MZ}_{w,+,\leq n} + \mathsf{MZ}_{w,\leq n-2}$ *we have:*

$$\begin{aligned}
&\zeta^{\natural,\mathcal{S}}(s_1,\dots,s_n)\\
&\equiv (-1)^{s_1} \sum_{\substack{\ell_2,\dots,\ell_n\geq 0\\ \ell_2+\cdots+\ell_n=s_n}} \prod_{j=2}^{n} \binom{s_j+\ell_j-1}{\ell_j} \zeta_*(s_2+\ell_2,\dots,s_n+\ell_n)\\
&-(-1)^{s_n} \sum_{\substack{\ell_1,\dots,\ell_{n-1}\geq 0\\ \ell_1+\cdots+\ell_{n-1}=s_1}} \prod_{j=1}^{n-1} \binom{s_j+\ell_j-1}{\ell_j} \zeta_*(s_1+\ell_1,\dots,s_{n-1}+\ell_{n-1}).
\end{aligned} \tag{6.24}$$

Proof. Comparing the coefficients of $x_1^{s_1-1}\cdots x_n^{s_n-1}$ on both sides of Eq. (6.23), we have

$$\begin{aligned}
&2\zeta_*^{\natural}(s_1,\dots,s_n)\\
&\equiv (-1)^{s_1} \sum_{\substack{\ell_2,\dots,\ell_n\geq 0\\ \ell_2+\cdots+\ell_n=s_1}} \prod_{j=2}^{n} \binom{s_j+\ell_j-1}{\ell_j} \zeta_*^{\natural}(s_2+\ell_2,\dots,s_n+\ell_n)\\
&-(-1)^{s_n} \sum_{\substack{\ell_1,\dots,\ell_{n-1}\geq 0\\ \ell_1+\cdots+\ell_{n-1}=s_n}} \prod_{j=1}^{n-1} \binom{s_j+\ell_j-1}{\ell_j} \zeta_*^{\natural}(s_1+\ell_1,\dots,s_{n-1}+\ell_{n-1}).
\end{aligned}$$

modulo $\mathsf{MZ}_{w,+,\leq n} + \mathsf{MZ}_{w,\leq n-2}$. Since $\zeta_*^{\natural}(\mathbf{t}) \equiv \zeta_*(\mathbf{t})$ modulo $\mathsf{MZ}_{k,\leq n-2}$ for all $\mathbf{t} \in \mathbb{N}^{n-1}$ and

$$\zeta^{\natural,\mathcal{S}}(\mathbf{s}) \equiv \zeta_*^{\natural}(\mathbf{s}) + (-1)^w \zeta_*^{\natural}(\overleftarrow{\mathbf{s}}) \equiv \zeta_*^{\natural}(\mathbf{s}) + (-1)^{w+n-1}\zeta_*^{\natural}(\mathbf{s}) \equiv 2\zeta_*^{\natural}(\mathbf{s})$$

modulo $\mathsf{MZ}_{w,+,\leq n}$, the proposition follows immediately. □

Remark 6.4.8. In view of Remark 6.4.6 we may replace ζ_* by $\zeta_{ш}$ on the right-hand side of Eq. (6.24). This also follows from the fact that no index of the form $(1,\dots,1)$ appears there.

6.4.4 *Diagonal translation invariant polynomials*

In order to prove the main result Thm. 6.4.19 of this section, we need to study a class of polynomials which are called diagonal translation invariant.

Lemma 6.4.9. *Let* $f(x_1,\dots,x_n) \in \mathbb{Q}[x_1,\dots,x_n]$*. Then the following conditions are equivalent:*

(i). $f(x_1+t,\dots,x_n+t) = f(x_1,\dots,x_n)$ *for all* $t \in \mathbb{R}$*.*

(ii). *There exists a polynomial g in $n-1$ variables such that*

$$f(x_1,\dots,x_n) = g(x_1-x_n,\dots,x_{n-1}-x_n).$$

(iii). $\displaystyle\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} = 0.$

Proof. (i)$\Longleftrightarrow$(iii). Let $h(t) = f(x_1+t,\dots,x_n+t)$. Then by the Chain Rule we have

$$\text{(i)} \iff h(t) = h(0) \iff h'(t) = 0 \iff \text{(iii)}.$$

(i)$\Longrightarrow$(ii). Take $t = -x_n$ and $g(x_1,\dots,x_{n-1}) = f(x_1,\dots,x_{n-1},0)$.
(ii)$\Longrightarrow$(i). Trivial.

□

Definition 6.4.10. A polynomial $f(x_1,\dots,x_n) \in \mathbb{Q}[x_1,\dots,x_n]$ is called *diagonal translation invariant* (DTI) if the equivalent conditions in Lemma 6.4.9 are satisfied.

Definition 6.4.11. A polynomial $f(x_1,\dots,x_n) \in \mathbb{Q}[x_1,\dots,x_n]$ is called *pseudo-symmetric* if the $f|\sigma = \pm f$ for all permutations $\sigma \in \mathfrak{S}_n$.

Lemma 6.4.12. *Suppose $f, f_1,\dots,f_m \in \mathbb{Q}[x_1,\dots,x_n]$ are all DTI. Then*

(i). *$f|\sigma$ is DTI for every permutation $\sigma \in \mathfrak{S}_n$.*
(ii). *Every $\mathbb{Q}$-linear combination of $f_1,\dots,f_m$ is DTI.*

Proof. By assumption $h(x_1+t,\dots,x_n+t) = h(x_1,\dots,x_n)$ for all $t \in \mathbb{R}$ and $h = f, f_1,\dots,f_m$. So the lemma is obvious. □

Lemma 6.4.13. *Let $f(x_1,\dots,x_n)$ be a pseudo-symmetric polynomial in n variables. Suppose that $\frac{\partial}{\partial x_n}(x_n f)$ is DTI. Then f is constant.*

Proof. We may assume that f is homogeneous of degree d. Let σ be the transposition (j,n). Then by pseudo-symmetry and Lemma 6.4.12

$$\frac{\partial}{\partial x_j}(x_j f) = \pm \left(\frac{\partial}{\partial x_n}(x_n f) \right) \Bigg| \sigma$$

is DTI for every positive integer $j \le n$. Hence, by Euler's Homogeneous Function Theorem and Lemma 6.4.12(ii)

$$f = \frac{1}{n+d} \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(x_i f)$$

is DTI. For any positive integer $j \leq n$, by direct computation

$$\frac{\partial f}{\partial x_j} = \sum_i \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}(x_j f) - \frac{\partial}{\partial x_j}\left(x_j \sum_i \frac{\partial f}{\partial x_i}\right) = 0.$$

Thus f is constant, as desired. □

Lemma 6.4.14. *Let $f(x_1, \ldots, x_n)$ be a homogeneous polynomial of degree $d \geq 2$. Suppose that*

$$\frac{1}{x_{n+1}}\big(f(x_1 - x_{n+1}, \ldots, x_n - x_{n+1}) - f(x_1, \ldots, x_n)\big)$$

is a pseudo-symmetric polynomial in $n+1$ variables. Then f is DTI.

Proof. We set $h(x_1, \ldots, x_{n+1}) = \frac{1}{x_{n+1}}\big(f(x_1 - x_{n+1}, \ldots, x_n - x_{n+1}) - f(x_1, \ldots, x_n)\big)$. By substition $x_j \to x_j + t$ one checks easily that

$$\frac{\partial}{\partial x_{n+1}}(x_{n+1}h) = \frac{\partial}{\partial x_{n+1}} f(x_1 - x_{n+1}, \ldots, x_n - x_{n+1})$$

is DTI. Further, h is a pseudo-symmetric by assumption, Lemma 6.4.13 then shows h is a constant of degree $d - 1 \geq 1$. So we must have $h = 0$. Hence f is DTI by defnition (setting $x_{n+1} = t$). □

Recall that $\mathsf{DSh}_n(d)$ is the degree d part of the double shuffle space

$$\mathsf{DSh}_n = \left\{ f \in V_n \;\middle|\; f|\operatorname{sh}_j^{(n)} = f^\sharp|\operatorname{sh}_j^{(n)} = 0 \text{ for } j = 1, \ldots, n-1 \right\}$$

defined by Definition 3.5.1.

Lemma 6.4.15. *Let $d \geq 2$ and $f \in \mathsf{DSh}_n(d)$. If*

$$\frac{1}{x_{n+1}}\big(f(x_1 - x_{n+1}, \ldots, x_n - x_{n+1}) - f(x_1, \ldots, x_n)\big)$$

is invariant under the cyclic permutations then f is DTI.

Proof. Set

$$h(x_1, \ldots, x_{n+1}) := \frac{1}{x_{n+1}}\big(f(x_1 - x_{n+1}, \ldots, x_n - x_{n+1}) - f(x_1, \ldots, x_n)\big).$$

Step (1). Since $f \in \mathsf{DSh}_n(d)$ by the definition of h we have

$$\begin{aligned} h|T_0 &= \frac{1}{x_1}\big(f(x_{n+1} - x_1, \ldots, x_2 - x_1) - f(x_{n+1}, \ldots, x_2)\big) \\ &\overset{(3.32)}{=\!=} \frac{(-1)^{n-1}}{x_1}\big(f(x_2 - x_1, \ldots, x_{n+1} - x_1) - f(x_2, \ldots, x_{n+1})\big) \\ &= (-1)^{n-1} h(x_2, \ldots, x_{n+1}, x_1) \qquad \text{(by defn. of } h\text{)} \\ &= (-1)^{n-1} h(x_1, x_2, \ldots, x_{n+1}) \qquad \text{(by cyclic invariance).} \end{aligned}$$

Step (2). By assumption, h is invariant under the action of

$$\gamma_0 := \iota_{n+1}(C_{n+1}) = \begin{pmatrix} & & 1 \\ & I_n & \\ & & \end{pmatrix} \in \mathrm{GL}_{n+1}(\mathbb{Z}).$$

Step (3). By setting

$$\gamma_1 := \begin{pmatrix} -1 & -\mathbf{1} \\ \mathbf{0}^t & I_n \end{pmatrix} \in \mathrm{GL}_{n+1}(\mathbb{Z})$$

we get

$$\begin{aligned} h|\gamma_1 &= h\big((x_1,\dots,x_{n+1})\gamma_1^{-1}\big) = h\big((x_1,\dots,x_{n+1})\gamma_1\big) \\ &= h(-x_1, x_2-x_1,\dots,x_{n+1}-x_1) \\ &= h(x_2-x_1,\dots,x_{n+1}-x_1,-x_1) && \text{(by cyclic invariance)} \\ &= \frac{f(x_2-x_1,\dots,x_{n+1}-x_1)-f(x_2,\dots,x_{n+1})}{x_1} && \text{(by defn. of } h\text{)} \\ &= h(x_2,\dots,x_{n+1},x_1) && \text{(by defn. of } h\text{)} \\ &= h(x_1,x_2,\dots,x_{n+1}) && \text{(by cyclic invariance).} \end{aligned}$$

Step (4). Recall that the image under ι_{n+1} of the permutation

$$T'_{n+1} = \begin{pmatrix} 1 & 2 & \cdots & n & n+1 & n+2 \\ n & n-1 & \cdots & 1 & n+2 & n+1 \end{pmatrix} \in \mathfrak{S}_{n+2}$$

is equal to $\gamma_2 := \nu T_0 T_0^{\sharp} T_0 = T_0\gamma_1\gamma_0$. Hence

$$h|\gamma_2 = h|T_0\gamma_1\gamma_0 \overset{Step(1)}{=\!=} \pm h|\gamma_1\gamma_0 = \pm h \tag{6.25}$$

since h is invariant under both γ_0 and γ_1. So $h|\sigma = \pm h$ for all $\sigma \in \mathfrak{S}_{n+2}$ by the following fact (see Exercise 6.8):

$$\mathfrak{S}_{n+2} = \langle T_{n+1,0}, C_{n+1}, T'_{n+1}\rangle.$$

In particular, $h|\sigma = \pm h$ for all $\sigma \in \mathfrak{S}_{n+1}$ which means h is pseudo-symmetric. Finally, the proof is completed by applying Lemma 6.4.14. □

Proposition 6.4.16. *For any even integer $d \geq 0$ and $f \in \mathsf{DSh}_n(d)$, if*

$$\begin{aligned} &\frac{1}{x_1}\Big(f(x_2-x_1,\dots,x_{n+1}-x_1)-f(x_2,\dots,x_{n+1})\Big) \\ &\qquad = \frac{1}{x_{n+1}}\Big(f(x_1-x_{n+1},\dots,x_n-x_{n+1})-f(x_1,\dots,x_n)\Big) \end{aligned} \tag{6.26}$$

then $f = 0$.

Proof. We may assume $d \geq 2$ without loss of generality. Notice that the condition Eq. (6.26) eaily implies that $\frac{1}{x_{n+1}}\big(f(x_1 - x_{n+1}, \dots, x_n - x_{n+1}) - f(x_1, \dots, x_n)\big)$ is invariant under the cyclic permutations so Lemma 6.4.15 yields that f is DTI. By taking $t = -x_n$ we get

$$f(-x_n, x_1 - x_n, \dots, x_{n-1} - x_n)) = f(0, x_1, \dots, x_{n-1}). \tag{6.27}$$

On the other hands, the matrix $Q_n = \nu T_0^\sharp T_0$ acts on f by $(-1)^d = \mathrm{id}$. Hence we have

$$\begin{aligned} f(x_1, \dots, x_n) &= f|Q(x_1, \dots, x_n) \\ &= f(-x_n, x_1 - x_n, \dots, x_{n-1} - x_n) = f(0, x_1, \dots, x_{n-1}) \end{aligned}$$

by Eq. (6.27). Repeatedly applying this we get

$$f(x_1, \dots, x_n) = f(0, 0, x_1, \dots, x_{n-2}) = \cdots = f(0, \dots, 0).$$

But f has degree $d \geq 2$ so we must have $f = 0$, as desired. □

6.4.5 *Symmetrized MZVs generate whole MZV space*

For any two integers $w, n \geq 0$ we set

$$\mathsf{MZ}'_{w,n} = \mathsf{MZ}_{w,\leq n-1} + \mathsf{MZ}_{w,+,\leq n}.$$

Let $\mathsf{MZ}''_{w,n}$ denote the $\mathbb{Q}$-linear subspace of $\mathsf{MZ}_{w,\leq n}$ generated by the set

$$\left\{ \text{RHS of Eq. (6.24) for}(s_1, \dots, s_{n+1}) \;\middle|\; \begin{array}{l} s_1, \dots, s_{n+1} \geq 1, \\ s_1 + \cdots + s_{n+1} = w \end{array} \right\}.$$

Proposition 6.4.17. *For any two integers $w, n \geq 0$ we have*

$$\mathsf{MZ}_{w,\leq n} = \mathsf{MZ}'_{w,n} + \mathsf{MZ}''_{w,n}. \tag{6.28}$$

Proof. Not that if w and n have different parity, then the proposition follows from the Parity Principle (see Cor. 3.5.7).

We now assume $d = w - n$ is even. By simple algebra argument we see that it suffices to show the restriction map $\mathrm{Hom}_{\mathbb{Q}}(\mathsf{MZ}_{w,\leq n}, \mathbb{Q}) \to \mathrm{Hom}_{\mathbb{Q}}(\mathsf{MZ}'_{w,n} + \mathsf{MZ}''_{w,n}, \mathbb{Q})$ is injective. Suppose $\alpha : \mathsf{MZ}_{k,\leq n} \to \mathbb{Q}$ is a $\mathbb{Q}$-linear map vanishing on $\mathsf{MZ}'_{w,n} + \mathsf{MZ}''_{w,n}$. To prove $\alpha = 0$ it is enough to show that

$$f := \sum_{\substack{s_1, \dots, s_n \geq 1 \\ s_1 + \cdots + s_n = w}} \alpha(\zeta_*(s_1, \dots, s_n)) x_1^{s_1 - 1} \cdots x_n^{s_n - 1}$$

vanishes. Note that d is the degree of f. Since α vanishes on $\mathsf{MZ}_{w,+}$ we see that $f \in \mathsf{DSh}_n(d)$. Finally, α is 0 on $\mathsf{MZ}''_{w,n}$ so that Eq. (6.26) holds by Eq. (6.23) and the proof of Prop. 6.4.7 since now $n+1$ and w have different parities. Hence $\alpha = 0$ by Prop. 6.4.16. We have finished the proof of the proposition. □

For any compositions of positive integers $\mathbf{s}$ and $\mathbf{t}$ let $\mathrm{St}(\mathbf{s},\mathbf{t})$ be the set of the stuffle products of $\mathbf{s}$ and $\mathbf{t}$ with multiplicity. For example,

$$\mathrm{St}((3),(3,2)) = \{(3,3,2),(3,3,2),(3,2,3),(6,2),(3,5)\}.$$

Lemma 6.4.18. *Let* $\mathbf{s} \in \mathbb{N}^m$ *and* $\mathbf{t} \in (\mathbb{Z}_{\geq 2})^n$. *Then we have*

$$\zeta^{\mathcal{S}}_{ш}(\mathbf{s})\zeta^{\mathcal{S}}(\mathbf{t}) = \sum_{\mathbf{r}\in \mathrm{St}(\mathbf{s},\mathbf{t})} \zeta^{\mathcal{S}}_{ш}(\mathbf{r}).$$

Proof. For $\mathbf{s} = (s_1,\dots,s_m)$ we set

$$I(\mathbf{s}) := \{((s_i,\dots,s_1),(s_{i+1},\dots,s_m)) \mid 0 \leq i \leq m\}$$

and similarly for $\mathbf{t}$. Then by the definition

$$\zeta^{\mathcal{S}}_{|||}(\mathbf{s})\zeta^{\mathcal{S}}(\mathbf{t}) = \sum_{(\mathbf{s}_1,\mathbf{s}_2)\in I(\mathbf{s})} \sum_{(\mathbf{t}_1,\mathbf{t}_2)\in I(\mathbf{t})} (-1)^{|\mathbf{s}_1|+|\mathbf{t}_1|}\zeta_{ш}(\mathbf{s}_1)\zeta(\mathbf{t}_1)\cdot\zeta_{ш}(\mathbf{s}_2)\zeta(\mathbf{t}_2).$$

It follows from Cor 3.3.23 that

$$\zeta_{ш}(\mathbf{s}_j)\zeta(\mathbf{t}_j) = \sum_{\mathbf{r}_j\in \mathrm{St}(\mathbf{s}_j,\mathbf{t}_j)} \zeta_{ш}(\mathbf{r}_j)$$

for $j = 1, 2$. So the lemma follows easily from the identity

$$\coprod_{\mathbf{r}\in\mathrm{St}(\mathbf{s},\mathbf{t})} I(\mathbf{r}) = \coprod_{(\mathbf{s}_1,\mathbf{s}_2)\in I(\mathbf{s})} \coprod_{(\mathbf{t}_1,\mathbf{t}_2)\in I(\mathbf{t})} \mathrm{St}(\mathbf{s}_1,\mathbf{t}_1)\times\mathrm{St}(\mathbf{s}_2,\mathbf{t}_2)$$

where both sides are regarded as multisets of pairs compositions. □

Recall that MZ_w, $\mathsf{MZ}^{\mathcal{S}}_{*,w}$ and $\mathsf{MZ}^{\mathcal{S}}_{ш,w}$ are the $\mathbb{Q}$-vector space generated by all weight w MZVs, $*$-SMZVs and ш-SMZVs, respectively.

Theorem 6.4.19. *For all positive interger* $w \geq 2$, *we have*

$$\mathsf{MZ}_w = \mathsf{MZ}^{\mathcal{S}}_{*,w} = \mathsf{MZ}^{\mathcal{S}}_{ш,w}.$$

Proof. First, since $\mathsf{MZ}^{\mathcal{S}}_{*,w} \subseteq \mathsf{MZ}_w$ by Prop. 6.1.3 we can prove $\mathsf{MZ}_w = \mathsf{MZ}^{\mathcal{S}}_{*,w}$ by showing the claim

$$\mathsf{MZ}_{w,\leq n} \subseteq \mathsf{MZ}^{\mathcal{S}}_{*,w} \qquad \text{for all} n \geq 0. \tag{6.29}$$

We now proceed by induction on n. When $n = 0$ this is trivial for all w. Assume $n \geq 1$ and the inclusion holds for all w. If $w-n$ is even then w and

$n+1$ have different parities and Eq. (6.29) follows from the Parity Principle and the induction. If $w-n$ is odd then by Prop. 6.4.17 we have

$$\begin{aligned}\mathsf{MZ}_{w,\le n+1} &= \mathsf{MZ}'_{w,n+1} + \mathsf{MZ}''_{w,n+1}\\ &\subseteq \mathsf{MZ}_{w,\le n} + \mathsf{MZ}_{w,+,\le n+1} \qquad \text{(by Prop. 6.4.7)}\\ &\subseteq \mathsf{MZ}^{\mathcal{S}}_{*,w} + \mathsf{MZ}^{\mathcal{S}}_{*,w,+} \qquad \text{(by induction)}\\ &\subseteq \mathsf{MZ}^{\mathcal{S}}_{*,w} + \sum_{w',w''\ge 1,\ w'+w''=w} \mathsf{MZ}^{\mathcal{S}}_{*,w'}\cdot \mathsf{MZ}^{\mathcal{S}}_{*,w''} \subseteq \mathsf{MZ}^{\mathcal{S}}_{*,w}\end{aligned}$$

by Cor. 6.1.5.

Turning to the ш-version, by Thm. 6.3.2 and Prop. 6.1.3, we see that

$$\mathsf{MZ}^{\mathcal{S}}_{*,w} \subseteq \mathsf{MZ}^{\mathcal{S}}_{ш,w} + \zeta(2)\mathsf{MZ}^{\mathcal{S}}_{ш,w-2} \subseteq \mathsf{MZ}_w = \mathsf{MZ}^{\mathcal{S}}_{*,w}$$

by what we have just proved. Thus, to prove $\mathsf{MZ}^{\mathcal{S}}_{*,w} = \mathsf{MZ}^{\mathcal{S}}_{ш,w}$ it suffices to show $\zeta(2)\mathsf{MZ}^{\mathcal{S}}_{ш,w-2} \subset \mathsf{MZ}^{\mathcal{S}}_{ш,w}$ for all $w \ge 2$, which follows immediately from Lemma 6.4.18. The theorem is now proved. □

6.5 Historical notes

Under a suggestion of Kontsevich the SMZVs were first introduced by D. Zagier around 2013 to establish a crucial bridge between the ordinary MZVs and their finite version which will be studied in Chap. 8. Moreover, the SMZVs are nicer versions of MZVs in the sense that they are defined for all positive integers and satisfy stuffle relations and shuffle relations. Furthermore, their shuffle relations are simpler to state than those for MZVs.

Section 6.3 follows Jarossay's notes [306] closely. We also present some results such as Thm. 6.2.2 and Thm. 6.2.4, proved by Murahara [420], that are very similar to the ones for the finite MZVs in Chap. 8. The guiding principle here is a mysterious relation between SMZVs and the finite MZVs which we will see in Conjecture 8.5.15 of Chap. 8.

Thm. 6.4.19 is proved by Yasuda [564]. We will generalize this result to Euler sums in Chap. 14 using essentially the same machinary.

Exercises

6.1. Show that

$$\zeta^{\mathcal{S}}_{ш}(2,1) = 3\zeta(2,1)$$

by using the fact that $b ш ab = bab + 2abb$.

6.2. Complete the proof of Prop. 6.1.3 by showing that $\zeta^{\mathcal{S}}_*$ are all finite real numbers.

6.3. Let the map inv be defined by Eq. (6.6). Show that if f_* is group-like under Δ_* then so is f_*^{inv}. Namely,

$$\Delta_*(f_*) = f_* \otimes f_* \Longrightarrow \Delta_*(f_*^{\mathrm{inv}}) = f_*^{\mathrm{inv}} \otimes f_*^{\mathrm{inv}}.$$

6.4. Prove Eq. (6.15) by mimicking the proof of Eq. (6.14) and using an argument similar to the proof of Eq. (6.3).

6.5. Show that under the embedding $\mathbb{Z}^n \hookrightarrow \mathbb{Z}^{(n+1)}$, $(a_1, \ldots, a_n) \to (a_1, \ldots, a_n, -(a_1 + \cdots + a_n))$, the image $\iota_n(T')$ is given by the matrix γ_2 as defined by Eq. 6.16.

6.6. For any $n \in \mathbb{N}$, define

$$\angle_{n,\mathbb{R}} = \left\{ (t_1, \ldots, t_n) \in (\mathbb{R}^\times)^n \;\middle|\; \frac{1}{t_1} > \cdots > \frac{1}{t_n} \right\}.$$

For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ let $B(\mathbf{x}, \varepsilon)$ be the open ball in $\mathbb{R}^n$ of radius ε centered at $\mathbf{x}$. Set

$$v(\mathbf{x}) = \lim_{\varepsilon \to 0} \frac{\mathrm{vol}(B(\mathbf{x}, \varepsilon) \cap \angle_{n,\mathbb{R}})}{\mathrm{vol}(B(\mathbf{x}, \varepsilon))}.$$

Show that for any $\mathbf{s} \in \mathbb{N}^n$,

$$\zeta^{\natural,\mathcal{S}}(\mathbf{s}) = (-1)^{|\mathbf{s}|} \lim_{M \to \infty} \sum_{\substack{\frac{1}{k_1} > \cdots > \frac{1}{k_d} \\ 0 \neq |k_j| \in \mathbb{Z}_{\leq M} \ \forall j = 1, \ldots, d}} \frac{v\big((k_1, \ldots, k_d)\big)}{k_1^{s_1} \cdots k_d^{s_d}}.$$

6.7. Prove Eq. (6.20) by applying

$$\mathrm{sh}_1^{(n)} = \sum_{i=1}^{n} \begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & n \\ i & 1 & \cdots & i-1 & i+1 & \cdots & n \end{pmatrix},$$

to both sides of Eq. (6.19).

6.8. Show that the symmetry group $\mathfrak{S}_{n+1}$ is generated by the following three elements: $C_n, T_{n,0} \in \mathfrak{S}_n$ and T'_n defined by Eq. (3.33).

Chapter 7

Multiple harmonic sums and the alternating version

A multiple harmonic sum (MHS) is a finite partial sums of the infinite series defined by Eq. (1.8) defining the multiple zeta values $\zeta(s_1, \ldots, s_d)$ for all positive integer arguments $s_1, \ldots, s_d$. In particular, $s_1 = 1$ is allowed. An alternating MHS (AMHS) is a finite partial sum of an infinite Euler sum, an alternating version of a MZV. All of the partial sums are rational numbers so it makes perfect sense to consider congruence properties modulo sufficiently large primes. This will be our first goal in this chapter. The second goal is to derive some identities using binomial coefficients which will be useful to evaluate MZVs and similar objects when the number of terms of the partial sums goes to infinity.

7.1 Definitions

We start by defining a sort of double cover of the set $\mathbb{N}_0$ of nonnegative integers.

Definition 7.1.1. Let $\mathbb{D}$ be the set of *signed numbers* $\mathbb{N}_0 \cup \overline{\mathbb{N}_0}$ where

$$\overline{\mathbb{N}_0} = \{\bar{k} : k \in \mathbb{N}_0\}.$$

Define the absolute value function $|\cdot|$ on $\mathbb{D}$ by $|k| = |\bar{k}| = k$ for all $k \in \mathbb{N}_0$ and the sign function by $\mathrm{sgn}(k) = 1$ and $\mathrm{sgn}(\bar{k}) = -1$ for all $k \in \mathbb{N}_0$. Numbers in $\overline{\mathbb{N}_0}$ are called *signed numbers* and numbers in $\mathbb{N}_0$ are called *unsigned numbers*. We make $\mathbb{D}$ a semi-group by defining a commutative and associative binary operation $\oplus$ (called *O-plus*) as follows: for all $a, b \in \mathbb{D}$

$$a \oplus b = \begin{cases} \overline{|a| + |b|}, & \text{if } \mathrm{sgn}(a) \neq \mathrm{sgn}(b); \\ |a| + |b|, & \text{if } \mathrm{sgn}(a) = \mathrm{sgn}(b). \end{cases} \tag{7.1}$$

Definition 7.1.2. For any $d \in \mathbb{N}$ and $\mathbf{s} = (s_1, s_2, \dots, s_d) \in \mathbb{D}^d$ we define the *alternating multiple harmonic sums* by

$$H_n(s_1, s_2, \dots, s_d) = \sum_{n \geq k_1 > k_2 > \dots > k_d \geq 1} \prod_{j=1}^{d} \frac{\operatorname{sgn}(s_j)^{k_j}}{k_j^{|s_j|}}, \tag{7.2}$$

$$H_n^\star(s_1, s_2, \dots, s_d) = \sum_{n \geq k_1 \geq k_2 \geq \dots \geq k_d \geq 1} \prod_{j=1}^{d} \frac{\operatorname{sgn}(s_j)^{k_j}}{k_j^{|s_j|}}. \tag{7.3}$$

When $\mathbf{s} \in \mathbb{N}^d$ these are called *multiple harmonic sums* (MHSs) and *multiple harmonic star sums* (MH*Ss), respectively. Conventionally, we call $\operatorname{dp}(\mathbf{s}) := d$ the depth and $|\mathbf{s}| := \sum_{j=1}^{d} |s_j|$ the weight. For convenience we set $H_n(\mathbf{s}) = 0$ if $n < \operatorname{dp}(\mathbf{s})$, $H_n(\emptyset) = H_n^\star(\emptyset) = 1$ for all $n \geq 0$.

For instance, the harmonic numbers and the alternating harmonic numbers have weight and depth both equal to 1:

$$H_n(1) = \sum_{k=1}^{n} \frac{1}{k}, \qquad H_n(\bar{1}) = \sum_{k=1}^{n} \frac{(-1)^k}{k}.$$

7.2 Stuffle relations

Throughout this section we let n be a fixed positive integer and only consider AMHSs of the form H_n or $H_n^\star$.

Recall the Hoffman algebra $\mathfrak{A}^1 = \mathbb{Q}\mathsf{e} \oplus \mathfrak{A}b$ given in Definition 3.2.2 is a noncommutative polynomial algebra on generators $\mathtt{y}_k := a^{k-1}b$ for all $k \geq 1$. We further equip it with the stuffle product $*$ and regard it as a subalgebra of $\mathfrak{A}_*$ defined in Definition 3.2.8. To study both MHSs and AMHSs we need to generalize this algebra as sketched below. We omit the details because we will generalize this even further to arbitrary level N (by replacing ± 1 by Nth roots of unity) in Chap. 13 (see Definition 13.3.2 there).

Definition 7.2.1. For $N = \in \mathbb{N}$ we define Hoffman-Racinet algebra of level N, denoted by $\mathfrak{A}_N$, to be the (weight) graded noncommutative polynomial $\mathbb{Q}$-algebra generated by $N+1$ letters $\mathtt{a}, \mathtt{b}_0, \dots, \mathtt{b}_{N-1}$. The number of letters in a word $\mathbf{w}$ is called its *weight* and the number of $\mathtt{b}_0, \dots, \mathtt{b}_{N-1}$, its *depth*. For every positive integer k define

$$\mathtt{z}_k := \mathtt{a}^{k-1}\mathtt{b}_0 \quad \text{and, if } N = 2, \quad \mathtt{z}_{\bar{k}} := \mathtt{a}^{k-1}\mathtt{b}_1. \tag{7.4}$$

Denote by $\mathfrak{A}_N^1$ the subalgebra of $\mathfrak{A}_N$ generated by words not ending with a.

Note that when $N = 2$, $\mathfrak{A}_2^1$ is the subalgebra of $\mathfrak{A}_2$ generated by words $\mathtt{z}_k$ with $k \in \mathbb{D}$.

Remark 7.2.2. For future reference when we deal with Euler sums we will need to consider the subalgebra $\mathfrak{A}_2^0$ of $\mathfrak{A}_2$ generated by words beginning with $\mathtt{a}$ or $\mathtt{b}_1$ and ending with $\mathtt{b}_0$ or $\mathtt{b}_1$. The words in $\mathfrak{A}_2^0$ are called *admissible words.*

For computational purposes we now describe a way to express every AMHS $H_n(s_1, \ldots, s_d)$ using a unique word $\mathbf{w} \in \mathfrak{A}_2^1$ (and write $H_n(\mathbf{w}) = H_n(s_1, \ldots, s_d)$) and vice versa:

Definition 7.2.3. Converting rule from words in $\mathfrak{A}_2^1$ to signed indices. Define

$$\mathrm{Tog} : \mathfrak{A}_2^1 \longrightarrow \mathbb{D}^d$$

by the following procedure. Given $\mathtt{z}_{s_1} \ldots \mathtt{z}_{s_d} \in \mathfrak{A}_2^1$, first we set $t_i = |s_i|$ for all $i = 1, \ldots, d$ and write $\mathbf{t} = (t_1, \ldots, t_d)$. Then go down from $\mathtt{z}_{s_1}$ to $\mathtt{z}_{s_d}$. As soon as we see the first signed number subscript, say the ith one is $\mathtt{z}_{\overline{t_i}}$, we change t_i in $\mathbf{t}$ to $\overline{t_i}$. Then keep moving on towards $\mathtt{z}_{s_d}$ until we see the first **unsigned** number, say $\mathtt{z}_{t_k}$. We then change t_k in $\mathbf{t}$ to $\overline{t_k}$. Then we move on and look for the next **signed** number and change the corresponding index in $\mathbf{t}$. Carry on this alternating search pattern till the end.

It is not hard to see that the number of the signed indices in the final resulting $\mathbf{t}$ should be equal to the number of sign changes in $\mathtt{z}_{s_1} \ldots \mathtt{z}_{s_d}$, assuming the first sign before $\mathtt{z}_{s_1}$ is positive. For example:

$$H_n(\mathtt{b}_1\mathtt{a}\mathtt{b}_1\mathtt{a}\mathtt{b}_1\mathtt{a}^3\mathtt{b}_0\mathtt{a}^2\mathtt{b}_0\mathtt{a}^4\mathtt{b}_1\mathtt{a}^5\mathtt{b}_0) = H_n(\mathtt{z}_{\bar{1}}\mathtt{z}_{\bar{2}}\mathtt{z}_{\bar{2}}\mathtt{z}_4\mathtt{z}_3\mathtt{z}_{\bar{5}}\mathtt{z}_6) = H_n(\bar{1}, 2, 2, \bar{4}, 3, \bar{5}, \bar{6}). \tag{7.5}$$

Definition 7.2.4. Converting rule from signed indices to words in $\mathfrak{A}_2^1$. Define

$$\mathrm{Tog}^{-1} : \mathbb{D}^d \longrightarrow \mathfrak{A}_2^1$$

by the following procedure. Given $(t_1, \ldots, t_d) \in \mathbb{D}^d$ we set $s_i = |t_i|$ for all $i = 1, \ldots, d$ and first write $w = \mathtt{z}_{s_1} \ldots \mathtt{z}_{s_d}$. Then we put bars on top of some of the s_j's according to the following rule. Going down from t_1 to t_d, as soon as we see the first signed number, say the ith component is $\overline{s_i}$, we change in w every s_j'safter $\mathtt{z}_{s_i}$ (inclusive) to $\mathtt{z}_{\overline{s_j}}$ until the next signed number $\bar{s}_k$ is encountered. We then leave alone all the s_j's in w after $\mathtt{z}_{s_k}$ (again inclusive) until we see the next signed number when we start to toggle again. Carry on this toggling till the end of w.

Imaginatively we can think the bars as switches between $\mathtt{z}_s$'s and $\mathtt{z}_{\bar{s}}$'s. So we get a reversal of Eq. (7.5):

$$\mathrm{Tog}^{-1}(\bar{1},2,2,\bar{4},3,\bar{5},\bar{6}) = \mathtt{z}_{\bar{1}}\mathtt{z}_{\bar{2}}\mathtt{z}_{\bar{2}}\mathtt{z}_4\mathtt{z}_3\mathtt{z}_{\bar{5}}\mathtt{z}_6 = \mathtt{b}_1\mathtt{a}\mathtt{b}_1\mathtt{a}\mathtt{b}_1\mathtt{a}^3\mathtt{b}_0\mathtt{a}^2\mathtt{b}_0\mathtt{a}^4\mathtt{b}_1\mathtt{a}^5\mathtt{b}_0.$$

Remark 7.2.5. The two converting rules above are inverses to each other (see Exercise 7.3). It is a special case of the more general correspondence for arbitrary levels in Sec. 13.3.1.

To define the stuffle relations among AMHSs, a very simple example is already illuminating:

$$H_n(\bar{1})H_n(2) = H_n(\bar{1},2) + H_n(2,\bar{1}) + H_n(\bar{3}).$$

This follows immediately from the expansion

$$\sum_{n\geq j>0}\ \sum_{n\geq k>0} = \sum_{n\geq j>k>0} + \sum_{n\geq k>j>0} + \sum_{n\geq j=k>0}.$$

Definition 7.2.6. For positive integer k define the *toggling operator* $\tau_{\mathtt{z}_k}$ as follows: $\forall \mathbf{w} \in \mathfrak{A}_2^1$ we set $\tau_{\mathtt{z}_k}(\mathbf{w}) = \mathbf{w}$, and define $\tau_{\mathtt{z}_{\bar{k}}}(\mathbf{w})$ to be the word produced from $\mathbf{w}$ by exchanging $\mathtt{z}_l$ and $\mathtt{z}_{\bar{l}}$ for all l. We define a multiplication $*$ on $\mathfrak{A}_2^1$ by requiring that $*$ distribute over addition, that $\mathtt{e} * \mathbf{w} = \mathbf{w} * \mathtt{e} = \mathbf{w}$ for each word $\mathbf{w} \in \mathfrak{A}_2^1$, and that, for all words $\mathbf{w}_1, \mathbf{w}_2$ and special words x and y equaling either $\mathtt{z}_k$ or $\mathtt{z}_{\bar{k}}$ for some k,

$$\begin{aligned} x\mathbf{w}_1 * y\mathbf{w_2} = \mathbf{x}\Big(\tau_{\mathbf{x}}\big(\tau_{\mathbf{x}}(\mathbf{w_1}) * \mathbf{y}\mathbf{w_2}\big)\Big) + \mathbf{y}\Big(\tau_{\mathbf{y}}\big(\mathbf{x}\mathbf{w_1} * \tau_{\mathbf{y}}(\mathbf{w_2})\big)\Big) \\ + [x,y]\Big(\tau_{[x,y]}\big(\tau_x(\mathbf{w}_1) * \tau_y(\mathbf{w}_2)\big)\Big) \end{aligned} \quad (7.6)$$

where

$$[\mathtt{z}_k, \mathtt{z}_l] = [\mathtt{z}_{\bar{k}}, \mathtt{z}_{\bar{l}}] = \mathtt{z}_{k+l}, \quad [\mathtt{z}_{\bar{k}}, \mathtt{z}_l] = [\mathtt{z}_k, \mathtt{z}_{\bar{l}}] = \mathtt{z}_{\overline{k+l}}.$$

Using Eq. (7.6) we can further define a multiplication $*^\star$ on $\mathfrak{A}_2^1$: simply replace $*$ by $*^\star$ and substitute $-[x,y]$ for $+[x,y]$ in Eq. (7.6).

The proof of the following stuffle relations is similar to that of MZV using induction on depth and is thus omitted here.

Theorem 7.2.7. *For any word* $\mathbf{w} \in \mathfrak{A}_2^1$ *let* $\mathrm{Tog}(\mathbf{w}) = (t_1, \ldots, t_d)$, $H_n(\mathbf{w}) = H_n(t_1, \ldots, t_d)$ *and* $H_n^\star(\mathbf{w}) = H_n^\star(t_1, \ldots, t_d)$, *and then linearly extend both to* $\mathfrak{A}_2^1$ *(cf. Definition* 3.2.2*). Then for any words* $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_2^1$ *we have*

$$H_n(\mathbf{u})H_n(\mathbf{v}) = H_n(\mathbf{u} * \mathbf{v}), \qquad H_n^\star(\mathbf{u})H_n^\star(\mathbf{v}) = H_n^\star(\mathbf{u} *^\star \mathbf{v}). \quad (7.7)$$

7.3 Quasi-symmetric functions with signed powers

To derive more relations between MHSs, MH*Ss and AMHSs we turn to the theory of quasi-symmetric functions. To treat AMHSs with signed indices in $\mathbb{D}$ we have to allow the powers in these quasi-symmetric functions to be signed numbers.

Definition 7.3.1. We denote by $\mathbb{Z}[x_1,\dots,x_n;\mathbb{D}]$ the set of polynomials in $x_1,\dots,x_n$ with signed powers, namely,

$$\mathbb{Z}[x_1,\dots,x_n;\mathbb{D}] = \left\{ \sum_{e_1=\overline{d_1}}^{d_1} \cdots \sum_{e_n=\overline{d_n}}^{d_n} c_{e_1,\dots,e_n} x_1^{e_1}\cdots x_n^{e_n} \,\middle|\, d_1,\dots,d_n \in \mathbb{N}_0, c_{e_1,\dots,e_n} \in \mathbb{Z} \right\}.$$

Here, $\overline{0}=0$ and $\sum_{e=\overline{d}}^{d}$ means e runs through the set $\{\overline{d},\dots,\overline{1},0,1,\dots,d\}$. Furthermore, $x_j^e x_j^{e'} = x_i^{e\oplus e'}$ for any $j \le n$ and $e,e' \in \mathbb{D}$. Also we set

$$\deg(x_1^{e_1}\cdots x_n^{e_n}) = |e_1| + \cdots + |e_n|.$$

For positive integer n we define a graded algebra homomorphism

$$\begin{aligned} \phi_n : \quad \mathfrak{A}_2^1 \quad &\longrightarrow \quad \mathbb{Z}[x_1,\dots,x_n;\mathbb{D}] \\ \mathsf{z}_{s_1}\cdots\mathsf{z}_{s_d} &\longmapsto \sum_{n\ge k_1>k_2>\cdots>k_d\ge 1} x_{k_1}^{t_1}\cdots x_{k_d}^{t_d}, \quad \forall d \le n, \end{aligned}$$

where $(t_1,\cdots,t_d) = \mathrm{Tog}(\mathsf{z}_{s_1}\cdots\mathsf{z}_{s_d})$ by the converting rule Definition 7.2.3, and set $\phi_n(\mathsf{e}) = 1$ and $\phi_n(\mathbf{w}) = 0$ if the length of $\mathbf{w}$ is greater than n. Here, we set the degree of all x_i's equal to 1. It is easy to see we can make $(\phi_n)_{n\ge1}$ into a compatible system to obtain a homomorphism ϕ from Hoffman-Racinet algebra $\mathfrak{A}_2^1$ to the algebra QSym_2 of quasi-symmetric functions which are defined below.

Definition 7.3.2. Let $\mathbf{x} = (x_j)_{j\ge1}$. An element of finite degree $F(\mathbf{x})$ in $\mathbb{Z}[\![\mathbf{x};\mathbb{D}]\!]$ is called a *quasi-symmetric function* if for any $i_1 > i_2 > \cdots > i_d$ and $j_1 > j_2 > \cdots > j_d$ and any signed powers $e_1,\dots,e_d \in \mathbb{D}$ the coefficients of the monomials $x_{i_1}^{e_1}\cdots x_{i_d}^{e_d}$ and $x_{j_1}^{e_1}\cdots x_{j_d}^{e_d}$ are the same. The set of all such quasi-symmetric functions is denoted by QSym_2.

For example, $\sum_{i>j\ge1} x_i x_j^{\bar2} \in \mathsf{QSym}_2$ but is *not* a symmetric function since the monomial $x_2x_1^{\bar2}$ appears but $x_1x_2^{\bar2}$ does not. We may choose an

integral basis for QSym_2 using

$$E_{\mathbf{t}} = E_{t_1,\dots,t_d} := \sum_{k_1\geq k_2\geq\cdots\geq k_d} x_{k_1}^{t_1}\cdots x_{k_d}^{t_d}, \quad \text{or}$$

$$M_{\mathbf{t}} = M_{t_1,\dots,t_d} := \sum_{k_1> k_2>\cdots> k_d} x_{k_1}^{t_1}\cdots x_{k_d}^{t_d} = \phi(\mathtt{z}_{s_1}\cdots\mathtt{z}_{s_d}), \tag{7.8}$$

where $(\mathtt{z}_{s_1}\cdots\mathtt{z}_{s_d}) = \mathrm{Tog}^{-1}(t_1,\dots,t_d)$ (see Exercise 7.5). This yields the following theorem which can be compared to [279, Thm. 2.2].

Theorem 7.3.3. *ϕ provides an isomorphism $(\mathfrak{A}_2^1, *) \cong \mathsf{QSym}_2$.*

We will see that the MHSs satisfy a different kind of duality from that of MZVs. To derive it, we first define the power set of $\mathbf{s}$ to be the partial sum sequence: $P(\mathbf{s}) = (s_1, s_1+s_2, \dots, s_1+\cdots+s_{d-1}, w)$ as an (increasingly ordered) d-subset (*i.e.*, subset of d elements) of $(1,2,\dots,w)$. We see that P provides a one-to-one correspondence between the compositions of weight w and length d and the d-subsets of $(1,2,\dots,w)$. Let

$$\mathfrak{R}_w(a_1,\dots,a_\ell) = (w+1-a_\ell,\dots,w+1-a_1).$$

The v-dual of $\mathbf{s}$ is the composition of weight w corresponding to the complimentary subset of $\mathfrak{R}_w P(\mathbf{s})$ in $(1,2,\cdots,w)$. Namely,

$$\mathbf{s}^\vee = P^{-1}\big((1,2,\cdots,w)\setminus \mathfrak{R}_w P(\mathbf{s})\big). \tag{7.9}$$

It's easy to see that $(\mathbf{s}^\vee)^\vee = \mathbf{s}$ so the v-dual is an involution. Another equivalent definition is the following (see Excercise 7.1).

Definition 7.3.4. For positive integers $r_1,\dots,r_\ell,t_1,\dots,t_\ell$ let

$$\mathbf{s} = (r_1,\{1\}^{t_1-1},r_2+1,\{1\}^{t_2-1},\dots,r_\ell+1,\{1\}^{t_\ell-1}).$$

We define the v-dual of $\mathbf{s}$ by

$$\mathbf{s}^\vee = (\{1\}^{r_1-1},t_1+1,\{1\}^{r_2-1},t_2+1,\dots,t_{\ell-1}+1,\{1\}^{r_\ell-1},t_\ell). \tag{7.10}$$

From the definition we clearly have

$$\mathrm{dp}(\mathbf{s}) + \mathrm{dp}(\mathbf{s}^\vee) = |\mathbf{s}| + 1. \tag{7.11}$$

Remark 7.3.5. The v-dual can be easily explained using the conjugation on the ribbons (a kind of skew-Young diagrams). See Exercise 7.2.

Recall that the φ-dual is defined on $\mathfrak{A}_1^1$ by Eq. (6.11): $\mathtt{x}_0^\varphi = \mathtt{x}_0+\mathtt{x}_1, \mathtt{x}_1^\varphi = -\mathtt{x}_1$. We transport it to QSym_1 by Eq. 7.8. Namely, we set

$$M_{\mathbf{s}}^\varphi := \phi(\mathtt{z}_{s_1}^\varphi\cdots\mathtt{z}_{s_d}^\varphi).$$

In the next theorem, the two kinds of dualities, v-dual and φ-dual, are related to each other.

Theorem 7.3.6. *Let* $\mathbf{s}$ *be a composition of positive integers. Then we have*

(i) $M^{\varphi}_{\mathbf{s}} = (-1)^{\mathrm{dp}(\mathbf{s})} \sum_{\mathbf{t} \succeq \mathbf{s}} M_{\mathbf{t}}$;
(ii) $E^{\varphi}_{\mathbf{s}} = -E_{\mathbf{s}^\vee}$.

Proof. (i) Suppose a word $\mathbf{w} \in \mathfrak{A}_1^1$ corresponds to the composition $\mathbf{s}$. Then a replacement of some particular $\mathbf{x}_0$ by $\mathbf{x}_1$ in $\mathbf{w}$ corresponds to a splitting of $\mathbf{s}$. Thus (i) follows.

(ii) Define $F_{\mathbf{s}} := \sum_{\mathbf{t} \succeq \mathbf{s}} M_{\mathbf{t}} = (-1)^{\mathrm{dp}(\mathbf{s})} M^{\varphi}_{\mathbf{s}}$ by (i). Then, by the definition

$$E^{\varphi}_{\mathbf{s}} = \sum_{\mathbf{t} \preceq \mathbf{s}} M^{\varphi}_{\mathbf{t}} = \sum_{\mathbf{t} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{t})} F_{\mathbf{t}}. \tag{7.12}$$

Dualizing [224, Cor. 3.16] we get $S(F_{\mathbf{t}}) = (-1)^{|\mathbf{t}|} F_{\overleftarrow{\mathbf{t}}^\vee}$ where S is the antipode of QSym_2 as defined in Thm. A.15. Thus, applying S to Eq. (7.12) we get

$$S(E^{\varphi}_{\mathbf{s}}) = \sum_{\mathbf{t} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{t})+|\mathbf{t}|} F_{\overleftarrow{\mathbf{t}}^\vee} = -\sum_{\mathbf{t} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{t}^\vee)} F_{\overleftarrow{\mathbf{t}}^\vee} = -\sum_{\mathbf{r} \succeq \overleftarrow{\mathbf{s}}^\vee} (-1)^{\mathrm{dp}(\mathbf{r})} F_{\mathbf{r}}$$

because of Eq. 7.11. By Möbius inversion formula we see that $S(E^{\varphi}_{\mathbf{s}}) = -(-1)^{\mathrm{dp}(\overleftarrow{\mathbf{s}}^\vee)} M_{\overleftarrow{\mathbf{s}}^\vee}$. Applying S again, we finally get

$$E^{\varphi}_{\mathbf{s}} = -(-1)^{\mathrm{dp}(\overleftarrow{\mathbf{s}}^\vee)} S(M_{\overleftarrow{\mathbf{s}}^\vee}) = -E_{\mathbf{s}^\vee}$$

by Thm. A.15(i). □

Corollary 7.3.7. *For any composition* $\mathbf{s}$ *of positive integers,*

$$-E_{\mathbf{s}^\vee} = \sum_{\mathbf{r} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{r})} \sum_{\mathbf{t} \succeq \mathbf{r}} M_{\mathbf{t}}. \tag{7.13}$$

Proof. By Thm. 7.3.6(ii) and the definition of $E_{\mathbf{s}}$ we get

$$-E_{\mathbf{s}^\vee} = \sum_{\mathbf{r} \preceq \mathbf{s}} M^{\varphi}_{\mathbf{r}}.$$

So Eq. (7.13) follows from Thm. 7.3.6(i). □

We now apply the above results concerning quasi-symmetric functions to AMHSs. In order to do so, for any $n \in \mathbb{N}$, we define the algebra homomorphism $\rho_n : (\mathfrak{A}_2^1, *) \to \mathbb{R}$ such that

$$\rho_n(M_{\mathbf{s}}) = H_n(\mathbf{s}). \tag{7.14}$$

Theorem 7.3.8. *For all* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{D}^d$ *and positive integer* n

$$H_n(\mathbf{s}) = \sum_{\mathbf{t} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{s}) - \mathrm{dp}(\mathbf{t})} H_n^\star(\mathbf{t}), \quad H_n^\star(\mathbf{s}) = \sum_{\mathbf{t} \preceq \mathbf{s}} H_n(\mathbf{t}), \tag{7.15}$$

$$H_n^\star(\overleftarrow{\mathbf{s}}) = (-1)^{\mathrm{dp}(\mathbf{s})} \sum_{\bigsqcup_{j=1}^{l} \mathbf{s}_j = \mathbf{s}} (-1)^r \prod_{j=1}^{r} H_n(\mathbf{s}_j), \tag{7.16}$$

$$H_n(\overleftarrow{\mathbf{s}}) = (-1)^{\mathrm{dp}(\mathbf{s})} \sum_{\bigsqcup_{j=1}^{l} \mathbf{s}_j = \mathbf{s}} (-1)^r \prod_{j=1}^{r} H_n^\star(\mathbf{s}_j). \tag{7.17}$$

Moreover, for all $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$

$$H_n^\star(\mathbf{s}^\vee) = -\sum_{\mathbf{r} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{r})} \sum_{\mathbf{t} \succeq \mathbf{r}} H_n(\mathbf{t}). \tag{7.18}$$

Proof. The four equations (7.15), (7.16), (7.17), and (7.18) follow from the definition of $E_{\mathbf{s}}$, Thm. A.15(ii) and (iii), and Cor. 7.3.7, respectively, after we apply ρ_n. □

Definition 7.3.9. We will call the relations in Eqs. (7.16) and (7.17) the *concatenation relations* between MHSs and MH*Ss.

7.4 Symmetric sums

For a partition $\Pi = \{P_1, \dots, P_l\}$ of the set $\{1, 2, \dots, d\}$, let

$$c(\Pi) = \prod_{i=1}^{l} (\sharp P_i - 1)! \quad \text{and} \quad \tilde{c}(\Pi) = (-1)^{d-l} \prod_{i=1}^{l} (\sharp P_i - 1)!,$$

Let $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{D}^d$ be a composition of length d. Set

$$H_n(\mathbf{s}; \Pi) = \prod_{i=1}^{l} H_n(\lambda_i), \quad H_n^\star(\mathbf{s}; \Pi) = \prod_{i=1}^{l} H_n^\star(\lambda_i),$$

where $\lambda_i = \oplus_{j \in P_i} s_j$. Further, let $\sigma \in \mathfrak{S}_d$ act on $\mathbf{s}$ by permuting the subscripts.

Theorem 7.4.1. *For any* $d, n \in \mathbb{N}$ *and any compositions* $\mathbf{s} \in \mathbb{D}^d$,

$$\sum_{\sigma \in \mathfrak{S}_d} H_n^\star\big(\sigma(\mathbf{s})\big) = \sum_{\Pi\text{: partitions of } \{1,2,\dots,d\}} \tilde{c}(\Pi) H_n(\mathbf{s}; \Pi) \tag{7.19}$$

$$\sum_{\sigma \in \mathfrak{S}_d} H_n\big(\sigma(\mathbf{s})\big) = \sum_{\Pi\text{: partitions of } \{1,2,\dots,d\}} c(\Pi) H_n(\mathbf{s}; \Pi). \tag{7.20}$$

Proof. We prove the theorem under the assumption that each of the components $s_1, \dots, s_d$ of $\mathbf{s}$ satisfies that $|s_j|$ is a positive real numbers greater than 1 (of course we should first extend $\mathbb{D}$ to a double cover of $\mathbb{R}$ in the obvious way). Then by taking limit we can prove the theorem easily. Further we may assume all $s_1, \dots, s_d$ are pairwise distinct since again we can take limit.

By the definition

$$\sum_{\sigma\in\mathfrak{S}_d} H_n\big(\sigma(\mathbf{s})\big) = \sum_{\sigma\in\mathfrak{S}_d} \sum_{k_1\geq\cdots\geq k_d} \frac{\operatorname{sgn}(s_1)^{k_{\sigma(1)}}\cdots\operatorname{sgn}(s_d)^{k_{\sigma(d)}}}{k_{\sigma(1)}^{|s_1|}\cdots k_{\sigma(d)}^{|s_d|}}.$$

Let $\mathfrak{S}_d$ be the symmetry group of order d. Let $\Xi(\mathbf{k})$ be the set of equivalence classes of $\{1,\dots,d\}$ with $i \overset{\mathbf{k}}{\sim} j$ if and only if $k_i = k_j$. Let $\mathfrak{S}_d(\mathbf{k}) = \{\sigma \in \mathfrak{S}_d | \sigma(i) \overset{\mathbf{k}}{\sim} i \ \forall i = 1,\dots,d\}$ be the isometry group of $\mathbf{k} = (k_1,\dots,k_d)$. Then

$$\frac{\operatorname{sgn}(s_1)^{k_{\sigma(1)}}\cdots\operatorname{sgn}(s_d)^{k_{\sigma(d)}}}{k_{\sigma(1)}^{|s_1|}\cdots k_{\sigma(d)}^{|s_d|}}$$

appears on the left-hand side of Eq. (7.19) exactly $\sharp\mathfrak{S}_d(\mathbf{k})$ times. On the right-hand side it appears exactly in those terms corresponding to partitions Π refining $\Xi(\mathbf{k})$, denoted by $\Pi \succeq \Xi(\mathbf{k})$, namely, $\Xi(\mathbf{k})$ can be obtained from Π by combining some of its equivalence classes (see Thm. A.15). Hence, to prove Eq. (7.19) it suffices to prove that for all $\mathbf{k} = (k_1,\dots,k_d)$

$$\sharp\mathfrak{S}_d(\mathbf{k}) = \sum_{\Pi\succeq\Xi(\mathbf{k})} c(\Pi).$$

Indeed, $c(\Pi)$ counts the permutations with the cycle type uniquely determined by Π while on the left-hand side every element of $\mathfrak{S}_d(\mathbf{k})$ determines a unique cycle type prescribed by some $\Pi \succeq \Xi(\mathbf{k})$.

We leave the proof of Eq. (7.20) as an exercise (see Exercise 7.6) for the interested reader. □

7.5 Binomial identities

Binomial identities play very important roles in combinatorics and discrete mathematics in general. In this section we will consider some of these in the context of MHSs and MH*Ss.

We start with an old formula known to Euler already.

Proposition 7.5.1. *For all positive integer n*

$$H_n(1) = -\sum_{k=1}^{n} \frac{(-1)^k}{k}\binom{n}{k}. \tag{7.21}$$

Proof. Let

$$f(x) = \sum_{k=1}^{n} \frac{x^k}{k}.$$

Then setting $y = 1 - x$

$$f'(x) = \sum_{k=1}^{n} x^{k-1} = \frac{1-x^n}{1-x} = \frac{1-(1-y)^n}{y} = -\sum_{k=1}^{n} (-1)^k y^{k-1} \binom{n}{k}.$$

Integrating from 1 to x we get

$$f(x) - f(1) = \sum_{k=1}^{n} (-1)^k \frac{(1-x)^k}{k} \binom{n}{k}.$$

So Eq. (7.21) follows by setting $x = 0$ as $f(1) = \sum_{k=1}^{n} \frac{1}{k} = H_n(1)$. □

We will see that more binomial identities come from Apéry's famous series

$$\zeta(2) = 3 \sum_{k=1}^{\infty} \frac{1}{k^2 \binom{2k}{k}}, \qquad \zeta(3) = \frac{5}{2} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k^3 \binom{2k}{k}}, \tag{7.22}$$

which were used in his irrationality proof of $\zeta(2)$ and $\zeta(3)$ (see [16] or [462]). These have been generalized by Leshchiner [389] to all Riemann zeta values and Dirichlet beta values $\beta(n) = \sum_{k=0}^{\infty} (-1)^k/(2k+1)^n$. For example, he showed

$$\begin{aligned}
\Big(1 - \frac{1}{2^{2m+1}}\Big)\zeta(2m+2) &= \frac{3}{2} \sum_{k=1}^{\infty} \frac{(-1)^m H_{k-1}(\{2\}^m)}{k^2 \binom{2k}{k}} \\
&\quad + 2 \sum_{j=1}^{m} \sum_{k=1}^{\infty} \frac{(-1)^{m-j} H_{k-1}(\{2\}^{m-j})}{k^{2j+2} \binom{2k}{k}}, \\
(-1)^{m-1} \cdot \zeta(2m+3) &= \frac{5}{2} \sum_{k=1}^{\infty} \frac{(-1)^k H_{k-1}(\{2\}^m)}{k^3 \binom{2k}{k}} \\
&\quad + 2 \sum_{j=1}^{m} \sum_{k=1}^{\infty} \frac{(-1)^{k-j} H_{k-1}(\{2\}^{m-j})}{k^{2j+3} \binom{2k}{k}}.
\end{aligned}$$

In the rest of this section we will state a few families of binomial identities for MH*Ss similar to Prop. 7.5.1. We will apply some of these results to prove a few congruences of MHSs and MH*Ss in Chap. 8 and to derive some identities of multiple zeta star values in Chap. 10.

Theorem 7.5.2. *Let $n, c \in \mathbb{N}$ and $a, b \in \mathbb{N}_0$. If $c \geq 2$ then*

$$\begin{aligned} H_n^{\star}(\{2\}^a, c, \{2\}^b) = & 2\sum_{k=1}^{n} \frac{(-1)^{k-1}\binom{n}{k}}{k^{2a+2b+c}\binom{n+k}{k}} \\ & + 4\sum_{i,j,\mathbf{x}} 2^{\mathrm{dp}(\mathbf{x})} \sum_{k=1}^{n} \frac{H_{k-1}(\mathbf{x}, i+2b)(-1)^{k-1}\binom{n}{k}}{k^{2a+j}\binom{n+k}{k}}, \end{aligned} \tag{7.23}$$

where in the second sum $i \geq 1$, $j \geq 2$ and $\mathbf{x} \in \mathbb{N}^r$ runs through all possible compositions of positive integers (or the empty set when $r = 0$) satisfying the restriction $i + j + |\mathbf{x}| = c$.

Remark 7.5.3. If we take $c = 3$ then we obtain the so-called *Two-three formula*. Here "Two" is capitalized to denote strings of 2's.

Theorem 7.5.4. *For a positive integer n and non-negative integers a, b, c,*

$$\begin{aligned} H_n^{\star}(\{1\}^a, c, \{1\}^b) = & \sum_{k=1}^{n} \frac{(-1)^{k-1}\binom{n}{k}}{k^{a+b+c}} \\ & + \sum_{\substack{i+j+|\mathbf{s}|=c \\ i\geq 1, j\geq 1, |\mathbf{s}|\geq 0}} \sum_{k=1}^{n} \frac{H_{k-1}(b+i, \mathbf{s})(-1)^{k-1}\binom{n}{k}}{k^{a+j}}. \end{aligned} \tag{7.24}$$

The following formulas are special cases of the so-called *Two-one formula*.

Theorem 7.5.5. *Let $a \in \mathbb{N}_0$ and $b \in \mathbb{N}$. Then for any $n \in \mathbb{N}$*

$$H_n^{\star}(\{2\}^a, 1) = 2\sum_{k=1}^{n} \frac{\binom{n}{k}}{k^{1+2a}\binom{n+k}{k}}, \tag{7.25}$$

$$H_n^{\star}(\{2\}^a, 1, \{2\}^b) = -2\sum_{k=1}^{n} \frac{(-1)^k\binom{n}{k}}{k^{1+2a+2b}\binom{n+k}{k}} - 4\sum_{k=1}^{n} \frac{H_{k-1}(\overline{2b})\binom{n}{k}}{k^{1+2a}\binom{n+k}{k}}. \tag{7.26}$$

The proofs of the above theorems will not be given here since we will generalize them to the case with argument strings having arbitrary times of repetitions of certain patterns (for example, $(\{2\}^{a_1}, c_1, \ldots, \{2\}^{a_r}, c_r)$ with $r \geq 1$) in Chap. 9, and moreover, generalize further to their q-analogs, which will be used in Chap. 10 where the multiple zeta star values are treated.

7.6 Historical notes

Even before the appearance of renewed interest in MZVs in the early 1990s, physicists already started to study the MHSs and AMHSs. In [157], Devoto

and Duke listed many concrete computations involving such sums. Later on, Blümlein and Kurth [56, 60] and Vermaseren [540] discovered that they can be realized as the general basis for the Mellin transforms of all individual functions of the momentum fraction x emerging in the quantities of massless quantum electrodynamics (QED) and quantum chromodynamics (QCD), as Mellin transforms and the inverse Mellin transforms for functions that are encountered in Feynman diagram calculations, and some other related functions having significant physical meanings. Incidentally, our binary operation $\oplus$ for the signed numbers is denoted by & by Vermaseren in [540]. Later, it is also noted that the MHS structure is closely related to high energy physics (see [57, 58]) and the relations between these sums can be used to simplify results of higher-order calculations in QED and QCD.

The first systematic study of MHSs from the number theoretical perspective was initiated by the author in the arxiv paper [589] in early 2003, and independently by Hoffman in the arxiv paper [279] in early 2004. In particular, the approach using algebras of words and quasi-symmetric functions to study MZVs was carried over to this new setting by Hoffman. Note that the star-version $H^\star$ is also denoted by S in [279] but it seems to be more appropriate to use $H^\star$ in this book due to its close connection with multiple zeta star values to be studied in Chap. 10.

Quasi-symmetric functions are generalizations of symmetric functions introduced by Gessel and Stanley in early 1980s to study the combinatorics of P-partitions and the counting of permutations with given descent sets, see [225, 226, 507]. They can be regarded as the dual algebra over the integers of the Leibniz-Hopf algebra (see [268]). In [276], Hoffman first studied MZVs from the point of view of two different product structures and later he extended these to the Hopf algebra on both MZVs and MHSs in [277, 280].

Our proof of Thm. 7.4.1 is adopted from that of Thms. 2.1 and 2.2 of [275]. Notice that [275, Thm. 2.2] was stated as Prop. 9.4 in [355] whose proof is different from that contained in [275] but also works here in the alternating setting.

We have seen many different types of sum relations among MZVs in Chap. 5. Thm. 7.4.1 is an analog of such formulas in the AMHS setting which is a generalization of Thms. 2.1 and 2.2 of [275] concerning only MHSs. Hessami Pilehroods and Tauraso discovered many similar results contained in Sec. 7.5 while trying to prove some van Hamme type congruences in [266] for MH*Ss.

Thm. A.15 generalizes [281, Thm. 6.2]. Prop. 7.5.1 was first proved

by Euler [197]. Incidentally, this is also the starting point from where Wolstenholme (see [553]), in 1862, began to prove the famous result now bearing his name (see Thm. 8.2.1). Our proof of Thm. 7.3.6 follows that of of [281, Thm. 6.3].

Finally, we note that Thms. 7.5.4 and 7.5.5 were both proved in [272] which motivates a complete proof of the Two-one formula in [608].

Exercises

7.1. Show that Eq. (7.9) and Eq. (7.10) are equivalent.

7.2. In this exercise, we re-interprete the v-dual using the following operation on the ribbons, *i.e.*, skew-Young diagrams. For a composition $\mathbf{s} = (s_1, \ldots, s_d)$ the ribbon $R_{\mathbf{s}}$ is defined to be the skew-Young diagram of d rows whose jth row starts below the last box of $(j-1)$st row and has exactly s_j boxes. Recall that the conjugate of a (skew-)Young diagram is the mirror image about the diagonal line going from the south-west corner to north-east. For e.g., the following two diagrams give the ribbon $R_{1,3,2}$ and its conjugate:

Show that in general the ribbon $R_{\mathbf{s}^\vee}$ is exactly the conjugate of the ribbon $R_{\overline{\mathbf{s}}}$. So $(2,3,1)^\vee = (1,2,1,2)$.

7.3. Show that the two converting rules in Definition 7.2.3 and Definition 7.2.4, Tog and Tog^{-1}, are indeed inverses to each other.

7.4. Prove Thm. 7.2.7.

7.5. Check that the set $\{M_{t_1,\ldots,t_d} : d \in \mathbb{N}, (t_1, \ldots, t_d) \in \mathbb{D}^d\}$ forms an integral basis for QSym_2.

7.6. Prove Eq. (7.20) by modifying the argument for the proof of Eq. (7.19).

Chapter 8

Finite multiple zeta values and finite Euler sums

A general philosophy in studying congruences is to consider the so-called van Hamme type congruences. One starts with an interesting infinite series whose summands are given by rational numbers and then, for suitable primes p, looks at the $(p-1)$-st partial sum modulo p, or even modulo higher powers p which leads to super congruences.

A classical result in this spirit is a variant of Wolstenholme's Theorem dating back to the mid nineteenth century: for all prime $p \geq 5$ we have

$$\sum_{k=1}^{p-1} \frac{1}{k^2} \equiv 0 \pmod{p}, \qquad \sum_{k=1}^{p-1} \frac{1}{k} \equiv 0 \pmod{p^2}. \tag{8.1}$$

These can be proved rather easily as follows. First, the map $f(x) = x^{-2}$ is an automorphism of $(\mathbb{Z}/p\mathbb{Z})^\times$ so the first congruence is obvious. For the second, consider the equations in the ring $\mathbb{Z}_p$ of p-adic integers:

$$\begin{aligned} 2\sum_{k=1}^{p-1} \frac{1}{k} = \sum_{k=1}^{p-1} \left(\frac{1}{k} + \frac{1}{p-k} \right) &= \sum_{k=1}^{p-1} \left(\frac{1}{k} - \frac{1}{k} \cdot \left(1 - \frac{p}{k}\right)^{-1} \right) \\ &\equiv -p \sum_{k=1}^{p-1} \frac{1}{k^2} \equiv 0 \pmod{p^2} \end{aligned}$$

by the first congruence.

These were improved to super congruences later (see Sec. 8.7). In the proof of these congruences one finds that Bernoulli numbers play the key roles by virtue of the following Faulhaber's formula: (see [5, § 23.1.4-7])

$$\sum_{j=1}^{n-1} j^d = \sum_{r=0}^{d} \binom{d+1}{r} \frac{B_r}{d+1} n^{d+1-r}, \quad \forall n, d \geq 1. \tag{8.2}$$

This chapter gives a brief introduction to the van Hamme type (super) congruences for multiple harmonic (star) sums and alternating Euler sums.

A general phenomenon is that in order to study weight w MHSs (or MH*S) modulo p^{k+1} one must first study those of weight $w+k$ objects modulo p. A few important ideas have been used to investigate this structure by a number of mathematicians. Generally speaking, they all use some kind of combinations of the following four approaches:

(1) Rearrange the sum
(2) Apply stuffle, shuffle, reversal, concatenation, and duality relations,
(3) Change to sums of positive powers and then use Bernoulli numbers/polynomials, and
(4) Exploit binomial identities.

Throughout this chapter we will follow the above approaches closely.

8.1 Definitions

Let $\mathcal{P}$ be the set of primes and let ℓ be a positive integer. Define

$$\mathcal{A}_\ell := \prod_{p\in\mathcal{P}} (\mathbb{Z}/p^\ell\mathbb{Z}) \Big/ \bigoplus_{p\in\mathcal{P}} (\mathbb{Z}/p^\ell\mathbb{Z}) \tag{8.3}$$

with componentwise addition and multiplication. Clearly the elements of $\mathcal{A}_\ell$ are represented by $(a_p)_{p\in\mathcal{P}}$ and two elements (a_p) and (b_p) represent the same element in $\mathcal{A}_\ell$ if $a_p = b_p$ for all but finitely many primes p. For every nonzero $r\in\mathbb{Q}$ by the fundamental Theorem of Arithmetic

$$r = \prod_{p\in\mathcal{P}} p^{\mathrm{ord}_p(r)}.$$

It is straightforward to see that $\mathcal{A}_\ell$ is a $\mathbb{Q}$-algebra after embedding $\mathbb{Q}$ in $\mathcal{A}_\ell$ "diagonally" using the following map:

$$\begin{aligned} \iota_\ell : \mathbb{Q} &\longrightarrow \mathcal{A}_\ell \\ r &\longrightarrow \big(\iota_{p,\ell}(r)\big)_{p\in\mathcal{P}} \end{aligned} \tag{8.4}$$

where $\iota_\ell(0) = (0)_{p\in\mathcal{P}}$ and for all nonzero $r\in\mathbb{Q}$

$$\iota_{p,\ell}(r) := \begin{cases} 0, & \text{if } \mathrm{ord}_p(r) < 0; \\ r \pmod{p^\ell}, & \text{if } \mathrm{ord}_p(r) \geq 0. \end{cases}$$

Proposition 8.1.1. *The map ι_ℓ is a monomorphism of algebras. Namely, ι_ℓ gives an embedding of $\mathbb{Q}$ into $\mathcal{A}_\ell$ as an algebra.*

Proof. Suppose $r,s\in\mathbb{Q}$ with $\iota_\ell(r) = \iota_\ell(s)$. We notice that $\mathrm{ord}_p(r) < 0$ or $\mathrm{ord}_p(s) < 0$ for only finitely many primes p. Thus $\iota_\ell(r) = \iota_\ell(s)$ implies that $r-s$ is divisible by infinitely many primes which immediately yields $r = s$. This shows ι_ℓ is injective. Since we clearly have $\iota_\ell(r)\iota_\ell(s) = \iota_\ell(rs)$ for all $r,s\in\mathbb{Q}$ the injection ι_ℓ is an algebra homomorphism. □

Remark 8.1.2. It is easy to see that the cardinality of $\mathcal{A}_\ell$ is $\aleph_1$, namely, the same as that of the real numbers. See Exercise 8.1. It is not too hard to construct an embedding of $\mathbb{R}$ into $\mathcal{A}_\ell$ as sets. But we don't know if there is an embedding of $\mathbb{R}$ as an algebra.

For convenience we abuse our notation by writing a_p for $(a_p)_{p\in\mathcal{P}}$ if no confusion arises. In particular, whenever p appears in an equation in $\mathcal{A}_\ell$ ($\ell \geq 2$) it means the element $(p)_{p\in\mathcal{P}}$. Also by writing $(a_p)_{p\geq k} \in \mathcal{A}_\ell$ we mean the element $(a_p)_{p\in\mathcal{P}}$ with $a_p = 0$ for all $p < k$.

Definition 8.1.3. We define a number $a \in \mathcal{A}_\ell$ to be *algebraic over* $\mathbb{Q}$ if there is a nontrivial polynomial $f(t) \in \mathbb{Q}[t]$ such that $f(a) = 0$. Non-algebraic numbers of $\mathcal{A}_\ell$ over $\mathbb{Q}$ are called *transcendental*. For finitely many numbers $a_1, \ldots, a_n \in \mathcal{A}_\ell$, we say they are *algebraically independent over* $\mathbb{Q}$ if for any nontrivial polynomial $f(t_1, \ldots, t_n) \in \mathbb{Q}[t_1, \ldots, t_n]$ we have $f(a_1, \ldots, a_n) \neq 0$. We call the elements in an infinite subset S of $\mathcal{A}_\ell$ *algebraically independent over* $\mathbb{Q}$ if any finitely many elements of S are.

Definition 8.1.4. For any non-negative integer k, we define the $\mathcal{A}_1$-*Bernoulli numbers* $\beta_k := \Big(-\frac{B_{p-k}}{p-k} \equiv \frac{B_{p-k}}{k} \pmod p \Big)_{p>k}$. For all $k \geq 2$ define the kth ($\mathcal{A}_1$-)Fermat quotient

$$\mathfrak{q}_k := \Big(\frac{k^{p-1}-1}{p} \pmod p \Big)_{p>k}.$$

We have $\beta_{2k} = 0$ for all $k \geq 1$ since all Bernoulli-number $B_{2j+1} = 0$ when $j \geq 1$. As for the $\mathcal{A}_1$-Fermat quotient $\mathfrak{q}_2$, according to [338], we know that for all primes less than 1.25×10^{15} the super congruence $2^{p-1} \equiv 1 \pmod{p^2}$ is satisfied by only two primes 1093 and 3511 which are called Wieferich primes. It is also known that $\mathfrak{q}_2 \neq 0$ in $\mathcal{A}_1$ under *abc*-conjecture (see [496]). Moreover, note that we have $\mathfrak{q}_{k\ell} = \mathfrak{q}_k + \mathfrak{q}_\ell$ for all k, ℓ and $\mathfrak{q}_1 = 0$. So $\mathfrak{q}_k$ is an $\mathcal{A}_1$-analog of the logarithm value $\log k$.

By considering the mysterious relation between Euler sums and their finite versions (see Sec. 8.6.3) we come to the following hypothesis.

Conjecture 8.1.5. *Put $\beta_1 = 1$ by abuse of notation. Suppose $n_1, \ldots, n_r \in \mathbb{N}$ such that $\log n_1, \ldots, \log n_r$ are $\mathbb{Q}$-linearly independent. Then the numbers in the set*

$$\bigcup_{k=0}^{\infty} \Big\{ \beta_{2k+1}\, \mathfrak{q}_{n_1}, \ldots, \beta_{2k+1}\, \mathfrak{q}_{n_r} \Big\}$$

are algebraically independent.

One should compare this with the following

Conjecture 8.1.6. *Put* $\zeta(1) = 1$ *by abuse of notation. Suppose* $n_1, \dots, n_r \in \mathbb{N}$ *such that* $\log n_1, \dots, \log n_r$ *are* $\mathbb{Q}$*-linearly independent. Then the numbers in the set*

$$\bigcup_{k=0}^{\infty} \Big\{ \zeta(2k+1) \log n_1, \dots, \zeta(2k+1) \log n_r \Big\}$$

are algebraically independent.

By the same token, every object similarly defined for each prime can be put into $\mathcal{A}_1$, for examples,

- Wilson quotients $\frac{(p-1)!+1}{p}$,
- p-adic Γ values at any rational point $\Gamma_p(a/b)$,
- p-adic zeta values $\zeta_p(k)$
- $F_{p-(\frac{5}{p})}/p$: the Fibonacci quotient (OEIS A092330)
- $U_{p-(\frac{2}{p})}(2,-1)/p$: the Pell quotient (OEIS A000129)
- $U_{p-(\frac{3}{p})}(4,1)/p$: a quotient related to the Lucas sequence 1, 4, 15, 56, 209, ... (OEIS A001353)
- $U_{p-(\frac{6}{p})}(10,1)/p$: a quotient related to the Lucas sequence 1, 10, 99, 980, 9701, ... (OEIS A004189)

Definition 8.1.7. For $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{D}^d$ (see Definition 7.1.1), we define

$$\zeta_{\mathcal{A}_\ell}(\mathbf{s}) = \sum_{p > k_1 > \cdots > k_d \geq 1} \frac{\operatorname{sgn}(s_1)^{k_1} \cdots \operatorname{sgn}(s_d)^{k_d}}{k_1^{|s_1|} \cdots k_d^{|s_d|}} \in \mathcal{A}_\ell,$$

$$\zeta^\star_{\mathcal{A}_\ell}(\mathbf{s}) = \sum_{p > k_1 \geq \cdots \geq k_n \geq 1} \frac{\operatorname{sgn}(s_1)^{k_1} \cdots \operatorname{sgn}(s_d)^{k_d}}{k_1^{|s_1|} \cdots k_d^{|s_d|}} \in \mathcal{A}_\ell.$$

These are called *finite Euler (star) sums*, abbreviated as FES (or FE*S). If all the s_j's are positive then they are called *finite multiple zeta (star) values*, abbreviated as FMZV (or FMZ*V). The number ℓ is called the *superbity*.

8.2 Reversal and concatenation relations

The famous Wolstenholme's Theorem stated in Eq. (8.1) can be rephrased in the following way.

Theorem 8.2.1. *For any prime* $p > 3$ *we have*

$$H_{p-1}(1) \equiv 0 \pmod{p^2} \qquad \text{and} \qquad H_{p-1}(2) \equiv 0 \pmod{p}.$$

In order to extend these results to general MHSs, we start in this section with some simple properties of MHSs, namely, the reversal and stuffle relations.

For any $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{D}^d$, denote its *reversal* by $\overleftarrow{\mathbf{s}} = (s_d, \ldots, s_1)$ and set $\operatorname{sgn}(\mathbf{s}) = \prod_{j=1}^{d} \operatorname{sgn}(s_j)$. The following results are called *reversal relations.*

Theorem 8.2.2. *Let* $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{D}^d$. *Then*

$$\zeta_{\mathcal{A}_1}(\overleftarrow{\mathbf{s}}) = (-1)^{|\mathbf{s}|} \operatorname{sgn}(\mathbf{s}) \zeta_{\mathcal{A}_1}(\mathbf{s}), \quad \zeta^{\star}_{\mathcal{A}_1}(\overleftarrow{\mathbf{s}}) = (-1)^{|\mathbf{s}|} \operatorname{sgn}(\mathbf{s}) \zeta^{\star}_{\mathcal{A}_1}(\mathbf{s}). \tag{8.5}$$

Proof. This follows easily from the index substitution $k \to p - k$ and the following congruence which is valid for all $s \in \mathbb{D}$:

$$\begin{aligned} \frac{\operatorname{sgn}(s)^{p-k}}{(p-k)^{|s|}} &= (-1)^{|s|} \operatorname{sgn}(s) \frac{\operatorname{sgn}(s)^k}{k^{|s|}} \left(1 - \frac{p}{k}\right)^{|s|} \\ &\equiv (-1)^s \operatorname{sgn}(s)^{k+1} \left(\frac{1}{k^{|s|}} + \frac{|s|p}{k^{|s|+1}} + \binom{|s|+1}{2} \frac{p^2}{k^{|s|+2}} \right) \pmod{p^3}. \end{aligned} \tag{8.6}$$

The rest is straight-forward and is left to the interested reader. □

Sometimes this idea can be applied to obtain super congruences, namely, with moduli given by higher powers of p.

Theorem 8.2.3. *Let* s *be any positive integer. Then*

$$\zeta_{\mathcal{A}_1}(2s) = 0, \qquad \text{and} \qquad \zeta_{\mathcal{A}_2}(2s+1) = 0. \tag{8.7}$$

Proof. Suppose the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^{\times}$ has a generator g_p. Then in $\mathcal{A}_1$

$$H_{p-1}(s) = \sum_{k=0}^{p-2} g_p^{sk} = \frac{1 - g_p^{s(p-1)}}{1 - g_p^n} = 0.$$

This argument fails if $s = p - 1$ in which case $H_{p-1}(s) = -1$ in $\mathbb{Z}/p\mathbb{Z}$. But in $\mathcal{A}_1$ this is negligible since s is fixed and we can neglect any number of finitely many primes. Therefore $\zeta_{\mathcal{A}_1}(s) = 0$ for all positive integer s. Now if $s = 2m + 1$ is odd then by taking reversal we see that in $\mathbb{Z}_p$

$$\begin{aligned} 2H_{p-1}(2m+1) &= \sum_{k=1}^{p-1} \left(\frac{1}{k^{2m+1}} + \frac{1}{(p-k)^{2m+1}} \right) \\ &\equiv \sum_{k=1}^{p-1} \frac{(2m+1)p}{k^{2m+2}} \equiv 0 \pmod{p^2}, \end{aligned}$$

because of the expansion in Eq. (8.6). This completes the proof of the theorem. □

We now lift Thm. 8.2.2 to superbity two by exactly the same idea .

Theorem 8.2.4. *Let* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{D}^d$. *Let* $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ *where* 1 *appears at the* i*th component. Then*

$$(-1)^{|\mathbf{s}|} \operatorname{sgn}(\mathbf{s}) \zeta_{\mathcal{A}_2}(\overleftarrow{\mathbf{s}}) = \zeta_{\mathcal{A}_2}(\mathbf{s}) + p \sum_{i=1}^{d} |s_i| \zeta_{\mathcal{A}_2}(\mathbf{s} \oplus \mathbf{e}_i), \tag{8.8}$$

$$(-1)^{|\mathbf{s}|} \operatorname{sgn}(\mathbf{s}) \zeta^{\star}_{\mathcal{A}_2}(\overleftarrow{\mathbf{s}}) = \zeta^{\star}_{\mathcal{A}_2}(\mathbf{s}) + p \sum_{i=1}^{d} |s_i| \zeta^{\star}_{\mathcal{A}_2}(\mathbf{s} \oplus \mathbf{e}_i), \tag{8.9}$$

where the binary operation $\oplus$ *is carried out componentwise by Eq.* (7.1).

Proof. Take reversal and then apply Eq. (8.6). □

To find a result corresponding to Thm. 8.2.3 for signed numbers the Euler polynomials and the Euler numbers are indispensable tools. Recall that the Euler polynomials $E_n(x)$ are defined by the generating function

$$\frac{2e^{tx}}{e^t + 1} = \sum_{n=0}^{\infty} E_n(x) \frac{t^n}{n!}.$$

Lemma 8.2.5. *Let* $n \in \mathbb{N}_0$. *Then we have*

$$\sum_{i=1}^{d-1} (-1)^i i^n = \frac{1}{2}\Big((-1)^{d-1} E_n(d) + E_n(0)\Big) = \sum_{a=0}^{n} \binom{n}{a} F_{n,d,a} d^{n-a}, \tag{8.10}$$

where

$$F_{n,d,a} = \begin{cases} (-1)^{d-1} E_a(0)/2, & \text{if } a < n; \\ (1 - (-1)^d) E_n(0)/2, & \text{if } a = n > 0; \\ -(1 + (-1)^d)/2, & \text{if } a = n = 0. \end{cases}$$

Moreover, $E_0(0) = 1$ *and for all* $a \in \mathbb{N}$

$$E_a(0) = \frac{2^{a+1}}{a+1}\Big(B_{a+1}\Big(\frac{1}{2}\Big) - B_{a+1}\Big) = \frac{2}{a+1}(1 - 2^{a+1}) B_{a+1}. \tag{8.11}$$

Proof. Consider the generating function

$$\begin{aligned} \sum_{n=0}^{\infty} \left(\sum_{i=1}^{d-1} (-1)^i i^n \right) \frac{t^n}{n!} &= \sum_{i=1}^{d-1} (-1)^i e^{ti} = \frac{(-e^t)^d - 1}{-e^t - 1} - 1 \\ &= \frac{(-1)^{d-1} e^{dt} + 1}{e^t + 1} - 1 \\ &= \frac{1}{2} \sum_{n=0}^{\infty} \Big((-1)^{d-1} E_n(d) + E_n(0)\Big) \frac{t^n}{n!} - 1. \end{aligned} \tag{8.12}$$

Now Eq. (8.10) follows from the notorious equation (see, for e.g., [5, § 23.1.7])

$$E_n(x) = \sum_{a=0}^{n} \binom{n}{a} E_a(0) x^{n-a} \tag{8.13}$$

for all $n > 0$. For Eq. (8.11), see, for e.g., § 23.1.20 on p. 805 of loc. cit. □

Remark 8.2.6. The classical Euler numbers E_k is defined by

$$\frac{2}{e^t + e^{-t}} = \sum_{k=0}^{\infty} E_k \frac{t^k}{k!}.$$

They are related to $E_k(0)$ by the formula

$$E_m(0) = \sum_{k=0}^{m} \binom{m}{k} \frac{E_k}{2^k} \left(-\frac{1}{2}\right)^{m-k}.$$

See § 23.1.7 on p. 805 of loc. cit.

Theorem 8.2.7. *Let $s \in \mathbb{N}$. Then*

$$\zeta_{\mathcal{A}_1}(\bar{s}) = \begin{cases} -2\mathfrak{q}_2, & \text{if } s = 1; \\ -2(1-2^{1-s})\beta_s, & \text{if } s > 2 \text{ is odd;} \end{cases} \tag{8.14}$$

$$\zeta_{\mathcal{A}_2}(\bar{s}) = s(1-2^{-s})p\beta_{s+1} \qquad \text{if } s \text{ is even.} \tag{8.15}$$

Proof. Taking $d = p$ and $n = p(p-1) - s$ in the Lemma we see that

$$\begin{aligned} H_{p-1}(\bar{s}) &\equiv F_{p(p-1)-s,p,p(p-1)} + p(p(p-1)-s)F_{p(p-1)-s,p,p(p-1)-1-s} \\ &\equiv E_{p(p-1)-s}(0) - \frac{1}{2}psE_{p(p-1)-1-s}(0) \pmod{p^2}, \end{aligned}$$

since all the coefficients in Eq. (8.10) are p-integral by Eq. (8.11) and the property of Bernoulli numbers: B_m is not p-integral if and only if $p-1$ divides $m > 0$. Then the corollary follows directly from Eq. (8.11) and the Kummer congruences

$$\frac{B_{p(p-1)-s}}{p(p-1)-s} \equiv \frac{B_{p-1-s}}{p-1-s}, \quad \frac{B_{p(p-1)-s+1}}{p(p-1)-s+1} \equiv \frac{B_{p-s}}{p-s} \pmod{p}. \qquad \square$$

Using the theory of quasi-symmetric functions one can get

Theorem 8.2.8. *For all $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{D}^d$ and positive integer ℓ*

$$\zeta^{\star}_{\mathcal{A}_\ell}(\mathbf{s}) = \sum_{\mathbf{t} \preceq \mathbf{s}} \zeta_{\mathcal{A}_\ell}(\mathbf{t}), \qquad \zeta_{\mathcal{A}_\ell}(\mathbf{s}) = \sum_{\mathbf{t} \preceq \mathbf{s}} (-1)^{\mathrm{dp}(\mathbf{s}) - \mathrm{dp}(\mathbf{t})} \zeta^{\star}_{\mathcal{A}_\ell}(\mathbf{t}). \tag{8.16}$$

When $\ell = 1$ we have

$$(-1)^{\mathrm{dp}(\mathbf{s})}\zeta^{\star}_{\mathcal{A}_1}(\overleftarrow{\mathbf{s}}) = \sum_{\bigsqcup_{j=1}^{r}\mathbf{s}_j=\mathbf{s}} (-1)^r \prod_{j=1}^{r}\zeta_{\mathcal{A}_1}(\mathbf{s}_j), \tag{8.17}$$

$$(-1)^{\mathrm{dp}(\mathbf{s})}\zeta_{\mathcal{A}_1}(\overleftarrow{\mathbf{s}}) = \sum_{\bigsqcup_{j=1}^{r}\mathbf{s}_j=\mathbf{s}} (-1)^r \prod_{j=1}^{r}\zeta^{\star}_{\mathcal{A}_1}(\mathbf{s}_j). \tag{8.18}$$

Proof. These follow immediately from Thm. 7.3.8. □

Definition 8.2.9. We will call the relations in (8.17) and (8.18) the *concatenation relations* between FMZVs and FESs.

Example 8.2.10. By (8.16) and the concatenation relation (8.17) we have

$$\zeta_{\mathcal{A}_1}(1,\bar{2}) + \zeta_{\mathcal{A}_1}(\bar{3}) = \zeta^{\star}_{\mathcal{A}_1}(1,\bar{2}) = \zeta_{\mathcal{A}_1}(\bar{2})\zeta_{\mathcal{A}_1}(1) - \zeta_{\mathcal{A}_1}(\bar{2},1).$$

Thus we get $\zeta_{\mathcal{A}_1}(\bar{3}) = -2\zeta_{\mathcal{A}_1}(\bar{2},1)$ since $\zeta_{\mathcal{A}_1}(1) = 0$ and $\zeta_{\mathcal{A}_1}(1,\bar{2}) = \zeta_{\mathcal{A}_1}(\bar{2},1)$ by the reversal relation (8.5).

8.3 Stuffle relations

We may obtain more congruences by combining the reversal relations with the stuffle relations in Thm. 7.2.7.

Definition 8.3.1. To find as many $\mathbb{Q}$-linear relations as possible in weight w we may choose all the known relations in weight $k < w$, multiply them by $\zeta_{\mathcal{A}_\ell}(\mathbf{s})$ for all $\mathbf{s} \in \mathbb{D}$ of weight $w-k$, and then expand all the products using the stuffle relations (7.6). Such $\mathbb{Q}$-linear relations among FESs of the same weight produced in the way are called *linear stuffle relations of FESs*.

For example, if k and l have the same parity then

$$0 = \zeta_{\mathcal{A}_1}(k)\zeta_{\mathcal{A}_1}(l) = \zeta_{\mathcal{A}_1}(k,l) + \zeta_{\mathcal{A}_1}(l,k) + \zeta_{\mathcal{A}_1}(k \oplus l) = 2\zeta_{\mathcal{A}_1}(k,l)$$

by Eqs. (8.7), (7.7) and (8.5). When $a, b > 0$ this is a special case of the more general Thm. 8.5.12.

An application of the stuffle relation yields

Theorem 8.3.2. *Let s and t be two positive integers.*

(i) *If $s+t$ is odd then*

$$\zeta^{\star}_{\mathcal{A}_2}(s,t) = \zeta_{\mathcal{A}_2}(s,t). \tag{8.19}$$

(ii) *If t is odd. Then*

$$\zeta^{\star}_{\mathcal{A}_2}(s,t,s) = \zeta_{\mathcal{A}_2}(s,t,s). \tag{8.20}$$

Proof. Part (i) follows easily from Thm. 8.2.3 since

$$\zeta^\star_{\mathcal{A}_2}(s,t) = \zeta_{\mathcal{A}_2}(s,t) + \zeta_{\mathcal{A}_2}(s+t).$$

For part (ii), by the stuffle relation we get

$$\zeta^\star_{\mathcal{A}_2}(s,t,s) = \zeta_{\mathcal{A}_2}(s,t,s) + \zeta_{\mathcal{A}_2}(s+t)\zeta_{\mathcal{A}_2}(s) - \zeta_{\mathcal{A}_2}(2s+t) = \zeta_{\mathcal{A}_2}(s,t,s)$$

by Thm. 8.2.3, as desired. □

8.4 Shuffle relations

It is not hard to see that Euler sums can be expressed by iterated integrals, just like MZVs. Suppose $s_j \in \mathbb{D}$ and $\mathrm{sgn}(s_j) = \eta_j \in \{1,-1\}$ for all $j = 1,\dots,d$. Then

$$\zeta(s_1,\dots,s_n) = \int_0^1 \left(\frac{\mathrm{d}t}{t}\right)^{|s_1|} \left(\frac{\mathrm{d}t}{\xi_1 - t}\right) \cdots \left(\frac{\mathrm{d}t}{t}\right)^{|s_d|} \left(\frac{\mathrm{d}t}{\xi_d - t}\right), \quad (8.21)$$

where $\xi_i = \prod_{j=1}^i \eta_i$, $i = 1,\dots,d$.

Recall from Definition 7.2.1 that $\mathfrak{A}_2^1$ is the algebra generated by $z_s, s \in \mathbb{D}$. Write $z_s = \mathtt{y}_{|s|,\mathrm{sgn}(s)}$. Then we define $\mathrm{tog}, \mathrm{tog}^{-1} : \mathfrak{A}_2^1 \to \mathfrak{A}_2^1$ (see) by

$$\mathrm{tog}(\mathtt{y}_{s_1,\xi_1}\cdots\mathtt{y}_{s_d,\xi_d}) := \mathtt{y}_{s_1,\eta_1}\cdots\mathtt{y}_{s_d,\eta_d},$$

where $\eta_i = \prod_{j=1}^i \xi_j$ and

$$\mathrm{tog}^{-1}(\mathtt{y}_{s_1,\eta_1}\cdots\mathtt{y}_{s_d,\eta_d}) := \mathtt{y}_{s_1,\xi_1}\cdots\mathtt{y}_{s_d,\xi_d},$$

where $\xi_j = \eta_{j-1}^{-1}\eta_j$ (setting $\eta_0 = 1$).

For all $s_1,\dots,s_d \in \mathbb{D}$, we define one-variable multiple polylog

$$\mathrm{L}_{s_1,\dots,s_d}(z) = \sum_{k_1>k_2>\dots>k_d\geq 1} z^{k_1} \prod_{j=1}^d \frac{\mathrm{sgn}(s_j)^{k_j}}{k_j^{|s_j|}}, \quad 0<|z|<1,$$

which generalizes Eq. (3.19) to the larger index set $\mathbb{D}$. It satisfies

$$\mathrm{L}_{s_1,\dots,s_d}(z) = \int_0^z \left(\frac{\mathrm{d}t}{t}\right)^{|s_1|-1} \left(\frac{\mathrm{d}t}{\xi_1 - t}\right) \cdots \left(\frac{\mathrm{d}t}{t}\right)^{|s_d|-1} \left(\frac{\mathrm{d}t}{\xi_d - t}\right), \quad (8.22)$$

where $\xi_i = \mathrm{sgn}(s_{i-1})\,\mathrm{sgn}(s_i)$, $i = 1,\dots,d$. Here we have set $\mathrm{sgn}(s_0) = 1$.

Define $W : \bigcup_{d\geq 1}\mathbb{D}^d \to \mathfrak{A}_2^1$ (see Definition 7.2.1) by setting $W(s_1,\dots,s_d) = \mathtt{z}_{s_1}\dots\mathtt{z}_{s_d}$. Further define the map $\mathrm{L}(-;z) : (\mathfrak{A}_2^1, *) \to \mathbb{Q}[\![z]\!]$ and $\mathrm{L}(\mathbf{w};z) = \mathrm{L}_{\mathbf{s}}(z)$ for all $\mathbf{w} = W(\mathbf{s})$.

Proposition 8.4.1. *Define the map* $\mathrm{L}_{ш}(-;z) : (\mathfrak{A}_2^1, ш) \to \mathbb{Q}[\![z]\!]$ *by setting* $\mathrm{L}_{ш}(\mathbf{w};z) = \mathrm{L}(\mathrm{tog}(\mathbf{w});z)$. *Then* $\mathrm{L}_{ш}(-;z) : (\mathfrak{A}_2^1, ш) \to \mathbb{Q}[\![z]\!]$ *is an algebra homomorphism.*

Proof. This follows from (8.22) and the shuffle relation satisfied by iterated integrals. □

Example 8.4.2. Suppose $\mathbf{s} = (\bar{2}, 1)$. Then $W(\mathbf{s}) = \mathtt{y}_{2,-1}\mathtt{y}_{1,1}$ and

$$\mathrm{L}(\mathtt{y}_{2,-1}\mathtt{y}_{1,1}; z) = \int_0^1 \left(\frac{\mathrm{d}t}{t}\right)^2 \left(\frac{-dt}{1+t}\right)^2 = \mathrm{L}_{\bar{2},1}(z).$$

Set $\zeta_{\mathcal{A}_1,\sqcup\!\sqcup} = \zeta_{\mathcal{A}_1,*} \circ \mathrm{tog} : \mathfrak{A}_2^1 \to \mathcal{A}_1$. Although $\zeta_{\mathcal{A}_1,\sqcup\!\sqcup}$ is not an algebra homomorphism from $(\mathfrak{A}_2^1, \sqcup\!\sqcup)$ we will see in the next theorem that it does provide a kind of shuffle relation. It is the FES analog of Thm. 6.3.4 for SMZVs.

Recall that $\tau : \mathfrak{A}_1^1 \to \mathfrak{A}_1^1$ is defined by (6.9).

Theorem 8.4.3. *For all words* $\mathbf{w}, \mathbf{u} \in \mathfrak{A}_1^1$ *and* $\mathbf{v} \in \mathfrak{A}_2^1$*, we have*

(i) $\zeta_{\mathcal{A}_1,\sqcup\!\sqcup}(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) = \zeta_{\mathcal{A}_1,\sqcup\!\sqcup}(\tau(\mathbf{u})\mathbf{v})$,
(ii) $\zeta_{\mathcal{A}_1,\sqcup\!\sqcup}((\mathbf{w}\mathbf{u}) \sqcup\!\sqcup \mathbf{v}) = \zeta_{\mathcal{A}_1,\sqcup\!\sqcup}(\mathbf{u} \sqcup\!\sqcup \tau(\mathbf{w})\mathbf{v})$,
(iii) *For all* $s \in \mathbb{N}$ $\zeta_{\mathcal{A}_1,\sqcup\!\sqcup}((\mathtt{x}_0^{s-1}\mathtt{x}_1\mathbf{u}) \sqcup\!\sqcup \mathbf{v}) = (-1)^s \zeta_{\mathcal{A}_1,\sqcup\!\sqcup}(\mathbf{u} \sqcup\!\sqcup (\mathtt{x}_0^{s-1}\mathtt{x}_1\mathbf{v}))$.

Proof. Taking $\mathbf{u} = \emptyset$ and then $\mathbf{w} = \mathbf{u}$ we see that (ii) implies (i). Decomposing $\mathbf{w}$ into strings of the type $\mathtt{x}_0^{s-1}\mathtt{x}_1$ we see that (iii) implies (ii). Now we only need to prove (iii).

For simplicity let $\mathtt{a} = \mathtt{x}_0$ and $\mathtt{b} = \mathtt{x}_1$ in the rest of this chapter. Let $s_j \in \mathbb{D}$ with $\mathrm{sgn}(s_j) = \eta_j$ for $j = 1, \ldots, d$. Let $\mathbf{u} = W(\mathbf{s})$ and $\mathbf{v} = \mathrm{tog}(\mathbf{u}) = \mathtt{y}_{t_1,\xi_1} \cdots \mathtt{y}_{t_d,\xi_d} \in \mathbb{D}^d$. Then $\mathrm{tog}(\mathtt{b}\mathbf{u}) = \mathtt{b}\mathbf{v}$.

For any prime $p > 2$, the coefficient of z^p in $\mathrm{L}_{\sqcup\!\sqcup}(\mathtt{b}\mathbf{u}; z)$ is given by

$$\mathrm{Coeff}_{z^p}\big[\mathrm{L}_{\sqcup\!\sqcup}(b\mathbf{u}; z)\big] = \frac{1}{p} \sum_{p>k_1>\cdots>k_d>0} \frac{\xi_1^{k_1} \cdots \xi_d^{k_d}}{k_1^{t_1} \cdots k_d^{t_d}} = \frac{1}{p} H_{p-1}(\mathrm{tog}(\mathbf{u})).$$

Observe that

$$\mathtt{b}\Big((\mathtt{a}^{s-1}\mathtt{b}\mathbf{u}) \sqcup\!\sqcup \mathbf{v} - (-1)^s \mathbf{u} \sqcup\!\sqcup (\mathtt{a}^{s-1}\mathtt{b}\mathbf{v})\Big) = \sum_{i=0}^{s-1} (-1)^i (\mathtt{a}^{s-1-i}\mathtt{b}\mathbf{u}) \sqcup\!\sqcup (\mathtt{a}^i\mathtt{b}\mathbf{v}).$$

By first applying $\mathrm{L}_{\sqcup\!\sqcup}(-; z)$ to the above and then extracting the coefficients of z^p from both sides we get

$$\begin{aligned}
&\frac{1}{p}\Big(H_{p-1} \circ \mathrm{tog}\left((\mathtt{a}^{s-1}\mathtt{b}\mathbf{u}) \sqcup\!\sqcup \mathbf{v}\right) - (-1)^s H_{p-1} \circ \mathrm{tog}\left(\mathbf{u} \sqcup\!\sqcup (\mathtt{a}^{s-1}\mathtt{b}\mathbf{v})\right)\Big) \\
&= \sum_{i=0}^{s-1} (-1)^i \mathrm{Coeff}_{z^p}\big[\mathrm{L}_{\sqcup\!\sqcup}(\mathtt{a}^{s-1-i}\mathtt{b}\mathbf{u}; z)\, \mathrm{L}_{\sqcup\!\sqcup}(\mathtt{a}^i\mathtt{b}\mathbf{v}; z)\big] \\
&= \sum_{i=0}^{s-1} (-1)^i \sum_{j=1}^{p-1} \mathrm{Coeff}_{z^j}\big[\mathrm{L}_{\sqcup\!\sqcup}(\mathtt{a}^{s-1-i}\mathtt{b}\mathbf{u}; z)\big] \mathrm{Coeff}_{z^{p-j}}\big[\mathrm{L}_{\sqcup\!\sqcup}(\mathtt{a}^i\mathtt{b}\mathbf{v}; z)\big]
\end{aligned}$$

by Prop. 8.4.1. Now the last sum is p-integral since $p-j<p$ and $j<p$ and therefore we get

$$H_{p-1}\circ\mathrm{tog}((\mathtt{a}^{s-1}\mathtt{bu})\sqcup\!\sqcup\,\mathbf{v})\equiv(-1)^s H_{p-1}\circ\mathrm{tog}(\mathbf{u}\sqcup\!\sqcup\,(\mathtt{a}^{s-1}\mathtt{bv}))\pmod p$$

which completes the proof of (iii). □

Definition 8.4.4. A relation produced by Thm. 8.4.3 is called a *linear shuffle relation* of FESs. For each weight $w\geq 2$, by the *double shuffle relations* of FESs of weight w we mean all the linear shuffle relations of weight w and all the linear stuffle relations of w defined in Definition 8.3.1.

Restricting to FMZVs, we obtain the *linear shuffle relations* and *double shuffle relations* of FMZVs.

8.5 Finite multiple zeta values

We now apply the results from the previous sections to study the FMZVs.

8.5.1 *Homogeneous finite multiple zeta values*

The following homogeneous generalization of Thm. 8.2.3 will be very handy, even though we will extend it further.

Theorem 8.5.1. *Let s and d be two positive integers. Then*

$$\zeta_{\mathcal{A}_1}(\{s\}^d)=0 \quad \textit{if the weight } ds \textit{ is even;} \tag{8.23}$$

$$\zeta_{\mathcal{A}_2}(\{s\}^d)=0 \quad \textit{if the weight } ds \textit{ is odd.} \tag{8.24}$$

Proof. We may proceed by induction on d. The case $d=1$ is just Thm. 8.2.3. Assume the theorem holds for $d\geq 2$. Then

$$0=\zeta_{\mathcal{A}_2}(s)\zeta_{\mathcal{A}_2}(\{s\}^{d-1})=d\zeta_{\mathcal{A}_2}(\{s\}^d)+\frac{1}{(d-2)!}\sum_{\sigma\in\mathfrak{S}_d}\zeta_{\mathcal{A}_2}\big(\sigma(2s,\{s\}^{d-2})\big).$$

So the theorem follows from Thm. 7.4.1 after counting the weight. □

Our goal in this section is to prove an improvement of Thm. 8.5.1. First, we need some preliminary results.

Lemma 8.5.2. *Let p be an odd prime>3 and let r be a positive integer, $r\leq p-3$, then*

$$H_{p-1}(r)=\sum_{k=1}^{p-1}\frac{1}{k^r}\equiv\begin{cases}-\frac{r(r+1)}{2(r+2)}B_{p-r-2}p^2 \pmod{p^3} & \textit{if } 2\nmid r,\\ \frac{r}{r+1}B_{p-r-1}p \pmod{p^2} & \textit{if } 2|r.\end{cases}$$

The proof is standard so we leave it to Exercise 8.2.

Lemma 8.5.3. *Let $r, \alpha_1, \cdots, \alpha_n$ be positive integers, $r = \alpha_1 + \alpha_2 + \cdots + \alpha_n \leq p-3$, then*

$$\sum_{\substack{1\leq l_1,\cdots,l_n\leq p-1\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1}\cdots l_n^{\alpha_n}} \equiv \begin{cases} \frac{(-1)^n n! r(r+1)}{2n(r+2)} B_{p-r-2}p^2 \pmod{p^3} & \text{if } 2\nmid r, \\ \frac{(-1)^{n-1} n! r}{n(r+1)} B_{p-r-1}p \pmod{p^2} & \text{if } 2|r. \end{cases}$$

Proof. If $n = 1$, the congruence is obvious by Lemma 8.5.2. We assume that the formula holds when the number of variables is less than n. Then by induction

$$\begin{aligned} &\sum_{\substack{1\leq l_1,\cdots,l_n<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1}\cdots l_n^{\alpha_n}} \\ &= \sum_{\substack{1\leq l_1,\cdots,l_{n-1}<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1}\cdots l_{n-1}^{\alpha_{n-1}}} \Big(\sum_{l_n=1}^{p-1} \frac{1}{l_n^{\alpha_n}} - \frac{1}{l_1^{\alpha_n}} - \cdots - \frac{1}{l_{n-1}^{\alpha_n}}\Big) \\ &= \sum_{\substack{1\leq l_1,\cdots,l_{n-1}<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1}\cdots l_{n-1}^{\alpha_{n-1}}} \sum_{l_n=1}^{p-1} \frac{1}{l_n^{\alpha_n}} - \sum_{\substack{1\leq l_1,\cdots,l_{n-1}<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1+\alpha_n} l_2^{\alpha_2}\cdots l_{n-1}^{\alpha_{n-1}}} \\ &\quad - \cdots - \sum_{\substack{1\leq l_1,\cdots,l_{n-1}<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1} l_2^{\alpha_2}\cdots l_{n-1}^{\alpha_{n-1}+\alpha_n}} \end{aligned}$$

By Lemma 8.5.2 and the inductive assumption, we have

$$\sum_{\substack{1\leq l_1,\ldots,l_{n-1}<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1} l_2^{\alpha_2}\cdots l_{n-1}^{\alpha_{n-1}}} \sum_{l_n=1}^{p-1} \frac{1}{l_n^{\alpha_n}} \equiv \begin{cases} 0 \pmod{p^3} & \text{if } 2\nmid r, \\ 0 \pmod{p^2} & \text{if } 2|r. \end{cases}$$

If r is odd, then by the inductive assumption,

$$\begin{aligned} \sum_{\substack{1\leq l_1,\ldots,l_n<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1} l_2^{\alpha_2}\cdots l_n^{\alpha_n}} &\equiv -(n-1) \sum_{\substack{1\leq l_1,\cdots,l_{n-1}<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1+\alpha_n} l_2^{\alpha_2}\cdots l_{n-1}^{\alpha_{n-1}}} \\ &\equiv -(n-1)(-1)^{n-1}(n-2)! \frac{r(r+1)}{2(r+2)} B_{p-r-2}p^2 \\ &\equiv (-1)^n (n-1)! \frac{r(r+1)}{2(r+2)} B_{p-r-2}p^2 \pmod{p^3}. \end{aligned}$$

Similarly, if r is even, then by the inductive assumption,

$$\sum_{\substack{1\le l_1,\cdots,l_n<p\\ l_i\neq l_j}} \frac{1}{l_1^{\alpha_1}l_2^{\alpha_2}\cdots l_n^{\alpha_n}} \equiv (-1)^{n-1}(n-1)!\frac{r}{r+1}B_{p-r-1}p \pmod{p^2}.$$

The lemma is now proved. □

Theorem 8.5.4. *Let $d, s \in \mathbb{N}$. If ds is even then*

$$\zeta_{\mathcal{A}_2}(\{s\}^d) = (-1)^d s\beta_{ds+1}p. \tag{8.25}$$

If ds is odd then

$$\zeta_{\mathcal{A}_3}(\{s\}^d) = -(-1)^n\frac{s(ds+1)}{2}\beta_{ds+2}p^2. \tag{8.26}$$

Proof. We can permute $l_1, \ldots, l_d$ in the inequality $l_1 > \cdots > l_d$ in $d!$ ways and then apply Lemma 8.5.3. This proves the theorem. □

8.5.2 *Duality of finite multiple zeta values*

The FMZVs satisfy a different kind of duality from that of MZVs. Recall from Definition 7.3.4 that for all positive integers $r_1, \ldots, r_\ell, t_1, \ldots, t_\ell$ and

$$\mathbf{s} = (r_1, \{1\}^{t_1-1}, r_2+1, \{1\}^{t_2-1}, \ldots, r_\ell+1, \{1\}^{t_\ell-1}),$$

we define the v-dual of $\mathbf{s}$ by

$$\mathbf{s}^\vee = (\{1\}^{r_1-1}, t_1+1, \{1\}^{r_2-1}, t_2+1, \ldots, t_{\ell-1}+1, \{1\}^{r_\ell-1}, t_\ell).$$

Theorem 8.5.5. *If $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ then*

$$\zeta^\star_{\mathcal{A}_1}(\mathbf{s}) = -\zeta^\star_{\mathcal{A}_1}(\mathbf{s}^\vee). \tag{8.27}$$

Proof. These follow immediately from Eq. (7.18). □

Definition 8.5.6. For any sequence $\{f(n)\}_{n\ge 0}$ with $f(0) = 0$ the summation operator Σ and the *conjugation* operator ∇ are defined by (using umbral calculus notation $f^k = f(k)$)

$$\Sigma f(n) = \sum_{i=0}^{n} f(i), \quad \nabla f(n) = (1-f)^n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}f(k), \quad \forall n \ge 0.$$

For any function $g(n) = \sum_{k=0}^{n} f(k)$ we set $\Sigma^{-1}g = f$.

Lemma 8.5.7. *The two operators Σ and ∇ generate a dihedral group, i.e, $\nabla^2 = \mathrm{id}$, $\Sigma\nabla = \nabla\Sigma^{-l}$, and therefore $(\Sigma\nabla)^2 = \mathrm{id}$.*

Proof. Clearly $\nabla^2 f(n) = (1-(1-f))^n = f^n = f(n)$. Moreover by changing the orders of summation

$$\begin{aligned}
\Sigma\nabla\Sigma f(n) &= \sum_{k=0}^{n}\sum_{j=0}^{k}(-1)^j\binom{k}{j}\sum_{l=0}^{j} f(l)\\
&= \sum_{k=0}^{n}\sum_{l=0}^{k}\sum_{j=l}^{k}(-1)^j\binom{k}{j} f(l)\\
&= \sum_{k=0}^{n}\sum_{l=0}^{k}(-1)^l\binom{k-1}{l-1} f(l)\\
&= \sum_{l=0}^{n}\sum_{k=l}^{n}(-1)^l\binom{k-1}{l-1} f(l)\\
&= \sum_{l=0}^{n}(-1)^l\binom{n}{l} f(l) = \nabla f(n).
\end{aligned}$$

This implies $(\Sigma\nabla)^2 = \mathrm{id}$ easily and the lemma is proved. □

Lemma 8.5.8. *For any positive number n and composition $\mathbf{s}$ of positive integers we have*

$$\Sigma\nabla H_n^\star(\mathbf{s}) = -H_n^\star(\mathbf{s}^\vee).$$

Proof. For $\mathbf{s} = (s_1, s_2, \ldots, s_d)$, we see that

$$\begin{aligned}
\nabla H_n^\star(\mathbf{s}) &= \sum_{k=0}^{n}(-1)^k\binom{n}{k}\sum_{k_1=1}^{k}\frac{1}{k_1^{s_1}}\sum_{k_1\geq k_2\geq\cdots\geq k_d}\frac{1}{k_2^{s_2}\cdots k_d^{s_d}}\\
&= \sum_{k_1=1}^{n}\sum_{k=k_1}^{n}(-1)^k\binom{n}{k}\frac{1}{k_1^{s_1}}\sum_{k_1\geq k_2\geq\cdots\geq k_d}\frac{1}{k_2^{s_2}\cdots k_d^{s_d}}\\
&= \sum_{k_1=1}^{n}(-1)^{k_1}\binom{n-1}{k_1-1}\frac{1}{k_1^{s_1}}\sum_{k_1\geq k_2\geq\cdots\geq k_d}\frac{1}{k_2^{s_2}\cdots k_d^{s_d}}\\
&= \frac{1}{n}\nabla f(n), \qquad (8.28)
\end{aligned}$$

where

$$f(n) = \frac{1}{n^{s_1-1}}H_n^\star(s_2,\ldots,s_d) = \begin{cases} H_n^\star(s_2,\ldots,s_d), & \text{if } s_1 = 1;\\ \Sigma^{-1}H_n^\star(s_1-1,s_2,\ldots,s_d), & \text{otherwise.}\end{cases}$$

We now prove the lemma by induction on the weight of $\mathbf{s}$. The weight one case is $\Sigma\nabla H_n^\star(1) = \nabla\Sigma^{-1}H_n^\star(1) = -H_n^\star(1)$, which is just Eq. (7.21).

Assume the lemma holds if $|\mathbf{s}| < n$. Suppose now $\mathbf{s} = (s_1, \ldots, s_d)$ has weight n. We consider two cases: (i) $s_1 = 1$, and (ii) $s_1 > 1$. If $s_1 = 1$ then $\mathbf{s}^\vee = (j_1+1, j_2, \ldots, j_r)$. Putting $(s_2, \ldots, s_d)^\vee = \mathbf{t} = (j_1, \ldots, j_r)$, using Eq. (8.28) and the induction hypothesis we get

$$\Sigma\nabla H^\star_n(\mathbf{s}) = \Sigma\left(\frac{\nabla H^\star_n(\mathbf{t}^\vee)}{n}\right) = -\Sigma\left(\frac{\Sigma^{-1}H^\star_n(\mathbf{t})}{n}\right) = -H^\star_n(j_1+1, j_2, \ldots, j_r).$$

Hence the lemma holds in this case. If $s_1 > 1$, $\mathbf{s}^\vee = (1, j_1, \ldots, j_r)$. Putting $\mathbf{t} = (s_1 - 1, s_2, \ldots, s_k)^\vee = (j_1, \ldots, j_r)$ we have Then

$$\begin{aligned}\Sigma\nabla H^\star_n(\mathbf{s}) =& \Sigma\left(\frac{1}{n}\nabla\Sigma^{-1}H^\star_n(\mathbf{t}^\vee)\right) = \Sigma\left(\frac{1}{n}\Sigma\nabla H^\star_n(\mathbf{t}^\vee)\right)\\ &= -\Sigma\left(\frac{1}{n}H^\star_n(\mathbf{t})\right) = -H^\star_n(1, j_1, \ldots, j_r) = -H^\star_n(\mathbf{s}^\vee)\end{aligned}$$

as desired. The lemma is now proved. □

Recall from Eq. (7.14) that $\rho_n : (\mathfrak{A}^1_2, *) \to \mathbb{R}$ is defined by $\rho_n(M_\mathbf{s}) = H_n(\mathbf{s})$. Let $\rho = (\rho_n)_{n\geq 1}$ be restricted to $\mathfrak{A}^1_1$.

Corollary 8.5.9. *The following diagram commutes:*

$$\begin{array}{ccc}\mathsf{QSym}_1 & \xrightarrow{\ \varphi\ } & \mathsf{QSym}_1\\ \Big\downarrow{\scriptstyle \rho} & & \Big\downarrow{\scriptstyle \rho}\\ \mathbb{R}^{\mathbb{N}} & \xrightarrow{\ \Sigma\nabla\ } & \mathbb{R}^{\mathbb{N}}\end{array}$$

where the map φ defined by Eq. 6.11 *is an involution on $\mathfrak{A}^1_1$ such that $\mathtt{a}^\varphi = \mathtt{a} + \mathtt{b}$ and $\mathtt{b}^\varphi = -\mathtt{b}$.*

Proof. This follows immediately from Thm. 7.3.6(ii) and Lemma 8.5.8. □

Parallel to Thm. 6.3.5 for the SMZVs, the following result provides another evidence for Conjecture 8.5.15.

Theorem 8.5.10. *For all $\mathbf{w} \in \mathfrak{A}^1_1$ we have*

$$\zeta_{\mathcal{A}}(\mathbf{w}) = \zeta_{\mathcal{A}}(\mathbf{w}^\varphi).$$

Proof. This is clear from Cor. 8.5.9 by considering ρ_{p-1} for primes p. □

We now consider duality property in superbity two.

Theorem 8.5.11. *Let* $\mathbf{s}$ *be any composition of positive integers of weight* w*. Then*

$$-\zeta^{\star}_{\mathcal{A}_2}(\mathbf{s}^{\vee}) = \zeta^{\star}_{\mathcal{A}_2}(\mathbf{s}) + p\left(\sum_{\mathbf{t}\preceq\mathbf{s}} \zeta_{\mathcal{A}_2}(1,\mathbf{t})\right). \tag{8.29}$$

Proof. By Lemma 8.5.8 we have

$$\begin{aligned}
&-\zeta^{\star}_{\mathcal{A}_2}(\mathbf{s}^{\vee}) - \zeta^{\star}_{\mathcal{A}_2}(\mathbf{s}) \\
&= \Sigma\nabla H^{\star}_{p-1}(\mathbf{s}) - \zeta^{\star}_{\mathcal{A}_2}(\mathbf{s}) = \sum_{i=0}^{p-2} \binom{p}{i+1} (-1)^i H^{\star}_i(\mathbf{s}) \qquad (\text{in } \mathcal{A}_2) \\
&= p\left(\sum_{i=0}^{p-2} \frac{1}{i+1}\binom{p-1}{i}(-1)^i H^{\star}_i(\mathbf{s})\right) \\
&= p\left(\sum_{i=0}^{p-2} \frac{1}{i+1} H^{\star}_i(\mathbf{s})\right) = p\left(\sum_{i=0}^{p-2}\sum_{\mathbf{t}\preceq\mathbf{s}} \frac{H_i(\mathbf{t})}{i+1}\right) = p\left(\sum_{\mathbf{t}\preceq\mathbf{s}} \zeta_{\mathcal{A}_2}(1,\mathbf{t})\right),
\end{aligned}$$

as desired. □

8.5.3 *Finite multiple zeta values of lower depths*

In this section we use Fermat's Little Theorem and Bernoulli numbers to derive $\mathbb{Q}$-linear relations among FMZVs when the depth is two or three.

Theorem 8.5.12. *For all positive integers* s *and* t

$$\zeta_{\mathcal{A}_1}(s,t) = (-1)^s \binom{s+t}{s} \beta_{s+t}. \tag{8.30}$$

Proof. The Faulhaber's formula Eq. (8.2) together with Fermat's Little Theorem yields

$$\begin{aligned}
\zeta_{\mathcal{A}_1}(s,t) = \zeta^{\star}_{\mathcal{A}_1}(s,t) &= \sum_{a=1}^{p-1}\sum_{b=1}^{a-1} \frac{1}{a^s b^t} = \sum_{a=1}^{p-1} a^{-s} \sum_{b=1}^{a-1} b^{p-1-t} \\
&= \sum_{a=1}^{p-1} \frac{a^{-s}}{p-t} \sum_{k=0}^{p-1-t} \binom{p-t}{k} B_k a^{p-t-k} = \frac{1}{p-t} \sum_{k=0}^{p-1-t} \binom{p-t}{k} B_k \sum_{a=1}^{p-1} a^{p-s-t-k}.
\end{aligned}$$

In $\mathcal{A}_1$ we know $\sum_{a=1}^{p-1} a^{p-s-t-k} = 0$ unless $k = p-s-t$. Thus

$$\begin{aligned}
\zeta_{\mathcal{A}_1}(s,t) = \zeta^{\star}_{\mathcal{A}_1}(s,t) &= (s+t)\frac{p-1}{p-t}\binom{p-t}{p-s-t}\beta_{s+t} \\
&= \frac{-(s+t)(p-t-1)!}{s!(p-s-t)!}\beta_{s+t}.
\end{aligned}$$

Equation Eq. (8.30) follows from this easily since by Wilson's Theorem $\frac{1}{(p-m)!} = (-1)^m(m-1)!$ in $\mathcal{A}_1$ for any fixed integer m. □

Theorem 8.5.13. *Let* $(l,m,n) \in \mathbb{N}^3$. *If* $w = l+m+n$ *is* odd *then*

$$-\zeta_{\mathcal{A}_1}(l,m,n) = \zeta^\star_{\mathcal{A}_1}(l,m,n) = \left[(-1)^l\binom{w}{l} - (-1)^n\binom{w}{n}\right]\frac{\beta_w}{2}.$$

Proof. From Eq. (8.16) and the concatenation relation Eq. (8.18) we have

$$\zeta_{\mathcal{A}_1}(l,m,n) = \zeta^\star_{\mathcal{A}_1}(l,m,n) - \zeta^\star_{\mathcal{A}_1}(l+m,n) - \zeta^\star_{\mathcal{A}_1}(l,m+n) + \zeta^\star_{\mathcal{A}_1}(l+m+n),$$

and

$$\begin{aligned}\zeta_{\mathcal{A}_1}(l,m,n) = \zeta^\star_{\mathcal{A}_1}(n)\zeta^\star_{\mathcal{A}_1}(m)\zeta^\star_{\mathcal{A}_1}(l) - \zeta^\star_{\mathcal{A}_1}(n,m)\zeta^\star_{\mathcal{A}_1}(l)\\ - \zeta^\star_{\mathcal{A}_1}(n)\zeta^\star_{\mathcal{A}_1}(m,l) + \zeta^\star_{\mathcal{A}_1}(n,m,l),\end{aligned} \quad (8.31)$$

respectively. Thus by Thm. 8.2.3 we have

$$\zeta^\star_{\mathcal{A}_1}(l,m,n) - \zeta^\star_{\mathcal{A}_1}(l+m,n) - \zeta^\star_{\mathcal{A}_1}(l,m+n) = \zeta^\star_{\mathcal{A}_1}(n,m,l).$$

Notice the weight is odd so the reversal relation Eq. (8.5) and Thm. 8.5.12 imply that

$$2\zeta^\star_{\mathcal{A}_1}(l,m,n) = \zeta^\star_{\mathcal{A}_1}(l+m,n)+\zeta^\star_{\mathcal{A}_1}(l,m+n) = \left[(-1)^l\binom{w}{l} - (-1)^n\binom{w}{n}\right]\beta_w.$$

This yields the second equation in the theorem. The first equation follows immediately from Eq. (8.31) by reversal relation. This completes the proof of the theorem. □

8.5.4 *Structure of finite multiple zeta values*

We denote by $\mathsf{FMZ}_{w,\ell}$ be the $\mathbb{Q}$-vector subspace of $\mathcal{A}_\ell$ generated by all weight w and superbity ℓ FMZV (resp. FMZ*V). Further, we write $\mathsf{FMZ}_w = \mathsf{FMZ}_{w,1}$. From overwhelming numerical evidence, we can formulate the following conjecture.

Conjecture 8.5.14. *Let* w *be any positive integer.*

(i) *Set* $d_0 = 1$ *and* $d_w = \dim \mathsf{FMZ}_w$ *for all* $w \geq 1$. *Then* $d_1 = d_2 = 0$ *and*

$$d_w = d_{w-2} + d_{w-3} \quad \forall w \geq 3.$$

(ii) *For all* $w \geq 3$, FMZ_w *has a basis*

$$\{\zeta_{\mathcal{A}_1}(1,2,a_1,\ldots,a_r) : a_1,\ldots,a_r = 2 \text{ or } 3, a_1+\cdots+a_r = w-3\}.$$

(iii) *All the $\mathbb{Q}$-linear relations among FMZVs can be produced by the double shuffle relations.*

In fact, Kaneko and Zagier has proposed the following conjecture providing a relation between FMZVs and MZVs.

Conjecture 8.5.15. *There is an $\mathbb{Q}$-algebra isomorphism*

$$f_{\mathrm{KZ}} : \mathsf{MZ}_w/\zeta(2)\mathsf{MZ}_{w-2} \longrightarrow \mathsf{FMZ}_{w,1}$$
$$\zeta^{\mathcal{S}}_{ш}(\mathbf{s}) \longmapsto \zeta_{\mathcal{A}_1}(\mathbf{s})$$

where $\zeta^{\mathcal{S}}_{ш}(\mathbf{s})$ is defined in Eq. (6.1).

It follows from this conjecture that $\dim \mathsf{FMZ}_{w,1} = \dim \mathsf{MZ}_w - \dim \mathsf{MZ}_{w-2} = \dim \mathsf{MZ}_{w-3}$ by Conjecture 3.4.2 of Zagier. However, even in weight 3, it seems impossible to prove the map f_{KZ} is injective at the moment. Indeed, we have

$$\zeta_{\mathcal{A}_1}(2,1) = 3\beta_3 = (B_{p-3})_p.$$

But we know $\zeta^{\mathcal{S}}_{ш}(2,1) = 3\zeta(2,1)$ (see Exercise 6.1 of Chap. 6) is nonzero on the left-hand side so that if f_{KZ} is injective then $\beta_3 \neq 0$ which implies that $B_{p-3} \not\equiv 0 \pmod p$ for infinitely many primes p, a statement far from proved. Note, however, Lemma 6.4.18 and Thm. 6.4.19 implies that the map in Conjecture 8.5.15 is well-defined and surjective since $\zeta^{\mathcal{S}}(2) = 2\zeta(2)$ so that $f_{\mathrm{KZ}}(\zeta(2)\mathsf{MZ}_{w-2}) = \frac{1}{2}\zeta_{\mathcal{A}_1}(2)\mathsf{FMZ}_{w-2,1} = 0$.

A useful way to compare the structure of FMZV and MZV is that β_k in the finite setting should have weight k and correspond to Riemann zeta value $\zeta(k)$ for odd k. It is conjectured that not every MZV can be expressed using Riemann zeta values, the first such instance should be $\zeta(6,2)$ (or $\zeta(2,6)$) in weight 8. On the FMZV side, it is likely that the first such value is also in weight 8. In Table E.3 except for the two values $\zeta_{\mathcal{A}_1}(1,3,1,3) = 0$ and $\zeta_{\mathcal{A}_1}(1,3,2,2) = (45/2)\beta_3\beta_5$ I expect none of the others can be expressed as a $\mathbb{Q}$ multiple of $\beta_3\beta_5$.

In higher weight cases, there should be linear independence among FMZVs which are expressible as $\mathbb{Q}$-linear combinations of weight k products of Bernoulli numbers, as implied by Conjecture 8.1.5.

As a further support of Conjecture 8.5.14,

Theorem 8.5.16. *For all $3 \leq w \leq 12$, we have*

$$\mathsf{FMZ}_w = \big\langle \zeta_{\mathcal{A}_1}(1,2,a_1,\ldots,a_r) : a_1,\ldots,a_r \in \{2,3\}, a_1+\cdots+a_r = w-3 \big\rangle.$$

Moreover, all the the relations in these weights can be proved by the double shuffle relations of FMZVs defined as in Definition 8.4.4.

Proof. This theorem can be proved with the help of Maple (see Appendix D.1). So we leave it to the interested reader. □

8.5.5 *Finite multiple zeta values of superbity 2*

We now move up to superbity 2, namely, we consider congruences modulo p^2. First, we improve on Thm. 8.5.12.

Theorem 8.5.17. *Suppose $s,t\in\mathbb{N}$ and $w=s+t$ is even. Then*

$$\zeta_{\mathcal{A}_2}(s,t)=p\left[(-1)^t s\binom{w+1}{t}-(-1)^t t\binom{w+1}{s}-w\right]\frac{\beta_{w+1}}{2},$$
$$\zeta^{\star}_{\mathcal{A}_2}(t,s)=p\left[(-1)^t s\binom{w+1}{t}-(-1)^t t\binom{w+1}{s}+w\right]\frac{\beta_{w+1}}{2}.$$

Proof. By the conditions on s and t we know from Thm. 8.2.3 and 8.5.4

$$\zeta_{\mathcal{A}_2}(s)\cdot\zeta_{\mathcal{A}_2}(t)=0,\qquad \zeta_{\mathcal{A}_2}(w)=pw\beta_{w+1}.$$

Therefore by the stuffle relation

$$\zeta_{\mathcal{A}_2}(s,t)+\zeta_{\mathcal{A}_2}(t,s)=\zeta_{\mathcal{A}_2}(s)\cdot\zeta_{\mathcal{A}_2}(t)-\zeta_{\mathcal{A}_2}(w)=-pw\beta_{w+1}. \tag{8.32}$$

Moreover, by superbity 2 reversal relation Eq. (8.8)

$$\begin{aligned}\zeta_{\mathcal{A}_2}(s,t)&=\sum_{1\le j<i<p}\frac{1}{(p-i)^s(p-j)^t}\\&=\zeta_{\mathcal{A}_2}(t,s)+ps\zeta_{\mathcal{A}_2}(t,s+1)+pt\zeta_{\mathcal{A}_2}(t+1,s)\\&=\zeta_{\mathcal{A}_2}(t,s)+p\left[(-1)^{s+1}s\binom{w+1}{t}+(-1)^s t\binom{w+1}{t+1}\right]\beta_{w+1}.\end{aligned}$$

Combined with Eq. (8.32) this completes the proof of the equation for $\zeta_{\mathcal{A}_1}$. Then the $\zeta^{\star}_{\mathcal{A}_1}$ part follows from the identity $\zeta^{\star}_{\mathcal{A}_2}(s,t)=\zeta_{\mathcal{A}_2}(s,t)+\zeta_{\mathcal{A}_2}(w)$ and Thm. 8.5.4. □

So up to weight 4 we have the following table (see [596, Prop. 3.7] for detailed computation in weight 4) Notice that $\zeta_{\mathcal{A}_2}(1)=\zeta_{\mathcal{A}_2}(3)=\zeta_{\mathcal{A}_2}(5)=\zeta_{\mathcal{A}_2}(\{1\}^3)=\zeta_{\mathcal{A}_2}(\{1\}^5)=0$. For higher weight we have the next two theorems.

Theorem 8.5.18. *In weight 5 we have*

$$\zeta^{\star}_{\mathcal{A}_2}(1,3,1)=\zeta_{\mathcal{A}_2}(1,3,1)=0, \tag{8.33}$$
$$\zeta^{\star}_{\mathcal{A}_2}(2,1,2)=\zeta_{\mathcal{A}_2}(2,1,2)=-9\frac{1}{3}p\beta_3^2, \tag{8.34}$$
$$\zeta^{\star}_{\mathcal{A}_2}(2,3)=\zeta_{\mathcal{A}_2}(2,3)=2\zeta^{\star}_{\mathcal{A}_2}(4,1)=2\zeta_{\mathcal{A}_2}(4,1), \tag{8.35}$$
$$2\zeta^{\star}_{\mathcal{A}_2}(2,2,1)-3\zeta_{\mathcal{A}_2}(4,1)=9p\beta_3^2, \tag{8.36}$$

$\zeta_{\mathcal{A}_2}(2)$	$2p\beta_3$	$\zeta_{\mathcal{A}_2}(1,1)$	$-p\beta_3$	$\zeta_{\mathcal{A}_2}(1,2)$	$\zeta_{\mathcal{A}_2}(1,2)$
$\zeta_{\mathcal{A}_2}(2,1)$	$-\zeta_{\mathcal{A}_2}(1,2)$	$\zeta_{\mathcal{A}_2}(4)$	$4p\beta_5$	$\zeta_{\mathcal{A}_2}(1,3)$	$\frac{1}{2}p\beta_5$
$\zeta_{\mathcal{A}_2}(3,1)$	$-\frac{9}{2}p\beta_5$	$\zeta_{\mathcal{A}_2}(1,1,2)$	$3\beta_5$	$\zeta_{\mathcal{A}_2}(1,2,1)$	$-\frac{9}{2}p\beta_5$
$\zeta_{\mathcal{A}_2}(2,1,1)$	$\frac{11}{2}\beta_5$	$\zeta_{\mathcal{A}_2}(\{1\}^4)$	$-\beta_5$		

Table 8.1: FMZV of superbity 2 in weight up to 4.

and

$$2\zeta^\star_{\mathcal{A}_2}(1,1,3) = -2\zeta^\star_{\mathcal{A}_2}(3,1,1) = -2\zeta_{\mathcal{A}_2}(1,1,3) = 2\zeta_{\mathcal{A}_2}(3,1,1) = \zeta_{\mathcal{A}_2}(4,1). \tag{8.37}$$

Proof. First, by applying Eq. (8.29) to $(1,1,3)$ we get

$$\begin{aligned}-\zeta^\star_{\mathcal{A}_2}(3,1,1) = \zeta^\star_{\mathcal{A}_2}(1,1,3) + p\Big(&\zeta_{\mathcal{A}_1}(1,1,1,3) + \zeta_{\mathcal{A}_1}(1,2,3)\\ &+ \zeta_{\mathcal{A}_1}(1,1,4) + \zeta_{\mathcal{A}_1}(1,5)\Big) = \zeta^\star_{\mathcal{A}_2}(1,1,3),\end{aligned} \tag{8.38}$$

which yields the first "=" in Eq. (8.37) by looking up Table E.1 on page 557 and easy computation with Thm. 8.5.12 for weight 6 FMZV. Notice by Eq. (8.20) $\zeta^\star_{\mathcal{A}_2}(1,3,1) = \zeta_{\mathcal{A}_2}(1,3,1)$ and by reversal relations Eq. (8.8) in superbity 2

$$-\zeta_{\mathcal{A}_2}(1,1,3) - \zeta_{\mathcal{A}_2}(3,1,1) = p(\zeta_{\mathcal{A}_1}(2,1,3) + \zeta_{\mathcal{A}_1}(1,2,3) + 3\zeta_{\mathcal{A}_1}(1,1,4))$$

by Table E.2 in Appendix E.2. Thus Eq. (8.38) together with stuffle relations yields

$$\begin{aligned}0 = \zeta_{\mathcal{A}_2}(1,1)\zeta_{\mathcal{A}_2}(3) = \zeta_{\mathcal{A}_2}(1,3,1) + \zeta_{\mathcal{A}_2}(1,1,3)&\\ + \zeta_{\mathcal{A}_2}(3,1,1) + \zeta_{\mathcal{A}_2}(2)\zeta_{\mathcal{A}_2}(3) - \zeta_{\mathcal{A}_2}(5) = \zeta_{\mathcal{A}_2}(1,3,1).&\end{aligned}$$

which gives Eq. (8.33) using Eq. (8.20). Applying Eq. (8.29) to $\mathbf{s} = (2,1,2)$ we get

$$\begin{aligned}-\zeta^\star_{\mathcal{A}_2}(1,3,1) = p\Big(&\zeta_{\mathcal{A}_1}(1,2,1,2) + \zeta_{\mathcal{A}_1}(1,3,2) + \zeta_{\mathcal{A}_1}(1,2,3) + \zeta_{\mathcal{A}_1}(1,5)\Big)\\ &+ \zeta^\star_{\mathcal{A}_2}(2,1,2) = \zeta^\star_{\mathcal{A}_2}(2,1,2) + 3p\beta_3^2\end{aligned}$$

using Table E.1 on page 557. Thus Eq. (8.34) follows immediately from Eq. (8.33).

Turning to the proof of Eq. (8.35) we notice that for any positive integer $k<p$ (setting $h_k=H_{k-1}$)

$$\begin{aligned}
&\frac{\binom{p-1}{k}}{\binom{p-1+k}{k}}=\frac{(p-1)(p-2)\cdots(p-k)}{p(p+1)\cdots(p+k-1)}=\frac{(-1)^k k}{p}\prod_{j=1}^{k}\left(1-\frac{p}{j}\right)\prod_{j=1}^{k-1}\left(1+\frac{p}{j}\right)^{-1}\\
\equiv&\frac{(-1)^k k}{p}\Big(1-pH_k(1)+p^2H_k(1,1)\Big)\Big(1-ph_k(1)+p^2h_k(2)+p^2h_k(1,1)\Big)\\
\equiv&\frac{(-1)^k k}{p}\left(1-2ph_k(1)-\frac{p}{k}+\frac{2p^2}{k}h_k(1)+2p^2h_k(2)+4p^2h_k(1,1)\right).
\end{aligned}\tag{8.39}$$

By Eq. (7.23) of Thm. 7.5.2 we obtain (with abuse of notation $H=H_{p-1}$)

$$\begin{aligned}
&H^{\star}_{p-1}(\{2\}^a,3,\{2\}^b)\\
\equiv&-\frac{2}{p}\sum_{k=1}^{p-1}\frac{1}{k^{2b+2a+2}}\left(1-2ph_k(1)-\frac{p}{k}+\frac{2p^2}{k}h_k(1)+2p^2h_k(2)+4p^2h_k(1,1)\right)\\
&-\frac{4}{p}\sum_{k=1}^{p-1}\frac{h_k(2b+1)}{k^{2a+1}}\left(1-2ph_k(1)-\frac{p}{k}+\frac{2p^2}{k}h_k(1)+2p^2h_k(2)+4p^2h_k(1,1)\right)\\
=&-\frac{2}{p}H(2b+2a+2)-\frac{4}{p}H(2a+1,2b+1)+4H(2b+2a+2,1)\\
&+2H(2b+2a+3)+8H(2a+1,1,2b+1)+8H(2a+1,2b+1,1)\\
&+8H(2a+1,2b+2)+4H(2a+2,2b+1)\\
&-4p\Big[H(2a+2b+3,1)+H(2a+2b+2,2)+2H(2a+2b+2,1,1)]\\
&-8p\Big[H(2a+2,1,2b+1)+H(2a+2,2b+1,1)+H(2a+2,2b+2)\\
&\quad+H(2a+1,2,2b+1)+H(2a+1,2b+1,2)+H(2a+1,2b+3)\\
&\quad+2H(2a+1,1,2b+2)+2H(2a+1,2b+2,1)+2H(2a+1,1,1,2b+1)\\
&\quad+2H(2a+1,1,2b+1,1)+2H(2a+1,2b+1,1,1)\Big].
\end{aligned}$$

Taking $a=1$ and $b=0$ we get (here $H=H_{p-1}$)

$$\begin{aligned}
&H^{\star}_{p-1}(2,3)=-\frac{2}{p}H(4)-\frac{4}{p}H(3,1)+16H(3,1,1)+8H(3,2)+8H(4,1)\\
&-4p\Big[H(5,1)+3H(4,2)+6H(4,1,1)+6H(3,2,1)+6H(3,1,2)+12H(3,1,1,1)\Big].
\end{aligned}\tag{8.40}$$

By stuffle relation and using Eq. (8.33) we see that

$$2\zeta_{\mathcal{A}_2}(3,1,1)+\zeta_{\mathcal{A}_2}(3,2)+\zeta_{\mathcal{A}_2}(4,1)=\zeta_{\mathcal{A}_2}(3,1)\zeta_{\mathcal{A}_2}(1)-\zeta_{\mathcal{A}_2}(1,3,1)=0.$$

On the other hand

$$\frac{1}{p}\Big(H(4)+2H(3,1)\Big)=\frac{1}{p}\Big(H(1)H(3)+H(3,1)-H(1,3)\Big)$$
$$\equiv-\Big(3H(4,1)+H(3,2)+3pH(4,2)+6pH(5,1)\Big)$$

by Eq. (8.8). Thus Eq. (8.40) is reduced to

$$H^{\star}_{p-1}(2,3)\equiv 6H(4,1)+2H(3,2)+6pH(4,2)+12pH(5,1)$$
$$-4p\Big[H(5,1)+3H(4,2)+6H(4,1,1)+6H(3,2,1)+6H(3,1,2)+12H(3,1,1,1)\Big]$$
$$\equiv 6H(4,1)+2H(3)H(2)-2H(2,3)\equiv 6H(4,1)-2H(2,3)$$

by using the reversal relation and then looking up Table E.1 for weight 6 FMZV on page 557. This together with Eq. (8.19) implies Eq. (8.35). Further, Eq. (8.35) shows that

$$\zeta^{\star}_{\mathcal{A}_2}(2,2,1)=\zeta_{\mathcal{A}_2}(2,2,1)+\zeta_{\mathcal{A}_2}(2,3)+\zeta_{\mathcal{A}_2}(4,1)+\zeta_{\mathcal{A}_2}(5)$$
$$=\zeta_{\mathcal{A}_2}(2,2,1)+3\zeta_{\mathcal{A}_2}(4,1).$$

Combining this with Thm. 8.5.4 and Eq. (8.35) we have

$$-6\beta_3^2=\zeta_{\mathcal{A}_2}(2)\zeta_{\mathcal{A}_2}(2,1)=2\zeta_{\mathcal{A}_2}(2,2,1)+\zeta_{\mathcal{A}_2}(4,1)+\zeta_{\mathcal{A}_2}(2,3)+\zeta_{\mathcal{A}_2}(2,1,2)$$
$$=2\zeta^{\star}_{\mathcal{A}_2}(2,2,1)-3\zeta_{\mathcal{A}_2}(4,1)+\zeta_{\mathcal{A}_2}(2,1,2)=2\zeta^{\star}_{\mathcal{A}_2}(2,2,1)-3\zeta_{\mathcal{A}_2}(4,1)-3\beta_3^2$$

by Eq. (8.34). This clearly implies Eq. (8.36).

Finally, the second "=" of Eq. (8.37) is an easy consequence of Eq. (8.31) since $\zeta_{\mathcal{A}_2}(3)=\zeta_{\mathcal{A}_2}(1)=0$. Therefore

$$\zeta^{\star}_{\mathcal{A}_2}(3,1,1)=\zeta_{\mathcal{A}_2}(3,1,1)+\zeta_{\mathcal{A}_2}(3,2)+\zeta_{\mathcal{A}_2}(4,1)+\zeta_{\mathcal{A}_2}(5)$$
$$=-\zeta^{\star}_{\mathcal{A}_2}(3,1,1)-\zeta_{\mathcal{A}_2}(4,1)$$

by Eq. (8.35). So the last "=" of Eq. (8.37) is verified. This finishes the proof of the theorem. □

Similarly we can obtain the following results in weight 6. We leave the detailed computation to the interested reader and provide essentially all the other values in Table E.7.

Theorem 8.5.19. *In weight* 6 *we have*

$$\zeta^{\star}_{\mathcal{A}_2}(1,1,4)=\zeta^{\star}_{\mathcal{A}_2}(4,1,1)=\zeta_{\mathcal{A}_2}(1,1,4)=\zeta_{\mathcal{A}_2}(4,1,1),$$
$$\zeta^{\star}_{\mathcal{A}_2}(1,4,1)=\zeta_{\mathcal{A}_2}(1,4,1)=-2\zeta_{\mathcal{A}_2}(4,1,1)+6p\beta_7,$$
$$\zeta^{\star}_{\mathcal{A}_2}(2,3,1)=\zeta_{\mathcal{A}_2}(1,3,2)=3\zeta_{\mathcal{A}_2}(4,1,1)-\frac{65}{4}p\beta_7,$$

$$\begin{aligned}
\zeta^\star_{\mathcal{A}_2}(1,2,3) = \zeta_{\mathcal{A}_2}(3,2,1) &= -2\zeta_{\mathcal{A}_2}(4,1,1) + \frac{17}{4}p\beta_7,\\
\zeta^\star_{\mathcal{A}_2}(2,1,3) = \zeta_{\mathcal{A}_2}(3,1,2) &= -\zeta_{\mathcal{A}_2}(4,1,1) + 11p\beta_7,\\
\zeta^\star_{\mathcal{A}_2}(1,3,2) = \zeta_{\mathcal{A}_2}(2,3,1) &= 3\zeta_{\mathcal{A}_2}(4,1,1) - \frac{9}{4}p\beta_7,\\
\zeta^\star_{\mathcal{A}_2}(3,2,1) = \zeta_{\mathcal{A}_2}(1,2,3) &= -2\zeta_{\mathcal{A}_2}(4,1,1) - \frac{11}{4}p\beta_7,\\
\zeta^\star_{\mathcal{A}_2}(3,1,2) = \zeta_{\mathcal{A}_2}(2,1,3) &= -\zeta_{\mathcal{A}_2}(4,1,1) + 18p\beta_7.
\end{aligned}$$

Corollary 8.5.20. *The weight* 6 *superbity* 2 *the FMZV space is given by*

$$\mathsf{FMZ}_{6,2} = \langle \zeta_{\mathcal{A}_2}(4,1,1), p\beta_7\rangle.$$

Proof. By Thm. 8.5.4 and Thm. 8.5.17 all depth 1 and 2 values are $\mathbb{Q}$ multiples of $p\beta_7$. All depth 3 values are presented in Thm. 8.5.19. By duality Eq. (8.29) and Table E.2 for weight 7 values (which are all $\mathbb{Q}$ multiples of β_7) we see that all values of larger depths are also linear combinations of $\zeta_{\mathcal{A}_2}(4,1,1)$ and $p\beta_7$. This completes the proof of the corollary. □

Using Maple one can compute FMZV of depth 2 up to weight 9 which are listed in the next theorem. For essentially complete tables of values of weight up to 6 (inclusive) see Tables 8.1, Tables E.6 and Tables E.7.

