Multiple zeta values and modular forms

Henrik Bachmann
Nagoya University
Topics in Mathematical Science III, Spring semester 2020

Contact: Math. Building Room 457, henrik.bachmann@math.nagoya-u.ac.jp
Lecture notes and videos are available at: https://www.henrikbachmann.com/mzv2020.html

These notes are under construction and therefore may contain mistakes and change without notice. If you find any typos/errors or have any suggestion, please let me know! Already a big thanks to: Ulf Kühn, Nils Matthes, Yuta Suzuki and Can Turan.

Contents

1 Overview & Basics 3
   1.1 The values of the Riemann zeta function 3
   1.2 Multiple zeta values 5
   1.3 Modular forms and the Broadhurst-Kreimer conjecture 9
   1.4 q-analogues of multiple zeta values 12

2 Algebraic setup 19
   2.1 Multiple polylogarithms, iterated integrals and duality 19
   2.2 The shuffle & stuffle product and finite double shuffle relations 23
   2.3 Quasi-shuffle algebras 27
   2.4 Regularizations 38

3 Families of linear relations and their q-relatives 43
   3.1 Extended double shuffle relations 43
   3.2 Seki-Yamamoto’s connected sums and Ohno’s relation 45
   3.3 The zoo of relations 50

4 Double zeta values and modular forms 57
   4.1 The formal double zeta space 57
   4.2 Period polynomial relations 69
   4.3 Double Eisenstein series 69

5 Multiple Eisenstein series and q-analogues of MZV 69
   5.1 The Fourier expansion of multiple Eisenstein series 69
   5.2 q-analogues of MZV 69
   5.3 Double indexed q-analogues of MZV 69
   5.4 Double shuffle relations for double indexed q-analogues 69
   5.5 Combinatorial approach to modular forms 69

References 70

Exercises 70

Version 3.9 (June 26, 2020)
Introduction

In this course, we are interested in a multiple version of the Riemann zeta function and in the connection of their values at positive integer points to modular forms. We will mainly deal with three types of objects: Multiple zeta values (real numbers), modular forms (holomorphic functions in the complex upper half plane) and $q$-analogues of multiple zeta values ($q$-series with rational coefficients). The goal of this lecture is to describe some of the relationships between these objects, which were initiated by the beautiful work [GKZ]. There are various points of views from which one can study multiple zeta values and we will just be able to cover a small part of their exciting aspects. For more details on multiple zeta values we refer in particular to the books/lecture notes of Arakawa-Kaneko [AK], Burgos-Frésan [BF], Waldschmidt [W], Zhao [Zh1] and to the collection of Hoffman on research papers of multiple zeta values [H0].

§1 Overview & Basics

1.1 Riemann zeta function

\[ \zeta(k_1) \zeta(k_2) \]

\[ \zeta(3) = \zeta(2, 1) \]

\[ \int_0^1 dx \int_0^x \frac{dy}{1-y} \zeta(s) \rightarrow \mathbb{R} \]

1.2 Multiple zeta values $\zeta(k) \in \mathbb{R}$

\[ \zeta(v \uplus w - v \ast w) = 0 \]

\[ \zeta(\tau(w)) = \zeta(\tau) \]

1.3 $G_k \in \mathcal{O}(\mathbb{H})$

\[ \dim_{\mathbb{C}} S_k \]

1.4 $g(k) \in \mathbb{Q}[[q]]$

\[ g(k_1) g(k_2) \]

A rough overview of the structure of these notes and the course.

The plan of this course is as follows: We start in Section 1 with an overview of almost all the objects we will deal with. First, we discuss the Riemann zeta function in 1.1, then introduce multiple zeta values (MZVs) in 1.2 and discuss some of their conjectures, algebraic structure and linear relations, which we then discuss in detail in Section 2 and 3. After this, we recall some basic facts on modular forms in 1.3 and state the Broadhurst-Kreimer conjecture, which gives the first indication of a connection of cusp forms and multiple zeta values. This connection will then be made precise in Section 4, where we present the main results of [GKZ]. In Section 1.4, we will introduce (some) $q$-analogues of multiple zeta values and show that they have a similar algebraic structure as multiple zeta values. This algebraic structure and its consequences will then be further described in Section 2 and 5.
§1 Overview & Basics

In this section, we will give an overview of the values of the Riemann zeta function, the definition of multiple zeta values, some of the main conjectures concerning their structure, and a glimpse of the connection of \((q\text{-analogues of})\) multiple zeta values and modular forms. The general picture of these concepts will be discussed in detail in the later Sections.

1.1 The values of the Riemann zeta function

The Riemann zeta function is defined for a complex variable \(s \in \mathbb{C}\) with \(\text{Re}(s) > 1\) by

\[
\zeta(s) = \sum_{m>0} \frac{1}{m^s}.
\]

This function appears in various fields of mathematics and theoretical physics and it can be studied from various points of views. It plays a pivotal role in analytic number theory and has applications in physics, probability theory, and applied statistics.

For example, it is well-known that the Riemann zeta function can be analytically continued to the whole complex plane with a simple pole at \(s = 1\). Even though \(\zeta(s)\) was already considered by L. Euler (1707 – 1783), it was named after B. Riemann (1826 – 1866), who proved its meromorphic continuation and functional equation and established a relation between its zeros and the distribution of prime numbers.

In particular, he gave his famous conjecture on the location of the zeros of the Riemann zeta function, stating that besides the trivial zeros at \(s = -2, -4, -6, \ldots\) all other zeros have real part \(\frac{1}{2}\).

The connection to prime numbers is given by the following product formula, which, to make things fair again, was named after Euler (Euler product formula)

\[
\zeta(s) = \prod_{p\text{ prime}} \frac{1}{1-p^{-s}}. \quad (\text{Re}(s) > 1)
\]

In this course, we will not study these analytic aspects but rather will be interested in the values of \(\zeta(k)\), when \(k \in \mathbb{Z}_{\geq 2}\) is a positive integer. The first result is the famous formula by Euler for \(\zeta(2)\), which states that

\[
\zeta(2) = \sum_{m>0} \frac{1}{m^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots = \frac{\pi^2}{6}.
\]
In general Euler proved that $\zeta(2m)$ is always a rational multiple of $\pi^{2m}$ and he gave the following explicit formula in terms of Bernoulli numbers.

**Proposition 1.1** (Euler, 1734). For all $m \in \mathbb{Z}_{\geq 1}$ we have

$$\zeta(2m) = \frac{B_{2m}}{2(2m)!} (2\pi i)^{2m} \in \mathbb{Q}\pi^{2m},$$

where $B_n$ denotes the $n$-th Bernoulli number defined by the Taylor expansion

$$\sum_{n=0}^{\infty} B_n \frac{x^n}{n!} = \frac{x}{e^x - 1}. \quad (1.2)$$

**Proof.** There are various ways to prove this fact and we will give the original approach due to Euler. First, consider the Weierstrass product of the sine function

$$\frac{\sin(\pi x)}{\pi x} = \prod_{n \geq 1} \left(1 - \frac{x^2}{n^2}\right). \quad (1.3)$$

For $x \in \mathbb{C}\setminus\mathbb{Z}$ we can take its logarithmic derivative to obtain the partial fraction expansion of the cotangent

$$\pi \cot(\pi x) = \frac{1}{x} + \sum_{n \geq 1} \left(\frac{1}{x+n} + \frac{1}{x-n}\right).$$

Expanding the right hand side in a geometric series gives

$$\frac{1}{x} + \sum_{n \geq 1} \left(\frac{1}{x+n} + \frac{1}{x-n}\right) = \frac{1}{x} - \sum_{m=1}^{\infty} 2\zeta(2m)x^{2m-1}.$$ 

On the other hand the left hand side can be evaluated as

$$\pi \cot(\pi x) = \pi i \frac{e^{i\pi x} + e^{-i\pi x}}{e^{i\pi x} - e^{-i\pi x}} = \pi i \left(1 + \frac{2}{e^{2i\pi x} - 1}\right) \frac{1}{x} + \sum_{m=1}^{\infty} \frac{B_{2m}(2\pi i)^{2m} x^{2m-1}}{(2m)!},$$

where in the last equality we used $B_1 = -\frac{1}{2}$.

The first explicit values for $\zeta(2m)$ are given by the following

$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945}, \quad \zeta(8) = \frac{\pi^8}{9450}, \quad \zeta(10) = \frac{\pi^{10}}{93555}, \quad \zeta(12) = \frac{691\pi^{12}}{638512875}.$$ 

Since $\pi$ is transcendental (Lindemann, 1882), Proposition 1.1 gives the only family of polynomial relations among even zeta values. On the other hand, one does not expect polynomial relations among odd zetas. This is part of the following folklore conjecture.

**Conjecture 1.2.** The numbers $1, \pi^2, \zeta(3), \zeta(5), \zeta(7), \ldots$ are algebraically independent.

So far there is not much known towards this conjecture. For the odd zeta values the following theorem gives an overview over the known facts.

**Theorem 1.3.**  i) $\zeta(3)$ is irrational. (Apéry, 1978)

ii) For $m \geq 1$ we have

$$\text{dim}_\mathbb{Q}\langle 1, \zeta(3), \ldots, \zeta(2m+1) \rangle \geq \frac{1}{3} \log(2m+1).$$

In particular infinitely many of the values $\zeta(2m+1)$ are irrational. (Ball–Rivoal, 2001)

iii) At least one of the values $\zeta(5), \zeta(7), \zeta(9)$ and $\zeta(11)$ is irrational. (Zudilin, 2001)
1.2 Multiple zeta values

Due to Conjecture 1.2 we do not expect any linear relations among the values \( \zeta(k_1)\zeta(k_2) \) if one of the \( k_i \) is odd. But it turns out that certain parts of these products satisfy numerous relations among each other. Splitting the product \( \zeta(k_1)\zeta(k_2) \) into the following three parts leads us to the definition of the double zeta values \( \zeta(k_1, k_2) \)

\[
\zeta(k_1)\zeta(k_2) = \sum_{m_1 > 0} \frac{1}{m_1} \sum_{m_2 > 0} \frac{1}{m_2} = \left( \sum_{m_1 > 0} \frac{1}{m_1} \sum_{m_2 > m_1 > 0} \frac{1}{m_2} + \sum_{m_1 > m_2 > 0} \frac{1}{m_1 m_2} \right) \frac{1}{m_1 m_2} \quad (k_1, k_2 \geq 2)
\]

\[
=: \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2).
\]

For example we have the following expressions for the products of Riemann zeta values

\[
\zeta(2)\zeta(5) = \zeta(2, 5) + \zeta(5, 2) + \zeta(7), \quad \zeta(3)\zeta(4) = \zeta(3, 4) + \zeta(4, 3) + \zeta(7).
\]

Even though we expect that there are no linear relations among \( \zeta(2)\zeta(5) \) and \( \zeta(3)\zeta(4) \), their "building blocks" given by \( \zeta(7), \zeta(2, 5), \zeta(3, 4), \zeta(4, 3) \) and \( \zeta(5, 2) \) satisfy various relations among each other. For example we will see [Exercise 1] that

\[
\zeta(7) = 4\zeta(3, 4) + 3\zeta(4, 3) - 2\zeta(5, 2).
\]

Considering product of more than just two zeta values and using the same idea as in (1.4) leads us for integers \( k_1, \ldots, k_r \) to sums of the form

\[
\zeta(k_1, \ldots, k_r) = \sum_{m_1 > \ldots > m_r > 0} \frac{1}{m_1^{k_1} \ldots m_r^{k_r}}.
\]

**Proposition 1.4.** For integers \( k_1 \geq 2, k_2, \ldots, k_r \geq 1 \) the sum (1.6) converges.

**Proof.** It is enough to show the convergence for \( k_1 = 2 \) and \( k_2 = \cdots = k_r = 1 \) for any \( r \), since this gives an estimate for the other cases. Using the well-known inequality \( \sum_{n=1}^{m} \frac{1}{n} \leq 1 + \log(m) \) we obtain

\[
\sum_{m_1 > m_2 > \cdots > m_r > 0} \frac{1}{m_1^{2} m_2 \cdots m_r} = \sum_{m=1}^{\infty} \frac{1}{m^2} \sum_{m_2 > m_1 > 0} \frac{1}{m_2 \cdots m_r} \leq \sum_{m=1}^{\infty} \frac{1}{m^2} (1 + \log(m))^{r-1}
\]

and since \( (1 + \log(m))^{r-1} = o(\sqrt{m}) \) as \( m \to \infty \) for any \( r \), the above sum converges.

The multiple sum (1.6) will give the definition of the multiple zeta values which we will give after introducing the following notation.

**Definition 1.5.**

i) For \( r \geq 0 \) we call a tuple \( k = (k_1, \ldots, k_r) \in \mathbb{Z}_{>1}^r \) of positive integers an index. For \( r = 0 \) we write \( k = \emptyset \) and refer to it as the empty index.

ii) An index \( k = (k_1, \ldots, k_r) \) is called admissible if \( k_1 \geq 2 \) or \( k = \emptyset \).

iii) For an index \( k = (k_1, \ldots, k_r) \) we call \( \text{wt}(k) = k_1 + \cdots + k_r \) its weight and \( \text{dep}(k) = r \) its depth.

We set \( \text{wt}(\emptyset) = \text{dep}(\emptyset) = 0 \).

**Definition 1.6.** For an admissible index \( k = (k_1, \ldots, k_r) \) we define the multiple zeta value \( \zeta(k) \) by

\[
\zeta(k) = \zeta(k_1, \ldots, k_r) = \sum_{m_1 > \ldots > m_r > 0} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \in \mathbb{R}
\]

and \( \zeta(\emptyset) = 1 \). In the case \( r = 1 \) (resp. \( r = 2 \)) we refer to these as single (resp. double) zeta values.
Remark 1.7. i) By Proposition 1.4, the \( \zeta(k) \) gives for every admissible index \( k \) a real number. Even though the notion of weight and depth for these real numbers might not be well defined (and indeed we will see already in Proposition 1.8 below that this is not the case for the depth), we also say that \( \zeta(k) = \zeta(k_1, \ldots, k_r) \) has weight \( \text{wt}(k) = k_1 + \cdots + k_r \) and depth \( \text{dep}(k) = r \).

ii) For \( r = 1 \) the multiple zeta values are given by the values of the Riemann zeta function. One can also define multiple zeta functions \( \zeta(s_1, \ldots, s_r) \) for complex variables \( s_1, \ldots, s_r \in \mathbb{C} \) and consider their analytic properties similar to the classical case. See for example the thesis of Onozuka [On] for a nice detailed survey or [Zh1].

As we have seen before in an example, multiple zeta values satisfy various linear relations. The first one appears in weight 3 and is originally due to Euler. During the course we will see several ways to prove it and the interested reader can find 32 ways of doing so in [BB].

Proposition 1.8. We have \( \zeta(3) = \zeta(2, 1) \).

Proof. The shortest proof known to the author is the following: Consider the following sum

\[
S = \sum_{m, n > 0} \frac{1}{mn(m+n)} = \sum_{m, n > 0} \frac{1}{n^2} \left( \frac{1}{m} - \frac{1}{m+n} \right) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{m=1}^{n} \frac{1}{m} = \zeta(3) + \zeta(2, 1).
\]

This sum can also be evaluated as follows

\[
S = \sum_{m, n > 0} \left( \frac{1}{n} + \frac{1}{m} \right) \frac{1}{(m+n)^2} = \sum_{m, n > 0} \frac{1}{n(m+n)^2} + \sum_{m, n > 0} \frac{1}{m(m+n)^2} = 2\zeta(2, 1)
\]

and therefore the relation \( \zeta(3) = \zeta(2, 1) \) follows.

Another way to obtain relations among multiple zeta values is to evaluate the product \( \zeta(k_1)\zeta(k_2) \) in two different ways. In \([1.4]\) we saw that for \( k_1, k_2 \geq 2 \)

\[
\zeta(k_1)\zeta(k_2) = \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2), \tag{1.7}
\]

which is often called the stuffle product (also called harmonic product). But we also have the following expression for the product, which is called the shuffle product.

Proposition 1.9. For \( k_1, k_2 \geq 2 \) we have

\[
\zeta(k_1)\zeta(k_2) = \sum_{j=2}^{k_1+k_2-1} \left( \binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) \zeta(j, k_1 + k_2 - j), \tag{1.8}
\]

where we use the usual convention \( \binom{n}{k} = 0 \) for \( n < k \).

Proof. This is Exercise 1 and can be done by using partial fraction decomposition.

Comparing the right hand sides of (1.7) and (1.8) gives for \( k_1, k_2 \geq 2 \) a linear relation among multiple zeta values, which is an example for a so-called (finite) double shuffle relation. We will consider the stuffle/harmonic & shuffle product and the resulting double shuffle relations in detail for arbitrary depth in Section 2 and 3.
Example 1.10. For \(k_1 = 2, k_2 = 3\) equations (1.7) and (1.8) give
\[
\begin{align*}
\zeta(2)\zeta(3) &= \zeta(2, 3) + \zeta(3, 2) + \zeta(5), \\
\zeta(2)\zeta(3) &= \zeta(2, 3) + 3\zeta(3, 2) + 6\zeta(4, 1),
\end{align*}
\]
from which we deduce the linear relation \(\zeta(5) = 2\zeta(3, 2) + 6\zeta(4, 1)\).

We will denote the \(\mathbb{Q}\)-vector space spanned by all multiple zeta values by
\[
\mathcal{Z} = \langle \zeta(k) \mid k \text{ admissible} \rangle_{\mathbb{Q}}.
\]

For a fixed weight \(k \geq 0\) we also define the space of weight \(k\) multiple zeta values by
\[
\mathcal{Z}_k = \langle \zeta(k) \mid k \text{ admissible}, \text{wt}(k) = k \rangle_{\mathbb{Q}}.
\]

Clearly we have \(\mathcal{Z} = \sum_{k \geq 0} \mathcal{Z}_k\). With the same idea as in (1.4), where we showed that \(\zeta(k_1)\zeta(k_2)\) is a linear combination of multiple zeta values of weight \(k_1 + k_2\), we will see in Section 2 that this is true for arbitrary products of multiple zeta values and we will show the following (see Corollary 2.10).

Proposition 1.11. The space \(\mathcal{Z}\) is a \(\mathbb{Q}\)-subalgebra of \(\mathbb{R}\) and we have \(\mathcal{Z}_{k_1} \cdot \mathcal{Z}_{k_2} \subseteq \mathcal{Z}_{k_1 + k_2}\) for \(k_1, k_2 \geq 0\).

All of the relations we saw so far: \(\zeta(3) = \zeta(2, 1)\), the relation (1.5), and the finite double shuffle relations are relations among multiple zeta values of the same weight. Indeed it is expected that there exist no \(\mathbb{Q}\)-linear relations among multiple zeta values of different weights, which is part of the following conjecture.

Conjecture 1.12. The space \(\mathcal{Z}\) is graded by weight, i.e.
\[
\mathcal{Z} = \bigoplus_{k \geq 0} \mathcal{Z}_k.
\]

This conjecture is very strong as it implies (Exercise 2) the transcendence of every multiple zeta value of non-zero weight. One of the main interest in the theory of multiple zeta values is to understand all of their \(\mathbb{Q}\)-linear relations. There are several families of relations which conjecturally give all linear relations among multiple zeta values in a fixed weight. We will describe some of them in Section 3. In particular, we have a conjecture for the dimension of the spaces \(\mathcal{Z}_k\), which was first observed by Zagier based on extensive numerical calculations. To state the conjecture we first introduce the integers \(d_k\) given by the following generating series
\[
\sum_{k \geq 0} d_k X^k = \frac{1}{1 - X^2 - X^3},
\]
i.e. they are given by \(d_0 = 1, d_1 = 0, d_2 = 1\) and the recursion \(d_k = d_{k-2} + d_{k-3}\) for \(k \geq 3\).

Conjecture 1.13 (Zagier, 1994). We have \(\dim_{\mathbb{Q}} \mathcal{Z}_k = d_k\) for all \(k \geq 0\).

This conjecture shows that multiple zeta values satisfy a lot of linear relations. For example in weight \(k = 14\) there are \(2^{12} = 4096\) admissible indices (i.e. generators of \(\mathcal{Z}_{14}\)) and the conjectured dimension is \(d_{14} = 21\). In the following we give a table for the number of admissible indices, the conjectured number of linearly independent relations and the numbers \(d_k\).

<table>
<thead>
<tr>
<th>weight (k)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td># of adm. ind.</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>64</td>
<td>128</td>
<td>256</td>
<td>512</td>
<td>1024</td>
<td>2048</td>
<td>4096</td>
</tr>
<tr>
<td># of relations (\geq)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>29</td>
<td>60</td>
<td>123</td>
<td>249</td>
<td>503</td>
<td>1012</td>
<td>2032</td>
<td>4075</td>
</tr>
<tr>
<td>(\dim_{\mathbb{Q}} \mathcal{Z}_k \geq) (d_k)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>12</td>
<td>16</td>
<td>21</td>
</tr>
</tbody>
</table>
Remark 1.14. One easy way to do numerical experiments for multiple zeta values is to use PARI/GP, which was actually the tool used by Zagier to come up with Conjecture 1.13 (It is available for free at https://pari.math.u-bordeaux.fr/). There the multiple zeta value \( \zeta(k_1, \ldots, k_r) \) is implemented as `zetamult([k_1, \ldots, k_r])` and the Riemann zeta function \( \zeta(s) \) as `zetamult([s])`, which is of course the same as `zetamult([1])`. Together with the function `lindep([v_1, \ldots, v_l])` one can search for linear relations among the values \( v_1, \ldots, v_l \). For example to check if there is a relation between \( \zeta(2,1) \) and \( \zeta(3) \) one enters

\[
\text{input: } \text{lindep([zetamult([2,1]),zeta(3)])} \\
\text{output: } [-1, 1]~
\]

The output gives the coefficient of the relation \(-1 \cdot \zeta(2,1) + 1 \cdot \zeta(3) = 0\). It is unknown if Euler also used PARI/GP to come up with \( \zeta(2) = \frac{\pi^2}{6} \), but he could have done so by using

\[
\text{input: } \text{lindep([zeta(2),Pi^2])} \\
\text{output: } [-6, 1]~
\]

Of course these numerical calculations will not give a proof of any relations, since it is just a check up to a certain precision. But in any case an interested student should play around a little bit with these tools and maybe try to find nice relations and patterns, which he/she then could try to prove using the machinery we learn during this course.

If the coefficients in the output are extremely large, then this is an indication for the fact there are no \( \mathbb{Q} \)-linear relations among the values in the input. For example to check if there is a \( \mathbb{Q} \)-linear relation among \( \zeta(3), \pi^3 \) and 1 one enters

\[
\text{input: } \text{lindep([zeta(3),Pi^3,1])} \\
\text{output: } [-5229795329281686, 216810578846293, -435977217249266]~
\]

which indicates that there are (as expected) no relations among these values. The size of the numbers here depends on the current precision used, which can be changed by `\p 50` to, for example, set the precision to 50 significant digits.

The Conjecture 1.13 is out of reach at the moment and so far there is no weight \( k \), for which we can actually prove that \( \dim_{\mathbb{Q}} \mathcal{Z}_k > 1 \), since for example it is not even known (but expected) that \( \zeta(5) \) and \( \zeta(2,3) \) are linearly independent and that \( \dim_{\mathbb{Q}} \mathcal{Z}_5 = 2 \). Even though it seems to be impossible to give lower bounds for \( \dim_{\mathbb{Q}} \mathcal{Z}_k \) so far, we know that the \( d_k \) give upper bounds:

**Theorem 1.15** (Terasoma (2002), Deligne–Goncharov (2005)). For all \( k \geq 0 \) we have \( \dim_{\mathbb{Q}} \mathcal{Z}_k \leq d_k \).

There also is a conjecture on an explicit basis for \( \mathcal{Z} \) due to Hoffman.

**Conjecture 1.16** (Hoffman [H1], 1997). For \( k \geq 0 \) the multiple zeta values

\[ \{\zeta(k_1, \ldots, k_r) \mid r \geq 0, k_1 + \cdots + k_r = k, k_1, \ldots, k_r \in \{2, 3\} \} \]

form a basis of \( \mathcal{Z}_k \).

Notice that this conjecture would imply (Exercise 2) Zagiers dimension conjecture (Conjecture 1.13) since \( d_k \) counts exactly the number of indices of weight \( k \) with only 2’s and 3’s. Multiple zeta values with only 2’s and 3’s in their index will be called **Hoffman elements**. The linear independence of the Hoffman elements is unknown so far, but we know that these generate the whole space due to the following deep result of Brown.

**Theorem 1.17** (Brown [Br], 2012). For all \( k \geq 0 \) we have

\[ \mathcal{Z}_k = \{\zeta(k_1, \ldots, k_r) \mid r \geq 0, k_1 + \cdots + k_r = k, k_1, \ldots, k_r \in \{2, 3\}\} \mathbb{Q} \].

Version 3.9 (June 26, 2020)
The only known proofs of Theorem 1.15 and 1.17 use deep concepts from algebraic geometry, particularly the theory of mixed Tate motives.  

**Remark 1.18.** In his work [Br], Brown shows that all the above conjectures hold for so-called motivic multiple zeta values $\mathcal{Z}^m$, which are conjecturally isomorphic as a $\mathbb{Q}$-algebra to $\mathbb{Z}$. Using the surjective period map $\text{per} : \mathcal{Z}^m \to \mathcal{Z}$ Theorem 1.17 is then just a consequence of his more general results. For details on this, we refer to the excellent book [BF].

### 1.3 Modular forms and the Broadhurst-Kreimer conjecture

In this section we want to give a glimpse of the connection of modular forms and multiple zeta values and state the Broadhurst-Kreimer conjecture, which is a refinement of Zagiers dimensions conjecture (Conjecture 1.13). For this we will give a naive argument why cusp forms give rise to relations among double zeta values, which we will make precise later in Section 4. We will not give a complete introduction to modular forms and just state the main structure theorems and definitions. For a complete introduction to the theory of modular forms we refer the reader to [Za1] and [Za2].

#### 1.3.1 Basics of modular forms

Let $\mathbb{H} = \{ x + iy \in \mathbb{C} \mid x, y \in \mathbb{R}, y > 0 \}$ denote the complex upper half plane. A holomorphic function $f \in \mathcal{O}(\mathbb{H})$ is called a **modular form of weight** $k \in \mathbb{Z}$ if it satisfies

$$f(\tau + 1) = f(\tau), \quad f(-\frac{1}{\tau}) = \tau^k f(\tau)$$

for all $\tau \in \mathbb{H}$ and if it has a Fourier expansion of the form

$$f(\tau) = \sum_{n=0}^{\infty} a_n q^n \quad (a_n \in \mathbb{C}), \quad q = q(\tau) = e^{2\pi i \tau}. \quad (1.9)$$

The map $q : \tau \to \exp(2\pi i \tau)$ is the holomorphic map, which sends $\mathbb{H}$ to the punctured unit disc. The existence of (1.9) states that $f$, as a function in $q$, can be analytically continued to $q = 0$, and therefore give rise to a holomorphic function in the whole open unit disc $\{ q \in \mathbb{C} \mid |q| < 1 \}$. This is also equivalent to the fact that $f(\tau)$ is bounded as $\tau \to i\infty$.

By $\mathcal{M}_k$ we denote the **space of all modular forms of weight** $k$ and we write $\mathcal{M} = \sum_{k \geq 0} \mathcal{M}_k$ for the space of all modular forms. It is easy to see that $\mathcal{M}$ equipped with the usual multiplication of holomorphic functions forms a $\mathbb{C}$-algebra. The coefficients $a_n \in \mathbb{C}$ in (1.9) are called the **Fourier coefficients** of $f$ and a modular form for which $a_0 = 0$ (i.e. where the sum (1.9) starts at $n = 1$) is called a **cusp form**. We write

$$\mathcal{S}_k = \{ f \in \mathcal{M}_k \mid f = \sum_{n=1}^{\infty} a_n q^n \}$$

for the **space of all cusp forms of weight** $k$. The first non-trivial examples of modular forms are given by **Eisenstein series**, which are for even $k \geq 4$ defined by

$$G_k(\tau) = \frac{1}{2} \sum_{m,n \in \mathbb{Z}, (m,n) \neq (0,0)} \frac{1}{(m\tau + n)^k}. \quad (1.10)$$

For all even $k \geq 4$ we have $G_k \in \mathcal{M}_k$. Notice that the above sum would vanish for odd $k$. In fact for all odd $k$ we have $\mathcal{M}_k = 0$. The Fourier expansion of Eisenstein series, which we will calculate in Section 5.1 for a more general object, is given by

$$G_k(\tau) = \zeta(k) + \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n, \quad (1.10)$$
where \( \sigma_k(n) = \sum_{d|n} d^{k-1} \) is the divisor sum. The expression (1.10) makes sense for any \( k \geq 2 \) and therefore we define \( G_k(\tau) \) for all \( k \geq 2 \) by (1.10). Though for \( k = 2 \) and odd \( k \) these are not modular forms of weight \( k \) anymore.

The Eisenstein series are the building blocks for all modular forms, and we summarize the main properties of modular forms in the following Theorem.

**Theorem 1.19.**  

i) For even \( k \geq 4 \) we have \( M_k = \mathbb{C} \cdot G_k \oplus S_k \).

ii) For odd, negative \( k \) or \( k = 2 \) we have \( M_k = 0 \) and \( M_0 = \mathbb{C} \).

iii) For all \( k_1, k_2 \geq 0 \) we have \( M_{k_1} \cdot M_{k_2} \subset M_{k_1+k_2} \).

iv) The Eisenstein series \( G_4 \) and \( G_6 \) are algebraically independent (over \( \mathbb{C} \)).

v) We have

\[
M = \bigoplus_{k=0}^{\infty} M_k = \mathbb{C}[G_4, G_6],
\]

i.e. \( M \) is a graded \( \mathbb{C} \)-algebra, which is isomorphic to the polynomial ring in two variables.

vi) The space \( M_k \) is generated by \( G_k \) and products of two Eisenstein series, i.e. for even \( k \geq 4 \)

\[
M_k = \mathbb{C} \cdot G_k + \langle G_{k_1} G_{k_2} \mid k_1, k_2 \geq 4 \text{ even}, k_1 + k_2 = k \rangle_\mathbb{C}.
\]

**Proof.** The statements i), ii) & iii) follow almost immediately from the definition. Statement iv) and v) need some complex analysis and can be found in any standard book of modular forms (e.g. [B3], [Za2], [Za1]). The statement vi) follows from work of Rankin and can be found in [KZ]. □

The first non-trivial cusp form is the **discriminant function** \( \Delta \) (a.k.a Ramanujan Delta function)

\[
\Delta(\tau) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = q - 24q^2 + 252q^3 - 1472q^4 + \ldots , \tag{1.11}
\]

which is a cusp form \( \Delta \in S_{12} \) of weight 12. Since \( \Delta \) has no zero in \( \mathbb{H} \) and a zero of order one in \( q = 0 \), one can show that every cusp form of weight \( k \) can be written as a product of a modular form of weight \( k - 12 \) and \( \Delta \). Together with Theorem 1.19 this gives the following well-known theorem.

**Theorem 1.20.**  

i) For \( k \geq 0 \) the map \( M_k \to S_{k+12} \) given by \( f \mapsto \Delta \cdot f \) is an isomorphism of \( \mathbb{C} \)-vector spaces.

ii) The generating series for the dimension of cusp forms of weight \( k \) is given by

\[
S(X) = \sum_{k \geq 0} \dim_{\mathbb{C}} S_k X^k = X^{12} \sum_{k \geq 0} \dim_{\mathbb{C}} M_k X^k = \frac{X^{12}}{(1 - X^4)(1 - X^6)}.
\]

### 1.3.2 The Broadhurst-Kreimer Conjecture

We now want to state a refinement of Conjecture 1.13 given by Broadhurst and Kreimer, which indicates a connection of cusp forms and the dimension of the depth graded spaces of multiple zeta values. The depth gives a filtration on the space \( \mathcal{Z} \) and we write

\[
\text{Fil}_r(D, k) = \langle \zeta(k) \mid k \text{ admissible, } wt(k) = k, \text{dep}(k) \leq r \rangle_{\mathbb{Q}}
\]

for its depth \( r \) part and denote the associated graded part by \( \text{gr}_r(D, k) \). In other words elements in \( \text{gr}_r(D, k) \) are multiple zeta values of weight \( k \) and depth \( r \) modulo multiple zeta values of lower depth. For example the class of \( \zeta(2,1) \) in \( \text{gr}_2(D, 3) \) is zero, since \( \zeta(2,1) = \zeta(3) \).
Conjecture 1.21 (Broadhurst-Kreimer, 1997 [BroK]). The generating series of the dimensions of the weight- and depth-graded parts of multiple zeta values is given by

\[ \sum_{k,r \geq 0} \dim_0(\text{gr}^D_kZ)X^kY^r = \frac{1 + E(X)Y}{1 - O(X)Y + S(X)Y^2 - S(X)Y^4}, \]

where

\[ E(X) = \frac{X^2}{1 - X^2}, \quad O(X) = \frac{X^3}{1 - X^2}, \quad S(X) = \sum_{k \geq 0} \dim_C S_kX^k = \frac{X^{12}}{(1 - X^2)(1 - X^4)}. \]

This conjecture reduces to Zagiers dimension conjecture (Exercise 2) by setting \( Y = 1 \). Observe that

\[ \frac{1 + E(X)Y}{1 - O(X)Y + S(X)Y^2 - S(X)Y^4} = 1 + (E(X) + O(X))Y + ((E(X) + O(X))O(X) - S(X))Y^2 + \cdots, \]

which indicates that cusp forms give rise to relation in depth 2. Before we make this more precise, we first want to give a naive reason why cusp forms give rise to linear relations among double zeta values. Let \( f = \sum_{n=1}^{\infty} a_nq^n \in S_k \) be a cusp form. By Theorem 1.19 vi) we know that \( f \) can be written for some \( \alpha, \beta_{a,b} \in \mathbb{C} \) (we will just be interested in \( \mathbb{Q} \)) as

\[ f = \alpha \mathbb{G}_k + \sum_{a,b \geq 4 \text{ even} \atop a+b=k} \beta_{a,b} \mathbb{G}_a \mathbb{G}_b. \quad (1.13) \]

Considering the constant term in the Fourier expansion of both sides then yields the relation

\[ 0 = \alpha \zeta(k) + \sum_{a,b \geq 4 \text{ even} \atop a+b=k} \beta_{a,b} \zeta(a) \zeta(b). \]

The products on the right hand side can now be evaluated by using either (1.7) or (1.8) to obtain a linear relation among double zeta values. This approach is not really interesting, since the representation of a cusp form (1.13) is not unique and also the choice of expanding the product is arbitrary. Therefore we can not really relate a cusp form to a single relation among double zeta values. But there is a surprising 1:1 correspondence between certain relations and cusp forms, which we will explain now.

For even weight \( k \), the Broadhurst-Kreimer conjecture predicts by (1.12) that \( \dim_0(\text{gr}^D_kZ) \) is given by the coefficient of \( X^k \) in \( O(X)O(X) - S(X) \). The coefficient of \( O(X)O(X) \) counts the number of indices \( (k_1, k_2) \) with \( k_1, k_2 \geq 3 \) odd and \( k_1 + k_2 = k \) for which we will write "(odd,odd)" in the following. If this would be the only contribution to \( \dim_0(\text{gr}^D_kZ) \) then a naive guess would be that the \( \zeta(\text{odd},\text{odd}) \) give a basis of \( \text{gr}^D_kZ \). Indeed, we will see in Section 4 that the \( \zeta(\text{odd},\text{odd}) \) span \( \text{gr}^D_kZ \). But the factor \( S(X) \) in the Broadhurst-Kreimer conjecture indicates that there are relations in weight \( k \) between these values whenever there exist cusp forms of weight \( k \). The first relation between \( \zeta(k_1, k_2) \), where both \( k_1 \) and \( k_2 \) are odd, appears in weight \( k_1 + k_2 = 12 \) and is given by

\[ -10394\zeta(3,9) + 47650\zeta(5,7) + 41431\zeta(7,5) - 720\zeta(9,3) - 10394\zeta(11,1) = 0. \quad (1.14) \]

As shown in [GKZ] (See section 1) we have for all even \( k \geq 4 \) the relation

\[ \sum_{k_1, k_2 \geq 3, k_1 \geq 3, k_2 \geq 1 \text{ odd} \atop k_1 + k_2 = k} \zeta(k_1, k_2) = \frac{1}{4} \zeta(k). \quad (1.15) \]
Combining the relations (1.14) and (1.15) gives

\[168\zeta(5,7) + 150\zeta(7,5) + 28\zeta(9,3) = \frac{5197}{691}\zeta(12).\]

From this we get the following relation among \(\zeta(\text{odd, odd})\) in \(\text{gr}^D_{2}\mathbb{Z}_{12}\):

\[168\zeta(5,7) + 150\zeta(7,5) + 28\zeta(9,3) \equiv 0 \mod \zeta(12).

In [GKZ] Gangl-Kaneko-Zagier give an explicit construction of such relations for a given cusp form \(f \in S_k\). For this they consider its period polynomial \(p_f(X,Y) \in \mathbb{C}[X,Y]\) and show that one can obtain a relation among double zeta values explicitly from the coefficients of this polynomial (see Section 4). For example the above relation can be obtain by taking for \(f\) a certain multiple of the cusp form \(\Delta\).

A consequence of their results is the following.

**Theorem 1.22.** (Gangl-Kaneko-Zagier, 2006) For even \(k \geq 4\) the number of (independent) \(\mathbb{Q}\)-linear relations among \(\zeta(2a + 1, 2b - 1)\) with \(a, b \geq 1\) and \(k = 2(a + b)\) is at least \(\dim_{\mathbb{C}} S_k\). Conjecturally the number of these relations are exactly \(\dim_{\mathbb{C}} S_k\).

Due to a recent result of Tasaka [Tas], also a somehow converse statement is known: Given a relation among \(\zeta(\text{odd, odd})\) (which follows from a certain set of relations) of weight \(k\), we can construct explicitly a cusp form of weight \(k\). His result uses double Eisenstein series which we will discuss in Section 5.

### 1.4 \(q\)-analogues of multiple zeta values

A \(q\)-analogue of a theorem, identity or expression is a generalization involving a new parameter \(q\) that returns the original theorem, identity or expression in the limit as \(q \to 1^1\). The easiest example is the \(q\)-analogue of a natural number \(m\) given by the \(q\)-integer

\[ [m]_q = \frac{1 - q^m}{1 - q} = 1 + q + \cdots + q^{m-1}, \quad \lim_{q \to 1} [m]_q = m. \quad (1.16)\]

There are various different models of \(q\)-analogues of multiple zeta values in the literature. We will consider a few of them in this course and start with the most common model which was first independently studied by Bradley [Bra] and Zhao [Zh2]. For an admissible index \(k = (k_1, \ldots, k_r)\) these are defined by

\[ \zeta_B^q(k) = \zeta_B^q(k_1, \ldots, k_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{q^{(k_1 - 1)m_1} \cdots q^{(k_r - 1)m_r}}{[m_1]_q^{k_1} \cdots [m_r]_q^{k_r}}. \quad (1.17)\]

By (1.16), together with a justification that one can interchange summation and taking the limit (see proof of Proposition 1.26), it is easy to see that we have

\[ \lim_{q \to 1} \zeta_B^q(k) = \zeta(k). \]

In Section 3 we will see that these \(q\)-series satisfy a lot of relations which are satisfied by multiple zeta values. In particular, this is the unique model of \(q\)-analogues (in the sense we will define later) which satisfies the duality relation, e.g. \(\zeta_B^q(2,1) = \zeta_B^q(3)\).

---

1. Here and in the following we mean by \(q \to 1\) the limit of \(q\) to 1 on the real axis with \(|q| < 1\).
In this course, we will mostly be interested in another model of \( q \)-analogues which is inspired by Eisenstein series and which was introduced by the author in his PhD-thesis [B2] and further studied in [BK1]. These object are not \( q \)-analogues in the above sense, but are called often modified \( q \)-analogues. By a modified \( q \)-analogue of weight \( k \) we mean, that we first need to multiply by \((1 - q)^k\) before taking the limit \( q \to 1 \). One motivation to consider modified \( q \)-analogues is the following.

**Proposition 1.23.** Let \( f(q) = \sum_{n=0}^{\infty} a_n q^n \in \mathcal{M}_k \) be a modular form of weight \( k \). Then \( f \) is, up to the factor \((2\pi i)^k\), a modified \( q \)-analogue of weight \( k \) of its constant term \( a_0 \), i.e. we have

\[
\lim_{q \to 1} (1 - q)^k f(q) = (2\pi i)^k a_0.
\]

**Proof.** This is a consequence of Proposition 1.26 below together with the fact that every modular form is a polynomial in \( G_4 \) and \( G_6 \). Another way to see this is by using the transformation property \( f \left( -\frac{1}{q} \right) = \tau^k f(\tau) \). Taking the limit \( q \to 1 \) on the real axis corresponds to the limit \( \tau \to 0 \) on the positive imaginary axis, since \( q = e^{2\pi i \tau} \). Together with \( \lim_{\tau \to i\infty} f(\tau) = a_0 \) we obtain

\[
\lim_{q \to 1} (1 - q)^k f(q) = \lim_{\tau \to 0} ((2\pi i \tau)^k + O(\tau^{k+1})) f(\tau) = \lim_{\tau \to 0} (2\pi i)^k f\left( \frac{-1}{\tau} \right) = \lim_{\tau \to i\infty} (2\pi i)^k f(\tau) = (2\pi i)^k a_0.
\]

\(\square\)

### 1.4.1 A modified \( q \)-analogue of multiple zeta values coming from Eisenstein series

For any \( k \geq 2 \) we defined in the previous section the Eisenstein series \( G_k(\tau) \) by its Fourier expansion

\[
G_k(\tau) = \zeta(k) + \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n.
\]

Clearly we can obtain \( \zeta(k) \) from these series in the case \( q \to 0 \), but as indicated in Proposition 1.23 this will also be possible by considering \( q \to 1 \) after some modification. From now on we will always consider \( q \) as a formal variable or a fixed complex number with \( |q| < 1 \). Define for even \( k \geq 2 \) the \( q \)-series \( G_k = (2\pi i)^{-k} G_k(\tau) \), then we have by Eulers formula \( \zeta(2m) = -\frac{B_{2m}}{2(2m)!} (2\pi i)^{2m} \) (Proposition 1.1)

\[
G_k = -\frac{B_k}{2k!} + \frac{1}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n \in \mathbb{Q}[[q]]. \tag{1.18}
\]

The right-hand side of 1.18 makes sense for any \( k \geq 1 \) and we will use it to define \( G_k \) for all \( k \geq 2 \). We will denote the constant term of \( G_k \) by \( \beta(k) = -\frac{B_k}{2k!} \in \mathbb{Q} \) and denote the rest by

\[
g(k) = \frac{1}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n,
\]

i.e. \( G_k = \beta(k) + g(k) \). The \( g(k) \) can be rewritten in the following way

\[
g(k) = \frac{1}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n = \sum_{m,d>0} \frac{d^{k-1}}{(k-1)!} q^{md} = \sum_{m>0} \frac{P_k(q^m)}{(1 - q^m)^k}, \tag{1.19}
\]
where for \( k \geq 1 \) the \( P_k(X) \in \mathbb{Q}[X] \) are the \textbf{Eulerian polynomials} defined by

\[
\frac{P_k(X)}{(1 - X)^k} = \sum_{d > 0} \frac{d^{k-1}}{(k-1)!} X^d.
\]

For \( k = 1, \ldots, 6 \) these are given by

\[
\begin{align*}
P_1(X) &= P_2(X) = X, \\
P_3(X) &= \frac{1}{2} X(X^2 + 1), \\
P_4(X) &= \frac{1}{6} X(X^2 + 4X + 1), \\
P_5(X) &= \frac{1}{24} X(X + 1)(X^2 + 10X + 1), \\
P_6(X) &= \frac{1}{120} X(X^4 + 26X^3 + 66X^2 + 26X + 1).
\end{align*}
\]

\textbf{Lemma 1.24.} For all \( k \geq 1 \) we have \( P_k(0) = 0 \) and \( P_k(1) = 1 \).

\textit{Proof.} This is Exercise 4 i).

As a multiple version of the \( q \)-series \( g(k) \) in (1.19) we define the following.

\textbf{Definition 1.25.} For any index \( k = (k_1, \ldots, k_r) \in \mathbb{Z}_{\geq 1}^r \) we define

\[
g(k) = g(k_1, \ldots, k_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{P_{k_1}(q^{m_1}) \cdots P_{k_r}(q^{m_r})}{(1 - q^{m_1})^{k_1} \cdots (1 - q^{m_r})^{k_r}} \in \mathbb{Q}[[q]].
\]

Notice that this is a well-defined \( q \)-series for any index (even when \( k_1 = 1 \)) since \( P_k(X) \in \mathbb{Q}[X] \) (Lemma 1.24). These series can be seen as modified \( q \)-analogues of multiple zeta values, where by modified we mean that we need to multiply by a power of \((1 - q)\) before taking the limit \( q \to 1 \).

\textbf{Proposition 1.26.} For any admissible index \( k \) we have

\[
\lim_{q \to 1} (1 - q)^{\text{wt}(k)} g(k) = \zeta(k).
\]

\textit{Proof.} First we need to justify the interchange of summation and taking the limit, which follows from the fact that for \(|q| < 1\) the sum inside the limit converges uniformly. We will skip the details and refer to [BK1, Lemma 6.6] for a precise proof. Then this is an easy consequence of Lemma 1.24 since for \( k \geq 1 \) we have

\[
\lim_{q \to 1} (1 - q)^k \frac{P_k(q^m)}{(1 - q^m)^k} = \lim_{q \to 1} \frac{P_k(q^m)}{m^k} = \frac{P_k(1)}{m^k} = \frac{1}{m^k}.
\]

We will denote the space spanned by all \( g(k) \) for any (not necessarily admissible!) index \( k \) by

\[
\mathcal{G} = \langle g(k) \mid k \text{ index} \rangle_{\mathbb{Q}} \subset \mathbb{Q}[[q]],
\]

where we also use the convention \( g(\emptyset) = 1 \). In Section 2 we will see (as a consequence of Lemma 2.18) that this space is also closed under multiplication and we will proof the following.

\textbf{Proposition 1.27.} The space \( \mathcal{G} \) is a \( \mathbb{Q} \)-subalgebra of \( \mathbb{Q}[[q]] \).

\footnotesize
\textsuperscript{2}In the literature usually \((k − 1)!X^{−1}P_k(X)\) is called Eulerian polynomial.
\textsuperscript{3}These \( q \)-series are called brackets in [BK1] and are denoted by \([k_1, \ldots, k_r]\) there.

\normalsize
Similar as we did for multiple zeta values, we will show the lowest depth case of this proposition now. For \( k_1, k_2 \geq 1 \) we have, similar as in (1.4)

\[
g(k_1)g(k_2) = \sum_{m_1 > 0} \frac{P_{k_1}(q^{m_1})}{(1 - q^{m_1})^{k_1}} \sum_{m_2 > 0} \frac{P_{k_2}(q^{m_2})}{(1 - q^{m_2})^{k_2}} \\
= \left( \sum_{m_1 > m_2 > 0} + \sum_{m_2 > m_1 > 0} + \sum_{m_1 = m_2 > 0} \right) \frac{P_{k_1}(q^{m_1})}{(1 - q^{m_1})^{k_1}} \frac{P_{k_2}(q^{m_2})}{(1 - q^{m_2})^{k_2}} \\
= g(k_1, k_2) + g(k_2, k_1) + \sum_{m > 0} \frac{P_{k_1}(q^{m})}{(1 - q^{m})^{k_1}} \frac{P_{k_2}(q^{m})}{(1 - q^{m})^{k_2}}.
\]

That this is again an element in \( \mathcal{G} \) follows now from the following lemma, which can be proven by using generating series together with the definition of the Bernoulli numbers (1.2).

**Lemma 1.28.** For \( k \geq 1 \) we set \( R_k(X) = \frac{P_k(X)}{(1-X)^2} = \sum_{d > 0} \frac{q^{d-1}}{(d-1)!} X^d \). Then for all \( k_1, k_2 \geq 1 \) we have

\[
R_{k_1}(X) \cdot R_{k_2}(X) = R_{k_1+k_2}(X) + \sum_{j=1}^{k_1+k_2-1} \left( \lambda^j_{k_1,k_2} + \lambda^j_{k_2,k_1} \right) R_j(X)
\]

where the rational numbers \( \lambda^j_{k_1,k_2} \) are given by

\[
\lambda^j_{k_1,k_2} = (-1)^{k_2-1} \binom{k_1+k_2-1-j}{k_1-j} \frac{B_{k_1+k_2-j}}{(k_1+k_2-j)!},
\]

and where we use the convention \( \binom{n}{k} = 0 \) for \( k < 0 \).  

**Proof.** This is Exercise 5 ii).

Lemma 1.28 together with (1.20) gives the following analogue for the \( q \)-series \( g \) of the stuffle product formula \( \zeta(k_1)\zeta(k_2) = \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2) \) of multiple zeta values.

**Proposition 1.29.** For \( k_1, k_2 \geq 1 \) and we have

\[
g(k_1)g(k_2) = g(k_1, k_2) + g(k_2, k_1) + g(k_1 + k_2) + \sum_{j=1}^{k_1+k_2-1} \left( \lambda^j_{k_1,k_2} + \lambda^j_{k_2,k_1} \right) g(j).
\]

**Proof.** This follows immediately by plugging (1.21) into (1.20).

We see that the extra terms in the right-hand side are of lower weight and therefore these will vanish when multiplying both sides with \((1 - q)^{k_1+k_2}\) and taking the limit \( q \to 1 \). In particular we see that this, together with Proposition 1.26 gives back the stuffle product formula for multiple zeta values. In contrast to \( \mathcal{Z} \) (which is conjecturally graded by weight) the space \( \mathcal{G} \) is therefore not graded by weight. We treat both of these products (for \( \zeta \) and \( g \)) simultaneously in Section 3 as examples for a quasi-shuffle product. The series \( g \) also satisfy an analogue of the shuffle product formula

\[
\zeta(k_1)\zeta(k_2) = \sum_{j=2}^{k_1+k_2-1} \left( \left( \frac{j-1}{k_1-1} \right) + \left( \frac{j-1}{k_2-1} \right) \right) \zeta(j, k_1 + k_2 - j),
\]

which we saw in Proposition 1.30. This formula involves not just the series \( g \) but also its derivative with respect to the differential operator \( \frac{d}{dq} \).
Proposition 1.30. For \( k_1, k_2 \geq 1 \) and \( k = k_1 + k_2 \) we have
\[
g(k_1) g(k_2) = \sum_{j=1}^{k-1} \left( \binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) g(j, k-j) \\
+ \binom{k-2}{k_1-1} \left( q \frac{d}{dq} g(k_2) - g(k-1) \right) + \delta_{k_1,1} \delta_{k_2,1} g(2),
\]
(1.22)
where \( \delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \) denotes the Kronecker delta.

We will give a proof of this formula by using generating series below. In general depth we will need, besides the derivative with respect to \( q \frac{d}{dq} \), even more extra terms to get an analogue of the shuffle product. This will be discussed in Section 5.

Example 1.31. i) Similar to the Example 1.10 where we showed
\[
\zeta(2) \zeta(3) = \zeta(2, 3) + \zeta(3, 2) + \zeta(5) = \zeta(2, 3) + 3 \zeta(3, 2) + 6 \zeta(4, 1),
\]
(1.23)
we can use Propositions 1.29 and 1.30 to get the following formulas for the \( q \)-series \( g \)
\[
g(2) g(3) = g(2, 3) + g(3, 2) + g(5) - \frac{1}{12} g(3), \\
g(2) g(3) = g(2, 3) + 3 g(3, 2) + 6 g(4, 1) - 3 g(4) + q \frac{d}{dq} g(3). 
\]
(1.24)
From this we deduce the linear relation \( g(5) = 2 g(3, 2) + 6 g(4, 1) + \frac{1}{12} g(3) - 3 g(4) + q \frac{d}{dq} g(3) \). Multiplying equation (1.24) with \((1 - q)^5\) and taking the limit \( q \to 1 \) gives the equation (1.23). This will be explained in detail in Section 5.

ii) Since Propositions 1.29 and 1.30 are also valid for non admissible indices we get
\[
g(1) g(2) = g(1, 2) + g(2, 1) + g(3) - \frac{1}{2} g(2), \\
g(1) g(2) = g(1, 2) + 2 g(2, 1) + q \frac{d}{dq} g(1) - g(2)
\]
by using \( k_1 = 1, k_2 = 2 \). This gives the following analogue of the relation \( \zeta(3) = \zeta(2, 1) \)
\[
g(3) = g(2, 1) + q \frac{d}{dq} g(1) - \frac{1}{2} g(2).
\]
Since the \( g(k) \) are essentially, up to a constant, the Eisenstein series of weight \( k \), above formulas can be used to give purely combinatorial proofs of identities among modular forms. One simple example is the identity \( G_4^2 = \frac{1}{3} G_8 \), which is a consequence of Proposition 1.29 and 1.30 (Exercise 4). In Section 5.5 we will elaborate on this combinatorial approach to modular forms.

1.4.2 Generating series

We now want to illustrate how the proofs of Proposition 1.29 and 1.30 can be done by using generating series. This will be done in more generality in Section 5, but we want to satisfy the curious reader.
The key point is, that there are two different ways to write the generating series of \( g(k) \). Multiplying one of them leads to the shuffle product and the other one to the shuffle product. For \( r \geq 1 \) we will denote the generating series of \( g(k_1, \ldots, k_r) \) by

\[
g(X_1, \ldots, X_r) = \sum_{k_1, \ldots, k_r \geq 1} g(k_1, \ldots, k_r)X_1^{k_1-1} \cdots X_r^{k_r-1}.
\]

**Lemma 1.32.** We have

\[
g(X_1, \ldots, X_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{e^{X_1 q^{m_1}}}{1 - e^{X_1 q^{m_1}}} \cdots \frac{e^{X_r q^{m_r}}}{1 - e^{X_r q^{m_r}}} \tag{1.25}
\]

\[
= \sum_{m_1 > \cdots > m_r > 0} \frac{e^{m_1 X_r q^{m_1}} e^{m_2 (X_{r-1} - X_r) q^{m_2}} \cdots e^{m_r (X_1 - X_2) q^{m_r}}}{1 - q^{m_1}} \frac{1}{1 - q^{m_2}} \cdots \frac{1}{1 - q^{m_r}} \tag{1.26}
\]

**Proof.** This is Exercise 3.1i). The proof of (1.25) follows directly from the definition. For (1.26) a suitable change of summation variables is needed.

Propositions 1.29 and 1.30 are a consequence of the following proposition by considering the coefficients of \( X^{k_1-1} Y^{k_2-1} \). (Exercise 3.iii)

**Proposition 1.33.** We have

\[
g(X)g(Y) = g(X,Y) + g(Y,X) + \frac{1}{e^{X-Y} - 1} g(X) + \frac{1}{e^{Y-X} - 1} g(Y) \tag{1.27}
\]

\[
= g(X + Y, X) + g(X + Y, Y) - g(X + Y) + g(X) + g(Y) \frac{d}{dk} \sum_{k \geq 1} g(k) \frac{(X + Y)^k}{k} + g(2) \tag{1.28}
\]

**Proof.** Using (1.25) and (1.26) in the smallest depth case together with the usual splitting of the summation of \( m_1, m_2 > 0 \) into the cases \( m_1 > m_2 > 0, m_2 > m_1 > 0 \) and \( m_1 = m_2 = m > 0 \) gives

\[
g(X)g(Y) \overset{(1.25)}{=} g(X,Y) + g(Y,X) + \sum_{m>0} \frac{e^{X q^m}}{1 - e^{X q^m}} \frac{e^{Y q^m}}{1 - e^{Y q^m}},
\]

\[
g(X)g(Y) \overset{(1.26)}{=} g(X + Y, X) + g(X + Y, Y) + \sum_{m>0} e^{m(X+Y)} \left( \frac{q^m}{1 - q^m} \right)^2.
\]

It remains to evaluate the third term in both equations. For the first equation one can check by direct calculation that

\[
\frac{e^{X q^m}}{1 - e^{X q^m}} \frac{e^{Y q^m}}{1 - e^{Y q^m}} = \frac{1}{e^{X-Y} - 1} \frac{e^{X q^m}}{1 - e^{X q^m}} + \frac{1}{e^{Y-X} - 1} \frac{e^{Y q^m}}{1 - e^{Y q^m}},
\]

which then gives

\[
g(X)g(Y) = g(X,Y) + g(Y,X) + \frac{1}{e^{X-Y} - 1} g(X) + \frac{1}{e^{Y-X} - 1} g(Y).
\]

For the second equation one uses first \((\frac{q^m}{1 - q^m})^2 = \frac{q^{2m}}{(1 - q^m)^2} - \frac{q^m}{1 - q^m}\), which gives

\[
\sum_{m>0} e^{m(X+Y)} \left( \frac{q^m}{1 - q^m} \right)^2 = \sum_{m>0} e^{m(X+Y)} \frac{q^m}{(1 - q^m)^2} - g(X + Y).
\]
The first sum on the right can then be evaluated as (Exercise 3(ii))
\[
\sum_{m>0} e^{m(X+Y)} \frac{q^m}{(1-q^m)^2} = g(2) + q \frac{d}{dq} \sum_{k \geq 1} g(k) \frac{(X + Y)^k}{k},
\]
from which the claimed formula follows.

### 1.4.3 General modified $q$-analogues of multiple zeta values

As mentioned before, there are several models of $q$-analogues of multiple zeta values, and in [Zh3] you can find a nice overview of some of them. Most of these have similar definitions as the $g(k)$ with the difference that the Eulerian polynomials get replaced by other polynomials. Also the modified version can find a nice overview of some of them. Most of these have similar definitions as the $g(k)$ type of [Zh3].

**Definition 1.34.** For $k_1, \ldots, k_r \geq 1$ and polynomials $Q_1(X) \in X \mathbb{Q}[X]$ and $Q_2(X) \ldots, Q_r(X) \in \mathbb{Q}[X]$ we define
\[
\zeta_q(k_1, \ldots, k_r; Q_1, \ldots, Q_r) = \sum_{m_1, \ldots, m_r > 0} \frac{Q_1(q^{m_1}) \ldots Q_r(q^{m_r})}{(1-q^{m_1})k_1 \ldots (1-q^{m_r})k_r}.
\]

Similar as in Proposition 1.26 these series can be seen as (modified) $q$-analogues of $\zeta(k_1, \ldots, k_r)$, since we have for $k_1 \geq 2$
\[
\lim_{q \to 1} (1-q)^{k_1 + \cdots + k_r} \zeta_q(k_1, \ldots, k_r; Q_1, \ldots, Q_r) = Q_1(1) \ldots Q_r(1) \cdot \zeta(k_1, \ldots, k_r).
\]

We only consider the case where $\deg(Q_j) \leq k_j$ and consider the following $\mathbb{Q}$-vector space:
\[
\mathcal{Z}_q := \left\langle \zeta_q(k_1, \ldots, k_r; Q_1, \ldots, Q_r) \mid r \geq 0, k_1, \ldots, k_r \geq 1, \deg(Q_j) \leq k_j \right\rangle_{\mathbb{Q}},
\]
where again $\zeta_q(0; 0) = 1$. It is again not hard to see that $\mathcal{Z}_q$ is a $\mathbb{Q}$-algebra, since it is again an example for a quasi-shuffle algebra (see Section 2.3.2), and we have for example
\[
\zeta_q(k_1; Q_1)\zeta_q(k_2; Q_2) = \zeta_q(k_1, k_2; Q_1, Q_2) + \zeta_q(k_2, k_1; Q_2, Q_1) + \zeta_q(k_1 + k_2; Q_1 \cdot Q_2).
\]
Since $g(k_1, \ldots, k_r) = \zeta_q(k_1, \ldots, k_r; P_{k_1}, \ldots, P_{k_r})$ and $\deg(P_k) \leq k$ we have $\mathcal{G} \subset \mathcal{Z}_q$. As we will see in the next section, we can describe the analogue of the stuffle product for elements in $\mathcal{G}$ and $\mathcal{Z}_q$ as examples of quasi-shuffle products. The reason to introduce the a priori bigger space $\mathcal{Z}_q$ is that we can describe the higher depth analogue of the shuffle product in this space. This we will do in Section 5. Even though we will not be able to describe the shuffle product analogue in the space $\mathcal{G}$ explicitly, we have the following surprising conjecture, which was discovered by the author during his PhD thesis.