Theorem 8.5.21. *Setting* $\zeta_2 = \zeta_{\mathcal{A}_2}$ *we have*

$$\begin{aligned}
&\mathsf{FMZ}_{1,2} = \langle 0\rangle, \quad \mathsf{FMZ}_{2,2} = \langle \zeta_2(2)\rangle, \quad \mathsf{FMZ}_{3,2} = \big\langle \zeta_2(1,2)\big\rangle,\\
&\mathsf{FMZ}_{4,2} = \langle \zeta_2(2,2)\rangle, \quad \mathsf{FMZ}_{5,2} = \big\langle \zeta_2(2,3), \zeta_2(1,2,2)\big\rangle,\\
&\mathsf{FMZ}_{6,2} = \big\langle \zeta_2(2,2,2), \zeta_2(1,2,3)\big\rangle,\\
&\mathsf{FMZ}_{7,2} = \left\langle \begin{array}{c} \zeta_2(\{1\}^5,2), \zeta_2(\{1\}^3,4), \zeta_2(1,6),\\ \boldsymbol{\zeta}_2(2,3,1,1), \boldsymbol{\zeta}_2(3,1,1,2), \boldsymbol{\zeta}_2(3,3,1)\end{array}\right\rangle,\\
&\mathsf{FMZ}_{8,2} = \left\langle \begin{array}{c} \zeta_2(\{1\}^6,2), \zeta_2(\{1\}^4,4), \zeta_2(1,1,6), \zeta_2(1,2,5),\\ \boldsymbol{\zeta}_2(5,1,1,1)\end{array}\right\rangle,\\
&\mathsf{FMZ}_{9,2} = \left\langle \begin{array}{c} \zeta_2(\{1\}^7,2), \zeta_2(\{1\}^5,4), \zeta_2(\{1\}^3,6), \zeta_2(1,8), \zeta_2(1,2,6),\\ \boldsymbol{\zeta}_2(1,3,2,3), \boldsymbol{\zeta}_2(2,1,3,3), \boldsymbol{\zeta}_2(2,3,1,3), \boldsymbol{\zeta}_2(1,2,5,1)\end{array}\right\rangle.
\end{aligned}$$

We expect that the bold-faced FMZVs are not really needed for generating the corresponding spaces because of the conjectured relations in Exercise 8.6. In general, we have the following dimension conjecture parallel to Conjecture 3.4.2 for MZVs. See Table 8.2 for numerical support.

Conjecture 8.5.22. *Let w be any positive integer. Set $d_{0,2} = 1$ and $d_{w,2} = \dim \mathsf{FMZ}_{w,2}$. Then $d_{1,2} = 0, d_{2,2} = 1$ and for all $w \geq 3$ we have*

$$d_{w,2} = d_{w-2,2} + d_{w-3,2}.$$

Moreover, for all $w \geq 1$ $\mathsf{FMZ}_{w,2}$ has a basis $\mathfrak{B}_{1,2}(w) \cup \mathfrak{B}_2(w)$ where

$$\mathfrak{B}_{\mathbf{s}}(w) = \{\zeta_{\mathcal{A}_2}(\mathbf{s}, a_1, \ldots, a_r) : a_1, \ldots, a_r = 2 \text{ or } 3, a_1 + \cdots + a_r = w - |\mathbf{s}|\}.$$

8.5.6 *Finite multiple zeta values of superbities ≤ 5*

First, let us look at some numerical data given in Table 8.2.

w	0	1	2	3	4	5	6	7	8	9	10	11	12
$\dim \mathsf{MZ}_w$	1	0	1	1	1	2	2	3	4	5	7	9	12
$\dim \mathsf{FMZ}_{w,1}$	1	0	0	1	0	1	1	1	2	2	3	4	5
$\dim \mathsf{FMZ}_{w,2}$	1	0	1	1	1	2	2	3	4	5	7	9	12
$\dim \mathsf{FMZ}_{w,3}$	1	1	1	2	2	3	4	5	7	9	12	16	21
$\dim \mathsf{FMZ}_{w,4}$	2	1	**1**	3	3	5	6	8	11	14	19	25	
$\dim \mathsf{FMZ}_{w,5}$	2	**1**	**2**	**3**	5	7	9	12	16	21			

Table 8.2: Numerically verified conjectural dimensions of MZ_w and $\mathsf{FMZ}_{w,\ell}$ for $\ell \leq 5$. Bold-faced numbers are all one less than they're supposed to be by the conjectured recurrence formula.

Notice that at every superbity the conjectured recurrence relation is always $f_w = f_{w-2} + f_{w-3}$ in general. The only exceptions occur at superbities at least 4 but only for very small weights where we have "too few" values. For example, at superbity 4 and weight 2 the dimension is supposed to be 2, but because of the trivial relation

$$\zeta_{\mathcal{A}_4}(2) + 2\zeta_{\mathcal{A}_4}(1,1) = \zeta_{\mathcal{A}_2}(1)^2 = 0 \in \mathcal{A}_4$$

the dimension is decreased to only 1. For another example, at superbity $\ell = 5$ or 6 we have another trivial relation

$$6\zeta_{\mathcal{A}_\ell}(1,1,1) + \zeta_{\mathcal{A}_\ell}(1,2) + \zeta_{\mathcal{A}_\ell}(2,1) + \zeta_{\mathcal{A}_\ell}(3) = \zeta_{\mathcal{A}_\ell}(1)^3 = 0.$$

8.5.7 *Finite multiple zeta values of arbitrary depths*

The next result is a kind of sum formula for FMZVs which is an analog of Thm. 6.2.2 for SMZVs. For $w, d, i \in \mathbb{Z}$ with $1 \leq i \leq d \leq w - 1$, set

$$I_{w,d,i} = \{\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d : |\mathbf{s}| = w,\ s_i \geq 2\}.$$

Theorem 8.5.23. *Fix $w, d, i \in \mathbb{Z}$ with $1 \leq i \leq d \leq w-1$ and write $\mathbf{s} = (s_1, \ldots, s_d)$. Then we have*

$$\sum_{\mathbf{s} \in I_{w,d,i}} \zeta_{\mathcal{A}_1}(\mathbf{s}) = (-1)^{i-1}\left(\binom{w-1}{i-1} + (-1)^d\binom{w-1}{d-i}\right)\beta_w, \tag{8.41}$$

$$\sum_{\mathbf{s} \in I_{w,d,i}} \zeta^\star_{\mathcal{A}}(\mathbf{s}) = (-1)^{i-1}\left((-1)^d\binom{w-1}{i-1} + \binom{w-1}{d-i}\right)\beta_w. \tag{8.42}$$

Proof. To save space we let $\mathcal{A} = \mathcal{A}_1$ and

$$S^\star_{w,d,i} = \sum_{\mathbf{s} \in I_{w,d,i}} \zeta^\star_{\mathcal{A}}(\mathbf{s}).$$

throughout this proof. First, by the stuffle relations we get

$$\begin{aligned} 0 = \zeta^\star_{\mathcal{A}}(s_1, \ldots, s_{d-1})\zeta^\star_{\mathcal{A}}(l) &= \sum_{j=1}^{d} \zeta^\star_{\mathcal{A}}(s_1, \ldots, s_{j-1}, l, s_j, \ldots, s_{d-1}) \\ &\quad - \sum_{j=1}^{d-1} \zeta^\star_{\mathcal{A}}(s_1, \ldots, s_{j-1}, s_j + l, s_{j+1}, \ldots, s_{d-1}). \end{aligned}$$

Adding the above equation when $(s_1, \ldots, s_{d-1}, l)$ ranges over $I_{w,d,i}$ we get

$$(d-i)S^\star_{w,d,i} + iS^\star_{w,d,i+1} - (k-d)S^\star_{w,d-1,i} = 0. \tag{8.43}$$

We now prove Eq. (8.42) by backward induction on d. When $d = w-1$ by the v-duality Eq.(8.27), we have

$$S^\star_{w,w-1,i} = \zeta^\star_{\mathcal{A}}\big(\{1\}^{i-1}, 2, \{1\}^{w-i-1}\big) = -\zeta^\star_{\mathcal{A}}(i, w-i) = (-1)^{i-1}\binom{w}{i}\beta_w$$

by Eqs. (8.45) and (8.50). Now assume that Eq. (8.42) holds for d. Then by Eq. (8.43)

$$\begin{aligned} (w-d)S^\star_{w,d-1,i} &= (d-i)S^\star_{w,d,i} + iS^\star_{w,d,i+1} \\ &= (d-i)(-1)^{i-1}\left((-1)^d\binom{w-1}{i-1} + \binom{w-1}{d-i}\right)\beta_w \\ &\quad + i(-1)^i\left((-1)^d\binom{w-1}{i} + \binom{w-1}{d-i-1}\right)\beta_w \\ &= (-1)^{i-1}\left((d-i)(-1)^d\binom{w-1}{i-1} + (w-d+i)\binom{w-1}{d-i-1}\right)\beta_w \\ &\quad + (-1)^i\left((w-i)(-1)^d\binom{w-1}{i-1} + i\binom{w-1}{d-i-1}\right)\beta_w \\ &= (w-d)(-1)^{i-1}\left((-1)^{d-1}\binom{w-1}{i-1} + \binom{w-1}{d-i-1}\right)\beta_w. \end{aligned}$$

Therefore Eq. (8.42) holds for $d-1$. This completes the proof of Eq. (8.42). We leave the proof of Eq. (8.41) to the interested reader since the idea is basically the same. □

Theorem 8.5.24. *When $w < 14$ the upper bound for $\dim \mathsf{FMZ}_{w,1}$ in Table 8.2 holds. Moreover, for all $3 \leq w \leq 13$, we have*

$$\mathsf{FMZ}_{w,1} = \left\langle \zeta_{\mathcal{A}_1}(1, 2, a_1, \ldots, a_r) : a_1, \ldots, a_r \in \{2, 3\}, a_1 + \cdots + a_r = w - 3 \right\rangle.$$

Proof. All these can be proved by the double shuffle relations of FMZVs defined as in Definition 8.4.4. □

8.6 Finite Euler sums

We now turn to the study of FESs. Most of the results for FMZVs generalize to these values although some, such as the duality relation, does not.

8.6.1 *Two reduction formulas*

We now prove two reduction formulas of $H(\mathbf{s})$ for arbitrary composition $\mathbf{s}$ of signed numbers, corresponding to the two cases where $\mathbf{s}$ begins with a positive or a negative number.

Theorem 8.6.1. *Let $a, d \in \mathbb{N}$, $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{D}^d$ and $\mathbf{s}' = (s_1, \ldots, s_{d-1})$. For any prime $p \geq a + 2$ write $H(-) = H_{p-1}(-)$. Then*

$$\begin{aligned} H(\mathbf{s}, a) \equiv & -\frac{1}{a} H\big(\mathbf{s}', (a-1) \oplus s_d\big) \\ & + \sum_{k=1}^{p-1-a} \binom{p-a}{k} \frac{B_k}{p-a} H\big(\mathbf{s}', (k+a-1) \oplus s_d\big) \pmod{p}. \end{aligned} \tag{8.44}$$

Proof. By the definition and Fermat's Little Theorem

$$\begin{aligned} H(\mathbf{s}, a) &\equiv \sum_{p > j_1 > \cdots > j_d} \frac{\operatorname{sgn}(s_1)^{j_1} \cdots \operatorname{sgn}(s_d)^{j_d}}{j_1^{|s_1|} \cdots j_d^{|s_d|}} \sum_{j=1}^{j_d - 1} j^{p-1-a} \\ &\equiv \sum_{k=0}^{p-1-a} \binom{p-a}{k} \frac{B_k}{p-a} \sum_{p > j_1 > \cdots > j_d \geq 1} \frac{\operatorname{sgn}(s_1)^{j_1} \cdots \operatorname{sgn}(s_d)^{j_d}}{j_1^{|s_1|} \cdots j_d^{|s_d|}} j_d^{p-a-k} \end{aligned}$$

which is exactly the right-hand side of Eq. (8.44). □

Theorem 8.6.2. *Let $a, d \in \mathbb{N}$, $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{D}^d$ and $\mathbf{s}' = (s_1, \dots, s_{d-1})$. For any prime $p \geq a+2$ write $H(-) = H_{p-1}(-)$. Then*

$$H(\mathbf{s}, \overline{a}) \equiv \frac{(1-2^{p-a})B_{p-a}}{p-a}\Big(H(\mathbf{s}) - H(\mathbf{s}', \overline{s_d})\Big) - \sum_{k=0}^{p-2-a}\binom{p-1-a}{k}\frac{(1-2^{k+1})B_{k+1}}{k+1}H\big(\mathbf{s}', (k+a)\oplus\overline{s_d}\big) \pmod p,$$

where $\overline{s_d} = t$ if $s_d = \overline{t} \in \overline{\mathbb{N}}$.

Proof. By the definition and Lemma 8.2.5

$$\begin{aligned} H(\mathbf{s}, a) &\equiv \sum_{p>j_1>\cdots>j_d} \frac{\operatorname{sgn}(s_1)^{j_1}\cdots\operatorname{sgn}(s_d)^{j_d}}{j_1^{|s_1|}\cdots j_d^{|s_d|}} \sum_{j=1}^{j_d-1}(-1)^j j^{p-1-a} \\ &\equiv \sum_{k=0}^{p-2-a}\binom{p-1-a}{k}\frac{E_k(0)}{2}\sum_{p>j_1>\cdots>j_d}\frac{\operatorname{sgn}(s_1)^{j_1}\cdots\operatorname{sgn}(s_{d-1})^{j_{d-1}}\operatorname{sgn}(s_d)^{j_d}}{(-1)^{j_d-1}j_1^{|s_1|}\cdots j_{d-1}^{|s_{d-1}|}j_d^{|s_d|+k+a}} \\ &\quad + \frac{E_{p-1-a}(0)}{2}\sum_{p>j_1>\cdots>j_d}\frac{\operatorname{sgn}(s_1)^{j_1}\cdots\operatorname{sgn}(s_d)^{j_d}}{j_1^{|s_1|}\cdots j_d^{|s_d|}}\Big(1-(-1)^{j_d}\Big) \\ &\equiv \frac{E_{a'}(0)}{2}\Big(H(\mathbf{s}) - H(\mathbf{s}', \overline{s_d})\Big) - \sum_{k=0}^{a'-1}\binom{a'}{k}\frac{E_k(0)}{2}H\Big(\mathbf{s}', (k+a)\oplus\overline{s_d}\Big), \end{aligned}$$

where $a' = p-1-a$. The theorem follows from Eq. (8.11) easily. □

8.6.2 *Finite Euler sums of small depths*

In this section we study FESs in depth 2 and 3. All but one case are given by very concise values involving Bernoulli numbers or Euler numbers (which are closely related by the identity Eq. (8.11)).

Lemma 8.6.3. *Let $a, b \in \mathbb{N}$ and prime $p \geq w+3$. Let $w = a+b$. Then*

$$H_{p-1}(\bar{a}, \bar{b}) \equiv \begin{cases} (-1)^a \frac{2^{p-w}-1}{w}\binom{w}{a}B_{p-w} \pmod p, & \text{if } w \text{ is odd;} \\ \frac{(2^a-2)(2^b-2)}{2^{w-1}ab}B_{p-a}B_{p-b} \pmod p, & \text{if } w \text{ is even.} \end{cases}$$

Proof. By the change of variable $k_j \to p - k_j$ for odd indices k_j $(j = 1, 2)$, we have

$$\begin{aligned} &H_{p-1}(\bar{a}, \bar{b}) + H_{p-1}(a, b) \\ &\equiv 2\sum_{p>k_1>k_2>0, 2|k_1, 2|k_2}\frac{1}{k_1^a k_2^b} + 2\sum_{p>k_1>k_2>0, 2\nmid k_1, 2\nmid k_2}\frac{1}{k_1^a k_2^b} \pmod p \\ &\equiv \frac{2}{2^w}H_{\frac{p-1}{2}}(a, b) + \frac{2}{(-2)^w}H_{\frac{p-1}{2}}(b, a) \pmod p. \end{aligned}$$

Similarly,

$$H_{p-1}(a,b)+H_{p-1}(a,b)\equiv H_{\frac{p-1}{2}}(a,b)+(-1)^w H_{\frac{p-1}{2}}(b,a) \\ +(-1)^b H_{\frac{p-1}{2}}(a)H_{\frac{p-1}{2}}(b) \pmod p.$$

By combining the above two congruences we get

$$H_{p-1}(\bar{a},\bar{b})\equiv -(1-2^{1-w})H_{p-1}(a,b)-\frac{(-1)^b}{2^{w-1}}H_{\frac{p-1}{2}}(a)H_{\frac{p-1}{2}}(b) \pmod p.$$

Observe

$$H_{p-1}(a)+H_{p-1}(\bar{a})\equiv \frac{2}{2^w}H_{\frac{p-1}{2}}(a) \pmod p.$$

The lemma follows from Thms. 8.2.3 and 8.2.7 immediately. □

Theorem 8.6.4. *Let $a,b\in\mathbb{N}$ and a prime $p\geq a+b+2$. If the weight $w=a+b$ is odd then we have*

$$\zeta^\star_{\mathcal{A}_1}(a,b)=\zeta_{\mathcal{A}_1}(a,b)=(-1)^a\binom{w}{a}\beta_w, \tag{8.45}$$

$$\zeta_{\mathcal{A}_1}(\overline{a},\overline{b})=(-1)^a(2^{p-w}-1)\binom{w}{a}\beta_w, \tag{8.46}$$

$$\zeta^\star_{\mathcal{A}_1}(\overline{a},\overline{b})=(2^{1-w}-1)\left(2+(-1)^a\binom{w}{a}\right)\beta_w, \tag{8.47}$$

$$\zeta_{\mathcal{A}_1}(\overline{a},b)=\zeta_{\mathcal{A}_1}(a,\overline{b})=(1-2^{1-w})\beta_w, \tag{8.48}$$

$$\zeta^\star_{\mathcal{A}_1}(\overline{a},\overline{b})=\zeta^\star_{\mathcal{A}_1}(a,\overline{b})=(2^{1-w}-1)\beta_w. \tag{8.49}$$

If $a+b$ is even then we have

$$\zeta^\star_{\mathcal{A}_1}(a,b)=\zeta_{\mathcal{A}_1}(a,b)=0, \tag{8.50}$$

$$\zeta^\star_{\mathcal{A}_1}(\overline{a},\overline{b})=\zeta_{\mathcal{A}_1}(\overline{a},\overline{b})=\frac{(2^a-2)(2^b-2)}{2^{w-1}}\beta_a\beta_b. \tag{8.51}$$

Moreover,

$$\zeta_{\mathcal{A}_2}(\overline{2a},\overline{2b})=\left(\frac{(b-a)(1-2^{-2w})}{2w+2}\binom{2w+2}{2a+1}+w\right)p\beta_{2w+1}. \tag{8.52}$$

Proof. Equations (8.45) and (8.50) follow from Thm. 8.5.12. Equations (8.46) and (8.51) follow from Lemma 8.6.3. To prove (8.48), by stuffle relation we see that for $(\alpha,\beta)=(\bar{a},b)$ or $(a,\bar{b})$ we have

$$\begin{aligned}0=\zeta_{\mathcal{A}_1}(\alpha)\zeta_{\mathcal{A}_1}(\beta)&=\zeta_{\mathcal{A}_1}(\alpha,\beta)+\zeta_{\mathcal{A}_1}(\beta,\alpha)+\zeta_{\mathcal{A}_1}(\alpha\oplus\beta)\\&=\zeta_{\mathcal{A}_1}(\alpha,\beta)+\mathrm{sgn}(\alpha\beta)(-1)^w\zeta_{\mathcal{A}_1}(\alpha,\beta)+\zeta_{\mathcal{A}_1}(\alpha\oplus\beta)\\&=2\zeta_{\mathcal{A}_1}(\alpha,\beta)+\zeta_{\mathcal{A}_1}(\alpha\oplus\beta).\end{aligned}$$

Thus

$$\zeta_{\mathcal{A}_1}(a,b)=\zeta_{\mathcal{A}_1}(a,\overline{b})=-\frac{1}{2}\zeta_{\mathcal{A}_1}(\overline{w})=(1-2^{1-w})\beta_w$$

from Thm. 8.2.7. Finally, all $\zeta^{\star}_{\mathcal{A}_1}$ identities are easy consequence of those of $\zeta_{\mathcal{A}_1}$ by the relation $\zeta^{\star}_{\mathcal{A}_1}(\alpha,\beta)=\zeta_{\mathcal{A}_1}(\alpha,\beta)+\zeta_{\mathcal{A}_1}(\alpha\oplus\beta)$.

We now turn to Eq. (8.52). By the definition

$$\begin{aligned}\zeta_{\mathcal{A}_2}(\overline{2a},\overline{2b})&=\sum_{p>j>k>0}\frac{(-1)^{j+k}}{j^{2a}k^{2b}}=\sum_{p>k>j>0}\frac{(-1)^{p-j+p-k}}{(p-j)^{2a}(p-k)^{2b}}\\&=\sum_{p>k>j>0}\frac{(-1)^{k+j}}{k^{2b}j^{2a}}\left(1+\frac{2bp}{k}\right)\left(1+\frac{2ap}{j}\right)\\&=\zeta_{\mathcal{A}_2}(\overline{2b},\overline{2a})+2bp\zeta_{\mathcal{A}_1}(\overline{2b+1},\overline{2a})+2ap\zeta_{\mathcal{A}_1}(\overline{2b},\overline{2a+1}).\end{aligned}$$

By Eq. (8.47)

$$\zeta_{\mathcal{A}_2}(\overline{2a},\overline{2b})=\zeta_{\mathcal{A}_2}(\overline{2b},\overline{2a})+\frac{2(b-a)(1-2^{-2w})}{2w+2}\binom{2w+2}{2a+1}p\beta_{2w+1}. \quad (8.53)$$

On the other hand, we have

$$\zeta_{\mathcal{A}_2}(\overline{2a},\overline{2b})+\zeta_{\mathcal{A}_2}(\overline{2b},\overline{2a})=\zeta_{\mathcal{A}_2}(\overline{2a})\zeta_{\mathcal{A}_2}(\overline{2b})-\zeta_{\mathcal{A}_2}(2w)$$

and therefore by Eq. (8.15) and Eq. (8.25),

$$\zeta_{\mathcal{A}_2}(\overline{2a},\overline{2b})+\zeta_{\mathcal{A}_2}(\overline{2b},\overline{2a})=-\zeta_{\mathcal{A}_2}(2w)=2pw\beta_{2w+1}. \quad (8.54)$$

Equation Eq. (8.52) now follows easily from Eq. (8.53) and Eq. (8.54). □

Turning to FES of depth 3, we first observe that for any $\alpha,\beta,\gamma\in\mathbb{Z}$

$$\begin{aligned}H(\alpha,\beta,\gamma)=&H(\alpha)H(\beta)H(\gamma)-H(\gamma)H(\beta,\alpha)-H(\gamma)H(\beta\oplus\alpha)\\&-H(\gamma,\beta)H(\alpha)-H(\gamma\oplus\beta)H(\alpha)+H(\gamma,\beta,\alpha)\\&+H(\gamma\oplus\beta,\alpha)+H(\gamma,\beta\oplus\alpha)+H(\gamma\oplus\beta\oplus\alpha).\end{aligned}$$

This can be easily checked by the stuffle relations but the idea is hidden in the general framework set up by Hoffman [280]. Combining with the reversal relations one can obtain the following results without much difficulty. We leave its proof to the interested reader.

Theorem 8.6.5. *Let $a,b,c\in\mathbb{N}$. If the weight $w=a+b+c$ is even then*

$$2\zeta_{\mathcal{A}_1}(a,\overline{b},c)=\zeta_{\mathcal{A}_1}(\overline{b+c},a)+\zeta_{\mathcal{A}_1}(c,\overline{a+b}), \quad (8.55)$$

$$2\zeta_{\mathcal{A}_1}(\overline{a},b,c)=-\zeta_{\mathcal{A}_1}(\overline{a})\zeta_{\mathcal{A}_1}(c,b)+\zeta_{\mathcal{A}_1}(c,\overline{a+b})+\zeta_{\mathcal{A}_1}(b+c,\overline{a}), \quad (8.56)$$

$$\begin{aligned}2\zeta_{\mathcal{A}_1}(\overline{a},\overline{b},\overline{c})=&-\zeta_{\mathcal{A}_1}(\overline{c})\zeta_{\mathcal{A}_1}(\overline{b},\overline{a})-\zeta_{\mathcal{A}_1}(\overline{c},\overline{b})\zeta_{\mathcal{A}_1}(\overline{a})\\&+\zeta_{\mathcal{A}_1}(b+c,\overline{a})+\zeta_{\mathcal{A}_1}(\overline{c},a+b). \quad (8.57)\end{aligned}$$

If the weight w is odd then

$$2\zeta_{\mathcal{A}_1}(\overline{a},\overline{b},c) = \zeta_{\mathcal{A}_1}(c,a+b) + \zeta_{\mathcal{A}_1}(\overline{b+c},\overline{a}) - \zeta_{\mathcal{A}_1}(\overline{a})\zeta_{\mathcal{A}_1}(c,\overline{b}), \tag{8.58}$$

$$\begin{aligned}2\zeta_{\mathcal{A}_1}(\overline{a},b,\overline{c}) = &-\zeta_{\mathcal{A}_1}(\overline{c})\zeta_{\mathcal{A}_1}(b,\overline{a}) - \zeta_{\mathcal{A}_1}(\overline{c},b)\zeta_{\mathcal{A}_1}(\overline{a})\\ &+\zeta_{\mathcal{A}_1}(\overline{a+b+c}) + \zeta_{\mathcal{A}_1}(\overline{c},\overline{a+b}).\end{aligned} \tag{8.59}$$

Because of the reversal relations when the weight is even there remains essentially only one more case to consider in depth 3. This is given by the next result.

Theorem 8.6.6. *Let a,b,c be positive integers such that $w := a+b+c$ is even. Then for any prime $p \geq w+3$ we have*

$$\begin{aligned}H(\overline{a},\overline{b},c) \equiv &-\sum_{k=2}^{p-w+1}\binom{p-a}{p-w-k+1}\binom{k+c-1}{c}\frac{(1-2^k)B_kB_{p-w-k+1}}{ak}\\ &-\sum_{k=p+1-b-c}^{p-c}\binom{p-a}{2p-w-k}\binom{k+c-1}{c}\frac{(1-2^k)B_kB_{2p-w-k}}{ak}\\ &-\frac{(1-2^{p-c})(1-2^{p-a-b})B_{p-a-b}B_{p-c}}{(a+b)c}.\end{aligned}$$

Proof. The proof is basically a repeated application of Thm. 8.6.1. But we spell out all the details below because there are some subtle details that we need to attend to.

By Eq. (8.2) and Fermat's Little Theorem we have modulo p

$$\begin{aligned}H(\overline{a},\overline{b},c) &\equiv \sum_{i=1}^{p-1}\frac{(-1)^i}{i^c}\sum_{j=1}^{i-1}(-1)^j j^{p-1-b}\sum_{k=1}^{j-1}k^{p-1-a}\\ &\equiv \sum_{i=1}^{p-1}\frac{(-1)^i}{i^c}\sum_{k=0}^{p-1-a}\binom{p-a}{k}\frac{B_k}{p-a}\sum_{j=1}^{i-1}(-1)^j j^{\rho(k)+p-a-b-k},\end{aligned}$$

where $\rho(k)=0$ if $k<p-a-b$ and $\rho(k)=p-1$ if $k\geq p-a-b$ (to make sure all exponents are positive in the sum of the second line above). By Lemma 8.2.5

$$\begin{aligned}H(\overline{a},\overline{b},c) &\equiv \sum_{k=0}^{p-1-a}\binom{p-a}{k}\frac{B_k}{p-a}\sum_{r=0}^{n}\binom{n}{r}\sum_{i=1}^{p-1}(-1)^i i^{n-r-c}F_{n,i,r}\\ &\equiv -\sum_{k=0}^{p-1-a}\binom{p-a}{k}\sum_{r=1}^{n-1}\binom{n}{r}\frac{(1-2^{r+1})B_kB_{r+1}}{(p-a)(r+1)}\sum_{i=1}^{p-1}i^{n-c-r}\\ &\quad+\sum_{k=0}^{p-1-a}\binom{p-a}{k}\frac{(1-2^{n+1})B_kB_{n+1}}{(p-a)(n+1)}\sum_{i=1}^{p-1}((-1)^i-1)i^{-c},\end{aligned}$$

where $n = \rho(k) + p - a - b - k$. Here we have used the fact that when $r = 0$ we have $F_{n,i,r} = (-1)^{i-1}$ and thus the inner sum is $\sum_{i=1}^{p-1} i^{n-c} \equiv 0 \pmod{p}$ except when $p-1|n-c$, *i.e.*, except when $n = c$ and $k = p - w$. But then $B_k = 0$ since w is even by assumption. Thus

$$H(\overline{a}, \overline{b}, c) \equiv \sum_{\substack{0 \le k < p-a \\ c < n}} \binom{p-a}{k} \binom{n}{n-c} \frac{(1-2^{n-c+1})B_k B_{n-c+1}}{(p-a)(n-c+1)} + H(\overline{c}) \sum_{k=0}^{p-1-a} \binom{p-a}{k} \frac{(1-2^{n+1})B_k B_{n+1}}{(p-a)(n+1)}.$$

Now if c is even then $H(\overline{c}) \equiv 0 \pmod{p}$. So we may assume c is odd in the last line above. Then $k + n + 1$ is always odd so that $B_k B_{n+1} \neq 0$ if and only if $k = 1$ and $n = p - a - b - 1$. Hence

$$H(a, \overline{b}, \overline{c}) \equiv \sum_{k=0}^{p-w-1} \binom{p-a}{k} \binom{p-a-b-k}{p-w-k} \frac{(1-2^{p-w-k+1})B_k B_{p-w-k+1}}{(p-a)(p-w-k+1)} + \sum_{k=p-a-b}^{p-1-a} \binom{p-a}{k} \binom{2p-1-a-b-k}{2p-1-w-k} \frac{(1-2^{2p-w-k})B_k B_{2p-w-k}}{(p-a)(2p-w-k)} - \frac{(1-2^{p-c})(1-2^{p-a-b})B_{p-a-b}B_{p-c}}{(a+b)c}.$$

After substitutions $k \to p - w + 1 - k$ in the first sum and $k \to 2p - w - k$ in the second sum the theorem follows immediately from Thm. 8.2.7, □

Remark 8.6.7. The condition $p \geq w+3$ in Thm. 8.6.6 cannot be weakened since

$$H(\overline{3}, \overline{2}, 1) \equiv \mathrm{RHS} + 5 \not\equiv \mathrm{RHS} \pmod{7}.$$

8.6.3 *Finite Euler sums of superbity 1 and 2*

According the Conjecture 14.2.3 the dimensions of Euler sums (see Sec. 14.2) are given by the Fibonacci numbers. From what we have found so far we get the following table. Let ES be the $\mathbb{Q}$-space spanned by of alternating Euler sums.

Here, we provide $\dim \mathsf{ES}_w$ for comparison purpose. Recall that these have been proved by Deligne [155] under the assumption of a variant Grothendieck period conjecture. Without assuming the conjecture he proved these numbers provide upper bounds.

w	0	1	2	3	4	5	6	7
$\dim \mathsf{ES}_w$	1	1	2	3	5	8	13	21
$\dim \mathsf{FES}_{w,1}$	1	1	1	2	3	5	8	13
$\dim \mathsf{FES}_{w,2}$	1	1	2	4	7	12	20	33

Table 8.3: Numerically verified conjectural dimensions of ES_w and $\mathsf{FES}_{w,\ell}$ for $\ell \leq 2$.

We conclude this section by a conjecture analogous to Conjecture 8.5.15. Define the symmetrized version of the alternating Euler sums

$$\zeta^{\mathcal{S}}_{ш}(s_1, s_2, \ldots, s_d) = \sum_{i=0}^{d} \left(\prod_{j=1}^{i} (-1)^{s_j} \operatorname{sgn}(s_j) \right) \zeta_{ш}(s_i, \ldots, s_1) \zeta_{ш}(s_{i+1}, \ldots, s_d),$$

$$\zeta^{\mathcal{S}}_{*}(s_1, s_2, \ldots, s_d) = \sum_{i=0}^{d} \left(\prod_{j=1}^{i} (-1)^{s_j} \operatorname{sgn}(s_j) \right) \zeta_{*}(s_i, \ldots, s_1) \zeta_{*}(s_{i+1}, \ldots, s_d),$$

where, similar to the MZV case, $\zeta_{ш}$ and ζ_* are regularized values given by Prop. 13.3.8. The sign in the above is necessary since we have the reversal relation for finite ES given as follows:

Proposition 8.6.8. *Let* $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{D}^d$. *We have*

$$(-1)^{\mathrm{wt}(\mathbf{s})} \operatorname{sgn}(\mathbf{s}) \zeta_{\mathcal{A}_2}(s_d, \ldots, s_1) = \zeta_{\mathcal{A}_2}(\mathbf{s}) + p \sum_{j=1}^{d} |s_j| \zeta_{\mathcal{A}_2}(\mathbf{s} \oplus \mathbf{e}_j),$$

where the binary operation $\oplus$ *is defined by Eq.* (7.1) *in Definition* 7.1.

Conjecture 8.6.9. *There is an isomorphism*

$$f : \mathsf{FES}_{w,1} \longrightarrow \frac{\mathsf{ES}_w}{\zeta(2)\mathsf{ES}_{w-2}},$$
$$\zeta_{\mathcal{A}_1}(\mathbf{s}) \longmapsto \zeta^{\mathcal{S}}_{ш}(\mathbf{s}).$$

Furthermore,

- $\zeta^{\mathcal{S}}_{ш}(\mathbf{s})$ *and* $\zeta^{\mathcal{S}}_{*}(\mathbf{s})$ *are all finite, and*
- $\zeta^{\mathcal{S}}_{ш}(\mathbf{s}) - \zeta^{\mathcal{S}}_{*}(\mathbf{s}) \in \zeta(2)\mathsf{ES}_{w-2}$ *for all* $|\mathbf{s}| = w$.

We have checked up to weight 5 that the map is well-defined under Conjecture 8.1.6 concerning the $\mathcal{A}_1$-Bernoulli numbers. We also verified the second part numerically up to weight 5.

8.6.4 *Two applications of MHS binomial identities*

In this section we derive a equations in $\mathcal{A}_1$ using the MHS binomial identities in Chap. 7. The first result is an application of the Two-three formula (see Remark 7.5.3).

Theorem 8.6.10. *Let a, b be non-negative integers. Then*

$$\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 3, \{2\}^b) = \frac{2(b-a)}{a+1}\binom{2a+2b+3}{2a+1}\beta_{2a+2b+3}, \tag{8.60}$$

$$\zeta_{\mathcal{A}_1}(\{2\}^a, 3, \{2\}^b) = \frac{2(-1)^{a+b}(a-b)}{a+1}\binom{2a+2b+3}{2a+1}\beta_{2a+2b+3}. \tag{8.61}$$

Proof. By Thm. 7.5.2 we have

$$\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 3, \{2\}^b) = \left(2\sum_{k=1}^{p-1}\frac{(-1)^{k-1}\binom{p-1}{k}}{k^{2a+2b+3}\binom{p-1+k}{k}} - 4\sum_{k=1}^{p-1}\frac{H_{k-1}(2b+1)(-1)^k\binom{p-1}{k}}{k^{2a+2}\binom{p-1+k}{k}}\right)_p.$$

Since

$$\begin{aligned}\frac{\binom{p-1}{k}}{\binom{p-1+k}{k}} &= \frac{(p-1)(p-2)\cdots(p-k)}{p(p+1)\cdots(p+k-1)}\\ &= \frac{(-1)^k k}{p}\prod_{j=1}^{k}\left(1-\frac{p}{j}\right)\prod_{j=1}^{k-1}\left(1+\frac{p}{j}\right)^{-1}\\ &\equiv \frac{(-1)^k k}{p}(1-pH_k(1))(1-pH_{k-1}(1))\\ &\equiv \frac{(-1)^k k}{p}\left(1-2pH_{k-1}(1)-\frac{p}{k}\right) \pmod p,\end{aligned} \tag{8.62}$$

we obtain $(a' = 2a+1, b' = 2b+1)$

$$\begin{aligned}&\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 3, \{2\}^b)\\ &\quad= \left(-\frac{2}{p}\sum_{k=1}^{p-1}\left(\frac{1}{k^{a'+b'}}+\frac{2H_{k-1}(b')}{k^{a'}}\right)\left(1-2pH_{k-1}(1)-\frac{p}{k}\right)\right)_p\\ &= -\frac{2}{p}\zeta_{\mathcal{A}_2}(2a+2b+2) - \frac{4}{p}\zeta_{\mathcal{A}_2}(a',b') + 4\zeta_{\mathcal{A}_1}(a'+b',1) + 4\zeta_{\mathcal{A}_1}(a'+1,b')\\ &\qquad + 8\zeta_{\mathcal{A}_1}(a',1,b') + 8\zeta_{\mathcal{A}_1}(a',b',1) + 8\zeta_{\mathcal{A}_1}(a',b'+1).\end{aligned}$$

Here we have applied the stuffle relation
$H_{k-1}(2b+1)H_{k-1}(1) = H_{k-1}(1,2b+1) + H_{k-1}(2b+1,1) + H_{k-1}(2b+2)$,
Now Eq. (8.60) follows from Thms. 8.5.4, 8.5.13, 8.5.17 easily. To prove Eq. (8.61), we take $\mathbf{s} = (\{2\}^a, 3, \{2\}^b)$ in the concatenation relation Eq. (8.18) then one of $\mathbf{s}_j = (\{2\}^\ell)$ for some positive ℓ and therefore $\zeta_{\mathcal{A}_1}(\mathbf{s}_j) = 0$. Thus

$$(-1)^{a+b}\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 3, \{2\}^b) = \zeta_{\mathcal{A}_1}(\{2\}^b, 3, \{2\}^a). \tag{8.63}$$

On the other hand, by reversal relation Eq. (8.5)

$$\zeta_{\mathcal{A}_1}(\{2\}^a, 3, \{2\}^b) = -\zeta_{\mathcal{A}_1}(\{2\}^b, 3, \{2\}^a). \tag{8.64}$$

Now from Eqs. (8.63) and (8.64) we get the required congruence and the proof is complete. □

The second result is an application of the Two-one formula.

Theorem 8.6.11. *Let $a, b \in \mathbb{N}_0$ and $w = 2a + 2b + 1$. Then*

$$\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 1, \{2\}^b) = \frac{4(b-a)(1-2^{1-w})}{w+1}\binom{w+1}{2a+1}\beta_w, \tag{8.65}$$

$$\zeta_{\mathcal{A}_1}(\{2\}^a, 1, \{2\}^b) = \frac{4(a-b)(1-2^{1-w})}{(-1)^{a+b}(w+1)}\binom{w+1}{2a+1}\beta_w. \tag{8.66}$$

Proof. Equation Eq. (8.66) follows easily from the Eq. (8.65) by using the concatenation relation Eq. (8.18). For the latter, we begin by considering the case $a \geq 1$. By Eq. (7.26) with $n = p-1$ and congruence Eq. (8.62) we have

$$\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 1, \{2\}^b) = \left(-\frac{2}{p}\sum_{k=1}^{p-1}\frac{1}{k^{2a+2b}}\Big(1 - 2pH_{k-1}(1) - \frac{p}{k}\Big)\right)_p$$
$$- \frac{4}{p}\sum_{k=1}^{p-1}\frac{H_{k-1}(\overline{2b})(-1)^k}{k^{2a}}\Big(1 - 2pH_{k-1}(1) - \frac{p}{k}\Big).$$

Applying the stuffle relation

$$H_{k-1}(1)H_{k-1}(\overline{2b}) = H_{k-1}(1,\overline{2b}) + H_{k-1}(\overline{2b},1) + H_{k-1}(\overline{2b+1}),$$

we get

$$\zeta^\star_{\mathcal{A}_1}(\{2\}^a, 1, \{2\}^b) = -\frac{2}{p}\zeta_{\mathcal{A}_2}(2a+2b) - \frac{4}{p}\zeta_{\mathcal{A}_2}(\overline{2a},\overline{2b}) + 4\zeta_{\mathcal{A}_1}(2a+2b,1)$$
$$+ 4\zeta_{\mathcal{A}_1}(\overline{2a+1},\overline{2b}) + 8\zeta_{\mathcal{A}_1}(\overline{2a},1,\overline{2b}) + 8\zeta_{\mathcal{A}_1}(\overline{2a},\overline{2b},1) + 8\zeta_{\mathcal{A}_1}(\overline{2a},\overline{2b+1}).$$

Now we can prove the case $a \geq 1$ by applying Eqs. (8.25), (8.30), (8.47), (8.52), (8.58), and (8.59).

The case $a = 0$ can be proved similarly (see Exercise 8.3). This finishes the proof of the theorem. □

Corollary 8.6.12. *Let p be a prime greater than 7. Then*

$$\zeta^\star_{\mathcal{A}_1}(2,2,2,1) = -\frac{189}{16}\beta_7, \quad \zeta^\star_{\mathcal{A}_1}(2,2,2,2,1) = -\frac{255}{16}\beta_9.$$

Proof. Both are consequences of Thm. 8.6.11 with $(a,b) = (0,3)$ and $(a,b) = (0,4)$, respectively. □

Proposition 8.6.13. *Let $a, b \in \mathbb{N}_0$, $a' = 2a+1$ and $b' = 2b+3$. Then we have*

$$\begin{aligned}\zeta^\star_{\mathcal{A}_1}(\{2\}^a,3,\{2\}^b,3) &= 2\zeta_{\mathcal{A}_1}(a',1,b'+1) - 6\zeta_{\mathcal{A}_1}(b'+1,a',1)\\ &+(8a-6)\zeta_{\mathcal{A}_1}(a'+1,b',1) + 16\zeta_{\mathcal{A}_1}(a'+b',1,1)\\ &+(8b+14)\zeta_{\mathcal{A}_1}(a',b'+1,1) - 16\zeta_{\mathcal{A}_1}(1,a',b',1)\\ &-2\zeta_{\mathcal{A}_1}(a'+1,1,b') - 2\zeta_{\mathcal{A}_1}(b',a'+1,1) + 4\zeta_{\mathcal{A}_1}(a',b',2).\end{aligned} \tag{8.67}$$

Proof. Using Eq. (8.39) we thus have

$$\begin{aligned}&\zeta^\star_{\mathcal{A}_1}(\{2\}^a,3,\{2\}^b,3)\\ &= \frac{-2}{p}\sum_{k=1}^{p-1} C_{n,k}\left\{\frac{1}{k^{a'+b'+1}} + \frac{2H_{k-1}(1)}{k^{a'+b'}} + \frac{2H_{k-1}(b'+1)}{k^{a'}} + \frac{4H_{k-1}(b',1)}{k^{a'}}\right\},\end{aligned}$$

where

$$C_{n,k} = 1 - 2pH_{k-1}(1) - \frac{p}{k}.$$

Hence (setting $H = H_{p-1}$)

$$\begin{aligned}\zeta^\star_{\mathcal{A}_1}(\{2\}^a,3,\{2\}^b,3) = \frac{-2}{p}&\Big(H(a'+b'+1) + 2H(a'+b',1)\\ &+2H(a',b'+1) + 4H(a',b',1)\Big)\\ +4&\Big(H(a'+b'+1,1) + 2H(a'+b',1*1)\\ &+2H(a',1*(b'+1)) + 4H(a',1*(b',1))\Big)\\ +2&\Big(H(a'+b'+1) + 2H(a'+b'+1,1)\\ &+2H(a'+1,b'+1) + 4H(a'+1,b',1)\Big),\end{aligned}$$

where $*$ is the stuffle operation. By stuffle relation we get

$$\begin{aligned}&H(a'+b',1) + H(a',b'+1) + 2H(a',b',1)\\ &\qquad = H(b')H(a',1) + H(a',b',1) - H(b',a',1) - H(a',1,b').\end{aligned}$$

By concatenation relation, when x and z are odd integers we get

$$\begin{aligned}\zeta_{\mathcal{A}_1}(x,y,z) &= \zeta^\star_{\mathcal{A}_2}(z,y,x) = \zeta_{\mathcal{A}_2}(z,y,x) + \zeta_{\mathcal{A}_2}(z+y,x) + \zeta_{\mathcal{A}_2}(z,y+x)\\ &= -\zeta_{\mathcal{A}_1}(x,y,z) + \zeta_{\mathcal{A}_2}(z+y,x) + \zeta_{\mathcal{A}_2}(z,y+x)\\ &\quad - p[x\zeta_{\mathcal{A}_1}(x+1,y,z) + y\zeta_{\mathcal{A}_1}(x,y+1,z) + z\zeta_{\mathcal{A}_1}(x,y,z+1)].\end{aligned}$$

So

$$\begin{aligned}2\zeta_{\mathcal{A}_1}(a',b',1) &= +\zeta_{\mathcal{A}_2}(a'+b',1) + \zeta_{\mathcal{A}_2}(a',4)\\ &\quad -p[(a')\zeta_{\mathcal{A}_1}(a'+1,b',1) + (b')\zeta_{\mathcal{A}_1}(a',b'+1,1) + \zeta_{\mathcal{A}_1}(a',b',2)],\\ 2\zeta_{\mathcal{A}_1}(b',a',1) &= +\zeta_{\mathcal{A}_2}(a'+b',1) + \zeta_{\mathcal{A}_2}(b',a'+1)\\ &\quad -p[(b')\zeta_{\mathcal{A}_1}(b'+1,a',1) + (a')\zeta_{\mathcal{A}_1}(b',a'+1,1) + \zeta_{\mathcal{A}_1}(b',a',2)],\\ 2\zeta_{\mathcal{A}_1}(a',1,b') &= +\zeta_{\mathcal{A}_2}(a'+1,b') + \zeta_{\mathcal{A}_2}(a',b'+1)\\ &\quad -p[(a')\zeta_{\mathcal{A}_1}(a'+1,1,b') + \zeta_{\mathcal{A}_1}(a',2,b') + (b')\zeta_{\mathcal{A}_1}(a',1,b'+1)].\end{aligned}$$

We thus get modulo p^2

$$\begin{aligned}&-2H(a'+b',1) - 2H(a',b'+1) - 4H(a',b',1)\\ &= -p[(b')\zeta_{\mathcal{A}_1}(b'+1,a',1) + (a')\zeta_{\mathcal{A}_1}(b',a'+1,1) + \zeta_{\mathcal{A}_1}(b',a',2)]\\ &\quad -p[(a')\zeta_{\mathcal{A}_1}(a'+1,1,b') + \zeta_{\mathcal{A}_1}(a',2,b') + (b')\zeta_{\mathcal{A}_1}(a',1,b'+1)]\\ &\quad +p[(a')\zeta_{\mathcal{A}_1}(a'+1,b',1) + (b')\zeta_{\mathcal{A}_1}(a',b'+1,1) + \zeta_{\mathcal{A}_1}(a',b',2)].\end{aligned}$$

since $\zeta_{\mathcal{A}_2}(b',a'+1)+\zeta_{\mathcal{A}_2}(a'+1,b') = \zeta_{\mathcal{A}_2}(a'+1)\zeta_{\mathcal{A}_2}(b') - \zeta_{\mathcal{A}_2}(a'+b'+1) = 0$. Simplifying we will finally arrive at Eq. (8.67). □

8.6.5 *Finite Euler sums of arbitrary depths*

The signed version of Thm. 8.5.1, *i.e.*, the case for the homogeneous FESs, is considerably more involved, as can be seen already in Thm. 8.2.7 for depth 1. To state it we denote by $\mathrm{Odd}(d)$ the subset of odd partitions $\Lambda = (\lambda_1,\dots,\lambda_r)$ of d (*i.e.*, $\sum_{j=1}^r \lambda_j = d$ and $\lambda_j > 0$ are odd for all j) and define

$$c(\Pi) = \prod_{j=1}^r (\lambda_j - 1)! \quad \text{and} \quad \tilde{c}(\Pi) = (-1)^{d-r} \prod_{j=1}^r (\lambda_j - 1)!.$$

Theorem 8.6.14. *Let $s,d \in \mathbb{N}$. For an odd partition $\Lambda = (\lambda_1,\dots,\lambda_r) \in \mathrm{Odd}(d)$ we put $\zeta_{\mathcal{A}_1}(\overline{s};\Lambda) = \prod_{i=1}^r \zeta_{\mathcal{A}_1}(\oplus_{j=1}^{\lambda_i}\overline{s})$ (notice $\oplus_{j=1}^{\lambda}\overline{s}$ equals $s\lambda$ for even λ and $\overline{s\lambda}$ for odd λ). Then*

$$\begin{aligned}\zeta_{\mathcal{A}_1}\big(\{\overline{s}\}^d\big) &= \frac{1}{d!}\sum_{\Lambda\in\mathrm{Odd}(d)} c(\Pi)\,\zeta_{\mathcal{A}_1}(\overline{s};\Lambda),\\ \zeta^\star_{\mathcal{A}_1}\big(\{\overline{s}\}^d\big) &= \frac{1}{d!}\sum_{\Lambda\in\mathrm{Odd}(d)} \tilde{c}(\Pi)\,\zeta_{\mathcal{A}_1}(\overline{s};\Lambda).\end{aligned} \tag{8.68}$$

In particular, if s is even then both sums in Eq. (8.68) *are* 0 *in* $\mathcal{A}_1$. *If $s > 1$ is odd then both sums in Eq.* (8.68) *are $\mathbb{Q}$-linear combinations of $\beta_{\lambda_1 s}\cdots\beta_{\lambda_r s}$ for odd partitions $\Lambda = (\lambda_1,\ldots,\lambda_r)$. If $s = 1$ then Eq.* (8.68) *is a $\mathbb{Q}$-linear combination of $\mathrm{q}_2^{\ell}\beta_{\lambda_{\ell+1}}\cdots\beta_{\lambda_r}$ for odd partitions $\lambda = (\{1\}^{\ell},\lambda_{\ell+1},\ldots,\lambda_r)$ ($\lambda_j > 1$ for all $j > \ell$), where $\mathrm{q}_2 = ((2^{p-1}-1)/p)_{p\in\mathcal{P}}$ is an $\mathcal{A}_1$-Fermat quotient.*

Proof. Apply Thm. 7.4.1 and then use Thm. 8.2.7. All are straightforward except in the case $s = 1$ we need the famous Claussen-von Staudt Theorem which says $p\beta_1 = -1$ in $\mathcal{A}_1$ (see [303, p. 233]). □

For example, by [518, Thm. 2.1] and [509, Thm. 5.2] we have

$$\zeta_{\mathcal{A}_3}(\bar{1}) = -\zeta_{\mathcal{A}_3}(1) + H_{\frac{p-1}{2}}(1) = -2\mathrm{q}_2 + p\mathrm{q}_2^2 - \frac{2}{3}p^2\mathrm{q}_2^3 - \frac{3}{4}p^2\beta_3. \quad (8.69)$$

Observe that $\mathrm{Odd}(2) = \{(1,1)\}, \mathrm{Odd}(3) = \{(1,1,1),(3)\}, \mathrm{Odd}(4) = \{(1,1,1,1),(1,3)\}$. It is obvious that $c_{(1,\cdots,1)} = 1$, $c_{(3)} = 2$, $c_{(1,3)} = 8$, $c_{(1,1,3)} = 20$, and $c_{(1,1,1,3)} = c_{(3,3)} = 40$, which implies that

$$\zeta_{\mathcal{A}_1}\big(\{\bar{1}\}^2\big) = 2\mathrm{q}_2^2, \quad \zeta_{\mathcal{A}_1}\big(\{\bar{1}\}^3\big) = -\frac{4}{3}\mathrm{q}_2^3 - \frac{1}{2}\beta_3, \quad \zeta_{\mathcal{A}_1}\big(\{\bar{1}\}^4\big) = \frac{2}{3}\mathrm{q}_2^4 + \mathrm{q}_2\beta_3,$$
$$\zeta_{\mathcal{A}_1}\big(\{\bar{1}\}^5\big) = -\frac{4}{15}\mathrm{q}_2^5 - \mathrm{q}_2^2\beta_3, \quad \zeta_{\mathcal{A}_1}\big(\{\bar{1}\}^6\big) = \frac{4}{45}\mathrm{q}_2^6 + \frac{2}{3}\mathrm{q}_2^3\beta_3 + \frac{1}{8}\beta_3^2.$$

The examples above are closely related to Conjecture 8.1.5.

8.7 Historical notes

The following identity is one of the series related to π that Ramanujan listed in [464]:

$$\sum_{k=0}^{\infty}(4k+1)\binom{-1/2}{k}^3 = \frac{2}{\pi}.$$

This was actually discovered earlier by Bauer in 1859. In his seminal paper [266] of 1997 van Hamme studied the $(p-1)$-st partial sums of many such series for odd primes p and made the following conjecture:

$$\sum_{k=0}^{p-1}(4k+1)\binom{-1/2}{k}^3 \equiv \left(\frac{-1}{p}\right)p \pmod{p^3}.$$

This was proved by Mortenson [417] using the Gamma function and the hypergeometric functions. Since then this kind of congruences has been said of van Hamme type.

The Wolstenholme's Theorem has been improved by Glaisher [228] to

$$\sum_{k=1}^{p-1}\frac{1}{k}\equiv -\frac{1}{3}p^2B_{p-3}\pmod{p^3},\ \text{for all prime } p>3, \tag{8.70}$$

and by Z.-W. Sun [509] as follows:

$$\sum_{k=1}^{p-1}\frac{1}{k}\equiv p^2\left(\frac{2B_{p-3}}{p-3}-\frac{B_{2p-4}}{2p-4}\right)\pmod{p^4},\ \text{for all prime } p>5. \tag{8.71}$$

Ever since Wolstenholme published his work on the congruences Eq. (8.1) a plethora of generalizations are given along various approaches, the majority of which concern with changing to composite moduli or changing the summation ranges, see, for example, the work by Leudesdorf [370], Glaisher [227, 228], Vandiver [538], Bayat [40], Pan [452] and so on. For a more complete history please refer to the survey article [411] by Mestrovic. Virtually none of these work, however, treated multiple variable cases. In an unpublished work R. Bruck seems to be the first who studied the congruence properties of the homogeneous multiple harmonic sums.

Generalizing Wolstenholme's Theorem to the more general MHSs (*i.e.*, the partial sums of MZVs) was the initial motivation of [589, 596] by the author starting around 2002. For example, Thm. 8.5.12, 8.5.13, and 8.5.17 were proved in [596] as Thm. 3.1, 3.5, and 3.2, respectively. A little later, Hoffman independently studied these sums by his powerful algebraic setup using word algebras. In order to find a suitable theoretical framework in which this kind of results can be organized and further investigated, the author first defined in [606] an adele-like structure essentially the same as Eq. (8.3) which was later defined by Kaneko and Zagier who call it poorman's adele when the superbity is 1.

The first systematic study of congruences involving alternating MHSs is carried out in the joint work of [520]. We recast these as finite alternating Euler sums in this chapter due to their close relations to the classical alternating Euler sums.

Note that [553] contains one more statement which is equivalent to the two congruences in Wolstenholme's Thm. 8.2.1:

$$\binom{2p-1}{p-1}\equiv 1\pmod{p^3},\quad \textit{i.e.},\quad \binom{2p}{p}\equiv 2\pmod{p^3}.$$

Wolstenholme's Theorem was rediscovered by Gardiner [221] who also showed that the three statements are equivalent.

Thm. 8.2.7 can also be obtained from [518, Thm. 2.1] combined with [509, Thm. 5.2]. Notice that both terms in [518, Thm. 2.1] contribute

nontrivially when s is even since the modulus is p^2. If s is even then the theorem is given also by [119, (6.2)]. For FMZVs, as an improvement of Thm. 8.5.1 Zhou and Cai [615] proved Thm. 8.5.4.

Notice that the dualities in the FMZV setting given by both Thm. 8.5.10 (φ-dual) and Thm. 8.5.11 (v-dual) are different from that of the MZVs. These were discovered by Hoffman first in [279,281]. Definition 8.5.6 comes from Hoffman's modification of an idea of Vermaseren described in [540, Appendix B]. Our proof of Lemma 8.5.8 here follows that of [279, Thm. 4.2] while Thm. 8.5.11 is [596, Thm. 2.11]. J. Rosen [471] extended this result to an asymptotic type formula with weights related to the powers of p. This hints at a relation to the p-adic MZVs subsequently studied by Jarossay [306–308]. In particular, the shuffle relation of the symmetrized MZVs stated in Thm. 6.3.4 corresponds to the shuffle relation satisfied the FMZVs. These are given precisely by Thm. 8.4.3, which also includes the case for FESs.

In the spring of 2013 after the author gave an Ober-Seminar talk at the Max Planck Institute for Mathematics at Bonn, Prof. D. Zagier mentioned that he had come to the dimension part of Conjecture 8.5.14 some time earlier [580]. He also informed the author of Conjecture 8.5.15 which was to appear in a joint paper with Kaneko (see [326]). Notice that Thm. 6.4.19, proved by Yasuda [564], shows that the map in Conjecture 8.5.15 is well-defined and surjective.

Thm. 8.5.23, an FMZV analog of the sum formula for SMZVs (see Thm. 6.2.2) was proved by Saito and Wakabayashi in [478,479], who also found the results in Exercise 8.7 which are the sum formulas for FMZVs in the spirit of Conjecture 5.9.2 for MZVs and Thm. 10.9.2 for MZSVs. The FMZV analog of the Ohno's relation of MZVs (see Thm. 5.4.9) can be found in Exercise 8.8, which was conjectured by Kaneko and proved by Oyama [448].

Exercises

8.1. For every positive integer ℓ, prove that the following procedure provides an embedding $\iota_\ell : \mathbb{R} \hookrightarrow \mathcal{A}_\ell$ as sets. If $r \in \mathbb{Q}$ we set $\iota_\ell(r)$ is defined as in Eq. (8.4). Assume $r = n.d_1d_2\cdots$ is irrational where n is its integral part and $d_1, d_2, \ldots$ are its decimal digits. Let $r_k = n.d_1d_2\ldots d_k$ be the kth approximation of r. We define $\iota_\ell(r) = \big(\iota_{p_j,\ell}(r)\big)_{j\geq 1}$ by the following.

$\iota_{p_j,\ell}(r) = \iota_{p_j,\ell}(r_k)$ for all $j = k(k-1)+1, k(k-1)+2, \cdots, k(k+1)$.

Roughly speaking, this means that for each k we put the kth approximation r_k of r at $2k$ consecutive components.

8.2. Prove Lemma 8.5.2 using Eq. (8.2) and Fermat's little Theorem:

$$k^{p^2(p-1)} \equiv 1 \pmod{p^3}$$

for all k not divisible by p. You also need the fact that pB_{2m} is p-integral for all $m \in \mathbb{N}$ by Von Staudt-Clausen Theorem [303, p. 233].

8.3. Prove the case $a = 0$ of Eq. (8.65) using Eq. (7.25), Thm. 8.5.1, and Thm. 8.6.4.

8.4. Prove Eq. (8.42) using the same idea as in the proof of Eq. (8.41).

8.5. Prove Eq. (8.42) by backward induction on d by first proving the following key identity:

$$S_{w,d,1} = \sum_{j=0}^{d-1} (-1)^j \binom{w-d+j-1}{j} S^{\star}_{w,d-j,1} = \left(1 + (-1)^d \binom{w-1}{d-1}\right) \beta_w.$$

8.6. With your favorite computer algebra system, verify numerically the following identities in $\mathcal{A}_2$: (setting $\zeta_2 = \zeta_{\mathcal{A}_2}$)

$$\begin{aligned}
68\boldsymbol{\zeta}_2(2,3,1,1) =& -480\zeta_2(\{1\}^5,2) + 716\zeta_2(\{1\}^3,4) - 843\zeta_2(1,6)\\
34\boldsymbol{\zeta}_2(3,1,1,2) =& -585\zeta_2(\{1\}^5,2) + 998\zeta_2(\{1\}^3,4) - 1029\zeta_2(1,6)\\
68\boldsymbol{\zeta}_2(3,3,1) =& \,1730\zeta_2(\{1\}^5,2) - 2480\zeta_2(\{1\}^3,4) + 2319\zeta_2(1,6)\\
12\boldsymbol{\zeta}_2(5,1,1,1) =& -73\zeta_2(\{1\}^6,2) + 12\zeta_2(1,1,6) - 3\zeta_2(1,2,5)\\
84\boldsymbol{\zeta}_2(3,2,1,3) =& -995\zeta_2(1,8) + 952\zeta_2(\{1\}^3,6) - 1288\zeta_2(\{1\}^5,4)\\
&- 437\zeta_2(\{1\}^7,2) - 624\zeta_2(1,2,6)\\
924\boldsymbol{\zeta}_2(2,1,3,3) =& \,2509\zeta_2(1,8) - 6356\zeta_2(\{1\}^3,6) + 5432\zeta_2(\{1\}^5,4)\\
&- 180\zeta_2(\{1\}^7,2) + 801\zeta_2(1,2,6)\\
924\boldsymbol{\zeta}_2(2,3,1,3) =& -9424\zeta_2(1,8) - 5824\zeta_2(\{1\}^3,6) + 11368\zeta_2(\{1\}^5,4)\\
&- 3807\zeta_2(\{1\}^7,2) + 852\zeta_2(1,2,6)\\
36\boldsymbol{\zeta}_2(1,2,5,1) =& \,358\zeta_2(1,8) - 248\zeta_2(\{1\}^3,6) + 464\zeta_2(\{1\}^5,4)\\
&+ 5\zeta_2(\{1\}^7,2) + 147\zeta_2(1,2,6).
\end{aligned}$$

Can you prove any of these rigorously?

8.7. Suppose a and b are odd positive integers and c is an even positive integer. Show that for all nonnegative integers m and n with $(m,n) \neq (0,0)$,

we have

$$\sum_{\substack{\sum_{i=0}^{2m} n_i=n \\ n_0,\ldots,n_{2m}\geq 0}} \zeta_{\mathcal{A}_1}(\{c\}^{n_0},a,\{c\}^{n_1},b,\{c\}^{n_2},\ldots,a,\{c\}^{n_{2m-1}},b,\{c\}^{n_{2m}}) = 0,$$

$$\sum_{\substack{\sum_{i=0}^{2m} n_i=n \\ n_0,\ldots,n_{2m}\geq 0}} \zeta^{\star}_{\mathcal{A}_1}(\{c\}^{n_0},a,\{c\}^{n_1},b,\{c\}^{n_2},\ldots,a,\{c\}^{n_{2m-1}},b,\{c\}^{n_{2m}}) = 0.$$

8.8. Let $\mathbf{s} = (s_1,\ldots,s_d) \in \mathbb{N}^d$ and suppose its v-dual $\mathbf{s}^\vee$ has depth d'. Then for any $n \in \mathbb{N}_0$, we have

$$\sum_{\mathbf{e}\in(\mathbb{N}_0)^d,|\mathbf{e}|=n} \zeta_{\mathcal{A}}(\mathbf{s}+\mathbf{e}) = \sum_{\mathbf{e}'\in(\mathbb{N}_0)^{d'},|\mathbf{e}'|=n} \zeta_{\mathcal{A}}(\mathbf{s}^\vee+\mathbf{e}').$$

Chapter 9

q-analog of multiple harmonic (star) sums

Throughout this chapter let m and n denote nonnegative integers and q a real number with $0 < q < 1$. For any real number a, we have the quantum factorial and quantum Pochhammer symbol defined by:

$$(a)_0 := (a;q)_0 := 1, \qquad (a)_n := (a;q)_n := \prod_{k=0}^{n-1}(1-aq^k), \quad n \geq 1.$$

As a convention throughout the chapter we always use [ ] to represent q-analog objects. For example, the q-analog of a positive integer n is defined as

$$[n] = [n]_q := \sum_{k=0}^{n-1} q^k = \frac{1-q^n}{1-q},$$

and the Gaussian q-binomial coefficient

$$\begin{bmatrix} n \\ m \end{bmatrix} := \begin{cases} \dfrac{(q)_n}{(q)_m(q)_{n-m}}, & \text{if } 0 \leq m \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

For $m \in \mathbb{N}_0$, $\mathbf{t} = (t_1, \ldots, t_m) \in \mathbb{Z}^m$ $\mathbf{s} = (s_1, \ldots, s_m) \in \mathbb{D}_0^m$ we set $\mathbf{s} = \emptyset$ if $m = 0$ and define the q-analogs of multiple harmonic (star) sums (q-MHS or q-MH*S)

$$H_n^{\mathbf{t}}[\mathbf{s}] := \sum_{n \geq k_1 > \cdots > k_m \geq 1} \prod_{j=1}^{m} \frac{\mathrm{sgn}(s_j)^{k_j} q^{k_j t_j}}{[k_j]^{|s_j|}},$$

$$H_n^{\star,\mathbf{t}}[\mathbf{s}] := \sum_{n \geq k_1 \geq \cdots \geq k_m \geq 1} \prod_{j=1}^{m} \frac{\mathrm{sgn}(s_j)^{k_j} q^{k_j t_j}}{[k_j]^{|s_j|}},$$

with the convention that $H_n^{\mathbf{t}}[\mathbf{s}] = 0$ if $n < m$, and $H_n^{\star,\mathbf{t}}[\emptyset] = H_n^{\mathbf{t}}[\emptyset] = 1$ for all $n \geq 0$. Notice that we allow s_j to be 0 or $\bar{0}$ in these sums. For

the star-version, since we will only consider $H_n^{\star,\mathbf{1}}[\mathbf{s}]$ in this chapter, we will simply write $H_n^{\star}[\mathbf{s}] = H_n^{\star,\mathbf{1}}[\mathbf{s}]$.

Now we fix a symbol θ and define for any $r \in \mathbb{Z} \cup \{\theta\}$ and $k \in \mathbb{N}$ the function

$$Q(r,k) := \begin{cases} rk(k-1)/2, & \text{if } r > 0; \\ rk(k-1)/2 - k, & \text{if } r \leq 0; \\ 0, & \text{if } r = \theta. \end{cases}$$

For $\mathbf{s} = (s_1, \ldots, s_m) \in \mathbb{D}_0^m$, and *mollifiers* $\mathbf{t} = (t_1, \ldots, t_m) \in (\mathbb{N}_0)^m$, and $\mathbf{r} = (r_1, \ldots, r_m) \in (\mathbb{Z} \cup \{\theta\})^m$, we define the *mollified companion of* $H_n^{\mathbf{1}}[\mathbf{s}]$ by

$$\mathcal{H}_n[\mathbf{s};\mathbf{t};\mathbf{r}] := \sum_{n \geq k_1 > \cdots > k_m \geq 1} \frac{\begin{bmatrix} n \\ k_1 \end{bmatrix}}{\begin{bmatrix} n+k_1 \\ k_1 \end{bmatrix}} \prod_{j=1}^{m} \frac{q^{t_j k_j + Q(r_j,k_j)}(1+q^{k_j})}{\operatorname{sgn}(s_j)^{k_j}[k_j]^{|s_j|}}. \tag{9.1}$$

We call $[\mathbf{s};\mathbf{t};\mathbf{r}]$ an *admissible triple of mollifiers* if the limit $\lim_{n\to\infty} \mathcal{H}_n[\mathbf{s};\mathbf{t};\mathbf{r}]$ exists.

Definition 9.0.1. Let $\boxplus$ be a binary operation on $\mathbb{Z} \cup \{\theta\}$, called $\square$-*plus*, such that

- $\theta \boxplus a = a \boxplus \theta = a$ for all $a \in \mathbb{Z} \cup \{\theta\}$,
- $a \boxplus b = a + b$ for all $a, b \in \mathbb{Z}$ with $(a,b) \neq (1,-1), (-1,1)$, and
- $1 \boxplus (-1) = \theta$ and $(-1) \boxplus 1 = 0$.

Lemma 9.0.2. *Let $r \in \{\theta\} \cup \mathbb{Z} \setminus \{0\}$, $d \in \{0,-1\}$. Then for any $k \in \mathbb{Z}$*

$$Q(r,k) + Q(1,k) = Q(r \boxplus 1, k), \tag{9.2}$$

$$k^2 + Q(d,k) = Q(2 \boxplus d, k). \tag{9.3}$$

Moreover, the projection

$$\begin{aligned} \pi : (\mathbb{Z} \cup \{\theta\}, \boxplus) &\longrightarrow (\mathbb{Z}, +) \\ a &\longmapsto a \quad \forall a \in \mathbb{Z}, \\ \theta &\longmapsto 0, \end{aligned}$$

is a homomorphism of semi-groups and its restriction to $\mathbb{Z}^*$ is injective.

Proof. Clear. □

For an admissible triple of mollifiers $[\mathbf{s};\mathbf{t};\mathbf{r}]$, we define

$$\{s_1 \circ \cdots \circ s_m; t_1 \circ \cdots \circ t_m; r_1 \circ \cdots \circ r_m\}$$

to be the set of triples of strings produced by replacing every $\circ$ in $\mathbf{s}$ by either comma "," or O-plus "$\oplus$" (see Eq. (7.1)), replacing every $\circ$ in $\mathbf{t}$ by either comma "," or the usual plus "+", and replacing every $\circ$ in $\mathbf{r}$ by either comma "," or $\square$-plus "$\boxplus$". Moreover, the commas should be at the same positions for all $\mathbf{s}$, $\mathbf{t}$ and $\mathbf{r}$. To save space, we define

$$a^{\circ k} = \underbrace{a \circ \ldots \circ a}_{k \text{ times}}, \quad \text{and} \quad \Big\lfloor\!\circ\!\Big\rfloor_{j=1}^{m} a_j = a_1 \circ \cdots \circ a_m.$$

9.1 Some q-combinatorial identities

We start with a preliminary combinatorial lemma.

Lemma 9.1.1. *For integers $n \geq 1$, $l \geq 0$ we have*

$$\sum_{k=l+1}^{n} (1+q^k) \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} (-1)^k q^{\binom{k}{2}} = \frac{[l]-[n]}{[n]} \frac{\begin{bmatrix} n \\ l \end{bmatrix}}{\begin{bmatrix} n+l \\ l \end{bmatrix}} (-1)^l q^{\binom{l}{2}}, \tag{9.4}$$

$$\sum_{k=l+1}^{n} (1+q^k)[k] \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} q^{k(k-1)} = ([n]-[l]) \frac{\begin{bmatrix} n \\ l \end{bmatrix}}{\begin{bmatrix} n+l \\ l \end{bmatrix}} q^{l^2}, \tag{9.5}$$

$$\sum_{k=1}^{n} \frac{1+q^k}{[k]} \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} q^{k^2} = \sum_{m=1}^{n} \frac{q^m}{[m]}. \tag{9.6}$$

Moreover, if $l \geq 1$ then

$$\sum_{k=l}^{n} \frac{q^k}{[k]^2} \frac{\begin{bmatrix} k \\ l \end{bmatrix}}{\begin{bmatrix} k+l \\ l \end{bmatrix}} = \frac{q^l}{[l]^2} \frac{\begin{bmatrix} n \\ l \end{bmatrix}}{\begin{bmatrix} n+l \\ l \end{bmatrix}}. \tag{9.7}$$

Proof. The general idea of the proof is to use WZ-method. First we set for all positive integers k, l and n

$$F_1(n,k) = (1+q^k) \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} (-1)^{k-1} q^{\binom{k}{2}}, \qquad G_1(n,k) = \frac{q^{n+k}-1}{q^k+1} F_1(n,k),$$

$$F_2(n,k) = \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} (1-q^{2k}) q^{k(k-1)}, \qquad G_2(n,k) = \frac{1-q^{n+k}}{1-q^{2k}} F_2(n,k),$$

$$F_3(l,k) = \frac{\begin{bmatrix} l \\ k \end{bmatrix}}{\begin{bmatrix} l+k \\ k \end{bmatrix}} \frac{1+q^k}{1-q^k} q^{k^2}, \qquad G_3(l,k) = \frac{q^{l-k+1}}{1+q^k} \frac{1-q^k}{1-q^{l-k+1}} F_3(l,k),$$

$$F_4(l,k) = \frac{q^{k-l}}{[k]^2} \frac{\begin{bmatrix} l \\ k \end{bmatrix}}{\begin{bmatrix} l+k \\ k \end{bmatrix}}, \qquad G_4(l,k) = \frac{(1-q^{k-l})(1-q^{k+l})}{q^{k-l}} F_4(l,k).$$

Set the difference operator $\Delta_n f(n) = f(n+1) - f(n)$. Easy computation shows that

$$(1-q^n) F_1(n,k) = \Delta_k G_1(n,k), \qquad F_2(n,k) = -\Delta_k G_2(n,k),$$
$$(1-q^l)^2 F_4(l,k) = \Delta_k G_4(l,k), \qquad \Delta_l F_3(l,k) = -\Delta_k G_3(l,k).$$

Therefore

$$(1-q^n)\sum_{k=l+1}^{n} F_1(n,k) = \sum_{k=l+1}^{n} \Delta_k G_1(n,k) = -G_1(n,l+1)$$
$$= \frac{1-q^{n+l+1}}{1+q^{l+1}}\, F_1(n,l+1) = (1-q^{n-l})\,\frac{\begin{bmatrix} n \\ l \end{bmatrix}}{\begin{bmatrix} n+l \\ l \end{bmatrix}}\,(-1)^l q^{l(l+1)/2},$$
$$\sum_{k=l+1}^{n} F_2(n,k) = G_2(n,l+1) = \frac{\begin{bmatrix} n \\ l \end{bmatrix}}{\begin{bmatrix} n+l \\ l \end{bmatrix}}\,(1-q^n)q^{l(l+1)},$$
$$\sum_{l=0}^{n-1} \Delta_k G_3(l,k) = -\sum_{l=0}^{n-1} \Delta_l F_3(l,k) = F_3(0,k) - F_3(n,k) = -F_3(n,k),$$
$$(1-q^l)^2 \sum_{k=l}^{n} F_4(l,k) = \sum_{k=l}^{n} \Delta_k G_4(l,k) = G_4(l,n+1) = (1-q)^2 \frac{\begin{bmatrix} n \\ l \end{bmatrix}}{\begin{bmatrix} n+l \\ l \end{bmatrix}}.$$

So Eq. (9.4), Eq. (9.5) and Eq. (9.7) are proved and moreover,

$$\sum_{k=1}^{n} F_3(n,k) = -\sum_{m=0}^{n-1}\sum_{k=1}^{n} \Delta_k G_3(m,k) = \sum_{m=0}^{n-1} (G_3(m,1) - G_3(m,n+1))$$
$$= \sum_{m=0}^{n-1} G_3(m,1) = \sum_{m=0}^{n-1} \frac{q^{m+1}}{1-q^{m+1}},$$

which implies Eq. (9.6). This completes the proof of the lemma. □

Lemma 9.1.2. *Let $a \in \mathbb{D}_0$, $b \in \mathbb{N}_0$, $c \in \mathbb{N}$, $r \in \{\theta\} \cup \mathbb{Z} \setminus \{0\}$, and $[\mathbf{x};\mathbf{y};\mathbf{z}]$ an admissible triple of mollifiers. Then for any positive integer n,*

$$\frac{1}{[n]^c}\mathcal{H}_n[a,\mathbf{x};b,\mathbf{y};r \boxplus 1,\mathbf{z}] = \sum_{\substack{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}) \in \{\overline{0}\circ 1^{\circ(c-1)} \circ (a \oplus \overline{1});\\ 0^{\circ c} \circ b;\, 1 \circ \theta^{\circ(c-1)} \circ r\}}} \mathcal{H}_n[\mathbf{p},\mathbf{x};\widetilde{\mathbf{p}},\mathbf{y};\widetilde{\widetilde{\mathbf{p}}},\mathbf{z}].$$

Proof. We prove the lemma by induction on c. Suppose all the mollifiers $\mathbf{x},\mathbf{y},\mathbf{z}$ have length m. Set

$$A_{n,k} = (-1)^k(1+q^k)q^{k(k-1)/2}\frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}}. \tag{9.8}$$

Then by Eq. (9.4),

$$\frac{(1+q^l)}{[l]}\sum_{k=l+1}^{n} A_{n,k} = \left(\frac{1}{[n]} - \frac{1}{[l]}\right) A_{n,l} \tag{9.9}$$

which, together with Eq. (9.2) yields

$$\begin{aligned}
&\mathcal{H}_n[\overline{0}, a\oplus\overline{1}, \mathbf{x}; 0, b, \mathbf{y}; 1, r, \mathbf{z}]\\
&= \sum_{n\geq k>k_0>k_1>\cdots>k_m\geq 1} \frac{A_{n,k}q^{bk_0+Q(r,k_0)}(1+q^{k_0})}{(-\operatorname{sgn}(a))^{k_0}[k_0]^{|a|+1}} \prod_{j=1}^{m} \frac{q^{y_jk_j+Q(z_j,k_j)}(1+q^{k_j})}{\operatorname{sgn}(x_j)^{k_j}[k_j]^{|x_j|}}\\
&= \sum_{n\geq k_0>\cdots>k_m\geq 1} \frac{q^{bk_0+Q(r,k_0)}(1+q^{k_0})}{(-\operatorname{sgn}(a))^{k_0}[k_0]^{|a|+1}} \prod_{j=1}^{m} \frac{q^{y_jk_j+Q(z_j,k_j)}(1+q^{k_j})}{\operatorname{sgn}(x_j)^{k_j}[k_j]^{|x_j|}} \sum_{k=k_0+1}^{n} A_{n,k}\\
&= \frac{1}{[n]}\mathcal{H}_n[a, \mathbf{x}; b, \mathbf{y}; r\uplus 1, \mathbf{z}] - \mathcal{H}_n[a\oplus 1, \mathbf{x}; b, \mathbf{y}; r\uplus 1, \mathbf{z}].
\end{aligned}$$

This proves the lemma for $c = 1$. Now suppose $c > 1$. By the case $c = 1$ we have just proved,

$$\begin{aligned}
&\frac{1}{[n]^c}\mathcal{H}_n[a, \mathbf{x}; b, \mathbf{y}; r\uplus 1, \mathbf{z}] = \frac{1}{[n]^{c-1}}\left(\frac{1}{[n]}\mathcal{H}_n[a, \mathbf{x}; b, \mathbf{y}; r\uplus 1, \mathbf{z}]\right)\\
&= \frac{1}{[n]^{c-1}}\mathcal{H}_n[a\oplus 1, \mathbf{x}; b, \mathbf{y}; r\uplus 1, \mathbf{z}] + \frac{1}{[n]^{c-1}}\mathcal{H}_n[\overline{0}, a\oplus\overline{1}, \mathbf{x}; 0, b, \mathbf{y}; 1, r, \mathbf{z}].
\end{aligned}$$

For the first summand we now apply induction assumption using case $c-1$ with a replaced by $a\oplus 1$. For the second summand we apply Lemma 9.1.2 case $c-1$ with $a = \overline{0}$, $b = 0$, and $r = \theta$. Then we see using the above is equal to

$$\begin{aligned}
&\sum_{(\mathbf{p},\widetilde{\mathbf{p}},\widetilde{\widetilde{\mathbf{p}}})\in\{\overline{0}\circ 1^{\circ(c-2)}\circ\left((a\oplus 1)\oplus\overline{1}\right);\, 0^{\circ(c-1)}\circ b;\, 1\circ\theta^{\circ(c-2)}\circ r\}} \mathcal{H}_n[\mathbf{p}, \mathbf{x}; \widetilde{\mathbf{p}}, \mathbf{y}; \widetilde{\widetilde{\mathbf{p}}}, \mathbf{z}]\\
&+ \sum_{(\mathbf{p},\widetilde{\mathbf{p}},\widetilde{\widetilde{\mathbf{p}}})\in\{\overline{0}\circ 1^{\circ(c-1)};\, 0^{\circ c};\, 1\circ\theta^{\circ(c-1)}\}} \mathcal{H}_n[\mathbf{p}, a\oplus\overline{1}, \mathbf{x}; \widetilde{\mathbf{p}}, b, \mathbf{y}; \widetilde{\widetilde{\mathbf{p}}}, r, \mathbf{z}]\\
&= \sum_{(\mathbf{p},\widetilde{\mathbf{p}},\widetilde{\widetilde{\mathbf{p}}})\in\{\overline{0}\circ 1^{\circ(c-1)}\circ(a\oplus\overline{1});\, 0^{\circ c}\circ b;\, 1\circ\theta^{\circ(c-1)}\circ r\}} \mathcal{H}_n[\mathbf{p}, \mathbf{x}; \widetilde{\mathbf{p}}, \mathbf{y}; \widetilde{\widetilde{\mathbf{p}}}, \mathbf{z}],
\end{aligned}$$

since $(a\oplus 1)\oplus\overline{1} = 1\oplus(a\oplus\overline{1})$ for all $a\in\mathbb{D}_0$. We have now completed the proof of the lemma. □

The next corollary is the degenerate case of the proceeding lemma.

Corollary 9.1.3. *Let* $c\in\mathbb{N}_0$ *and* $\Pi(c) = \{\overline{0}\circ 1^{\circ c};\, 0^{\circ(c+1)};\, 1\circ\theta^{\circ c}\}$. *Then*

$$\frac{1}{[n]^c} = -\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi(c)} \mathcal{H}_n[\mathbf{p}; \widetilde{\mathbf{p}}; \widetilde{\widetilde{\mathbf{p}}}].$$

Proof. The case $c = 0$ follows from by Eq. (9.4) by setting $l = 0$. For $c \geq 1$, using the $c = 0$ case we get

$$\frac{1}{[n]^c} = -\frac{1}{[n]^c}\mathcal{H}_n[\bar{0}; 0; 1] = -\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi(c)} \mathcal{H}_n[\mathbf{p}; \widetilde{\mathbf{p}}; \widetilde{\widetilde{\mathbf{p}}}]$$

by taking $a = \bar{0}$, $b = 0,$, $r = \theta$ and $\mathbf{x} = \mathbf{y} = \mathbf{z} = \emptyset$ in Lemma 9.1.2. Hence the corollary is proved. □

9.2 q-MH*S identities

In this section we prove some general rules which explain what to expect for the q-MH*S $H^\star_n[\mathbf{s}]$ when we add a component in front to a positive composition $\mathbf{s}$. This allows us to extend expansion formulas from the two base cases of depth one Eq. (9.10) and Eq. (9.11) to arbitrary positive compositions $\mathbf{s}$.

For simplicity, throughout this section we always set $\underline{\mathbf{p}} = (\mathbf{p}; \widetilde{\mathbf{p}}; \widetilde{\widetilde{\mathbf{p}}})$.

Theorem 9.2.1. *Let $n, c \in \mathbb{N}$ and $c \geq 2$. Then*

$$H^\star_n[1] = \mathcal{H}_n[1; 1; 2] = \sum_{k=1}^{n} \frac{1+q^k}{[k]} \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} q^{k^2}. \tag{9.10}$$

$$H^\star_n[c] = -\sum_{\underline{\mathbf{p}}\in\{(\overline{2})\circ 1^{\circ(c-2)}; 1\circ 0^{\circ(c-2)}; 1\circ\theta^{\circ(c-2)}\}} \mathcal{H}_n[\underline{\mathbf{p}}]. \tag{9.11}$$

Proof. Equation (9.10) is just Eq. (9.6).

We proceed by induction on n to prove Eq. (9.11). When $n = 1$ we have $H^\star_1(c) = q$. On the other hand,

$$\sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_1[\underline{\mathbf{p}}] = \mathcal{H}_1[\overline{c}; 1; 1] = -q,$$

and therefore the formula is true. Suppose the statement is true for $n-1$. Then by the definition

$$H^\star_n[c] = H^\star_{n-1}[c] + \frac{q^n}{[n]}.$$

Let $\Pi_1(c) = \{(\overline{2}) \circ 1^{\circ(c-2)}; 1 \circ 0^{\circ(c-2)}; 1 \circ \theta^{\circ(c-2)}\}$. Applying inductive hypothesis, we obtain

$$H^\star_n[c] = -\sum_{\underline{\mathbf{p}}\in\Pi_1(c)} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] + \frac{q^n}{[n]^c}. \tag{9.12}$$

Set $\Pi' = \{0 \circ 1^{\circ(c-2)};\, 0^{\circ(c-1)};\, \theta^{\circ(c-1)}\}$. To save space, for any string $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)$ we write the substring $\boldsymbol{\lambda}_{\widehat{1}} = (\lambda_2, \dots, \lambda_m)$. Then the sum on the right-hand side of Eq. (9.12) becomes

$$\sum_{\underline{\mathbf{p}} \in \Pi_1(c)} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] = \sum_{\underline{\mathbf{p}} \in \Pi'} \mathcal{H}_{n-1}[\overline{2+p}_1, \mathbf{p}_{\widehat{1}}; 1+\widetilde{p}_1, \widetilde{\mathbf{p}}_{\widehat{1}}; 1 \boxplus \widetilde{\widetilde{p}}_1, \widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}]$$

$$= \sum_{\underline{\mathbf{p}} \in \Pi'} \sum_{n > k_1 > \dots > k_m \geq 1} \frac{q^{k_1(1+\widetilde{p}_1)} A_{n-1,k_1}}{[k_1]^{2+p_1}} \prod_{j=2}^{m} \frac{q^{\widetilde{p}_j k_j} (1+q^{k_j})}{[k_j]^{p_j}},$$

where $A_{n,k}$ is defined in Eq. (9.8). Plugging this into Eq. (9.12) and using the formula

$$A_{n-1,k} = A_{n,k} \left(1 - \frac{[k]^2}{[n]^2}\, q^{n-k}\right), \tag{9.13}$$

we obtain

$$H_n^{\star}[c] = -\sum_{\underline{\mathbf{p}} \in \Pi'} \sum_{n \geq k_1 > \cdots > k_m \geq 1} \frac{q^{k_1(\widetilde{p}_1+1)} A_{n,k_1}}{[k_1]^{2+p_1}} \prod_{j=2}^{m} \frac{q^{\widetilde{p}_j k_j} (1+q^{k_j})}{[k_j]^{p_j}}$$

$$+ \frac{q^n}{[n]^2} \sum_{\underline{\mathbf{p}} \in \Pi'} \sum_{n \geq k_1 > \cdots > k_m \geq 1} \frac{q^{k_1 \widetilde{p}_1} A_{n,k_1}}{[k_1]^{p_1}} \prod_{j=2}^{m} \frac{q^{\widetilde{p}_j k_j} (1+q^{k_j})}{[k_j]^{p_j}} + \frac{q^n}{[n]^c},$$

which implies

$$H_n^{\star}[c] = -\sum_{\underline{\mathbf{p}} \in \Pi_1(c)} \mathcal{H}_n[\underline{\mathbf{p}}] + \frac{q^n}{[n]^2} \sum_{\underline{\mathbf{p}} \in \Pi(c-2)} \mathcal{H}_n[\underline{\mathbf{p}}] + \frac{q^n}{[n]^2},$$

where $\Pi(c-2) = \{\overline{0} \circ 1^{\circ(c-2)};\, 0^{\circ(c-1)};\, 1 \circ \theta^{\circ(c-2)}\}$. Hence the theorem follows from Cor. 9.1.3 immediately by replacing c by $c-2$ there. □

We now turn to the general rules. For any string $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)$, we set $\boldsymbol{\lambda}^{\circ} = \lambda_1 \circ \cdots \circ \lambda_m$ and $\boldsymbol{\lambda}_{\widehat{1}} = (\lambda_2, \dots, \lambda_m)$.

Theorem 9.2.2. *Let $n \in \mathbb{N}$ and $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$. Set $\delta(\mathbf{s}) = 1$ if $s_d = 1$ and $\delta(\mathbf{s}) = -1$ if $s_d > 1$. Suppose there are triple of mollifiers $[\boldsymbol{\lambda}; \widetilde{\boldsymbol{\lambda}}; \widetilde{\widetilde{\boldsymbol{\lambda}}}] = [\lambda_1, \dots, \lambda_m; \widetilde{\lambda}_1, \dots, \widetilde{\lambda}_m; \widetilde{\widetilde{\lambda}}_1, \dots, \widetilde{\widetilde{\lambda}}_m]$ satisfying*

$$\widetilde{\widetilde{\lambda}}_1 \boxplus \cdots \boxplus \widetilde{\widetilde{\lambda}}_j \in \{1, 2\} \quad \forall j \geq 1, \tag{9.14}$$

such that there is an expansion of the form

$$H_n^{\star}[\mathbf{s}] = \delta(\mathbf{s}) \sum_{\underline{\mathbf{p}} \in \{\boldsymbol{\lambda}^{\circ}; \widetilde{\boldsymbol{\lambda}}^{\circ}; \widetilde{\widetilde{\boldsymbol{\lambda}}}^{\circ}\}} \mathcal{H}_n[\underline{\mathbf{p}}].$$

Then for any integer $c \geq 3$ we have

$$H_n^\star[1,\mathbf{s}] = \delta(\mathbf{s}) \sum_{\underline{\mathbf{p}} \in \Pi_1} \mathcal{H}_n[\underline{\mathbf{p}}], \tag{9.15}$$

where $\Pi_1 = \{1 \circ \boldsymbol{\lambda}^\circ;\ 1 \circ \widetilde{\boldsymbol{\lambda}}^\circ;\ 2 \circ (\widetilde{\widetilde{\lambda}}_1 \boxplus -2) \circ (\widetilde{\widetilde{\boldsymbol{\lambda}}}_{\widehat{1}})^\circ\}$,

$$H_n^\star[2,\mathbf{s}] = \delta(\mathbf{s}) \sum_{\underline{\mathbf{p}} \in \Pi_2} \mathcal{H}_n[\underline{\mathbf{p}}], \tag{9.16}$$

where $\Pi_2 = \{(2 \oplus \lambda_1) \circ \boldsymbol{\lambda}_{\widehat{1}}^\circ;\ (1 + \widetilde{\lambda}_1) \circ \widetilde{\boldsymbol{\lambda}}_{\widehat{1}}^\circ;\ \widetilde{\widetilde{\boldsymbol{\lambda}}}^\circ\}$, *and*

$$H_n^\star[c,\mathbf{s}] = \delta(\mathbf{s}) \sum_{\underline{\mathbf{p}} \in \Pi_c} \mathcal{H}_n[\underline{\mathbf{p}}], \tag{9.17}$$

where $\Pi_c = \{(\overline{2}) \circ 1^{\circ(c-3)} \circ (\lambda_1 \oplus \overline{1}) \circ \boldsymbol{\lambda}_{\widehat{1}}^\circ;\ 1 \circ 0^{\circ(c-3)} \circ \widetilde{\boldsymbol{\lambda}}^\circ;\ 1 \circ \theta^{\circ(c-3)} \circ (\widetilde{\widetilde{\lambda}}_1 \boxplus -1) \circ (\widetilde{\widetilde{\boldsymbol{\lambda}}}_{\widehat{1}})^\circ\}$. *Moreover, in all the sets* Π_1, Π_2, *and* Π_c *the third components still satisfy Eq.* (9.14).

Notice that the condition in Eq. (9.14) essentially guarantees that all the triples of mollifiers considered in this section are admissible.

Proof. Set $\Pi(\mathbf{s}) = \{\boldsymbol{\lambda}^\circ;\ \widetilde{\boldsymbol{\lambda}}^\circ;\ \widetilde{\widetilde{\boldsymbol{\lambda}}}^\circ\}$. We prove the three identities in the theorem by induction on n. When $n = 1$ the theorem is clear. Assume the formulas in Eqs. (9.15), (9.16) and (9.17) are true for $n - 1$.

First, we prove Eq. (9.15). By the definition, we have

$$H_n^\star[1,\mathbf{s}] = H_{n-1}^\star[1,\mathbf{s}] + \frac{q^n}{[n]} H_n^\star[\mathbf{s}].$$

Applying induction assumption, we obtain

$$\delta(\mathbf{s}) H_n^\star[1,\mathbf{s}] = \sum_{\underline{\mathbf{p}} \in \Pi_1} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] + \frac{q^n}{[n]} \sum_{\underline{\mathbf{p}} \in \Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}]. \tag{9.18}$$

Setting $\Pi_0 = \{\boldsymbol{\lambda}^\circ; \widetilde{\boldsymbol{\lambda}}^\circ; (\widetilde{\widetilde{\lambda}}_1 \boxplus -2) \circ (\widetilde{\widetilde{\boldsymbol{\lambda}}}_{\widehat{1}})^\circ\}\}$. To save space, for $\underline{\mathbf{p}} \in \Pi_0$ and $\mathbf{k} = (k_1, \ldots, k_m)$ we write

$$f_j(\underline{\mathbf{p}}, \mathbf{k}) = \frac{\operatorname{sgn}(p_j)^{k_j} q^{\widetilde{p}_j k_j + Q(\widetilde{\widetilde{p}}_j, k_j)} (1 + q^{k_j})}{[k_j]^{|p_j|}}.$$

Then

$$\sum_{\underline{\mathbf{p}} \in \Pi_1} \mathcal{H}_{n-1}[\underline{\mathbf{p}}]$$

$$= \sum_{\underline{\mathbf{p}}\in\Pi_0} \Big(\mathcal{H}_{n-1}[1,\mathbf{p};1,\widetilde{\mathbf{p}};2,\widetilde{\widetilde{\mathbf{p}}}] + \mathcal{H}_{n-1}[1\oplus p_1,\mathbf{p}_{\widehat{1}};1+\widetilde{p}_1,\widetilde{\mathbf{p}}_{\widehat{1}};2\boxplus\widetilde{\widetilde{p}}_1,\widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}] \Big)$$

$$= \sum_{\underline{\mathbf{p}}\in\Pi_0} \Bigg(\sum_{n>k_0>k_1>\cdots>k_m\geq 1} \frac{\begin{bmatrix} n-1 \\ k_0 \end{bmatrix}}{\begin{bmatrix} n-1+k_0 \\ k_0 \end{bmatrix}} \frac{(1+q^{k_0})q^{k_0^2}}{[k_0]} \prod_{j=1}^{m} f_j(\underline{\mathbf{p}},\mathbf{k})$$

$$+ \sum_{n>k_1>\cdots>k_m\geq 1} \frac{\begin{bmatrix} n-1 \\ k_1 \end{bmatrix}}{\begin{bmatrix} n-1+k_1 \\ k_1 \end{bmatrix}} \frac{(1+q^{k_1})q^{(1+\widetilde{p}_1)k_1+Q(2\boxplus\widetilde{\widetilde{p}}_1,k_1)}}{\mathrm{sgn}(p_1)^{k_1}[k_1]^{1+|p_1|}} \prod_{j=2}^{m} f_j(\underline{\mathbf{p}},\mathbf{k}) \Bigg).$$

Substituting the above expression into Eq. (9.18) and using Eq. (9.13) we obtain

$$\delta(\mathbf{s})H_n^{\star}[1,\mathbf{s}] - \frac{q^n}{[n]} \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}]$$

$$= \sum_{\underline{\mathbf{p}}\in\Pi_0} \Bigg(\sum_{n\geq k_0>k_1>\cdots>k_m\geq 1} \frac{\begin{bmatrix} n \\ k_0 \end{bmatrix}}{\begin{bmatrix} n+k_0 \\ k_0 \end{bmatrix}} \frac{(1+q^{k_0})q^{k_0^2}}{[k_0]} \prod_{j=1}^{m} f_j(\underline{\mathbf{p}},\mathbf{k})$$

$$- \frac{q^n}{[n]^2} \sum_{n\geq k_0>k_1>\cdots>k_m\geq 1} \frac{\begin{bmatrix} n \\ k_0 \end{bmatrix}}{\begin{bmatrix} n+k_0 \\ k_0 \end{bmatrix}} (1+q^{k_0})q^{k_0(k_0-1)}[k_0] \prod_{j=1}^{m} f_j(\underline{\mathbf{p}},\mathbf{k})$$

$$+ \sum_{n\geq k_1>\cdots>k_m\geq 1} \frac{\begin{bmatrix} n \\ k_1 \end{bmatrix}}{\begin{bmatrix} n+k_1 \\ k_1 \end{bmatrix}} \frac{(1+q^{k_1})q^{(1+\widetilde{p}_1)k_1+Q(2\boxplus\widetilde{\widetilde{p}}_1,k_1)}}{\mathrm{sgn}(p_1)^{k_1}[k_1]^{2a+1+|p_1|}} \prod_{j=2}^{m} f_j(\underline{\mathbf{p}},\mathbf{k})$$

$$- \frac{q^n}{[n]^2} \sum_{n\geq k_1>\cdots>k_m\geq 1} \frac{\begin{bmatrix} n \\ k_1 \end{bmatrix}}{\begin{bmatrix} n+k_1 \\ k_1 \end{bmatrix}} \frac{(1+q^{k_1})q^{\widetilde{p}_1k_1+Q(2\boxplus\widetilde{\widetilde{p}}_1,k_1)}}{\mathrm{sgn}(p_1)^{k_1}[k_1]^{|p_1|-1}} \prod_{j=2}^{m} f_j(\underline{\mathbf{p}},\mathbf{k}) \Bigg).$$

Noticing that the first and third sums on the right-hand side add up to $\sum_{\underline{\mathbf{p}}\in\Pi_1} \mathcal{H}_n[\underline{\mathbf{p}}]$, we have

$$\delta(\mathbf{s})H_n^{\star}[1,\mathbf{s}] = \frac{q^n}{[n]} \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}] + \sum_{\underline{\mathbf{p}}\in\Pi_1} \mathcal{H}_n[\underline{\mathbf{p}}]$$

$$- \frac{q^n}{[n]^2} \sum_{\underline{\mathbf{p}}\in\Pi_0} \sum_{n\geq k_1>\cdots>k_m\geq 1} \Bigg(\prod_{j=1}^{m} f_j(\underline{\mathbf{p}},\mathbf{k}) \sum_{k_0=k_1+1}^{n} \frac{\begin{bmatrix} n \\ k_0 \end{bmatrix}}{\begin{bmatrix} n+k_0 \\ k_0 \end{bmatrix}} \frac{(1+q^{k_0})q^{2\binom{k_0}{2}}}{[k_0]^{-1}}$$

$$+ \frac{\begin{bmatrix} n \\ k_1 \end{bmatrix}}{\begin{bmatrix} n+k_1 \\ k_1 \end{bmatrix}} \frac{(1+q^{k_1})q^{\widetilde{p}_1k_1+Q(2\boxplus\widetilde{\widetilde{p}}_1,k_1)}}{\mathrm{sgn}(p_1)^{k_1}[k_1]^{|p_1|-1}} \prod_{j=2}^{m} f_j(\underline{\mathbf{p}},\mathbf{k}) \Bigg).$$

Summing the multiple sum in the above over k_0 by Eq. (9.5) and noticing that for $\underline{\mathbf{p}} \in \Pi_0$, by Eq. (9.14), the first component $\widetilde{\widetilde{p}}_1$ can take only values

-1 and 0, with the help of Eq. (9.3) we obtain

$$\delta(\mathbf{s})H^\star_n[1,\mathbf{s}] = \sum_{\underline{\mathbf{p}}\in\Pi_1} \mathcal{H}_n[\underline{\mathbf{p}}].$$

This proves Eq. (9.15) by induction.

Similarly, for the second identity we have by the definition

$$H^\star_n[2,\mathbf{s}] = H^\star_{n-1}[\mathbf{s}] + \frac{q^n}{[n]^2} H^\star_n[\mathbf{s}].$$

Applying induction assumption, we obtain

$$\delta(\mathbf{s})H^\star_n[2,\mathbf{s}] = \sum_{\underline{\mathbf{p}}\in\Pi_2} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] + \frac{q^n}{[n]^2} \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}]. \tag{9.19}$$

Expanding the first sum on the right-hand side

$$\begin{aligned}\sum_{\underline{\mathbf{p}}\in\Pi_2} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] &= \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_{n-1}[2\oplus p_1, \mathbf{p}_{\widehat{1}}; 1+\widetilde{p}_1, \widetilde{\mathbf{p}}_{\widehat{1}}; \widetilde{\widetilde{\mathbf{p}}}] \\ &= \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \sum_{n>k_1>\cdots>k_r\geq 1} \frac{\begin{bmatrix} n-1 \\ k_1 \end{bmatrix}}{\begin{bmatrix} n-1+k_1 \\ k_1 \end{bmatrix}} \frac{q^{k_1}}{[k_1]^2} \prod_{j=1}^r f_j(\underline{\mathbf{p}}, \mathbf{k}),\end{aligned}$$

we obtain

$$\begin{aligned}\delta(\mathbf{s})H^\star_n[2,\mathbf{s}] = \frac{q^n}{[n]^2} &\sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}] \\ &+ \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \sum_{n\geq k_1>\cdots>k_r\geq 1} \frac{\begin{bmatrix} n \\ k_1 \end{bmatrix}}{\begin{bmatrix} n+k_1 \\ k_1 \end{bmatrix}} \left(\frac{q^{k_1}}{[k_1]^2} - \frac{q^n}{[n]^2} \right) \prod_{j=1}^r f_j(\underline{\mathbf{p}}, \mathbf{k})\end{aligned}$$

which implies Eq. (9.16) by the definition and straightforward cancelation.

Finally, we turn to Eq. (9.17). By the definition we have

$$H^\star_n[c,\mathbf{s}] = H^\star_{n-1}[c,\mathbf{s}] + \frac{q^n}{[n]^c} H^\star_n[\mathbf{s}].$$

By the induction assumption, we see that

$$\delta(\mathbf{s})H^\star_n[c,\mathbf{s}] = \sum_{\underline{\mathbf{p}}\in\Pi_c} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] + \frac{q^n}{[n]^c} \sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}]. \tag{9.20}$$

Setting $\Pi' = \{0\circ 1^{\circ(c-3)}\circ(\lambda_1\oplus\overline{1})\circ\boldsymbol{\lambda}^\circ_{\widehat{1}}; 0^{\circ(c-2)}\circ\widetilde{\boldsymbol{\lambda}}^\circ; \theta^{\circ(c-2)}\circ(\widetilde{\widetilde{\lambda}}_1\boxplus -1)\circ(\widetilde{\widetilde{\boldsymbol{\lambda}}}_{\widehat{1}})^\circ\}$ we have

$$\begin{aligned}\sum_{\underline{\mathbf{p}}\in\Pi_c} \mathcal{H}_{n-1}[\underline{\mathbf{p}}] &= \sum_{\underline{\mathbf{p}}\in\Pi'} \mathcal{H}_{n-1}[\overline{2}\oplus p_1, \mathbf{p}_{\widehat{1}}; 1+\widetilde{p}_1, \widetilde{\mathbf{p}}_{\widehat{1}}; 1\boxplus\widetilde{\widetilde{p}}_1, \widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}] \\ &= \sum_{\underline{\mathbf{p}}\in\Pi'} \sum_{n>k_1>\cdots>k_r\geq 1} \frac{\begin{bmatrix} n-1 \\ k_1 \end{bmatrix}}{\begin{bmatrix} n-1+k_1 \\ k_1 \end{bmatrix}} \frac{q^{k_1(1+\widetilde{p}_1)+Q(1\boxplus\widetilde{\widetilde{p}}_1,k_1)}}{(-\operatorname{sgn}(p_1))^{k_1}\,[k_1]^{2+|p_1|}} \prod_{j=2}^r f_j(\underline{\mathbf{p}}, \mathbf{k}).\end{aligned}$$

Plugging this into Eq. (9.20) and using Eq. (9.13) we obtain

$$\delta(\mathbf{s})H_n^\star[c,\mathbf{s}] = \sum_{\underline{\mathbf{p}}\in\Pi'} \mathcal{H}_n[\overline{2}\oplus p_1, \mathbf{p}_{\widehat{1}}; 1+\widetilde{p}_1, \widetilde{\mathbf{p}}_{\widehat{1}}; 1\boxplus\widetilde{\widetilde{p}}_1, \widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}] + \frac{q^n}{[n]^c}\sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})}\mathcal{H}_n[\underline{\mathbf{p}}] - \frac{q^n}{[n]^2}\sum_{\underline{\mathbf{p}}\in\Pi'}\mathcal{H}_n[\overline{p}_1, \mathbf{p}_{\widehat{1}}; \widetilde{\mathbf{p}}; 1\boxplus\widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}],$$

which implies

$$\delta(\mathbf{s})H_n^\star[c,\mathbf{s}] = \sum_{\underline{\mathbf{p}}\in\Pi_c}\mathcal{H}_n[\underline{\mathbf{p}}] - \frac{q^n}{[n]^2}\sum_{\underline{\mathbf{p}}\in\Pi''}\mathcal{H}_n[\underline{\mathbf{p}}] + \frac{q^n}{[n]^c}\sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})}\mathcal{H}_n[\underline{\mathbf{p}}], \quad (9.21)$$

where $\Pi'' = \{\overline{0}\circ 1^{\circ(c-3)}\circ(\lambda_1\oplus\overline{1})\circ\boldsymbol{\lambda}_{\widehat{1}}^\circ; 0^{\circ(c-2)}\circ\widetilde{\boldsymbol{\lambda}}^\circ; 1\circ\theta^{\circ(c-3)}\circ(\widetilde{\widetilde{\lambda}}_1\boxplus(-1))\circ(\widetilde{\widetilde{\boldsymbol{\lambda}}}_{\widehat{1}})^\circ\}$. Expanding the second sum from Eq. (9.21), we have

$$\sum_{\underline{\mathbf{p}}\in\Pi''}\mathcal{H}_n[\underline{\mathbf{p}}] = \sum_{\underline{\mathbf{p}}\in\Pi'''}\sum_{\substack{\mathbf{w}=\overline{0}\circ 1^{\circ(c-3)}\circ(p_1\oplus\overline{1})\\ \widetilde{\mathbf{w}}=0^{\circ(c-2)}\circ\widetilde{p}_1;\,\widetilde{\widetilde{\mathbf{w}}}=1\circ\theta^{\circ(c-3)}\circ\widetilde{\widetilde{p}}_1}}\mathcal{H}_n[\mathbf{w},\mathbf{p}_{\widehat{1}};\widetilde{\mathbf{w}},\widetilde{\mathbf{p}}_{\widehat{1}};\widetilde{\widetilde{\mathbf{w}}},\widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}],$$

where $\Pi''' = \{\boldsymbol{\lambda}^\circ; \widetilde{\boldsymbol{\lambda}}^\circ; (\widetilde{\widetilde{\lambda}}_1\boxplus(-1))\circ(\widetilde{\widetilde{\boldsymbol{\lambda}}}_{\widehat{1}})^\circ\}$. Applying Lemma 9.1.2 to the inner sum with $a=p_1$, $b=\widetilde{p}_1$, $r=\widetilde{\widetilde{p}}_1$, c replaced by $c-2$, and $\mathbf{x}=\mathbf{p}_{\widehat{1}}$, $\mathbf{y}=\widetilde{\mathbf{p}}_{\widehat{1}}$, $\mathbf{z}=\widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}$, we obtain

$$\sum_{\underline{\mathbf{p}}\in\Pi''}\mathcal{H}_n[\underline{\mathbf{p}}] = \sum_{\underline{\mathbf{p}}\in\Pi'''}\frac{1}{[n]^{c-2}}\mathcal{H}_n[p_1,\mathbf{p}_{\widehat{1}};\widetilde{\mathbf{p}};\widetilde{\widetilde{p}}_1\boxplus 1,\widetilde{\widetilde{\mathbf{p}}}_{\widehat{1}}] = \frac{1}{[n]^{c-2}}\sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})}\mathcal{H}_n[\underline{\mathbf{p}}]. \quad (9.22)$$

To justify the last equality above we need to show that for the components of $\widetilde{\widetilde{\boldsymbol{\lambda}}}$ satisfying Eq. (9.14) we have $\widetilde{\widetilde{\lambda}}_1\boxplus(-1)\boxplus 1=\widetilde{\widetilde{\lambda}}_1$ and for any $j\geq 2$,

$$\widetilde{\widetilde{\lambda}}_1\boxplus(-1)\boxplus\widetilde{\widetilde{\lambda}}_2\boxplus\cdots\boxplus\widetilde{\widetilde{\lambda}}_j\boxplus 1 = \widetilde{\widetilde{\lambda}}_1\boxplus\widetilde{\widetilde{\lambda}}_2\boxplus\cdots\boxplus\widetilde{\widetilde{\lambda}}_j.$$

These can be proved by using the projection π of Lemma 9.0.2 and the fact that $\pi(\widetilde{\widetilde{\lambda}}_1\boxplus(-1)\boxplus\cdots\boxplus\widetilde{\widetilde{\lambda}}_j\boxplus 1)=\pi(\widetilde{\widetilde{\lambda}}_1\boxplus\cdots\boxplus\widetilde{\widetilde{\lambda}}_j)\in\{1,2\}$ by Eq. (9.14). Now by Eq. (9.21) and Eq. (9.22) we see that Eq. (9.17) is true.

We have completed the proof of the theorem. □

We now state the main q-MH⋆S identity that will lead to many applications later. First, we need some extra notation. For any two compositions $\mathbf{a}=(a_1,\ldots,a_r)$, $\mathbf{b}=(b_1,\ldots,b_t)$ we define

$$\mathbf{a}\wedge\mathbf{b} = (a_1,\ldots,a_r,b_1,\ldots,b_t), \quad \mathbf{a}*\mathbf{b} = (a_1,\ldots,a_{r-1},a_r*b_1,b_2,\ldots,b_t)$$

where $* = \oplus, +, \boxplus$. For $\mathbf{s} \in \mathbb{N}^\ell$ we define a sequence of admissible triple mollifiers $\underline{\mathbf{m}}_i(\mathbf{s}) := (\mathbf{m}_i, \widetilde{\mathbf{m}}_i, \widetilde{\widetilde{\mathbf{m}}}_i)$ $(i = \ell, \dots, 2, 1)$ (backward) inductively as follows.

$$\mathbf{m}_\ell = \begin{cases} (1), & \text{if } s_\ell = 1; \\ (\overline{2}, \{1\}^{s_\ell - 2}), & \text{if } s_\ell \geq 2, \end{cases} \quad \mathbf{m}_i = \begin{cases} (1) \wedge \mathbf{m}_{i+1}, & \text{if } s_i = 1; \\ (2) \oplus \mathbf{m}_{i+1}, & \text{if } s_i = 2; \\ (\overline{2}, \{1\}^{s_i - 3}, \overline{1}) \oplus \mathbf{m}_{i+1}, & \text{if } s_i \geq 3, \end{cases}$$

$$\widetilde{\mathbf{m}}_\ell = \begin{cases} (1), & \text{if } s_\ell = 1; \\ (1, \{0\}^{s_\ell - 2}), & \text{if } s_\ell \geq 2, \end{cases} \quad \widetilde{\mathbf{m}}_i = \begin{cases} (1) \wedge \widetilde{\mathbf{m}}_{i+1}, & \text{if } s_i = 1; \\ (1) + \widetilde{\mathbf{m}}_{i+1}, & \text{if } s_i = 2; \\ (1, \{0\}^{s_i - 3}) \wedge \widetilde{\mathbf{m}}_{i+1}, & \text{if } s_i \geq 3, \end{cases}$$

$$\widetilde{\widetilde{\mathbf{m}}}_\ell = \begin{cases} (2), & \text{if } s_\ell = 1; \\ (1, \{\theta\}^{s_\ell - 2}), & \text{if } s_\ell \geq 2, \end{cases} \quad \widetilde{\widetilde{\mathbf{m}}}_i = \begin{cases} (2, -2) \boxplus \widetilde{\widetilde{\mathbf{m}}}_{i+1}, & \text{if } s_i = 1; \\ \widetilde{\widetilde{\mathbf{m}}}_{i+1}, & \text{if } s_i = 2; \\ (1, \{0\}^{s_i - 3}, -1) \boxplus \widetilde{\widetilde{\mathbf{m}}}_{i+1}, & \text{if } s_i \geq 3, \end{cases}$$

for all $i < \ell$.

Theorem 9.2.3. *For all* $\mathbf{s} = (s_1, \dots, s_\ell) \in \mathbb{N}^\ell$ *let* $\underline{\mathbf{m}}_1(\mathbf{s}) = (\mathbf{m}; \widetilde{\mathbf{m}}; \widetilde{\widetilde{\mathbf{m}}})$. *Then we have*

$$H_n^\star[\mathbf{s}] = \delta(\mathbf{s}) \sum_{\underline{\mathbf{p}} \in \{\mathbf{m}^\circ; \widetilde{\mathbf{m}}^\circ; \widetilde{\widetilde{\mathbf{m}}}^\circ\}} \mathcal{H}_n[\underline{\mathbf{p}}], \quad \textit{where } \delta(\mathbf{s}) = \begin{cases} 1, & \textit{if } s_\ell = 1; \\ -1, & \textit{if } s_\ell \geq 2. \end{cases}$$

Proof. The theorem follows from Thms. 9.2.1 and 9.2.2 easily. □

9.3 Some applications

By repeatedly applying Thm. 9.2.3 we can now obtain many different types of identities.

The following two results gives us the general q-analog of the so-called **2**-1-**2** formula (or simply *Two-one formula*). Here the underline means the string type $(\{2\}^a, 1)$ may be repeated an arbitrary number of times where a may vary in the repetitions.

Corollary 9.3.1. *Suppose* $\ell \in \mathbb{N}_0$ *and* $a_1, \dots, a_\ell \in \mathbb{N}_0$. *Then*

$$H_n^\star[\{2\}^{a_1}, 1, \dots, \{2\}^{a_\ell}, 1] = \sum_{\underline{\mathbf{p}} \in \Pi} \mathcal{H}_n[\underline{\mathbf{p}}] \tag{9.23}$$

where $\Pi = \left\{ \ \boldsymbol{\lfloor} \circ \boldsymbol{\rfloor}_{j=1}^{\ell} (2a_j + 1); \ \boldsymbol{\lfloor} \circ \boldsymbol{\rfloor}_{j=1}^{\ell} (a_j + 1); 2 \circ 0^{\circ(\ell-1)} \right\}$.

Corollary 9.3.2. *Suppose* $\ell \in \mathbb{N}_0$, $a_1, \dots, a_\ell \in \mathbb{N}_0$ *and* $a_{\ell+1} \in \mathbb{N}$. *Then*

$$H_n^\star[\{2\}^{a_1}, 1, \dots, \{2\}^{a_\ell}, 1, \{2\}^{a_{\ell+1}}] = -\sum_{\underline{\mathbf{p}} \in \Pi} \mathcal{H}_n[\underline{\mathbf{p}}]. \tag{9.24}$$

where $\Pi = \Big\{ \Big(\big\lfloor\circ\big\rfloor_{j=1}^{\ell}(2a_j+1)\Big) \circ \overline{2a_{\ell+1}}; \Big(\big\lfloor\circ\big\rfloor_{j=1}^{\ell}(a_j+1)\Big) \circ a_{\ell+1}; 2 \circ 0^{\circ(\ell-1)} \circ (-1)\Big\}$.

We can also get the following general q-analog of the **2**-c-**2** ($c \geq 3$) formula.

Corollary 9.3.3. *Suppose* $\ell, a_j, c_j \in \mathbb{N}_0$ *and* $c_j \geq 3$ *for all* $1 \leq j \leq \ell+1$. *Set* $\mathbf{s} = (\{2\}^{a_1}, c_1, \ldots, \{2\}^{a_\ell}, c_\ell, \{2\}^{a_{\ell+1}})$. *Then*

$$H_n^{\star}[\mathbf{s}] = -\sum_{\underline{\mathbf{p}} \in \Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}], \tag{9.25}$$

where $\Pi(\mathbf{s}) = \Big\{ (\overline{2a_1+2}) \circ 1^{\circ(c_1-3)} \circ \Big(\big\lfloor\circ\big\rfloor_{j=1}^{\ell}(2a_j+3) \circ 1^{\circ(c_j-3)}\Big) \circ (2a_{\ell+1}+1); \Big(\big\lfloor\circ\big\rfloor_{j=1}^{\ell}(a_j+1) \circ 0^{\circ(c_j-3)}\Big) \circ a_{\ell+1};\ 1 \circ \big\lfloor\circ\big\rfloor_{j=1}^{\ell} \theta^{\circ(c_j-2)}\Big\}$.

Remark 9.3.4. If we take $c_j = 3$ for all j then we obtain the q-analog of the so-called *Two-three formula*.

9.4 Congruences of q-MHSs

From last chapter we see that it is quite fruitful to study van Hamme type congruence of the MHSs. In this section, abiding by the same philosophy we will consider some congruences of the q-MHSs of the same type. But this time we treat q-MHSs as rational functions in q.

Notice that $[p] = [p]_q$ is an irreducible polynomial in $\mathbb{Z}[q]$ when p is a prime. For example, we have the q-analog of Wolstenhomle's Theorem Eq. (8.1) given below.

Theorem 9.4.1. *For any* $n \in \mathbb{N}$ *and any primes* $p \geq 5$

$$\sum_{j=1}^{p-1} \frac{1}{[j]_q} \equiv \frac{p-1}{2}(1-q) + \frac{p^2-1}{24}(1-q)^2[p] \pmod{[p]_q^2},$$

$$\sum_{j=1}^{p-1} \frac{1}{[j]_q^2} \equiv -\frac{(p-1)(p-5)}{12}(1-q)^2 \pmod{[p]}.$$

Or, equivalently,

$$\sum_{j=1}^{p-1} \frac{1}{1-q^j} \equiv \frac{p-1}{2} + \frac{p^2-1}{24}(1-q)[p] \pmod{[p]_q^2}, \tag{9.26}$$

$$\sum_{j=1}^{p-1} \frac{1}{(1-q^j)^2} \equiv -\frac{(p-1)(p-5)}{12} \pmod{[p]}. \tag{9.27}$$

Actually we will prove some generalizations of the above congruences. As in the previous sections we will suppress the dependence on q and write $[j] = [j]_q$ for all $j \in \mathbb{N}$. As before, for $\mathbf{s} := (s_1, \ldots, s_d) \in \mathbb{N}^d$, $\mathbf{t} := (t_1, \ldots, t_d) \in \mathbb{N}^d$ we set

$$H_n^{(\mathbf{t})}[\mathbf{s}] := \sum_{n \geq k_1 > \cdots > k_d \geq 1} \frac{q^{k_1 t_1 + \cdots + k_d t_d}}{[k_1]^{s_1} \cdots [k_d]^{s_d}}, \qquad \mathfrak{h}_n^{(\mathbf{t})}[\mathbf{s}] = H_n^{(\mathbf{t})}[\mathbf{s}]/(1-q)^{|\mathbf{s}|}. \tag{9.28}$$

Note that $H_n^{(s_1-1,\ldots,s_d-1)}[\mathbf{s}]$ are the partial sums of the first version of the q-analogs of multiple zeta functions. For trivial $\mathbf{t}$ we set

$$H_n^{\mathbf{0}}[\mathbf{s}] := H_n^{(0,\ldots,0)}[\mathbf{s}], \qquad \mathfrak{h}_n^{\mathbf{0}}[\mathbf{s}] = H_n^{\mathbf{0}}[\mathbf{s}]/(1-q)^{|\mathbf{s}|}.$$

To state the first theorem we define the degenerate Bernoulli numbers $\beta_n(\lambda)$ by the generating series

$$\frac{x}{(1+\lambda x)^{1/\lambda} - 1} = \sum_{n=0}^{\infty} \beta_n(\lambda) \frac{x^n}{n!}.$$

Clearly we have $\lim_{\lambda \to 0} \beta_n(\lambda) = B_n$.

Remark 9.4.2. The degenerate Bernoulli numbers $\beta_n(\lambda)$ are polynomials of λ of degree n. More precisely, Howard proves in [292, Thm. 3.1] that $\beta_0(\lambda) = 1$, $\beta_1(\lambda) = (\lambda - 1)/2$ and

$$\beta_n(\lambda) = b_n \lambda^n + n \sum_{j=1}^{[n/2]} \frac{B_{2j}}{2j} s(n-1, 2j-1) \lambda^{n-2j} \quad \forall n \geq 2, \tag{9.29}$$

where $s(n,j)$ are the Stirling numbers of the first kind (cf. Eq. (5.37)) defined by

$$x(x-1)(x-2)\cdots(x-n+1) = \sum_{j=0}^{n} s(n,j) x^j,$$

and b_n are the Bernoulli numbers of the second kind defined by

$$\frac{x}{\log(1+x)} = \sum_{n=0}^{\infty} \tilde{b}_n \frac{x^n}{n!} = 1 + \frac{1}{2}\frac{x}{1!} - \frac{1}{6}\frac{x^2}{2!} + \frac{1}{4}\frac{x^3}{3!} - \frac{19}{24}\frac{x^4}{4!} + \cdots.$$

Theorem 9.4.3. *If $p > 3$ is a prime, then for all integers $n > 1$ we have*

$$\mathfrak{h}_{p-1}^{(1)}[n] = \sum_{j=1}^{p-1} \frac{q^j}{(1-q^j)^n} \equiv K_n(p) \pmod{[p]},$$

where

$$K_n(p) = \frac{(-1)^{n+1}}{n!} p^n \beta_n\Big(\frac{1}{p}\Big) = (-1)^{n+1}\frac{\tilde{b}_n}{n!} - \frac{(-1)^n}{(n-1)!}\sum_{j=1}^{[n/2]} \frac{B_{2j}}{2j} s(n-1,2j-1)p^{2j}. \quad (9.30)$$

Proof. Fix p throughout this proof. Set $f_n(q) = \mathfrak{h}^{(1)}_{p-1}[n]_q$. Let $\mu = \mu_p = \exp(2\pi\sqrt{-1}/p)$ be the pth primitive root of unity. Then as a polynomial in $\mathbb{Z}[q]$ we have

$$[p]_q = \prod_{k=1}^{p-1}(q-\mu^k).$$

Hence we only need to show that $f_n(\mu^k) = K_n(p)$ for all j. But

$$f_n(\mu^k) = \sum_{j=1}^{p-1} \frac{\mu^{jk}}{(1-\mu^{jk})^n} = \sum_{j=1}^{p-1} \frac{\mu^j}{(1-\mu^j)^n} = f_n(\mu).$$

Now one can check that (see Exercise 9.6) that

$$\frac{p}{1-z^p} - \frac{1}{1-z} = \sum_{j=1}^{p-1} \frac{\mu^j}{\mu^j - z} = \sum_{j=1}^{p-1} \frac{\mu^j}{\mu^j - 1}\frac{1}{1-\frac{z-1}{\mu^j-1}} \quad (9.31)$$

$$= \sum_{j=1}^{p-1} \mu^j \sum_{n=0}^{\infty} \frac{(z-1)^n}{(\mu^j-1)^{n+1}} = \sum_{n=0}^{\infty} (-1)^{n+1} f_{n+1}(\mu)(z-1)^n.$$

On the other hand, by the definition

$$\sum_{n=0}^{\infty}(-1)^{n+1}K_{n+1}(p)(z-1)^n = \frac{1}{1-z}\sum_{n\geq 0}\frac{\beta_{n+1}(1/p)}{(n+1)!}\big(p(z-1)\big)^{n+1} = \frac{p}{1-z^p} - \frac{1}{1-z}.$$

Comparing the above two generating functions we finish the proof of the theorem. Notice that the second equation in Eq. (9.30) follows from Remark 9.4.2. □

Corollary 9.4.4. *If $p > 3$ is a prime, then for every integer $n \geq 0$ we have*

$$\mathfrak{h}^{(0)}_{p-1}[n] \equiv p + \sum_{j=0}^{n} K_j(p) \pmod{[p]}.$$

Proof. If $n = 0$ then $\mathfrak{h}^{(0)}_{p-1}[n] = p-1$ while $K_0(p) = -1$. Thus the corollary is true. Suppose $n \geq 1$. Then we have

$$\mathfrak{h}^{(0)}_{p-1}[n] = \sum_{j=1}^{p-1} \frac{1}{(1-q^j)^n} = \sum_{j=1}^{p-1} \frac{1-q^j+q^j}{(1-q^j)^n} \equiv \mathfrak{h}^{(0)}_{p-1}[n-1] + K_n(p) \pmod{[p]}.$$

Hence the corollary follows by an easy induction. □

Corollary 9.4.5. *If $p > 3$ is a prime, then for every integer $n > t \geq 1$ we have*

$$\mathfrak{h}^{(t)}_{p-1}[n] = \sum_{j=1}^{p-1} \frac{q^{tj}}{(1-q^j)^n} \equiv \sum_{i=0}^{t-1} \binom{t-1}{i} (-1)^i K_{n-i}(p) \pmod{[p]}. \quad (9.32)$$

Moreover,

$$\sum_{j=1}^{p-1} \frac{1}{(1-q^j)^n} \equiv \left(\frac{p-1}{2} + \sum_{j=2}^{n} K_j(p) \right) \pmod{[p]}. \quad (9.33)$$

Proof. If $t > 1$ it is clear that

$$q^{tj} = q^j \big(1-(1-q^j)\big)^{t-1} = q^j \sum_{i=0}^{t-1} \binom{t-1}{i} (-1)^i (1-q^j)^i.$$

So Eq. (9.32) follows from Thm. 9.4.3 and Eq. (9.33) follows from Cor. 9.4.4 and Eq. (9.29). □

Theorem 9.4.6. *Let s be a positive integer and $\mu_s = \exp(2\pi\sqrt{-1}/s)$ the sth primitive root of unity. . Then*

$$\sum_{d=0}^{\infty} \mathfrak{h}^{\mathbf{0}}_{p-1}\big[\{s\}^d\big] x^d \equiv \frac{(-1)^s}{p^s x} \prod_{n=0}^{s-1} \Big(1-(1-\mu_s^n(-x)^{1/s})^p\Big) \pmod{[p]}.$$

Proof. Let $\mu = \exp(2\pi\sqrt{-1}/p)$ be the primitive pth root of unity and set

$$P_n = \sum_{j=1}^{p-1} \frac{1}{(1-\mu^j)^n} - p\delta_{0,n} \quad \forall n \geq 0, \quad (9.34)$$

where $\delta_{0,n}$ is the Kronecker symbol. Similar to the argument in the proof of Thm. 9.4.3 we see that $\mathfrak{h}^{\mathbf{0}}(n) \equiv P_n \pmod{[p]}$ for all $n \geq 1$. By partial

fraction expansion in Eq. (9.31):

$$\begin{aligned} g(x) &:= \sum_{n=0}^{\infty} P_n x^n = \sum_{j=1}^{p-1} \frac{1-\mu^j}{1-x-\mu^j} - p \\ &= \sum_{j=1}^{p-1} \frac{\mu^{-j}}{z\mu^{-j}-1} - \sum_{j=1}^{p-1} \frac{\mu^j}{z-\mu^j} - p \quad (z = 1-x) \\ &= -z^{-1}\left(\frac{p}{z^{-p}-1} - \frac{1}{z^{-1}-1}\right) - \left(\frac{p}{z^p-1} - \frac{1}{z-1}\right) - p \\ &= -\frac{px(x-1)^{p-1}}{1-(1-x)^p}. \end{aligned}$$

Let $a_d = \mathfrak{h}^{\mathbf{0}}_{p-1}\big[\{s\}^d\big]$ for all $d \geq 0$. Let $w(x) = \sum_{d=0}^{\infty} a_d x^d$ be its generating function. By

$$\mathfrak{h}^{\mathbf{0}}_{p-1}\big[\{s\}^d\big] = \frac{1}{d}\sum_{k=0}^{d-1}(-1)^{d-k-1}\mathfrak{h}^{\mathbf{0}}_{p-1}\big[(d-k)s\big]\cdot\mathfrak{h}^{\mathbf{0}}_{p-1}\big[\{s\}^k\big] \tag{9.35}$$

(see Exercise 9.5), we get

$$w(x) = \sum_{d=0}^{\infty} a_d x^d \equiv 1 + \sum_{d=1}^{\infty}\frac{1}{d}\sum_{k=0}^{d-1}(-1)^{d-k-1}P_{(d-k)s}a_k x^d \pmod{[p]}.$$

Differentiating both sides and changing index $d \to d+1$ we get, modulo $[p]$,

$$w'(x) \equiv \sum_{d=0}^{\infty}\sum_{k=0}^{d}(-1)^{d-k}P_{(d-k+1)s}a_k x^d \equiv \sum_{k=0}^{\infty}\sum_{d=k}^{\infty}(-1)^{d-k}P_{(d-k+1)s}a_k x^d.$$

Changing index $d \to d+k$ and then exchanging the order of summation we get

$$\begin{aligned} w'(x) \equiv& w(x)\sum_{d=0}^{\infty}P_{(d+1)s}(-x)^d \equiv \frac{w(x)}{-x}\left(\sum_{d=0}^{\infty}P_{ds}(-x)^d + 1\right) \\ \equiv& \frac{w(x)}{-sx}\left(s + \sum_{n=0}^{s-1}\sum_{d=0}^{\infty}P_d(\mu_s^n(-x)^{1/s})^d\right) \\ \equiv& \frac{w(x)}{-sx}\Big(s + \sum_{n=0}^{s-1} g\big(\mu_s^n(-x)^{1/s}\big)\Big) \\ \equiv& \frac{w(x)}{-sx}\left(s - \sum_{n=0}^{s-1}\frac{p\mu^n(-x)^{1/s}(\mu^n(-x)^{1/s}-1)^{p-1}}{1-(1-\mu_s^n(-x)^{1/s})^p}\right). \end{aligned}$$

Thus

$$(\ln w(x))' = \left(-(\ln x)' + \sum_{n=0}^{s-1} \frac{(1-(1-\mu^n(-x)^{1/s})^p)'}{1-(1-\mu_s^n(-x)^{1/s})^p} \right).$$

Therefore by comparing the constant term we get

$$w(x) \equiv \frac{(-1)^s}{p^s x} \prod_{n=0}^{s-1} \left(1-(1-\mu_s^n(-x)^{1/s})^p\right) \pmod{[p]}$$

as desired. □

Corollary 9.4.7. *For every positive integer $d < p$ we have*

$$\mathfrak{h}_{p-1}^{\mathbf{0}}\left[\{1\}^d\right] \equiv \frac{1}{d+1}\binom{p-1}{d} \pmod{[p]_q}.$$

Proof. By the theorem we get

$$\begin{aligned}\sum_{d=0}^{\infty} \mathfrak{h}_{p-1}^{\mathbf{0}}\left[\{1\}^d\right] x^d \equiv \frac{(1+x)^p-1}{px} &\equiv \frac{1}{px}\sum_{d=0}^{\infty}\binom{p}{d+1}x^{d+1} \\ &\equiv \sum_{d=0}^{\infty}\frac{1}{d+1}\binom{p-1}{d}x^d \pmod{[p]_q}.\end{aligned}$$

The corollary follows immediately. □

Corollary 9.4.8. *For every positive integer $d < p$ we have*

$$\begin{aligned}&\mathfrak{h}_{p-1}^{\mathbf{0}}\left[\{2\}^d\right] \\ &\equiv \frac{(-1)^d}{p^2}\left\{ \sum_{\substack{j+k=d \\ 0\le j,k<p/2}} \binom{p}{2j+1}\binom{p}{2k+1} - \sum_{\substack{j+k=d+1 \\ 1\le j,k<p/2}} \binom{p}{2j}\binom{p}{2k} \right\} \\ &\equiv (-1)^d \frac{2\cdot d!}{(2d+2)!}\binom{p-1}{d}\cdot F_{2,d}(p)\cdot(1-q)^{2d} \pmod{[p]_q}, \qquad (9.36)\end{aligned}$$

where $F_{2,d}(p)$ is a monic polynomial in p of degree d.

Proof. By Thm. 9.4.6 we have modulo $[p]_q$

$$\begin{aligned}&\sum_{d=0}^{\infty}\mathfrak{h}_{p-1}^{\mathbf{0}}\left[\{2\}^d\right]x^d \equiv \frac{1}{p^2x}\left(1-(1-\sqrt{-1}\sqrt{x})^p\right)\left(1-(1+\sqrt{-1}\sqrt{x})^p\right) \\ &\equiv \frac{1}{p^2x}\left|\sum_{j=1}^{(p-1)/2}\binom{p}{2j}(-1)^j x^j + \sqrt{-1}\sqrt{x}\sum_{j=0}^{(p-1)/2}\binom{p}{2j+1}(-1)^j x^j\right|^2,\end{aligned}$$

which easily yields Eq. (9.36) easily. In the first sum of Eq. (9.36) if $j+k=d+1$ and $1\le j,k<p/2$ then we may assume $j>d/2$. Then $(d+1)!\binom{p}{d+1}$ is a factor of $(2j+1)!\binom{p}{2j+1}$ as a polynomial of p, so is $d!\binom{p-1}{d}$. Similarly we can see that $d!\binom{p-1}{d}$ is a factor of the second sum.

In order to determine the leading coefficient we set

$$C_1(x)=\sum_{j=0}^{d}\frac{(2d+2)!\,x^{2j+1}}{(2j+1)!(2l-2j+1)!}=\frac{(x+1)^{2d+2}-(x-1)^{2d+2}}{2},$$

$$C_2(x)=\sum_{j=0}^{d+1}\frac{(2d+2)!\,x^{2j}}{(2j)!(2l-2j+2)!}=\frac{(x+1)^{2d+2}+(x-1)^{2d+2}}{2}.$$

Hence

$$\sum_{\substack{j+k=d\\0\le j,k<p/2}}\frac{1}{(2j+1)!(2k+1)!}-\sum_{\substack{j+k=d+1\\1\le j,k<p/2}}\frac{1}{(2j)!(2k)!}$$
$$=\frac{C_1(1)-(C_2(1)-2)}{(2d+2)!}=\frac{2}{(2d+2)!}.$$

This finishes the proof of the corollary. □

Corollary 9.4.9. *Let d be a positive integer. Set $\delta_d=(1+(-1)^d)$ and $L=3d+3$. Then for every prime $p\ge L$ we have modulo $[p]_q$*

$$\mathfrak{h}^{\mathbf{0}}_{p-1}\big[\{3\}^d\big]\equiv\frac{1}{p^3}\left\{3\delta_d\binom{p}{L}+(-1)^d\cdot 3\sum_{k\ge1}\binom{p}{L-k}\binom{L-k}{k}+\binom{p}{d+1}\right\} \tag{9.37}$$

$$\equiv\begin{cases}\dfrac{-3\cdot d!}{(3d+1)!}\binom{p-1}{d}\cdot F_{3,d}(p) & \text{, if } d \text{ is odd,}\\[2ex] \dfrac{6\cdot d!}{(3d+3)!}\binom{p-1}{d}\cdot F_{3,d}(p) & \text{, if } d \text{ is even,}\end{cases} \tag{9.38}$$

where $F_{3,d}(p)$ is a monic polynomial in p of degree $2d-1$ if d is odd and of degree $2d$ if d is even.

Proof. Let $\mu=\exp(2\pi i/3)$. Then $\mu^2+\mu+1=0$. By Thm. 9.4.6 we have

$$\sum_{d=0}^{\infty}\mathfrak{h}^{\mathbf{0}}_{p-1}\big[\{3\}^d\big]x^d\equiv\frac{-1}{p^3x}\prod_{a=0}^{2}\Big(1-(1-\mu^a\sqrt[3]{-x})^p\Big). \tag{9.39}$$

We now use two ways to expand this. Set $y = \sqrt[3]{-x}$. First, the product on the right-hand side of Eq. (9.39) can be expressed as

$$\begin{aligned}
&1 - \sum_{a=0}^{2}(1-\mu^a y)^p + \sum_{a=0}^{2}(1-\mu^a y)^p(1-\mu^{a+1}y)^p - \prod_{a=0}^{2}(1-\mu^a y)^p \\
=&1 - \sum_{j=0}^{p}\binom{p}{j}\sum_{a=0}^{2}\mu^{aj}y^j + \sum_{a=0}^{2}(1+\mu^a y+\mu^{a+1}y^2)^p - (1+x)^p \\
=&1 - 3\sum_{j=0}^{[p/3]}\binom{p}{3j}x^j + 3\sum_{\substack{j,k\geq 0, j+k<p \\ 2j+k\equiv 0(3)}}\frac{p!\,(-x)^{(j+2k)/3}}{j!k!(p-j-k)!} - (1+x)^p.
\end{aligned}$$

This leads to Eq. (9.37) for $d > 0$. Note that if d is odd then the degree of the polynomial is reduced to $3d-1$ with leading coefficient given by

$$(-1)^d \cdot 3\frac{1}{(L-1)!}\binom{L-1}{1} = \frac{-3}{(L-2)!} = \frac{-3}{(3d+1)!}$$

as we wanted.

Now to prove $d!\binom{p}{d}$ is a factor we use the following expansion of Eq. (9.39):

$$\sum_{d=0}^{\infty}\frac{1}{p^3 x}\sum_{j,k,n\geq 1}(-1)^{j+k+n}\binom{p}{j}\binom{p}{k}\binom{p}{n}x^{(j+k+n)/3}\mu^{k+2n}.$$

Thus

$$\mathfrak{h}_{p-1}^{\mathbf{0}}\big[\{3\}^d\big] \equiv \frac{1}{p^3}\sum_{\substack{1\leq j,k,n\leq p \\ j+k+n=3d+3}}(-1)^{j+k+n}\binom{p}{j}\binom{p}{k}\binom{p}{n}\mu^{k+2n} \qquad (\mathrm{mod}\ [p]_q).$$

Notice that $j+k+n = 3d+3$ implies one of the indices, say j, is at least $d+1$. Then clearly $\binom{p}{j}$ contains $d!\binom{p}{d}$ as a factor, therefore so does $\mathfrak{h}_{p-1}^{\mathbf{0}}\big[\{3\}^d\big]$ (mod $[p]_q$). This completes the proof of the corollary. □

One can also consider the non-homogeneous cases.

Theorem 9.4.10. *Let m, n be two positive integers. For every prime p we have*

$$\mathfrak{h}_{p-1}^{(m,n)}[2m, 2n] \equiv \frac{1}{2}\left\{f(m;p)f(n;p) - f(m+n;p)\right\} \quad (\mathrm{mod}\ [p]),$$

where

$$f(N;p) = \sum_{i=0}^{N-1}\binom{N-1}{i}(-1)^i K_{2N-i}(p).$$

Proof. By the definition and substitution $i \to p-i$ and $j \to p-j$ we have

$$\begin{aligned}
\mathfrak{h}_{p-1}^{(m,n)}[2m,2n] &= \sum_{1\le j<i<p} \frac{q^{mi+nj}}{(1-q^i)^{2m}(1-q^j)^{2n}} \\
&= \sum_{1\le i<j<p} \frac{q^{pm+pn-mi-nj}}{(1-q^{p-i})^{2m}(1-q^{p-j})^{2n}} \\
&\equiv \sum_{1\le i<j<p} \frac{q^{mi+nj}}{(q^i-q^p)^{2m}(q^j-q^p)^{2n}} \qquad (\mathrm{mod}\ [p]) \\
&\equiv \sum_{1\le i<j<p} \frac{q^{mi+nj}}{(1-p^i)^{2m}(1-p^j)^{2n}} \qquad (\mathrm{mod}\ [p]) \\
&\equiv \mathfrak{h}_{p-1}^{(n,m)}[2n,2m] \qquad (\mathrm{mod}\ [p]), \quad (9.40)
\end{aligned}$$

since $q^p \equiv 1 \pmod{[p]}$. By shuffle relation we have

$$\mathfrak{h}_{p-1}^{(m)}[2m]\mathfrak{h}_{p-1}^{(n)}[2n] = \mathfrak{h}_{p-1}^{(m,n)}[2m,2n] + \mathfrak{h}_{p-1}^{(n,m)}[2n,2m] + \mathfrak{h}_{p-1}^{(m+n)}[2m+2n].$$

Together with Eq. (9.40)this yields

$$2\mathfrak{h}_{p-1}^{(m,n)}[2m,2n] \equiv \mathfrak{h}_{p-1}^{(m)}[2m]\mathfrak{h}_{p-1}^{(n)}[2n] - \mathfrak{h}_{p-1}^{(m+n)}[2m+2n] \pmod{[p]}.$$

Our theorem follows from Eq. (9.32) quickly. □

In the study of the q-MZVs the following value appears naturally (see Eq. (12.80)):

$$\varphi_q[n] = \sum_{k=1}^{\infty}(k-1)\frac{q^{(n-1)k}}{[k]^n} = \sum_{k=1}^{\infty}\frac{kq^{(n-1)k}}{[k]^n} - \zeta_q^I[n],$$

where $\zeta_q^I[n] = \sum_{k=1}^{\infty}\frac{q^{(n-1)k}}{[k]^n}$ is the q-Riemann zeta value.

Proposition 9.4.11. *For every prime p we have*

$$\sum_{k=1}^{p-1}\frac{kq^k}{[k]^2} \equiv -\frac{p(p-1)(p+1)}{24}(1-q)^2 \pmod{[p]}. \qquad (9.41)$$

Proof. To save space, all congruences are modulo $[p]$ throughout this proof.

We can check the congruence Eq. (9.41) for $p=2$ and $p=3$ easily by hand. Now we assume $p \ge 5$. By the definition we have

$$\mathfrak{h}_{p-1}^{\mathbf{0}}[1,2] = \sum_{p>i>j<1}\frac{1}{(1-q^i)(1-q^j)^2}.$$

With substitution $i \to p-i$ and $j \to p-j$ we get

$$
\begin{aligned}
&-\mathfrak{h}^{\mathbf{0}}_{p-1}[1,2] = -\sum_{1\le j<i<p} \frac{q^{2i}\cdot q^j}{(q^i-q^p)^2(q^j-q^p)} \\
&\equiv -\sum_{1\le j<i<p} \frac{q^{2i}\cdot q^j}{(q^i-1)^2(q^j-1)} \\
&\equiv -\sum_{1\le j<i<p} \frac{(q^i-1)^2+2(q^i-1)+1}{(q^i-1)^2}\cdot\frac{1-q^j-1}{1-q^j} \\
&\equiv \mathfrak{h}[2,1]-2\mathfrak{h}[1,1]+\sum_{k=1}^{p-1}\frac{p-3+k}{1-q^k}-\sum_{k=1}^{p-1}\frac{k-1}{(1-q^k)^2}-\binom{p-1}{2} \\
&\equiv \mathfrak{h}^{\mathbf{0}}_{p-1}[2,1]-2\mathfrak{h}^{\mathbf{0}}_{p-1}[1,1]+(p-3)\mathfrak{h}^{\mathbf{0}}_{p-1}[1]+\mathfrak{h}^{\mathbf{0}}_{p-1}[2]-\binom{p-1}{2}-\sum_{k=1}^{p-1}\frac{kq^k}{(1-q^k)^2}.
\end{aligned}
$$

Notice that we have the stuffle relations

$$
\begin{aligned}
\mathfrak{h}^{\mathbf{0}}_{p-1}[2,1]+\mathfrak{h}^{\mathbf{0}}_{p-1}[1,2] &= \mathfrak{h}^{\mathbf{0}}_{p-1}[1]\mathfrak{h}^{\mathbf{0}}_{p-1}[2]-\mathfrak{h}^{\mathbf{0}}_{p-1}[3], \\
2\mathfrak{h}^{\mathbf{0}}_{p-1}[1,1] &= \mathfrak{h}[1]^2-\mathfrak{h}^{\mathbf{0}}_{p-1}[2].
\end{aligned}
$$

Hence

$$
\begin{aligned}
\sum_{k=1}^{p-1}\frac{kq^k}{(1-q^k)^2} \equiv{}& (\mathfrak{h}^{\mathbf{0}}_{p-1}[1]+2)\mathfrak{h}^{\mathbf{0}}_{p-1}[2] \\
&-\mathfrak{h}^{\mathbf{0}}_{p-1}[3]-\mathfrak{h}^{\mathbf{0}}_{p-1}[1]^2+(p-3)\mathfrak{h}^{\mathbf{0}}_{p-1}[1]-\binom{p-1}{2}.
\end{aligned}
$$

Taking $d=1$ in Eq. (9.37) of Cor. 9.4.9 we get

$$
\mathfrak{h}^{\mathbf{0}}_{p-1}[3] \equiv -\frac{(p-1)(p-3)}{8}.
$$

The proposition now follows from Eqs. (9.26) and (9.27) immediately. □

9.5 Historical notes

The study of q-series and q-analogs of number theoretical objects has been an active and fruitful area in mathematics for quite a long time. In recent years, q-calculus has gradually developed into an interdisciplinary subject, so much so that some would call it the q-disease [193]. It is related to many other branches of mathematics such as the partitions, elliptic curves and quantum group theory. Many classical identities have q-analogs some of

which can even have combinatorial proofs. Note that the notation of quantum factorial and quantum Pochhammer symbol has only become semi-standard (for e.g., see [9, 193]).

The results of Lemma 9.1.1, due to Hessami Pilehroods (see [271]), are generalizations of the ordinary combinatorial identities contained in [272]. Lemma 9.1.2 is the q-analog of [608, Lemma 3.1]. Like in the study of classical combinatorics, the WZ method (more precisely, its q-analog, see [353]) often plays an important role in the proof of these combinatorial results.

Thm. 9.4.1 is due to Pan [452] who generalized an earlier result of Andrews [10]. In [160] Dilcher extended it again to higher weight in Thm. 9.4.3 by using the degenerate Bernoulli numbers $\beta_n(\lambda)$ which were first defined by Carlitz [111]. The author proved Cor. 9.4.5 and the homogeneous case, Thm. 9.4.6, first in [609]. We note that congruence (9.33) is a variation of [160, (5.11)].

Comparing to the results from Chap. 8 we see that there is a key missing link here: the q-analog of Bernoulli numbers. Even though there are several candidates in the literature (see, for e.g., Eq. (11.23) or [110, 530]), they have not shown to be related to the q-MHSs via any congruence relations. It would be very interesting to work out a solution to this problem. One possible approach is to consider the p-adic objects in [213] and suitable congruences.

A certain modified form (type I form in the terminology of Chap. 11) of Cors. 9.4.8 and 9.4.9 has been improved in [273] (see Exercises 9.7 and 9.8). Moreover, Hessami Pilehroods and Tauraso obtained some "quasi-homogeneous" results for both H- and $H^\star$-versions (see Exercises 9.9 and 9.10). They also found a general duality type result (see Exercises 9.11) which is related to the work of Bradley [84] by the reversal relations.

Exercises

9.1. For integers $n \geq 1$, $l \geq 0$ show that

$$2\sum_{k=l+1}^{n}(-1)^k\frac{\binom{n}{k}}{\binom{n+k}{k}}=\frac{l-n}{n}(-1)^l\frac{\binom{n}{l}}{\binom{n+l}{l}},$$

$$2\sum_{k=l+1}^{n}\frac{k\binom{n}{k}}{\binom{n+k}{k}}=(n-l)\frac{\binom{n}{l}}{\binom{n+l}{l}},$$

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

Moreover, if $l \geq 1$ then

$$\sum_{k=l}^{n} \frac{\binom{k}{l}}{k^2\binom{k+l}{l}} = \frac{\binom{n}{l}}{l^2\binom{n+l}{l}}.$$

9.2. Prove Cor. 9.3.1 and Cor. 9.3.3.

9.3. Suppose $\ell \in \mathbb{N}_0$, $n \in \mathbb{N}$ and $a_0, a_j, b_j, c_j - 3 \in \mathbb{N}_0$ for all $j \geq 1$. Consider the following two possible types of compositions:

$$\begin{aligned}
&(\underline{\mathbf{2}}\text{-}\underline{c}\text{-}\underline{\mathbf{2}}\text{-}\underline{1}) : \mathbf{s} = \Big(\Big\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\Big\}_{j=1}^{\ell}\Big), \ \ \ell \in \mathbb{N},\\
&\Pi(\mathbf{s}) = \Big\{ \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big(\overline{2b_j+2} \circ 1^{\circ(c_j-3)} \circ \overline{2a_j+2}\Big);\\
&\qquad\qquad \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big((b_j+1) \circ 0^{\circ(c_j-3)} \circ (a_j+1)\Big);\\
&\qquad\qquad 1 \circ \theta^{\circ(c_1-3)} \circ 1 \circ \Big\lfloor\circ\Big\rfloor_{j=2}^{\ell} \Big((-1) \circ \theta^{\circ(c_j-3)} \circ 1\Big)\Big\}\\
&(\mathbf{2}\text{-}1\text{-}\underline{\mathbf{2}}\text{-}\underline{c}\text{-}\underline{\mathbf{2}}\text{-}\underline{1}) : \mathbf{s} = \Big(\{2\}^{a_0}, 1, \Big\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\Big\}_{j=1}^{\ell}\Big), \ \ \ell \in \mathbb{N}_0,\\
&\Pi[\mathbf{s}] = \Big\{(2a_0+1) \circ \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big(\overline{2b_j+2} \circ 1^{\circ(c_j-3)} \circ \overline{2a_j+2}\Big);\\
&\qquad\qquad (a_0+1) \circ \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big((b_j+1) \circ 0^{\circ(c_j-3)} \circ (a_j+1)\Big);\\
&\qquad\qquad 2 \circ \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big((-1) \circ \theta^{\circ(c_j-3)} \circ 1\Big)\Big\}.
\end{aligned}$$

Then in each case show that

$$H_n^{\star}[\mathbf{s}] = \sum_{\underline{\mathbf{p}} \in \Pi(\mathbf{s})} \mathcal{H}_n[\underline{\mathbf{p}}].$$

9.4. Suppose $\ell \in \mathbb{N}_0$, $n, a_{\ell+1} \in \mathbb{N}$, and $a_0, a_j, b_j, c_j - 3 \in \mathbb{N}_0$ for all $1 \leq j \leq \ell$. Consider the following two possible types of compositions:

$$\begin{aligned}
&(\underline{\mathbf{2}}\text{-}\underline{c}\text{-}\underline{\mathbf{2}}\text{-}\underline{1}\text{-}\mathbf{2}) : \mathbf{s} = \Big(\Big\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\Big\}_{j=1}^{\ell}, \{2\}^{a_{\ell+1}}\Big), \ \ \ell \in \mathbb{N},\\
&\Pi[\mathbf{s}] = \Big\{ \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big(\overline{2b_j+2} \circ 1^{\circ(c_j-3)} \circ \overline{2a_j+2}\Big) \circ \overline{2a_{\ell+1}};\\
&\qquad\qquad \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell} \Big((b_j+1) \circ 0^{\circ(c_j-3)} \circ (a_j+1)\Big) \circ a_{\ell+1};\\
&\qquad\qquad 1 \circ \theta^{\circ(c_1-3)} \circ 1 \circ \Big\lfloor\circ\Big\rfloor_{j=2}^{\ell} \Big((-1) \circ \theta^{\circ(c_j-3)} \circ 1\Big) \circ (-1)\}\\
&(\mathbf{2}\text{-}1\text{-}\underline{\mathbf{2}}\text{-}\underline{c}\text{-}\underline{\mathbf{2}}\text{-}\underline{1}\text{-}\mathbf{2}) : \mathbf{s} = \Big(\{2\}^{a_0}, 1, \Big\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\Big\}_{j=1}^{\ell}, \{2\}^{a_{\ell+1}}\Big), \ \ \ell \in \mathbb{N}_0,
\end{aligned}$$

$$\Pi[\mathbf{s}] = \Big\{(2a_0+1)\circ\Big\lfloor\circ\Big\rfloor_{j=1}^{\ell}\Big(\overline{2b_j+2}\circ 1^{\circ(c_j-3)}\circ\overline{2a_j+2}\Big)\circ\overline{2a_{\ell+1}};$$
$$(a_0+1)\circ\Big\lfloor\circ\Big\rfloor_{j=1}^{\ell}\Big((b_j+1)\circ 0^{\circ(c_j-3)}\circ(a_j+1)\Big)\circ a_{\ell+1};$$
$$2\circ\Big\lfloor\circ\Big\rfloor_{j=1}^{\ell}\Big((-1)\circ\theta^{\circ(c_j-3)}\circ 1\Big)\circ(-1)\Big\}.$$

Then in each case show that

$$H_n^{\star}[\mathbf{s}] = -\sum_{\underline{\mathbf{p}}\in\Pi(\mathbf{s})}\mathcal{H}_n[\underline{\mathbf{p}}].$$

9.5. Prove Eq. (9.35) by induction on the depth d.

9.6. Verify the partial fraction expansion in Eq. (9.31).

In the rest of the exercises we set $\mathfrak{h}_{p-1}^{\mathrm{I}}[\mathbf{s}] = \mathfrak{h}_{p-1}^{\mathbf{s-1}}[\mathbf{s}]$ and similarly for $\mathfrak{h}_{p-1}^{\star,\mathrm{I}}[\mathbf{s}]$. Also, all the congruences are modulo $[p]$.

9.7. For any positive integer n and for any prime $p > n$,

$$\mathfrak{h}_{p-1}^{\mathrm{I}}\big[\{2\}^n\big] \equiv \frac{(-1)^n}{p(n+1)}\binom{p+n}{2n+1}.$$

Moreover,

$$\mathfrak{h}_{p-1}^{\mathbf{0}}\big[\{2\}^n\big] \equiv \frac{1}{p(n+1)}\left((-1)^n\binom{p-1}{2n+1}+\binom{p-1}{n}\right).$$

9.8. For any positive integer n and for any prime $p > n$,

$$\mathfrak{h}_{p-1}^{\mathrm{I}}\big[\{3\}^n\big] \equiv \frac{1}{p^2(n+1)}\left(\binom{p+2n+1}{3n+2}+\binom{-p+2n+1}{3n+2}\right).$$

9.9. Suppose $a, b \in \mathbb{N}_0$ and $t = a+b+1$. For any prime $p > 2t+1$, we have

$$\mathfrak{h}_{p-1}^{\mathrm{I}}\big[\{2\}^a,3,\{2\}^b\big]+\mathfrak{h}_{p-1}^{\mathbf{0}}\big[\{2\}^b,3,\{2\}^a\big] \equiv \frac{(-1)^{t-1}}{p(t+1)}\binom{p+t}{2t+1}.$$

9.10. Suppose $a, b \in \mathbb{N}_0$ and $w = a+b+2$. For any prime $p \geq w+1$, we have

$$\mathfrak{h}_{p-1}^{\mathrm{I}}\big[\{1\}^a,2,\{1\}^b\big]+\mathfrak{h}_{p-1}^{\mathrm{I}}\big[\{1\}^b,2,\{1\}^a\big] \equiv -\frac{1}{p}\binom{p+1}{w+1}.$$

9.11. For any prime p and composition $\mathbf{s}$, we have

$$H_{p-1}^{\star,\mathrm{I}}\big[\overleftarrow{\mathbf{s}}\big] \equiv -(-1)^{|\mathbf{s}|}H_{p-1}^{\star,\mathrm{I}}\big[\mathbf{s}^{\vee}\big].$$

where the v-dual $\mathbf{s}^{\vee}$ is defined by Eq. (7.9).

9.12. Let $e_0, e_1, \ldots, e_n \in \mathbb{N}_0$. Then

(i) For any prime $p > w+1$, where $w = 2n + \sum_{i=0}^{n} e_i$,

$$\sum_{i=0}^{n} \mathfrak{h}^{\mathrm{I}}_{p-1}\left[\{1\}^{e_i}, 2, \{1\}^{e_{i+1}}, 2, \ldots, 2, \{1\}^{e_{i+n}}\right] \equiv \frac{(-1)^n}{p}\binom{p+n}{w+1}. \quad (9.42)$$

ii) For any prime $p > w+1$, where $w = 3n + \sum_{i=0}^{n} 2e_i$,

$$\sum_{i=0}^{n} \mathfrak{h}^{\mathrm{I}}_{p-1}\left[\{2\}^{e_i}, 3, \{2\}^{e_{i+1}}, 3, \ldots, 3, \{2\}^{e_{i+n}}\right] \equiv \frac{1}{p} f(p, w, n), \quad (9.43)$$

where f is independent of q.

In both congruences we assume $e_i = e_j$ if $i \equiv j \pmod{t+1}$.

Chapter 10

Multiple zeta star values

Multiple zeta star values are a kind of variation of multiple zeta values.

Definition 10.0.1. For any fixed positive integer d and $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$, $s_1 \geq 2$, the *multiple zeta star function* is defined by

$$\zeta^\star(\mathbf{s}) = \sum_{k_1 \geq \cdots \geq k_d > 0} k_1^{-s_1} \cdots k_d^{-s_d} \tag{10.1}$$

where $s_1 + \cdots + s_j > j$ for all $j = 1, \ldots, d$. Again, the number $d = \mathrm{dp}(\mathbf{s})$ is called the *depth* (or *length*), and $|\mathbf{s}| = s_1 + \cdots + s_d$ the *weight* when $\mathbf{s} \in \mathbb{N}^d$, and in this case the values $\zeta^\star(\mathbf{s})$ are called the *multiple zeta star values* (MZ*Vs).

In this chapter we present some of the most important results concerning MZ*V.

10.1 An integral expression

Recall from section 5.1 the power set of $\mathbf{s} = (s_1, \ldots, s_d)$ of weight $w = |\mathbf{s}|$ is $P(\mathbf{s}) = (s_1, s_1 + s_2, \ldots, s_1 + \cdots + s_{d-1}, w)$.

Theorem 10.1.1. *Let* $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ *and*

$$\Delta(\mathbf{s}) = \left\{ (t_1, \ldots, t_w) \in [0,1]^w \,\middle|\, \begin{array}{l} t_j < t_{j+1}, \text{ if } j \notin P(\mathbf{s}); \\ t_j > t_{j+1}, \text{ if } j \in P(\mathbf{s}). \end{array} \right\}$$

Then for every $N \in \mathbb{N}$ *the MH***S* $H_N^\star(\mathbf{s})$ *(see Definition* (7.3)*) is given by the iterated integral*

$$H_N^\star(s_1, \ldots, s_d) = \int_{\Delta(\mathbf{s})} \omega_{\delta(w)} \cdots \omega_{\delta(2)} (t^{N-1} \mathrm{d}t)$$

where $\omega_0 = \frac{\mathrm{d}t}{t}$, $\omega_1 = \frac{\mathrm{d}t}{1-t}$ *and*

$$\delta(j) = \delta(j, \mathbf{s}) \begin{cases} 0, \text{ if } j - 1 \notin P(\mathbf{s}); \\ 1, \text{ if } j - 1 \in P(\mathbf{s}). \end{cases}$$

Proof. For $j = 1, \dots, d$ we set $p_j = s_1 + \cdots + s_j$. We have

$$\Delta(\mathbf{s}) = \big\{0 < t_1 < \cdots < t_{p_1} > t_{p_1+1} < t_{p_1+2} < \cdots \\ \cdots < t_{p_{d-1}} > t_{p_{d-1}+1} < \cdots < t_w < 1\big\}$$

Now we prove the theorem by induction on d. When $d = 1$ it is not hard to see the theorem is true. Now suppose $d \geq 2$ and the theorem is true for depth less than d. Then

$$\begin{aligned}
&\int_{t_{p_1+1}}^1 \omega_{\delta(p_1+1)} \cdots \omega_{\delta(2)}(t^{N-1}\mathrm{d}t) \\
=&\left(\int_t^1 \left(\int_0^{t_{p_1}} \left(\frac{\mathrm{d}t}{t}\right)^{p_1-2} (t^{N-1}\mathrm{d}t)\right) \frac{\mathrm{d}t_{p_1}}{t_{p_1}}\right) \frac{\mathrm{d}t_{p_1+1}}{1-t_{p_1+1}} \\
=&\left(\int_{t_{p_1+1}}^1 \frac{t_{p_1}^{N-1}\mathrm{d}t_{p_1}}{N^{p_1-1}}\right) \frac{\mathrm{d}t_{p_1+1}}{1-t_{p_1+1}} \\
=&\frac{1}{N^{p_1}} \frac{(1-t_{p_1+1}^N)dt_{p_1+1}}{1-t_{p_1+1}} \\
=&\sum_{N\geq m} \frac{1}{N^{p_1}} t_{p_1+1}^{m-1} dt_{p_1+1}.
\end{aligned}$$

Setting $\mathbf{s}' = (s_2, \dots, s_d)$ we see that

$$\begin{aligned}
\int_{\Delta(\mathbf{s})} \omega_{\delta(w)} \cdots \omega_{\delta(2)}(t^{N-1}dt) &= \sum_{N\geq m} \frac{1}{N^{s_1}} \int_{\Delta(\mathbf{s}')} \omega_{\delta(w)} \cdots \omega_{\delta(p_1+2)}(t^{m-1}dt) \\
&= \sum_{N\geq m} \frac{1}{N^{s_1}} \int_{\Delta(\mathbf{s}')} \omega_{\delta'(w-s_1)} \cdots \omega_{\delta'(2)}(t^{m-1}dt)
\end{aligned}$$

where $\delta'(j) = \delta(j, \mathbf{s}')$. Now the theorem following immediately from the induction assumption. □

10.2 Relation between MZ*V and MZV

For compositions $\mathbf{s}$ and $\mathbf{t}$ we say $\mathbf{s}$ refines $\mathbf{t}$ (denoted $\mathbf{s} \succ \mathbf{t}$) if $\mathbf{t}$ can be obtained from $\mathbf{s}$ by combining some of its parts.

Proposition 10.2.1. *For all admissible* $\mathbf{s} \in \mathbb{N}^d$ *we have*

$$\zeta^\star(\mathbf{s}) = \sum_{\mathbf{s}\succeq\mathbf{t}} \zeta(\mathbf{t}), \tag{10.2}$$

$$\zeta(\mathbf{s}) = \sum_{\mathbf{s}\succeq\mathbf{t}} (-1)^{\mathrm{dp}(\mathbf{s})-\mathrm{dp}(\mathbf{t})} \zeta^\star(\mathbf{t}). \tag{10.3}$$

Proof. The first formula Eq. (10.2) is clear by the definition. The second formula Eq. (10.3) follows from the Inclusion-Exclusion Principle. □

Fix a positive integer d. For a partition $\Pi = \{P_1, \dots, P_\ell\}$ of the set $\{1, 2, \cdots, d\}$, we denote $\ell(\Pi) = \ell$ its number of parts and let

$$c(\Pi) = \prod_{k=1}^{\ell} (\sharp P_k - 1)!,$$

where $\sharp S$ denotes the cardinality of a set S. The next result expresses some combination of MZ*V by using only the Riemann zeta values.

Theorem 10.2.2. *Let* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{R}^d$ *where* $s_1 > 1, \dots, s_d > 1$. *Then*

$$\sum_{\sigma \in \mathfrak{S}_d} \zeta^\star(s_{\sigma(1)}, \dots, s_{\sigma(d)}) = \sum_{\Pi\text{: partition of } \{1,\dots,d\}} c(\Pi)\zeta(\mathbf{i}, \Pi), \tag{10.4}$$

$$\sum_{\sigma \in \mathfrak{S}_d} \zeta(s_{\sigma(1)}, \dots, s_{\sigma(d)}) = \sum_{\Pi\text{: partition of } \{1,\dots,d\}} (-1)^{d-\ell(\Pi)} c(\Pi)\zeta(\mathbf{i}, \Pi), \tag{10.5}$$

where $\mathfrak{S}_d$ *is the symmetric group of* d *letters and* $\zeta(\mathbf{i}, \Pi) = \prod_{k=1}^{\ell} \zeta\Big(\sum_{j \in P_k} s_j\Big)$.

Proof. Notice that all the components in $\mathbf{s}$ are real numbers so we may assume they are all distinct since otherwise we may take limits.

Let us prove Eq. (10.4) first. By the definition it is easy to see

$$\sum_{\sigma \in \mathfrak{S}_d} \zeta^\star(s_{\sigma(1)}, \dots, s_{\sigma(d)}) = \sum_{\sigma \in \mathfrak{S}_d} \sum_{k_1 \geq \cdots \geq k_d > 0} \frac{1}{k_{\sigma(1)}^{s_1} \cdots k_{\sigma(d)}^{s_d}}.$$

For each d-tuple $\mathbf{k} = (k_1, \dots, k_d)$ let $\mathfrak{S}_d(\mathbf{k})$ be its stabilizer and $\Pi(\mathbf{k})$ the associated partition of $\{1, \dots, d\}$. By considering each term $1/k_{\sigma(1)}^{s_1} \cdots k_{\sigma(d)}^{s_d}$ we only need to show for every $\mathbf{k}$

$$\sharp \mathfrak{S}_d(\mathbf{k}) = \sum_{\Pi \succeq \Pi(\mathbf{k})} c(\Pi).$$

Now every element in $\mathfrak{S}_d$ can be written as a product of disjoint cycles (called its cycle type) uniquely, up to permutation of the cycles. So $\sharp \mathfrak{S}_d(\mathbf{k})$ counts the number of those elements whose cycle types are specified by a partition of $\{1, \dots, d\}$ that refines $\Pi(\mathbf{k})$. But by the definition $c(\Pi)$ is exactly the number of cycle types that are specified by Π. This concludes the proof of the theorem.

Similarly, to prove Eq. (10.5) we need to show that for every $\mathbf{k}$

$$\sum_{\Pi \succeq \Pi(\mathbf{k})} (-1)^{d-\ell(\Pi)} c(\Pi) = \begin{cases} 0, \text{ if } \ell(\Pi) < d; \\ 1, \text{ if } \ell(\Pi) = d. \end{cases} \tag{10.6}$$

This is because on the left of Eq. (10.5) all of the components of $\mathbf{k}$ are distinct. A moment reflection leads to the conclusion that $d - \ell(\Pi)$ is even if and only if the permutations of cycle type Π is even. For example, $\ell(\Pi) = d$ if and only if the permutations are trivial. Hence the left-hand side of Eq. (10.6) is the signed sum of the number of even and odd permutations in the stabilizer $\mathfrak{S}_d(\Pi)$. If $\mathfrak{S}_d(\Pi)$ is not trivial then there must be at least one transposition which is an odd permutation. Hence the numbers of even and odd permutations are the same. After cancelation the only nontrivial contribution is from $\Pi = \{\{1\}, \{2\}, \ldots, \{d\}\}$ which produces 1. This finishes the proof of Eq. (10.5) and the theorem. □

10.3 Sum relations

We have see the sum relation of the MZVs in Cor. 5.2.4. For MZ*Vs we have a similar result.

Theorem 10.3.1. *For every pair of positive integers $d \leq w$ we have*

$$\sum_{k_1+\cdots+k_d=w} \zeta^\star(k_1, \ldots, k_d) = \binom{w-1}{d-1} \zeta(w). \tag{10.7}$$

Proof. First, for each fixed $\mathbf{s}$ of depth ℓ and weight w we want to count how many copies of $\zeta(\mathbf{s})$ can the left-hand side of Eq. (10.7) produce. For fixed positive number d, this number turns out to be

$$\sharp\{\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{N}^d : k_1 > 1, \mathbf{k} \succeq \mathbf{s}\} = \binom{w-d-1}{d-\ell}.$$

This can be verified as follows. In the sum $s_1 + \cdots + s_\ell = w$ one can think of each s_j as a string of 1's of length s_j . Then each refinement $\mathbf{k} \succeq (s_1, \ldots, s_\ell)$ corresponds to choosing from $(s_1 - 2) + (s_2 - 1) + \cdots + (s_\ell - 1) = w - \ell - 1$ positions between the 1's to insert $d - \ell$ '+' signs (notice $s_1, k_1 \geq 2$). Therefore

$$\sum_{k_1+\cdots+k_d=w} \zeta^\star(k_1, \ldots, k_d) = \sum_{\ell=1}^{d} \binom{w-\ell-1}{d-\ell} \sum_{|\mathbf{s}|=w, \mathrm{dp}(\mathbf{s})=\ell} \zeta(\mathbf{s}). \tag{10.8}$$

Since we have proved the following sum relation for MSVs (see Eq. (5.12))

$$\sum_{|\mathbf{s}|=w, \mathrm{dp}(\mathbf{s})=\ell} \zeta(\mathbf{s}) = \zeta(w),$$

it follows from Eq. (10.8) and an easy combinatorial formula that

$$\sum_{k_1+\cdots+k_d=w} \zeta^\star(k_1,\ldots,k_d) = \sum_{\ell=1}^{d} \binom{w-\ell-1}{d-\ell} \zeta(w) = \binom{w-1}{d-1}\zeta(w).$$

as desired. This finishes the proof the theorem. □

10.4 Cyclic sum relations

The cyclic sum relation for MZV was considered in Sec. 5.2. With the same notation we have the following identity.

Theorem 10.4.1. *For any index set $S \in \mathcal{C}(w,d)$ with $0 < d < w$, we have*

$$\sum_{(k_1,k_2,\ldots,k_d)\in S} \sum_{i=0}^{k_1-2} \zeta^\star(k_1-i,k_2,\ldots,k_d,i+1) = \frac{w\sharp S}{d}\zeta(w+1) \qquad (10.9)$$

where the inner sum on the left-hand side is treated as 0 whenever $k_1 = 1$.

Proposition 10.4.2. *The cyclic sum relation Eq.* (10.9) *implies the sum relation Eq.* (10.7).

Proof. For each divisor $\ell|d$ we let $\mathcal{C}_\ell(w,d)$ be the set of those cyclic equivalent classes S generated by a composition $\big(\{s_1,\ldots,s_\ell\}^{d/\ell}\big)$ where $(s_1,\ldots,s_\ell)$ itself is not a nontrivial repetition of another composition of smaller depth (such $(s_1,\ldots,s_\ell)$ is called primitive). Note that $\sharp S = \ell$ and

$$\sum_{\ell|d} \ell\sharp\mathcal{C}_\ell(w,d) = \sharp\{\mathbf{s}\in\mathbb{N}^d : |\mathbf{s}| = w\} = \binom{w-1}{d-1}.$$

Now every composition $\mathbf{s}$ of weight $w+1$ and depth $d+1$ can be written uniquely in the form of $(k_1-i,k_2,\ldots,k_d,i+1)$ for some $\mathbf{k}$ of weight w and some i such that $0 \le i \le k_1-2$. Moreover, such $\mathbf{k}$ lies in a unique cyclic equivalence class of some $\mathcal{C}_\ell(w,d)$ for a unique ℓ. Here we also need the fact that if $\mathbf{k}$ is a permutation of $\big(\{s_1,\ldots,s_\ell\}^{d/\ell}\big)$ with primitive $(s_1,\ldots,s_\ell)$ then it must have the form $\big(\{s_j,\ldots,s_\ell,s_1,\ldots,s_{j-1}\}^{d/\ell}\big)$ for some j and $(s_j,\ldots,s_\ell,s_1,\ldots,s_{j-1})$ is again primitive. Thus by Eq. (10.9)

$$\begin{aligned}\sum_{k_1+\cdots+k_{d+1}=w+1} \zeta^\star(k_1,\ldots,k_{d+1}) &= \sum_{\ell|d} \sharp\mathcal{C}_\ell(w,d)\frac{w\ell}{d}\zeta(w+1)\\ &= \frac{w}{d}\binom{w-1}{d-1}\zeta(w+1) = \binom{w}{d}\zeta(w+1).\end{aligned}$$

This is equivalent to the sum relation Eq. (10.7). The proposition is now proved. □

Example 10.4.3. Applying the cyclic sum formula Eq. (10.9) to $\big(\{2\}^n\big)$ and $\big(3,\{2\}^{n-2}\big)$ respectively, we get

$$\zeta^\star(\{2\}^n,1)=2\zeta(2n+1), \tag{10.10}$$

$$\zeta^\star(\{2\}^n)+\sum_{i+j=n-2}\zeta^\star\big(\{2\}^i,3,\{2\}^j,1\big)=(2n-1)\zeta(2n). \tag{10.11}$$

Notice Eq. (10.11) is obviously true even for $n=1$.

10.5 Weighted sum relations

Similar to the MZV case we have the following result.

Theorem 10.5.1. *For any integer $w\geq 3$, we have*

$$\sum_{i=2}^{w-1}2^i\zeta^\star(i,w-i)=(2^w+w-3)\zeta(w). \tag{10.12}$$

Proof. By Eq. (10.2) we have $\zeta^\star(i,w-i)=\zeta(i,w-i)+\zeta(w)$. It follows from the weighted sum relations of MZV Eq. (5.15) that

$$\sum_{i=2}^{w-1}2^i\zeta^\star(i,w-i)=\left(w+1+\sum_{i=2}^{w-1}2^i\right)\zeta(w)=\big(2^w+w-3\big)\zeta(w).$$

Hence Eq. (10.12) is proved. □

However, a parallel result corresponding to Thm. 5.10.2 by Gue and Xie for arbitrary depth is not available. In fact we have the following equation.

Example 10.5.2. When the depth $d=3$ and $w=5$ or 6 we have

$$\sum_{\substack{a\geq 2,b\geq 1\\ a+b<5}}\Big(2^{a+b-1}+2^{a-1}-2^{b-1}\Big)\zeta^\star(a,b,5-a-b)=6\zeta(2)\zeta(3)+25\zeta(5),$$

$$\sum_{\substack{a\geq 2,b\geq 1\\ a+b<6}}\Big(2^{a+b-1}+2^{a-1}-2^{b-1}\Big)\zeta^\star(a,b,6-a-b)=9\zeta(3)^2+86\zeta(6).$$

Conjecturally, $\zeta(2)\zeta(3)$ and $\zeta(5)$ are linearly independent over $\mathbb{Q}$, and so are $\zeta(3)^2$ and $\zeta(6)$.

10.6 Fixed weight, height and depth relations

In this section we will consider the sum formula of MZ*Vs analogous to that of the MZVs (see Sec. 5.5). The basic idea is still to use generating functions. However, it turns out that these generating functions can be expressed succinctly by some hypergeometric functions. Naturally, the proof uses transformation properties of hypergeometric functions heavily so the reader is advised to take a brief look at Appendix C before reading this section.

Recall from Definition 5.5.1, for non-negative integers w, d and h, we set

$$\begin{aligned} I(w,d,h) =& \big\{\mathbf{s}=(s_1,\dots,s_d)\in\mathbb{N}^d : |\mathbf{s}|=w, \mathrm{ht}(\mathbf{s})=h\big\},\\ I_0(w,d,h) =& \big\{\mathbf{s}=(s_1,\dots,s_d)\in\mathbb{N}^d : |\mathbf{s}|=w, \mathrm{ht}(\mathbf{s})=h, s_1\geq 2\big\}. \end{aligned}$$

We first consider the sum

$$G_0^\star(w,d,h)=\sum_{\mathbf{s}\in I_0(w,d,h)}\zeta^\star(\mathbf{s}).$$

In fact, similar to the proof of Thm. 5.5.2, we will treat this as a special point in a one parameter family $X_0(w,d,h;t)$ and consider the whole family's generating function. So we define the one variable multiple polylog star function

$$\widetilde{\mathrm{Li}}^\star_{\mathbf{s}}(t)=\widetilde{\mathrm{Li}}^\star_{s_1,\dots,s_d}(t):=\sum_{k_1\geq k_2\geq\cdots\geq k_d\geq 1}\frac{t^{k_1}}{k_1^{s_1}\cdots k_d^{s_d}},$$

which converges uniformly for all $|t|<1$. Moreover, it is easy to see that for all admissible $\mathbf{s}$ we have $\widetilde{\mathrm{Li}}^\star_{\mathbf{s}}(1)=\zeta^\star(\mathbf{s})$. Since $\widetilde{\mathrm{Li}}^\star_{\mathbf{s}}(t)$ makes sense for all positive compositions $\mathbf{s}$ we define

$$G^\star(w,d,h;t)=\sum_{\mathbf{s}\in I(w,d,h)}\widetilde{\mathrm{Li}}^\star_{\mathbf{s}}(t),\quad G_0^\star(w,d,h;t)=\sum_{\mathbf{s}\in I_0(w,d,h)}\widetilde{\mathrm{Li}}^\star_{\mathbf{s}}(t)$$

and their generating functions

$$\begin{aligned} \Psi^\star(t)=\Psi^\star(x,y,z;t)&:=\sum_{w\geq d+h,d\geq h\geq 0}G^\star(w,d,h;t)x^{w-d-h}y^{d-h}z^{2h},\\ \Psi_0^\star(t)=\Psi_0^\star(x,y,z;t)&:=\sum_{w\geq d+h,d\geq h\geq 1}G_0^\star(w,d,h;t)x^{w-d-h}y^{d-h}z^{2h-2}. \end{aligned}$$

Lemma 10.6.1. *The formal power series $\Psi_0^\star(t)$ is a unique power series solution of*

$$t^2(1-t)\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}+t[(1-t)(1-x)-y]\frac{\mathrm{d}f}{\mathrm{d}t}+(xy-z^2)f=t,\quad f(0)=0. \tag{10.13}$$

Proof. It is straightforward to check that (see Exercise 10.1)

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathrm{Li}}_1^\star(t)=\frac{1}{1-t},\quad \frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathrm{Li}}_{\mathbf{s}}^\star(t)=\begin{cases}\dfrac{1}{t}\widetilde{\mathrm{Li}}^\star_{s_1-1,\ldots,s_d}(t), & \text{if } s_1\geq 2;\\ \dfrac{1}{t(1-t)}\widetilde{\mathrm{Li}}^\star_{s_2,\ldots,s_d}(t), & \text{if } s_1=1, d\geq 2.\end{cases}\tag{10.14}$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}G_0^\star(w,d,h;t)=\frac{1}{t}\Big(G^\star(w-1,d,h-1;t)-G_0^\star(w-1,d,h-1;t)+G_0^\star(w-1,d,h;t)\Big)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(G^\star(w,d,h;t)-G_0^\star(w,d,h;t)\Big)=\frac{1}{t(1-t)}G^\star(w-1,d-1,h;t).$$

Therefore the generating functions satisfy

$$\frac{\mathrm{d}\Psi_0^\star}{\mathrm{d}t}=\frac{1}{yt}\Big(\Psi^\star-1-z^2\Psi_0^\star\Big)+\frac{x}{t}\Psi_0^\star$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\Psi^\star-z^2\Psi_0^\star\Big)=\frac{y}{t(1-t)}\Big(\Psi^\star-1\Big)+\frac{y}{1-t}.$$

Solving $\Psi^\star$ from the first equation and plugging it into the second yields the lemma. □

Theorem 10.6.2. *Let α and β be determined by $\alpha+\beta=x+y$ and $\alpha\beta=xy-z^2$. We have*

$$\begin{aligned}&\Psi_0^\star(x,y,z;t)\\ &=\frac{\Gamma(\beta-\alpha)\Gamma(1-y)}{\Gamma(1-\alpha)\Gamma(1+x-\alpha)}\int_0^t v^{-\beta}(1-v)^{y-1}{}_2F_1(\alpha,\alpha-x;\alpha-\beta+1;v)\,dv\\ &+\frac{\Gamma(\alpha-\beta)\Gamma(1-y)}{\Gamma(1-\beta)\Gamma(1+x-\beta)}\int_0^t v^{-\alpha}(1-v)^{y-1}{}_2F_1(\beta,\beta-x;\beta-\alpha+1;v)\,dv;\\ &\Psi_0^\star(x,y,z)\\ &=\frac{\Gamma(\beta-\alpha)\Gamma(1-\beta)\Gamma(y)\Gamma(1-y)}{\Gamma(1-\alpha)\Gamma(1+x-\alpha)\Gamma(1+\alpha-x)}\sum_{n=0}^\infty\frac{(\alpha)_n(1-\beta)_n}{n!(1+\alpha-\beta)_n}\frac{\alpha-x}{n+\alpha-x}\\ &+\frac{\Gamma(\alpha-\beta)\Gamma(1-\alpha)\Gamma(y)\Gamma(1-y)}{\Gamma(1-\beta)\Gamma(1+x-\beta)\Gamma(1+\beta-x)}\sum_{n=0}^\infty\frac{(\beta)_n(1-\alpha)_n}{n!(1+\beta-\alpha)_n}\frac{\beta-x}{n+\beta-x}.\end{aligned}$$

Proof. Since $\Psi_0^\star(x,y,z;t)$ satisfies the second order ordinary differential equation (10.13) we only need to solve it and then specialize at $t=1$. By

Lemma C.28 one checks easily that the two solutions for the corresponding homogeneous differential equation are given by

$$\varphi_j = t^{\alpha_j} {}_2F_1(\alpha_j, \alpha_j - x; \alpha_j - \alpha_{1-j} + 1; t), \quad j = 0, 1, \tag{10.15}$$

where α_0 and α_1 are the two roots of the quadratic equation

$$t^2 - (x + y)t + xy - z^2 = 0.$$

By variations of parameters (see Exercise 10.2) one can find a solution of Eq. (10.13):

$$\Psi_0^\star(x, y, z; t) = u_0(t)\varphi_0(t) + u_1(t)\varphi_1(t), \tag{10.16}$$

where

$$u_j(t) = \frac{1}{\alpha_{1-j} - \alpha_j} \int_0^t v^{-\alpha_{1-j}} (1 - v)^{y-1} {}_2F_1(\alpha_j, \alpha_j - x; \alpha_j - \alpha_{1-j} + 1; v)\, dv$$

for $j = 0, 1$. Finally, taking $t = 1$ we get

$$\Psi_0^\star(x, y, z) = f(1) = u_0(1)\varphi_0(1) + u_1(1)\varphi_1(1)$$

where

$$\varphi_j(1) = \frac{\Gamma(\alpha_j - \alpha_{1-j} + 1)\Gamma(1 - y)}{\Gamma(1 - \alpha_{1-j})\Gamma(1 + x - \alpha_{1-j})} \quad \text{for } j = 0, 1,$$

by the Gauss summation formula Eq. (C.2). This proves the first equation in the theorem. For the second, we can expand the ${}_2F_1$ by the definition and then use the beta integral to evaluate. This completes the proof of the theorem. □

Corollary 10.6.3. *Using the same notation as in the above theorem, we have*

$$\Psi_0^\star(x, y, z) = \frac{1}{(1 - \beta)(1 - y)} {}_3F_2(1 - \beta, 1 - \beta + x, 1; 2 - y, 2 - \beta; 1).$$

Proof. Using linear transformations (see Exercise 10.3) we see that

$$\begin{aligned}
\Psi_0^\star(x, y, z) &= \frac{1}{1 - y} \int_0^1 v^{-\beta} {}_2F_1(1 - \beta, 1 - \beta + x, 2 - y; 1 - v)\, dv \\
&= \frac{1}{1 - y} \int_0^1 (1 - v)^{-\beta} {}_2F_1(1 - \beta, 1 - \beta + x, 2 - y; v)\, dv \qquad (10.17) \\
&= \frac{1}{1 - y} \sum_{n=0}^{\infty} \int_0^1 (1 - v)^{-\beta} v^n \frac{(1 - \beta)_n (1 - \beta + x)_n}{(2 - y)_n n!}\, dv \\
&= \frac{1}{1 - y} \sum_{n=0}^{\infty} \frac{(1 - \beta)_n (1 - \beta + x)_n \Gamma(1 - \beta)}{\Gamma(2 - \beta + n)(2 - y)_n} \quad \text{(beta integral)} \\
&= \frac{1}{(1 - \beta)(1 - y)} \sum_{n=0}^{\infty} \frac{(1 - \beta)_n (1 - \beta + x)_n}{(2 - \beta)_n (2 - y)_n} \quad \text{(by Eq. (1.13))} \\
&= \frac{1}{(1 - \beta)(1 - y)} {}_3F_2(1 - \beta, 1 - \beta + x, 1; 2 - y, 2 - \beta; 1),
\end{aligned}$$

as desired. □

We have another useful power series expansion of the generating function.

Theorem 10.6.4. *We have*

$$\Psi_0^\star(x,y,z;t)=\sum_{n=1}^{\infty}\frac{t^n}{n(n-x)}\prod_{j=1}^{n}\frac{j(j-x)}{(j-x)(j-y)-x^2}$$

Proof. Assume the power-series solution at $t=0$ is given by

$$\Phi_0(x,y,z;t)=\sum_{n=0}^{\infty}a_nt^n.$$

Then by using the differential equation (10.13) we find

$$a_0=0,\ a_1=\frac{1}{(1-x)(1-y)-z^2},\ a_n=\frac{(n-1)(n-1-x)a_{n-1}}{(n-x)(n-y)-z^2}\ \forall n\geq 2.$$

The yields the theorem easily. □

Let $I_0(w,*,h)=\bigcup_{d\geq 1}I_0(w,d,h)$. The next result is the sum formula of MZ*Vs of fixed weight and height.

Corollary 10.6.5. *Let h and w be integers such that $h\geq 1$ and $w\geq 2h$. Then*

$$\sum_{\mathbf{s}\in I_0(w,*,h)}\zeta^\star(\mathbf{s})=2\binom{w-1}{2h-1}(1-2^{1-w})\zeta(w).$$

Proof. By Thm. 10.6.4, we have

$$\Psi_0^\star(x,x,z;t)=\sum_{n=1}^{\infty}\frac{t^n}{n(n-x)}\prod_{j=1}^{n}\frac{j(j-x)}{(j-x)^2-z^2}:=\sum_{n=1}^{\infty}a_nt^n.$$

It is easy to see that

$$a_n=\frac{\Gamma(n)\Gamma(n-x)\Gamma(1-x-z)\Gamma(1-x+z)}{\Gamma(1-x)\Gamma(1-x-z+n)\Gamma(1-x+z+n)}.$$

By partial fraction expansion we may assume

$$a_n=\sum_{\ell=1}\left(\frac{A_{n,\ell}^+}{x+z-\ell}+\frac{A_{n,\ell}^-}{x-z-\ell}\right)$$

where

$$A_{n,\ell}^+=\lim_{z\to\ell-x}(x+z-\ell)a_n=(-1)^\ell\binom{n-1}{\ell-1}\frac{\Gamma(n-x)}{\Gamma(1-x)(1+\ell-2x)_n}.$$

Similar computation leads to $A^-_{n,\ell} = A^+_{n,\ell}$. Hence specializing to $t = 1$ gives

$$\Psi^\star_0(x,x,z;1) = \sum_{n=1}^{\infty} a_n. = \sum_{\ell=1}\sum_{n=\ell}^{\infty}\left(\frac{A^+_{n,\ell}}{x+z-\ell} + \frac{A^-_{n,\ell}}{x-z-\ell}\right)$$
$$= \sum_{\ell=1}(-1)^\ell\left(\frac{1}{x+z-\ell}+\frac{1}{x-z-\ell}\right)\sum_{n=0}^{\infty}\frac{\binom{n+\ell-1}{\ell-1}\Gamma(n+\ell-x)}{\Gamma(1-x)(1+\ell-2x)_{n+\ell}}.$$

Notice that the sum over n can be rewritten as

$$\frac{(1-x)\cdots(\ell-1-x)}{(1+\ell-2x)\cdots(2\ell-2x)}{}_2F_1(\ell,\ell-x,1+2\ell-2x;1)$$
$$= \frac{(1-x)\cdots(\ell-1-x)}{(1+\ell-2x)\cdots(2\ell-2x)}\frac{\Gamma(1+2\ell-2x)\Gamma(1-x)}{\Gamma(1+\ell-2x)\Gamma(1+\ell-x)} = \frac{1}{\ell-x}$$

by Gauss summation formula Eq. (C.28). Therefore

$$\Psi^\star_0(x,x,z;1) = \sum_{\ell=1}\frac{(-1)^\ell}{\ell-x}\left(\frac{1}{x+z-\ell}+\frac{1}{x-z-\ell}\right)$$
$$= \frac{1}{z}\sum_{\ell=1}(-1)^\ell\left(\frac{1}{x+z-\ell}-\frac{1}{x-z-\ell}\right).$$

Expanding and taking the coefficient of $x^{w-2h}z^{2h-2}$ we get

$$2\binom{w-1}{2h-1}\sum_{\ell=1}^{\infty}\frac{(-1)^{\ell-1}}{\ell^w} = 2\binom{w-1}{2h-1}(1-2^{1-w})\zeta(w).$$

This finishes the proof of the corollary. □

Let $I_0(w,d,*) = \bigcup_{h\geq 1} I_0(w,d,h)$. The next result is the sum formula of MZ*Vs of fixed weight and depth.

Corollary 10.6.6. *Let d and w be integers such that $d \geq 1$ and $w \geq d+1$. Then*

$$\sum_{\mathbf{s}\in I_0(w,d,*)}\zeta^\star(\mathbf{s}) = \binom{w-1}{d-1}\zeta(w).$$

Proof. Take $z^2 = xy$, $\alpha = 0$, $\beta = x+y$, by Eq. (10.17), we have

$$\Psi^\star_0(x,y,z) = \frac{1}{1-y}\int_0^1 (1-v)^{-x-y}{}_2F_1(1-x-y,1-y,2-y;v)\,dv$$
$$= \sum_{n=0}^{\infty}\frac{(1-x-y)_n}{n!(n+1-y)}\int_0^1(1-v)^{-x-y}v^n\,dv$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{1}{(n+1-y)(n+1-x-y)} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n-y)(n-x-y)} \\
&= \frac{1}{x} \sum_{n=1}^{\infty} \left(\frac{1}{n-x-y} - \frac{1}{n-y} \right) \\
&= \frac{1}{x} \sum_{n=1}^{\infty} \sum_{w=1}^{\infty} \frac{1}{n^{w+1}} \left((x+y)^w - y^w \right) \\
&= \frac{1}{x} \sum_{w=1}^{\infty} \zeta(w+1) \sum_{d=0}^{w-1} \binom{w}{d} x^{w-d} y^d \\
&= \sum_{w \geq d+1 \geq 2}^{\infty} \zeta(w+1) \binom{w-1}{d-1} x^{w-d-1} y^{d-1}.
\end{aligned}
$$

Comparing the coefficient of $x^{w-d-1}y^{d-1}$ we complete the proof of the corollary. □

10.7 Duality relations

Unlike MZVs the duality for MZ*Vs is not a simple equation, instead, the duality says the difference between a MZ*V and its dual value can be expressed as a polynomial of Riemann zeta values with rational coefficients.

The notation from the last section is still in force. Define a and b by the conditions $a+b=-x+y$ and $ab=-xy-z^2$. Since we do not use the word algebra so in this section a and b are not used as alphabet. Set

$$
\begin{aligned}
&F_1(x,y,a,b) \\
&= \frac{(x+a)\Gamma(b-a)\Gamma(1+a-b)}{(x+b)\Gamma(1-x-a)\Gamma(1+x+a)} \left(\frac{x\Gamma(y)\Gamma(1-y)}{b\Gamma(a)\Gamma(1-a)} - \frac{y\Gamma(x)\Gamma(1-x)}{a\Gamma(b)\Gamma(1-b)} \right), \\
&F_2(x,y,a,b) \\
&= \frac{xy(a-b)\Gamma(b-a)}{(x+b)\Gamma(1-x-a)\Gamma(1+x+a)} \times \\
&\times \left(\frac{\Gamma(y)\Gamma(1-y)\Gamma(-b)}{\Gamma(1-a)} - \frac{\Gamma(x)\Gamma(1-x)\Gamma(a)}{\Gamma(1+b)} \right) {}_3F_2 \left(\begin{matrix} a, -b, x+a \\ a-b, 1+x+a \end{matrix} \right),
\end{aligned}
$$

where we put ${}_3F_2(\cdots) = {}_3F_2(\cdots;1)$ to save space. Further we define

$$\begin{aligned} &A(x,y,a,b) = A(x,y,b,a) \\ &= \frac{1}{2\pi}\left\{\frac{\cos\pi x}{\sin\pi y} - \frac{\cos\pi y}{\sin\pi x} + \cos\pi(a-b)(\cot\pi x - \cot\pi y)\right\} \qquad (10.18) \\ &= \frac{1}{\Gamma(x+a)\Gamma(1-x-a)}\left(\frac{\Gamma(y)\Gamma(1-y)}{\Gamma(a)\Gamma(1-a)} + \frac{\Gamma(x)\Gamma(1-x)}{\Gamma(b)\Gamma(1-b)}\right), \qquad (10.19) \end{aligned}$$

which can be checked easily using the well-known reflection formula

$$\Gamma(\gamma)\Gamma(1-\gamma) = \frac{\pi}{\sin(\pi\gamma)}.$$

Lemma 10.7.1. *We have*

$$\begin{aligned} &F_1(x,y,a,b) + F_1(x,y,b,a) \\ &\qquad = \frac{x-y}{ab} + \frac{(a-b)xy}{ab(x+a)(x+b)}\Gamma(b-a)\Gamma(1+a-b)A(x,y,a,b). \end{aligned}$$

Proof. Using the reflection formula for the gamma function, we see that

$$\begin{aligned} F_1(x,y,a,b) + F_1(x,y,b,a) &= \Gamma(b-a)\Gamma(1+a-b)\times \\ &\times\left\{\frac{\sin\pi(x+a)}{\pi(x+b)}\left(\frac{x\sin\pi a}{b\sin\pi y} - \frac{y\sin\pi b}{a\sin\pi x}\right)\right. \\ &\left. -\frac{\sin\pi(x+b)}{\pi(x+a)}\left(\frac{x\sin\pi b}{a\sin\pi y} - \frac{y\sin\pi a}{b\sin\pi x}\right)\right\}. \end{aligned}$$

With $x+a = y-b$ the terms within the curly brackets can be reduced to

$$\begin{aligned} &\frac{1}{2\pi(x+b)}\left(\frac{x(\cos\pi x - \cos\pi y\cos\pi(a-b) + \sin\pi y\sin\pi(a-b))}{b\sin\pi y}\right. \\ &\left. -\frac{y(\cos\pi x\cos\pi(a-b) - \sin\pi x\sin\pi(a-b) - \cos\pi y)}{a\sin\pi x}\right) \\ &-\frac{1}{2\pi(x+a)}\left(\frac{x(\cos\pi x - \cos\pi y\cos\pi(b-a) + \sin\pi y\sin\pi(b-a))}{a\sin\pi y}\right. \\ &\left. -\frac{y(\cos\pi x\cos\pi(b-a) - \sin\pi x\sin\pi(b-a) - \cos\pi y)}{b\sin\pi x}\right). \end{aligned} \qquad (10.20)$$

By using the identities

$$\begin{aligned} \frac{1}{b(x+b)} - \frac{1}{a(x+a)} &= \frac{y(a-b)}{ab(x+a)(x+b)}, \\ \frac{1}{a(x+b)} - \frac{1}{b(x+a)} &= \frac{x(b-a)}{ab(x+a)(x+b)}, \\ \frac{x}{b(x+b)} + \frac{y}{a(x+b)} + \frac{x}{a(x+a)} + \frac{y}{b(x+a)} &= \frac{2(y-x)}{ab}, \end{aligned}$$

we see that the expression Eq. (10.20) becomes

$$\frac{xy(a-b)}{ab(x+a)(x+b)}A(x,y,a,b)+\frac{(y-x)\sin\pi(a-b)}{ab\pi}$$

by Eq. (10.18). We have completed the proof of the lemma. □

Lemma 10.7.2. *We have*

$$\begin{aligned}&F_2(x,y,a,b)+F_2(x,y,b,a)\\&\quad=\frac{(b-a)xy}{ab(x+a)(x+b)}\Gamma(b-a)\Gamma(1+a-b)A(x,y,a,b)\\&\qquad+A(x,y,a,b)\frac{\Gamma(a)\Gamma(1-a)\Gamma(b)\Gamma(1-b)\Gamma(x+a)\Gamma(x+b)}{\Gamma(x)\Gamma(y)}.\end{aligned}$$

Proof. Applying Prop. C.29 to $F_2(x,y,a,b)$, we get

$$\begin{aligned}&F_2(x,y,a,b)\\&=\frac{(a-b)\Gamma(a-b)\Gamma(b-a)\Gamma(x+b)}{\Gamma(x)\Gamma(y)\Gamma(1-x-a)}\left(\frac{\Gamma(y)\Gamma(1-y)}{\Gamma(a)\Gamma(1-a)}-\frac{\Gamma(x)\Gamma(1-x)}{\Gamma(-b)\Gamma(1+b)}\right)\times\\&\qquad\times\left(\psi(1+y)-\psi(a)-\psi(-b)-\gamma-\sum_{n=1}^{\infty}\frac{(a)_n(1+b)_n}{nn!(1+y)_n}\right)\\&=A(x,y,a,b)\frac{\Gamma(b-a)\Gamma(1+a-b)\Gamma(x+a)\Gamma(x+b)}{\Gamma(x)\Gamma(y)}\times\\&\qquad\times\left(\psi(1+y)-\psi(a)-\psi(-b)-\gamma-\sum_{n=1}^{\infty}\frac{(a)_n(1+b)_n}{nn!(1+y)_n}\right)\end{aligned}$$

using definition Eq. (10.19). By $A(x,y,a,b)=A(x,y,b,a)$ we see that

$$\begin{aligned}&F_2(x,y,a,b)+F_2(x,y,b,a)\\&=A(x,y,a,b)\frac{\Gamma(b-a)\Gamma(1+a-b)\Gamma(x+a)\Gamma(x+b)}{\Gamma(x)\Gamma(y)}\times\\&\times\left\{\sum_{n=1}^{\infty}\left(\frac{(1+a)_n(b)_n}{n(1+y)_nn!}-\frac{(a)_n(1+b)_n}{n(1+y)_nn!}\right)+\psi(b)-\psi(-b)+\psi(-a)-\psi(a)\right\}.\end{aligned}$$

Further,

$$\begin{aligned}
&\sum_{n=1}^{\infty}\left(\frac{(1+a)_n(b)_n}{nn!(1+y)_n}-\frac{(a)_n(1+b)_n}{n(1+y)_n n!}\right)\\
=&\frac{b-a}{ab}\sum_{n=1}^{\infty}\frac{(a)_n(b)_n}{n!(1+y)_n}\\
=&\frac{b-a}{ab}\left(\frac{\Gamma(1+x)\Gamma(1+y)}{\Gamma(1+x+a)\Gamma(1+x+b)}-1\right)\\
=&\frac{(b-a)\Gamma(1+x)\Gamma(1+y)}{ab\Gamma(1+x+a)\Gamma(1+x+b)}-\frac{1}{a}+\frac{1}{b}
\end{aligned}$$

by Gaussian summation formula Eq. (C.2). So the lemma follows from the identities

$$\pi\cot\pi a-\pi\cot\pi b=\frac{\pi\sin\pi(b-a)}{\sin(\pi a)\sin(\pi b)}=\frac{\Gamma(a)\Gamma(1-a)\Gamma(b)\Gamma(1-b)}{\Gamma(b-a)\Gamma(1+a-b)},$$

and the functional equation of the digamma function $\psi(-x)-\psi(x)-1/x=\pi\cot\pi x$ which can be derived easily from Eq. (1.6). □

Theorem 10.7.3. *We have*

$$\begin{aligned}
&x\Psi_0^\star(-x,y,z)-y\Psi_0^\star(-y,x,z)\\
&\quad=\frac{x-y}{ab}+A(x,y,a,b)\frac{\Gamma(a)\Gamma(1-a)\Gamma(b)\Gamma(1-b)\Gamma(x+a)\Gamma(x+b)}{\Gamma(x)\Gamma(y)}.
\end{aligned}\tag{10.21}$$

Proof. Replacing x by $-x$ in the second equation in Thm. 10.6.2 we see that

$$\begin{aligned}
&\Psi_0^\star(-x,y,z)\\
=&\frac{\Gamma(b-a)\Gamma(1-b)\Gamma(y)\Gamma(1-y)}{\Gamma(1-a)\Gamma(1-x-a)\Gamma(1+a+x)}\sum_{n=0}^{\infty}\frac{(a)_n(1-b)_n}{n!(1+a-b)_n}\frac{a+x}{n+a+x}\\
+&\frac{\Gamma(a-b)\Gamma(1-a)\Gamma(y)\Gamma(1-y)}{\Gamma(1-b)\Gamma(1-x-b)\Gamma(1+b+x)}\sum_{n=0}^{\infty}\frac{(b)_n(1-a)_n}{n!(1+b-a)_n}\frac{b+x}{n+b+x}.
\end{aligned}$$

Thus

$$x\Psi_0^\star(-x,y,z)-y\Psi_0^\star(-y,x,z)=F(x,y,a,b)+F(x,y,b,a),\tag{10.22}$$

where

$$\begin{aligned}
&F(x,y,a,b)\\
&=\frac{x\Gamma(y)\Gamma(1-y)\Gamma(1-b)\Gamma(b-a)}{\Gamma(1-a)\Gamma(1-x-a)\Gamma(1+x+a)}\sum_{n=0}^{\infty}\frac{(a)_n(1-b)_n}{n!(1+a-b)_n}\frac{x+a}{n+x+a}\\
&-\frac{y\Gamma(x)\Gamma(1-x)\Gamma(1+a)\Gamma(b-a)}{\Gamma(1+b)\Gamma(1-x-a)\Gamma(1+x+a)}\sum_{n=0}^{\infty}\frac{(1+a)_n(-b)_n}{n!(1+a-b)_n}\frac{x+a}{n+x+a}.
\end{aligned}$$

But

$$\sum_{n=0}^{\infty}\frac{(a)_n(1-b)_n}{n!(1+a-b)_n}\frac{x+a}{n+x+a}=\sum_{n=0}^{\infty}\frac{(a)_n(-b)_n}{n!(a-b)_n}\frac{(a-b)(x+a)(n-b)}{(-b)(n+a-b)(n+x+a)},$$

and

$$\frac{-(n-b)}{(n+a-b)(n+x+a)}=\frac{a}{x+b}\frac{1}{n+a-b}-\frac{y}{x+b}\frac{1}{n+x+a},$$

we see that

$$\begin{aligned}
&\sum_{n=0}^{\infty}\frac{(a)_n(1-b)_n}{n!(1+a-b)_n}\frac{x+a}{n+x+a}\\
&=\frac{a(x+a)}{b(x+b)}{}_2F_1\left(\begin{matrix}a,-b\\1+a-b\end{matrix}\right)-\frac{y(a-b)}{b(x+b)}{}_3F_2\left(\begin{matrix}a,-b,x+a\\a-b,1+x+a\end{matrix}\right)\\
&=\frac{a(x+a)\Gamma(1+a-b)}{b(x+b)\Gamma(1+a)\Gamma(1-b)}-\frac{y(a-b)}{b(x+b)}{}_3F_2\left(\begin{matrix}a,-b,x+a\\a-b,1+x+a\end{matrix}\right)
\end{aligned}$$

by the Gauss summation formula Eq. (C.2). Exchanging $(a,x)\leftrightarrow(-b,y)$ and noticing that $x+a=y-b$, we get

$$\begin{aligned}
&\sum_{n=0}^{\infty}\frac{(1+a)_n(-b)_n}{n!(1+a-b)_n}\frac{x+a}{n+x+a}\\
&\qquad=\frac{b(x+a)\Gamma(1+a-b)}{a(x+b)\Gamma(1+a)\Gamma(1-b)}+\frac{x(a-b)}{a(x+b)}{}_3F_2\left(\begin{matrix}a,-b,x+a\\a-b,1+x+a\end{matrix}\right).
\end{aligned}$$

Therefore

$$F(x,y,a,b)=F_1(x,y,a,b)+F_2(x,y,a,b).$$

So the theorem follows from this together with Eq. (10.22), Lemma 10.7.2 and Lemma 10.7.1. □

The theorem has an immediate application to the sum formulas of MZ*Vs.

Corollary 10.7.4. *For any w, d, h with $w, d \geq h$, the difference*
$D_{w,d,h} := (-1)^w X_0^\star(w+d+1, d+1, h) - (-1)^d X_0^\star(w+d+1, w+1, h)$
can be expressed as a polynomial of Riemann zeta values with rational coefficients.

Proof. We have

$$x\Psi_0^\star(-x,y,z) - y\Psi_0^\star(-y,x,z) = \sum_{w,d\geq h\geq 1} (-1)^h D_{w,d,h} x^{w+1-h} y^{d+1-h} z^{2h-2} + \sum_{d\geq h\geq 1} (-1)^{d+h} X_0^\star(d+h,h,h)(x^{d+1-h} - y^{d+1-h}) z^{2h-2}.$$

So the corollary follows from the expansion of Gamma function Eq. (3.17). □

Corollary 10.7.5. *Let $\psi(X)$ be the digamma function (see Eq.* (1.5)*). Then*

$$x\Psi_0^\star(-x,y,0) - x\Psi_0^\star(-x,0,0) - y\Psi_0^\star(-y,x,0) + y\Psi_0^\star(-y,0,0) = -\psi(x) + \psi(y) - \pi\big(\cot(\pi x) - \cot(\pi y)\big)\frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}.$$

Proof. Setting $\alpha = x$ and $\beta = y$ in Thm. 10.6.2 we find

$$\Psi_0^\star(x,y,0) = \frac{\Gamma(y-x)\Gamma(1-y)}{\Gamma(1-x)} \int_0^t v^{-y}(1-v)^{y-1} dv + \frac{1}{x-y}\int_0^t v^{-x}(1-v)^{y-1}\, {}_2F_1(y, y-x; y-x+1; v)\, dv.$$

Clearly we have

$$y\Psi_0^\star(-y,0,0) - x\Psi_0^\star(-x,0,0) = \int_0^1 \frac{v^x}{1-v} - \frac{v^y}{1-v}\, dv = \psi(y+1) - \psi(x+1) = \psi(y) + \frac{1}{y} - \psi(x+1) - \frac{1}{x} \quad (10.23)$$

by Eq. (1.5) and Eq. (1.6). To compute $x\Psi_0^\star(-x,y,0) - y\Psi_0^\star(-y,x,0)$ we take $z = 0, (a,b) = (-x,y)$ in Thm. 10.7.3. Thus we have

$$\begin{aligned} & x\Psi_0^\star(-x,y,0) - y\Psi_0^\star(-y,x,0) \\ &= \frac{1}{x} - \frac{1}{y} + \left(\frac{\Gamma(y)\Gamma(1-y)}{\Gamma(a)\Gamma(1-a)} + \frac{\Gamma(x)\Gamma(1-x)}{\Gamma(b)\Gamma(1-b)}\right)\frac{\Gamma(a)\Gamma(1-a)\Gamma(b)\Gamma(1-b)}{\Gamma(x)\Gamma(y)\Gamma(x+y)^{-1}} \\ &= \frac{1}{x} - \frac{1}{y} - \left(\frac{\sin\pi y}{\sin\pi x} - \frac{\sin\pi x}{\sin\pi y}\right)\frac{\pi\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)\sin\pi(x+y)} \\ &= \frac{1}{x} - \frac{1}{y} - \big(\cot(\pi x) - \cot(\pi y)\big)\frac{\pi\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)} \end{aligned}$$

by an easy manipulation of trigonometric identities. This completes the proof of the corollary. □

Corollary 10.7.6. *Let $\psi(X)$ be the digamma function. Then*

$$\sum_{\ell,d\geq 2}\Big((-1)^{\ell}\zeta^{\star}(\ell,\{1\}^{d-1})-(-1)^{d}\zeta^{\star}(d,\{1\}^{\ell-1})\Big)x^{\ell-1}y^{d-1}$$
$$=-\psi(x)+\psi(y)-\pi(\cot(\pi x)-\cot(\pi y))\frac{\Gamma(1-x)\Gamma(1-y)}{\Gamma(1-x-y)}.$$

Consequently, $(-1)^{\ell}\zeta^{\star}(\ell,\{1\}^{d-1})-(-1)^{d}\zeta^{\star}(d,\{1\}^{\ell-1})$ can always be expressed as a $\mathbb{Q}$-linear combination of products of Riemann zeta values.

Proof. The first part follows from Cor. 10.7.5 immediately. For the last sentence, it suffices to notice that

$$\psi(x)=-\frac{1}{x}-\sum_{\ell\geq 2}\zeta(\ell)x^{\ell-1}-\gamma,\quad\text{(by Eq. (1.5) and Eq. (1.6))}$$
$$\pi\cot(\pi x)=\frac{1}{x}-\sum_{\ell\geq 1}\zeta(2\ell)x^{2\ell-1},\qquad\text{(see, for e.g., [5, § 4.3.70]),}$$

and the expansion (6.5) of $\Gamma(1-x)\Gamma(1-y)/\Gamma(1-x-y)$. □

10.8 Some applications of binomial identities of q-MZ*Vs

We will prove some binomial identities of q-analog of MZ*Vs in Chap. 12 by using results from Chap. 9. By taking $q\to 1$ in Thm. 12.9.4 we obtain the following result.

Theorem 10.8.1. *For all $\mathbf{s}=(s_1,\dots,s_\ell)\in\mathbb{N}^{\ell}$ we define a sequence of composition $\mathbf{m}_i(\mathbf{s})$ with entries in $\mathbb{D}$ (see Definition*7.1.1*) for $i=\ell,\dots,2,1$ (backward) inductively as follows.*

$$\mathbf{m}_{\ell}(\mathbf{s})=\begin{cases}(1), & \text{if } s_{\ell}=1;\\ (\overline{2},\{1\}^{s_{\ell}-2}), & \text{if } s_{\ell}\geq 2,\end{cases}$$
$$\mathbf{m}_{i}(\mathbf{s})=\begin{cases}(1)\wedge\mathbf{m}_{i+1}(\mathbf{s}), & \text{if } s_i=1;\\ (2)\oplus\mathbf{m}_{i+1}(\mathbf{s}), & \text{if } s_i=2;\\ (\overline{2},\{1\}^{s_i-3},\overline{1})\oplus\mathbf{m}_{i+1}(\mathbf{s}), & \text{if } s_i\geq 3,\end{cases}$$

If $s_1>1$ then we have

$$\zeta^{\star}(\mathbf{s})=\delta(\mathbf{s})\sum_{\mathbf{p}\in\{\mathbf{m}(\mathbf{s})^{\circ}_{1}\}}2^{\mathrm{dp}(\mathbf{p})}\zeta(\mathbf{p}),\quad\text{where }\delta(\mathbf{s})=\begin{cases}1, & \text{if } s_{\ell}=1;\\ -1, & \text{if } s_{\ell}\geq 2.\end{cases}$$

The following result is called the *Two-one formula*. For ease of notation, we write $\mathbf{p} \preceq \mathbf{t}$ to mean $\mathbf{p}$ is obtained from $\mathbf{t}$ by combining some parts using $\oplus$.

Proposition 10.8.2. *For any positive integers $s_1, \dots, s_\ell$, define the composition $\mathbf{s}' = (2s_1+1, \dots, 2s_\ell+1)$. Then we have*

$$\zeta^\star(\{2\}^{s_1}, 1, \dots, \{2\}^{s_\ell}, 1) = \sum_{\mathbf{p} \preceq \mathbf{s}'} 2^{\mathrm{dp}(\mathbf{p})} \zeta(\mathbf{p}) \tag{10.24}$$

$$= \sum_{\mathbf{p} \preceq \mathbf{s}'} (-1)^{\ell - \mathrm{dp}(\mathbf{p})} 2^{\mathrm{dp}(\mathbf{p})} \zeta^\star(\mathbf{p}). \tag{10.25}$$

Proof. Equation Eq. (10.24) follows easily from Thm. 10.8.1. By Eq. (10.2) we have

$$\sum_{\mathbf{p} \preceq \mathbf{s}'} (-1)^{\ell - \mathrm{dp}(\mathbf{p})} 2^{\mathrm{dp}(\mathbf{p})} \zeta^\star(\mathbf{p}) = \sum_{\mathbf{p} \preceq \mathbf{s}'} (-1)^{\ell - \mathrm{dp}(\mathbf{p})} 2^{\mathrm{dp}(\mathbf{p})} \sum_{\mathbf{t} \preceq \mathbf{p}} \zeta(\mathbf{t})$$

Now for each fixed $\mathbf{t} \preceq \mathbf{s}'$ and fixed number r between $d := \mathrm{dp}(\mathbf{t})$ and $\ell = \mathrm{dp}(\mathbf{s}')$ we need to count how many ways to choose $\mathbf{p}$ so that $\mathbf{t} \preceq \mathbf{p} \preceq \mathbf{s}'$ and $\mathrm{dp}(\mathbf{p}) = r$. Easy combinatorial consideration gives this number as $\binom{\ell-d}{\ell-r}$. Thus

$$\sum_{\mathbf{p} \preceq \mathbf{s}'} (-1)^{\ell - \mathrm{dp}(\mathbf{p})} 2^{\mathrm{dp}(\mathbf{p})} \sum_{\mathbf{t} \preceq \mathbf{p}} \zeta(\mathbf{t}) = \sum_{\mathbf{t} \preceq \mathbf{s}'} (-1)^{\ell} \sum_{r=\mathrm{dp}(\mathbf{t})}^{\ell} (-2)^r \binom{\ell - \mathrm{dp}(\mathbf{t})}{\ell - r} \zeta(\mathbf{t})$$

$$= \sum_{\mathbf{t} \preceq \mathbf{s}'} 2^{\mathrm{dp}(\mathbf{t})} \zeta(\mathbf{t}).$$

Hence the second equation is proved. □

By the same idea we can get from Thm. 10.8.1 the following **2**-1-**2** formula.

Proposition 10.8.3. *Suppose $\ell, a_j \in \mathbb{N}_0$ for all $j \le \ell$ and $a_{\ell+1} \in \mathbb{N}$. Set $\mathbf{s} = (\{2\}^{a_1}, 1, \dots, \{2\}^{a_\ell}, 1, \{2\}^{a_{\ell+1}})$. Then we have*

$$\zeta^\star{}_n(\mathbf{s}) = - \sum_{\mathbf{p} \preceq (2a_1+1, \dots, 2a_\ell+1, \overline{2a_{\ell+1}})} \zeta(\mathbf{p}).$$

For example,

$$\zeta^\star(\{2\}^n) = -2\zeta(\overline{2n}) \quad \text{for all } n \ge 1. \tag{10.26}$$

Similarly, using Cor. 9.3.3 we can obtain the following **2**-c-**2** ($c \ge 3$) formula. When all $c = 3$ we get the *Two-three formula*.

Proposition 10.8.4. *Suppose* $\ell, a_j, c_j \in \mathbb{N}_0$ *and* $c_j \geq 3$ *for all* $1 \leq j \leq \ell+1$. *Set* $\mathbf{s} = (\{2\}^{a_1}, c_1, \ldots, \{2\}^{a_\ell}, c_\ell, \{2\}^{a_{\ell+1}})$. *Then*

$$\zeta^\star(\mathbf{s}) = -\sum_{\mathbf{p} \preceq \mathbf{s}'} 2^{\mathrm{dp}(\mathbf{p})} \zeta(\mathbf{p}), \tag{10.27}$$

where $\mathbf{s}' = \big(\overline{2a_1+2}, \{1\}^{c_1-3}, \big\{2a_j+3, \{1\}^{c_j-3}\big\}_{j=2}^{\ell}, 2a_{\ell+1}+1\big)$.

The $\underline{\mathbf{2}}$-$\underline{c}$-$\mathbf{2}$ formula can be proved similarly using Thm. 10.8.1.

Proposition 10.8.5. *Suppose* $\ell, c_j, a_j \in \mathbb{N}_0$ *and* $c_j \geq 3$ *for all* $j \leq \ell$ *and* $a_{\ell+1} \in \mathbb{N}$. *Set* $\mathbf{s} = (\{2\}^{a_1}, c_1, \ldots, \{2\}^{a_\ell}, c_\ell, \{2\}^{a_{\ell+1}})$. *Then we have*

$$\zeta^\star{}_n(\mathbf{s}) = -\sum_{\mathbf{p} \preceq (\overline{2a_1+2}, \{1\}^{c_1-3}, 2a_2+1, \ldots, \{1\}^{c_\ell-3}, 2a_{\ell+1}+1)} \zeta(\mathbf{p}).$$

For example, for all $m, n \geq 0$ and $n \geq 1$, we have the $\underline{\mathbf{2}}$-$\underline{3}$-$\mathbf{2}$ formula

$$\zeta^\star(\{2\}^m, 3, \{2\}^n) = -4\zeta(\overline{2m+2}, 2n+1) - 2\zeta(\overline{2m+2n+3}). \tag{10.28}$$

We leave the proof of the next two propositions to the interested reader.

Proposition 10.8.6. *Let* $\ell \in \mathbb{N}$ *and* $a_j, b_j, c_j - 3 \in \mathbb{N}_0$ *for all* $j \geq 1$. *Then*

($\underline{\mathbf{2}}$-$\underline{c}$-$\underline{\mathbf{2}}$-$\underline{1}$)*: For* $\mathbf{s} = \big(\big\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\big\}_{j=1}^{\ell}\big)$, $\ell \geq 1$, *we have*

$$\zeta^\star(\mathbf{s}) = \sum_{\mathbf{p} \preceq \big(\big\{\overline{2b_j+2}, \{1\}^{c_j-3}, \overline{2a_j+2}\big\}_{j=1}^{\ell}\big)} 2^{\ell(\mathbf{p})} \zeta(\mathbf{p}). \tag{10.29}$$

($\mathbf{2}$-1-$\underline{\mathbf{2}}$-$\underline{c}$-$\underline{\mathbf{2}}$-$\underline{1}$)*: For* $\mathbf{s} = \big(\{2\}^{a_0}, 1, \big\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\big\}_{j=1}^{\ell}\big)$, $\ell \geq 0$ *and* $a_0 \geq 1$, *we have*

$$\zeta^\star(\mathbf{s}) = \sum_{\mathbf{p} \preceq \big(2a_0+1, \big\{\overline{2b_j+2}, \{1\}^{c_j-3}, \overline{2a_j+2}\big\}_{j=1}^{\ell}\big)} 2^{\ell(\mathbf{p})} \zeta(\mathbf{p}). \tag{10.30}$$

For example,we have

$$\zeta^\star(\{2\}^b, 3, \{2\}^a, 1) = 2\zeta(2a+2b+4) + 4\zeta(\overline{2b+2}, \overline{2a+2}), \tag{10.31}$$

and

$$\begin{aligned}\zeta^\star(\{2\}^a, 1, \{2\}^b, 3, \{2\}^c, 1) = {} & 2\zeta(2(a+b+c)+5) + 4\zeta(2a+1, 2b+2c+4) \\ & + 4\zeta(\overline{2a+2b+3}, \overline{2c+2}) + 8\zeta(2a+1, \overline{2b+2}, \overline{2c+2}).\end{aligned}$$

Proposition 10.8.7. *Let $\ell \in \mathbb{N}$ and $a_j, b_j, c_j - 3 \in \mathbb{N}_0$ for all $j \geq 1$. Then*

(**2**-1-**2**-c-**2**). *For $\mathbf{s} = \left(\{\{2\}^{a_j}, 1, \{2\}^{b_j}, c_j\}_{j=1}^{\ell}, \{2\}^t\right)$ with $a_1 \geq 1$ and $t \geq 0$, we have*

$$\zeta^\star(\mathbf{s}) = - \sum_{\mathbf{p} \preceq \left(2a_1+1, \overline{2b_1+2}, \{1\}^{c_1-3}, \left\{\overline{2a_j+2}, \overline{2b_j+2}, \{1\}^{c_j-3}\right\}_{j=2}^{\ell}, 2t+1\right)} 2^{\ell(\mathbf{p})}\zeta(\mathbf{p}). \tag{10.32}$$

(**2**-c-**2**-1-**2**-c-**2**). *For $\mathbf{s} = \left(\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\}_{j=1}^{\ell}, \{2\}^{b_{\ell+1}}, c_{\ell+1}, \{2\}^t\right)$ with $t \geq 0$, we have*

$$\zeta^\star(\mathbf{s}) = - \sum_{\mathbf{p} \preceq \left(\left\{\overline{2b_j+2}, \{1\}^{c_j-3}, \overline{2a_j+2}\right\}_{j=1}^{\ell}, \overline{2b_{\ell+1}+2}, \{1\}^{c_{\ell+1}-3}, 2t+1\right)} 2^{\ell(\mathbf{p})}\zeta(\mathbf{p}). \tag{10.33}$$

For example, taking $r = 1$ and $t = 0$ in case (**2**-1-**2**-c-**2**) we get

$$\begin{aligned}\zeta^\star(\{2\}^a, 1, \{2\}^b, 3) = &- 2\zeta(\overline{2a+2b+4}) - 4\zeta(1+2a, \overline{2b+3}) \\ &-4\zeta(\overline{2a+2b+3}, 1) - 8\zeta(2a+1, \overline{2b+2}, 1).\end{aligned}$$

and in case (**2**-c-**2**-1-**2**-c-**2**) we get

$$\begin{aligned}&\zeta^\star(\{2\}^{b_1}, 3, \{2\}^a, 1, \{2\}^{b_2}, 3) \\ =&- 2\zeta(\overline{2(b_1+a+b_2)+7}) - 4\zeta(2b_1+2a+4, \overline{2b_2+3}) \\ &- 4\zeta(\overline{2b_1+2}, 2a+2b_2+5) - 4\zeta(\overline{2b_1+2a+2b_2+6}, 1) \\ &- 8\zeta(\overline{2b_1+2}, \overline{2a+2}, \overline{2b_2+3}) - 8\zeta(\overline{2b_1+2}, 2a+2b_2+4, 1) \\ &- 8\zeta(2b_1+2a+4, \overline{2b_2+2}, 1) - 16\zeta(\overline{2b_1+2}, \overline{2a+2}, \overline{2b_2+2}, 1).\end{aligned}$$

An interesting corollary is the following formula:

$$\zeta^\star(\{2\}^a, 3, \{2\}^b, 1) + \zeta^\star(\{2\}^b, 3, \{2\}^a, 1) = \zeta^\star(\{2\}^{a+1})\zeta^\star(\{2\}^{b+1}). \tag{10.34}$$

The interested reader can verify all the formulas similar to the above in this section numerically for small values of a_1, a_2, a, b, t using EZ-face (see [75]).

10.9 Three-Two-One formulas

In this section we collect some results concerning MZ*V $\zeta^\star(\mathbf{s})$ with the condition that every component in $\mathbf{s}$ is no larger than 3.

Lemma 10.9.1. *Let $n_1, \dots, n_\ell \in \mathbb{D}$ such that $|n_j|$ is even for every j. Set $m = |n_1| + \cdots + |n_\ell|$. Then*

$$\sum_{g\in\mathfrak{S}_\ell} \zeta\big(n_{g(1)}, \dots, n_{g(\ell)}\big) =$$

$$\sum_{e_1+\cdots+e_p=\ell} (-1)^{\ell-p} \prod_{s=1}^{p} (e_s-1)! \sum \zeta\left(\bigoplus_{k\in\pi_1} n_k\right) \dots \zeta\left(\bigoplus_{k\in\pi_p} n_k\right) \in \mathbb{Q}\pi^m$$

where the sum in the right is taken over all the possible unordered partitions of the set $\{1, \dots, \ell\}$ into p subsets $\pi_1, \dots, \pi_p$ with $e_1, \dots, e_p$ elements respectively.

Proof. When all the arguments $n_1, \dots, n_\ell$ are positive the lemma becomes [275, Thm. 2.2]. Its proof there can be used here almost word for word. Notice the same result is also stated as [355, Prop. 9.4] whose proof is different from that of [275] but also works here. Thus we leave the details to the interested reader. □

Theorem 10.9.2. *Let r be a positive integer, and $e_1, \dots, e_{2r+1}$ nonnegative integers.*

(i) *Put $m = e_1 + \cdots + e_{2r}$. Then the sum*

$$\sum_{\tau\in\mathfrak{S}_{2r}} \zeta^\star(\{2\}^{e_{\tau(1)}}, 3, \{2\}^{e_{\tau(2)}}, 1, \{2\}^{e_{\tau(3)}}, \dots, 3, \{2\}^{e_{\tau(2r)}}, 1)$$

is a rational multiple of π^{2m+4r}.

(ii) *Put $m = e_1 + \cdots + e_{2r+1}$. Then the sum*

$$\sum_{\tau\in\mathfrak{S}_{2r+1}} \zeta^\star(\{2\}^{e_{\tau(1)}}, 3, \{2\}^{e_{\tau(2)}}, 1, \dots, 3, \{2\}^{e_{\tau(2n)}}, 1, \{2\}^{e_{\tau(2r+1)}+1})$$

is a rational multiple of $\pi^{2m+4r+2}$.

Proof. We start with (i) first. When $r = 1$ this follows quickly from Eq. (10.31) by the stuffle relation. For general r let $a_j = e_{2j}$ and $b_j = e_{2j-1}$ for all $j \le r$ and let $A_j = 2e_j + 2$ for all $j \le 2r$. Then we can apply Eq. (10.29) of Thm. 10.8.6 to the string $\mathbf{s} = (\{2\}^{b_1}, 3, \{2\}^{a_1}, 1, \dots, \{2\}^{b_\ell}, 3, \{2\}^{a_\ell}, 1)$ and get

$$\zeta^\star(\mathbf{s}) = \sum_{\mathbf{p}\in\{\overline{A_1}\circ\cdots\circ\overline{A_{2r}}\}} 2^{\ell(\mathbf{p})}\zeta(\mathbf{p}).$$

For $\mathbf{s} = (\{2\}^{e_1}, 3, \{2\}^{e_2}, 1, \dots, \{2\}^{e_{2r-1}}, 3, \{2\}^{e_{2r}}, 1)$ and any permutation $\tau \in \mathfrak{S}_{2r}$ we define

$$\mathbf{s}^\tau = (\{2\}^{e_{\tau(1)}}, 3, \{2\}^{e_{\tau(2)}}, 1, \dots, \{2\}^{e_{\tau(2r-1)}}, 3, \{2\}^{e_{\tau(2r)}}, 1)$$

Let $\mathbf{A} = (\overline{A_1}, \dots, \overline{A_{2r}})$, $\mathbf{A}^\tau = (\overline{A_{\tau(1)}}, \dots, \overline{A_{\tau(2r)}})$, and $P_\ell(2r)$ be the set of all partitions of $[2r] := \{1, 2, \dots, 2r\}$ into ℓ consecutive subsets. If $\lambda = (\lambda_1, \dots, \lambda_\ell) \in P_\ell(2r)$ then we set $\lambda_j(\mathbf{A}^\tau) = (\overline{A_{\tau(i)}})_{i\in\lambda_j}$ so that the concatenation $\bigsqcup_{j=1}^{\ell} \lambda_j(\mathbf{A}^\tau) = \mathbf{A}^\tau$. Because of the permutation we see that

$$\begin{aligned}\sum_{\tau\in\mathfrak{S}_{2r}} \zeta^\star(\mathbf{s}^\tau) &= \sum_{\tau\in\mathfrak{S}_{2r}} \sum_{\ell=1}^{2r} 2^\ell \sum_{\lambda\in P_\ell(2r)} \zeta\big(\oplus\lambda_1(\mathbf{A}^\tau), \dots, \oplus\lambda_\ell(\mathbf{A}^\tau)\big)\\ &= \sum_{\tau\in\mathfrak{S}_{2r}} \sum_{\ell=1}^{2r} \frac{2^\ell}{\ell!} \sum_{\lambda\in P_\ell(2r)} \sum_{g\in\mathfrak{S}_\ell} \zeta\big(\oplus\lambda_{g(1)}(\mathbf{A}^\tau), \dots, \oplus\lambda_{g(\ell)}(\mathbf{A}^\tau)\big),\end{aligned}$$

where $\oplus\mathbf{t}$ is the $\oplus$-sum of all the components of $\mathbf{t}$ for any composition $\mathbf{t}$. Hence Thm. 10.9.2(i) follows readily from the Lemma 10.9.1 since all A_j's are even numbers.

Thm. 10.9.2 (ii) follows from Thm. 12.9.9 in a similar fashion so we leave the details to the interested reader. □

Theorem 10.9.3. *For any positive integers n and m, we have*

$$\zeta^\star(\{2\}^m, 1)\zeta^\star(\{2\}^n, 1) = \zeta^\star(\{2\}^m, 1, \{2\}^n, 1) + \zeta^\star(\{2\}^n, 1, \{2\}^m, 1), \tag{10.35}$$

$$\zeta^\star(\{2\}^m, 1)\zeta^\star(\{2\}^n) = \zeta^\star(\{2\}^m, 1, \{2\}^n) + \zeta^\star(\{2\}^{n-1}, 3, \{2\}^m), \tag{10.36}$$

$$\zeta^\star(\{2\}^m)\zeta^\star(\{2\}^n) = \zeta^\star(\{2\}^{m-1}, 3, \{2\}^{n-1}, 1) + \zeta^\star(\{2\}^{n-1}, 3, \{2\}^{m-1}, 1). \tag{10.37}$$

Proof. By Prop. 10.8.2 and Eq. (10.26) we have

$$\begin{aligned}\frac{1}{4}\zeta^\star(\{2\}^m, 1, \{2\}^n, 1) &= \zeta(2m+1, 2n+1) + \frac{1}{2}\zeta(2m+2n+2),\\ \frac{1}{4}\zeta^\star(\{2\}^m, 1)\cdot\zeta^\star(\{2\}^n, 1) &= \zeta(2m+1)\zeta(2n+1)\\ &= \zeta(2m+1, 2n+1) + \zeta(2m+2n+2) + \zeta(2n+1, 2m+1),\\ -\frac{1}{4}\zeta^\star(\{2\}^m, 1, \{2\}^n) &= \zeta(2m+1, \overline{2n}) + \frac{1}{2}\zeta(\overline{2m+2n+1}),\\ -\frac{1}{4}\zeta^\star(\{2\}^m, 1)\cdot\zeta^\star(\{2\}^n) &= \zeta(2m+1)\zeta(\overline{2n})\\ &= \zeta(2m+1, \overline{2n}) + \zeta(\overline{2m+2n+1}) + \zeta(\overline{2n}, 2m+1),\end{aligned}$$

$$\begin{aligned}\frac{1}{4}\zeta^\star(\{2\}^m)\cdot\zeta^\star(\{2\}^n) &= \zeta(\overline{2m})\zeta(\overline{2n})\\ &= \zeta(\overline{2m},\overline{2n})+\zeta(\overline{2m+2n})+\zeta(\overline{2n},\overline{2m}),\end{aligned}$$

by the stuffle relations. On the other hand, by Eq. (10.31) and Eq. (10.28), we get

$$\begin{aligned}\frac{1}{4}\zeta^\star(\{2\}^{n-1},3,\{2\}^{m-1},1) &= \frac{1}{2}\zeta(2m+2n)+\zeta(\overline{2n},\overline{2m}),\\ \frac{1}{4}\zeta^\star(\{2\}^{m-1},3,\{2\}^{n-1},1) &= \frac{1}{2}\zeta(2m+2n)+\zeta(\overline{2m},\overline{2n}),\\ -\frac{1}{4}\zeta^\star(\{2\}^{n-1},3,\{2\}^{m}) &= \zeta(\overline{2n},2m+1)+\frac{1}{2}\zeta(\overline{2m+2n+1}).\end{aligned}$$

The rest of the proof is straight-forward and is thus left to the interested reader. □

Theorem 10.9.4. *Let m and n be two nonnegative integers.*

(i) *We have*

$$(2n+1)\zeta^\star(\{3,1\}^n,2)=\sum_{j+k=n}\zeta^\star(\{3,1\}^j)\zeta^\star(\{2\}^{2k+1}).$$

(ii) *If $n\geq 1$ then we have*

$$\begin{aligned}&\sum_{\substack{e_1+e_2+\cdots+e_{2n}=1\\ e_1,e_2,\ldots,e_{2n}\geq 0}}\zeta^\star(\{2\}^{e_1},3,\{2\}^{e_2},1,\ldots,\{2\}^{e_{2n-1}},3,\{2\}^{e_{2n}},1)\\ &\qquad=\sum_{j+k=n-1}\zeta^\star(\{3,1\}^j,2)\zeta^\star(\{2\}^{2k+2}).\end{aligned}$$

Remark 10.9.5. Notice that Thm. 10.9.4(ii) is related to Thm. 10.9.11 but is not a consequence of it. Also notice there is no string of 2's at the end of any MZ*V in (i) while there is at least one 2 at the end of every MZ*V in (ii). Further, the statements in Thm. 10.9.4 is generally false if $\zeta^\star$ is replaced by ζ.

Proof. (i). By removing all the 2's in **2**-3-**2**-1 formula we get

$$\zeta^\star(\{3,1\}^j)=\sum_{\mathbf{p}_{2j}\preceq\left(\{\overline{2}\}^{2j}\right)}2^{\ell(\mathbf{p}_{2j})}\zeta(\mathbf{p}_{2j}).$$

All of the components a_j of $\{\overline{2}\circ\cdots\circ\overline{2}\}$ must satisfy the following sign rule:

$$a_j>0 \text{ if and only if } 4|a_j. \tag{10.38}$$

On the other hand, by removing all the 2's in **2**-*c*-**2**-1-**2** formula except the last 2, we have

$$\zeta^\star(\{3,1\}^n,2) = -\sum_{\mathbf{p}_{2n+1}\preceq\left(\{\overline{2}\}^{2n+1}\right)} 2^{\ell(\mathbf{p}_{2n+1})}\zeta(\mathbf{p}_{2n+1}).$$

Hence by Eq. (10.26) we need to show that

$$(2n+1)\sum_{\mathbf{p}_{2n+1}\preceq\left(\{\overline{2}\}^{2n+1}\right)} 2^{\ell(\mathbf{p}_{2n+1})}\zeta(\mathbf{p}_{2n+1}) = \sum_{j=0}^{n}\sum_{\mathbf{p}_{2j}\preceq\left(\{\overline{2}\}^{2j}\right)} 2^{\ell(\mathbf{p}_{2j})}\zeta(\mathbf{p}_{2j})\cdot 2\zeta(\overline{4(n-j)+2}). \quad (10.39)$$

Suppose an index $\mathbf{p}_{2n+1}\preceq\left(\{\overline{2}\}^{2n+1}\right)$ has length t $(1\le t\le 2n+1)$ given as

$$(a_1,\dots,a_t),\quad a_i\in\mathbb{D}\ \forall i=1,\dots,t.$$

We now show that there are exactly $2^t(2n+1)$ copies of such term produced by the stuffle product on the right-hand side of Eq. (10.39). Indeed, for each $i=1,\dots,t$ the entry a_i has two possibilities:

(1). $a_i=4b_i>0$. Then for each $k=1,\dots,b_i$ we may produce such a term on the right-hand side of Eq. (10.39) by stuffing

$$2^t\zeta(a_1,\dots,a_{i-1},\overline{4k-2},a_{i+1},\dots,a_t)$$

from $\mathbf{p}_{2j}$ having length t with the term $2\zeta(\overline{4(b_i-k)+2})$ at the right end of Eq. (10.39). Notice no shuffle is possible since $4(n-j)+2$ is not a multiple of 4. Hence these contribute to $2^{t+1}b_i=2^{t-1}a_i$ copies of $\zeta(a_1,\dots,a_t)$.

(2). $a_i=\overline{4b_i+2}$. Then for each $k=1,\dots,b_i$ we may produce such a term on the right-hand side of Eq. (10.39) by stuffing

$$2^t\zeta(a_1,\dots,a_{i-1},\overline{4k},a_{i+1},\dots,a_t)$$

from $\mathbf{p}_{2j}$ having length t with the term $2\zeta(\overline{4(b_i-k)+2})$ at the right end of Eq. (10.39). Further, there is exactly one possible shuffle given by

$$2^{t-1}\zeta(a_1,\dots,a_{i-1},a_{i+1},\dots,a_{t+1})\sqcup\!\sqcup\left\{2\zeta(\overline{4b_i+2})\right\},$$

since the index $(a_1,\dots,a_{i-1},a_{i+1},\dots,a_t)$ has only length $t-1$. Altogether these produce $2^{t+1}b_i+2^t=2^{t-1}|a_i|$ copies of $\zeta(a_1,\dots,a_t)$.

By combining (1) and (2) we see that the right-hand side of Eq. (10.39) produces exactly

$$\sum_{i=1}^{t}2^{t-1}|a_i| = 2^{t-1}\cdot|(a_1,\dots,a_t)| = 2^t(2n+1)$$

copies of $\zeta(a_1,\ldots,a_t)$ since the weight is $4n+2$. This proves (i).

(ii). We use the same analysis as above and see that we need to prove the following identity:

$$\sum_{\mathbf{q}_{2n}} 2^{\ell(\mathbf{q}_{2n})}\zeta(\mathbf{q}_{2n}) = \sum_{j=0}^{n-1} \sum_{\mathbf{p}_{2j+1}\preceq\left(\{\overline{2}\}^{2j+1}\right)} 2^{\ell(\mathbf{p}_{2j+1})}\zeta(\mathbf{p}_{2j+1})\cdot 2\zeta(\overline{4(n-j)}), \tag{10.40}$$

where $\mathbf{q}_{2n}$ runs through all indices of the form $A_1\circ\cdots\circ A_{2n}$ with one of the A_j's (say A_{j_0}) equal to $\overline{4}$ and all the other A_j's equal to $\overline{2}$. For each choice of $2^t\zeta(a_1,\ldots,a_{t+1})$ with length $t+1$ from the left-hand side of Eq. (10.40), all but one of the argument components $a_1,\ldots,a_{t+1}$ must satisfy the sign rule Eq. (10.38). The only exceptional component, say a_i, must involve a merge with the special entry $A_{j_0}=\overline{4}$. Now there are two possibilities:

(1). $a_i = 4b_i+2>0$. Then for each $k=0,\ldots,b_i-1$ we may produce such a term on the right-hand side of Eq. (10.40) by stuffing

$$2^t\zeta(a_1,\ldots,a_{i-1},\overline{4k+2},a_{i+1},\ldots,a_{t+1})$$

from $\mathbf{p}_{2j+1}$ having length $t+1$ with the term $2\zeta(\overline{4(b_i-k)})$ at the right end of (10.40). Notice no shuffle is possible since $4(n-j)$ is a multiple of 4. Hence these contribute to $2^{t+1}b_i$ copies of $\zeta(a_1,\ldots,a_{t+1})$. On the left-hand side, such a term must be produced by setting all $2b_i-1$ consecutive $\circ$'s around $A_{j_0}=\overline{4}$ to $\oplus$:

$$\ldots,\underbrace{A_i\oplus A_{i+1}\oplus\cdots\oplus A_{j_0}\oplus\cdots\oplus A_\ell}_{2b_i\text{ entries}},\ldots.$$

But A_{j_0} can be at any one of the $2b_i$ possible positions, thus producing $2^{t+1}b_i$ copies of $\zeta(a_1,\ldots,a_{t+1})$ which match exactly the right-hand side of (10.40).

(2). $a_i=\overline{4b_i}$. Then for each $k=1,\ldots,b_i-1$ we may produce such a term on the right-hand side of (10.40) by stuffing

$$2^t\zeta(a_1,\ldots,a_{i-1},4k,a_{i+1},\ldots,a_{t+1})$$

from $\mathbf{p}_{2j}$ having length $t+1$ with the term $2\zeta(\overline{4(b_i-k)})$ at the right end of Eq. (10.40). Further, there is exactly one possible shuffle given by

$$2^{t-1}\zeta(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_{t+1})\sqcup\!\sqcup\left\{2\zeta(\overline{4b_i})\right\}.$$

Hence these contribute to $2^{t+1}(b_i-1)+2^t = 2^t(2b_i-1)$ copies of $\zeta(a_1,\ldots,a_{t+1})$. Similar to (1), on the left-hand side, such a term must be produced by setting all $2b_i-2$ consecutive $\circ$'s around A_{j_0} to $\oplus$. And

A_{j_0} can be at any one of the $2b_i - 1$ possible positions, thus producing $2^t(2b_i - 1)$ copies of $\zeta(a_1, \ldots, a_{t+1})$ which match exactly the right-hand side of Eq. (10.40).

This concludes the proof of theorem. □

Next we will prove a more general formula similar to Thm. 10.9.3. First, we introduce some more notation. For $\infty \geq A \geq B \geq 1$, we put

$$S_{-1}(A, B) = \delta_{A,B} A^2, \qquad S_0(A, B) = 1,$$

$$S_j(A, B) = \sum_{A \geq a_1 \geq \cdots \geq a_j \geq B} \frac{1}{a_1^2 \cdots a_j^2} \qquad (j = 1, 2, 3, \ldots),$$

where $\delta_{A,B}$ is the Kronecker symbol. Then we have

$$S_j(A, B) = \sum_{p=B}^{A} \frac{1}{p^2} S_{j-1}(p, B) = \sum_{p=B}^{A} S_{j-1}(A, p) \frac{1}{p^2} \tag{10.41}$$

for any $j \geq 0$. We now prove a lemma by using partial fractions.

Lemma 10.9.6. *For any $j \geq -1$ and $1 \leq p, q \leq \infty$, we have*

$$\sum_{p_0=1}^{p} S_j(p, p_0) \frac{q}{p_0(p_0 + q)} = \sum_{q_0=1}^{q} S_j(q, q_0) \frac{p}{q_0(q_0 + p)}. \tag{10.42}$$

Here, $\frac{q}{p_0(p_0+q)}$ (resp. $\frac{p}{q_0(q_0+p)}$) means $\frac{1}{p_0}$ (resp. $\frac{1}{q_0}$) if $q = \infty$ (resp. $p = \infty$), and so on.

Proof. We use induction on j. When $j = -1$, both sides are equal to $\frac{pq}{p+q}$. Thus we assume $j \geq 0$. By Eq. (10.41), we can rewrite the left-hand side of Eq. (10.42) as

$$\begin{aligned}\sum_{p_0=1}^{p} S_j(p, p_0) \frac{q}{p_0(p_0 + q)} &= \sum_{p_0=1}^{p} \sum_{p_1=p_0}^{p} S_{j-1}(p, p_1) \frac{1}{p_1^2} \left(\frac{1}{p_0} - \frac{1}{p_0 + q} \right) \\ &= \sum_{p_1=1}^{p} S_{j-1}(p, p_1) \frac{1}{p_1^2} \sum_{p_0=1}^{p_1} \left(\frac{1}{p_0} - \frac{1}{p_0 + q} \right).\end{aligned}$$

By using the identity

$$\begin{aligned}&\sum_{p_0=1}^{p_1} \left(\frac{1}{p_0} - \frac{1}{p_0 + q} \right) \\ &= \sum_{p_0=1}^{\infty} \left\{ \left(\frac{1}{p_0} - \frac{1}{p_0 + q} \right) - \left(\frac{1}{p_0 + p_1} - \frac{1}{p_0 + p_1 + q} \right) \right\} \\ &= \sum_{p_0=1}^{q} \left(\frac{1}{p_0} - \frac{1}{p_0 + p_1} \right) = \sum_{q_1=1}^{q} \frac{p_1}{q_1(q_1 + p_1)},\end{aligned}$$

we get

$$\sum_{p_1=1}^{p}\frac{S_{j-1}(p,p_1)}{p_1^2}\sum_{p_0=1}^{p_1}\left(\frac{1}{p_0}-\frac{1}{p_0+q}\right)=\sum_{p_1=1}^{p}\frac{S_{j-1}(p,p_1)}{p_1^2}\sum_{q_1=1}^{q}\frac{p_1}{q_1(q_1+p_1)}$$
$$=\sum_{q_1=1}^{q}\frac{1}{q_1^2}\sum_{p_1=1}^{p}\frac{q_1S_{j-1}(p,p_1)}{p_1(p_1+q_1)}.$$

Finally, using the induction hypothesis and Eq. (10.41) again, we obtain

$$\sum_{q_1=1}^{q}\frac{1}{q_1^2}\sum_{p_1=1}^{p}\frac{q_1S_{j-1}(p,p_1)}{p_1(p_1+q_1)}=\sum_{q_1=1}^{q}\frac{1}{q_1^2}\sum_{q_0=1}^{q_1}\frac{pS_{j-1}(q_1,q_0)}{q_0(q_0+p)}$$
$$=\sum_{q_0=1}^{q}\frac{pS_j(q,q_0)}{q_0(q_0+p)}.$$

Thus we have shown Eq. (10.42). □

For $n\in\mathbb{N}$, let $\mathbf{j}=(j_1,\dots,j_n)\in(\mathbb{N}_0)^n$ and $\mathbf{e}=(e_1,\dots,e_n)$ where $e_j=1$ or 3. Define

$$\mathcal{C}(\mathbf{j};\mathbf{e}):=\zeta^\star\big(\{2\}^{j_1},e_1,\{2\}^{j_2},e_2,\dots,e_{n-1},\{2\}^{j_n}\big). \tag{10.43}$$

For $n=0$, we simply put $\mathcal{C}=1$. Note that the right-hand side of Eq. (10.43) diverges if and only if $n\geq 2$, $j_1=0$ and $e_1=1$. Otherwise, $(\mathbf{j},\mathbf{e})$ is called an *admissible pair*.

For convenience we put $j_{a\uparrow b}=(j_a,j_{a+1},\dots,j_b)$ and $j_{b\downarrow a}=(j_b,j_{b-1},\dots,j_a)$ for any $a\leq b$.

Theorem 10.9.7. *Let n be a positive integer, $\mathbf{j}=(j_1,\dots,j_n)\in(\mathbb{N}_0)^n$ and $\mathbf{e}=(e_1,\dots,e_n)$ with $e_j=1$ or 3 for all j. Assume that both $(\mathbf{j},\mathbf{e})$ and $(\overleftarrow{\mathbf{j}},\overleftarrow{\mathbf{e}})$ are admissible pairs. For $k=0,\dots,n$, if $e_k=1$ define*

$$X(k):=\mathcal{C}(j_{1\uparrow k},0;e_{1\uparrow k})\cdot\mathcal{C}(j_{n\downarrow(k+1)};e_{(n-1)\downarrow k})$$

and if $e_k=3$ define

$$X(k):=\mathcal{C}(j_{1\uparrow(k-1)},j_k+1;e_{1\uparrow(k-1)})\cdot\mathcal{C}(j_{n\downarrow(k+2)},j_{k+1}+1;e_{(n-1)\downarrow(k+1)}).$$

Then we have

$$\sum_{k=0}^{n}(-1)^kX(k)=0.$$

Proof. Set $e_0 = e_n = 1$. For $k = 0, \dots, n$, put

$$E(k) = \sum_{\substack{p_0 \geq \cdots \geq p_k \geq 1 \\ q_{n+1} \geq q_n \geq \cdots \geq q_{k+1} \geq 1}} \prod_{\alpha=1}^{k} \frac{S_{j_\alpha}(p_{\alpha-1}, p_\alpha)}{p_\alpha^{e_\alpha}} \prod_{\beta=k+1}^{n} \frac{S_{j_\beta}(q_{\beta+1}, q_\beta)}{q_\beta^{e_{\beta-1}}} \cdot \frac{p_k^{e_k-1} q_{k+1}}{p_k + q_{k+1}},$$

$$F(k) = \sum_{\substack{p_0 \geq \cdots \geq p_k \geq 1 \\ q_{n+1} \geq q_n \geq \cdots \geq q_{k+1} \geq 1}} \prod_{\alpha=1}^{k} \frac{S_{j_\alpha}(p_{\alpha-1}, p_\alpha)}{p_\alpha^{e_\alpha}} \prod_{\beta=k+1}^{n} \frac{S_{j_\beta}(q_{\beta+1}, q_\beta)}{q_\beta^{e_{\beta-1}}} \cdot \frac{p_k q_{k+1}^{e_k-1}}{p_k + q_{k+1}}.$$

Here we allow $p_0 = \infty$ and $q_{n+1} = \infty$. In particular, we have $E(0) = 0$ (resp. $F(n) = 0$) because of the factor $\frac{q_1}{\infty + q_1}$ (resp. $\frac{p_n}{p_n + \infty}$). For general k, we see that $E(k) + F(k) = X(k)$ since

$$\frac{p_k^{e_k-1} q_{k+1}}{p_k + q_{k+1}} + \frac{p_k q_{k+1}^{e_k-1}}{p_k + q_{k+1}} = \begin{cases} 1, & \text{if } e_k = 1; \\ p_k q_{k+1}, & \text{if } e_k = 3. \end{cases}$$

Moreover, when $1 \leq k \leq n$, one has

$$\sum_{p_k=1}^{p_{k-1}} \frac{S_{j_k}(p_{k-1}, p_k)}{p_k^{e_k}} \frac{p_k^{e_k-1} q_{k+1}}{p_k + q_{k+1}} = \sum_{q_k=1}^{q_{k+1}} \frac{S_{j_k}(q_{k+1}, q_k)}{q_k^{e_{k-1}}} \frac{p_{k-1} q_k^{e_{k-1}-1}}{p_{k-1} + q_k}$$

by Lemma 10.9.6, hence $E(k) = F(k-1)$. Therefore,

$$\begin{aligned} \sum_{k=0}^{n} (-1)^k X(k) &= \sum_{k=0}^{n} (-1)^k \big(E(k) + F(k)\big) \\ &= \sum_{k=1}^{n} (-1)^k F(k-1) + \sum_{k=0}^{n-1} (-1)^k F(k) = 0 \end{aligned}$$

as desired. □

Example 10.9.8. Taking $\mathbf{e} = (3,3)$ or $\mathbf{e} = (3,1)$ in Thm. 10.9.7, one has

$$\begin{aligned} &\zeta^\star(\{2\}^{j_1}, 3, \{2\}^{j_2}, 3, \{2\}^{j_3}, 1) + \zeta^\star(\{2\}^{j_1+1}) \cdot \zeta^\star(\{2\}^{j_3}, 3, \{2\}^{j_2+1}) \\ &\quad = \zeta^\star(\{2\}^{j_1}, 3, \{2\}^{j_2+1}) \cdot \zeta^\star(\{2\}^{j_3+1}) + \zeta^\star(\{2\}^{j_3}, 3, \{2\}^{j_2}, 3, \{2\}^{j_1}, 1) \end{aligned}$$

or

$$\begin{aligned} &\zeta^\star(\{2\}^{j_1}, 3, \{2\}^{j_2}, 1, \{2\}^{j_3}, 1) + \zeta^\star(\{2\}^{j_1+1}) \cdot \zeta^\star(\{2\}^{j_3}, 1, \{2\}^{j_2+1}) \\ &\quad = \zeta^\star(\{2\}^{j_1}, 3, \{2\}^{j_2}, 1) \cdot \zeta^\star(\{2\}^{j_3}, 1) + \zeta^\star(\{2\}^{j_3}, 1, \{2\}^{j_2}, 3, \{2\}^{j_1}, 1), \end{aligned}$$

respectively. Of course, these two relations can by proved by applying Thm. 10.8.1 to each of the MZ*Vs and then using stuffle relations.

Corollary 10.9.9. *Let $n > 0$ be an integer. Then*

(i) *For any* $a_1, a_2, \ldots, a_n \in \mathbb{N}_0$ *such that* $a_1, a_n \geq 1$, *we have*

$$\sum_{k=0}^{n}(-1)^k\zeta^\star\Big(\big\{\{2\}^{a_j},1\big\}_{j=1}^{k}\Big)\cdot\zeta^\star\Big(\big\{\{2\}^{a_j},1\big\}_{j=n}^{k+1}\Big)=0. \tag{10.44}$$

Here $\zeta^\star\Big(\big\{\{2\}^{a_j},1\big\}_{j=1}^{0}\Big)=\zeta^\star\Big(\big\{\{2\}^{a_j},1\big\}_{j=n}^{n+1}\Big)=1.$

(ii) *For any* $a_1, a_2, \ldots, a_{2n} \in \mathbb{N}_0$, *we have*

$$\begin{aligned}&\sum_{k=0}^{n}\zeta^\star(\{2\}^{a_1},3,\{2\}^{a_2},1,\ldots,3,\{2\}^{a_{2k}},1)\times\\ &\qquad\qquad\times\zeta^\star(\{2\}^{a_{2n}},3,\{2\}^{a_{2n-1}},1,\ldots,3,\{2\}^{a_{2k+1}},1)\\ =&\sum_{k=1}^{n}\zeta^\star(\{2\}^{a_1},3,\{2\}^{a_2},1,\ldots,1,\{2\}^{a_{2k-1}+1})\times\\ &\qquad\qquad\times\zeta^\star(\{2\}^{a_{2n}},3,\{2\}^{a_{2n-1}},1,\ldots,1,\{2\}^{a_{2k}+1}).\end{aligned} \tag{10.45}$$

Proof. Taking $\mathbf{e}=(\{1\}^n)$ we get

$$\begin{aligned}X(k)&=\mathcal{C}\big(j_{1\uparrow k},0;\{1\}^k\big)\cdot\mathcal{C}\big(j_{n\downarrow(k+1)},0;\{1\}^{n-k}\big)\\&=\zeta^\star(\{2\}^{j_1},1,\ldots,\{2\}^{j_k},1)\cdot\zeta^\star(\{2\}^{j_n},1,\ldots,\{2\}^{j_{k+1}},1).\end{aligned}$$

Hence Thm. 10.9.7 implies the identity Eq. (10.44). Similarly, for $\mathbf{e}=(\{3,1\}^n)$, we have

$$\begin{aligned}X(k)=&\zeta^\star(\{2\}^{j_1},3,\{2\}^{j_2},1,\ldots,3,\{2\}^{j_k},1)\\&\times\zeta^\star(\{2\}^{j_{2n}},3,\{2\}^{j_{2n-1}},1,\ldots,3,\{2\}^{j_{k+1}},1)\end{aligned}$$

if k is even, and

$$\begin{aligned}X(k)=&\zeta^\star(\{2\}^{j_1},3,\{2\}^{j_2},1,\ldots,3,\{2\}^{j_{k-1}},1,\{2\}^{j_k+1})\\&\times\zeta^\star(\{2\}^{j_{2n}},3,\{2\}^{j_{2n-1}},1,\ldots,3,\{2\}^{j_{k+2}},1,\{2\}^{j_{k+1}+1})\end{aligned}$$

if k is odd. This proves Eq. (10.45). □

Proposition 10.9.10. *For positive integers* m *and* n, *we have*

$$\zeta^\star\big(\{2m\}^n\big)\;=\;\pi^{2mn}\sum_{\substack{\sum_{j=0}^{m-1}n_j=mn\\ n_i\geq 0}}(-1)^{mn}\exp\left(\frac{2\pi i}{m}\sum_{l=0}^{m-1}ln_l\right)\prod_{k=0}^{m-1}\beta(n_k)$$

where $\beta(n)=(2-2^{2n})B_{2n}/(2n)!$.

Proof. Using the infinite product for the sine function, we have

$$\csc \pi x e^{\frac{\pi i}{m}k} = \left(\pi x e^{\frac{\pi i}{m}k} \prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2}e^{\frac{2\pi i}{m}k}\right)\right)^{-1}.$$

Substituting $k = 0, 1, \ldots, m-1$ and multiplying both sides, we obtain

$$\left(\pi^m x^m \prod_{k=0}^{m-1} e^{\frac{\pi i}{m}k}\right) \prod_{k=0}^{m-1} \csc \pi x e^{\frac{\pi i}{m}k} = \prod_{n=1}^{\infty}\left(1-\frac{x^{2m}}{n^{2m}}\right)^{-1}. \tag{10.46}$$

The right-hand side of Eq. (10.46) equals

$$1 + \left(\sum_{n_1>0} \frac{1}{n_1^{2m}}\right) x^{2m} + \cdots = 1 + \sum_{n=1}^{\infty} \zeta^\star(\{2m\}^n) x^{2mn}.$$

On the other hand, the left-hand side of Eq. (10.46) equals

$$\left(\pi^m x^m \prod_{k=0}^{m-1} e^{\frac{\pi i}{m}k}\right) \prod_{k=0}^{m-1} \sum_{n_k=0}^{\infty} (-1)^{n_k-1} \frac{(2^{2n_k}-2)B_{2n_k}}{(2n_k)!} \times$$

$$\times \pi^{2n_k-1} x^{2n_k-1} e^{\frac{\pi i}{m}k(2n_k-1)}$$

$$= \prod_{k=0}^{m-1} \sum_{n_k=0}^{\infty} (-1)^{n_k-1} \frac{(2^{2n_k}-2)B_{2n_k}}{(2n_k)!} \pi^{2n_k} x^{2n_k} e^{\frac{2\pi i}{m}kn_k}$$

$$= 1 + \sum_{n=1}^{\infty} \sum_{\sum_{j=0}^{m-1} n_j = mn} (-1)^{m(n-1)} (\pi x)^{2mn} \exp\left(\frac{2\pi i}{m} \sum_{j=0}^{m-1} j n_j\right) \prod_{j=0}^{m-1} \beta(n_j).$$

Comparing coefficients of both sides, we obtain the desired identity. □

Theorem 10.9.11. *For all nonnegative integers m and n we have*

$$\zeta^\star(\{2\}^m \sqcup\!\sqcup \{3,1\}^n) \in \mathbb{Q}\pi^{2m+4n}.$$

[347]

10.10 Historical notes

Even at the early stage of Euler's study of MZV he in fact considered the not MZV but the so-called multiple zeta star values (MZ*V). Notice sometimes these values were also called non-strict MZV in some older literature (for e.g., [433]). Instead of Eq. (1.23) Euler actually showed

$$2\zeta^\star(n,1) = (n+2)\zeta(n+1) - \sum_{i=1}^{n-2} \zeta(n-i)\zeta(i+1), \tag{10.47}$$

where $\zeta^\star(n,1)$ is a double zeta star value. See [198].

In [275] Hoffman established many fundamental results concerning both MZVs and MZ*Vs. Since then MZ*Vs frequently appear in the literature although they do not seem to enjoy the same fame as their cousins, MZVs. Most often one can reduce the study of MZ*Vs to that of MZVs but sometimes MZ*V version is nicer and easier while other times the opposite is true. To make things both more complicated and more interesting, the symmetrized MZVs (see Chap. 6) are even more elegant than both MZVs and MZ*Vs in some situations.

The integral expression of MZ*Vs in Thm. 10.1.1 is first stated in Yamamoto's paper [562]. Using this idea, he gave a few very nice relations between the MZ*Vs and Arakawa-Kaneko zeta values (see [19]).

In [275] Hoffman applied some combinatorial theory to MZ*V and proved Thm. 10.2.2. Machide [397] then generalized the MZV part of this theorem to regularized MZVs of weight ≤ 4. The sum relation stated as Eq. (10.7) is first conjectured by C. Moen in an unpublished manuscript according to [275] by Hoffman. It was initially verified for depth up to three by Moen and by Hoffman for weight up to six. Our proof above closely follows the argument of [275]. Ohno and Wakabayashi [435] proved the cyclic sum relation of Thm. 10.4.1 for MZ*V by using partial fractions. In [515], Tanaka and Wakabayashi used Hoffman's algebra machinery to prove some more results related to the cyclic sum formula.

In [12] Aoki and Ohno used generating functions to prove the sum formula of MZ*Vs of fixed weight and height, which is the content of Cor. 10.6.5. Their method was extended to handle the case with the depth also fixed in [11]. This is the source of Thm. 10.6.2.

The identity Eq. (10.10) was proved by Zlobin [618]. He also mentioned it was essentially proved by Vasilev [539]. Cor. 10.7.6 was first proved by Kaneko and Ohno [323, Thm. 1.3]. Another proof was given by Yamazaki [563]. It was later generalized to arbitrary height by Z. Li [380] whose proof is used here to show Thm. 10.7.3.

Prop. 10.8.2 was discovered by Ohno and Zudilin [437] numerically first. They called the identity "Two-one formula", denoted by **2**-1 formula here. In [561] Yamamoto considers some algebraic structures depending on a variable t which reflect the properties of MZV and MZSV when $t = 0$ and $t = 1$, respectively. As he pointed out (see [560, Conjecture 4.4]) the validity of the **2**-1 formula implies that the algebra structure of MZSVs of the form $\zeta^\star(\{2\}^{a_1}, 1, \ldots, \{2\}^{a_\ell}, 1)$ is reflected by setting $t = 1/2$.

Equation (10.26) was first proved by Zlobin [616]. Equation (10.35) was

proved by Ohno and Zudilin [437] by partial fractions. Thm. 10.9.7 is due to Tasaka and Yamamoto [517].

Thm. 10.9.2 was first conjectured by Imatomi *et al.* as Conjectures 4.1 and 4.3 of [301]. Some special cases were proved as [301, Thm. 1.1] by Imatomi *et al.* and also in [517] by Tasaka and Yamamoto. Later, Yamamoto [560] proved a more precise version. Here we have given a different and concise proof. We notice that in [560, Thm. 1.1] Yamamoto obtains a more precise formula by using partial sums and generating functions:

$$\begin{aligned}&\sum_{\substack{e_0,e_1,\ldots,e_{2r}\geq 0\\ e_0+e_1+\cdots+e_{2r}=m}} \zeta^\star(\{2\}^{e_0},3,\{2\}^{e_1},1,\{2\}^{e_2},3,\ldots,3,\{2\}^{e_{2r-1}},1,\{2\}^{e_{2r}})\\ &=\sum_{\substack{2i+k+u=2r\\ j+l+v=m}}(-1)^{j+k}\binom{k+l}{k}\binom{u+v}{u}\binom{2i+j}{j}\frac{\beta'_{k+l}\beta'_{u+v}\pi^{4r+2m}}{(2i+1)(4i+2j+1)!},\end{aligned} \tag{10.48}$$

where $\beta'_n=(-1)^n(2-2^{2n})B_{2n}/(2n)!$. It is possible to modify the proof of Thm. 10.9.2 to give this more quantified version.

The identities in Thm. 10.9.4 were first conjectured by Imatomi *et al.* as [301, Conjecture 4.5]. Note that Thm. 10.9.4(i) is the more precise version of the $n=1$ case of Thm. 10.9.2(i). And Thm. 10.9.4(iii) can be written more compactly as

$$\zeta^\star\big(\{2\}\sqcup\{3,1\}^n\big)=\sum_{k=0}^{n}\zeta^\star(\{3,1\}^{n-k},2)\zeta^\star(\{2\}^{2k}),$$

which is the more precise version of the $m=1$ case of Thm. 10.9.11, a result of Kondo *et al.* [347]. The case $m=0$ case has the following precise formulation by Muneta [418]:

$$\zeta^\star\big(\{3,1\}^n\big)=\sum_{i=0}^{n}\left\{\frac{2}{(4i+2)!}\sum_{\substack{n_0+n_1=2(n-i)\\ n_0,n_1\geq 0}}(-1)^{n_1}\beta(n_0)\beta(n_1)\right\}\pi^{4n},$$

where $\beta(n)=(2-2^{2n})B_{2n}/(2n)!$. Muneta also found precise form in case $m=1$ and Prop. 10.9.10. Of course, these are all special cases of Yamamoto's general formula Eq. (10.48). Tasaka and Yamamoto proved Thm. 10.9.3 in [517].

Exercises

10.1. Prove Eq. (10.14) by using geometric series.

10.2. Using Lemma C.28 verify that the two functions in Eq. (10.15) are indeed the solutions of the corresponding homogeneous differential equation (10.13). Then verify that Eq. (10.16) gives a solution of Eq. (10.13). Notice that the homogeneous equation is a special case of Riemann' differential equations with the singularities at $0, 1, \infty$ (see Sec. C.1).

10.3. Prove the transformation Eq. (10.17) using the linear transformation Eq. (C.5) with $(a, b, c) = (1-\beta, 1-\beta+x, 2-y)$ followed by the linear transformation Eq. (C.4).

10.4. Prove Prop. 10.8.7 and Prop. 10.8.6 by repeatedly applying Thm. 10.8.1.

10.5. Prove that for $\mathbf{s} = \left(\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\}_{j=1}^{\ell}\right)$, $\ell \geq 1$, we have

$$\zeta^{\star}(\mathbf{s}) = \sum_{\mathbf{p} \preceq \left(\{\overline{2b_j+2}, \{1\}^{c_j-3}, \overline{2a_j+2}\}_{j=1}^{\ell}\right)} 2^{\ell(\mathbf{p})}\zeta(\mathbf{p}).$$

10.6. Let $t, \ell \in \mathbb{N}$ and $a_j, b_j, c_j - 3 \in \mathbb{N}_0$ for all $j \geq 1$. Prove the following two general formulas.

(**2**-*c*-**2**-1-**2**). For $\mathbf{s} = \left(\{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\}_{j=1}^{\ell}, \{2\}^{t}\right)$ with $t, \ell \geq 1$, show that

$$\zeta^{\star}(\mathbf{s}) = -\sum_{\mathbf{p} \preceq \left(\{\overline{2b_j+2}, \{1\}^{c_j-3}, \overline{2a_j+2}\}_{j=1}^{\ell}, \overline{2t}\right)} 2^{\ell(\mathbf{p})}\zeta(\mathbf{p}). \tag{10.49}$$

(**2**-1-**2**-*c*-**2**-1-**2**): For $\mathbf{s} = \left(\{2\}^{a_0}, 1, \{\{2\}^{b_j}, c_j, \{2\}^{a_j}, 1\}_{j=1}^{\ell}, \{2\}^{t}\right)$, with $\ell \geq 0$ and $a_0, t \geq 1$, show that

$$\zeta^{\star}(\mathbf{s}) = -\sum_{\mathbf{p} \preceq \left(2a_0+1, \{\overline{2b_j+2}, \{1\}^{c_j-3}, \overline{2a_j+2}\}_{j=1}^{\ell}, \overline{2t}\right)} 2^{\ell(\mathbf{p})}\zeta(\mathbf{p}). \tag{10.50}$$

Then verify the following special cases:

$$\begin{aligned}\zeta^{\star}(\{2\}^{b}, 3, \{2\}^{a}, 1, \{2\}^{t}) = &-2\zeta(\overline{2b+2a+2t+4}) - 4\zeta(2b+2a+4, \overline{2t}) \\ &-4\zeta(\overline{2b+2}, 2a+2t+2) - 8\zeta(\overline{2b+2}, \overline{2a+2}, \overline{2t}),\end{aligned}$$

and

$$\begin{aligned}&\zeta^{\star}(\{2\}^{a_1}, 1, \{2\}^{b}, 3, \{2\}^{a_2}, 1, \{2\}^{t}) \\ &= -2\zeta(\overline{2a_1+2b+2a_2+2t+5}) - 4\zeta(2a_1+2b+2a_2+5, \overline{2t}) \\ &\quad -4\zeta(\overline{2a_1+2b+3}, 2a_2+2t+2) - 8\zeta(\overline{2a_1+2b+3}, \overline{2a_2+2}, \overline{2t}) \\ &\quad -4\zeta(2a_1+1, \overline{2b+2a_2+2t+4}) - 8\zeta(2a_1+1, 2b+2a_2+4, \overline{2t}) \\ &\quad -8\zeta(2a_1+1, \overline{2b+2}, 2a_2+2t+2) - 16\zeta(2a_1+1, \overline{2b+2}, \overline{2a_2+2}, \overline{2t}).\end{aligned}$$

10.7. For any positive integer n, show that

$$\zeta^\star(\{2\}^{n-1},3)=\sum_{l=1}^{n}\zeta^\star(\{2\}^{l},1,\{2\}^{n-l})-\sum_{l=0}^{n-1}\zeta^\star(\{2\}^{l},3,\{2\}^{n-1-l}).$$

Chapter 11

q-analogs of multiple zeta functions

In this chapter we define some q-analogs of multiple zeta functions and study their analytic properties. We will also consider some special values at non-positive integers. The study of special values at positive integers will be carried out in the next chapter. As before, throughout this chapter let q be a real number with $0 < q < 1$.

11.1 Definitions

First we define an auxiliary function of $2d$ complex variables $s_1, \dots, s_d, t_1, \dots, t_d \in \mathbb{C}$

$$\zeta_q^{(t_1,\dots,t_d)}[s_1, \dots, s_d] = \sum_{k_1 > \cdots > k_d > 0} \frac{q^{k_1 t_1 + \cdots + k_d t_d}}{[k_1]^{s_1} \cdots [k_d]^{s_d}}. \tag{11.1}$$

Proposition 11.1.1. *The function $\zeta_q^{(t_1,\dots,t_d)}[s_1, \dots, s_d]$ converges if $\mathrm{Re}(t_1 + \cdots + t_j) > 0$ for all $j = 1, \dots, d$. It can be analytically continued to a meromorphic function over $\mathbb{C}^{2d}$ via the series expansion*

$$\zeta_q^{(t_1,\dots,t_d)}[s_1, \dots, s_d] = (1-q)^{|\mathbf{s}|} \sum_{r_1,\dots,r_d=0}^{\infty} \prod_{j=1}^{d} \left[\binom{s_j + r_j - 1}{r_j} \frac{q^{(d+1-j)(r_j+t_j)}}{1 - q^{r_1+t_1+\cdots+r_j+t_j}} \right], \tag{11.2}$$

where $|\mathbf{s}| = s_1 + \cdots + s_d$. It has the following (simple) poles: $t_1 + \cdots + t_j \in \mathbb{Z}_{\le 0} + \frac{2\pi i}{\log q}\mathbb{Z}$ for $j = 1, \dots, d$.

Proof. Assume $|\mathrm{Re}(s_j)| < N_j$ and let $\tau_j = \mathrm{Re}(t_j)$ for all $j = 1, \dots, d$. Since

$$1 < [k] = 1 + q + \cdots + q^{k-1} < k \tag{11.3}$$

we have

$$\begin{aligned}\left|\zeta_q^{(t_1,\dots,t_d)}[s_1,\dots,s_d]\right| &< \sum_{k_1>\cdots>k_d>0}\prod_{j=1}^{d}(k_j)^{N_j}q^{k_j\tau_j}\\ &= \sum_{m_1,\dots,m_d=1}^{\infty}\prod_{j=1}^{d}(m_j+\cdots+m_d)^{N_j}q^{(m_j+\cdots+m_d)\tau_j}\\ &< \prod_{j=1}^{d}j^{N_j}\sum_{m_1,\dots,m_d=1}^{\infty}\prod_{j=1}^{d}m_j^{N_1+\cdots+N_j}q^{(\tau_1+\cdots+\tau_j)m_j}.\end{aligned}$$

By root test $\sum_{m>0}m^Nq^{m\tau}$ converges if $\tau>0$. This proves the first part of the proposition.

We now prove Eq. (11.2). By binomial expansion

$$(1-x)^{-s}=\sum_{r=0}^{\infty}\binom{s+r-1}{r}x^r$$

we get

$$\zeta_q^{(t_1,\dots,t_d)}[s_1,\dots,s_d]=(1-q)^{|\mathbf{s}|}\sum_{k_1>\cdots>k_d>0}\sum_{r_1,\dots,r_d=0}^{\infty}\left(\prod_{j=1}^{d}\binom{s_j+r_j-1}{r_j}q^{k_j(r_j+t_j)}\right).$$

As $0<q<1$ the series converges absolutely by Stirling's formula so we can exchange the summations. The proposition follows immediately from the next lemma by taking $x_j=q^{t_j+r_j}$ for $j=1,\dots,d$. □

Lemma 11.1.2. *Let $x_j\in\mathbb{C}$ such that $|x_j|<1$ for $j=1,\dots,d$. Then*

$$\sum_{k_1>\cdots>k_d>0}\prod_{j=1}^{d}x_j^{k_j}=\prod_{j=1}^{d}\frac{x_1\cdots x_j}{1-x_1\cdots x_j}=\prod_{j=1}^{d}\frac{x_j^{d+1-j}}{1-x_1\cdots x_j}. \tag{11.4}$$

Proof. By re-indexing $k_j=m_1+\cdots+m_j$ we have

$$\sum_{k_1>\cdots>k_d>0}\prod_{j=1}^{d}x_j^{k_j}=\sum_{m_1=1}^{\infty}\cdots\sum_{m_d=1}^{\infty}x_1^{m_1+\cdots+m_d}\cdots x_{d-1}^{m_{d-1}+m_d}x_d^{m_d}=\prod_{j=1}^{d}\frac{x_1\cdots x_j}{1-x_1\cdots x_j}.$$

The second equation in the lemma follows immediately. □

Now we define the first version of the q-analog of multiple zeta functions.

Definition 11.1.3. For all $\mathbf{s}=(s_1,\dots,s_d)\in\mathbb{C}^d$ we set

$$\zeta_q^{\mathrm{I}}[\mathbf{s}]:=\zeta_q^{\mathbf{s-1}}[\mathbf{s}]=\sum_{k_1>\cdots>k_d\geq 1}\prod_{j=1}^{d}\frac{q^{k_j(s_j-1)}}{[k_j]^{s_j}}. \tag{11.5}$$

The following corollary follows easily from Prop. 11.1.1.

Theorem 11.1.4. *Set*

$$\mathsf{S}'_d = \left\{ (s_1,\dots,s_d)\in\mathbb{C}^d \;\middle|\; \begin{array}{l} s_1 \in 1+\dfrac{2\pi i}{\log q}\mathbb{Z},\ \text{or } s_1\in\mathbb{Z}_{\le 0}+\dfrac{2\pi i}{\log q}\mathbb{Z}_{\neq 0},\\ \text{or } s_1+\cdots+s_j\in\mathbb{Z}_{\le j}+\dfrac{2\pi i}{\log q}\mathbb{Z},\ j>1 \end{array} \right\}.$$

Then $\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]$ *defined by Eq.* (11.5) *can be extended to a meromorphic function with simple poles lying along* S'_d *and it can be written as*

$$\zeta_q^{\mathrm{I}}[s_1,\dots,s_d] = (1-q)^{|\mathbf{s}|}\sum_{r_1,\dots,r_d=0}^{+\infty}\prod_{j=1}^{d}\left[\binom{s_j+r_j-1}{r_j}\frac{q^{(d+1-j)(r_j+s_j-1)}}{1-q^{r_1+s_1+\cdots+r_j+s_j-j}}\right]. \tag{11.6}$$

It is easy to see that as $q\to 1$ the set S'_d becomes the set defined by Eq. (1.15):

$$\mathsf{S}_d = \left\{ (s_1,\dots,s_d)\in\mathbb{C}^d \;\middle|\; \begin{array}{l} s_1 = 1, s_1+s_2 = 2,1,-2m\ \forall m\in\mathbb{Z}_{\ge 0},\\ \text{and } s_1+\cdots+s_j\in\mathbb{Z}_{\le j}\ \forall j\ge 3. \end{array} \right\} \tag{11.7}$$

which is exactly the singular set of the multiple zeta function $\zeta(s_1,\dots,s_d)$.

To see the effect of taking different specializations of t_j in $\zeta_q^{\mathbf{t}}[\mathbf{s}]$ we define the shifting operators $\mathcal{S}_j$ $(1\le j\le d)$ on any function $f_q[s_1,\dots,s_d]$ as follows:

$$\mathcal{S}_j f_q[s_1,\dots,s_d] = f_q[s_1,\dots,s_d] + (1-q)f_q[s_1,\dots,s_j-1,\dots,s_d]. \tag{11.8}$$

It is obvious that these operators commute.

Proposition 11.1.5. *Let* $n_1,\dots,n_d$ *be non-negative integers. Then we have*

$$\begin{aligned}\zeta_q^{(s_1-1-n_1,\dots,s_d-1-n_d)}[s_1,\dots,s_d] &= \mathcal{S}_1^{n_1}\circ\cdots\circ\mathcal{S}_d^{n_d}\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]\\ &= \sum_{r_1=0}^{n_1}\cdots\sum_{r_d=0}^{n_d}\left(\prod_{j=1}^{d}\binom{n_j}{r_j}(1-q)^{r_j}\right)\zeta_q^{\mathrm{I}}[s_1-r_1,\dots,s_d-r_d].\end{aligned}$$

Proof. We only sketch the proof in the case $n_2=\cdots=n_d=0$. The general

case is completely similar. Suppose $n_1 = n = 1$. Then

$$\begin{aligned}
&\zeta_q^{(s_1-2,s_2-1,\ldots,s_d-1)}[s_1,\ldots,s_d] \\
=&\sum_{k_1>\cdots>k_d>0} \frac{q^{k_1(s_1-2)+k_2(s_2-1)+\cdots+k_d(s_d-1)}}{[k_1]^{s_1}\cdots[k_d]^{s_d}} \\
=&\sum_{k_1>\cdots>k_d>0} \frac{q^{k_1(s_1-2)}(1-q^{k_1})+q^{k_1(s_1-1)}}{[k_1]^{s_1}} \cdot \frac{q^{k_2(s_2-1)+\cdots+k_d(s_d-1)}}{[k_2]^{s_2}\cdots[k_d]^{s_d}} \\
=&\mathcal{S}\zeta(s_1,\ldots,s_d).
\end{aligned}$$

The rest follows easily by induction on n and some binomial identities. See Exercise 11.1. □

Corollary 11.1.6. *Let n be a positive integer. Then*

$$\zeta_q^{(s-1-n)}[s] = \mathcal{S}^n\zeta_q^{\mathrm{I}}[s] = \sum_{r=0}^{n}\binom{n}{r}(1-q)^r\zeta_q^{\mathrm{I}}[s-r].$$

Observe that one effect of the shifting operator is to bring in more poles. Essentially, $\mathcal{S}^n$ shifts all the poles of $\zeta_q^{\mathrm{I}}[s]$ by n to the right on the complex plane.

11.2 Analytic continuation

As a golden rule, the correct q-analog of any classical mathematical object should have the property that when $q \to 1$ the q-analog becomes its classical counterpart. Although it is obvious that when $\mathrm{Re}(s_1+\cdots+s_j) > j$ we have

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[s_1,\ldots,s_d] = \zeta(s_1,\ldots,s_d),$$

it is highly nontrivial to verify this still holds for all $(s_1,\ldots,s_d)$ outside the singular set Eq. (1.15) of the analytic continuation of the multiple zeta function $\zeta(s_1,\ldots,s_d)$. We will prove this fact in this section.

Similar to the classical case (see Sec. 1.3) the basic idea is to use the classical Euler-Maclaurin summation formula Eq. (1.12). Let $f(x)$ be any (complex-valued) C^∞ function on $[1,\infty)$ and let m and M be two positive integers. Recall that

$$\begin{aligned}
\sum_{n=1}^{m} f(n) = &\int_1^m f(x)\,dx + \frac{1}{2}(f(1)+f(m)) + \sum_{r=1}^{M}\frac{B_{r+1}}{(r+1)!}\left(f^{(r)}(m)-f^{(r)}(1)\right) \\
&-\frac{(-1)^{M+1}}{(M+1)!}\int_1^m \widetilde{B}_{M+1}(x)f^{(M+1)}(x)\,dx. \qquad (11.9)
\end{aligned}$$

where the periodic Bernoulli polynomial satisfies

$$\widetilde{B}_k(x) = -k! \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{e^{2\pi inx}}{(2\pi in)^k}. \tag{11.10}$$

We first consider the depth 1 case.

Theorem 11.2.1. *The q-analog of Riemann zeta function $\zeta_q^{\mathrm{I}}[s]$ can be extended to a meromorphic function on $\mathbb{C}\setminus \mathsf{S}_1'$ with simple poles along S_1'. Further, for all $s\in\mathbb{C}\setminus\mathsf{S}_1$*

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[s] = \zeta(s).$$

Proof. We set

$$\begin{aligned} F(x) &= \frac{q^{x(s-1)}}{(1-q^x)^s}. \quad F'(x) = (\log q)q^{x(s-1)}\frac{s-1+q^x}{(1-q^x)^{s+1}},\\ F''(x) &= (\log q)^2 q^{x(s-1)}\frac{s(s+1)-3s(1-q^x)+(1-q^x)^2}{(1-q^x)^{s+2}}. \end{aligned}$$

Taking $f(x)=F(x)$ in Eq. (11.9) and letting $m\to\infty$ we get

$$\begin{aligned} \sum_{n=1}^{\infty}\frac{q^{n(s-1)}}{(1-q^n)^s} &= \int_1^{\infty}\frac{q^{x(s-1)}}{(1-q^x)^s}\,dx + \frac{1}{2}\cdot\frac{q^{s-1}}{(1-q)^s} - \frac{1}{12}(\log q)q^{s-1}\frac{s-1+q}{(1-q)^{s+1}}\\ &\quad - \frac{(\log q)^2}{2}\int_1^{\infty}\widetilde{B}_2(x)q^{x(s-1)}\frac{s(s+1)-3s(1-q^x)+(1-q^x)^2}{(1-q^x)^{s+2}}\,dx \end{aligned}$$

for $\mathrm{Re}(s)>1$. Hence (see Exercise 11.2)

$$\begin{aligned} \zeta_q^{\mathrm{I}}[s] &= \frac{q^{s-1}}{s-1}\cdot\frac{q-1}{\log q} + \frac{q^{s-1}}{2} + \frac{q^{s-1}}{12}\frac{\log q}{q-1}(s-1+q)\\ &\quad - (1-q)^s\frac{(\log q)^2}{2}\int_1^{\infty}\widetilde{B}_2(x)q^{x(s-1)}\frac{s(s+1)-3s(1-q^x)+(1-q^x)^2}{(1-q^x)^{s+2}}\,dx. \end{aligned} \tag{11.11}$$

By Eq. (11.10)

$$\begin{aligned} \zeta_q^{\mathrm{I}}[s] &= \frac{q^{s-1}}{s-1}\cdot\frac{q-1}{\log q} + \frac{q^{s-1}}{2} + \frac{q^{s-1}}{12}\frac{\log q}{q-1}(s-1+q)\\ &\quad + \sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{(1-q)^s(\log q)^2}{(2\pi in)^2}\int_1^{\infty}e^{2\pi inx}q^{x(s-1)}\frac{s(s+1)-3s(1-q^x)+(1-q^x)^2}{(1-q^x)^{s+2}}\,dx. \end{aligned}$$

Setting $q^x = u$ we find

$$\zeta_q^{\mathrm{I}}[s] = \frac{q^{s-1}}{s-1}\cdot\frac{q-1}{\log q} + \frac{q^{s-1}}{2} + \frac{q^{s-1}}{12}\frac{\log q}{q-1}(s-1+q) - \sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{(1-q)^s\log q}{(2\pi i n)^2}\sum_{\nu=\pm1,0}a_\nu(s)b_q(s-1+\delta n,-s+\nu), \quad (11.12)$$

where $\delta = 2\pi i/\log q$, $a_{-1}(s) = s(s+1)$, $a_0(s) = -3s$, $a_1(s) = 1$, and

$$b_t(\alpha,\beta) = \int_0^t u^{\alpha-1}(1-u)^{\beta-1}\,du. \quad (11.13)$$

Observe that $\sigma = \mathrm{Re}(s) > 1$ and for all $\nu = -1, 0, 1$

$$\left|b_q(s-1+\delta n,-s+\nu)\right| \le \int_0^q u^{\sigma-2}(1-u)^{-\sigma+\nu-1}du.$$

Hence the sum in Eq. (11.12) converges absolutely. For Eq. (11.13), we have (see Exercise 11.3)

$$b_q(\alpha,\beta) = \sum_{r=1}^{M-1}(-1)^{r-1}\frac{(1-\beta)_{r-1}}{(\alpha)_r}q^{\alpha+r-1}(1-q)^{\beta-r} + (-1)^{M-1}\frac{(1-\beta)_{M-1}}{(\alpha)_{M-1}}b_q(\alpha+M-1,\beta-M+1) \quad (11.14)$$

for any $M \ge 2$. Applying this to $b_q(s-1+\delta n,-s-1)$ and using $q^{\delta n} = 1$ we get

$$b_q(s-1+\delta n,-s+\nu) = \sum_{r=1}^{M-1}(-1)^{r-1}\frac{(s+1-\nu)_{r-1}}{(s-1+\delta n)_r}q^{s+r-2}(1-q)^{-s+\nu-r} + (-1)^{M-1}\frac{(s+1-\nu)_{M-1}}{(s-1+\delta n)_{M-1}}b_q(s-2+M+\delta n,-s+\nu+1-M). \quad (11.15)$$

For all $\nu = -1, 0, 1$ this provides the analytic continuation of $b_q(s-1+\delta n,-s+\nu)$ as functions of s into the region $\mathrm{Re}(s) > 2-M$. From this we have

$$\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{s(s+1)}{(2\pi i n)^2}b_q(s-1+\delta n,-s-1) = \sum_{r=1}^{M-1}(-1)^{r-1}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{(s)_{r+1}\cdot q^{s+r-2}(1-q)^{-s-1-r}}{(2\pi i n)^2(s-1+\delta n)_r} - (-1)^{M-1}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{(s)_{M+1}\cdot\log q}{(2\pi i n)^2(s-1+\delta n)_{M-1}}\int_1^\infty e^{2\pi i n x}q^{x(s-2+M)}(1-

Observe that

$$\lim_{q\to 1}\frac{\log q}{1-q}=-1,\quad \lim_{q\to 1}(1-q)^r(s-1+\delta n)_r=(-2\pi i)^r,\quad \lim_{q\to 1}\frac{1-q^x}{1-q}=x.$$

Thus for $\mathrm{Re}(s)>2-M$, by Eq. (11.10) (and its special case when $x=1$) we get

$$\begin{aligned}
&\lim_{q\to 1}(1-q)^s\log q\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{s(s+1)}{(2\pi in)^2}b_q(s-1+\delta n,-s-1)\\
&=\sum_{r=1}^{M-1}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{(s)_{r+1}}{(2\pi in)^{r+2}}-\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{(s)_{M+1}}{(2\pi in)^{M+1}}\int_1^\infty e^{2\pi inx}x^{-s-M-1}\,dx\\
&=-\sum_{r=1}^{M-1}\frac{B_{r+2}}{(r+2)!}(s)_{r+1}+\frac{(s)_{M+1}}{(M+1)!}\int_1^\infty \widetilde{B}_{M+1}(x)x^{-s-M-1}\,dx.
\end{aligned}$$

The same procedure applies to the sums involving $b_q(s-1+\delta n,-s)$ and $b_q(s-1+\delta n,-s+1)$. See Exercise 11.4. Consequently, for $\mathrm{Re}(s)>2-M$, we have

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[s]=\frac{1}{s-1}+\frac{1}{2}+\sum_{r=1}^{M}\frac{B_{r+1}}{(r+1)!}(s)_r-\frac{(s)_{M+1}}{(M+1)!}\int_1^\infty \widetilde{B}_{M+1}(x)x^{-s-M-1}\,dx=\zeta(s)$$

by Eq. (1.16). Since M is arbitrary this finishes the proof of the theorem. □

Theorem 11.2.2. *The multiple q-zeta function $\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]$ can be extended to a meromorphic function on $\mathbb{C}^d\setminus \mathsf{S}'_d$ with simple poles along S'_d. Further, for all $(s_1,\dots,s_d)\in\mathbb{C}^d\setminus\mathsf{S}_d$*

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]=\zeta(s_1,\dots,s_d).$$

Proof. For simplicity, only in this proof we write $\zeta[s_1,\dots,s_d]=\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]$.

We proceed by induction on d. If $d=1$ the theorem is reduced to Theorem 11.2.1. Now we assume $d\geq 2$. In definition Eq. (11.5) we replace s_1, k_1, and k_2 by s, n and k, respectively. Taking $f(x)=F(x+k-1)$ and

letting $m \to \infty$ in Eq. (1.12) we get

$$\begin{aligned}
&\sum_{n>k} \frac{q^{n(s-1)}}{[n]^s} = (1-q)^s \left(-F(k) + \sum_{n=1}^{\infty} f(n) \right) \\
&= (1-q)^s \left(\int_k^{\infty} F(x)\,dx - \frac{1}{2}F(k) - \frac{1}{12}F'(k) - \frac{1}{2}\int_1^{\infty} \widetilde{B}_2(x) f''(x)\,dx \right) \\
&= \frac{(q-1)}{(s-1)\log q} \frac{q^{k(s-1)}}{[k]^{s-1}} - \frac{1}{2}\frac{q^{k(s-1)}}{[k]^s} + \frac{1}{12}\frac{\log q}{q-1}\frac{q^{k(s-1)}(s+q^k-1)}{[k]^{s+1}} \\
&- \frac{(1-q)^s(\log q)^2}{2} \int_k^{\infty} \widetilde{B}_2(x) q^{x(s-1)} \frac{s(s+1) - 3s(1-q^x) + (1-q^x)^2}{(1-q^x)^{s+2}}\,dx
\end{aligned} \tag{11.16}$$

because $\widetilde{B}_2(x+k-1) = \widetilde{B}_2(x)$ by periodicity. By exactly the same argument as in the proof of depth 1 case, we can obtain from Eq. (11.10) the following expression for the last term involving the integral in Eq. (11.16):

$$-\sum_{n \in \mathbb{Z}\setminus\{0\}} \frac{(1-q)^s \log q}{(2\pi i n)^2} \sum_{\nu=\pm 1,0} a_\nu(s) b_{q^k}(s-1+\delta n, -s+\nu). \tag{11.17}$$

Set $\mathbf{s}' = (s_3, \ldots, s_d)$ if $d \geq 3$ and $\mathbf{s}' = \emptyset$ if $d = 2$. Replacing q by q^k in Eq. (11.15), plugging it into Eq. (11.17) which is used to replace the last line of Eq. (11.16), and then applying Prop. 11.1.5, we finally get

$$\begin{aligned}
&\zeta[s_1, s_2, \mathbf{s}'] = \frac{(q-1)}{(s_1-1)\log q} \zeta[s_1+s_2-1, \mathbf{s}'] - \frac{1}{2}\mathcal{S}\zeta[s_1+s_2, \mathbf{s}'] \\
&+ \frac{s_1}{12}\frac{\log q}{q-1}\mathcal{S}^2\zeta[s_1+s_2+1, \mathbf{s}'] + \frac{\log q}{12}\mathcal{S}\zeta[s_1+s_2, \mathbf{s}'] - \sum_{\nu=\pm 1,0} (C_\nu + D_\nu),
\end{aligned} \tag{11.18}$$

where C_ν and D_ν are the contributions from the sum involving $b_{q^k}(\ldots, -s+\nu)$ (note that $s = s_1$). Explicitly they are computed as follows. Write

$$T(q, s, n, r) = \prod_{j=0}^{r-1} \big(2\pi i n + (s-1+j)\log q\big)^{-1}. \tag{11.19}$$

Then

$$C_{-1} = \sum_{r=1}^{M-1} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{T(q,s,n,r)}{(2\pi i n)^2} \left(\frac{\log q}{q-1}\right)^{r+1} (s)_{r+1} \cdot \mathcal{S}^3\zeta[s_1+s_2+r+1,\mathbf{s}'],$$

$$D_{-1} = - \sum_{k_2>\cdots>k_d>0} \frac{q^{k_2(s_2-1)+\cdots+k_d(s_d-1)}}{[k_2]^{s_2}\cdots[k_d]^{s_d}} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{T(q,s,n,M-1)}{(2\pi i n)^2} \cdot$$
$$\cdot \left(\frac{\log q}{q-1}\right)^{M+1} (s)_{M+1} \cdot \int_{k_2}^{\infty} e^{2\pi i n x} q^{x(s-2+M)} \left(\frac{1-q^x}{1-q}\right)^{-s-M-1} dx$$
$$=: \sum_{k_3>\cdots>k_d>0} \frac{q^{k_3(s_3-1)+\cdots+k_d(s_d-1)}}{[k_3]^{s_3}\cdots[k_d]^{s_d}} \sum_{k=k_3+1}^{\infty} R(M,q,k,s_2,s),$$

where we replace the index k_2 by k. Similarly,

$$C_0 = 3\log q \sum_{r=1}^{M-1} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{T(q,s,n,r)}{(2\pi i n)^2} \left(\frac{\log q}{q-1}\right)^{r} (s)_r \cdot \mathcal{S}^2\zeta[s_1+s_2+r,\mathbf{s}'],$$

$$D_0 = - 3\log q \sum_{k_2>\cdots>k_d>0} \frac{q^{k_2(s_2-1)+\cdots+k_d(s_d-1)}}{[k_2]^{s_2}\cdots[k_d]^{s_d}} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{T(q,s,n,M-1)}{(2\pi i n)^2} \cdot$$
$$\cdot \left(\frac{\log q}{q-1}\right)^{M} (s)_M \cdot \int_{k_2}^{\infty} e^{2\pi i n x} q^{x(s-2+M)} \left(\frac{1-q^x}{1-q}\right)^{-s-M} dx,$$

and

$$C_1 = (\log q)^2 \sum_{r=1}^{M-1} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{T(q,s,n,r)}{(2\pi i n)^2} \left(\frac{\log q}{q-1}\right)^{r-1} (s)_{r-1} \cdot \mathcal{S}\zeta[s_1+s_2+r-1,\mathbf{s}'],$$

$$D_1 = - (\log q)^2 \sum_{k_2>\cdots>k_d>0} \frac{q^{k_2(s_2-1)+\cdots+k_d(s_d-1)}}{[k_2]^{s_2}\cdots[k_d]^{s_d}} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{T(q,s,n,M-1)}{(2\pi i n)^2} \cdot$$
$$\cdot \left(\frac{\log q}{q-1}\right)^{M-1} (s)_{M-1} \cdot \int_{k_2}^{\infty} e^{2\pi i n x} q^{x(s-2+M)} \left(\frac{1-q^x}{1-q}\right)^{-s-M+1} dx.$$

The crucial step next is to control the summations $\sum_{k=k_3+1}^{\infty}$ and show that they converge uniformly with respect to q. When $0 < q \le 1/2$ this is clear. The only non-trivial part is when $q \to 1$. So we assume $1/2 < q < 1$. Note that

$$\lim_{q\to 1} T(q,s,n,r) = \frac{1}{(2\pi i n)^r}, \quad \lim_{q\to 1} \frac{\log q}{q-1} = 1.$$

Lemma 11.2.3. *Let $s_1 = s = \sigma + i\tau$. Let $q_0 = \max\{1/2, e^{(6-2\pi)/\tau}\}$ if $\tau > 0$ and let $q_0 = 1/2$ if $\tau \le 0$. Then for all $1 > q > q_0$ and positive integer k we have*

$$\left|\frac{\log q}{q-1}\right| < 2, \quad \text{and} \quad |T(q,s,n,r)| < \frac{1}{(6n)^r}.$$

Proof. Let $f(q) = 2(1-q) + \log q$. Then $f'(q) = -2 + 1/q < 0$ whenever $q > q_0$. So $f(q) > f(1) = 0$ whenever $1 > q > q_0$. This implies that $2(1-q) > -\log q$ whence $\log q/(q-1) < 2$.