**Conjecture 1.35.** ([BZ], [BK2]) We have $\mathcal{G} = \mathcal{Z}_q$.

If time permits we will discuss application and motivations of this conjecture in Section 5.
§2 Algebraic setup

In this section we want to explain the algebraic structure of the spaces \( Z \) (multiple zeta values), \( G \) (the space of the \( q \)-analogues \( q(k) \)) and \( Z_q \) (\( q \)-analogues of multiple zeta values). In particular, we will see why these spaces are all \( \mathbb{Q} \)-algebras. For this we will introduce the algebraic setup of Hoffman, which was first introduced in [H1] and then later generalized to quasi-shuffle algebras in [HI].

2.1 Multiple polylogarithms, iterated integrals and duality

In the following we want to introduce the iterated integrals expression for multiple zeta values. This will be used in the next subsection to give another explanation of the shuffle product formula in (1.8).

We start by calculating one simple example by hand, before giving a general formula afterwards.

Consider the following iterated integral
\[
\int_0^1 \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{1-t_2} = \int_0^1 \frac{dt_1}{t_1} \int_0^{t_1} \sum_{n=0}^{\infty} t_2^ndt_2 = \int_0^1 \frac{dt_1}{t_1} \left[ \sum_{n=0}^{\infty} \frac{t_2^{n+1}}{n+1} \right]_0^{t_1} = \sum_{n=0}^{\infty} \frac{1}{n+1} = \zeta(2)
\]

(2.1)

With the same idea one can also show that we have (Exercise 6 i))
\[
\zeta(2,3) = \int_0^1 \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{1-t_2} \int_0^{t_2} \frac{dt_3}{t_3} \int_0^{t_3} \frac{dt_4}{t_4} \int_0^{t_4} \frac{dt_5}{1-t_5}.
\]

(2.2)

In general we will see that an index \( k = (k_1, \ldots, k_r) \) corresponds to an iterated integral of length \( wt(k) \), where each \( k_j \) gives a block of \( k_j - 1 \) integrals over \( \frac{dt}{t} \) and one integral over \( \frac{dt}{1-t} \). To prove these iterated integrals in general we will introduce multiple polylogarithms, which can be seen as a simultaneous generalization of the polylogarithm \((r=1)\) and multiple zeta values \((z=1)\).

**Definition 2.1.** For \(|z| < 1\) and \( k = (k_1, \ldots, k_r) \in \mathbb{Z}_{>1}^r \) we define the multiple polylogarithm by
\[
\operatorname{Li}_k(z) = \operatorname{Li}_{k_1, \ldots, k_r}(z) = \sum_{m_1 > \cdots > m_r > 0} \frac{z^{m_1}}{m_1^{k_1} \cdots m_r^{k_r}}
\]

and set \( \operatorname{Li}_0(z) = 1 \).

For an arbitrary index \( k \) the \( \operatorname{Li}_k(z) \) are holomorphic functions in the open unit disc, but clearly when \( k \) is admissible \( \operatorname{Li}_k(z) \) is also defined for \( z = 1 \) and we have
\[
\operatorname{Li}_k(1) = \zeta(k).
\]

Multiple polylogarithms also have an iterated integral expression, and for example using the same calculation as in (2.1) we see for example that
\[
\operatorname{Li}_2(z) = \int_0^z \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{1-t_2}.
\]

It becomes clear that we will deal with iterated integrals of two different differential forms. To describe the iterated integrals and the shuffle product, we will therefore introduce the following algebraic setup.
2.1.1 The spaces $\mathfrak{H}_1$, $\mathfrak{H}_0$ and $\mathfrak{H}^0$

We denote by $\mathfrak{H} = \mathbb{Q}(x, y)$ the polynomial ring in the two non-commutative variables $x$ and $y$. A monomial in $x$ and $y$ will also be called a word, and $\mathfrak{H}$ is therefore the $\mathbb{Q}$-vector space spanned by all words in the letters $x$ and $y$. Further, we define the subspace $\mathfrak{H}_1 = \mathbb{Q} + x\mathfrak{H}y$, which is spanned by the empty word $1$ and all words in $x$ and $y$ which end in $y$. For $k \geq 1$ we define

$$z_k = x^{k-1}y.$$ 

With this we see that $\mathfrak{H}_1 = \mathbb{Q}(z_1, z_2, \ldots)$, i.e. we could say that $\mathfrak{H}_1$ is spanned by all words in the letters $z_k$. For an index $k = (k_1, \ldots, k_r) \in \mathbb{Z}_{\geq 1}^r$ we define

$$z_k = z_{k_1}z_{k_2} \cdots z_{k_r} \in \mathfrak{H}_1$$

and set $z_0 = 1$. Now define the space $\mathfrak{H}^0 = \mathbb{Q} + x\mathfrak{H}y$, which is the subspace of $\mathfrak{H}_1$ generated by all words which start in $x$ and end in $y$. In other words, $\mathfrak{H}^0$ is spanned by all $z_k$ with admissible indices $k$. Summarizing everything we have

$$\mathfrak{H}^0 = \langle z_k \mid k \text{ admissible index} \rangle_{\mathbb{Q}} \subset \mathfrak{H}_1 = \langle z_k \mid k \text{ index} \rangle_{\mathbb{Q}} \subset \mathfrak{H} = \mathbb{Q}(x, y).$$

2.1.2 Iterated integral expression for $\zeta$ and $\zeta$

From now on we will restrict to real $|z| < 1$ and consider integrals on the real axis. Since $\zeta_k(z)$ is defined for any $k$, we can view $\zeta_k$ as a $\mathbb{Q}$-linear map from $\mathfrak{H}_1$ to the space of real valued continuous functions on $(0, 1)$, i.e. $C((0, 1); \mathbb{R})$, defined on the generators by

$$\zeta_k(z) = \zeta_{k_1}(z_1) \zeta_{k_2}(z_2) \cdots \zeta_{k_r}(z_r).$$

(2.3)

By abuse of notation we write $\zeta_k(z) = \zeta_k(z)$ for any $w \in \mathfrak{H}_1$, which is defined by linearly extending the definition on the generators $z_k$. For example for $w = xyyxy + 2xxxyy = z_2z_3 + 2z_5 \in \mathfrak{H}_1$ we have $\zeta(w)(z) = \zeta(z_2,z_3)(z) + 2 \zeta(z_5)(z)$. Now we want to describe the iterated integral expression for the multiple polylogarithm using this setup.

**Lemma 2.2.** Let $w \in \mathfrak{H}_1$ be a linear combination of words all starting with the letter $a \in \{x, y\}$, i.e. $w = au$ for some $u \in \mathfrak{H}_1$. Then we have

$$\frac{d}{dz} \zeta_u(z) = \frac{d}{dz} \zeta_{au}(z) = \begin{cases} \frac{1}{z} \zeta_u(z), & a = x \\ \frac{1}{1-z} \zeta_u(z), & a = y. \end{cases}$$

**Proof.** Since $\zeta_k$ is linear it suffices to proof the statement for a word $w$. Assuming $w = z_k$ for $k = (k_1, \ldots, k_r)$, we have

$$\frac{d}{dz} \zeta_k(z) = \frac{d}{dz} \zeta_{k_1,z_2} \cdots \zeta_{k_r}(z) = \sum_{m_1, \ldots, m_r > 0} \frac{z^{m_1}}{m_1^{k_1-1}m_2^{k_2} \cdots m_r^{k_r}} = \sum_{m_1, \ldots, m_r > 0} \frac{z^{m_1-1}}{m_1^{k_1-1}m_2^{k_2} \cdots m_r^{k_r}}.$$

Let $a = x$, which is equivalent to $k_1 > 1$. In this case we obtain

$$\frac{d}{dz} \zeta_u(z) = \frac{d}{dz} \zeta_{xu}(z) = \frac{1}{z} \sum_{m_1, \ldots, m_r > 0} \frac{z^{m_1}}{m_1^{k_1-1}m_2^{k_2} \cdots m_r^{k_r}} = \frac{1}{z} \zeta_u(z).$$

Version 3.9 (June 26, 2020)
If \(a = y\), then we have \(k_1 = 1\) and
\[
\frac{d}{dz} \text{Li}_w(z) = \frac{d}{dz} \text{Li}_{yu}(z) = \sum_{m_1 > \cdots > m_r > 0} \frac{z^{m_1 - 1}}{m_2^{m_2} \cdots m_r^{m_r}} = \sum_{m_2 > \cdots > m_r > 0} \frac{1}{m_2^{m_2} \cdots m_r^{m_r}} \sum_{m_1 = m_2 + 1}^{\infty} \frac{z^{m_1 - 1}}{m_1^{m_1}}.
\]

Motivated by the iterated integrals (2.1), (2.2), and the above Lemma, we define
\[
\omega_x(t) = \frac{dt}{t}, \quad \omega_y(t) = \frac{dt}{1 - t}.
\]

With these differential forms we can write the multiple polylogarithms as the following iterated integral.

**Proposition 2.3.** For any word \(w = a_1 \ldots a_k \in \mathcal{S}^1\), with \(a_1, \ldots, a_k \in \{x, y\}\) and \(0 \leq z < 1\) we have
\[
\text{Li}_w(z) = \int_0^z \omega_{a_1}(t_1) \int_0^{t_1} \omega_{a_2}(t_2) \cdots \int_0^{t_{k-1}} \omega_{a_k}(t_k).
\]

**Proof.** This follows from Lemma 2.2 by induction on \(k\). In the case \(k = 1\) we have \(w = y = z_1\), i.e.
\[
\text{Li}_w(z) = \text{Li}_1(z) = \sum_{m > 0} \frac{z^m}{m} = \int_0^z \frac{dt}{1 - t} = \int_0^z \omega_y(t).
\]

The induction step is then exactly the statement of Lemma 2.2 since \(\text{Li}_w(0) = 0\) for non-empty \(w\). \(\square\)

For a real \(z\) we will also use the following simplified notation for iterated integrals for \(a_1 \ldots a_k \in \mathcal{S}^1\)
\[
\int_{z > t_1 > \cdots > t_k > 0} \omega_{a_1}(t_1) \cdots \omega_{a_k}(t_k) := \int_0^z \omega_{a_1}(t_1) \int_0^{t_1} \omega_{a_2}(t_2) \cdots \int_0^{t_{k-1}} \omega_{a_k}(t_k).
\]

Since \(\zeta(k)\) is just defined for admissible indices, we can, similar to (2.3), define a \(\mathbb{Q}\)-linear map from \(\mathcal{S}^0\) to the space of multiple zeta values \(\mathcal{Z}\), defined on the generators by
\[
\zeta: \mathcal{S}^0 \rightarrow \mathcal{Z} \\
z_k \mapsto \zeta(k).
\]

Also here we write \(\zeta: w \mapsto \zeta(w)\) for any \(w \in \mathcal{S}^0\). Since \(\text{Li}_w(1) = \zeta(w)\) for any \(w \in \mathcal{S}^0\) we also get an iterated integral expression for multiple zeta values as a consequence of Proposition 2.3.

**Corollary 2.4.** For any word \(w = a_1 \ldots a_k \in \mathcal{S}^0\), with \(a_1, \ldots, a_k \in \{x, y\}\) we have
\[
\zeta(w) = \int_{1 > t_1 > \cdots > t_k > 0} \omega_{a_1}(t_1) \cdots \omega_{a_k}(t_k).
\]
2.1.3 Duality relation

We now give an direct consequence of the iterated integral expression. Making the change of variables $s_j = 1 - t_{k-j+1}$ in the iterated integral expression gives a linear relation among multiple zeta values, which is called the duality relation. For example if $k = 3$ we can make the change of variables $s_1 = 1 - t_1$, $s_2 = 1 - t_2$, $s_3 = 1 - t_1$ in the following iterated integral

$$\zeta(3) = \int_0^1 \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{t_2} \int_0^{t_2} \frac{dt_3}{1-t_3} = \int_0^1 \frac{-ds_3}{1-s_3} \int_1^{s_3} \frac{-ds_2}{1-s_2} \int_1^{s_2} \frac{-ds_1}{s_1} = \int_0^1 \frac{ds_3}{s_1} \int_0^{s_1} \frac{ds_2}{1-s_2} \int_0^{s_2} \frac{ds_1}{1-s_1} = \zeta(2, 1).$$  \hspace{1cm} (2.6)

from which we again get the relation $\zeta(3) = \zeta(2, 1)$ in Proposition 1.8. This change of variables can be described nicely in terms of an anti-automorphism on the space $\mathfrak{H}$. For this we denote by $\tau$ the anti-automorphism of $\mathfrak{H}$ which interchanges $x$ and $y$. Here we view $\mathfrak{H}$, and all its subspaces, as $\mathbb{Q}$-algebras where the product is given by the usual non-commutative product in $\mathbb{Q}[x,y]$. That $\tau$ is an anti-automorphism just means that $\tau(ww) = \tau(w)\tau(u)$ for $w, u \in \mathfrak{H}$ and $\tau(1) = 1$. For example if $w = z_3 = xxy$, then

$$\tau(z_3) = \tau(xxy) = \tau(y)\tau(xx) = \tau(y)\tau(x)\tau(x) = xxy = z_2z_1.$$  

Notice that $\tau(\mathfrak{H}^0) \subset \mathfrak{H}^0$, since any non-empty word $w \in \mathfrak{H}^0$ is of the form $w = xuy$ for some $u \in \mathfrak{H}$ and therefore $\tau(w) = \tau(xuy) = \tau(y)\tau(u)\tau(x) = x\tau(u)y \in \mathfrak{H}^0$. Further notice that $\tau$ is an involution, i.e. $\tau^2 = \text{id}_\mathfrak{H}$ and $\tau(\mathfrak{H}^0) = \mathfrak{H}^0$.

**Proposition 2.5 (Duality relation).** For all $w \in \mathfrak{H}^0$ we have

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

**Proof.** This is just a generalization of the variable change $s_j = 1 - t_{k-j+1}$ in the iterated integral expression in Corollary 2.4 similar to (2.6). Interchanging $x$ and $y$ corresponds to $\omega_a(t_{k-j+1}) = -\omega_{\tau(a)}(1-s_j)$ for $a \in \{x,y\}$. The property of $\tau$ being an anti-automorphism corresponds to changing the order/directions of the integrals, which also gets rid of the minus signs.

A few explicit examples of the duality relations are given by the following Corollary, which both can be seen as a generalization of the formula $\zeta(3) = \zeta(2, 1)$. Here we use the common notation $\{k_1, \ldots, k_r\}^n = k_1^{n_1} \cdots k_r^{n_r}$ for $n$ copies of the string $k_1, \ldots, k_r$.

**Corollary 2.6.** i) For all $k \geq 3$ we have

$$\zeta(k) = \zeta(2, 1, \ldots, 1) = \zeta(2, \{1\}^{k-2}).$$

ii) For all $n \geq 1$ we have

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

**Proof.** Both statements are immediate consequences of the duality relations, since $\tau(z_k) = \tau(x^{k-1}y) = xy^{k-1} = z_2z_1 \cdots z_1$ and $\tau((z_2z_1)^n) = \tau(z_2z_1)^n = z_2^n$.

**Remark 2.7.** In Section 3 we will see another proof of the duality relation, which is not using the iterated integral expression. This new proof is based on so-called connected sums, which were just recently introduced by Seki and Yamamoto in [SY]. There we will also see that the duality is true for the $q$-analogue model of Bradley-Zhao (3.7) and that we have $e_q^{\text{BZ}}(\tau(w)) = e_q^{\text{BZ}}(w)$, when considering $\zeta_q^{\text{BZ}}$ as a map from $\mathfrak{H}^1$ to $\mathbb{Q}[q]$. 


2.2 The shuffle & stuffle product and finite double shuffle relations

In this subsection we will introduce the shuffle product \( \shuffle \) and stuffle product \( \ast \) on the spaces \( \mathcal{H} \), \( \mathcal{H}^1 \) and \( \mathcal{H}^0 \). We will then show that the space \( \mathcal{Z} \) is a \( \mathbb{Q} \)-algebra (i.e. give a proof of Proposition 1.11) and see that the map \( \zeta \) in (2.5) is an algebra homomorphism from \( \mathcal{H}^0 \) to \( \mathbb{R} \) with respect to both products \( \shuffle \) and \( \ast \). This will then lead to families of linear relations, which are called finite double shuffle relations.

2.2.1 The shuffle product

The iterated integral expressions give another way to obtain the shuffle product formula

\[
\zeta(k_1)\zeta(k_2) = \sum_{j=2}^{k_1+k_2-1}\left(\frac{j-1}{k_1-1}\right)\left(\frac{j-1}{k_2-1}\right)\zeta(j,k_1+k_2-j)
\]

in Proposition 1.9 which was proved by using partial fraction decomposition. For example we have

\[
\zeta(2)\zeta(3) = \zeta(2, 3) + 3\zeta(3, 2) + 6\zeta(4, 1).
\]

We will now describe how this relation can also be obtained from the iterated integral expression in Corollary 2.4. Using the iterated integral expression of \( \zeta(2) \) and \( \zeta(3) \) we get

\[
\zeta(2)\zeta(3) = \int_{1>s_1>s_2>0} \omega_x(t_1) \omega_y(t_2) \int_{1>s_1>s_2>0} \omega_x(s_1) \omega_x(s_2) \omega_y(s_3) \omega_y(s_3).
\]

By blue we indicate the variables which correspond to the differential form \( \omega_x \) and by red the ones corresponding to \( \omega_y \). This makes it easier to translate the iterated integrals below back to multiple zeta values. Using Fubini’s theorem the right-hand side is the iterated integral of \( \omega_x(t_1) \omega_y(t_2) \omega_x(s_1) \omega_x(s_2) \omega_y(s_3) \omega_y(s_3) \) over the domain where \( 1 > t_1 > t_2 > 0 \) and \( 1 > s_1 > s_2 > s_3 > 0 \). This can be decomposed into the following iterated integrals, where we can neglect the non-trivial intersections \( t_j = s_i \) since they have measure zero.

\[
\zeta(2)\zeta(3) = \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right) + \left(\begin{array}{c}
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2) \\
\int_{1>s_1>s_2>t_1>t_2>0} \omega_x(t_1) \omega_y(t_2)
\end{array}\right)
\]

The above example shows the origin of the name shuffle product, since one can interpret a multiple zeta value as a deck of blue and red cards, which correspond to the differential forms \( \omega_x \) and \( \omega_y \). Taking the product of two multiple zeta values then corresponds just to the shuffle of these two decks of cards. We will now describe this product on the space \( \mathcal{H} \) and its subspaces.
Definition 2.8. We define the shuffle product $\shuffle$ on $\mathcal{H}$ as the $\mathbb{Q}$-bilinear product, which satisfies
\[ 1 \shuffle w = w \shuffle 1 = w \text{ for any word } w \in \mathcal{H} \]
and
\[ a_1 w_1 \shuffle a_2 w_2 = a_1 (w_1 \shuffle a_2 w_2) + a_2 (a_1 w_1 \shuffle w_2) \]
for any letters $a_1, a_2 \in \{x, y\}$ and words $w_1, w_2 \in \mathcal{H}$.

By induction on the lengths of words, one can show that $\shuffle$ is a commutative and associative product and $\mathcal{H}_{\shuffle} = (\mathcal{H}, \shuffle)$ is therefore a commutative $\mathbb{Q}$-algebra. One can also see that the subspaces $\mathcal{H}_1^\shuffle$ and $\mathcal{H}_0^\shuffle$ are both closed under $\shuffle$ and therefore we have subalgebras $\mathcal{H}_0^\shuffle \subset \mathcal{H}_1^\shuffle \subset \mathcal{H}_{\shuffle}$. You can check that this definition corresponds exactly to multiplying iterated integrals as above, i.e. we have
\[ z_2 \shuffle z_3 = xy \shuffle xyz = xyz + 3xxyxy + 6xxxyy = z_2 z_3 + 3z_4 z_2 + 6z_4 z_1. \]
That this is true in general will be proven now.

Proposition 2.9. For any $w, u \in \mathcal{H}_1^\shuffle$ we have
\[ L_{iw}(z) L_{iu}(z) = L_{iw\shuffle u}(z), \]
i.e. the map $L_i$ is an algebra homomorphism from $\mathcal{H}_1^\shuffle$ to $C((0, 1); \mathbb{R})$.

Proof. It is sufficient to prove the statement for words $w, u \in \mathcal{H}_1^\shuffle$. We will do this by induction on the sum of the lengths of $w$ and $u$. If one of them equals the empty word 1, the statement is clear. So let's assume that $w = aw'$ and $u = bu'$ for words $w', u' \in \mathcal{H}_1^\shuffle$ and letters $a, b \in \{x, y\}$. Then we have
\[ \frac{d}{dz} (L_{iw}(z) L_{iu}(z)) = \frac{d}{dz} (L_{iw'}(z) L_{ib'}(z)) = \left( \frac{d}{dz} L_{iw'}(z) \right) L_{ib'}(z) + L_{iw'}(z) \left( \frac{d}{dz} L_{ib'}(z) \right). \]
Using now Lemma 2.2 we get $\frac{d}{dz} L_{iw'}(z) = f_a(z) L_{iw'}(z)$ with $f_x(z) = \frac{1}{z}$ and $f_y(z) = \frac{1}{1-z}$. Using this together with the induction hypothesis we have
\[ \frac{d}{dz} (L_{iw}(z) L_{iu}(z)) = f_a(z) L_{iw'}(z) L_{ib'}(z) + f_b(z) L_{iw'}(z) L_{iu}(z) = f_a(z) L_{iw'\shuffle bu'}(z) + f_b(z) L_{iw'\shuffle u'}(z). \]

Applying Lemma 2.2 again gives
\[ \frac{d}{dz} (L_{iw}(z) L_{iu}(z)) = \frac{d}{dz} L_{i(aw'bu')}(z) + \frac{d}{dz} L_{ib(aw'bu')}(z) = \frac{d}{dz} L_{iw\shuffle u}(z), \]
i.e. $L_{iw}(z) L_{iu}(z) = L_{iw\shuffle u}(z) + c$ for some constant $c$. But since both sides vanish for $z = 0$, we conclude $c = 0$.

For $w, u \in \mathcal{H}_0^\shuffle$ we can also set $z = 1$ in the Proposition above and obtain the following.

Corollary 2.10. For any $w, u \in \mathcal{H}_0^\shuffle$ we have
\[ \zeta(w) \zeta(u) = \zeta(w \shuffle u). \]
In particular the space $\mathcal{Z}$ is a $\mathbb{Q}$-subalgebra of $\mathbb{R}$ and $\zeta$ is an algebra homomorphism from $\mathcal{H}_0^\shuffle$ to $\mathcal{Z}$. 

Version 3.9 (June 26, 2020)
2.2.2 The stuffle product

In Section 2.1 we saw that for \( k_1, k_2 \geq 2 \) we have the stuffle product formula

\[
\zeta(k_1) \zeta(k_2) = \left( \sum_{m_1 > m_2 > 0} + \sum_{m_2 > m_1 > 0} + \sum_{m_1 = m_2 > 0} \right) \frac{1}{m_1^{k_1} m_2^{k_2}} = \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2).
\]

With the same argument, i.e. splitting up the summation, we get for \( k_1, k_2 \geq 2, k_3 \geq 1 \)

\[
\zeta(k_1) \zeta(k_2, k_3) = \zeta(k_1, k_2, k_3) + \zeta(k_2, k_1, k_3) + \zeta(k_2, k_3, k_1) + \zeta(k_1 + k_2, k_3) + \zeta(k_2, k_1 + k_3).
\]

Similar to the shuffle product we will now define the stuffle product \( \ast \) on the space \( \mathcal{H}^1 \) and \( \mathcal{H}^0 \) and then show that \( \zeta \) is also an algebra homomorphism with respect to this product. Recall that \( \mathcal{H}^1 = \mathbb{Q} \langle z_1, z_2, \ldots \rangle \), i.e. every element in \( \mathcal{H}^1 \) can be viewed as a linear combination of words in the letters \( z_j \) instead of the letters \( x, y \). Here and in the following we will use the terminology ‘word’ in these two different ways and it will be clear from context if we talk about words in the \( z_j \) or \( x, y \). In the next subsection we will see that the shuffle and the stuffle product are both examples of quasi-shuffle products over different alphabets.

**Definition 2.11.** We define the stuffle product \( \ast \) on \( \mathcal{H}^1 \) as the \( \mathbb{Q} \)-bilinear product, which satisfies \( 1 \ast w = w \ast 1 = w \) for any word \( w \in \mathcal{H}^1 \) and

\[
z_i w_1 \ast z_j w_2 = z_i (w_1 \ast z_j w_2) + z_j (z_i w_1 \ast w_2) + z_{i+j} (w_1 \ast w_2)
\]

for any \( i, j \geq 1 \) and words \( w_1, w_2 \in \mathcal{H}^1 \).

Notice that this product also replicated the above product formula of multiple zeta values, since

\[
z_{k_1} \ast z_{k_2} = z_{k_1} z_{k_2} + z_{k_2} z_{k_1} + z_{k_1+k_2}.
\]

This product is also called the harmonic product and one can check (see [H1]) that it is commutative and associate and therefore \( \mathcal{H}^1 = (\mathcal{H}^1, \ast) \) is a commutative \( \mathbb{Q} \)-algebra. By definition it is easy to check that \( \mathcal{H}^0 \) is also closed under \( \ast \) and we get a subalgebra \( \mathcal{H}^0 \subseteq \mathcal{H}^1 \).

In the case of the shuffle product we used the polylogarithm to prove that the product of multiple zeta values satisfy the shuffle product formula by considering \( z = 1 \). In the case of the stuffle product, we will consider the **truncated multiple zeta values** (also often called multiple harmonic sums), which are for an integer \( M \geq 1 \) and any index \( k = (k_1, \ldots, k_r) \in \mathbb{Z}_{\geq 1}^r \) defined by

\[
\zeta_M(k) = \zeta_M(k_1, \ldots, k_r) = \sum_{M > m_1 > \cdots > m_r > 0} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \in \mathbb{Q}.
\]

Clearly if \( k \) is admissible we have \( \lim_{M \to \infty} \zeta_M(k) = \zeta(k) \). For a fixed \( M \) we can view \( \zeta_M \) as a \( \mathbb{Q} \)-linear map from \( \mathcal{H}^1 \) to \( \mathbb{Q} \), defined on the generators by \( \zeta_M : z_k \mapsto \zeta_M(k) \).

**Proposition 2.12.** For any \( w, u \in \mathcal{H}^1 \) and \( M \geq 1 \) we have

\[
\zeta_M(w) \zeta_M(u) = \zeta_M(w \ast u),
\]

i.e. the map \( \zeta_M \) is an algebra homomorphism from \( \mathcal{H}^1 \) to \( \mathbb{Q} \).

**Proof.** This can be done by induction on the depths or \( M \) (Exercise 6 ii). We will prove this in more general form for quasi-shuffle algebras in the next section (Lemma 2.18). \( \square \)

For \( w, u \in \mathcal{H}^0 \) we can also take the limit \( M \to \infty \) in the Proposition above and obtain the following.

**Corollary 2.13.** For any \( w, u \in \mathcal{H}^0 \) we have

\[
\zeta(w) \zeta(v) = \zeta(w \ast v),
\]

i.e. \( \zeta \) is an algebra homomorphism from \( \mathcal{H}^0 \) to \( \mathbb{Z} \).

Version 3.9 (June 26, 2020)
2.2.3 Finite double shuffle relations

Since the map $\zeta: H^0 \to \mathbb{Z}$ is an algebra homomorphism with respect to the shuffle product $\shuffle$ and the stuffle product $\ast$, we get a large family of linear relations among multiple zeta values.

**Proposition 2.14** (Finite double shuffle relations). For $w, u \in H^0$ we have

$$\zeta(w \shuffle u - w \ast u) = 0.$$ 

But it is also clear, that these do not give all linear relations among multiple zeta values. For example the relation $\zeta(3) = \zeta(2, 1)$ is not a consequence of the above Proposition. Counting the finite double shuffle relations, we get the following table, which comes from the survey article [118]. In this article, you can also find the numbers of other families of relations, such as the duality relation.

<table>
<thead>
<tr>
<th>weight $k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td># all conjectured relations</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>29</td>
<td>60</td>
<td>123</td>
<td>249</td>
<td>503</td>
<td>1012</td>
</tr>
<tr>
<td># finite double shuffle relations</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>16</td>
<td>40</td>
<td>92</td>
<td>200</td>
<td>429</td>
<td>902</td>
</tr>
</tbody>
</table>

We see that the first possible finite double shuffle relations appears in weight 4 by choosing $w = u = z_2$, which gives

$$w \shuffle u - w \ast u = (2z_2z_2 + 4z_3z_1) - (2z_2z_2 + z_4) = 4z_3z_1 - z_4$$

i.e. $4\zeta(3, 1) = \zeta(4)$. This relation is a special case of the following family of linear relations which is a consequence of finite double shuffle relations.

**Proposition 2.15.** For all $n \geq 1$ we have

$$4^n \zeta(\{3, 1\}^n) = \zeta(\{4\}^n).$$

**Proof.** This can be done by proving the following equations in $\mathcal{H}^0$ (Exercise 7)

$$\sum_{j=-n}^n (-1)^j z_2^{-j} \shuffle z_2^{n+j} = 4^n(z_3z_1)^n, \quad \sum_{j=-n}^n (-1)^j z_2^{-j} \ast z_2^{n+j} = z_4^n.$$ 

(Notice: Here $z_k^n$ means $z_k z_k \ldots z_k$, i.e. the usual non-commutative product in $\mathcal{H}$ and not the shuffle or stuffle product.)

Together with the explicit formula (Exercise 7)

$$\zeta(\{2\}^n) = \frac{\pi^{2n}}{(2n+1)!}$$

(2.7)

the proof of Proposition 2.15 can also be used to show

$$\zeta(\{4\}^n) = \frac{4^n2\pi^{4n}}{(4n+2)!}, \quad \zeta(\{3, 1\}^n) = \frac{2\pi^{4n}}{(4n+2)!},$$

where the second equation here is known as the 3-1 formula for multiple zeta values.
2.3 Quasi-shuffle algebras

In this section, we want to generalize what we did in the previous section. On $\mathcal{S}_j$, we defined the shuffle product on words in letters $x$ and $y$, and on $\mathcal{S}_1^1$, we defined the shuffle product on words in letters $z_j$ for $j \geq 1$. This idea will be generalized now by allowing an arbitrary set of letters $A$ and then define a product on the space of words in these letters. We will mainly follow the definitions and theorems in [HI] and [IKZ], but will also introduce some definitions and theorems, which can not be found in the literature.