To bound $T(q, s, n, r)$ we consider each of its factors in definition Eq. (11.19). For each $0 \le j < r$ we have $|2\pi in + (s-1+j)\log q|^2 = \big((\sigma - 1 + j)\log q\big)^2 + (2\pi n + \tau \log q)^2 \ge (2\pi n + \tau \log q)^2$ which is independent j. If $\tau \le 0$ then clearly $|2\pi in + (s-1+j)\log q| > 6n$. If $\tau > 0$ then it follows from $q > e^{(6-2\pi)\tau}$ that

$$2\pi n + \tau \log q > 2\pi n + 6 - 2\pi \ge 6n,$$

as desired. □

Next we want to bound the integral terms in D_{-1}. Let $|\sigma_1| < N$ and $|\sigma_2| < N'$ for some positive integers N and N'. Fix an arbitrary $x > k$ and a positive integer

$$M > 16 + 2N + 6\sum_{j=2}^{d}(N_j + 1). \tag{11.20}$$

Then

$$q^{-k(M/6-N'-1)}\left|q^{k(s'-1)}q^{x(s-2+M)}\left(\frac{1-q^x}{1-q}\right)^{-s-M-1}\right| < q^{x(M-N-2)-kM/6}\left(\frac{1-q}{1-q^x}\right)^{M-N+1}.$$

Denote by $g(q)$ the right-hand side of the above inequality. It can be proved that for $1/2 < q < 1$, $g(q)$ is increasing as a function of q so that (see Exercise 11.5)

$$g(q) \le \lim_{q\to 1} g(q) = \frac{1}{x^{M-N+1}}. \tag{11.21}$$

We now can bound the innermost sum of D_{-1}. From Lemma 11.2.3 and Eq. (11.21) we have (if $d = 2$ then take $k_3 = 0$)

$$\begin{aligned}
&\sum_{k=1+k_3}^{\infty} |R(M, q, k, s', s)| \\
<&\sum_{k=1+k_3}^{\infty} (2k)^{N'}\frac{2\zeta(M+1)}{4\pi^2 6^{M-1}}2^{M+1}(N)_{M+1}q^{k(M/6-N'-1)}\int_k^{\infty}\frac{\mathrm{d}x}{x^{M-N+1}} \\
<&\frac{(M+1)!}{M-N}\binom{M+N}{M+1}\sum_{k=1+k_3}^{\infty}\frac{q^{k(M/6-N'-1)}}{k^{M-N-N'}} \\
<&(M+1)!(2M)^{M+1}\sum_{k=1+k_3}^{\infty} q^{k(M/6-N'-1)}
\end{aligned}$$

since $2^{N'+M+2}\zeta(M+1) < 4\pi^2 6^{M-1}$ and $M-N>2$. Therefore by Eq. (11.3)

$$\left|D_{-1}\right| < (2M)^{2M} \sum_{k_3>\cdots>k_d} \left(\prod_{j=3}^{d} k_j^{N_j} q^{k_j(-N_j-1)} \right) q^{k_3(M/6-N_2-1)} \quad (11.22)$$

which converges by the condition Eq. (11.20) on M (see the proof of Prop. 11.1.1).

Exactly the same argument applies to the integral terms in D_0 and D_1 (see Exercise 11.7). These convergence results imply two things. First we can show by induction on d that Eq. (11.18) gives rise to an analytic continuation of $\zeta[s_1,\dots,s_d]$ as a meromorphic function on $\mathbb{C}^d \setminus \mathsf{S}'_d$. Second, also by induction on d, we now can conclude that it's legitimate to take the limit $q\to 1$ inside the sums of C_ν and D_ν to get (note that $\lim_{q\to 1} \mathcal{S}^n\zeta[\mathbf{s}] = \zeta(\mathbf{s})$ for any $\mathbf{s}\in\mathbb{C}^{d-1}\setminus \mathsf{S}'_{d-1}$ and any positive integer n)

$$\begin{aligned}
\lim_{q\to 1}\zeta[s_1,s_2,\mathbf{s}'] =& \frac{1}{s-1}\zeta(s_1+s_2-1,\mathbf{s}') - \frac{1}{2}\zeta(s_1+s_2,\mathbf{s}'] + \frac{s}{12}\zeta(s_1+s_2+1,\mathbf{s}') \\
&- \sum_{r=1}^{M-1} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{(2\pi i n)^{r+2}} (s_1)_{r+1}\cdot \zeta(s_1+s_2+r+1,\mathbf{s}') \\
&+ \sum_{k_2>\cdots>k_d} \frac{1}{k_2^{s_2}\cdots k_d^{s_d}} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{(s)_{M+1}}{(2\pi i n)^{M+1}} \int_{k_2}^{\infty} e^{2\pi i n x} x^{-s_1-M-1}\,dx \\
=& \sum_{r=0}^{M+1} \frac{B_r}{r!}(s_1)_{r-1}\cdot\zeta(s_1+s_2+r-1,s_3,\dots,s_d) \\
&- \frac{1}{(M+1)!} \sum_{k_2>\cdots>k_d>0} \frac{1}{k_2^{s_2}\cdots k_d^{s_d}} \int_{k_2}^{\infty} \widetilde{B}_{M+1}(x) \frac{(s)_{M+1}}{x^{s_1+M+1}}\,dx
\end{aligned}$$

by Eq. (11.10) and its specialization with $x=1$

$$\sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{(2\pi i n)^{r+2}} = -\frac{\widetilde{B}_{r+2}(1)}{(r+2)!} = -\frac{B_{r+2}}{(r+2)!}.$$

The theorem now mostly follows from Theorem 1.3.2 (by replacing M by $M+1$ there). The poles at $s_1 = m - \frac{2\pi i}{\log q} n$ are given by the first term in formula Eq. (11.18) when $m=1$ and $n=0$ and by the terms $T(q,s_d,n,r)$ as defined in Eq. (11.19) if $m\le 1$ and $n\neq 0$. The location of the other poles are obtained by induction using those poles of the q-Riemann zeta function presented in Theorem 11.2.1 for the initial step. This completes the proof of the theorem. □

11.3 Special values at non-positive integers

In this section, we will consider some special values of $\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]$ at non-positive integers.

Theorem 11.3.1. *For any nonnegative integer m, the limiting value $\zeta_q^{\mathrm{I}}[-m]=\lim_{s\to -m}\zeta_q^{\mathrm{I}}[s]$ exists and is given explicitly by*

$$\zeta_q^{\mathrm{I}}[-m]=(1-q)^{-m}\left\{\sum_{r=0}^{m}(-1)^r\binom{m}{r}\frac{1}{q^{m+1-r}-1}+\frac{(-1)^{m+1}}{(m+1)\log q}\right\}.$$

Moreover,

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[-m]=-\frac{B_{m+1}}{m+1}. \tag{11.23}$$

Proof. Note the terms with $r\geq m+2$ in the sum Eq. (11.6) vanishes when $s\to -m$ since $\binom{-m+r-1}{r}=0$ and $1-q^{-m+r-1}\neq 0$. On the other hand, for $r=m+1$ we have $\lim_{s\to -m}(s+m)/(1-q^{s+m})=-1/\log q$ and hence

$$\lim_{s\to -m}\binom{s+m}{m+1}\frac{q^{s+m}}{1-q^{s+m}}=\frac{(-1)^m m!}{(m+1)!}\left(-\frac{1}{\log q}\right)=\frac{(-1)^{m+1}}{(m+1)\log q}.$$

To prove Eq. (11.23) we apply Theorem 11.2.1 to get

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[-m]=\zeta(-m)=-\frac{B_{m+1}}{m+1}.$$

This completes the proof of the theorem. □

Remark 11.3.2. Thanks to the theorem, we can regard $m\zeta_q^{\mathrm{I}}[1-m]$ as the q-analog of the Bernoulli number B_m.

In general, for integers $n_1,\dots,n_d$ we set

$$\zeta_q^{\mathrm{I}}[n_1,\dots,n_d]=\lim_{s_d\to n_d}\cdots\lim_{s_1\to n_1}\zeta_q^{\mathrm{I}}[s_1,\dots,s_d],$$
$$\zeta_q^{R}[n_1,\dots,n_d]=\lim_{s_1\to n_1}\cdots\lim_{s_d\to n_d}\zeta_q^{\mathrm{I}}[s_1,\dots,s_d]$$

if the limits exist. Similarly we can define

$$\zeta(n_1,\dots,n_d)=\lim_{s_d\to n_d}\cdots\lim_{s_1\to n_1}\zeta(s_1,\dots,s_d),$$
$$\zeta^{R}(n_1,\dots,n_d)=\lim_{s_1\to n_1}\cdots\lim_{s_d\to n_d}\zeta(s_1,\dots,s_d)$$

if the limits exist. Interesting phenomena occur already in the case $d = 2$ and these can be generalized to arbitrary depth. By Theorem 11.1.4 we get

$$\begin{aligned}\zeta_q^{\mathrm{I}}[s_1, s_2] = &\sum_{r_1,r_2=0}^{+\infty} \binom{s_1+r_1-1}{r_1}\binom{s_2+r_2-1}{r_2} \frac{(1-q)^{s_1+s_2} q^{2s_1+2r_1+s_2+r_2-3}}{(1-q^{s_1+r_1-1})(1-q^{s_1+r_1+s_2+r_2-2})} \\ =&(1-q)^{s_1+s_2}\left\{\frac{q^{2s_1+s_2-3}}{(1-q^{s_1-1})(1-q^{s_1+s_2-2})} + \frac{s_2 q^{2s_1+s_2-2}}{(1-q^{s_1-1})(1-q^{s_1+s_2-1})}.\right. \\ &+\frac{s_1 q^{2s_1+s_2-1}}{(1-q^{s_1})(1-q^{s_1+s_2-1})} + \frac{s_1 s_2 q^{2s_1+s_2}}{(1-q^{s_1})(1-q^{s_1+s_2})} \\ &\left.+ \frac{s_2(s_2+1)q^{2s_1+s_2-1}}{2(1-q^{s_1-1})(1-q^{s_1+s_2})} + \frac{s_1(s_1+1)q^{2s_1+s_2+1}}{2(1-q^{s_1+1})(1-q^{s_1+s_2})} + \cdots\right\}\end{aligned} \tag{11.24}$$

Clearly we have

$$\begin{aligned}\zeta_q^{\mathrm{I}}[0,0] &= \lim_{s_2\to 0}\lim_{s_1\to 0}\zeta_q^{\mathrm{I}}[s_1,s_2] = \frac{1}{(q^2-1)(q-1)} - \frac{3}{2(q-1)\log q} + \frac{1}{\log^2 q}, \\ \zeta_q^{R}(0,0) &= \lim_{s_1\to 0}\lim_{s_2\to 0}\zeta_q^{\mathrm{I}}[s_1,s_2] = \frac{1}{(q^2-1)(q-1)} - \frac{1}{(q-1)\log q} + \frac{q}{2(q-1)\log q}.\end{aligned}$$

It is not too hard to find that

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[0,0] = \frac{1}{3}, \quad \lim_{q\to 1}\zeta_q^{R}(0,0) = \frac{5}{12}.$$

This is consistent with the behavior of double zeta functions near $(0,0)$ (see Exercise 1.6). It has the following asymptotic expansion:

$$\zeta(s_1,s_2) = \frac{5s_1+4s_2}{12(s_1+s_2)} + R(s_1,s_2)$$

where $R(s_1,s_2)$ is analytic at $(0,0)$ and $\lim_{(s_1,s_2)\to(0,0)} R(s_1,s_2) = 0$.

Let n, k be two non-negative integers, and $m = k-n-2$. We now consider the double zeta function around $(s_1,s_2) = (-n,-m)$ which has the following expression by Theorem 1.3.2(ii):

$$\begin{aligned}&\frac{-1}{n+1}\zeta(s_1+s_2-1) - \frac{1}{2}\zeta(s_1+s_2) - \sum_{r=2}^{n+1}\frac{B_r}{r}\binom{n}{r-1}\zeta(s_1+s_2+r-1) \\ &\qquad + \alpha(m)(-1)^n\frac{B_k}{k!}n!(k-n-2)!(s_1+n)\zeta(s_1+s_2+k-1)\end{aligned} \tag{11.25}$$

where $\alpha(m) = 0$ if $m \le -1$ and $\alpha(m) = 1$ if $m \ge 0$. Note that the last term is zero when computing

$$\zeta(-n,-m) = \lim_{s_2\to -m}\lim_{s_1\to -n}\zeta(s_1,s_2)$$

$k,\ n$	pole, residue	$\overset{\text{indeterminacy}}{\zeta=\zeta^R(-n,n+2-k)}$
$k=0$	$-1/(n+1)$	NO
$k=1$	$-1/2$	NO
$2\nmid k, 3\le k\le n+1$	NO	$a_k+\displaystyle\sum_{r=k-1}^{n+1}\frac{B_r}{r}\binom{n}{r-1}\zeta(r+1-k)$
$2\nmid k, k=n+2$	NO	$B_{k-1}/(k-1)$
$2\nmid k, k>n+2$	NO	$B_{k-1}/2(k-1)$
$2\mid k, 2\le k\le n+1$	$(-1)^{k+1}\dfrac{B_k}{k}\binom{n}{k-1}$	NO

Table 11.1: Poles and indeterminacy of double zeta function, $a_k = B_{k-1}/(2k-2)$.

while it has possibly nontrivial contribution for $\zeta^R(-n,-m)$ since

$$\lim_{s_1\to -n}\lim_{s_2\to -m}(s_1+n)\zeta(s_1+s_2+k-1)=1.$$

When $m\ge 0$ and $2|k$ the values of $\zeta(-n,-m)$ and $\zeta^R(-n,-m)$ are different in general:

$$\zeta(-n,-m)=\frac{B_k}{k(n+1)}+\sum_{r=1}^{n+1}\frac{B_r}{r}\binom{n}{r-1}\frac{B_{k-r}}{k-r} \tag{11.26}$$

$$\zeta^R(-n,-m)=\zeta(-n,-m)+(-1)^n\frac{B_k}{k!}n!(k-n-2)!. \tag{11.27}$$

Note that the term corresponding to $r=1$ is non-zero if and only if $k=2$ (and $n=0$). From this observation we again recover that

$$\zeta(0,0)=\frac{1}{3},\quad \zeta^R(0,0)=\frac{5}{12}. \tag{11.28}$$

We now consider the double q-zeta function.

Theorem 11.3.3. *Let k, n be two non-negative integers, and $m=k-n-2$. If $m\le -1$ then the double q-zeta function $\zeta_q^{\mathrm{I}}[s_1,s_2]$ has a pole at $(-n,-m)$ with residue given by:*

$$\frac{\operatorname*{Res}_{(s_1,s_2)=(-n,-m)}\zeta_q^{\mathrm{I}}[s_1,s_2]}{-(1-q)^{2-k}(\log q)^{-1}}$$
$$=\begin{cases}\displaystyle\sum_{r=0}^{k}(-1)^r\binom{n+1-r}{k-r}\binom{n}{r}/(q^{n+1-r}-1) & \text{if } k\le n,\\ \displaystyle\sum_{r=0}^{n}(-1)^r\binom{n}{r}/(q^{n+1-r}-1)-\frac{(-1)^n}{(n+1)\log q} & \text{if } k=n+1.\end{cases}$$

Proof. This follows directly from Eq. (11.24). □

Corollary 11.3.4. *Let n be a non-negative integer. Then*

$$\operatorname*{Res}_{(s_1,s_2)=(-n,1)} \zeta_q^{\mathrm{I}}[s_1,s_2] = \frac{q-1}{\log q}\zeta_q^{\mathrm{I}}[-n]. \tag{11.29}$$

Setting $B_1^* = 1/2$ then we have for all $n \geq 0$

$$\lim_{q\to 1} \zeta_q^{\mathrm{I}}[-n] = -\frac{B_{n+1}^*}{n+1} = \operatorname*{Res}_{(s_1,s_2)=(-n,1)} \zeta(s_1,s_2). \tag{11.30}$$

Proof. Equation Eq. (11.29) follows from the case $m=-1$ in the above theorem and Theorem 11.3.1. The first equality in Eq. (11.30) is Eq. (11.23) and the second equality follows from Table 11.1. □

Corollary 11.3.5. *Let k, n be two non-negative integers, $m = k-n-2 \leq -2$. Then*

$$\lim_{q\to 1} \operatorname*{Res}_{(s_1,s_2)=(-n,-m)} \zeta_q^{\mathrm{I}}[s_1,s_2] = \operatorname*{Res}_{(s_1,s_2)=(-n,-m)} \zeta(s_1,s_2). \tag{11.31}$$

Proof. By Theorem 11.3.3 and Table 11.1 we only need to prove

$$\lim_{q\to 1} \sum_{r=0}^{k} \frac{(q-1)^{1-k}(-1)^r}{q^{n+1-r}-1} \binom{n+1-r}{k-r}\binom{n}{r} = \begin{cases} \dfrac{(-1)^k}{n+1} & \text{if } k=0, \\ \dfrac{B_k}{k}\binom{n}{k-1} & \text{if } 1\leq k\leq n. \end{cases} \tag{11.32}$$

First by generating function of the Bernoulli numbers

$$\frac{1}{q^{n+1-r}-1} = \frac{1}{e^{(n+1-r)\log q}-1} = \sum_{l=0}^{\infty} \frac{B_l}{l!}\big((n+1-r)\log q\big)^{l-1}.$$

Plugging this into the left-hand side of Eq. (11.32), replacing $1-q$ by $-\log q$, and exchanging the summation we get

$$\begin{aligned} &(\log q)^{1-k} \sum_{r=0}^{k} (-1)^r \binom{n+1-r}{k-r}\binom{n}{r} \big/ (q^{n+1-r}-1) \\ &\quad = \sum_{l=0}^{\infty} \frac{B_l}{l!} (\log q)^{l-k} \sum_{r=0}^{k} (-1)^r \binom{n+1-r}{k-r}\binom{n}{r} (n+1-r)^{l-1}. \end{aligned} \tag{11.33}$$

Then the inner sum over r is the coefficient of x^k of the following polynomial

$$\begin{aligned} f_l(x) &= \sum_{k=0}^{n+1}\sum_{r=0}^{k}(-1)^r\binom{n+1-r}{k-r}\binom{n}{r}(n+1-r)^{l-1}x^k \\ &= \sum_{r=0}^{n}(-1)^r\binom{n}{r}(n+1-r)^{l-1}\sum_{k=r}^{n+1}\binom{n+1-r}{k-r}x^k \\ &= (x+1)^{n+1}\sum_{r=0}^{n}(-y)^r\binom{n}{r}(n+1-r)^{l-1}, \end{aligned} \tag{11.34}$$

where $y = x/(x+1)$. When $l = 0$ this expression becomes

$$\begin{aligned} f_0(x) &= \frac{(x+1)^{n+1}}{n+1}\sum_{r=0}^{n}(-y)^r\binom{n+1}{r} \\ &= \frac{(x+1)^{n+1}}{n+1}\Big[(1-y)^{n+1}-(-y)^{n+1}\Big] = \frac{1}{n+1}\Big[1-(-x)^{n+1}\Big]. \end{aligned}$$

Note that $k \le n$ we see the coefficient of x^k in $f_0(x)$ is 0 if $k > 0$ and it's $1/(n+1)$ if $k = 0$. If $k = 0$ then only the constant term $-1/(n+1)$ in Eq. (11.33) remains when $q \to 1$ which proves the corollary in this case. So we can assume $l, k > 0$. Then

$$\begin{aligned} f_l(x) &= \left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)^{l-1}\left\{(x+1)^{n+1}\sum_{r=0}^{n}(-y)^r\binom{n}{r}z^{n+1-r}\right\}\Big|_{z=1} \\ &= \left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)^{l-1}\left\{(x+1)^{n+1}z(z-y)^n\right\}\Big|_{z=1}. \end{aligned}$$

Note that highest degree term in $f_l(x)$ is contained in

$$\begin{aligned} (x+1)^{n+1}\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^{l-1}(z-y)^n\Big|_{z=1} &= n(n-1)\cdots(n-l+2)(x+1)^{n+1}(1-y)^{n-l+1} \\ &= n(n-1)\cdots(n-l+2)(x+1)^l. \end{aligned}$$

If $l = 1$ one can easily modify this to get just $x+1$. If $l < k$ then the coefficient of x^k in $f_l(x)$ is 0. If $l = k$ it is equal to

$$n(n-1)\cdots(n-k+2) = (k-1)!\binom{n}{k-1}.$$

The last express is valid even for $k = l = 1$. Thus the range of l in the outer sum of Eq. (11.33) starts from k. Moreover, the first term of Eq. (11.33) is

$$\frac{B_k}{k}\binom{n}{k-1}$$

as desired. This completes the proof of the corollary. □

Proposition 11.3.6. *Let k, n be two non-negative integers such that $n \geq k$ and k is even. Let $m = k - n - 2$. Then*

$$\operatorname*{Res}_{(s_1,s_2)=(-n,-m)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \frac{-f(q)(q-1)/\log q}{D(q)}, \quad D(q) = \prod_{j=n+1-k}^{n+1} F(q,j)^{\epsilon_j}$$

where $f(q) \in \mathbb{Z}[q]$ is a palindrome with leading coefficient $\binom{n}{k}$, $F(q,j) \in \mathbb{Z}[q]$ is a factor of $(q^j-1)/(q-1)$, $\epsilon_j = 0$ or 1, and $\deg_q D(q) = n + \deg_q f(q)$, such that

$$\lim_{q\to 1} \operatorname*{Res}_{(s_1,s_2)=(-n,-m)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \operatorname*{Res}_{(s_1,s_2)=(-n,-m)} \zeta(s_1, s_2).$$

Proof. The computational proof is left as an exercise for the interested readers. See Exercise 11.6. □

Example 11.3.7. By Theorem 11.3.3 we find with the help of Maple

$$\operatorname*{Res}_{(s_1,s_2)=(-4,4)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \frac{-2q^3(3q^2+4q+3)(q-1)/\log q}{P_1(q,2)P_1(q,3)P_1(q,4)}.$$

where $P_a(q,m) = \sum_{j=0}^{m} q^{aj}$. Moreover we can check that

$$\lim_{q\to 1} \operatorname*{Res}_{(s_1,s_2)=(-4,4)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \operatorname*{Res}_{(s_1,s_2)=(-4,4)} \zeta(s_1, s_2) = -\frac{1}{3}$$

by Table 11.1 with $k = 2$ and $n = 4$.

Example 11.3.8. By Theorem 11.3.3 we get

$$\operatorname*{Res}_{(s_1,s_2)=(-8,6)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \frac{-14q^5 g(q)(q-1)/\log q}{P_1(q,4)P_1(q,5)P_1(q,6)P_2(q,3)P_3(q,2)}$$

where $g(q)$ is a polynomial in q of degree 14 satisfying

$q^{14} f(1/q) = f(q) = 5q^{14}+6q^{13}+8q^{12}+7q^{11}-q^{10}-20q^9-30q^8-34q^7-\cdots$.

Then we can compute with Maple

$$\lim_{q\to 1} \operatorname*{Res}_{(s_1,s_2)=(-8,6)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \operatorname*{Res}_{(s_1,s_2)=(-8,6)} \zeta(s_1, s_2) = \frac{7}{15}$$

by Table 11.1 with $k = 4$ and $n = 8$.

Example 11.3.9. Consider the point $(s_1, s_2) = (-9, 5)$. We have

$$\operatorname*{Res}_{(s_1,s_2)=(-9,5)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \frac{-42q^4 g(q)(q-1)/\log q}{P_1(q,4)P_1(q,6)P_1(q,7)P_2(q,2)P_3(q,2)A(q,4)}$$

where $A(q,m) = \sum_{j=0}^{m} (-1)^j q^j$ and $g(q)$ is a polynomial in q of degree 18 satisfying

$q^{18} g(1/q) = g(q) = 2q^{18}-q^{17}-7q^{15}-11q^{14}-16q^{13}-4q^{12}+9q^{11}+28q^{10}+30q^9+\cdots$

so that we again have the equality

$$\lim_{q\to 1} \operatorname*{Res}_{(s_1,s_2)=(-9,5)} \zeta_q^{\mathrm{I}}[s_1, s_2] = \operatorname*{Res}_{(s_1,s_2)=(-9,5)} \zeta(s_1, s_2) = -\frac{1}{2}$$

by Table 11.1 with $k = 6$ and $n = 9$.

Theorem 11.3.10. *Let m, n be two non-negative integers and $k = m + n + 2$. Then $\zeta_q^{\mathrm{I}}[s_1, s_2]$ has indeterminacy at $(-n, -m)$ such that*

$$\begin{aligned}\zeta_q^{\mathrm{I}}[-n,-m] =&(1-q)^{2-k}\left\{\frac{(-1)^k}{(m+1)(n+1)(\log q)^2}+\sum_{r=0}^{m}\frac{(-1)^{r+n+1}}{(n+1)\log q}\binom{m}{r}\frac{1}{q^{m+1-r}-1}\right.\\&+\sum_{r=0}^{n}\frac{(-1)^{r+m+1}}{\log q}\frac{m!(n+1-r)!}{(k-r)!}\binom{n}{r}\frac{1}{q^{n+1-r}-1}\\&\left.+\sum_{r_1=0}^{m}\sum_{r_2=0}^{n}(-1)^{r_1+r_2}\binom{m}{r_1}\binom{n}{r_2}\frac{1}{q^{n+1-r_2}-1}\frac{1}{q^{k-r_1-r_2}-1}\right\},\end{aligned}$$

and

$$\begin{aligned}\zeta_q^{R}[-n,-m] =&(1-q)^{2-k}\left\{\sum_{r=0}^{m}\frac{(-1)^{r+n+1}}{(n+1)\log q}\binom{m}{r}\frac{1}{q^{m+1-r}-1}\right.\\&+\sum_{r=0}^{m}\frac{(-1)^{r+n}}{\log q}\binom{k-n-2}{r}\frac{n!(m+1-r)!}{(k-r)!}\frac{q^{m+1-r}}{q^{m+1-r}-1}\\&\left.+\sum_{r_1=0}^{m}\sum_{r_2=0}^{n}(-1)^{r_1+r_2}\binom{m}{r_1}\binom{n}{r_2}\frac{1}{q^{n+1-r_2}-1}\frac{1}{q^{k-r_1-r_2}-1}\right\}.\end{aligned}$$

Proof. This follows from Eq. (11.24) by straightforward computation. □

Similar to Cor. 11.3.4 and Cor. 11.3.5 we have

Corollary 11.3.11. *Let m and n be two non-negative integers. Then*

$$\lim_{q\to 1}\zeta_q^{\mathrm{I}}[-n,-m]=\zeta(-n,-m),\quad \lim_{q\to 1}\zeta_q^{R}[-n,-m]=\zeta^{R}(-n,-m). \tag{11.35}$$

Proof. Set $k = m + n + 2$. We consider $\zeta_q^R[-n, -m]$ first. From Theorem 11.3.10 we get

$$\zeta_q^{R}[-n,-m]=(q-1)^{-k}\left(\frac{q-1}{\log q}\right)^2(A+B+C),$$

where

$$\begin{aligned}A=&\sum_{j=0}^{\infty}\frac{B_j}{j!}(\log q)^j\sum_{r=0}^{m}\frac{(-1)^{m+r+1}}{n+1}\binom{m}{r}(m+1-r)^{j-1},\\B=&\sum_{i=0}^{\infty}\frac{B_i}{i!}(\log q)^i\sum_{r=0}^{m}(-1)^{r+m+i}\binom{m}{r}\frac{n!(m+1-r)!}{(k-r)!}(m+1-r)^{i-1},\\C=&\sum_{i,j=0}^{\infty}\frac{B_iB_j}{i!j!}(\log q)^{i+j}\sum_{r_1=0}^{m}\sum_{r_2=0}^{n}(-1)^{k+r_1+r_2}\binom{m}{r_1}\binom{n}{r_2}\times\\&\qquad\times(n+1-r_2)^{i-1}(k-r_1-r_2)^{j-1}.\end{aligned}$$

We first compute B as follows. Write

$$B = \sum_{i=0}^{\infty} \frac{B_i}{i!} (\log q)^i W(m,n,i),$$

where

$$W(m,n,i) = \sum_{r=0}^{m} (-1)^{r+k+n+i} \frac{m!n!}{r!(k-r)!} (m+1-r)^i.$$

If $i = 0$ then we can prove by decreasing induction on n that

$$W(m,n,0) = \sum_{r=0}^{m} (-1)^{r+k+n} \frac{m!n!}{r!(k-r)!} = \frac{1}{k(n+1)}. \tag{11.36}$$

This is trivial if $n = k-2$. Suppose Eq. (11.36) is true for $n \geq 1$ then we have

$$W(m,n-1,0) = -\frac{k-n-1}{n} W(m,n,0) + \frac{(n-1)!}{(n+1)!} = \frac{1}{kn},$$

as desired.

Similarly, we can compute C as follows. Put

$$C = \sum_{i,j=0}^{\infty} \frac{B_i B_j (\log q)^{i+j}}{i!j!(n+1)} \sum_{r_1=0}^{m} \sum_{r_2=0}^{n} (-1)^{k+r_1+r_2} \binom{m}{r_1} \binom{n+1}{r_2} (n+1-r_2)^i (k-r_1-r_2)^{j-1}.$$

We now change the upper limit of r_2 from n to $n+1$ in the above. The extra terms correspond to those by setting $i = 0$ and $r_2 = n+1$, which produce exactly A. Therefore,

$$C = \sum_{i,j=0}^{\infty} \frac{B_i B_j (\log q)^{i+j}}{i!j!} V(m,n,i,j) - A$$

where

$$V(m,n,i,j) = \sum_{r_1=0}^{m} \sum_{r_2=0}^{n+1} \frac{(-1)^{k+r_1+r_2}}{n+1} \binom{m}{r_1} \binom{n+1}{r_2} (n+1-r_2)^i (k-r_1-r_2)^{j-1}.$$

For $j \geq 1$ we have

$$\begin{aligned} V(m,n,i,j) &= \left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{j-1} \left\{ \left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i \Big\{ \sum_{r_1=0}^{m} \sum_{r_2=0}^{n+1} \frac{(-1)^{k+r_1+r_2}}{n+1} \times \right. \\ &\qquad \left. \times \binom{m}{r_1} \binom{n+1}{r_2} y^{n+1-r_2} x^{k-r_1-r_2} \Big\} \Big|_{y=1} \right\} \Big|_{x=1} \\ &= \frac{(-1)^k}{n+1} \left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{j-1} \left\{ \left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i \Big\{ x(xy-1)^{n+1}(x-1)^{k-n-2} \Big\} \Big|_{y=1} \right\} \Big|_{x=1} \end{aligned} \tag{11.37}$$

$$= \begin{cases} 0 & \text{if } i+j < k, \\ \dfrac{(-1)^k n!}{(n+1-i)!} (k-i-1)! & \text{if } i+j = k, i \leq n+1. \end{cases}$$

We have used the fact that if $i+j=k$ and $i>n+1$ then $j<k-n-1$ and by exchanging the two operators $x(d/dx)$ and $y(d/dy)$ we can easily show that Eq. (11.37) is zero. So if $l<k$ the total contribution to the coefficient of $(\log q)^l$ from $V(m,n,i,j)$ with $j>0$ is trivial and if $l=k$ it is equal to

$$\sum_{i=0}^{n+1} B_i B_{k-i} \frac{(-1)^k n!(k-i-1)!}{i!(k-i)!(n+1-i)!} = \begin{cases} \dfrac{B_{k-1}}{2(k-1)} & \text{if } 2 \nmid k, \\ \dfrac{B_k}{k(n+1)} + \displaystyle\sum_{i=1}^{n+1} \frac{B_i B_{k-i}}{i(k-i)} \binom{n}{i-1} & \text{if } 2 \mid k, \end{cases} \tag{11.38}$$

because $k \geq n+2 \geq 2$ and $B_k = 0$ if k is odd.

To deal with $V(m,n,i,0)$ note that equation Eq. (11.37) still makes sense if we interpret the operator $x(d/dx)^{-1}$ as follows:

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{-1} \Big\{F(x)\Big\}\Big|_{x=1} = \int_0^1 \frac{F(x)}{x} dx$$

whenever $F(0)=0$. Thus we get

$$V(m,n,i,0) = \frac{(-1)^k}{n+1} \left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i \left\{ \int_0^1 (xy-1)^{n+1}(x-1)^m \, dx \right\}\Bigg|_{y=1}.$$

Therefore if $i=0$ then we get

$$V(m,n,0,0) = \frac{(-1)^k}{n+1} \int_0^1 (x-1)^{k-1} \, dx = -\frac{1}{k(n+1)} = -W(m,n,0) \tag{11.39}$$

from equation Eq. (11.36). If $i \geq 1$ then integrating by parts we get

$$\begin{aligned}
&\int_0^1 (xy-1)^{n+1}(x-1)^m \, dx \\
=&\frac{(xy-1)^{n+2}}{y(n+2)}(x-1)^m\Big|_0^1 - \frac{m}{y(n+2)} \int_0^1 (xy-1)^{n+2}(x-1)^{m-1} \, dx \\
=&\cdots\cdots \\
=&(-1)^{k+1}\left(\frac{1}{y(n+2)} - \frac{m}{y^2(n+2)(n+3)} + \cdots + (-1)^{m-1}\frac{m!(n+1)!}{y^m(m+n+1)!}\right. \\
&\left. +(-1)^m \frac{m!(n+1)!}{y^m(m+n+1)!} \int_0^1 (xy-1)^{m+n+1} \, dx, \right) \\
=&(-1)^{k+n}\frac{m!(n+1)!}{k!}\frac{(y-1)^k}{y^{m+1}} + \sum_{r=0}^{m}(-1)^{r+k+1}\frac{m!(n+1)!}{y^{r+1}(m-r)!(n+2+r)!}.
\end{aligned}$$

It follows from changing the index r to $m-r$ that

$$V(m,n,i,0) = (-1)^{k+n}\frac{m!n!}{k!}\left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i\left\{\frac{(y-1)^k}{y^{m+1}}\right\}\Bigg|_{y=1} \tag{11.40}$$

$$+\sum_{r=0}^{m}(-1)^{r+n+i+1}\frac{m!n!}{r!(k-r)!}(m+1-r)^i$$

$$=\begin{cases}(-1)^{k+1}W(m,n,0) & \text{if } 0<i<k,\\ (-1)^{k+1}W(m,n,0)+(-1)^{k+n}n!(k-n-2)! & \text{if } i=k.\end{cases} \tag{11.41}$$

Thus when $0<i<k$ and k is even we have $V(m,n,i,0) = -W(m,n,i)$. It follows from Eq. (11.38), Eq. (11.39) and Eq. (11.41) that

$$\lim_{q\to 1}\zeta_q^R(-m,-n) = \zeta^R(s_1,s_2)$$

since $B_k = 0$ if $k>2$ is odd.

Let's turn to prove the first equality in Eq. (11.35). By Theorem 11.3.10 we have

$$\zeta_q^{\mathrm{I}}[-n,-m] = (q-1)^{-k}\left(\frac{q-1}{\log q}\right)^2(D+A+E+C)$$

where A and C are as above and

$$D = \frac{1}{(m+1)(n+1)},\quad E = \sum_{i=0}^{\infty}\frac{B_i}{i!}(\log q)^i U(m,n,i),$$

where

$$U(m,n,i) = -\sum_{r=0}^{n}(-1)^{r+n}\frac{m!n!}{r!(k-r)!}(n+1-r)^i.$$

Hence

$$U(m,n,0) = W(n,m,0) = \frac{-1}{k(m+1)} = \frac{-1}{k(k-n-1)} = \frac{1}{k(n+1)} - D. \tag{11.42}$$

We only need to show that

$$U(m,n,i) = \begin{cases}W(m,n,i) & \text{if } 0<i<k,\\ W(m,n,i)-(-1)^n n!(k-n-2)! & \text{if } i=k.\end{cases} \tag{11.43}$$

Indeed when $i > 0$ we have

$$\begin{aligned}
U(m,n,i) &= (-1)^{n+1}\left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i\left\{\sum_{r=0}^{n}(-1)^r\frac{m!n!}{k!}\binom{k}{r}y^{n+1-r}\right\}\Bigg|_{y=1}\\
&=(-1)^{n+1}\frac{m!n!}{k!}\left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i\left\{y^{n+1-k}(y-1)^k-\sum_{r=n+1}^{k}(-1)^r\binom{k}{r}y^{n+1-r}\right\}\Bigg|_{y=1}\\
&=(-1)^{n+1}\frac{m!n!}{k!}\left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^i\left\{\frac{(y-1)^k}{y^{m+1}}\right\}+\sum_{r=n+1}^{k}(-1)^{r+n}\frac{m!n!}{r!(k-r)!}(n+1-r)^i.
\end{aligned}$$

Then first term is 0 if $i < k$ and it's $(-1)^{n+1}m!n!$ if $i = k$. When $r = n+1$ the summand in the second term is zero since $i > 0$. So we can let r range only from $n+2$ to k. Then change the index r to $k-r$ (and let r run from 0 to m) we can see immediately that the second term is exactly $W(m,n,i)$. This proves Eq. (11.43) which together with Eq. (11.42) implies the first equation in Eq. (11.35). We thus finish the proof of the corollary. □

We conclude this section by remarking that by the stuffle relation Eq. (12.23) from Chap. 12 we can also analyze $\zeta_q^{\mathrm{I}}[s_1,s_2]$ at $(-n,n+2-k)$ for any non-negative integers k and n. For example, it's easy to compute directly that

$$\begin{aligned}
\operatorname*{Res}_{(s_1,s_2)=(2,-3)}\zeta_q^{\mathrm{I}}[s_1,s_2] &= -\frac{1}{(1-q)\log q}\sum_{r=0}^{3}(-1)^r\binom{3}{r}(r+1)\frac{q^{r+1}}{1-q^{r+1}}\\
&=\frac{-q(q-1)^2}{(q+1)(q^2+1)(q^2+q+1)\log q},
\end{aligned}\tag{11.44}$$

which can be obtained also by the shuffle relation Eq. (12.23) and the expression

$$\operatorname*{Res}_{(s_1,s_2)=(-3,2)}\zeta_q^{\mathrm{I}}[s_1,s_2]=\frac{q(q-1)^2}{(q+1)(q^2+1)(q^2+q+1)\log q}$$

by taking $k = n = 3$ in Theorem 11.3.3. Thus $(2,-3)$ is a simple pole of the double q-zeta function $\zeta_q^{\mathrm{I}}[s_1,s_2]$. On the other hand the ordinary double zeta function $\zeta(s_1,s_2)$ does not have a pole along $s_1+s_2=-1$. Indeed from Eq. (11.44) we find that

$$\lim_{q\to 1}\operatorname*{Res}_{(s_1,s_2)=(2,-3)}\left\{\zeta_q^{\mathrm{I}}[s_1,s_2]\right\}=0.$$

11.4 Historical notes

The q-analog of Riemann zeta function is studied in [321]. In particular, the depth one case of Theorem 11.2.1 and Theorem 11.3.1 are proved there. The general multiple q-zeta function is treated systematically in [594] from the analytic point of view. Besides the analytic continuation, the author also studied the special values of the double q-zeta functions at non-negative integers.

In [321], Kawagoe *et al.* showed that if $\varphi(s)$ is a meromorphic function of s then the limit $\lim_{q\to 1}\zeta_q^{\varphi(s)}[s] = \zeta(s)$ if and only if $\varphi(s) = s - \nu$ for some $\nu \in \mathbb{N}$. We don't know in depth $d > 1$ case what one should expect of $\varphi_j(\mathbf{s})$ if $\lim_{q\to 1}\zeta_q^{(\varphi_1(\mathbf{s}),\ldots,\varphi_d(\mathbf{s}))}[\mathbf{s}] = \zeta[\mathbf{s}]$.

Exercises

11.1. Prove Prop. 11.1.5 completely using induction and the binomial identity

$$\binom{n}{r} + \binom{n}{r-1} = \binom{n+1}{r}.$$

11.2. Show by substitution that

$$\int_1^\infty \frac{q^{x(s-1)}}{(1-q^x)^s}\,dx = -\frac{q^{s-1}(1-q)^{1-s}}{(s-1)\log q}.$$

11.3. Prove Eq. (11.14) by repeatedly using integration by parts.

11.4. Show that

$$\lim_{q\to 1}(1-q)^s \log q \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{s}{(2\pi i n)^2} b_q(s-1+\delta n, -s) = 0,$$

$$\lim_{q\to 1}(1-q)^s \log q \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{(2\pi i n)^2} b_q(s-1+\delta n, -s+1) = 0.$$

11.5. Let $M > 2N + 16$ and $x, k > 0$. Prove that

$$g(q) = q^{x(M-N-2)-kM/6}\left(\frac{1-q}{1-q^x}\right)^{M-N+1}$$

is an increasing function for $1/2 < q < 1$.

11.6. Prove Prop. 11.3.6 using Eq. (11.24).

11.7. Show that D_0 and D_1 satisfy the same bound as Eq. (11.22) if $1/2 < q < 1$ and $M > 16 + 2N_1 + 6\sum_{j=2}^d (N_j + 1)$ where the positive integers $N_j > |\mathrm{Re}(s_j)|$ for all j.

Chapter 12

q-analogs of multiple zeta (star) values

There are quite a few different ways to define q-analogs of MZVs and MZ*Vs. A good analog should not only deform into its classical counterpart when $q \to 1$ but also satisfy the key relations that hold in the classical setting. Therefore, we seek for the q-analog of MZVs (q-MZVs) that obey some kind of double shuffle relations (DBSF). On the other hand, for q-analogs and MZ*Vs (q-MZ*Vs), we choose those which enjoy interesting identities similar to the non-q version. In this chapter we present some of these and consider their important properties.

12.1 Various definitions of q-MZVs

We have seen from last chapter that a good q-analog of the multiple zeta function is defined by

$$\zeta_q^{\mathrm{I}}[\mathbf{s}] := \zeta_q^{\mathbf{s-1}}[\mathbf{s}] = \sum_{k_1>\cdots>k_d\geq 1} \frac{q^{k_1(s_1-1)}\cdots q^{k_d(s_d-1)}}{[k_1]^{s_1}\cdots k_d]^{s_d}}.$$

We can define two more analogs as follows:

$$\zeta_q^{\mathrm{II}}[\mathbf{s}] := \zeta_q^{\mathbf{s}}[\mathbf{s}] = \sum_{k_1>\cdots>k_d\geq 1} \prod_{j=1}^{d} \frac{q^{k_1 s_1}\cdots q^{k_d s_d}}{[k_1]^{s_1}\cdots k_d]^{s_d}}, \tag{12.1}$$

$$\zeta_q^{\mathrm{III}}[\mathbf{s}] := \zeta_q^{(1,0,\ldots,0)}[\mathbf{s}] = \sum_{k_1>\cdots>k_d\geq 1} \frac{q^{k_1}}{[k_1]^{s_1}\cdots k_d]^{s_d}}. \tag{12.2}$$

By Prop. 11.1.1 we obtain the following corollary immediately.

Corollary 12.1.1. *Let* $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{Z}^d$. *Then*

(i) $\zeta_q^{\mathrm{I}}[\mathbf{s}]$ *converges if* $s_1 + \cdots + s_j > j$ *for all* $j = 1, \ldots, d$.
(ii) $\zeta_q^{\mathrm{II}}[\mathbf{s}]$ *converges if* $s_1 + \cdots + s_j > 0$ *for all* $j = 1, \ldots, d$.
(iii) $\zeta_q^{\mathrm{III}}[\mathbf{s}]$ *always converges.*

In this chapter we will also consider two more special types of q-MZVs that are more complicated than the above three types but have some advantages by themselves. Essentially both of them have the form

$$\sum_{k_1>\cdots>k_d\geq 1} \frac{P_{s_1}(q^{k_1})\cdots P_{s_d}(q^{k_d})}{[k_1]^{s_1}\cdots k_d]^{s_d}}$$

where $P_n(t)$ are polynomials of t without constant. The first class, called the Bachmann-Kühn brackets and denoted by $\zeta_q^{\mathrm{BK}}[\mathbf{s}]$, uses the Eulerian polynomial $E_n(t)$ for $P_n(t)$, where

$$E_n(t)=\sum_{n=0}^{s-1} E_{s,n}t^n, \tag{12.3}$$

and the Eulerian numbers $E_{s,n}$ are defined by

$$E_{s,n}=\sum_{i=0}^{n}(-1)^i\binom{s+1}{i}(n+1-i)^s.$$

The second class, called Okounkov's q-MZVs and denoted by $\zeta_q^{\mathrm{O}}[\mathbf{s}]$, is defined using the following palindrome polynomials:

$$P_n(t)=t^{n^+}+t^{n^-} \tag{12.4}$$

where, for any $n\in\mathbb{N}$, n^- and n^+ are the two nonnegative integers such that

$$\frac{n-1}{2}\leq n^-\leq\frac{n}{2}\leq n^+\leq\frac{n+1}{2}.$$

Clearly we have $n^++n^-=n$ always, $n^+=n^-$ if n is even, and $n^+=n^-+1$ if n is odd.

All of the above q-MZVs can be expressed using the most general *type G* q-MZVs $\zeta_q^{(t_1,\ldots,t_d)}[s_1,\ldots,s_d]$ defined by Eq. (11.1) where $1\leq t_1\leq s_1$, $0\leq t_j\leq s_j$ for all $j\geq 2$. They are all convergent by Prop. 11.1.1.

12.2 Double shuffle relations

In this section, we will use a few algebras of words to encode the structure of various analogs of q-MZVs. Motivated by the iterated Jackson q-integrals we can apply Rota-Baxter algebra to study the shuffle algebra. This enables us to obtain the DBSF of these q-MZVs.

12.2.1 *Rota-Baxter algebra*

We first use Rota-Baxter algebra to study the q-analog of shuffle relations for different versions of q-MZVs.

Definition 12.2.1. Fix an algebra A over a commutative ring R and an element $\lambda \in R$. We call A a Rota-Baxter R-algebra and $\mathcal{P}$ a Rota-Baxter operator of weight λ if the operator $\mathcal{P}$ satisfies the following Rota-Baxter relation of weight λ:

$$\mathcal{P}(x)\mathcal{P}(y) = \mathcal{P}(\mathcal{P}(x)y) + \mathcal{P}(x\mathcal{P}(y)) + \lambda\mathcal{P}(xy) \quad \forall x, y \in A. \tag{12.5}$$

Recall that for any continuous function $f(x)$ on $[\alpha, \beta]$ Jackson q-integral is defined by

$$\int_\alpha^\beta f(x)\, d_q x := \sum_{k\geq 0} f\big(\alpha + q^k(\beta-\alpha)\big)(q^k - q^{k+1})(\beta-\alpha). \tag{12.6}$$

Taking $\alpha = 0$ and $\beta = t$ in Eq. (12.6) we now set

$$\begin{aligned}\mathbf{J}[f](t) := (1-q)\sum_{k\geq 0} f(q^k t)q^k t = (1-q)\sum_{k\geq 0} \mathbf{E}^k\big[\mathbf{I}\cdot f\big](t)\\ = (1-q)\mathbf{P}[\mathbf{I}\cdot f](t).\end{aligned} \tag{12.7}$$

where $\mathbf{I}(t) = t$ is the identity function,

$$\begin{aligned}\mathbf{E}[f](t) :=& \mathbf{E}_q[f](t) := f(qt),\\ \mathbf{P}[f](t) :=& \mathbf{P}_q[f](t) := f(t) + f(qt) + f(q^2t) + \cdots\end{aligned}$$

are the *q-expanding* and the (*principle*) *q-summation operators*, respectively. We also need to define the (*remainder*) *q-summation operator*

$$\mathbf{R}[f](t) := \mathbf{R}_q[f](t) := f(qt) + f(q^2t) + \cdots = (\mathbf{P}[f] - [f])(t).$$

So, $\mathbf{P}$ is the principle part (*i.e.*, the whole thing) while $\mathbf{R}$ is the remainder (*i.e.*, without the first term). Clearly, $\mathbf{P} = \mathbf{R} + \mathbf{I}$ where, as an operator, $\mathbf{I}[f] = f$.

Let $t\mathbb{Q}[\![t, q]\!]$ be the space of formal series in two variables with $t > 0$. Then $\mathbf{J}$, $\mathbf{E}$, $\mathbf{P}$ and $\mathbf{R}$ are all $\mathbb{Q}[[q]]$-linear endomorphism of $t\mathbb{Q}[\![t, q]\!]$. We can further define the inverse to $\mathbf{P}$ which is called the *q-difference operator*:

$$\mathbf{D} := \mathbf{I} - \mathbf{E}. \tag{12.8}$$

Proposition 12.2.2. *For any $f, g \in t\mathbb{Q}[\![t,q]\!]$ we have*

$$\begin{aligned}
\mathbf{P}[f]\mathbf{P}[g] &= \mathbf{P}\big[\mathbf{P}[f]g\big] + \mathbf{P}\big[f\mathbf{P}[g]\big] - \mathbf{P}[fg], && (12.9)\\
\mathbf{R}[f]\mathbf{R}[g] &= \mathbf{R}\big[\mathbf{R}[f]g\big] + \mathbf{R}\big[f\mathbf{R}[g]\big] + \mathbf{R}[fg], && (12.10)\\
\mathbf{R}[f]\mathbf{P}[g] &= \mathbf{R}\big[\mathbf{R}[f]g\big] + \mathbf{R}\big[f\mathbf{R}[g]\big] + \mathbf{R}[f]g + \mathbf{R}[fg], && (12.11)\\
\mathbf{J}[f]\mathbf{J}[g] &= \mathbf{J}\big[\mathbf{J}[f]g\big] + \mathbf{J}\big[f\mathbf{J}[g]\big] - (1-q)\mathbf{J}\big[\mathbf{I}fg\big], && (12.12)\\
&= \mathbf{J}\big[f\mathbf{J}[g]\big] + q\mathbf{J}\big[\mathbf{J}\big[E[f]\big]g\big], && (12.13)\\
\mathbf{D}[f]\mathbf{D}[g] &= \mathbf{D}[f]g + f\mathbf{D}[g] - \mathbf{D}[fg], && (12.14)\\
\mathbf{D}[f]\mathbf{P}[g] &= \mathbf{D}\big[f\mathbf{P}[g]\big] + \mathbf{D}[f]g - fg, && (12.15)\\
\mathbf{D}[f]\mathbf{R}[g] &= \mathbf{D}\big[f\mathbf{R}[g]\big] + \mathbf{D}[fg] - fg, && (12.16)\\
\mathbf{DP} &= \mathbf{PD} = \mathbf{I}. && (12.17)
\end{aligned}$$

Proof. For the first identity Eq. (12.9), we have

$$\begin{aligned}
\Big(\mathbf{P}[f]\mathbf{P}[g]\Big)(t) &= \sum_{j\geq 0} f(q^j t) \sum_{k\geq 0} g(q^k t)\\
&= \sum_{j\geq k\geq 0} f(q^j t)g(q^k t) + \sum_{k\geq j\geq 0} f(q^j t)g(q^k t) - \sum_{j\geq 0} f(q^j t)g(q^j t)\\
&= \sum_{\ell,k\geq 0} f(q^{\ell+k} t)g(q^k t) + \sum_{\ell,j\geq 0} f(q^j t)g(q^{\ell+j} t) - \mathbf{P}[fg]\\
&= \Big(\mathbf{P}\big[\mathbf{P}[f]g\big] + \mathbf{P}\big[f\mathbf{P}[g]\big] - \mathbf{P}[fg]\Big)(t)
\end{aligned}$$

All the other identities follow from the same type straight-forward computation which is left as an exercise (see Exercise12.7). □

By Prop. 12.2.2 we see that $\mathbf{P}$ and $\mathbf{R}$ are both Rota-Baxter operators on $t\mathbb{Q}[\![t,q]\!]$ (of weight -1 and 1, respectively) but $\mathbf{D}$ is not. In fact, $\mathbf{D}$ satisfies the condition Eq. (12.14) of a differential Rota-Baxter operator (see [251]). Moreover, it is *invertible* in the sense that Rota-Baxter operator $\mathbf{P}$ and the differential $\mathbf{D}$ are mutually inverse by Eq. (12.17).

We end this section with an identity which will be used to derive the duality relations among the q-MZVs. For any $n \in \mathbb{N}$, set

$$\mathbf{P}^n = \underbrace{\mathbf{P}\circ\cdots\circ\mathbf{P}}_{n \text{ times}} \quad \text{and} \quad \mathbf{R}^n = \underbrace{\mathbf{R}\circ\cdots\circ\mathbf{R}}_{n \text{ times}}.$$

Theorem 12.2.3. *Let $d \in \mathbb{N}$ and $\alpha_j, \beta_j \in \mathbb{N}$ for all $j = 1, \dots, d$. Let $\mathbf{y}(t) = \frac{t}{1-t}$. Then we have*

$$\begin{aligned}&\mathbf{R}^{\alpha_1}\mathbf{y}^{\beta_1}\cdots\mathbf{R}^{\alpha_\ell}\mathbf{y}^{\beta_\ell}(t)\\ &\quad= \sum_{\substack{j_1\geq\beta_1,\dots,\,j_\ell\geq\beta_\ell\\ k_1\geq\alpha_1,\dots,\,k_\ell\geq\alpha_\ell}} \prod_{r=1}^{\ell}\left[\binom{j_r-1}{\beta_r-1}\binom{k_r-1}{\alpha_r-1}q^{k_r\sum_{s=r}^{\ell}j_s}t^{j_r}\right]. \quad (12.18)\end{aligned}$$

Proof. First we show that

$$\mathbf{R}^{\alpha}(t^j) = \frac{q^{\alpha j}t^j}{(1-q^j)^{\alpha}} \quad (12.19)$$

In deed, if $\alpha = 1$ then

$$\mathbf{R}(t^j) = \sum_{k\geq 1} q^{kj}t^j = \frac{q^j t^j}{1-q^j}.$$

So Eq. (12.19) can be proved easily by induction.

Now we proceed to prove that for any integer $m \geq 0$

$$\begin{aligned}&\mathbf{R}^{\alpha_1}\mathbf{y}^{\beta_1}\cdots\mathbf{R}^{\alpha_\ell}\left(\mathbf{y}^{\beta_\ell}(t)\cdot t^m\right)\\ &= \sum_{\substack{j_1\geq\beta_1,\dots,\,j_\ell\geq\beta_\ell\\ k_1\geq\alpha_1,\dots,\,k_\ell\geq\alpha_\ell}} t^m\prod_{r=1}^{\ell}\left[\binom{j_r-1}{\beta_r-1}\binom{k_r-1}{\alpha_r-1}q^{k_r(m+\sum_{s=r}^{\ell}j_s)}t^{j_r}\right]. \quad (12.20)\end{aligned}$$

If $\ell = 1$ then we have

$$\begin{aligned}\mathbf{R}^{\alpha}\left(\mathbf{y}^{\beta}(t)\cdot t^m\right) &= \mathbf{R}^{\alpha}\left(\left(\frac{t}{1-t}\right)^{\beta}t^m\right) = \mathbf{R}^{\alpha}\sum_{j\geq 0}\binom{\beta+j-1}{j}t^{m+\beta+j}\\ &= \mathbf{R}^{\alpha}\sum_{j\geq\beta}\binom{j-1}{\beta-1}t^{m+j}\\ &= \sum_{j\geq\beta}\binom{j-1}{\beta-1}\frac{q^{\alpha(m+j)}t^{m+j}}{(1-q^{m+j})^{\alpha}} \qquad \text{(by Eq. (12.19))}\\ &= \sum_{j\geq\beta}\binom{j-1}{\beta-1}\sum_{k\geq 0}\binom{\alpha+k-1}{k}q^{(\alpha+k)(m+j)}t^{m+j}\\ &= \sum_{j\geq\beta}\sum_{k\geq\alpha}\binom{j-1}{\beta-1}\binom{k-1}{\alpha-1}q^{k(m+j)}t^{m+j}.\end{aligned}$$

This proves Eq. (12.20) when $\ell = 1$. In general

$$\begin{aligned}&\mathbf{R}^{\alpha_1}\mathbf{y}^{\beta_1}\cdots\mathbf{R}^{\alpha_{\ell-1}}\Big(\mathbf{y}^{\beta_{\ell-1}}(t)\Big(\mathbf{R}^{\alpha_\ell}\Big(\mathbf{y}^{\beta_\ell}(t)\cdot t^m\Big)\Big)\Big)\\ &=\sum_{\substack{j_\ell\geq\beta_\ell\\ k_\ell\geq\alpha_\ell}}\binom{j_\ell-1}{\beta_\ell-1}\binom{k_\ell-1}{\alpha_\ell-1}q^{k_\ell(m+j_\ell)}\mathbf{R}^{\alpha_1}\mathbf{y}^{\beta_1}\cdots\mathbf{R}^{\alpha_{\ell-1}}\Big(\mathbf{y}^{\beta_{\ell-1}}(t)\cdot t^{m+j_\ell}\Big).\end{aligned}$$

So Eq. (12.20) follows immediately by induction. We can now finish the proof of the theorem by taking $m = 0$. □

Corollary 12.2.4. *Let $d \in \mathbb{N}$ and $\alpha_j, \beta_j \in \mathbb{N}$ for all $j = 1, \ldots, d$. Then we have*

$$\mathbf{R}^{\alpha_1}\mathbf{y}^{\beta_1}\cdots\mathbf{R}^{\alpha_\ell}\mathbf{y}^{\beta_\ell}(1) = \mathbf{R}^{\beta_\ell}\mathbf{y}^{\alpha_\ell}\cdots\mathbf{R}^{\beta_1}\mathbf{y}^{\alpha_1}(1). \tag{12.21}$$

Proof. In Eq. (12.18) we use the substitutions $j_r \leftrightarrow k_{\ell+1-r}$ for all $r = 1, \ldots, \ell$. Then we have

$$\begin{aligned}\sum_{r=1}^{\ell}\sum_{s=r}^{\ell}j_sk_r \longrightarrow \sum_{r=1}^{\ell}\sum_{s=r}^{\ell}j_{\ell+1-r}k_{\ell+1-s} &= \sum_{s=1}^{\ell}\sum_{r=1}^{s}j_{\ell+1-r}k_{\ell+1-s}\\ = \sum_{s=1}^{\ell}\sum_{r=1}^{\ell+1-s}j_{\ell+1-r}k_s = \sum_{s=1}^{\ell}\sum_{r=s}^{\ell}j_rk_s &= \sum_{r=1}^{\ell}k_r\sum_{s=r}^{\ell}j_s.\end{aligned}$$

which follows from $s \leftrightarrow \ell+1-s$ followed by $r \leftrightarrow \ell+1-r$ and $r \leftrightarrow s$. This proves the corollary. □

12.2.2 *q-analogs of Hoffman algebras*

We have seen that DBSF lead to many (and conjecturally all) $\mathbb{Q}$-linear relations among MZVs. To study similar relations of the q-MZVs we can modify the Hoffman algebra in this new setting. First we consider some algebras which will be used to define the stuffle relations later.

Definition 12.2.5. on the alphabet $X_\theta = \{a, b, \theta\}$. Denote by $\mathfrak{A}_\theta = \mathbb{Q}\langle a, b, \theta\rangle$ the (weight) filtered noncommutative polynomial $\mathbb{Q}$-algebra of words from X_θ^*. The weights of a, b and θ are all equal to one. Set

$$\gamma := b - \theta, \qquad \mathsf{z}_s := a^{s-1}b, \quad \mathsf{z}'_s := a^{s-1}\theta, \quad s \in \mathbb{Z}.$$

Let $\mathsf{Y}_{\tilde{\mathrm{I}}} := \{\theta\} \cup \{\mathsf{z}_k\}_{k\geq 1}$, $\mathsf{Y}_{\mathrm{II}} := \{\mathsf{z}'_k\}_{k\geq 0}$, and $\mathsf{Y}_{\mathrm{III}} := \{\mathsf{z}_k\}_{k\in\mathbb{Z}}$. We put a tilde on top of I since we need to consider some kind of regularization due to convergence issues involved in type I q-MZVs. This is realized by the

introduction of the letter θ. Again, we use e for the empty word and Y_τ^* to denote the set of words generated on Y_τ for any type τ.

Let $\mathfrak{A}_{\tilde{\mathrm{I}}}^1$, $\mathfrak{A}_{\mathrm{II}}^1$ and $\mathfrak{A}_{\mathrm{III}}$ be the subalgebra of $\mathfrak{A}_\theta$ freely generated by the sets $\mathsf{Y}_{\tilde{\mathrm{I}}}$, Y_{II} and $\mathsf{Y}_{\mathrm{III}}$, respectively. Set

$$\mathfrak{A}_{\mathrm{III}}^1 = \sum_{k\in\mathbb{Z}} \mathsf{z}'_k \mathfrak{A}_{\mathrm{III}} \not\subset \mathfrak{A}_{\mathrm{III}}.$$

Here, all integer subscripts are allowed in $\mathsf{Y}_{\mathrm{III}}$ because type III q-MZVs converge for all integer arguments. Further, we define the following subalgebras corresponding to the convergent values:

$$\mathfrak{A}_{\mathrm{I}}^0 := \mathbb{Q}\mathsf{e} + \sum_{k\geq 2} \mathsf{z}_k \mathfrak{A}_{\tilde{\mathrm{I}}}^1, \qquad \mathfrak{A}_{\tilde{\mathrm{I}}}^0 := \mathbb{Q}\mathsf{e} + \theta\mathfrak{A}_{\tilde{\mathrm{I}}}^1 + \sum_{k\geq 2} \mathsf{z}_k \mathfrak{A}_{\tilde{\mathrm{I}}}^1 \subsetneq \mathfrak{A}_{\tilde{\mathrm{I}}}^1,$$

$$\mathfrak{A}_{\mathrm{II}}^0 := \mathbb{Q}\mathsf{e} + \sum_{k\geq 1} \mathsf{z}'_k \mathfrak{A}_{\mathrm{II}}^1 \subsetneq \mathfrak{A}_{\mathrm{II}}^1, \qquad \mathfrak{A}_{\mathrm{III}}^0 := \mathfrak{A}_{\mathrm{III}}^1.$$

For each type τ the words in $\mathfrak{A}_\tau^0$ are called *type τ-admissible*. This is consistent with Prop. 11.1.1 since we only consider non-negative compositions $\mathbf{s}$.

To define the stuffle product for type $\tau = \tilde{\mathrm{I}}$ and II, similar to the MZV case we define a commutative product $[-,-]_\tau$ first:

$$[\mathsf{z}_k, \mathsf{z}_l]_{\tilde{\mathrm{I}}} = \mathsf{z}_{k+l} + \mathsf{z}_{k+l-1}, \quad [\theta, \mathsf{z}_k]_{\tilde{\mathrm{I}}} = \mathsf{z}_{k+1}, \quad [\theta,\theta]_{\tilde{\mathrm{I}}} = \mathsf{z}_2 - \theta, \quad [\mathsf{z}'_k, \mathsf{z}'_l]_{\mathrm{II}} = \mathsf{z}'_{k+l} \tag{12.22}$$

for all $k, l \geq 1$. Now we define the stuffle product $*_\tau$ on $\mathfrak{A}_\tau^1$ inductively as follows. For any words $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_\tau^1$ and letters $\alpha, \beta \in \mathsf{Y}_\tau$, we set $\mathsf{e} *_\tau \mathbf{u} = \mathbf{u} = \mathbf{u} *_\tau \mathsf{e}$ and

$$(\alpha\mathbf{u}) *_\tau (\beta\mathbf{v}) = \alpha(\mathbf{u} *_\tau \beta\mathbf{v}) + \beta(\alpha\mathbf{u} *_\tau \mathbf{v}) + [\alpha,\beta]_\tau(\mathbf{u} * \mathbf{v}). \tag{12.23}$$

Remark 12.2.6. (i). The definition for $*_{\tilde{\mathrm{I}}}$ is the same as in [512].

(ii). One can check that $*_\tau$ is well-defined for $\tau = \tilde{\mathrm{I}}$ and II. Namely, $\mathbf{u} *_\tau \mathbf{v} \in \mathfrak{A}_\tau^1$ if $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_\tau^1$.

(iii). It is not hard to check that for $\tau = \tilde{\mathrm{I}}$ and II, $(\mathfrak{A}_\tau^0, *_\tau) \subset (\mathfrak{A}_\tau^1, *_\tau)$ as subalgebras.

Definition 12.2.7. Define the injective shifting operator $\mathcal{S}_-$ on any word of $\mathfrak{A}_{\mathrm{III}}^0$ by acting on the first letter:

$$\mathcal{S}_-(\mathsf{z}'_n \mathbf{w}) := \mathsf{z}_n \mathbf{w} - \mathsf{z}_{n-1}\mathbf{w} \quad \text{for all } n \in \mathbb{Z} \text{ and } \mathbf{w} \in \mathsf{Y}_{\mathrm{III}}^*. \tag{12.24}$$

For any $k, l \in \mathbb{Z}$ and any $\mathbf{u}, \mathbf{v} \in \mathsf{Y}_{\mathrm{III}}^*$ define the stuffle product $*_{\mathrm{III}}$ by

$$\mathsf{z}'_k\mathbf{u} *_{\mathrm{III}} \mathsf{z}'_l\mathbf{v} = \mathsf{z}'_k\big(\mathbf{u} *_{\mathrm{II}} \mathcal{S}_-(\mathsf{z}'_l\mathbf{v})\big) + \mathsf{z}'_l\big(\mathcal{S}_-(\mathsf{z}'_k\mathbf{u}) *_{\mathrm{II}} \mathbf{v}\big) + (\mathsf{z}'_{k+l} - \mathsf{z}'_{k+l-1})(\mathbf{u} *_{\mathrm{II}} \mathbf{v}).$$

Here $*_{\mathrm{II}}$ is the ordinary stuffle with $[\mathsf{z}_r, \mathsf{z}_s]_{\mathrm{II}} = \mathsf{z}_{r+s}$ for all $r, s \in Z$.

Lemma 12.2.8. *The stuffle products $*_{\mathbb{III}}$ is well-defined. Namely, if $\mathbf{u}, \mathbf{v} \in \mathfrak{A}^0_{\mathbb{III}}$ then $\mathbf{u} *_{\mathbb{III}} \mathbf{v} \in \mathfrak{A}^0_{\mathbb{III}}$.*

Proof. First we notice that $k + l - 1 \geq 1$ if $k, l \geq 1$. So the first word of each of the terms of $\mathbf{u} *_{\widetilde{\mathbb{IV}}} \mathbf{v}$ has the right form. We need to show that after truncating the first word each term lies in $\mathfrak{A}_{\widetilde{\mathbb{IV}}}$. Notice that $\mathcal{S}_+(\mathbf{z}_l\mathbf{v}), \mathcal{S}_+(\mathbf{z}_k\mathbf{u}) \in \mathfrak{A}_{\widetilde{\mathbb{IV}}}$ and $*_{\mathbb{II}}$ does not decrease the size the subscripts (which are all non-negative). The lemma is now proved. □

Proposition 12.2.9. *Let $\tau = \tilde{\mathbb{I}}, \mathbb{II}$ or $\mathbb{III}$. Then the stuffle algebras $(\mathfrak{A}^1_\tau, *_\tau)$ are all commutative and associative.*

Proof. This follows from the fact that the product $[-,-]_\tau$ are all commutative and associative which can be verified easily. □

We now turn to the shuffle algebra which is an analog of the corresponding algebra for MZVs reflecting the properties of their representations using iterated integrals.

Definition 12.2.10. Let $X_\pi = \{\pi, \delta, y\}$ be an alphabet and X^*_π be the set of words generated by X_π. Define $\mathfrak{A}_\pi = \mathbb{Q}\langle\pi, \delta, y\rangle$ to be the noncommutative polynomial $\mathbb{Q}$-algebra of words of X^*_π. We may embed $\mathfrak{A}_\theta$ defined by Definition 12.2.5 as a subalgebra of $\mathfrak{A}_\pi$ in two different ways: put $\rho = \pi - \mathsf{e}$ and let

$$\begin{aligned} &(A) \quad a := \pi, \quad a^{-1} := \delta, \quad b := \pi y, \quad \theta = \rho y \quad \Longrightarrow \quad \gamma := y, \\ &(B) \quad a := \rho, \quad a^{-1} := -, \quad b := \pi y, \quad \theta = \rho y \quad \Longrightarrow \quad \gamma := y. \end{aligned}$$

We denote the image of the embedding by $\mathfrak{A}^{(A)}_\theta$ and $\mathfrak{A}^{(B)}_\theta$, respectively. The dash $-$ for the image of a^{-1} in (B) means it does not matter what image we choose since a^{-1} only appears when we consider type $\mathbb{III}$ q-MZVs using (A). We will use embedding (B) for the other types for which a^{-1} will not be utilized essentially because of convergence issues.

12.2.3 *q-stuffle relations*

We now move on to the second ingredient of the DBSF, the stuffle relations. Define the $\mathbb{Q}$-linear realization maps $\mathfrak{z}_q : \mathfrak{A}^0_\tau \to \mathbb{C}$ $(\tau = \tilde{\mathbb{I}}, \mathbb{II})$ by $\mathfrak{z}_q[\mathsf{e}] = 1$ and

$$\mathfrak{z}_q[y^\tau_1 \dots y^\tau_d] := \sum_{k_1 > \dots > k_d > 0} M^\tau_{k_1}(y^\tau_1) \dots M^\tau_{k_d}(y^\tau_d),$$

where $y_1^\tau \dots y_d^\tau \in \mathfrak{A}_\tau^0$ and the $\mathbb{Q}$-linear maps

$$M_k^{\tilde{\mathrm{I}}}(\theta) := \frac{q^k}{(1-q^k)}, \quad M_k^{\tilde{\mathrm{I}}}(\mathsf{z}_s) := \frac{q^{(s-1)k}}{(1-q^k)^s}, \quad M_k^{\mathrm{II}}(\mathsf{z}'_s) := \frac{q^{sk}}{(1-q^k)^s}.$$

Note that $M_k^{\tilde{\mathrm{I}}}(\gamma) = M_k^{\tilde{\mathrm{I}}}(\mathsf{z}_1 - \theta) = 1$. For example, we have

$$\mathfrak{z}_q[\mathsf{z}_2\mathsf{z}_5\gamma^2\mathsf{z}_1] = \mathfrak{z}_q^{(1,4,0,0,0)}[2,5,0,0,1], \quad \mathfrak{z}_q[\theta\mathsf{z}_7\theta\mathsf{z}_4] = \mathfrak{z}_q^{(1,6,1,3)}[1,7,1,4],$$

which are not q-MZVs of type I.

For type III, we similarly define the $\mathbb{Q}$-linear realization maps $\mathfrak{z}_q : \mathfrak{A}_{\mathrm{III}}^0 \to \mathbb{C}$ by $\mathfrak{z}_q[\mathsf{e}] = 1$ and

$$\mathfrak{z}_q[y_1 \dots y_d] := \sum_{k_1 > \dots > k_d > 0} M_{k_1}^{1,\mathrm{III}}(y_1) M_{k_2}^{\mathrm{III}}(y_2) \dots M_{k_d}^{\mathrm{III}}(y_d),$$

where $y_1 \dots y_d \in \mathfrak{A}_{\mathrm{III}}^0$ and the $\mathbb{Q}$-linear maps

$$M_k^{1,\mathrm{III}}(\mathsf{z}'_s) := \frac{q^k}{(1-q^k)^s}, \quad M_k^{\mathrm{III}}(\mathsf{z}_s) := \frac{1}{(1-q^k)^s}.$$

Theorem 12.2.11. *Let $\tau = \tilde{\mathrm{I}}$, II or III. For any $\mathbf{u}_\tau, \mathbf{v}_\tau \in \mathfrak{A}_\tau^0$ we have*

$$\mathfrak{z}_q[\mathbf{u}_\tau *_\tau \mathbf{v}_\tau] = \mathfrak{z}_q[\mathbf{u}_\tau]\mathfrak{z}_q[\mathbf{v}_\tau]. \tag{12.25}$$

Proof. Let consider type $\tilde{\mathrm{I}}$ first. For a positive integer N we define the $\mathbb{Q}$-linear map $F_N : \mathfrak{A}_{\mathrm{I}}^1 \to \mathbb{C}$ by $F_N(\mathsf{e}) = 1$ and

$$F_N(\alpha_1 \dots \alpha_d) = \sum_{N > k_1 > \dots > k_d > 0} M_{k_1}^{\tilde{\mathrm{I}}}(\alpha_1) \dots M_{k_d}^{\tilde{\mathrm{I}}}(\alpha_d), \tag{12.26}$$

where $\alpha_j \in \mathsf{Y}_{\tilde{\mathrm{I}}}$ for all $j = 1, \dots, d$. Since $F_N(\alpha\mathbf{u}) = \sum_{N>m>0} M_m^{\tilde{\mathrm{I}}}(\alpha) F_m(\mathbf{u})$ for any $\alpha \in \mathsf{Y}_{\tilde{\mathrm{I}}}$ and $\mathbf{u} \in \mathfrak{A}_\theta^1$. we see that $\mathfrak{z}_q^{\mathrm{I}}(\mathbf{u}) = \lim_{N\to\infty} F_N(\mathbf{u})$. Thus we only need to show that

$$F_N(\mathbf{u} * \mathbf{v}) = F_N(\mathbf{u}) F_N(\mathbf{v}) \tag{12.27}$$

for all words $\mathbf{u}$, $\mathbf{v}$ in Y_θ not starting with z_1. Indeed, Eq. (12.27) can be proved by using induction on $|\mathbf{u}| + |\mathbf{v}|$. Note that Eq. (12.27) is trivial if $\mathbf{u}$ or $\mathbf{v}$ is empty word e. So we assume $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_\theta^0$ and $\alpha, \beta \in \mathsf{Y}_\theta \setminus \{\mathsf{z}_1\}$. Then

$$\begin{aligned} F_N(\alpha\mathbf{u})F_N(\beta\mathbf{v}) &= \sum_{N>m>0} \Big(M_m^{\tilde{\mathrm{I}}}(\alpha) F_m(\mathbf{u}) F_m(\beta\mathbf{v}) + M_m^{\tilde{\mathrm{I}}}(\beta) F_m(\alpha\mathbf{u}) F_m(\mathbf{v}) \\ &\qquad + M_m^{\tilde{\mathrm{I}}}(\alpha) M_m^{\tilde{\mathrm{I}}}(\beta) F_m(\mathbf{u}) F_m(\mathbf{v}) \Big) \\ &= \sum_{N>m>0} \Big(M_m^{\tilde{\mathrm{I}}}(\alpha) F_m(\mathbf{u} * \beta\mathbf{v}) + M_m^{\tilde{\mathrm{I}}}(\beta) F_m(\alpha\mathbf{u} * \mathbf{v}) \\ &\qquad + M_m^{\tilde{\mathrm{I}}}(\alpha) M_m^{\tilde{\mathrm{I}}}(\beta) F_m(\mathbf{u} * \mathbf{v}) \Big) \end{aligned}$$

by inductive assumption. Hence Eq. (12.27) follows from Eq. (12.22) and

$$M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_k)M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_l) = M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_{k+l}) + M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_{k+l-1}), \tag{12.28}$$

$$M_m^{\tilde{\mathrm{I}}}(\theta)M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_k) = M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_{k+1}), \tag{12.29}$$

$$M_m^{\tilde{\mathrm{I}}}(\theta)^2 = M_m^{\tilde{\mathrm{I}}}(\mathbf{z}_2) - M_m^{\tilde{\mathrm{I}}}(\theta) \tag{12.30}$$

for all $k, l \geq 2$ and $m \geq 1$.

The same argument works for type $\mathbb{II}$ and $\mathbb{III}$ if we observe that

$$M_m^{\mathbb{II}}(\mathbf{z}'_k)M_m^{\mathbb{II}}(\mathbf{z}'_l) = M_m^{\mathbb{II}}(\mathbf{z}'_{k+l}), \quad M_m^{1,\mathbb{III}}(\mathbf{z}'_k)M_m^{\mathbb{III}}(\mathbf{z}_l) = M_m^{\mathbb{III}}(\mathbf{z}_{k+l} - \mathbf{z}_{k+l-1}),$$
$$M_m^{\mathbb{III}}(\mathbf{z}_k)M_m^{\mathbb{III}}(\mathbf{z}_l) = M_m^{\mathbb{III}}(\mathbf{z}_{k+l}), \quad M_m^{1,\mathbb{III}}(\mathbf{z}'_k)M_m^{1,\mathbb{III}}(\mathbf{z}'_l) = M_m^{1,\mathbb{III}}(\mathbf{z}'_{k+l} - \mathbf{z}'_{k+l-1}),$$

for all $k, l \geq 0,\ m \geq 1$. This completes the proof of the theorem. □

12.2.4 *Iterated Jackson q-integrals*

As we have seen in section 3.2, the iterated integral representations of the MZVs and its shuffle product structure is the driving force in defining Hoffman's shuffle algebra of words. Set

$$x_0 := x_0(t) = \frac{1}{t}, \quad x_1 := x_1(t) = \frac{1}{1-t}, \quad \mathbf{y} := \mathbf{y}(t) = \frac{t}{1-t}.$$

Recall that for $a = x_0(t)dt$ and $b = x_1(t)dt$, we can express MZVs by Chen's iterated integrals:

$$\zeta(s_1, \ldots, s_d) = \int_0^1 a^{s_1-1}b \cdots a^{s_d-1}b.$$

In the current q-analog setting, replacing the Riemann integrals by the Jackson q-integrals Eq. (12.7) one gets the following two generalizations after slight modification. First, for any $(a_1, \ldots, a_d), (b_1, \ldots, b_d) \in (\mathbb{Z}_{\geq 0})^d$, we define

$$\begin{aligned}\mathfrak{z}_q[\rho^{a_1}\pi^{b_1}y \ldots \rho^{a_d}\pi^{b_d}y; t] \\ = \mathbf{R}^{a_1}\big[\mathbf{P}^{b_1}[\mathbf{y}\mathbf{R}^{a_2}[\mathbf{P}^{b_2}[\mathbf{y}\cdots\mathbf{R}^{a_d}[\mathbf{P}^{b_d}[\mathbf{y}]]\cdots]]]\big](t).\end{aligned} \tag{12.31}$$

Theorem 12.2.12. *Let* $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ *and* $\mathbf{a} = (a_1, \ldots, a_d) \in (\mathbb{Z}_{\geq 0})^d$. *Put* $w = |\mathbf{s}|$ *and* $\mathbf{w} = \mathbf{w}^{\mathbf{a}}(\mathbf{s}) = \rho^{a_1}\pi^{s_1-a_1}y \ldots \rho^{a_d}\pi^{s_d-a_d}y$. *If* $a_1 + \cdots + a_j > 0$ *for all* $j = 1, \ldots, d$, *then we have*

$$\zeta_q^{\mathbf{a}}[\mathbf{s}] = (1-q)^w \mathfrak{z}_q[\mathbf{w}^{\mathbf{a}}(\mathbf{s}); 1], \quad \mathfrak{z}_q^{\mathbf{a}}[\mathbf{s}] = \mathfrak{z}_q[\mathbf{w}^{\mathbf{a}}(\mathbf{s})] := \mathfrak{z}_q[\mathbf{w}^{\mathbf{a}}(\mathbf{s}); 1]. \tag{12.32}$$

Proof. First we observe three important facts: for any $k \geq 1$ we have

$$\mathbf{P}(t^k) = \sum_{j\geq 0} q^{kj} t^k = \frac{t^k}{1-q^k}, \quad \mathbf{R}(t^k) = \sum_{j\geq 1} q^{kj} t^k = \frac{q^k t^k}{1-q^k}$$

and $\mathbf{D}(t^k) = t^k(1-q^k)$, by the definition of the two summation operators and the difference operator. Repeatedly applying this we get

$$\mathbf{P}^m(t^k) = \sum_{j\geq 0} q^{kj} t^k = \frac{t^k}{(1-q^k)^m}, \qquad \forall m \in \mathbb{Z}, \tag{12.33}$$

$$\mathbf{R}^m(t^k) = \sum_{j\geq 1} q^{kj} t^k = \frac{q^{mk} t^k}{(1-q^k)^m} \qquad \forall m \in \mathbb{Z}_{\geq 0}. \tag{12.34}$$

Thus

$$\mathbf{P}\big(\mathbf{y}(t)\cdot t^k\big) = \sum_{j\geq 0} \frac{q^{j(k+1)} t^{k+1}}{1-q^j t} = \sum_{j\geq 0}\sum_{\ell\geq 0} q^{j(k+\ell+1)} t^{k+\ell+1} = \sum_{\ell>k} \frac{t^\ell}{1-q^\ell}.$$

Similarly, we have

$$\mathbf{D}\big(\mathbf{y}(t)\cdot t^k\big) = \frac{t^{k+1}}{1-t} - \frac{q^{k+1}t^{k+1}}{1-qt} = \sum_{\ell\geq 0}(1-q^{k+\ell+1})t^{k+\ell+1} = \sum_{\ell>k}(1-q^\ell)t^\ell,$$

and

$$\mathbf{R}\big(\mathbf{y}(t)\cdot t^k\big) = \sum_{j\geq 1} \frac{q^{j(k+1)} t^{k+1}}{1-q^j t} = \sum_{j\geq 1}\sum_{\ell\geq 0} q^{j(k+\ell+1)} t^{k+\ell+1} = \sum_{\ell>k} \frac{q^\ell t^\ell}{1-q^\ell}.$$

It follows from Eq. (12.33) and Eq. (12.34) that

$$\mathbf{P}^m\big(\mathbf{y}(t)\cdot t^k\big) = \sum_{\ell>k} \frac{t^\ell}{(1-q^\ell)^m} \qquad \forall m \in \mathbb{Z}, \tag{12.35}$$

$$\mathbf{R}^m\big(\mathbf{y}(t)\cdot t^k\big) = \sum_{\ell>k} \frac{q^{m\ell} t^\ell}{(1-q^\ell)^m} \qquad \forall m \in \mathbb{Z}_{\geq 0}. \tag{12.36}$$

We now prove by induction on the the depth d that for all $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$,

$$\mathfrak{z}_q[\mathbf{w}^{\mathbf{a}}(\mathbf{s}); t] = \sum_{k_1>\cdots>k_d>0} \frac{t^{k_1} q^{k_1 a_1} \cdots q^{k_d a_d}}{(1-q^{k_1})^{s_1}\cdots(1-q^{k_d})^{s_d}}. \tag{12.37}$$

When $d = 1$, *i.e.*, $\mathbf{s} = s$, then by Eq. (12.35) followed by Eq. (12.34)

$$\mathfrak{z}_q[\mathbf{w}^a(s); t] = \mathbf{R}^a\mathbf{P}^{s-a}[\mathbf{y}](t) = \sum_{k>0} \frac{\mathbf{R}^a\big(t^k\big)}{(1-q^k)^{s-a}} = \sum_{k>0} \frac{q^{ak} t^k}{(1-q^k)^s} = \mathfrak{z}_q^{(a)}[s; t].$$

This proof works even when $s = a$ because of Eq. (12.36) (take $k = 0$ and $m = a$ there).

In general, assuming $d \geq 2$ and Eq. (12.37) is true for smaller depths. Then by the inductive assumption

$$\begin{aligned}
\mathfrak{z}_q[\mathbf{w}^{\mathbf{a}}(\mathbf{s}); t] &= \mathbf{R}^{a_1}\mathbf{P}^{s_1-a_1}\big[\mathbf{y}\mathbf{R}^{a_2}\mathbf{P}^{s_2-a_2}[\mathbf{y}\cdots\mathbf{R}^{a_d}\mathbf{P}^{s_d-a_d}[\mathbf{y}]\cdots]\big](t) \\
&= \sum_{k_2>\cdots>k_d>0} \frac{\mathbf{R}^{a_1}\mathbf{P}^{s_1-a_1}\big(\mathbf{y}(t)\cdot t^{k_2}\big)q^{k_2a_2}\dots q^{k_da_d}}{(1-q^{k_2})^{s_2}\cdots(1-q^{k_d})^{s_d}} \\
&\overset{(12.35)}{=\!=} \sum_{k_1>\cdots>k_d>0} \frac{\mathbf{R}^{a_1}\big(t^{k_1}\big)q^{k_2a_2}\dots q^{k_da_d}}{(1-q^{k_1})^{s_1-a_1}(1-q^{k_2})^{s_2}\cdots(1-q^{k_d})^{s_d}} \\
&\overset{(12.34)}{=\!=} \sum_{k_1>\cdots>k_d>0} \frac{t^{k_1}q^{k_1a_1}\dots q^{k_da_d}}{(1-q^{k_1})^{s_1}\cdots(1-q^{k_d})^{s_d}}.
\end{aligned}$$

Again, if $s_1 = a_1$ the proof is still valid. This completes the proof of Eq. (12.37). Setting $t = 1$ we arrive at Eq. (12.32). □

By change of variables $a_j \to s_j - a_j$ for all $j = 1, \dots, d$ we immediately obtain the next result.

Theorem 12.2.13. *For* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ *and* $\mathbf{a} = (a_1, \dots, a_d) \in (\mathbb{Z}_{\geq 0})^d$, *we set* $\mathbf{s} - \mathbf{a} = (s_1 - a_1, \dots, s_d - a_d)$, $w = |\mathbf{s}|$ *and* $\mathbf{w}^{\mathbf{s}-\mathbf{a}}(\mathbf{s}) = \rho^{s_1-a_1}\pi^{a_1}y\dots\rho^{s_d-a_d}\pi^{a_d}y$. *Define*

$$\mathfrak{z}_q[\mathbf{w}^{\mathbf{s}-\mathbf{a}}(\mathbf{s}); t] = \mathbf{R}^{s_1-a_1}\big[\mathbf{P}^{a_1}[\mathbf{y}\mathbf{R}^{s_2-a_2}[\mathbf{P}^{a_2}[\mathbf{y}\cdots\mathbf{R}^{s_d-a_d}[\mathbf{P}^{a_d}[\mathbf{y}]]\cdots]]]\big](t).$$

If $s_1 + \cdots + s_j > a_1 + \cdots + a_j$ *for all* $j = 1, \dots, d$, *then we have*

$$\begin{aligned}
\zeta_q^{\mathbf{s}-\mathbf{a}}[\mathbf{s}] &= (1-q)^w \mathfrak{z}_q[\mathbf{w}^{\mathbf{s}-\mathbf{a}}(\mathbf{s}); 1], \\
\mathfrak{z}_q^{\mathbf{s}-\mathbf{a}}[\mathbf{s}] &= \mathfrak{z}_q[\mathbf{w}^{\mathbf{s}-\mathbf{a}}(\mathbf{s})] := \mathfrak{z}_q[\mathbf{w}^{\mathbf{s}-\mathbf{a}}(\mathbf{s}); 1].
\end{aligned} \tag{12.38}$$

By specializing the proceeding two theorems we quickly find the following corollary. For future reference, we will say $\mathbf{w}_\tau$ has the *typical type* τ *form* for each type τ.

Corollary 12.2.14. *For* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$, *we set*

$$\begin{aligned}
&\mathbf{w}_{\mathrm{I}} = \rho^{s_1-1}\pi y\dots\rho^{s_d-1}\pi y = \mathbf{z}_{s_1}\dots\mathbf{z}_{s_d} \in \mathfrak{A}_\theta^{(B)} \subset \mathfrak{A}_\pi y \qquad (s_1 \geq 2), \\
&\mathbf{w}_{\mathrm{II}} = \rho^{s_1}y\dots\rho^{s_d}y = \mathbf{z}'_{s_1}\dots\mathbf{z}'_{s_d} \in \mathfrak{A}_\theta^{(B)} \subset \mathfrak{A}_\pi y, \\
&\mathbf{w}_{\mathrm{III}} = \pi^{s_1-1}\rho y\pi^{s_2}y\dots\pi^{s_d}y = \mathbf{z}'_{s_1}\mathbf{z}_{s_2}\dots\mathbf{z}_{s_d} \in \mathfrak{A}_\theta^{(A)} \subset \mathfrak{A}_\pi y,
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{z}_q[\mathbf{w}_{\mathrm{I}};t] &= \mathbf{R}^{s_1-1}\big[\mathbf{P}[\mathbf{y}\mathbf{R}^{s_2-1}[\mathbf{P}[\mathbf{y}\cdots\mathbf{R}^{s_d-1}[\mathbf{P}[\mathbf{y}]]\cdots]]]\big](t),\\
\mathfrak{z}_q[\mathbf{w}_{\mathrm{II}};t] &= \mathbf{R}^{s_1}\big[\mathbf{y}\mathbf{R}^{s_2}[\mathbf{y}\cdots\mathbf{R}^{s_d}[\mathbf{y}]\cdots]\big](t),\\
\mathfrak{z}_q[\mathbf{w}_{\mathrm{III}};t] &:= \mathbf{P}^{s_1-1}\Big[\mathbf{R}\big[\mathbf{y}[\mathbf{P}^{s_2}[\mathbf{y}[\mathbf{P}^{s_3}[\mathbf{y}\cdots\mathbf{P}^{s_d}[\mathbf{y}]\cdots]]]]\big]\Big](t).
\end{aligned}$$

Then for all the types $\tau = \mathrm{I}$, II and III, we have

$$\zeta_q^{\tau}[\mathbf{s}] = (1-q)^w \mathfrak{z}_q[\mathbf{w}_{\tau};\,1], \qquad \mathfrak{z}_q^{\tau}[\mathbf{s}] = \mathfrak{z}_q[\mathbf{w}_{\tau}] := \mathfrak{z}_q[\mathbf{w}_{\tau};\,1].$$

Moreover, similar results hold for type $\tilde{\mathrm{I}}$ q-MZVs. We may replace any of the consecutive strings $\rho^{s_j-1}\pi$ by a single ρ in $\mathbf{w}_{\tilde{\mathrm{I}}}$, and replace the corresponding operator string $\mathbf{P}^{s_j-1}\mathbf{R}$ by a single $\mathbf{R}$.

We now apply the above to Okounkov's q-MZVs. Let $\mathbf{s} \in (\mathbb{Z}_{\geq 2})^d$. Then using Eq. (12.4) we define

$$\begin{aligned}
\zeta_q^{\mathrm{O}}[\mathbf{s}] &:= \sum_{k_1>\cdots>k_d>0} \prod_{j=1}^{d} \frac{q^{k_j^+ s_j} + q^{k_j^- s_j}}{[k_j]^{s_j}}\\
&= (1-q)^{|\mathbf{s}|} \sum_{k_1>\cdots>k_d>0} \prod_{j=1}^{d} \frac{q^{k_j^+ s_j} + q^{k_j^- s_j}}{(1-q^{k_j})^{s_j}}.
\end{aligned}$$

Again, its modified form is:

$$\mathfrak{z}_q^{\mathrm{O}}[\mathbf{s}] := \sum_{k_1>\cdots>k_d>0} \prod_{j=1}^{d} \frac{q^{k_j^+ s_j} + q^{k_j^- s_j}}{(1-q^{k_j})^{s_j}}.$$

Remark 12.2.15. The above variation is equal to Okounkov's original q-MZVs up to a suitable 2-power. More precisely, the power is given by the number of even arguments in $\mathbf{s}$.

Corollary 12.2.16. *For $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$, we set*

$$\begin{aligned}
\mathbf{w}_{\mathrm{O}} &= (\rho^{s_1^-}\pi^{s_1^+} + \rho^{s_1^+}\pi^{s_1^-})y \ldots (\rho^{s_d^-}\pi^{s_d^+} + \rho^{s_d^+}\pi^{s_d^-})y\\
&= \mathbf{z}_{s_1}\ldots\mathbf{z}_{s_d} \in \mathfrak{A}_{\theta}^{(B)} \subset \mathfrak{A}_{\pi}y
\end{aligned}$$

and

$$\mathfrak{z}_q[\mathbf{w}_{\mathrm{O}};t] = (\mathbf{R}^{s_1^-}\mathbf{P}^{s_1^+} + \mathbf{R}^{s_1^+}\mathbf{P}^{s_1^-})\big[\mathbf{y}\cdots(\mathbf{R}^{s_d^-}\mathbf{P}^{s_d^+} + \mathbf{R}^{s_d^+}\mathbf{P}^{s_d^-})[\mathbf{y}]\cdots\big](t).$$

Then we have

$$\zeta_q^{\mathrm{O}}[\mathbf{s}] = (1-q)^w \mathfrak{z}_q[\mathbf{w}_{\mathrm{O}};\,1], \qquad \mathfrak{z}_q^{\mathrm{O}}[\mathbf{s}] = \mathfrak{z}_q[\mathbf{w}_{\mathrm{O}}] := \mathfrak{z}_q[\mathbf{w}_{\mathrm{O}};\,1].$$

It is possible to obtain the shuffle relations among $\mathfrak{z}_q^{\mathrm{O}}[\mathbf{s}]$-values using Cor. 12.2.16. The stuffle relations among $\mathfrak{z}_q^{\mathrm{O}}[\mathbf{s}]$ is mentioned implicitly in Okounkov's original paper. For our modified version, they can be derived from the following fact. Let $F_n^{\mathrm{O}}(t) = (t^{n^+} + t^{n^-})/(1-t)^n$ for all $n \geq 2$. Then for all $r, s \in \mathbb{Z}_{\geq 2}$, we have

$$F_r^{\mathrm{O}}(t) \cdot F_s^{\mathrm{O}}(t) = \begin{cases} 2F_{r+s}^{\mathrm{O}}(t), & \text{if } r \text{ or } s \text{ is even;} \\ 2F_{r+s}^{\mathrm{O}}(t) + \frac{1}{2}F_{r+s-2}^{\mathrm{O}}(t), & \text{if } r \text{ and } s \text{ are odd.} \end{cases}$$

For example,

$$\mathfrak{z}_q^{\mathrm{O}}[2,3]\mathfrak{z}_q^{\mathrm{O}}[2] = 2\mathfrak{z}_q^{\mathrm{O}}[2,2,3] + \mathfrak{z}_q^{\mathrm{O}}[2,3,2] + 2\mathfrak{z}_q^{\mathrm{O}}[4,3] + 2\mathfrak{z}_q^{\mathrm{O}}[2,5],$$
$$\mathfrak{z}_q^{\mathrm{O}}[2,3]\mathfrak{z}_q^{\mathrm{O}}[3] = 2\mathfrak{z}_q^{\mathrm{O}}[2,3,3] + \mathfrak{z}_q^{\mathrm{O}}[3,2,3] + 2\mathfrak{z}_q^{\mathrm{O}}[5,3] + 2\mathfrak{z}_q^{\mathrm{O}}[2,6] + \frac{1}{2}\mathfrak{z}_q^{\mathrm{O}}[2,4].$$

12.2.5 *q-shuffle relations*

In contrast to the MZV case, the q-shuffle product is much more difficult to define than the q-stuffle product. In this section we will use the Rota-Baxter algebra approach to define this for type $\tilde{\mathrm{I}}$, $\mathbb{I}$, and $\mathbb{III}$ q-MZVs.

The *q-shuffle product* on $\mathfrak{A}_\pi$ is defined recursively as follows: for any words $\mathbf{u}, \mathbf{v} \in X_\pi^*$ we define $\mathsf{e} \sqcup\!\sqcup \mathbf{u} = \mathbf{u} \sqcup\!\sqcup \mathsf{e} = \mathbf{u}$ and

$$(y\mathbf{u}) \sqcup\!\sqcup \mathbf{v} = \mathbf{u} \sqcup\!\sqcup (y\mathbf{v}) = y(\mathbf{u} \sqcup\!\sqcup \mathbf{v}), \tag{12.39}$$
$$\pi\mathbf{u} \sqcup\!\sqcup \pi\mathbf{v} = \pi(\mathbf{u} \sqcup\!\sqcup \pi\mathbf{v}) + \pi(\pi\mathbf{u} \sqcup\!\sqcup \mathbf{v}) - \pi(\mathbf{u} \sqcup\!\sqcup \mathbf{v}), \tag{12.40}$$
$$\delta\mathbf{u} \sqcup\!\sqcup \delta\mathbf{v} = \mathbf{u} \sqcup\!\sqcup \delta\mathbf{v} + \delta\mathbf{u} \sqcup\!\sqcup \mathbf{v} - \delta(\mathbf{u} \sqcup\!\sqcup \mathbf{v}), \tag{12.41}$$
$$\delta\mathbf{u} \sqcup\!\sqcup \pi\mathbf{v} = \pi\mathbf{v} \sqcup\!\sqcup \delta\mathbf{u} = \delta(\mathbf{u} \sqcup\!\sqcup \pi\mathbf{v}) + \delta\mathbf{u} \sqcup\!\sqcup \mathbf{v} - \mathbf{u} \sqcup\!\sqcup \mathbf{v} \tag{12.42}$$

for any words $\mathbf{u}, \mathbf{v} \in X_\pi^*$. The first equation reflects the fact that when $\mathbf{y}(t)$ is multiplied in front of either of the two factors in a product, it can be multiplied after taking the product. The other equations formalize Eq. (12.10),Eq. (12.9), Eq. (12.14), Eq. (12.15), and Eq. (12.16), respectively.

Corollary 12.2.17. *For any words $\mathbf{u}, \mathbf{v} \in X_\pi^*$, we have*

$$\rho\mathbf{u} \sqcup\!\sqcup \rho\mathbf{v} = \rho(\mathbf{u} \sqcup\!\sqcup \rho\mathbf{v}) + \rho(\rho\mathbf{u} \sqcup\!\sqcup \mathbf{v}) + \rho(\mathbf{u} \sqcup\!\sqcup \mathbf{v}), \tag{12.43}$$
$$\begin{aligned} \rho\mathbf{u} \sqcup\!\sqcup \pi\mathbf{v} &= \pi\mathbf{v} \sqcup\!\sqcup \rho\mathbf{u} \\ &= \rho(\rho\mathbf{u} \sqcup\!\sqcup \mathbf{v}) + \rho(\mathbf{u} \sqcup\!\sqcup \rho\mathbf{v}) + \rho\mathbf{u} \sqcup\!\sqcup \mathbf{v} + \rho(\mathbf{u} \sqcup\!\sqcup \mathbf{v}), \end{aligned} \tag{12.44}$$
$$\begin{aligned} \delta\mathbf{u} \sqcup\!\sqcup \rho\mathbf{v} &= \rho\mathbf{v} \sqcup\!\sqcup \delta\mathbf{u} \\ &= \delta(\mathbf{u} \sqcup\!\sqcup \pi\mathbf{v}) - \mathbf{u} \sqcup\!\sqcup \mathbf{v} = \delta(\mathbf{u} \sqcup\!\sqcup \rho\mathbf{v}) + \delta(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) - \mathbf{u} \sqcup\!\sqcup \mathbf{v}. \end{aligned} \tag{12.45}$$

Proof. These follows easily from Eq. (12.39)–Eq. (12.42) and the relation $\rho = \pi - \mathsf{e}$. □

Corollary 12.2.18. *For $j = 1, 2$ let $X_\theta^{(j)}$ and $X_\theta^{(j),*}$ be the embedding of X_θ and X_θ^* into X_π^*, respectively, by Definition* 12.2.10. *For any $\alpha, \beta \in X_\theta^{(j)}$ and $\mathbf{u}, \mathbf{v} \in X_\theta^{(j),*}$, we have $1 \sqcup\!\sqcup \mathbf{u} = \mathbf{u} \sqcup\!\sqcup 1 = \mathbf{u}$ and*

$$\alpha\mathbf{u} \sqcup\!\sqcup \beta\mathbf{v} = \alpha(\mathbf{u} \sqcup\!\sqcup \beta\mathbf{v}) + \beta(\alpha\mathbf{u} \sqcup\!\sqcup \mathbf{v}) + [\alpha, \beta]_j(\mathbf{u} \sqcup\!\sqcup \mathbf{v}),$$

where $[\alpha, \beta]_j$ is determined by $[a, b]_1 = [b, a]_1 = -b$, $[a, b]_2 = [b, a]_2 = 0$ and

$$[a, a]_j = (-1)^j a, \quad [b, b]_j = -b\gamma, \quad [\alpha, \gamma]_j = [\gamma, \alpha]_j = -\alpha\gamma. \tag{12.46}$$

Proof. All of these identities follow from straight-forward computation using Eq. (12.39)–Eq. (12.45). For example,

$$\begin{aligned} b\mathbf{u} \sqcup\!\sqcup b\mathbf{v} &= \pi y\mathbf{u} \sqcup\!\sqcup \pi y\mathbf{v} \\ &= \pi(y\mathbf{u} \sqcup\!\sqcup \pi y\mathbf{v}) + \pi(\pi y\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) - \pi(y\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) \\ &= \pi y(\mathbf{u} \sqcup\!\sqcup \pi y\mathbf{v}) + \pi y(\pi y\mathbf{u} \sqcup\!\sqcup \mathbf{v}) - \pi yy(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) \\ &= b(\mathbf{u} \sqcup\!\sqcup b\mathbf{v}) + b(b\mathbf{u} \sqcup\!\sqcup \mathbf{v}) - b\gamma(\mathbf{u} \sqcup\!\sqcup \mathbf{v}). \end{aligned} \tag{12.47}$$

Similarly,

$$\begin{aligned} \theta\mathbf{u} \sqcup\!\sqcup \theta\mathbf{v} &= \rho y\mathbf{u} \sqcup\!\sqcup \rho y\mathbf{v} \\ &= \rho(y\mathbf{u} \sqcup\!\sqcup \rho y\mathbf{v}) + \rho(\rho y\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) + \rho(y\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) \\ &= \rho y(\mathbf{u} \sqcup\!\sqcup \rho y\mathbf{v}) + \rho y(\rho y\mathbf{u} \sqcup\!\sqcup \mathbf{v}) + \rho yy(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) \\ &= \theta(\mathbf{u} \sqcup\!\sqcup \theta\mathbf{v}) + \theta(\theta\mathbf{u} \sqcup\!\sqcup \mathbf{v}) + \theta\gamma(\mathbf{u} \sqcup\!\sqcup \mathbf{v}). \end{aligned} \tag{12.48}$$

The rest of the proof is left to the interested reader. □

Proposition 12.2.19. *The algebra $(\mathfrak{A}_\pi, \sqcup\!\sqcup)$ is commutative and associative.*

Proof. Proving commutativity is routine and is thus left to the interested reader as an exercise. The associativity relation $(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) \sqcup\!\sqcup \mathbf{w} = \mathbf{u} \sqcup\!\sqcup (\mathbf{v} \sqcup\!\sqcup \mathbf{w})$ is much harder. We now prove it by an induction on $|\mathbf{u}| + |\mathbf{v}| + |\mathbf{w}|$.

Without loss of generality we write $\mathbf{u} = \alpha\mathbf{a}$, $\mathbf{v} = \beta\mathbf{b}$, and $\mathbf{w} = \gamma\mathbf{c}$ where $\alpha, \beta, \gamma \in \{\pi, \delta, y\}$ and $\mathbf{a}, \mathbf{b}, \mathbf{c} \in X_\pi^*$. The 27 different cases can all be reduced to the following 5 cases by commutativity:

(1) **One of the letters α, β or γ is a y.** By inductive assumption and Eq. (12.39):

$$\begin{aligned} (y\mathbf{a} \sqcup\!\sqcup \mathbf{v}) \sqcup\!\sqcup \mathbf{u} &= \big(y(\mathbf{a} \sqcup\!\sqcup \mathbf{v})\big) \sqcup\!\sqcup \mathbf{u} = y\big((\mathbf{a} \sqcup\!\sqcup \mathbf{v}) \sqcup\!\sqcup \mathbf{u}\big) \\ &= y\big(\mathbf{a} \sqcup\!\sqcup (\mathbf{v} \sqcup\!\sqcup \mathbf{u})\big) = y\mathbf{a} \sqcup\!\sqcup (\mathbf{v} \sqcup\!\sqcup \mathbf{u}). \end{aligned}$$

Hence associativity in this case is proved.

(2) **One π and two δ's.** By commutativity we may assume $\alpha = \beta = \delta$ and $\gamma = \pi$. By Eq. (12.41) and Eq. (12.42)

$$
\begin{aligned}
&(\delta\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b}) \sqcup\!\sqcup \pi\mathbf{c} \\
&= \big(\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} + \delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b})\big) \sqcup\!\sqcup \pi\mathbf{c} \\
&= \mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c} + \mathbf{b} \sqcup\!\sqcup \delta\mathbf{a} \sqcup\!\sqcup \pi\mathbf{c} - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \pi\mathbf{c} \\
&= \mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) + \mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c} - \mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} + \mathbf{b} \sqcup\!\sqcup \delta(\mathbf{a} \sqcup\!\sqcup \pi\mathbf{c}) \\
&+ \mathbf{b} \sqcup\!\sqcup \delta\mathbf{a} \sqcup\!\sqcup \mathbf{c} - \mathbf{b} \sqcup\!\sqcup \mathbf{a} \sqcup\!\sqcup \mathbf{c} - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \mathbf{c} + \mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} \\
&= \underbrace{\mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c})}_{1} + \underbrace{\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c}}_{2} - \underbrace{\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c}}_{3} + \underbrace{\mathbf{b} \sqcup\!\sqcup \delta(\mathbf{a} \sqcup\!\sqcup \pi\mathbf{c})}_{4} \\
&+ \underbrace{\mathbf{b} \sqcup\!\sqcup \delta\mathbf{a} \sqcup\!\sqcup \mathbf{c}}_{5} - \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \pi\mathbf{c})}_{6} - \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \mathbf{c}}_{7}.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\delta\mathbf{a} \sqcup\!\sqcup (\delta\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) \\
&= \delta\mathbf{a} \sqcup\!\sqcup \big(\delta(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) + \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c} - \mathbf{b} \sqcup\!\sqcup \mathbf{c}\big) \\
&= \delta\mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) + \delta\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c} - \delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} \\
&= \mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) + \delta\mathbf{a} \sqcup\!\sqcup \pi\mathbf{c} \sqcup\!\sqcup \mathbf{b} - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \pi\mathbf{c}) + \mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c} \\
&+ \delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \mathbf{c} - \delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} \\
&= \underbrace{\mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{c})}_{1} + \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \pi\mathbf{c}) \sqcup\!\sqcup \mathbf{b}}_{4} + \underbrace{\delta\mathbf{a} \sqcup\!\sqcup \mathbf{c} \sqcup\!\sqcup \mathbf{b}}_{5} - \underbrace{\mathbf{a} \sqcup\!\sqcup \mathbf{c} \sqcup\!\sqcup \mathbf{b}}_{3} \\
&- \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \pi\mathbf{c})}_{6} + \underbrace{\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c}}_{2} - \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \mathbf{c}}_{7}.
\end{aligned}
$$

Hence associativity in case (2) is proved.

(3) **Two π's and one δ.** By commutativity we may assume $\alpha = \beta = \pi$ and $\gamma = \delta$. By Eq. (12.40) and Eq. (12.42)

$$
\begin{aligned}
&(\pi\mathbf{a} \sqcup\!\sqcup \pi\mathbf{b}) \sqcup\!\sqcup \delta\mathbf{c} \\
&= \pi(\mathbf{a} \sqcup\!\sqcup \pi\mathbf{b}) \sqcup\!\sqcup \delta\mathbf{c} + \pi(\pi\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \delta\mathbf{c} - \pi(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \delta\mathbf{c} \\
&= \delta\big(\mathbf{c} \sqcup\!\sqcup \pi(\mathbf{a} \sqcup\!\sqcup \pi\mathbf{b})\big) + \delta\mathbf{c} \sqcup\!\sqcup \mathbf{a} \sqcup\!\sqcup \pi\mathbf{b} - \mathbf{a} \sqcup\!\sqcup \pi\mathbf{b} \sqcup\!\sqcup \mathbf{c} \\
&+ \delta\big(\mathbf{c} \sqcup\!\sqcup \pi(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{a})\big) + \delta\mathbf{c} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \pi\mathbf{a} - \mathbf{b} \sqcup\!\sqcup \pi\mathbf{a} \sqcup\!\sqcup \mathbf{c} \\
&- \delta\big((\mathbf{c} \sqcup\!\sqcup \pi(\mathbf{a} \sqcup\!\sqcup \mathbf{b})\big) - \delta\mathbf{c} \sqcup\!\sqcup \mathbf{a} \sqcup\!\sqcup \mathbf{b} + \mathbf{c} \sqcup\!\sqcup \mathbf{a} \sqcup\!\sqcup \mathbf{b} \\
&= \delta\big(\mathbf{c} \sqcup\!\sqcup \pi(\mathbf{a} \sqcup\!\sqcup \pi\mathbf{b})\big) + \mathbf{a} \sqcup\!\sqcup \delta(\mathbf{c} \sqcup\!\sqcup \pi\mathbf{b}) + \mathbf{a} \sqcup\!\sqcup \delta\mathbf{c} \sqcup\!\sqcup \mathbf{b} - \mathbf{a} \sqcup\!\sqcup \mathbf{c} \sqcup\!\sqcup \mathbf{b} \\
&- \mathbf{a} \sqcup\!\sqcup \pi\mathbf{b} \sqcup\!\sqcup \mathbf{c} + \delta\big(\mathbf{c} \sqcup\!\sqcup \pi(\mathbf{b} \sqcup\!\sqcup \pi\mathbf{a})\big) + \mathbf{b} \sqcup\!\sqcup \delta(\mathbf{c} \sqcup\!\sqcup \pi\mathbf{a}) + \mathbf{b} \sqcup\!\sqcup \delta\mathbf{c} \sqcup\!\sqcup \mathbf{a} \\
&- \mathbf{b} \sqcup\!\sqcup \mathbf{c} \sqcup\!\sqcup \mathbf{a} - \mathbf{b} \sqcup\!\sqcup \pi\mathbf{a} \sqcup\!\sqcup \mathbf{c} - \delta\big((\mathbf{c} \sqcup\!\sqcup \pi(\mathbf{a} \sqcup\!\sqcup \mathbf{b})\big) - \delta\mathbf{c} \sqcup\!\sqcup \mathbf{a} \sqcup\!\sqcup \mathbf{b} + \mathbf{c} \sqcup\!\sqcup \mathbf{a} \sqcup\!\sqcup \mathbf{b}.
\end{aligned}
$$

$$
\begin{aligned}
&= \underbrace{\delta\big(\mathbf{c} \sqcup \pi(\mathbf{a} \sqcup \pi\mathbf{b})\big)}_{1} + \underbrace{\mathbf{a} \sqcup \delta(\mathbf{c} \sqcup \pi\mathbf{b})}_{2} - \underbrace{\mathbf{a} \sqcup \mathbf{c} \sqcup \mathbf{b}}_{3} \\
&- \underbrace{\mathbf{a} \sqcup \pi\mathbf{b} \sqcup \mathbf{c}}_{4} + \underbrace{\delta\big(\mathbf{c} \sqcup \pi(\mathbf{b} \sqcup \pi\mathbf{a})\big)}_{5} + \underbrace{\mathbf{b} \sqcup \delta(\mathbf{c} \sqcup \pi\mathbf{a})}_{6} \\
&+ \underbrace{\mathbf{b} \sqcup \delta\mathbf{c} \sqcup \mathbf{a}}_{7} - \underbrace{\mathbf{b} \sqcup \pi\mathbf{a} \sqcup \mathbf{c}}_{8} - \underbrace{\delta\big(\mathbf{c} \sqcup \pi(\mathbf{a} \sqcup \mathbf{b})\big)}_{9}.
\end{aligned}
$$

On the other hand:

$$
\begin{aligned}
&\pi\mathbf{a} \sqcup (\pi\mathbf{b} \sqcup \delta\mathbf{c}) \\
&= \pi\mathbf{a} \sqcup \big(\delta(\mathbf{c} \sqcup \pi\mathbf{b}) + \delta\mathbf{c} \sqcup \mathbf{b} - \mathbf{c} \sqcup \mathbf{b}\big) \\
&= \delta(\mathbf{c} \sqcup \pi\mathbf{b} \sqcup \pi\mathbf{a}) + \delta(\mathbf{c} \sqcup \pi\mathbf{b}) \sqcup \mathbf{a} - \mathbf{c} \sqcup \pi\mathbf{b} \sqcup \mathbf{a} \\
&+ \delta(\mathbf{c} \sqcup \pi\mathbf{a}) \sqcup \mathbf{b} + \delta\mathbf{c} \sqcup \mathbf{a} \sqcup \mathbf{b} - \mathbf{a} \sqcup \mathbf{c} \sqcup \mathbf{b} - \pi\mathbf{a} \sqcup \mathbf{c} \sqcup \mathbf{b} \\
&= \underbrace{\delta\big((\mathbf{c} \sqcup \pi(\mathbf{b} \sqcup \pi\mathbf{a})\big)}_{5} + \underbrace{\delta\big(\mathbf{c} \sqcup \pi(\pi\mathbf{b} \sqcup \mathbf{a})\big)}_{1} - \underbrace{\delta\big(\mathbf{c} \sqcup \pi(\mathbf{b} \sqcup \mathbf{a})\big)}_{9} \\
&+ \underbrace{\delta(\mathbf{c} \sqcup \pi\mathbf{b}) \sqcup \mathbf{a}}_{2} - \underbrace{\mathbf{c} \sqcup \pi\mathbf{b} \sqcup \mathbf{a}}_{4} + \underbrace{\delta(\mathbf{c} \sqcup \pi\mathbf{a}) \sqcup \mathbf{b}}_{6} \\
&+ \underbrace{\delta\mathbf{c} \sqcup \mathbf{a} \sqcup \mathbf{b}}_{7} - \underbrace{\mathbf{a} \sqcup \mathbf{c} \sqcup \mathbf{b}}_{3} - \underbrace{\pi\mathbf{a} \sqcup \mathbf{c} \sqcup \mathbf{b}}_{8}.
\end{aligned}
$$

Hence associativity in case (3) is proved.

(4) $\alpha = \beta = \gamma = \delta$. By inductive assumption, commutativity, and Eq. (12.41)

$$
\begin{aligned}
&(\delta\mathbf{a} \sqcup \delta\mathbf{b}) \sqcup \delta\mathbf{c} \\
&= \big(\mathbf{a} \sqcup \delta\mathbf{b} + \delta\mathbf{a} \sqcup \mathbf{b} - \delta(\mathbf{a} \sqcup \mathbf{b})\big) \sqcup \delta\mathbf{c} \\
&= \mathbf{a} \sqcup \delta\mathbf{b} \sqcup \delta\mathbf{c} + \mathbf{b} \sqcup \delta\mathbf{a} \sqcup \delta\mathbf{c} - \delta(\mathbf{a} \sqcup \mathbf{b}) \sqcup \delta\mathbf{c} \\
&= \mathbf{a} \sqcup \big(\mathbf{b} \sqcup \delta\mathbf{c} + \delta\mathbf{b} \sqcup \mathbf{c} - \delta(\mathbf{b} \sqcup \mathbf{c})\big) - \mathbf{a} \sqcup \mathbf{b} \sqcup \delta\mathbf{c} \\
&+ \mathbf{b} \sqcup \big(\mathbf{a} \sqcup \delta\mathbf{c} + \delta\mathbf{a} \sqcup \mathbf{c} - \delta(\mathbf{a} \sqcup \mathbf{c})\big) - \delta(\mathbf{a} \sqcup \mathbf{b}) \sqcup \mathbf{c} + \delta(\mathbf{a} \sqcup \mathbf{b} \sqcup \mathbf{c}) \\
&= \underbrace{\mathbf{a} \sqcup \delta\mathbf{b} \sqcup \mathbf{c}}_{1} - \underbrace{\mathbf{a} \sqcup \delta(\mathbf{b} \sqcup \mathbf{c})}_{2} + \underbrace{\mathbf{b} \sqcup \mathbf{a} \sqcup \delta\mathbf{c}}_{3} + \underbrace{\mathbf{b} \sqcup \delta\mathbf{a} \sqcup \mathbf{c}}_{4} \\
&- \underbrace{\mathbf{b} \sqcup \delta(\mathbf{a} \sqcup \mathbf{c})}_{5} - \underbrace{\delta(\mathbf{a} \sqcup \mathbf{b}) \sqcup \mathbf{c}}_{6} + \underbrace{\delta(\mathbf{a} \sqcup \mathbf{b} \sqcup \mathbf{c})}_{7}.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\delta\mathbf{a} \sqcup (\delta\mathbf{b} \sqcup \delta\mathbf{c}) \\
&= \delta\mathbf{a} \sqcup \big(\mathbf{b} \sqcup \delta\mathbf{c} + \delta\mathbf{b} \sqcup \mathbf{c} - \delta(\mathbf{b} \sqcup \mathbf{c})\big)
\end{aligned}
$$

$$
\begin{aligned}
&= \delta\mathbf{a} \sqcup\!\sqcup \delta\mathbf{c} \sqcup\!\sqcup \mathbf{b} + \delta\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c} - \delta\mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \mathbf{c}) \\
&= \mathbf{a} \sqcup\!\sqcup \delta\mathbf{c} \sqcup\!\sqcup \mathbf{b} + \delta\mathbf{a} \sqcup\!\sqcup \mathbf{c} \sqcup\!\sqcup \mathbf{b} - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{c}) \sqcup\!\sqcup \mathbf{b} + \mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c} + \delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} \\
&\quad - \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \mathbf{c} - \mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \mathbf{c}) - \delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c} + \delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c}). \\
&= \underbrace{\mathbf{a} \sqcup\!\sqcup \delta\mathbf{c} \sqcup\!\sqcup \mathbf{b}}_{3} - \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{c}) \sqcup\!\sqcup \mathbf{b}}_{5} + \underbrace{\mathbf{a} \sqcup\!\sqcup \delta\mathbf{b} \sqcup\!\sqcup \mathbf{c}}_{1} + \underbrace{\delta\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c}}_{4} \\
&\quad - \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b}) \sqcup\!\sqcup \mathbf{c}}_{6} - \underbrace{\mathbf{a} \sqcup\!\sqcup \delta(\mathbf{b} \sqcup\!\sqcup \mathbf{c})}_{2} + \underbrace{\delta(\mathbf{a} \sqcup\!\sqcup \mathbf{b} \sqcup\!\sqcup \mathbf{c})}_{7}.
\end{aligned}
$$

Hence associativity in case (4) is proved.

(5) $\alpha = \beta = \gamma = \pi$. This is the ordinary stuffle case. We leave it as an exercise.

This concludes the proof of the proposition. □

Corollary 12.2.20. *For $j = 1$ or 2 the algebras $(\mathfrak{A}_\theta^{(j)}, \sqcup\!\sqcup)$ are commutative and associative.*

Proof. This follows immediately from Prop. 12.2.19 since $(\mathfrak{A}_\theta^{(j)}, \sqcup\!\sqcup)$ are subalgebras of $(\mathfrak{A}_\pi, \sqcup\!\sqcup)$ if $\sqcup\!\sqcup$ for $\mathfrak{A}_\theta^{(j)}$ is defined as in Cor. 12.2.18. □

Our next theorem shows that we may use the shuffle algebra structure defined above to describe the shuffle relations among different types of q-MZVs. Before doing so, we need to show that for each type the shuffle product really makes sense.

Proposition 12.2.21. *Embed $\mathfrak{A}_{\tilde{\mathrm{I}}}^0, \mathfrak{A}_{\mathrm{II}}^0 \subset \mathfrak{A}_\theta^{(B)}$ and $\mathfrak{A}_{\mathrm{III}}^0 \subset \mathfrak{A}_\theta^{(A)}$. Then for each type τ, if the two words $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_\tau^0$ have the typical type τ form listed in Cor.* 12.2.14 *then there is an algorithm to express $\mathbf{u} \sqcup\!\sqcup \mathbf{v}$ using only those words in the same form.*

Proof. For type $\tilde{\mathrm{I}}$ we prove in general that for $k, l \geq 1$ and $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_{\tilde{\mathrm{I}}}^1$

$$
\mathbf{z}_k\mathbf{u} \sqcup\!\sqcup \mathbf{z}_l\mathbf{v} \in \sum_{j \geq \min(k,l)} \mathbf{z}_j \mathfrak{A}_\theta^1. \tag{12.49}
$$

If $k = l = 1$ Eq. (12.49) follows from $b = \mathbf{z}_1$ and $\alpha(b, b) = -b\gamma$. If $k = 1$ and $l \geq 2$, we can prove Eq. (12.49) by induction on l using

$$
b\mathbf{u} \sqcup\!\sqcup \mathbf{z}_l\mathbf{v} = b(\mathbf{u} \sqcup\!\sqcup \mathbf{z}_l\mathbf{v}) + a(b\mathbf{u} \sqcup\!\sqcup \mathbf{z}_{l-1}\mathbf{v})
$$

since $a\mathbf{z}_j = \mathbf{z}_{j+1}$ for all $j \geq 1$. Now we assume $k, l \geq 2$. Then

$$
\mathbf{z}_k\mathbf{u} \sqcup\!\sqcup \mathbf{z}_l\mathbf{v} = a(\mathbf{z}_{k-1}\mathbf{u} \sqcup\!\sqcup \mathbf{z}_l\mathbf{v} + \mathbf{z}_k\mathbf{u} \sqcup\!\sqcup \mathbf{z}_{l-1}\mathbf{v} + \mathbf{z}_{k-1}\mathbf{u} \sqcup\!\sqcup \mathbf{z}_{l-1}\mathbf{v}) \in \sum_{j \geq \min(k,l)} a\mathbf{z}_{j-1} \mathfrak{A}_{\tilde{\mathrm{I}}}^1
$$

by induction assumption. Thus Eq. (12.49) holds. Finally we prove that

$$\theta\mathbf{u} \sqcup\!\sqcup \mathbf{z}_k\mathbf{v}, \theta\mathbf{u} \sqcup\!\sqcup \theta\mathbf{v} \in \mathfrak{A}^0_{\tilde{\mathrm{I}}}$$

for all $\mathbf{u}, \mathbf{v} \in \mathfrak{A}^1_{\tilde{\mathrm{I}}}$ and $k \geq 2$. This follows easily from Eq. (12.49) and

$$\begin{aligned} \theta w \sqcup\!\sqcup \mathbf{z}_k\mathbf{v} &= \theta(\mathbf{u} \sqcup\!\sqcup \mathbf{z}_k\mathbf{v}) + a(bw \sqcup\!\sqcup \mathbf{z}_{k-1}\mathbf{v}), \\ \theta w \sqcup\!\sqcup \theta\mathbf{v} &= \theta\big(\mathbf{u} \sqcup\!\sqcup \theta\mathbf{v} + \theta\mathbf{u} \sqcup\!\sqcup \mathbf{v} - \gamma(\mathbf{u} \sqcup\!\sqcup \mathbf{v})\big). \end{aligned} \tag{12.50}$$

Type $\mathbb{II}$ is in fact the easiest since we can restrict ourselves to use only Eq. (12.39) and Eq. (12.43) to compute the shuffle and therefore π never comes into the picture. Clearly all such words must start with ρ and end with y.

For type $\mathbb{III}$ let's assume $\mathbf{u} = \pi^{s_1-1}\rho y\pi^{s_2}y \dots \pi^{s_d}y$ and $\mathbf{v} = \pi^{t_1-1}\rho y\pi^{t_2}y \dots \pi^{t_d}y$. If we use the definition Eq. (12.40) repeatedly then in each word appearing in $\mathbf{u} \sqcup\!\sqcup \mathbf{v}$ the first ρ always appears before all the y's. Such a word can be written in the form $\pi^s\rho^r y \cdots$ for some $s \in \mathbb{Z}$ and $r \geq 1$ (notice that if ρ and π are commutative). Now we can rewrite this as $\pi^s(\pi - \mathsf{e})^{r-1}\rho y \cdots$ and replace all the ρ's after the first y by $\pi - \mathsf{e}$. This produces a word of typical type $\mathbb{III}$ form.

We have finished the proof of the proposition. □

Theorem 12.2.22. *Embed* $\mathfrak{A}^0_{\mathrm{I}}, \mathfrak{A}^0_{\mathbb{II}} \subset \mathfrak{A}^{(B)}_\theta$ *and* $\mathfrak{A}^0_{\mathbb{III}} \subset \mathfrak{A}^{(A)}_\theta$. *Then for each type* τ *and for any* $\mathbf{u}_\tau, \mathbf{v}_\tau \in \mathfrak{A}^0_\tau$, *we have*

$$\mathfrak{z}_q[\mathbf{u}_\tau]\mathfrak{z}_q[\mathbf{v}_\tau] = \mathfrak{z}_q[\mathbf{u}_\tau \sqcup\!\sqcup \mathbf{v}_\tau]. \tag{12.51}$$

Proof. For each type τ we observe that $\mathfrak{z}_q[\mathbf{w}_\tau; t]$ satisfies Eq. (12.51) because of the identities in Prop. 12.2.2. Then the theorem follows from the fact that $\mathfrak{z}_q[\mathbf{w}_\tau] = \mathfrak{z}_q[\mathbf{w}_\tau; 1]$ for any word $\mathbf{w}_\tau \in \mathfrak{A}^0_\tau$ by Cor. 12.2.14. □

12.3 Duality Relations

The DBSFs do not contain all linear relations among the various types of q-MZVs. Some of the missing ones can be proved by the following "duality" relations because of their similarity to the duality relations for the ordinary MZVs.

Theorem 12.3.1. *For a positive integer* k, *set*

$$\varphi_k := (-1)^k \left(\sum_{j=2}^{k} (-1)^j \mathbf{z}_j - \theta \right),$$

where $\varphi_1 = \theta = \rho y \in \mathfrak{A}^{(B)}_\theta$. *Let* $\ell \in \mathbb{N}$ *and* $\alpha_j, \beta_j \in \mathbb{Z}_{\geq 0}$ *for all* $j = 1, \dots, \ell$. *Then we have*

$$\zeta^{\tilde{\mathrm{I}}}_q[\varphi_{\alpha_1+1}\gamma^{\beta_1} \cdots \varphi_{\alpha_\ell+1}\gamma^{\beta_\ell}] = \zeta^{\tilde{\mathrm{I}}}_q[\varphi_{\beta_\ell+1}\gamma^{\alpha_\ell} \cdots \varphi_{\beta_1+1}\gamma^{\alpha_1}]. \tag{12.52}$$

We can use the Rota-Baxter algebra approach to give a new proof of this result.

Proof. Notice that $\gamma = y$, $\mathtt{z}_j = \rho^{j-1}\pi y$ and $\theta = \rho y$ with the embedding $\mathfrak{A}_{\mathrm{I}}^0 \subset \mathfrak{A}_\theta^{(B)}$. Since $\pi = \rho + \mathtt{e}$, for all $k \geq 1$, we have

$$\begin{aligned}\varphi_k =&(-1)^k \left(\sum_{j=2}^{k} (-1)^j \rho^{j-1}(\rho + \mathtt{e})y - \rho y \right) \\ =&(-1)^k \left(\sum_{j=2}^{k} (-1)^j \rho^j y + \sum_{j=2}^{k} (-1)^j \rho^{j-1} y - \rho y \right) = \rho^k y.\end{aligned} \tag{12.53}$$

Thus the theorem follows from Cor. 12.2.4 and Cor. 12.2.14 easily. □

Notice that in the expression of φ_k the letter θ appears. However, q-MZVs of the form such as $\zeta_q^{\mathrm{I}}[\theta\gamma\mathtt{z}_2\gamma] = \zeta_q^{\mathrm{I}}[\rho y^2 \rho^2 \pi y]$ is not really defined. In fact, it should be denoted by $\zeta_q^{\tilde{\mathrm{I}}}[\theta\gamma\mathtt{z}_2\gamma] = \zeta_q^{(1,0,1,0)}[1,0,2,0]$ (and such values always converge by Prop. 11.1.1 because of the leading 1 in the auxiliary variable $\mathbf{t}$). But, suitable $\mathbb{Q}$-linear combinations of Eq. (12.52) may lead to identities in which only $\mathtt{z}_k$'s appear. Then all terms can be written as honest ζ_q^{I}-values. This explains the use of two admissible structures $\mathfrak{A}_I^0$ and $\mathfrak{A}_{\mathrm{I}}^0$.

Similar relations for type $\mathbb{II}$ q-MZVs have the most aesthetic appeal and is the primary reason why we prefer to call it by "duality".

Theorem 12.3.2. *Let $\ell \in \mathbb{N}$ and $\alpha_j, \beta_j \in \mathbb{N}$ for all $j = 1, \ldots, \ell$. Then we have*

$$\zeta_q^{\mathbb{II}}[\rho^{\alpha_1} y^{\beta_1} \cdots \rho^{\alpha_\ell} y^{\beta_\ell}] = \zeta_q^{\mathbb{II}}[\rho^{\beta_\ell} y^{\alpha_\ell} \cdots \rho^{\beta_1} y^{\alpha_1}].$$

Proof. This follows from Cor. 12.2.4 and Cor. 12.2.14 immediately. □

Of course we may apply the same idea to type $\mathbb{III}$ q-MZVs.

Theorem 12.3.3. *Let $\ell \in \mathbb{N}$ and $\alpha_j, \beta_j \in \mathbb{N}$ for all $j = 1, \ldots, \ell$. Then we have*

$$\begin{aligned}&\zeta_q^{\mathbb{III}}[(\pi - \mathtt{e})^{\alpha_1 - 1} \rho y^{\beta_1} (\pi - \mathtt{e})^{\alpha_2} y^{\beta_2} \cdots (\pi - \mathtt{e})^{\alpha_\ell} y^{\beta_\ell}] \\ &\qquad = \zeta_q^{\mathbb{III}}[(\pi - \mathtt{e})^{\beta_\ell - 1} \rho y^{\alpha_\ell} (\pi - \mathtt{e})^{\beta_{\ell-1}} y^{\alpha_{\ell-1}} \cdots (\pi - \mathtt{e})^{\beta_1} y^{\alpha_1}].\end{aligned}$$

Proof. Since $\rho = \pi - \mathtt{e}$ this follows from Cor. 12.2.4 and Cor. 12.2.14 easily. □

12.4 The P-R and R-P relations

We observe that there is often more than one way to express a type a q-MZV using words because of the relation $\pi = \rho + \mathsf{e}$. For example, replacing π by $\rho + \mathsf{e}$ in the identity

$$\pi^2\rho y = \pi\rho^2 y + \pi\rho y = \rho^3 y + 2\rho^2 y + \rho y$$

we get immediately the relations

$$\mathfrak{z}_q^{(1)}[3] = \mathfrak{z}_q^{(2)}[3] + \mathfrak{z}_q^{(1)}[2] = \mathfrak{z}_q^{(3)}[3] + 2\mathfrak{z}_q^{(2)}[2] + \mathfrak{z}_q^{(1)}[1].$$

We call all such relations **P-R** relations. On the level of q-MZVs this is a consequence of the following identity

$$\frac{q^{nt}}{(1-q^n)^s} = \frac{q^{nt} - q^{(n+1)t} + q^{(n+1)t}}{(1-q^n)^s} = \frac{q^{nt}}{(1-q^n)^{s-1}} + \frac{q^{(n+1)t}}{(1-q^n)^s}.$$

Now we consider the general type G q-MZVs $\zeta_q^{(t_1,\ldots,t_d)}[s_1,\ldots,s_d]$ defined by Eq. (11.1) where $1 \le t_1 \le s_1$, $0 \le t_j \le s_j$ for all $j \ge 2$. Repeatedly applying the **P-R** relations we can see these values can be reduced to type $\mathbb{I}$ q-MZVs since in the word expression we only need to use ρ and y.

On the other hand, if we use ρ by $\pi - \mathsf{e}$ repeatedly we may convert all type G q-MZVs $\zeta_q^{(t_1,\ldots,t_d)}[s_1,\ldots,s_d]$ to type $\mathbb{III}$ values. We call these relations **R-P** relations. For example, from

$$\rho^3 y = (\pi^2 - 2\pi + \mathsf{e})\rho y$$

we get

$$\zeta_q^{\mathbb{II}}[3] = \zeta_q^{\mathbb{III}}[3] - 2\zeta_q^{\mathbb{III}}[2] + \zeta_q^{\mathbb{III}}[1].$$

12.5 The general type G q-MZVs

All of the q-MZVs of type $\tilde{\mathrm{I}}$, $\mathbb{I}$ and $\mathbb{III}$ considered earlier in the chapter are some special forms of the q-MZVs $\zeta_q^{(t_1,\ldots,t_d)}[s_1,\ldots,s_d]$ defined by Eq. (11.1) where $1 \le t_1 \le s_1$, $0 \le t_j \le s_j$ for all $j \ge 2$, all of which are convergent by Prop. 11.1.1. Let $\mathfrak{A}_{\mathrm{G}}^0$ be the set of all such values which we call *(admissible) type G* q-MZVs. Similar to the above three types, we may use words to encode these values according to Thm. 12.2.12 by setting $a_j = t_j$ there. Namely, setting

$$\mathbf{w}^{\mathbf{t}}(\mathbf{s}) = \rho^{t_1}\pi^{s_1-t_1}y\cdots\rho^{t_d}\pi^{s_d-t_d}y \in X_\pi^*,$$

we have

$$\mathfrak{z}_q^{(t_1,\ldots,t_d)}[s_1,\ldots,s_d] = \mathfrak{z}_q[\mathbf{w}^{\mathbf{t}}(\mathbf{s})]$$

where the right-hand side is defined by Eq. (12.31). The shuffle product structure are reflected by $(X_\pi^*, \sqcup\!\sqcup)$ where the $\sqcup\!\sqcup$ is defined by Eq. (12.39), Eq. (12.40), Eq. (12.43) and Eq. (12.44).

Proposition 12.5.1. *For all* $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_{\mathrm{G}}^0$, *we have* $\mathbf{u} \sqcup\!\sqcup \mathbf{v} \in \mathfrak{A}_{\mathrm{G}}^0$.

Proof. Notice that admissible words in $\mathfrak{A}_{\mathrm{G}}^0$ must end with y and have at least one ρ before the first y. Moreover, the converse is also true. This is rather straight-forward if we use the **P-R** relations repeatedly to get rid of all the π's.

Now, by using the definition of $\sqcup\!\sqcup$ it is not hard to see that $\mathbf{u} \sqcup\!\sqcup \mathbf{v}$ ends with y and and has at least one ρ before the first y if both $\mathbf{u}$ and $\mathbf{v}$ are admissible. So $\mathbf{u} \sqcup\!\sqcup \mathbf{v} \in \mathfrak{A}_{\mathrm{G}}^0$ and the proposition is proved. □

To define the stuffle product we let

$$\mathsf{Y}_{\mathrm{G}} = \{\mathbf{z}_{t,s} \mid t, s \in \mathbb{Z}_{\geq 0}, t \leq s\},$$

and let $\mathfrak{A}_{\mathrm{G}}$ be the the noncommutative polynomial $\mathbb{Q}$-algebra of words of $\mathsf{Y}_{\mathrm{G}}^*$ built on the alphabet Y_{G}. Define the type G-admissible words as those in

$$\mathfrak{A}_{\mathrm{G}}^0 = \bigcup_{1 \leq t \leq s} \mathbf{z}_{t,s} \mathfrak{A}_{\mathrm{G}}.$$

We can regard $\mathfrak{A}_{\mathrm{G}}$ as a subalgebra of X_π^* by setting $\mathbf{z}_{t,s} = \rho^t \pi^{s-t} y$. Then stuffle product $*_{\mathrm{G}}$ on $\mathfrak{A}_{\mathrm{G}}^0$ can be defined inductively as follows. For any words $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_{\mathrm{G}}^0$ and letters $\mathbf{z}_{t,s}, \mathbf{z}_{t',s'} \in \mathsf{Y}_{\mathrm{G}}$ with $1 \leq t \leq s$ and $1 \leq t' \leq s'$ we set $\mathsf{e} *_{\mathrm{G}} \mathbf{u} = \mathbf{u} = \mathbf{u} *_{\mathrm{G}} \mathsf{e}$ and

$$(\mathbf{z}_{t,s}\mathbf{u}) *_{\mathrm{G}} (\mathbf{z}_{t',s'}\mathbf{v}) = \mathbf{z}_{t,s}(\mathbf{u} *_{\mathrm{G}} \mathbf{z}_{t',s'}\mathbf{v}) + \mathbf{z}_{t',s'}(\mathbf{z}_{t,s}\mathbf{u} *_{\mathrm{G}} \mathbf{v}) + \mathbf{z}_{t+t',s+s'}(\mathbf{u} *_{\mathrm{G}} \mathbf{v}).$$

It is easy to show that $(\mathfrak{A}_{\mathrm{G}}^0, *_{\mathrm{G}})$ is a commutative and associative algebra.

We leave the proof of the following theorems to the interested reader. The first result clearly provides the DBSFs of type G q-MZVs.

Theorem 12.5.2. *For any* $\mathbf{u}, \mathbf{v} \in \mathfrak{A}_{\mathrm{G}}^0 \subset X_\pi^*$ *we have*

$$\mathfrak{z}_q[\mathbf{u} *_{\mathrm{G}} \mathbf{v}] = \mathfrak{z}_q[\mathbf{u} \sqcup\!\sqcup \mathbf{v}] = \mathfrak{z}_q[\mathbf{u}]\mathfrak{z}_q[\mathbf{v}]. \tag{12.54}$$

The duality relations are given in the cleanest form by Thm. 12.3.3 which can be translated into the following.

Theorem 12.5.3. *Let* $\ell \in \mathbb{N}$ *and* $\alpha_j, \beta_j \in \mathbb{N}$ *for all* $j = 1, \ldots, \ell$. *Set*

$$\mathbf{s} = (\alpha_1, 0^{\beta_1 - 1}, \alpha_2, 0^{\beta_2 - 1}, \ldots, \alpha_\ell, 0^{\beta_\ell - 1}),$$
$$\mathbf{s}^\vee = (\beta_\ell, 0^{\alpha_\ell - 1}, \beta_{\ell-1}, 0^{\alpha_{\ell-1} - 1}, \ldots, \beta_1, 0^{\alpha_1 - 1}).$$

Then we have

$$\zeta_q^{\mathbf{s}}[\mathbf{s}] = \zeta_q^{\mathbf{s}^\vee}[\mathbf{s}^\vee].$$

By numerical computation with the help of some computer algebra system one can show that for all of the type $\tilde{\mathrm{I}}$, II and III q-MZVs duality relations are necessary to generate some $\mathbb{Q}$-linear relations among q-MZVs that are missed by the DBSFs, at least when the weight is large enough. However, the combination of all the DBSFs and dualities are often not exhaustive yet. Sometimes, this difficulty can be overcome by increasing the weight and depth. But this seems to fail in some other cases, for example, for type $\tilde{\mathrm{I}}$ q-MZVs of weight 4.

We can improve the above situation by considering the more general type G values. The advantage is that we have the new **P-R** relations which provide a lot of new relations between type G q-MZVs, much more than the DBSFs and duality combined. The disadvantage is that there are too many type G values so that even when the weight is 5 the current computer power is too week to produce all the necessary relations. However, by using **P-R** relations all type G values can be converted to $\mathbb{Q}$-linear combinations of type II values which can be handled by computers a lot easier.

12.6 Application to Okounkov's conjecture

Using Cor. 12.2.16 we may regard Okounkov's q-MZVs as $\mathbb{Q}$-linear combinations of the q-MZVs $\mathfrak{z}_q^{\mathbf{t}}[\mathbf{s}]$ for suitable auxiliary variable $\mathbf{t}$. Further by using the **P-R** relations we may further reduce this to type II q-MZVs where we don't need the letter π.

w	2	3	4	5	6	7	8	9	10	11	12
$\sharp(\mathsf{W})^{\mathrm{O}}_{\leq w}$	1	2	4	7	12	20	33	54	88	143	232
$\dim \mathsf{Z}^{\mathrm{O}}_{\leq w}$	1	2	4	7	11	18	27	42	63	95	142
$\dim \mathrm{All}$	0	0	0	0	1	2	6	12	25	48	90

Table 12.1: Dimension of type O q-MZVs, proved rigorously for $w \leq 6$ and numerically for $w \leq 12$. $\mathrm{All} = \mathsf{DS}^{\mathrm{O}}_{\leq w} \cup \mathsf{DU}^{\mathrm{G}}_{\leq w}$

Applying the same idea as above it is possible to verify the following Okounkov's dimension conjecture, at least when the weight is small.

Conjecture 12.6.1. *Let* $\mathbf{Z}^{\mathrm{O}}_w$ *be the* $\mathbb{Q}$*-vector space generated by* $\mathfrak{z}^{\mathrm{O}}_q[\mathbf{s}]$, $|\mathbf{s}| \leq$

w. Then

$$\sum_{w=0}^{\infty} t^w \dim \mathbf{Z}^{\mathrm{O}}_{\leq w} = \frac{1}{1-t-t^2+t^6+t^8-t^{13}} - \frac{1}{1-t}$$
$$= t^2+2t^3+4t^4+7t^5+11t^6+18t^7+27t^8+42t^9+63t^{10}+95t^{11}+142t^{12}+O(t^{13}).$$

For example, one can verify all of the following $\mathbb{Q}$-linearly independent relations in the lower weight cases up to q^{100}. One can also rigorously prove the first identity given by Eq. (12.55) involving only weight 4 and 6 values by using the relations we have found for type $\mathbb{I}\!\mathbb{I}$ q-MZVs:

$$\begin{aligned}
4\mathfrak{z}[6] &= \mathfrak{z}[2,2] + 12\mathfrak{z}[3,3] - 6\mathfrak{z}[4,2], \qquad (12.55)\\
4\mathfrak{z}[7] &= \mathfrak{z}[2,3] + \mathfrak{z}[3,2] + 8\mathfrak{z}[3,4] + 6\mathfrak{z}[4,3] - 4\mathfrak{z}[5,2],\\
\mathfrak{z}[8] &= \mathfrak{z}[2,4] - \mathfrak{z}[6] + 2\mathfrak{z}[3,3] + 6\mathfrak{z}[4,4],\\
9\mathfrak{z}[8] &= \mathfrak{z}[6] - 6\mathfrak{z}[3,3] + 3\mathfrak{z}[4,2] + 20\mathfrak{z}[3,5] + 16\mathfrak{z}[5,3] - 10\mathfrak{z}[6,2],\\
\mathfrak{z}[8] &= 2\mathfrak{z}[2,6] - \mathfrak{z}[6] + 2\mathfrak{z}[3,3] + 4\mathfrak{z}[3,5] - 16\mathfrak{z}[5,3]\\
&\quad -6\mathfrak{z}[2,3,3] + 3\mathfrak{z}[2,4,2] - 6\mathfrak{z}[3,2,3] - 3\mathfrak{z}[4,2,2],\\
4\mathfrak{z}[3,6] &= \mathfrak{z}[2,5] + 4\mathfrak{z}[5,2] + 3\mathfrak{z}[3,4] + 6\mathfrak{z}[4,5] + 8\mathfrak{z}[5,4] + 2\mathfrak{z}[7,2],\\
8\mathfrak{z}[9] &= \mathfrak{z}[3,4] - 5\mathfrak{z}[2,5] - 8\mathfrak{z}[5,2] - 30\mathfrak{z}[4,5] - 2\mathfrak{z}[4,3] - 36\mathfrak{z}[5,4] - 10\mathfrak{z}[6,3],\\
6\mathfrak{z}[4,2] &= 10\mathfrak{z}[6] + 42\mathfrak{z}[8] - 60\mathfrak{z}[2,6] - 12\mathfrak{z}[3,3] - 120\mathfrak{z}[3,5] + 312\mathfrak{z}[5,3]\\
&\quad -15\mathfrak{z}[2,2,2] + 180\mathfrak{z}[2,3,3] - 90\mathfrak{z}[2,4,2] + 180\mathfrak{z}[3,2,3] + 60\mathfrak{z}[3,3,2],\\
72\mathfrak{z}[9] &= 62\mathfrak{z}[5,2] + 40\mathfrak{z}[2,5] - 4\mathfrak{z}[3,4] + 40\mathfrak{z}[3,6] - 2\mathfrak{z}[4,3] + 240\mathfrak{z}[4,5]\\
&\quad +264\mathfrak{z}[5,4] - 5\mathfrak{z}[2,2,3] - 60\mathfrak{z}[3,3,3] - 30\mathfrak{z}[4,2,3],\\
16\mathfrak{z}[9] &= 2\mathfrak{z}[3,4] - 10\mathfrak{z}[2,5] - 12\mathfrak{z}[2,7] - 8\mathfrak{z}[5,2] - 60\mathfrak{z}[4,5] - 24\mathfrak{z}[5,4]\\
&\quad +4\mathfrak{z}[2,3,2] + 4\mathfrak{z}[3,2,2] + 3\mathfrak{z}[2,2,3] + 24\mathfrak{z}[2,3,4] + 18\mathfrak{z}[2,4,3]\\
&\quad +12\mathfrak{z}[3,3,3] - 12\mathfrak{z}[2,5,2] + 24\mathfrak{z}[3,2,4] + 6\mathfrak{z}[4,3,2],\\
64\mathfrak{z}[9] &= 40\mathfrak{z}[2,5] + 20\mathfrak{z}[2,7] - 8\mathfrak{z}[3,4] + 44\mathfrak{z}[5,2] + 240\mathfrak{z}[4,5] + 168\mathfrak{z}[5,4]\\
&\quad +20\mathfrak{z}[3,6] - 4\mathfrak{z}[4,3] - 5\mathfrak{z}[2,3,2] - 5\mathfrak{z}[2,2,3] - 40\mathfrak{z}[2,3,4] - 30\mathfrak{z}[2,4,3]\\
&\quad +20\mathfrak{z}[2,5,2] - 5\mathfrak{z}[3,2,2] - 40\mathfrak{z}[3,2,4] - 100\mathfrak{z}[3,3,3] + 10\mathfrak{z}[3,4,2],\\
56\mathfrak{z}[9] &= 30\mathfrak{z}[2,5] + 20\mathfrak{z}[2,7] + 26\mathfrak{z}[5,2] - \mathfrak{z}[3,4] + 180\mathfrak{z}[4,5] + 112\mathfrak{z}[5,4]\\
&\quad +40\mathfrak{z}[3,6] - 6\mathfrak{z}[4,3] - 5\mathfrak{z}[2,2,3] - 5\mathfrak{z}[2,3,2] - 5\mathfrak{z}[3,2,2] - 40\mathfrak{z}[2,3,4]\\
&\quad +20\mathfrak{z}[5,2,2] - 40\mathfrak{z}[3,2,4] - 30\mathfrak{z}[2,4,3] + 20\mathfrak{z}[2,5,2] - 140\mathfrak{z}[3,3,3].
\end{aligned}$$

Therefore, Conjecture 12.6.1 is proved rigorously up to weight 6 (inclusive), and verified numerically up to weight 12 (inclusive). The list of relations for weight 10 to 12 is too long to be presented here.

12.7 The q-analog of Drinfeld associator

We have seen that the MZVs can be regarded as the coefficients of the Drinfeld associator related to the KZ equation given by Eq. (4.4). Its q-analog has the following form

$$\frac{1}{t}\frac{G(t)-G(qt)}{1-q}=\left(\frac{\pi}{t}+\frac{\pi y}{1-t}\right)G(t). \tag{12.56}$$

Notice that

$$\lim_{q\to 1}\frac{1}{t}\frac{G(t)-G(qt)}{1-q}=G'(t).$$

Theorem 12.7.1. *The solution to Eq.* (12.56) *is given by*

$$G(t)=\sum_{\mathbf{w}\in X^*_\pi}\mathfrak{Z}^{\text{Ш}}_q[\mathbf{w};t]\mathbf{w}\in t\mathbb{Q}[\![t,q]\!]\langle\langle X^*_\pi\rangle\rangle. \tag{12.57}$$

where $\mathfrak{Z}^{\text{Ш}}_q[\pi^{s_1}y\ldots\pi^{s_d}y;t]=(1-q)^{|\mathbf{s}|}\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};t]$ *as defined in Eq.* (12.2).

Proof. Let's consider $\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};qt]$. There are two cases. If $s_1=1$ then by writing $q^{k_1}t^{k_1}=t^{k_1}-(1-q^{k_1})t^{k_1}$ we have

$$\begin{aligned}\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};qt]&=\sum_{k_1>\cdots>k_d>0}\frac{q^{k_1}t^{k_1}}{(1-q^{k_1})(1-q^{k_1})^{s_2}\cdots(1-q^{k_d})^{s_d}}\\&=\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};t]-\sum_{k_2>\cdots>k_d>0}\sum_{k_1>k_2}\frac{t^{k_1}}{(1-q^{k_2})^{s_2}\cdots(1-q^{k_d})^{s_d}}\\&=\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};t]-\frac{t}{1-t}\sum_{k_2>\cdots>k_d>0}\frac{t^{k_2}}{(1-q^{k_2})^{s_2}\cdots(1-q^{k_d})^{s_d}}\\&=\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};t]-\frac{t}{1-t}\mathfrak{z}^{\text{Ш}}_q[s_2,\ldots,s_d;t].\end{aligned}$$

Similar but simpler computation shows that if $s_1>1$ then

$$\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};qt]=\mathfrak{z}^{\text{Ш}}_q[\mathbf{s};t]-\mathfrak{z}^{\text{Ш}}_q[s_1-1,s_2,\ldots,s_d;t].$$

Therefore,

$$G(qt)=G(t)-(1-q)\left(\pi+\pi y\frac{t}{1-t}\right)G(t).$$

This is obviously the same as Eq. (12.56). The proof is now complete. □

Definition 12.7.2. We define the q-analog of the Drinfeld associator as

$$\Phi_q:=\sum_{\mathbf{w}\in X^*_\pi}\mathfrak{z}^{\text{Ш}}_q[\mathbf{w};q]\mathbf{w}\in t\mathbb{Q}[\![t,q]\!]\langle\langle X^*_\pi\rangle\rangle \tag{12.58}$$

12.8 q-MZVs by Bachmann and Kühn

12.8.1 *The mono-bracket*

It is well-known that the Eisenstein series of weight $s \geq 4$

$$G_s(\tau) := \frac{1}{2} \sum_{(m,n)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m+n\tau)^s} = \zeta(s) + \frac{(2\pi i)^s}{(s-1)!} \sum_{n=1}^{\infty} \sigma_{s-1}(n) q^n,$$

where $q = e^{2\pi i\tau}$ and $\sigma_k(n) = \sum_{d|n} d^k$ is the divisor sum function. To define the multiple Eisenstein series, we first generalize $\sigma_k(n)$ to the multiple divisor sums as follows:

$$\sigma_{r_1,\ldots,r_l}(n) = \sum_{\substack{u_1v_1+\cdots+u_lv_l=n \\ u_1>\cdots>u_l\geq 1}} v_1^{r_1}\ldots v_l^{r_l}. \tag{12.59}$$

Definition 12.8.1. For any $s_1,\ldots,s_l \in \mathbb{N}$ the generating function for the multiple divisor sum $\sigma_{s_1-1,\ldots,s_l-1}$ is defined by the formal power series

$$[s_1,\ldots,s_l] := \frac{1}{(s_1-1)!\ldots(s_l-1)!} \sum_{n\geq 1} \sigma_{s_1-1,\ldots,s_l-1}(n) q^n \ \in \mathbb{Q}[\![q]\!],$$

which is called a *mono-bracket*. Here we call $|\mathbf{s}|$ the *weight* and l the *depth*. For convenience, we set $[\emptyset] = 1$.

It is straight-forward to see that

$$[s_1,\ldots,s_l] = \frac{1}{(s_1-1)!\cdots(s_l-1)!} \sum_{\substack{m_1,\ldots,m_l\geq 1 \\ d_1,\ldots,d_l\geq 1}} d_1^{s_1-1}\cdots d_l^{s_l-1} \times \\ \times\, q^{(m_1+\cdots+m_l)d_1+(m_2+\cdots+m_l)d_2+\cdots+m_ld_l}. \tag{12.60}$$

By the Prop. 12.8.3, we see that the mono-bracket $[\mathbf{s}]$ also provides a q-analog of MZV $\zeta(\mathbf{s})$. First, we define the s-th (modified) Eulerian polynomial by

$$\frac{P_s(z)}{(1-z)^{s+1}} = \left(z\frac{\mathrm{d}}{\mathrm{d}z}\right)^s \frac{z}{1-z} = \sum_{d=1}^{\infty} d^s z^d. \tag{12.61}$$

where $P_s(z) = z\tilde{P}_s(z)$. We need to find its special value $P_s(1) = \tilde{P}_s(1)$ even though we do not need the full strength of the following lemma.

Lemma 12.8.2. *For any $s \in \mathbb{N}$ we have*

$$\tilde{P}_s(X) = \sum_{n=0}^{s-1} A_{s,n} X^n,$$

where the Eulerian numbers $A_{s,n}$ are defined by

$$A_{s,n} = \sum_{i=0}^{n} (-1)^i \binom{s+1}{i} (n+1-i)^s.$$

Moreover, $P_s(1) = \tilde{P}_s(1) = s!$.

Proof. By the definition of Eq. (12.61) we see easy to see that $\tilde{P}_0(z) = \tilde{P}_1(z) = 1$ and $\mathcal{P}_1(z) = 1$ where $\mathcal{P}_s(X) = \sum_{n=0}^{s-1} A_{s,n} X^n$. Now it is not hard to verify that $\mathcal{P}_s(X)$ satisfy the recurrence

$$\mathcal{P}_{s+1}(z) = \mathcal{P}_s(z)(1+sz) + z(1-z)\mathcal{P}_s'(z),$$

the same as that of $\tilde{P}_s(z)$, yielding $\tilde{P}_s(z) = \mathcal{P}_s(z)$. This recurrence further implies that $\tilde{P}_s(1) = s!$, as desired. □

Proposition 12.8.3. *For any $\mathbf{s} \in \mathbb{N}^l$ with $s_1 \geq 2$, we have*

$$\lim_{q \to 1^-} (1-q)^{|\mathbf{s}|} [\mathbf{s}] = \zeta(\mathbf{s}) := \sum_{n_1 > \cdots > n_l \geq 1} \frac{1}{n_1^{s_1} \cdots n_l^{s_l}}. \tag{12.62}$$

Proof. Clear by Lemma 12.8.2. □

Definition 12.8.4. We define a *normalized polylog* by

$$\mathrm{Ln}_s(z) := \frac{\mathrm{Li}_{1-s}(z)}{\Gamma(s)},$$

where for $z \in \mathbb{C}$, $|z| < 1$ the polylog $\mathrm{Li}_s(z)$ is extended to complex weight s:

$$\mathrm{Li}_s(z) = \sum_{n \geq 1} \frac{z^n}{n^s}.$$

The normalized polylog extends to an entire function in s and to a holomorphic function in z where $|z| < 1$.

Proposition 12.8.5. *For $q \in \mathbb{C}$ with $|q| < 1$ and for all $s_1, \ldots, s_l \in \mathbb{N}$ we can write the mono-bracket as*

$$\begin{aligned}[s_1, \ldots, s_l] &= \frac{1}{(s_1-1)! \ldots (s_l-1)!} \sum_{n_1 > \cdots > n_l > 0} \prod_{j=1}^{l} \frac{P_{s_j-1}(q^{n_j})}{(1-q^{n_j})^{s_j}} \\ &= \sum_{n_1 > \cdots > n_l \geq 1} \mathrm{Ln}_{s_1}(q^{n_1}) \ldots \mathrm{Ln}_{s_l}(q^{n_l})\end{aligned}$$

where $P_k(t)$ is the k-th Eulerian polynomial.

Proof. We have by the definition Eq. (12.61)

$$\sum_{n_1>\cdots>n_l>0}\prod_{j=1}^{l}\frac{P_{s_j-1}(q^{n_j})}{(1-q^{n_j})^{s_j}}=\sum_{n_1>\cdots>n_l>0}\prod_{j=1}^{l}\sum_{v_j\geq 1}v_j^{s_j-1}q^{v_jn_j}$$
$$=\sum_{n\geq 1}\sigma_{s_1-1,\ldots,s_l-1}(n)q^n.$$

This yields the first equality in the proposition. The second follows from Definition 12.8.4. □

Corollary 12.8.6. *If $s_1\geq 2$ then*

$$\lim_{q\to 1^-}(1-q)^{|\mathbf{s}|}[\mathbf{s}]=\zeta(\mathbf{s}).$$

Proof. This follows from $P_s(1)=s!$ by Lemma 12.8.2 immediately. □

Definition 12.8.7. Let MB to be the $\mathbb{Q}$-vector space generated by $[\emptyset]=1\in\mathbb{Q}[\![q]\!]$ and all the mono-brackets $\Big\{[s_1,\ldots,s_l]:l\in\mathbb{N},s_1,\ldots,s_l\in\mathbb{N}\Big\}$.

In order to study the product structure of the mono-brackets we need the following lemma.

Lemma 12.8.8. *For $a,b\in\mathbb{N}$ we have*

$$\mathrm{Ln}_a(z)\cdot\mathrm{Ln}_b(z)=\sum_{j=1}^{a}\lambda_{a,b}^{j}\mathrm{Ln}_j(z)+\sum_{j=1}^{b}\lambda_{b,a}^{j}\mathrm{Ln}_j(z)+\mathrm{Ln}_{a+b}(z),$$

where

$$\lambda_{a,b}^{j}=(-1)^{b-1}\binom{a+b-j-1}{a-j}\frac{B_{a+b-j}}{(a+b-j)!}\in\mathbb{Q}. \tag{12.63}$$

Proof. Set

$$L(X):=\sum_{k\geq 1}\mathrm{Ln}_k(z)X^{k-1}=\sum_{k\geq 1}\sum_{n\geq 1}\frac{n^{k-1}z^n}{(k-1)!}X^{k-1}=\sum_{n\geq 1}e^{nX}z^n=\frac{e^Xz}{1-e^Xz}.$$

Then

$$L(X)\cdot L(Y)=\frac{1}{e^{X-Y}-1}L(X)+\frac{1}{e^{Y-X}-1}L(Y)$$
$$=\sum_{n\geq 1}\frac{B_n}{n!}\Big((X-Y)^{n-1}L(X)+(Y-X)^{n-1}L(Y)\Big)+\frac{L(X)-L(Y)}{X-Y}.$$

We can now derive the lemma by comparing the coefficient of $X^{a-1}Y^{b-1}$. □

It is straight-forward to see that the product of two mono-brackets can be found by using Prop. 12.8.5 and Lemma 12.8.8. Furthermore, similar to MZVs, mono-brackets together with their products can be formalized by using some algebra of words.

Definition 12.8.9. Let alphabet $A_{\mathsf{MB}} = \{z_j : j \in \mathbb{N}\}$, A^*_{MB} the set of words over A_{MB}, and $\mathbb{Q}\langle A_{\mathsf{MB}}\rangle$ the noncommutative polynomial algebra over $\mathbb{Q}$ generated by A^*_{MB}. Define $\diamond : A^2_{\mathsf{MB}} \to \mathbb{Q}\langle A_{\mathsf{MB}}\rangle$ by

$$z_a \diamond z_b = z_{a+b} + \sum_{j=1}^{a} \lambda^j_{a,b} z_j + \sum_{j=1}^{b} \lambda^j_{b,a} z_j \qquad \forall a, b \in \mathbb{N}.$$

For all $x, y \in A_{\mathsf{MB}}$ and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in A^*_{\mathsf{MB}}$ we define a multiplication on $\mathbb{Q}\langle A_{\mathsf{MB}}\rangle$ recursively by $\mathsf{e} * \mathbf{w} = \mathbf{w} * \mathsf{e} = \mathbf{w}$ and

$$x\mathbf{u} \star y\mathbf{v} := x(\mathbf{u} \star y\mathbf{v}) + y(x\mathbf{u} \star \mathbf{v}) + (x \diamond y)(\mathbf{u} \star \mathbf{v}).$$

The *weight* of a word $z_{s_1} \dots z_{s_l}$ is defined to be $s_1 + \cdots + z_{s_l}$ and the *depth*, l.

Proposition 12.8.10. *The algebra $(\mathbb{Q}\langle A^*_{\mathsf{MB}}\rangle, \star)$ is a commutative and associative $\mathbb{Q}$-algebra.*

Proof. The proof is routine by using induction on the sum of of depths of the words involved. For example, see the proof of [286, Thm. 2.1]. □

Proposition 12.8.11. *Define a $\mathbb{Q}$-linear map $[\,.\,] : (\mathbb{Q}\langle A_{\mathsf{MB}}\rangle, *) \to (\mathsf{MB}, \cdot)$ so that $[z_{s_1} \dots z_{s_l}] := [s_1, \dots, s_l]$. Then we have*

$$[\mathbf{u} \star \mathbf{v}] = [\mathbf{u}] \cdot [\mathbf{v}] \qquad \forall \mathbf{u}, \mathbf{v} \in A^*_{\mathsf{MB}}.$$

Thus $[\,.\,]$ a $\mathbb{Q}$-algebra homomorphism.

Proof. This follows by the same argument as in the MZV case by Lemma 12.8.8 and Prop. 12.8.5. □

Example 12.8.12. For any $r, s, t \in \mathbb{N}$ we have

$$[s] \cdot [t] = [s,t] + [t,s] + [s+t] + \sum_{j=1}^{s} \lambda^j_{s,t}[j] + \sum_{j=1}^{t} \lambda^j_{t,s}[j]$$

and

$$\begin{aligned}[r] \cdot [s,t] = {} & [r,s,t] + [s,r,t] + [s,t,r] + [s,r+t] + [r+s,t] \\ & + \sum_{j=1}^{r} \lambda^j_{r,t}[s,j] + \sum_{j=1}^{t} \lambda^j_{t,r}[s,j] + \sum_{j=1}^{r} \lambda^j_{r,s}[j,t] + \sum_{j=1}^{s} \lambda^j_{s,r}[j,t].\end{aligned}$$

Considering the analogy between MZVs and the mono-brackets we see that a missing link is the shuffle product structure. We will turn to this when we study the bi-brackets which are generalizations of mono-brackets.

We know an admissible MZV $\zeta(s_1,\dots,s_l)$ must have $s_1 \geq 2$ so we define $q\mathsf{MZ}$ to be the subset of MB generated by the mono-brackets $[s_1,\dots,s_l]$ with $s_1 \geq 2$. We expect that mono-brackets $[1,\dots]$ should behave quite differently from those in $q\mathsf{MZ}$ which is reflected in the next result.

Proposition 12.8.13. *The mono-bracket* $[1]$ *is transcendental over* $q\mathsf{MZ}$.

Proof. By Cor. 12.8.6 we see that if $s_1 \geq 2$ then near $q = 1$ we have $[\mathbf{s}] \sim \zeta(\mathbf{s})(1-q)^{-|\mathbf{s}|}$. We now claim that

$$[1] \asymp -\frac{\log(1-q)}{1-q}. \tag{12.64}$$

Indeed, by the definition

$$[1] = \sum_{n=1}^{\infty} \sigma_0(n) q^n = \sum_{m,n=1}^{\infty} q^{mn} = \sum_{n=1}^{\infty} \frac{q^n}{1-q^n}.$$

Thus

$$[1] = \frac{1}{1-q} \sum_{n=1}^{\infty} \frac{q^n}{[n]} \geq -\frac{\log(1-q)}{1-q}. \tag{12.65}$$

On the other hand, let $n \in S_N$ if $N = \lfloor 1 + q + \cdots + q^n \rfloor$. Then

$$[1] = \frac{1}{1-q} \sum_{N=1}^{\lfloor 1/(1-q) \rfloor} \frac{1}{N} \sum_{n \in S_N} q^n.$$

Let $N_* = \min\{n : n \in S_N\}$ and $N^* = \max\{n : n \in S_N\}$. Then

$$\lfloor 1 + q + \cdots + q^{N_*} \rfloor = \lfloor 1 + q + \cdots + q^{N_*} + q^{N_*+1} + \cdots + q^{N^*} \rfloor.$$

Thus $q^{N_*+1} + \cdots + q^{N^*} < 1$. Therefore $\sum_{n \in S_N} q^n < 2$ and

$$[1] < \frac{2}{1-q} \sum_{N=1}^{\lfloor 1/(1-q) \rfloor} \frac{1}{N} < \frac{2}{1-q} \left(1 + \int_1^{1/(1-q)} \frac{\mathrm{d}x}{x} \right) = \frac{2\big(1 - \log(1-q)\big)}{1-q}.$$

Together with Eq. (12.65) this yields Eq. (12.64) which implies that $[1]$ is transcendental over $q\mathsf{MZ}$. □

Proposition 12.8.14. MB *can be regarded as a polynomial ring over* $q\mathsf{MZ}$ *in the variable* $[1]$, *i.e.,* $\mathsf{MB} = q\mathsf{MZ}[[1]]$.

Proof. Let $f = [\{1\}^n, s_1, \dots, s_{l-n}]$, with $s_1 > 1$. We now prove the proposition by induction m. The case $n = 0$ is trivial so we assume $n \geq 1$. But by the stuffle product given by Prop. 12.8.11 and the induction

$$[1] \cdot [\{1\}^{n-1}, s_1, \dots, s_{l-n}] - nf \in q\mathsf{MZ}[[1]].$$

So $f \in q\mathsf{MZ}[[1]]$. This finishes the proof of the proposition. □

12.8.2 *A derivation*

In this subsection we turn to the operator $\mathfrak{D} = q\frac{\mathrm{d}}{\mathrm{d}q}$ on MB and show it is a derivation and can provide some useful information of MB. For this purpose, we need to consider some finer structures of MB.

Definition 12.8.15. Let $\mathrm{Fil}^{\mathrm{W}}_{\bullet}$ denote the increasing weight filtration on MB and $\mathrm{Fil}^{\mathrm{D}}_{\bullet}$ the increasing depth filtration, namely,

$$\mathrm{Fil}^{\mathrm{W}}_{k}(\mathsf{MB}) := \Big\langle [\mathbf{s}] : \mathbf{s} \in \mathbb{N}^r, r \geq 1, |\mathbf{s}| \leq k \Big\rangle_{\mathbb{Q}},$$
$$\mathrm{Fil}^{\mathrm{D}}_{l}(\mathsf{MB}) := \Big\langle [\mathbf{s}] : \mathbf{s} \in \mathbb{N}^r, r \leq l \Big\rangle_{\mathbb{Q}}.$$

If we consider both the depth and weight filtrations at the same time we write $\mathrm{Fil}^{\mathrm{W,D}}_{k,l} := \mathrm{Fil}^{\mathrm{W}}_{k}\,\mathrm{Fil}^{\mathrm{D}}_{l}$.

Theorem 12.8.16. *The operator $\mathfrak{D} = q\frac{\mathrm{d}}{\mathrm{d}q}$ is a derivation on MB, and*

$$\mathfrak{D}\left(\mathrm{Fil}^{\mathrm{W,D}}_{k,l}(\mathsf{MB})\right) \subseteq \mathrm{Fil}^{\mathrm{W,D}}_{k+2,l+1}(\mathsf{MB}).$$

More precisely, let $\oplus$ be componentwise addition and $\mathbf{e}_j = [\ldots, 0, 1, 0, \ldots]$ where 1 only occurs at the jth component. Then for all $\mathbf{s} = [s_1, \ldots, s_l] \in \mathsf{MB}$ we have (setting $s_{l+1} = 1$ and $\tilde{\mathbf{s}} = [s_1, \ldots, s_{l+1}]$)

$$\begin{aligned}\mathfrak{D}[s_1, \ldots, s_l] &= [2] \cdot [s_1, \ldots, s_l] - \sum_{j=1}^{l+1} s_j(\tilde{\mathbf{s}} + \mathbf{e}_j) + \sum_{k=1}^{l}(l-k+1)s_k(\mathbf{s} + \mathbf{e}_k) \\ &\quad - \sum_{j=1}^{l}\sum_{k=1}^{j-1} s_k \sum_{\alpha+\beta=s_j+1} ([s_1, \ldots, s_{j-1}, \alpha, \beta, s_{j+1}, \ldots, s_l] + \mathbf{e}_k) \\ &\quad - \sum_{j=1}^{l}(\alpha-1) \sum_{\alpha+\beta=s_j+2} [s_1, \ldots, s_{j-1}, \alpha, \beta, s_{j+1}, \ldots, s_l].\end{aligned}$$

Proof. For any fixed $l \in \mathbb{N}$, define the generating series

$$\begin{aligned}T(X_1, \ldots, X_l) &:= \sum_{s_1, \ldots, s_l \geq 1} [s_1, \ldots, s_l] \prod_{j=1}^{l} X_j^{s_j - 1} \\ &= \sum_{n_1, \ldots, n_l \geq 1} \prod_{j=1}^{l} \frac{e^{n_j X_j} q^{n_1 + \cdots + n_j}}{1 - q^{n_1 + \cdots + n_j}}.\end{aligned}$$

Then we have

$$\begin{aligned} &T(X)\cdot T(Y_1,\dots,Y_l)\\ &\quad=\sum_{m,n_1,\dots,n_l\geq 1} e^{mX+n_1Y_1+\cdots+n_lY_l}\frac{q^m}{1-q^m}\frac{q^{n_1}}{1-q^{n_1}}\cdots\frac{q^{n_1+\cdots+n_l}}{1-q^{n_1+\cdots+n_l}}\\ &=T(X+Y_1,\dots,X+Y_l,X)+\sum_{j=1}^{l}T(X+Y_1,\dots,X+Y_j,Y_j,\dots,Y_l)\\ &\qquad\qquad+R_l-\sum_{j=1}^{l}T(X+Y_1,\dots,X+Y_j,Y_{j+1},\dots,Y_l),\quad (12.66)\end{aligned}$$

where the last line comes from those terms when $m=n_1+\cdots+n_j$ for some j so that

$$R_l=\sum_{j=1}^{l}\left(\sum_{n_1,\dots,n_l\geq 1}e^{\sum_{k=1}^{j}n_k(X+Y_k)+\sum_{k=j+1}^{l}n_kY_k}\prod_{i=1}^{l}\frac{q^{n_1+\cdots+n_i}}{(1+q^{n_1+\cdots+n_i})^{\delta_{i,j}+1}}\right)$$

since

$$\left(\frac{q^n}{1-q^n}\right)^2=\frac{q^n}{(1-q^n)^2}-\frac{q^n}{1-q^n}$$

The first line of comes from the terms whose indices satisfy $m>n_1+\cdots+n_l$ (put $j=l+1$) or $n_1+\cdots+n_{j-1}<m<n_1+\cdots+n_j$ for some j. Setting $m=n_1+\cdots+n_{j-1}+m'$ and $n_j=m'+n_j'$ for these terms we see the sum is now over $m',n_1,\dots,n_j',\dots,n_l$ which then gives $T(X+Y_1,\dots,X+Y_j,Y_j,\dots,Y_l)$.

Define $D(f)=\left(\frac{\mathrm{d}}{\mathrm{d}X}f\right)\big|_{X=0}$. Then

$$D(R_l)=\sum_{j=1}^{l}\left(\sum_{n_1,\dots,n_l\geq 1}(n_1+\cdots+n_j)e^{\sum_{k=1}^{l}n_kY_k}\prod_{i=1}^{l}\frac{q^{n_1+\cdots+n_i}}{(1+q^{n_1+\cdots+n_i})^{\delta_{i,j}+1}}\right)$$

and this is exactly $\mathfrak{D}\,T(Y_1,\dots,Y_l)$ which can be seen by induction on l and the Leibniz rule. The explicit expression of $\mathfrak{D}$ is a straight-forward computation and is thus left to the interested reader. This finishes the proof of the theorem. □

Corollary 12.8.17. *The space qMZ is closed under $\mathfrak{D}$.*

Proof. Clear. □

Example 12.8.18. We know that the Eulerian polynomials $P_1(t) = t$ and $P_2(t) = t(t+1)$. Therefore we have

$$\frac{t^2}{(1-t)^3} = \frac{P_2(t)}{2(1-t)^3} - \frac{1}{2}\frac{P_1(t)}{(1-t)^2}.$$

This quickly leads to

$$\mathfrak{z}_q^{\mathrm{I}}[3] = [3] - \frac{1}{2}[2] = [2,1] - [2] + \mathfrak{D}[1]$$

using the first identity in Exercise 12.17.

A surprising property of the derivation $\mathfrak{D}$ is that it can provide many $\mathbb{Q}$-linear relations between ordinary MZVs. To see this we define for any $\kappa \geq 1$ the following spaces

$$\mathsf{F}_\kappa = \Big\{ \sum_{n\geq 1} a_n q^n \in \mathbb{R}[\![q]\!] : a_n = O(n^{\kappa-1}) \Big\}$$
$$\mathsf{F}'_\kappa = \Big\{ \sum_{n\geq 1} a_n q^n \in \mathbb{R}[\![q]\!] : a_n = O(n^{\kappa-1}\log^M(n)) \text{ for some } M > 0 \Big\}$$
$$\mathsf{F}_{<\kappa} = \Big\{ \sum_{n\geq 1} a_n q^n \in \mathbb{R}[\![q]\!] : \exists\, \varepsilon > 0 \text{ with } a_n = O(n^{\kappa-1-\varepsilon}) \Big\}.$$

Lemma 12.8.19. *Let $\kappa > 1$. Then*

(i) *All of $\mathsf{F}_{<\kappa}$, F'_κ and F_κ are $\mathbb{R}$ vector spaces.*
(ii) *We have $\mathsf{F}_{\kappa-1} \subset \mathsf{F}_{<\kappa} \subset \mathsf{F}_\kappa$.*
(iii) *Let $r, s \geq 1$. Then*

$$\mathsf{F}_{<r}\cdot\mathsf{F}_{<s} \subset \mathsf{F}_{<r+s}, \quad \mathsf{F}_{<r}\cdot\mathsf{F}_s \subset \mathsf{F}_{<r+s}, \quad \mathsf{F}_r\cdot\mathsf{F}_s \subset \mathsf{F}_{r+s},$$
$$\mathsf{F}_{<r}\cdot\mathsf{F}'_s \subset \mathsf{F}_{<r+s}, \quad \mathsf{F}_r\cdot\mathsf{F}'_s \subset \mathsf{F}'_{r+s}, \quad \mathsf{F}'_r\cdot\mathsf{F}'_s \subset \mathsf{F}'_{r+s}.$$

Proof. (i) and (ii) are straight-forward. For (iii) let $f = \sum_{n\geq 1} a_n q^n \in \mathsf{F}_r$, $g = \sum_{n\geq 1} b_n q^n \in \mathsf{F}_s$. Assume $fg = \sum_{n\geq 1} c_n q^n$. Then by the definition

$$c_n = \sum_{n_1+n_2=n} a_{n_1} b_{n_2} = \sum_{n_1+n_2=n} O(n_1^{r-1})O(n_2^{s-1}) = O(n^{r+s-1}).$$

Thus $fg \in \mathsf{F}_{r+s}$. The other five cases can be proved similarly. □

For $\kappa > 1$ define the map $Z_\kappa : \mathbb{R}[\![q]\!] \to \mathbb{R}$ by

$$Z_\kappa\Big(\sum_{n\geq 1} a_n q^n\Big) = \lim_{q\to 1^-} (1-q)^\kappa \sum_{n\geq 1} a_n q^n.$$

Proposition 12.8.20. *We have*

(i) Z_κ *is a linear map from* F_κ *to* $\mathbb{R}$
(ii) $\mathsf{F}_{<\kappa} \subset \ker Z_\kappa$.
(iii) $\mathfrak{D}\,\mathsf{F}_{<\kappa-1} \subset \ker(Z_\kappa)$.

Proof. By [146, Cor. 1.2], if $-s \notin \mathbb{N}$ and $e^{-2\pi} < q < 1$ then we have

$$\mathrm{Li}_{-s}(q) = \Gamma(1+s)(-\log q)^{-s-1} + \sum_{n=0}^{\infty} \frac{\zeta(-s-n)}{n!} (\log q)^n$$

where $\zeta(-s-n)$ is the analytic continuation of the Riemann zeta-function. Hence for $\kappa > 1$ and $\varepsilon \geq 0$

$$\begin{aligned}
&\lim_{q\to 1^-} (1-q)^\kappa \sum_{n\geq 1} n^{\kappa-1-\varepsilon} q^n \\
=& \lim_{q\to 1^-} (1-q)^\kappa \,\mathrm{Li}_{\varepsilon+1-\kappa}(q) \\
=& \lim_{q\to 1^-} (1-q)^\kappa \left(\Gamma(\kappa-\varepsilon)(-\log q)^{\varepsilon-\kappa} + \sum_{n=0}^{\infty} \frac{\zeta(\varepsilon-\kappa-n)}{n!} (\log q)^n \right) \\
=& \Gamma(\kappa-\varepsilon) \lim_{q\to 1^-} (-\log q)^\varepsilon \left(\frac{q-1}{\log q} \right)^\kappa \\
=& \begin{cases} \Gamma(\kappa), & \text{if } \varepsilon = 0; \\ 0, & \text{if } \varepsilon > 0. \end{cases}
\end{aligned}$$

Now assume that for a $\varepsilon \geq 0$ we have $f = \sum_{n\geq 1} a_n q^n$ with $a_n = O(n^{\kappa-1-\varepsilon})$, i.e. $f \in \mathsf{F}_\kappa$ for $\varepsilon = 0$ and $f \in \mathsf{F}_{<\kappa}$ for $\varepsilon > 0$, then the calculation above gives

$$|Z_\kappa(f)| \ll Z_\kappa\Big(\sum_{n\geq 1} n^{\kappa-1-\varepsilon} q^n \Big) = \begin{cases} C \cdot \Gamma(\kappa), & \text{if } \varepsilon = 0; \\ 0, & \text{if } \varepsilon > 0. \end{cases}$$

Hence (i) and (ii) are proved.

For (iii) notice that $\mathfrak{D} \sum_{n\geq 1} a_n q^n = \sum_{n\geq 1} n a_n q^n$. So by (i) we get $\mathfrak{D}\,(\mathsf{F}_{<\kappa-1}) \subset \mathsf{F}_{<\kappa} \subset \ker(Z_\kappa)$. □

The brackets $[s_1, \ldots, s_l]$ can be considered as elements in the spaces we studied above.

Proposition 12.8.21. *Let* $\mathbf{s} = (s_1, \ldots, s_l) \in \mathbb{N}^l$. *Then*

(i) *If all* $s_1, \ldots, s_l > 2$, *then* $[s_1, \ldots, s_l] \in \mathsf{F}_{|\mathbf{s}|}$.
(ii) *We have* $[s_1, \ldots, s_l] \in \mathsf{F}'_{<|\mathbf{s}|} \subset \mathsf{F}_{<|\mathbf{s}|+1}$.
(iii) *We have*

$$[s_1, \ldots, s_l] \subset \ker(Z_{|\mathbf{s}|+1}) \quad \textit{and} \quad \mathfrak{D}[s_1, \ldots, s_l] \subset \ker(Z_{|\mathbf{s}|+2}).$$

Proof. If $s > 2$ the divisor sum

$$\sigma_{s-1}(n) < \sum_{d=1}^{n} \left(\frac{n}{d}\right)^{s-1} < \zeta(s-1)n^{s-1}$$

and therefore $[s] \in \mathsf{F}_s$. Then by Lemma 12.8.19 (iii) we have

$$\sum_{n\geq 1} a_{s_1,\ldots,s_l}(n)q^n := [s_1]\ldots[s_l] \in \mathsf{F}_{s_1+\cdots+s_l}$$

Hence

$$\begin{aligned}\sigma_{s_1-1,\ldots,s_l-1}(n) &= \sum_{\substack{u_1v_1+\cdots+u_lv_l=n\\ u_1>\cdots>u_l\geq 1}} v_1^{s_1-1}\ldots v_l^{s_l-1}\\ &\leq \sum_{u_1v_1+\cdots+u_lv_l=n} v_1^{s_1-1}\ldots v_l^{s_l-1} = a_{s_1,\ldots,s_l}(n)\,,\end{aligned}$$

which implies $[s_1,\ldots,s_l] \in \mathsf{F}_{s_1+\cdots+s_l}$.

In order to show (ii) we can use the same argument as above except that we have $\sigma_0(n) = O(\log(n))$ and $\sigma_1(n) = O(n\log(n))$ so that $[1] \in \mathsf{F}'_1$ and $[2] \in \mathsf{F}'_2$. By Lemma 12.8.19 (iii) we thus get $[s_1,\ldots,s_l] \in \mathsf{F}'_{|\mathbf{s}|} \subset \mathsf{F}_{<|\mathbf{s}|+1}$ for $s_1,\ldots,s_l \geq 1$.

Finally, (iii) is an easy consequence of Prop. 12.8.20 (ii) and (iii). □

Since $\mathsf{MB} = q\mathsf{MZ}[[1]]$ by Prop. 12.8.14 we can define a map

$$\begin{aligned}Z_k^{alg} : \mathrm{Fil}_k^{\mathrm{W}}(\mathsf{MB}) &\longrightarrow \mathbb{R}[T],\\ \sum_{j=0}^{k} g_j[1]^{k-j} &\longmapsto \sum_{j=0}^{k} Z_j(g_j)T^{k-j},\end{aligned}$$

where $g_j \in \mathrm{Fil}_j^{\mathrm{W}}(q\mathsf{MZ})$.

Proposition 12.8.22. *For all $f \in \mathrm{Fil}_k^{\mathrm{W}}(\mathsf{MB})$ we have $Z_{k+2}^{alg}(\mathfrak{D}\, f) = 0$.*

Proof. An element in $\mathrm{Fil}_k^{\mathrm{W}}(\mathsf{MB})$ can be written as $\sum_{j=0}^{k} g_j[1]^{k-j}$ with $g_j \in \mathrm{Fil}_j^{\mathrm{W}}(q\mathsf{MZ})$. The map Z_k^{alg} is linear, it therefore suffices to prove the statement for a $f \in \mathrm{Fil}_k^{\mathrm{W}}(\mathsf{MB})$ of the form $f = g_j[1]^{k-j}$. Now

$$\mathfrak{D}\, f = \mathfrak{D}\, g_j \cdot [1]^{k-j} + (k-j)g_j\, \mathfrak{D}[1]\cdot[1]^{k-j-1}.$$

Since $\mathfrak{D}[1] = [3] + \frac{1}{2}[2] - [2,1] \in q\mathsf{MZ}$ (see Exercise 12.17) and $\mathfrak{D}\, g_j \in q\mathsf{MZ}$ since $\mathfrak{D}$ is closed on $q\mathsf{MZ}$ by Cor. 12.8.17. The map Z_{k+2}^{alg} is therefore given on $\mathfrak{D}\, f$ by

$$Z_{k+2}^{alg}(\mathfrak{D}\, f) = Z_{j+2}\left(\mathfrak{D}\, g_j\right)T^{k-j} + (k+j)Z_{j+3}\left(g_j\, \mathfrak{D}[1]\right)T^{k-j-1}.$$

We get $Z_{j+3}\left(g_j\, \mathfrak{D}[1]\right) = Z_j(g_j)\cdot Z_3(\mathfrak{D}[1])$ and by Prop. 12.8.21 we obtain $Z_3\left(\mathfrak{D}[1]\right) = Z_{j+2}\left(\mathfrak{D}\, g_j\right) = 0$ from which the statement follows. □

12.8.3 *Multiple Eisenstein series*

We have seen in § 5.8.3 that Eisenstein series plays a crucial role in the study of modular forms. It is well-known (see [158, Chap. 4]) that the Eisenstein series of weight $s \in 2\mathbb{Z}_{\geq 2}$ has the following expansion at $i\infty$:

$$G_s(\tau) := \frac{1}{2} \sum_{(m,n)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m+n\tau)^s} = \zeta(s) + \frac{(2\pi i)^s}{(s-1)!} \sum_{n=1}^{\infty} \sigma_{s-1}(n) q^n$$

where $q = e^{2\pi i\tau}$. Using mono-brackets we get

$$G_s(\tau) = \zeta(s) + (2\pi i)^s [s].$$

As we have the multiple variable version of both $\zeta(s)$ and $[s]$ it is natural to generalize $G_s(\tau)$ as follows.

Definition 12.8.23. For any $m, c \in \mathbb{Z}$ and $\tau \in \mathbb{H}$ (upper half plane), we write $m\tau + c \succ 0$ if $m \geq 1$ or $m = 0$ and $c \geq 1$ and $m\tau + c \succ n\tau + d$ if $m\tau + c - (n\tau + d) \succ 0$. For any $\mathbf{a} = (c_1, \ldots, c_d) \in (\mathbb{Z}/N\mathbb{Z})^d$ and $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ we define the multiple Eisenstein series by

$$G_{\mathbf{s}}(\tau) := \sum_{\substack{m_1\tau+c_1 \succ \cdots \succ m_d\tau+c_d \succ 0 \\ m_j, c_j \in \mathbb{Z},\ \forall j}} \frac{1}{(m_1\tau + c_1)^{s_1} \cdots (m_d\tau + c_d)^{s_d}}. \tag{12.67}$$

We will call d the *depth* and the sum $s_1 + \cdots + s_d$ the *weight*.

Remark 12.8.24. It is not too hard to show that the series in Eq. (12.67) converges absolutely when $s_1 \geq 3$ and $s_j \geq 2$ for all $j \geq 2$. It converges conditionally if $s_j \geq 2$ for all $j \geq 1$. See [25, Prop. 2.2].

By the definition, multiple Eisenstein series clearly satisfy the stuffle relation. For example,

$$G_2(\tau)G_3(\tau) = G_{2,3}(\tau) + G_{3,2}(\tau) + G_5(\tau).$$

However, it is much harder to define the shuffle product structure since no integral expressions for multiple Eisenstein series are unavailable. Bachmann and Tasaka partially developed such a theory by using the Hopf algebra structure of the (generic) iterated integrals introduced by Goncharov. We refer the interested reader to their original paper [25] for details.

12.8.4 *Bi-brackets and double shuffle relations*

To find the DBSF for the brackets, Bachmann introduced the *bi-brackets* which are generalizations of the mono-brackets. For any $\mathbf{s} = (s_1, \dots, s_l), \mathbf{r} = (r_1, \dots, r_l) \in \mathbb{N}^l$, we modify his definition by

$$\begin{aligned}\begin{bmatrix}\mathbf{s}\\ \mathbf{r}\end{bmatrix} &:= \prod_{j=1}^{l} \sum_{\substack{n_1>\cdots>n_l\geq 1\\ d_1,\dots,d_l\geq 1}} \prod_{j=1}^{l} \frac{n_j^{r_j-1}}{(r_j-1)!}\frac{\mathrm{d}_j^{s_j-1}}{(s_j-1)!} q^{n_j d_j}\\ &= \frac{1}{\prod_{j=1}^{l}(r_j-1)!(s_j-1)!} \sum_{n_1>\cdots>n_l\geq 1} \prod_{j=1}^{l} \frac{n_j^{r_j-1} P_{s_1-1}(q^{n_j})}{(1-q^{n_j})^{s_j}}. \end{aligned} \tag{12.68}$$

Here we call $|\mathbf{s}| + |\mathbf{r}|$ the *weight* and l the *depth*. Let BB be $\mathbb{Q}$-vector space generated by $\left[\begin{smallmatrix}\emptyset\\ \emptyset\end{smallmatrix}\right] = 1 \in \mathbb{Q}[\![q]\!]$ and all the bi-brackets.

As before, to investigate the product structures of BB we resort to an algebra of words.

Definition 12.8.25. Let alphabet $A_{\mathsf{BB}} = \{z_{s,r} : s, r \in \mathbb{N}\}$, A_{BB}^* the set of words over A_{BB}, and $\mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ the noncommutative polynomial algebra over $\mathbb{Q}$ generated by A_{BB}^*. Define $\diamond : A_{\mathsf{BB}}^2 \to \mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ by

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

where λ's are defined by Eq. (12.63). For all $x, y \in A_{\mathsf{BB}}$ and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in A_{\mathsf{BB}}^*$ we define a multiplication on $\mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ recursively by $\mathsf{e} \star \mathbf{w} = \mathbf{w} \star \mathsf{e} = \mathbf{w}$ and

$$x\mathbf{u} \star y\mathbf{v} := x(\mathbf{u} \star y\mathbf{v}) + y(x\mathbf{u} \star \mathbf{v}) + (x \diamond y)(\mathbf{u} \star \mathbf{v}). \tag{12.69}$$

Proposition 12.8.26. *For all $x, y \in A_{\mathsf{BB}}$ and $\mathbf{u}, \mathbf{v} \in A_{\mathsf{BB}}^*$ we have*

$$\mathbf{u}x \star \mathbf{v}y := (\mathbf{u} \star \mathbf{v}y)x + (\mathbf{u}x \star \mathbf{v})y + (\mathbf{u} \star \mathbf{v})(x \diamond y).$$

Proof. This is straight-forward. See Exercise 12.18. □

It is clear that every mono-bracket $[\mathbf{s}]$ is equal to the bi-bracket $\left[\begin{smallmatrix}\mathbf{s}\\ \mathbf{1}\end{smallmatrix}\right]$. The following two results are similar to those for MB so we leave the proofs to the interested readers.

Proposition 12.8.27. *The $\mathbb{Q}$-algebra $(\mathbb{Q}\langle A_{\mathsf{BB}}\rangle, \star)$ is commutative and associative.*

Proposition 12.8.28. *The $\mathbb{Q}$-linear map*

$$[\,\cdot\,]\colon \mathbb{Q}\langle A_{\mathsf{BB}}\rangle \longrightarrow \mathsf{BB}$$
$$z_{s_1,r_1}\dots z_{s_l,r_l} \longmapsto \begin{bmatrix} s_1,\dots,s_l \\ r_1,\dots,r_l \end{bmatrix}$$

satisfies $[w \star v] = [w]\cdot[v]$, so that it is a $\mathbb{Q}$-algebra homomorphism.

To define the shuffle product, we need define a new class of objects which are closely related to bi-brackets.

Definition 12.8.29. Let $\mathbf{s} = (s_1,\dots,s_l), \mathbf{r} = (r_1,\dots,r_l) \in \mathbb{N}^l$. The *multiple q-zeta brackets* are defined by

$$\mathfrak{Z}_q\begin{bmatrix}\mathbf{s}\\ \mathbf{r}\end{bmatrix} := \sum_{\substack{m_1,\dots,m_l>0\\ d_1,\dots,d_l>0}} \prod_{j=1}^{l} \frac{m_j^{r_j-1} d_j^{s_j-1} q^{(m_j+\cdots+m_l)d_j}}{(r_j-1)!(s_j-1)!}$$
$$= \sum_{m_1,\dots,m_l>0} \prod_{j=1}^{l} \frac{m_j^{r_j-1} P_{s_j-1}(q^{m_j+\cdots+m_l})}{(r_j-1)!(s_j-1)!(1-q^{m_j+\cdots+m_l})^{s_j}}. \tag{12.70}$$

Clearly we have

$$\begin{bmatrix} s_1 \\ r_1 \end{bmatrix} = \mathfrak{Z}_q\begin{bmatrix} s_1 \\ r_1 \end{bmatrix} \qquad \text{and} \qquad [\mathbf{s}] = \begin{bmatrix}\mathbf{s}\\ \mathbf{1}\end{bmatrix} = \mathfrak{Z}_q\begin{bmatrix}\mathbf{s}\\ \mathbf{1}\end{bmatrix}.$$

In general, every bi-bracket can be expressed using the multiple q-zeta brackets, and vice versa.

Proposition 12.8.30. *Let $\mathbf{s} = (s_1,\dots,s_l), \mathbf{r} = (r_1,\dots,r_l) \in \mathbb{N}^l$. Then we have*

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

and

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

Proof. The equations follow easily from the identities

$$\frac{(m+n)^{r_1-1}}{(r_1-1)!}\,\frac{n^{r_2-1}}{(r_2-1)!}=\sum_{j=1}^{r_1+r_2-1}\binom{j-1}{r_2-1}\frac{m^{r_1+r_2-j-1}}{(r_1+r_2-j-1)!}\,\frac{n^{j-1}}{(j-1)!}$$

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 leave the details to the interested reader. □

The next result provides a kind of duality for bi-brackets which is quite similar to Thm. 12.2.3 in nature.

Proposition 12.8.31. *Let* $\mathbf{s}=(s_1,\ldots,s_l),\mathbf{r}=(r_1,\ldots,r_l)\in\mathbb{N}^l$. *Then we have*

$$\mathfrak{Z}_q\begin{bmatrix}\mathbf{s}\\ \mathbf{r}\end{bmatrix}=\mathfrak{Z}_q\begin{bmatrix}\overleftarrow{\mathbf{r}}\\ \overleftarrow{\mathbf{s}}\end{bmatrix}.$$

Proof. This follows from the identity

$$\sum_{i=1}^{l}d_i\sum_{j=i}^{l}m_j=\sum_{i=1}^{l}d'_i\sum_{j=i}^{l}m'_j$$

where $d'_i=m_{l+1-i}$ and $m'_j=d_{l+1-j}$. □

If we define a $\mathbb{Q}$-linear map

$$\begin{aligned}\mathfrak{Z}_q\colon\mathbb{Q}\langle A_{\mathsf{BB}}\rangle&\longrightarrow\mathsf{BB}\\ z_{s_1,r_1}\ldots z_{s_l,r_l}&\longmapsto\mathfrak{Z}_q\begin{bmatrix}s_1,\ldots,s_l\\ r_1,\ldots,r_l\end{bmatrix}\end{aligned}$$

then the duality can be also written as

$$\mathfrak{Z}_q[\mathbf{w}]=\mathfrak{Z}_q[\tau\mathbf{w}]\quad\forall\mathbf{w}\in A^*_{\mathsf{BB}},$$

where τ is the anti-automorphism on $\mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ such that $\tau(z_{s_1,r_1}\ldots z_{s_l,r_l})=z_{r_l,s_l}\ldots z_{r_1,s_1}$. Moreover, Prop. 12.8.30 provides an isomorphism φ of $\mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ so that

$$[\mathbf{w}]=\mathfrak{Z}_q[\varphi\mathbf{w}]\quad\text{and}\quad\mathfrak{Z}_q[\mathbf{w}]=[\varphi^{-1}\mathbf{w}]\qquad\forall\mathbf{w}\in A^*_{\mathsf{BB}}.$$

However, φ acts trivially on A^*_{MB} which is regarded as a subset of A^*_{BB} by identifying z_s with $z_{s,1}$.

Definition 12.8.32. For all $\mathbf{u}, \mathbf{v} \in \mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ we define the shuffle product

$$\mathbf{u} \sqcup\!\sqcup \mathbf{v} := \varphi^{-1}\tau(\tau\varphi\mathbf{u} \mathbin{\text{☆}} \tau\varphi\mathbf{v}). \tag{12.71}$$

Theorem 12.8.33. *For all* $\mathbf{u}, \mathbf{v} \in \mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ *we have*

$$[\mathbf{u} \sqcup\!\sqcup \mathbf{v}] = [\mathbf{u}] \cdot [\mathbf{v}].$$

Hence $[-]$ *provides a homomorphism of the* $\mathbb{Q}$*-algebras* $(\mathbb{Q}\langle Z\rangle, \sqcup\!\sqcup) \longrightarrow (\mathsf{BB}, \cdot)$.

Proof. This follows quickly from Prop. 12.8.28 and Prop. 12.8.31. □

To see that $\sqcup\!\sqcup$ corresponds to the shuffle product in MZV setting when $q \to 1^-$ we define an alphabet of three letters $A = \{a, b, c\}$ and regard $A_{\mathsf{BB}} \subset A^*$ by setting $z_{s,r} = a^{s-1}bc^{r-1}$ for all $r, s \in \mathbb{N}$. Clearly $A_{\mathsf{BB}}*$ can be identified with the subset of $\mathcal{A} = A^*$ of words not starting with c, not ending with a and not containing the subword ac. The action of τ on $\mathcal{A}$ becomes the anti-automorphism that exchanges a and c but fixes b.

Theorem 12.8.34. *Modulo the lower weights, we have*

$$x\mathbf{u} \sqcup\!\sqcup y\mathbf{v} \equiv x(\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) + y(x\mathbf{u} \sqcup\!\sqcup \mathbf{v}) \tag{12.72}$$

for arbitrary words $\mathbf{u}, \mathbf{v} \in \mathbb{Q} + \mathbb{Q}\langle a, b\rangle b$ *and* $x, y \in \{a, b\}$.

Proof. By Prop. 12.8.26, modulo lower weights, we have

$$\begin{aligned}
\mathbf{u}z_{s_1,r_1} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2} &\equiv (\mathbf{u} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2})z_{s_1,r_1} + (\mathbf{u}z_{s_1,r_1} \mathbin{\text{☆}} \mathbf{v})z_{s_2,r_2},\\
\mathbf{u}z_{s_1,r_1+1} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2} &\equiv (\mathbf{u}z_{s_1,r_1} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2})c + (\mathbf{u}z_{s_1,r_1+1} \mathbin{\text{☆}} \mathbf{v})z_{s_2,r_2},\\
\mathbf{u}z_{s_1,r_1+1} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2+1} &\equiv (\mathbf{u}z_{s_1,r_1} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2+1})c + (\mathbf{u}z_{s_1,r_1+1} \mathbin{\text{☆}} \mathbf{v}z_{s_2,r_2})c.
\end{aligned}$$

Hence

$$\mathbf{u}x \mathbin{\text{☆}} \mathbf{u}y \equiv (\mathbf{u} \mathbin{\text{☆}} \mathbf{v}y)x + (\mathbf{u}x \mathbin{\text{☆}} \mathbf{v})y \tag{12.73}$$

for any $\mathbf{u}, \mathbf{v} \in \mathbb{Q} + \mathbb{Q}\langle A_{\mathsf{BB}}\rangle$ and $x, y \in A_{\mathsf{BB}} \cup \{c\}$.

Since φ acts trivially on $\mathbb{Q}\langle A_{\mathsf{MB}}\rangle$, applying $\tau\varphi$ to both sides of Eq. (12.71) we get

$$\tau(\mathbf{u} \sqcup\!\sqcup \mathbf{v}) \equiv \tau\mathbf{u} \mathbin{\text{☆}} \tau\mathbf{v} \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{Q}\langle A_{\mathsf{MB}}\rangle.$$

By Eq. (12.73) we have

$$\begin{aligned}
\tau(x\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) &\equiv \tau(x\mathbf{u}) \mathbin{\text{☆}} \tau(y\mathbf{v}) \equiv (\tau\mathbf{u})\tau x \mathbin{\text{☆}} (\tau\mathbf{v})\tau y\\
&\equiv (\tau\mathbf{u} \mathbin{\text{☆}} (\tau\mathbf{v})\tau y)\tau x + ((\tau\mathbf{u})\tau x \mathbin{\text{☆}} \tau\mathbf{v})\tau y\\
&\equiv (\tau\mathbf{u} \mathbin{\text{☆}} \tau(y\mathbf{v}))\tau x + (\tau(x\mathbf{u}) \mathbin{\text{☆}} \tau\mathbf{v})\tau y\\
&\equiv (\tau(\mathbf{u} \sqcup\!\sqcup y\mathbf{v}))\tau x + (\tau(x\mathbf{u} \sqcup\!\sqcup \mathbf{v}))\tau y\\
&\equiv \tau(x(\mathbf{u} \sqcup\!\sqcup y\mathbf{v}) + y(x\mathbf{u} \sqcup\!\sqcup \mathbf{v})),
\end{aligned}$$

which implies the desired result. □

12.9 Some applications of MZ*V binomial identities

For the sake of convenience, we define more generally the q-analog of Euler (star) sums. For $\mathbf{s} = (s_1, s_2, \ldots, s_d) \in \mathbb{D}^d$ (see Eq. (7.1.1) for the definition of $\mathbb{D}$) we set

$$\zeta_q^{\mathbf{t}}[\mathbf{s}] := \sum_{k_1 > \cdots > k_d \geq 1} \prod_{j=1}^{d} \frac{\operatorname{sgn}(s_j)^{k_j} q^{k_j t_j}}{[k_j]^{|s_j|}}$$

and

$$\zeta_q^{\star,\mathbf{t}}[\mathbf{s}] := \sum_{k_1 \geq \cdots \geq k_d \geq 1} \prod_{j=1}^{d} \frac{\operatorname{sgn}(s_j)^{k_j} q^{k_j t_j}}{[k_j]^{|s_j|}}.$$

As we will only use $\zeta_q^{\star,\mathbf{1}}[\mathbf{s}]$ for the star-type we will simply write $\zeta^\star[\mathbf{s}] = \zeta_q^{\star,\mathbf{1}}[\mathbf{s}]$ in the rest of this section.

12.9.1 *Mollified companion of q-MZVs and q-MZ*Vs*

Adopting the notation from Chap. 9 we define the *mollified companion* of $\zeta_q^{\mathbf{1}}[\mathbf{s}]$ associated with the admissible triple of mollifiers $[\mathbf{s};\mathbf{t};\mathbf{r}]$ to be

$$\begin{aligned}\mathfrak{Z}[\mathbf{s};\mathbf{t};\mathbf{r}] &:= \mathfrak{Z}[s_1,\ldots,s_m;t_1,\ldots,t_m;r_1,\ldots,r_m] \\ &= \sum_{k_1>\cdots>k_m\geq 1} \prod_{j=1}^{m} \frac{\operatorname{sgn}(s_j)^{k_j} q^{t_j k_j + Q(r_j,k_j)}(1+q^{k_j})}{[k_j]^{|s_j|}}.\end{aligned}$$

If $m = 0$, we put $\zeta^\star[\emptyset] = \mathfrak{Z}[\emptyset;\emptyset;\emptyset] = 1$. Throughout the paper the triples of mollifiers $[\mathbf{s};\mathbf{t};\mathbf{r}]$ are chosen in such a way that the above multiple series always converges. Notice that in [271] the mollified companions of $\zeta[\mathbf{s}]$ are defined similarly.

Although our primary goal is to prove q-MZSV identities, throughout the chapter we will always work with binomial identities for q-MHS first. In order to obtain the corresponding q-MZSV identities we need the next result.

Lemma 12.9.1. *Let $0 < q < 1$, $c, c_1, c_2 \in \mathbb{R}$, $c > 0$, and let R_k be a sequence of real numbers satisfying $|R_k| < k^{c_1} q^{c_2 k}$ for all $k = 1, 2, \ldots$. Then*

$$\lim_{n\to\infty} \sum_{k=1}^{n} q^{ck^2} \left(1 - \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}}\right) R_k = 0.$$

Proof. For $1 \leq k < n/2$ and n sufficiently large,

$$1-\frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} = 1-\frac{(q;q)_n^2}{(q;q)_{n-k}(q;q)_{n+k}} = 1-\frac{(1-q^{n-k+1})(1-q^{n-k+2})\cdots(1-q^n)}{(1-q^{n+1})(1-q^{n+2})\cdots(1-q^{n+k})} = O(q^{c_3 n}),$$

where c_3 is some positive constant independent of n. Therefore, we have

$$\left|\sum_{k=1}^{n/2} q^{ck^2}\left(1-\frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}}\right) R_k\right| < O(q^{c_3 n})\sum_{k=1}^{n/2} q^{ck^2}|R_k| < O(q^{c_3 n})\sum_{k=1}^{n/2} k^{c_2} q^{ck^2+c_1 k}. \tag{12.74}$$

On the other side, for $n/2 \leq k \leq n$, we can apply the trivial inequality

$$0 < 1-\frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}} = 1-\frac{(1-q^{n-k+1})(1-q^{n-k+2})\cdots(1-q^n)}{(1-q^{n+1})(1-q^{n+2})\cdots(1-q^{n+k})} < 1$$

to obtain

$$\left|\sum_{k=n/2}^{n} q^{ck^2}\left(1-\frac{\begin{bmatrix} n \\ k \end{bmatrix}}{\begin{bmatrix} n+k \\ k \end{bmatrix}}\right) R_k\right| < q^{cn^2/4}\sum_{k=n/2}^{n}|R_k| < n^{c_4} q^{cn^2/4+c_5 n}, \tag{12.75}$$

where c_4, c_5 are some real constants independent of n. We now finish the proof by taking $n \to \infty$ using Eq. (12.74) and Eq. (12.75). □

Using the above lemma we may take limit $n \to \infty$ in Thm. 9.2.1 and Thm. 9.2.2 to obtain identities for q-MZSV.

Theorem 12.9.2. *Let $n, c \in \mathbb{N}$ and $c \geq 2$. Then*

$$\zeta^\star[c] = -\sum_{\mathbf{p}\in\{(\overline{2})\circ 1^{\circ(c-2)};1\circ 0^{\circ(c-2)};1\circ\theta^{\circ(c-2)}\}} \mathfrak{Z}[\mathbf{p}]. \tag{12.76}$$

Theorem 12.9.3. *Keep the same notation as in Thm.* 9.2.2. *Then for any integers $c \geq 3$ we have*

$$\begin{aligned}
\zeta^\star[2,\mathbf{s}] =& \delta(\mathbf{s})\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi_2} \mathfrak{Z}[\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}],\\
\zeta^\star[1,\mathbf{s}] =& \delta(\mathbf{s})\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi_1} \mathfrak{Z}[\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}],\\
\zeta^\star[c,\mathbf{s}] =& \delta(\mathbf{s})\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi_c} \mathfrak{Z}[\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}].
\end{aligned}$$

Using the same notation as on page 310 we get the following general result concerning MZ*Vs.

Theorem 12.9.4. *For all* $\mathbf{s}=(s_1,\ldots,s_\ell)\in\mathbb{N}^\ell$ *let* $\underline{\mathbf{m}}(\mathbf{s})_1=(\mathbf{m};\widetilde{\mathbf{m}};\widetilde{\widetilde{\mathbf{m}}})$. *If* $s_1>1$ *then we have*

$$\zeta^\star[\mathbf{s}]=-\sum_{\underline{\mathbf{p}}\in\{\mathbf{m}^\circ;\widetilde{\mathbf{m}}^\circ;\widetilde{\widetilde{\mathbf{m}}}\circ\}}\mathfrak{Z}[\underline{\mathbf{p}}],\quad \textit{where}\quad \delta(\mathbf{s})=\begin{cases}1, & \textit{if } s_\ell=1;\\ -1, & \textit{if } s_\ell\geq 2.\end{cases}$$

By letting $q\to 1$ in Thm. 12.9.4 we can finally prove Thm. 10.8.1 which gives the corresponding general rule for the classical MZSV in Chap. 10. Of course, to guarantee convergence we need to restrict $a\geq 1$ there.

12.9.2 *Applications to MZVs and MZ*Vs*

By applying Lemma 12.9.1 to Cor. 9.3.3 we immediately get

Corollary 12.9.5. *With the same notation as in Cor.* 9.3.3*, we have*

$$\zeta^\star[\mathbf{s}]=-\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi(\mathbf{s})}\mathfrak{Z}[\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}].$$

In particular, if $c_1=c_2=\ldots=c_\ell=3$*, then*

$$\zeta^\star[\mathbf{s}]=\sum_{(\mathbf{p},\widetilde{\mathbf{p}})}\sum_{k_1>\cdots>k_m\geq 1}(-1)^{k_1-1}q^{\binom{k_1}{2}}\prod_{j=1}^{m}\frac{q^{\widetilde{p}_jk_j}(1+q^{k_j})}{[k_j]^{p_j}},$$

where $(\mathbf{p},\widetilde{\mathbf{p}})$ *runs through* $\Big\{(2a_1{+}2)\circ\Big\lfloor\circ\Big\rfloor_{j=2}^{\ell}(2a_j{+}3)\circ(2a_{\ell+1}{+}1),\ \Big\lfloor\circ\Big\rfloor_{j=1}^{\ell}(a_j{+}1)\circ(a_{\ell+1})\Big\}$.

Remark 12.9.6. When $q\to 1$ one can recover all the MZSV identities contained in [385].

The following result is a corollary of Exercise 9.3.

Corollary 12.9.7. *With the same notation as in Exercise* 9.3 *we have*

$$\zeta^\star[\mathbf{s}]=\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi(\mathbf{s})}\mathfrak{Z}[\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}].$$

For example, taking $\ell=1$ and $c_1=3$, we get (cf. [608, (25)] and the identity after it)

$$\zeta^\star[\{2\}^b,3,\{2\}^a,1]=\mathfrak{Z}[2a{+}2b{+}4;a{+}b{+}2;2]{+}\mathfrak{Z}[\overline{2b+2},\overline{2a+2};b{+}1,a{+}1;1,1]$$

and

$$\begin{aligned}\zeta^\star[\{2\}^{a_0},1,\{2\}^b,3,\{2\}^{a_1},1] &= \mathfrak{Z}[2(a_0+b+a_1)+5;a_0+b+a_1+3;2]\\ &+\mathfrak{Z}[2a_0+1,2a_1+2b+4;a_0+1,a_1+b+2;2,0]\\ &+\mathfrak{Z}[\overline{2a_0+2b+3},\overline{2a_1+2};a_0+b+2,a_1+1;1,1]\\ &+\mathfrak{Z}[2a_0+1,\overline{2b+2},\overline{2a_1+2};a_0+1,b+1,a_1+1;2,-1,1].\end{aligned} \tag{12.77}$$

We can also get the following identity which is the q-analog of Eq. (10.34).

Corollary 12.9.8. *Let a,b be two nonnegative integers. Then*

$$\begin{aligned}&\zeta^\star[\{2\}^a,3,\{2\}^b,1]+\zeta^\star[\{2\}^b,3,\{2\}^a,1]\\ &\quad=\zeta^\star[\{2\}^{a+1}]\zeta^\star[\{2\}^{b+1}]+(1-q)\mathfrak{Z}[2a+2b+3;a+b+2;2].\end{aligned}$$

Proof. By taking $n\to\infty$ in Eq. (9.16) and using Lemma 12.9.1 we get

$$\zeta^\star[\{2\}^a]=\mathfrak{Z}[2a;a;1]=\sum_{k=1}^{\infty}(-1)^{k-1}\frac{q^{\binom{k}{2}+ak}(1+q^k)}{[k]^{2a}}.$$

Thus

$$\begin{aligned}&\zeta^\star[\{2\}^a,3,\{2\}^b,1]+\zeta^\star[\{2\}^b,3,\{2\}^a,1]-\zeta^\star[\{2\}^{a+1}]\zeta^\star[\{2\}^{b+1}]\\ =&\sum_{k=1}^{\infty}\frac{q^{(a+b+2)k+Q(2,k)}\big(2(1+q^k)-(1+q^k)^2\big)}{[k]^{2a+2b+4}}\\ =&\sum_{k=1}^{\infty}\frac{q^{(a+b+2)k+Q(2,k)}(1+q^k)(1-q^k)}{[k]^{2a+2b+4}}\\ =&(1-q)\mathfrak{Z}[2a+2b+3;a+b+2;2]\end{aligned}$$

as desired. □

By taking $n\to\infty$ in Exercise 9.4 we have

Corollary 12.9.9. *Let notation be the same as in Exercise* 9.4. *Then*

$$\zeta^\star[\mathbf{s}]=-\sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}})\in\Pi(\mathbf{s})}\mathfrak{Z}[\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}].$$

For example, taking $\ell=1$ and $c_1=3$, we get in case (**2**-*c*-**2**-1-**2**)

$$\begin{aligned}&\zeta^\star[\{2\}^b,3,\{2\}^{a_1},1,\{2\}^{a_2}]\\ =&-\mathfrak{Z}[\overline{2a_1+2b+2a_2+4};a_1+b+a_2+2;1]\\ &-\mathfrak{Z}[\overline{2b+2},2a_1+2a_2+2;b+1,a_1+a_2+1;1,\theta]\\ &-\mathfrak{Z}[2a_1+2b+4,\overline{2a_2};a_1+b+2,a_2;2,-1]\\ &-\mathfrak{Z}[\overline{2b+2},\overline{2a_1+2},\overline{2a_2};b+1,a_1+1,a_2;1,1,-1].\end{aligned}$$

By taking $q \to 1$ this yields the identity immediately after [608, Thm. 6.2].

As a non-trivial example of Thm. 12.9.3 we may attach a string of type $(2^a, 1)$ to the front of the already treated type $(\{2\}^b, 1, \{2\}^c, 3, \{2\}^d, 1)$ given by Eq. (12.77) and get the following q-MZSV identity: for any nonnegative integers a, b, c, d

$$\begin{aligned}
&\zeta^\star[\{2\}^a, 1, \{2\}^b, 1, \{2\}^c, 3, \{2\}^d, 1] \\
&\quad = \mathfrak{Z}[2a+2b+2c+2d+6; a+b+c+d+4; 2] \\
&\quad + \mathfrak{Z}[2a+1, 2b+2c+2d+5; a+1, b+c+d+3; 2, 0] \\
&\quad + \mathfrak{Z}[2a+2b+2, 2c+2d+4; a+b+2, c+d+2; 2, 0] \\
&\quad + \mathfrak{Z}[\overline{2a+2b+2c+4}, \overline{2d+2}; a+b+c+3, d+1; 1, 1] \\
&\quad + \mathfrak{Z}[2a+1, 2b+1, 2c+2d+4; a+1, b+1, c+d+2; 2, 0, 0] \\
&\quad + \mathfrak{Z}[2a+1, \overline{2b+2c+3}, \overline{2d+2}; a+1, b+c+2, d+1; 2, -1, 1] \\
&\quad + \mathfrak{Z}[2a+2b+2, \overline{2c+2}, \overline{2d+2}; a+b+2, c+1, d+1; 2, -1, 1] \\
&\quad + \mathfrak{Z}[2a+1, 2b+1, \overline{2c+2}, \overline{2d+2}; a+1, b+1, c+1, d+1; 2, 0, -1, 1].
\end{aligned} \tag{12.78}$$

By taking $q \to 1$ in Eq. (12.78) we discover the classical MZSV identity:

$$\begin{aligned}
&\zeta^\star(\{2\}^a, 1, \{2\}^b, 1, \{2\}^c, 3, \{2\}^d, 1) \\
&= 2\zeta(2a+2b+2c+2d+6) + 4\zeta(2a+1, 2b+2c+2d+5) \\
&+ 4\zeta(2a+2b+2, 2c+2d+4) + 4\zeta(\overline{2a+2b+2c+4}, \overline{2d+2}) \\
&+ 8\zeta(2a+1, 2b+1, 2c+2d+4) + 8\zeta(2a+1, \overline{2b+2c+3}, \overline{2d+2}) \\
&+ 8\zeta(2a+2b+2, \overline{2c+2}, \overline{2d+2}) + 16\zeta(2a+1, 2b+1, \overline{2c+2}, \overline{2d+2}).
\end{aligned}$$

By applying Lemma 12.9.1 to Cor. 9.3.3 we immediately get

Theorem 12.9.10. *Suppose* $\ell \in \mathbb{N}_0$. *Let* $\mathbf{s} = \left(\left\{\{2\}^{a_j}, c_j\right\}_{j=1}^{\ell}, \{2\}^{a_{\ell+1}}\right)$ *with* $a_j, c_j \in \mathbb{N}_0$ *and* $c_j \geq 3$ *for all* $j \geq 1$. *Then*

$$\zeta^\star[\mathbf{s}] = - \sum_{(\mathbf{p};\widetilde{\mathbf{p}};\widetilde{\widetilde{\mathbf{p}}}) \in \Pi(\mathbf{s})} \mathfrak{Z}[\mathbf{p}; \widetilde{\mathbf{p}}; \widetilde{\widetilde{\mathbf{p}}}], \tag{12.79}$$

where

$$\begin{aligned}
\Pi(\mathbf{s}) \Big\{ &(\overline{2a_1+2}) \circ 1^{\circ(c_1-3)} \circ \Big\lfloor \circ \Big\rfloor_{j=2}^{\ell} \left((2a_j+3) \circ 1^{\circ(c_j-3)}\right) \circ (2a_{\ell+1}+1); \\
&\Big\lfloor \circ \Big\rfloor_{j=1}^{\ell} \left((a_j+1) \circ 0^{\circ(c_j-3)}\right) \circ a_{\ell+1};\ 1 \circ \Big\lfloor \circ \Big\rfloor_{j=1}^{\ell} \theta^{\circ(c_j-2)} \Big\}.
\end{aligned}$$

12.10 Historical notes

There is a long history of studying q-analogs of number theoretical objects in general. Concerning the multiple zeta values, Kaneko *et al.* first studied in detail the analytic properties and some special values of the type I q-analog of Riemann zeta functions in [321] in early 2000, which was generalized to multiple zeta setting by the author [594] (first version was posted on arxiv in April 2003). In [594], we considered both the analytical and algebraic properties of type I q-analog of the multiple zeta functions. We then defined their stuffle and shuffle products using Jackson q-iterated integrals. As the shuffle is quite complicated only very few DBSF could be found. In 2004, Bradley independently studied type I q-MZVs in [82] by concentrating on the generalization of classical identities of MZVs to this q-analog setting, and later, Okuda and Takeyama also studied some of the relations among this type of q-MZVs in [444].

Before type I q-MZVs, slightly different definitions of q-MZVs were already presented in [481, 620]. Additionally, it is not hard to see that Schlesinger's version diverges when $|q| < 1$ but can converge if $|q| > 1$. In fact, for $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{Z}^d$

$$\mathfrak{z}_{q^{-1}}^{(0,\ldots,0)}[\mathbf{s}] = (-1)^{|\mathbf{s}|}\mathfrak{z}_q^{(\mathbf{s})}[\mathbf{s}] = (-1)^{|\mathbf{s}|}\mathfrak{z}_q^{\mathrm{II}}[\mathbf{s}].$$

So it suffices to consider type II first defined in [620] in order to understand Schlesinger's q-MZVs.

The analytic properties of type I q-MZVs are studied in [594] which also includes the stuffle relation and the shifting principle. However, the shuffle structure was not defined until Takeyama's work in Ref. [512]. Unfortunately, his approach to the shuffle relations relies on some auxiliary multiple polylogarithm functions and consequently it is very hard to see why these relations should hold.

The situation looks much better with the appearance of the joint work [116] by Castillo Medina *et al.* who generalized Chen's iterated integrals to iterated Jackson q-integrals to study type III q-MZVs by using Rota-Baxter algebra techniques, which is modified and generalized in section 12.2.

In 1958, influenced by Chen's study of iterated integrals, Ree first considered algebras in which the product is expressed in terms of shuffles in [468]. Later, G. Baxter constructed algebras with an operator $\mathcal{P}$ satisfying Eq. (12.5) with $\lambda = 1$. In general, Eq. (12.5) was defined and popularized by Rota and his school. See [472–474]. Linear operators satisfying Eq. (12.5) in the context of Lie algebras were introduced in the

1980s and were related to the solutions of the classical YangBaxter equation (after C. Yang and R. Baxter). Along this line, Connes and Kreimer applied the Hopf algebra technique and developed a renormalization theory in perturbative quantum field theory (see [140, 141]) which led to some related work on renormalizations of multiple zeta values in [258] and their q-analogs in [597]. Note that Jackson q-integral was first defined in [305] is used. However, his difference operator in [304] is different from ours in this chapter.

In [512], Takeyama also noticed the set of type I q-MZVs needs to enlarged to type $\tilde{\mathrm{I}}$ q-MZVs which are a kind of "regularized" q-MZVs in the sense that one needs to consider some convergent versions of q-MZVs when $s_1 = 1$ by modifying the auxiliary variables of $\mathbf{t}$. But for these type $\tilde{\mathrm{I}}$ q-MZVs themselves, DBSFs are insufficient to provide all the $\mathbb{Q}$-linear relations and a certain "Resummation Identity" defined by Takeyama is required. In this book, we adopt the term "duality" due to its similarity to the duality relations of the ordinary MZVs.

Concerning the individual results, Thm. 12.2.11 is parallel to [116, Prop. 9] and includes [512, Thm. 1] as a special type $\tilde{\mathrm{I}}$ case. Prop. 12.2.19 is [116, Thm. 7]. Cor. 12.2.20 generalizes [512, Prop. 1]. The case for type $\tilde{\mathrm{I}}$ of Prop. 12.2.21 is proved by Takeyama as [512, Prop. 2.4]. Note that Thm. 12.2.22 generalizes [512, Thm. 2] but it does not contain [116, Thm. 7] since our word representation of type $\mathbb{III}$ q-MZVs in this book is different from that of [116]. Thm. 12.3.1 was first proved by Takeyama as [512, Thm. 4].

Turning to the more complicated q-MZVs, Okounkov stated Conjecture 12.6.1 in [442] after he relates his version of q-MZVs to some q-series invariants of the Hilbert schemes parameterizing 0-dimensional subschemes of finite lengths on a nonsingular quasi-projective surface.

All the results in Sec. 12.9 are consequences of the corresponding results of q-MH*S obtained in Chap. 9. The key is Lemma 12.9.1 obtained in Ref. [271] which generalizes a similar result in the classical non-q setting appearing first as [385, Lemma 4.2].

Exercises

12.1. There are many interesting relations among type I q-MZVs. The following is the q-analog of the cyclic sum formula (cf. Thm. 5.2.3). Let

$\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$, $s_1 \geq 2$. Then

$$\sum_{i=1}^{d} \zeta_q^{\mathrm{I}}[s_i + 1, s_{i+1}, \dots, s_d, s_1, \dots, s_{i-1}] = \sum_{i=1}^{d} \sum_{j=0}^{s_i - 2} \zeta_q^{\mathrm{I}}[s_i - j, s_{i+1}, \dots, s_d, s_1, \dots, s_{i-1}, j+1].$$

12.2. The following result gives the q-analog of Ohno's Relations (cf. Thm. 5.4.9). Let $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$, $s_1 \geq 2$ and its dual $\mathbf{s}^* = (s'_1, \dots, s'_{d'})$ (see Eq. (5.1)). Then for any non-negative integer l

$$\sum_{\substack{e_1 + \cdots + e_d = l \\ e_i \geq 0}} \zeta_q^{\mathrm{I}}[s_1 + e_1, \dots, s_d + e_d] = \sum_{\substack{e'_1 + \cdots + e'_{d'} = l \\ e'_i \geq 0}} \zeta_q^{\mathrm{I}}[k'_1 + e'_1, \dots, k'_{d'} + e'_{d'}].$$

12.3. Prove the following q-analog of Euler's Decomposition Formula. For all $a, b \in \mathbb{N}$

$$\begin{aligned} \mathfrak{z}_q^{\mathrm{I}}[s]\mathfrak{z}_q^{\mathrm{I}}[t] &= \sum_{a=0}^{s-1} \sum_{b=0}^{s-1-a} \binom{a+t-1}{t-1} \binom{t-1}{b} \mathfrak{z}_q^{\mathrm{I}}[t+a, s-a-b] \\ &\quad + \sum_{a=0}^{t-1} \sum_{b=0}^{t-1-a} \binom{a+s-1}{s-1} \binom{s-1}{b} \mathfrak{z}_q^{\mathrm{I}}[s+a, t-a-b] \\ &\quad - \sum_{j=1}^{\min(s,t)} \frac{(s+t-j-1)!}{(s-j)!\,(t-j)!} \cdot \frac{1}{(j-1)!} \frac{1}{(1-q)^{s+t-j}} \varphi_q[s+t-j]. \end{aligned}$$

where φ_q is defined by

$$\varphi_q[n] = \sum_{k=1}^{\infty} (k-1) \frac{q^{(n-1)k}}{[k]^n}. \tag{12.80}$$

12.4. Verify Eq. (12.28), Eq. (12.29) and Eq. (12.30) by direct computation.

12.5. Equation Eq. (12.13) is commonly considered to be the q-analog for the classical integration by parts rule:

$$R(f)R(g) = R\big(fR(g)\big) + R\big(R(f)g\big), \tag{12.81}$$

which can be seen as dual to Leibniz rule for usual derivations. For any fixed constant a, show that Eq. (12.81) holds if we set $R(f)(t) = \int_a^t f(x)\,dx$.

12.6. Prove that $\lim_{q \to 1} J_q[f](t) = \int_0^t f(x)\,dx$.

12.7. Prove Prop. 12.2.2 by direct computation.

12.8. Prove Cor. 12.2.18.

12.9. Show that the shifting operator defined by Eq. (12.24) is injective.

12.10. Prove the following q-analog of Euler's Decomposition Formula. For all $a, b \in \mathbb{N}$

$$\begin{aligned}\mathfrak{z}_q^{\text{Ш}}[a]\mathfrak{z}_q^{\text{Ш}}[b] &= \sum_{l=0}^{a-1}\sum_{k=0}^{a-1-l}(-1)^k\binom{l+b-1}{b-1}\binom{b}{k}\mathfrak{z}_q^{\text{Ш}}[b+l, a-l-k] \\ &+ \sum_{l=0}^{b-1}\sum_{k=0}^{\min(a,b-1-l)}(-1)^k\binom{l+a-1}{a-1}\binom{a}{k}\mathfrak{z}_q^{\text{Ш}}[a+l, b-l-k] \\ &- \sum_{k=1}^{a}\beta_{a-k}\mathfrak{z}_q^{\text{Ш}}[a+b-k] + \sum_{j=1}^{a-1}\alpha_{a+b-1-j}\,\mathfrak{D}\,\mathfrak{z}_q^{\text{Ш}}[a+b-1-j].\end{aligned}$$

where $\mathfrak{D} := q\frac{\mathrm{d}}{\mathrm{d}q}$ and

$$\begin{aligned}\alpha_k &= \sum_{j=b}^{k}\frac{(-1)^{a+b-j}}{1-j}\binom{j-1}{j-b, j-a, a+b-j-1}, \\ \beta_j &= (-1)^{a-j}\binom{j+b-1}{j, j+b-a, a-j-1}.\end{aligned}$$

12.11. Prove the commutativity of the q-shuffle product by showing that $\mathbf{u} \sqcup\!\sqcup \mathbf{v} = \mathbf{v} \sqcup\!\sqcup \mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in \widetilde{Y}^*$ using induction on the sum $|\mathbf{u}| + |\mathbf{v}|$.

12.12. Show that $(\mathfrak{A}_{\mathrm{G}}^0, *_{\mathrm{G}})$ is a commutative and associative algebra.

12.13. Define another type of q-MZVs by

$$\mathfrak{z}_q^{(s_1-1,s_2,\ldots,s_d)}[s_1,\ldots,s_d], \quad s_1 \geq 1, s_2, \ldots, s_d \geq 0.$$

Define the shuffle and stuffle structures of these q-MZVs by mimicking the treatment of type I q-MZVs.

12.14. Complete the proof of Prop. 12.8.11.

12.15. For s_1, s_2 with $s_1 + s_2 > 2$ and $s = s_1 + s_2 - 2$ show that

$$\begin{aligned}\binom{s}{s_1-1}\frac{\mathfrak{D}[s]}{s} &= [s_1]\cdot[s_2] + \binom{s}{s_1-1}[s+1] \\ &\quad - \sum_{a+b=s+2}\left(\binom{a-1}{s_1-1} + \binom{a-1}{s_2-1}\right)[a,b].\end{aligned}$$

12.16. For $s_1, s_2 \in \mathbb{N}$ show that

$$\begin{aligned}\mathfrak{D}[s_1,s_2] &= [2]\cdot[s_1,s_2] - s_1[s_1+1,s_2,1] - s_2[s_1,s_2+1,1] - [s_1,s_2,2] \\ &\quad + 2s_1[s_1+1,s_2] + s_2[s_1,s_2+1] - \Bigg(\sum_{a+b=s_1+2}(a-1)[a,b,s_2] \\ &\quad + \sum_{a+b=s_2+1} s_1[s_1+1,a,b] + \sum_{a+b=s_2+2}(a-1)[s_1,a,b]\Bigg).\end{aligned}$$

12.17. Show that

$$\begin{aligned}\mathfrak{D}[1] &= [3] + \frac{1}{2}[2] - [2,1], \\ \mathfrak{D}[2] &= [4] + 2[3] - \frac{1}{6}[2] - 4[3,1], \\ \mathfrak{D}[2] &= 2[4] + [3] + \frac{1}{6}[2] - 2[2,2] - 2[3,1], \\ \mathfrak{D}[1,1] &= [3,1] + \frac{3}{2}[2,1] + \frac{1}{2}[1,2] + [1,3] - 2[2,1,1] - [1,2,1], \\ \mathfrak{D}[1,2] &= -\frac{1}{6}[1,2] + 2[1,3] + [1,4] + \frac{3}{2}[2,2] + [3,2] \\ &\quad - 4[1,3,1] - [2,1,2] - 2[2,2,1].\end{aligned}$$

12.18. Prove Prop. 12.8.26 by induction on the sum of the weights of the words involved. In fact, if we change the definition of bi-brackets by using the increasing order on the indices in Eq. (12.68) then this is equivalent to the definition in Eq. (12.69).

Chapter 13

Colored multiple zeta values

In Chap. 2 we have studied mainly the analytic properties of the the multiple polylog functions. In the following two chapters we turn to the algebraic side of the special values of these functions.

13.1 Definitions

Recall from Definition 2.3.1 that for any positive integers $n_1, \ldots, n_d$, the multiple polylog of complex variables is defined as follows:

$$\mathrm{Li}_{n_1,\ldots,n_d}(z_1,\ldots,z_d) = \sum_{k_1>\cdots>k_d>0} \frac{z_1^{k_1}\cdots z_d^{k_d}}{k_1^{n_1}\cdots k_d^{n_d}}, \tag{13.1}$$

where $|z_1\cdots z_j|<1$ for $j=1,\ldots,d$. It is analytically continued to a multi-valued meromorphic function on $\mathbb{C}^d$ in section 2.4.

Definition 13.1.1. Fix an Nth root of unity $\mu=\mu_N := \exp(2\pi\sqrt{-1}/N)$. A *colored multiple zeta value (CMZV) of level* N is a number of the form

$$L_N(s_1,\ldots,s_d|i_1,\ldots,i_d) := \mathrm{Li}_{s_1,\ldots,s_d}(\mu^{i_1},\ldots,\mu^{i_d}). \tag{13.2}$$

We will always identify $(i_1,\ldots,i_d)$ with $(i_1,\ldots,i_d) \pmod N$. The Nth roots of unity are called the *colors*. As before, d is called the *depth* and $s_1+\cdots+s_d$ the *weight*.

It is easy to see from Eq. (13.1) that a CMZV converges if and only if $(s_1,\mu^{i_1})\neq(1,1)$. Clearly, all CMZVs of level N are automatically of level kN for every positive integer k. For example, when $i_1=\cdots=i_d=0$ or $N=1$ one gets the MZV $\zeta(s_1,\ldots,s_d)$. When $N=2$ one recovers the so-called Euler sums which will be treated in section 14.2 and whose partial sums have been studied in Chap. 7. We also have the following integral

expression by Eq. (2.8):

$$L_N(s_1,\dots,s_d|i_1,\dots,i_d)=\int_0^1\left(\frac{\mathrm{d}t}{t}\right)^{s_1-1}\frac{\mathrm{d}t}{a_1-t}\cdots\left(\frac{\mathrm{d}t}{t}\right)^{s_d-1}\frac{\mathrm{d}t}{a_d-t} \tag{13.3}$$

where $a_i=\prod_{j=1}^{i}\mu^{-i_j}$.

13.2 Dimension upper bound by Deligne and Goncharov

Similar to the MZV case standard conjectures in arithmetic algebraic geometry imply that $\mathbb{Q}$-linear relations among CMZVs can only exist between those of the same weight. Let $\mathsf{CMZV}(w,N)$ be the $\mathbb{Q}$-span of all the CMZVs of weight w and level N. Let $d(w,N)$ denote its dimension. In general, it is very difficult to determine $d(w,N)$ because any nontrivial lower bound would provide some nontrivial irrational/transcendental result which is related to a variant of Grothendieck's period conjecture. For example, one can show easily that $\mathsf{CMZV}(2,4)=\langle\log^2 2,\pi^2,\pi\sqrt{-1}\log 2,(K-1)\sqrt{-1}\rangle$, where $K=\sum_{n\geq 0}(-1)^n/(2n+1)^2$ is the Catalan's constant. From a variant of Grothendieck's period conjecture we know $d(2,4)=4$ (see [155]) but we don't have an unconditional proof yet. Namely, we cannot prove that the four number $\log^2 2,\pi^2,\pi\sqrt{-1}\log 2,(K-1)\sqrt{-1}$ are linearly independent over $\mathbb{Q}$. Consequently, a nontrivial lower bound of $d(w,N)$ is hard to come by.

On the other hand, one may obtain upper bounds of $d(w,N)$ by finding as many $\mathbb{Q}$-linear relations in $\mathsf{CMZV}(w,N)$ as possible. As in the cases of MZVs and the Euler sums the double shuffle relations will play important roles in revealing the relations among CMZVs. In such a relation if all the CMZVs involved are convergent then it is called a *finite double shuffle relation* (finite DBSF) similar to the case of MZVs. In general we need to regularize the divergent CMZVs to obtain the *regularized double shuffle relations* (regularized DBSFs). We shall recall this theory in Sec. 13.3.1.

The proof of the following important result of Deligne and Goncharov is far beyond the scope of this book, so we omit it here.

Theorem 13.2.1. *For any weight $w\geq 1$ and level $N\geq 1$, we have $d(w,N)\leq D(w,N)$ where $D(w,N)$ are defined by the formal power series*

$$1+\sum_{w=1}^{\infty}D(w,N)t^w=\begin{cases}(1-t^2-t^3)^{-1}, & \text{if } N=1;\\ (1-t-t^2)^{-1}, & \text{if } N=2;\\ (1-at+bt^2)^{-1}, & \text{if } N\geq 3,\end{cases} \tag{13.4}$$

where $a=a(N):=\varphi(N)/2+\nu(N)$, $b=b(N):=\nu(N)-1$, φ is the Euler's totient function and $\nu(N)$ is the number of distinct prime factors of N.

If $N > 2$ then we have

$$\sum_{w=1}^{\infty} D(w,N)t^w = at + (a^2-b)t^2 + (a^3-2ab)t^3 + (a^4-3a^2b+b^2)t^4 + \cdots$$

In particular, if p is a prime then for any positive integer n

$$D(w,p^n) = a(p^n)^w = \left(\frac{p^{n-1}(p-1)}{2}+1\right)^w. \tag{13.5}$$

We will compare the bound $D(w,N)$ of Deligne and Goncharov to the bound obtained by the standard relations to be defined next.

13.3 Standard relations

It turns out that in higher level cases, the DBSFs are far from enough to produce all the $\mathbb{Q}$-linear relations between CMZVs. In this section, we list some more relations that can be defined without too much difficulty.

Definition 13.3.1. We call a $\mathbb{Q}$-linear relation between CMZVs *standard*[1] if it can be produced by $\mathbb{Q}$-linear combinations of the following four families of relations: regularized DBSFs (Sec. 13.3.1), regularized distribution relations (regularized DISTs) (Sec. 13.3.4), weight one relations (Sec. 13.3.3), and lifted relations from the above three types (Sec. 13.3.5). Otherwise, it is called a *non-standard* relation.

It is commonly believed that all linear relations among MZVs (*i.e.*, level one CMZVs) are consequences of regularized DBSFs. But when level $N \geq 2$ we don't know exactly how to produce all linear relations among the CMZVs. When $N = 2$ (*i.e.*, for Euler sums) we don't know if the regularized DBSFs, regularized DISTs, the doubling relations and the generalized doubling relations (see Sec. 14.2.5) are independent or not.

13.3.1 *Regularized double shuffle relations*

Similar to the cases of MZVs, we first define some algebras which reflects the behavior of multiplications of CMZVs.

Set

$$a = \frac{\mathrm{d}t}{t}, \qquad b_i = \frac{\mu^i \mathrm{d}t}{1-\mu^i t} \quad \text{for } i = 0,1,\ldots,N-1.$$

For every positive integer $k \geq 1$ define the word of length k

$$\mathbf{z}_{k,i} := a^{k-1}b_i.$$

[1]This term was suggested by P. Deligne in a letter to Goncharov and Racinet dated Feb. 25, 2008.

Then it is straight-forward to verify using Eq. (13.3) that if $(s_1, \mu^{i_1}) \neq (1,1)$ then

$$L_N(s_1,\dots,s_n|i_1,i_2,\dots,i_n) = \int_0^1 \mathsf{z}_{s_1,i_1}\mathsf{z}_{s_2,i_1+i_2}\cdots\mathsf{z}_{s_n,i_1+i_2+\cdots+i_n}. \quad (13.6)$$

One can now define an algebra of words as follows (cf. Definition 7.2.1):

Definition 13.3.2. The set of alphabet X consists of $N+1$ letters a and b_i for $i \equiv 0,\dots,N-1 \pmod N$. The *weight* of a word $\mathbf{w}$ (*i.e.*, a monomial in the letters in X) is the number of letters contained in $\mathbf{w}$ and its *depth* is the number of b_i's contained in $\mathbf{w}$. Define Hoffman-Racinet algebra of level N, denoted by $\mathfrak{A}_N$, to be the (weight) graded noncommutative polynomial $\mathbb{Q}$-algebra generated by words over the alphabet X. Let $\mathfrak{A}_N^0$ be the subalgebra of $\mathfrak{A}_N$ generated by words not beginning with b_0 and not ending with a. The words in $\mathfrak{A}_N^0$ are called *admissible words.*

By Eq. (13.6) every CMZV can be expressed as an iterated integral over the closed interval $[0,1]$ of an admissible word w in $\mathfrak{A}_N^0$. This is denoted by

$$Ł(\mathbf{w}) := \int_0^1 \mathbf{w}. \quad (13.7)$$

We also extend Ł to $\mathfrak{A}_N$ by $\mathbb{Q}$-linearity. We remark that the length $\lg(\mathbf{w})$ of $\mathbf{w}$ is equal to the weight of Ł$(\mathbf{w})$. Therefore in general one has

$$L_N(s_1,\dots,s_n|i_1,i_2,\dots,i_n) = Ł(\mathsf{z}_{s_1,i_1}\mathsf{z}_{s_2,i_1+i_2}\cdots\mathsf{z}_{s_n,i_1+i_2+\cdots+i_n}), \quad (13.8)$$

$$Ł(\mathsf{z}_{s_1,i_1}\mathsf{z}_{s_2,i_2}\cdots\mathsf{z}_{s_n,i_n}) = L_N(s_1,\dots,s_n|i_1,i_2-i_1,\dots,i_n-i_{n-1}). \quad (13.9)$$

For example $L_3(1,2,2|1,0,2) = Ł(\mathsf{z}_{1,1}\mathsf{z}_{2,1}\mathsf{z}_{2,0})$. Therefore, we can apply Chen's theory on iterated integrals to compute the product of two CMZVs. For example, by Lemma 2.1.2 from Chap. 2, one has

$$\begin{aligned}
&L_N(1|1)L_N(2,3|1,2) = Ł(\mathsf{z}_{1,1})\,Ł(\mathsf{z}_{2,1}\mathsf{z}_{3,3}) = Ł(b_1 \sqcup\!\sqcup (ab_1a^2b_3))\\
=&Ł(b_1ab_1a^2b_3 + 2ab_1^2a^2b_3 + (ab_1)^2ab_3 + ab_1a^2b_1b_3 + ab_1a^2b_3b_1)\\
=&Ł(\mathsf{z}_{1,1}\mathsf{z}_{2,1}\mathsf{z}_{3,3} + 2\mathsf{z}_{2,1}\mathsf{z}_{1,1}\mathsf{z}_{3,3} + \mathsf{z}_{2,1}^2\mathsf{z}_{2,3} + \mathsf{z}_{2,1}\mathsf{z}_{3,1}\mathsf{z}_{1,3} + \mathsf{z}_{2,1}\mathsf{z}_{3,3}\mathsf{z}_{1,1})\\
=&L_N(1,2,3|1,0,2) + 2L_N(2,1,3|1,0,2) + L_N(2,2,2|1,0,2)\\
&\qquad + L_N(2,3,1|1,0,2) + L_N(2,3,1|1,2,N-2).
\end{aligned}$$

Let $\mathfrak{A}_{N,\sqcup\!\sqcup}$ be the algebra of $\mathfrak{A}_N$ where the multiplication is defined by the shuffle product $\sqcup\!\sqcup$. Denote the subalgebra $\mathfrak{A}_N^0$ by $\mathfrak{A}_{N,\sqcup\!\sqcup}^0$ when one considers the shuffle product. Then one can easily prove

Proposition 13.3.3. *The map* $Ł : \mathfrak{A}^0_{N,ш} \longrightarrow \mathbb{C}$ *is an algebra homomorphism.*

Proof. See Exercise 13.2. □

On the other hand, CMZVs are known to satisfy the series stuffle relations. For example

$$L_N(2|5)L_N(3|4) = L_N(2,3|5,4) + L_N(3,2|4,5) + L_N(5|9).$$

To study such relations in general one has the following definition.

Definition 13.3.4. Denote by $\mathfrak{A}^1_N$ the subalgebra of $\mathfrak{A}_N$ which is generated by words $\mathtt{z}_{s,i}$ with $s \in \mathbb{N}$ and $i \equiv 0, \ldots, N-1 \pmod N$. Equivalently, $\mathfrak{A}^1_N$ is the subalgebra of $\mathfrak{A}_N$ generated by words not ending with a. For any word $\mathbf{w} = \mathtt{z}_{s_1,i_1}\mathtt{z}_{s_2,i_2}\cdots\mathtt{z}_{s_n,i_n} \in \mathfrak{A}^1_N$ and positive integer j one defines the exponent shifting operator τ_j by

$$\tau_j(\mathbf{w}) = \mathtt{z}_{s_1,j+i_1}\mathtt{z}_{s_2,j+i_2}\cdots\mathtt{z}_{s_n,j+i_n}.$$

For convenience, on the empty word we adopt the convention that $\tau_j(\mathtt{e}) = \mathtt{e}$. We then define another multiplication $*$ on $\mathfrak{A}^1_N$ by requiring that $*$ distribute over addition, that $\mathtt{e} * \mathbf{w} = \mathbf{w} * \mathtt{e} = \mathbf{w}$ for any word $\mathbf{w}$, and that, for any words $\mathbf{u}, \mathbf{v}$,

$$\begin{aligned}\mathtt{z}_{s,j}\mathbf{u} * \mathtt{z}_{t,k}\mathbf{v} = \mathtt{z}_{s,j}\Big(\tau_j\big(\tau_{-j}(\mathbf{u}) * \mathtt{z}_{t,k}\mathbf{v}\big)\Big) + \mathtt{z}_{t,k}\Big(\tau_k\big(\mathtt{z}_{s,j}\mathbf{u} * \tau_{-k}(\mathbf{v})\big)\Big)\\ + \mathtt{z}_{s+t,j+k}\Big(\tau_{j+k}\big(\tau_{-j}(\mathbf{u}) * \tau_{-k}(\mathbf{v})\big)\Big). \quad (13.10)\end{aligned}$$

This multiplication is called the *stuffle product*.

If we denote by $\mathfrak{A}^1_{N,*}$ the algebra $(\mathfrak{A}^1_N, *)$ then it is not hard to prove the next proposition.

Proposition 13.3.5. *The polynomial algebra* $\mathfrak{A}^1_{N,*}$ *is a commutative weight graded* $\mathbb{Q}$*-algebra.*

Proof. See Exercise 13.3. □

Now we can define the subalgebra $\mathfrak{A}^0_{N,*}$ similar to $\mathfrak{A}^0_{N,ш}$ by replacing the shuffle product by the stuffle product. Then by induction on the lengths and using the series definition one can quickly check that for any $\mathbf{w}_1, \mathbf{w}_2 \in \mathfrak{A}^0_{N,*}$

$$Ł(\mathbf{w}_1)\,Ł(\mathbf{w}_2) = Ł(\mathbf{w}_1 * \mathbf{w}_2).$$

This implies that

Proposition 13.3.6. *The map* $Ł : \mathfrak{A}^0_{N,*} \longrightarrow \mathbb{C}$ *is an algebra homomorphism.*

Proof. See Exercise 13.3. □

Definition 13.3.7. Let w be an integer such that $w \geq 2$. For nontrivial words $\mathbf{w}_1, \mathbf{w}_2 \in \mathfrak{A}^0_N$ with $|\mathbf{w}_1| + |\mathbf{w}_2| = w$ we say that the equation

$$Ł(\mathbf{w}_1 \sqcup\!\sqcup \mathbf{w}_2 - \mathbf{w}_1 * \mathbf{w}_2) = 0 \tag{13.11}$$

provides a *finite double shuffle relation* (finite DBSF) of weight w.

It is known that even in level one (*i.e.*, the MZV case) these relations are not enough to provide all the relations among MZVs. For example, the weight of the product of two MZVs is at least 4. So the finite DBSFs cannot imply the well-known identity $\zeta(2,1) = \zeta(3)$. However, it is believed that one can remedy this by considering *regularized double shuffle relation* (regularized DBSF) produced by the following mechanism.

First, combining Prop. 13.3.6 and Prop. 13.3.3 we can easily prove the following algebraic result.

Proposition 13.3.8. *We have two algebra homomorphisms:*

$$Ł_* : (\mathfrak{A}^1_{N,*}, *) \longrightarrow \mathbb{C}[T], \quad \textit{and} \quad Ł_{\sqcup\!\sqcup} : (\mathfrak{A}^1_{N,\sqcup\!\sqcup}, \sqcup\!\sqcup) \longrightarrow \mathbb{C}[T]$$

which are uniquely determined by the properties that they both extend the evaluation map $Ł : \mathfrak{A}^0_N \longrightarrow \mathbb{C}$ *by sending* $b_0 = \mathbf{z}_{1,1}$ *to* T.

Second, in order to establish the crucial relation between $Ł_*$ and $Ł_{\sqcup\!\sqcup}$ one can adopt the machinery similar to that of the MZVs developed in section 3.2 as follows. For any $(\mathbf{s}|\mathbf{i}) = (s_1, \ldots, s_n | i_1, \ldots, i_n)$ where i_j's are integers and s_j's are positive integers, let the image of the corresponding words in $\mathfrak{A}^1_N$ under $Ł_*$ and $Ł_{\sqcup\!\sqcup}$ be denoted by $Ł_*(\mathbf{s}|\mathbf{i}; T)$ and $Ł_{\sqcup\!\sqcup}(\mathbf{s}|\mathbf{i}; T)$ respectively.

Theorem 13.3.9. *Define a* $\mathbb{C}$*-linear map* $\rho : \mathbb{C}[T] \to \mathbb{C}[T]$ *by*

$$\rho(e^{Tu}) = \exp\left(\sum_{n=2}^{\infty} \frac{(-1)^n}{n} \zeta(n) u^n\right) e^{Tu}, \qquad |u| < 1.$$

Then for any index set $(\mathbf{s}|\mathbf{i})$ *one has*

$$Ł_{\sqcup\!\sqcup}(\mathbf{s}|\mathbf{i}; T) = \rho\big(Ł_*(\mathbf{s}|\mathbf{q}(\mathbf{i}); T)\big), \tag{13.12}$$

where $\mathbf{q}(i_1, \ldots, i_n) = (i_1, i_2 - i_1, \ldots, i_n - i_{n-1})$.

Thm. 13.3.9 is a generalization of Thm. 3.3.21 to the higher level CMZV cases, but the proof is much more complicated, which will be given in the next section using some advanced results from Lie algebra.

Definition 13.3.10. Let w be a positive integer such that $w \geq 2$. Let $(\mathbf{s}|\mathbf{i})$ be any index set with the weight of $\mathbf{s}$ equal to w. Then every weight w CMZV relation produced by Eq. (13.12) is called a *regularized double shuffle relation* (regularized DBSF) of weight w. This is obtained by formally setting $T = 0$ in Eq. (13.12).

The above steps can be easily transformed to computer codes which have been used in some Maple programs. For example, one gets by the stuffle product

$$\begin{aligned} TL_N(2|3) &= \text{Ł}_*(1|0;T)\,\text{Ł}_*(2|3;T) = \text{Ł}_*(\mathtt{z}_{1,0} * \mathtt{z}_{2,3}) \\ &= \text{Ł}_*(1,2|0,3;T) + \text{Ł}_*(2,1|3,3;T) + \text{Ł}_*(3|3;T), \end{aligned}$$

while using shuffle product one has

$$\begin{aligned} TL_N(2|3) &= \text{Ł}_{\sqcup\!\sqcup}(1|0;T)\,\text{Ł}_{\sqcup\!\sqcup}(2|3;T) = \text{Ł}_{\sqcup\!\sqcup}(\mathtt{z}_{1,0} \sqcup\!\sqcup \mathtt{z}_{2,3}) = \text{Ł}_{\sqcup\!\sqcup}(b_0 \sqcup\!\sqcup ab_3) \\ &= \text{Ł}_{\sqcup\!\sqcup}(1,2|0,3;T) + \text{Ł}_{\sqcup\!\sqcup}(2,1|0,3;T) + \text{Ł}_{\sqcup\!\sqcup}(2,1|3,0;T). \end{aligned}$$

We thus obtains the following regularized DBSF if we compare the above two expressions using Thm. 13.3.9:

$$L_N(2,1|3,0) + L_N(3|3) = L_N(2,1|3,N-3) + L_N(2,1|0,3).$$

13.3.2 *Proof of regularized double shuffle relations*

When $N = 1$ many results in this subsection have been proved in Chap. 3 already. But the proofs of these result for general levels N contained here are often different because more advanced tools are needed.

Recall that we have defined the alphabet $\mathsf{X} = \mathsf{X}_N = \{a, b_j : j \in \mathbb{Z}/N\mathbb{Z}\}$ and chosen $\mu = \mu_N = \exp(2\pi\sqrt{-1}/N)$. Let $\Gamma_N := \{\mu^j : j \in \mathbb{Z}/N\mathbb{Z}\}$ be the set of Nth roots of unity. For convenience, we also use the letters $\mathtt{x}_\xi$, $\xi \in \Gamma_N \cup \{0\}$, such that $\mathtt{x}_0 = a$ and $\mathtt{x}_{\mu^j} = b_j$. Thus $\mathsf{X} = \{\mathtt{x}_\xi : \xi \in \Gamma_N \cup \{0\}\}$.

We now denote by Y the subset of X of all the words $\mathtt{y}_{k,\xi} = \mathtt{x}_0^{k-1}\mathtt{x}_\xi$ for $k \in \mathbb{N}$ and $\xi \in \Gamma_N$.

Remark 13.3.11. Observe that $\mathtt{y}_{n,\mu^j} = \mathtt{z}_{n,j}$ used in the proceeding subsection. It turns out that the notation $\mathtt{y}_{n,\xi}$ adopted throughout this subsection has some theoretical advantage while it is much easier to use $\mathtt{z}_{n,j}$ in computer aided symbolic computations which have been implicitly used in the examples of the proceeding subsection.

Let R be any $\mathbb{Q}$-algebra. The subalgebra of $R\langle \mathsf{X}\rangle$ generated by Y is denoted by $R\langle \mathsf{Y}\rangle$ and is free on Y. As a vector space, it is generated by the words over X not ending with $\mathtt{x}_0$ thus it could be identified with the quotient $R\langle \mathsf{X}\rangle/R\langle \mathsf{X}\rangle \mathtt{x}_0$. The corresponding projection is denoted by π_{Y}. The vector subspace $R\langle \mathsf{X}\rangle_{\rm cv}$ of $R\langle \mathsf{X}\rangle$ generated by words over X that are not ending with $\mathtt{x}_0$ and not starting with $\mathtt{x}_1$ is also a graded algebra. It may be identified with the quotient $R\langle \mathsf{X}\rangle/(\mathtt{x}_1 R\langle \mathsf{X}\rangle + R\langle \mathsf{X}\rangle \mathtt{x}_0)$. We denote by $\pi_{\rm cv}$ the corresponding projection. The algebra $R\langle \mathsf{X}\rangle_{\rm cv}$ is same as the subalgebra of $R\langle \mathsf{Y}\rangle$ generated by words over Y that are not starting with $\mathtt{y}_{1,1} = \mathtt{x}_1$. (Note that the second subscript 1 in $\mathtt{y}_{1,1}$ is an element in Γ_N.) Thus it can be identified with the quotient $R\langle \mathsf{Y}\rangle/\mathtt{y}_{1,1} R\langle \mathsf{Y}\rangle$, which is denoted by $R\langle \mathsf{Y}\rangle_{\rm cv}$. Again its corresponding projection is denoted by $\pi_{\rm cv}$. Similarly one defines $R\langle\langle \mathsf{X}\rangle\rangle_{\rm cv}$ and $R\langle\langle \mathsf{Y}\rangle\rangle_{\rm cv}$ as the completion of $R\langle \mathsf{X}\rangle_{\rm cv}$ and $R\langle \mathsf{Y}\rangle_{\rm cv}$ with respect to the natural grading, respectively.

Remark 13.3.12. The algebras $\mathfrak{A}$, $\mathfrak{A}^0$ and $\mathfrak{A}^1$ in Sec. 13.3.1 are essentially equal to $\mathbb{Q}\langle \mathsf{X}\rangle$, $\mathbb{Q}\langle \mathsf{X}\rangle_{\rm cv}$ and $\mathbb{Q}\langle \mathsf{Y}\rangle$, respectively.

Now we take $(\mathbb{C}\langle \mathsf{X}\rangle, \bullet, \Delta)$ to be the graded bialgebra defined by Example B.18 in Appendix B and $(\mathbb{C}\langle \mathsf{Y}\rangle, \bullet, \Delta_*)$ the graded bialgebra defined by Example B.19. In particular, all $\mathtt{x}_\xi$ are primitive elements in $(\mathbb{C}\langle \mathsf{X}\rangle, \bullet, \Delta)$. Similarly, $\mathbb{C}\langle\langle \mathsf{X}\rangle\rangle$, $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle$, $\mathbb{C}\langle\langle \mathsf{X}\rangle\rangle_{\rm cv}$ and $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle_{\rm cv}$ are all graded bialgebras (see Exercise 13.4).

Recall that for any $a_1, \ldots, a_k \in \mathbb{C}$ and any path $\gamma : [0,1] \to \mathbb{C}$, the iterated integral $\mathrm{I}_\gamma(a_1, \ldots, a_k)$ is defined by Eq. (2.14). In particular, for $\xi_1, \ldots, \xi_d \in \Gamma_N \cup \{0\}$, $\xi_1 \neq 1$ and $\xi_d \neq 0$, we have

$$I_{[0,1]}(\xi_1, \ldots, \xi_d) = \int_{1>t_1>\cdots>t_d>0} \bigwedge_{j=1}^{d} \omega_{\xi_j}(t_j),$$

where for all $j = 1, \ldots, d$

$$\omega_{\xi_j}(t_j) = \begin{cases} \dfrac{\mathrm{d}t_j}{\xi_j^{-1} - t_j}, \text{ if } \xi_j \neq 0; \\ \dfrac{\mathrm{d}t_j}{t_j}, \quad \text{if } \xi_j = 0. \end{cases}$$

Recall that an element x in any coalgebra (C, ϵ, Δ) is called *group-like* (see Definition A.6) if $\Delta(x) = x \otimes x$ and $\epsilon(x) = 1$.

Proposition 13.3.13. *For any path $\gamma \in \mathbb{C}^* \setminus \Gamma_N$, we may define*

$$\mathbf{I}_\gamma := \sum_{d \in \mathbb{N}_0, \xi_1, \ldots, \xi_d \in \Gamma_N \cup \{0\}} \mathrm{I}_\gamma(\xi_1, \ldots, \xi_d) \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_d}.$$

Then $\mathbf{I}_\gamma$ is a group-like element in $(\mathbb{C}\langle \mathsf{X}\rangle, \bullet, \Delta)$.

Proof. Let X^* be the set of words over X. For any word $\mathbf{w} = l_1 \dots l_n \in \mathsf{X}^*$ and any subset $J = \{i_1, \dots, i_k\}$ of $\{1, \dots, n\}$ in increasing order, we define $\mathbf{w}|J = l_{i_1} \dots l_{i_n}$. Further set $\mathbf{w}|\emptyset = \mathtt{e}$. Since Δ is an algebra homomorphism we have

$$\Delta(\mathbf{w}) = \Delta(l_1) \dots \Delta(l_n) = \prod_{j=1}^{n} (l_j \otimes 1 + 1 \otimes l_j)$$

$$= \sum_{J_1 \cup J_2 = \{1,\dots,n\}, J_1 \cap J_2 = \emptyset} (\mathbf{w}|J_1) \otimes (\mathbf{w}|J_2) = \sum_{\mathbf{u}_1, \mathbf{u}_2 \in \mathsf{X}^*} (\mathbf{w}, \mathbf{u}_1 \sqcup\!\sqcup \mathbf{u}_2) \mathbf{u}_1 \otimes \mathbf{u}_2$$

by the definition of $\sqcup\!\sqcup$, where $(\mathbf{w}, \mathbf{u}_1 \sqcup\!\sqcup \mathbf{u}_2)$ is the coefficient of $\mathbf{w}$ in $\mathbf{u}_1 \sqcup\!\sqcup \mathbf{u}_2$. This implies that

$$\begin{aligned} \Delta(\mathbf{I}_\gamma) &= \sum_{\substack{n \in \mathbb{N}_0, \xi_1, \dots, \xi_n \in \Gamma_N \cup \{0\} \\ m \in \mathbb{N}_0, \tau_1, \dots, \tau_m \in \Gamma_N \cup \{0\}}} \mathrm{I}_\gamma\Big(\boldsymbol{\xi} \sqcup\!\sqcup \boldsymbol{\tau}\Big) \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_n} \otimes \mathtt{x}_{\tau_1} \cdots \mathtt{x}_{\tau_m} \\ &= \sum_{\substack{n \in \mathbb{N}_0, \xi_1, \dots, \xi_n \in \Gamma_N \cup \{0\} \\ m \in \mathbb{N}_0, \tau_1, \dots, \tau_m \in \Gamma_N \cup \{0\}}} \mathrm{I}_\gamma(\boldsymbol{\xi}) \mathrm{I}_\gamma(\boldsymbol{\tau}) \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_n} \otimes \mathtt{x}_{\tau_1} \cdots \mathtt{x}_{\tau_m}, \end{aligned}$$

by Lemma 2.1.2 (iv), where $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$ and $\boldsymbol{\tau} = (\tau_1, \dots, \tau_m)$. Thus $\Delta(\mathbf{I}_\gamma) = \mathbf{I}_\gamma \otimes \mathbf{I}_\gamma$ which yields the proposition. □

For any $0 < \alpha < z < 1$ and $\xi_1, \dots, \xi_n \in \Gamma_N \cup \{0\}$ we can also prove $I_{[\alpha,z]}(\xi_1, \dots, \xi_n)$ is group-like in the following way. Fixing α and regarding $\mathbf{I}_{[\alpha,z]}$ as a function of z, we see that it provides a solution of the differential equation

$$d\,\mathbf{I} = \Omega\,\mathbf{I}, \qquad \text{where } \Omega = \sum_{\xi \in \Gamma_N \cup \{0\}} \omega_\xi(z) \mathtt{x}_\xi.$$

Notice that Ω is primitive for the coproduct Δ so both $\Delta\,\mathbf{I}_{[\alpha,z]}$ and $\mathbf{I}_{[\alpha,z]} \otimes \mathbf{I}_{[\alpha,z]}$ are the solutions to the differential equation $d\mathcal{J} = \Delta(\Omega)\mathcal{J}$ satisfying the initial condition $\mathcal{J}(a) = 1 \otimes 1$. Therefore

$$\Delta\,\mathbf{I}_{[\alpha,z]} = \mathbf{I}_{[\alpha,z]} \otimes \mathbf{I}_{[\alpha,z]}\,.$$

Proposition 13.3.14. *For all $z \in (0,1)$ the series*

$$\mathbf{I}_{[0,z]} := \lim_{\alpha \to 0} \pi_Y(\mathbf{I}_{[\alpha,z]}) = \sum_{\substack{\xi_1, \dots, \xi_d \in \Gamma_N \cup \{0\} \\ \xi_d \neq 0, d \in \mathbb{N}_0}} \mathrm{I}_{[0,z]}(\xi_1, \dots, \xi_d) \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_d} \quad (13.13)$$

is a group-like element in $(\mathbb{C}\langle\langle \mathsf{Y} \rangle\rangle, \Delta)$.

Proof. Regarding α as variable in $(0,1)$, by Prop. 13.3.13 we see that the series $\pi_Y(\mathbf{I}_{[\alpha,z]})$ is a group-like element in $(\mathbb{C}^{(0,1)}\langle\langle Y\rangle\rangle, \Delta)$ since π_Y is a coalgebra homomorphism. Taking $\alpha \to 0$ we see that the limit $\pi_Y(\mathbf{I}_{[0,z]})$ exists by Prop. 2.3.8. This completes the proof of the proposition since taking limit is also a coalgebra homomorphism. □

For any word $\mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2}\cdots\mathtt{y}_{s_d,\xi_d}$ over Y, we define the endomorphism $\mathbf{q}$ of $\mathbb{C}\langle\mathsf{Y}\rangle$ as

$$\mathbf{q}(\mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2}\cdots\mathtt{y}_{s_d,\xi_d}) = \mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2\xi_1^{-1}}\cdots\mathtt{y}_{s_d,\xi_d\xi_{d-1}^{-1}}. \tag{13.14}$$

It has an inverse defined by

$$\mathbf{p}(\mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2}\cdots\mathtt{y}_{s_d,\xi_d}) = \mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_1\xi_2}\cdots\mathtt{y}_{s_d,\xi_1\xi_2\cdots\xi_d}. \tag{13.15}$$

Remark 13.3.15. The map $\mathbf{q}$ originates from the formula

$$I_{[0,1]}(\{0\}^{s_1-1},\xi_1,\ldots,\{0\}^{s_d-1}\xi_d) = L_N(s_1,\ldots,s_d|i_1,i_2-i_1,\ldots,i_d-i_{d-1})$$

where $(s_1,\mu^{i_1}) \neq (1,1)$ and $\xi_j = \mu^{i_j}$ for all $j = 1,\ldots,d$.

Proposition 13.3.16. *For any $z \in (0,1)$, we have the following identity in $\mathbb{C}\langle\langle\mathsf{Y}\rangle\rangle$:*

$$\mathbf{L}_z := \sum_{d\in\mathbb{N}_0,\mathbf{s}\in\mathbb{N}^d,\xi_1,\ldots,\xi_d\in\Gamma_N} \mathrm{Li}_{\mathbf{s}}(z\xi_1,\xi_2,\ldots,\xi_d)\mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2}\cdots\mathtt{y}_{s_d,\xi_d} = \mathbf{q}(\mathbf{I}_{[0,z]}).$$

Proof. This is just a restatement of Prop. 2.3.8. □

In the next result we need to apply the projection π_{cv} to remove the divergent terms from Eq. (13.13).

Proposition 13.3.17. *The series*

$$\mathbf{I}_{\mathrm{cv}} := \lim_{z\to 1^-} \pi_{\mathrm{cv}}\,\mathbf{I}_{[0,z]} = \sum_{\xi_1,\ldots,\xi_d\in\Gamma_N\cup\{0\},\xi_1\neq 1,\xi_d\neq 0} \mathrm{I}_{[0,1]}(\xi_1,\ldots,\xi_d)\mathtt{x}_{\xi_1}\cdots\mathtt{x}_{\xi_d}$$

is a group-like element in $\mathbb{C}\langle\langle\mathsf{X}\rangle\rangle_{\mathrm{cv}}$.

Proof. Prop. 13.3.16 shows that as $z \to 1^-$ the asymptotic expansion of every coefficient of the series $\pi_{\mathrm{cv}}(\mathbf{I}_{[0,z]})$ is given by the associated power series (see Eq. (2.9)) of $\left(\mathrm{Li}_{\mathbf{s}}^{(M)}(\boldsymbol{\xi})\right)_{M\geq 1} \in \mathsf{DivLog}_{\mathbb{N}}$ for some $\mathbf{s} \in \mathbb{N}^d$ and $\boldsymbol{\xi} \in (\Gamma_N)^d$ ($(s_1,\xi_1) \neq (1,1)$), which converges by Cor. 2.3.10. The proposition now follows from Abel's lemma on the convergence of a series at the boundary of its convergence interval (*i.e.*, Lemma 2.3.7). □

By applying π_{cv} we can remove the divergent terms as follows:

$$\lim_{z\to 1^-} \pi_{\mathrm{cv}}\, \mathbf{L}_z = \sum_{\substack{d\in\mathbb{N}_0,\mathbf{s}=(s_1,\dots,s_d)\in\mathbb{N}^d\\ \xi_1,\dots,\xi_d\in\Gamma_N,(s_1,\xi_1)\neq(1,1)}} \mathrm{Li}_{\mathbf{s}}(\xi_1,\dots,\xi_d)\mathtt{y}_{s_1,\xi_1}\cdots\mathtt{y}_{s_d,\xi_d}.$$

By straight-forward computation we find the following crucial relation between the above two elements.

Proposition 13.3.18. *In* $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle_{\mathrm{cv}}$ *one has*

$$\mathbf{L}_{\mathrm{cv}} := \lim_{z\to 1^-} \pi_{\mathrm{cv}}\, \mathbf{L}_z = \mathbf{q}(\mathbf{I}_{\mathrm{cv}}).$$

Proof. Noticing that $\mathbf{q}$ and π_{cv} commute, by the definition of π_{cv} we get

$$\begin{aligned}\mathbf{L}_{\mathrm{cv}} = \lim_{z\to 1^-} \pi_{\mathrm{cv}}\, \mathbf{L}_z &= \lim_{z\to 1^-} \pi_{\mathrm{cv}}\mathbf{q}(\mathbf{I}_{[0,z]}) \quad \text{(by Prop. 13.3.16)}\\ &= \mathbf{q}(\lim_{z\to 1^-} \pi_{\mathrm{cv}}\, \mathbf{I}_{[0,z]}) = \mathbf{q}(\mathbf{I}_{\mathrm{cv}})\end{aligned}$$

by Prop. 13.3.17 □

Proposition 13.3.19. *The series* $\mathbf{L}_{\mathrm{cv}}$ *is a group-like element in* $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle_{\mathrm{cv}}$.

Proof. First we observe that

$$\mathcal{Y} := \sum_{n\in\mathbb{N}_0,\xi\in\Gamma_N} \mathtt{y}_{n,\xi} \tag{13.16}$$

is a group-like element in $(\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle, \bullet, \Delta_*)$ (see Exercise 13.5). Next it can be checked easily that for any $\lambda\in\mathbb{C}$, both of the maps $\mathtt{y}_{n,\xi}\mapsto\lambda^n\mathtt{y}_{n,\xi}$ and $\mathtt{y}_{n,\xi}\mapsto\xi\mathtt{y}_{n,\xi}$ are homomorphisms of the coalgebra $(\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle, \Delta_*)$, meaning they commute with Δ_*. Therefore by applying the first map once and the second m times to the sum in Eq. (13.16) we get a group-like element for Δ_* (see Exercise 13.7):

$$\mathcal{Y}_m(\lambda) := \sum_{n\in\mathbb{N}_0,\xi\in\Gamma_N} \xi^m\lambda^n\mathtt{y}_{n,\xi}.$$

Now for any $M\in\mathbb{N}$, we may define the Mth partial sum of CMZVs by (cf. Eq. (2.13))

$$\mathrm{Li}^{(M)}_{s_1,\dots,s_d}(\xi_1,\dots,\xi_d) := \sum_{M\geq k_1>\cdots>k_d>0} \frac{\xi_1^{k_1}\cdots\xi_d^{k_d}}{k_1^{s_1}\cdots k_d^{s_d}}.$$

Then it is straight-forward to check that its generating series

$$\mathbf{L}^{(M)} := \sum_{\substack{d\in\mathbb{N}_0,\xi_1,\dots,\xi_d\in\Gamma_N\\ s_1,\dots,s_d\in\mathbb{N}}} \mathrm{Li}^{(M)}_{s_1,\dots,s_d}(\xi_1,\dots,\xi_d)\mathtt{y}_{s_1,\xi_1}\cdots\mathtt{y}_{s_d,\xi_d} \tag{13.17}$$

can be expressed as (see Exercise 13.8)

$$\mathbf{L}^{(M)} = \mathcal{Y}_M\Big(\frac{1}{M}\Big)\mathcal{Y}_{M-1}\Big(\frac{1}{M-1}\Big)\cdots\mathcal{Y}_2\Big(\frac{1}{2}\Big)\mathcal{Y}_1\Big(\frac{1}{1}\Big). \tag{13.18}$$

Since each factor on the right-hand side is group-like we see that $\mathbf{L}^{(M)}$ must be group-like, too.

We now can check that $\mathtt{y}_{1,1}\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle$ is a co-ideal for Δ_* since $\Delta_*(\mathtt{y}_{1,1}) = \mathtt{y}_{1,1}\otimes 1 + 1\otimes \mathtt{y}_{1,1}$ by the definition. Therefore, by Prop. A.10 in Appendix A, the coproduct Δ_* descends onto the quotient $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle/\mathtt{y}_{1,1}\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle$ which is identified with $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle_{\mathrm{cv}}$. As the projection π_{cv} is clearly a coalgebra homomorphism (*i.e.*, commutes with Δ_*) we see immediately that

$$\mathbf{L}_{\mathrm{cv}} = \lim_{M\to\infty} \pi_{\mathrm{cv}}(\mathbf{L}^{(M)})$$

is a group-like element in $(\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle, \Delta_*)$. This completes the proof of the proposition. □

The following two lemmas are immediate consequences of Cor. B.26 in Appendix B (see Exercise 13.10).

Lemma 13.3.20. *Every group-like element Φ_{cv} in $(\mathbb{Q}\langle \mathsf{X}\rangle_{\mathrm{cv}}, \Delta)$ is the image of some group-like elements in $(\mathbb{Q}\langle \mathsf{X}\rangle, \Delta)$ under π_{cv}. Moreover, any two such elements Φ_1 and Φ_1 in $(\mathbb{Q}\langle \mathsf{X}\rangle_{\mathrm{cv}}, \Delta)$ are related by*

$$\Phi_2 = \exp((\lambda_2-\lambda_1)\mathtt{x}_1)\Phi_1\exp((\rho_2-\rho_1)\mathtt{x}_0)$$

where λ_j and ρ_j are the coefficients of $\mathtt{x}_1$ and $\mathtt{x}_0$ of Φ_j ($j=1,2$), respectively.

Proof. Apply Cor. B.26 twice. □

Lemma 13.3.21. *Any group-like series Φ_{cv} of $(\mathbb{Q}\langle\langle \mathsf{Y}\rangle\rangle_{\mathrm{cv}}, \Delta_*)$ is the image of a group-like element of $(\mathbb{Q}\langle\langle \mathsf{Y}\rangle\rangle, \Delta_*)$. Two such elements Φ_1 and Φ_2 are related by*

$$\Phi_2 = \exp((\lambda_2-\lambda_1)\mathtt{y}_{1,1})\Phi_1,$$

where λ_1 and λ_2 are the coefficients of $\mathtt{y}_{1,1}$ in Φ_1 and Φ_2 respectively.

With the above lemma available, we can define

Definition 13.3.22. Let $\mathbf{L}$ (resp. $\mathbf{I}$) be the unique group-like element in $(\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle, \Delta_*)$ (resp. $(\mathbb{C}\langle\langle \mathsf{X}\rangle\rangle, \Delta)$) such that $\pi_{\mathrm{cv}}(\mathbf{L}) = \mathbf{L}_{\mathrm{cv}}$ (resp. $\pi_{\mathrm{cv}}(\mathbf{I}) = \mathbf{I}_{\mathrm{cv}}$) and in which the coefficients of $\mathtt{y}_{1,1}$ is 0 (resp. in which the coefficients of $\mathtt{x}_0$ and $\mathtt{x}_1$ are 0). We also call $\mathbf{L}$ (resp. $\mathbf{I}$) the regularization of $\mathbf{L}_{\mathrm{cv}}$ (resp. $\mathbf{I}_{\mathrm{cv}}$) .

Remark 13.3.23. The coefficients of $\mathbf{L}$ (resp. $\mathbf{I}$) may be viewed as regularized values of the series $\mathrm{Li}_{s_1,\ldots,s_n}(\xi_1,\ldots,\xi_n)$ (resp. the iterated integral $I_{[0,1]}(\xi_1,\ldots,\xi_n)$). Note that Racinet's choice is different from that given by Ihara, Kaneko, and Zagier in which the coefficient of $\mathtt{y}_{1,1}$ is T instead of 0.

In what follows we will derive a relation between the two series $\mathbf{L}$ and $\mathbf{I}$. The main tools are contained in the three subsections of Sec. 2.3.

Suppose $f : k_1 \longrightarrow k_2$ is a $\mathbb{C}$-linear map between two rings over $\mathbb{C}$. If $\Phi \in k_1\langle\langle \mathsf{X}\rangle\rangle$ then we denote by $\bar{f}(\Phi)$ the element in $k_2\langle\langle \mathsf{X}\rangle\rangle$ obtained by applying f on the coefficients of Φ. By $\mathbb{C}$-linearity we see that $\bar{f}$ is a homomorphism of $\mathbb{C}\langle\langle \mathsf{X}\rangle\rangle$-bimodules. Further, if f is an algebra homomorphism and Φ is group-like for Δ then so is $\bar{f}(\Phi)$. The same is true for Y and Δ_*.

By Cor. 2.3.10 all the coefficients of the series $\mathbf{L}^{(M)}$ defined by Eq. (13.17) are in $\mathsf{DivLog}_{\mathbb{N}}$. By Remark 2.3.6 we see that As is an algebra homomorphism, so $\mathbf{L}_t := \overline{\mathrm{As}}(\mathbf{L}^{(M)})$ is group-like in $(\mathbb{C}[t]\langle\langle \mathsf{Y}\rangle\rangle, \Delta_*)$, where we have extended As from $\mathbb{C}^{\mathbb{N}}$ to $\mathbb{C}^{\mathbb{N}}\langle\langle \mathsf{Y}\rangle\rangle$. Clearly, $\pi_{\mathrm{cv}}\,\mathbf{L}_t = \mathbf{L}_{\mathrm{cv}}$ since all convergent coefficients result in constant functions. Hence we can regard the space $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle$ as a subspace of $\mathbb{C}[t]\langle\langle \mathsf{Y}\rangle\rangle$.

Lemma 13.3.24. *In* $\mathbb{C}[t]\langle\langle \mathsf{Y}\rangle\rangle$*, we have*

$$\mathbf{L}_t = \exp\left((t+\gamma)\mathtt{y}_{1,1}\right)\mathbf{L}\,. \tag{13.19}$$

Proof. Noticing that the coefficient of $\mathtt{y}_{1,1}$ in $\mathbf{L}_t$ is $\mathrm{As}_{\mathbf{h}} = t + \gamma$ (see Example 2.3.3), we can readily prove the lemma by applying Lemma 13.3.21. □

Lemma 13.3.25. *The group-like element* $\exp\left(T\mathtt{y}_{1,1}\right)\mathbf{L}$ *in* $\mathbb{C}[T]\langle\langle \mathsf{Y}\rangle\rangle$ *is the generating series of* $Ł_*(\mathbf{s}|\mathbf{i};T)$*. Namely,*

$$\exp\left(T\mathtt{y}_{1,1}\right)\mathbf{L} = \sum_{\substack{d\in\mathbb{N}_0,\mathbf{s}=(s_1,\ldots,s_d)\in\mathbb{N}^d\\ \mathbf{i}=(i_1,\ldots,i_d)\in(\mathbb{Z}/N\mathbb{Z})^d}} Ł_*(\mathbf{s}|\mathbf{i};T)\mathtt{y}_{s_1,\mu^{i_1}}\cdots\mathtt{y}_{s_d,\mu^{i_d}}.$$

Proof. By the proof of Thm. 4.1.6 we know the generating series of $Ł_*(\mathbf{s}|\mathbf{i};T)$ is group-like for Δ_* since $Ł_*(\mathbf{s}|\mathbf{i};T)$ satisfies stuffle relations. Then we can apply Lemma 13.3.21 to deduce our result since $Ł_*(1|0;T) = T$ by Prop. 13.3.8. □

Corollary 13.3.26. *Let* $\mathbf{h}(k) = \left(\mathrm{Li}^{(M)}_{\{1\}^k}\left(\{1\}^k\right)\right)_{M\geq 1}$*. Then the associated power series* $\mathrm{Fp}(\mathbf{h}(k)) = (\mathrm{Li}_1)^k/k!$ *where* $\mathrm{Li}_1(z) = -\log(1-z)$*. Moreover,* $\mathrm{As}_{\mathbf{h}(k)}(t)$ *has degree* k*.*

Proof. By the definition, the coefficient of $\mathtt{y}_{1,1}^k$ in $\mathbf{L}^{(M)}$ is $\mathrm{Li}^{(M)}_{\{1\}^k}\big(\{1\}^k\big)$. So the last sentence of the corollary follows immediately from Lemma 13.3.24.

By the definition of Eq. (2.9), the function $\mathrm{Fp}(\mathbf{h}(k))$ is given by the power series

$$f_k(z) := \sum_{n_1>n_2>\cdots>n_k} \frac{z^{n_1}}{n_1 n_2 \cdots n_k}.$$

A simple differentiation yields the recurrence relation $f_k'(z) = f_{k-1}(z)/(1-z)$ for all $k \geq 1$ and $f_0(z) = 1$. Now by induction and the fact that $f_k(0) = 0$ for all $k > 0$ we get immediately that $f(z) = (\mathrm{Li}_1(z))^k/k!$, as desired. □

Proposition 13.3.27. *The coefficient of* $\mathtt{x}_1^k$ *in the group-like element* $\mathbf{I}_{[0,z]}$ *is* $(\mathrm{Li}_1(z))^k/k!$.

Proof. By the definition, the coefficient is

$$\int_0^z \left(\frac{\mathrm{d}t}{1-t}\right)^k = \int_1^{1-z} \left(-\frac{\mathrm{d}t}{t}\right)^k$$

by substitution $t \to 1-t$. This clearly evaluates to $(\mathrm{Li}_1(z))^k/k!$. □

Proposition 13.3.28. *There exists* $\mathbf{S}(\mathtt{y}_{1,1}) \in \mathbb{C}[\![\mathtt{y}_{1,1}]\!]$ *such that*

$$\mathbf{L} = \mathbf{S}(\mathtt{y}_{1,1})\mathbf{q}\pi_Y(\mathbf{I}).$$

Proof. Define $\mathbf{I}_t = \overline{\mathrm{As}}(\mathbf{I}_{[0,z]})$ which is a group-like element of $(\mathbb{C}[t]\langle\langle \mathsf{Y}\rangle\rangle, \Delta)$ so that $\pi_{\mathrm{cv}}\,\mathbf{I}_t = \mathbf{I}_{\mathrm{cv}}$. Since the coefficient of $\mathtt{x}_1$ in $\mathbf{I}_t$ is $\mathrm{As}_{\mathrm{Li}_1}(t) = t$ we conclude from Lemma 13.3.21 that

$$\mathbf{I}_t = \exp(\mathtt{x}_1 t)\pi_Y(\mathbf{I}). \tag{13.20}$$

On the other hand, by Prop. 13.3.16 $\mathbf{L}_z = \mathbf{q}(\mathbf{I}_{[0,z]})$ we find $\overline{\mathrm{cp}}(\mathbf{L}_t) = \mathbf{q}\,\mathbf{I}_t$ by using the commutative diagram on page 43. Since $\overline{\mathrm{cp}}$ is an endomorphism of $\mathbb{C}\langle\langle \mathsf{Y}\rangle\rangle$-modules we get from Eq. (13.19) and Eq. (13.20) that $\mathbf{L} = \overline{\mathrm{cp}}\big(\exp\big((2t+\gamma)\mathtt{y}_{1,1}\big)\big)\mathbf{q}\pi_Y(\mathbf{I})$. Now that all the coefficients of $\mathbf{L}$ and $\mathbf{I}$ are in $\mathbb{C}$ we see that $\overline{\mathrm{cp}}\big(\exp\big((2t+\gamma)\mathtt{y}_{1,1}\big)\big)$ must be in $\mathbb{C}[\![\mathtt{y}_{1,1}]\!]$. This completes the proof of the proposition. □

Lemma 13.3.29. *The group-like element* $\exp\big(T\mathtt{x}_1\big)\,\mathbf{I}$ *in* $\mathbb{C}[T]\langle\langle \mathsf{X}\rangle\rangle$ *is the generating series of* $\mathrm{L}_{ш}(\mathbf{s}|\mathbf{i};T)$. *Namely,*

$$\exp\big(T\mathtt{x}_1\big)\,\mathbf{I} = \sum_{\substack{d\in\mathbb{N}_0,\mathbf{s}=(s_1,\ldots,s_d)\in\mathbb{N}^d \\ \mathbf{i}=(i_1,\ldots,i_d)\in(\mathbb{Z}/N\mathbb{Z})^d}} \mathrm{L}_{ш}(\mathbf{s}|\mathbf{i};T)\mathtt{x}_0^{s_1-1}\mathtt{x}_{\mu^{i_1}}\cdots\mathtt{x}_0^{s_d-1}\mathtt{x}_{\mu^{i_d}}.$$

Proof. By the proof of Prop. 13.3.13 we know the generating series of $Ł_{ш}(\mathbf{s}|\mathbf{i};T)$ is group-like for Δ since $Ł_{ш}(\mathbf{s}|\mathbf{i};T)$ satisfies the shuffle relations. Then we can apply Lemma 13.3.20 to deduce our result since $Ł_{ш}(1|0;T)=T$ by Prop. 13.3.8. □

Next, we want to determine S explicitly.

Theorem 13.3.30. *The relation between* $\mathbf{L}$ *and* $\mathbf{I}$ *is given by*

$$\mathbf{L}=\mathbf{S}(\mathtt{y}_{1,1})\cdot\mathbf{q}\pi_{\mathsf{Y}}(\mathbf{I}),$$

where $\mathbf{S}(u)=\exp\left(\sum_{n=2}^{\infty}\frac{(-1)^{n-1}\zeta(n)}{n}u^n\right)$.

Proof. This follows easily from Thm. B.22 in Appendix B. □

Let's recall the linear map defined in Eq. (3.3.17):

$$\rho:\mathbb{R}[T]\to\mathbb{R}[T],$$
$$\rho(e^{Tu})\mapsto A(u)e^{Tu},$$

where $A(u)=\mathbf{S}(u)^{-1}$. Let $\mathbf{L}':=\exp(T\mathtt{y}_{1,1})\cdot\mathbf{L}$ and $\mathbf{I}':=\exp(T\mathtt{x}_1)\cdot\mathbf{I}$. By Lemma 13.3.21, the coefficients of $\mathtt{y}_{1,1}$ and $\mathtt{x}_1$ in $\mathbf{L}'$ and $\mathbf{I}'$ are both equal to T. Note that $\mathtt{y}_{1,1}=\mathtt{x}_1$. For any $\mathbf{s}=(s_1,\dots,s_n)\in\mathbb{N}^n$ and $\mathbf{i}=(i_1,\dots,i_n)\in(\mathbb{Z}/N\mathbb{Z})^n$, by Lemma 13.3.25 (resp. Lemma 13.3.29) $Ł_*(\mathbf{s}|\mathbf{i};T)$ (resp. $Ł_{ш}(\mathbf{s}|\mathbf{i};T)$) is the coefficient of $\mathtt{y}_{s_1,\mu^{i_1}}\cdots\mathtt{y}_{s_n,\mu^{i_n}}$ (resp. $\mathtt{x}_0^{s_1-1}\mathtt{x}_{\mu^{i_1}}\cdots\mathtt{x}_0^{s_n-1}\mathtt{x}_{\mu^{i_n}}$, in $\mathbf{L}'$ (resp. $\mathbf{I}'$). Finally we prove the regularized DBSFs as given in Thm. 13.3.9 which can be reformulated as follows.

Theorem 13.3.31. *For any indices* $s_1,\dots,s_n\in\mathbb{N}$ *and* $i_1,\dots,i_n\in\mathbb{Z}/N\mathbb{Z}$, *we have*

$$\rho\big(Ł_*(s_1,\dots,s_n|i_1,i_2-i_1,\dots,i_n-i_{n-1};T)\big)=Ł_{ш}(s_1,\dots,s_n|i_1,\dots,i_n;T).$$

Proof. Dividing both sides of the equation in Thm. 13.3.30 by $\mathbf{S}(\mathtt{y}_{1,1})$, we obtain

$$\begin{aligned}A(\mathtt{y}_{1,1})\cdot\mathbf{L}'&=\exp(T\mathtt{y}_{1,1})\cdot A(\mathtt{y}_{1,1})\cdot\mathbf{L}\\&=\exp(T\mathtt{y}_{1,1})\cdot\mathbf{q}\pi_{\mathsf{Y}}(\mathbf{I})\\&=\mathbf{q}\pi_{\mathsf{Y}}(\exp(T\mathtt{x}_1)\,\mathbf{I})=\mathbf{q}\pi_{\mathsf{Y}}(\mathbf{I}').\end{aligned}$$

Hence

$$\rho(\mathbf{L}')=\rho(\exp(T\mathtt{y}_{1,1})\cdot\mathbf{L})=A(\mathtt{y}_{1,1})\cdot\exp(T\mathtt{y}_{1,1})\cdot\mathbf{L}=\mathbf{q}\pi_{\mathsf{Y}}(\mathbf{I}').$$

Finally, comparing the coefficient of $\mathtt{y}_{s_1,\mu^{i_1}}\mathtt{y}_{s_1,\mu^{i_2-i_1}}\cdots\mathtt{y}_{s_n,\mu^{i_n-i_{n-1}}}$ on both sides of the equation yields the theorem quickly. □

13.3.3 *Weight one relations*

If the level $N > 3$ then there are many non-trivial linear relations in $\mathsf{CMZV}(1, N)$ of weight one whose structure is known to us, which we now review briefly.

Fix the Nth root of unity $\mu = \exp(2\pi i/N)$ as above. We know the cyclotomic units can be constructed using $1 - \mu^j$ for $1 \leq j < N$ and we have

$$-\log(1-\mu) = \sum_{k>0} \frac{\mu^k}{k} = \mathrm{Li}_1(\mu).$$

Here we have chosen the branch cut of the logarithm by deleting the non-positive real numbers $(-\infty, 0]$ from the complex plane. The next lemma is straight-forward.

Lemma 13.3.32. *All of the following relations hold of* Li_1.

(i) *(Symmetry relations). For all* $1 < j < N/2$

$$\mathrm{Li}_1(\mu^j) - \mathrm{Li}_1(\mu^{-j}) = \frac{(N-2j)\pi i}{N} = \frac{N-2j}{N-2}\big(\mathrm{Li}_1(\mu) - \mathrm{Li}_1(\mu^{-1})\big).$$

(ii) *(Distribution relations). For any divisor* d *of* N *and* $1 \leq a < d$ *we have*

$$\sum_{0 \leq j < N/d} \mathrm{Li}_1(\mu^{a+dj}) = \mathrm{Li}_1(\mu^{aN/d}).$$

Proof. (i) is straight-forward by Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ and the double angle formula. (ii) follows from the identity

$$\sum_{j=1}^{N} \mu^{jM} = \begin{cases} N, \text{ if } N|M; \\ 0, \text{ otherwise.} \end{cases}$$

We leave the details as Exercise 13.1. □

Definition 13.3.33. For any level N, one may multiply any relation in Lemma 13.3.32 by any level N CMZV of weight $w-1$ to produce a non-trivial linear relation among some CMZVs of weight w (after using either stuffle or shuffle product to expand the products). We call such a relation a *weight one relation.*

Notice that the weight one relations have the following forms in words: for all $w \in \mathfrak{A}^0$

$$\begin{cases} (N-2)\,\mathrm{Ł}(\mathrm{y}_{1,j} * w - \mathrm{y}_{1,-j} * w) = (N-2j)\,\mathrm{Ł}(\mathrm{y}_{1,1} * w - \mathrm{y}_{1,-1} * w), \\ \displaystyle\sum_{0 \leq j < d} \mathrm{Ł}(\mathrm{y}_{1,a+jN/d} * w) = \mathrm{Ł}(\mathrm{y}_{1,ad} * w). \end{cases} \tag{13.21}$$

where $\mathtt{y}_{n,j} = a^{n-1}b_j$ as in Sec. 13.3.1

The proof of the following theorem of Bass [35] (corrected by Ennola [186]) will take us too far afield so we have to refer the interested reader to their original papers. We will not use it in the rest of the book.

Theorem 13.3.34. *All $\mathbb{Q}$-linear relations among $\{\mathrm{Li}_1(\mu^j) : j = 1, \ldots, N-1\}$ are consequences of relations I and II.*

13.3.4 *Regularized distribution relations*

It is well-known that MPLs satisfy the following distribution formula: for all $d \in \mathbb{N}$, $\mathbf{s} \in \mathbb{N}^d$ and $\xi_1, \ldots, \xi_n \in \Gamma_N$

$$\mathrm{Li}_{\mathbf{s}}(\xi_1, \ldots, \xi_n) = d^{|\mathbf{s}|-n} \sum_{z_j^d = \xi_j, 1 \le j \le n} \mathrm{Li}_{\mathbf{s}}(z_1, \ldots, z_n). \tag{13.22}$$

When $s_1 = 1$ one has to exclude the divergent case $\xi_1 = 1$. We call these *finite distribution relations* (finite DISTs).