2.3.1 The algebra of letters and words

In the following, we assume that $k$ is a field containing $\mathbb{Q}$, and $A$ is a countable set to which we refer as the set of letters. Let $kA$ be the $k$-vector space generated by $A$ and let $\circ$ be a $k$-bilinear, associative and commutative product on $kA$. We obtain a (non-unital) $k$-algebra $(kA, \circ)$, to which we refer as the algebra of letters. Notice that in the work [HI] the authors just assume that $\circ$ is associative and commutative and do not consider $kA$ as an algebra, but it will be useful for our purposes (e.g., Lemma 2.18). For such a product $\circ$ on letters, we want to assign a product $\ast_\circ$ on the space of words $k\langle A \rangle$, which generalizes the shuffle and stuffle product we have seen before. Here a monomial $w = a_1 \cdots a_l$ in $k\langle A \rangle = k\langle a_1, a_2, \ldots \rangle$ will again be called a word and the unit $(l = 0)$, denoted by $1$, is again called the empty word. By $\ell(w) = l$ we denote the length of the word $w$.

**Definition 2.16.** Let $\circ$ be a product on $kA$ as above. Then we define the quasi-shuffle product $\ast_\circ$ on $k\langle A \rangle$ as the $k$-bilinear product, which satisfies $1 \ast_\circ w = w \ast_\circ 1 = w$ for any word $w \in k\langle A \rangle$ and

$$aw \ast_\circ bv = a(w \ast_\circ bv) + b(aw \ast_\circ v) + (a \circ b)(w \ast_\circ v) \quad (2.8)$$

for any letters $a, b \in A$ and words $w, v \in k\langle A \rangle$.

**Theorem 2.17.** The space $k\langle A \rangle$ equipped with the product $\ast_\circ$ becomes a commutative $k$-algebra.

**Proof.** This is Theorem 2.1 in [HI]. It suffices to show that $\ast_\circ$ is commutative and associative, which can be done straightforward by induction on the lengths of words.

We will call $(k\langle A \rangle, \ast_\circ)$ a quasi-shuffle algebra or algebra of words. Notice that this generalized the $\mathbb{Q}$-algebra $\mathcal{S}_0$ by choosing $k = \mathbb{Q}$, $A = \{x, y\}$ and $a \circ b = 0$ for any letters $\{x, y\}$ and it generalizes the $\mathbb{Q}$-algebra $\mathcal{S}_1$, by choosing $k = \mathbb{Q}$, $A = \{z_1, z_2, \ldots\}$ and $z_i \circ z_j = z_{i+j}$ (See section 2.3.2 for details).

Our purpose to introduce quasi-shuffle products is to also describe the product structure of the $q$-series $g(k)$ and the (modified) $q$-analogues $\zeta_q$.

**Lemma 2.18.** Let $R$ be a $k$-algebra and $f_m : (kA, \circ) \to R$ be $k$-algebra homomorphisms for $m \geq 1$. Then for all $M \geq 1$ the $k$-linear map $F_M : k\langle A \rangle \to R$ defined on a word $w = a_1 \cdots a_r \in k\langle A \rangle$ by

$$F_M(w) = \sum_{M > m_1 > \cdots > m_r > 0} f_{m_1}(a_1) \cdots f_{m_r}(a_r)$$

and $F_M(1) = 1$ is a $k$-algebra homomorphism from $(k\langle A \rangle, \ast_\circ)$ to $R$.

**Proof.** It suffices to show that for any $M \geq 1$ and words $w, v \in k\langle A \rangle$ we have

$$F_M(w)F_M(v) = F_M(w \ast_\circ v).$$
We will prove this by induction on $M$. The case $M = 1$ is trivial, since $F_1(w) = 0$ for all non-empty $w$ and $1 = F_1(1)F_1(1) = F_1(1\ast_s 1)$. Notice that we have $F_M(aw) = \sum_{M > m > 0} f_m(a)F_m(w)$ for a letter $a$ and a word $w$. For $w = aw'$, $v = bv'$ with letters $a, b \in A$ and words $w', v' \in k\langle A \rangle$ we therefore get

$$F_M(w)F_M(v) = \sum_{M > m > 0} f_m(a)F_m(w') \sum_{M > n > 0} f_n(b)F_n(v')$$

$$= \left( \sum_{M > m > n > 0} + \sum_{M > m > n > 0} + \sum_{M > m = n > 0} \right) f_m(a)F_m(w')f_n(b)F_n(v')$$

$$= \sum_{M > m > 0} f_m(a)F_m(w')F_m(bv') + \sum_{M > n > 0} f_n(b)F_n(aw')F_n(v') + \sum_{M > m > n > 0} f_m(a)f_n(b)F_m(w')F_n(v')$$

$$= \sum_{M > m > 0} f_m(a)F_m(w' \ast_s bv') + \sum_{M > n > 0} f_n(b)F_n(aw' \ast \ast v') + \sum_{M > m > n > 0} f_m(a)f_n(b)F_m(w' \ast \ast v')$$

$$= F_M(a(w' \ast_s bv')) + F_M(b(aw' \ast \ast v')) + F_M((a \ast b)(w' \ast \ast v')) = F_M(aw \ast_s v).$$

Here we used that $f_m$ is an algebra homomorphism together with the induction hypothesis in the fourth equation. □

### 2.3.2 Examples of quasi-shuffle products & (sub)algebras

In the following, we give a few explicit examples for quasi-shuffle products and algebras which appear in this course. We will also consider certain subalgebras, and the first statement we want to show now is that subalgebras of the algebra of letters gives subalgebras of the algebra of words.

**Proposition 2.19.** If $B \subset A$ is a subset of letters such that $(kB, \circ)$ is a subalgebra of $(kA, \circ)$, then $(k\langle B \rangle, \ast_\circ)$ is a subalgebra of $(k\langle A \rangle, \ast_\circ)$.

**Proof.** We need to show that $k\langle B \rangle$ is closed under $\ast_\circ$. But this follows directly from the definition of $\ast_\circ$, since if $kB$ is closed under $\circ$ then we have for $a, b \in B$ that $a \circ b \in B$. Therefore all elements on the right-hand side of (2.8) are in $k\langle B \rangle$ if $w, v \in k\langle B \rangle$ and $a, b \in B$. □

Most of our objects depend on some index $k = (k_1, \ldots, k_r)$ and therefore most of our examples use the set of letters, the ”$z$-alphabet”, defined by

$$A_z := \{z_1, z_2, \ldots \}.$$

Notice that we have a abuse of notion here, since $z_k = x^{k-1}y$ denoted elements in $\mathcal{H}$ previously. But from the context it should always be clear if we talk about the formal elements $z_k$ in $A_z$ or the elements in $\mathcal{H}$.

#### i) Shuffle product

For any field $k$ and any set of letters $A$, one can define the trivial product $a \circ b = 0$ for $a, b \in A$. The resulting quasi-shuffle product is then just the shuffle product $\shuffle = \ast_\circ$.

As a special case we considered $k = \mathbb{Q}$, $A = \{x, y\}$ before but will also deal with the shuffle product on $A_z$ later, which is sometimes also called the ”index-shuffle product”.

#### ii) Stuffle product

Another example we considered before is the stuffle product. Choosing $k = \mathbb{Q}$, $A_z = \{z_1, z_2, \ldots \}$ and $z_i \circ z_j = z_{i+j}$ we write $\ast = \ast_\circ$. Here the $z_j$ are considered as variables itself, but we see that $(\mathbb{Q}[A_z], \ast)$ is isomorphic to $\mathcal{H}_k^1$ as a $\mathbb{Q}$-algebra, when sending $z_k$ to $x^{k-1}y$.

Notice that the algebra of letters $(\mathbb{Q}[A_z], \circ)$ is isomorphic to $X\mathbb{Q}[X]$, by sending $z_k$ to $x^k$. Defining for $m \geq 1$ the algebra homomorphisms $f_m : X\mathbb{Q}[X] \to \mathbb{Q}$ by $p \mapsto p(m^{-1})$ gives the truncated
multiple zeta values $\zeta_M$ as the function $F_M$ in Lemma 2.18. In particular we obtain Proposition 2.12 as a consequence. Considering the subsets $A^{\geq 2}_z \subset A_z$ and $A^{ev}_z \subset A_z$ defined by

$$A^{\geq 2}_z := A_z \setminus \{z_1\} = \{z_2, z_3, z_4, \ldots\}, \quad A^{ev}_z := \{z_2, z_4, z_6, \ldots\},$$

we clearly have that $(Q A^{\geq 2}_z, \circ)$ and $(Q A^{ev}_z, \circ)$ are subalgebras of $(Q A_z, \circ)$. As a consequence of Proposition 2.19 we see that

$$Z^{\geq 2} = \langle \zeta(k_1, \ldots, k_r) \mid r \geq 0, k_1, \ldots, k_r \geq 2 \rangle_Q,$$

$$Z^{ev} = \langle \zeta(k_1, \ldots, k_r) \mid r \geq 0, k_1, \ldots, k_r \geq 2 \text{ even} \rangle_Q$$

are subalgebras of $Z$. Notice that by Brown’s Theorem 1.17 we actually have $Z^{\geq 2} = Z$.

iii) The $q$-series $g(k)$: Recall that we defined for an index $k = (k_1, \ldots, k_r)$ the modified $q$-analogues

$$g(k) = g(k_1, \ldots, k_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{P_{k_1}(q^{m_1})}{(1 - q^{m_1})^{k_1}} \cdots \frac{P_{k_r}(q^{m_r})}{(1 - q^{m_r})^{k_r}} \in \mathbb{Q}[[q]].$$

Inspired by Lemma 1.28 we define on $QA_z$ the product

$$z_{k_1} \hat{\circ} z_{k_2} = z_{k_1+k_2} + \sum_{j=1}^{k_1+k_2-1} \left( \lambda^j_{k_1,k_2} + \lambda^j_{k_2,k_1} \right) z_j$$

(2.9)

where the rational numbers $\lambda^j_{k_1,k_2}$ are given by

$$\lambda^j_{k_1,k_2} = (-1)^{k_2-1} \binom{k_1+k_2-1-j}{k_1-j} \frac{B_{k_1+k_2-j}}{(k_1+k_2-j)!}.$$ 

It is easy to see that this product is indeed associative and commutative. By Lemma 1.28 we see that for $m \geq 1$ the map $f_m : QA_z \to \mathbb{Q}[[q]]$ defined on the generators by

$$f_m(z_k) = \frac{P_k(q^m)}{(1 - q^m)^k}$$

is a $\mathbb{Q}$-algebra homomorphism from $(QA_z, \hat{\circ})$ to $\mathbb{Q}[[q]]$. We will denote the corresponding quasi-shuffle product by $\hat{\circ} = *_{\circ}$. Using Lemma 2.18 we see that, after taking the limit $M \to \infty$, that the space $G$, spanned by all $g(k)$, is a $\mathbb{Q}$-subalgebra of $\mathbb{Q}[[q]]$ and we can view $G$ as an algebra homomorphism from $(QA_z, \hat{\circ})$ to $G$. This proofs Proposition 1.27.

**Proposition 2.20.** The subspaces $G^{ev} \subset G^{\geq 2} \subset G$, defined by

$$G^{\geq 2} = \langle g(k_1, \ldots, k_r) \mid r \geq 0, k_1, \ldots, k_r \geq 2 \rangle_Q,$$

$$G^{ev} = \langle g(k_1, \ldots, k_r) \mid r \geq 0, k_1, k_2, \ldots, k_r \geq 2 \text{ even} \rangle_Q,$$

are $\mathbb{Q}$-subalgebras of $G$. 


2.3.3 Words of repeating letters

iv) Bradley-Zhao q-MZV: We defined for an admissible index \( k = (k_1, \ldots, k_r) \)
\[
\mathcal{S}^\text{BZ}_q(k_1, \ldots, k_r) = \sum_{m_1 > \ldots > m_r > 0} \frac{q^{(k_1-1)m_1} \cdots q^{(k_r-1)m_r}}{[m_1]_q^{k_1} \cdots [m_r]_q^{k_r}}.
\]

Since we have for \( m \geq 1 \) and \( k_1, k_2 \geq 2 \)
\[
\frac{q^{(k_1-1)m} q^{(k_2-1)m}}{[m]_q^{k_1} [m]_q^{k_2}} = \frac{q^{(k_1+k_2-2)m}}{[m]_q^{k_1+k_2-1}} + (1-q) \frac{q^{(k_1+k_2-2)m}}{[m]_q^{k_1+k_2-1}}
\]
we choose \( k = \mathbb{Q}(1-q) \) and define on \( kA_z \) the product
\[
z_{k_1} \circ z_{k_2} = z_{k_1+k_2} + (1-q)z_{k_1+k_2-1}.
\]

As before we get a quasi-shuffle algebra \((kA_z, *)_q\) and for a \( M \geq 1 \) an algebra homomorphism
\( F_M \) to \( \mathbb{Q}[[q]] \) by sending \( z_{k_1}, \ldots, z_{k_r} \) to the truncated version \( \mathcal{S}^\text{BZ}_M(k_1, \ldots, k_r) \) defined in the obvious way. Notice that one could also consider the modified version \( \mathcal{S}'_q(k) := (1-q)^{-\text{wt}(k)} \mathcal{S}^\text{BZ}_q(k) \). With this one can choose again \( k = \mathbb{Q} \) and use the product \( z_{k_1} \circ z_{k_2} = z_{k_1+k_2} + z_{k_1+k_2-1} \). The algebraic structure of these modifies version are for example studied in [Tak].

v) Generalized modified q-MZV: Take \( A_{z,Q} = \{z_Q^k \mid k \geq 1, Q \in \mathbb{Q}[X], \deg(Q) \leq k \} \) and define
\[
z_{k_1} \circ z_{k_2} = z_{k_1+k_2}.
\]

Then clearly \( f_m : (AQ_1, \circ) \to \mathbb{Q}[[q]] \) given by \( f_m(z_Q^k) = \frac{Q(q^m)}{(1-q^m)^k} \) is an \( \mathbb{Q} \)-algebra homomorphism. Again as a consequence of Lemma 1.28 we see that the space \( Z_q \) is a \( \mathbb{Q} \)-subalgebra of \( \mathbb{Q}[[q]] \).

2.3.3 Words of repeating letters

In the following subsection, we want to state some standard facts on quasi-shuffle algebras, which were established in [HI], [H2], and [IKZ] (without using the notion of quasi-shuffle algebras). Some of these will be given without proofs and we refer the reader to the above references for details.
Let \( f = \sum_{n=0}^{\infty} c_n T^n \in k[[T]] \) and \( \bullet \in \{ \diamond, *_{\diamond} \} \), then we define for \( a \in A \)

\[
f_*(aX) = \sum_{n=0}^{\infty} c_n a \bullet \cdots \bullet a X^n = \sum_{n=0}^{\infty} c_n a^n X^n \in k\langle A \rangle[[X]].
\]

In other words \( f_*(aX) \) means, that we plug \( aX \) into the power series \( f \), and then use the product \( \bullet \) to evaluate the products of \( a \) in \( (aX)^n \). Therefore it also makes sense to consider \( f_*(zX) \) for any \( z \in kA[[X]] \) and then evaluate \( (zX)^n \) as an element in \( k\langle A \rangle, *_{\diamond}[[X]] \) or \( k\langle A \rangle, \diamond[[X]] \). This we will do in Proposition 2.21 below. Notice that we have two different products on \( k\langle A \rangle \) given by the quasi-shuffle product \( *_{\diamond} \) and given by the usual non-commutative multiplication. When calculating with elements in \( k\langle A \rangle[[X]] \), we will always do it with respect to the usual non-commutative multiplication in \( k\langle A \rangle \).

**Proposition 2.21.** For all \( z \in k\langle A \rangle[[X]] \) we have

\[
\exp_{*_{\diamond}}(\log_{\diamond}(1 + zX)) = \frac{1}{1 - zX}.
\]

**Proof.** This is [HI, Corollary 5.1], but it will also be a consequence of Proposition 2.29 below. \( \square \)

Asume we have an algebra homomorphism \( \varphi : (k\langle A \rangle, \diamond_{\diamond}) \to R \) in some \( k \)-algebra \( R \). Applying \( \varphi \) to (2.12) with \( z = a \in A \) gives the following equation in \( R[[X]] \)

\[
\exp\left( \sum_{n=1}^{\infty} (-1)^{n-1} \varphi(a^n) \frac{X^n}{n} \right) = 1 + \sum_{n=1}^{\infty} \varphi(a^n) X^n.
\]

(2.13)

In particular we have

\[
\varphi(a^n) \in k \left[ \varphi(a^j) \mid 1 \leq j \leq n \right],
\]

i.e. for \( a \in A \) the \( \varphi(a^n) = \varphi(aa \cdots a) \) is a polynomial in \( \varphi(a^j) \) with \( 1 \leq j \leq n \).

**Corollary 2.22.**

**i)** For all \( k, M \geq 1 \) we have

\[
\exp\left( \sum_{n=1}^{\infty} (-1)^{n-1} \zeta_M(nk) \frac{X^n}{n} \right) = 1 + \sum_{n=1}^{\infty} \zeta_M([k]^n) X^n
\]

and therefore \( \zeta_M([k]^n) \in \mathbb{Q}[z_M([jk]) \mid 1 \leq j \leq n] \) for all \( n \geq 1 \). In particular, for \( k \geq 2 \) these statements also hold by replacing \( \zeta_M \) with \( \zeta \).

**ii)** For all \( k, n \geq 1 \) we have

\[
g([k]^n) \in \mathbb{Q}[g(j) \mid 1 \leq j \leq kn].
\]

In addition if \( k \geq 2 \) is even, then \( g([k]^n) \in \mathbb{Q}[g(j) \mid 2 \leq j \leq kn, j \text{ even}] \)

**Proof.** Statement i) follows directly from (2.13) by using the algebra homomorphism \( \zeta_M : (k\langle A \rangle, *) \to \mathbb{Q}, a = z_k \) and the fact that \( z_k^n = z_{kn} \) if \( z_i \diamond z_j = z_{i+j} \). For ii) we use the algebra homomorphism \( g : (k\langle A \rangle, \diamond) \to \mathbb{Q}[g] \) together with (2.14). The second part of ii) follows since \( kA_{even} \) is closed under \( \diamond \) as we saw in Proposition 2.20 and therefore \( z_k^n \in kA_{even} \) if \( k \) is even. \( \square \)
2.3.4 Application: Quasi-modular forms

By Proposition 1.1 we know that for \( m \geq 1 \) we have \( \zeta(2m) \in \mathbb{Q}\pi^{2m} \) and with Corollary 2.22 we get
\[
\zeta(2m, \ldots, 2m) \in \mathbb{Q}[\zeta(2)] = \mathbb{Q}[\pi^2].
\]

For example as we have already seen before, we have for all \( n \geq 1 \)
\[
\zeta(\{2\}^n) = \frac{\pi^{2n}}{(2n+1)!}, \quad \zeta(\{4\}^n) = \frac{4^n2^{4n}}{(4n+2)!}.
\]

A similar statement can also be shown for the \( q \)-series \( g \) and the \( q \)-analogon of \( \mathbb{Q}[\zeta(2)] \) is given by the ring of quasi-modular forms (with rational coefficients), defined by
\[
\tilde{\mathcal{M}}^Q := \mathbb{Q}[g(2), g(4), g(6)].
\]

As an anlogon from Euler formula for \( \zeta(2m) \) we get the following.

**Theorem 2.23.** For all \( m \geq 1 \) we have \( g(2m) \in \tilde{\mathcal{M}}^Q \).

**Proof.** We will proof this in Section 5 and also give an explicit formula similar to the one of Euler for \( \zeta(2m) \) in terms of \( \zeta(2) \). But with some work this also follows from Proposition 1.29 and 1.30 (Similar to Exercise 1). \( \square \)

**Corollary 2.24.** For all \( m \geq 1 \) we have
\[
g(2m, \ldots, 2m) \in \tilde{\mathcal{M}}^Q.
\]

**Proof.** This now follows directly from Theorem 2.23 and Corollary 2.22 ii). \( \square \)

**Proposition 2.25.**

i) The space \( \tilde{\mathcal{M}}^Q \) is closed under \( q\frac{d}{dq} \).

ii) We have
\[
\tilde{\mathcal{M}}^Q = \mathbb{Q}[g(2), g'(2), g''(2)].
\]

where \( g' \) denotes the derivative with respect to \( q\frac{d}{dq} \).

iii) Let \( k = (k_1, \ldots, k_r) \) be an index with \( k_1, \ldots, k_r \geq 2 \) even. Then we have
\[
g^{sym}(k) := \sum_{\sigma \in S_r} g(k_{\sigma(1)}, \ldots, k_{\sigma(r)}) \in \tilde{\mathcal{M}}^Q,
\]

where \( S_r \) denotes the set of all permutations of \( \{1, \ldots, r\} \).

iv) We have
\[
\tilde{\mathcal{M}}^Q = \left\langle g^{sym}(k_1, \ldots, k_r) \mid r \geq 0, k_1 \geq k_2 \geq \cdots \geq k_r \geq 2 \text{ even} \right\rangle_{\mathbb{Q}},
\]

where we set \( g^{sym}(0) = 1 \).

**Proof.** This is Exercise 8. The first statement can be proven with Proposition 1.29 and 1.30 by giving explicit formulas for \( q\frac{d}{dq} g(2), q\frac{d}{dq} g(4) \) and \( q\frac{d}{dq} g(6) \) as polynomials in \( g(2), g(4) \) and \( g(6) \). From this one also deduces ii). Statement iii) and iv) can be proven by induction on \( r \) and the weight respectively. \( \square \)
2.3.5 Linear maps induced by power series

In this section, we want to illustrate an important tool for quasi-shuffle algebras, which was first established in \( [H2] \) and later generalized in [HII] Section 3. Also, there is a recent work of Yamamoto [Y], which generalizes this construction even more and which gives a nice reinterpretation of some of the results we will mention here. One motivation to study these maps is the following: For a given quasi-shuffle algebra \((k(A), \ast_0)\), one can always construct an explicit isomorphism (of \(k\)-algebras) to \((k\langle A \rangle, \cup)\). In other words, all quasi-shuffle algebras over the same alphabet are isomorphic.

We will first illustrate this the basic idea on multiple zeta values, which are the image of an algebra homomorphism from \((\mathbb{Q}A_z, \ast)\) to \(\mathbb{R}\). We will ignore convergence issues for now, since everything we are going to do could also be done for the truncated version \(\zeta_M\). Recall that shuffle product formulas in small depths are given by

\[
\zeta(k_1)\zeta(k_2) = \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2)
\]

\[
\zeta(k_1)\zeta(k_2, k_3) = \zeta(k_1, k_2, k_3) + \zeta(k_2, k_1, k_3) + \zeta(k_3, k_2, k_1) + \zeta(k_1 + k_2, k_3) + \zeta(k_2 + k_3, k_1) + \zeta(k_1, k_2 + k_3).
\]

(2.15)

Now one could ask for a new object \(S(k_1, \ldots, k_r) \in \mathbb{R}\), which does not satisfy the stuffle product formula, but the index shuffle product formula, i.e. which is an image of an algebra homomorphism from \((\mathbb{Q}A_z, \cup)\) to \(\mathbb{R}\). By this we mean we want to construct something out of the multiple zeta values which satisfies in low depths

\[
S(k_1)S(k_2) = S(k_1, k_2) + S(k_2, k_1),
\]

\[
S(k_1)S(k_2, k_3) = S(k_1, k_2, k_3) + S(k_2, k_1, k_3) + S(k_3, k_2, k_1).
\]

(2.16)

One can check easily that \(S(k) = \zeta(k)\) and

\[
S(k_1, k_2) = \zeta(k_1, k_2) + \frac{1}{2}\zeta(k_1 + k_2),
\]

\[
S(k_1, k_2, k_3) = \zeta(k_1, k_2, k_3) + \frac{1}{2}\zeta(k_1 + k_2, k_3) + \frac{1}{2}\zeta(k_1, k_2 + k_3) + \frac{1}{6}\zeta(k_1 + k_2 + k_3),
\]

(2.17)

satisfy the index-shuffle product formula (2.16) as a consequence of the shuffle product formula (2.15). One might guess, that this works in arbitrary depths and that the coefficients are given by \(\frac{1}{n!}\), whenever one adds together \(n\) indices. Since \(\frac{1}{n!}\) is the coefficient of \(T^n\) in \(\exp(T)\) one might think of the above constructions as some exponential map \(\exp: (\mathbb{Q}A_z, \cup) \to (\mathbb{Q}A_z, \ast)\), which gives us the algebra homomorphism \(S: (\mathbb{Q}A_z, \cup) \to \mathbb{R}\) by setting \(S = \zeta \circ \exp\). In this section we want to make this precise, by associating to some power series \(f \in T_k[\{T\}]\) a linear map \(\Psi_f : k(A) \to k(A)\). In the case \(f(T) = \exp(T) - 1\) we will get the map in the example above.

Now let \(w = a_1a_2 \cdots a_n\) be a word of length \(\ell(w) = n\) with letters \(a_1, \ldots, a_n \in A\). Let \(I = (i_1, \ldots, i_m)\) be a composition of \(n\), i.e. \(i_1 + \cdots + i_m = n\) with \(m \geq 1, i_1, \ldots, i_m \geq 1\). For such an \(I\) we define

\[
I[w] = (a_1 \circ \cdots \circ a_{i_1})(a_{i_1+1} \circ \cdots \circ a_{i_1+i_2}) \cdots (a_{i_1+\cdots+i_{m-1}} \circ \cdots \circ a_n).
\]

For example for \(w = a_1a_2a_3\) a composition of \(n = 3\) is given by \(I = (1, 2)\), and we get \(I[w] = a_1(a_2 \circ a_3)\). By \(\mathcal{C}(n)\) we denote the set of all compositions of \(n\) and usually \(n\) will be given by the length \(n = \ell(w)\) of some word \(w\).

**Definition 2.26.** For a formal power series \(f = \sum_{i=1}^{\infty} c_i T^i \in T_k[\{T\}]\) we define the \(k\)-linear map \(\Psi_f : k(A) \to k(A)\) by \(\Psi_f(1) = 1\) and

\[
\Psi_f(w) = \sum_{I = (i_1, \ldots, i_m) \in \mathcal{C}(\ell(w))} c_{i_1} \cdots c_{i_m} I[w]
\]

for a nonempty word \(w\).
For example for words of length 2 and 3 we have $C(2) = \{(1, 1), (2)\}$ and $C(3) = \{(1, 1, 1), (2, 1), (1, 2), (3)\}$ which gives for $a_1, a_2, a_3 \in A$

\[
\Psi_f(a_1a_2) = c_1^2 a_1 a_2 + c_2 a_1 \circ a_2, \\
\Psi_f(a_1a_2a_3) = c_1^3 a_1 a_2 a_3 + c_2 c_3 ((a_1 \circ a_2) a_3 + a_1 (a_2 \circ a_3)) + c_3 a_1 \circ a_2 \circ a_3.
\]

Observe the similarity of this to (2.17) when $A = A_2$, $z_i \circ z_j = z_{i+j}$ and $c_i = \frac{1}{i!}$, i.e. $f(T) = \exp(T) - 1$. In the case $f = T$, we obtain the identity map on $k\langle A \rangle$, i.e. $\Psi_f(w) = w$ for all $w \in k\langle A \rangle$. Now for $f, g \in Tk[[T]]$ denote by $f \circ g \in Tk[[T]]$ the usual composition of formal power series, i.e. the power series given by $(f \circ g)(T) = f(g(T))$.

**Theorem 2.27.** For $f, g \in Tk[[T]]$ we have

\[
\Psi_f \Psi_g = \Psi_{f \circ g}.
\]

**Proof.** This can be found in [HI, Theorem 3.1] and follows from a straightforward but tedious calculation, which we will omit here. \hfill \square

One interesting case of the linear maps $\Psi_f$ and $\Psi_g$ is given by $f(T) = \exp(T) - 1$ and $g(T) = \log(1+T)$. In these cases we have $(f \circ g)(T) = (g \circ f)(T) = T$ and therefore $\Psi_f$ and $\Psi_g$ are inverse to each other by Theorem 2.27. As in [HI] we write $\exp := \Psi_f$ and $\log := \Psi_g$, i.e. $\exp, \log \in \text{Hom}(k\langle A \rangle, k\langle A \rangle)$. As illustrated in the introduction of this subsection we have the following result, which was first proven in [HI].

**Theorem 2.28.** The map

\[
\exp : (k\langle A \rangle, \cup) \longrightarrow (k\langle A \rangle, \ast)
\]

is an $k$-algebra isomorphism with inverse

\[
\log : (k\langle A \rangle, \ast) \longrightarrow (k\langle A \rangle, \cup).
\]

**Proof.** This is [HI, Theorem 2.5]. \hfill \square

Theorem 2.28 will be used in Section 5 in the construction of shuffle regularized multiple Eisenstein series, which were introduced in [HT]. Another application will be the proof of Proposition 2.29 which we will do now. For this, we will first give a more general statement on the linear maps $\Psi_f$.

**Proposition 2.29.** For $f \in Tk[[T]]$ and $z \in kA[[X]]$ we have the following equality in $k\langle A \rangle[[X]]$

\[
\Psi_f \left( \frac{1}{1-zX} \right) = \frac{1}{1-f_0(zX)}.
\]

Here $\Psi_f$ acts on $k\langle A \rangle[[X]]$ component wise.

**Proof.** Let $f(T) = \sum_{i=1}^{\infty} c_i T^i$, then the left-hand side is given by

\[
\Psi_f \left( \frac{1}{1-zX} \right) = \Psi_f \left( 1 + zX + z^2X^2 + z^3X^3 + \ldots \right) = 1 + \sum_{n \geq 1} \sum_{(i_1, \ldots, i_m) \in \mathbb{C}(n)} c_{i_1} \cdots c_{i_m} I[z^n]X^n
\]

\[
= 1 + \sum_{n \geq 1} \sum_{i_1=1}^{\infty} c_{i_1} z \circ \cdots \circ z X^{i_1} \sum_{l=(i_2, \ldots, i_m) \in \mathbb{C}(n-i_1)} c_{i_2} \cdots c_{i_m} I[z^{n-i_1}]X^{n-i_1},
\]

Version 3.9 (June 26, 2020)
where the last sum on the right needs to be interpreted as 1 in the case $n = i_1$ and 0 if $i_1 > n$. Also notice that $I$ has been extended linearly to $k\langle A \rangle$ and it acts on $k\langle A \rangle[[X]]$ component wise. Now recall that $f_0(zX) = \sum_{i=1}^{\infty} c_i z \circ \cdots \circ z X^i$. With this the above equation gives

$$\Psi_f \left( \frac{1}{1 - zX} \right) = 1 + f_0(zX) \Psi_f \left( \frac{1}{1 - zX} \right),$$

from which the statement follows.