We now consider the regularized version of Eq. (13.22). For any divisor d of N let $\imath_d : \Gamma_{N/d} \hookrightarrow \Gamma_N$ and let $\rho^d : \Gamma_N \twoheadrightarrow \Gamma_{N/d}$ be the dth power map. They induce two algebra homomorphisms, respectively:

$$\begin{aligned} \rho_*^d : \mathbb{Q}\langle \mathsf{X}_N \rangle &\longrightarrow \mathbb{Q}\langle \mathsf{X}_{N/d} \rangle \\ \mathtt{x}_\xi &\longmapsto \begin{cases} d\mathtt{x}_0, & \text{if } \xi = 0, \\ \mathtt{x}_{\xi^d}, & \text{if } \xi \in \Gamma_N, \end{cases} \end{aligned}$$

and

$$\begin{aligned} \imath_d^* : \mathbb{Q}\langle \mathsf{X}_N \rangle &\longrightarrow \mathbb{Q}\langle \mathsf{X}_{N/d} \rangle \\ \mathtt{x}_\xi &\longmapsto \begin{cases} \mathtt{x}_\xi, & \text{if } \xi \in \Gamma_{N/d} \cup \{0\}, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

It is easy to see that both $\imath_d^*$ and ρ_*^d are Δ-coalgebra homomorphisms such that $\imath_d^*(\mathbf{I})$ and $\rho_*^d(\mathbf{I})$ have the same image under the map π_{cv}.

Theorem 13.3.35. *For every divisor d of N*

$$\rho_*^d(\mathbf{I}) = \exp\left(\sum_{\xi^d = 1, \xi \ne 1} \mathrm{Li}_1(\xi)\mathtt{x}_1 \right) \imath_d^*(\mathbf{I}). \tag{13.23}$$

Proof. It is not hard to see that Eq. (13.22) is equivalent to $\imath_d^*(\mathbf{L}_{\mathrm{cv}}) = \pi_{\mathrm{cv}}\rho_*^d(\mathbf{L})$. Since both $\imath_d^*$ and ρ_*^d commute with $\mathbf{p}$ and $\mathbf{q}$, we have $\imath_d^*(\mathbf{I}_{\mathrm{cv}}) =$

$\pi_{\mathrm{cv}}\mathfrak{p}_*^d(\mathbf{I})$. Notice that both $\mathfrak{i}_d^*$ and $\mathfrak{p}_*^d$ are coalgebra homomorphisms with respect to Δ_* we know $\mathfrak{i}_d^*(\mathbf{I})$ and $\mathfrak{p}_*^d(\mathbf{I})$ are both group-like. Further,

$$\pi_{\mathrm{cv}}\mathfrak{i}_d^*(\mathbf{I}) = \mathfrak{i}_d^*(\mathbf{I}_{\mathrm{cv}}) = \pi_{\mathrm{cv}}\mathfrak{p}_*^d(\mathbf{I}).$$

The relation Eq. (13.23) now follows immediately from Lemma 13.3.20. □

Combined with Prop. 13.3.42 the above result provides the so-called *regularized distribution relations* (regularized DISTs) which of course include all the finite DISTs of CMZVs given by Eq. (13.22).

However, sometimes finite DISTs are not independent of the other relations. In the next theorem one sees that when the weight $w = 2$ and level N is a prime, all the DISTs in Eq. (13.22), where $\mathtt{x}_j = 1$ for all j, are consequences of the regularized DBSFs of CMZVs of level N.

Theorem 13.3.36. *For any prime p write $L(i,j) = L_p(1,1|i,j)$ and $D(i) = L_p(2|i)$. Define for $1 \le i,j < p$:*

$$\mathrm{FDT} := -D(0) + p\sum_{j=0}^{p-1} D(j), \qquad \mathrm{RDS}(i) := D(i) + L(i,0) - L(i,-i),$$

$$\mathrm{FDS}(i,j) := D(i+j) + L(i,j) + L(j,i) - L(i,j-i) - L(j,i-j).$$

Then one has

$$\mathrm{FDT} = \sum_{1\le i<p} \mathrm{FDS}(i,i) + 2\sum_{1\le j<i<p} \mathrm{FDS}(i,j) + 2\sum_{i=1}^{p-1} \mathrm{RDS}(i). \tag{13.24}$$

Proof. When $p = 2$ the second term on the right-hand side of Eq. (13.24) is vacuous. Then it is easy to see that both sides of Eq. (13.24) are equal to $D(0) + 2D(1)$.

We now assume $p \ge 3$. Changing the order of summation yields that

$$\begin{aligned}
&2\sum_{1\le j<i<p} D(i+j) = \sum_{i=2}^{p-1}\sum_{j=1}^{i-1} D(i+j) + \sum_{j=1}^{p-2}\sum_{i=j+1}^{p-1} D(i+j)\\
=&\sum_{i=2}^{p-2}\sum_{i\ne j=1}^{p-1} D(i+j) + \sum_{j=1}^{p-2} D(j-1) + \sum_{i=2}^{p-1} D(i+1)\\
=&(p-3)\sum_{j=0}^{p-1} D(j) - \sum_{i=2}^{p-2} D(i) - \sum_{i=1}^{p-1} D(2i) + \sum_{j=1}^{p-2} D(j) + \sum_{j=2}^{p-1} D(j) + 2D(0)\\
=&(p-1)D(0) + (p-3)\sum_{j=1}^{p-1} D(j)
\end{aligned}$$

since $\sum_{j=0}^{p-1} D(i+j) = \sum_{j=0}^{p-1} D(j)$ for all i and $\sum_{i=1}^{p-1} D(2i) = \sum_{i=1}^{p-1} D(i)$. This implies that the dilogarithms on the right-hand side of Eq. (13.24) exactly add up to a finite DIST. Thus one only needs to show that all the double logarithms on the right-hand side of Eq. (13.24) cancel out.

First observe that $L(i,0)$ in $\mathrm{FDS}(i,i)$ and $\mathrm{RDS}(i)$ cancel out each other. Now let us consider the lattice points (i,j) of $\mathbb{Z}^2$ corresponding to $L(i,j)$. The points (i,j) corresponding to $L(i,j)$ with positive signs fill in exactly the area inside the square $[1,p-1]\times[1,p-1]$ (including boundary): $L(i,i)$ in $\mathrm{FDS}(i,i)$ provides the diagonal $y=x$, $\sum_{1\le j<i<p} L(i,j)$ (resp. $\sum_{1\le j<i<p} L(j,i)$) form the lower right (resp. upper left) triangular region.

For the negative terms of the double logs, $L(i,-i)$ in $\mathrm{RDS}(i)$ provides the diagonal $x+y=p$, $\sum_{1\le j<i<p} L(i,j-i) = \sum_{i=2}^{p-1}\sum_{j=p+1-i}^{p-1} L(i,j)$ form the upper right triangular region. Similarly, by changing the order of summation $\sum_{1\le j<i<p} L(j,i-j) = \sum_{i=1}^{p-2}\sum_{j=i+1}^{p-1} L(i,j-i) = \sum_{i=1}^{p-2}\sum_{j=1}^{p-1-i} L(i,j)$ fills the lower left region. □

To conclude this section we remark that numerical evidence up to level $N=169$ supports the following

Conjecture 13.3.37. *In weight two, all regularized DISTs are consequences of the weight one relations, the regularized DBSFs and the depth two finite DISTs.*

13.3.5 *Lifted relations from lower weights*

When we deal with MZVs (resp. Euler sums) we hope that all the linear relations come from the regularized DBSFs (resp. regularized DBSFs and DISTs). Since there are no weight one relations when level $N=3$ it is natural to ask if the regularized DBSFs and DISTs of weight w are enough to produce all the linear relations among weight w CMZVs at level 3. Surprisingly, the answer is no.

The first counterexample is in weight four, *i.e.*, $(w,N)=(4,3)$. Easy computation shows that there are 144 CMZVs in this case among which there are 239 nontrivial regularized DBSFs of weight four which include 191 finite DBSFs of weight four (see Eq. (13.11) and Eq. (13.12)). Furthermore, it is easy to verify that all the seven regularized DISTs (including four finite DISTs) can be derived from regularized DBSFs. Using these relations we get 127 independent linear relations among the 144 CMZVs. But we expect to have $d(4,3)\le 16$ by Thm. 13.2.1, so there must be at least one more

relation linearly independent from the above. Where else can we find it? The answer is the so-called *lifted relations*.

We know that a product of two weight two CMZVs is of weight four. So on each of the five regularized DBSFs (including two finite DBSFs) of weight two in $\mathsf{CMZV}(2,3)$ we can multiply any one of the nine CMZVs of $(w,N)=(2,3)$ to get a relation in $\mathsf{CMZV}(4,3)$. For instance, we have a finite DBSF

$$Ł(\mathtt{y}_{1,1} * \mathtt{y}_{1,1} - \mathtt{y}_{1,1} \sqcup\!\sqcup \mathtt{y}_{1,1}) = L_3(2|2) + 2L_3(1,1|1,1) - L_3(1,1|1,0) = 0.$$

Here, as we mentioned in Remark 13.3.11 earlier, $\mathtt{y}_{1,1} = \mathtt{x}_\mu = b_1$ (since we are using it in a concrete computation with the aid of a computer). Multiplying by $L_3(1,1|1,1) = Ł(\mathtt{y}_{1,1}\mathtt{y}_{1,2})$ we obtain a new relation which is linearly independent from the regularized DBSFs of weight four in $\mathsf{CMZV}(4,3)$:

$$\begin{aligned}
&\zeta\big(\mathtt{y}_{1,1}\mathtt{y}_{1,2} \sqcup\!\sqcup (\mathtt{y}_{2,0} + 2\mathtt{y}_{1,1}\mathtt{y}_{1,2} - 2\mathtt{y}_{1,1}\mathtt{y}_{1,0})\big)\\
=&L_3(1,1,2|1,1,0) + 2L_3(1,2,1|1,1,0) + 2L_3(2,1,1|1,1,0)\\
+&L_3(2,1,1|2,2,1) + 4L_3(\{1\}^4|1,1,2,1) + 8L_3(\{1\}^4|1,0,1,0)\\
-&6L_3(\{1\}^4|1,0,0,1) - 4L_3(\{1\}^4|1,0,1,2) - 2L_3(\{1\}^4|1,1,2,0) = 0.
\end{aligned}$$

Such relations coming from the lower weights are called *lifted relations.* In this way, when $(w,N)=(4,3)$ we can produce 45 lifted regularized DBSFs from weight two, 58 from weight three. We may also lift regularized DISTs and obtain nine and six relations from weight two and three, respectively. However, all the lifted relations together only produce one new linearly independent relation, as expected. Hence we find totally 128 linearly independent relations among the 144 CMZVs with $(w,N)=(4,3)$. This implies that $d(4,3) \leq 16$ which is the same bound obtained by Thm. 13.2.1 and is proved to be exact under a variant of Grothendieck's period conjecture by Deligne (see [155]).

For levels $N \geq 4$ one may lift not only regularized DBSFs and DISTs but also the weight one relations. But by a moment reflection one sees that the lifted weight one relations are still weight one relations by themselves so one does not really need to consider them after all.

Definition 13.3.38. We call a $\mathbb{Q}$-linear relation among CMZVs *standard* if it can be produced by a $\mathbb{Q}$-linear combination of the following four families of relations: the regularized DBSFs, the regularized DISTs, the lifted relations from the above two types, and the weight one relations. Otherwise, it is called a *non-standard* relation.

For general levels N and weights w, there are no strict inclusion relations among the four families of the standard relations. But for specific choice of N and w there may exist some. For an earlier example, see Thm. 13.3.36.

Computation in small weight cases supports the following conjecture.

Conjecture 13.3.39. *Suppose $N = 3$ or 4. Every CMZV of level N is a linear combination of CMZVs of the form $L(\{1\}^w|t_1,\dots,t_w)$ with $t_j \in \{1,2\}$. Consequently, the $\mathbb{Q}$-dimension of the CMZVs of weight w and level N is given by $d(w,N) = 2^w$ for all $w \geq 1$.*

Remark 13.3.40. The data in Table 14.1 shows that one cannot produce enough relations by using only the standard relations when $(w,N) = (3,4)$. In fact, even though one has $d(3,4) \leq 8$ and $d(4,4) \leq 16$ by [156, 5.25], one can only show that $d(3,4) \leq 9$ and $d(4,4) \leq 21$ by using only the standard relations. However, thanks to the octahedral symmetry of $\mathbb{P}^1 - (\{0,\infty\} \cup \boldsymbol{\mu}_4)$ one can find (presumably all) the non-standard relations in these two cases (see Thm. 14.6.1).

Remark 13.3.41. Let $N = 2,3,4$ or 8. Assuming a variant of Grothendieck's period conjecture, Deligne constructed explicitly a set of basis for $\mathsf{CMZV}(w,N)$ in [155]. His results would also imply that $d(w,2)$ is given by the Fibonacci numbers, $d(w,3) = d(w,4) = 2^w$, and $d(w,8) = 3^w$ under Grothendieck's period conjecture.

13.3.6 *The group-like element* $\mathbf{I}$

As in Sec. 13.3.2 let $\mathbf{I}$ be the unique group-like element in $(\mathbb{C}\langle\langle \mathsf{X} \rangle\rangle, \Delta)$ whose coefficients of a and b_1 are 0 such that $\pi_{\mathrm{cv}}(\mathbf{I}) = \mathbf{I}_{\mathrm{cv}}$. In order to do the numerical computation one needs to determine explicitly the coefficients for $\mathbf{I}$. Put

$$\mathbf{I} = \sum_{k \in \mathbb{N}_0, \xi_1,\dots,\xi_k \in \Gamma_N \cup \{0\}} C(\xi_1,\dots,\xi_k) \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_k}. \tag{13.25}$$

Proposition 13.3.42. *Let k, m and n be three non-negative integers. If $k > 0$ then we assume $\xi_1 \neq 1$ and $\xi_k \neq 0$. Set $(\xi_1,\dots,\xi_k,\{0\}^n) =$*

$(\xi_1, \dots, \xi_q)$. *Then*

$$C(\{1\}^m, \xi_1, \dots, \xi_k, \{0\}^n) = \begin{cases} 0, & \text{if } mn = k = 0; \\ Ł\left(\pi_{\mathrm{cv}}(\mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_k})\right), & \text{if } m = n = 0; \\ -\dfrac{1}{m} \displaystyle\sum_{i=1}^{q} C(\{1\}^{m-1}, \xi_1, \dots, \xi_i, 1, \xi_{i+1}, \dots, \xi_q), & \text{if } m > 0; \\ -\dfrac{1}{n} \displaystyle\sum_{i=1}^{k} C(\xi_1, \dots, \xi_{i-1}, 0, \xi_i, \dots, \xi_k, \{0\}^{n-1}), & \text{if } m = 0, n > 0. \end{cases} \tag{13.26}$$

Here Ł *is defined by Eq.* (13.7) *after using the identification given by Remark* 13.3.12.

Remark 13.3.43. This proposition provides the recursive relations one may use to compute all the coefficients of $\mathbf{I}$.

Proof. Since $\mathbf{I}$ is group-like one has

$$\Delta\, \mathbf{I} = \mathbf{I} \otimes \mathbf{I}\,. \tag{13.27}$$

The first case follows from this immediately since $C(0) = C(1) = 0$. The second case is essentially the definition Eq. (13.7) of Ł. If $m > 0$ then one can compare the coefficient of

$$\mathtt{x}_1 \otimes \mathtt{x}_1^{m-1} \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_q} \tag{13.28}$$

of the two sides of Eq. (13.27) and find the relation Eq. (13.26). Indeed, the coefficient on the right-hand side comes from only one possible tensor product which leads to the coefficient $C(1)C(\{1\}^{m-1}, \xi_1, \dots, \xi_k, \{0\}^n) = 0$ since $C(1) = 1$ by Definition 13.3.22 of $\mathbf{I}$. On the left-hand side, the term Eq. (13.28) can be produced by $\Delta(\mathtt{x}_1^m \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_q})$ which gives m copies, and produced by $\Delta(\mathtt{x}_1^{m-1} \mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_i} \mathtt{x}_{\xi_1} \mathtt{x}_{\xi_{i+1}} \cdots \mathtt{x}_{\xi_q})$ for $i = 1, \dots, q$ each of which gives only one copy.

Finally, if $m = 0$ and $n > 0$ then one may consider the coefficient of $\mathtt{x}_{\xi_1} \cdots \mathtt{x}_{\xi_k} a^{n-1} \otimes a$ in Eq. (13.27). This finishes the proof of the proposition. □

Theorem 13.3.44. *Let* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ *and* $\xi_1, \dots, \xi_d \in \Gamma_N$. *Then*

$$C(\{0\}^{s_1-1}, \xi_1, \dots, \{0\}^{s_d-1}, \xi_d, \{0\}^k) = (-1)^k \sum_{\substack{t_1+\cdots+t_d=k, \\ \mathbf{t}=(t_1,\dots,t_d)\in(\mathbb{N}_0)^d}} \prod_{j=1}^{d} \binom{s_j + t_j - 1}{t_j} Ł_{ш}\left((\mathbf{s}+\mathbf{t})|\mathbf{i}; 0\right)$$

where $\mathbf{i} = (i_1, \dots, i_d)$ *such that* $\xi_j = \mu^{i_j}$ *for all* $j = 1, \dots, d$.

Proof. By induction and Prop. 13.3.42 it is not hard to reduce the proof of the theorem to the case when $k = 0$:

$$C(\{0\}^{s_1-1}, \xi_1, \dots, \{0\}^{s_d-1}, \xi_d) = \text{Ł}_{\sqcup\!\sqcup}(\mathbf{s}|\mathbf{i}; 0)$$

which follows from Lemma 13.3.29 by setting $T = 0$. □

By taking dual $t \to 1 - t$ we get

Theorem 13.3.45. *Let* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ *and* $\xi_1, \dots, \xi_d \in (\Gamma_N \setminus \{1\}) \cup \{0\}$. *Then*

$$C(\{1\}^k, \xi_1, \{1\}^{s_1-1}, \dots, \xi_d, \{1\}^{s_d-1}) = (-1)^k \sum_{\substack{t_1,\dots,t_d \in \mathbb{N}_0 \\ t_1+\cdots+t_d=k}} C\Big(\big\{\xi_j, \{1\}^{s_j+t_j-1}\big\}_{j=1}^d\Big) \prod_{j=1}^d \binom{s_j + t_j - 1}{t_j}.$$

13.4* Generalized Drinfeld associators

It is possible to generalize the setup in Definition 4.1.2 to the higher levels. For $N \in \mathbb{N}$, let $\Psi_N(\mathtt{x}_0, \mathtt{x}_\xi : \xi \in \Gamma_N)$ be the generalized associator defined as the renormalized holonomy from 0 to 1 of

$$\frac{\mathrm{d}G}{\mathrm{d}z} = \left(\frac{\mathtt{x}_0}{z} + \sum_{\xi \in \Gamma_N} \frac{\mathtt{x}_\xi}{\xi^{-1} - z}\right) G, \tag{13.29}$$

i.e., $\Psi_N = G_1^{-1} G_0$, where G_0, G_1 are the solutions of Eq. (13.29) on the open interval $(0, 1)$ such that $G_0(z) \simeq z^{\mathtt{x}_0}$ when $z \to 0^+$, $G_1(z) \simeq (1 - z)^{\mathtt{x}_1}$ when $z \to 1^-$ and $\mathtt{x}_0$, $\mathtt{x}_\xi$ $(\xi \in \Gamma_N)$ are free variables. When $N = 1$ $\Psi_1(\mathtt{x}_0, \mathtt{x}_1) = \Phi(\mathtt{x}_0, \mathtt{x}_1)$ is the KZ associator defined as in Definition 4.1.2. Similar to $\Phi(\mathtt{x}_0, \mathtt{x}_1)$ it is not very difficult to see that Ψ_N is the generating series for CMZVs.

Theorem 13.4.1. *The generalized Drinfeld associator* $\Psi_N(\mathtt{x}_0, \mathtt{x}_\xi : \xi \in \Gamma_N)$ *is the only element in Hopf algebra* $(\mathfrak{R}, \Delta_{\sqcup\!\sqcup}, \epsilon_{\sqcup\!\sqcup})$ *such that*

(i) $\Delta_{\sqcup\!\sqcup}(\Psi_N \Phi) = \Psi_N \otimes \Psi_N$ *(*i.e. Ψ_N *is group-like),*
(ii) $\Psi_N[\mathtt{x}_0] = \Psi_N[\mathtt{x}_1] = 0$,
(iii) $\Psi_N[\mathbf{w}] = (-1)^d \operatorname{Li}_{s_1,\dots,s_d}(\xi_1, \xi_2\xi_1^{-1}, \dots, \xi_d\xi_{d-1}^{-1})$ $\forall \mathbf{w} = \mathtt{x}_0^{s_1-1}\mathtt{x}_{\xi_1} \cdots \mathtt{x}_0^{s_d-1}\mathtt{x}_{\xi_d} \in \mathfrak{A}^0$.

Proof. This will follow from Lemma 13.3.20 immediately. □

To have a more thorough understanding of these generalized Drinfeld associators we need to study some complicated Lie algebras as follows. This is an extension of the objects defined in Definitions 4.1.16 to 4.1.21.

Definition 13.4.2. Let $\mathbb{K}$ be a field of characteristic 0. For $n \geq 2$, denote by $\mathfrak{t}_{n,N}$ the completed $\mathbb{K}$-Lie algebra with generators

$$t^{1i}\ (2 \leq i \leq n), \text{ and } t(\alpha)^{ij}\ (i \neq j, 2 \leq i, j \leq n, \alpha \in \mathbb{Z}/N\mathbb{Z})$$

subject to the following relations

$$t(\alpha)^{ij} = t(-\alpha)^{ji}, \quad [t(\alpha)^{ij}, t(\alpha+\beta)^{ik} + t(\beta)^{jk}] = 0,$$
$$[t^{1i}, t(\alpha)^{jk}] = 0, \quad [t^{1i}, t^{1j} + \sum_{\gamma \in \mathbb{Z}/N\mathbb{Z}} t(\gamma)^{ij}] = 0,$$
$$[t(\alpha)^{ij}, t(\beta)^{kl}] = 0, \quad [t(\alpha)^{ij}, t^{1i} + t^{1j} + \sum_{\gamma \in \mathbb{Z}/N\mathbb{Z}} t(\gamma)^{ij}] = 0,$$

for all α, $\beta \in \mathbb{Z}/N\mathbb{Z}$ and all distinct i, j, k, l $(2 \leq i, j, k, l \leq n)$.

We note that $\mathfrak{t}_{n,1}$ is equal to $\mathfrak{t}_n$ for $n \geq 2$ (see Definition 4.1.16).

For a partially defined map $f : \{1, \ldots, m\} \to \{1, \ldots, n\}$ such that $f(1) = 1$, the Lie algebra morphism $\mathfrak{t}_{n,N} \to \mathfrak{t}_{m,N}$: $x \mapsto x^f = x^{f^{-1}(1),\ldots,f^{-1}(n)}$ is uniquely defined by

$$(t(\alpha)^{ij})^f = \sum_{i' \in f^{-1}(i), j' \in f^{-1}(j)} t(\alpha)^{i'j'} \quad (i \neq j, 2 \leq i, j \leq n)$$

and $(2 \leq j \leq n)$

$$(t^{1j})^f = \sum_{j' \in f^{-1}(j)} t^{1j'} + \frac{1}{2} \sum_{j', j'' \in f^{-1}(j)} \sum_{c \in \mathbb{Z}/N\mathbb{Z}} t(c)^{j'j''} + \sum_{i' \neq 1 \in f^{-1}(1), j' \in f^{-1}(j)} \sum_{c \in \mathbb{Z}/N\mathbb{Z}} t(c)^{i'j'}.$$

Again for a partially defined map $g : \{2, \ldots, m\} \to \{1, \ldots, n\}$, the Lie algebra morphism $\mathfrak{t}_n \to \mathfrak{t}_{m,N}$: $x \mapsto x^g = x^{g^{-1}(1),\ldots,g^{-1}(n)}$ is uniquely defined by

$$(t^{ij})^g = \sum_{i' \in g^{-1}(i), j' \in g^{-1}(j)} t(0)^{i'j'} \qquad (i \neq j, 1 \leq i, j \leq n).$$

Definition 13.4.3. Let σ be the cyclic permutation so that $\sigma(a_1, \ldots, a_N) = (a_2, \ldots, a_N, a_1)$. For $N \geq 1$, the Lie algebra $\mathfrak{grtm}_{(\bar{1},1)}(N, \mathbb{K})$ is defined to be the set of pairs $(\varphi, \psi) \in \mathfrak{grt}_1(\mathbb{K}) \times \mathfrak{t}^0_{3,N}$ satisfying the *mixed pentagon relation* in $\mathfrak{t}^0_{4,N}$

$$\psi^{1,2,34} + \psi^{12,3,4} = \varphi^{2,3,4} + \psi^{1,23,4} + \psi^{1,2,3}, \tag{13.30}$$

the *octagon relations* in $\mathfrak{t}^0_{3,N}$

$$\psi\big(a,\mathbf{b}\big)-\psi\big(a,\sigma(\mathbf{b})\big)+\psi\big(c,\sigma^{-2}(\overleftarrow{\mathbf{b}})\big)-\psi\big(c,\sigma^{-1}(\overleftarrow{\mathbf{b}})\big)=0 \tag{13.31}$$

with $\mathbf{b}=(b_0,b_1,\dots,b_{N-1})$ and $a+\sum_{\alpha\in\mathbb{Z}/N\mathbb{Z}}b_\alpha+c=0$, the *special derivation condition* in $\mathfrak{t}^0_{3,N}$

$$\sum_{a\in\mathbb{Z}/N\mathbb{Z}}\big[\psi\big(a,\sigma^\alpha(\mathbf{b})\big),b_\alpha\big]\;+\;\big[\psi\big(a,\mathbf{b}\big)\;-\;\psi\big(c,\sigma^{-1}(\overleftarrow{\mathbf{b}})\big),c\big]\;\;=\;\;0 \tag{13.32}$$

and $\psi[b_0]=0$. The Lie bracket is defined in Exercise 13.17.

Definition 13.4.4. For $N\geq 1$, the group $\mathrm{GRTM}_{(\bar{1},1)}(N,\mathbb{K})$ is defined to be the set of pairs $(g,h)\in\mathrm{GRT}_1(\mathbb{K})\times\exp\mathfrak{t}^0_{3,N}$ with $h[b_0]=0$ satisfying the *mixed pentagon relation* in $\exp\mathfrak{t}^0_{4,N}$

$$h^{1,2,34}h^{12,3,4}=g^{2,3,4}h^{1,23,4}h^{1,2,3}, \tag{13.33}$$

the *octagon relations* in $\exp\mathfrak{t}^0_{3,N}$

$$h\big(c,\sigma^{-2}(\overleftarrow{\mathbf{b}})\big)h\big(c,\sigma^{-1}(\overleftarrow{\mathbf{b}})\big)^{-1}h\big(a,\mathbf{b}\big)=h\big(a,\sigma(\mathbf{b})\big) \tag{13.34}$$

with $a+\sum_{\alpha\in\mathbb{Z}/N\mathbb{Z}}b_\alpha+c=0$ and the *special action condition* in $\exp\mathfrak{t}^0_{3,N}$

$$a+\sum_{\alpha\in\mathbb{Z}/N\mathbb{Z}}\mathrm{Ad}(\tau_\alpha h^{-1})(b_\alpha)+\mathrm{Ad}\Big(h^{-1}\cdot h\big(c,\sigma^{-1}(\overleftarrow{\mathbf{b}})\big)\Big)(c)=0 \tag{13.35}$$

where τ_α is the automorphism defined by $a\mapsto a$ and $b_\gamma\mapsto b_{\gamma+\alpha}$ for all $\gamma\in\mathbb{Z}/N\mathbb{Z}$. The corresponding group multiplication is defined in Exercise 13.18.

Definition 13.4.5. The torsor $\mathrm{Pseudo}_{(\bar{1},1)}(N,\mathbb{K})$ is defined to be the set of pairs $(g,h)\in\mathrm{M}_1(\mathbb{K})\times\exp\mathfrak{t}^0_{3,N}$ with $h[b_0]=0$ satisfying the mixed pentagon equation (13.33) and the following variant of octagon equation in $\exp\mathfrak{t}^0_{3,N}$

$$e^{\frac{b_1}{2}}h\big(c,\sigma^{-2}(\overleftarrow{\mathbf{b}})\big)e^{\frac{c}{N}}h\big(c,\sigma^{-1}(\overleftarrow{\mathbf{b}})\big)^{-1}e^{\frac{b_0}{2}}h\big(a,\mathbf{b}\big)e^{\frac{a}{N}}=h\big(a,\sigma(\mathbf{b})\big). \tag{13.36}$$

It can be shown that $\mathrm{Pseudo}_{(\bar{1},1)}(N,\mathbb{K})$ is a right $\mathrm{GRTM}_{(\bar{1},1)}(N,\mathbb{K})$-torsor. See Exercise 13.19.

13.5 Historical notes

Most of the recent studies of MPLarithms are prompted by their intimate relations to theoretical physics (the section 14.7 for the relevant history).

In the literature, there exist many different names for CMZVs, such as MPL values at roots of unity [599], the MZVs rel. μ_N [156], or the multiple L-values [18]. Our use of the word "color" to denote the roots of unity

is influenced by [48, 49] and has been used by some other authors under similar context (for e.g., [357, 412, 603]). CMVZs are also closley related to the cyclotomic harmonic polylogs studied in [3].

In previous chapters, we have studied the level one case of the CMZVs, *i.e.*, the MZVs in details. In this chapter, we have turned to the algebraic structures of CMZVs of higher levels, paying more attention to the $\mathbb{Q}$-linear relations between these values. Since the linear dependence problem is directly related to some very hard transcendental number theory we are content with the upper bound of the dimensions for the moment. For the interested reader, a precise formulation of this problem is contained in the statement of a variant of Grothendieck's period conjecture which can be found in [155] or in § 5.27(c) of [156].

From the point of view of Lyndon words and quasi-symmetric functions Bigotte *et al.* have studied CMZVs (they call them *colored MZVs*) primarily by using DBSFs and monodromy argument (cf. [48, Thm. 5.1]). However, when the level $N \geq 2$, these DBSFs often are not complete, as we have seen in this chapter (for level two, see also [59]).

Thm. 13.2.1 was proved by Deligne and Goncharov in their seminal joint paper [156, § 5.25] using the theory of motivic fundamental groups of $\mathbb{P}^1 - (\{0, \infty\} \cup \boldsymbol{\mu}_N)$. Their machinery has since become the driving force for a series of new discoveries by many followers such as Brown [104–106] and his student Jarossay [306–308].

It is Kontsevich who first noticed that MZVs can be represented by iterated integrals (see [348]). Regularized DBSFs of MZVs are also called EDS (extended DBSFs) in [298]. Their approach was generalized to treat CMVZs by Arakawa and Kaneko [19]. Earlier, Racinet [465] extended DBSFs to CMZVs in the more abstract Lie and Hopf algebra setting. Section 13.3.1 was thus built on the results of [298, 465]. The main result, Thm. 13.3.9, is essentially [465, Cor. 2.24]. The presentation of the standard relations in Sec. 13.3 follows mostly that of [465] except for the lifted relations.

When $N \geq 4$ there exist linear relations among CMZVs of weight one Li_1 by a theorem of Bass [35]. Using the general weight one relations induced by them is one of the key ideas in finding the formula of [156, § 5.25] concerning $d(w, N)$. Further, it follows from the motivic theory of classical polylogs developed by Beilinson and Deligne [41] and the Borel's theorem (see [240, Thm. 2.1]) that the $\mathbb{Q}$-dimension of $\mathsf{CMZV}(1, N)$ is

$$d(1, N) = \dim K_1(\mathbb{Z}[\mu_N][1/N]) \otimes \mathbb{Q} + 1 = \varphi(N)/2 + \nu(N),$$

where φ is the Euler's totient function and $\nu(N)$ is the number of distinct

prime factors of N.

In [164] Drinfeld first showed that there exists a formal Lie series P with $\mathbb{Q}$-coefficients such that $\exp P(ht^{12}, ht^{23})$ is essentially $\Phi(a, b)$. More important is that his proof utilizes the properties of a remarkable group called the Grothendieck-Teichmüller group GT. He further showed that there is a canonical homomorphism from the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ into the groups $\mathrm{GT}(\mathbb{Q}_l)$ for all primes l. Many results were later generalized to higher level cases in [187, 188].

Enriquez showed in [188] that the special derivation condition (13.32) for ψ implies the octagon equation (13.31). However, in contrast to the level one case where the special derivation relations follows from duality and hexagon equations it is unknown if the level N special derivation relations (13.32) and (13.35) can be derived from the rest of the equations (see [188, Remark 7.8] and [191]).

In [209] Furusho showed that the mixed pentagon equation (13.33) implies the regularized DBSF among CMZVs.

Exercises

13.1. (i). Prove the symmetry relations hold for Li_1, *i.e.*, for all $1 < j < N/2$

$$\mathrm{Li}_1(\mu^j) - \mathrm{Li}_1(\mu^{-j}) = \frac{N-2j}{N-2}\big(\mathrm{Li}_1(\mu) - \mathrm{Li}_1(\mu^{-1})\big).$$

(ii). Complete the proof of the distribution relations: for any divisor d of N and $1 \le a < d$

$$\sum_{0 \le j < N/d} \mathrm{Li}_1(\mu^{a+dj}) = \mathrm{Li}_1(\mu^{aN/d}).$$

13.2. Prove Prop. 13.3.3by mimicking the proof of 3.2.4.

13.3. Prove Prop. 13.3.5 and Prop. 13.3.6 by mimicking the proof of 3.2.9.

13.4. Show that $\mathbb{C}\langle\langle \mathsf{X} \rangle\rangle$, $\mathbb{C}\langle\langle \mathsf{Y} \rangle\rangle$, $\mathbb{C}\langle\langle \mathsf{X} \rangle\rangle_{\mathrm{cv}}$ and $\mathbb{C}\langle\langle \mathsf{Y} \rangle\rangle_{\mathrm{cv}}$ are all graded bialgebras.

13.5. Show that

$$\mathcal{Y} := \sum_{n \in \mathbb{N}_0, \xi \in \Gamma_N} \mathtt{y}_{n,\xi}$$

is a group-like element in $(\mathbb{C}\langle\langle \mathsf{Y} \rangle\rangle, \bullet, \Delta_*)$.

13.6. Prove that $\mathbf{q}$ defined by Eq. (13.14) and $\mathbf{p}$ defined by Eq. (13.15) are inverses to each other.

13.7. For any $\lambda \in \mathbb{C}$, define the $\mathbb{C}$-linear maps $\phi_\lambda, \psi : \mathbb{C}\langle\langle Y \rangle\rangle \longrightarrow \mathbb{C}\langle\langle Y \rangle\rangle$ by the following:

$$\phi_\lambda(\mathtt{y}_{n,j}) = \lambda^n \mathtt{y}_{n,j}, \qquad \psi(\mathtt{y}_{n,j}) = \mu^j \mathtt{y}_{n,j}.$$

Show that both ϕ_λ and ψ commute with Δ_* defined by Eq. (B.1).

13.8. Prove Eq. (13.18) by comparing the coefficient of $\mathtt{y}_{s_1,\xi_1} \cdots \mathtt{y}_{s_d,\xi_d}$ on both sides.

13.9. Verify that $\mathtt{y}_{1,1}\mathbb{C}\langle\langle \mathsf{Y} \rangle\rangle$ is a co-ideal for Δ_* and that the projection π_{cv} is a coalgebra homomorphism.

13.10.). Prove Lemma 13.3.21 by taking $\mathfrak{g}_1 = \mathbb{C}\mathtt{y}_{1,1}$ and applying Cor. B.26.

13.11. Prove the FDT stated in Eq. (13.22). Then show that it is equivalent to $\imath_d^*(\mathbf{L}_{\mathrm{cv}}) = \pi_{\mathrm{cv}} \boldsymbol{p}_a^d st(\mathbf{L})$.

13.12. Prove that both $\imath_d^*$ and $\boldsymbol{p}_a^d st$ commute with $\mathbf{p}$ and $\mathbf{q}$.

13.13. Prove that both $\imath_d^*$ and $\boldsymbol{p}_a^d st$ are coalgebra homomorphisms with respect to Δ_*.

13.14. Show that there is a natural injection $\mathfrak{t}_{n-1,N} \hookrightarrow \mathfrak{t}_{n,N}$ by $x \mapsto x^{0,\ldots,n-1}$.

13.15. Show that the Lie subalgebra $\mathfrak{f}_{n,N}$ of $\mathfrak{t}_{n,N}$ generated by t^{1n} and $t(\alpha)^{in}$ $(2 \le i \le n-1,\ \alpha \in \mathbb{Z}/N\mathbb{Z})$ is free of rank $(n-2)N+1$ and forms an ideal of $\mathfrak{t}_{n,N}$. Further, show that $\mathfrak{t}_{n,N}$ is a semi-direct product of $\mathfrak{f}_{n,N}$ and $\mathfrak{t}_{n-1,N}$.

13.16. Prove that the element $z_{n,N} = \sum_{1 \le i < j \le n} t^{ij}$ with $t^{ij} = \sum_{a \in \mathbb{Z}/N\mathbb{Z}} t(\alpha)^{ij}$ $(2 \le i < j \le n)$ is central in $\mathfrak{t}_{n,N}$. Let $\mathfrak{t}^0_{n,N}$ be the Lie subalgebra of $\mathfrak{t}_{n,N}$ with the same generators except t^{1n} and the same relations as $\mathfrak{t}_{n,N}$. Show that

$$\mathfrak{t}_{n,N} = \mathfrak{t}^0_{n,N} \oplus \mathbb{K} \cdot z_{n,N}.$$

In particular, when $n = 3$ $\mathfrak{t}^0_{3,N}$ is the free Lie algebra $\mathfrak{F}_{N+1}$ of rank $N+1$ with generators $a := t^{12}$ and $b_\alpha = t(\alpha)^{23}$ $(\alpha \in \mathbb{Z}/N\mathbb{Z})$.

13.17. Let $U\mathfrak{F}_{N+1} = \mathbb{K}\langle !\langle a, b_\alpha \rangle ! \rangle_{\alpha \in \mathbb{Z}/N\mathbb{Z}}$. For any $\mathbb{K}$-algebra homomorphism $\iota : U\mathfrak{F}_{N+1} \to S$ the image $\iota(\varphi) \in S$ is denoted by $\varphi\big(\iota(a), \iota(b_0), \ldots, \iota(b_{N-1})\big)$. The Lie algebra bracket of $\mathfrak{grtm}_{(\bar{1},1)}(N, \mathbb{K})$ is given by

$$\langle (\varphi_1, \psi_1), (\varphi_2, \psi_2) \rangle = (\langle \varphi_1, \varphi_2 \rangle, \langle \psi_1, \psi_2 \rangle) \tag{13.37}$$

with

$$\begin{aligned}\langle\varphi_1,\varphi_2\rangle &= [\varphi_1,\varphi_2]+D_{\varphi_2}(\varphi_1)-D_{\varphi_1}(\varphi_2),\\ \langle\psi_1,\psi_2\rangle &= [\psi_1,\psi_2]+\bar{D}_{\psi_2}(\psi_1)-\bar{D}_{\psi_1}(\psi_2).\end{aligned}$$

where $\bar{D}_\psi$ is the derivation of $\mathfrak{t}^0_{3,N}$ defined by

$$\bar{D}_\psi(a)=[\psi,a],\ \bar{D}_\psi(b_\alpha)=[\psi-\psi\big(a,b_\alpha,b_{\alpha+1},\dots,b_{\alpha-1}\big),b_\alpha]$$

for $\alpha\in\mathbb{Z}/N\mathbb{Z}$ and

$$\bar{D}_\psi(c)=[\psi\big(c,b_0,b_{N-1},\dots,b_1\big),c].$$

Show that the Lie bracket is well-defined. Furthermore, show that the Lie algebra structure is realized by the embedding

$$\mathfrak{grtm}_{(\bar{1},1)}(N,\mathbb{K})\hookrightarrow \mathrm{Der}(\mathfrak{t}^0_3)\times\mathrm{Der}(\mathfrak{t}^0_{3,N})$$

sending $(\varphi,\psi)\mapsto(D_\varphi,\bar{D}_\psi)$.

13.18. Show that $\mathrm{GRTM}_{(\bar{1},1)}(N,\mathbb{K})$ forms a group is we define the multiplication by

$$\begin{aligned}(g_1,h_1)\circ(g_2,h_2)=\Big(&g_2(a,b)\cdot g_1(a,\mathrm{Ad}(g_2^{-1})(b)),h_2\big(a,b_0,b_1,\dots,b_{N-1}\big)\cdot\\ &h_1\big(a,\mathrm{Ad}(h_2^{-1})b_0,\mathrm{Ad}(\tau_1h_2^{-1})b_1,\dots,\mathrm{Ad}(\tau_{N-1}h_2^{-1})b_{N-1}\big)\Big).\end{aligned}\tag{13.38}$$

Further, show that the group structure is realized by the embedding (with opposite homomorphism)

$$\mathrm{GRTM}_{(\bar{1},1)}(N,\mathbb{K})\hookrightarrow \mathrm{Aut}\,\mathfrak{t}^0_3\times\mathrm{Aut}\,\mathfrak{t}^0_{3,N}$$

sending (g,h) to the automorphism $(A_g,\bar{A}_h)$ where $\bar{A}_h$ is defined by $A\mapsto A$ and $b_\alpha\mapsto\mathrm{Ad}(\tau_ah^{-1})\,(b_\alpha)$ for $\alpha\in\mathbb{Z}/N\mathbb{Z}$. Finally, prove that its associated Lie algebra is $\mathfrak{grtm}_{(\bar{1},1)}(\mathbb{K})$.

13.19. Show that $\mathrm{Pseudo}_{(\bar{1},1)}(N,\mathbb{K})$ forms a right $\mathrm{GRTM}_{(\bar{1},1)}(N,\mathbb{K})$-torsor by (13.38) with $(g_1,h_1)\in\mathrm{Pseudo}_{(\bar{1},1)}(N,\mathbb{K})$ and $(g_2,h_2)\in\mathrm{GRTM}_{(\bar{1},1)}(N,\mathbb{K})$.

Chapter 14

Colored multiple zeta values at lower levels

In this chapter, we concentrate on the special cases of CMZVs when the levels are two, four and six. We start with a very useful online tool which can be used to check numerically if Euler sum identities are correct and also to detect $\mathbb{Q}$-linear relations between Euler sums.

14.1 EZ-face: an online computation resource

Recall the one variable MPLs defined in Sec. 4.2. Let $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$. Then for all $|z| < 1$, we have

$$\widetilde{\mathrm{Li}}_{\mathbf{s}}(z) = \int_0^z \left(\frac{\mathrm{d}t}{t}\right)^{s_1-1} \frac{\mathrm{d}t}{1-t} \cdots \left(\frac{\mathrm{d}t}{t}\right)^{s_d-1} \frac{\mathrm{d}t}{1-t} = \sum_{k_1>\cdots>k_d\geq 1} \frac{z^{k_1}}{k_1^{s_1}\cdots k_d^{s_d}}. \tag{14.1}$$

For convenience we set $\lambda_z(\mathbf{s}) = \widetilde{\mathrm{Li}}_{\mathbf{s}}(1/z)$ for any $|z| \geq 1$. For $\mathbf{s}, \mathbf{r} \in \mathbb{N}^d$ and $1 \leq j \leq d$, define the concatenation products

$$\begin{aligned}\vec{a}_j = \vec{a}_j(\mathbf{s};\mathbf{r}) &:= = \{s_j+1, \{1\}^{r_j-1}, \ldots, s_d+1, \{1\}^{r_d-1}\},\\ \vec{b}_j = \vec{b}_j(\mathbf{s};\mathbf{r}) &:= = \{r_j+1, \{1\}^{s_j-1}, \ldots, r_1+1, \{1\}^{s_1-1}\},\end{aligned}$$

and $\vec{a}_{d+1} := \vec{b}_0 := \{\}$.

Theorem 14.1.1. *Suppose $|u|, |v| \geq 1$ such that $1/u + 1/v = 1$. Then we have the Hölder convolution for the MZVs:*

$$\begin{aligned}\zeta(\vec{a}_1) = \zeta(\overleftarrow{b}_d) = \sum_{j=1}^{d} \Bigg\{ &\sum_{k=0}^{s_j} \lambda_u\big(s_j+1-k, \{1\}^{r_j-1}, \vec{a}_{j+1}\big)\lambda_v\big(\{1\}^k, \overleftarrow{b}_{j-1}\big)\\ &+ \sum_{k=1}^{r_j-1} \lambda_u\big(\{1\}^k, \vec{a}_{j+1}\big)\lambda_v\big(r_j+1-k, \{1\}^{s_j-1}, \overleftarrow{b}_{j-1}\big)\Bigg\} + \lambda_v\big(\overleftarrow{b}_d\big).\end{aligned}$$

Proof. Suppose $|\mathbf{s}|+|\mathbf{r}| = w$. In the iterated integral representation of MZV by Eq. (3.1)

$$\zeta(\vec{a}_1) = \int_0^1 a^{s_1} b^{r_1} \cdots a^{s_d} b^{r_d}, \quad a = \frac{\mathrm{d}t}{t},\ b = \frac{\mathrm{d}t}{1-t},$$

we may regard the region as given by $0 < t_{j+1} < t_j < 1$ for $j = 1, \ldots, w-1$. Then we can break this region into $w+1$ parts each of which is a product of regions $0 < t_{j+1} < t_j < 1/2$ for the last n variables, and $1/2 < t_{j+1} < t_j < 1$ for the remaining $w-n$ variables, where $n = 0, 1, \ldots, w$. Then we can apply the substitution $t_j \to 1-t_j$ and convert the latter type to one of the former type. Now the theorem follows immediately. □

Remark 14.1.2. The Hölder convolution has important applications.

(i) Notice that $\vec{a}_1$ and $\overleftarrow{b}_d$ are dual to each other by the duality relation Eq. (5.1).

(ii) Take $p = q = 2$. Thm. 14.1.1 provides us with a fast numerical computational tool for MZVs with very high precisions because the λ_2 converges geometrically. This is crucial when we want to find integer relations among MZVs by using LLL or PSLQ algorithms (see Appendix D.3).

(iii) The Hölder convolution can be extended to compute other special values of MPLs. A free online resource [75], called EZ-face, is available based on this idea which can be used to handle both MZVs and alternating Euler sums.

Example 14.1.3. Take $\mathbf{s} = \mathbf{r} = (1)$. Then we have

$$\zeta(2) = 2\lambda_2(2) + \lambda_2(1)^2 = 2\sum_{n\geq 1} \frac{1}{2^n n^2} + \left(\sum_{n\geq 1} \frac{1}{2^n n}\right)^2$$

which gives a rather fast way to compute $\pi^2 = 6\zeta(2)$. In fact, by taking only 28 terms in the sums above we already obtain 8 valid digits after the decimal point: $\pi^2 \approx 9.86960440$. To achieve the same precision, we would have to add almost 1.4 billion terms in the series definition $6\sum_{n\geq 1} \frac{1}{n^2}$.

Example 14.1.4. Take $\mathbf{s} = (2,1)$ and $\mathbf{r} = (3,2)$. Then we have

$$\begin{aligned}
&\zeta(3,1,1,2,1)\\
&= \lambda_2(3,1,1,2,1) + \lambda_2(2,1,1,2,1)\lambda_2(1) + \lambda_2(1,1,1,2,1)\lambda_2(1,1)\\
&+ \lambda_2(1,2,1)\lambda_2(3,1) + \lambda_2(1,1,2,1)\lambda_2(2,1) + \lambda_2(2,1)\lambda_2(4,1)\\
&+ \lambda_2(1,1)\lambda_2(1,4,1) + \lambda_2(1)\lambda_2(2,4,1) + \lambda_2(3,4,1) = \zeta(3,4,1).
\end{aligned}$$

14.2 Euler sums

In the case of $N = 2$ we have called the CMZV the Euler sums. These values attracted a lot attention when they first appeared prominently in the computation of Feymann integral in the 1990s. Many interesting rational relations have been found among these values and research in this area is still quite active up to this day.

Previously, we have set $a = dt/t$ and $b = dt/(1-t)$. For level $n = 2$ we need to introduce $c := -dt/(1+t)$ which will be in force throughout the rest of this chapter.

14.2.1 *Dimension conjectures*

If an Euler sum can be expressed by linear combination of the products of Euler sums with lower weights then the Euler sum is called *reducible*. Otherwise, it is called *irreducible*.

Let $E_{n,d}$ (resp. $M_{n,d}$) be the minimum number of irreducible Euler sums of weight n and depth d, such that every Euler sum (resp. MZVs) of weight n and depth d can be written as a $\mathbb{Q}$-linear combination of these irreducibles, augmented by products of irreducibles of lower weights and depths.

Conjecture 14.2.1. *We have*

$$\prod_{n\geq 3}\prod_{d\geq 1}(1-x^n y^d)^{E_{n,d}} = 1 - \frac{x^3 y}{1-x^2}\,\frac{1}{1-xy}, \tag{14.2}$$

$$\prod_{n\geq 3}\prod_{d\geq 1}(1-x^n y^d)^{M_{n,d}} = 1 - \frac{x^3 y}{1-x^2}, . \tag{14.3}$$

Explicitly, these conjectural numbers can be given in terms of the following quantities[1]:

$$T(m,n) := \frac{1}{m+n}\sum_{d|m,d|n}\mu(d)\binom{(m+n)/d}{m/d}, \tag{14.4}$$

where μ is the Möbius function defined by Eq. (3.12). When $n \equiv d \pmod 2$ and $n > d$, one can show that

$$E_{n,d} = T((n-d)/2, d). \tag{14.5}$$

Similarly, when $n \equiv d \pmod 2$ and $n > 3d$, one has (see Exercise 14.1)

$$M_{n,d} = T((n-3d)/2, d). \tag{14.6}$$

[1]Broadhurst calls these numbers *Euler's triangle* in [86, 87].

In order to find the finer structures of Euler sums we need to use the Lyndon words on the alphabet of all positive integers with its natural order (see Definition 3.3.3).

Conjecture 14.2.2. *Let $w > d$ be two positive integers. Then one can choose the irreducibles by using $\mathrm{L}(\mathbf{w})$ where $\mathbf{w} = \delta$ or $\mathbf{w} = a^{n_1-1}ca^{n_2-1}c\dots a^{n_d-1}c$ subject to the following conditions:*

(i) $w = \sum_{j=1}^{d} n_j$,
(ii) $n_j \geq 2$,
(iii) $(n_d, \dots, n_1)$ *is a Lyndon word.*

For example, the words corresponding to the first few irreducible Euler sums are given below:

$$c,\ ac,\ a^2c,\ a^3c,\ a^4c,\ a^2cac,\ a^5c, \dots$$

Using this we can produce the conjectured basis of MDVs in lower weights with following words:

$w = 1 : c$;
$w = 2 : c^{\sqcup\!\sqcup 2},\ ac$;
$w = 3 : c \sqcup\!\sqcup ac,\ c^{\sqcup\!\sqcup 3},\ a^2c$;
$w = 4 : c^{\sqcup\!\sqcup 2} \sqcup\!\sqcup ac,\ c^{\sqcup\!\sqcup 4},\ c \sqcup\!\sqcup a^2c,\ ac \sqcup\!\sqcup ac$;
$w = 5 : c^{\sqcup\!\sqcup 3} \sqcup\!\sqcup ac,\ c^{\sqcup\!\sqcup 5},\ c^{\sqcup\!\sqcup 2} \sqcup\!\sqcup a^2c,\ c \sqcup\!\sqcup ac \sqcup\!\sqcup ac,\ a^2c \sqcup\!\sqcup ac,\ a^4c,\ a^2cac; \dots$

When considering only the linear independence of Euler sums Broadhurst conjectures that the $\mathbb{Q}$-dimension d_n of the space ES_n of weight n Euler sum sums is given by the Fibonacci numbers: $d_1 = 1$, $d_2 = 2$, $d_3 = 3$, $d_4 = 5$, $d_5 = 8$, and so on. Zlobin further proposes another set of basis.

Conjecture 14.2.3. *The $\mathbb{Q}$-space ES_w has the following basis:*

$$\Big\{\zeta(\bar{b}_1, b_2, \dots, b_r) : b_j = 1 \text{ or } 2, \sum_{j=0}^{r} b_j = w\Big\}.$$

Consequently, $\dim_{\mathbb{Q}} \mathsf{CMZV}(w, 2)$ *are given by Fibonacci numbers.*

14.2.2 *A possible integral structure*

Our numerical evidence suggests there may exist even subtler **integral** structures. So we propose

Conjecture 14.2.4. *Let n be a positive integer. Then there are $\mathbb{Q}$-linearly independent Euler sums of weight n such that every Euler sum of weight n is a $\mathbb{Z}$-linear combination of these sums.*

This conjecture can be compared to the corresponding integral structure of the group of cyclotomic units studied by Conrad [142, 143]. The MZVs seems to have an integral structure over $\mathbb{Z}[1/2]$.

We shall now use regularized DBSF to compute the relations between Euler sums of weight two, three and four. Most of the computations in this section are carried out by Maple. We have checked the consistency of these relations with the many known ones and verified numerically all the identities in the paper by EZ-face (see [75]) with error smaller than 10^{-45}. From these numerical results we derived our Conjecture 14.2.4.

Proposition 14.2.5. *All the weight two Euler sums can be expressed as $\mathbb{Z}$-linear combinations of $\zeta(\bar{2})$ and $\zeta(\bar{1},1)$:*

$$\zeta(2) = -2\zeta(\bar{2}), \quad \zeta(\bar{1},\bar{1}) = \zeta(\bar{2}) + \zeta(\bar{1},1)$$

Proof. It is easy to see from regularized DBSF that

$$\zeta(2) = -2\zeta(\bar{1},\bar{1}) + 2\zeta(\bar{1},1), \qquad \zeta(\bar{2}) = -\zeta(\bar{1},1) + \zeta(\bar{1},\bar{1}).$$

□

Remark 14.2.6. From the proposition and a stuffle relation we get

$$2\zeta(\bar{1},1) = 2\zeta(\bar{1},\bar{1}) - 2\zeta(\bar{2}) = \zeta(\bar{1})^2 = \ln(2)^2.$$

Hence it is apparent that $\zeta(2)$ and $\zeta(\bar{1},1)$ are linearly independent which verifies the Main Conjecture 1 in this case.

Proposition 14.2.7. *We can express all weight three Euler sums as $\mathbb{Z}$-linear combinations of $\zeta(\bar{2},1)$, $\zeta(\bar{1},1,1)$ and $\zeta(\bar{1},2)$:*

$$\begin{aligned}
\zeta(3) &= \ 8\zeta(\bar{2},1),\\
\zeta(\bar{3}) &= -6\zeta(\bar{2},1),\\
\zeta(2,1) &= \ 8\zeta(\bar{2},1),\\
\zeta(2,\bar{1}) &= \ 2\zeta(\bar{2},1) - 3\zeta(\bar{1},2),\\
\zeta(\bar{2},\bar{1}) &= \ 3\zeta(\bar{1},2) - 7\zeta(\bar{2},1),\\
\zeta(\bar{1},\bar{2}) &= -2\zeta(\bar{1},2) + \zeta(\bar{2},1),\\
\zeta(\bar{1},1,\bar{1}) &= \ \ \zeta(\bar{2},1) + \zeta(\bar{1},1,1),\\
\zeta(\bar{1},\bar{1},1) &= \ \ \zeta(\bar{1},2) - 5\zeta(\bar{2},1) + \zeta(\bar{1},1,1),\\
\zeta(\bar{1},\bar{1},\bar{1}) &= \ \ \zeta(\bar{1},2) + \zeta(\bar{1},1,1).
\end{aligned}$$

Proof. When weight is three, by only finite DBSF we have

$$\begin{aligned}
\zeta(\bar{1},1,\bar{1})+2\zeta(\bar{1},\bar{1},1)+\zeta(\bar{1},\bar{2})+\zeta(2,1)-3\zeta(\bar{1},1,1)&=0,\\
2\zeta(\bar{1},\bar{1},\bar{1})+\zeta(\bar{1},2)+\zeta(2,\bar{1})-2\zeta(\bar{1},1,\bar{1})&=0,\\
\zeta(\bar{2},\bar{1})+\zeta(\bar{1},\bar{2})+\zeta(3)-2\zeta(\bar{2},1)-\zeta(\bar{1},2)&=0,\\
\zeta(\bar{1},2)+\zeta(\bar{3})-\zeta(\bar{2},\bar{1})-\zeta(\bar{1},\bar{2})&=0.
\end{aligned}$$

These are far from enough to prove the proposition. But by regularized DBSF we have five more relations:

$$\begin{aligned}
\zeta(\bar{3})+2\zeta(\bar{2},1)+\zeta(\bar{1},2)+2\zeta(\bar{1},1,1)-\zeta(2,\bar{1})+\zeta(\bar{1})\zeta(2)-2\zeta(\bar{1},\bar{1},1)&=0,\\
\zeta(\bar{1},1,\bar{1})-\zeta(\bar{2},1)-\zeta(\bar{1},2)-2\zeta(\bar{1},1,1)+\zeta(\bar{1},\bar{1},\bar{1})&=0,\\
\zeta(\bar{1},\bar{1},1)-\zeta(\bar{2},\bar{1})-\zeta(\bar{1},\bar{2})-\zeta(\bar{1},1,\bar{1})&=0,\\
\zeta(\bar{2},\bar{1})-\zeta(\bar{3})-\zeta(\bar{2},1)+\zeta(2,\bar{1})&=0,\\
\zeta(2,1)-\zeta(3)&=0.
\end{aligned}$$

Now the proposition follows from the stuffle relation: $\zeta(\bar{1})\zeta(2)=\zeta(\bar{3})+\zeta(2,\bar{1})+\zeta(\bar{1},2)$. □

Remark 14.2.8. By our standard conjectures there should be no further linear relations among $\zeta(\bar{2},1)$, $\zeta(\bar{1},1,1)$ and $\zeta(\bar{1},2)$ which gives $\dim_{\mathbb{Q}}\mathsf{CMZV}(3,2)=3$. This is easy to see to be equivalence to the linear independence of $\zeta(3)$, $\zeta(\bar{1})\zeta(2)$ and $\zeta(\bar{1},\bar{1},1)$.

The proceeding two propositions and the following results show that in weight ≤ 5 both Broadhurst-Zlobin Conjecture and our integral structure Conjecture 14.2.4 are true.

Proposition 14.2.9. *All weight four Euler sums are $\mathbb{Z}$-linear combinations of $A=\zeta(\bar{2},1,1)$, $B=\zeta(\bar{2},2)$, $C=\zeta(\bar{1},2,1)$, $D=\zeta(\bar{1},1,2)$, and $E=\zeta(\bar{1},1,1,1)$.*

The next proposition also shows that the $\mathbb{Q}$-basis conjectured by Zlobin can not be chosen as the $\mathbb{Z}$-linear basis in general.

Proposition 14.2.10. *All weight five Euler sums are $\mathbb{Q}$-linear combinations of $\zeta(\bar{1},1,1,1,1)$, $\zeta(\bar{1},1,2,1)$, $\zeta(\bar{2},1,1,1)$, $\zeta(\bar{1},1,1,2)$, $\zeta(\bar{1},2,1,1)$, $\zeta(\bar{2},1,2)$, $\zeta(\bar{2},2,1)$ and $\zeta(\bar{1},2,2)$.*

We put the detailed tables of the above in Appendix E.1.

14.2.3 *A few useful integral substitutions*

Recall that the Hoffman algebra $\mathfrak{A}_1 = \mathbb{Q}\langle a, b\rangle$ (resp. Hoffman-Racinet algebra $\mathfrak{A}_2 = \mathbb{Q}\langle a, b, c\rangle$ by setting $b = b_0$ and $c = b_1$ in Definition 13.3.2) is the graded noncommutative polynomial $\mathbb{Q}$-algebra generated by words $\mathsf{X}_1 = \{a, b\}$ (resp. $\mathsf{X}_2 = \{a, b, c\}$). Let $\mathfrak{A}_2^0$ be the subalgebra generated by words not beginning with b and not ending with a. Such words are called *admissible words*.

It is possible to prove that the only nontrivial changes of variable that preserve the interval $[0, 1]$ and send $\mathfrak{A}_1$ into $\mathfrak{A}_2$ are given by

$$\mathbf{w}(a, b) \mapsto \overleftarrow{\mathbf{w}}(b, a) \qquad \text{if } t \mapsto 1 - t, \tag{14.7}$$

$$\mathbf{w}(a, b) \mapsto \mathbf{w}(2a, b + c) \qquad \text{if } t \mapsto t^2, \tag{14.8}$$

$$\mathbf{w}(a, b) \mapsto \mathbf{w}(a + c, b - c) \qquad \text{if } t \mapsto \frac{2t}{1 + t}, \tag{14.9}$$

$$\mathbf{w}(a, b) \mapsto \mathbf{w}(a + 2c, 2b - 2c) \qquad \text{if } t \mapsto \frac{4t}{(1 + t)^2}, \tag{14.10}$$

and their compositions, where $\overleftarrow{\mathbf{w}}(b, a)$ denotes the word formed by switching a and b and then reversing the order of the letters in the word $\mathbf{w}(a, b)$. For example,

$$\mathbf{w}(a, b) \mapsto \mathbf{w}(2a, b + c) \qquad \text{if } t \mapsto 1 - 2t/(1 + t) = (1 - t)/(1 + t). \tag{14.11}$$

Example 14.2.11. We now show that

$$\zeta(3) = 8\zeta(\overline{2}, 1). \tag{14.12}$$

By direct expansion in $\mathfrak{A}_2$ we obtain

$$2[ab^2 - 2a(b+c)^2] + 8[ab^2 - (a+c)(b-c)^2] + [(a+2c)(2b-2c)^2 - ab^2] = ab^2 - 8ac^2.$$

Using substitutions Eq. (14.8), Eq. (14.9) and Eq. (14.10) respectively for the three square brackets on the left-hand side, we get

$$\zeta(3) - 8\zeta(\overline{2}, 1) = Ł_{⧢}(ab^2 - 8ac^2) = 0.$$

14.2.4 *A family of identities* $\boldsymbol{\zeta(\{3\}^n) = 8^n\zeta(\{\overline{2}, 1\}^n)}$

In this subsection we generalize Eq. (14.12) as follows.

Theorem 14.2.12. *For every positive integer n*

$$\zeta(\{3\}^n) = 8^n\zeta(\{\overline{2}, 1\}^n). \tag{14.13}$$

The key idea of the proof is to find some suitable combinations of finite DBSFs and DISTs in level 2.

In order to encode the stuffle relation we need to define the sub-algebra $\mathfrak{A}_2^1 = \mathbb{Q}\langle \mathsf{Y} \rangle$ of $\mathfrak{A}_2 = \mathbb{Q}\langle \mathsf{X} \rangle$ where Y is the alphabet consisting of the letters $\{\beta_n, \gamma_n : n \in \mathbb{N}\}$, where

$$\beta_n := a^{n-1}b \quad \text{and} \quad \gamma_n := a^{n-1}c.$$

For any word $\mathbf{w} \in \mathsf{Y}^*$ and positive integer n define the toggling operator $\tau_{\beta_n}(\mathbf{w}) = \mathbf{w}$, and $\tau_{\gamma_n}(\mathbf{w})$ to be the word with β and γ exchanged. For example $\tau_{\gamma_1}(\gamma_2\beta_4) = \beta_2\gamma_4$. Then the stuffle product defined in Eq. (13.10) can be rephrased as follows. For any $u, v \in \mathsf{Y}$

$$u\mathbf{w}_1 * v\mathbf{w}_2 = u\Big(\tau_u\big(\tau_u(\mathbf{w}_1) * v\mathbf{w}_2\big)\Big) + v\Big(\tau_v\big(u\mathbf{w}_1 * \tau_v(\mathbf{w}_2)\big)\Big) + [u, v]\Big(\tau_{[u,v]}\big(\tau_u(\mathbf{w}_1) * \tau_v(\mathbf{w}_2)\big)\Big) \quad (14.14)$$

where

$$[\beta_m, \beta_n] = [\gamma_m, \gamma_n] = \beta_{m+n}, \quad [\gamma_m, \beta_n] = [\beta_m, \gamma_n] = \gamma_{m+n}.$$

For a reformulation of the converting rule between words and the corresponding Euler sums (implicitly given by Eq. (13.15) and Eq. (13.14)) we refer the reader to Definition 7.2.3 and Definition 7.2.4.

For any $l_1, l_2 \in \{a, b, c\}$ define the $\diamond$-concatenation by setting $l_1 \diamond l_2 = l_1 l_2$ except that

$$b \diamond b = bc, \quad \text{and} \quad c \diamond c = cb. \quad (14.15)$$

For any positive integer i define the i-th cut of a word $l_1 \diamond \cdots \diamond l_m$ ($l_i \in \{a, b, c\}$) to be a pair of words given by

$$\mathrm{Cut}_i(l_1 \diamond l_2 \diamond \cdots \diamond l_m) := \begin{cases} \big(l_1 \diamond l_2 \diamond \cdots \diamond l_{i-1} \diamond l_i,\ l_{i+1} \diamond l_{i+2} \diamond \cdots \diamond l_m\big), & \text{if } i \text{ is odd,} \\ \big(l_i \diamond l_{i-1} \diamond \cdots \diamond l_2 \diamond l_1,\ l_{i+1} \diamond l_{i+2} \diamond \cdots \diamond l_m\big), & \text{if } i \text{ is even,} \end{cases}$$

for $i = 0, \cdots, m$. For any two words $\mathbf{w}_1$ and $\mathbf{w}_2$ we define the two bilinear maps $\mathfrak{A}_2^0 \times \mathfrak{A}_2^0 \to \mathfrak{A}_2^0$ by

$$\sqcup\!\sqcup(\mathbf{w}_1, \mathbf{w}_2) = \mathbf{w}_1 \sqcup\!\sqcup \mathbf{w}_2, \quad \text{and} \quad *(\mathbf{w}_1, \mathbf{w}_2) = \mathbf{w}_1 * \mathbf{w}_2.$$

Let $[r]$ and $\{r\}$ be the integral part and the fractional part of any $r \in \mathbb{R}$, respectively.

Theorem 14.2.13. *For every positive integer n the following holds in $\mathfrak{A}_2^0$:*

$$2^n(ac^2ab^2)^{[n/2]}(ac^2)^{2\{n/2\}} = (a^2(b+c))^n + \sum_{i=0}^{2n}(-1)^{n-i}\Big(\sqcup\!\sqcup \circ \mathrm{Cut}_i\left((c\diamond d)^{\diamond n}\right) - * \circ \mathrm{Cut}_i\left((c\diamond d)^{\diamond n}\right)\Big). \quad (14.16)$$

Here to compute $\sqcup\!\sqcup \circ \mathrm{Cut}_i\left((c\diamond d)^{\diamond n}\right)$ we use the following steps: (i) apply Cut_i to $(c\diamond d)^{\diamond n}$ by using the $\diamond$-concatenation, (ii) replace the letter d by $a(b+c)$, (iii) expand the $\diamond$-concatenation by the definition Eq. (14.15)*, and (iv) apply $\sqcup\!\sqcup$. We can compute $* \circ \mathrm{Cut}_i\left((c\diamond d)^{\diamond n}\right)$ similarly.*

Remark 14.2.14. Observe that $(b+c)\diamond c = (b+c)\diamond b = bc+cb$ and therefore

$$\mathrm{L}((c\diamond d)^{\diamond n}) = \sum_{t_1,\dots,t_n\in\{2,\bar{2}\}} \zeta\big(\bar{1},t_1,\bar{1},t_2,\dots,\bar{1},t_n\big).$$

The $\diamond$-concatenation appears because neither c^2 nor b^2 can appear in any of the Euler sums in the above sum. We should keep this in mind because the operator Cut_i will lead to some order reversals which also should obey this condition.

To prove Thm. 14.2.13 we need two separate identities involving stuffles and shuffles respectively.

Proposition 14.2.15. *For every positive integer n*

$$\sum_{i=0}^{2n}(-1)^i * \circ \mathrm{Cut}_i\left((c\diamond d)^{\diamond n}\right) = (-1)^n(a^2(b+c))^n. \quad (14.17)$$

Proof. We proceed by induction on n. When $n=1$ the left-hand side of Eq. (14.17) is

$$\begin{aligned} cd - c*d + d\diamond c =& \gamma_1(\gamma_2+\beta_2) - \gamma_1*(\gamma_2+\beta_2) + a(bc+cb) \\ =& -(\gamma_2+\beta_2)\tau\gamma_1 - \beta_3 - \gamma_3 + \gamma_2\beta_1 + \beta_2\gamma_1 = -\beta_3-\gamma_3. \end{aligned}$$

This is exactly the right-hand side $-a^2(b+c)$. Now assume that identity Eq. (14.17) holds up to $n-1$ for some $n\geq 2$. Set $\gamma=\gamma_1=c$, $d:=\beta_2+\gamma_2$ and $d_3=\beta_3+\gamma_3$. Then

$$\Delta := d\diamond\gamma = \beta_2\gamma+\gamma_2\beta_1.$$

In the rest of the paper we set $\tau=\tau_\gamma$. Note that $\tau(d)=d$ and $\tau(\Delta)=\Delta$. Hence by the recursive definition of the stuffle product Eq. (14.14)

$$\sum_{i=0}^{2n}(-1)^i * \circ \mathrm{Cut}_i\left((c\diamond d)^{\diamond n}\right)$$

$$
\begin{aligned}
&= \sum_{j=0}^{n} \Delta^j * (\gamma d)^{\diamond(n-j)} - \sum_{j=1}^{n} \Big(\gamma\Delta^{j-1}\Big) * \Big(d \diamond (\gamma d)^{n-j}\Big) \\
&= \sum_{j=0}^{n} \Big(\beta_2\gamma\Delta^{j-1}\Big) * (\gamma d)^{\diamond(n-j)} - \sum_{j=1}^{n} \Big(\gamma\Delta^{j-1}\Big) * \Big(\beta_2\gamma\Delta^{n-j} d\Big) \\
&+ \sum_{j=0}^{n} \Big(\gamma_2\beta_1\Delta^{j-1}\Big) * (\gamma d)^{\diamond(n-j)} - \sum_{j=1}^{n} \Big(\gamma\Delta^{j-1}\Big) * \Big(\gamma_2\beta_1\Delta^{n-j} d\Big) \\
&= \sum_{j=1}^{n-1} \Big\{\beta_2\Big(\big(\gamma\Delta^{j-1}\big) * (\gamma d)^{\diamond(n-j)}\Big) + \gamma\tau\Big(\big(\beta_2\gamma\Delta^{j-1}\big) * \big(\Delta^{n-j-1} d\big)\Big) \\
&\qquad + \gamma_3\tau\Big(\big(\gamma\Delta^{j-1}\big) * \big((d \diamond \gamma)^{n-j-1} d\big)\Big)\Big\} + (\gamma d)^{\diamond n} + \Delta^n \\
&- \sum_{j=1}^{n} \Big\{\gamma\tau\Big(\Delta^{j-1} * \big(\beta_2\gamma\Delta^{n-j} d\big)\Big) + \beta_2\Big(\big(\gamma\Delta^{j-1}\big) * \big(\gamma\Delta^{n-j} d\big)\Big) \\
&\qquad + \gamma_3\tau\Big(\Delta^{j-1} * \big(\gamma(d \diamond \gamma)^{n-j} d\big)\Big)\Big\} \\
&+ \sum_{j=1}^{n-1} \Big\{\gamma_2\tau\Big(\big(\gamma\Delta^{j-1}\big) * (\gamma d)^{\diamond(n-j)}\Big) + \gamma\tau\Big(\big(\gamma_2\beta_1\Delta^{j-1}\big) * \big(\Delta^{n-j-1} d\big)\Big) \\
&\qquad + \beta_3\Big(\big(\gamma\Delta^{j-1}\big) * \big(\Delta^{n-j-1} d\big)\Big)\Big\} \\
&- \sum_{j=1}^{n} \Big\{\gamma\tau\Big(\Delta^{j-1} * \big(\gamma_2\beta_1\Delta^{n-j} d\big)\Big) + \gamma_2\tau\Big(\big(\gamma\Delta^{j-1}\big) * \big(\gamma\Delta^{n-j} d\big)\Big) \\
&\qquad + \gamma_3\tau\Big(\Delta^{j-1} * \big(\gamma\Delta^{n-j} d\big)\Big)\Big\}.
\end{aligned}
$$

Converting $\beta_2\gamma + \gamma_2\beta_1$ back to Δ and canceling all the terms not involving γ_3 or β_3 we get

$$
\begin{aligned}
&\sum_{i=0}^{2n} (-1)^i * \circ \mathrm{Cut}_i\left((c \diamond d)^{\diamond n}\right) \\
&\qquad = d_3 \left\{ \sum_{j=1}^{n-1} \big(\gamma\Delta^{j-1}\big) * \big(\Delta^{n-j-1} d\big) - \sum_{j=1}^{n} \Delta^{j-1} * \big(\gamma\Delta^{n-j} d\big) \right\}
\end{aligned}
$$

By induction assumption the expression in the last big curly bracket is $(-1)^n(a^2(b+c))^{n-1}$. This proves the proposition since $d_3 = a^2(b+c)$. □

Proposition 14.2.16. *For every positive integer n*

$$
\sum_{i=0}^{2n} (-1)^i \sqcup\!\sqcup \circ \mathrm{Cut}_i\left((c \diamond d)^{\diamond n}\right) = (-2)^n (ac^2 ab^2)^{[n/2]} (ac^2)^{2\{n/2\}} \qquad (14.18)
$$

and

$$\sum_{i=0}^{2n}(-1)^i \sqcup\!\sqcup \circ \mathrm{Cut}_i\left((b\diamond d)^{\diamond n}\right) = (-2)^n (ab^2ac^2)^{[n/2]}(ab^2)^{2\{n/2\}}. \quad (14.19)$$

Here we set $d\diamond b = d\diamond c = a(cb+bc)$.

Proof. We can proceed by induction on n similarly to that of Prop. 14.2.15. We leave the details to the interested reader. □

Proof of Thm. 14.2.13. This follows easily from Prop. 14.2.15 and Eq. (14.18) of Prop. 14.2.16. □

Proof of Thm. 14.2.12. For any positive integer n, we have

$$\zeta(\{\bar{2},1\}^n) = Ł\left((ac^2ab^2)^{[n/2]}(ac^2)^{2\{n/2\}}\right),$$

and

$$\zeta(\{3\}^n) = Ł\left((a^2b)^n\right).$$

On the other hand, by Eq. (14.8) an integral substitution $t\to t^2$ yields

$$\zeta(\{3\}^n) = 4^n\, Ł\left((a^2(b+c))^n\right). \quad (14.20)$$

This also follows quickly from the following special case of the DIST Eq. (13.22):

$$\begin{aligned} Ł\left((a^2(b+c))^n\right) &= \sum_{1\le i_1<\cdots<i_j\le n}\ \sum_{k_1>\cdots>k_n>0} \frac{(-1)^{k_{i_1}+\cdots+k_{i_j}}}{(k_1\cdots k_n)^3} \\ &= \sum_{k_1>\cdots>k_n>0} \frac{(1+(-1)^{k_1})\cdots(1+(-1)^{k_n})}{(k_1\cdots k_n)^3} = \frac{1}{4^n}\zeta(\{3\}^n). \end{aligned}$$

Now multiplying 4^n on both sides of Eq. (14.16) in Thm. 14.2.13 and then applying Ł we can see that Thm. 14.13 follows immediately from Eq. (13.11), Eq. (14.17) and Eq. (14.18). □

Remark 14.2.17. Maple computation shows that Eq. (14.20) cannot be derived from finite DBSF in general.

14.2.5 *Generalized doubling relations at level 2*

The finite DISTs at level 2 are often called *the doubling relations* in the literature. It is possible to generalize this idea by manipulation of the Euler sum series. What one obtains then should be closely related to the regularized DISTs. For example, by using partial fraction expansion Eq. (5.14) we get (assuming $s > 1$, $w = s + t$, and $A_i^{s,t} = \binom{s+t-i-1}{t-1}$)

$$
\begin{aligned}
&E(s,t) = \frac{1}{2}(\zeta(s,t) + \zeta(\bar{s},\bar{t}))\\
&= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1}{(m+n)^s\ n^t}\frac{1+(-1)^m}{2} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{1}{(2m+n)^s\ n^t}\\
&= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left[\sum_{i=1}^{s} \frac{A_i^{s,t}}{(2m+2n)^{w-i}\ (2m+n)^i} + \sum_{i=1}^{t} \frac{A_i^{t,s}}{(2m+2n)^{w-i}\ n^i}\right]\\
&= \sum_{i=1}^{s} \frac{1}{2^{w-i}} \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{A_i^{s,t}}{(m+n)^{w-i}\ (2m+n)^i} + \sum_{i=1}^{t} \frac{A_i^{t,s}}{2^{w-i}}\zeta(w-i,i)\\
&= \sum_{i=1}^{t} \frac{A_i^{t,s}}{2^{w-i}}\zeta(w-i,i) + \sum_{i=1}^{s}\sum_{m=1}^{\infty}\sum_{n=m+1}^{\infty} \frac{A_i^{s,t}2^{i-w}}{(m+n)^i n^{w-i}}\\
&= \sum_{i=1}^{t} \frac{A_i^{t,s}}{2^{w-i}}\zeta(w-i,i) + \sum_{i=1}^{s}\sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty} - \sum_{n=1}^{m}\right) \frac{A_i^{s,t}2^{i-w}}{(m+n)^i n^{w-i}}\\
&= \sum_{i=1}^{t} \frac{A_i^{t,s}}{2^{w-i}}\zeta(w-i,i) + \sum_{i=1}^{s} \frac{A_i^{s,t}}{2^{w-i}}\zeta(i,w-i) - \sum_{i=1}^{s}\sum_{n=1}^{\infty}\sum_{m=n}^{\infty} \frac{A_i^{s,t}2^{i-w}}{(m+n)^i n^{w-i}}\\
&= \sum_{i=1}^{t} A_i^{t,s}2^{i-w}\zeta(w-i,i) + \sum_{i=1}^{s} A_i^{s,t}2^{i-w}\zeta(i,w-i)\\
&\qquad - \sum_{i=1}^{s}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{A_i^{s,t}2^{i-w}}{(m+2n)^i n^{w-i}} - \sum_{i=1}^{s}\sum_{n=1}^{\infty} \frac{A_i^{s,t}2^{i-w}}{(2n)^i n^{w-i}}\\
&= \sum_{i=1}^{t} A_i^{t,s}2^{i-w}\zeta(w-i,i) + \sum_{i=1}^{s} A_i^{s,t}2^{i-w}\zeta(i,w-i)\\
&- \sum_{i=1}^{s}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{1+(-1)^n}{2}\frac{A_i^{s,t}}{(m+n)^i n^{w-i}} - \sum_{i=1}^{s}\sum_{n=1}^{\infty} \frac{A_i^{s,t}2^{-w}}{n^w}\\
&= \sum_{i=1}^{t} A_i^{t,s}2^{i-w}\zeta(w-i,i) + \sum_{i=1}^{s} A_i^{s,t}2^{i-w}\zeta(i,w-i)
\end{aligned}
$$

$$-\sum_{i=1}^{s} A_i^{s,t}\frac{1}{2}\Big(\zeta(i,w-i)+\zeta(i,\overline{w-i})\Big)-\frac{(w-1)!}{(s-1)!\,t!}2^{-w}\zeta(w).$$

For larger depths, one can obtain similar relations as above. We call these *the generalized doubling relations*.

Conjecture 14.2.18. *The regularized DBSFs, finite DISTs and the generalized doubling relations are sufficient to reduce the Euler sums of a given weight and depth to a minimal set that is in agreement with Conjecture* 14.2.1*, both in weight and in depth.*

14.3 Nonstandard octahedral relations at level 4

In this section, we study the level four case where there are explicit nonstandard relations occurring already in weight 3:

$$\begin{aligned}5\,\mathrm{Li}_{1,2}(-1,-i) &= 46\,\mathrm{Li}_{1,1,1}(i,1,1)-7\,\mathrm{Li}_{1,1,1}(-1,-1,i)\\ &\quad-13\,\mathrm{Li}_{1,1,1}(i,i,i)+13\,\mathrm{Li}_{1,2}(-i,i)-\mathrm{Li}_{1,1,1}(-i,-1,1)\\ &\quad+25\,\mathrm{Li}_{1,1,1}(-i,1,1)-8\,\mathrm{Li}_{1,1,1}(i,i,-1)+18\,\mathrm{Li}_{2,1}(-i,1).\end{aligned}\tag{14.21}$$

Notice that there are $N-1$ weight one CMZVs of level N and there are $N^2(N+1)^{w-2}$ weight $w\geq 2$ CMZVs of level N. Therefore there are 80 CMZVs of weight 3 and level 4. In order to show the nonstandard relation (14.21), one may proceed as follows.

(1) Find all the finite DBSFs by multiplying each CMZV of weight one and each CMZV of weight two: there are 48 such relations.
(2) Find all the regularized non-finite DBSFs: there are 19 such relations.
(3) Find all the regularized DISTs: there are 14 such relations.
(4) Find all the weight one relations: there are 16 such relations.

Put all of these relations together and use Maple to solve the system we find that among the 80 formal variables (*i.e.*, weight 3 CMZVs of level 4) there are 9 free ones given by the CMZVs appearing in Eq. (14.21). But we know the upper bound of $\dim \mathsf{CMZV}(3,4)$ is 8 according to Thm. 13.2.1. Moreover, with Maple it can be shown that the system with all the above relations together with Eq. (14.21) has only 8 free variable. Therefore, if Eq. (14.21) is valid then it is indeed a *non-standard relation*.

Remark 14.3.1. In fact, the computation is a little simpler if we use the Lie algebra approach. However, this would require us to use the motivic language which is beyond the scope of this book. For the interested reader, please refer to the original paper [599].

To prove Eq. (14.21) we need to study the geometry of the Riemann sphere $\mathbb{P}^1$ punctured at $\{0, \infty, \pm 1, \pm i\}$ which clearly possesses an octahedral symmetry. However, one should use a system of tangent vectors stable (up to multiplication by roots of unity) by the octahedral group to use this symmetry: one should use the straight line path from 1_0 to $(-2)_1$. We now explain the meaning of this in details.

Recall that in § 13.3.2 we considered the Racnet algebra $\mathbb{Q}\langle\langle \mathtt{x}_0, \mathtt{x}_\xi \rangle\rangle_{\xi\in\Gamma_4}$. We define $\mathtt{x}_\infty = -\sum_{\xi\in\Gamma_4\cup\{0\}} \mathtt{x}_\xi$. Define the $\mathbb{C}$-linear map $\rho : \mathbb{Q}\langle\langle \mathtt{x}_0, \mathtt{x}_\xi \rangle\rangle_{\xi\in\Gamma_4} \longrightarrow \mathbb{Q}\langle\langle \mathtt{x}_0, \mathtt{x}_\xi \rangle\rangle_{\xi\in\Gamma_4}$ such that

$$\rho : \mathtt{x}_0 \to \mathtt{x}_1 \to \mathtt{x}_i \to \mathtt{x}_0, \qquad \mathtt{x}_\infty \to \mathtt{x}_{-1} \to \mathtt{x}_{-i} \to \mathtt{x}_\infty.$$

We also need the Möbius transformation

$$R(z) := \frac{i-z}{i+z}, \qquad \text{and its inverse} \qquad R^{-1}(z) = \frac{i(1-z)}{1+z}.$$

Notice that both ρ and R have order 3. The key fact is that for any path $\mathfrak{p}$

$$\int_{R(\mathfrak{p})} \Omega = \rho\left(\int_{\mathfrak{p}} \Omega\right). \tag{14.22}$$

Let $0 < \varepsilon < 1/3$ and C_ε be the path $A_1, \ldots, A_6$ in the complex plane shown in Figure 14.1, where A_2, A_4 and A_6 are the quarter circles of radii ε, 2ε and 2ε, respectively, oriented clockwise.

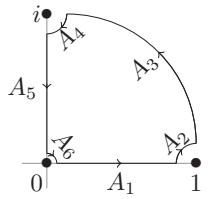


Figure 14.1: The contour of the integral for producing a non-standard relations in weight 3 and level 4.

Let $\Omega = \sum_{\xi\in\Gamma_4\cup\{0\}} \frac{\mathtt{x}_\xi}{z-\xi} \mathrm{d}z$. Then by Definition 13.3.22 the regularized iterated integral

$$\mathfrak{I} = \sum_{n=0}^{\infty} \mathrm{Reg} \int_0^1 \Omega^{\circ n}$$

is a group-like element in $(\mathbb{C}\langle\langle \mathsf{X} \rangle\rangle, \Delta)$. Since $\int_{\mathfrak{p}} \Omega^{\circ 0} = 1$ for any path $\mathfrak{p}$, we have

$$1 = \int_{C_\varepsilon} \sum_{n=0}^{\infty} \Omega^{\circ n} = \int_{A_6} \sum_{n=0}^{\infty} \Omega^{\circ n} \int_{A_5} \sum_{n=0}^{\infty} \Omega^{\circ n} \cdots \int_{A_1} \sum_{n=0}^{\infty} \Omega^{\circ n}. \tag{14.23}$$

We now compute Eq. (14.23) explicitly. First, by the shuffle property of iterated integrals

$$\sum_{n=0}^{\infty}\int_{A_6}\Omega^{\circ n}=\sum_{n=0}^{\infty}\int_{A_6}\left(\frac{\mathrm{d}t}{t}\mathbf{x}_0\right)^{\circ n}=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\int_{A_6}\frac{\mathrm{d}t}{t}\mathbf{x}_0\right)^{n}=e^{(-2\pi i/4)\mathbf{x}_0}$$

Similarly we have

$$\sum_{n=0}^{\infty}\int_{A_2}\Omega^{\circ n}=e^{(-2\pi i/4)\mathbf{x}_1},\qquad \sum_{n=0}^{\infty}\int_{A_4}\Omega^{\circ n}=e^{(-2\pi i/4)\mathbf{x}_i}.$$

Now A_1 is the straight path $\varepsilon\to 1-\varepsilon$ followed by the straight path $1-\varepsilon\to 1-2\varepsilon$. Thus

$$\begin{aligned}\int_{A_1}\sum_{n=0}^{\infty}\Omega^{\circ n}&=\int_{1-\varepsilon}^{1-2\varepsilon}\sum_{n=0}^{\infty}\Omega^{\circ n}\cdot\int_{\varepsilon}^{1-\varepsilon}\sum_{n=0}^{\infty}\Omega^{\circ n}\\&=\int_{1-\varepsilon}^{1-2\varepsilon}\sum_{n=0}^{\infty}\left(\frac{\mathrm{d}z}{z-1}\mathbf{x}_1\right)^{\circ n}\cdot\int_{\varepsilon}^{1-\varepsilon}\sum_{n=0}^{\infty}\Omega^{\circ n}+O(\varepsilon).\end{aligned}$$

Regularizing by using Definition 13.3.22 we have

$$\operatorname{Reg}\sum_{n=0}^{\infty}\int_{A_1}\Omega^{\circ n}=e^{\mathbf{x}_1\log 2}\mathcal{I}.$$

It follows from (14.22) that

$$\int_{A_3}\Omega=\rho\left(\int_{R^2(A_3)}\Omega\right)=\rho\left(\int_{A_1}\Omega\right)+O(\varepsilon),$$

$$\int_{A_5}\Omega=\rho^2\left(\int_{R(A_5)}\Omega\right)=\rho^2\left(\int_{A_1}\Omega\right)+O(\varepsilon).$$

Thus

$$\operatorname{Reg}\sum_{n=0}^{\infty}\int_{A_3}\Omega^{\circ n}=e^{\mathbf{x}_i\ln 2}\rho(\mathcal{I}),\quad \operatorname{Reg}\sum_{n=0}^{\infty}\int_{A_5}\Omega^{\circ n}=e^{\mathbf{x}_0\ln 2}\rho^2(\mathcal{I}).$$

Remark 14.3.2. Using tangential base points the above computation can be simplified a lot. For example, replacing the straight path $1_0\to(-1)_1$ by the straight path $1_0\to(-2)_1$ changes $\mathcal{I}$ to $\mathcal{I}'=e^{\mathbf{x}_1\log 2}\mathcal{I}$. So the regularized integral I over the path A_1 followed by A_2 gives $I=e^{-2\pi i\mathbf{x}_1/4}\mathcal{I}'=e^{-2\operatorname{Li}_1(i)\mathbf{x}_1}\mathcal{I}$.

Hence (14.23) yields $\rho^2(I)\rho(I)I=1$, i.e.,

$$e^{-2\operatorname{Li}_1(i)\mathbf{x}_0}\rho^2(\mathcal{I})e^{-2\operatorname{Li}_1(i)\mathbf{x}_i}\rho(\mathcal{I})e^{-2\operatorname{Li}_1(i)\mathbf{x}_1}\mathcal{I}=1. \tag{14.24}$$

Notice that all the coefficients in the above equation can be written explicitly as a $\mathbb{Q}$-linear combination of the nine CMZVs appearing in Eq. (14.21). By comparing the coefficient of $\mathbf{x}_2\mathbf{x}_1^2$ in (14.24) one finally arrives at (14.21).

Similarly in weight 4, there are five $\mathbb{Q}$-linearly independent non-standard octahedral relations at level 4. See the computation in Sec. 14.6.

14.4 Multiple Clausen and Deligne values at level 6

In this section, we consider some special forms of CMZVs at level 6. Throughout this section, we let $\mu = \mu_6 := (1 + i\sqrt{3})/2$ be a primitive 6th root of unity and put $\theta = \pi/3$.

14.4.1 *Definitions*

Set $a = d\log(t)$, $b = d\log(1-t)$ and $\delta = d\log(1-\mu t)$. Recall that we have the evaluation map $\zeta : \mathfrak{A}_6^1 \to \mathbb{C}$ where $\mathfrak{A}_6$ is the level 6 Hoffman-Racinet algebra defined by Definition 13.3.2.

Definition 14.4.1. A *multiple Deligne value* (MDV) is given by $(-1)^d \mathrm{L}(\mathbf{w})$ for some admissible word $\mathbf{w} \in \{a, b, \delta\}^*$ of depth d. Here, the evaluation map L is defined by Prop. 13.3.8.

Recall the one variable MPL $\widetilde{\mathrm{Li}}_{\mathbf{s}}(z) = \mathrm{Li}_{\mathbf{s}}(z, 1, \dots, 1)$ by (14.1).

Definition 14.4.2. Let $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$. We set $\lambda(\mathbf{s}) = \widetilde{\mathrm{Li}}_{\mathbf{s}}(\mu)$ and call it a *multiple Clausen value* (MCV). Notice that MCVs are complex numbers in general.

Numerical evidence shows that every MDV can be expressed as a $\mathbb{Q}$-linear combination of MCVs of the same weight. See Conjecture 14.4.27.

Lemma 14.4.3. *For all* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$*, we have*

$$\lambda(s_1 + 1, \dots, s_d) = \int_0^\mu \frac{\mathrm{Li}_{s_1,\dots,s_d}(z)}{z} dz \tag{14.25}$$

$$\lambda(1, s_1, \dots, s_d) = \int_0^\mu \frac{\mathrm{Li}_{s_1,\dots,s_d}(z)}{1-z} dz. \tag{14.26}$$

Proof. By the series definition in Eq. (14.1) we easily see that

$$\frac{\mathrm{d}}{\mathrm{d}z} \widetilde{\mathrm{Li}}_{s_1+1,s_2,\dots,s_n}(z) = \frac{1}{z} \widetilde{\mathrm{Li}}_{s_1,\dots,s_n}(z), \tag{14.27}$$

$$\frac{\mathrm{d}}{\mathrm{d}z} \widetilde{\mathrm{Li}}_{1,s_2,\dots,s_n}(z) = \frac{1}{1-z} \widetilde{\mathrm{Li}}_{s_2,\dots,s_n}(z). \tag{14.28}$$

Integrating we obtain Eqs. (14.25) and (14.26), respectively. □

Lemma 14.4.4. *For all* $|z| \leq 1$ *but* $z \neq 1$*, we have*

$$\widetilde{\mathrm{Li}}_{\{1\}^n}(z) = \frac{(-\log(1-z))^n}{n!}. \tag{14.29}$$

Proof. Repeated application of (14.28). □

The following formula is straight-forward.

Lemma 14.4.5. *For any* $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ *we have*

$$\lambda(\mathbf{s}) = (-1)^d \mathrm{L}(a^{s_1-1}\delta \cdots a^{s_d-1}\delta).$$

Hence every MCV is a MDVs.

Definition 14.4.6. Let $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ and $|\mathbf{s}| = w$. A *multiple Glaishers* (MGl) is the value $\mathrm{mgl}(\mathbf{s}) := \mathrm{Re}\big(i^w\lambda(\mathbf{s})\big)$. On the other hand, we call $\mathrm{mcl}(\mathbf{s}) := \mathrm{Im}\big(i^w\lambda(\mathbf{s})\big)$ a *multiple Clausens* (MCl). We call $|\mathbf{w}|$ their *weight*. Notice that both MGl and MCl are real numbers.

By the definition, the MGl at $\mathbf{s}$ has the following expansion

$$\mathrm{mgl}(\mathbf{s}) = \begin{cases} (-1)^{(w+1)/2} \displaystyle\sum_{k_1>\cdots>k_d>0} \frac{\sin(k_1\theta)}{k_1^{s_1}\cdots k_d^{s_d}}, & \text{if } w| \text{ is odd;} \\ (-1)^{w/2} \displaystyle\sum_{k_1>\cdots>k_d>0} \frac{\cos(k_1\theta)}{k_1^{s_1}\cdots k_d^{s_d}}, & \text{if } w \text{ is even.} \end{cases}$$

The MCl at $\mathbf{s}$ has the following expansion

$$\mathrm{mcl}(\mathbf{s}) = \begin{cases} (-1)^{(w-1)/2} \displaystyle\sum_{k_1>\cdots>k_d>0} \frac{\cos(k_1\theta)}{k_1^{s_1}\cdots k_d^{s_d}}, & \text{if } w \text{ is odd;} \\ (-1)^{w/2} \displaystyle\sum_{k_1>\cdots>k_d>0} \frac{\sin(k_1\theta)}{k_1^{s_1}\cdots k_d^{s_d}}, & \text{if } w \text{ is even.} \end{cases}$$

In the depth one case, we have the classical results.

Proposition 14.4.7. *For any positive integer* n*, we have*

$$\mathrm{mcl}(2n+1) = \frac{(-1)^n}{2}(1-2^{-2n})(1-3^{-2n})\zeta(2n+1), \tag{14.30}$$

$$\mathrm{mgl}(n) = -\frac{1}{2}B_n\Big(\frac{1}{6}\Big)\frac{(-2\pi)^n}{n!}. \tag{14.31}$$

Here $B_n(x)$ *is the* n*th Bernoulli polynomial defined by the generating series* $te^{xt}/(e^t-1) = \sum_{n\geq 0} B_n(x)t^n/n!$.

Proof. See Exercise 14.5. □

Remark 14.4.8. In depth one case when $\mathbf{s} = n$ is a positive integer, then we have the relations

$$\mathrm{mgl}(n) = (-1)^{\lceil n/2\rceil}\mathrm{Gl}_n(\theta) \quad \text{and} \quad \mathrm{mcl}(n) = (-1)^{\lfloor n/2\rfloor}\mathrm{Cl}_n(\theta)$$

using the notation in Lewin's book [373].

Turning to MDVs, let $\overline{\mathrm{L}}(\mathbf{w})$ be the complex conjugate of $\mathrm{L}(\mathbf{w})$ and extend it $\mathbb{Q}$-linearly to $\mathfrak{A}^1$. Define the *dual* of a word $\mathbf{w}$, denoted by, $\widetilde{\mathbf{w}}$, to be the word obtained by writing $\mathbf{w}$ backwards and then interchanging a and b. For example, $\widetilde{a^3b\delta^2} = \delta^2 a b^3$.

Lemma 14.4.9. *For any admissible word $\mathbf{w} \in \{a, b, \delta\}^*$, we have*

$$\overline{\mathrm{L}}(\mathbf{w}) = (-1)^{|\mathbf{w}|}\,\mathrm{L}(\widetilde{\mathbf{w}}).$$

Proof. Suppose $\sharp\{a : a \in \mathbf{w}\} = \alpha$, $\sharp\{b : b \in \mathbf{w}\} = \beta$ and $\sharp\{\delta : \delta \in \mathbf{w}\} = \gamma$. Since

$$d\log(1 - \overline{\mu}(1-t)) = d\log(1-\mu t)$$

we get

$$\begin{aligned}\overline{\mathrm{L}}(\mathbf{w}) = (-1)^{\beta+\gamma}\int_0^1 \overline{\mathbf{w}} \overset{t\to 1-t}{=\!=\!=} (-1)^{\alpha+\gamma}\int_1^0 \mathbf{w}(a \leftrightarrow b)\\ = (-1)^{\beta}\int_0^1 \widetilde{\mathbf{w}} = (-1)^{|\mathbf{w}|}\,\mathrm{L}(\widetilde{\mathbf{w}})\end{aligned}$$

because $\mathrm{dp}(\widetilde{\mathbf{w}}) = \alpha + \gamma$ and $|\mathbf{w}| = \alpha + \beta + \gamma$. □

Corollary 14.4.10. *All of MCVs, MGls and MCls of weight w are in the $\mathbb{Q}$-vector space MDV_w generated by MDVs of weight w.*

Proof. By Lemmas 14.4.5 and 14.4.9 both $\lambda(\mathbf{s}) + \overline{\lambda(\mathbf{s})}$ and $\lambda(\mathbf{s}) - \overline{\lambda(\mathbf{s})}$ are in MDV. □

Lemma 14.4.11. *For any positive integer n set $P_n := (-\theta)^n/n!$. Then we have*

$$\mathrm{L}(\delta^n) = (-1)^n\lambda(\{1\}^n) = i^n P_n$$

Proof. Observe that by the shuffle relations $n\,\mathrm{L}(\delta^n) = \mathrm{L}(\delta)\,\mathrm{L}(\delta^{n-1}) = -i\theta\,\mathrm{L}(\delta^{n-1})$ since $\mathrm{L}(\delta) = \log(1-\mu) = -i\theta$. The lemma follows easily by induction. □

Proposition 14.4.12. *For any integer $w \geq 2$ define $G_w := \sum_{w>n>0} \delta^n b \delta^{w-1-n}$. Then we have*

$$2\mathrm{Re}\,\mathrm{L}(G_w/i^w) = (w-1)P_w.$$

Proof. By shuffle relations we have

$$Ł(b)\,Ł(\delta^{w-1}) = Ł(b\delta^{w-1}) + G_w.$$

By stuffle relations we have

$$Ł(b)\,Ł(\delta^{w-1}) = Ł(b\delta^{w-1}) + (w-1)\,Ł(\delta^{w}) - \sum_{w>n>0} Ł(\delta^{w-1-n}a\delta^{n}).$$

This yields $Ł(G_w) + Ł(\widetilde{G_w}) = (w-1)\,Ł(\delta^w)$. Hence

$$Ł(G_w) + (-1)^w \overline{Ł}(G_w) = (w-1)i^w P_w$$

by Lemma 14.4.9 and 14.4.11. The proposition follows easily. □

14.4.2 *Central binomial sums*

In this subsection we consider the following *central binomial sums*:

$$\mathrm{CB}(k) := \sum_{n=1}^{\infty} \frac{1}{n^k \binom{2n}{n}} \tag{14.32}$$

We define

$$j(s,t) := \int_0^\theta \log^s\left(2\sin\frac{\tau}{2}\right)\tau^t d\tau, \tag{14.33}$$

$$\begin{aligned} r(s,t) :=& \frac{i^{t+1}}{s!t!}\int_0^\theta \log^s(1-e^{i\tau})\tau^t d\tau \\ =& \frac{i^{t+1}}{s!t!}\int_0^\theta \left(\log\left(2\sin\frac{\tau}{2}\right) + \frac{i}{2}(\tau-\pi)\right)^s \tau^t d\tau. \end{aligned} \tag{14.34}$$

Lemma 14.4.13. *For any $s,t\in\mathbb{N}_0$ the real number $j(s,t)$ can be written as a $\mathbb{Q}[i]$-linear combination of $r(c,d)$ with $c\le s$ and $c+d\le s+t$.*

Proof. By the definition $j(0,t) = (-i)^{t+1}t!r(0,t)$. By binomial expansion we see that

$$r(s,t) = \frac{i^{t+1}}{s!t!}j(s,t) + \sum_{c,d\in\mathbb{N},c<s,c+d\le s+t} q(c,d)j(c,d)$$

where $q(c,d)\in\mathbb{Q}[i]$. The lemma quickly follows from an induction on s. □

Example 14.4.14. We have

$$\begin{aligned} r(1,t) =& \frac{i^{t+1}}{t!}\left(j(1,t) + \frac{i}{2}j(0,t+1) - \frac{i\pi}{2}j(0,t)\right) \\ \Longrightarrow j(1,t) =& \frac{(-i)^{t+1}}{2}t!\Big(2r(1,t) - (t+1)r(0,t+1) + i\pi r(0,t)\Big) \end{aligned}$$

Lemma 14.4.15. *For all positive integers $k \geq 2$, we have*

$$\mathrm{CB}(k) = \frac{(-2)^{k-2}}{(k-2)!} j(k-2,1). \tag{14.35}$$

Proof. Note that Gamma function satisfies $s^{-n}\Gamma(n) = \int_0^1 y^{s-1}(-\log y)^{n-1}dy$. Hence

$$\begin{aligned}\mathrm{CB}(k) &= \sum_{n=1}^{\infty} \frac{1}{\binom{2n}{n} n^k} = \sum_{n=1}^{\infty} \frac{1}{\binom{2n}{n}\Gamma(n)} \int_0^1 (-\log x)^{k-1} x^{n-1} dx \\ &= \sum_{n=1}^{\infty} \frac{1}{\binom{2n}{n}\Gamma(n)} \int_0^1 (-2\log y)^{k-1} y^{2n-2} (2y) dy\end{aligned}$$

by substitution $x = y^2$. Using integration by part we get

$$\begin{aligned}\mathrm{CB}(k) &= \frac{(-2)^{k-1}}{(k-2)!} \int_0^1 \frac{(\log y)^{k-2}}{y} \sum_{n=1}^{\infty} \frac{y^{2n}}{n\binom{2n}{n}} dy \\ &= \frac{(-2)^{k-1}}{(k-2)!} \int_0^1 \frac{(\log y)^{k-2}}{y} \frac{y \sin^{-1}(y/2)}{\sqrt{1-(y/2)^2}} dy \\ &= \frac{(-2)^{k-2}}{(k-2)!} \int_0^{\theta} \log^{k-2}\left(2\sin\frac{\tau}{2}\right) \tau d\tau \quad (\text{by } y = 2\sin\frac{\tau}{2}).\end{aligned}$$

Here we have used the Maclaurin series expansion of $\sin^{-1}(t)/\sqrt{1-t^2}$ (see [65, p. 384]). □

Lemma 14.4.16. *For all positive integers n and k, we have $r(0,k) = \lambda(\{1\}^{k+1})$ and*

$$(-1)^{k+n} r(n,k) = \zeta(k+2,\{1\}^{n-1}) - \sum_{\ell=0}^{k} (-i)^{\ell} P_{\ell} \lambda(k+2-\ell,\{1\}^{n-1}). \tag{14.36}$$

Proof. The expression of $r(0,k)$ is easy to verify by using the definition Eq. (14.34) and Lemma 14.4.11. By Eq. (14.27) we have

$$\lambda(k+2,\{1\}^{n-1}) = \zeta(k+2,\{1\}^{n-1}) + \int_1^{\mu} \frac{\widetilde{\mathrm{Li}}_{k+1,\{1\}^{n-1}}(z)}{z} dz.$$

Applying integration by parts repeatedly we get

$$\begin{aligned}&\lambda(k+2,\{1\}^{n-1}) = \zeta(k+2,\{1\}^{n-1}) - \\ &\sum_{\ell=1}^{k} \frac{(-1)^{\ell}\lambda(k+2-\ell,\{1\}^{n-1})\log^{\ell}(\mu)}{\ell!} + \frac{(-1)^k}{k!} \int_1^{\mu} \frac{\widetilde{\mathrm{Li}}_{\{1\}^n}(z)\log^k(z)}{z} dz.\end{aligned}$$

Let $z = e^{i\tau}$. Since $\log(\mu) = i\theta$ by Eq. (14.29) we have

$$\lambda(k+2,\{1\}^{n-1}) = \zeta(k+2,\{1\}^{n-1}) - \sum_{\ell=1}^{k} \frac{(-i\theta)^{\ell}\lambda(k+2-\ell,\{1\}^{n-1})}{\ell!} + \frac{i(-1)^k}{k!n!}\int_0^{\theta} (-\log(1-e^{i\tau}))^n (i\tau)^k d\tau,$$

which readily reduces to Eq. (14.36). □

Theorem 14.4.17. *For any positive integer $k \geq 2$ the central binomial sum* $\mathrm{CB}(k)$ *lies in the weight k piece of the $\mathbb{Q}$-algebra generated by the MZVs, MGls and MCls.*

Proof. Notice that $\pi = -3\,\mathrm{mgl}(1)$ so the theorem follows from Lemmas 14.4.13, 14.4.15 and 14.4.16. □

14.4.3 *Dualities of multiple Clausen values*

We present two different duality results for MCVs in this subsection, depending on whether the first argument is equal to or greater than 1.

Theorem 14.4.18. *For all positive integers $s_1, s_2, \ldots, s_n$*

$$\begin{aligned}\lambda(1,s_1,\ldots,s_n) &= \lambda(1)\lambda(s_1,\ldots,s_n) - \lambda(2)\lambda(s_1-1,\ldots,s_n) + \cdots \\ &- (-1)^{s_1}\lambda(s_1)\lambda(1,s_2,\ldots,s_n) + (-1)^{s_1}\lambda(1,s_1)\lambda(s_2,\ldots,s_n) - \cdots \\ &- (-1)^{|\mathbf{s}|}\lambda(s_n,...,s_1)\lambda(1) + (-1)^{|\mathbf{s}|}\lambda(1,s_n,...,s_1).\end{aligned} \tag{14.37}$$

Proof. This can be proved by repeated integration by parts using Lemma 14.4.3. We leave the details to the interested reader. □

Example 14.4.19. In general, there are $|\mathbf{s}|+1$ terms on the right-hand side of Eq. (14.37). In weight 7 we have

$$\begin{aligned}\lambda(1,2,1,3) &= \lambda(1)\lambda(2,1,3) - \lambda(2)\lambda(1,1,3) + \lambda(1,2)\lambda(1,3) \\ &\quad - \lambda(1,1,2)\lambda(3) + \lambda(2,1,2)\lambda(2) - \lambda(3,1,2)\lambda(1) + \lambda(1,3,1,2).\end{aligned}$$

Theorem 14.4.20. *For $\mathbf{s},\mathbf{r} \in \mathbb{N}^d$ and $1 \leq j \leq d$, define the compositions*

$$\vec{a}_j = \vec{a}_j(\mathbf{s};\mathbf{r}) := (s_j+1,\{1\}^{r_j-1},\ldots,s_d+1,\{1\}^{r_d-1}),$$
$$\overleftarrow{b}_j = \overleftarrow{b}_j(\mathbf{s};\mathbf{r}) := (r_j+1,\{1\}^{s_j-1},\ldots,r_1+1,\{1\}^{s_1-1}),$$

and $\vec{a}_{d+1} := \overleftarrow{b}_0 := \{\}$. *Then we have*

$$\zeta(\vec{a}_1) = \zeta(\overleftarrow{b}_d) = \sum_{j=1}^{d} \left\{ \sum_{k=0}^{s_j} \overline{\lambda\big(s_j + 1 - k, \{1\}^{r_j-1}, \vec{a}_{j+1}\big)} \lambda\big(\{1\}^k, \overleftarrow{b}_{j-1}\big) + \sum_{k=1}^{r_j-1} \overline{\lambda\big(\{1\}^k, \vec{a}_{j+1}\big)} \lambda\big(r_j + 1 - k, \{1\}^{s_j-1}, \overleftarrow{b}_{j-1}\big) \right\} + \lambda\big(\overleftarrow{b}_d\big). \quad (14.38)$$

Proof. Apply the Hölder convolution as presented in Thm. 14.1.1 with $p = \mu$ and $q = 1 - \mu = \overline{\mu}$. Note that in Thm. 14.1.1 $\lambda_p = \overline{\lambda}$ and $\lambda_q = \lambda$. □

Example 14.4.21. In general, there are $|\mathbf{s}| + |\mathbf{r}| + 1$ λ-product terms on the right-hand of Eq. (14.38). Take $\mathbf{s} = (2, 1)$ and $\mathbf{r} = (3, 2)$ in Eq. (14.38). Then we have

$$\begin{aligned}\zeta(3,1,1,2,1) = &\overline{\lambda(3,1,1,2,1)} + \overline{\lambda(2,1,1,2,1)}\lambda(1) + \overline{\lambda(1,1,1,2,1)}\lambda(1,1) \\ &+ \overline{\lambda(1,2,1)}\lambda(3,1) + \overline{\lambda(1,1,2,1)}\lambda(2,1) + \overline{\lambda(2,1)}\lambda(4,1) \\ &+ \overline{\lambda(1,1)}\lambda(1,4,1) + \overline{\lambda(1)}\lambda(2,4,1) + \lambda(3,4,1) = \zeta(3,4,1).\end{aligned}$$

14.4.4 *A* 1-2-1 *formula of multiple Glaishers*

Our goal in this subsection is to prove a **1-2-1** type formula for the MGl $\mathrm{mgl}(\{1\}^m, 2, \{1\}^n)$.

First, recall that $\lambda_z(\mathbf{s}) = \widetilde{\mathrm{Li}}_{\mathbf{s}}(1/z)$ for any $|z| \geq 1$ where the $\widetilde{\mathrm{Li}}$ is the one variable MPLs defined by Eq. (14.1).

Proposition 14.4.22. *For all* $|z| \geq 1$ *we have*

$$F_z(x, y) := \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} x^{m+1} y^{n+1} \lambda_z(m + 2, \{1\}^n) = 1 - {}_2F_1\left(\begin{matrix} y, -x \\ 1 - x \end{matrix} \middle| \frac{1}{z} \right).$$

Proof. By the definition,

$$\begin{aligned}&\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} x^{m+1} y^{n+1} \lambda_z(m + 2, \{1\}^n) \\ =&y \sum_{m=0}^{\infty} x^{m+1} \sum_{n=0}^{\infty} \frac{1}{(n+1)^{m+2} z^k} \prod_{j=1}^{n} \left(1 + \frac{y}{j}\right) = \sum_{m=0}^{\infty} x^{m+1} \sum_{k=1}^{\infty} \frac{(y)_k}{k^{m+1} k! z^k} \\ =&\sum_{k=1}^{\infty} \frac{(y)_k}{k! z^k} \frac{x}{k - x} = -\sum_{k=1}^{\infty} \frac{(y)_k (-x)_k}{k! z^k (1 - x)_k} = 1 - {}_2F_1\left(\begin{matrix} y, -x \\ 1 - x \end{matrix} \middle| \frac{1}{z} \right)\end{aligned}$$

as desired. □

Lemma 14.4.23. *Define*

$$\begin{aligned} M(x,y) &:= F_{\overline{\mu}}(ix,-iy) \\ &= \sum_{m,n\geq 0} (-1)^{n+1}\Big(\operatorname{mgl}(m+2,\{1\}^n) + i\operatorname{mcl}(m+2,\{1\}^n)\Big)x^{m+1}y^{n+1}. \end{aligned}$$

Then we have

$$e^{\theta x}M(x,y) + e^{\theta y}\overline{M(y,x)} + (e^{\theta x}-1)(e^{\theta y}-1) = F_1(ix,-iy). \tag{14.39}$$

Moreover,

$$F_1(x,y) = 1 - \exp\left(\sum_{k\geq 2}\frac{x^k+y^k-(x+y)^k}{k}\zeta(k)\right). \tag{14.40}$$

Proof. Taking $d=1$ in Thm. 14.4.20 we get (see Exercise 14.6)

$$\begin{aligned} &\zeta(m+2,\{1\}^n) + \frac{(-1)^m(i\theta)^{m+n+2}}{(m+1)!(n+1)!} = \\ &\qquad \sum_{r=0}^{m}\frac{(-i\theta)^r}{r!}\lambda(m+2-r,\{1\}^n) + \sum_{r=0}^{n}\frac{(i\theta)^r}{r!}\overline{\lambda(n+2-r,\{1\}^m)}. \end{aligned}$$

Multiplying the above by $x^{m+1}y^{n+1}$ and summing over all $m,n\geq 0$, we get

$$e^{-ix\theta}F_{\overline{\mu}}(x,y) + e^{iy\theta}\overline{F_{\overline{\mu}}(y,x)} + (e^{-ix\theta}-1)(e^{iy\theta}-1) = F_1(x,y), \tag{14.41}$$

by Prop. 14.4.22. Clearly Eq. (14.41) yields Eq. (14.39). Further, by the definition

$$F_1(x,y) = 1 - {}_2F_1\left(\begin{matrix} y,-x \\ 1-x \end{matrix}\middle|\, \frac{1}{z}\right) = 1 - \frac{\Gamma(1-x)\Gamma(1-x)}{\Gamma(1-x-y)}.$$

by applying the Gauss summation formula C.2. Finally, Eq. (14.40) can be proved by Eq. (3.16). □

Lemma 14.4.24. *For every non-negative integer n*

$$\operatorname{mgl}(2,\{1\}^n) = (-1)^{n+1}\frac{\theta^{n+2}}{2(n+2)!}. \tag{14.42}$$

Proof. Multiplying by y/x on both sides of Eq. (14.39)) and then considering the real parts after taking $x\to 0$, we get

$$B(y) + e^{\theta y}C(y) + \theta y(e^{\theta y}-1) = D(y) \tag{14.43}$$

where

$$B(y) := \sum_{n=0}^{\infty} (-1)^{n+1} \operatorname{mgl}(2, \{1\}^n) y^{n+2} \quad (m = 0),$$
$$C(y) := -\sum_{m=0}^{\infty} \operatorname{mgl}(m+2) y^{m+2} \quad (n = 0),$$
$$D(y) := \sum_{m=1}^{\infty} (-1)^{m-1} \zeta(2m) y^{2m} = \frac{\pi y}{e^{2\pi y} - 1} - \frac{1}{2} + \frac{\pi y}{2}.$$

by $(-1)^{m-1}\zeta(2m) = (1/2)(2\pi)^{2m} B_{2m}/(2m)!$. By Eq. (14.31)

$$\begin{aligned} C(y) =& \frac{1}{2} \sum_{m=0}^{\infty} B_{m+2}\Big(\frac{1}{6}\Big) \frac{(-2\pi y)^{m+2}}{(m+2)!} \\ =& \frac{1}{2} \sum_{k=2}^{\infty} B_k\Big(\frac{1}{6}\Big) \frac{(-2\pi y)^k}{k!} \\ =& \frac{-\pi y e^{-\theta y}}{e^{-2\pi y} - 1} - \frac{1}{2}\Big(1 - B_1\Big(\frac{1}{6}\Big) 2\pi y\Big) \\ =& \frac{\pi y e^{2\pi y - \theta y}}{e^{2\pi y} - 1} - \frac{1}{2} - \theta y \end{aligned}$$

After simplification Eq. (14.43) yields

$$B(y) = \frac{e^{\theta y}}{2} - \frac{\theta y}{2} - \frac{1}{2}$$

which is equivalent to Eq. (14.42). □

Finally, we can prove the following **1-2-1** type formula.

Theorem 14.4.25. *For non-negative integers m and n*

$$\operatorname{mgl}(\{1\}^m, 2, \{1\}^n) = -\frac{1}{2} \cdot \frac{(-\theta)^w}{w!}, \tag{14.44}$$

where $w = m + n + 2$.

Proof. By repeated integration by parts, and using $\operatorname{Li}_{\{1\}^m}(\mu) = (i\theta)^n/n!$

and Eq. (14.26), we get

$$\begin{aligned}\lambda(\{1\}^n,2,\{1\}^m) &= \int_0^\mu \frac{\mathrm{Li}_{\{1\}^{n-1},2,\{1\}^m}(z)}{1-z}dz\\ &= i\theta\lambda(\{1\}^{n-1},2,\{1\}^m)+\cdots+(-1)^{n+1}\frac{(i\theta)^n}{n!}\lambda(2,\{1\}^m)\\ &\quad+\int_0^\mu \frac{(\log(1-z))^n}{n!}\cdot\frac{\mathrm{Li}_{\{1\}^{m+1}}(z)}{z}dz\\ &= i\theta\lambda(\{1\}^{n-1},2,\{1\}^m)+\cdots+(-1)^{n+1}\frac{(i\theta)^n}{n!}\lambda(2,\{1\}^m)\\ &\quad+(-1)^n\binom{m+n+1}{n}\lambda(2,\{1\}^{m+n})\end{aligned}$$

using Eqs. (14.25) and (14.29). Multiplying by i^{m+n+2} and taking the real parts, we get

$$\begin{aligned}&\mathrm{mgl}(\{1\}^n,2,\{1\}^m)+\theta\lambda(\{1\}^{n-1},2,\{1\}^m)+\cdots+\frac{\theta^n}{n!}\lambda(2,\{1\}^m)\\ &\qquad=(-1)^n\binom{m+n+1}{n}\lambda(2,\{1\}^{m+n})\end{aligned}$$

Using Lemma 14.4.24 and induction we see that the theorem follows immediately from the following binomial identity

$$\sum_{j=0}^{n}(-1)^j\binom{m+n+2}{j}=(-1)^n\binom{m+n+1}{n}.$$

But this clearly follows from the well-known identity

$$\binom{m+n+2}{j}=\binom{m+n+1}{j}+\binom{m+n+1}{j-1}.$$

We have completed the proof of the theorem. □

Proposition 14.4.26. *For every integer $w\geq 2$, we have*

$$(-1)^w\sum_{j=0}^{w-2}\frac{\theta^j}{j!}\,\mathrm{mcl}(w-j)=\begin{cases}\mathrm{mcl}(2,\{1\}^{w-2}), & \text{if } 2|w;\\ \mathrm{mcl}(2,\{1\}^{w-2})+(-1)^{(w-1)/2}\zeta(w), & \text{if } 2\nmid w.\end{cases}$$

In particular, every MCl $\mathrm{mcl}(2,\{1\}^n)$ *is a $\mathbb{Q}[\pi]$-linear combination of depth one MCls and at most one Riemann zeta value.*

Proof. Multiplying by y/x on both sides of Eq. (14.39)) and then considering the imaginary parts after taking $x\to 0$, we get

$$E(y)+e^{\theta y}F(y)=G(y)$$

where

$$E(y) := \sum_{n=0}^{\infty} (-1)^{n+1} \operatorname{mcl}(2, \{1\}^n) y^{n+2} \quad (m = 0),$$
$$F(y) := \sum_{m=0}^{\infty} \operatorname{mcl}(m+2) y^{m+2} \quad (n = 0),$$
$$G(y) := \sum_{m=1}^{\infty} (-1)^{m-1} \zeta(2m+1) y^{2m+1}.$$

The proposition follows immediately. □

[88, (127)] and [89]

In [29] a precise evaluation of CB(20) is given.

14.4.5 *Dimension conjectures*

Let MGl_w, MCl_w, MCV_w and MDV_w be the $\mathbb{Q}$-vector space generated by weight w MGls, MCls, MCVs and MDVs respectively. First, since $\operatorname{Li}_1(\mu) = i\theta$ and $\operatorname{Li}_{1,1}(\mu) = -\theta^2/2$ we have

$$\operatorname{mgl}(1) = -\theta,\ \operatorname{mcl}(1) = 0,\ \operatorname{mgl}(1,1) = \theta^2/2,\ \operatorname{mcl}(1,1) = 0.$$

Further $\operatorname{mgl}(2) = -\theta^2/4$ by Eq. (14.42) and

$$\operatorname{mcl}(2) = -\frac{\sqrt{3}}{2}\left(1 + \frac{1}{2^2} - \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} + \frac{1}{8^2} - \frac{1}{10^2} - \frac{1}{11^2} + \cdots\right) < 0.$$

Thus we know $\mathsf{MGl}_1 = \langle \pi \rangle$, $\mathsf{MCl}_1 = \emptyset$, $\mathsf{MCV}_1 = \langle i\pi \rangle$, $\mathsf{MGl}_2 = \langle \pi^2 \rangle$, $\mathsf{MCl}_2 = \langle \operatorname{mcl}(2) \rangle$, and $\mathsf{MCV}_2 = \langle \pi^2, i \operatorname{mcl}(2) \rangle$. We see easily that MCVs are automatically MDVs. In weight one and two cases, the only possible extra MDV is

$$\mathrm{L}(\delta b) = \overline{\mathrm{L}(a\delta)} = -\overline{\operatorname{Li}_2(\mu)} = -\frac{1}{4}\theta^2 - i \operatorname{mcl}(2)$$

by Lemma 14.4.9. So $\mathsf{MDV}_1 = \mathsf{MCV}_1$ and $\mathsf{MDV}_2 = \mathsf{MCV}_2$.

The following conjecture are supported by extensive numerical data. In particular, it says the dimensions of the MCVs are given by Fibonacci numbers.

Conjecture 14.4.27. *Set* $\dim \mathsf{MCl}_0 =$ *and* $\dim \mathsf{MGl}_0 = \dim \mathsf{MCV}_0 = 1$. *For any positive integer* $w \geq 2$, *we have*

$$\dim \mathsf{MGl}_w = \dim \mathsf{MGl}_{w-1} + \dim \mathsf{MCl}_{w-2},$$
$$\dim \mathsf{MCl}_w = \dim \mathsf{MCl}_{w-1} + \dim \mathsf{MGl}_{w-2},$$
$$\dim \mathsf{MCV}_w = \dim \mathsf{MCV}_{w-1} + \dim \mathsf{MCV}_{w-2}.$$

Moreover, for all $w \geq 1$, *we have* $\mathsf{MDV}_w = \mathsf{MCV}_w$

Definition 14.4.28. An *irreducible* MDV is a MDV which is not a $\mathbb{Q}$-linear combination of products of MDVs of lower depths and weights.

Similar to the situation for MZVs and Euler sums, we may define $p_{n,d}$ to be be the minimum number of irreducible MDVs of weight n and depth d, such that a rational basis at this weight and depth is formed from these irreducibles, augmented by products of irreducibles of lower weights and depths. The following conjecture is an analog of Conjecture 14.2.1.

Conjecture 14.4.29. *We have*

$$\prod_{n\geq 2}\prod_{d\geq 1}(1-x^n y^d)^{p_{n,d}} = 1-\frac{x^2 y}{1-x}. \tag{14.45}$$

In order to find the finer structures of MDVs we need to use the Lyndon words on the alphabet of all positive integers with its natural order (see Definition 3.3.3).

Conjecture 14.4.30. *Let $w > d$ be two positive integers. Then one can choose the irreducibles by using $\mathrm{L}(\mathbf{w})$ where $\mathbf{w} = \delta$ or $\mathbf{w} = a^{n_1-1}\delta a^{n_2-1}\delta\ldots a^{n_d-1}\delta$ subject to the following conditions:*

(i) $w=\sum_{j=1}^{d} n_j$,
(ii) $n_j \geq 2$,
(iii) $(n_d,\ldots,n_1)$ *is a Lyndon word.*

For example, the words corresponding to the first few irreducible MDVs are given below:

$$\delta,\ a\delta,\ a^2\delta,\ a^3\delta,\ a^4\delta,\ a^2\delta a\delta,\ a^5\delta,\ldots$$

Using this we can produce the conjectured basis of MDVs in lower weights with following words:

$w=1:\delta$;
$w=2:\delta^{\sqcup\!\sqcup 2},\ a\delta$;
$w=3:\delta\sqcup\!\sqcup a\delta,\ \delta^{\sqcup\!\sqcup 3},\ a^2\delta$;
$w=4:\delta^{\sqcup\!\sqcup 2}\sqcup\!\sqcup a\delta,\ \delta^{\sqcup\!\sqcup 4},\ \delta\sqcup\!\sqcup a^2\delta,\ a\delta\sqcup\!\sqcup a\delta$;
$w=5:\delta^{\sqcup\!\sqcup 3}\sqcup\!\sqcup a\delta,\ \delta^{\sqcup\!\sqcup 5},\ \delta^{\sqcup\!\sqcup 2}\sqcup\!\sqcup a^2\delta,\ \delta\sqcup\!\sqcup a\delta\sqcup\!\sqcup a\delta,\ a^2\delta\sqcup\!\sqcup a\delta,\ a^4\delta,\ a^2\delta a\delta;\ldots$

We end this subsection with a conjecture which plays an important role in Broadhurst's discovery of a significant basis in [89] which is different from the one given in Conjecture 14.4.30. It can be used to express numerically the counterterm $P_{7,11}$ of the following 7-loop subdivergence-free Feynman

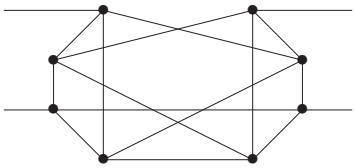


Figure 14.2: The unique 7-loop subdivergence-free Feynman diagram.

diagram in ϕ^4 theory with well-behaved rational coefficients whose denominators don't have large prime factors:

Conjecture 14.4.31. *At odd weight $w > 1$, there exists a* unique $\mathbb{Z}$*-linear reducible combination*

$$\sum_{k=1}^{(w-1)/2} C_{w,k} \mathrm{Im}\, \text{Ł}\left(a^{w-2k-1}\delta a^{2k-1} b\right) \in \pi^w \mathbb{Q}, \tag{14.46}$$

with $C_{w,1} > 0$ and $\gcd\{C_{w,k} : k = 1, \ldots, (w-1)/2\} = 1$. At even weight w, the depth-2 real part $\mathrm{Re}\,\text{Ł}(a^{w-2}\delta^2)$ is reducible to MZVs.

14.5* Level p and p^2 cases ($p \geq 5$ a prime)

In this optional section, we assume to reader to be familiar with the basic theory of the motivic fundamental group laid out in [156]. Let $p \geq 5$ be prime throughout this section. From Thm. 13.2.1 we see that $d(2,p) \leq (p+1)^2/4$. In this section, we will improve this bound in Thm. 14.5.1 by using the motivic setup.

Let $\mathbb{K}$ be the field over $\mathbb{Q}$ generated by a primitive Nth root of unity and $\mathcal{O}$ its ring of integers. Let Γ be the $\mathbb{Q}$-vector subspace of $\mathcal{O}[1/N]^* \otimes \mathbb{Q}$ generated by $1-\xi$ for $1 \neq \xi \in \Gamma_N$. We refer to [156] for the definition of the Tannakian category of mixed Tate motives $\mathrm{MT}(k)_\Gamma$, the definition of the fibre functor $\omega(M) := \bigoplus_n \mathrm{Hom}(\mathbb{Q}(n), \mathrm{Gr}^W_{-2n}(M))$, and of the corresponding motivic fundamental group G. Let $\mathfrak{L}$ be the Lie algebra generated by the generators $\mathtt{x}_\xi, \xi \in \{0,\infty\} \cup \Gamma_N$, with the only relation $\sum \mathtt{x}_\xi = 0$. Set the degree of $\mathtt{x}_\xi$ to be 1. Then $\mathfrak{L}$ is freely generated by the $\mathtt{x}_\xi, \xi \neq \infty$. Let Π be the pro-nilpotent group $\lim_n \exp(\mathfrak{L}/\deg^{>n})$. The completed enveloping algebra of $\mathfrak{L}$ is nothing but $\mathbb{Q}\langle\langle \mathtt{x}_\xi \rangle\rangle_{\xi \in \Gamma_N \cup \{0\}}$ with the coproduct $\Delta : \mathtt{x}_\xi \mapsto \mathtt{x}_\xi \otimes 1 + 1 \otimes \mathtt{x}_\xi$, and Π is the subscheme of the group-like elements under Δ.

Let H_ω be the group of Γ_N-equivariant automorphisms of Π defined in [156, § 5.8]. Set $G_\omega := \omega(G)$, $U_\omega := \ker(G_\omega \to \mathbb{G}_m)$ and $V_\omega := \ker(H_\omega \to$

$\mathbb{G}_m$). There are morphisms of group schemes $G_\omega \to H_\omega$ and $U_\omega \to V_\omega$ compatible with the semi-direct product structure

$$\iota : G_\omega = \mathbb{G}_m \ltimes U_\omega \to H_\omega = \mathbb{G}_m \ltimes V_\omega. \tag{14.47}$$

By [156, § 6.13] one may safely replace $\mathcal{G}_N^{(\ell)}$ by $\mathcal{G}$ throughout [240]. In particular, Goncharov's results can now be used to study the structure of Galois Lie algebra $\mathcal{G}_{\bullet,\bullet}(\Gamma_N) = \sum_{w\geq 1, l\geq 1} \mathcal{G}_{-w,-l}(\Gamma_N)$ graded by the weight w and depth l. The weight filtration is defined via the action of $\mathbb{G}_m$ and the depth filtration comes from the descending central series of $\ker f_*$, where $f_* : \pi_1(\mathbb{G}_m - \Gamma_N, x) \longrightarrow \pi_1(\mathbb{G}_m)$ for any base point x is induced from the embedding $\mathbb{G}_m - \Gamma_N \hookrightarrow \mathbb{G}_m$.

Theorem 14.5.1. *Let $p \geq 5$ be a prime. Then*

$$d(2,p) \leq \frac{(5p+7)(p+1)}{24}. \tag{14.48}$$

If Grothendieck's period conjecture (see [156, § 5.27(c)]*) is true then the equality holds and all the $\mathbb{Q}$-linear relations in $\mathsf{CMZV}(2,p)$ follow from the standard relations.*

Proof. Let $\{\ ,\ \}$ be the Lie bracket in $\mathcal{G}_{\bullet,\bullet}(\Gamma_N)$. Now in depth 1 one has $\dim \mathcal{G}_{-2,-1}(\Gamma_p) = (p-1)/2$ (see [240, Thm. 2.1]). By [240, Cor. 2.16]

$$\dim \mathcal{G}_{-2,-2}(\Gamma_p) = \frac{(p-1)(p-5)}{12}. \tag{14.49}$$

For any positive integer N define

$$\beta_N : \bigwedge^2 \mathcal{G}_{-2,-1}(\Gamma_N) \longrightarrow \mathcal{G}_{-2}(\Gamma_N), \quad a \wedge b \longmapsto \{a,b\}. \tag{14.50}$$

Then

$$\dim(\ker \beta_p) = \frac{1}{2}\frac{p-1}{2}\left(\frac{p-1}{2} - 1\right) - \frac{(p-1)(p-5)}{12} = \frac{p^2-1}{24}. \tag{14.51}$$

Let SR be the affine subgroup of Π defined by only those polynomial equations satisfied by the coefficients of $\mathcal{I}$ which are deduced from the standard relations, plus "$2\pi i = 0$", as explained in [156, § 5.22]. This is the group DMRD_0 studied by Racinet in [465, § 3.2, Thm. I]. Its Lie algebra $\operatorname{Lie} SR$ is graded by weight and depth, independent of the embedding by Prop. 4.1 of op. cit. and contains $\mathcal{G}$ by [156, § 5.22]. Further, Goncharov shows in [240, § 7.7] that the standard relations provide a complete list of constraints on the diagonal part of the Lie algebra $\mathcal{G}$ in depth ≤ 2, yielding $(\operatorname{Lie} SR)_{m,m} = \mathcal{G}_{m,m}$ for $m = -1, -2$. Together with Eq. (14.51) this

means in the proof of [156, § 5.25] one can decrease the bound $D(2,p)$ in Thm. 13.2.1 by $(p^2-1)/24$ and arrive at the bound $(5p+7)(p+1)/24$.

Grothendieck's conjecture implies that $\mathcal{G} = \operatorname{Lie} R$ where R is the affine subgroup of Π defined by all the polynomial equations satisfied by the coefficients of $\mathcal{I}$ plus "$2\pi i = 0$". Therefore one obtains both the equality in Eq. (14.48) and the statement about the completeness of the standard relations. This concludes the proof of the theorem. □

Remark 14.5.2. (a) Notice that when $p \geq 11$ the vector space $\ker \beta_p$ contains a subspace isomorphic to the space of cusp forms of weight two on $X_1(p)$ which has dimension $(p-5)(p-7)/24$ (see [240, Lemma 2.3 & Thm. 7.8]). So it must contain another piece which has dimension $(p-3)/2$ by (14.51). What is this missing piece?

(b) When $w > 2$ it's not too hard to improve the bound of $d(w,p)$ given in [156, § 5.24] by the same idea as used in the proof of the theorem (for example, decrease the bound by $(p^2-1)/24$). But they are often not the best.

Turn now to the case of level $N = p^2$. Notice that $\varphi(p^2)(\varphi(p^2)-2)/4$ is a trivial upper bound of $\dim \mathcal{G}_{-2,-2}(\Gamma_{p^2})$ (see [240, Thm. 2.1]). If one had any nontrivial upper bound of $\dim \mathcal{G}_{-2,-2}(\Gamma_{p^2})$ one would have a nontrivial lower bound of $\dim(\ker \beta_{p^2})$ where β_{p^2} is defined by (14.50). This would in turn provide an improved upper bound of $d(2,p^2)$ similar to Thm. 14.5.1. However, at present we can only show that

Theorem 14.5.3. *If $p \geq 5$ is a prime then* $\ker \beta_{p^2} \neq 0$ *and* $d(2,p^2) < p^2(p-1)^2/4$.

Proof. Fix a primitive p^2th root of unity μ. Put $e(a) = \mathtt{x}_{\mu^a}$ for all integer a. Define

$$g_{k,j} = e(pk+j) + e(p^2-pk-j) + e(pj) + e(p^2-pj)$$

for $0 \leq k < (p-1)/2$, $1 \leq j \leq p-1$, and for $k = (p-1)/2$, $1 \leq j \leq (p-1)/2$. One only needs to prove the following

Claim. Let $h = (p-3)/2$. Then one has

$$\begin{aligned}
&\sum_{k=0}^{h}\sum_{l=k}^{h}\sum_{j=2}^{p-2}\{g_{k,1}, g_{l,j}\} + \sum_{k=0}^{h+1}\sum_{j=2}^{h+1}\{g_{k,1}, g_{h+1,j}\} \\
&+\sum_{k=0}^{h}\sum_{l=k+1}^{h}\sum_{j=2}^{p-2}\{g_{k,p-1}, g_{l,j}\} + \sum_{k=0}^{h}\sum_{j=2}^{h+1}\{g_{k,p-1}, g_{h+1,j}\} \\
&-\sum_{k=0}^{h}\sum_{l=k}^{h}\sum_{j=2}^{p-2}\{g_{k,j}, g_{l,p-1}\} - \sum_{k=0}^{h}\sum_{l=k}^{h}\sum_{j=2}^{p-2}\{g_{k,j}, g_{l+1,1}\} = 0.
\end{aligned}$$

There are $h(2h+3)^2 = hp^2$ distinct terms on the left, each with coefficient ± 1.

The proof of the claim is straight-forward by a little tedious change of indices and regrouping.

$$-\sum_{k=0}^{h}\sum_{l=k}^{h}\sum_{j=2}^{p-2}\{g_{k,j}, g_{l+1,1}\} = \sum_{k=0}^{h}\sum_{l=0}^{k}\sum_{j=2}^{p-2}\{g_{k+1,1}, g_{l,j}\} = \sum_{k=1}^{h+1}\sum_{l=0}^{k-1}\sum_{j=2}^{p-2}\{g_{k,1}, g_{l,j}\}.$$

Then the expression in the claim becomes

$$\begin{aligned}
&\sum_{k=1}^{h}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,1}, g_{l,j}\} + \sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{0,1}, g_{l,j}\} + \sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{h+1,1}, g_{l,j}\} \\
&+\sum_{k=0}^{h+1}\sum_{j=2}^{h+1}\{g_{k,1}, g_{h+1,j}\} + \sum_{k=0}^{h}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,p-1}, g_{l,j}\} + \sum_{k=0}^{h}\sum_{j=2}^{h+1}\{g_{k,p-1}, g_{h+1,j}\} \\
&= \sum_{k=0}^{h+1}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,1}, g_{l,j}\} + \sum_{k=0}^{h+1}\sum_{j=2}^{h+1}\{g_{k,1}, g_{h+1,j}\} \\
&+\sum_{k=0}^{h}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,p-1}, g_{l,j}\} + \sum_{k=0}^{h}\sum_{j=2}^{h+1}\{g_{k,p-1}, g_{h+1,j}\}
\end{aligned}$$

By abuse of notation, we use $\{a,b\}=\{e(a),e(b)\}$. By the definition

$$\begin{aligned}
&\{g_{k,1},g_{l,j}\}\\
=&\{pk+1,pl+j\}+\{-pk-1,pl+j\}+\{p,pl+j\}+\{-p,pl+j\}\\
&+\{pk+1,-pl-j\}+\{-pk-1,-pl-j\}+\{p,-pl-j\}\\
&+\{-p,-pl-j\}+\{pk+1,pj\}+\{-pk-1,pj\}+\{p,pj\}+\{-p,pj\}\\
&+\{pk+1,-pj\}+\{-pk-1,-pj\}+\{p,-pj\}+\{-p,-pj\}\\
=&\{pk+1,pl+j\}+\{p(p-k)-1,pl+j\}+\{p,pl+j\}+\{-p,pl+j\}\\
&+\{pk+1,p(p-1-l)+p-j\}+\{p(p-k)-1,p(p-1-l)+p-j\}\\
&+\{p,p(p-1-l)+p-j\}+\{-p,p(p-1-l)+p-j\}+\{pk+1,pj\}\\
&+\{p(p-k)-1,pj\}+\{p,pj\}+\{-p,pj\}+\{pk+1,p(p-j)\}\\
&+\{p(p-k)-1,p(p-j)\}+\{p,p(p-j)\}+\{-p,p(p-j)\}
\end{aligned}$$

Then

$$\begin{aligned}
&\sum_{k=0}^{h+1}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,1},g_{l,j}\}\\
=&\sum_{k=0}^{h+1}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{pk+1,pl+j\}+\{p(p-k)-1,pl+j\}\\
&+\{pk+1,p(p-1-l)+j\}+\{p(p-k)-1,p(p-1-l)+j\}\\
&+\{p,pl+j\}+\{-p,pl+j\}+\{p,p(p-1-l)+j\}+\{-p,p(p-1-l)+j\}\\
&+2\{pk+1,pj\}+2\{p(p-k)-1,pj\}+2\{p,pj\}+2\{-p,pj\}\\
=&\sum_{k=0}^{h+1}\sum_{l=0,l\neq h+1}^{p-1}\sum_{j=2}^{p-2}\{pk+1,pl+j\}+\sum_{k=h+2}^{p}\sum_{l=0,l\neq h+1}^{p-1}\sum_{j=2}^{p-2}\{pk-1,pl+j\}\\
&+\sum_{k=0}^{h+1}\sum_{l=0,l\neq h+1}^{p-1}\sum_{j=2}^{p-2}\Big(\{p,pl+j\}+\{-p,pl+j\}\Big)+2(h+1)\sum_{k=0}^{h+1}\sum_{j=2}^{p-2}\{pk+1,pj\}\\
&+2(h+1)\sum_{k=h+2}^{p}\sum_{j=2}^{p-2}\{pk-1,pj\}+2(h+2)(h+1)\sum_{j=2}^{p-2}\Big(\{p,pj\}+\{-p,pj\}\Big)
\end{aligned}$$

Thus

$$\begin{aligned}
&\sum_{k=0}^{h+1}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,1},g_{l,j}\}+\sum_{k=0}^{h+1}\sum_{j=2}^{h+1}\{g_{k,1},g_{h+1,j}\}\\
=&\sum_{k=0}^{h+1}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\{pk+1,pl+j\}+\sum_{k=h+2}^{p}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\{pk-1,pl+j\}
\end{aligned}$$

$$+\frac{p+1}{2}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\Big(\{p,pl+j\}+\{-p,pl+j\}\Big)+p\sum_{k=0}^{h+1}\sum_{j=2}^{p-2}\{pk+1,pj\}$$
$$+p\sum_{k=h+2}^{p}\sum_{j=2}^{p-2}\{pk-1,pj\}+\frac{p^2-1}{2}\sum_{j=2}^{p-2}\Big(\{p,pj\}+\{-p,pj\}\Big)$$

Similarly,

$$\sum_{k=0}^{h}\sum_{l=0}^{h}\sum_{j=2}^{p-2}\{g_{k,p-1},g_{l,j}\}+\sum_{k=0}^{h}\sum_{j=2}^{h+1}\{g_{k,p-1},g_{h+1,j}\}$$
$$=\sum_{k=1}^{h+1}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\{pk-1,pl+j\}+\sum_{k=h+2}^{p-1}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\{pk+1,pl+j\}$$
$$+\frac{p-1}{2}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\Big(\{p,pl+j\}+\{-p,pl+j\}\Big)+p\sum_{k=1}^{h+1}\sum_{j=2}^{p-2}\{pk-1,pj\}$$
$$+p\sum_{k=h+2}^{p-1}\sum_{j=2}^{p-2}\{pk+1,pj\}+\frac{(p-1)^2}{2}\sum_{j=2}^{p-2}\Big(\{p,pj\}+\{-p,pj\}\Big)$$

Altogether the expression in the claim is reduced to

$$\sum_{k=0}^{p-1}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\{pk+1,pl+j\}+\sum_{k=1}^{p}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\{pk-1,pl+j\}$$
$$+p\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\Big(\{p,pl+j\}+\{-p,pl+j\}\Big)+p\sum_{k=0}^{p-1}\sum_{j=2}^{p-2}\{pk+1,pj\}$$
$$+p\sum_{k=1}^{p}\sum_{j=2}^{p-2}\{pk-1,pj\}+p(p-1)\sum_{j=2}^{p-2}\Big(\{p,pj\}+\{-p,pj\}\Big)$$

This is equal to

$$\sum_{k=0}^{p-1}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\Big([pk+1,pl+j]-[p(k+l)+j+1,pl+j]$$
$$+[p(k+l)+j+1,pk+1]\Big)$$
$$\sum_{k=1}^{p}\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\Big([pk-1,pl+j]-[p(k+l)+j-1,pl+j]$$
$$+[p(k+l)+j-1,pk+1]\Big)$$

$$
\begin{aligned}
&+p\sum_{l=0}^{p-1}\sum_{j=2}^{p-2}\Big([p,pl+j]-[p(l+1)+j,pl+j]+[p(l+1)+j,p]\\
&\qquad\qquad+[-p,pl+j]-[p(l-1)+j,pl+j]+[p(l-1)+j,-p]\Big)\\
&+p\sum_{k=0}^{p-1}\sum_{j=2}^{p-2}\Big([pk+1,pj]-[p(j+k)+1,pj]+[p(j+k)+1,pk+1]\Big)\\
&+p\sum_{k=1}^{p}\sum_{j=2}^{p-2}\Big([pk-1,pj]-[p(j+k)-1,pj]+[p(j+k)-1,pk-1]\Big)\\
&+p(p-1)\sum_{j=2}^{p-2}\Big([p,pj]-[p(j+1),pj]+[p(j+1),p]\\
&\qquad\qquad+[-p,pj]-[p(j-1),pj]+[p(j-1),-p]\Big)=0
\end{aligned}
$$

after many cancelations. □

14.6 Computational results in weight ≤ 5

In this last section we briefly discuss our results in weight three, four and five. Since the computational complexity increases exponentially with the weight we cannot do as many cases as we have done in weight two.

Combining the FDS in weight 3 to 5 and the weight one relations Eq. (13.21) and using Maple we have verified that $d(3,1)=1, d(3,2)\leq 3, d(3,3)\leq 8$....

N	1	2	3	4	5	6	7	8	9	10	11
$SR(3)$	1	3	8	9	22	23	50	38	67	70	157
$D(3)$	1	3	8	8	27	21	64	27	64	56	216
$SR(4)$	1	5	16	21	61	69					
$D(4)$	1	5	16	16	81	55	256				
$SR(5)$	2	8	32								
$D(5)$	2	8	32	32	243	144	1024				

Table 14.1: Upper bounds of $d(w,N)$ by the standard relations and [156, § 5.25].

We have done similar computation in other small weight and low level cases and listed the results in Table 14.1. The values of Deligne and Gon-

charov's bound $D(w) = D(w, N)$ in Table 14.1 should be compared with the bound $SR(w) = SR(w, N)$ obtained by the standard relations.

Note that $SR(3, 4) = D(3, 4) + 1$ as we saw in Sec. 14.3. In weight 5 and level 4, there are five non-standard relations (see Exercise 14.8). We now can prove this by using the octahedral symmetry of $\mathbb{P}^1 - (\{0, \infty\} \cup \Gamma_4)$ (see Remark 13.3.40).

Theorem 14.6.1. *If the conjecture in* [156, § 5.27(c)] *is true then all the linear relations among CMZVs of level* 4 *and weight* 3 *(resp. weight* 4*) are the consequences of the standard relations and the octahedral relation Eq.* (14.21) *(resp. the five octahedral relations Eq.* (14.52)*-Eq.* (14.56)*).*

Proof. We have dealt with the weight 3 case in Sec. 14.3. Weight 4 case is similar. See also [599, § 3]. □

From the available data in Table 14.1 we can formulate the following conjecture.

Conjecture 14.6.2. *Suppose the level $N = p$ is a prime ≥ 5. Then*

$$d(3, p) \leq \frac{p^3 + 4p^2 + 5p + 14}{12}.$$

Moreover, equality holds if standard relations produce all the linear relations.

We formulated this conjecture under the belief that the upper bound of $d(3, p)$ produced by the standard relations should be a polynomial of p of degree 3. Then we find the coefficients by the bounds of $d(3, p)$ for $p = 5, 7, 11, 13$ in Table 14.1.

When $w > 2$ it's not too hard to improve the bound of $d(w, p)$ given in [156, § 5.25] by the same idea as used in the proof of [156, § 5.24] (for example, decrease the bound by $(p^2 - 1)/24$). But they are often not the best. We conclude our paper with the following conjecture.

Conjecture 14.6.3. *If N is a standard level then the standard relations always provide the sharp bounds of $d(w, N)$, namely, all linear relations can be derived from the standard ones, if further $N > 3$ then the bound $D(w, N)$ in Eq.* (13.4) *by Deligne and Goncharov can be lowered. If N is a non-standard level then the bound $D(w, N)$ is sharp and there exists a positive integer $w_0(N)$ so that at least one non-standard relation exists in $\mathsf{CMZV}(w, N)$ for each $w \geq w_0(N)$.*

It is likely that one can take $w_0(4) = w_0(6) = w_0(9) = 3$ and $w_0(N) = 2$ for all the other non-standard levels N.

14.7 Historical notes

As we mentioned in the historical notes concerning MPLs, the Euler sums and MDVs attracted the attention of physicists in the early 1990s with the advent of high performance computers which enable them to compute many Feynman integrals numerically with high accuracy. D. J. Broadhurst and his collaborators first noticed the intimate relations between these objects, see for example, [86–89,93,94]. Many dimension conjectures in this chapter belong to them. Some of these have been partially proved by Deligne who showed in [155] that those conjectured dimensions at least provide the upper bounds.

The integral substitutions presented in Sec. 14.2.3 are adopted from [67] in which many different proof of the identities $\zeta(3) = \zeta(2,1)$ and $\zeta(3) = 8\zeta(\bar{2},1)$ can be found. One of these is the elegant proof given in Example 14.2.11.

Doubling relations at level 2 of Sec. 14.2.5 were first investigated in [59] where Blümlein *et al.* collected a lot of data and facts concerning MZVs and Euler sums.

The identity family of Sec. 14.2.5 were first conjectured around 1996 by Borwein, Bradley and Broadhurst [68]. Besides this one, they numerically discovered many other striking identities involving Euler sums. In the following years a large majority of them have been rigorously proven exception Eq. (14.13). This identity is remarkable in that it was the only conjectured family of identities relating alternating Euler sums to MZVs, neither of which have closed evaluations. Several proofs of the case $n = 1$ can be found in [67]. In an email dated Oct. 08, 2006, Prof. J. Borwein informed the author that the only known proof of Eq. (14.13) for $n = 2$ was by computer verification. Within a year, the general case was solved and the proof was given in [600]. Thm. 14.6.1 was proved in [599].

The generalized doubling relations at level 2 is explained in details in [59] where Conjecture 14.2.18 can be found on p. 596. Broadhurst pointed out some relations at level 2 in [88, (127)] should be related to the doubling relations.

Broadhurst formulated Conjecture 14.2.1 in [87], which provided a great insight into the intimate relations between MZVs and Euler sums. Note that Deligne proved the dimensions of ES_n are correct under the assumption of a variant Grothendieck's period conjecture (see [155]). However, his basis

(see Remark 13.3.41) is different from the one given in Conjecture 14.2.3.

The non-standard octahedral relations in level 4 first appeared in [599]. Okuda discovered some relations in [443, Example 4.2] similar to the octahedral relations.

It was Broadhurst who first noticed the close connection between Feymann integral and knot theory (see [88]. He and his collaborators carried out extensive computer aided computations in the late 1990s and discovered the detailed structures MCls and MGls in level 6 (see Refs. [68, 72]). Motivated by some ingenious computations by Panzer [455] and Schnetz [488], Broadhurst also considers the MCVs and posed many conjectures in [89]. Their results are contained in Sec. 14.4.

Exercises

14.1. Prove the conjectural expression (14.5) and (14.6) assuming Conjecture 14.2.1.

14.2. Prove that the only nontrivial changes of variable that preserve the interval $[0, 1]$ and send $\mathfrak{A}_1$ into $\mathfrak{A}_2$ are given by (14.7) to (14.10) and their compositions.

14.3. By induction and Prop. 13.3.42 reduce the proof of Thm.13.3.44 to the $k = 0$ case.

14.4. Verify the following values of the central binomial sums:

$$\begin{aligned}
S(2) &= \frac{\zeta(2)}{3} \\
S(3) &= -\frac{2\pi}{3}\,\mathrm{mcl}(2) - \frac{4}{3}\zeta(3) \\
S(4) &= \frac{17}{36}\zeta(4) \\
S(5) &= 2\pi\,\mathrm{mcl}(4) - \frac{19}{3}\zeta(5) + \frac{2}{3}\zeta(3)\zeta(2).
\end{aligned}$$

See p. 25 of [72] for more values of central binomial sums.

14.5. Show that for all positive integers n

$$\begin{aligned}
\mathrm{mcl}(2n+1) &= \frac{(-1)^n}{2}(1-2^{-2n})(1-3^{-2n})\zeta(2n+1), \\
\mathrm{mgl}(n) &= \frac{(-1)^{n+1}2^{n-1}\pi^n B_n(\frac{1}{6})}{n!}.
\end{aligned}$$

Here $B_n(x)$ is the nth Bernoulli polynomial defined by the generating series $te^{xt}/(e^t-1) = \sum_{n\geq 0} B_n(x)t^n/n!$.

14.6. Let m, n be two positive integers. Prove that

$$(-1)^{m+n}\zeta(m+1,\{1\}^{n-1}) = (-1)^n P_m P_n + \sum_{r=0}^{m-1} P_r \lambda(m+1-r,\{1\}^{n-1}) + \sum_{r=0}^{n-1}(-1)^r P_r \overline{\lambda(n+1-r,\{1\}^{m-1})}.$$

14.7. Use Thm. 14.4.25 to show that

$$\mathrm{mgl}(2,1) = \frac{\pi^3}{324}.$$

Then use the Exercise 14.6 to show

$$\mathrm{mgl}(3,1) = \frac{-23}{19440}\pi^4.$$

14.8. For Maple users, check the following five relations in weight 3 and level 4 are $\mathbb{Q}$-linearly independent non-standard relations:

$$\begin{aligned}
0 &= -255608l_1 - 265360l_2 - 219216l_3 - 19306179l_4 - 214008l_5 + 45560l_6 \\
&\quad - 148296l_7 - 1117280l_8 - 677152l_9 + 86512l_{10} - 239320l_{11} - 50032l_{12} \\
&\quad - 121008l_{13} - 96944l_{14} + 202328l_{15} - 1178499l_{16} + 98944l_{17} \\
&\quad + 1565754l_{18} + 23071580l_{19} + 363568l_{20} - 3310177l_{21}, \qquad (14.52)\\
0 &= 29752l_1 + 23312l_2 + 10960l_3 + 6123413l_4 + 16440l_5 - 12408l_6 \\
&\quad + 7144l_7 + 58272l_8 + 86976l_9 - 15952l_{10} + 41144l_{11} + 13552l_{12} \\
&\quad + 29552l_{13} + 9840l_{14} - 36696l_{15} + 375805l_{16} - 41760l_{17} \\
&\quad - 477366l_{18} - 7196900l_{19} - 62128l_{20} + 1048983l_{21}, \qquad (14.53)\\
0 &= 477444l_1 + 431352l_2 + 268168l_3 + 98404710l_4 + 308964l_5 - 233140l_6 \\
&\quad + 130028l_7 + 1563872l_8 + 1516032l_9 - 296664l_{10} + 702308l_{11} + 190136l_{12} \\
&\quad + 506440l_{13} + 141592l_{14} - 636468l_{15} + 6027441l_{16} - 701600l_{17} \\
&\quad - 7683609l_{18} - 115803282l_{19} - 1063768l_{20} + 16877562l_{21}, \qquad (14.54)\\
0 &= -5976l_1 + 1776l_2 + 8496l_3 - 2132671l_4 + 3176l_5 + 1752l_6 \\
&\quad + 3832l_7 + 50976l_8 - 2688l_9 + 2320l_{10} - 10264l_{11} - 5808l_{12} \\
&\quad - 6128l_{13} + 2320l_{14} + 8120l_{15} - 132307l_{16} + 13856l_{17} \\
&\quad + 162614l_{18} + 2487604l_{19} + 12720l_{20} - 368485l_{21}, \qquad (14.55)\\
0 &= -474064l_1 - 405248l_2 - 243520l_3 - 54556373l_4 - 283952l_5 + 84368l_6 \\
&\quad - 170640l_7 - 1033056l_8 - 994784l_9 + 174880l_{10} - 540432l_{11} - 156544l_{12} \\
&\quad - 240512l_{13} - 49344l_{14} + 411152l_{15} - 3357683l_{16} + 292256l_{17}
\end{aligned}$$

$$+ 4291792l_{18} + 64572648l_{19} + 743136l_{20} - 9470695l_{21}. \tag{14.56}$$

where by setting $L = L_4$, $1^4 = \{1\}^4$, ...

$$\begin{array}{lll}
l_1 = L(1^4|2,1,0,1), & l_2 = L(1^4|2,1^2,0), & l_3 = L(1^4|2,0,3,1),\\
l_4 = L(1^4|2,0^3), & l_5 = L(1^4|1,2,0,3), & l_6 = L(1^4|3^2,0,3),\\
l_7 = L(1^4|3,1,3,2), & l_8 = L(1^4|3,0^3), & l_9 = L(1^4|3,0,1,0),\\
l_{10} = L(1^4|3,0,1^2), & l_{11} = L(2,1^2|0,3,0), & l_{12} = L(3,1|0,3),\\
l_{13} = L(1^4|2,2,3,0), & l_{14} = L(2,1^2|3,1^2), & l_{15} = L(2,1^2|3,0,3),\\
l_{16} = L(1^2,2|2^3), & l_{17} = L(1^4|2,0,1,0), & l_{18} = L(2,1^2|2^2,0),\\
l_{19} = L(1^4|\{2,0\}^2), & l_{20} = L(2^2|3,0), & l_{21} = L(1^4|2^4).
\end{array}$$

Chapter 15

Application to Feynman integrals

In this last chapter we shift our attention to one of the most active research areas in mathematical physics nowadays: the study of Feynman integrals in the perturbative quantum field theory (QFT). For simplicity, we will only consider the four-dimensional case in Sections 15.1-15.4.

15.1 Why Feynman integrals?

Generally speaking, a Feynman diagram represents a perturbative contribution to the probability amplitude (given by the corresponding Feynman integral) of a quantum transition from some initial quantum state to some final quantum state. So at high energies this is a valuable tool to obtain predictions of all coupling constants from the QFT. For example, the diagram on the book cover represents an electron-positron collision producing a pair of muon and antimuon, or producing some hadrons. Then the Feynman integral can be used to compute the ratio

$$\frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}$$

of the total cross section (*i.e.*, probability) for the production of hadron to the total cross section for the production of $\mu^+\mu^-$ pair. In fact, the calculation using the diagram on the book cover (reproduced below) already gives pretty good numerical result fitting the experimental data. To reach more accuracy, one needs to consider Z boson and taking quarks (in other words, quantum chromodynamics or QCD) further into account by including not only gluon emission as shown by the diagrams in Fig. 15.1 but also gluon loop corrections which leads the Feynman diagrams with loops shown in Fig. 15.2. More gluons lead to more loops and difference in the order of emissions and absorptions of them corresponds to different configurations of the loops.

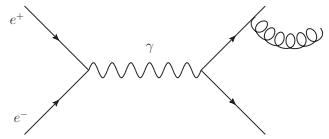


Figure 15.1: A Feynman diagram for electron-positron annihilation.

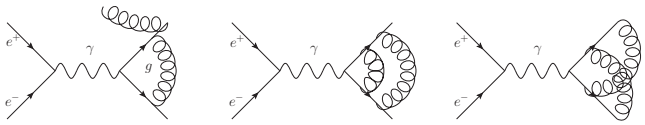


Figure 15.2: Several one- and two-loop Feynman diagrams contributing to the electron-positron annihilation process.

15.2 Feynman diagrams

We first recall some basic definitions in the graph theory.

Definition 15.2.1. A graph $G = (V_G, E_G)$ consists of the following data: V_G is the set of vertices, E_G is the set of edges whose two ends are vertices in V_G. We usually set $n_G := \sharp(E_G)$. The number of edges incident in a vertex v is called the degree (or valance) of v. A graph G is in ϕ^4 if the degree of every vertex of G is at most 4.

It is known that the most basic and important Feynman integrals are those associated with graphs in ϕ^4.

If we make G a directed graph and denote by $s(e)$ and $t(e)$ the source and the target of an edge $e \in E_G$ respectively, then the homology complex $H_\bullet(G, \mathbb{Z})$ of G with integral coefficients is defined by

$$0 \longrightarrow \mathbb{Z}[E_G] \xrightarrow{d} \mathbb{Z}[E_G] \longrightarrow 0$$

where $d(e) = t(e) - s(e)$ and $\mathbb{Z}[E_G]$ sits at degree 0. Let $h_0(G)$ be the number of connected components of G and $h_1(G)$ be the number of independent loops of G. It is straight-forward to check that (see Exercise 15.1)

$$H_0(G, \mathbb{Z}) \cong \mathbb{Z}^{h_0(G)}, \qquad H_1(G, \mathbb{Z}) \cong \mathbb{Z}^{h_1(G)}. \tag{15.1}$$

The loop number $h_G := h_1(G)$ is often called the *first Betti number* of G.

Definition 15.2.2. A *Feynman graph*

$$G = (V_G, E_G, E_G^{\text{ext}}, \{m_e\}_{e\in E_G}, \{q_i\}_{i\in E_G^{\text{ext}}})$$

consists of the following: V_G is the set of vertices, E_G is the set of labeled (internal) edges whose two ends are vertices in V_G, E_G^{ext} is the set of external half-edges (or legs), together with the following *kinematics data*:

(i) A particle mass $m_e \in \mathbb{R}$ for all $e \in E_G$,
(ii) An incoming momentum $q_i \in \mathbb{R}^D$ for all $i \in E_G^{\text{ext}}$ such that the *momentum conservation* holds:

$$\sum_{i\in E_G^{\text{ext}}} q_i = 0.$$

It is a convention to draw massive edges (i.e., $m_e \neq 0$) as thickened lines and draw a negative incoming momentum as a positive outgoing momentum. It is also helpful sometimes to think each internal edge as a union of two half-edges. For example, by closing the external legs of the two-point

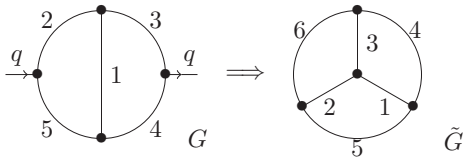


Figure 15.3: Left: 2-point 2-loop massless diagram. Right: wheel with three spokes.

(i.e., two incoming momenta) two-loop graph on the left of Fig. 15.3 we get a three-loop vacuum graph (i.e., no incoming momentum) on the right.

15.3 Feynman integrals in momentum space

We will omit the derivation and the subsequent transformation of the Feynman integrals from the Standard Model of the particle physics into their momentum space form and then to the multi-loop Euclidean parametric form (see [427] for more details).

Suppose a Feynman graph G has m external legs and $n := n_G$ internal edges. For each internal edge e_j the corresponding propagator in the integrand can be raised to an integer power a_j so that the integral also depends on the numbers a_1,...,a_n. The associated scalar (*i.e.*, no numerator in the

integrand) $h := h_G$-loop integral $I(G)$, called its *period*, is defined by

$$I(G) = \int_{(\mathbb{R}^D)^h} \prod_{r=1}^{h} \frac{\mathrm{d}^D k_r}{i\pi^{\frac{D}{2}}} \prod_{j=1}^{n} \frac{1}{(-p_j^2 + m_j^2)^{a_j}}. \tag{15.2}$$

Here $k_r = (k_{r,1}, \ldots, k_{r,D})$ is a D-dimensional vector variable with $\mathrm{d}^D k_r = \mathrm{d}k_{r,1} \ldots \mathrm{d}k_{r,D}$, and p^2 means the square of the length of p if p is a D-dimensional momentum vector. The independent (virtual) loop momenta are labeled $k_1, \ldots, k_h$. The momenta flowing through the propagators are then given as a linear combination of the external momenta p and the loop momenta k determined by momentum conservation at every vertex:

$$p_i = \sum_{j=1}^{h} c_{ij} k_j + \sum_{j=1}^{m} d_{ij} q_j, \quad c_{ij}, d_{ij} \in \{-1, 0, 1\}.$$

Notice also by this conservation the number of independent internal momenta must be equal to the loop number h.

15.3.1 *Divergence types*

Let us first consider the simple example of a scalar two-point one-loop integral with zero external momentum:

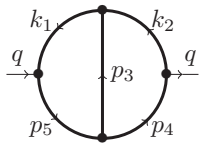
$$q = 0 \quad \text{—}\bullet\!\!\left(\begin{matrix} k \\ k \end{matrix}\right)\!\!\bullet\text{—} \quad = \frac{1}{(2\pi)^4} \int_{\mathbb{R}^4} \frac{\mathrm{d}^4 k}{(k^2)^2} = \frac{1}{8\pi^2} \int_0^\infty \frac{\mathrm{d}r}{r}$$

by using 4-dimensional spherical coordinates. Here $k = (k_1, \ldots, k_4)$ is a 4-dimensional vector, $\mathrm{d}^4 k = \mathrm{d}k_1 \cdots \mathrm{d}k_4$, and $k^2 = k_1^2 + \cdots + k_4^2$. This integral diverges when $r \to \infty$ as well as when $r \to 0$. The former divergence is called *ultraviolet (UV) divergence*, the later is called *infrared (IR) divergence*. Another physics term often appears in the literature: if the 4-momentum p and the mass m of a particle satisfies $p^2 = m^2$ then it is said to be *on shell*; otherwise it is said to be *off shell*.

One of the key ideas to handle UV divergence is to use the *dimension renormalization*, namely, one can regard $D = 4 - 2\varepsilon$ as a continuous real variable and then expand the integral as a Laurent series of ε. To handle IR divergence coming from massless situation one can take a small mass first and then take the limit when the mass goes to 0.

15.3.2 *A short example*

Consider the Feynman graph in Fig. 15.4. The massive Feynman integral



$$p_3 = k_1 - k_2,$$
$$p_4 = q + k_2,$$
$$p_5 = q + k_1.$$

Figure 15.4: The 2-point 2-loop massive diagram.

corresponding to it is:

$$\int\int \frac{\mathrm{d}^D k_1 \mathrm{d}^D k_2}{(m_1^2 - k_1^2)^{a_1}(m_2^2 - k_2^2)^{a_2}(m_3^2 - p_3^2)^{a_3}(m_4^2 - p_4^2)^{a_4}(m_5^2 - p_5^2)^{a_5}}, \quad (15.3)$$

where q is the incoming momentum entering from the left and $D = 4 - 2\varepsilon$ is the number of dimensions. If the graph is massless (*i.e.*, $m_j = 0$ for all $j = 1, \ldots, 5$) then it is proved [47] that MZVs are sufficient for the Laurent expansion of the integral in Eq. (15.3) if all powers of the propagators are of the form $a_j = n_j + c_j\varepsilon$ with $n_j \in \mathbb{N}$ and $c_j \in \mathbb{R}_{\geq 0}$.

15.3.3 *Massive and massless integrals*

When allowing the Feynman integrals to depend on kinematical invariants and particle masses, we know MPLs are not sufficient to express the result. Here the first cases show up already in the massive sunrise graphs $\ominus$ (see [55]). Because our main interest in this book is on MPLs and MZVs we will only consider massless Feynman integrals in the rest of this chapter if we do not point out otherwise.

15.4 Graphical functions in position space

In this section we consider loop integrals associated with ladder graphs and some generalizations. These integrals can be expressed using single-valued polylogs satisfying certain differential equations. Since these equations reduce the loop level by one, they can be solved iteratively in the loop order.

We begin with the box ladder graphs in Fig. 15.5. Let $p = p_1 + p_4 + \sum_{j=1}^{h} k_j$ and $y_j = k_1 + \cdots + k_j$ for $j = 1, \ldots, h$. Then the period of the h-loop box ladder with incoming momenta $p_1, \ldots, p_4$ is given by the period integral (up to some prefactor depending on p_j's)

$$I(\mathrm{B}_h(p_1, p_2, p_3, p_4)) = \frac{1}{\pi^{2h}} \int \frac{(p_1 + p_4)^2}{(y_h + p_1 + p_4)^2} \prod_{i=1}^{h} \frac{(p_3 + p_4)^2 \mathrm{d}^4 k_j}{k_j^2 (y_j + p_1)^2 (y_j - p_2)^2}.$$

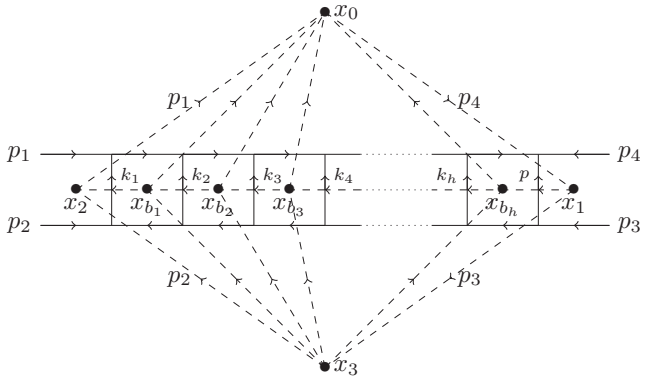


Figure 15.5: The box ladder graph and its dual graph.

Let $x_{ij} = x_j - x_i$ be the vector from x_i to x_j where x's are points in 4-dimensional space such that $p_1 = x_{20}$, $p_2 = x_{32}$, $p_3 = x_{13}$, $p_4 = x_{01}$ (note that $\sum p_j = 0$). Then by the change of variables $k_1 = x_{b_1,2}$ and $k_j = x_{b_j,b_{j-1}}$ for $j = 2, \dots, h$ we get

$$I(\mathrm{B}_h(p_1,p_2,p_3,p_4)) = \frac{x_{12}^2}{\pi^{2h}} \int \frac{1}{x_{b_1,2}^2} \prod_{j=1}^{h-1} \left(\frac{x_{03}^2 \cdot \mathrm{d}^4 x_{b_j}}{x_{b_j,b_{j+1}}^2 x_{b_j,0}^2 x_{b_j,3}^2} \right) \frac{x_{03}^2 \cdot \mathrm{d}^4 x_{b_h}}{x_{b_h,0}^2 x_{b_h,1}^2 x_{b_h,3}^2}.$$

Since the integral is conformally covariant we can take $x_3 \to \infty$ and find that there is some function $\mathcal{B}^{(h)}$ such that (see Eq. (15.9))

$$\begin{aligned} I(\mathrm{B}_h(p_1,p_2,p_3,p_4)) &= \frac{x_{12}^2}{\pi^{2h}} \int \frac{1}{x_{b_1,2}^2} \prod_{i=1}^{h-1} \left(\frac{\mathrm{d}^4 x_{b_i}}{x_{b_i,b_{i+1}}^2 x_{b_i,0}^2} \right) \frac{\mathrm{d}^4 x_{b_h}}{x_{b_h,1}^2 x_{b_h,0}^2} \\ &= \mathcal{B}^{(h)} \left(\frac{x_{02}^2}{x_{12}^2}, \frac{x_{01}^2}{x_{12}^2} \right). \end{aligned}$$

15.4.1 *Sequential graphs and graphical functions*

We now generalize the (planar dual of the) box ladder graphs to the following setting.

Definition 15.4.1. Let $\mathbf{w}$ be a word in $\{0,1,2\}^*$. We denote by $\lg(\mathbf{w})$ the number of letters in $\mathbf{w}$. The *sequential graph* $G_{\mathbf{w}}$ is the graph with two distinguished vertices 0 and 1 and a horizontal chain of vertices that connect to either 0 or 1, or both. We can obtain the word $\mathbf{w}$ by recording

the connections for each vertex from left to right sequentially: 0 (resp. 1) for a single connection to 0 (resp. 1), or 2 for a double connection to both 0 and 1.

Example 15.4.2. The left graph in Fig. 15.6 is the sequential graph $G_{20111202}$ if z and its edge are both removed. The right one corresponds to the dual graph of the zig-zag graph Z_5 if $z \to 0$[1].

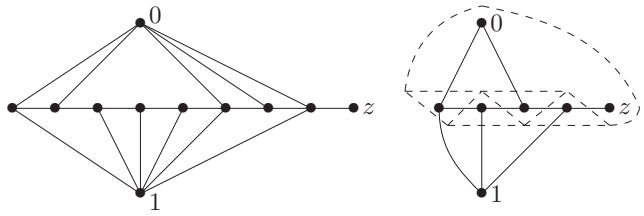


Figure 15.6: Left: the sequential graph $G_{20111202}$. Right: the zig-zag Z_5 (dashed graph) and its dual G_{2101} (solid graph).

Given a word $\mathbf{w} = a_1 \dots a_{h-1}$ with letters $a_i \in \{0, 1\}$, we can define a generalized h-loop (planar dual of) the three-point ladder integral by

$$I_{\mathbf{w}}(x_0, x_1, x_2) = \frac{1}{\pi^{2h}} \int \frac{1}{x_{b_1,2}^2} \prod_{i=1}^{h-1} \left(\frac{\mathrm{d}^4 x_{b_i}}{x_{b_i,b_{i+1}}^2 x_{b_i,a_i}^2} \right) \frac{\mathrm{d}^4 x_{b_h}}{x_{b_h,0}^2 x_{b_h,1}^2} \,. \tag{15.4}$$

Here the integration vertices are labeled $x_{b_1}, \dots, x_{b_h}$ while x_{ij}^2 denotes the square distance $(x_i - x_j)^2$ in 4-dimensional Euclidean space. We now show that these integrals can be expressed in terms of classical polylogs. In particular we will derive the formula for $\mathcal{B}^{(h)}$ in terms the single-valued polylogs.

Definition 15.4.3. Let vertex 0 correspond to the origin in $\mathbb{R}^4$ and $1 = (1, 0, 0, 0) \in \mathbb{R}^4$. Let $z \neq 0, 1$ be a complex variable embedded in $\mathbb{R}^4$ by $(\mathrm{Re}(z), \mathrm{Im}(z), 0, 0)$. Given a graph G with three distinguished vertices $0, 1, z$ with position space Feynman rules, we define its *graphical function*

$$f_G(z) = \left(\prod_{v \in V_G \setminus \{0,1,z\}} \int_{\mathbb{R}^4} \frac{\mathrm{d}^4 x_v}{\pi^2} \right) \frac{1}{\prod_{e \in E_G} Q_e} \tag{15.5}$$

[1]Strictly speaking, in the dual graph 0 and 1 should be joint by a solid curve. But this doesn't change its Feynman integral.

where $Q_e = \|x-y\|^2$ for every edge e connecting the two positions x and y in $\mathbb{R}^4$.

Remark 15.4.4. By definition, $\|z-1\|^2$ in $\mathbb{R}^4$ is equal to $(z-1)(\bar{z}-1) \in \mathbb{R}$. Similarly, $\|z\|^2 = z\bar{z}$.

Example 15.4.5. Let I be the graph $\underset{0}{\bullet}\!\!-\!\!\underset{z}{\bullet}\!\!-\!\!\underset{1}{\bullet}$. Then we have.

$$f_{\mathrm{I}}(z) = \frac{1}{Q_{0,z}Q_{1,z}} = \frac{1}{z\bar{z}(z-1)(\bar{z}-1)}. \tag{15.6}$$

In general, we can compute the period integrals for the box ladder graphs B_n, the wheel graphs WS_n (see Thm. 15.4.19), and the zig-zag graphs Z_n (see Thm. 15.4.20), by using some sequential graphical functions:

$$\begin{aligned} I(\mathrm{B}_n) &= f_{2\{0\}^{n-1}}(1), \\ I(\mathrm{Z}_{2n+3}) &= f_{2\{01\}^n1}(0) = f_{2\{01\}^n0}(1) = f_{2\{10\}^n1}(0) = f_{2\{10\}^n0}(1), \quad (15.7) \\ I(\mathrm{Z}_{2n+4}) &= f_{2\{01\}^{n+1}}(0) = f_{2\{01\}^n00}(1) = f_{2\{10\}^n11}(0) = f_{2\{10\}^{n+1}}(1). \quad (15.8) \end{aligned}$$

In particular, by the right picture in Fig. 15.6 we obtain $I(\mathrm{Z}_5) = f_{2101}(0)$.

Example 15.4.6. The dual of Fig. 15.5 for the box ladder graph is the sequential graph $G_{2\{0\}^h}$ if we set $x_0 = 0$, $x_2 = 1$ and $x_1 = z$. Thus $p_4^2 = x_{01}^2 = \|z\|^2$, $p_1^2 = x_{02}^2 = 1$, and $(p_1+p_4)^2 = (p_2+p_3)^2 = x_{12}^2 = \|1-z\|^2$.

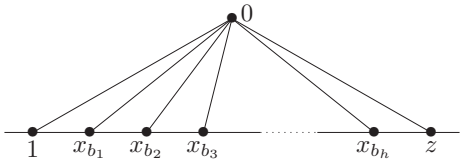


Figure 15.7: The dual graph of the box ladder graph as a sequential graph.

So we have

$$I(\mathrm{B}_h(p_1,p_2,p_3,p_4)) = \mathcal{B}^{(h)}\left(\frac{1}{\|z\|^2}, \frac{\|1-z\|^2}{\|z\|^2}\right) = f_{2\{0\}^{h-1}}(z).$$

Therefore period of the box ladder graph is given by

$$I(\mathrm{B}_h) = f_{2\{0\}^{h-1}}(1). \tag{15.9}$$

15.4.2 *Properties of graphical functions*

Recall that a complex variable function $f(z)$ is called *real analytic* in an open set $D \subseteq \mathbb{C}$ if $f(z)$ is real valued and is infinitely differentiable with respect to both ∂_z and $\partial_{\bar{z}}$.

Definition 15.4.7. A single-valued function $f : \mathbb{C} \setminus \{0,1\} \to \mathbb{R}$ is in $\mathsf{C}_{1,1}$ if for all $\epsilon > 0$

(i) $f_{G^\sharp}(z) = o(|z|^{-1-\epsilon})$ for $z \to 0$, *i.e.*, $\lim_{z\to 0} f_{G^\sharp}(z)|z|^{1+\epsilon} = 0$, and
(ii) $f_{G^\sharp}(z) = o(|z|^{-1+\epsilon})$ for $z \to \infty$, *i.e.*, $\lim_{z\to \infty} f_{G^\sharp}(z)|z|^{1-\epsilon} = 0$.

Proposition 15.4.8. *Let G be a graph with three distinguished vertices $0, 1, z$. Then its graphical function $f_G(z) : \mathbb{C} \setminus \{0,1\} \to \mathbb{R}$ satisfies*

(i) $f_G(z) = f_G(\bar{z})$,
(ii) f_G *is a single-valued,*
(iii) $f_{\mathbf{w}} \in \mathsf{C}_{1,1}$ *for every sequential graph* $G_{\mathbf{w}}$,
(iv) $f_{\mathbf{w}}$ *is real analytic in* $\mathbb{C} \setminus \{0,1\}$ *for every sequential graph* $G_{\mathbf{w}}$.

Proof. (i) follows from the change of coordinates $x_v = (x_v^1, x_v^2, x_v^3, x_v^4) \mapsto (x_v^1, -x_v^2, x_v^3, x_v^4)$.

(ii) An infinitesimal ε-deformation of the denominator factors avoiding all singularities shows that the monodromy for f_G is always trivial. So (ii) follows by taking the limit $\varepsilon \to 0$.

The proofs of (iii) and (iv) are long and very technical so we refer the interested reader to Schnetz' original paper where more general results can be found in [487, Prop. 3.14 and Prop. 3.15]. □

Definition 15.4.9. Let G be a graph with three marked vertices $0, 1, z$. We can produce a new graph $G^\sharp$ by appending an edge to the vertex z and relabel the new vertex by z. See Fig. 15.8.

Figure 15.8: Graph $G^\sharp$ is obtained from G by appending an edge to z.

In what follows we will assume that both G and $G^\sharp$ have well-defined graphical functions.

Proposition 15.4.10. *The following identity holds*

$$\mathfrak{D}_{z,\bar{z}} f_{G^\sharp}(z) := \frac{1}{z-\bar{z}}\,\partial_z\partial_{\bar{z}}\,(z-\bar{z}) f_{G^\sharp}(z) = -f_G(z). \tag{15.10}$$

Proof. By using D-dimensional spherical coordinates $(r, \phi^1, \ldots, \phi^{D-1})$ we see that, as a function in D-dimensional space

$$\begin{aligned} h(z) &:= \Delta_z^{(D)} \int \frac{F(y)}{\|y-z\|^2} \mathrm{d}^D y \\ &= 2(4-D) \int \frac{F(y)}{\|y-z\|^4} \mathrm{d}^D y = 2(4-D)(F * g)(z) \end{aligned} \tag{15.11}$$

where $*$ stands for the convolution and $g(x) = \|x\|^{-4}$ whose Fourier transform is given by

$$\hat{g}(\xi) = \int \frac{e^{2\pi i x\cdot\xi - \epsilon\|x\|}}{\|x\|^4} \mathrm{d}^D x = \frac{\pi^{4-D/2}}{\|\xi\|^{D-4}} \Gamma(D/2-2).$$

Thus by dimension regularization $D = 4-2\varepsilon$ we get from Eq. (15.11)

$$\begin{aligned} \hat{h}(\xi) = \hat{F}(\xi)\hat{g}(\xi) &= 2(4-D)\frac{\pi^{4-D/2}}{\|\xi\|^{D-4}} \Gamma(D/2-2)\hat{F}(\xi) \\ &= 4\varepsilon\pi^{2+\varepsilon}\|\xi\|^{2\varepsilon}\Gamma(-\varepsilon)\hat{F}(\xi) = -4\pi^{2+\varepsilon}\|\xi\|^{2\varepsilon}\Gamma(1-\varepsilon)\hat{F}(\xi). \end{aligned}$$

Taking $\varepsilon \to 0$ we obtain

$$\Delta^{(4)} \int \frac{F(y)}{\|y-z\|^2} \mathrm{d}^4 y = -4\pi^2 F(z).$$

Thus, for $F(y) = f_G(y)$, we have

$$\Delta^{(4)} f_{G^\sharp}(z) = -4 f_G(z). \tag{15.12}$$

Let $e_1 = (1,0,0,0)$ in $f_{G^\sharp}$ then $f_{G^\sharp}(z)$ depends only on r and ϕ^1 in spherical coordinates. Hence the Laplacian becomes (see Exercise 15.4)

$$\Delta^{(4)} f_{G^\sharp}(r,\phi^1) = \left(\partial_r^2 + \frac{3}{r}\partial_r + \frac{1}{r^2}\partial_{\phi^1}^2 + \frac{2\cos\phi^1}{r^2\sin\phi^1}\partial_{\phi^1}\right) f_{G^\sharp}(r,\phi^1)$$

which coincides with the following operator in 2-dimensional polar coordinates (r,ϕ^1):

$$\frac{1}{r\sin\phi^1}\Delta_z^{(2)} \circ M(r\sin\phi^1) = \partial_r^2 + \frac{3}{r}\partial_r + \frac{1}{r^2}\partial_{\phi^1}^2 + \frac{2\cos\phi^1}{r^2\sin\phi^1}\partial_{\phi^1},$$

where $M(r\sin\phi^1)$ just means multiplication by $r\sin\phi^1$ (see Exercise 15.5). By transition to $\mathbb{C}$ the proposition follows immediately from Eq. (15.12) since $\Delta^{(2)} = 4\partial_z\partial_{\bar{z}}$ and $z-\bar{z} = 2ir\sin\phi^1$. □

Recall that the important properties of SVMPLs are given by Thm. 4.3.2. Although they are not defined at 0, 1 and ∞ we can define their regularized values at these points. In particular, the SVMZVs are just the regularized values of SVMPLs at 1 (see Definition 4.4.3). Also recall from Sec. 4.4 we have defined the sets $\mathsf{O} = \mathbb{Q}\left[z, \frac{1}{z}, \frac{1}{z-1}\right]$, $\overline{\mathsf{O}} = \mathbb{Q}\left[\bar{z}, \frac{1}{\bar{z}}, \frac{1}{\bar{z}-1}\right]$. and the two bi-differential $(\partial_z, \partial_{\bar{z}})$-algebras $\mathsf{A} := \mathsf{O}\overline{\mathsf{O}}\mathsf{Ls}$ and $\mathsf{A}^{\mathrm{sv}} := \mathsf{O}\overline{\mathsf{O}}\mathsf{Ls}^{\mathrm{sv}}$, where Ls (resp. $\mathsf{Ls}^{\mathrm{sv}}$) is the shuffle-algebra spanned by SVMPLs over $\mathbb{C}$ (resp. over $\mathsf{MZ}_{\mathbb{Z}}^{\mathrm{sv}}$, the set of SVMZVs).

Definition 15.4.11. Let

$$\mathsf{B} := \left\{ \frac{f(z)}{z-\bar{z}} : f \in \mathsf{A}, f(z) = -f(\bar{z}) \right\}. \tag{15.13}$$

Similarly, $\mathsf{B}^0, \mathsf{B}^{0,\mathrm{sv}}$ are defined with $f \in \mathsf{Ls}, f \in \mathsf{Ls}^{\mathrm{sv}}$ in Eq. (15.13), respectively. Let $\mathsf{B}_n^\bullet$ be the (total) weight n piece of $\mathsf{B}^\bullet$.

The differential operator in equation (15.10) maps graphical functions in B into B. By integration in A the differential operator can be inverted in B, namely, we can produce $f_{G^\sharp}$ from f_G.

Theorem 15.4.12. *Assume the graphical function $f_{G^\sharp}(z)$ is real analytic in $\mathbb{C} \setminus \{0,1\}$ and for $z \in \mathbb{R}^4$, $f_{G^\sharp}(z)$, $f_{G^\sharp}(1-z) \in \mathsf{C}_{1,1}$. Moreover, let*

$$f_G(z) = \sum_{a,b\in\{0,1\}} \frac{g_{a,b}(z)}{(z-a)(\bar{z}-b)(z-\bar{z})} \in \mathsf{B} \tag{15.14}$$

for some SVMPL $g_{a,b}$ in Ls. Then

$$f_{G^\sharp}(z) = \mathfrak{I}_{z,\bar{z}} f_G(z) := \frac{-1}{2(z-\bar{z})} \left(\int_0 \mathrm{d}z \int_0 \mathrm{d}\bar{z} + \int_0 \mathrm{d}\bar{z} \int_0 \mathrm{d}z \right) (z-\bar{z}) f_G(z). \tag{15.15}$$

The function $f_{G^\sharp}$ is the unique solution of Eq. (15.10) in B^0.

Proof. It is easy to verify that the right-hand side of Eq. (15.15), denoted by $F(z)$, satisfies Eq. (15.10). The kernel of $\partial_z\partial_{\bar{z}}$ is a sum of a holomorphic h and an anti-holomorphic $\overline{h}$ (*i.e.*, $\partial_{\bar{z}} h = \partial_z \overline{h} = 0$). By Prop. 15.4.8(i) both f_G and $f_{G^\sharp}$ are invariant under $z \leftrightarrow \bar{z}$ so that

$$(z-\bar{z})\big(f_{G^\sharp}(z) - F(z)\big) = h(z) - h(\bar{z}).$$

Since $f_{G^\sharp}$ and F are real analytic on $\mathbb{C} \setminus \{0,1\}$ we get

$$h'(z) = \partial_z (z-\bar{z})(f_{G^\sharp}(z) - F(z)).$$

Now, h' is a single-valued real analytic function on $\mathbb{C}\setminus\{0,1\}$ with $\partial_{\bar{z}}h'=0$. Hence h' is holomorphic on $\mathbb{C}\setminus\{0,1\}$. We now claim that

$$h'(z)=c/z+d/(z-1)\quad\text{for some constants } c,d\in\mathbb{C}. \tag{15.16}$$

Indeed, by assumption $f_{G^\sharp}(z)\in\mathsf{C}_{1,1}$ we have $f_{G^\sharp}(z)=o(|z|^{-1-\epsilon})$ for $z\to 0$ so that $\partial_z(z-\bar{z})f_{G^\sharp}(z)=o(|z|^{-1-\epsilon})$. By Eq. (15.14), Prop. 4.4.9 and Prop. 4.4.12 we have $(z-\bar{z})F(z)\in\mathsf{Ls}$. Hence $(z-\bar{z})F(z)=o(|z|^{-\epsilon})$ by Exercise 4.22 and $\partial_z(z-\bar{z})F(z)=o(|z|^{-1-\epsilon})$. Thus $h'(z)=o(|z|^{-1-\epsilon})$ and $z^2h'(z)=o(|z|^{1-\epsilon})$. In particular $z^2h'(z)$ is continuous at $z=0$ and by Riemann's Theorem on Removable Singularities $z^2h'(z)$ is holomorphic at 0. It vanishes at 0 and therefore $zh'(z)$ is holomorphic at 0.

For $z=1$ the above argument can be applied to $f_{G^\sharp}(1-z)$ and $f_G(1-z)$ so that we get that $(z-1)h'(z)$ is holomorphic at 1. Hence

$$h'(z)=c/z+d/(z-1)+H(z)$$

for some $H(z)$ holomorphic on all of $\mathbb{C}$. To show $H(z)=0$, notice $(z-\bar{z})F(z)\in\mathsf{Ls}$ we get $\partial_z(z-\bar{z})F(z)=o(|z|^{-1+\epsilon})$ as $z\to\infty$ by Exercise 4.23. By assumption $f_{G^\sharp}(z)\in\mathsf{C}_{1,1}$ so $\partial_z(z-\bar{z})f_{G^\sharp}(z)=o(|z|^{-1+\epsilon})$. Thus $\lim_{z\to\infty}h'(z)|z|^{1-\epsilon}=0$. But this implies $\lim_{z\to\infty}H(z)=0$ so $H(z)$ is a bounded holomorphic function on $\mathbb{C}$. Therefore $H(z)$ is a constant equal to $\lim_{z\to\infty}H(z)=0$. The Claim Eq. (15.16) is now proved. Consequently,

$$f_{G^\sharp}(z)-F(z)=\frac{c(\log(z)-\log(\bar{z}))+d(\log(z-1)-\log(\bar{z}-1))}{z-\bar{z}}.$$

But $f_{G^\sharp}(z)-F(z)$ is single-valued so we must have $c=d=0$, namely, $f_{G^\sharp}=F$. Further, $f_{G^\sharp}\in\mathsf{B}^0$ follows from Eqs. (15.14), (15.15), Props. 4.4.9 and 4.4.12.

To prove the uniqueness in the last sentence, assume $f\in\mathsf{B}^0$ is another solution of Eq. (15.10). Since the kernel of $\partial_z\partial_{\bar{z}}$ in Ls is the constant function, $f_{G^\sharp}(z)=f(z)+c/(z-\bar{z})$. So $c=0$ by the invariance of $f_{G^\sharp}$ and f under $z\leftrightarrow\bar{z}$.

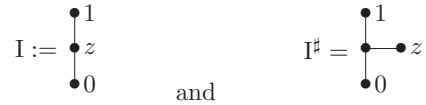
We now have completed the proof of the theorem. □

Remark 15.4.13. It is clear that the differential operator $\mathfrak{D}_{z,\bar{z}}$ in Eq. (15.10) can be rewritten as

$$\mathfrak{D}_{z,\bar{z}}:=\frac{1}{z-\bar{z}}\partial_z\partial_{\bar{z}}(z-\bar{z})=\partial_z\partial_{\bar{z}}-\frac{\partial_z-\partial_{\bar{z}}}{z-\bar{z}}.$$

This is sometimes useful in the computation because by the theorem if f_G is known and we can find a solution in B^0 of Eq. (15.10) (by whatever method) then it must be the unique solution of Eq. (15.15) as well.

Example 15.4.14. Consider the following graphs:



By Example 15.4.5 we have $f_{\mathrm{I}}(z) = 1/\big(z\bar{z}(z-1)(\bar{z}-1)\big)$. Thus $f_{\mathrm{I}^\sharp} = f_2$ is the sequential function of the word 2. We can check trivially that

$$\mathfrak{D}_{z,\bar{z}}\left(\frac{4i\,\mathscr{L}i_2(z)}{z-\bar{z}}\right) = -\frac{1}{z-\bar{z}}\left(\frac{1}{\bar{z}(z-1)} - \frac{1}{z(\bar{z}-1)}\right) = f_{\mathrm{I}}(z)$$

where $\mathscr{L}i_2(z)$ is the single-valued Bloch-Wigner dilog function defined by Eq. (2.46). So we get

$$f_2(z) = \frac{4i\,\mathscr{L}i_2(z)}{z-\bar{z}} \in \mathsf{B}^0 \tag{15.17}$$

by Remark 15.4.13. In fact, it is not difficulty to use the theorem to integrate and obtain directly that $f_2(z)$ is equal to

$$\frac{\mathrm{L}_{10}(z) + \mathrm{L}_1(z)\,\mathrm{L}_0(\bar{z}) + \mathrm{L}_{01}(\bar{z}) - \mathrm{L}_{10}(\bar{z}) - \mathrm{L}_1(\bar{z})\,\mathrm{L}_0(z) - \mathrm{L}_{01}(z)}{z-\bar{z}} \tag{15.18}$$

which, by the identities in Example 4.2.10, also implies Eq. (15.17).

Lemma 15.4.15. *Let* $\mathbf{w}$ *be a word in* $0,1,2$ *that begins with* 2 *then* $f_{\mathbf{w}}(e_1 - z)$ *is the sequential function of the word with* 0*'s and* 1*'s flipped.*

Proof. Note that for the graphical function defined on $\mathbb{R}^4$ we interpret $1-z$ as $e_1 - z$ for the unit vector $e_1 \in \mathbb{R}^4$. Denote by $\nu_{x,y}$ the number of edges connecting two vertices x and y. Assume

$$f_{\mathbf{w}}(z) = f_{0,1,z,\infty} = \int \frac{f^0_{0,1,z,\infty}}{\|z\|^{2\nu_{0,z}}\|e_1 - z\|^{2\nu_{1,z}}} \prod_{a=0,1} \prod_{\substack{x:\text{internal}\\ \nu(x,a)=1}} \frac{\mathrm{d}^4x}{\|x-a\|^2}, \tag{15.19}$$

where f^0 only involves propagators between internal vertices (i.e., those $\neq 0,1,z$). A change of variables $x_i \mapsto 1 - x_i$ leaves the integration measure invariant (since we are in even dimensions) and also keeps all the propagators between internal vertices intact so we get $f^0_{0,1,1-z,\infty} = f^0_{1,0,z,\infty}$. A propagator from an internal vertex x_i to an external vertex a is mapped to a propagator from x_i to $1-a$. Thus we have

$$\begin{aligned} f_{\mathbf{w}}(e_1 - z) &= f_{0,1,1-z,\infty} \\ &= \int \frac{f^0_{1,0,z,\infty}}{\|e_1 - z\|^{2\nu_{0,z}}\| - z\|^{2\nu_{1,z}}} \prod_{a=0,1} \prod_{\substack{x:\text{internal}\\ \nu(x,1-a)=1}} \frac{\mathrm{d}^4x}{\|x-a\|^2} = f_{\tau(\mathbf{w})}(z) \end{aligned}$$

where τ exchanges 0 and 1, as desired. □

Proposition 15.4.16. *In four dimensions, the sequential functions* $f_{\mathbf{w}}$*, where* $\mathbf{w} \in \{0,1,2\}^*$ *begin* 2*, provide the unique solutions in* B^0 *of*

$$\mathfrak{D}_{z,\bar{z}} f_{\mathbf{w}a}(z) = \begin{cases} \dfrac{-1}{(z-a)(\bar{z}-a)} f_{\mathbf{w}}(z), & \text{if } a = 0,1; \\ \dfrac{-1}{z\bar{z}(z-1)(\bar{z}-1)} f_{\mathbf{w}}(z), & \text{if } a = 2 \end{cases} \tag{15.20}$$

with the initial condition given by Eq. (15.17)*.*

Proof. By Lemma 15.4.15, Prop. 15.4.8(iii) and (iv) both $f_{\mathbf{w}}$ and $f_{\mathbf{w}}(e_1 - z)$ are real analytic on $\mathbb{C} \setminus \{0,1\}$ and in $\mathsf{C}_{1,1}$. After adding an edge from a to z we get a graph G with $f_G(z) = f_{\mathbf{w}}/\|z-a\|^2$ (see the first two pictures in Fig. 15.9). If adding both edges from 0 to z and from 1 to z we would get a graph G with $f_G(z) = f_{\mathbf{w}}/(\|z\|\|z-1\|)^2$. On the other hand, by the definition of the sequential graphs, appending an edge to z to $G_{\mathbf{w}}$ we immediately get the graph $G_{\mathbf{w}a}$ (see the last picture in Fig. 15.9). Equation (15.14) is obvious for $a = 0,1$. For $a = 2$, Eq. (15.14) follows

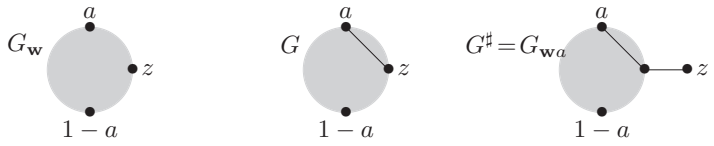


Figure 15.9: Adding edge a to z and applying $\sharp$ to $G_{\mathbf{w}}$ produces $G_{\mathbf{w}a}$.

from $1/(z(z-1)) = 1/(z-1) - 1/z$ and $1/(\bar{z}(\bar{z}-1)) = 1/(\bar{z}-1) - 1/\bar{z}$. The proposition now follows from Thm. 15.4.12. □

The solution of the recursive differential equation (15.20) can be given explicitly.

Proposition 15.4.17. *Let* $\mathbf{w} \in \{0,1,2\}^*$ *that begins with* 2*. Then the sequential function* $f_{\mathbf{w}}$ *is recursively given by Eq.* (15.17) *and*

$$f_{\mathbf{w}a}(z) = \begin{cases} \mathfrak{I}_{z,\bar{z}} \left(\dfrac{(z-\bar{z}) f_{\mathbf{w}}(z)}{(z-a)(\bar{z}-a)} \right), & \text{if } a = 0,1; \\ \mathfrak{I}_{z,\bar{z}} \left(\dfrac{(z-\bar{z}) f_{\mathbf{w}}(z)}{z\bar{z}(z-1)(\bar{z}-1)} \right), & \text{if } a = 2. \end{cases} \tag{15.21}$$

Moreover, $f_{\mathbf{w}} \in \mathsf{B}^{0,\mathrm{sv}}_{2\lg(\mathbf{w})}$.

Proof. Equation (15.21) is clear from Thm. 15.4.12 and Prop. 15.4.16. The nontrivial task is to prove that the coefficients of $f_{\mathbf{w}}$ are in $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$. We proceed by induction on $\lg(\mathbf{w})$. The length 1 case $\mathbf{w} = 2$ is Eq. (15.17). Assuming the proposition holds for $\mathbf{w}$ we now prove the case for $\mathbf{w}a$ where $a \in \{0,1,2\}$.

Case (i) $a \in \{0,1\}$. Assume $b \in \{0,1\}$. Then

$$Q_b(z) := \int_0 \frac{(z-\bar{z})f_{\mathbf{w}}(z)\mathrm{d}\bar{z}}{\bar{z}-b} \in \mathsf{Ls}^{\mathrm{sv}}_{2\lg(\mathbf{w})+1}$$

by Prop. 4.4.12. Thus

$$R_{a,b}(z) := \int_0 \frac{Q_b(z)\mathrm{d}z}{z-a} \in \mathsf{Ls}^{0,\mathrm{sv}}_{2\lg(\mathbf{w})+2} \tag{15.22}$$

by Prop. 4.4.9. For $\int_0 \mathrm{d}\bar{z}\int_0 \mathrm{d}z$, since

$$\overline{Q_a(z)} = \int_0 \frac{(\bar{z}-z)f_{\mathbf{w}}(\bar{z})\mathrm{d}z}{z-a} = -\int_0 \frac{(z-\bar{z})f_{\mathbf{w}}(z)\mathrm{d}z}{z-a}$$

we get by setting $P(z) = (z-\bar{z})f_{\mathbf{w}}(z)$ in Exercise 4.24

$$\int_0 \mathrm{d}\bar{z}\int_0 \mathrm{d}z \left(\frac{(z-\bar{z})f_{\mathbf{w}}(z)}{(z-a)(\bar{z}-b)}\right) = R_{a,b}(z) - (Q_b(a) + \overline{Q_a(b)})\,\mathcal{L}_1(z) \in \mathsf{Ls}^{0,\mathrm{sv}}_{2\lg(\mathbf{w})+2}$$

since $\overline{Q_a(b)} = Q_a(b) \in \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z},2\lg(\mathbf{w})+1}$ for $a,b \in \{0,1\}$. Combining with Eq. (15.22) we have

$$\mathfrak{I}_{z,\bar{z}}\left(\frac{(z-\bar{z})f_{\mathbf{w}}(z)}{(z-a)(\bar{z}-b)}\right) = \frac{-1}{z-\bar{z}}\left(R_{a,b}(z) - \frac{1}{2}(Q_b(a)+Q_a(b))\,\mathcal{L}_1(z)\right). \tag{15.23}$$

Consequently when $a = b \in \{0,1\}$ we get $f_{\mathbf{w}a} \in \mathsf{B}^{0,\mathrm{sv}}_{2\lg(\mathbf{w})+2}$.

Case (ii) $a = 2$. By partial fractions we get

$$\frac{1}{z\bar{z}(z-1)(\bar{z}-1)} = \left(\frac{1}{z-1}-\frac{1}{z}\right)\left(\frac{1}{\bar{z}-1}-\frac{1}{\bar{z}}\right) = \sum_{a,b=0}^{1}\frac{(-1)^{a+b}}{(z-a)(\bar{z}-b)}.$$

Then Eq. (15.23) yields

$$f_{\mathbf{w}2} = \frac{-1}{z-\bar{z}}\sum_{a,b=0}^{1}(-1)^{a+b}(R_{a,b}(z) - Q_a(b)\,\mathcal{L}_1(z)) \in \mathsf{B}^{0,\mathrm{sv}}_{2\lg(\mathbf{w})+2},$$

as desired. □

Corollary 15.4.18. *For every word* $\mathbf{w} \in 2\{0,1,2\}^*2$*, the graph period* $I(G_{\mathbf{w}})$ *of the sequential graph* $G_{\mathbf{w}}$ *lies in* $\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z},2|\mathbf{w}|-1}$.

Proof. Assume $\mathbf{w} = \mathbf{u}2$. Then we have $I(G_{\mathbf{u}2}) = f_{\mathbf{u}1}(0)$ with $f_{\mathbf{u}1}(z) = g(z)/(z-\bar{z})$ and $g \in \mathsf{Ls}^{\mathrm{sv}}_{2\,\mathrm{lg}(\mathbf{w})}$ by Prop. 15.4.17. But by setting $\mathbf{w} = \mathbf{u}$ in the proof of Prop. 15.4.17 we get

$$Q_1(z) := \int_0 \frac{(z-\bar{z})f_{\mathbf{u}}(z)\mathrm{d}\bar{z}}{\bar{z}-1}, \quad R_{1,1}(z) := \int_0 \frac{Q_1(z)\mathrm{d}z}{z-1}$$

and from Eq. (15.23)

$$g(z) = (z-\bar{z})f_{\mathbf{u}1}(z) = Q_1(1)\mathscr{L}_1(z) - \int_0 \frac{Q_1(z)\mathrm{d}z}{z-1}.$$

By Example 4.3.3 we know $\mathscr{L}_1(z) = \log(1-z) + \log(1-\bar{z})$. Thus

$$\partial_z g(z) = \frac{Q_1(1)}{z-1} - \frac{Q_1(z)}{z-1}.$$

From Exercise 15.6 we see that $I(G_{\mathbf{u}2}) = \frac{1}{2}(\partial_z g(0) - \partial_{\bar{z}} g(0)) = Q_1(0) - Q_1(1) \in \mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$ since $g(z) = -g(\bar{z})$. □

We are now ready to compute the period the box ladder graph and the *wheel graph* WS_{n+1} *with* $n+1$ *spokes*, $n \geq 1$. See Fig. 15.10.

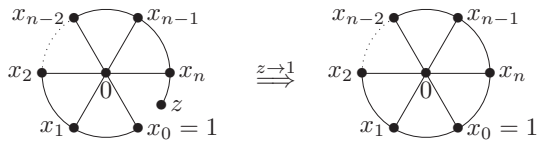


Figure 15.10: The Wheel Graph is the limit of $G_{2\{0\}^{n-1}}$ when $z \to 1$.

Theorem 15.4.19. *For any* $n \geq 2$ *we have*

$$I(\mathrm{WS}_{n+1}) = I(\mathrm{B}_n) = \binom{2n}{n}\zeta(2n-1).$$

Proof. We only need to compute the graphical function $f_{2\{0\}^{n-1}}(z)$ and then take the limit $z \to 1$. The main idea is repeated integration starting from the initial case f_2 which is given by Eq. (15.17), or, by Eq. (15.18) and Example 4.2.10

$$f_2(z) = \frac{\mathscr{L}_{10}(z) - \mathscr{L}_{01}(z)}{z-\bar{z}}. \tag{15.24}$$

Notice that an integration of $\mathscr{L}_{\mathbf{w}}$ with $\int_0 \mathrm{d}z/z$ or $\int_0 \mathrm{d}\bar{z}/\bar{z}$ is equivalent to prepending or appending one 0, respectively, according to Eq. (4.57) (notice

that Lemma 4.3.1 says that $\mathtt{y}_1 - \mathtt{x}_1$ contains at least two $\mathtt{x}_1$'s). Hence applying Prop. 15.4.17 with $a = 0$ exactly $(n-1)$-times we obtain

$$f_{2\{0\}^{n-1}}(z) = (-1)^n \frac{\mathcal{L}_{\{0\}^n 1\{0\}^{n-1}}(z) - \mathcal{L}_{\{0\}^{n-1}1\{0\}^n}(z)}{z - \bar{z}}.$$

If g is a single-valued function which vanishes at $z = 1$, we can take $z = 1+i\epsilon$ and use L'Hôpital's Rule to obtain

$$\lim_{z\to 0} \frac{g(z)}{z-\bar{z}} = \frac{1}{2}(\partial_z g(1) - \partial_{\bar{z}} g(1)).$$

Applying this to $f_{2\{0\}^{n-1}}(z)$ and using $\mathcal{L}_{\{0\}^{n-1}1\{0\}^n}(\bar{z}) = \mathcal{L}_{\{0\}^n 1\{0\}^{n-1}}(z)$ by Thm. 4.3.2(ii), we obtain

$$\begin{aligned} f_{2\{0\}^{n-1}}(1) &= \frac{(-1)^n}{2}(\partial_z - \partial_{\bar{z}})\Big(\mathcal{L}_{\{0\}^n 1\{0\}^{n-1}}(z) - \mathcal{L}_{\{0\}^{n-1}1\{0\}^n}(z)\Big)\Big|_{z=1} \\ &= (-1)^n \partial_z \Big(\mathcal{L}_{\{0\}^n 1\{0\}^{n-1}}(z) - \mathcal{L}_{\{0\}^{n-1}1\{0\}^n}(z)\Big)\Big|_{z=1}. \quad (15.25) \end{aligned}$$

It follows from Exercise 4.20 that

$$I(\mathrm{WS}_{n+1}) = f_{2\{0\}^{n-1}}(1) = \binom{2n}{n}\zeta(2n-1). \quad (15.26)$$

By combining Eqs. (15.9) and (15.26) we also get

$$I(\mathrm{B}_n) = \binom{2n}{n}\zeta(2n-1).$$

This completes the proof of the theorem. □

Theorem 15.4.20. *For any $n \geq 3$ the period of the graph Z_n is given by*

$$I(\mathrm{Z}_n) = \frac{4\cdot(2n-2)!}{n!(n-1)!}\Big(1 - \frac{1-(-1)^n}{2^{2n-3}}\Big)\zeta(2n-3). \quad (15.27)$$

Proof. By Eqs. (15.7) and (15.8) we only need to compute $I(\mathrm{Z}_{2n+3}) = f_{2\{10\}^n 1}(0)$ and $I(\mathrm{Z}_{2n+2}) = f_{2\{01\}^n}(0)$.

For all alternating words $\mathbf{w} \in \mathsf{H}^0$ we set $\mathbf{v} = \overleftarrow{\mathbf{w}}\mathtt{x}_1\mathtt{x}_0\mathbf{w}$. From Thms. 4.5.6 and 4.5.7 it is not hard to see that the functions $\tilde{f}_{2\mathbf{w}}$ defined by

$$\tilde{f}_{2\mathbf{w}}(z) = (-1)^{\mathrm{lg}(\mathbf{w})} \begin{cases} \dfrac{F_{\mathbf{v}}(z) - F_{\overleftarrow{\mathbf{v}}}(z)}{z-\bar{z}} & \text{if } \mathbf{w} = \mathtt{x}_0\mathbf{u}, \\ \dfrac{\hat{F}_{\mathbf{v}}(z) - \hat{F}_{\overleftarrow{\mathbf{v}}}(z)}{z-\bar{z}} & \text{if } \mathbf{w} = \mathtt{x}_1\mathbf{u}, \end{cases} \quad (15.28)$$

are single-valued, real-valued and satisfy the system of differential equations

$$\mathfrak{D}_{z,\bar{z}}\tilde{f}_{2\mathbf{w}\mathtt{x}_a}(z) = \frac{-1}{(z-a)(\bar{z}-a)} f_{2\mathbf{w}}(z) \quad (15.29)$$

for $a \in \{0,1\}$. Moreover, if $\mathbf{w} = \mathtt{e}$ is the empty word then going back to the $\{0,1\}^*$-word indexing for the MPLs we get

$$F_{\mathtt{x}_0\mathtt{x}_1}(z) = \hat{F}_{\mathtt{x}_0\mathtt{x}_1}(z) = \mathrm{L}_{01}(z) + \mathrm{L}_0(z)\,\mathrm{L}_1(\bar{z}) + \mathrm{L}_{01}(\bar{z}),$$
$$F_{\mathtt{x}_1\mathtt{x}_0}(z) = \hat{F}_{\mathtt{x}_1\mathtt{x}_0}(z) = \mathrm{L}_{10}(z) + \mathrm{L}_1(z)\,\mathrm{L}_0(\bar{z}) + \mathrm{L}_{10}(\bar{z}),$$

since the coefficients of $\mathtt{x}_0$, $\mathtt{x}_1$, $\mathtt{x}_0\mathtt{x}_1$ and $\mathtt{x}_1\mathtt{x}_0$ all vanish in the series S_{X} and $\hat{\mathsf{S}}_{\mathsf{X}}$ by Definitions 4.5.3 and 4.5.3. Thus

$$\begin{aligned}\tilde{f}_2(z) &= \frac{F_{\mathtt{x}_1\mathtt{x}_0}(z) - F_{\mathtt{x}_0\mathtt{x}_1}(z)}{z-\bar{z}}\\ &= \frac{\mathrm{L}_{10}(z) + \mathrm{L}_1(z)\,\mathrm{L}_0(\bar{z}) + \mathrm{L}_{01}(\bar{z}) - \mathrm{L}_{10}(\bar{z}) - \mathrm{L}_1(\bar{z})\,\mathrm{L}_0(z) - \mathrm{L}_{01}(z)}{z-\bar{z}}\\ &= f_2(z) = \frac{4i\,\mathscr{L}i_2(z)}{(z-\bar{z})}\end{aligned}$$

by (15.18). By the uniqueness in Thm. 15.4.12 we see that

$$\tilde{f}_{2\mathbf{w}}(z) = f_{2\mathfrak{b}(\mathbf{w})}(z)$$

where $\mathfrak{b} : \{\mathtt{x}_0, \mathtt{x}_1\}^* \to \{0,1\}^*$ is the homomorphism of the word algebras such that $\mathfrak{b}(\mathtt{x}_a) = a$ for $a = 0$ and 1.

Let us determine the regular value of $f_{2\mathbf{w}}(z)$ at $z = 0$ for $\mathbf{w} = \{10\}^n 1$ or $\{01\}^n$ (see Definition 4.4.1). Similar to the computation in Eq. (15.25) we can take $z = i\epsilon$ and use L'Hôpital's Rule to compute $\tilde{f}_{2\mathbf{w}}(0) = \mathrm{reg}_{z=0}\,\tilde{f}_{2\mathbf{w}}(z)$. Setting $\mathbf{v} = \overleftarrow{\mathbf{w}}\mathtt{x}_1\mathtt{x}_0\mathbf{w} = \mathtt{x}_1\mathbf{u}\,\mathtt{x}_1$, we thus obtain by Thms. 4.5.6 and 4.5.7

$$\tilde{f}_{2\mathbf{w}}(0) = \begin{cases} F_{\mathbf{u}\,\mathtt{x}_1}(0) - F_{\mathtt{x}_1\mathbf{u}}(0) & \text{if } \mathbf{w} = (\mathtt{x}_0\mathtt{x}_1)^n;\\ \hat{F}_{\mathtt{x}_1\mathbf{u}}(0) - \hat{F}_{\mathbf{u}\,\mathtt{x}_1}(0) & \text{if } \mathbf{w} = (\mathtt{x}_1\mathtt{x}_0)^n\mathtt{x}_1,\end{cases}$$

where we have used fact that $F_{\mathtt{x}_1\overleftarrow{\mathbf{u}}}(z) = F_{\mathbf{u}\,\mathtt{x}_1}(\bar{z})$ and $F_{\overleftarrow{\mathbf{u}}\,\mathtt{x}_1}(z) = F_{\mathtt{x}_1\mathbf{u}}(\bar{z})$. To calculate the regularized values above using the definitions in Eqs. (4.61) and (4.66), we note that $\mathrm{reg}_{z=0}\,\mathrm{L}_{\mathbf{t}} = 0$ for all nonempty words $\mathbf{t}$, so that by Lemma 4.4.2

$$\tilde{f}_{2\mathbf{w}}(0) = \begin{cases} \mathsf{S}_{\mathbf{u}\,\mathtt{x}_1} - \mathsf{S}_{\mathtt{x}_1\mathbf{u}} & \text{if } \mathbf{w} = (\mathtt{x}_0\mathtt{x}_1)^n;\\ \hat{\mathsf{S}}_{\mathtt{x}_1\,\mathbf{u}} - \hat{\mathsf{S}}_{\mathbf{u}\,\mathtt{x}_1} & \text{if } \mathbf{w} = (\mathtt{x}_1\mathtt{x}_0)^n\mathtt{x}_1.\end{cases}$$

Now, in the even weight case since $\mathbf{u} = (\mathtt{x}_0\mathtt{x}_1)^n\mathtt{x}_0(\mathtt{x}_0\mathtt{x}_1)^{n-1}\mathtt{x}_0$ we find that $\mathsf{S}_{\mathtt{x}_1\mathbf{u}}$ and $\mathsf{S}_{\mathbf{u}\,\mathtt{x}_1}$ are summands in $\mathsf{S}^1_{1,0}$ and $\mathsf{S}^1_{0,1}$, respectively. Hence by Eq. (4.58) we have

$$\begin{aligned}I(\mathrm{Z}_{2n+2}) &= \mathsf{S}^1_{1,0}\,(n,n) - \mathsf{S}^1_{1,0}\,(n+1,n-1)\\ &= 4\left[\binom{4n}{2n} - \binom{4n}{2n+2}\right]\zeta(4n+1) = \frac{4\cdot(4n+2)!}{(2n+2)!(2n+1)!}\zeta(4n+1).\end{aligned}$$

by Eq. (4.58). On the other hand, for Z_{2n+3}, we see that $\hat{\mathsf{S}}_{\mathbf{x}_1\mathbf{u}}$ and $\hat{\mathsf{S}}_{\mathbf{u}\,\mathbf{x}_1}$ are summands in $\hat{\mathsf{S}}^1_{1,0}$ and $\hat{\mathsf{S}}^1_{0,1}$, respectively. Thus we get

$$\begin{aligned} I(\mathrm{Z}_{2n+3}) &= \hat{\mathsf{S}}^1_{1,0}\,(n,n+1) - \hat{\mathsf{S}}^1_{1,0}\,(n+1,n) \\ &= 4\left(1-\frac{1}{2^{4n+2}}\right)\frac{(4n+4)!}{(2n+3)!(2n+2)!}\zeta(4n+3) \end{aligned}$$

by Eq. (4.59). Finally, straight-forward simplification finishes the proof of the zig-zag theorem. □

15.5 Feynman integrals in parametric form

The essential difficulty in evaluating Feynman integrals lies in the complexity of the corresponding Feynman graphs. To clarify this we now define some polynomials associated to these graphs. Later this will be used to express Feynman integrals.

15.5.1 *Graph polynomials in edge variables*

Definition 15.5.1. Suppose G is a Feynman graph as in Definition 15.2.2. Let $\mathcal{T}_1$ be the set of the spanning trees of G. Note that to get a spanning tree we need to remove exactly h_G edges, one from each loop. Cutting one more line of a spanning tree leads to two disconnected trees $(T1, T2)$, or a 2-tree (2-*forest* in graph theory). Let $\mathcal{T}_2$ be the set of all such pairs. To each $e \in E_G$ we associate a "Schwinger parameter" α_e. The *first* and *second Symanzik polynomials* are respectively defined by

$$\Psi_G := \sum_{T\in\mathcal{T}_1}\prod_{e\notin T}\alpha_e, \quad \Phi_G(\mathbf{q}) := \sum_{(T_1,T_2)\in\mathcal{T}_2}\left(\prod_{e\notin(T_1,T_2)}\alpha_e\right)\left(q^{T_1}\right)^2 \tag{15.30}$$

where $q^{T_1} = -q^{T_2}$ is the total momenta into T_1. Finally, we define

$$\Xi_G(\mathbf{m},\mathbf{q}) := \Phi_G(\mathbf{q}) + \left(\sum_{e\in E_G} m_e^2\alpha_e\right)\Psi_G \tag{15.31}$$

The following results are straight-forward.

Lemma 15.5.2. *The first (resp. second) Symanzik polynomial is homogeneous of degree h_G (resp. h_G+1) in the edge variables $\{\alpha_e : e \in E_G\}$.*

Example 15.5.3. Let G be the Feynman graph in Figure 15.11. Suppose $m_1 \neq 0$ (denoted by a thickened line) and $m_2 = m_3 = 0$. We have

$$\Psi_G = \alpha_1+\alpha_2+\alpha_3, \quad \Phi_G(\mathbf{q}) = q_1^2\alpha_2\alpha_3 + q_2^2\alpha_1\alpha_3 + q_3^2\alpha_1\alpha_2.$$

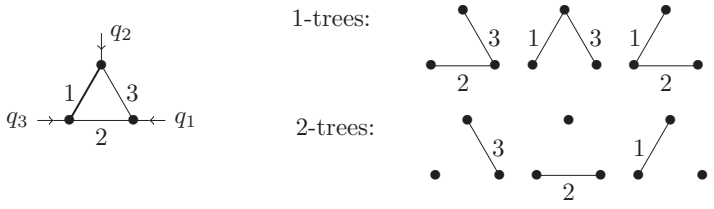


Figure 15.11: A 3-point 1-loop massive diagram and its spanning trees.

Hence

$$\Xi_G(\mathbf{m},\mathbf{q}) = q_1^2\alpha_2\alpha_3 + q_2^2\alpha_1\alpha_3 + q_3^2\alpha_1\alpha_2 + m_1^2\alpha_1(\alpha_1+\alpha_2+\alpha_3).$$

15.5.2 *Feynman integrals in projective space*

Let G be a Feynman graph as in Definition 15.2.2. Let n_G be the number of edges of G as before. Using Schwinger parameter (edge variables) we can define the Feynman integral associated with G by the integral in projective space $\mathbb{P}^{n_G-1}$:

$$I_G(\mathbf{m},\mathbf{q},D) := \Gamma\Big(n_G - h_G\frac{D}{2}\Big)\int_\sigma f_G(\mathbf{m},\mathbf{q},D)\cdot\Omega_G \tag{15.32}$$

where

$$\begin{aligned} f_G(\mathbf{m},\mathbf{q},D) &= \frac{1}{\Psi_G^{D/2}}\left(\frac{\Psi_G}{\Xi_G(\mathbf{m},\mathbf{q})}\right)^{n_G - h_G\frac{D}{2}}, \\ \Omega_G &= \sum_{j=1}^{n_G}(-1)^j\alpha_j d\alpha_1\cdots\widehat{d\alpha_j}\cdots d\alpha_{n_G}, \\ \sigma &= \Big\{(\alpha_1:\cdots:\alpha_{n_G}):\ \alpha_j\in\mathbb{R}_{\geq 0}\Big\}\subset\mathbb{P}^{n_G-1}(\mathbb{R}). \end{aligned}$$

Note that $f_G(\mathbf{m},\mathbf{q},D)$ is homogeneous of degree $-n_G$ so that by Euler's Homogeneous Function Theorem

$$\left(n_G + \sum_{e\in E_G}\alpha_e\frac{\partial}{\partial\alpha_e}\right) f_G(\mathbf{m},\mathbf{q},D) = 0$$

so $d(f_G(\mathbf{m},\mathbf{q},D)\cdot\Omega_G) = 0$. Therefore $I_G(\mathbf{m},\mathbf{q},D)$ is well-defined. One can always take $\alpha_j = 1$ and change the integral into an ordinary affine integral.

Example 15.5.4. Let G be the Feynman graph in Figure 15.11. Suppose further $m_1 = m_2 = m_3 = 0$ but $q_1^2 + q_2^2 + q_3^2 \neq 0$. We have

$$I_G := I_G(\mathbf{m}, \mathbf{q}, 4) = \int_\sigma \frac{\Omega_G}{(\alpha_1+\alpha_2+\alpha_3)(q_1^2\alpha_2\alpha_3 + q_2^2\alpha_1\alpha_3 + q_3^2\alpha_1\alpha_2)}.$$

If one of the incoming momenta vanishes then the integral becomes trivial (see Exercise 15.2). So we now assume $q_j^2 \neq 0$ for $j = 1, 2, 3$. By the change of variables $\alpha_j \to q_j^2\alpha_j$ we get

$$I_G = \int_\sigma \frac{\Omega_G}{(q_1^2\alpha_1 + q_2^2\alpha_2 + q_3^2\alpha_3)(\alpha_2\alpha_3 + \alpha_1\alpha_3 + \alpha_1\alpha_2)}.$$

First we assume

$$q_1^2 t^2 + (q_3^2 - q_1^2 - q_2^2)t + q_2^2 = 0 \tag{15.33}$$

has two distinct real roots z and $\bar{z}$. Then we get

$$\begin{aligned} q_1^2 I_G &= \int_\sigma \frac{\Omega_G}{(\alpha_1 + z\bar{z}\alpha_2 + (1-z)(1-\bar{z})\alpha_3)(\alpha_2\alpha_3 + \alpha_1\alpha_3 + \alpha_1\alpha_2)} \\ &= \int_{\substack{(\alpha_1:\alpha_2)\in\mathbb{P}^1 \\ \alpha_1,\alpha_2\geq 0}} \frac{\Omega_G}{(\alpha_1 + z\alpha_2)(\alpha_1 + \bar{z}\alpha_2)} \log \frac{(\alpha_1+\alpha_2)(\alpha_1 + z\bar{z}\alpha_2)}{(1-z)(1-\bar{z})\alpha_1\alpha_2} \end{aligned}$$

by integrating over $\alpha_3 \in (0, \infty)$ using the partial fraction expansion:

$$\frac{1}{(at+b)(ct+d)} = \det \begin{vmatrix} a & b \\ c & d \end{vmatrix}^{-1} \left(\frac{1}{t + b/a} - \frac{1}{t + d/c} \right)$$

where

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha_1 + \alpha_2 & \alpha_1\alpha_2 \\ (1-z)(1-\bar{z}) & \alpha_1 + z\bar{z}\alpha_2 \end{bmatrix}.$$

By messy computation (see Exercise 15.3) we finally obtain

$$I_G = \frac{1}{q_1^2(z-\bar{z})} \left(2\operatorname{Li}_2(z) - 2\operatorname{Li}_2(\bar{z}) + \log(z\bar{z}) \log \frac{1-z}{1-\bar{z}} \right) = \frac{2\mathcal{L}i_2(z)}{q_1^2 \operatorname{Im}(z)}$$

where $\mathcal{L}i_2(z) = \operatorname{Im}\big(\operatorname{Li}_2(z)\big) + \arg(1-z)\log|z|$ is the single-valued Bloch-Wigner dilog function (2.46).

If Eq. 15.33 has real roots, then similarly we can see that I_G can be expressed in terms of dilogs and logarithms.

The following result can be proved by complicated reduction process which we refer the interest reader to [150]. The motivic explanation is given by [54, Prop. 8.3].

Theorem 15.5.5. *Every one-loop Feynman integral can be expressed in terms of dilogs and logarithms evaluated at algebraic functions of the kinematic invariants (momenta and masses).*

15.6 Historical notes

In particle physics, Z boson was found in 1983 at CERN using photon anti-photon collisions. However, by far the best and most comprehensive measurements of the Z boson were made in the 1980s (before its discovery) and 1990s by electron-positron experiments at CERN's Large Electron-Positron Collider.

The structure of ϕ^4 periods was first studied systematically by Broadhurst and Kreimer in 1995 [92] (see also [93]). Usually, four equivalent ways can be used to define a graph period: position space, momentum space, parametric space, and dual parametric space. Hence different methods have be utilized: Mellin-Barnes integrals, hypergeometric functions, differential equation, difference equations and so on. To apply techniques from algebraic geometry one often uses the Feynman (Schwinger) parametric form.

The use of differential equations in the study of loop integrals has a long history (see [501]). The equations considered in Sec. 15.4 are for off-shell integrals and are second-order inhomogeneous equations. Similar equations for on-shell integrals can be found in [165].

[259] www.americanscientist.org/issues/pub/g-ology/6

According to [91], the Feynman integral corresponding to the Feynman diagram $K_{3,4}$ in Fig. 15.12 is given by

$$\frac{27}{5}\zeta(5,3)+\frac{45}{4}\zeta(5)\zeta(3)-\frac{261}{20}\zeta(8).$$

Note that this graph is the simplest graph that does not have Hamiltonian circles.

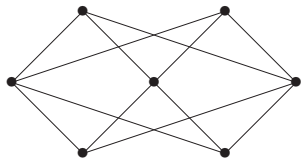


Figure 15.12: The first irreducible double zeta $\zeta(5,3)$ appears in this Feynman diagram.

Until 2012 only eight ϕ^4 periods have been calculated: The trivial period 1 ($= P(\mathrm{Z}_1)$), the wheels with three [133] and with four [134] spokes which are Z_3 and Z_4, respectively. The period of the zig-zag graph Z_5 with 5 (see

Figure 15.6) loops was calculated by Kazakov in 1983 [333], while Z_6 by Broadhurst in 1985 [85] and Ussyukina in 1991 [536].

In the mid 1990s it was realized that there is a close connection between perturbative QFT/QCD and knot theory. Kreimer [351] described a way to connect the Feynman integrals with UV-divergences to knot invariants where the knots are extracted from the topology of the corresponding Feynman diagrams.

Until 2012 the only calculated non zig-zag periods were $G(3,1,0)$ in [92] which is $P_{6,2}$ in [484], $P_{6,3}$ by Panzer, and the bipartite graph $K_{4,4}$ in [483] which is $P_{6,4}$ in [484]. By exact numerical methods twenty-four more periods were determined up to eight loops [484], all of which are MZVs.

The sector decomposition method (see [63] and its references) was developed to compute numerically rather complicated multi-scale Feynman integrals. But this method typically produces a large number of convergent parametric integrals which are difficulty to analyze precisely and rigorously by analytical tools in mathematics. Moreover, it often produces spurious intermediate quantities which are canceled out at the end. In [453,454] Panzer discovered a new approach to computing these integrals by applying an appropriate sequence of partial integrations which ultimately render all Feynman parameter integrations in terms of CMZVs when the integrands are linear reducible.

The computation of the periods of wheel graphs and the zig-zag graphs in Sec. 15.4.1 by using sequential graphs and graphical functions follows mostly that of Schnetz [487]. Sequential graphs were also independently considered in by Drummond [166].

The period equation of the boxed ladder graph in Exercise 15.7 was first derived by N. Ussyukina and A. Davydychev in [537].

Exercises

15.1. Show that the homological groups of a graph are given by (15.1).

15.2. Show that

$$\int_0^\infty \int_0^\infty \frac{\mathrm{d}\alpha_1 d\alpha_2}{(\alpha_1+\alpha_2+q^2)(\alpha_1+\alpha_2+\alpha_1\alpha_2)}$$

can be expressed using logarithms evaluated at some algebraic functions of q^2.

15.3. We can use the following steps to compute I_G in Example 15.5.4.

(i) Let $\alpha_2 = 1$. We get

$$q_1^2 I_G = \int_0^\infty \frac{\mathrm{d}\alpha}{(\alpha+z)(\alpha+\bar{z})} \log \frac{(\alpha+1)(\alpha+z\bar{z})}{(1-z)(1-\bar{z})\alpha}.$$

Let $f(z,\bar{z})$ be the above integral over $(0,1)$. Then

$$q_1^2 I_G = f(z,\bar{z}) - f(z^{-1},\bar{z}^{-1}).$$

(ii) By partial fractions

$$(z-\bar{z}) f(z,\bar{z}) = g(\bar{z},z) - g(z,\bar{z})$$

where

$$g(z,\bar{z}) = \int_0^1 \frac{\mathrm{d}\alpha}{\alpha+z} \log \frac{(\alpha+1)(\alpha+z\bar{z})}{(1-z)(1-\bar{z})\alpha}.$$

(iii) Evaluate $g(z,\bar{z})$ by using the formula

$$\int_0^x \frac{\log(t+b)}{t+a}\,dt = \mathrm{Li}_2\left(\frac{a+x}{a-b}\right) - \mathrm{Li}_2\left(\frac{a}{a-b}\right) + \log(b-a)\log\frac{a+x}{a}.$$

(iv) Simplify the above results by using the five-term relation a few times:

$$\mathrm{Li}_2(xy) = \mathrm{Li}_2(x)+\mathrm{Li}_2(y)+\mathrm{Li}_2\left(\frac{xy-x}{1-x}\right)+\mathrm{Li}_2\left(\frac{xy-y}{1-y}\right)+\frac{1}{2}\log^2\left(\frac{1-x}{1-y}\right).$$

15.4. Suppose the D-dimensional spherical coordinates are given by

$$\begin{aligned}
x^1 &= r\cos\phi^1,\\
x^2 &= r\sin\phi^1\cos\phi^2,\\
&\vdots\\
x^{D-1} &= r\sin\phi^1\cdots\sin\phi^{D-2}\cos\phi^{D-1},\\
x^D &= r\sin\phi^1\cdots\sin\phi^{D-2}\sin\phi^{D-1},
\end{aligned}$$

where $r\in[0,\infty)$, $\phi^1,\ldots,\phi^{D-2}\in[0,\pi]$, and $\phi^{D-1}\in[0,2\pi]$. Show that the Laplacian is

$$\begin{aligned}
\Delta^{(D)} &:= \sum_{k=1}^D \partial_{x^k}^2\\
&= \frac{1}{r^{D-1}}\partial_r r^{D-1}\partial_r + \frac{1}{r^2(\sin\phi^1)^{D-2}}\partial_{\phi^1}(\sin\phi^1)^{D-2}\partial_{\phi^1}\\
&\quad + \frac{1}{(r\sin\phi^1)^2(\sin\phi^2)^{D-3}}\partial_{\phi^2}(\sin\phi^2)^{D-3}\partial_{\phi^2}\\
&\quad + \cdots\\
&\quad + \frac{1}{(r\sin\phi^1\cdots\sin\phi^{D-3})^2\sin\phi^{D-2}}\partial_{\phi^{D-2}}\sin\phi^{D-2}\partial_{\phi^{D-2}}\\
&\quad + \frac{1}{(r\sin\phi^1\cdots\sin\phi^{D-2})^2}\partial_{\phi^{D-1}}^2.
\end{aligned}$$

15.5. By using Exercise 15.4 show that

$$\frac{1}{r\sin\phi^1}\Delta^{(2)}\circ M(r\sin\phi^1)=\partial_r^2+\frac{3}{r}\partial_r+\frac{1}{r^2}\partial_{\phi^1}^2+\frac{2\cos\phi^1}{r^2\sin\phi^1}\partial_{\phi^1},$$

15.6. Assume $f_{\mathbf{v}1}(z)=g(z)/(z-\bar{z})$. By taking $z=i\epsilon$ and using L'Hôpital's Rule to show that $I(G_{\mathbf{v}2})=\frac{1}{2}(\partial_z g(0)-\partial_{\bar{z}}g(0))$.

15.7. By using Exercise 4.20 show that

$$f_{2\{0\}^{n-1}}(z)=\sum_{k=n}^{2n}(-1)^k\binom{k}{n}\frac{(\log z\bar{z})^{2n-k}}{(2n-k)!}\frac{\mathrm{Li}_k(z)-\mathrm{Li}_k(\bar{z})}{z-\bar{z}}.$$

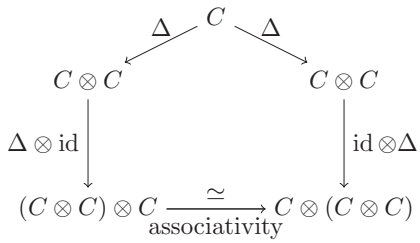
Appendix A

Key concepts of Hopf algebras

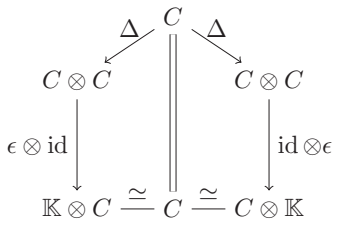
In this appendix we provide a quick review of the basic definitions and properties of Hopf algebras that we have used in this book.

Definition A.1. A *coalgebra* (C, Δ, ϵ) over a field $\mathbb{K}$ is a $\mathbb{K}$-vector space C with two $\mathbb{K}$-linear maps $\Delta : C \to C \otimes C$ (called *coproduct*) and $\epsilon : C \to \mathbb{K}$ (called *counit*) such that the following two diagrams commutate:

(i) Coassociativity



(ii) ϵ and Δ are compatible.



Dual to the concept of commutativity, we say C is *cocommutative* if $\tau \circ \Delta = \Delta$ where $\tau(x \otimes y) = y \otimes x$ for all $x, y \in C$.

Example A.2. Let $\mathbb{K}[x]$ be the polynomial ring over $\mathbb{K}$. We can define

$$\Delta(1) = 1 \otimes 1, \ \Delta(x) = 1 \otimes x + x \otimes 1 \ \text{ and } \ \epsilon(x^n) = \delta_{0,n}$$

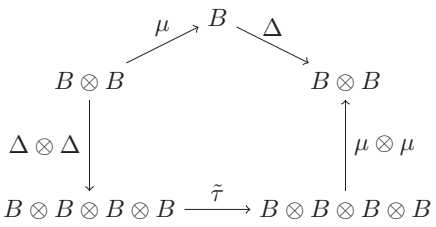
where $\delta_{0,n}$ is the Kronecker symbol. Then it is not hard to see that

$$\Delta(x^n) = \sum_{j=0}^{n} x^j \otimes x^{n-j}$$

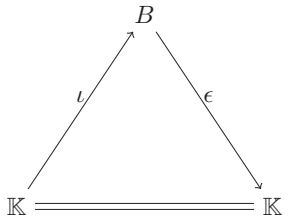
and that $(\mathbb{K}[x], \Delta, \epsilon)$ fomrs a coalgebra.

Definition A.3. A *bialgebra* $(B, \mu, \iota, \Delta, \epsilon)$ over a field $\mathbb{K}$ is a $\mathbb{K}$-vector space B with four $\mathbb{K}$-linear maps $\mu : B \otimes B \to B$ (the *product*, also denoted by $\bullet$ sometimes), $\iota : \mathbb{K} \to B$ (the *unit*, or *identity*), $\Delta : B \to B \otimes B$ (the coproduct) and $\epsilon : B \to \mathbb{K}$ (the counit), such that (B, μ, ι) is an associative algebra and (C, Δ, ϵ) is a coassociative coalgebra satisfying the following four compatibility conditions:

(iii) μ and Δ

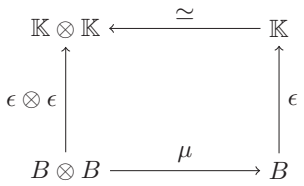


(iv) ι and ϵ

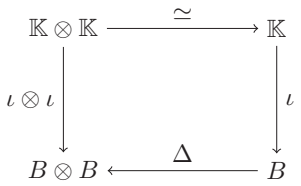


where $\tilde{\tau} = \mathrm{id} \otimes \tau \otimes \mathrm{id}$ and

(v) μ and ϵ

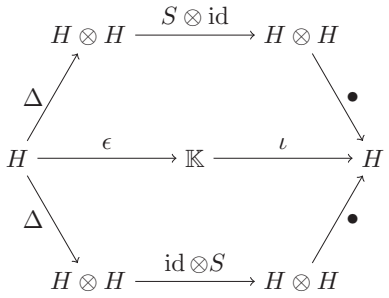


(vi) ι and Δ



Condition (iii) implies the unit ι is a coalgebra homomorphism and condition (vi) means the counit ϵ is an algebra homomorphism. We say B is *commutative* if it is commutative as an algebra. We say B is *cocommutative* if it is cocommutative as a coalgebra.

Definition A.4. A *Hopf algebra* is a bialgebra H over a field $\mathbb{K}$ together with a $\mathbb{K}$-linear anti-homomorphism $S : H \to H$ (called the *antipode*) such that (vii) S is compatible with all the other four structure maps, *i.e.*, the following diagram commutes:



Example A.5. The simplest nontrivial example of a Hopf algebra is given by the group algebra

$$k[G] := \left\{ \sum_{j=1}^{n} a_j g_j : n \in \mathbb{N}, a_j \in \mathbb{K}, g_j \in G \ \forall j = 1, \cdots, n \right\}$$

for any group G and field $\mathbb{K}$. The algebra structure is straight-forward. For the coalgebra structure, we define $\Delta(g) = g \otimes g$, $\epsilon(g) = 1$ and $S(g) = g^{-1}$ for all $g \in G$. It is easy to check all the Hopf algebra axioms are satisfied.

Because of this example, we have the following definition.

Definition A.6. An element x in any coalgebra (C, ϵ, Δ) is called *group-like* (or *diagonal*) if $\Delta(x) = x \otimes x$ and $\epsilon(x) = 1$.

Example A.7. In the study of Lie algebras $\mathfrak{g}$ over a field $\mathbb{K}$ (see Appendix B), we can give the universal enveloping algebra $U\mathfrak{g}$ a natural Hopf algebra structure by imposing some conditions on the structure maps as follows. For all $x \in \mathfrak{g}$, we set $\Delta(x) = x \otimes 1 + 1 \otimes x$, $\epsilon(x) = 0$ and $S(x) = -x$, where 1 is the unit in $\mathbb{K}$.

Definition A.8. An element x in any bialgebra B over a field $\mathbb{K}$ with coproduct Δ is called *primitive* if $\Delta(x) = x \otimes 1 + 1 \otimes x$, where 1 is the unit in $\mathbb{K}$.

Definition A.9. A subset I of a coalgebra (C, Δ, ϵ) is called a *co-ideal* if $I \subseteq \ker(\epsilon)$ and $\Delta(I) \subseteq I \otimes C + C \otimes I$.

Proposition A.10. If I is a co-ideal of a coalgebra (C, Δ, ϵ) then the quotient space C/I has a natural coalgebra structure.

Proof. It is straight-forward to check that

$$\big((\Delta\otimes\mathrm{id})\circ\Delta\big)(I)\subseteq I\otimes C\otimes C+C\otimes I\otimes C+C\otimes C\otimes I,$$
$$\big((\mathrm{id}\otimes\Delta)\circ\Delta\big)(I)\subseteq I\otimes C\otimes C+C\otimes I\otimes C+C\otimes C\otimes I,$$

which vanishes in $(C/I)\otimes(C/I)\otimes(C/I)$. Similar,

$$\big((\epsilon\otimes\mathrm{id})\circ\Delta\big)(I)\subseteq I,\quad \big((\mathrm{id}\otimes\epsilon)\circ\Delta\big)(I)\subseteq I,$$

which vanishes in C/I. The proposition follows easily from these facts. □

We now consider bialgebras with some grading.

Definition A.11. A bialgebra B over a field $\mathbb{K}$ is *graded* if $B=\oplus_{n\geq 0}B_n$ where the elements in B_n all have weight n such that

(1) $\dim_k B_n<0$ as a $\mathbb{K}$-vector space for all $n\geq 0$
(2) The product $B_m\bullet B_n\subseteq B_{m+n}$ for all $m,n\geq 0$
(3) The coproduct $\Delta(B_n)\subseteq\oplus_{j=0}^{n}B_j\otimes B_{n-j}$ for all $n\geq 0$

In this book we always consider *connected* graded bialgebras, *i.e.*, $\dim_k B_0=1$.

Example A.12. Let Γ_N be the set of Nth roots of unity. Let $\mathsf{X}=\mathsf{X}_N$ be an alphabet with $N+1$ letters $\{x_0,x_\xi:\xi\in\Gamma_N\}$ and X^* the set of words over X. All the letters in X have weight one. The empty word e has weight zero. Let $\mathbb{Q}\langle\mathsf{X}\rangle$ be the $\mathbb{Q}$-linear vector space generated by e and X^*. It becomes a graded Hopf algebra when we define the product by concatenation, the unit by $\iota(\alpha)=\alpha\mathsf{e}$ for all $\alpha\in\mathbb{Q}$, the coproduct by $\Delta(x)=1\otimes x+x\otimes 1$ for all $x\in\mathsf{X}$, the counit by $\epsilon(\mathbf{w})=0$ for all nontrivial word $\mathbf{w}$ and $\epsilon(\mathsf{e})=1$, and the antipode by $S(\mathsf{e})=\mathsf{e}$ and $S(x)=-x$ for all $x\in\mathsf{X}$.

For any field $\mathbb{K}$ containing $\mathbb{Q}$ we may define $\mathbb{K}\langle\mathsf{X}\rangle$ easily by extension of scalars.

Theorem A.13. For any word $\mathbf{w}\in\mathsf{X}^*$, we have

$$S(\mathbf{w})=(-1)^{|\mathbf{w}|}\overleftarrow{\mathbf{w}},$$

where $\overleftarrow{\mathbf{w}}$ is the reversal of $\mathbf{w}$.

Proof. Clear by anti-homomorphism. □

Example A.14. Another important example of Hopf algebra is the set of quasi-symmetric functions QSym_2 defined in Definition 7.3.2. We can make QSym_2 a commutative Hopf algebra by setting its coproduct Δ by

$$\Delta(M_{s_1,\ldots,s_d}) = \sum_{j=0}^{d} M_{s_1,\ldots,s_j} \otimes M_{s_{j+1},\ldots,s_d},$$

and counit ε with $\varepsilon(u) = 0$ for all elements u of positive degree. Its antipode S is an involution of QSym_2, *i.e.*, an automorphism with $S^2 = \mathrm{id}$.

Theorem A.15. The antipode S of QSym_2 is given by the followings: for every composition $\mathbf{s}$ of signed numbers,

(i). $S(M_{\mathbf{s}}) = (-1)^{\mathrm{dp}(\mathbf{s})} E_{\overleftarrow{\mathbf{s}}}$, where $E_{\mathbf{t}} = \sum_{\mathbf{s} \preceq \mathbf{t}} M_{\mathbf{s}}$ and $\mathbf{t} \preceq \mathbf{s}$ means $\mathbf{t}$ can be obtained from $\mathbf{s}$ by combining some of its parts using $\oplus$.

(ii.) $S(M_{\mathbf{s}}) = (-1)^{\mathrm{dp}(\mathbf{s})} E_{\overleftarrow{\mathbf{s}}} = \displaystyle\sum_{\bigsqcup_{j=1}^{r} \mathbf{s}_j = \mathbf{s}} (-1)^r M_{\mathbf{s}_1} M_{\mathbf{s}_2} \cdots M_{\mathbf{s}_r}$, where $\bigsqcup_{j=1}^{r} \mathbf{s}_j$ is the concatenation of $\mathbf{s}_1$ to $\mathbf{s}_r$.

(iii). $(-1)^{\mathrm{dp}(\mathbf{s})} M_{\overleftarrow{\mathbf{s}}} = \displaystyle\sum_{\bigsqcup_{j=1}^{r} \mathbf{s}_j = \mathbf{s}} (-1)^r E_{\mathbf{s}_1} E_{\mathbf{s}_2} \cdots E_{\mathbf{s}_r}$.

Proof. Assume first $\mathbf{s}$ is a composition of positive integers. Then part (i) is given by [178, Prop. 3.4] or [399, Cor. 2.3]. Part (ii) follow from the general theory of Hopf algebras (see [178, p. 7]). Applying antipode S to (ii) and use (i) we can derive (iii) quickly. A careful reading reveals that all the proofs are still valid when all the components of $\mathbf{s}$ are signed number. □

Appendix B

Some useful results of Lie algebras

In this appendix we give a very concise review of the basic results of Lie algebras that we need in this book.

A *Lie algebra* is an algebra $(\mathfrak{g}, [\cdot,\cdot])$ where $\mathfrak{g}$ is a vector field over some field $\mathbb{K}$ together with a $\mathbb{K}$-bilinear binary operation $[\cdot,\cdot]$, called *Lie bracket*, that satisfies $[x,x]=0$ for all $x \in \mathfrak{g}$ and the following *Jacobi identity*:

$$[x,[y,z]]+[z,[x,y]]+[y,[z,x]]=0 \quad \forall x,y,z \in \mathfrak{g}.$$

Note that $[x,x]=0$ implies the skew-symmetry (or anti-commutativity) $[x,y]=-[y,x]$ for all $x \in \mathfrak{g}$.

Theorem B.16. If B is a bi-algebra (see Appendix A), then the set of primitive elements form a Lie algebra with the usual commutator bracket $[x,y]=xy-yx$.

Proof. We only need to show $[x,y]$ is primitive if both x and y are primitive. Indeed,

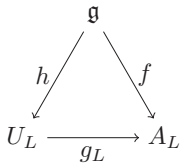
$$\begin{aligned}\Delta([x,y]) &= \Delta(x)\Delta(y)-\Delta(y)\Delta(x)\\ &= (x\otimes 1+1\otimes x)(y\otimes 1+1\otimes y)-(y\otimes 1+1\otimes y)(x\otimes 1+1\otimes x)\\ &= xy\otimes 1+1\otimes xy-yx\otimes 1-1\otimes yx\\ &= [x,y]\otimes 1+1\otimes [x,y],\end{aligned}$$

as desired. □

Example B.17. Let A be an associative algebra over the field $\mathbb{K}$. Then A becomes a Lie algebra over $\mathbb{K}$, denoted by A_L, if we define the Lie bracket by

$$[a,b]=ab-ba \quad \forall a,b \in A.$$

Let $\mathfrak{g}$ be any Lie algebra over $\mathbb{K}$. Given a unital associative $\mathbb{K}$-algebra U and a Lie algebra homomorphism: $h : \mathfrak{g} \longrightarrow U_L$ (notation as in Example B.17), we say that U is the *universal enveloping algebra* of $\mathfrak{g}$, denoted by $U\mathfrak{g}$, if it satisfies the following universal property: for any unital associative $\mathbb{K}$-algebra A and Lie algebra homomorphism $f : \mathfrak{g} \longrightarrow A_L$ there exists a unique unital algebra homomorphism $g : U \longrightarrow A$ such that the following diagram commutes:



A standard way to construct the universal enveloping algebra of $\mathfrak{g}$ is to take $T\mathfrak{g}/I$ where $T\mathfrak{g}$ is the tensor algebra of $\mathfrak{g}$ and I is the two-sided ideal generated by $a \otimes b - b \otimes a - [a, b]$ for all $a, b \in \mathfrak{g}$. The universal enveloping algebra $U\mathfrak{g}$ has a natural way to become a Hopf algebra (see Example A.7).

Example B.18. For any set X and a $\mathbb{Q}$-algebra R, let $R\langle\mathsf{X}\rangle$ (resp. $\mathfrak{Lib}_R(\mathsf{X})$) be the free[2] noncommutative associative algebra (resp. the free Lie algebra constructed as in Example B.17) on X with coefficients in R, where the product is given by concatenation. Note that $R\langle\mathsf{X}\rangle$ can be viewed as the universal enveloping algebra of $\mathfrak{Lib}_k(\mathsf{X})$. When all elements in X are given weight one $R\langle\mathsf{X}\rangle$ has a graded Hopf algebra structure as defined in Example A.12. We denote by $R\langle\langle\mathsf{X}\rangle\rangle$ the set of all formal series on X over the field $\mathbb{K}$.

The above example and the next one have played important roles in the study of regularized double shuffle relations of multiple polylog values.

Example B.19. Let N be a positive integer and $\mu = \exp(2\pi\sqrt{-1}/N)$. Let $\Gamma_N := \{\mu^j : j \in \mathbb{Z}/N\mathbb{Z}\}$ be the set of Nth roots of unity. Denote by Y the infinite alphabet consisting of letters $\mathtt{y}_{n,\xi}$ for $n \in \mathbb{N}$ and $\xi \in \Gamma_N$, and Y^* the set of the words over Y. For any fixed $n \in \mathbb{N}$ and $\xi \in \boldsymbol{\mu}_N$, the letters $\mathtt{y}_{n,\xi}$ has colored weight (n, ξ). The empty word $\mathtt{e}$ has weight zero. We extend the definition of $\mathtt{y}_{n,\xi}$ to the case where $n = 0$ by setting

$$\mathtt{y}_{0,\xi} = \begin{cases} 1, \text{ if } \xi = 1; \\ 0, \text{ if } \xi \neq 1. \end{cases}$$

[2] $\mathfrak{Lib}$ comes from the French *libre* for free.

Let $\mathbb{K}\langle\langle \mathsf{Y} \rangle\rangle$ be the graded noncommutative associative algebra on Y with coefficients in $\mathbb{K}$, where the product is given by concatenation. Put on $\mathbb{K}\langle\langle \mathsf{Y} \rangle\rangle$ a grading using *colored weights* as follows:

$$k\langle\langle \mathsf{Y} \rangle\rangle = \bigoplus_{\substack{n\in\mathbb{N}_0 \\ \xi\in\boldsymbol{\mu}}} B_{n,\xi}, \quad \text{where } B_{n,\xi} = \sum_{\substack{d\in\mathbb{N},\ \xi_1\xi_2\cdots\xi_d=\xi \\ s_1+\cdots+s_d=n}} \mathtt{y}_{s_1,\xi_1}\cdots\mathtt{y}_{s_d,\xi_d},$$

Similar to Example A.12 we can equip $\mathbb{K}\langle\langle \mathsf{Y} \rangle\rangle$ with a graded bialgebra structure when we define the unit by $\iota(\alpha) = \alpha\mathtt{e}$ for all $\alpha \in \mathbb{K}$, the coproduct by

$$\Delta_*(\mathtt{y}_{n,\xi}) = \sum_{k+\ell=n,\lambda\nu=\xi} \mathtt{y}_{n,\xi} \otimes \mathtt{y}_{n,\xi} \quad \forall n \in \mathbb{N}_0, \xi \in \boldsymbol{\mu}, \tag{B.1}$$

the counit by $\epsilon(\mathbf{w}) = 0$ for all nontrivial word $\mathbf{w}$ and $\epsilon(\mathtt{e}) = 1$.

Remark B.20. It is crucial to understand that the graded bialgebra $(R\langle \mathsf{Y} \rangle, \Delta_*)$ is not the universal enveloping algebra of the free Lie algebra $\mathfrak{Lib}_k(\mathsf{Y})$.

Put

$$\mathcal{Y} := \sum_{n\in\mathbb{N}_0,\xi\in\boldsymbol{\mu}} \mathtt{y}_{n,\xi}.$$

Since it is group-like for Δ_* (see Exercise 13.5) we know $\log(\mathcal{Y})$ is primitive by [469, Thm. 3.2]. Let $u_{n,\xi}$ be the component of $\log(\mathcal{Y})$ in $B_{n,\xi}$ (*i.e.*, the colored weight (n,ξ) piece). It is obvious that $u_{0,\xi} = 0$ for all $\xi \in \boldsymbol{\mu}$. Let

$$U = \{u_{n,\xi} : n \in \mathbb{N}, \xi \in \boldsymbol{\mu}\}.$$

Then all the elements in U are primitive since $\log(\mathcal{Y})$ is primitive and Δ_* respects the colored weight grading. So by Example A.12 we see that $(\mathbb{Q}\langle U \rangle, \bullet, \Delta)$ is the universal enveloping algebra of the free Lie algebra $\mathfrak{Lib}_{\mathbb{Q}}(U)$.

Proposition B.21. The associative $\mathbb{Q}$-algebra $\mathbb{Q}\langle\langle \mathsf{Y} \rangle\rangle$ is freely generated by U. Moreover, as bialgebras

$$(\mathbb{Q}\langle\langle \mathsf{Y} \rangle\rangle, \bullet, \Delta_*) \cong (\mathbb{Q}\langle U \rangle, \bullet, \Delta)$$

Proof. Since $-\sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} \mathtt{y}_{n,\xi} = 1 - \mathcal{Y}$, by the definition

$$\sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} u_{n,\xi} = \log(\mathcal{Y}) = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \Bigg(\sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} \mathtt{y}_{n,\xi} \Bigg)^j.$$

By considering the depth filtration we see immediately that

$$u_{n,\xi} = \mathtt{y}_{n,\xi} + (\text{terms with depth} > 1) \tag{B.2}$$

and more generally,

$$u_{s_1,\xi_1} \cdots u_{s_d,\xi_d} = \mathtt{y}_{s_1,\xi_1} \cdots \mathtt{y}_{s_d,\xi_d} + (\text{terms with depth} > d).$$

This yields the proposition directly. □

We now take $\mathsf{X} = \mathsf{X}_N$ as in Example A.12 and use the notation in Example B.18. Further we regard Y as a subset of X^* by defining $\mathtt{y}_{n,\xi} = x_0^{n-1}x_\xi$ and set $\pi_{\mathsf{Y}} : R\langle\langle \mathsf{X} \rangle\rangle \to R\langle\langle \mathsf{Y} \rangle\rangle$ by sending all words ending with x_0 to 0. Let

$$\mathbf{q}(\mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2}\cdots\mathtt{y}_{s_d,\xi_d}) = \mathtt{y}_{s_1,\xi_1}\mathtt{y}_{s_2,\xi_2\xi_1^{-1}}\cdots\mathtt{y}_{s_d,\xi_d\xi_{d-1}^{-1}}.$$

Then we have

Theorem B.22. Let $\mathbb{K}$ be a field extension of $\mathbb{Q}$. Let Φ be a group-like element in $(R\langle\langle \mathsf{X} \rangle\rangle, \Delta)$ whose coefficient of x_1 is 0. Then

(i) $\log(\Phi)$ is a primitive element.

(ii) $\overleftarrow{\Phi}(-x_0, -x_1) = \Phi(x_0, x_1)^{-1}$ where $\leftarrow$ means the reversal of all the words in the coefficients.

(iii) Suppose $S \in \mathbb{K}[\![t]\!]$ such that

$$\Phi_* := S(\mathtt{y}_{1,1}) \cdot \mathbf{q}\pi_{\mathsf{Y}}(\Phi) \tag{B.3}$$

in $R\langle\langle \mathsf{Y} \rangle\rangle$ is diagonal for Δ_* whose coefficient of $\mathtt{y}_{1,1}$ is 0. Then we have

$$S(t) = \exp\left(\sum_{n=2}^{\infty} \frac{(-1)^{n-1}}{n} (\Phi | x_0^{n-1}x_1) t^n\right).$$

Proof. (i) follows easily from [469, Thm. 3.2].

(ii) follows from the formula for the antipode (see Thm. A.13) of the Hopf algebra $(R\langle\langle \mathsf{X} \rangle\rangle, \Delta)$ (*i.e.*, signed reversal of words) and [469, Prop. 1.10].

For (iii), regard the bialgebra $(k[\![t]\!], \Delta)$ as the completion (with respect to the natural grading) of the universal enveloping algebra of the free Lie-algebra $\{\alpha t : \alpha \in \mathbb{K}\}$ with $\Delta(t) = 1 \otimes t + t \otimes 1$. Then

$$\begin{aligned} \pi_t : (R\langle\langle \mathsf{X} \rangle\rangle, \Delta) &\longrightarrow (k[\![t]\!], \Delta) \\ x_\alpha &\longmapsto \delta_{\alpha,1} t \end{aligned}$$

is a coalgebra homomorphism. Clearly it descends to $(R\langle\langle \mathsf{Y} \rangle\rangle, \Delta) \to (k[\![t]\!], \Delta)$ so that $\pi_t(\mathtt{y}_{1,1}) = t$ and $\pi_t(\mathtt{y}_{n,\xi}) = 0$ if $(n, \xi) \neq (1, 1)$.

By assumption, Φ is group-like so $\pi_t(\Phi)$ is group-like and therefore $\pi_t(\Phi) = \exp(\alpha t)$ for some $\alpha \in \mathbb{K}$. But the coefficient of x_1 is 0 which is mapped to that of t in $\pi_t(\Phi)$ so we must have $\alpha = 0$. Hence by Eq. (B.3)

$$\pi_t(\Phi_*) = \pi_t(S(\mathtt{y}_{1,1}))\pi_t(\mathbf{q}\pi_{\mathsf{Y}}(\Phi)) = S(t)\pi_{\mathsf{Y}}\pi_t(\Phi) = S(t)$$

since $\mathbf{q}$ fixes all depth 1 elements of $R\langle\langle \mathsf{Y} \rangle\rangle$.

On the other hand, by Prop. B.21 we see that

$$\Phi_* = \exp\Big(\sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} \alpha_{n,\xi} u_{n,\xi} + (\text{terms of depth} > 1) \Big) \tag{B.4}$$

where all the larger depth terms comes from the Lie brackets. But $\mathbb{K}[\![t]\!]$ is commutative so that π_t sends all Lie brackets to 0. Hence

$$\log S(t) = \log \pi_t(\Phi_*) = \sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} \alpha_{n,\xi} \pi_t(u_{n,\xi}).$$

If $\boldsymbol{\mu} \ni \xi = \xi_1\xi_2\dots\xi_d \neq 1$ then there must be $\xi_j \neq 1$ for some j. Hence if the color weight $\xi \neq 1$ then $\pi_t(u_{n,\xi}) = 0$. Thus we get

$$\sum_{n\in\mathbb{N}} \pi_t(u_{n,1}) = \pi_t\Big(\sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} u_{n,\xi} \Big) \overset{\text{defn}}{=\!=} \pi_t \log\Big(\sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} \mathtt{y}_{n,\xi} \Big) = \log(1+t).$$

This implies that $\pi_t(u_{n,1}) = (-1)^{n-1}t^n/n$ and therefore

$$\log S(t) = \log \pi_t(\Phi_*) = \sum_{n\in\mathbb{N}} \frac{(-1)^{n-1}}{n} \alpha_{n,1} t^n.$$

Combining Eq. (B.2) and Eq. (B.4) we see that

$$\Phi_* = 1 + \sum_{n\in\mathbb{N},\xi\in\boldsymbol{\mu}} \alpha_{n,\xi} \mathtt{y}_{n,\xi} + (\text{terms of depth} > 1).$$

Thus by Eq. (B.3) we get $\alpha_{n,1} = (\Phi | x_0^{n-1} x_1)$ if $n \geq 2$. Finally, $\alpha_{1,1} = 0$ by the assumption of Φ_*. This completes the proof of the theorem. □

One of the fundamental results of Lie algebra is the following Poincaré-Birkhoff-Witt Theorem (PBW Theorem). From the definition of universal enveloping algebra we know that there is a canonical map $h : \mathfrak{g} \to U\mathfrak{g}$.

Theorem B.23. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$ and X a totally ordered basis of $\mathfrak{g}$. A canonical monomial over X is a finite sequence $(x_1, x_2, \dots, x_n)$ of elements of X which is non-decreasing in the order $\leq$, that is, $x_1 \leq x_2 \leq \cdots \leq x_n$. Extend h to all canonical monomials as follows: If $(x_1, x_2, \dots, x_n)$ is a canonical monomial, let

$$h(x_1, x_2, \dots, x_n) = h(x_1) \cdot h(x_2) \cdots h(x_n).$$

Then h is injective on the set of the canonical monomials and its range forms a basis of the $\mathbb{K}$-vector space $U\mathfrak{g}$.

The PBW Theorem has the following immediately corollary.

Corollary B.24. If $\mathfrak{g}$ is a Lie algebra over a field then the canonical map $\mathfrak{g} \to U\mathfrak{g}$ is injective. In particular, any Lie algebra over a field is isomorphic to a Lie subalgebra of an associative algebra.

The PBW Theorem has a coalgebra version when the base field contains $\mathbb{Q}$.

Theorem B.25. Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be two sub-Lie algebra of $\mathfrak{g}$ over $\mathbb{K}$ containing $\mathbb{Q}$ such that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ as vector spaces. Let $\widehat{U\mathfrak{g}}$ denote the completion of $U\mathfrak{g}$. Then the map

$$\begin{aligned} \mathrm{PBW}: \ \widehat{U\mathfrak{g}_1} \otimes \widehat{U\mathfrak{g}_2} &\longrightarrow \widehat{U\mathfrak{g}} \\ a \otimes b &\longmapsto ab \end{aligned}$$

is a coalgebra isomorphism.

Keeping the same notation, Thm. B.25 has the following corollary.

Corollary B.26. Let $\pi : \widehat{U\mathfrak{g}} \to \widehat{U\mathfrak{g}}/\mathfrak{g}_1\widehat{U\mathfrak{g}}$ be the projection. Then every diagonal element of $\widehat{U\mathfrak{g}}/\mathfrak{g}_1\widehat{U\mathfrak{g}}$ is the image of some element in $\exp(\mathfrak{g})$ by π. Moreover, if $f_1, f_2 \in \exp(\mathfrak{g})$ with $\pi(f_1) = \pi(f_2)$ then $f_1 = \exp(\lambda) f_2$ for some $\lambda \in \mathfrak{g}_1$.

Proof. Let ϵ be the counit of $\widehat{U\mathfrak{g}_1}$. The map $\epsilon \otimes \mathrm{id} : \widehat{U\mathfrak{g}_1} \otimes \widehat{U\mathfrak{g}_2} \to \otimes\widehat{U\mathfrak{g}_2}$ is a coalgebra homomorphism whose kernel is given by $\mathfrak{g}_1\widehat{U\mathfrak{g}_1} \otimes \widehat{U\mathfrak{g}_2} \simeq \mathfrak{g}_1\widehat{U\mathfrak{g}}$ by Thm. B.25. Thus the coalgebra quotient $\widehat{U\mathfrak{g}}/\mathfrak{g}_1\widehat{U\mathfrak{g}} \simeq \widehat{U\mathfrak{g}_2}$. Therefore we have the following coalgebra commutative diagram:

$$\begin{array}{ccc} \widehat{U\mathfrak{g}_1} \otimes \widehat{U\mathfrak{g}_2} & \xrightarrow{\quad\mathrm{PBW}\quad} & \widehat{U\mathfrak{g}} \\ {\scriptstyle \epsilon \otimes \mathrm{id}} \downarrow \uparrow {\scriptstyle 1 \otimes \mathrm{id}} & & \downarrow {\scriptstyle \pi} \\ \widehat{U\mathfrak{g}_2} & \xrightarrow{\quad\simeq\quad} & \widehat{U\mathfrak{g}}/\mathfrak{g}_1\widehat{U\mathfrak{g}} \end{array}$$

The corollary now follows from the fact that a diagonal element of $\widehat{U\mathfrak{g}}$ must have the form $\exp(\lambda)$ form some $\lambda \in \mathfrak{g}$ by [469, Thm. 3.2]. □

Another deep result from Lie algebra is the following theorem whose proof is too long to be presented here. See the proof of Prop. 5.7 of [164].

Definition B.0.1. A *Lie monomial* in the Lie algebra $\mathbb{K}\langle x, y\rangle$ freely generated by x and y is an expression formed from x and y by using only brackets, but not addition. A *Lie polynomial* over $\mathbb{K}$ is a $\mathbb{K}$-linear combination of Lie monomials.

For instance $[[x, y], [x, [[x, y], x]]]$ and $[x, [x, \ldots, [x, y] \ldots]]$ are both Lie monomials.

Theorem B.27. Let $\mathfrak{g}$ be a Lie algebra containing two generators x and y. Suppose $s(x, y) = s(y, x)$ is a Lie polynomial in $U\mathfrak{g}$ (hence $\Delta s = 1 \otimes s + s \otimes 1$). If

$$s(x, y) - s(x, y + z) + s(x + y - z) - s(y, z) = 0$$

for any generators $x, y, z \in \mathfrak{g}$ then $s = 0$.

Appendix C

Basics of hypergeometric functions

Given any n parameters $b_1, \ldots, b_n$ which are not negative integers, m parameters $a_1, \ldots, a_m$, and a complex variable t, the (generalized) hypergeometric function is defined by

$$
\begin{aligned}
&{}_mF_n(a_1, \ldots, a_m; b_1, \ldots, b_n; t) \\
&\qquad = {}_mF_n\left(\begin{matrix} a_1, \ldots, a_m \\ b_1, \ldots, b_n \end{matrix}; t\right) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_m)_k}{(b_1)_k (b_2)_k \cdots (b_n)_k} \frac{t^k}{k!}, \quad \text{(C.1)}
\end{aligned}
$$

where we have use the Pochhammer symbols defined by Eq. (1.13). In this Appendix, we collect some results on hypergeometric functions which we have used in the book. The interested reader is strongly urged to read old but authoritative classic by Bailey in Ref. [31] and Chapter 15 of the comprehensive handbook [5] for more detailed information.

C.1 Solutions of some ordinary differential equations

The classical hypergeometric function usually refers to the ${}_2F_1$-series. The first influential result concerning these series was discovered by Gauss (see § 15.1.20 of Ref. [5]). He showed that

$$
{}_2F_1\left(\begin{matrix} a, b \\ c \end{matrix}\middle| 1\right) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \ \mathrm{Re}(c) > \mathrm{Re}(a+b), \quad \text{(C.2)}
$$

which is now called the Gauss summation formula/theorem. These series are important in many branches of mathematics, mainly because they satisfy some ubiquitous differential equations. For example, we have the following important result.

Lemma C.28. The hypergeometric function ${}_2F_1(a, b; c; t)$ provides a solution of the following homogeneous differential equation:

$$
t(1-t)\frac{d^2 f}{dt^2} + \Big(c - (1+a+b)t\Big)\frac{df}{dt} - abf = 0. \quad \text{(C.3)}
$$

Proof. This can be proved in a straightforward manner using the definition. □

It turns out that a more general type second order ordinary differential equations, called the Riemann's (or Papperitz) differential equations, can always be solved using ${}_2F_1$-series. These equations have the following general form

$$\frac{d^2f}{dt^2}+\left[\frac{1-\alpha-\alpha'}{t-a}+\frac{1-\beta-\beta'}{t-b}+\frac{1-\gamma-\gamma'}{t-c}\right]\frac{df}{dt}$$
$$+\left[\frac{\alpha\alpha' g(a,b,c)}{t-a}+\frac{\beta\beta' g(b,c,a)}{t-b}+\frac{\gamma\gamma' g(c,a,b)}{t-c}\right]\frac{f}{(t-a)(t-b)(t-c)}=0,$$

where $g(a,b,c)=(a-b)(a-c)$, subject to the condition

$$\alpha+\alpha'+\beta+\beta'+\gamma+\gamma'=1.$$

We notice that these equations have regular singular points at a, b and c. It has the general solution

$$f(t)=\left(\frac{t-a}{t-b}\right)^{\alpha}\left(\frac{t-c}{t-b}\right)^{\gamma}{}_2F_1\left(\begin{matrix}\alpha+\beta+\gamma,\alpha+\beta'+\gamma\\ 1+\alpha-\alpha'\end{matrix}\middle|\,\frac{(t-a)(c-b)}{(t-b)(c-a)}\right).$$

C.2 Transformation formulas

There are many functional relations among the hypergeometric functions which are usually called the transformation formulas. In this section we collect some of those that we have used in this book.

C.2.1 *Linear transformations*

Combining the integral representation

$${}_2F_1\left(\begin{matrix}a,b\\ c\end{matrix}\middle|\,t\right)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 v^{b-1}(1-v)^{c-b-1}(1-vt)^{-a}\,dv$$

where $\mathrm{Re}(c)>\mathrm{Re}(b)>0$, with some Mellin-Barnes type integrals we have

$${}_2F_1\left(\begin{matrix}a,b\\ c\end{matrix}\middle|\,1-t\right)=(1-t)^{c-a-b}{}_2F_1\left(\begin{matrix}c-a,c-b\\ c\end{matrix}\middle|\,t\right),\tag{C.4}$$

and

$${}_2F_1\left(\begin{matrix}a,b\\ c\end{matrix}\middle|\,1-t\right)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}{}_2F_1\left(\begin{matrix}a,b\\ a+b-c+1\end{matrix}\middle|\,t\right)$$
$$+t^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}{}_2F_1\left(\begin{matrix}c-a,c-b\\ a+b-c+1\end{matrix}\middle|\,t\right),\tag{C.5}$$

for all t with $|\arg(z)|<\pi$.

C.2.2 *Some transformations of ${}_3F_2$*

Let $\alpha_{i,j} = \alpha_i + \alpha_j$ and similarly for β. The first transformation formula of the ${}_3F_2$-series is (see § 3.8, Eq. (1) of Ref. [31])

$$\begin{aligned}{}_3F_2\left(\begin{matrix}\alpha_1,\alpha_2,\alpha_3\\ \beta_1,\beta_2\end{matrix}\right) &= \frac{\Gamma(\beta_1)\Gamma(\beta_1-\alpha_{1,2})}{\Gamma(\beta_1-\alpha_1)\Gamma(\beta_1-\alpha_2)}{}_3F_2\left(\begin{matrix}\alpha_1,\alpha_2,\beta_2-\alpha_3\\ \alpha_{1,2}-\beta_1+1,\beta_2\end{matrix}\right)\\ &+\frac{\Gamma(\beta_1)\Gamma(\beta_2)\Gamma(\alpha_{1,2}-\beta_1)\Gamma(\beta_{1,2}-\alpha_{1,2,3})}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\beta_2-\alpha_3)\Gamma(\beta_{1,2}-\alpha_{1,2})}{}_3F_2\left(\begin{matrix}\beta_1-\alpha_1,\beta_1-\alpha_2,\beta_{1,2}-\alpha_{1,2,3}\\ \beta_1-\alpha_{1,2}+1,\beta_{1,2}-\alpha_{1,2}\end{matrix}\right),\end{aligned} \tag{C.6}$$

provided that $\mathrm{Re}(\beta_{1,2}-\alpha_{1,2,3})>0$ and $\mathrm{Re}(\alpha_3-\beta_1+1)>0$. The second one is (see Ex. 7 on p. 98 of Ref. [31])

$${}_3F_2\left(\begin{matrix}\alpha_1,\alpha_2,\alpha_3\\ \beta_1,\beta_2\end{matrix}\right) = \frac{\Gamma(\beta_2)\Gamma(\beta_{1,2}-\alpha_{1,2,3})}{\Gamma(\beta_2-\alpha_3)\Gamma(\beta_{1,2}-\alpha_{1,2})}{}_3F_2\left(\begin{matrix}\beta_1-\alpha_1,\beta_1-\alpha_2,\alpha_3\\ \beta_1,\beta_{1,2}-\alpha_{1,2}\end{matrix}\right), \tag{C.7}$$

provided that $\mathrm{Re}(\beta_{1,2}-\alpha_{1,2,3})>0$ and $\mathrm{Re}(\beta_2-\alpha_3)>0$.

We now use the above two formulas to prove the following result (due to Z. Li in Ref. [380]) used in the proof of Thm. 10.7.3.

Proposition C.29. Let $a,b,c\in\mathbb{C}$ with sufficiently small real parts. Then

$$\begin{aligned}{}_3F_2\left(\begin{matrix}a,b,c\\ a+b,1+c\end{matrix}\right) &= \frac{\Gamma(a+b)\Gamma(1+c)\Gamma(1+c-a-b)}{\Gamma(a)\Gamma(b)\Gamma(1+c-a)\Gamma(1+c-b)}(\psi(1+c-b)-\psi(a)-\psi(b)-\gamma)\\ &-\frac{\Gamma(a+b)\Gamma(1+c)\Gamma(1+c-a-b)}{\Gamma(a)\Gamma(b)\Gamma(1+c-a)\Gamma(1+c-b)}\sum_{n=1}^{\infty}\frac{(a)_n(1-b)_n}{n(1+c-b)_n n!}.\end{aligned}$$

Proof. Taking a parameter ε with $|\varepsilon|$ sufficient small, we have

$${}_3F_2\left(\begin{matrix}a,b,c\\ a+b,1+c\end{matrix}\right) = \lim_{\varepsilon\to 0}{}_3F_2\left(\begin{matrix}a,b,c\\ a+b+\varepsilon,1+c-\varepsilon\end{matrix}\right).$$

Now we consider the series ${}_3F_2\left(\begin{smallmatrix}a,b,c\\ a+b+\varepsilon,1+c-\varepsilon\end{smallmatrix}\right)$. Applying Eq. (C.6), we get

$${}_3F_2\left(\begin{matrix}a,b,c\\ a+b+\varepsilon,1+c-\varepsilon\end{matrix}\right) = \frac{\Gamma(a+b+\varepsilon)\Gamma(\varepsilon)}{\Gamma(a+\varepsilon)\Gamma(b+\varepsilon)}{}_3F_2\left(\begin{matrix}a,b,1-\varepsilon\\ 1-\varepsilon,1+c-\varepsilon\end{matrix}\right) \tag{C.8}$$

$$+\frac{\Gamma(a+b+\varepsilon)\Gamma(1+c-\varepsilon)\Gamma(-\varepsilon)}{\Gamma(a)\Gamma(b)\Gamma(1-\varepsilon)\Gamma(1+c)}{}_3F_2\left(\begin{matrix}a+\varepsilon,b+\varepsilon,1\\ 1+\varepsilon,1+c\end{matrix}\right). \tag{C.9}$$

Applying Eq. (C.2) to the ${}_3F_2$-series in Eq. (C.8), we get

$${}_3F_2\left(\begin{matrix}a,b,1-\varepsilon\\ 1-\varepsilon,1+c-\varepsilon\end{matrix}\right) = {}_2F_1\left(\begin{matrix}a,b\\ 1+c-\varepsilon\end{matrix}\right) = \frac{\Gamma(1+c-\varepsilon)\Gamma(1+c-\varepsilon-a-b)}{\Gamma(1+c-\varepsilon-a)\Gamma(1+c-\varepsilon-b)}$$

for $\mathrm{Re}(1+c-\varepsilon-a-b)>0$. Applying Eq. (C.7) to the ${}_3F_2$-series in Eq. (C.9), we have

$$\begin{aligned}
&{}_3F_2\left(\begin{matrix}a,b,c\\ a+b+\varepsilon,1+c-\varepsilon\end{matrix}\right)=\frac{\Gamma(1+\varepsilon)\Gamma(a+b+\varepsilon)\Gamma(1+c-\varepsilon)\Gamma(1+c-a-b-\varepsilon)}{\varepsilon\Gamma(a+\varepsilon)\Gamma(b+\varepsilon)\Gamma(1+c-a-\varepsilon)\Gamma(1+c-b-\varepsilon)}\\
&-\frac{\Gamma(a+b+\varepsilon)\Gamma(1+c-\varepsilon)\Gamma(1+c-a-b-\varepsilon)}{\varepsilon\Gamma(a)\Gamma(b)\Gamma(1+c-a-\varepsilon)\Gamma(1+c-b)}{}_3F_2\left(\begin{matrix}\varepsilon,1-b,a+\varepsilon\\ 1+\varepsilon,1+c-b\end{matrix}\right)\\
&\quad=\frac{1}{\varepsilon}\left(\frac{\Gamma(1+\varepsilon)\Gamma(a+b+\varepsilon)\Gamma(1+c-\varepsilon)\Gamma(1+c-a-b-\varepsilon)}{\Gamma(a+\varepsilon)\Gamma(b+\varepsilon)\Gamma(1+c-a-\varepsilon)\Gamma(1+c-b-\varepsilon)}\right.\\
&\qquad\left.-\frac{\Gamma(a+b+\varepsilon)\Gamma(1+c-\varepsilon)\Gamma(1+c-a-b-\varepsilon)}{\Gamma(a)\Gamma(b)\Gamma(1+c-a-\varepsilon)\Gamma(1+c-b)}\right)\\
&-\frac{\Gamma(a+b+\varepsilon)\Gamma(1+c-\varepsilon)\Gamma(1+c-a-b-\varepsilon)}{\Gamma(a)\Gamma(b)\Gamma(1+c-a-\varepsilon)\Gamma(1+c-b)}\sum_{n=1}^{\infty}\frac{(a+\varepsilon)_n(1-b)_n}{(n+\varepsilon)(1+c-b)_n n!}.
\end{aligned}$$

by expanding the ${}_3F_2$-series on the right. Using L'Hôpital's rule and the fact that $\psi(1)=-\gamma$ we can now finish the proof of the proposition by taking $\varepsilon\to 0$. □

The following result will be used in the proof of Thm. 5.6.4.

Proposition C.30. Suppose $\mathrm{Re}(\alpha-2\beta)>-2$. Then

$$\begin{aligned}
&{}_3F_2\left(\begin{matrix}1,\alpha,\beta\\ 1+\alpha,2+\alpha-\beta\end{matrix};1\right)\\
&=\frac{\alpha(1+\alpha-\beta)}{\beta-1}\Big[\psi(\alpha)-\psi(2+\alpha-2\beta)-\psi\Big(\frac{\alpha+1}{2}\Big)+\psi\Big(\frac{\alpha+3}{2}-\beta\Big)\Big].
\end{aligned}\tag{C.10}$$

Proof. Taking $m=1$ and $k=l=0$ in (4) of Ref. [390], we get

$$\begin{aligned}
{}_3F_2\left(\begin{matrix}1,\alpha,\beta\\ 1+\alpha,2+\alpha-\beta\end{matrix};1\right)&=\frac{-\alpha(1+\alpha-\beta)}{2(\beta-1)}[\psi((\alpha+1)/2)-\psi(\alpha/2)\\
&\quad+\psi(\alpha/2-\beta+1)-\psi((\alpha+1)/2-\beta+1)].
\end{aligned}$$

An application of the duplication relation 1.6 yields the proof of the proposition. □

Appendix D

Sample computer codes

In this appendix, we provide the pseudo-codes for a few key computer programs we have used in the book. The readers who like to explore by themselves should be able to modify these easily using their favorite computer algebra systems.

D.1 Double shuffle relations

Both the shuffle and the stuffle relations can be computed using the recursive definitions.

(i). **The stuffle program.**

```
stuffle:=proc(s, t)
    d =depth(s); l =depth(t);
    if d = 0 or l = 0 then
        RETURN([s,t]);
    else
        s' = [s_2, ..., s_d]; t' = [t_2, ..., t_l];
        a =stuffle(s', t); b =stuffle(s, t'); c =stuffle(s', t');
        RETURN([s_1, a] + [t_1, b] + [s_1 + t_1, c]);
    end if;
end proc:
```

(ii). **The shuffle program.**

We can modify program (i) by removing the term **c** from the output.

(iii). Other variations of the double shuffle relations such as the q-analogs can be obtained by further modifying the above programs.

D.2 Multiple harmonic sums

We use the formula

$$H_n(s_1,\ldots,s_d)=\sum_{k=d}^{n}\frac{H_k(s_2,\ldots,s_d)}{k^{s_1}}$$

(i). **The MHS program.**

```
H:=proc(t, n)
    d =depth(s);
    S_0 := 1; # S_k stands for the partial sums.
    for i to d do
        S_i := 0; # Initialize the partial sums
    end do;
    for k from 0 to n − d do
        for i from 1 to d do
                        S_i := S_i + S_{i−1}/(k + i)^{t_i};
        end do;
    end do;
    RETURN(S_d);
end proc:
```

(ii). **The alternating MHS program and the q-analogs.**

These can be written down by modifying (i).

D.3 Finite multiple zeta values

In Chapter 8, we found many linear relations between FMZVs and FESs experimentally and formed a few conjectures based on these numerical results. The follow pseudo-code gives the idea to detach such relations. The key idea is to use the Chinese Remainder's Theorem and the PSQL algorithm [199].

Suppose $\mathbf{s}_1,\ldots,\mathbf{s}_k$ are k different compositions of positive integers (or signed numbers for AMHS). We want to know if there are rational numbers $c_1,\ldots,c_k$ such that

$$c_1H_{p-1}(\mathbf{s}_1)+\cdots+c_kH_{p-1}(\mathbf{s}_k)\equiv 0 \pmod p$$

for all sufficiently large primes p.

Take a large number M, say $M=1000$. Suppose $p_1,\ldots,p_M$ are M sufficiently large primes. By the Chinese Remainder's Theorem there are positive integers $A_1,\ldots,A_k$ (independent of $p_1,\ldots,p_M$) such that

$$H_{p_j-1}(\mathbf{s}_1)\equiv A_1,\ldots,H_{p_j-1}(\mathbf{s}_k)\equiv A_k \pmod{p_j}$$

for all $j = 1, \ldots, M$. This can be found by the command `chrem` in Maple and `ChineseRemainder` in Mathematica. Assume $P = \prod_{j=1}^{M}$. Then we can use PSQL algorithm to find rational numbers $c_0', c_1', \ldots, c_k' \in \mathbb{Q}$ such that

$$c_0' P + c_1' A_1 + \cdots + c_k' A_k = 0.$$

This algorithm has been implement in Maple simply as `PQSL`. Mathematica users can consult [147]. Or, one can use the `LLL` algorithm in PARI/GP.

Now, if all of $c_0', c_1', \ldots, c_k'$ stay relatively small and stabilize for all integers M beyond a certain large value, then we can assume that we have found a linear relation

$$c_1 \zeta_{\mathcal{A}_1}(\mathbf{s}_1) + \cdots + c_k \zeta_{\mathcal{A}_1}(\mathbf{s}_k) = 0.$$

The interested reader should have no trouble modifying the above to treat higher superbities, i.e., the super congruences.

Appendix E

Tables of special values

E.1 Euler sums of lower weights

We list all the $\mathbb{Q}$-linear independent relations obtained by FDS and RDS between Euler sums of weight < 6. Most of the computations are carried out by Maple. We have checked the consistency of these relations with the many known ones and verified numerically all these identities by EZ-face [75] with error smaller than 10^{-50}.

All weight four Euler sums are $\mathbb{Z}$-linear combinations of $A = \zeta(\bar{2}, 1, 1)$, $B = \zeta(\bar{2}, 2)$, $C = \zeta(\bar{1}, 2, 1)$, $D = \zeta(\bar{1}, 1, 2)$, and $E = \zeta(\bar{1}, 1, 1, 1)$:

$$
\begin{aligned}
\zeta(4) &= 64A + 16B,\\
\zeta(\bar{4}) &= -56A - 14B,\\
\zeta(3,1) &= 16A + 4B,\\
\zeta(3,\bar{1}) &= 118A + 19B + 14C,\\
\zeta(2,2) &= 48A + 12B,\\
\zeta(\bar{3},1) &= 10A + 2B,\\
\zeta(\bar{3},\bar{1}) &= -140A - 24B - 14C,\\
\zeta(2,\bar{2}) &= -24A - 7B,\\
\zeta(\bar{2},\bar{2}) &= -12A - 3B,\\
\zeta(\bar{1},3) &= -38A - 5B - 6C,\\
\zeta(\bar{1},\bar{3}) &= 58A + 8B + 8C,\\
\zeta(2,1,1) &= 64A + 16B,\\
\zeta(2,1,\bar{1}) &= 16A + 2B + 6C + 3D,\\
\zeta(2,\bar{1},1) &= 22A + 3B + C - 3D,\\
\zeta(2,\bar{1},\bar{1}) &= 100A + 13B + 9C - 6D,\\
\zeta(\bar{2},1,\bar{1}) &= 91A + 14B + 8C - 3D,\\
\zeta(\bar{2},\bar{1},1) &= -161A - 26B - 15C + 3D,\\
\zeta(\bar{2},\bar{1},\bar{1}) &= -101A - 14B - 9C + 6D,\\
\zeta(\bar{1},2,\bar{1}) &= -102A - 14B - 8C + 6D,
\end{aligned}
$$

$$
\begin{aligned}
\zeta(\bar{1},\bar{2},1) &= 69A+11B+8C,\\
\zeta(\bar{1},\bar{2},\bar{1}) &= 63A+8B+3C-6D,\\
\zeta(\bar{1},1,\bar{2}) &= 21A+3B+C-2D,\\
\zeta(\bar{1},\bar{1},\bar{2}) &= A+2B\phantom{+C}+D,\\
\zeta(\bar{1},1,1,\bar{1}) &= A\phantom{+2B+C}+E,\\
\zeta(\bar{1},1,\bar{1},1) &= 11A+2B+C\phantom{+D}+E,\\
\zeta(\bar{1},1,\bar{1},\bar{1}) &= \phantom{11A+2B+}C\phantom{+D}+E,\\
\zeta(\bar{1},\bar{1},1,1) &= -83A-16B-5C+D+E,\\
\zeta(\bar{1},\bar{1},1,\bar{1}) &= -38A-5B-5C+D+E,\\
\zeta(\bar{1},\bar{1},\bar{1},1) &= \phantom{38A-5B-5C+}D+E,\\
\zeta(\bar{1},\bar{1},\bar{1},\bar{1}) &= A+B\phantom{+C}+D+E.
\end{aligned}
$$

All weight five Euler sums are $\mathbb{Z}$-linear combinations of the following

$$
\begin{aligned}
&A=\zeta(\bar{1},\bar{1},\bar{1},2),\quad B=\zeta(\bar{2},1,\bar{1},\bar{1}),C=\zeta(\bar{1},1,\bar{1},\bar{2}),\quad D=\zeta(\bar{2},1,1,1),\\
&E=\zeta(\bar{1},\bar{1},\bar{1},1,1),F=\zeta(2,2,\bar{1}),\quad G=\zeta(\bar{1},1,\bar{1},1,\bar{1}),H=\zeta(\bar{1},\bar{1},\bar{1},\bar{1},\bar{1}).
\end{aligned}
$$

$$
\begin{aligned}
\zeta(5) &= -13504A+1856B-1344C+26880D-18752E-640F-31552G+50304H,\\
\zeta(\bar{5}) &= 12660A-1740B+1260C-25200D+17580E+600F+29580G-47160H,\\
\zeta(4,1) &= -9808A+1344B-944C+19632D-13648E-464F-22848G+36496H,\\
\zeta(4,\bar{1}) &= -14918A+2044B-1434C+29862D-20758E-704F-34748G+55506H,\\
\zeta(\bar{4},1) &= 3638A-498B+346C-7296D+5066E+172F+8466G-13532H,\\
\zeta(\bar{4},\bar{1}) &= 19862A-2722B+1914C-39744D+27634E+938F+46274G-73908H,\\
\zeta(3,2) &= 22672A-3104B+2160C-45456D+31568E+1072F+52768G-84336H,\\
\zeta(3,\bar{2}) &= 4562A-626B+446C-9108D+6342E+216F+10642G-16984H,\\
\zeta(\bar{3},2) &= -6552A+898B-632C+13110D-9116E-310F-15266G+24382H,\\
\zeta(\bar{3},\bar{2}) &= -17848A+2444B-1704C+35772D-24848E-844F-41548G+66396H,\\
\zeta(2,3) &= -26368A+3616B-2560C+52704D-36672E-1248F-61472G+98144H,\\
\zeta(2,\bar{3}) &= 6792A-934B+680C-13506D+9428E+322F+15878G-25306H,\\
\zeta(\bar{2},\bar{3}) &= 24902A-3412B+2394C-49854D+34654E+1178F+58004G-92658H,\\
\zeta(\bar{2},3) &= -8622A+1182B-834C+17244D-11994E-408F-20094G+32088H,\\
\zeta(\bar{1},4) &= 5266A-720B+494C-10582D+7338E+248F+12240G-19578H,\\
\zeta(\bar{1},\bar{4}) &= -8990A+1230B-850C+18044D-12522E-424F-20910G+33432H,\\
\zeta(3,1,1) &= -9808A+1344B-944C+19632D-13648E-464F-22848G+36496H,\\
\zeta(3,1,\bar{1}) &= -5314A+725B-500C+10677D-7402E-250F-12339G+19741H,\\
\zeta(3,\bar{1},1) &= -2257A+312B-225C+4489D-3137E-108F-5290G+8427H,\\
\zeta(3,\bar{1},\bar{1}) &= -7299A+1005B-713C+14566D-10151E-347F-17057G+27208H,\\
\zeta(\bar{3},1,1) &= 4482A-614B+430C-8974D+6238E+212F+10438G-16676H,
\end{aligned}
$$

$$
\begin{aligned}
\zeta(\bar{3},1,\bar{1}) &= 9570A-1308B+908C-19204D+13328E+452F+22250G-35578H,\\
\zeta(\bar{3},\bar{1},1) &= 12462A-1710B+1204C-24924D+17338E+590F+29056G-46394H,\\
\zeta(\bar{3},\bar{1},\bar{1}) &= -4288A+582B-396C+8646D-5978E-201F-9922G+15900H,\\
\zeta(2,2,1) &= 22672A-3104B+2160C-45456D+31568E+1072F+52768G-84336H,\\
\zeta(2,\bar{2},1) &= 3025A-414B+287C-6065D+4213E+143F+7038G-11251H,\\
\zeta(2,\bar{2},\bar{1}) &= 6421A-881B+627C-12818D+8927E+303F+14977G-23904H,\\
\zeta(2,1,2) &= -26368A+3616B-2560C+52704D-36672E-1248F-61472G+98144H,\\
\zeta(2,\bar{1},2) &= 2206A-302B+210C-4428D+3074E+104F+5134G-8208H,\\
\zeta(2,1,\bar{2}) &= 7958A-1093B+786C-15861D+11056E+377F+18581G-29637H,\\
\zeta(2,\bar{1},\bar{2}) &= 23513A-3221B+2255C-47094D+32727E+1113F+54757G-87484H,\\
\zeta(\bar{2},2,1) &= -12813A+1755B-1227C+25664D-17835E-606F-29835G+47670H,
\end{aligned}
$$

$$
\begin{aligned}
\zeta(\bar{2},2,\bar{1}) &= -20468A+2804B-1964C+40988D-28488E-968F-47668G+76156H,\\
\zeta(\bar{2},\bar{2},1) &= -5477A+750B-523C+10977D-7625E-259F-12750G+20375H,\\
\zeta(\bar{2},\bar{2},\bar{1}) &= 12308A-1686B+1180C-24654D+17132E+582F+28662G-45794H,\\
\zeta(\bar{2},1,2) &= 12622A-1729B+1210C-25281D+17568E+597F+29393G-46961H,\\
\zeta(\bar{2},1,\bar{2}) &= -3065A+420B-295C+6135D-4265E-145F-7140G+11405H,\\
\zeta(\bar{2},\bar{1},2) &= -14047A+1923B-1337C+28170D-19561E-665F-32691G+52252H,\\
\zeta(\bar{2},\bar{1},\bar{2}) &= -9411A+1290B-909C+18831D-13095E-445F-21930G+35025H,\\
\zeta(\bar{1},3,1) &= 123A-17B+13C-242D+171E+6F+289G-460H,\\
\zeta(\bar{1},3,\bar{1}) &= -11820A+1614B-1120C+23726D-16460E-557F-27466G+43926H,\\
\zeta(\bar{1},\bar{3},1) &= 6380A-874B+612C-12776D+8880E+302F+14858G-23738H,\\
\zeta(\bar{1},\bar{3},\bar{1}) &= 12610A-1722B+1194C-25312D+17560E+594F+29302G-46862H,\\
\zeta(\bar{1},2,2) &= -190A+26B-18C+384D-266E-9F-442G+708H,\\
\zeta(\bar{1},2,\bar{2}) &= 13726A-1880B+1314C-27494D+19106E+649F+31960G-51066H,\\
\zeta(\bar{1},\bar{2},2) &= -13631A+1867B-1305C+27302D-18973E-644F-31739G+50712H,\\
\zeta(\bar{1},\bar{2},\bar{2}) &= 599A-82B+57C-1203D+835E+28F+1394G-2229H,\\
\zeta(\bar{1},1,3) &= -3186A+435B-300C+6399D-4438E-150F-7401G+11839H,\\
\zeta(\bar{1},1,\bar{3}) &= -2732A+376B-268C+5448D-3798E-130F-6384G+10182H,\\
\zeta(\bar{1},\bar{1},3) &= 20431A-2799B+1969C-40888D+28427E+966F+47591G-76018H,\\
\zeta(\bar{1},\bar{1},\bar{3}) &= -7808A+1070B-758C+15608D-10858E-369F-18196G+29054H,\\
\zeta(2,1,1,1) &= -13504A+1856B-1344C+26880D-18752E-640F-31552G+50304H,\\
\zeta(2,1,1,\bar{1}) &= -11109A+1518B-1044C+22320D-15477E-523F-25812G+41289H,\\
\zeta(2,1,\bar{1},1) &= 1174A-158B+101C-2395D+1642E+54F+2691G-4333H,\\
\zeta(2,1,\bar{1},\bar{1}) &= 14927A-2044B+1431C-29899D+20773E+705F+34745G-55518H,\\
\zeta(2,\bar{1},1,1) &= -2712A+371B-258C+5439D-3777E-128F-6306G+10083H,\\
\zeta(2,\bar{1},1,\bar{1}) &= -14828A+2030B-1419C+29709D-20641E-701F-34517G+55158H,\\
\zeta(2,\bar{1},\bar{1},1) &= -7120A+977B-681C+14262D-9911E-337F-16585G+26496H,
\end{aligned}
$$

$$
\begin{aligned}
\zeta(2,\bar{1},\bar{1},\bar{1}) &= 11204A - 1534B + 1074C - 22440D + 15595E + 530F + 26096G - 41691H,\\
\zeta(\bar{2},\bar{1},1,1) &= 8717A - 1197B + 847C - 17415D + 12122E + 412F + 20334G - 32456H,\\
\zeta(\bar{2},\bar{1},1,\bar{1}) &= -8511A + 1162B - 806C + 17085D - 11852E - 401F - 19775G + 31627H,\\
\zeta(\bar{2},\bar{1},\bar{1},1) &= 3432A - 470B + 327C - 6882D + 4779E + 162F + 7980G - 12759H,\\
\zeta(\bar{2},\bar{1},\bar{1},\bar{1}) &= -652A + 89B - 66C + 1296D - 905E - 31F - 1531G + 2436H,\\
\zeta(\bar{2},1,\bar{1},1) &= 8659A - 1183B + 822C - 17376D + 12059E + 409F + 20134G - 32193H,\\
\zeta(\bar{2},1,1,\bar{1}) &= 3571A - 490B + 344C - 7145D + 4969E + 169F + 8322G - 13291H,
\end{aligned}
$$

$$
\begin{aligned}
\zeta(\bar{1},2,1,1) &= 190A - 26B + 18C - 384D + 265E + 9F + 442G - 707H,\\
\zeta(\bar{1},2,1,\bar{1}) &= 190A - 25B + 18C - 385D + 265E + 9F + 442G - 707H,\\
\zeta(\bar{1},2,\bar{1},1) &= -27776A + 3801B - 2654C + 55667D - 38668E - 1313F - 64648G + 103316H,\\
\zeta(\bar{1},2,\bar{1},\bar{1}) &= -19006A + 2604B - 1828C + 38048D - 26452E - 900F - 44288G + 70740H,\\
\zeta(\bar{1},\bar{2},1,1) &= -2407A + 330B - 233C + 4812D - 3347E - 113F - 5610G + 8957H,\\
\zeta(\bar{1},\bar{2},1,\bar{1}) &= -202A + 32B - 34C + 353D - 274E - 12F - 533G + 807H,\\
\zeta(\bar{1},\bar{2},\bar{1},1) &= 6507A - 893B + 631C - 13009D + 9054E + 309F + 15184G - 24238H,\\
\zeta(\bar{1},\bar{2},\bar{1},\bar{1}) &= 31628A - 4333B + 3038C - 63328D + 44021E + 1497F + 73681G - 117702H,\\
\zeta(\bar{1},1,2,1) &= -122A + 17B - 13C + 242D - 170E - 6F - 288G + 458H,\\
\zeta(\bar{1},1,2,\bar{1}) &= 3310A - 453B + 314C - 6641D + 4609E + 156F + 7692G - 12301H,\\
\zeta(\bar{1},1,\bar{2},\bar{1}) &= 6195A - 850B + 600C - 12383D + 8619E + 294F + 14454G - 23073H,\\
\zeta(\bar{1},1,\bar{2},1) &= 2888A - 394B + 272C - 5803D + 4023E + 136F + 6706G - 10729H,\\
\zeta(\bar{1},\bar{1},2,1) &= -7433A + 1019B - 711C + 14888D - 10348E - 352F - 17315G + 27663H,\\
\zeta(\bar{1},\bar{1},2,\bar{1}) &= 6793A - 932B + 658C - 13586D + 9453E + 322F + 15848G - 25301H,\\
\zeta(\bar{1},\bar{1},\bar{2},1) &= -313A + 43B - 30C + 626D - 437E - 15F - 730G + 1167H,\\
\zeta(\bar{1},\bar{1},\bar{2},\bar{1}) &= -18914A + 2592B - 1822C + 37855D - 26321E - 895F - 44073G + 70394H,\\
\zeta(\bar{1},1,1,2) &= 191A - 26B + 18C - 384D + 267E + 9F + 442G - 709H,\\
\zeta(\bar{1},1,1,\bar{2}) &= -2521A + 345B - 240C + 5054D - 3510E - 119F - 5864G + 9374H,\\
\zeta(\bar{1},1,\bar{1},2) &= 13126A - 1798B + 1257C - 26291D + 18271E + 621F + 30567G - 48838H,\\
\zeta(\bar{1},\bar{1},1,2) &= 13312A - 1826B + 1295C - 26595D + 18512E + 630F + 31043G - 49555H,\\
\zeta(\bar{1},\bar{1},1,\bar{2}) &= -4812A + 661B - 475C + 9593D - 6687E - 228F - 11237G + 17924H,\\
\zeta(\bar{1},\bar{1},\bar{1},\bar{2}) &= -13127A + 1798B - 1258C + 26291D - 18271E - 621F - 30565G + 48836H,\\
\zeta(\bar{1},1,1,1,1) &= -191A + 26B - 18C - 442G + 384D - 266E - 9F + 709H,\\
\zeta(\bar{1},1,1,1,\bar{1}) &= -191A + 26B - 18C + 385D - 266E - 9F - 442G + 709H,\\
\zeta(\bar{1},1,1,\bar{1},1) &= 4481A - 614B + 430C - 8973D + 6237E + 212F + 10438G - 16674H,\\
\zeta(\bar{1},1,1,\bar{1},\bar{1}) &= -A - E + 2H,\\
\zeta(\bar{1},1,\bar{1},1,1) &= -4693A + 643B - 451C + 9395D - 6531E - 222F - 10930G + 17462H,\\
\zeta(\bar{1},1,\bar{1},\bar{1},1) &= -313A + 43B - 31C - 730G + 626D - 436E - 15F + 1167H,
\end{aligned}
$$

$$\begin{aligned}
\zeta(\bar{1},1,\bar{1},\bar{1},\bar{1}) &= -13126A + 1798B - 1258C + 26291D - 18271E - 621F - 30565G + 48837H,\\
\zeta(\bar{1},\bar{1},1,1,1) &= 7496A - 1031B + 747C - 14915D + 10408E + 355F + 17522G - 27929H,\\
\zeta(\bar{1},\bar{1},1,1,\bar{1}) &= 2081A - 285B + 194C - 4183D + 2901E + 98F + 4840G - 7740H,\\
\zeta(\bar{1},\bar{1},1,\bar{1},1) &= -3308A + 452B - 313C + 6641D - 4607E - 156F - 7689G + 12297H,\\
\zeta(\bar{1},\bar{1},1,\bar{1},\bar{1}) &= -12121A + 1660B - 1164C + 24269D - 16868E - 573F - 28225G + 45094H,\\
\zeta(\bar{1},\bar{1},\bar{1},1,\bar{1}) &= 12622A - 1729B + 1210C - 25280D + 17569E + 597F + 29393G - 46961H,\\
\zeta(\bar{1},\bar{1},\bar{1},\bar{1},1) &= 10552A - 1445B + 1008C - 21144D + 14690E + 499F + 24565G - 39254H.
\end{aligned}$$

E.2 Finite multiple zeta values

We now list the FMZVs in lower weight. Notice that by reversal relation in fact we only need to list non-decreasing entries using lexicographic ordering. Moreover, by duality relation given by Eq. (8.29) we only need to list entries of depth no greater than $(w+1)/2$. In odd weight all palindromic entries vanish by reversal relation. Furthermore, we omit all depth one and two elements, and depth 3 odd weight elements because we may use Thm. 8.5.12 and Thm. 8.5.13 to compute them fairly quickly.