We can use this Proposition together with Theorem 2.28 to prove Proposition 2.21.

**Proof of Proposition 2.21.** First notice that for $f(T) = \exp(T) - 1$ the left-hand side of Proposition 2.29 is given by

$$\Psi_f \left( \frac{1}{1 - zX} \right) = \exp \left( \frac{1}{1 - zX} \right).$$

On the other hand we have

$$\frac{1}{1 - zX} = 1 + zX + z^2 X^2 + z^3 X^3 + \ldots$$

$$= 1 + zX + \frac{1}{2!} (z \shuffle z) X^2 + \frac{1}{3!} (z \shuffle z \shuffle z) X^3 + \cdots = \exp_{\shuffle}(zX).$$

By Theorem 2.28 $\exp : (k\langle A \rangle, \shuffle) \rightarrow (k\langle A \rangle, \ast_\circ)$ is an algebra homomorphism and we get

$$\exp \left( \frac{1}{1 - zX} \right) = \exp(\exp_{\shuffle}(zX)) = \exp_{\ast_\circ}(zX),$$

which gives for any $z \in kA[[X]]$ by Proposition 2.29

$$\exp_{\ast_\circ}(zX) = \frac{1}{1 - \exp_{\circ}(zX)}.$$

Since this holds for any $z \in kA[[X]]$, the $zX$ can be replaced by any power series in $XkA[[X]]$, i.e. in particular we can choose $\log_{\circ}(1 + zX) \in XkA[[X]]$ for any $z \in kA[[X]]$, to get

$$\exp_{\ast_\circ}(\log_{\circ}(1 + zX)) = \frac{1}{1 - \exp_{\circ}(\log_{\circ}(1 + zX))} = \frac{1}{1 - zX},$$

which is exactly the statement of Proposition 2.21.

### 2.3.6 Subalgebras of words with restricted first and last letters.

Now we will present a general statement for quasi-shuffle algebras, which we will use to regularize multiple zeta values in the next section. Since multiple zeta values $\zeta(k_1, \ldots, k_r)$ are just defined for indices with $k_1 \geq 2$, the map $\zeta$ was just defined on $\mathcal{H}^0$. The subspace $\mathcal{H}^0$ is spanned by words starting in $x$ and ending in $y$, or, when viewed as words in $A_z$, spanned by words not starting in the letter $z_1$. We want to extend this map to all indices, i.e. to the space $\mathcal{H}^1$. For example to make sense of $\zeta$ for the element $z_1 z_2$ one first notices that

$$z_2 \ast z_1 = z_2 z_1 + z_1 z_2 + z_3,$$

- 35 -
i.e. \( z_1 z_2 = z_2^* z_1 z_2 z_1 - z_3 \) is a polynomial in \( z_1 \) (with respect to \( * \)) with coefficients in \( \mathcal{F}_0 \). We will then view \( z_1 \) as a variable \( T \) and define the stuffle regularized multiple zeta value as the polynomial 
\[
\zeta^*(1,2;T) = \zeta(2)T - \zeta(2,1) - \zeta(3),
\]
which will give us an algebra homomorphism \( \zeta^* : \mathcal{F}_1 \rightarrow \mathbb{Z}[T] \). This we will do in the next section after proving a general statement for quasi-shuffle algebras in the following, which assures that a polynomial representation as above is possible in certain cases.

For subsets \( S, E \subset A \) we define the following subspace of our quasi-shuffle algebra \( Q = k\langle A \rangle \)
\[
Q^E_S = Q + \langle a_1 a_2 \ldots a_n \mid a_1 \in S, a_2, \ldots, a_{n-1} \in A, a_n \in E, n \geq 1 \rangle_Q
\]
\[
= Q + SQE,
\]
i.e. this is the subspace of \( k\langle A \rangle \) of words starting with letters in \( S \) and ending with letters in in \( E \). In particular we have \( Q^A_A = Q \) and we omit writing \( S \) or \( E \) if they equal \( A \), i.e. \( Q^A_S = Q^A_A, Q^E_S = Q^E_A \).

**Proposition 2.30.** If \( kS, \circ \) and \( kE, \circ \) are subalgebras of \( kA, \circ \), then \( (Q^E_S, \circ) \) is a subalgebra of \( (Q, \circ) \).

**Proof.** This is again a direct consequence of the definition of the quasi-shuffle product
\[
aw \circ bv = a(w \circ bv) + b(aw \circ v) + (a \circ b)(w \circ v)
\]
since the first letters (resp. last letters) of the elements in this product just come from the first letters (resp. last letters) and their \( \circ \) products.

**Theorem 2.31.** Assume \( kS, \circ \), \( kE, \circ \) are subalgebras of \( kA, \circ \) and we have an \( a \in A \), such that \( A \circ A \{a\} \subset k\langle A \{a\} \rangle \).

i) If \( a \in E \) we have \( Q^E = Q^E_{A\{a\}}[a] \) and the the map
\[
\text{pol}_a : Q^E \rightarrow Q^E_{A\{a\}}[T]
\]
\[
a \mapsto T
\]
is an isomorphism of \( k \)-algebras.

ii) If \( a \in S \) we have \( Q_S = Q^A_{A\{a\}}[a] \) and the the map
\[
\text{pol}^a : Q_S \rightarrow Q^A_{A\{a\}}[T]
\]
\[
a \mapsto T
\]
is an isomorphism of \( k \)-algebras.

**Proof.** For i) we first we want to show that \( Q^E = Q^E_{A\{a\}}[a] \), so we need to show that any word \( w \in Q^E \) is a polynomial in \( a \) (with respect to the product \( \circ \)) with coefficients given by linear combination of words not starting in \( a \). We write \( w = a^m v \) for \( m \geq 0 \) and \( v = b_1 \cdots b_l \in Q^E_{A\{a\}} \) and prove the statement by induction on \( m \), where the \( m = 0 \) case is clear since \( v = v \in Q^E_{A\{a\}} \). By the definition of the quasi-shuffle product we obtain
\[
a \circ a^{m-1} v = ma^m v + a^{m-1} b_1 (a \cup b_2 \cdots b_l) + \sum_{j=0}^{m-2} a^j (a \circ a) a^{m-2-j} v + a^{m-1} \sum_{i=1}^l b_1 \cdots (a \circ b_i) \cdots b_l.
\]

\[\text{In other words, the polynomial } \sum_{j=0}^{m} w_j \circ a^{* \circ j} \text{ with } w_j \in Q^E_{A\{a\}} \text{ gets send to } \sum_{j=0}^{m} w_j T^j.\]
The last three terms are linear combinations of words starting with $a^j$ with $0 \leq j < m$. Here we used for the last sum the condition $a \circ b_1 \in kA \setminus \{a\}$. Also all words end with letters in $E$, since $a, b_1 \in E$ and therefore $a \circ b_1 \in kE$. Therefore by the induction hypothesis we get that $w = a^m v$ is also an element in $Q_E^{\{a\}}[a]$. To show that the given maps are isomorphism we therefore just need to show that the representation as such a polynomial is unique, i.e. the given maps are injective. But this follows from the fact that there are no linear relations among the elements in $Q_E^{\{a\}}$. Since assuming that $0 = \sum_{j=0}^m w_j \circ a^m$ with $w_j \in Q_E^{\{a\}}$ we immediately get that $w_m = 0$ since it is the only part which gives words starting with $a^m$. We omit the proof of ii), since the argument for ii) is exactly the same, except that we consider words ending in $a^m$ instead of starting in $a^m$.  

Theorem 2.31 can be used the prove the following statements, for the spaces $\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}$ and the products $\bigcup$, $\ast$ and $\ast_\circ$. The first two statements are classical results and the last one can be found in [BK1, Theorem 2.14].

**Corollary 2.32.** We have

1. $\mathcal{H}_{\bigcup} = \mathcal{H}_0[y]$ and $\mathcal{H}_{\bigcup} = \mathcal{H}_0[x] = \mathcal{H}_{\bigcup} [x, y]$.
2. $\mathcal{H}_1 = \mathcal{H}_0[z_1]$.
3. $\mathcal{H}_1 = \mathcal{H}_0[z_1]$. 

**Proof.** In all cases we have $k = \mathbb{Q}$. For i) we choose $A = \{x, y\}$, $S = \{x\}$, $E = \{y\}$ and use the trivial product for $\circ$. With this we have $\mathcal{H}_0 = \mathcal{H}_0 = \mathcal{H}_0$. The statement then follows from Theorem 2.31 (i) and ii). For ii) and iii) we choose $A = A_1, a = z_1$ and the usual shuffle product for i) and the quasi-shuffle product $\ast$ for iii). The latter one was defined by using the following product $\circ$ on $kA_1$

$$z_{k_1} \ast z_{k_2} = z_{k_1+k_2} + \sum_{j=1}^{k_1+k_2-1} \binom{\lambda_{k_1,k_2}^j + \lambda_{k_2,k_1}^j}{k_1,k_2} z_j$$

where the rational numbers $\lambda_{k_1,k_2}^j$ are given by

$$\lambda_{k_1,k_2}^j = (-1)^{k_2-1} \binom{k_1+k_2-1-j}{k_1-j} \frac{B_{k_1+k_2-j}}{(k_1+k_2-j)!}.$$

By the definition of the $\lambda_{k_1,k_2}^j$ one checks that $\lambda_{k_1,k_2}^j + \lambda_{k_2,k_1}^j = 0$ whenever $k_1 + k_2 \geq 2$ (By using the equation (2.10) in the proof of Proposition 2.20). This shows that $z_1 \circ A \setminus \{z_1\} \subset QA \setminus \{z_1\}$ and we can apply Theorem 2.31.

**Example 2.33.** As an example of Corollary (2.4), i) ii) and iii) we give the following expressions of $z_1 z_2$ as a polynomial in $z_1 = y$ having coefficients in $\mathcal{H}_0$ with respect to the products $\bigcup, \ast$ and $\ast_\circ$

$$z_1 z_2 = \frac{1}{2} \mathcal{H}_0 \bigcup z_1 z_2 = 2 z_2 z_1 \bigcup z_1 + 3 z_2 z_1 z_1, \quad z_1 z_2 = \frac{1}{2} \ast_\circ z_1 z_2 = (z_2 z_1 + z_3) \ast z_1 + \left(2 z_2 z_1 z_1 + z_3 z_1 + \frac{1}{2} z_4\right),$$

$$z_1 z_2 = \frac{1}{2} z_1 \ast_\circ z_1 z_2 = (z_2 z_1 + z_3) \ast z_1 + \left(z_2 z_1 z_1 + z_3 z_2 + \frac{1}{2} z_4 - z_3 - \frac{3}{2} z_2 z_1 + \frac{5}{12} z_2\right).$$
Notice that the product \( \sqcup \) here is with respect to the alphabet \( A = \{x, y\} \) and not \( A_z \)! So for the first statement one should rewrite \( z_1 z_2 z_3 = y y x y \), \( z_2 z_1 = x y y \), etc. If you want to create more examples and play around with \( \sqcup, \ast \) and \( \hat{\ast} \) you can use the following online tool: [https://www.henrikbachmann.com/shuffle.html](https://www.henrikbachmann.com/shuffle.html), where these three products are implemented on the space \( A_z \). There one could check the above examples by entering (sh = \( \sqcup \), st = \( \ast \), gst = \( \hat{\ast} \))

\[
\begin{align*}
\frac{1}{2} [1] \text{st} [1] \text{st} [2] - [1] \text{st} ([2, 1] + [3]) + [2, 1, 1] + [3, 1] + 1/2 [4] \\
\end{align*}
\]

which all give the output \((1, 1, 2)\).

### 2.4 Regularizations

As mentioned already at the beginning of the last section we now want to make sense of multiple zeta values for non-admissible indices by using the previous results on quasi-shuffle algebras.

#### 2.4.1 Stuffle and shuffle regularized multiple zeta values

As a consequence of Theorem 2.31 (Corollary 2.4.3) we have for \( \bullet \in \{\sqcup, \ast\} \) isomorphism of \( \mathbb{Q} \)-algebras

\[
\text{reg}_T: \mathcal{F}_1 \to \mathcal{F}_0[T],
\]

which send an element \( w = \sum_{j=0}^{m} w_j \bullet z_1^m \) with \( w_j \in \mathcal{F}_0 \) to \( \text{reg}_T(w) = \sum_{j=0}^{m} w_j T^m \). This enables us to extend the algebra homomorphism \( \zeta: \mathcal{F}_0 \to \mathbb{Z} \) to an algebra homomorphism \( \zeta^\bullet: \mathcal{F}_1 \to \mathbb{Z}[T] \) by extending \( \zeta \) to \( \mathcal{F}_0[T] \) and setting \( \zeta^\bullet = \zeta \circ \text{reg}_T \), i.e. we have the following commutative diagram of \( \mathbb{Q} \)-algebra homomorphisms

\[
\begin{array}{ccc}
\mathcal{F}_1 & \xrightarrow{\text{reg}_T} & \mathcal{F}_0[T] \\
\zeta^\bullet \downarrow & & \downarrow \zeta \\
\mathbb{Z}[T] & & \mathbb{Z}[T]
\end{array}
\]

**Definition 2.34.** Let \( k = (k_1, \ldots, k_r) \in \mathbb{Z}_{\geq 1}^r \) be any index.

i) We define the **shuffle regularized multiple zeta value**

\[
\zeta^\sqcup(k; T) = \zeta^\sqcup(k_1, \ldots, k_r; T) := \zeta^\sqcup(z_k) \in \mathbb{Z}[T]
\]

In the case \( T = 0 \) we just write \( \zeta^\sqcup(k) = \zeta^\sqcup(k; 0) \).

ii) We define the **shuffle regularized multiple zeta value**

\[
\zeta^\ast(k; T) = \zeta^\ast(k_1, \ldots, k_r; T) := \zeta^\ast(z_k) \in \mathbb{Z}[T]
\]

In the case \( T = 0 \) we just write \( \zeta^\ast(k) = \zeta^\ast(k; 0) \).

From Example 2.33 we obtain

\[
\begin{align*}
\zeta^\sqcup(1, 1, 2; T) &= \frac{1}{2} \zeta(2) T^2 - 2 \zeta(2, 1) T + 3 \zeta(2, 1, 1), \\
\zeta^\ast(1, 1, 2; T) &= \frac{1}{2} \zeta(2) T^2 - (\zeta(2, 1) + \zeta(3)) T + \zeta(2, 1, 1) + \zeta(3, 1) + \frac{1}{2} \zeta(4).
\end{align*}
\]

*In the literature often “shuffle/shuffle regularized multiple zeta value” refers to the \( T = 0 \) case.*
Even though the coefficients of $T^2$ and $T$ are the same (because we know $\zeta(2, 1) = \zeta(3)$), the constant terms differ. In general $\zeta^{(i)}(k; T)$ and $\zeta^{(i)}(k; T)$ are different if $k$ is non-admissible. But we will see in the next section that there is an explicit relationship between these polynomials. Further, we will see in Section 3 that we have for $w \in S^0, v \in S^1$ and $\bullet \in \{\shuffle, *\}$ the extended double shuffle relations

$$
\zeta^*(w \shuffle v - w * v) = 0.
$$

Since $\zeta^{(i)}$ and $\zeta^*$ differ on $S^1$ this is not obvious at all. For example, we have (Exercise 9)

$$
\zeta^*(z_2 \shuffle z_1 z_1 - z_2 * z_1 z_1) = 0,
$$

which implies linear relations among multiple zeta values in weight four.

### 2.4.2 Comparison of $\zeta^{(i)}$ and $\zeta^*$

As we saw before, the two regularization $\zeta^{(i)}(k; T)$ and $\zeta^*(k; T)$ differ as elements in $R[T]$. In this section, we will present the exact relationship between these two regularizations as it was done in [IKZ]. For this first consider the following series

$$
A(u) = \exp \left( \sum_{n=2}^{\infty} \frac{(-1)^n}{n} \zeta(n) u^n \right)
= 1 + \frac{\zeta(2)}{2} u^2 - \frac{\zeta(3)}{3} u^3 + \left( \frac{\zeta(4)}{4} + \frac{\zeta(2)^2}{8} \right) u^4 - \left( \frac{\zeta(5)}{5} + \frac{\zeta(2)\zeta(3)}{6} \right) u^5 + \ldots
=: \sum_{k \geq 0} \gamma_k u^k.
$$

Here the $\gamma_k \in \mathbb{Q}[\zeta(j) \mid j \geq 2]$ are polynomials of single zeta values which, considered as multiple zeta values, have homogeneous weight $k$. Using this we define the $R$-linear map $\rho : R[T] \to R[T]$ by

$$
\rho(e^{Tu}) := A(u)e^{Tu}.
$$

Notice that this defines the linear map $\rho$ uniquely by comparing the coefficients of $u^m$ on both sides. Since $\rho$ is linear, we get

$$
\rho(e^{Tu}) = \rho(1) + \rho(T)u + \frac{1}{2!} \rho(T^2)u^2 + \frac{1}{3!} \rho(T^3)u^3 + \ldots
=: \left( 1 + \gamma_1 u + \gamma_2 u^2 + \ldots \right) \left( 1 + \frac{1}{2!} u^2 + \frac{1}{3!} u^3 + \ldots \right) = A(u)e^{Tu},
$$

and therefore we obtain for $m \geq 0$ the explicit formula

$$
\rho(T^m) = m! \sum_{k=0}^{m} \gamma_k \frac{T^{m-k}}{(m-k)!}.
$$

Version 3.9 (June 26, 2020)
Also notice that $\rho$ is bijective. For the first values of $m$ we get

$$
\begin{align*}
\rho(1) &= 1, \\
\rho(T) &= T, \\
\rho(T^2) &= T^2 + \zeta(2), \\
\rho(T^3) &= T^3 + 3\zeta(2)T - 2\zeta(3), \\
\rho(T^4) &= T^4 + 6\zeta(2)T^2 - 8\zeta(3)T + 6\zeta(4) + 3\zeta(2)^2, \\
\rho(T^5) &= T^5 + 10\zeta(2)T^3 - 20\zeta(3)T^2 + (30\zeta(4) + 15\zeta(2)^2)T - 24\zeta(5) - 20\zeta(2)\zeta(3).
\end{align*}
$$

The stuffle regularized multiple zeta values are elements in $\mathbb{R}[T]$ and for example we saw that

$$
\zeta^*(1, 1, 2; T) = \frac{1}{2}\zeta(2)T^2 - (\zeta(2, 1) + \zeta(3))T + \zeta(2, 1, 1) + \zeta(3, 1) + \frac{1}{2}\zeta(4).
$$

Applying the linear map $\rho$ to this, we see that we just get an additional contribution of $\frac{1}{2}\zeta(2)^2$, i.e.

$$
\rho(\zeta^*(1, 1, 2; T)) = \frac{1}{2}\zeta(2)T^2 - (\zeta(2, 1) + \zeta(3))T + \zeta(2, 1, 1) + \zeta(3, 1) + \frac{1}{2}\zeta(4) + \frac{1}{2}\zeta(2)^2.
$$

Using the known relations $\zeta(3) = \zeta(2, 1)$ and $\zeta(4) = \zeta(2, 1, 1)$ (duality), $\zeta(3, 1) = \frac{1}{4}\zeta(4)$ (finite double shuffle) and $\zeta(2)^2 = \frac{5}{2}\zeta(4)$ (Euler), we get

$$
\rho(\zeta^*(1, 1, 2; T)) = \frac{1}{2}\zeta(2)T^2 - 2\zeta(2, 1)T + 3\zeta(2, 1, 1)
\quad = \zeta^{\mu}(1, 1, 2; T),
$$

i.e. the $\rho$ sends the stuffle regularized multiple zeta value to the shuffle regularized multiple zeta value.

In general the map $\rho$ has this property and we have the following

**Theorem 2.35.** For all $k \in \mathbb{Z}_{\geq 1}$, we have

$$
\zeta^{\mu}(k; T) = \rho(\zeta^*(k; T)).
$$

Or equivalently, when viewed as maps from $\mathcal{H}^1$ to $\mathbb{R}[T]$, we have $\zeta^{\mu} = \rho \circ \zeta^*$.

**Proof.** This is [IKZ] Theorem 1 and we just give a sketch of the proof here. A really detailed version of this proof can also be found in the book of Zhao [Zh1, Section 3.3.2]. The main idea is to compare the behavior of the truncated multiple zeta values $\zeta_M(k)$ (which satisfy the stuffle product formula for all $k$) and the multiple polylogarithm $\text{Li}_k(z)$ (which satisfy the shuffle product formula for all $k$) as $M \to \infty$ and $z \to 1$. We have the classical formula

$$
\zeta_M(1) = 1 + \frac{1}{2} + \cdots + \frac{1}{M-1} = \log(M) + \gamma + O\left(\frac{1}{M}\right),
$$

as $M \to \infty$, where $\gamma = 0.57721\ldots$ denotes the Euler–Mascheroni constant. As a consequence of the stuffle product formula one can show by induction that for some $J$ we have (see [Zh1, Lemma 3.3.19])

$$
\zeta_M(k) = \zeta^*(k; \log(M) + \gamma) + O(M^{-1}\log^J(M)) \quad (as \ M \to \infty).
$$

(2.19)
Similiarly by \( \text{Li}_1(z) = \log \left( \frac{1}{1-z} \right) \) and using the shuffle product formula for \( \text{Li} \) one can show, together with the fact that for admissible \( k \) we have \( \text{Li}_k(1) - \text{Li}_k(z) = O(1-z) \), that (see [Zh1] Lemma 3.3.20)

\[
\text{Li}_k(z) = \zeta^{\mu} \left( k, \log \left( \frac{1}{1-z} \right) \right) + O \left( (1-z) \log^J \left( \frac{1}{1-z} \right) \right) \quad \text{(as } z \to 1) .
\]

The connection of the multiple polylogarithm to the truncated multiple zeta values is, that these basically give the taylor coefficients of \( \text{Li}_k \):

\[
\text{Li}_k(z) = \text{Li}_{k_1, \ldots, k_r}(z) = \sum_{m_1 > \cdots > m_r > 0} \frac{z^{m_1}}{m_1^{k_1} \cdots m_r^{k_r}} = \sum_{m=1}^{\infty} \left( \sum_{m_{m_1} > \cdots > m_{m_r} > 0} \frac{1}{m_1^{k_1} m_2^{k_2} \cdots m_r^{k_r}} \right) z^m
\]

\[
= \sum_{m=1}^{\infty} \left( \zeta_{m+1}(k) - \zeta_m(k) \right) z^m = (1-z) \sum_{m=1}^{\infty} \zeta_m(k) z^{m-1} .
\]

The statement then follows from the following general fact, that for any polynomial \( P(T) \in \mathbb{R}[T] \) and \( Q(T) = \rho(P(T)) \) one has (see [IKZ, Lemma 1] or [Zh1, Lemma 3.3.17])

\[
(1-z) \sum_{m=1}^{\infty} P(m) z^m = Q \left( \log \left( \frac{1}{1-z} \right) \right) + O \left( (1-z) \log^J \left( \frac{1}{1-z} \right) \right) \quad \text{(as } z \to 1) .
\]

for \( J = \deg(P) - 1 \) and for \( l \geq 0 \)

\[
\sum_{m=1}^{\infty} \frac{\log^{l+1}(m)}{m} z^{m-1} = O \left( \log^{l+1} \left( \frac{1}{1-z} \right) \right) \quad \text{(as } z \to 1) .
\]

Choosing \( P(T) = \zeta^*(k; T) \) and combining all the equations above gives \( Q(T) = \zeta^{\mu}(k; T) \) and therefore \( \zeta^{\mu}(k; T) = \rho(\zeta^*(k; T)) \).

\[ \Box \]

2.4.3 The \( g \)-series \( g \) for admissible indices

We defined the \( q \)-series \( g(k) \) for any index \( k = (k_1, \ldots, k_r) \) by

\[
g(k) = g(k_1, \ldots, k_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{P_{k_1}(q^{m_1})}{(1-q^{m_1})^{k_1}} \cdots \frac{P_{k_r}(q^{m_r})}{(1-q^{m_r})^{k_r}}
\]

\[
= \sum_{m_1 > \cdots > m_r > 0} \frac{d_1^{k_1-1}}{(k_1-1)!} \cdots \frac{d_r^{k_r-1}}{(k_r-1)!} q^{m_1 d_1 + \cdots + m_r d_r} .
\]

Only for admissible indices we could consider the limit \( q \to 1 \) to obtain multiple zeta values, i.e for an admissible index \( k \) we have

\[
\lim_{q \to 1} (1-q)^{\text{wt}(k)} g(k) = \zeta(k) . \tag{2.20}
\]

Motivated by this, we define for \( k \geq 0 \) the following spaces

\[
\mathcal{G}_0^k = \left\{ g(k) \mid k \text{ admissible index} \right\}_Q ,
\]

\[
\mathcal{G}_{\leq k}^0 = \left\{ g(k) \mid k \text{ admissible index, } \text{wt}(k) \leq k \right\}_Q .
\]
With this we define for $k \geq 0$ the linear map $Z_k$

$$Z_k : G_{\leq k}^0 \rightarrow Z_k$$

$$f \mapsto \lim_{q \rightarrow 1} (1 - q)^k f. \quad (2.21)$$

By (2.20) this map is surjective. We now want to extend the map $Z_k$ to all $G$, i.e. to the space

$$G_{\leq k} = \{ g(k) \mid k \text{ index, wt}(k) \leq k \}_{\mathbb{Q}}.$$

Considered as a $\mathbb{Q}$-linear map $g : \mathcal{S}^1 \rightarrow G$, where $g(z_k) = g(k)$, we saw that this gives an algebra homomorphism from $H_1^\ast$ to $G$ and the above space is given by $G_0 = g(\mathcal{S}^0)$. By Corollary 2.4.3 we know that $H_1^\ast = H_0^\ast [z_1]$, which gives the following.

**Proposition 2.36.**

i) We have $G = G_0^0 [g(1)]$.

ii) $g(1)$ is algebraically independent over $G^0$.

**Proof.** Statement i) is a direct consequence of Corollary 2.4.3. For ii) the argument is that $g(1) = \sum_{m,d \geq 1} q^{dm} \approx - \log(1 - q)$ as $q \rightarrow 1$ and $g(k) \approx 1(1 - q)^{wt(k)}$ for admissible $k$. Here $f(q) \approx g(q)$ as $q \rightarrow 1$ means that there exist constants $A, B \in \mathbb{R}_{>0}$, such that $A \leq |f(q) - g(q)| \leq B$ as $q \rightarrow 1$. Both statements can also be found in [BK1, Theorem 2.14] (see also [P, Lemma 2]).

Using Proposition 2.36 we can extend for $k \geq 0$ the map $Z_k$ to $G_{\leq k}$. Since any $f \in G_{\leq k}$ can be written uniquely as $f = \sum_{j=0}^k f_j g(1)^{k-j}$, with $f_j \in G_{\leq j}^0$. With this we define

$$Z_k^T : G_{\leq k} \rightarrow Z[T]$$

$$\sum_{j=0}^k f_j g(1)^{k-j} \mapsto \sum_{j=0}^k Z_j(f_j) T^{k-j}$$

where $Z_j(f_j)$ for $f_j \in G_{\leq j}^0$ is defined by (2.21).

**Proposition 2.37.** For any index $k$ we have

$$Z_{\text{wt}(k)}^T (g(k)) = \zeta^\ast(k; T).$$

**Proof.** Let $w \in \mathcal{S}^1$. By Corollary we have $\mathcal{S}_1^1 = \mathcal{S}_1^0 [z_1]$ and $\mathcal{S}_1^1 = \mathcal{S}_2^0 [z_1]$ and therefore there exist $u_j, v_j \in \mathcal{S}_k$, such that

$$w = \sum_{j=0}^m u_j * z_1^{j} = \sum_{j=0}^m v_j * z_1^{j}.$$

Since we have

$$z_{k_1} \cdot z_{k_2} = z_{k_1+k_2} + \sum_{j=1}^{k_1+k_2-1} \left( \lambda_{j,k_1,k_2} + \lambda_{j,k_2,k_1} \right) z_j$$

we see, by following the proof of Theorem 2.31, that $u_j$ and $v_j$ just differ by linear combination of words $z_{k'}$ with indices satisfying $\text{wt}(k') < j$. Since $G_{\leq j-1} \subset \ker(Z_j)$ the statement follows.
§3 Families of linear relations and their $q$-relatives

In this section, we want to discuss several families of linear relations among multiple zeta values and also some of their $q$-analogues. Asking for linear relations among multiple zeta values is equivalent in asking for the kernel of the map $\zeta : \mathcal{H}^0 \to \mathbb{Z}$.

3.1 Extended double shuffle relations

As an extension of the finite double shuffle relations (Proposition 2.14) we will now present a family of linear relations, which give conjecturally all linear relations among multiple zeta values. We define for $w, u \in \mathcal{H}^1$ the element

$$ds(w, u) := w \shuffle u - w \ast u \in \mathcal{H}^1.$$  

The statement of Proposition 2.14 then was, that $ds(w, u) \in \ker \zeta$ if $w, u \in \mathcal{H}^0$. The extended version states, that one of the words $w$ and $u$ is allowed to be in $\mathcal{H}^1$. In this case $ds(w, u)$ is not necessary in $\mathcal{H}^0$ anymore, but after projecting to $\mathcal{H}^0$ by the map $\text{reg}_T : \mathcal{H}^1 \to \mathcal{H}^0[T]$ and then comparing the coefficients of $T$ (or setting $T = 0$, for which we write $\text{reg}_* := \text{reg}_0$) one still obtains a relation among multiple zeta values. In other words $ds(w, u)$ is in the kernel of the regularized multiple zeta value maps.

**Theorem 3.1** (Extended double shuffle relations). For $w \in \mathcal{H}^1$, $u \in \mathcal{H}^0$ and $\bullet \in \{\shuffle, \ast\}$ we have

$$\zeta^\bullet(w \shuffle u - w \ast u; T) = 0,$$

i.e. in particular $\text{reg}_* (ds(w, u)) \in \ker \zeta$.

**Proof.** By Theorem 2.35 we have for all $w \in \mathcal{H}^1$

$$\zeta^\shuffle(w; T) = \rho(\zeta^*(w; T)).$$

Multiplying both sides with $\zeta^\shuffle(u) = \zeta^*(u) = \zeta(u) \in \mathbb{R}$, for $u \in \mathcal{H}^0$, we get by the $\mathbb{R}$-linearity of $\rho$

$$\zeta^\shuffle(w \shuffle u; T) = \rho(\zeta^*(w \ast u; T)) = \zeta^\shuffle(w \ast u; T).$$

This gives $\zeta^\shuffle(w \shuffle u - w \ast u; T) = 0$ and $\zeta^*(w \shuffle u - w \ast u; T) = 0$ by applying the inverse of $\rho$.  

**Conjecture 3.2.** The kernel of $\zeta : \mathcal{H}^0 \to \mathbb{Z}$ is given by

$$\ker \zeta = \langle \text{reg}_\shuffle (w \shuffle u - w \ast u) \mid w \in \mathcal{H}^1, u \in \mathcal{H}^0 \rangle_Q = \langle \text{reg}_* (w \shuffle u - w \ast u) \mid w \in \mathcal{H}^1, u \in \mathcal{H}^0 \rangle_Q,$$

i.e. the extended double shuffle relations give all $\mathbb{Q}$-linear relations among multiple zeta values.

Since it is expected that the extended double shuffle relations give all relations among multiple zeta values, one could obtain upper bounds for the dimension of $\mathbb{Z}_k$, i.e., an alternative proof of Theorem 1.15 by counting these relations. This is still an open problem.
Open problem 3.3. Count the number of linearly independent extended double shuffle relations. Define for \( k \geq 0 \) and \( \bullet \in \{ \text{reg}, * \} \) the spaces
\[
eds_{\bullet}^k = \langle \text{reg}_{\bullet}^* (w \sqcup \text{reg} u - w * u) \mid w \in \mathcal{F}^1, u \in \mathcal{F}^0, \text{wt}(w) + \text{wt}(u) = k \rangle_{\mathbb{Q}}.
\]
Show that for all \( k \geq 2 \) and any \( \bullet \in \{ \text{reg}, * \} \) we have
\[
\dim_{\mathbb{Q}} \edss_{\bullet}^k = 2^{k-2} - d_k.
\] (3.1)
Here the \( d_k \) are the conjectured dimensions (Conjecture 1.13) of \( \mathcal{Z}_k \), defined by
\[
\sum_{k \geq 0} d_k X^k = \frac{1}{1 - X^2 - X^3},
\]
and \( 2^{k-2} \) is the number of admissible indices of weight \( k \).

So far the equation (3.1) has been checked up to \( k = 21 \) (T. Machide, T. Sonobe, 2020+) by extensive computer calculations. All the relations we obtained so far should be a consequence of the extended double shuffle relations. For the finite double shuffle relations this is obvious, but for the duality relation this is actually also an open problem. Recall that the duality relations (Proposition 2.5) stated that for all \( v \in \mathcal{F}^0 \) we have
\[
\zeta(\tau(v)) = \zeta(v),
\]
where \( \tau \) was the anti-automorphism defined on \( \mathcal{F}^0 \) with \( \tau(y) = x \) and \( \tau(y) = x \).

Open problem 3.4. Show that for any \( v \in \mathcal{F}^0 \) we have for \( \bullet \in \{ \text{reg}, * \} \)
\[
\tau(v) - v \in \langle \text{reg}_{\bullet}^* (w \sqcup \text{reg} u - w * u) \mid w \in \mathcal{F}^1, u \in \mathcal{F}^0 \rangle_{\mathbb{Q}},
\]
i.e. the duality relation is a consequence of the extended double shuffle relations.

We now want to discuss a refinement of the extended double shuffle relations. For this, we first consider the following special case.

Proposition 3.5 (Hoffman’s relation ([H1])). For an admissible index \( k = (k_1, \ldots, k_r) \) we have
\[
\sum_{i=1}^{r} \zeta(k_1, \ldots, k_{i-1}, k_i + 1, k_{i+1}, \ldots, k_r) = \sum_{1 \leq i \leq r} \sum_{j=0}^{k_{i+1}} \zeta(k_1, \ldots, k_{i-1}, k_i - j, j + 1, k_{i+1}, \ldots, k_r).
\]

Proof. This is a special case of the extended double shuffle relation by choosing \( w = z_1 = y \) and \( u = z_k \). In particular, we have \( ds(z_1, z_k) \in \mathcal{F}^0 \) (Exercise 10) and therefore no regularization is necessary. Another proof of this relation, using partial fraction decomposition, can be found in [Zn2] Theorem 1.

Hoffman’s relation is a special case of the extended double shuffle relations, which are not a consequence of the finite double shuffle relations. Surprisingly, it seems that this is the only part of the extended double shuffle relations we need.

Conjecture 3.6 (H. N. Minh, M. Petitot et al. ([MJOP])). We have
\[
\ker \zeta = \langle w \sqcup u - w * u \mid w \in \mathcal{F}^0 \cup \{ z_1 \}, u \in \mathcal{F}^0 \rangle_{\mathbb{Q}},
\]
i.e. Hoffman’s relation and the finite double shuffle relations give all linear relations among multiple zeta values.
But still, this set of relations seems to be too much, and the following gives an even better refinement.

**Conjecture 3.7** (M. Kaneko, M.Noro and K.Tsurumaki ([KNT])). We have
\[
\ker \zeta = \langle w \cup u - w \ast u \mid w \in \{z_1, z_2, z_3, z_2z_1\}, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}}.
\]

The equivalence of these conjectures is not known and in particular we have the the following open problem.

**Open problem 3.8.** Show that
\[
\langle \text{reg}_{\infty} (w \cup u - w \ast u) \mid w \in \mathcal{F}_1^1, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}} = \langle \text{reg}_{\infty} (w \cup u - w \ast u) \mid w \in \mathcal{F}_1^1, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}}
= \langle w \cup u - w \ast u \mid w \in \mathcal{F}_0^0 \cup \{z_1\}, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}}
= \langle w \cup u - w \ast u \mid w \in \{z_1, z_2, z_3, z_2z_1\}, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}}.
\]

Having an element \( w \in \ker \zeta \) and \( u \in \mathcal{F}_0^0 \), then clearly also \( w \ast u, w \cup u \in \ker \zeta \). As far as the author knows, it is also still an open problem to show that the space of extended double shuffle relations is closed under the multiplication with elements in \( \mathcal{F}_0^0 \).

**Open problem 3.9.** Show that for \( \bullet_1, \bullet_2 \in \{\cup, \ast\} \) we have
\[
\mathcal{F}_0^0 \bullet_1 \langle \text{reg}_{\bullet_2} (w \cup u - w \ast u) \mid w \in \mathcal{F}_1^1, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}} \subseteq \langle \text{reg}_{\bullet_2} (w \cup u - w \ast u) \mid w \in \mathcal{F}_1^1, u \in \mathcal{F}_0^0 \rangle_{\mathbb{Q}}.
\]

### 3.2 Seki-Yamamoto’s connected sums and Ohno’s relation

In this section, we want to present an alternative proof of the duality relation (Proposition 2.5), which was given by Seki and Yamamoto in [SY]6. In this nice work, they introduce the notion of connected sums, which recently also found their ways into various other proofs of families of relations among multiple zeta values and some their variants, such as finite multiple zeta values. In [S] you can find an overview of different applications of connected sums. We will use this setup to present a proof of Ohno’s relation for \( q \)-analogues of multiple zeta values as it was given in [SY].

#### 3.2.1 Connected sums and the duality relation for multiple zeta values

We start by reformulating the duality relations on the level of indices. Recall that \( \tau : \mathcal{F}_0^0 \to \mathcal{F}_0^0 \) was defined as the anti-automorphism (with respect to the usual multiplication in \( \mathbb{Q}(x, y) \)) satisfying \( \tau(x) = y \) and \( \tau(y) = x \). Now let \( k = (k_1, \ldots, k_s) \) be an admissible index, i.e. \( k_1 \geq 2 \) and \( z_k \in \mathcal{F}_0^0 \).

Then there exist numbers \( a_1, b_1, \ldots, a_s, b_s \geq 1 \), such that
\[
k = (a_1 + 1, \{1\}^{b_1-1}, a_2 + 1, \{1\}^{b_2-1}, \ldots, a_s + 1, \{1\}^{b_s-1}).
\]

With these numbers we define the admissible index \( k^\dagger \) by
\[
k^\dagger := (b_s + 1, \{1\}^{a_s-1}, b_{s-1} + 1, \{1\}^{a_{s-1}-1}, \ldots, b_1 + 1, \{1\}^{a_1-1}).
\]

One can then see by induction on \( s \), that we have (Exercise 11)
\[
z_{k^\dagger} = \tau(z_k),
\]

---

6In the work [SY] the order of summation in the definition of multiple zeta values is reversed. Therefore one needs to be careful when comparing the results here and the ones in their work.

Version 3.9 (June 26, 2020)
i.e. the duality relation of multiple zeta values can be stated as \( \zeta(\mathbf{k}) = \zeta(\mathbf{k}^\dagger) \) for all admissible \( \mathbf{k} \).

Now we will introduce the connected sum for multiple zeta values. The name comes from the fact that these sums look like the product of two multiple zeta values, which get connected at someplace by a connector.

**Definition 3.10.** Let \( \mathbf{k} = (k_1, \ldots, k_r) \), \( \mathbf{l} = (l_1, \ldots, l_s) \) be two non-empty indices. Then we define the connected sum \( Z(\mathbf{k}; \mathbf{l}) \) by

\[
Z(\mathbf{k}; \mathbf{l}) = Z(k_1, \ldots, k_r; l_1, \ldots, l_s) = \sum_{m_1 > m_2 > \cdots > m_r > 0 \atop n_1 > n_2 > \cdots > n_s > 0} \frac{m_1! n_1!}{(m_1 + n_1)!} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \frac{1}{n_1^{l_1} \cdots n_s^{l_s}} \tag{3.3}
\]

and set \( Z(\mathbf{k}; \emptyset) = Z(\emptyset; \mathbf{k}) = \zeta(\mathbf{k}) \) for admissible \( \mathbf{k} \).

One can check that the sum (3.3) converges for all non-empty indices \( \mathbf{k} \) and \( \mathbf{l} \). Also notice that (3.3) is essentially the product \( \zeta(\mathbf{k})\zeta(\mathbf{l}) \), which gets connected by the **connector** \( c(m_1, n_1) = \frac{m_1! n_1!}{(m_1 + n_1)!} \) at the beginning. The relationship to the duality relations comes from the fact that one can show that \( Z(\mathbf{k}; \emptyset) = \cdots = Z(\emptyset; \mathbf{k}^\dagger) \), by using the following transport relations.

**Proposition 3.11** (Transport relations). Let \( (k_1, \ldots, k_r) \) and \( (l_1, \ldots, l_s) \) be two indices. If \( s > 0 \) then we have

\[
Z(1, k_1, \ldots, k_r; l_1, \ldots, l_s) = Z(k_1, \ldots, k_r; l_1 + 1, l_2, \ldots, l_s)
\]

and if \( r > 0 \), then we have

\[
Z(k_1 + 1, k_2, \ldots, k_r; l_1, \ldots, l_s) = Z(k_1, \ldots, k_r; 1, l_1, l_2, \ldots, l_s).
\]

**Proof.** To prove the first equality we use for \( m \geq 0 \) the telescoping sum

\[
\sum_{a = m + 1}^{\infty} \frac{1}{a} \frac{a! n!}{(a + n)!} = \frac{1}{n} \sum_{a = m + 1}^{\infty} \left( \frac{(a - 1)! n!}{(a - 1 + n)!} - \frac{a! n!}{(a + n)!} \right) = \frac{1}{n} \frac{m! n!}{(m + n)!}, \tag{3.4}
\]

from which we obtain

\[
Z(1, k_1, \ldots, k_r; l_1, \ldots, l_s) = \sum_{a > m_1 > m_2 > \cdots > m_r > 0 \atop n_1 > n_2 > \cdots > n_s > 0} \frac{1}{a} \frac{a! n_1!}{(a + n_1)!} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \frac{1}{n_1^{l_1} \cdots n_s^{l_s}} = \frac{1}{m_1 (m_1 + n_1)!} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \frac{1}{n_1^{l_1} \cdots n_s^{l_s}} = Z(k_1, \ldots, k_r; l_1 + 1, l_2, \ldots, l_s).
\]

Also notice that we obtain \( Z(1; l_1, \ldots, l_s) = Z(\emptyset; l_1 + 1, l_2, \ldots, l_s) = \zeta(l_1 + 1, l_2, \ldots, l_s) \) here in the case \( r = 0 \) by using the \( m = 0 \) case of (3.4). The second statement follows from the symmetry \( Z(\mathbf{k}; \mathbf{l}) = Z(\mathbf{l}; \mathbf{k}) \).

Using these transport relations, we can therefore always transport one index from the left to the right and vice-verse. For example, we have

\[
\zeta(3) = Z(3; \emptyset) = Z(2; 1) = Z(1; 1, 1) = Z(\emptyset; 2, 1) = \zeta(2, 1).
\]
In general the duality relations follows from this, since if we start with an admissible index \( k = (a_1 + 1, \{1\}^{b_1-1}, a_2 + 1, \{1\}^{b_2-1}, \ldots, a_s + 1, \{1\}^{b_s-1}) \), we can use \( a_1 \)-times the second transport relations and \( b_1 \)-times the first, etc., to get
\[
\zeta(k) = Z(a_1 + 1, \{1\}^{b_1-1}, a_2 + 1, \{1\}^{b_2-1}, \ldots, a_s + 1, \{1\}^{b_s-1}; 0) = \ldots
\]
\[
= Z((1)^{b_1-1}, a_2 + 1, \{1\}^{b_2-1}, \ldots, a_s + 1, \{1\}^{b_s-1}; 2, \{1\}^{a_1-1}) = \ldots
\]
\[
= Z(a_1 + 1, \{1\}^{b_1-1}, \ldots, a_s + 1, \{1\}^{b_s-1}; b_1 + 1, \{1\}^{a_1-1}) = \ldots
\]
\[
= \ldots
\]
\[
= Z(\emptyset; b_s + 1, \{1\}^{a_s-1}, b_{s-1} + 1, \{1\}^{a_{s-1}-1}, \ldots, b_1 + 1, \{1\}^{a_1-1}) = Z(\emptyset; k^\dagger) = \zeta(k^\dagger).
\]

### 3.2.2 The sum formula and Ohno’s relation

We will now present a family of linear relations, which are known as Ohno’s relation. These types of relations generalize the duality relation and the so-called sum formula, given by the following.

**Theorem 3.12** (Sum formula (Granville [G], Zagier)). For all \( k \geq 2 \) and \( 1 \leq r < k \) we have
\[
\sum_{\substack{k_1 + \ldots + k_r = k \\
k_1 \geq 2, k_2, \ldots, k_r \geq 1}} \zeta(k_1, \ldots, k_r) = \zeta(k). \tag{3.6}
\]

The sum formula therefore states that the sum over all multiple zeta values of weight \( k \) in any fixed depths always gives the Riemann zeta value \( \zeta(k) \). Notice again, that our first relation \( \zeta(2, 1) = \zeta(3) \) is also the first (non-trivial) example of the sum-formula. There a various generalization of the sum formula, which for example, also include weights and therefore are called weighted sum formulas.

To state Ohno’s relation we first define for an admissible index \( k = (k_1, \ldots, k_r) \) the **Ohno sum** by
\[
O^X(k) = \sum_{c \geq 0} O(k; c) X^c \in \mathbb{Z}[X],
\]
where we write for \( c \geq 0 \)
\[
O(k; c) = \sum_{\substack{c_1 + \ldots + c_r = c \\
c_1, \ldots, c_r \geq 0}} \zeta(k_1 + c_1, \ldots, k_r + c_r) \in \mathbb{Z}^{\text{wt}(k) + c}.
\]

Notice that for \( k = (2, \{1\}^{r-1}) \) and \( c = k - r - 1 \) we obtain the left-hand side of (3.6)
\[
O(2, \{1\}^{r-1}; k - r - 1) = \sum_{\substack{k_1 + \ldots + k_r = k \\
k_1 \geq 2, k_2, \ldots, k_r \geq 1}} \zeta(k_1, \ldots, k_r).
\]

Ohno’s relation now states, that the Ohno sum also satisfies the duality relation.

**Theorem 3.13** (Ohno’s relation (Ohno [Oh])). For any admissible index \( k \) we have
\[
O^X(k) = O^X(k^\dagger).
\]

Since \( O^0(k) = O(k; 0) = \zeta(k) \) we obtain the duality as a special case by considering the constant term in Ohno’s relation. Choosing for \( r \geq 1 \) the index \( k = (2, \{1\}^{r-1}) \) we have \( k^\dagger = (r + 1) \), and therefore the sum formula (3.6) follows by considering the coefficient of \( X^{k-r-1} \) in \( O^X(2, 1, \ldots, 1) = O^X(r + 1) \). The formulation of the Ohno relation in the original work of Ohno is a bit different and we use here the formulation of [HMOS], where the authors also prove further relations of the Ohno sums besides the duality relation.
Example 3.14. For $k = (2, 2, 1)$ we have $k^\dagger = (3, 2)$ and

$$
O^X_{(2, 2, 1)} = \zeta(2, 2, 1) + (\zeta(3, 2, 1) + \zeta(2, 3, 1) + \zeta(2, 2, 2))X
+ (\zeta(4, 2, 1) + \zeta(2, 4, 1) + \zeta(2, 2, 3) + \zeta(3, 3, 1) + \zeta(3, 2, 2) + \zeta(2, 3, 3))X^2 + \ldots,
$$

$$
O^X_{(3, 2)} = \zeta(3, 2) + (\zeta(4, 2) + \zeta(3, 3))X + (\zeta(5, 2) + \zeta(3, 4) + \zeta(4, 3))X^2 + \ldots
$$

which gives linear relation among multiple zeta values of weight $5 + c$ by comparing the coefficients of $X^c$ in $O^X_{(2, 2, 1)} = O^X_{(3, 2)}$.

The number of Ohno’s relations is given by the following table (calculated by Tanaka in [Tan2]).

<table>
<thead>
<tr>
<th>weight $k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td># all conjectured relations</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>29</td>
<td>60</td>
<td>123</td>
<td>249</td>
<td>503</td>
<td>1012</td>
</tr>
<tr>
<td># Ohno’s relations</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>23</td>
<td>46</td>
<td>98</td>
<td>199</td>
<td>411</td>
<td>830</td>
</tr>
</tbody>
</table>

We now want to state the $q$-analogue version of Ohno’s relation. Recall that we defined for an admissible index $k = (k_1, \ldots, k_r)$ the Bradley-Zhao $q$-analogues of multiple zeta values by

$$
\zeta_{BZ}^q(k) = \sum_{m_1 > \ldots > m_r > 0} q^{(k_1-1)m_1} \ldots q^{(k_r-1)m_r} \frac{1}{[m_1]_{q}^{k_1} \ldots [m_r]_{q}^{k_r}},
$$

(3.7)

where $[m]_q = \frac{1-q^m}{1-q}$ denotes the $q$-integer. It was first shown by Bradley ([Bra]) that these $q$-series also satisfy Ohno’s relation. Define for an admissible index $k = (k_1, \ldots, k_r)$ the $q$-Ohno sum by

$$
O^X_q(k) = \sum_{c \geq 0} O_q(k; c)X^c \in \mathbb{Q}[[q]][[X]],
$$

where we write for $c \geq 0$

$$
O_q(k; c) = \sum_{c_1 + \ldots + c_r = c} \zeta_{BZ}^q(k_1 + c_1, \ldots, k_r + c_r).
$$

The series $O^X_q(k)$ is a formal power-series in $X$ with coefficients given by formal power series in $q$. For real $0 < q, X < 1$ one can show that $O^X_q(k)$ converges and gives a well-defined real number.

Theorem 3.15 (Ohno’s relation for $q$-MZV ([Bra])). For any admissible index $k$ we have

$$
O^X_q(k) = O^X_q(k^\dagger).
$$

Notice that Theorem 3.15 implies that the $q$-analogues $\zeta_{BZ}^q$ also satisfy the sum-formula and the duality relation. Also sending $q \to 1$ we obtain Ohno’s relation for multiple zeta values.

The goal is now to give a proof of Theorem 3.15 by using Seki-Yamamoto’s concept of connected sums.
For this we first rewrite the $q$-Ohno sum as

$$O^X_q(k) = \sum_{c \geq 0} O_q(k;c) X^c = \sum_{c_1, \ldots, c_r \geq 0} \zeta_B^Z(k_1 + c_1, \ldots, k_r + c_r) X^{c_1 + \cdots + c_r}$$

$$= \sum_{m_1 > \cdots > m_r > 0, c_1, \ldots, c_r \geq 0} \frac{q^{(k_1+c_1-1)m_1} \cdots q^{(k_r+c_r-1)m_r}}{[m_1]_q^{k_1+c_1} \cdots [m_r]_q^{k_r+c_r}} X^{c_1 + \cdots + c_r}$$

$$= \sum_{m_1 > \cdots > m_r > 0} \frac{q^{(k_1-1)m_1}}{[m_1]_q - q^m X}[m_1]_q^{k_1-1} \cdots \frac{q^{(k_r-1)m_r}}{[m_r]_q - q^m X}[m_r]_q^{k_r-1}$$

$$= \sum_{m_1 > \cdots > m_r > 0} s^X_q(k_1, m_1) \cdots s^X_q(k_r, m_r),$$

where we set

$$s^X_q(k, m) = \frac{q^{(k-1)m}}{([m]_q - q^m X)[m]_q^{k-1}}. \ (3.8)$$

Notice that $s^X_q(k, m) = \frac{1}{m^r}$, i.e. the $q$-Ohno sum reduces to the multiple zeta value $O^Z_q(k) = \zeta(k)$ in this case. To define the connected sum we need to find the correct connector $c^X_q(m, n)$, which generalizes the connector $c^Z_q(m, n) = c(m, n) = \frac{m!n!}{(m+n)!}$ we used for multiple zeta values. This was done by Seki and Yamamoto in [SY], by choosing the connector

$$c^X_q(m, n) = \frac{q^{mn} f^X_q(m) f^X_q(n)}{f^X_q(m+n)}, \ (3.9)$$

where $f^X_q(m) = \prod_{j=1}^m ([j]_q - q^j X)$, which can be seen as a variant of the factorial, since $f^X_q(1) = m!$.

**Definition 3.16.** Let $k = (k_1, \ldots, k_r), l = (l_1, \ldots, l_s)$ be two non-empty indices. Then we define the connected $q$-Ohno sum $O^X_q(k;l)$ by

$$O^X_q(k;l) = \sum_{m_1 > \cdots > m_r > 0, n_1 > \cdots > n_s > 0} c^X_q(m_1, n_1) \prod_{i=1}^r s^X_q(k_i, m_i) \prod_{j=1}^s s^X_q(l_j, n_j) \ (3.10)$$

and set $O^X_q(k;0) = O^X_q(0;k) = O^X_q(k)$ for admissible $k$. The $s^X_q$ and $c^X_q$ are given by (3.8) and (3.9).

**Proposition 3.17** ($q$-Transport relations). Let $(k_1, \ldots, k_r)$ and $(l_1, \ldots, l_s)$ be two indices. If $s > 0$ then we have

$$O^X_q(1, k_1, \ldots, k_r; l_1, \ldots, l_s) = O^X_q(k_1, \ldots, k_r; 1, l_1, l_2, \ldots, l_s)$$

and if $r > 0$, then we have

$$O^X_q(k_1 + 1, k_2, \ldots, k_r; l_1, \ldots, l_s) = O^X_q(k_1, \ldots, k_r; 1, l_1, l_2, \ldots, l_s).$$

**Proof.** This proof is Exercise 11 and it is similar to the proof of Proposition 3.11.\[\square\]

With the same argument as in (3.3) we see that these transport relations imply Theorem 3.15 i.e. $O^X_q(k) = O^X_q(k)$.\[\square\]

**Remark 3.18.** We will see in Section 5 that Theorem 3.15 also implies a version of Ohno's relation for our $q$-series $g(k)$. More precisely, we will introduce a double-indexed version $g^q(k)$, which span the space $Z_q$, introduced in Section 1.4.3. We then obtain relations in this space since the coefficients of the modified version $(1 - q)^{-w(k)} O^q_{(1-q)^{-1}}(k)$ are also elements in $Z_q$.\[\square\]
3.3 The zoo of relations

We already studied various relations of multiple zeta values. But there are many more relations, which we will not be able to cover in this course. Instead, we will provide a small overview of some of them and their relationships. In Figure 1, we give an overview of families of linear and algebraic relations among multiple zeta values and some of their implications. This overview was done together with the help of T. Tanaka, and for a more detailed list of results on multiple zeta values, one should have a look at the list of research papers collected by Hoffman in [H0]. Some of the presented relations already appeared before, and for some of them, we will provide a few explanations. For the remaining ones, we refer to the literature. (I will try to add more details & examples in the future)

In the following, we provide references and some explanations of Figure 1.

1 Motivic relations: See [Br].

2 Associator relations: See [D] and [F].

3 Confluence relations: See [HS].

4 Integral-Series identity: See [KV].

5 Extended double shuffle relations: This is Theorem 3.1.

6 Kawashima’s relations: Define the automorphism \( \varphi \in \text{Aut}(\mathfrak{g}) \) (with respect to the concatenation) on the generators by \( \varphi(x) = x + y \) and \( \varphi(y) = -y \) and define for words \( v, w \in \mathfrak{g}y \) the operator \( z_p v \otimes z_q w = z_{p+q}(v \ast w) \). With this the Kawashima relations can be stated as follows:

**Theorem 3.19.** ([Kaw, Corollary 5.4]) For all \( v, w \in \mathfrak{g}y \) and \( m \geq 1 \) we have

\[
\sum_{i+j=m, i,j \geq 1} \zeta(\varphi(v) \otimes y^i) \zeta(\varphi(w) \otimes y^j) = \zeta(\varphi(v \ast w) \otimes y^m). 
\]  

(3.11)

It is expected that Theorem 3.19 gives all \( \mathbb{Q} \)-linear relations between multiple zeta values after evaluating the product on the left-hand side by the shuffle product formula. Moreover, numerical experiment suggests that the two cases \( m = 1, 2 \) are enough to obtain all linear relations.

7 Duality: This is Proposition 2.5.

8 Euler’s relations: This is Proposition 1.1.

9 Finite double shuffle relations: This is Proposition 2.14.

10 Shuffle product: This is Corollary 2.10.

11 Stuffle product: This is Corollary 2.13.
MZVs and modular forms • Families of linear relations and their $q$-relatives

Figure 1: Overview of some relations among multiple zeta values.
12. **Ohno’s relations:** This is Theorem 3.13.

13. **Linear part of Kawashima’s relations** This is the $m=1$ case of Theorem 3.19. Notice that in this case the sum on the left-hand side of (3.11) is zero and therefore we obtain the linear relation

$$\zeta(\varphi(v \ast w) \otimes y) = 0.$$  

The number of relations obtained from the linear part of Kawashima’s relations is given by the following table (calculated by Tanaka in [Tan2]).

<table>
<thead>
<tr>
<th>weight $k$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td># all conjectured relations</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>29</td>
<td>60</td>
<td>123</td>
<td>249</td>
<td>503</td>
<td>1012</td>
</tr>
<tr>
<td># Linear part of Kawashima’s relations</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>23</td>
<td>46</td>
<td>98</td>
<td>200</td>
<td>413</td>
<td>838</td>
</tr>
</tbody>
</table>

14. **Tanaka’s rooted tree maps:** Rooted tree maps were introduced by Tanaka in [Tan3]. To a rooted tree he assigns a map $f \in \text{End} H$, which gives an element in $f(w) \in \ker \zeta$, when evaluated at an admissible word $w \in H^0$. Before we can give the definition of the rooted tree maps, we need to recall some basics on rooted trees and the Connes-Kreimer coproduct.

A rooted tree is a finite graph which is connected has no cycles, and has a distinguished vertex called the root. We draw rooted trees with the root on top, and we just consider rooted trees without plane structure, which means that we do not distinguish between $\circ$ and $\ast$. A product (given by the disjoint union) of rooted trees will be called a (rooted) forest, and by $H$ we denote the $\mathbb{Q}$-algebra of forests generated by all trees. The unit of $H$, given by the empty forest, will be denoted by $I$. Since we just consider trees without plane structure, the algebra $H$ is commutative. Due to the work of Connes and Kreimer (CK), the space $H$ has the structure of a Hopf algebra. To define the coproduct on $H$, we first define the linear map $B_+$ on $H$, which connects all roots of the trees in a forest to a new root. For example we have $B_+ (\circ) = \ast$.

Clearly for every non-empty tree $t \in H$ there exists a unique forest $f_t \in H$ with $t = B_+(f_t)$, which is just given by removing the root of $t$. The coproduct on $H$ can then be defined recursively for a tree $t \in H$ by

$$\Delta(t) = t \otimes I + (\text{id} \otimes B_+) \circ \Delta(f_t)$$

and for a forest $f = gh$ with $g, h \in H$ multiplicatively by $\Delta(f) = \Delta(g) \Delta(h)$ and $\Delta(I) = I \otimes I$. For example we have

$$\Delta(\ast) = \ast \otimes I + \ast \otimes \ast + 2 \ast \otimes \ast \otimes I \otimes \ast.$$  

In [Tan3] Tanaka uses the coproduct $\Delta$ to assign to a forest $f \in H$ a $\mathbb{Q}$-linear map on the space $H$, called a rooted tree map, by the following:

**Definition 3.20.** The rooted tree map of the empty forest $I$ is given by the identity map on $H$. For any non-empty forest $f \in H$, we define a $\mathbb{Q}$-linear map on $H$, also denoted by $f$, recursively: For a word $w \in H$ and a letter $u \in \{x, y\}$ we set

$$f(wu) := M(\Delta(f)(w \otimes u)), \quad (3.12)$$

where $M(w_1 \otimes w_2) = w_1 w_2$ denotes the multiplication on $H$. This reduces the calculation to $f(u)$ for a letter $u \in \{x, y\}$, which is defined by the following:

i) If $f = \ast$, then $f(x) := xy$ and $f(y) := -xy$.  