$\zeta_{\mathcal{A}}(1,1,1,2)$	$-5\beta_5$	$\zeta_{\mathcal{A}}(1,1,2,1)$	$10\beta_5$	$\zeta_{\mathcal{A}}(1,2,3)$	$3\beta_3^2$
$\zeta_{\mathcal{A}}(1,3,2)$	$-\frac{9}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(1,4,1)$	$3\beta_3^2$	$\zeta_{\mathcal{A}}(1,1,4)$	$-\frac{3}{2}\beta_3^2$
$\zeta_{\mathcal{A}}(2,1,3)$	$\frac{3}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(2,3,1)$	$-\frac{9}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(3,1,2)$	$\frac{3}{2}\beta_3^2$
$\zeta_{\mathcal{A}}(3,2,1)$	$3\beta_3^2$	$\zeta_{\mathcal{A}}(4,1,1)$	$-\frac{3}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(1,1,1,3)$	$-\frac{3}{2}\beta_3^2$
$\zeta_{\mathcal{A}}(1,1,2,2)$	$3\beta_3^2$	$\zeta_{\mathcal{A}}(1,1,3,1)$	$\frac{3}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(3,1,1,1)$	$-\frac{3}{2}\beta_3^2$
$\zeta_{\mathcal{A}}(1,3,1,1)$	$\frac{3}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(2,2,1,1)$	$3\beta_3^2$	$\zeta_{\mathcal{A}}(1,2,2,1)$	$-\beta_3^2$
$\zeta_{\mathcal{A}}(1,2,1,2)$	$\frac{9}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(2,1,2,1)$	$\frac{9}{2}\beta_3^2$	$\zeta_{\mathcal{A}}(2,1,1,2)$	$-6\beta_3^2$

Table E.1: FMZVs of superbity 1 in weight 5 and weight 6.

$\zeta_{\mathcal{A}}(1,1,1,4)$	$-27\beta_7'$	$\zeta_{\mathcal{A}}(1,1,2,3)$	$69\beta_7'$	$\zeta_{\mathcal{A}}(1,1,3,2)$	$-27\beta_7'$
$\zeta_{\mathcal{A}}(1,1,4,1)$	$33\beta_7'$	$\zeta_{\mathcal{A}}(1,2,1,3)$	$-27\beta_7'$	$\zeta_{\mathcal{A}}(1,2,2,2)$	$-27\beta_7'$
$\zeta_{\mathcal{A}}(1,2,3,1)$	$-63\beta_7'$	$\zeta_{\mathcal{A}}(1,3,1,2)$	$-9\beta_7'$	$\zeta_{\mathcal{A}}(2,1,1,3)$	$33\beta_7'$
$\zeta_{\mathcal{A}}(2,1,2,2)$	$-63\beta_7'$				

Table E.2: FMZVs of superbity 1 in weight 7.

$\zeta_{\mathcal{A}}(1,2,5)$	$E+F$	$\zeta_{\mathcal{A}}(1,3,4)$	$-\frac{25}{2}E-\frac{9}{2}F$
$\zeta_{\mathcal{A}}(1,4,3)$	$10E+4F$	$\zeta_{\mathcal{A}}(1,5,2)$	$\frac{5}{2}E-\frac{1}{2}F$
$\zeta_{\mathcal{A}}(1,6,1)$	$-2E$	$\zeta_{\mathcal{A}}(2,1,5)$	$-\frac{7}{2}E-\frac{1}{2}F$
$\zeta_{\mathcal{A}}(2,2,4)$	$10E+2F$	$\zeta_{\mathcal{A}}(2,3,3)$	$10E+2F$
$\zeta_{\mathcal{A}}(2,4,2)$	$-20E-4F$	$\zeta_{\mathcal{A}}(3,1,4)$	$\frac{5}{2}E+\frac{1}{2}F$
$\zeta_{\mathcal{A}}(3,2,3)$	$-20E-4F$	$\zeta_{\mathcal{A}}(4,1,3)$	$\frac{5}{2}E+\frac{1}{2}F$
$\zeta_{\mathcal{A}}(1,1,1,5)$	$\frac{3}{4}E-\frac{1}{4}F$	$\zeta_{\mathcal{A}}(1,1,2,4)$	$\frac{3}{4}E+\frac{5}{4}F$
$\zeta_{\mathcal{A}}(1,1,3,3)$	$-\frac{21}{2}E-3F$	$\zeta_{\mathcal{A}}(1,1,4,2)$	$\frac{33}{4}E+\frac{9}{4}F$
$\zeta_{\mathcal{A}}(1,1,5,1)$	$-\frac{3}{4}E+\frac{1}{4}F$	$\zeta_{\mathcal{A}}(1,3,1,3)$	0
$\zeta_{\mathcal{A}}(1,3,2,2)$	$\frac{3}{2}F$	$\zeta_{\mathcal{A}}(1,3,3,1)$	$\frac{27}{2}E+\frac{9}{2}F$
$\zeta_{\mathcal{A}}(2,2,1,3)$	$-\frac{9}{4}F-\frac{45}{4}E$	$\zeta_{\mathcal{A}}(3,1,1,3)$	$\frac{15}{2}E+\frac{3}{2}F$
$\zeta_{\mathcal{A}}(1,2,1,4)$	$-\frac{9}{2}E-F$	$\zeta_{\mathcal{A}}(1,2,2,3)$	$\frac{9}{2}E-\frac{3}{2}F$
$\zeta_{\mathcal{A}}(1,2,3,2)$	$-\frac{27}{4}E-\frac{1}{4}F$		

Table E.3: FMZVs of superbity 1 in weight 8, depth 3 and 4. Here, $E = \zeta_{\mathcal{A}}(1,1,6)$ and $F = \beta_3\beta_5$.

$\zeta_{\mathcal{A}}(1,1,1,6)$	$-\frac{1889}{72}\beta_9-\frac{1}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,2,5)$	$\frac{5551}{72}\beta_9+\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,3,4)$	$-\frac{4697}{72}\beta_9-\frac{1}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,4,3)$	$\frac{2609}{36}\beta_9+\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,5,2)$	$-\frac{821}{36}\beta_9-\frac{5}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,6,1)$	$\frac{521}{24}\beta_9+\frac{1}{18}\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,1,5)$	$-\frac{209}{72}\beta_9+\beta_3^3$	$\zeta_{\mathcal{A}}(1,2,2,4)$	$-\frac{3857}{72}\beta_9-\frac{2}{27}\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,3,3)$	$-\frac{1591}{36}\beta_9-\frac{2}{27}\beta_3^3$	$\zeta_{\mathcal{A}}(1,2,4,2)$	$-\frac{401}{36}\beta_9+\frac{5}{27}\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,5,1)$	$-\frac{1399}{24}\beta_9-\frac{1}{9}\beta_3^3$	$\zeta_{\mathcal{A}}(1,3,1,4)$	$-\frac{127}{12}\beta_9-\frac{1}{18}\beta_3^3$
$\zeta_{\mathcal{A}}(1,3,2,3)$	$\frac{292}{3}\beta_9+\frac{2}{9}\beta_3^3$	$\zeta_{\mathcal{A}}(1,3,3,2)$	$-\beta_9-\frac{1}{6}\beta_3^3$
$\zeta_{\mathcal{A}}(1,3,4,1)$	$\frac{1567}{18}\beta_9+\frac{5}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,4,1,3)$	$-\frac{251}{9}\beta_9+\frac{1}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,4,2,2)$	$-\frac{122}{3}\beta_9-\frac{1}{9}\beta_3^3$	$\zeta_{\mathcal{A}}(1,5,1,2)$	$\frac{167}{72}\beta_9+\frac{2}{27}\beta_3^3$
$\zeta_{\mathcal{A}}(2,1,1,5)$	$\frac{749}{18}\beta_9+\frac{1}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(2,1,2,4)$	$-\frac{1813}{18}\beta_9-\beta_3^3$
$\zeta_{\mathcal{A}}(2,1,3,3)$	$\frac{1579}{36}\beta_9+\frac{1}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(2,1,4,2)$	$-\frac{2071}{36}\beta_9+\beta_3^3$
$\zeta_{\mathcal{A}}(2,2,1,4)$	$\frac{357}{8}\beta_9$	$\zeta_{\mathcal{A}}(2,2,2,3)$	$54\beta_9$
$\zeta_{\mathcal{A}}(2,2,3,2)$	$84\beta_9$	$\zeta_{\mathcal{A}}(2,3,1,3)$	$\frac{43}{3}\beta_9-\frac{1}{9}\beta_3^3$
$\zeta_{\mathcal{A}}(2,3,2,2)$	$-84\beta_9$	$\zeta_{\mathcal{A}}(3,1,1,4)$	$-\frac{3647}{72}\beta_9+\beta_3^3$
$\zeta_{\mathcal{A}}(3,1,2,3)$	$\frac{1567}{18}\beta_9-\frac{2}{27}\beta_3^3$		

Table E.4: FMZVs of superbity 1 in weight 9 and depth 4.

$\zeta_{\mathcal{A}}(1,1,1,1,5)$	$-\frac{3965}{144}\beta_9-\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,1,2,4)$	$\frac{11107}{144}\beta_9+2\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,1,3,3)$	$-\frac{2333}{144}\beta_9+\frac{1}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,1,4,2)$	$\frac{1709}{48}\beta_9-\frac{3}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,1,5,1)$	$\frac{1481}{72}\beta_9+2\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,2,1,4)$	$-\frac{2333}{144}\beta_9+\frac{1}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,2,2,3)$	$-\frac{13133}{144}\beta_9-4\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,2,3,2)$	$-\frac{1891}{48}\beta_9+3\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,2,4,1)$	$-\frac{5239}{72}\beta_9-4\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,3,1,3)$	$-\frac{85}{144}\beta_9+\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,3,2,2)$	$\frac{3485}{48}\beta_9+3\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,3,3,1)$	$\frac{1015}{24}\beta_9-\frac{3}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,1,4,1,2)$	$\frac{5185}{144}\beta_9+\frac{5}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,1,4,2,1)$	$-\frac{1865}{36}\beta_9-\frac{5}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,1,1,4)$	$\frac{1709}{48}\beta_9+3\beta_3^3$	$\zeta_{\mathcal{A}}(1,2,1,2,3)$	$-\frac{1891}{48}\beta_9-6\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,1,3,2)$	$\frac{223}{16}\beta_9+\frac{9}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,2,1,4,1)$	$-\frac{131}{8}\beta_9-\frac{9}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,2,1,3)$	$\frac{3485}{48}\beta_9-\frac{3}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,2,2,2,2)$	$\frac{255}{16}\beta_9$
$\zeta_{\mathcal{A}}(1,2,2,3,1)$	$-\frac{765}{8}\beta_9+\frac{1}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,2,3,1,2)$	$\frac{595}{16}\beta_9-\frac{1}{3}\beta_3^3$
$\zeta_{\mathcal{A}}(1,2,4,1,1)$	$\frac{1865}{36}\beta_9+\frac{5}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,3,1,1,3)$	$-\frac{5185}{144}\beta_9-2\beta_3^3$
$\zeta_{\mathcal{A}}(1,3,1,2,2)$	$\frac{595}{16}\beta_9+\frac{9}{2}\beta_3^3$	$\zeta_{\mathcal{A}}(1,3,2,1,2)$	$-\frac{935}{16}\beta_9+\frac{9}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(1,4,1,1,2)$	$\frac{4633}{144}\beta_9-4\beta_3^3$	$\zeta_{\mathcal{A}}(2,1,1,1,4)$	$\frac{1481}{72}\beta_9-\frac{5}{2}\beta_3^3$
$\zeta_{\mathcal{A}}(2,1,1,2,3)$	$-\frac{5239}{72}\beta_9+5\beta_3^3$	$\zeta_{\mathcal{A}}(2,1,1,3,2)$	$-\frac{131}{8}\beta_9-\frac{1}{3}\beta_3^3$
$\zeta_{\mathcal{A}}(2,1,2,1,3)$	$\frac{1015}{24}\beta_9+3\beta_3^3$	$\zeta_{\mathcal{A}}(2,1,2,2,2)$	$\frac{765}{8}\beta_9$
$\zeta_{\mathcal{A}}(2,1,3,2,1)$	$-\frac{595}{16}\beta_9+\frac{1}{3}\beta_3^3$	$\zeta_{\mathcal{A}}(2,2,1,1,3)$	$-\frac{1865}{36}\beta_9+2\beta_3^3$

Table E.5: FMZVs of superbity 1 in weight 9 and depth 5.

$\zeta_{\mathcal{A}_2}(1,4)$	$3p\beta_3^2-2z311$	$\zeta_{\mathcal{A}_2}(4,1)$	$-3p\beta_3^2+2z311$
$\zeta_{\mathcal{A}_2}(1,1,3)$	$-z311$	$\zeta_{\mathcal{A}_2}(2,3)$	$-3p\beta_3^2+4z311$
$\zeta_{\mathcal{A}_2}(1,2,2)$	$-\frac{3}{2}p\beta_3^2+3z311$	$\zeta_{\mathcal{A}_2}(2,1,2)$	$-3p\beta_3^2$
$\zeta_{\mathcal{A}_2}(3,2)$	$3p\beta_3^2-4z311$	$\zeta_{\mathcal{A}_2}(2,2,1)$	$\frac{9}{2}p\beta_3^2-3z311$
$\zeta_{\mathcal{A}_2}(1,3,1)$	0		

Table E.6: FMZVs of superbity 2 in weight 5, where $z311=\zeta_{\mathcal{A}_2}(3,1,1)$.

$\zeta_{\mathcal{A}_2}(6)$	$6\beta_7$	$\zeta_{\mathcal{A}_2}(2,1,3)$	$18\beta_7 - z411$
$\zeta_{\mathcal{A}_2}(5,1)$	$-10\beta_7$	$\zeta_{\mathcal{A}_2}(1,2,3)$	$-\frac{11}{4}\beta_7 - 2z411$
$\zeta_{\mathcal{A}_2}(1,5)$	$4\beta_7$	$\zeta_{\mathcal{A}_2}(1,3,2)$	$-\frac{65}{4}\beta_7 + 3z411$
$\zeta_{\mathcal{A}_2}(2,4)$	$-10\beta_7$	$\zeta_{\mathcal{A}_2}(2,3,1)$	$-\frac{9}{4}\beta_7 + 3z411$
$\zeta_{\mathcal{A}_2}(3,3)$	$-3\beta_7$	$\zeta_{\mathcal{A}_2}(3,1,2)$	$11\beta_7 - z411$
$\zeta_{\mathcal{A}_2}(4,2)$	$4\beta_7$	$\zeta_{\mathcal{A}_2}(3,2,1)$	$\frac{17}{4}\beta_7 - 2z411$
$\zeta_{\mathcal{A}_2}(2,2,2)$	$2\beta_7$	$\zeta_{\mathcal{A}_2}(1,4,1)$	$6\beta_7 - 2z411$
$\zeta_{\mathcal{A}_2}(1,1,4)$	$z411$		

Table E.7: FMZVs of superbity 2 in weight 6, where $z411 = \zeta_{\mathcal{A}_2}(4,1,1)$.

Bibliography

[1] N.H. Abel, Note sur la fonction $\psi x = x + \frac{x^2}{2^2} + \frac{x^3}{3^2} + \cdots + \frac{x^n}{n^2} + \cdot$, In Sylow, L. Lie, S. Œuvres complètes de Niels Henrik Abel–Nouvelle édition, Tome II (in French), 1881. Christiania [Oslo]: Grøndahl & Søn. pp. 189–193. (This was written in 1826 and published posthumously.)

[2] J. Ablinger and J. Blümlein, Harmonic sums, polylogarithms, special numbers, and their generalizations, in: Computer Algebra in Quantum Field Theory, C. Schneider, J. Bluemlein (ed.), Texts and Monographs in Symbolic Computation, Springer, 2013, pp. 1–32. arXiv:1304.7071

[3] J. Ablinger, J. Blümlein and C. Schneider, Harmonic sums and polylogarithms generated by cyclotomic polynomials, *J. Math. Phys.* **52** (10) (2011), p. 102301.

[4] M. Abouzahra and L. Lewin, Ladders in the trans-Kummer region, [374, p. 49–68].

[5] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, (1972) Dover, New York.

[6] T. Agoh and K. Dilcher, Convolution identities and lacunary recurrences for Bernoulli numbers, *J. Number Theory* **124** (2007), pp. 105–122.

[7] S. Akiyama, S. Egami and Y. Tanigawa, Analytic continuation of multiple zeta-functions and their values at non-positive integers, *Acta Arithmetica* **98** (2001), pp. 107–116.

[8] S. Akiyama and H. Ishikawa, An analytic continuation of multiple L-functions and related zeta-functions, in: Analytic Number Theory (Developments in Mathematics), C. Jia and K. Matsumoto eds., Springer, 2002, pp. 1–16.

[9] G.E. Andrews, The Theory of Partitions, Addison-Wesley Publishing Co., Reading, Mass.-LondonAmsterdam, 1976, *Encyclopedia Math. Appl.* Vol. 2.

[10] G.E. Andrews, q-analogs of the binomial coefficient congruences of Babbage, Wolstenholme and Glaisher, **204** (1999), pp. 15–25.

[11] T. Aoki, Y. Kombu and Y. Ohno, A generating function for sums of multiple zeta values and its applications, *Proc. Amer. Math. Soc.* **136**(2) (2007), pp. 387–395.

[12] T. Aoki and Y. Ohno, Sum relations for multiple zeta values and connection formulas for the Gauss hypergeometric functions, *Publ. Res. Inst. Math. Sci.* **41** (2005), pp. 329–337.
[13] T. Aoki, Y. Ohno and N. Wakabayashi, On generating functions of multiple zeta values and generalized hypergeometric functions, *Manuscripta Math.* **134** (2011), pp. 139–155.
[14] K. Aomoto, Addition theorem of Abel type for hyperlogarithms, *Nagoya Math. J.* **88** (1982), pp. 55–71.
[15] K. Aomoto, Special values of hyperlogarithms and linear difference schemes, *Illinois J. of Math.* **34**(2) (1990), pp. 191–216.
[16] R. Apéry, Irrationalité de $\zeta(2)$ et $\zeta(3)$, *Astérisque* **61** (1979), pp. 11-13. Coppo
[17] T.M. Apostol, and T.H. Vu, Dirichlet series related to the Riemann zeta function, *J. Number Theory* **19** (1984), pp. 85–102.
[18] T. Arakawa and M. Kaneko, On multiple L-values, *J. Math. Soc. Japan* **56** (2004), pp. 967–991.
[19] T. Arakawa and M. Kaneko, Multiple zeta values, poly-Bernoulli numbers, and related zeta functions, *Nagoya Math. J.* **153** (1999), pp. 189–209.
[20] F. V. Atkinson, The mean value of the Riemann zeta-function, *Acta Math.* **81** (1949), pp. 353–376.
[21] H. Bachmann, Multiple zeta-Werte und de Verbindung zu Modulformen durch Multiple Eisensteinreihen, master thesis, 2012, Hamburg Univesity.
[22] H. Bachmann, The algebra of bi-brackets and regularized multiple Eisenstein series. arXiv:1504.08138.
[23] H. Bachmann and U. Kühn, The algebra of multiple divisor functions and applications to multiple zeta values. arXiv:1309.3920.
[24] H. Bachmann and U. Kühn, A short note on a conjecture of Okounkov about a q-analogue of multiple zeta values. arXiv:1407.6796.
[25] H. Bachmann and K. Tasaka, The double shuffle relations for multiple Eisenstein series. arXiv:1502.03408.
[26] D.F. Bailey, Two p^3 variations of Lucas' theorem, *J. Number Theory* **35**(2) (1990), pp. 208–215.
[27] D.H. Bailey, J.M. Borwein, D. Broadhurst and W. Zudilin, Experimental mathematics and mathematical physics, *Contemp. Math.* **517** (2010), pp. 41–58.
[28] D.H. Bailey, J.M. Borwein, N. Calkin, R. Girgensohn, R, Luke and V. Moll. Experimental Mathematics in Action. A K Peters, Natick, MA, 2007.
[29] D.H. Bailey and D.J. Broadhurst, Parallel integer relation detection: Techniques and appli- cations, *Math. Comp.* **70** (2000), pp. 1719–1736.
[30] D.H. Bailey, D. Broadhurst, A seventeenth-order polylogarithm ladder. arXiv:math/9906134.
[31] W. N. Bailey, Generalited hypergeometric series, Cambridge Uniyersity Press, Cambridge, 1935.
[32] K. Ball, T. Rivoal, Irrationalité d'une infinité de valeurs de la fonction zêta aux entiers impairs, *Inv. Math.* **146**(1) (2001), pp. 193–207.
[33] D. Bar-Natan, On associators and the Grothendieck-Teichmüller group I,

Selecta Math. New Series **4** (1998), pp. 183–212.

[34] R. Barrieri, J.A. Mignaco and E. Remiddi, Electron Form Factors up to Fourth Order. II, *Il Nuovo Cimento* **11**(4) (1972), pp. 865–916.

[35] H. Bass, *Generators and relations for cyclotomic units, Nagoya Math. J.* **27**(2) (1966), pp. 401–407.

[36] A. Basu and T. Apostol, A new method of investigating Euler sums, *Ramanujan J.* **4** (2000), pp. 397–419.

[37] S. Baumard, Aspects modulaires et elliptiques des relations entre multizêtas, Number Theory. Université Pierre et Marie Curie - Paris VI, 2014. French. ¡NNT:2014PA066100¿.¡tel-01017022v2¿

[38] S. Baumard and L. Schneps, Period polynomial relations between double zeta values, *Ramanujan J.* **32** (1) (2013), pp. 83–100.

[39] G. Baxter, An analytic problem whose solution follows from a simple algebraic identity, *Pacific J. Math.* **10** (1960), pp. 731–742.

[40] M. Bayat, A generalization of Wolstenholme's Theorem, *Amer. Math. Monthly* **104**(6) (1997), pp. 557–560.

[41] A.A. Beilinson and P. Deligne, Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs, in: *Proc. Sym. Pure Math.* **55** part 2, Amer. Math. Soc. (1994), pp. 97–121.

[42] A.A. Beilinson, A. B. Goncharov, V.V. Schechtman and A. N. Varchenko, Aomoto dilogarithms, mixed Hodge structures and motivic cohomology of pairs of triangles in the plane, in: Grothendieck Festschrift II, *Prog. in Math.* **87**, Birkhäuser, Boston, 1991, pp. 78–131.

[43] A. A. Beilinson, A. N. Varchenko, A. B. Goncharov and V. V. Schechtman, Projective geometry and algebraic K-theory, *Leningrad Math. J.* **2** (1991), pp. 523–576.

[44] A.A. Beilinson, R. MacPherson and V. Schechtman, Notes on motivic cohomology, *Duke Math. J.* **54** (1987), pp. 679–710.

[45] A. Besser, Finite and p-adic polylogarithms, *Compos. Math.* **130**(2) (2002), pp, 215–223.

[46] A. Besser and H. Furusho, The double shuffle relations for p-adic multiple zeta values, Amer. Math. Soc. *Contemp. Math.* **416** (2006), pp. 9–29.

[47] I. Bierenbaum and S. Weinzierl, The massless two-loop two-point function, *Eur. Phys. J.* **C32** (2003), pp. 67–78.

[48] M. Bigotte, , G. Jacob, N.E. Oussous and M. Petitot, Generating power series of coloured polylogarithm functions and Drinfel'd associator. In: Computer mathematics (Chiang Mai, 2000), *Lecture Notes Ser. Comput.* **8**, pp. 39–48. World Sci. Publ., River Edge, NJ, 2000.

[49] M. Bigotte, , G. Jacob, N.E. Oussous and M. Petitot, Lyndon words and shuffle algebras for generating the coloured multiple zeta values relations tables, *Theoretical Computer Sci.* **273** (1-2) (2002), pp. 271–282.

[50] S. Bloch, Application of the dilogarithm function in algebraic K-theory and algebraic geometry, in: *Proc. Intl Symp. Alg. Geom.* (Kyoto Univ., Kyoto, 1977), pp. 103–114, Kinokuniya Book Store, Tokyo, 1978.

[51] S. Bloch, Algebraic cycles and higher K-theory, *Adv. Math.* **140** (1994), pp. 557–605.

[52] S. Bloch, Lectures on mixed motives given at Santa Cruz, 1995, available online `http://www.math.uchicago.edu/~bloch/publications.html`

[53] S. Bloch, Higher Regulators, Algebraic K-Theory, and Zeta Functions of Elliptic Curves, CRM Monograph Series, Vol. **11**, Amer. Math. Soc. Centre de Recherches Math., 2000.

[54] S. Bloch and D. Kreimer, Feynman amplitudes and Landau singularities for one-loop graphs, *Comm. Number Thy. Phys.* **4** (4) (2010), pp. 709–753.

[55] S. Bloch and P. Vanhove, The elliptic dilogarithm for the sunset graph, *J. Number Theory* **148** (3) (2015), pp. 328–364.

[56] J. Blümlein, Harmonic sums and Mellin transforms, *Nucl. Phys. Proc. Suppl.* **79** (1999), pp. 166–168.

[57] J. Blümlein, Reduction of multiple harmonic sums and harmonic polylogarithms, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **534** (1-2) (2004), pp. 279–283.

[58] J. Blümlein, Relations between harmonic sums in massless QCD calculations, *Few-Body Systems* **36** (2005), pp. 29–34.

[59] J. Blümlein, D.J. Broadhurst, and J.A.M. Vermaseren, The multiple zeta value data mine, *Comput. Phys. Commun.* **181** (2010), pp. 582–625,

[60] J. Blümlein and S. Kurth, Harmonic sums and Mellin transforms up to two-loop order, *Phys. Rev.* **D60** (1999) 014018.

[61] C. Bogner and F. Brown, Symbolic integration and multiple polylogarithms. arXiv:1209.6524.

[62] C. Bogner and M, Lüders, Multiple polylogarithms and linearly reducible Feynman graphs. arXiv:1302.6215.

[63] S. Borowka and G. Heinrich, Massive non-planar two-loop four-point integrals with SecDec 2.1, *Comput. Phys. Commun.* **184** (2013), pp. 2552. arXiv: hep-ph 1303.1157.

[64] J.M. Borwein and D.H. Bailey, Mathematics by Experiment, 2nd edi., A K Peters Ltd., Natick, MA, 2008.

[65] J.M. Borwein and P. Borwein, Pi and the AGM, John Wiley and Sons, Toronto and New York, 1987 (paperback 1998).

[66] D. Borwein, J.M. Borwein and R. Girgensohn, Explicit evaluation of Euler sums, *Proc. Edin. Math. Soc.* **38** (1995), pp. 273–294.

[67] J.M. Borwein and D.M. Bradley, Thirty-two Goldbach variations, *Intl. J. Number Theory* **2**(1) (2006), pp. 65–103.

[68] 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**(2) (1997), R5.

[69] J.M. Borwein, D.M. Bradley, D.J. Broadhurst and P. Lisoněk, Combinatorial aspects of multiple zeta values, *Electron. J. Combin.* **5** (1998), R38.

[70] J.M. Borwein, D.M. Bradley, D.J. Broadhurst, and P. Lisonek, Special values of multidimensional polylogarithms, *Trans. Amer. Math. Soc.* **353** (2001), pp. 907–941.

[71] J.M. Borwein and D.J. Broadhurst, Determinations of rational Dedekind-zeta invariants of hyperbolic manifolds and Feynman knots and links.

arXiv:hep-th/9811173.

[72] J.M. Borwein, D.J. Broadhurst, and J. Kamnitzer, Central binomial sums, multiple Clausen values, and zeta values, *Experiment. Math.* **10** (1) (2001), pp. 25–34.

[73] P. Borwein, S. Choi and B. Rooney, eds., The Riemann Hypothesis: A Resource for the Afficionado and Virtuoso Alike, CMS Books in Mathematics, New York: Springer, 2008, doi:10.1007/978-0-387-72126-2.

[74] J.M. Borwein and R. Girgensohn, Evaluation of triple Euler sums, *Electron. J. Combin.* **3** (1996), ♯ R23.

[75] J.M. Borwein, P. Lisonek and P. Irvine, An interface for evaluation of Euler sums, available online at `http://oldweb.cecm.sfu.ca/cgi-bin/EZFace/zetaform.cgi`

[76] O. Bouillot, The algebra of Hurwitz multizeta functions, *Comptes Rendus Math.* **352** (11)(2014), pp. 865–869.

[77] D.M. Bowman and D.F. Bailey, Multiple polylogarithms: a brief survey, in: Proc. Conf. on q-Series with Applications to Combinatorics, Number Theory and Physics, B.C. Berndt and K. Ono eds., *Contemp. Math.* **291** (2001), pp. 71–92.

[78] D. Bowman and D.M. Bradley, The algebra and combinatorics of shuffles and multiple zeta values, *J. Combin. Theory Ser. A* **97**(1) (2002), pp. 43–63

[79] K. Boyadzhiev, Evaluation of Euler-Zagier sums, *Internat. J. Math. Math. Sci.* **27** (2001), pp. 407-412.

[80] K. Boyadzhiev, Consecutive evaluation of Euler sums, *Internat. J. Math. Math. Sci.* **29** (2002), pp. 555-561.

[81] D.W. Boyd, A p-adic study of the partial sums of the harmonic series, *Experimental Math.* **4**(3) (1994), pp. 287–302.

[82] D. M. Bradley, Multiple q-zeta values, *J. Algebra* **283** (2005), pp. 752–798.

[83] D. M. Bradley, A q-analog of Euler's decomposition formula for the double zeta function, *Int. J. Math. and Math. Sci.*, **2005** (21) (2005), pp. 3453–3458.

[84] D. M. Bradley, Duality for finite multiple harmonic q-series, *Discrete Math.* **300** (2005), pp. 44–56.

[85] D.J. Broadhurst, Massless scalar Feynman diagrams: five loops and beyond, Open Univ. preprint OUT-4102-18 (1985).

[86] D.J. Broadhurst, On the enumeration of irreducible k-fold Euler sums and their roles in knot theory and field theory. arXiv:hep-th/9604128

[87] D.J. Broadhurst, Conjectured enumeration of irreducible multiple zeta values, from knots and Feynman diagrams. arXiv:hep-th/9612012.

[88] D.J. Broadhurst, Massive 3-loop Feynman diagrams reducible to SC^* primitives of algebras of the sixth root of unity, *European Phys. J. C (Fields)* **8** (1999), pp. 311–333.

[89] D.J. Broadhurst, Multiple Deligne values: a data mine with empirically tamed denominators. arXiv:1409.7204.

[90] D.J. Broadhurst, Multiple Landen values and the tribonacci numbers. arXiv:1504.05303.

[91] D.J. Broadhurst, J.A. Gracey, and D. Kreimer, Beyond the triangle and

uniqueness relations: Non-zeta counterterms at large N from positive knots, *Z. Phys. C* **75** (1997), pp. 559–574.
[92] D.J. Broadhurst and D. Kreimer, Knots and numbers in ϕ^4 Theory to 7 loops and beyond, *Int. J. Mod. Phys.* **C6** (1995), pp. 519–524.
[93] D.J. Broadhurst and D. Kreimer, Association of multiple zeta values with positive knots via Feynman diagrams up to 9 loops *Phys. Lett. B* **393**(34) (1997), pp. 403–412.
[94] D.J. Broadhurst and O. Schnetz, Algebraic geometry informs perturbative quantum field theory. arXiv:1409.5570.
[95] A. Brochier, A KohnoDrinfeld Theorem for the monodromy of cyclotomic KZ connections, *Comm. Math. Phys.* **311**(1) 2012, pp. 55–96.
[96] J. Broedel, C.R. Mafra, N. Matthes and O. Schlotterer, Elliptic multiple zeta values and one-loop superstring amplitudes. arXiv:1412.5535
[97] J. Broedel, N. Matthes and O. Schlotterer, Relations between elliptic multiple zeta values and a special derivation algebra, arXiv:1507.02254
[98] J. Broedel, O. Schlotterer and S. Stieberger, Polylogarithms, multiple zeta values and superstring amplitudes. arXiv:1304.7267
[99] F. Brown, Single-valued hyperlogarithms and unipotent differential equations, preprint, http://www.ihes.fr/ brown/RHpaper5.pdf.
[100] F. Brown, Périodes des espaces des modules $M_{0,n}$ et valeurs zêtas multiples, *C. R. Acad. Sci. Paris, Ser. I* **342** (2006), pp. 949-954.
[101] F. Brown, Multiple zeta values and and periods of moduli spaces $M_{0,n}$, *Ann. Sci. Ecole Norm. S.* **42**(3) (2009), pp. 371–489.
[102] F. Brown, Single-valued multiple polylogarithms in one variable, *C. R. Acad. Sci. Paris, Ser. I* **338** (2004), pp. 527–532.
[103] F. Brown, The massless higher-loop two-point function, *Commun. Math. Phys.* **287** (2009), pp. 925-958.
[104] F. Brown, Mixed Tate motives over Spec($\mathbb{Z}$), *Ann. Math.* **175**(2) (2012), pp. 949–976.
[105] F. Brown, On the decomposition of motivic multiple zeta values, in: Galois-Teichmüller theory and Arithmetic Geometry, *Adv. Studies Pure Math.* Vol. **63**, The Math. Soc. of Japan, 2012, pp. 3158.
[106] F. Brown, Depth-graded motivic multiple zeta values. arXiv:1301.3053.
[107] F. Brown, Irrationality proofs for zeta values, moduli spaces and dinner parties, arXiv:1412.6508.
[108] F. Brown and A. Levin, Multiple elliptic polylogarithms. arXiv:1110.6917.
[109] F. Brown and O. Schnetz, Single-valued multiple polylogarithms and a proof of the zigzag conjecture *J. Number Theory* **148** (3) (2015), pp. 478–506.
[110] L. Carlitz, q-Bernoulli numbers and polynomials, *Duke Math. J.* **15** (1948), pp. 987–1000.
[111] L. Carlitz, A degenerate Staudt-Clausen theorem, *Arch. Math.* **7** (1956), pp. 28–33.
[112] B.C. Carlson, New proof of the addition theorem for Gegenbauer polynomials, *SIAM J. Math. Anal.*, **2** (3) (1971), pp. 347–351.
[113] S. Carr, H. Gangl and L. Schneps, On the Broadhurst-Kreimer generating

series for multiple zeta values, to appear in the Proceedings of the Madrid-ICMAT conference on Multizetas, 2015.

[114] P. Cartier, Fonctions polylogarithms, nombres polyzetas et groupes prounipotents, Séminaire Bourbaki, 53ème année, 2000-2001, n° 885, *Asterisque* **282** (2002), pp. 137–173.

[115] J. Castillo Medina, K. Ebrahimi-Fard, D. Manchon, On Euler's decomposition formula for qMZVs, *Ramanujan J.* (2015), DOI 10.1007/s11139-014-9638-8.

[116] J. Castillo Medina, K. Ebrahimi-Fard, D. Manchon, Unfolding the double shuffle structure of q-multiple zeta values. arXiv:1310.1330

[117] J.L. Cathelineau, Remarques sur les différentielles des polylogarithmes uniformes, *Ann. Inst. Fourier (Grenoble)* **46**(5) (1996), pp. 1327-1347.

[118] E. Cattani, Introduction to variations of hodge structure summer school on hodge theory, Ictp, June 2010, available online: `www.math.umass.edu/~cattani/ICTP/cattani_vhs.pdf`.

[119] M. Chamberland and K. Dilcher, Divisibility properties of a class of binomial sums , *J. of Number Theory* **120** (2006), pp. 349–371.

[120] C.-Y. Chang, Linear independence of monomials of multizeta values in positive characteristic, *Compositio Math.* **150** (2014), pp. 1789–1808.

[121] C.-Y. Chang, M. A. Papanikolas and J. Yu, An effective criterion for Eulerian multizeta values in positive characteristic. arXiv:1411.0124.

[122] S. Charlton, $\zeta(2^m, 1, 2^m, 3^n, 2^m)/\pi^{4n+2m(2n+1)}$ is rational, *J. Number Theory* **148** (3) (2015), pp. 463-477.

[123] A. Chatzistamatiou, On integrality of p-adic iterated integrals. arXiv:1501.05760.

[124] F. Chavez and C. Duhr, Three-mass triangle integrals and single-valued polylogarithms, arXiv:1209:2722.

[125] H.-J. Chen, On shuffle of double zeta values over $\mathbb{F}_q[t]$, *J. Number Theory* **148** (3) (2015), pp. 153–163.

[126] K.-T. Chen, Iterated integrals and exponential homomorphisms, *Proc. London Math. Soc.* **4**(3) (1954), pp. 502–512.

[127] K.-T. Chen, Algebras of iterated path integrals and fundamental groups, *Trans. Amer. Math. Soc.* **156** (1971), pp. 359–379.

[128] K.-T. Chen, Iterated path integrals, *Bull. Amer. Math. Soc.* **83** (1977), pp. 831–879.

[129] K.-T. Chen, R.H. Fox and R.C. Lyndon, Free differential calculus. IV. The quotient groups of the lower central series, *Ann. Math.* **68**(1) (1958), pp. 81–95,

[130] K.-W. Chen, Sums of products of generalized Bernoulli polynomials, *Pacific J. Math.* **208**(1) (2003), pp. 39–52,

[131] K-W. Chen and M. Eie, Explicit evaluations of extended Euler sums, *J. Number Theory* **117** (2006), pp. 31–52.

[132] Y.-H. Chen, Multiple zeta values and application to the Lacunary recurrence formulas of Bernoulli numbers, *J. Physics: Conference Series* **96** (2008), paper 012212.

[133] K.G. Chetyrkin, A.L. Kataev and F.V. Tkachov, New approach to evalua-

tion of multiloop Feynman integrals: The Gegenbauer polynomial x-space technique, *Nucl. Phys.* **B174** (1980), pp. 345–377.
[134] K.G. Chetyrkin and F.V. Tkachov, Integration by parts: The algorithm to calculate β- functions in 4 Loops, *Nucl. Phys.* **B192** (1981), pp. 159–204.
[135] Y. Choie and K. Matsumoto, Functional equations for double series of Euler type with coefficients. arXiv:1306.0987.
[136] W. Chu, Symmetric functions and the Riemann zeta series, *Indian J. Pure Appl. Math.* **31** (2000), pp. 1677–1689.
[137] H. Cohen, Number Theory. Vol. **II**. Analytic and Modern Tools, Graduate Texts in Mathematics **240**, Springer, New York, 2007.
[138] H. Cohen, L. Lewin and D. Zagier, A sixteenth-order polylogarithm ladde, *Experimental Math.* **1**(1) (1992), pp. 25–34.
[139] G. Combariza, A few conjectures about the multiple zeta values. arXiv:1207.1735.
[140] A. Connes and D. Kreimer, Renormalization in quantum field theory and the RiemannHilbert problem. I. The Hopf algebra structure of graphs and the main theorem, *Comm. Math. Phys.* **210** (1) (2000), pp. 249–273.
[141] A. Connes and D. Kreimer, Renormalization in quantum field theory and the RiemannHilbert problem. II. The β-function, diffeomorphisms and the renormalization group, *Comm. Math. Phys.* **216** (1) (2001), pp. 215–241.
[142] M. Conrad, Basen vo Moduln mit Anwendung auf Kresenheiten und Stickelbergerelemente, Ph. D. thesis, University of Saarbrücken, 1997. `www.perisic.com/marc/PUBL/thesis.ps`
[143] M. Conrad, Construction of bases for the group of cyclotomic units, *J. Number Theory* **81** (2000), pp. 1–15.
[144] M-A. Coppo and B. Candelpergher, The Arakawa-Kaneko zeta function, *Ramanujan J.* **22** (2010), pp. 153–162.
[145] C. Costermans H.N. Minh, Noncommutative algebra, multiple harmonic sums and applications in discrete probability, *J. Symbolic Comput.* **44**(7) (2009), pp. 801-817.
[146] O. Costin and S. Garoufalidis, Resurgence of the fractional polylogarithms *Math. Res. Lett.* **16**(5) (2009), pp. 817826
[147] R. Crandall, Topics in Advanced Scientific Computation, Springer, 1st ed. 1996.
[148] R. Crandall and J. Buhler, On the evaluation of Euler sums, *Experiment. Math.* **3** (1994), pp. 275-285.
[149] C. Curry, K. Ebrahimi-Fard, S. Malham and A. Wiese, Levy processes and quasi-shuffle algebras, to appear in: Stochastics: An International Journal of Probability and Stochastic Processes. arXiv:1311.1683.
[150] A.I. Davydychev and R. Delbourgo, A geometrical angle on Feynman integrals, *J. Math. Phys.* **39** (1998), pp. 4299–4334. arXiv:hep-th/9709216.
[151] P. Deligne, Equations differentielles á points singuliers réguliers, *Lecture Notes in Math.*, Vol. **163**, Berlin-Heidelberg-New York, Springer-Verlag, 1970.
[152] P. Deligne, Letter to Spencer Bloch, April 3, 1984.
[153] P. Deligne, Le groupe fondamental de la droite projective moins trois points,

in: Galois groups over $\mathbb{Q}$, *Publ. MSRI* **16** (1989), Springer-Verlag, pp. 79–297.

[154] P. Deligne, Analytic continuation of multiple zeta functions, unpublished manuscript.

[155] P. Deligne, Le groupe fondamental de la $\mathbb{G}_m - \boldsymbol{\mu}_N$, pour $N = 2$, 3, 4, 6 ou 8, *Publ. Math. Inst. Hautes Etudes Sci.* **112** (2010), pp. 101–141.

[156] P. Deligne and A. Goncharov, Groupes fondamentaux motiviques de Tate mixte, *Ann. Sci. Ecole Norm. S.* **38**(1) (2005), pp. 1–56.

[157] A. Devoto, D.W. Duke, Table of integrals and formulae for Feynman diagram calculations, *Riv. Nuovo Cim.* **7**(6) (1984), pp. 1–39.

[158] F. Diamond and J. Shurman, A First Course in Modular Forms, *Graduate Texts in Mathematics*, Vol. **228**, Springer-Verlag, first ed., 2007.

[159] K. Dicher, Sums of products of Bernoulli numbers, *J. number theory* **60** (1996), pp. 23–41.

[160] K. Dilcher, Determinant expressions for q-harmonic congruences and degenerate Bernoulli numbers, *Electron. J. Combin.* **15**(1) (2008), R63.

[161] K. Dilcher, Kh. Hessami Pilehrood, and T. Hessami Pilehrood, On q-analogues of double Euler sums, *J. Math. Anal. Appl.* **420** (2014), pp. 979–988.

[162] L. Dixon, C. Duhr and J. Pennington, Single-valued harmonic polylogarithms and the multi-Regge limit, arXiv:1207.0186.

[163] V.G. Drinfeld, Quasi-Hopf algebras, *Leningrad Math J.* **1** (1989), pp. 1419–1457.

[164] V.G. Drinfeld, On quasitriangular quasi-Hopf algebras and a group closely connected with $Gal(\mathbb{Q}/Q)$, *Leningrad Math. J.* **2** (1991), pp. 829–860.

[165] J.M. Drummond, J.M. Henn and J. Trnka, New differential equations for on-shell loop integrals, *JHEP* **1104** (2011), 083.

[166] J.M. Drummond, Generalised ladders and single-valued polylogs, arXiv:hep-th/1207.3824.

[167] C. Dupont, The combinatorial Hopf algebra of motivic dissection polylogarithms, *Adv. Math.* **264** (10) (2014), pp. 646–699.

[168] K. Ebrahimi-Fard and L. Guo, Rota-Baxter algebras and multiple zeta values, *Integers* **8** (2008), A4.

[169] K. Ebrahimi-Fard, L. Guo and D. Kreimer, Integrable renormalization I: the ladder case, *J. Math. Phys.* **45** (2004), pp. 3758–3769.

[170] K. Ebrahimi-Fard, L. Guo and D. Kreimer, Integrable Renormalization II: the General case, *Ann. Henri Poincaré* **6** (2005), pp. 369–395.

[171] K. Ebrahimi-Fard and D. Manchon, A Magnus- and Fer-type formula in dendriform algebras, *Found. Comput. Math.* **9** (2009), pp. 295–316.

[172] K. Ebrahimi-Fard and D. Manchon, Twisted dendriform algebras and the pre-Lie Magnus expansion, *J. Pure Appl. Alg.* **215** (2011), pp. 2615–2627.

[173] K. Ebrahimi-Fard, D. Manchon, and J. Singer, 'The Hopf algebra of (q)multiple polylogarithms with nonpositive arguments,' preprint 1503.02977[NT].

[174] J. Ecalle, Les fonctions résurgentes, Vol. 2: Les fonctions résurgentes appliquées à l'itération. *Publ. Math. Orsay* 81.06 (1981), # 283 pp. Available

online: www.math.u-psud.fr/ ecalle/publi.html
[175] J. Ecalle, ARI/GARI, la dimorphie et l'arithmétique des multizetas: un premier bilan. *J. Théorie Nombres Bordeaux***15** (2003), pp. 411–478.
[176] R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne, The Analytic S-Matrix, 1st edition, Cambridge University Press, 1966.
[177] S. Egami and H. Matsumoto, Asymptotic expansions of multiple zeta functions and power mean values of Hurwitz zeta functions, *J. London Math. Soc.* **66**(2) (2002), pp. 41–60.
[178] R. Ehrenborg, On posets and Hopf algebras, *Adv. Math.* **119** (1996), pp, 1–25.
[179] M. Eie, A note on Bernoulli numbers and Shintani generalized Bernoulli polynomials, *Trans. Amer. Math. Soc.* **348** (1996), pp. 1117–1136.
[180] M. Eie, Topics in Number Theory, *Monographs in Number Theory* Vol. **2**, World Scientific, 2009.
[181] M. Eie and K.F. Lai, On Bernoulli identities and applications, *Rev. Mat. Iberoamericana* **14**(1) (1998), pp. 167–213.
[182] M. Eie, W. Liaw, Y.L. Ong, A restricted sum formula among multiple zeta values, *J. Number Theory* **129**(4) (2009), pp. 908–921.
[183] M. Eie and C.-S. Wei, Generalizations of Euler decomposition and their applications, *J. Number Theory* **133**(8) (2013), pp. 2475–2495.
[184] M. Eie, F-Y. Yang, and Y.L. Ong, Some alternating double sum formulae of multiple zeta values, *Comput. Appl. Math.* **29** (2010), pp. 375–391.
[185] P. Elbaz-Vincent and H. Gangl, On Poly(ana)logs I, *Compos. Math.* **130**(2) (2002), pp. 161–214,
[186] V. Ennola, On relations between cyclotomic units, *J. Number Theory* **4** (1972), pp. 236–247.
[187] B. Enriquez, On some relations between generalized associators. arXiv:math/0609440.
[188] B. Enriquez, Quasi-reflection algebras and cyclotomic associators, *Selecta Math. New Series* **13** (2007), pp. 391-463.
[189] B. Enriquez, Analogues elliptiques des nombres multizétas. arXiv:1301.3042
[190] B. Enriquez, Elliptic associators, *Selecta Math. New Series* **20** (2014), pp. 491–584.
[191] B. Enriquez and H. Furusho, Mixed pentagon, octagon and Broadhurst duality equation, *J. Pure Appl. Alg.* **216** (2012), pp. 982-995.
[192] B. Enriquez and F. Gavarini, A Formula for the logarithm of the KZ associator, *SIGMA* **2** (2006), #80.
[193] T. Ernst, A method for q-calculus, *J. Nonlinear Math. Phys.* **10** (4) (2003), pp. 487–525.
[194] D. Essouabri, Singularité de séries de Dirichlet associées à des polynômes de plusieurs variables et applications en théorie analytique des nombres, Thése, Univ. Henri Poincaré — Nancy I, 1995.
[195] D. Essouabri, Singularité de séries de Dirichlet associées à des polynômes de plusieurs variables et applications en théorie analytique des nombres, *Ann. Inst. Fourier Grenoble* **47** (2) (1997), pp. 429–483.
[196] D. Essouabri, K. Matsumoto and H. Tsumura, Multiple zeta-functions as-

sociated with linear recurrence sequences and the vectorial sum formula, *Canad. J. Math.* **63** (2) (2011), pp. 241-276.

[197] L. Euler, Demonstratio insignis theorematis numerici circa unicias potestatum binomialium, *Nova Acta Acad. Sci. Petropol.* **15** (1799/1802), 33–43; reprinted in *Opera Omnia*, Ser. I, Vol. **16** (2), B. Teubner, Leipzig, 1935, pp. 104-116.

[198] L. Euler, Meditationes circa singulare serierum genus, *Novi Comm. Acad. Sci. Petropol.* **20** (1776), pp. 140–186; reprinted in Opera Omnia, Ser. I, Vol. **15**, B. Teubner (ed.), Berlin, 1927, pp. 217–267.

[199] A polynomial time, numerically stable integer relation algorithm, H. Ferguson and D.H. Bailey; RNR Technical Report RNR-91-032; July 14, 1992.

[200] S. Fomin and A. Zelevinsky, Y-systems and generalized associahedra, *Ann. Math.* **158**(3) (2003), pp. 977–1018.

[201] E. Frenkel and A. Szenes, Thermodynamic Bethe ansatz and dilogarithm identities. I, *Math. Res. Lett.* **2**(6) (1995), pp. 677–693.

[202] S. Friedberg, D. Bump, D. Goldfeld and J. Hoffstein, Multiple Dirichlet Series, Automorphic Forms and Analytic Number Theory: Proc. of the Bretton Woods Workshop on Multiple Dirichlet Series, Amer. Math. Soc., 2005.

[203] H. Furusho, The multiple zeta value algebra and the stable derivation algebra, *Publ. Res. Inst. Math. Sci.* **39** (2003), pp. 695–720.

[204] H. Furusho, p-adic multiple zeta values I.: p-adic multiple polylogarithms and the p-adic KZ equation, *Inv. Math.* **155** (2004), pp. 253–286.

[205] H. Furusho, Multiple zeta values and Grothendick-Teichmller groups, in: Primes and Knots (Baltimore, 2003), T. Kohno and M. Morishita eds., *Contemp. Math.* **416**, Amer. Math. Soc., Providence, RI, 2006, pp. 49–82.

[206] H. Furusho, p-adic multiple zeta values II – tannakian interpretations, *Amer. J. Math.* **129** (2007), pp. 1105–1144.

[207] H. Furusho, Pentagon and hexagon equations, *Ann. Math.* **171** (2010), pp. 545–556.

[208] H. Furusho, Double shuffle relation for associators, *Ann. Math.* **174** (2011), pp. 341–360.

[209] H. Furusho, Geometric interpretation of double shuffle relation for multiple L-values, in: *Galois-Teichmller Theory and Arithmetic Geometry*, H. Nakamura et. al. (eds.), Adv. Studies in Pure Math. **68**, Math. Soc. Japan, Tokyo, 2012, pp. 163–187.

[210] H. Furusho, On relations among multiple zeta values obtained in knot theory. arXiv:1501.06638.

[211] H. Furusho and A. Jafari, Algebraic cycles and motivic generic iterated integrals, *Math. Res. Lett.* **14** (6) (2007), pp. 923-942.

[212] H. Furusho and A.Jafari, Regularization and generalized double shuffle relations for p-adic multiple zeta values, *Compos. Math.* **143** (2007), pp. 1089–1107.

[213] H. Furusho, Y. Komori, K. Matsumoto, and H. Tsumura, Desingularization of complex multiple zeta functions, fundamentals of p-adic multiple L-functions, and evaluation of their special values, arXiv:1309.3982.

[214] H. Gangl, Some Computations in weight 4 motivic cohomology, in: Regulators in Analysis, Geometry and Number Theory, *Progress in Math.*, Vol. **171**, pp. 117–125, Birkhuser, Boston, 2000.

[215] H. Gangl, Functional equations of higher logarithms, *Selecta Math.* **9** (2003), pp. 361–379.

[216] H. Gangl, Functional equations and ladders for polylogarithms. *Comm. Num. Theory Phys.* **7**(3) (2013), pp. 397–410.

[217] H. Gangl, A.B. Goncharov and A. Levin, Multiple logarithms, trees and algebraic cycles, in: *Frontiers in Number Theory, Physics, and Geometry II*, Cartier, Julia, Moussa and Vanhove (eds.), Springer (2007), pp. 759–776.

[218] H. Gangl, A.B. Goncharov and A. Levin, Multiple polylogarithms, polygons, trees and algebraic cycles, *Proc. Symp. Pure Math.* Vol. **80.2**, 2009, pp. 547–594.

[219] H. Gangl and S. Müller-Stach, Polylogarithmic identities in cubical higher Chow groups, in: *Algebraic K-theory* (Seattle, WA, 1997), Amer. Math. Soc., Providence, RI., 1999, pp. 25–40.

[220] H. Gangl, M. Kaneko, D. Zagier, Double zeta values and modular forms, in: *Automorphic Forms and Zeta Functions*, S. Böcherer et al. eds., World Scientific, Hackensack, NJ (2006), pp. 71–106.

[221] A. Gardiner, Four problems on prime power divisibility, *Amer. Math. Monthly* **95** (1988), pp. 926–931.

[222] L. Gegenbauer, Uber einige bestimmte Integrale, *Sitz. Math. Natur. Math.-Naturw. Kl.* **66** (2) (1875), pp. 433–443.

[223] L. Geissinger, Hopf algebras of symmetric functions and class functions, in: Combinatoire et reprbentation de groupe symme'trique (Strasbourg, 1976), *Lecture Notes in Math.*, Vol. **579**, Springer-Verlag, New York, 1977, pp. 168–181.

[224] I.M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V.S. Retakh and J.-Y. Thibon,, Noncommutative symmetric functions, *Adv. Math.* **112** (1995), pp. 218–348.

[225] I. Gessel, Multipartite P-partitions and inner product of skew Schur functions, in: *Contemp. Math.* Vol. **34**, pp. 289–301, Amer. Math. Soc., Providence, RI, 1984.

[226] I. Gesseland and C. Reutenauer, Counting permutations with given cycle-structure and descent set, *J. Combin. Theory Ser. A* **64** (1993), pp. 189–215.

[227] J.W.L. Glaisher, On the residues of the sums of products of the first $p-1$ numbers, and their powers, to modulus p^2 or p^3, *Quart. J. Math.* **31** (1900), pp. 321-353.

[228] J.W.L. Glaisher, On the residues of the sums of the inverse powers of numbers in arithmetical progression, *Quart. J. Math.* **32** (1900), pp. 271-288.

[229] C. Glanois, Motivic unipotent fundamental groupoid of $\mathbb{G}_m \setminus \mu_N$ for $N = 2, 3, 4, 6, 8$ and Galois descents. arXiv:1411.4947.

[230] J. Golden, A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, Motivic amplitudes and cluster coordinates. arXiv:1305.1617

[231] J. Golden, M. Paulos, M. Spradlin, and A. Volovich, Cluster polylogarithms for scattering amplitudes. arXiv:1401.6446.
[232] A.B. Goncharov, Polylogarithms and motivic Galois groups, in: *Symp. in Pure Math.* Vol. **55**, part 2, 1994, pp. 43–97.
[233] A.B. Goncharov, Polylogarithms in arithmetic and geometry, in: Proc. ICM, Vol. **1** (Zurich, 1994), pp. 374387, Birkhauser, Basel, 1995.
[234] A.B. Goncharov, Geometry of configurations, polylogarithms, and motivic cohomology, *Adv. Math.* **114** (1995), pp. 197-318.
[235] A.B. Goncharov, The double logarithm and Manin's complex for modular curves, *Math. Res. Letters* **4** (1997), pp. 617–636.
[236] A.B. Goncharov, Multiple polylogarithms, cyclotomy and modular complexes, *Math. Res. Letters* **5** (1998), pp. 497–516.
[237] A. B. Goncharov, Mixed elliptic motives, in: *London Math. Soc. Lecture Notes* Vol. **254**, pp. 147–221, Cambridge Univ. Press, Cambridge, 1998.
[238] A.B. Goncharov, Geometry of trilogarithm and the motivic Lie algebra of a field, in: Regulators in analysis, geometry and number theory, *Prog. in Math.* **171**, Birkhäuser, Boston, 2000, pp. 127–165.
[239] A.B. Goncharov, Multiple ζ-values, Galois groups, and geometry of modular varieties, in: European Congress of Mathematics (Barcelona, 2000), Vol. **I**, Progr. Math. **201**, Birkhäuser, Basel, 2001, pp. 361–392.
[240] A.B. Goncharov, The dihedral Lie algebaras and Galois symmetries of $\pi_1(\mathbb{P}^1 - \{0, \infty, \mu_N\})$, *Duke J. Math.* **110**(3) (2001), pp. 397–487,
[241] A.B. Goncharov, Galois symmetries of fundamental groupoids and noncommutative geometry, *Duke Math. J.* **128** (2005), pp. 209–284.
[242] A.B. Goncharov and Y.I. Manin, Multiple ζ-motives and moduli spaces $M_{0,n}$, *Compositio Math.* **140** (2004), pp. 1–14.
[243] A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, Classical polylogarithms for amplitudes and Wilson loops, *Phys. Rev. Lett.* **105** (10) (2010), 151605.
[244] A.B. Goncharov and J. Zhao, Grassmannian trilogarithms, *Compos. Math.* **127** (2001), pp. 83–108.
[245] J. Gonzalez-Lorca, Série de Drinfel'd, monodromie et algèbres de Hecke, preprint ENS, LMENS-98-12, www.math.ens.fr/EDITION/preprints/Index.98.html.
[246] D. Goss, Basic Structures of Function Field Arithmetic, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], Vol. **35**, Springer-Verlag, Berlin, 1996.
[247] A. Granville, Arithmetic properties of binomial coefficients I: Binomial coefficients modulo prime powers, *Canad. Math. Soc. Conf. Proc.* **20** (1997), pp. 253–275.
[248] A. Granville, A decomposition of Riemanns zeta-function, in: Analytic Number Theory, Y. Motohashi (ed.), *London Math. Soc. Lecture Note Ser.* **247** (1997), Cambridge, pp. 95-101.
[249] A. Granville, The square of the Fermat quotient, *Integers: Elec. J. of Combin. Num. Thy.* **4** (2004), #A22.
[250] A. Grothendieck, Esquisse d'un programme, in: Geometric Galois Actions,

I, L. Schneps and P. Lochak eds., *London Math. Soc. Lecture Note Ser.* Vol. **242**, Cambridge Univ. Press, 1997, pp. 5–48.

[251] L. Guo and W. Keigher, On differential Rota–Baxter algebras, *J. Pure and Appl. Alg.*, **212** (2008), pp. 522–540.

[252] L. Guo, P. Lei and B. Ma, Applications of shuffle product to restricted decomposition formulas for multiple zeta values, *J. Number Theory* **144** (2014), pp. 219–233.

[253] L. Guo, P. Lei and J. Zhao, Families of weighted sum formulas for multiple zeta values, *Intl. J. Number Theory* **11** (3) (2015), pp. 9971025.

[254] L. Guo and B. Xie, Weighted sum formula for multiple zeta values, *J. Number Theory* **129** (11) (2009), pp. 2747–2765.

[255] L. Guo and B. Xie, The shuffle relation of fractions from multiple zeta values, *Ramanujan J.* **25** (2011), pp. 307–317.

[256] L. Guo and B. Xie, Explicit double shuffle relations and a generalization of Euler's decomposition formula, *J. Algebra* **380** (2013), pp. 46–77.

[257] L. Guo and B. Xie, Structure theorems of mixable shuffle algebras and free commutative Rota-Baxter algebras, *Comm. Algebra* **41** (2013), pp. 2629–2649.

[258] L. Guo and B. Zhang, Renormalization of multiple zeta values, *J. Alg.* **319**(9) (2008), pp. 3770–3809.

[259] B. Hayes, *g*-ology, The decades-long quest for the elusive physical constant *g*, *Amer. Scientist* **92** May-June (3) (2004), pp. 212–214. available online www.americanscientist.org/issues/pub/g-ology/6.

[260] R.Ḣain, The geometry of the mixed Hodge structure on the fundamental group, *Proc. Sympos. Pure Math.* **46** (1987), pp. 247–282.

[261] R. Hain, Classical polylogarithms, in: Motives, *Proc. Symp. Pure Math.* **55** part 2 (1994), pp. 3–42.

[262] R. Hain, Classical polylogarithms, Lectures on the hodge-de rham theory of the fundamental group of $\mathbb{P}-\{0,1,\infty\}$, lecture notes dated March 10, 2005, available online: `www.math.duke.edu/~hain/aws/lectures_provisional.pdf`

[263] R. Hain and R. MacPherson, Higher logarithms, *Ill. J. Math.* **34**(1990), pp. 392–475.

[264] R. Hain and S. Zucker, Unipotent variations of mixed Hodge structure, *Inv. Math.* **88**(1987), pp. 83–124.

[265] R. Hain and S. Zucker, A guide to unipotent variations of mixed Hodge structure, in: Hodge theory (Sant Cugat, 1985), *Lecture Notes in Math.* **1246**(1987), Springer, Berlin-New York, pp. 92–106.

[266] L. van Hamme, Some conjectures concerning partial sums of generalized hypergeometric series, in: *p*-adic functional analysis (Nijmegen, 1996), pp. 223–236, *Lecture Notes in Pure and Appl. Math.* **192**, Dekker, 1997.

[267] M. Hanamura, Mixed motives and algebraic cycles. I. and III, *Math. Res. Lett.* **2** (1995), pp. 811–821 and **6** (1999), pp. 61–82.

[268] M. Hazewinkel, The algebra of quasi-symmetric functions is free over the integers, *Adv. Math.* **164**(2001), pp. 283–300.

[269] W. Heisenberg and H. Euler, Folgerungen aus der Diracschen Theo-

rie des Positrons, *Zeitschr. Phys. 98* (1936), 714. Translation available: arXiv:physics/0605038.

[270] Kh. Hessami Pilehrood and T. Hessami Pilehrood, Congruences arising from Apéry-type series for zeta values, *Adv. Appl. Math.* **49** (2012), pp. 218–238.

[271] Kh. Hessami Pilehrood and T. Hessami Pilehrood, On q-analogues of two-one formulas for multiple harmonic sums and multiple zeta star values. arXiv:1304.0269.

[272] Kh. Hessami Pilehrood, T. Hessami Pilehrood, and R. Tauraso, New properties of multiple harmonic sums modulo p and p-analogues of Leshchiner's series, *Trans. Amer. Math. Soc.*, **366**(6) (2014), pp. 3131–3159.

[273] Kh. Hessami Pilehrood, T. Hessami Pilehrood, and R. Tauraso, Some q-congruences for homogeneous and quasi-homogeneous multiple q-harmonic sums, arXiv:1406.4022.

[274] Kh. Hessami Pilehrood, T. Hessami Pilehrood and J. Zhao, On q-analogues of some families of multiple harmonic sum and multiple zeta star value identities. arXiv:1307.7985.

[275] M.E. Hoffman, Multiple harmonic series, *Pacific J. Math.* **152** (1992), pp. 275–290.

[276] M.E. Hoffman, The algebra of multiple harmonic series, *J. Alg.* **194** (1997), pp. 477-495.

[277] M.E. Hoffman, Quasi-shuffle products, *J. Alg. Combin.* **11** (2000), pp. 49–68.

[278] M.E. Hoffman, Periods of mirrors and multiple zeta values, *Proc. Amer. Math. Soc.* **130** (2002), pp. 971–974.

[279] M.E. Hoffman, Quasi-symmetric functions and mod p multiple harmonic sums, to appear in *Kyushu J. Math.* arXiv:math/0401319.

[280] M.E. Hoffman, The Hopf algebra structure of multiple harmonic sums, *Nuclear Phys. B (Proc. Suppl.)* **135** (2004), 214-219.

[281] M.E. Hoffman, Algebraic aspects of multiple zeta values, in: Zeta Functions, Topology and Quantum Physics, T. Aoki et al. (eds.), *Dev. Math.* **14**, Springer, New York, 2005, pp. 51–74.

[282] M.E. Hoffman, Algebra of multiple zeta values and Euler sums, in: Mini-Conference on Zeta Functions, Index and Twisted K-Theory: Interactions with Physics, Oberwolfach, Germany, May 2, 2006. `www.usna.edu/Users/math/meh`

[283] M.E. Hoffman, A character on the quasi-symmetric functions coming from multiple zeta values, *Electron. J. Combin.* **15**(1) (2008), R97.

[284] M. E. Hoffman, On multiple zeta values of even arguments. arXiv:1205.7051.

[285] M.E. Hoffman, References on Multiple Zeta values and Euler sums, http://www.usna.edu/Users/math/meh/biblio.html

[286] M. Hoffman and K. Ihara, Quasi-shuffle products revisited, Max-Planck-Institut fr Mathematik Preprint Series 2012 (16).

[287] M.E. Hoffman and C. Moen, Sums of triple harmonic series, *J. Number Theory* **60**(1996), pp. 329–331.

[288] M.E. Hoffman and Y. Ohno, Relations of multiple zeta values and their algebraic expression, *J. Algebra* **262** (2003), pp. 332-347.
[289] G. 't Hooft and M. Veltman, Scalar one loop integrals, *Nucl. Phys. B* **153** (1979), pp. 365–401.
[290] I. Horozov, Multiple zeta functions, modular forms and adeles. arXiv:math/0611849.
[291] I. Horozov, Multiple Dedekind zeta functions. arXiv:1101.1594.
[292] F. T. Howard, Explicit formulas for degenerate Bernoulli numbers, *Discrete Math.* **162** (1996), pp. 175–185.
[293] R. Hudson, The $\mathbb{Z}_2$-graded sticky shuffle product Hopf algebra, pp. 237–243, in: Quantum Probability, ed. M Bozejko, *Banach Centre Publ.* **73** (2006).
[294] M. Igarashi, A generalization of Ohno's relation for multiple zeta values, *J. Number Theory* **132** (2012), pp. 565–578.
[295] M. Igarashi, Note on relations among multiple zeta-star values. arXiv:1106.0481.
[296] Y. Ihara, Profinite braid groups, Galois representations and complex multiplications, *Ann. Math.* **123** (1986), pp. 43–106.
[297] Y. Ihara, The Galois representation arising from $\mathbb{P}^1 - \{0, 1, \infty\}$ and Tate twists of even degree, in: Galois groups over $\mathbb{Q}$, *Publ. MSRI* **16** (1989), Springer-Verlag, pp. 299–313.
[298] K. Ihara, M. Kaneko and D. Zagier, Derivation and double shuffle relations for multiple zeta values, *Compos. Math.* **142** (2006), pp. 307–338.
[299] K. Ihara, J. Kajikawa, Y. Ohno and J. Okuda, Multiple zeta values vs. multiple zeta-star values, *J. Algebra* **332** (2011), pp. 187–208.
[300] K. Ihara and H. Ochiai, Symmetry on linear relations for multiple zeta values, *Nagoya Math. J.* **189** (2008), pp. 49–62.
[301] K. Imatomi, T. Tanaka, K. Tasaka and N. Wakabayashi, On some combinations of multiple zeta-star values, *Acta Human. Sci. Univ. Sangio Kyotiensis, Natural science series* **42** (2013), pp. 1–20.
[302] K. Imatomi, T. Tanaka and M. Kaneko, Multi-poly-Bernoulli numbers and finite multiple zeta values, *J. Integer Sequences* **17** (2014), pp. 1–12.
[303] K. Ireland and M. Rosen, A Classical Introduction to Modern Number Theory, Graduate Texts in Math. Vol. **84**, 2nd Edition, Spring-Verlag, Berlin, New York, 1990.
[304] F.H. Jackson, On q-functions and a certain difference operator, *Trans. Roy. Soc. Edin.* **46** (1908), pp. 253–281.
[305] F.H. Jackson, On q-definite integrals, *Quart. J. Pure and Appl. Math.* **41** (1908), pp. 193–203.
[306] D. Jarossay, Double mélange des multizêtas finis et multizêtas symétrisés, *C. R. Acad. Sci. Paris, Ser. I* **352** (2014), pp. 767-771.
[307] D. Jarossay, Finite multiple zeta values, formal algebraic relations and the fundamental group of $M_{0,4}$ and $M_{0,5}$ - I. arXiv:1412.5099.
[308] D. Jarossay, p-adic multiple zeta values and multiple harmonic sums - II: p-adic decompositions of multiple harmonic sums. arXiv:1501.04893.
[309] C. Ji, A simple proof of a curious congruence by Zhao, *Proc. Amer. Math.*

Soc. **133** (2005), pp. 3469–3472.
[310] A. Jonquière, Note sur la srie $\sum_{n=1}^{\infty} \frac{x^n}{n^s}$, *Bull. SocṀath. France* (in French) **17**(1889), 142–152.
[311] S. Joyner, On an extension of the universal monodromy representation for $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. arXiv:1008.4087
[312] A.P. Juskevic and E. Winter, eds., Leonhard Euler und Christian Goldbach: Briefwechsel 1729–1764, Akademie-Verlag, Berlin, 1965.
[313] J. Kajikawa, Duality and double shuffle relations of multiple zeta values, *J. Number Theory* **121**(1) (2006), pp. 1–6.
[314] G. Kallen and A. Sabry, Fourth order vacuum polarization, *K. Dan. Vidensk. Selsk. Mat.-Fys. Medd.* **29** (1955), No. 17.
[315] M. Kaneko, *Double zeta values and modular forms*, proceedings of the Japan-Korea joint seminar on Number Theory (Kuju, Japan) (H. K. Kim and Y. Taguchi, eds.), October 2004.
[316] M. Kaneko, On an extension of the derivation relation for multiple zeta values, in: The Conference on L-Functions, World Sci. Publ., Hackensack, NJ, 2007, pp. 8994.
[317] M. Kaneko, A note on poly-Bernoulli numbers and multiple zeta values, in: Diophantine Analysis and Related Fields: DARF 2007/2008, T. Komatsu (ed.), *AIP Conf. Proc.* Vol. **976**, American Institute of Physics, Melville, NY, 2008, pp. 118–124.
[318] M. Kaneko, Poly-Bernoulli numbers and related zeta functions, in: *Algebraic and Analytic Aspects of Zeta Functions and L-Functions*, G. Bhowmik et. al. (eds.), MSJ Memoirs **21**, Mathematical Society of Japan (2010), pp. 73–85.
[319] M. Kaneko, Finite multiple zeta values (in Japanese), *RIMS Kôkyûroku Bessatsu* **Bx** (2014), pp. ?–?.
[320] M. Kaneko and T. Arakawa, On multiple L-values, *J. Math. Soc. Japan* **56** (2004), pp. 967–991.
[321] M. Kaneko, N. Kurokawa and M. Wakayama, A variation of Eulers approach to values of the Riemann zeta function, *Kyushu J. Math.* **57** (2003), pp. 175–192.
[322] M. Kaneko, M. Noro and K. Tsurumaki, On a conjecture for the dimension of the space of the multiple zeta values, in: *The IMA Volumes in Math. and Appl.*, Vol. **148**, 2010, Software for Algebraic Geometry, pp. 47–58.
[323] M. Kaneko and Y. Ohno, On a kind of duality of multiple zeta-star values, *Internat. J. Number Theory* **6**(8) (2010), pp. 1927–1932.
[324] M. Kaneko and K. Tasaka, Double zeta values, double Eisenstein series, and modular forms of level 2, *Math. Ann.* **357** (3)(2013), pp. 1091–1118.
[325] M. Kaneko and H. Tsumura, Multi-poly-Bernoulli numbers and related zeta functions. arXiv:1503.02156.
[326] M. Kaneko and D. Zagier, Finite multiple zeta values, in preparation.
[327] R. Karplus and N.M. Kroll, Fourth-order corrections in quantum electrodynamics and the magnetic moment of the electron, *Phys. Rev.* **76** (1949), pp. 846–847.
[328] R. Karplus and N.M. Kroll, Fourth-order corrections in quantum electro-

dynamics and the magnetic moment of the electron, *Phys. Rev.* **77** (1950), pp. 536–549.

[329] C. Kassel, Quantum Groups, Graduate Texts in Mathematics **155**, Springer-Verlag, New York, 1995.

[330] K. Kawagoe, M. Wakayama and Y. Yamasaki, q-Analogues of the Riemann zeta, the Dirichlet L-functions and a crystal zeta function, *Forum Math.* **20** (2008), pp. 1–26.

[331] G. Kawashima, A class of relations among multiple zeta values, *J. Number Theory* **129**(4) (2009), pp. 755–788.

[332] G. Kawashima, T. Tanaka and N. Wakabayashi, Cyclic sum formula for multiple L-values, *J. Alg.* **348** (1)(2011), pp. 336–349.

[333] D. Kazakov, The method of uniqueness, a new powerful technique for multiloop calculations, *Phys. Lett.* **B133** (6) (1983), pp. 406–410.

[334] G.S. Kazandzidis, Congruences on the binomial coefficients, *Bull. Soc. Math. Grce (N.S.)* **9** (1968), pp. 1–12.

[335] J.P. Kelliher and R. Masri, Analytic continuation of multiple Hurwitz zeta functions, *Math. Proc. Camb. Phil. Soc.* **145**(3) (2008), pp. 605–617.

[336] A.N. Kirillov, Dilogarithm identities, *Prog. Theor. Phys. Suppl.* **118** (1995), pp. 61–142.

[337] T. Klassen and E. Melzer, The thermodynamics of purely elastic scattering theories and conformal perturbation theory *Nucl. Phys. B 350* (1991), pp. 635–689.

[338] J. Knauer and J. Richstein, The continuing search for Wieferich primes, *Math. Comp.* **74** (2005), pp. 1559–1563.

[339] V.G. Knizhnik and A.B. Zamolodchikov, Current algebras and Wess-Zumino models in two dimensions, *Nucl. Phys. B* **247** (1984), pp. 83–103.

[340] W. Kohnen and D. Zagier, Modular forms with rational periods, in: *Modular Forms*, R.A. Rankin (ed.), Ellis Horwood, Chichechester (1984), pp. 197-249.

[341] T. Kohno, Monodromy representations of braid groups and Yang-Baxter equations, *Ann. Inst. Fourier* **37**(4) (1987), pp. 139–160.

[342] Y. Komori, K. Matsumoto and H. Tsumura, Functional relations for zeta-functions of root systems, in: Number Theory: Dreaming in Dreams, Proc. 5th China-Japan Seminar, T. Aoki, S. Kanemitsu and J.-Y. Liu (eds.), World Sci. Publ. 2009, pp. 135–183.

[343] Y. Komori, K. Matsumoto and H. Tsumura, On Witten multiple zeta-functions associated with semisimple Lie algebras II, *J. Math. Soc. Japan.* **62**(2) (2010), pp. 355–394

[344] Y. Komori, K. Matsumoto and H. Tsumura, On Witten multiple zeta-functions associated with semisimple Lie algebras III, in 'Multiple Dirichlet Series, L-functions and Automorphic Forms', D. Bump, S. Friedberg and D. Goldfeld (eds), Progress in Math., Vol. **300**, Birkhauser, 2012, pp. 215–276.

[345] Y. Komori, K. Matsumoto and H. Tsumura, On Witten multiple zeta-functions associated with semisimple Lie algebras IV, *Glasgow Math. J.* **53** (2011), pp. 185–206.

[346] Y. Komori, K. Matsumoto and H. Tsumura, Barnes multiple zeta-functions,

Ramanujan's formula, and relevant series involving hyperbolic functions. arXiv:1006.3336.

[347] Y. Kondo, S. Saito and T. Tanaka, The Bowman-Bradley theorem for multiple zeta-star values, *J. Number Theory* **132**(9) (2012), pp. 1984–2002.

[348] M. Kontsevich, *Vassiliev's knot invariants,* I. M. Gel'fand Seminar, Adv. Soviet Math., **16**(2) (1993), pp. 137–150.

[349] M. Kontsevich, The 1/2-logarithm, unpublished note, 1995, reproduced as an appendix [185].

[350] M. Kontsevich and D. Zagier: Periods, in: Mathematics Unlimited — 2001 and beyond, Engquist and Schmidt (eds.), pp. 771–808, Springer, 2001.

[351] D. Kreimer, Knots and divergences, *Phys. Lett.* **B354** (1?2) (1995), pp. 117–124.

[352] E. E. Kummer, Über die Transcendenten, welche aus wiederholten Integrationen rationaler Formeln entstehen, *J. Reine Angew. Math.* **21** (1840), pp. 74–90.

[353] T.H. Koornwinder, On Zeilbergers algorithm and its q-analogue, *J. Comput. Appl. Math.* **48** (1993), pp. 91111.

[354] N. Kurokawa, Multiple zeta functions: an example, in: Zeta Functions in Geometry, (Tokyo, 1990), *Adv. Stud. Pure. Math.* Vol. **21**, Kinokuniya, Tokyo, 1992, pp. 219–226.

[355] N. Kurokawa, M. Lalin, and H. Ochiai, Higher Mahler measures and zeta functions, *Acta Arith.* **135**(3) (2008), pp. 269–297.

[356] M.N. Lalin, Some examples of Mahler measures as multiple polylogarithms, *J. Number Theory* **103** (2003), pp. 85–108.

[357] M.N. Lalin, On a certain combination of colored multizeta values, *J. Ramanujan Math. Soc.* **20** (2006), pp. 115–127.

[358] J. Lambert, Mémoire sur quelques propriétés remarquables des quantités transcendantes circulaires et logarithmiques, in: Pi: A Source Book (3rd ed.), L. Berggren, J. Borwein and P. Borwein (eds.), New York: Springer-Verlag, pp. 129140, 2004,

[359] S. Lang, Introduction to Modular Forms, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Corr. 3rd printing 2001.

[360] J.A. Lara Rodrguez, Some conjectures and results about multizeta values over $\mathbb{F}_q[t]$, *J. Number Theory* **130** (2010), pp. 1013–1023.

[361] J.A. Lara Rodrguez, Relations between multizeta values in characteristic p, *J. Number Theory* **131** (2011), pp. 2081–2099.

[362] J.A. Lara Rodrguez, Special relations between multizeta values and parity results. arXiv:1108.4726.

[363] J.A. Lara Rodrguez and D. Thakur, Zeta-like multizeta values for $\mathbb{F}_q[t]$. arXiv:1312.4928.

[364] S. Laporta and E. Remiddi, The analytical value of the electron $(g-2)$ at order α^3 in QED, *Phys. Lett. B* **379** (1996), pp. 283–291.

[365] J. A. Lappo-Danilevsky, Théorie algorithmique des corps de Riemann, *Rec. Math. Moscou* **34** (6) (1927), pp. 113–146. URL http://mi.mathnet.ru/msb7440.

[366] J. A. Lappo-Danilevsky, Memoires sur la theorie des systemes des équations

différentieles linéaires, French Edition, Amer. Math. Soc. Chelsea Publishing, 1953.
[367] T.Q.T. Le and J. Murakami, Kontsevich's integral for the Homfly polynomial and relations between values of multiple zeta functions, *Topology Appl.* **62** (1995), pp. 193–206.
[368] T.Q.T. Le and J. Murakami, Kontsevich's integral for the Kauffman polynomial, *Nagoya Math. J.* **142** (1996), pp. 39–65.
[369] R. Lee and J. Schwermer, Cohomology of arithmetic subgroups of SL_3 at infinity, *J. Reine Angew. Math.* **330** (1982), pp. 100-131.
[370] C. Leudesdorf, Some results in the elementary theory of numbers, *Proc. London Math. Soc.* **20** (1889), pp. 199-212.
[371] S. Leurent and D. Volin, Multiple zeta functions and double wrapping in planar $N = 4$ SYM, arXiv: 1302.1135.
[372] L. Lewin, Dilogarithms and associated functions, Macdonald, London, 1958.
[373] L. Lewin, Polylogarithms and Associated Functions, Elsevier Sci. Publishers, New York, New York, 1981.
[374] L. Lewin, Structural properties of polylogarithms, Math. Surverys and Monographs, Vol. **37**, Amer. Math. Soc. 1991, Providence, RI.
[375] M. Levine, Mixed motives, Amer. Math. Soc., 1998.
[376] Z. Li, Sum of multiple zeta values of fixed weight, depth and i-height, *Math. Z.* **258** (2008), pp. 133–142.
[377] Z. Li, Sum of multiple q-zeta values, *Proc. Amer. Math. Soc.* **138** (2010), pp. 505-516.
[378] Z. Li, Gamma series associated to elements satisfying regularized double shuffle relations, *J. Number Theory* **130** (2010), pp. 213–231.
[379] Z. Li, Higher order shuffle regularization of multiple zeta values, *Proc. Amer. Math. Soc.* **138** (2010), pp. 2321–2333.
[380] Z. Li, On a conjecture of Kaneko and Ohno, *Pacific J. Math.* **257** (2012), pp. 419–430.
[381] Z. Li, Regularized double shuffle and Ohno-Zagier relations of multiple zeta values, *J. Number Theory* **133** (2013), pp. 596–610.
[382] Z. Li, Some identities in the harmonic algebra concerned with multiple zeta values, *Int. J. Number Theory* **9** (2013), pp. 783–798.
[383] Z. Li, Another proof of Zagier's evaluation formula of the multiple zeta values $\zeta(2,\ldots,2,3,2,\ldots,2)$, *Math. Research Lett.* **20** (2013), pp. 947–950.
[384] Z. Li, On harmonic sums and alternating Euler sums. arXiv:1012.5192.
[385] E. Linebarger and J. Zhao, A family of multiple harmonic sum and multiple zeta star value identities, *Mathematika* **61** (1) (2015), pp. 63–71.
[386] S. Lichtenbaum, The construction of weight two arithmetic cohomology, *Inv. Math.* **88** (1987), pp. 183–215.
[387] F. Lindemann, Über die Ludolph'sche Zahl, *Sitzungsber. Königl. Preuss. Akad. Wissensch. zu Berlin* **2** 1882, pp. 679-682.
[388] A.E. Lipstein, L. Mason, From dlogs to dilogs; the super Yang-Mills MHV amplitude revisited arXiv:1307.1443.
[389] D. Leshchiner, Some new identities for $\zeta(k)$, *J. Number Theory,* **13** (1981),

pp. 355–362.
[390] J. Lavoie, F. Grondin, A. Rathie and K. Arora, Generalizations of Dixon's theorem on the sum of a ${}_3F_2$, *Math. Comp.* **62** (1994), pp. 267–272.
[391] M. Lothaire, Combinatorics of Words, Encycl. Math. Appl., Vol. **17**, Addison-Wesley, 1983.
[392] D. Ma, Period polynomial relations between double zeta values of odd weight. arXiv:1409.7717.
[393] T. Machide, Weighted sums with two parameters of multiple zeta values and their formulas. arXiv:1112.2554.
[394] T. Machide, Extended double shuffle relations and the generating function of triple zeta values of any fixed weight. arXiv:1204.4085.
[395] T. Machide, Some restricted sum formulas for double zeta values, *Proc. Japan Acad. Ser.* **A 89** (2013), pp. 51–54.
[396] T. Machide, Generators for vector spaces consisting of double zeta values with even weight, *J. Number Theory* **133** (2013), pp. 2240–2246.
[397] T. Machide, Identities involving cyclic sums of regularized multiple zeta values each of depth less than 5. arXiv:1502.02505.
[398] T. Machide, Use of the generating funtion to generalize the sum formula for quadruple zeta values. arXiv:1503.01194.
[399] C. Malvenuto and C. Reutenauer, Duality between quasi-symmetric functions and the Solomon descent algebra, *J. Algebra* **177** (1994), pp. 967982.
[400] D. Manchon and S. Paycha, Nested sums of symbols and renormalised multiple zeta functions, *Int. Math. Res. Notices* (2010), pp. 4628–4697.
[401] C. Markett, Triple sums and the Riemann zeta function, *J. Number Theory* **48** (1994), pp. 113-132.
[402] R. Masri, The Herglotz-Zagier function, double zeta values, and values of L-series, *J. Number Theory* **106** (2004), pp. 219–237.
[403] R. Masri, Multiple Dedekind zeta functions and evaluations of extended multiple zeta values, *J. Number Theory* **115** (2005), pp. 295–309.
[404] R. Masri, Multiple zeta values over global function fields, in: Multiple Dirichlet Series, Automorphic Forms, and Analytic Number Theory (Bretton Woods, 2005), S. Friedberg et al. (eds.), *Proc. Symp. Pure Math.* **75**, Amer. Math. Soc., Providence, RI, 2006, pp. 157-175.
[405] K. Matsumoto, On the analytic continuation of various multiple zeta functions, in: Number Theory for the Millenium II, 2002, M.A. Bennett et al (eds.), A. K. Peters, pp. 417–440.
[406] K. Matsumoto, Functional equations for double zeta-functions. *Math. Proc. Camb. Phil. Soc.* **136**(1) (2004), pp. 1–7.
[407] K. Matsumoto, Analytic properties of multiple zeta-functions in several variables, in: Number Theory: Tradition and Modrnization, Proc. 3rd China-Japan Seminar, W. Zhang and Y. Tanigawa (eds.), Springer, 2006, pp. 153–173.
[408] K. Matsumoto, T. Nakamura and H. Tsumura, Functional relations and special values of Mordell-Tornheim triple zeta and L-functions, *Proc. Amer. Math. Soc.* **136**(3) (2008), pp. 2135–2145.
[409] K. Matsumoto and H. Tsumura, On Witten multiple zeta-functions asso-

ciated with semisimple Lie algebras I, Annales de l'institut Fourier, **56** (5) (2006), pp. 1457–1504.

[410] C. Mazza, V. Voevodsky and C.A. Weibel, Lecture Notes on Motivic Cohomology, Amer. Math. Soc. Bookstore, 2006.

[411] R. Mestrovic, Wolstenholme's theorem: its generalizations and extensions in the last hundred and fifty years (18622012). arXiv:1111.3057.

[412] H. N. Minh, Shuffle algebra and differential Galois group of colored polylogarithms, *Nuclear Phys.* **B135** (2004), pp. 220–224.

[413] H. N. Minh, G. Jacob, M. Petitot and N. Oussous, De l'algèbre des ζ de Riemann multivariées à l'algèbre des ζ de Hurwitz multivariées, *Sem. Lothar. Combin.* **44** (2000), Art. B44i, 21p.

[414] H. N. Minh and M. Petitot, Lyndon words, polylogarithms, and the Riemann ζ function, *Discrete Math.* **217** (2000), pp. 273-292.

[415] H. N. Minh, M. Petitot and J. Van Der Hoeven, Shuffle algebra and polylogarithms, *Disc. Math.* **225** (13) (2000), pp. 217–230.

[416] V. H. Moll, L. Jiu and C. Vignat A symbolic approach to multiple zeta values at the negative integers. arXiv:1503.04257.

[417] E. Mortenson, A p-adic supercongruence conjecture of van Hamme, *Proc. Amer. Math. Soc.* **136** (2008), pp. 4321–4328.

[418] S. Muneta, On some explicit evaluations of multiple zeta-star values, *J. Number Theory* **128**(9) (2008), pp. 2538–2548.

[419] S. Muneta, A note on evaluations of multiple zeta values *Proc. Amer. Math. Soc.* **137**(3) (2009), pp. 931–935.

[420] H. Murahara, A Note on finite real multiple zeta values. arXiv:1411.1123.

[421] M.R. Murty and K. Sinha, Multiple Dirichlet series, automorphic forms, and analytic number theory, pp. 135–156, Proc. Symp. Pure Math. **75**, Amer. Math. Soc., Providence, RI (2006).

[422] T. Nakamura, Bernoulli numbers and multiple zeta values, *Proc. Japan Acad. Ser. A* **81** (2005), pp. 21–22.

[423] T. Nakamura, Restricted and weighted sum formulas for double zeta values of even weight, *Šiauliai Math. Seminar* **4** (2009), pp. 151–155.

[424] T. Nakamura, Zeros and the universality for the EulerZagierHurwitz type of multiple zeta-functions, *Bull. London Math. Soc.* **41** (2009), pp. 691–700.

[425] T. Nakamura and Ł. Pankowski, On universality for linear combinations of L-functions, *Monatshefte Math.* **165**(3-4) (2012), pp. 433–446.

[426] T. Nakamura and K. Tasaka, Remarks on double zeta values of level 2, *J. Number Theory* **133** (2013), pp. 48–54.

[427] N. Nakanishi, Graph theory and Feynman integrals, Gordon and Breach Science Publishes, New York, London, Paris, 1971.

[428] A. Nakayashiki, A. Pakuliak and V. Tarasov, On solutions of the KZ and qKZ equations at level zero, *Ann. Inst. Henri Poincare* **71** (1999), pp. 459–496.

[429] W. Neumann and D. Zagier, Volumes of hyperbolic three-manifolds *Topology* **24** (1985), pp. 307–332.

[430] N. Nielsen, Handbuch der Theorie der Gammafunktion, Leipzig, 1906. Available online:

archive.org/stream/handbuchgamma00nielrich#page/n62/mode/1up
[431] Y. Ohno, A generalization of the duality and sum formulas on the multiple zeta values, *J. Number Thery* **74** (1999), pp. 39-43.
[432] Y. Ohno, Sum relations for multiple zeta values, in: Zeta Functions, Topology and Quantum Physics, *Dev. Math.* **14**, Springer, 2005, pp. 131-144.
[433] Y. Ohno and J. Okuda, On the sum formula for the -analogue of non-strict multiple zeta values, *Proc. Amer. Math. Soc.* **135** (2007), pp. 3029–3037.
[434] Y. Ohno, J. Okuda and W. Zudilin, Cyclic q-MZSV sum, *J. Number Theory* **132** (2012), pp. 144–155.
[435] Y. Ohno and N. Wakabayashi, Cyclic sum of multiple zeta values, *Acta Arith.* **123**(3) (2006), pp. 289–295.
[436] Y. Ohno and D. Zagier, Multiple zeta values of fixed weight, depth, and height, *Indag. Math.* **12** (2001), pp. 483-487.
[437] Y. Ohno and W. Zudilin, Zeta stars, *Commun. Number Theory Phys.* **2** (2008), pp. 325–347.
[438] T. Okamoto, Generalizations of Apostol-Vu and Mordell-Tornheim multiple zeta functions, *Acta Arith.* **140** (2009), pp. 169–187,
[439] T. Okamoto, Multiple zeta values related with the zeta-functions of the root system of type B_r, *JP J. Algebra Number Theory Appl.* **22** (2011), pp. 1–44.
[440] T. Okamoto, Multiple zeta values related with the zeta-function of the root system of type A_2, B_2 and G_2, *Comment. Math. Univ. St. Pauli* **61** (2012), pp. 9–27.
[441] T. Okamoto, On alternating analogues of the Mordell-Tornheim triple zeta values, *J. Ramanujan Math. Soc.* **28**(2) (2013), pp. 247–269.
[442] A. Okounkov, Hilbert schemes and multiple q-zeta values. arXiv:1404.3873.
[443] J. Okuda, Duality formulas of the special values of multiple polylogarithms, *Bull. London Math. Soc.* **37** (2005), pp. 230-242.
[444] J. Okuda and Y. Takeyama, On relations for the q-multiple zeta values, *Ramanujan J.* **14** (2007), pp. 379–387.
[445] J. Okuda and K. Ueno, Relations for multiple zeta values and Mellin transforms of multiple polylogarithms, *Publ. Res. Inst. Math. Sci.* **40** (2004), pp. 537-564.
[446] J. Okuda and K. Ueno, New approach to Ohno relations for multiple zeta values. arXiv:math/0106148.
[447] J. Oesterlé, Polylogarithmes, *Astérisque* **216** (1993), pp. 49–67.
[448] K. Oyama, Ohno's relation for finite multiple zeta values, arXiv:1506.00833.
[449] Y.L. Ong, M. Eie, and C-S. Wei, Explicit evaluations of quadruple Euler sums, *Acta Arith.* **144** (2010), pp. 213–230.
[450] The On-Line Encyclopedia of Integer Sequences, `http://oeis.org/`.
[451] Three-photon annihilation of an electron-positron pair, A. Ore and J.L. Powell, *Phys. Rev.* **75** (1949), pp. 1696–1699.
[452] H. Pan, A generalization of Wolstenholme's harmonic series congruence, *Rocky Mountain J. Math.* **38** (2008), pp. 1263–1269.
[453] E. Panzer, On hyperlogarithms and Feynman integrals with dive rgences and many scales, *JHEP* (71) (2014), pp. 1403. arXiv: hep-th 1401.4361.

[454] E. Panzer, Feynman integrals via hyperlogarithms, *Proc. Sci.* (LL2014)049 (2014), proceedings of "Loops & Legs in Quantum Field Theory", Weimar (Germany), April 27th — May 2nd.

[455] E. Panzer, Feynman integrals and hyperlogarithms, PhD thesis, Humboldt-Universität zu Berlin, submitted 26th of August 2014. Updated version at arXiv: 1506.07243

[456] D. Parker, A. Scherlis, M. Spradlin and A. Volovich, Hedgehog bases for A_n cluster polylogarithms and an application to six-point amplitudes, arXiv:1507.01950.

[457] A. Petermann, Fourth order magnetic moment of the electron, *Helv. Phys. Acta.* **30** (1957), pp. 407–408.

[458] M. Petitot, Table des multiple zeta values jusqu'au poids 16, `www.lifl.fr/~petitot/recherche/MZV`

[459] M. Petkovsek, H. Wilf and D. Zeilberger, A=B, A K Peters/CRC Press, 1996.

[460] W.F. Pickard, On polylogarithms, *Publ. Math. Debrecen* **15** (1968), pp. 33–43.

[461] H. Poincaré, Oevres, Vol. **2**, Paris, 1916.

[462] A. van der Poorten, A proof did Euler missed ... apery's proof of the irrationality of $\zeta(3)$. An informal report, *The Mathematical Intelligencer* **1** (December 1979), pp. 195–203.

[463] D. Ramakrishnan, On the monodromy of higher logarithms, *Proc. Amer. Math. Soc.* **85** (1982), pp. 596–599.

[464] S. Ramanujan, Modular equations and approximations to π, *Quart. J. Math. (Oxford)* **45** (1914), pp. 350–372.

[465] G. Racinet, Doubles mélanges des polylogarithmes multiples aux racines de l'unité, *Publ. Math. IHES* **95** (2002), pp. 185–231.

[466] H. Rademacher, Topics in Analytic Number Theory (1973), Springer-Verlag.

[467] T. Rivoal, La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs, *C. R. Acad. Sci. Ser. A. Math.* **331** (2000), pp. 267–270.

[468] R. Ree, Lie elements and an algebra associated with shuffles, *Ann. Math.* **68** (1958), pp. 210–220.

[469] C. Reutenauer, Free Lie Algebras, Oxford Science Publications, Oxford, 1993.

[470] B. Riemann, Über die Anzahl der Primzahlen unter einer gegebenen Grösse, *Monatsberichte der Berliner Akademie*, 1859. In: Gesammelte Werke, Teubner, Leipzig (1892). Reprinted by Dover, New York (1953). Original manuscript (with English translation) reprinted in [73].

[471] J. Rosen, Multiple harmonic sums and Wolstenholme's theorem, *Intl. J. Number Theory* **9** (8) (2013), pp. 2033-2052.

[472] G.-C. Rota, Baxter algebras and combinatorial identities I, *Bull. Amer. Math. Soc.* **5** (1969), pp. 325-329.

[473] G.-C. Rota and D. A. Smith, Fluctuation theory and Baxter algebras, *Symposia Math.* **IX** (1971), pp. 179–201.

[474] G.-C. Rota, Baxter operators, an introduction, In: *Gian-Carlo Rota on Combinatorics, Introductory papers and commentaries*, J. P. S. Kung Ed., Contemp. Mathematicians, Birkhäuser Boston, Boston, MA 1995.
[475] D. Rolfsen, Knots and Links, AMS Chelsea Publishing, 2003.
[476] B. Sadaoui, Multiple zeta values at the non-positive integers, *Comptes Rendus Math.* **352** (12) (2014), pp. 977-984.
[477] S. Saito, T. Tanaka and N. Wakabayashi, Combinatorial remarks on the cyclic sum formula for multiple zeta values, *J. Integer Sequences* **14** (2011), Article 11.2.4.
[478] S. Saito and N. Wakabayashi, The Bowman-Bradley type theorem for finite multiple zeta values, *J. Integer Sequences* **14** (2011), art. 11.2.4, 20 pp.
[479] S. Saito and N. Wakabayashi, Sum formula for finite multiple zeta values. *J. Math. Soc. Japan*. arXiv:1305.6529.
[480] A. Sankaranarayanan, An identity involving Riemann zeta function, *Indian J. Pure Appl. Math.* **18**(7) (1987), pp. 794–800.
[481] K.-G. Schlesinger, Some remarks on q-deformed multiple polylogarithms. arXiv:math/0111022.
[482] O. Schlotterer and S. Stieberger, Motivic multiple zeta values and superstring amplitudes, *J. Phys. A: Math. Theor.* **46** (2013), 475401 (37 pp).
[483] O. Schnetz, Calculation of the ϕ^4 6-loop non-zeta transcendental. arXiv:hep-th/9912149.
[484] O. Schnetz, Quantum periods: A census of ϕ^4 transcendentals, *Comm. in Number Theory Phys.* **4** (1) (2010), pp. 1–48.
[485] L. Schneps: On the Poisson bracket on the free Lie algebra in two generators, *J. Lie Theory* **16** (1)(2006), pp. 19–37.
[486] ARI, GARI, Zig and Zag: An introduction to Ecalle's theory of multiple zeta values, arXiv:1507.01534.
[487] O. Schnetz, Graphical functions and single-valued multiple polylogarithms. arXiv:1302.6445.
[488] O. Schnetz, Algebraic geometry informs perturbative quantum field theory. arXiv:1409.5570.
[489] J. Schwinger, Quantum electrodynamics. III. The electromagnetic properties of the electron — radiative corrections to scattering, *Phys. Rev.* **79** (1949), pp. 790–817.
[490] M.P. Schutzenberger, On a factorisation of free monoids, *Proc. Amer. Math. Soc.* **16** (1965), pp. 21–24,
[491] Z. Shen and T. Cai, Some identities for multiple Hurwitz-zeta values (in Chinese), *Sci. Sinica Math.* **41** (2011), 955-970.
[492] Z. Shen and T. Cai, Some identities for multiple zeta values, *J. Number Theory* **132** (2012), pp. 314–323.
[493] Z. Shen and T. Cai, Congruences for alternating triple harmonic sums, *Acta Mathematica Sinica, Chinese Series*, **55** (4) (2012), pp. 737–748.
[494] L.-L. Shi and H. Pan, A q-analogue of Wolstenholme's harmonic series congruence, *Amer. Math. Monthly*, **114**(6) (2007), pp. 529-531.
[495] G. Shimura, Elementary Dirichlet Series and Modular Forms, *Springer Monographs in Math.*, Springer, New York, 2007.

[496] J. H. Silverman, Wieferich's criterion and the *abc*-conjecture, *J. Number Theory* **30** (1988), pp. 226–237.
[497] R. Sitaramachandra Rao, A formula of S. Ramanujan, *J. Number Theory* **25**(1) (1987), pp. 1–19.
[498] R. Sitaramachandrarao Rao and B. Davis, Some identities involving the Riemann zeta function, II, *Indian J. Pure Appl. Math.* **17**(10) (1986), pp. 1175–1186.
[499] R. Sitaramachandra Rao and M. V. Subbarao, Transformation formulae for multiple series, *Pacific J. Math.* **113**(2) (1984), pp. 471–479.
[500] L. Slater, Generalized Hypergeometric Functions, Cambridge University Press, Cambridge (1966).
[501] V.A. Smirnov, Feynman integral calculus, Berlin, Springer, 2006.
[502] C.M. Sommerfield, Magnetic dipole moment of the electron, *Phys. Rev.* **107** (1957), pp. 328–329.
[503] I. Soudère, Motivic double shuffle, *Intl. J. Number Theory* **6** (2010), pp. 339–370.
[504] I. Soudère, Cycle complex over the projective line minus three points: toward multiple zeta values cycles. arXiv:1210.4653.
[505] I. Soudère, A relative basis for mixed Tate motives over the projective line minus three points. arXiv:1312.1849.
[506] H.M. Srivastava and J. Choi, Zeta and q-zeta functions and associated series and integrals, Elsevier, 1st edition (2011).
[507] R. Stanley, On the number of reduced decompositions of elements of Coxeter groups, *European J. Combin.* **5** (1984), pp. 359–372.
[508] J. Steenbrink and S. Zucker, Variation of mixed Hodge structure. I, *Inv. Math.* **80** (1985), pp. 489–542.
[509] Z.-W. Sun, Congruences concerning Bernoulli numbers and Bernoulli polynomial, *Disc. Applied Math.* **105** (2000), pp. 193–223.
[510] A.A. Suslin, K_3 of a field and the Bloch group, in: Galois Theory, Rings, Algebraic Groups and their Applications, (Russian), *Trudy Mat. Inst. Steklov.* **183** (1990), pp. 180–199, 229.
[511] T. Takamuki, The Kontsevich invariant and relations of multiple zeta values, *Kobe J. Math.* **16** (1999), pp. 27-43.
[512] Y. Takeyama, The algebra of a q-analogue of multiple harmonic series, *SIGMA* **9** (2013), Paper 0601, 15 pp.
[513] T. Tanaka, On the quasi-derivation relation for multiple zeta values, *J. Number Theory* **129**(9) (2009), pp. 2021–2034.
[514] T. Tanaka, Restricted sum formula and derivation relation for multiple zeta values. arXiv:1303.0398.
[515] T. Tanaka and N. Wakabayashi, An algebraic proof of the cyclic sum formula for multiple zeta values, *J. Algebra* **323**(3) (2010), pp. 766–778.
[516] B. Tangedal and P. Young, On p-adic multiple zeta and log gamma functions, *J. Number Theory* **131**(7) 2011, pp. 1240–1257.
[517] K. Tasaka and S. Yamamoto, On some multiple zeta-star values of one-two-three indices. *Intl. J. Num. Theory* **9** (5) (2013), pp. 1171–1184.
[518] R. Tauraso, Congruences involving alternating multiple harmonic sum,

Electron. J. Combin. **17**(1) (2010), R16.
[519] R. Tauraso, New harmonic number identities with applications, *Sém. Lothar. Combin.* **63** (2010), art. B63g (8 pp).
[520] R. Tauraso and J. Zhao, Congruences of alternating multiple harmonic sums, *J. Comb. Number Theory* **2**(2) (2010), pp. 129–159.
[521] D. Terhune, Evaluation of double L-values, *J. Number Theory* **105** (2004), pp. 275–301.
[522] D. Terhune, Evaluations of a class of double L-values *Proc. Amer. Math. Soc.* **134** (2006), pp. 1881–1889.
[523] D.S. Thakur, Function Field Arithmetic, World Scientific, 2004.
[524] T. Terasoma, Mixed Tate motives and multiple zeta values, *Inv. Math.* **149** (2002), pp. 339–369.
[525] T. Terasoma, Period integral and multiple zeta values, *Sugaku Expositions* **23** (2010), pp. 91–103. Translation of Japanese original, *Sugaku* **57** (2005), pp. 255–266.
[526] T. Terasoma, Geometry of multiple zeta values, in: *Proc. Intl. Congress of Mathematicians* (Madrid, 2006), Vol. **II**, European Mathe. Soc., Zürich, 2006, pp. 627–635.
[527] T. Terasoma, Brown-Zagier relation for associators. arXiv:1301.7474.
[528] I. Todorov, Polylogarithms and multizeta values in massless Feynman amplitudes, Bures-sur-Yvette preprint, IHES/P/14/10.
[529] L. Tornheim, Harmonic double series, *Amer. J. Math.* 72 (1950), pp. 303–314.
[530] H. Tsumura, A note on q-analogues of the Dirichlet series and q-Bernoulli numbers, *J. Number Theory***39** (3) (1991), pp. 251–256.
[531] H. Tsumura, Combinatorial relations for Euler-Zagier sums, *Acta Arith.* **111** (2004), pp. 27–42.
[532] H. Tsumura, On alternating analogues of Tornheim's double series II, *Ramanujan J.* **18** (2009), pp. 81-90.
[533] C. Truesdell, On a function which occurs in the theory of the structure of polymers, *Ann. Math.* **46** (1945), pp. 144157.
[534] H. Tsunogai, On ranks of the stable derivation algebra and Delignes problem, *Proc. Japan Acad. Ser. A Math. Sci.* **73** (1997) 29–31.
[535] E. A. Ulanskii, Identities for generalized polylogarithms, (Russian), Mat. Zametki **73**(4) (2003), pp. 613–624; English transl., Math. Notes **73**(3-4) (2003), pp. 571–581.
[536] N. Ussyukina, Calculation of multiloop diagrams in high orders of perturbation theory, *Phys. Lett.* **B267** (1991), pp. 382–388.
[537] N. Ussyukina, A Davydychev, Exact results for three- and four-point ladder diagrams with an arbitrary number of rungs, *Phys. Lett.* **B305** (1993), pp. 136–143.
[538] H. S. Vandiver, On development in an arithmetic theory of the Bernoulli and allied numbers, *Scripta Math.* **25** (1961), pp. 273–303.
[539] D. V. Vasilev, Some formulas for Riemann zeta-function at integral points, *Vestnik Moskov. Univ. Ser. I, Mat. Mekh.* **1** (1996), pp. 81–84; English translation in *Moscow Univ. Math. Bull.* **51**(1) (1996), pp. 41–43.

[540] J. A. M. Vermaseren, Harmonic sums, Mellin transforms and integrals, *Internat. J. Modern Phys. A* **14**(13) (1999), pp. 2037–2076.
[541] J. Vollinga and S. Weinzierl, Numerical evaluation of multiple polylogarithms, *Computer Phys. Comm.* **167**(3) (2005), pp. 177–194.
[542] V. Voevodsky, Triangulated Category of Motives over a field, in: Cycles, Transfers and Motivic Homology Theories, V. Voevodsky, A. Suslin and E. Friedlander eds., Annals of Math. Studies, 1999.
[543] M. Waldschmidt, Valeurs zêta multiples. Une introduction, *J. Théor. Nombres Bordeaux* **12** (2000), pp. 581–595.
[544] M. Waldschmidt, Multiple polylogarithms: an introduction, in: *Number Theory and Discrete Mathematics Trends in Mathematics*, 2002, pp. 1–12.
[545] J. Wan, Some notes on weighted sum formulae for double zeta values, in: Number Theory and Related Fields: In Memory of Alf van der Poorten, Springer Proc. in Math. & Stat., J. Borwein, I. Shparlinski and W. Zudilin (eds.), Vol. **43**, pp. 361–379, 2013.
[546] Q. Wang, Moduli spaces and multiple polylogarithm motives, *Adv. in Math.* **206** (2006), pp. 329–357.
[547] L. Wang and T. Cai, A curious congruence modulo prime powers, to appear in *J. Number Theory*.
[548] L.C. Washington, Introduction to Cyclotomic Fields, Graduate Texts in Mathematics Vol. **83**, Springer, 1997.
[549] S. Weinzierl, Feynman graphs. arXiv:1301.6918
[550] E. T. Whittaker and G. N. Watson, A Course of Modern Analysis, Cambridge University Press, 1996.
[551] J. Wildeshaus, Realizations of Polylogarithms, Lecture Notes in Math. Vol. **1650**, Springer-Verlag, 1997.
[552] E. Witten, On quantum gauge theories in two-dimensions, *Commun. Math. Phys.* **141**(1) (1991), pp. 153–209.
[553] J. Wolstenholme, On certain properties of prime numbers, *Quarterly J. Pure & Appl. Math.* **5** (1862), pp. 35–39.
[554] Z. Wojtkowiak, On l-adic iterated integrals. I. Analog of Zagier conjecture, *Nagoya Math. J.* **176** (2004), pp. 113–158.
[555] Z. Wojtkowiak, A construction of analogs of the Bloch-Wigner function. *Math. Scandinavica* **65** (1989), pp. 140–142.
[556] Z. Wojtkowiak, The basic structure of polylogarithmic functional equations, in: *Structural Properties of Polylogarithms*, Mathematicals Surveys and Monographs, L. Lewin (ed.), Amer. Math. Soc., Vol. **37**, 1992, pp. 205–231.
[557] Z. Wojtkowiak, On l-adic iterated integrals. II. Functional equations and l-adic iterated polylogarithms, *Nagoya Math. J.* **177** (2005), pp. 117–153.
[558] A.C.T. Wu, On the analytic properties of the 4-point function in perturbation theory, *Mat. Fys. Medd. Dan. Vid. Selsk.* **33**(3) (1961), pp. 1–88.
[559] B. Xia and T. Cai, Bernoulli numbers and congruences for harmonic sums, *Intl. J. of Number Theory* **6** (4) (2010), pp. 849855.
[560] S. Yamamoto, Explicit evaluation of certain sums of multiple zeta-star values, *Funct. Approx. Comment. Math.* **49** (2) (2013), pp. 283–289.

[561] S. Yamamoto, Interpolation of multiple zeta and zeta-star values, *J. Algebra*, **385** (2013), pp. 102–114.
[562] S. Yamamoto, Multiple zeta-star values and multiple integrals. arXiv:1405.6499.
[563] C. Yamazaki, On the duality for multiple zeta-star values of height one, *Kyushu J. Math.* **64**(1) (2010), pp. 145–152.
[564] S. Yasuda, Finite real multiple zeta values generate the whole space $\mathbb{Z}$. arXiv:1403.1134.
[565] A. J. Yee, A new shuffle convolution for multiple zeta values, *J. Algebraic Combin.* **21** (2005), pp. 55-69.
[566] J. Yu, Transcendence and special zeta values in characteristic p, *Ann. Math.* **134** (1991), pp. 1–23.
[567] H. Yuan and J. Zhao, Restricted sum formula of multiple zeta values, *Funct. Approx. Comment. Math.* **51** (1) (2014), pp. 111–119.
[568] H. Yuan and J. Zhao, New families of weighted sum formulas for multiple zeta values, *Rocky Mountain J. Math.* http://projecteuclid.org/euclid.rmjm/1407155212. arXiv:1303.3608.
[569] H. Yuan and J. Zhao, Double shuffle relations of double zeta values and double eisenstein series of level N. arXiv:1401.6699.
[570] H. Yuan and J. Zhao, Multiple divisor functions and multiple zeta values at level N. arXiv:1408.4983.
[571] D. Zagier, Hyperbolic manifolds and special values of Dedekind zeta-functions, *Inv. Math.* **83** (1986), pp. 285–301.
[572] D. Zagier, The remarkable dilogarithm, *J. Math. Phys. Sci.* **22** (1988), pp. 131–145.
[573] D. Zagier, The Block-Wigner-Ramakrishnan polylogarithm function, *Math. Ann.* **286** (1990), pp. 613–624.
[574] D. Zagier, Periods of modular forms and Jacobi theta functions, *Inv. Math.* **104** (1991), pp. 449-465.
[575] D. Zagier, Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields, in: Arithmetic Algebraic Geometry, *Prog. in Math.* Vol. **89** (1991), pp. 391–430, Birkhäuser, Boston.
[576] D. Zagier, Periods of modular forms, trace of Hecke operators, and multiple zeta values, *RIMS Kokyuroku* **843** (1993), pp. 162–170.
[577] D. Zagier, Values of zeta functions and their applications, in: First European Congress of Mathematics (Paris, 1992), Vol. **II**, A. Joseph et al. (eds.), Birkhäuser, Basel, 1994, pp. 497–512.
[578] D. Zagier, Quelques conséquences surprenantes de la cohomologie de $\mathrm{SL}(2,\mathbb{Z})$ in: *Leccons de Mathématiques d'aujourd'hui, Cassini, Paris* (2000), pp. 99–123.
[579] D. Zagier, The dilogarithm function, in: *Frontiers In Number Theory, Physics, And Geometry II*, 2007, Part I, pp. 3–65.
[580] D. Zagier, Multiple zêta values, (in French), lecture notes of a course at College de France, dated December 5, 2011.
[581] D. Zagier, Evaluation of the multiple zeta values $\zeta(2,\ldots,2,3,2,\ldots,2)$, *Ann. Math.* **175** (2012), pp. 977–1000.

[582] D. Zeilberger, The method of creative telescoping, *J. Symbolic Comput.* **11** (1991), pp. 195–204.

[583] D. Zeilberger, Maple Package EKHAD, available online `http://www.math.rutgers.edu/~zeilberg/tokhniot/EKHAD`

[584] W.P. Zhang, On the several identities of Riemann zeta-function, *Chinese Sci. Bull.* **36**(22) (1991), pp. 1852–1856.

[585] J. Zhao, Analytic continuation of multiple zeta functions, *Proc. Amer. Math. Soc.* **128**(1999), pp. 1275–1283.

[586] J. Zhao, Remarks on a Hopf algebra for defining motivic cohomology, *Duke Math. J.* **103**(2000), pp. 445–458.

[587] J. Zhao, Motivic complexes of weight three and pairs of simplices in projective 3-space, *Adv. Math.* **161**(2001), pp. 141–208.

[588] J. Zhao, Multiple polylogarithms: analytic continuation, monodromy, and variations of mixed Hodge structures, in: Contemporary Trends in Algebraic Geometry and Algebraic Topology, S.S. Chern, L. Fu and R. Hain eds., *Nankai Tracts in Math.,* Vol. **5**, pp. 167–193, World Scientific, 2002.

[589] J. Zhao, Multiple harmonic series I: generalizations of Wolstenholme's Theorem. arXiv:math/0301252. This is the unpublished version of [596].

[590] J. Zhao, Motivic cohomology of pairs of simplices, *Proc. London Math. Soc.* **88**(3) (2004), pp. 313–354.

[591] J. Zhao, Variations of mixed Hodge structures of multiple polylogarithms, *Canad. J. Math.* **56**(6) (2004), pp. 1308–1338.

[592] J. Zhao, Goncharov's relations in Bloch's higher Chow group $CH^3(F, 5)$. *J. Number Theory* **124** (2007), pp. 1–25.

[593] J. Zhao, Bernoulli numbers, Wolstenholme's Theorem, and p^5 variations of Lucas' Theorem, *J. Number Theory* **123** (2007), pp. 18–26.

[594] J. Zhao, Multiple q-zeta functions and multiple q-polylogarithms, *Ramanujan J.*, **14**(2) (2007), pp. 189–221. arXiv:math/0304448.

[595] J. Zhao, Analytic continuation of multiple polylogarithms, *Anal. Math.* **33** (2007), pp. 301–323.

[596] J. Zhao, Wolstenholme type Theorem for multiple harmonic sums, *Int. J. of Number Theory* **4**(1) (2008), pp. 73–106, arXiv:math/0301252.

[597] J. Zhao, Renormalization of multiple q-zeta values, *Acta Math. Sin. (Engl. Ser.)* **24**(10) (2008), pp. 1593–1616.

[598] J. Zhao, An exotic shuffle relation of multiple zeta values, *Arch. Math. (Basel)* **91**(5) (2008), pp. 409–415.

[599] J. Zhao, Multiple polylogarithm values at roots of unity, *C. R. Acad. Sci. Paris, Ser. I.* **346** (2008), pp. 1029–1032.

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

[601] J. Zhao, Standard relations of multiple polylogarithm values at roots of unity, *Doc. Math.* **15** (2010), pp. 1–34.

[602] J. Zhao, Multi-polylogs at twelfth roots of unity and special values of Witten multiple zeta function attached to the exceptional Lie algebra $\mathfrak{g}_2$, *J. Algebra Appl.* **9**(2) (2010), pp. 327–337.

[603] J. Zhao, A note on colored Tornheim's double series, *Integers* **10** (2010),

A59.

[604] J. Zhao, Witten volume formulas for semi-simple Lie algebras, *Integers* **11A** (2011), A22.

[605] J. Zhao, Alternating Euler sums and special values of Witten multiple zeta function attached to $\mathfrak{so}(5)$, *J. Aust. Math. Soc.* **89**(3) (2011), pp. 419–430.

[606] J. Zhao, Mod p structure of alternating and non-alternating multiple harmonic sums, *J. Théor. Nombres Bordeaux* **23**(1) (2011), pp. 259–268.

[607] J. Zhao, Finiteness of p-divisible sets of multiple harmonic sums, *Ann. Sci. Math. Québec.* **36** (2012), pp. 419–443.

[608] J. Zhao, Identity families of multiple harmonic sums and multiple zeta (star) families, to appear in *J. Math. Soc. Japan*. arXiv:1303.2227.

[609] J. Zhao, On q-analog of Wolstenholme type congruences for multiple harmonic sums, *Integers* **13** (2013), #A23 (11 pp).

[610] J. Zhao, Sum formula of multiple Hurwitz-zeta values, *Forum Mathematicum* **27** (2) (2015), pp. 929–936.

[611] J. Zhao, Restricted sum formula of alternating Euler sums, *Ramanujan J. Math* **36** (3) (2013), pp. 375-401.

[612] J. Zhao, Uniform Approach to double shuffle and duality relations of various q-analogs of multiple zeta values via Rota-Baxter algebras. arXiv:1412.8044.

[613] J. Zhao and X. Zhou, Witten multiple zeta values attached to $\mathfrak{sl}(4)$, *Tokyo J. Math.* **34** (2011), pp. 135–152.

[614] J. Zhao and X. Zhou, Reducibility of signed cyclic sums of Mordell-Tornheim zeta and L-values, *J. Ramanujan Math. Soc.* **26** (2011), pp. 383–414.

[615] X. Zhou and T. Cai, A generalization of a curious congruence on harmonic sums, *Proc. of Amer. Math. Soc.* **135** (2007), pp. 1329–1333.

[616] S. Zlobin, Generating functions for the values of a multiple zeta function, *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **2** (2005), pp. 55–59. English translation: *Moscow Univ. Math. Bull.* **60**(2) (2005), pp. 44–48.

[617] S. Zlobin, Special values of generalized polylogarithms. arXiv:0712.1656

[618] S. Zlobin, Relations for multiple zeta values, *Math. Notes* **84**(6) (2008), pp. 771–782.

[619] W. Zudilin, One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational, *Russian Math. Surveys* **56**(4) (2001), pp. 774-776.

[620] W. Zudilin, Algebraic relations for multiple zeta values, *Russian Math. Surveys* **58**(1) (2003), 1–29.

[621] W. Zudilin, Multiple q-zeta brackets, *Mathematics* **3** (2015), pp. 119–130.

List of Abbreviations

AMHS	alternating multiple harmonic sum
CMZV	colored multiple zeta value
DBSF	double shuffle relation
DIST	distribution relation
FES	finite Euler sum
FE⋆S	finite Euler star sum
FMZV	finite multiple zeta value
FMZ⋆V	finite multiple zeta star value
MCl	multiple Clausens
MCV	multiple Clausen value
MDV	multiple Deligne value
MGl	multiple Glaishers
MHS	multiple harmonic sum
MH⋆S	multiple harmonic star sum
MHS	mixed Hodge structures (only in Chap. 2)
MPL	multiple polylog(arithm)
MZV	multiple zeta value
MZ⋆V	multiple zeta star value
SVMPL	single-valued single-variable multiple polylog(arithm)
q-MHS	q-analog of multiple harmonic sum
q-MH⋆S	q-analog of multiple harmonic star sum
q-MZV	q-analog of multiple zeta value
q-MZ⋆V	q-analog of multiple zeta star value

List of Symbols

$(a)_n$, 299, 321
a, 153
$(a_p)_{p\in\mathcal{P}}$, 258
$(a;q)_n$, 299
A, 140, 515
A^{sv}, 140, 515
$\mathfrak{A}$, 86
$\mathfrak{A}^0$, **87**
$\mathfrak{A}^0_{\mathrm{G}}$, 405
$\mathfrak{A}^1$, 87
$\mathfrak{A}_{\sqcup\!\sqcup}$, 87
$\mathfrak{A}^1_{\sqcup\!\sqcup}$, 95
$\mathfrak{A}_*$, 90
$\mathfrak{A}^1_*$, 92
$\mathfrak{A}_\theta$, 390, 392
$\mathcal{A}_\ell$, 258
$A_n(F)$, 74
$A_{n,k}$, 302
$\aleph_1$, 259

b, 153
B_n, 26
$\widetilde{B}_k(x)$, 27
B_h, 509
$B_{[n]}$, 70
$\tilde{b}_n$, 312
β_k, **259**
$\beta_n(\lambda)$, 312

$\mathbb{C}$, xxiii
$\mathbb{C}\langle\langle\mathsf{X}\rangle\rangle$, 120
$\mathbb{C}^d$, 29
$\mathbb{CP}^1$, 39
$c(n,j)$, 168
C_{n+1}, 109

Δ_*, 539
$\mathrm{dp}(\mathbf{s})$, **27**, 153
$D_n(x)$, 39
DSh_n, 106, 237
Δ, 531, 535
$\Delta_{\sqcup\!\sqcup}$, 120
Δ_*, 120

e, **86**
$\mathbf{E}[f]$, 387
E_G, 506, 507
E_G^{ext}, 507
ϵ, 531, 539
$E_{\mathbf{t}}$, 535
$\widehat{\mathcal{E}}_w$, 199
$\epsilon_{\sqcup\!\sqcup}$, 120

∇f, 64
$f \mid M$, 105
$f \mid \sigma$, 106
$f^\sharp$, 106
$f_G(z)$, 511
$\nabla \mathcal{F}^p$, 62
$\mathcal{F}^{-k}$, 65
$F^{(n)}_{ш}(x_1, \ldots, x_n)$, 106, 229
$F^{(n)}_{ш}(x_1, \ldots, x_n)$, 106, 229
$\varphi_q[n]$, 319, **432**
Φ, 119
$\Phi(\mathbf{x}_0, \mathbf{x}_1)$, 119
ϕ_n, 247
ϕ^4, 506
$\Phi[\mathbf{w}]$, 119

$G_{\mathbf{w}}$, $\mathbf{w} \in \{0,1\}^*$, 510
$\Gamma(x)$, 26
$\Gamma_p(a/b)$, 260

$H^0(\mathbb{CP}^n, \Omega^1_{\mathbb{CP}^n}(\log(\mathcal{D}_n)))$, 64
$h_0(G)$, 506
$h_1(G) = h_G$, 506, 508
$H_n(s_1, s_2, \ldots, s_d)$, 244
$H^{\mathbf{t}}_n[\mathbf{s}]$, 299
$H^\star_n(s_1, s_2, \ldots, s_d)$, 244
$H^{\star,\mathbf{t}}_n[\mathbf{s}]$, 299
$\mathcal{H}_n[\mathbf{s}; \mathbf{t}; \mathbf{r}]$, 300
$\mathfrak{h}^{(\mathbf{t})}_n[\mathbf{s}]$, 312

$I(G)$, 508
$\mathcal{I}_n$, 46
$\int_\alpha w_1 \cdots w_r$, 37
$\mathbf{I}(t)$, 387
ι_ℓ, 258

$\mathbf{J}[f]$, 387

$K_n(p)$, 313

Ξ_G, 523, 524

$\mathrm{L}_{\mathbf{s}}(z)$, 100, 132, 265
$\mathrm{L}_{\mathbf{w}}(z)$, 134, 517, 522
$\mathscr{L}_{\mathbf{w}}(z)$, 137, 520
$\mathrm{L}_{ш}(\mathbf{w}; z)$, 265
$\mathbf{L}_{\mathrm{cv}}$, 123
$\mathrm{Li}_n(x)$, 38
$\mathscr{L}i_n(x)$, 39
$\mathrm{Li}_{s_1,\ldots,s_d}(\mathbf{x})$, 40
$\mathscr{L}i_{s_1,\ldots,s_n}(\mathbf{x})$, 70
$\mathrm{Ln}_s(z)$, **411**
$\mathfrak{L}_n(\mathbf{x})$, 46
Ls, Ls_n, 138, 515
$\mathsf{Ls}^{\mathrm{sv}}$, $\mathsf{Ls}^{\mathrm{sv}}_n$, 140
$\mathrm{LW}(\mathbf{w})$, 92
$\Lambda(u)$, **97**, 123
$\Lambda_n(L; M)$, 74

$\widetilde{\mathcal{M}}$, 106
$\mathrm{mcl}(\mathbf{s})$, 481
m_e, 507
$\mathrm{mgl}(\mathbf{s})$, 481
$\mathcal{M}_{[n]}(\mathbf{x})$, 49
$M_{\mathbf{t}}$, 248
M_w, 193
$\mathsf{MZ}_{\mathbb{Z}}$, 136
$\mathsf{MZ}^{\mathrm{sv}}_{\mathbb{Z}}$, 139, 515
MZ_w, 103, 229
$\mathsf{MZ}_{w,\leq d}$, 104, 229
$\mu(d)$, 95, 467
Γ_N, 441
μ_N, 313, 314, 435, 441

$\mathbb{N}$, xxiii
$[n]$, 299
$N_{\mathcal{D}}$, 65
n_G, 506, 507
$\left[{n \atop m}\right]$, 299

$\nu(x_j)$, 109

O, 136, 140, 515
$\overline{\mathsf{O}}$, 136, 140, 515
Ω_G, 525
$\boldsymbol{\omega}$, 63
ω_L, 74

$\mathcal{P}$, 258
$\mathbf{P}[f]$, 387
$\widehat{P}_f(X,Y)$, 198
$P_f(X,Y)$, 193
$\oplus$, 243
$\boxplus$, 300
π_{cv}, 123

$\mathbb{Q}$, xxiii
q_i, 507
$\mathfrak{q}_k$, 259
$Q(r,k)$, 300
$\mathbb{Q}\langle \mathtt{x}_0, \mathtt{x}_1 \rangle$, 86

$\mathbb{R}$, xxiii
$\mathcal{R}$, 120
$\mathcal{R}_{\sqcup\!\sqcup}$, 120
$\mathcal{R}_*$, 120
$\mathcal{R}_{\mathrm{cv}}$, 123
$\mathcal{R}^T_{\sqcup\!\sqcup}$, 122
$\mathcal{R}^T_*$, 121
$r(a,b,c,d)$, 81
$\mathrm{Re}(s)$, 27
$\mathbf{R}[f]$, 387
ρ, **96**, 124
$\rho_{\mathbf{i}}$, 48

S_d, 29
$\mathcal{S}_n$, 49
$\mathcal{S}'_n$, 47
$S^\sharp$, 106
$\mathbf{s}^\vee$, 248, 269
$\mathbf{s}^*$, 153
$\mathfrak{S}_n$, 106, 251
$(s)_M$, 28
$\mathcal{S}$, **363**
$\mathcal{S}_j$, 363
$\mathsf{ShC}_n(d)$, 108, **109**
Sh_j, 106
sh_j, 106
$s(n,j)$, 312
S_w, 193
$\sqcup\!\sqcup$, 87
$\psi(x)$, **26**, 223
$\sigma_k(n)$, 410

T', 110
$\tau_n(\lambda)$, 70
$\tau_{1,n+1}$, 110
$T_{\mathcal{D}}$, 65
T_j, 109
Tog, 245
tog, 265
θ, 300

V_G, 506, 507
V_n, 105
$V_{[n]}(\mathbf{x})$, 64
V_w^{ev}, 189
V_w^{od}, 189
$V_w^{\pm}$, 189

W_{-2k}, 65
WS_{n+1}, 520
W_w, 193
W_w^{ev}, 191
W_w^{od}, 191
$W_w^{\pm}$, **191**
$\widehat{W}_w$, 198
$\widehat{W}_w^{\pm}$, 198

$\mathtt{x}_0, \mathtt{x}_1$, 85, 266
X_π, 392
X_π^*, 392
X_θ, 390
X_θ^*, 390

$\mathtt{y}_k$, **87**, 91, 120, 244
$\mathtt{y}_{k,\xi}$, 441
Y_{G}, 406
$\mathsf{Y}_{\tilde{\mathrm{I}}}$, 390
Y_{II}, 390
$\mathsf{Y}_{\mathrm{III}}$, 390

$\mathbb{Z}$, xxiii
$\mathtt{z}_k$, 244, 390
$\mathtt{z}_{\bar{k}}$, 244
$\mathtt{z}'_k$, 390
$\mathtt{z}_{k,i}$, 437
$\zeta(s)$, 25
$\zeta(\mathbf{s})$, 27
$\zeta^\star(\mathbf{s})$, 27
$\zeta_{\sqcup\!\sqcup}$, 87
ζ_*, 90
$\zeta_*(\mathbf{w}; T)$, 95
$\zeta_{\sqcup\!\sqcup}(\mathbf{w}; T)$, 95
$\zeta^{\mathcal{S}}_{\sqcup\!\sqcup}(\mathbf{s})$, 217
$\zeta^{\mathcal{S}}_*(\mathbf{s})$, 217
$\zeta_{\mathcal{A}_\ell}(\mathbf{s})$, 260
$\zeta^\star_{\mathcal{A}_\ell}(\mathbf{s})$, 260
$\zeta_F(s)$, 40
ζ_M, 98
$\zeta_q^{\mathbf{t}}[\mathbf{s}]$, 361
$\zeta_q^{\mathrm{I}}[\mathbf{s}]$, 385
$\zeta_q^{\mathrm{II}}[\mathbf{s}]$, 385
$\zeta_q^{\mathrm{III}}[\mathbf{s}]$, 385
$\mathfrak{z}_q$, 392, 394
$\mathfrak{z}_q^{\mathrm{O}}[\mathbf{s}]$, 407
$\mathfrak{z}_q^{\mathbf{t}}[\mathbf{s}]$, 394
$\mathfrak{z}_q^{\tau}[\mathbf{s}]$, $(\tau = \mathrm{I}, \mathrm{II}, \mathrm{III})$, 397
Z_n, 511, 512, 521

Index

abc-conjecture, 259
admissible word
  Hoffman algebra, 87
  Hoffman-Racinet algebra, 245, 471

Bernoulli numbers, **26**, 160
  $\mathcal{A}_1$-Bernoulli numbers, 259
  $\mathcal{A}_1$-Bernoulli numbers, 259
  degenerate, 312
  of the second kind, **312**
  q-analog, 321, **372**
Bernoulli polynomials, 27, 160
  periodic, **27**, 365
bialgebra, 532
Bloch group, 77

Chen's iterated integrals, 37
coalgebra, 120, 531
  group-like, 120
colored multiple zeta values, **435**
  depth, 435
  finite double shuffle relations, 440
  level, 435
  lifted relations, **454**
  non-standard relations, 454, 477
  regularized distribution relations, 451
  regularized double shuffle relations, 436, **440**
  standard relations, 437, **454**, 493
  weight, 435
  weight one relations, 450
congruence, 257
  q-analog, **311**
  super congruence, 257
  van Hamme type, 254, **257**
cusp form, **193**
cusp forms, 209

digamma function, 26, 223, 339
  duplication relation, 26
  recurrence relation, 26
  reflection relation, 26
divisor sum, 410
Drinfeld associator, **119**, 120
  duality relation, 125
  generalized, **457**
    mixed pentagon relations,

147
octagon relations, 147
generalized Drinfeld associator, 147
group-like, 121, 148
hexagon relation, 125, 147
mixed pentagon relations, 458
octagon relations, 459
pentagon relation, 125, 131, 147
special action condition, **132**, 148
special derivation condition, 459

Eisenstein series, **193**, 410
Emultiple zeta values
finite version
double shuffle relations, 267
Euler sums, 206, 259
doubling relations, **476**
finite version, 259, **260**
double shuffle relations, 267
shuffle relations, 267
stuffle relations, 264
generalized doubling relations, **477**
Euler-Maclaurin summation formula, 28, 364
Eulerian polynomial, **410**

Fermat quotient, $\mathcal{A}_1$-, **259**, 293
Feynman diagram, **505**
Feynman graph, **507**
box ladder graph, **509**, 512
first Symanzik polynomial, 523
graphical function, 511
kinematics data, 507
loop number, 506
momentum conservation, 507, 508
period, **508**, 509
second Symanzik polynomial, 523
sequential graph, 510
vacuum graph, 507
virtual momentum, 508
wheel graph, **520**
zig-zag graph, 511
Feynman integral, **505**
dimension renormalization, 508
infrared (IR) divergence, 508
ultraviolet (UV) divergence, 508

Gamma function, 97
Grothendieck-Teichmüller group, 125
generalized, 459
special action condition, **126**
Grothendieck-Teichmüller Lie algebra, **125**
duality relation, 125
hexagon relation, 125
pentagon relation, 125
special derivation condition, 125

Hoffman algebra, **86**, 157, 244, 390, 471
admissible word, 87
derivation, 170, 171
Lyndon factorization, 92
q-analog, 390
shuffle algebra, 87

stuffle algebra, 90
word, **86**
admissible word, 87
depth, **87**
Lyndon word, **91**
weight, **87**
Hoffman-Racinet algebra, **244**, **438**, 471
admissible word, **245**, 438
level two, 245, 247
converting rule, 245, 246
toggling operator, 246
regularization, 446
shuffle product, 438
stuffle product, 439
Hopf algebra, 103, **532**, 535
antipode, 249, **532**, 535
bialgebra, **532**
co-ideal, 446, **533**
coalgebra, **531**
cocommutative, **532**
commutative, 103, **532**
coproduct, **531**, 534, 539
counit, 226, **531**, 535
graded, **534**
group-like, **533**
group-like element, 122, 457
primitive element, 226, **533**, 537

Jackson q-integral, 387

KZ differential equation, 118, 127, 148, 457

Lie algebra, 103, 104, 114, 127, 147, 170, **537**
Lie bracket, 125, 149, **537**
Lie element, 226
Poincaré-Birkhoff-Witt Theorem, 147, 149, 541

Möbius function, **95**, 114, 467
Möbius inversion formula, 114
modular forms, **192**, 209
mollified companion, 300
mollifiers, 300
admissible triple of, 300
mono-bracket, **410**
Mordell-Tornheim L-functions, 31
multiple Clausen value, **480**
multiple Clausens, **481**
multiple Dedekind zeta functions, 30
multiple Deligne value, **480**
multiple divisor sums, **410**
multiple Eisenstein series, **420**
multiple Glaishers, **481**
multiple harmonic star sums, **244**, 299
alternating version, 243, 250
symmetric sums, 243, 250
binomial identities, 253
concatenation relations, 250
multiple harmonic sums, **244**, 299
alternating version, **244**
alternating harmonic numbers, 244
converting rule, 245, 246
O-plus, 243
signed numbers, 243
stuffle relations, 246
symmetric sums, 243, 250
unsigned numbers, 243
binomial identities, 251
concatenation relations, 250

depth, 244
harmonic numbers, 244
q-analogs
□-plus, 300
weight, 244
multiple Hurwitz zeta functions, 31
multiple polylog, **40**, 435
Aomoto polylogs, 73
depth, **40**
double polylogs, 55
limit mixed Hodge structures, 67
single-valued version, 72
multiple logarithms, 46
analytic continuation, 47
limit mixed Hodge structures, 65
monodromy, 53
one-variable version, 465
single-valued version, 70
single-variable version, **100**, 132, 265
single-valued, 137
variation matrix, 54
variations of mixed Hodge structures, 61
multiple zeta functions, **26**
analytic continuation, 27
Barnes' version, 35
depth, **27**
Kurokawa's version, 35
of function fields over finite fields, 33
p-adic version, 30, 35
q-analog, **385**
Witten's version, 32
volume formulas, 32
multiple zeta star functions, **27**, **325**
multiple zeta star values, 208, 252, **325**, 355
cyclic sum relations, 329
depth, 325
duality relations, 339
finite version, **260**
concatenation relations, 264
reversal relations, 261
fixed weight, height and depth relations, 332
sum relations, 328
weight, 325
weighted sum relations, 330
multiple zeta value
Hoffman element, 143
multiple zeta values, 85
codepth, 169
cyclic sum relations, 156, 208
depth, **27**, 87, 153, 272
derivation relations, 172, 208
dimension conjecture, 105
double shuffle relations (DBSF), 91, 208
finite, 91
Hoffman relations, 102
regularized, 91, 101, **102**
double shuffle space, 106, 237
dual, **153**
duality relations, 153
Euler's decomposition formula, 85
finite version, 241, **260**
φ-dual, 271
concatenation relations, 264, 273, 290
reversal relations, 261
shuffle relations, 267

superbity, **260**
v-dual, **248**, 269
fixed weight, height and depth relations, 177
height, 210
Hoffman elements, 113
Hoffman's Basis Conjecture, 113
Ohno's relations, 176, 208
parity principle, 108, **110**, 114
q-analog
bi-brackets, **421**
mono-brackets, **410**
q-zeta brackets, **422**
Okounkov's, 386
type I, II and III, 385
regularization map, 173, 174
regularizations, 95
restricted sum relations, 159, 169
$\sqcup\!\sqcup$-symmetrized, 217
duality, 222
shuffle relations, 226
shuffle relations, 86, 121
single-valued, 139
$*$-symmetrized, 217
stuffle relations, 219, 225
sum formula, 220
stuffle relations, 90, 124, 162
harmonic shuffle, 112
quasi-shuffle, 112
sticky shuffle, 112
sum relations, 157, 208, 254
symmetrized, 225, 228, 241
φ-dual, 226, 227, 271
weight, **27**, 87
weighted sum relations, 158, 207, 208

p-adic Γ values, 260
Padovan numbers, 103
period polynomial, **193**
Pochhammer symbol, 28, 207
quantum, **299**
polylogarithm (polylog), **38**, 411
dilogarithm, 39
five term relations, 40
infinitesimal version, 81
Kontsevich's finite version, 81
normalized polylog, **411**
p-adic version, 81
Ramakrishnan's single-valued version, 39
trilogarithms, 40
22-term relations, 40
Zagier's single-valued version, 39

quantum factorial, **299**
quasi-symmetric functions, **247**
signed powers, 247

Riemann zeta function, 25

shifting operators, **363**
Stirling numbers of the first kind, **312**
unsigned, 168
symmetric group, 106

Two-one formula, 253, 290, **343**
q-analog, **310**
Two-three formula, **253**, 289, **343**
q-analog, **311**

Wieferich primes, 259
Wilson quotients, 260

Wolstenholme's Theorem, 255, **257**
q-analog, **311**

Zagier's Conjecture
of Dedekind zeta functions, 40
of MZVs, **103**, 274