Version 3.9 (June 26, 2020)
MZVs and modular forms

ii) For a tree \( t = B_+(f) \) we set \( t(u) := R_y R_{x+2y} R_y^{-1} f(u) \), where \( R_v \) is the linear map given by \( R_v(w) = vw \) \((v, w \in \mathbb{S})\).

iii) If \( f = gh \) is a forest with \( g, h \neq 1 \), then \( f(u) := g(h(u)) \).

**Theorem 3.21.** ([Tan3] Theorem 1.3) For any non-empty forest \( f \in \mathcal{H} \) we have 
\[
 f(\Delta) \subset \ker \zeta.
\]

**Example 3.22.** For the tree \( f = \Delta \) and the word \( w = xy \) we obtain for 
\[
\mathfrak{I}(xy) = M(\Delta(\mathfrak{I}(x \otimes y))) = M(\mathfrak{I}(x) \otimes y + \mathfrak{I}(y))
\]
Together with \( \mathfrak{I}(x) = xy \) and \( \mathfrak{I}(x) = R_y R_{x+2y} R_y^{-1} \mathfrak{I}(x) = x(y + 2y) \) we get 
\[
\mathfrak{I}(xy) = 2xyy - xyy - xxy - xyy = 2z_2z_1z_1 - z_2z_1 - z_4 - z_3z_1
\]

Ohno-Zagier relations: We define the **height** of an index \( k = (k_1, \ldots, k_j) \) by 
\[
\text{ht}(k) = \#\{i \mid k_i > 1\},
\]
i.e. the number of \( k_j \) not equal to 1. The Ohno-Zagier relations give an explicit formula for the sum of all multiple zeta values of a fixed weight, depth and height as a polynomial in single zeta values:

**Theorem 3.23 ([OZ]).** We have 
\[
\sum_{k \geq r + h \atop r \geq h \geq 1} \left( \sum_{\substack{k, \text{adm.} \atop \text{wt}(k) = k \atop \text{dep}(k) = r \atop \text{ht}(k) = h}} \zeta(k) \right) X^{k-r-h} Y^{r-h} Z^{h-1} = \frac{1}{XY - Z} \left( 1 - \exp \left( \sum_{n=2}^{\infty} \frac{\zeta(n)}{n} S_n(X, Y, Z) \right) \right),
\]

where the \( S_n(X, Y, Z) \in \mathbb{Z}[X, Y, Z] \) are defined by 
\[
S_n(X, Y, Z) = X^n + Y^n - \alpha^n - \beta^n, \quad \alpha, \beta = \frac{X + Y \pm \sqrt{(X + Y)^2 - 4Z}}{2}.
\]

For \( n = 2, \ldots, 5 \) the \( S_n(X, Y, Z) \) are given by 
\[
S_2(X, Y, Z) = -2XY + 2Z, \quad S_3(X, Y, Z) = -3X^2Y - 3XY^2 + 3YZ + 3YZ,
S_4(X, Y, Z) = -4X^3Y - 6X^2Y^2 - 4XY^3 + 4X^2Z + 8XYZ + 4Y^2Z - 2Z^2.
\]

There are also Ohno-Zagier type relations for the \( q \)-MZV \( \zeta_{q}^{\mathcal{BZ}} \) proven by Okuda and Takeyama in [OT].

**Quasi-derivation relations:** Quasi-derivation relations were first proposed in [Kan], and then it was shown in [Tan2] that they give linear relations among multiple zeta values.
Definition 3.24. Let \( c \in \mathbb{Q} \) and \( H \) the derivation on \( \mathcal{H} \) defined by \( H(w) = \deg(w)w \) for any \( w \in \mathcal{H} \). For an integer \( n \geq 1 \), the \( \mathbb{Q} \)-linear map \( \partial_n^{(c)} : \mathcal{H} \to \mathcal{H} \), called quasi-derivation, is defined by
\[
\partial_n^{(c)} = \frac{1}{(n-1)!} \text{ad} \left( \theta^{(c)} \right)^{n-1}(\partial_1),
\]
where \( \theta^{(c)} : \mathcal{H} \to \mathcal{H} \) is the \( \mathbb{Q} \)-linear map defined by \( \theta^{(c)}(x) = \frac{1}{2}(xz+zx) \), \( \theta^{(c)}(y) = \frac{1}{2}(yz+zy) \) with \( z = x + y \) and the rule
\[
\theta^{(c)}(w) = \theta^{(c)}(w)w' + w\theta^{(c)}(w') + c\partial_1(w)H(w')
\]
for any \( w, w' \in \mathcal{H} \).

In [Tan2] it was shown that \( \partial_n^{(c)} \) evaluated at admissible words gives linear relations among multiple zeta values. Further, it was shown that these relations are consequences of the linear part of Kawashima’s relations.

Theorem 3.25. (Quasi-derivation relations, [Tan2, Theorem 3]) For all \( n \geq 1 \) and \( c \in \mathbb{Q} \) we have
\[
\partial_n^{(c)}(\mathcal{H}^0) \subset \ker \zeta.
\]

Open problem 3.26. Show that the quasi-derivations \( \partial_n^{(c)} \) can be written in terms of rooted tree maps.

In [BTan2] it was shown, that the linear part of the Kawashima relation (\( m = 1 \) in Theorem 3.19) is equivalent to the rooted tree maps relations (Theorem 3.21). Since the proof of the quasi-derivation relations just uses the linear part of the Kawashima’s relations one might expect that there is an explicit relationship. For \( n = 1, 2, 3, 4 \) one can actually show that we have
\[
\partial_1^{(c)} = \mathcal{T},
\]
\[
3\partial_2^{(c)} = (2 \mathcal{I} - \mathcal{O}) + (\mathcal{I} + \mathcal{O}) c,
\]
\[
7\partial_3^{(c)} = (3 \mathcal{I} - 3 \mathcal{O} + \mathcal{T}) + \left( \frac{7}{3} \mathcal{A} + \frac{17}{6} - 2 \mathcal{O} - \mathcal{T} \right) c + \left( 2 \mathcal{A} + \frac{2}{3} + \frac{1}{2} \mathcal{O} + \mathcal{T} \right) c^2,
\]
\[
15\partial_4^{(c)} = (2 \mathcal{A} + 4 \mathcal{O} + 2 \mathcal{I} - 2 \mathcal{I} - 4 \mathcal{O} + 4 \mathcal{O} + \mathcal{T})
\]
\[
+ \left( \frac{139}{63} \mathcal{A} + \frac{53}{63} \mathcal{O} + \frac{52}{35} \mathcal{I} + \frac{37}{63} \mathcal{I} - \frac{52}{9} \mathcal{O} - \frac{22}{7} \mathcal{I} + \frac{58}{63} \mathcal{O} + \frac{53}{63} \mathcal{T} \right) c
\]
\[
+ \left( \frac{211}{63} \mathcal{A} + \frac{281}{126} \mathcal{O} + \frac{71}{18} \mathcal{I} + \frac{31}{14} \mathcal{I} - \frac{41}{18} \mathcal{O} - \frac{4}{21} \mathcal{I} + \frac{82}{63} \mathcal{O} + \frac{88}{63} \mathcal{T} \right) c^2
\]
\[
+ \left( \frac{173}{126} \mathcal{A} + \frac{52}{63} \mathcal{O} + \frac{7}{9} \mathcal{I} + \frac{2}{7} \mathcal{I} + \frac{1}{18} \mathcal{I} + \frac{4}{21} \mathcal{O} + \frac{58}{63} \mathcal{O} + \frac{53}{63} \mathcal{T} \right) c^3.
\]

But for \( n \geq 5 \) it is not clear how to express \( \partial_n^{(c)} \) in terms of rooted tree maps and for \( n \geq 4 \) this representation is also not unique, since there are relations among rooted tree maps. For example we have
\[
\mathcal{T} = 2 \mathcal{I} + \mathcal{O} - \mathcal{A} - \mathcal{A}.
\]

\[\text{Here } \text{ad}(\theta)(\partial) := \theta \partial - \partial \theta.\]
Derivation relations: Define for $n \geq 1$ the derivation $\partial_n$ on $\mathcal{F}$ by
$$\partial_n(x) = x(x+y)^{n-1}y$$
and $\partial_n(y) = -x(x+y)^{n-1}y$. This is a derivation on $\mathcal{F}$ with respect to the usual non-commutative multiplication and therefore is suffices to just define it on the generators $x$ and $y$. For any $w = uv$ with $u, v \in \mathcal{F}$ it is then defined by using Leibniz’s rule $\partial_n(uv) = \partial_n(u)v + u\partial_n(v)$.

**Theorem 3.27.** (Derivation relation, [IKZ, Corollary 6]) For all $n \geq 1$ we have
$$\partial_n(\mathcal{F}^0) \subset \ker \zeta.$$  
Also notice that this is the $c = 0$ case of the quasi-derivation relations (Theorem 3.25), since $\partial_n^{(0)} = \partial_n$.

**Example 3.28.** If $n = 2$ we have
$$\partial_2(x) = -\partial_2(y) = xxy + yyy,$$
i.e. we get for the admissible word $xy$:
$$\partial_2(xy) = \partial_2(x)y + x\partial_2(y) = xxyy + yyy - xxyy = xyy - xxyy = z_{2,1,1} - z_4.$$  
The derivation relation then gives the relation $\zeta(2,1,1) = \zeta(4)$.

Restricted sum formula: As a generalization of the sum formula Eie, Liaw, and Ong proved the following formula.

**Theorem 3.29** (Restricted sum formula, [ELO]). For integers $p \geq 0$ and $k > r \geq 1$ we have
$$\sum_{k \text{ adm.}, \frac{\text{wt}(k) = k}{\text{dep}(k) = r}} \zeta(k, \{1\}^p) = \sum_{l \text{ adm.}, \frac{\text{wt}(l) = k+p}{\text{dep}(l) = p+1}} \zeta(l).$$

Notice that this reduces to the sum formula (Theorem 3.12) in the case $p = 0$.

Cyclic sum formula: In [HO] Hoffman and Ohno proved the following relation.

**Theorem 3.30** (Cyclic sum formula). Let $k_1, \ldots, k_r \geq 1$ be integers with at least one $k_j \geq 2$. Then
$$\sum_{j=1}^{r} \zeta(k_j + 1, k_{j+1}, \ldots, k_r, k_1, \ldots, k_{j-1}) = \sum_{1 \leq j \leq r} \sum_{m=0}^{k_j-2} \zeta(k_j - m, k_{j+1}, \ldots, k_r, k_1, \ldots, k_{j-1}, m + 1).$$

**Example 3.31.** If we take $k_1 = 1, k_2 = 2$ and $k_3 = 3$ the cyclic sum formula gives
$$\zeta(2, 2, 3) + \zeta(3, 3, 1) + \zeta(4, 1, 2) = \zeta(2, 3, 1, 1) + \zeta(3, 1, 2, 1) + \zeta(2, 1, 2, 2).$$

Weighted sum formula: There are different types of weighted sum formulas, and one variant of them can be, for example, found in [Kad].
21 **Hoffman’s relations:** This is Proposition 3.5

22 **Sum formula:** This is Theorem 3.12

23 The extended double shuffle relations imply the Ohno-Zagier relations: See [L].

24 The associator relations imply the extended double shuffle relations: See [F2].

25 The associator relations and the confluence relations are equivalent: See [F3].

26 The extended double shuffle relations together with the duality imply Kawashima’s relations: See [Kaw].

27 The rooted tree maps relations are equivalent to the linear part of Kawashima’s relations: See [BTan1].

28 The rooted tree maps relations imply the derivation relation: This was shown in [BTan1] by writing the derivations $\partial_n$ explicitly as rooted tree maps. For this just trees without any branches are needed, i.e. just consider for $m \geq 1$ the ladder trees

$$\lambda_m = m$$

and set $\lambda_0 = I$. With this the main result of [BTan1] states.

**Theorem 3.32** ([BTan1]). For all $n \geq 1$ the derivation $\partial_n$ is given by

$$\partial_n = \frac{n}{2^n - 1} \sum_{d=1}^{n} \frac{(-1)^{d+1}}{d} \sum_{m_1 + \cdots + m_d = n} m_1 \cdots m_d \lambda_{m_1} \cdots \lambda_{m_d}. \quad (3.13)$$

By Definition 3.20 iii) we have $\lambda_{m_1} \lambda_{m_2} = \lambda_{m_2} \lambda_{m_1}$, so we get for the first few values of $n$ ($w \in H$)

$$\partial_1(w) = \bullet(w), \quad \partial_2(w) = \frac{2}{3} \bullet(w) - \frac{1}{3} \bullet\bullet(w), \quad \partial_3(w) = \frac{3}{7} \bullet\bullet\bullet(w) - \frac{3}{7} \bullet(w) + \frac{1}{7} \bullet\bullet\bullet\bullet(w).$$

29 The linear part of Kawashima’s relations imply the cyclic sum formula: See [TW].

30 **Ohno’s relations imply the weighted sum formula:** See [Kad].

31 The derivation relations imply the restricted sum formula: See [Tan4].

32 **Period polynomial relations:** These types of relations will be explained in the next section. See [GKZ].

33 **Double Ohno relations** See [HMOS] and [HSS].
§4 Double zeta values and modular forms

In this section, we will focus on the extended double shuffle relations in the smallest depth, i.e., on some of the relations among double zeta values. Almost all results in this section are contained in or inspired by the work [GKZ].

4.1 The formal double zeta space

The idea of the formal double zeta space is to consider formal symbols, which satisfy the same relations as double zeta values, which come from the double shuffle relations in the smallest depths. We, therefore, start by recalling these relations before defining the formal double zeta space.

4.1.1 Double zeta values

Recall that for \( k_1, k_2 \geq 2 \) we have the finite double shuffle relations
\[
\zeta(k_1)\zeta(k_2) = \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2)
\]
\[
= \sum_{j=2}^{k_1+k_2-1} \left( \left( \frac{j-1}{k_1-1} \right) + \left( \frac{j-1}{k_2-1} \right) \right) \zeta(j, k_1 + k_2 - j).
\]

Using the stuffle and shuffle regularized multiple zeta values, we have for all \( k_1, k_2 \geq 1 \)
\[
\zeta^\ast(k_1; T)\zeta^\ast(k_2; T) = \zeta^\ast(k_1, k_2; T) + \zeta^\ast(k_2, k_1; T) + \zeta^\ast(k_1 + k_2; T)
\]
\[
= \sum_{j=1}^{k_1+k_2-1} \left( \left( \frac{j-1}{k_1-1} \right) + \left( \frac{j-1}{k_2-1} \right) \right) \zeta^\ast(j, k_1 + k_2 - j; T). \tag{4.1}
\]

The comparison map \( \rho \) (Theorem 2.35), which gives \( \zeta^\ast(k; T) = \rho(\zeta^\ast(k; T)) \), satisfies \( \rho(1) = 1, \rho(T) = T \) and \( \rho(T^2) = T^2 + \zeta(2) \). Therefore the \( \zeta^\ast(k_1, k_2; T) \) and \( \zeta^\ast(k_1, k_2; T) \) just differ in the case \( k_1 = k_2 = 1 \) and we have
\[
\zeta^\ast(1, 1; T) = \zeta^\ast(1, 1; T) + \frac{1}{2} \zeta(2), \tag{4.2}
\]
and \( \zeta^\ast(1, 1; T) = \frac{1}{2} T^2, \zeta^\ast(1, 1; T) = \frac{1}{2} T^2 - \frac{1}{2} \zeta(2) \). Now define for \( \bullet \in \{\ast, \ast\} \) their generating series
\[
\mathcal{T}^\bullet(X) = \sum_{k\geq 1} \zeta^\bullet(k; T)X^{k-1}, \quad \mathcal{T}^\bullet(X, Y) = \sum_{k_1, k_2 \geq 1} \zeta^\bullet(k_1, k_2; T)X^{k_1-1}Y^{k_2-1}.
\]

Using \( \frac{X^{k_1-1}Y^{k_2-1}}{X-Y} = \sum_{k_1+k_2=k} X^{k_1-1}Y^{k_2-1} \) we see that \( \text{(4.1)} \) together with \( \text{(4.2)} \) can therefore be written as
\[
\mathcal{T}^\bullet(X)\mathcal{T}^\bullet(Y) = \mathcal{T}^\bullet(X, Y) + \mathcal{T}^\bullet(Y, X) + \frac{\mathcal{T}^\bullet(X) - \mathcal{T}^\bullet(Y)}{X-Y} - \delta_{\ast, \ast} \zeta(2) \tag{4.3}
\]
\[
= \mathcal{T}^\bullet(X + Y; Y) + \mathcal{T}^\bullet(X + Y; X) + \delta_{\ast, \ast} \zeta(2),
\]
where \( \delta \) denotes the Kronecker-delta.
4.1.2 The formal double zeta space

We will now define the formal double zeta space which is spanned by formal symbols $Z_k, Z_{k_1,k_2}$ and $P_{k_1,k_2}$ for $k, k_1, k_2 \geq 2$, which satisfy similar relations to the regularized versions of $\zeta(k), \zeta(k_1, k_2)$ and $\zeta(k_1)\zeta(k_2)$. The only difference will be, that we will ignore the correction term in the $k = 2$ case, which will not bring any problems but makes things a little bit cleaner.

**Definition 4.1.** We define for $k \geq 1$ the **formal double zeta space** in weight $k$ as

$$D_k = \left\langle Z_k, Z_{k_1,k_2}, P_{k_1,k_2} \mid k_1 + k_2 = k, k_1, k_2 \geq 1 \right\rangle_{\mathbb{Q}}$$

where we divide out the following relations for $k_1, k_2 \geq 1$

$$P_{k_1,k_2} = Z_{k_1,k_2} + Z_{k_2,k_1} + Z_{k_1+k_2}$$

$$= \sum_{j=1}^{k_1+k_2-1} \left( \binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) Z_{j,k_1+k_2-j}.$$

(4.4)

**Remark 4.2.** Notice that by definition the $P_{k_1,k_2}$ can always be expressed in terms of the $Z$ and it would therefore be equivalent to define the space $D_k$ by the span of elements $Z_{k_1,k_2}$ and $Z_k$ modulo the relations

$$Z_{k_1,k_2} + Z_{k_2,k_1} + Z_{k_1+k_2} = \sum_{j=1}^{k_1+k_2-1} \left( \binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) Z_{j,k_1+k_2-j}.$$

(4.5)

But it is convenient to also work with the $P_{k_1,k_2}$, since they correspond to something like the product in most of the realizations (Definition 4.3) later.

For small weights the formal double zeta space is given by the following relations and basis elements. Since $P_{k_1,k_2} = P_{k_2,k_1}$ is symmetric we will always just consider the case $k_1 \leq k_2$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>Relations in $D_k$</th>
<th>Basis of $D_k$</th>
<th>dim$_{\mathbb{Q}} D_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$Z_2 = 0$, $P_{1,1} = 2Z_{1,1}$</td>
<td>$Z_1$</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$Z_{2,1} = Z_3$, $Z_{1,2} = P_{1,2} - 2Z_3$</td>
<td>$Z_{1,1}$</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>$Z_4 = 4Z_{3,1}$, $Z_{2,2} = 3Z_{3,1}$, $P_{1,3} = Z_{1,3} + 5Z_{3,1}$, $P_{2,2} = 10Z_{3,1}$</td>
<td>$Z_{3}, P_{1,3}$</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>$Z_{4,1} = 2Z_5 - P_{2,3}$, $Z_{3,2} = \frac{-11}{2} Z_5 + 3P_{2,3}$, $Z_{2,3} = \frac{9}{2} Z_5 - 2P_{2,3}$, $Z_{1,4} = -3Z_5 + P_{1,4} + P_{2,3}$</td>
<td>$Z_5, P_{1,4}, P_{2,3}$</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>$Z_{1,5}, Z_{3,3}, Z_{5,1}$</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 2: Relations and bases for the formal double shuffle space in small weights.
Observe in Figure 2 that for even weight $k$ we seem to have $\dim_{\mathbb{Q}} D_k = \frac{k}{2}$ with a basis given by $Z_{\text{odd, odd}}$ and for odd weight $k$ it seems that $\dim_{\mathbb{Q}} D_k = \frac{k+1}{2}$ with a basis given by $Z_k$ and $P_{k_1, k_2}$. These observations are indeed correct and we will prove them in this section. Notice that lower bounds of $\dim_{\mathbb{Q}} D_k$, which coincides with the observed dimension, already follow from the definition. Since (4.5) is symmetric in $k_1$ and $k_2$ we obtain for $k$ even $\frac{k-1}{2}$ relations among the $k$ generators $Z_k, Z_{1,k}, \ldots, Z_{k-1,1}$ and therefore we have for even $k$

$$\dim_{\mathbb{Q}} D_k \geq \frac{k}{2}. \quad (4.6)$$

For $k$ odd we have $\frac{k-1}{2}$ relations and therefore

$$\dim_{\mathbb{Q}} D_k \geq \frac{k+1}{2}. \quad (4.7)$$

It is convenient to consider generating series when working with the formal double zeta space and therefore we define the following elements in $D_k[ X, Y ]$:

$$\mathcal{Z}_k(X, Y) = \sum_{k_1+k_2=k, k_1, k_2 \geq 1} Z_{k_1, k_2} X^{k_1-1} Y^{k_2-1},$$

$$\mathcal{P}_k(X, Y) = \sum_{k_1+k_2=k, k_1, k_2 \geq 1} P_{k_1, k_2} X^{k_1-1} Y^{k_2-1},$$

$$\mathcal{R}_k(X, Y) = Z_k \frac{X^{k-1} - Y^{k-1}}{X - Y}.$$

With this the double shuffle relations (4.4) can be written as

$$\mathcal{P}_k(X, Y) = \mathcal{Z}_k(X, Y) + \mathcal{Z}_k(Y, X) + \mathcal{R}_k(X, Y)$$

$$= \mathcal{Z}_k(X + Y, Y) + \mathcal{Z}_k(X + Y, X). \quad (4.8)$$

We will not only be interested in relations in the space $D_k$, but also in "real"-mathematical objects which satisfy these relations. Therefore we first introduce the following notation.

**Definition 4.3.** Let $A$ be a $\mathbb{Q}$-vector space. We define the $A$-valued points $D_k(A)$ for $D_k$ by

$$D_k(A) = \text{Hom}_{\mathbb{Q}}(D_k, A) = \{(Z_k, Z_{k_1, k_2}) \in A^k \mid \text{satisfying (4.5)} \}.$$

An element in $D_k(A)$, i.e. one particular choice of $Z_k, Z_{k_1, k_2} \in A$ for all $k_1 + k_2 = k$ which satisfy (4.5), will be called a **realization** of the double zeta space in $A$.

By comparing (4.3) and (4.8) we see that one realization is given by the shuffle regularized multiple zeta values: For $A = \mathbb{R}[T]$ a realization of $D_k$ with $k \geq 1$ is given for $k_1, k_2 \geq 1$ and $k_1 + k_2 = k$ by

$$Z_k \mapsto \begin{cases} \zeta^{(1)}(k; T) & k \neq 2 \\ 0 & k = 2 \end{cases},$$

$$Z_{k_1, k_2} \mapsto \zeta^{(1)}(k_1, k_2; T),$$

$$P_{k_1, k_2} \mapsto \zeta^{(1)}(k_1; T) \zeta^{(1)}(k_2; T).$$

We will refer to this realization as the **multiple zeta realization**. Later we will also introduce realizations in the cases $A = \mathbb{Q}, A = \mathbb{Q}[q]$ and $A = \mathcal{O}(\mathbb{H})$. But before doing so we will prove some results in $D_k$. Using the description in terms of generating series we obtain the following theorem.
Theorem 4.4.  
i) For all \( k \geq 2 \) we have
\[
\sum_{j=2}^{k-1} Z_{j,k-j} = Z_k.
\]

ii) For \( k \geq 2 \) even, we have
\[
\sum_{\substack{j=2 \\ j \text{ even}}}^{k-2} Z_{j,k-j} = \frac{3}{4} Z_k, \quad \sum_{\substack{j=2 \\ j \text{ odd}}}^{k-1} Z_{j,k-j} = \frac{1}{4} Z_k.
\]

Proof. By \([4,8]\) we have
\[
D(X,Y) := 3_k(X+Y,Y) + 3_k(X+Y,X) - 3_k(Y,X) - 3_k(Y,X) - 9_k(X,Y) = 0.
\]
The first statement now follows by taking the case \((X,Y) = (1,0)\), since
\[
0 = D(1,0) = 3_k(1,0) + 3_k(1,1) - 3_k(1,0) - 3_k(0,1) - Z_k = \sum_{j=1}^{k-1} Z_{j,k-j} - Z_{1,k-1} - Z_k.
\]

For the second statement first consider for even \( k \)
\[
0 = D(1,-1) = 3_k(0,-1) + 3_k(0,1) - 3_k(-1,0) - 3_k(-1,1) - Z_k = 2 \sum_{j=2}^{k-1} (-1)^j Z_{j,k-j} - Z_k.
\]
Taking \( D(1,0) \pm \frac{1}{2} D(1,-1) \) we therefore obtain
\[
0 = D(1,0) + \frac{1}{2} D(1,-1) = 2 \sum_{\substack{j=2 \\ j \text{ even}}}^{k-2} Z_{j,k-j} - \frac{3}{2} Z_k
\]
\[
0 = D(1,0) - \frac{1}{2} D(1,-1) = 2 \sum_{\substack{j=2 \\ j \text{ odd}}}^{k-1} Z_{j,k-j} - \frac{1}{2} Z_k,
\]
from which the second statement follows after dividing by 2.

These polynomials are all elements in \( D_k \otimes \mathbb{Q} V_k \), where \( V_k \subset \mathbb{Q}[X,Y] \) denotes the space of all homogeneous polynomials of degree \( k-2 \). On \( V_k \) we define a right-action of \( \text{GL}_2(\mathbb{Z}) \) for a \( \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathbb{Z}) \) and \( F \in V_k \) by
\[
(F|\gamma)(X,Y) = F(aX+bY,cX+dY).
\]
This action can then extended linearly to an action of the group ring \( \mathbb{Z}[\text{GL}_2(\mathbb{Z})] \) on \( D_k \otimes \mathbb{Q} V_k \). The following elements in \( \text{GL}_2(\mathbb{Z}) \) will be of importance when working with the above group action.
\[
\sigma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \epsilon = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \delta = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},
\]
\[
T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.
\]
With this we can rewrite (4.8) even simpler as
\[ P_k = 3_k \mid (1 + \epsilon) + \Re_k \]
\[ = 3_k \mid T(1 + \epsilon) . \] (4.9)

**Lemma 4.5.** For \( k \geq 1 \) and \( A = e U \epsilon \) we have
\[ 3_k \mid (1 - \sigma) = P_k \mid (1 - \delta)(1 + A - SA^2) - \Re_k \mid (1 + A + A^2) . \]

**Proof.** First notice that \( A = e U \epsilon = T \epsilon T^{-1} \epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \) and that we have \( A^3 = \sigma \). By (4.9) we get
\[ 3_k \mid \epsilon = -3_k + P_k - \Re_k \]
\[ 3_k \mid T \epsilon T^{-1} = -3_k + P_k \mid T^{-1} . \]
and therefore
\[ 3_k \mid = 3_k \mid (T \epsilon T^{-1}) \epsilon = (-3_k + P_k \mid T^{-1}) \epsilon = 3_k + P_k \mid (T^{-1} \epsilon - 1) + \Re_k \]

Iterating this identity two more times we get
\[ 3_k \mid A^3 = 3_k + \Re \mid (1 + A + A^2) . \]

By direct calculation one can check that \((T^{-1} \epsilon - 1)(1 + A + A^2) = -(1 - \delta)(1 + A - SA^2)\).

Since \((3_k \mid \sigma)(X, Y) = 3_k(-X, -Y) = (-1)^k 3_k(X, Y)\) we have
\[ 3_k \mid (1 - \delta) = \begin{cases} 2 3_k & , k \text{ odd} \\ 0 & , k \text{ even} . \end{cases} \]

Also notice that \( P_k \mid (1 - \delta) \) is the generating series of \( P_{ev, \ldots} \), for which we write
\[ P_{ev, \ldots}(X, Y) := \frac{1}{2} (P_k \mid (1 - \delta))(X, Y) = \sum_{j=2}^{k-1} \sum_{j \text{ even}} P_{j, k-j} X^{j-1} Y^{k-j-1} . \]

**Theorem 4.6 (Parity).** For odd \( k \geq 3 \), every \( Z_{k_1, k_2} \) with \( k_1, k_2 \geq 1 \) and \( k_1 + k_2 = k \) can be written as a linear combination of \( P_{ev, odd} \) and \( Z_k \). More precisely we have
\[ Z_{k_1, k_2} = (-1)^{k_1} \sum_{j=2}^{k-1} \left( \left( \begin{array}{c} k-j-1 \\ k_1-1 \end{array} \right) + \left( \begin{array}{c} k-j-1 \\ k_2-1 \end{array} \right) + \delta_{j, k_1} \right) P_{j, k-j} + \frac{1}{2} (-1)^{k_1} \left( \begin{array}{c} k_1 + k_2 \\ k_2 \end{array} \right) - 1 \right) Z_k . \]

**Proof.** This follows directly from Lemma 4.5 by considering the coefficient of \( X^{k_1-1} Y^{k_2-1} \) in
\[ 3_k(X, Y) = (P_{ev, \ldots} \mid (1 + A - SA^2))(X, Y) - \frac{1}{2} (\Re_k \mid (1 + A + A^2))(X, Y) , \] (4.10)
and checking that
\[ \sum_{j=1}^{k-1} \left( \left( \begin{array}{c} k-j-1 \\ k_1-1 \end{array} \right) + \left( \begin{array}{c} k-j-1 \\ k_2-1 \end{array} \right) \right) = \left( \begin{array}{c} k_1 + k_2 \\ k_1 \end{array} \right) . \]
Example 4.7. As a consequence of Theorem 4.6 we get the following relations

\[
\begin{align*}
Z_{1,2} &= P_{2,1} - 2Z_3, & Z_{2,3} &= -2P_{2,3} + \frac{9}{2}Z_5, \\
Z_{2,1} &= Z_3, & Z_{3,2} &= 3P_{2,3} - \frac{11}{2}Z_5, \\
Z_{1,4} &= P_{2,3} + P_{4,1} - 3Z_5, & Z_{4,1} &= -P_{2,3} + 2Z_5.
\end{align*}
\]

Using Theorem 4.6 we now can also prove the dimension formula for \(D_k\) in the odd weight \(k\) case.

Theorem 4.8. For odd \(k \geq 1\) we have \(\dim \mathcal{Q}D_k = \frac{k+1}{2}\) and the sets

\[
B_1 = \{Z_k, P_{2,k-3}, P_{4,k-4}, \ldots, P_{k-1,1}\}, \quad B_2 = \{Z_k, Z_{1,k-1}, Z_{3,k-3}, \ldots, Z_{k-2,2}\},
\]

are both bases of \(D_k\).

Proof. We already saw that \(\dim \mathcal{Q}D_k \geq \frac{k+1}{2}\) since there are just \(\frac{k-1}{2}\) different relations among the \(k\) generators of \(D\). To prove equality it suffices therefore to show that any element can be expressed in terms of elements in \(B_1\) or \(B_2\). By Theorem 4.6 we get that any element in \(D_k\) can be expressed as linear combinations of elements in \(B_1\). For the second basis we just need to show, by symmetry, that any \(P_{j,k-j}\) with \(j\) even can be expressed by \(Z_k\) and \(Z_{k_1,k_2}\) with \(k_1\) odd. The equation in Theorem 4.6 modulo \(Z_k\) for odd \(k = 1, 3, \ldots, k - 2\) reads

\[
Z_{k_1,k-k_1} \equiv -\sum_{j=2,\text{even}}^{k-1} \binom{k-j-1}{k_1-1} + \binom{k-j-1}{k_1-1} \mod \mathbb{Q}Z_k.
\]

We therefore need to show that the matrix \(\left(\binom{k-2j-1}{2i-2} + \binom{k-2j-1}{k_1-2i}\right)_{1 \leq i, j \leq \frac{k-1}{2}}\) is invertible. The binomial coefficient \(\binom{m}{n} = \frac{m(m-1)}{n(n-1)}\) is even, when \(m\) is even and \(n\) is odd. Therefore modulo 2 the factor \(\binom{k-2j-1}{k_1-2i}\) vanishes and we get a triangular matrix with 1 on the diagonal, from which we deduce that the matrix is invertible and therefore we can express \(P_{j,k-j}\) in terms of \(Z_{k,odd,even}\) and \(Z_k\).

We will now present consequences of Lemma 4.5 for the even weight case. In general the rest of the whole Section will be devoted to the even weight case and the connection to modular forms. For even \(k\) Lemma 4.5 implies relations among \(P_{ev, ev}\) and \(Z_{ev}\).

Theorem 4.9 (Relations among \(P_{ev, ev}\) and \(Z_{ev}\)). For all \(k_1, k_2 \geq 1\) with \(k = k_1 + k_2\) even we have

\[
\frac{1}{2} \left(\binom{k_1+k_2}{k_2} - (-1)^{k_1}\right) Z_k = \sum_{j=2, \text{even}}^{k-1} \left(\binom{k-j-1}{k_1-1} + \binom{k-j-1}{k_2-1} - \delta_{j,k_1}\right) P_{j,k-j}.
\]

Proof. This also follows from Lemma 4.5 considering the coefficient of \(X^{k_1-1}Y^{k_2-1}\) in

\[
(\Psi_k^{ev} | (1 + A - SA^2)) (X, Y) = \frac{1}{2} (\Re_k | (1 + A + A^2)) (X, Y).
\]
Example 4.10. As a consequence of Theorem 4.9 we get the following relations by considering the coefficients of $X^5Y$ and $X^4Y^2$ in (1.12):

\[6P_{2,6} + 3P_{4,4} = \frac{27}{2} Z_8, \quad 15P_{2,6} + 3P_{4,4} = \frac{57}{2} Z_8.\]

Combining these two relations we obtain

\[P_{4,4} = \frac{7}{6} Z_8.\]

Using the multiple zeta realization, this gives another proof of $\zeta(4)^2 = \frac{7}{6} \zeta(8)$.

Corollary 4.11. For even $k$ we have

\[\sum_{j=2}^{k-2} P_{j,k-j} = \frac{k+1}{2} Z_k.\]

Proof. This is the $(k_1, k_2) = (1, k-1)$ case in Theorem 4.9 but can also be obtained from the even sum formula in Theorem 4.4 ii) together with the relation $P_{j,k-j} = Z_{j,k-j} + Z_{k-j,j} + Z_k$. \hfill \Box

The even weight analogue of Theorem 4.8 is given by the following

Theorem 4.12. For even $k \geq 2$ we have $\dim_{\mathbb{Q}} D_k = k^2$ and the set of $Z_{\text{od}, \text{od}}$, i.e.

\[\{Z_{1,k-1}, Z_{3,k-3}, \ldots, Z_{k-1,1}\},\]

is a basis of $D_k$.

Proof. We will give a proof of this later and also, similar to the odd weight case, give explicit formulas to express the $P_{k_1,k_2}$ and in terms of $Z_{\text{od}, \text{od}}$. By Theorem 3.12 ii) this is already known for $Z_k$. \hfill \Box

From now on we will be interested in realizations in a $\mathbb{Q}$-algebra $A$ and we will consider all weights $k$ at the same time. Therefore we consider the following generating series

\[\mathcal{Z}(X,Y) = \sum_{k_1,k_2 \geq 1} Z_{k_1,k_2} X^{k_1-1} Y^{k_2-1},\]

\[\mathcal{Z}(X) = \sum_{k \geq 1} Z_k X^{k-1},\]

\[\Psi(X,Y) = \sum_{k_1,k_2 \geq 1} P_{k_1,k_2} X^{k_1-1} Y^{k_2-1},\]

\[\mathcal{R}(X,Y) = \sum_{k \geq 1} Z_k \frac{X^{k-1} - Y^{k-1}}{X - Y} = \frac{\mathcal{Z}(X) - \mathcal{Z}(Y)}{X - Y},\]

which will be viewed as elements in $D_{X,Y} = \prod_{k \geq 1} D_k \otimes_{\mathbb{Q}} V_k$. The action of $\mathbb{Z}[\text{GL}_2(\mathbb{Z})]$ on $V_k$ can be extended to $D_{X,Y}$ in the obvious way and therefore we can write the double shuffle relations (4.9) as

\[\Psi = \mathcal{Z} | (1 + \epsilon) + \mathcal{R} = 3 | T(1 + \epsilon).\quad (4.13)\]
In the following we will consider realization of $D_k$ for all $k \geq 1$ in some $\mathbb{Q}$-algebra $A$, i.e. $A$ is not just a vector space anymore. These can be viewed as an element in $\text{Hom}_\mathbb{Q}(D_{X,Y}, A[[X,Y]])$, which we will call a realization of $D$ in $A$. For a realization $\varphi \in \text{Hom}_\mathbb{Q}(D_{X,Y}, A[[X,Y]])$ we write $\varphi(P_{k_1,k_2})$, $\varphi(Z_{k_1,k_2})$ for the coefficients of $X^{k_1-1}Y^{k_2-1}$ in $\varphi(\mathfrak{P}(X,Y))$ and $\varphi(\mathfrak{Z}(X,Y))$. Similarly we denote by $\varphi(Z_k)$ the coefficient of $X^{k-1}$ in $\varphi(\mathfrak{Z}(X))$.

**Theorem 4.13.** Let $\varphi$ be a realization of $D$ in an $\mathbb{Q}$-algebra $A$. For $k \geq 1$ we write

$$Z(k) = \varphi(Z_k) + \delta_{k,2}Z(2).$$

for some fixed element $Z(2) \in A$.

i) Assume that for even $k_1, k_2 \geq 2$ we have

$$\varphi(P_{k_1,k_2}) = Z(k_1)Z(k_2).$$

Then for even $k \geq 2$ we have $Z(k) \in \mathbb{Q}[Z(2)]$ and more precisely we obtain for $m \geq 1$

$$Z(2m) = -\frac{B_{2m}}{2(2m)!} (-24Z(2))^m. \quad (4.14)$$

ii) Assume that there exist a derivation $\partial \in \text{Der}(A)$ such that for all even $k_1, k_2 \geq 2$

$$\varphi(P_{k_1,k_2}) = Z(k_1)Z(k_2) + \frac{\delta_{k_1,2}}{2k_2} \partial Z(k_2) + \frac{\delta_{k_2,2}}{2k_1} \partial Z(k_1).$$

Then for even $k \geq 2$ we have $Z(k) \in \mathbb{Q}[Z(2), Z(4), Z(6)]$. Moreover we get

$$\partial Z(2) = 5Z(4) - 2Z(2)^2,$$

$$\partial Z(4) = 8Z(6) - 14Z(2)Z(4),$$

$$\partial Z(6) = \frac{120}{7}Z(4)^2 - 12Z(2)Z(6), \quad (4.15)$$

and therefore the space $\mathbb{Q}[Z(2), Z(4), Z(6)] = \mathbb{Q}[Z(2), \partial Z(2), \partial^2 Z(2)]$ is closed under $\partial$.

iii) Assume that for even $k_1, k_2 \geq 4$

$$\varphi(P_{k_1,k_2}) = Z(k_1)Z(k_2).$$

Then for even $k \geq 4$ we have $Z(k) \in \mathbb{Q}[Z(4), Z(6)]$.

**Proof.** We first show ii). The explicit formulas in (4.15) can be obtained by combining the relations coming from Theorem 4.9 in the correct way. One of these relations (Corollary 4.11) gives for $k \geq 4$

$$Z(k) = \frac{2}{k+1} \sum_{j \geq 2}^{k-2} \varphi(P_{j,k-j}) = \frac{2}{k+1} \sum_{j \geq 2}^{k-2} Z(j)Z(k-j) + \frac{2}{(k+1)(k-2)} \partial Z(k-2).$$

Inductively we see, together with (4.15), that $Z(k) \in \mathbb{Q}[Z(2), Z(4), Z(6)]$. The first statement in i) is just a special case of ii) for $\partial \equiv 0$. The explicit formula (4.14) can be obtained by considering the generating series of the right hand side and show that it satisfies the same recursion as the generating series of the left hand side (Exercise 12). For ii) consider the equation (4.11) in the cases $(k_1,k_2) = (m,2)$ and $(k_1,k_2) = (m-1,3)$. Taking $m(m-1)$-times the first case and adding $m$-times the second gives for even $k = m + 2 \geq 8$ a relation of $Z_k$ as a linear combination of $P_{j,k-j}$ with $j,k-j \geq 4$ even. Recursively it follows that $Z(k) \in \mathbb{Q}[Z(4), Z(6)]$. \hfill $\square$
Since the multiple zeta realization gives a realization of $D$ in $\mathbb{R}$ which satisfies the condition in i) with $Z(2) = \zeta(2) = \frac{(2\pi i)^2}{2}$, we obtain from \[ \text{Eq. (1.14)} \] the Euler relation $\zeta(2m) = -\frac{B_{2m}}{2} (2\pi i)^{2m}$.

The first three equations in \[ \text{Eq. (4.15)} \] are known as Ramanujan’s differential equation and the last one is called Chazy equation (e.g. see \[ \text{Za2, Section 5.1} \]). One set of solutions for \[ \text{Eq. (4.14)} \] (and actually a description of almost all solutions for the Chazy equation) is given by the Eisenstein series $Z(k) = G_k$, which we defined for $k \geq 1$ by $G_k = -\frac{B_k}{2k!} + g(k) = \beta(k) + g(k)$.

In the following, we want to introduce a realization, which satisfies the condition in ii), and which is exactly given by the Eisenstein series in depth one. For this we will first introduce a realization with $Z(k) = g(k)$ and then a realization with $Z(k) = \beta(k)$. Since $\text{Hom}_\mathbb{Q}(D_{X,Y}, A[[X,Y]])$ is a group, we can add realizations and obtain another realization. Combining the above two realizations will then give us the realization of the Eisenstein series. We will see that this realization satisfies the conditions of ii) & iii) in Theorem \[ \text{4.13} \] and therefore we will obtain a proof of \[ \text{Eq. (4.14)} \] for Eisenstein series and the fact that every Eisenstein series $G_k$ with $k \geq 4$ even can be written as a polynomial in $G_4$ and $G_6$.

### 4.1.3 Combinatorial double Eisenstein series

In this section we want to introduce a realization $\varphi_G$ of $D$ in the space $A = \mathbb{Q}[[q]]$, which is given by the Eisenstein series $G(k) = \beta(k) + g(k)$ in depth one. As mentioned before we will introduce two realizations, one for the constant term, denoted by $\varphi_{\beta}$, and one for the ”g-part”, denoted by $\varphi_g$. The combinatorial double Eisenstein realization will then be their sum

$$\varphi_G = \varphi_{\beta} + \varphi_g.$$  

We will start with the realization $\varphi_g$, which in depth one will be given by

$$\varphi_g : \mathfrak{B}(X) \mapsto g(X) - g(2)X,$$

where $g$ was defined in Section 1 as the generating series for the $q$-series $g$, i.e.

$$g(X_1, \ldots, X_r) = \sum_{k_1, \ldots, k_r \geq 1} g(k_1, \ldots, k_r) X_1^{k_1-1} \cdots X_r^{k_r-1}.$$  

By Lemma \[ \text{1.25} \] we saw that we have the following two explicit expressions

$$g(X_1, \ldots, X_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{e^{X_1 q^{m_1}} \cdots e^{X_r q^{m_r}}}{1 - e^{X_1 q^{m_1}} \cdots 1 - e^{X_r q^{m_r}}} = \sum_{m_1 > \cdots > m_r > 0} \frac{e^{m_1 X_1 q^{m_1}} \cdots e^{m_r (X_1 - X_2) q^{m_r}}}{1 - q^{m_1} \cdots 1 - q^{m_r}}. \tag{4.16}$$

To express their product in a suitable way we introduce the following series

$$b(X) = \frac{1}{2} \left( \frac{1}{X} - \frac{1}{e^X - 1} - \frac{1}{2} \right) = \sum_{k \geq 1} \beta(k) X^{k-1} = -\sum_{k \geq 2} \frac{B_k}{2k!} X^{k-1} = \sum_{m \geq 1} \frac{\zeta(2m)}{(2\pi i)^{2m}} X^{2m-1},$$

which will also give the depth one part of the realization $\varphi_{\beta}$.
Lemma 4.14. We have
\[ g(X)g(Y) = g(X, Y) + g(Y, X) + \frac{g(X) - g(Y)}{X - Y} + \left( b(Y - X) - b(X - Y) \right) \left( g(X) - g(Y) \right) - \frac{1}{2} \left( g(X) + g(Y) \right) \]
\[ = g(X + Y, X) + g(X + Y, Y) - g(X + Y) + g'\left( X + Y \right) + g(2), \]
where we write \( g'(X) = q \frac{d}{dq} \sum_{k \geq 1} g(k) \frac{X^k}{k} \).

Proof. In Proposition 1.33 we showed by using (4.16) that
\[ g(X)g(Y) = \frac{1}{e^X - 1} g(X) + \frac{1}{e^Y - 1} g(Y) \]
\[ = g(X + Y, X) + g(X + Y, Y) - g(X + Y) + \frac{d}{dq} \sum_{k \geq 1} g(k) \frac{(X + Y)^k}{k} + g(2), \]
from which the statement follows by using the definition of \( b(X) \).

Theorem 4.15. Define the following generating series
\[ b(X) = g(X, Y) - g(2)X, \]
\[ b(X, Y) = g(X, Y) - b(Y - X) + \frac{1}{2} \left( g(X) + b(Y)g(X) + b(X - Y)g(Y) \right) + \frac{1}{2} g'(X), \]
where as before \( g'(X) = q \frac{d}{dq} \sum_{k \geq 1} g(k) \frac{X^k}{k} \). This gives a realization \( \varphi_g \) of \( D \) in \( \mathbb{Q}(\!\![q]\!) \) by
\[ \varphi_g : \mathcal{D}(X) \mapsto b(X), \]
\[ \varphi_g : \mathcal{D}(X, Y) \mapsto b(X, Y), \]
\[ \varphi_g : \mathcal{Q}(X, Y) \mapsto g(X)g(Y) + b(X)g(Y) + b(Y)g(X) + \frac{1}{2} \left( g'(X)Y + g'(Y)X \right). \]

Proof. This follows by a direct calculation from Lemma 4.14 but is also given in [GKZ, Theorem 7].
(The formula needs to be checked, \( \Box \))

We will now give the realization \( \varphi_\beta \), which will give the constant term for the combinatorial double Eisenstein series.

Theorem 4.16. With \( b(X) = \sum_{k \geq 1} \beta(k)X^{k-1} = \frac{1}{2} \left( 1 - \frac{1}{e^X - 1} \right) \) and
\[ b(X, Y) = \sum_{k_1, k_2 \geq 1} \beta(k_1, k_2)X^{k_1-1}Y^{k_2-1} \]
\[ := \frac{1}{3} \left( b(X) + b(Y) \right) b(Y) - \frac{5}{12} \frac{b(X) - b(Y)}{X - Y} + \frac{b(X) - b(X - Y)}{4Y} - \frac{b(Y) - b(Y - X)}{12X} - \frac{1}{96} \]
we have
\[ b(X)b(Y) = b(X, Y) + b(Y, X) + \frac{b(X) - b(Y)}{X - Y} - \frac{1}{2} \]
\[ = b(X + Y, X) + b(X + Y, X). \]
In particular this gives a realization \( \varphi_\beta \) of \( \mathcal{D} \) in \( \mathbb{Q} \) by

\[
\begin{align*}
\varphi_\beta : 3(X) &\mapsto b(X) - \beta(2)X, \\
\varphi_\beta : 3(X, Y) &\mapsto b(X, Y), \\
\varphi_\beta : \mathfrak{P}(X, Y) &\mapsto b(X)b(Y).
\end{align*}
\]

**Proof.** Again this can be checked explicitly by using the definition of \( b(X) \). A more systematic point of view, which we will make more explicit later, is that the hyperbolic cotangent

\[
F(X) = -\frac{1}{2} X + b(X) = -\frac{1}{4} \coth \left( \frac{X}{2} \right)
\]

satisfies the Fay identity

\[
F(X)F(Y) + F(X - Y)F(X) + F(-Y)F(X - Y) = \frac{1}{16}.
\] (4.17)

Writing \( G(X, Y) = F(X)F(Y) \) the equation (4.17) can be written as

\[
G(|1 + U + U^2|) = \frac{1}{16},
\]

which leads to a connection of period polynomials for modular forms. We will see later that basically any such \( G \) gives rise to a realization of \( \mathcal{D} \).

**Definition 4.17.** i) We define the combinatorial Eisenstein realization of \( \mathcal{D} \) in \( \mathbb{Q}[\![q]\!] \) by

\[
\varphi_G = \varphi_\beta + \varphi_\kappa,
\]

where the realization \( \varphi_\beta \) and \( \varphi_\kappa \) are given by Theorem 4.15 and 4.15.

ii) For \( k_1, k_2 \geq 1 \) the combinatorial double Eisensteins series \( G(k_1, k_2) \in \mathbb{Q}[\![q]\!] \) are defined by

\[
G(k_1, k_2) = \varphi_G(Z_{k_1, k_2}),
\]

i.e. they are explicitly given by

\[
\sum_{k_1, k_2 \geq 1} G(k_1, k_2)X^{k_1-1}Y^{k_2-1} := b(X, Y) + b(X, Y),
\]

where \( b(X, Y) \) and \( b(X, Y) \) are given in Theorem 4.14 and 4.15.

Notice that we have

\[
\varphi_G(P_{k_1, k_2}) = \beta(k_1, k_2) + g(k_1)g(k_2) + \beta(k_1)g(k_2) + \beta(k_2)g(k_1) + \frac{\delta_{k_1, 2}}{2k_2} \frac{d}{dq} g(k_2) + \frac{\delta_{k_2, 2}}{2k_1} \frac{d}{dq} g(k_1)
\]

\[
= G_{k_1}G_{k_2} + \frac{\delta_{k_1, 2}}{2k_2} \frac{d}{dq} G_{k_2} + \frac{\delta_{k_2, 2}}{2k_1} \frac{d}{dq} G_{k_1},
\]

and therefore the combinatorial Eisenstein realization satisfies the conditions for ii) with \( \partial = \frac{d}{dq} \) and iii) in Theorem 4.13. This gives a proof of Ramanujan’s differential equations and the fact that for any even \( k \geq 4 \) we have \( G_k \in \mathbb{Q}[G_4, G_6] \).

**Proposition 4.18.** The combinatorial double Eisenstein series are modified \( q \)-analogues of the double zeta values, i.e. for \( k_1 \geq 2, k_2 \geq 1 \) we have

\[
\lim_{q \to 1} (1 - q)^{k_1 + k_2} G(k_1, k_2) = \zeta(k_1, k_2).
\]

**Proof.**
Proposition 4.19. Any modular form with rational coefficients can be written as a linear combination of \( G(\text{odd}, \text{odd}) \), i.e for even \( k \geq 4 \) we have

\[
\mathcal{M}_k^Q \subset \langle G(j, k-j) \mid j = 3, 5, \ldots, k-1 \rangle_Q.
\]

Proof. □

Lemma 4.20. For even \( k \geq 4 \) the set

\[
\{G_k\} \cup \left\{ G_{2j}G_{k-2j} \mid \left\lfloor \frac{k-2}{6} \right\rfloor + 2 \leq j \leq \left\lfloor \frac{k}{4} \right\rfloor \right\}
\]

forms a basis of \( \mathcal{M}_k^Q \).

Proof. This is Corollary 1 in [HST]. □

By the Lemma we get a basis of \( S_k^Q \) by

\[
\left\{ G_{2j}G_{k-2j} - \frac{\beta(2j)\beta(k-2j)}{\beta(k)} G_k \mid \left\lfloor \frac{k-2}{6} \right\rfloor + 2 \leq j \leq \left\lfloor \frac{k}{4} \right\rfloor \right\}.
\]

Since we have for \( k_1, k_2 \geq 4 \) that \( \varphi_G(P_{k_1, k_2}) = G_{k_1}G_{k_2} \) we can use an explicit version of Theorem 4.12 to give explicit expressions of this basis in terms of \( G(\text{odd}, \text{odd}) \). Using Proposition 4.18 and the fact that modular forms are also modified \( q \)-analogues of their constant terms gives us then an explanation for the factor \( O(X)O(X) - S(X) \) in the Broadhurst-Kreimer conjecture (Conjecture 1.21).

Theorem 4.21. There exist an explicit basis of \( S_k^Q \) in terms of \( G(\text{odd}, \text{odd}) \). (Can be written down as a closed formula)

Example 4.22. i) A basis for \( S_{12}^Q \) is given by

\[
G(3,9) - 23825 \frac{5197}{10394} G(5,7) - 41431 \frac{10394}{5197} G(7,5) + 360 \frac{5197}{157934} G(9,3) + G(11,1).
\]

ii) A basis for \( S_{16}^Q \) is given by

\[
G(3,13) - 279116 \frac{5197}{78967} G(5,11) - 2125607 \frac{315868}{22562} G(7,9) - 154671 \frac{22562}{315868} G(9,7) - 1040507 \frac{315868}{157934} G(11,5) + 38573 G(13,3) + G(15,1).
\]
The following (sub)sections will be filled in the coming weeks.

4.2 Period polynomial relations
4.3 Double Eisenstein series

§5 Multiple Eisenstein series and q-analogues of MZV

5.1 The Fourier expansion of multiple Eisenstein series
5.2 q-analogues of MZV
5.3 Double indexed q-analogues of MZV
5.4 Double shuffle relations for double indexed q-analogues
5.5 Combinatorial approach to modular forms
Exercise 1.
i) Prove Proposition 1.9, i.e. show that for $k_1, k_2 \geq 2$ we have

$$\zeta(k_1) \zeta(k_2) = \sum_{j=2}^{k_1+k_2-1} \left( \binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) \zeta(j, k_1 + k_2 - j).$$

For this you can use (without a proof) the partial fraction expansion $(a,b \geq 1)$

$$\frac{1}{x^a y^b} = \sum_{j=1}^{a+b-1} \left( \binom{j-1}{a-1} \left( \frac{1}{(x+y)^j x^{a-j}} \right) + \binom{j-1}{b-1} \left( \frac{1}{(x+y)^j y^{b-j}} \right) \right),$$

which can be proven by induction on $a$ and $b$.

ii) Use i) together with $\zeta(k_1) \zeta(k_2) = \zeta(k_1 + k_2) + \zeta(k_1) + \zeta(k_2) + \zeta(1)$ to prove the relation

$$\zeta(7) = 4 \zeta(3,4) + 3 \zeta(4,3) - 2 \zeta(5,2).$$

Exercise 2.
i) Show that Conjecture 1.12 together with Proposition 1.11 would imply that all multiple zeta values (except for $\zeta(\emptyset) = 1$) are transcendental.

ii) Show that Conjecture 1.16 (Hoffman) would imply Conjecture 1.13 (Zagier).

iii) Show that Conjecture 1.21 (Broadhurst-Kreimer) would imply Conjecture 1.13 (Zagier).

Exercise 3. We defined for $k \geq 1$ the Eulerian polynomials $P_k(X)$ and the power series $R_k(X)$ by

$$R_k(X) = \frac{P_k(X)}{(1-X)^k} = \sum_{d>0} \frac{d^{k-1}}{(k-1)!} X^d.$$

i) Prove Lemma 1.24 i.e. show that we have $P_k(0) = 0$ and $P_k(1) = 1$ for all $k \geq 1$.

ii) Prove Lemma 1.28 i.e. show that for all $k_1, k_2 \geq 1$

$$R_{k_1}(X) \cdot R_{k_2}(X) = R_{k_1+k_2}(X) + \sum_{j=1}^{k_1+k_2-1} \left( \lambda_{k_1,k_2}^j + \lambda_{k_2,k_1}^j \right) R_j(X),$$

where the rational numbers $\lambda_{k_1,k_2}^j$ are given by

$$\lambda_{k_1,k_2}^j = (-1)^{k_2-1} \binom{k_1+k_2-1-j}{k_1-1} \frac{B_{k_1+k_2-j}}{(k_1+k_2-j)!},$$

and where we use the convention $\binom{n}{k} = 0$ for $k < 0$. 

Exercise version 0.11 (June 26, 2020)
Exercise 4. Define for even \( k \geq 4 \) the normalized Eisenstein series \( E_k \) by
\[
E_k = 1 - \frac{2k!}{B_k} g(k)
\]
and give two different proofs of the identity
\[
E_4^2 = E_8
\]
between the Eisenstein series
\[
E_4 = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n \quad \text{and} \quad E_8 = 1 + 480 \sum_{n=1}^{\infty} \sigma_7(n) q^n .
\]

i) Analytic proof of (5.1): Use the theory of modular forms, i.e. use Theorem 1.19 and 1.20.

ii) Combinatorial proof of (5.1): You are just allowed to use Proposition 1.29 and 1.30.

Exercise 5.

i) Prove Lemma 1.32 i.e. show that
\[
g(X_1, \ldots, X_r) = \sum_{k_1, \ldots, k_r \geq 1} g(k_1, \ldots, k_r) X_1^{k_1-1} \cdots X_r^{k_r-1}
\]
can be written in the following two ways
\[
g(X_1, \ldots, X_r) = \sum_{m_1 > \cdots > m_r > 0} e^{X_1 q^{m_1}} \cdots e^{X_r q^{m_r}} \frac{1}{1 - e^{X_1 q^{m_1}}} \cdots \frac{1}{1 - e^{X_r q^{m_r}}}
\]
\[
= \sum_{m_1 > \cdots > m_r > 0} e^{m_1 X_1 q^{m_1}} e^{m_2 (X_1 - X_2) q^{m_2}} q^{m_2} \cdots e^{m_r (X_1 - X_2) q^{m_r}} q^{m_r}.
\]

ii) Show equation (1.29), i.e. show that we have
\[
\sum_{m > 0} e^{mX} \frac{q^m}{(1 - q^m)^2} = g(2) + q \frac{d}{dq} \sum_{k \geq 1} g(k) \frac{X^k}{k}.
\]

iii) Show that Propositions 1.29 and 1.30 are a consequence of Proposition 1.33.

Exercise 6.

i) Calculate (2.2) by hand, i.e. show
\[
\zeta(2, 3) = \int_0^1 \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{1 - t_2} \int_0^{t_2} \frac{dt_3}{t_3} \int_0^{t_3} \frac{dt_4}{t_4} \int_0^{t_4} \frac{dt_5}{1 - t_5}.
\]
without using Corollary 2.4.
ii) Prove Proposition 2.12 i.e. show that for any $w, u \in H^1$ and $M \geq 1$ we have
\[ \zeta_M(w)\zeta_M(u) = \zeta_M(w \ast u). \]

**Exercise 7.**

i) Prove Proposition 2.15 i.e. show that for $n \geq 1$ we have
\[ 4^n \zeta(\{3, 1\}^n) = \zeta(\{4\}^n). \]

ii) Show (2.7), i.e. show that we have for $n \geq 1$
\[ \zeta(\{2\}^n) = \frac{\pi^{2n}}{(2n + 1)!}. \]

(Hint: Calculate the coefficient of $x^{2n}$ in $[1, 3].$)

**Exercise 8.** Prove Proposition 2.25 i.e. show the following:

i) The space $\hat{M}^Q$ is closed under $q \frac{d}{dq}$.

ii) We have
\[ \hat{M}^Q = Q[g(2), g'(2), g''(2)]. \]

where $g'$ denotes the derivative with respect to $q \frac{d}{dq}$.

iii) Let $k = (k_1, \ldots, k_r)$ be an index with $k_1, \ldots, k_r \geq 2$ even. Then we have
\[ g^\text{sym}(k) := \sum_{\sigma \in S_r} g(k_{\sigma(1)}, \ldots, k_{\sigma(r)}) \in \hat{M}^Q, \]

where $S_r$ denotes the set of all permutations of $\{1, \ldots, r\}$.

iv) We have
\[ \hat{M}^Q = \left\{ g^\text{sym}(k_1, \ldots, k_r) \mid r \geq 0, k_1 \geq k_2 \geq \cdots \geq k_r \geq 2 \text{ even} \right\}_Q, \]

where we set $g^\text{sym}(\emptyset) = 1$.

**Exercise 9.** Let $\bullet \in \{\cup, \ast\}$.

i) Calculate $\zeta^\bullet(1, 2, 1)$.

ii) Show that
\[ \zeta^\bullet(z_2 \cup z_1, z_1 - z_2 \ast z_1) = 0, \]

by using the finite double shuffle relations and/or the duality relation and/or Eulers formula.
Exercise 10.

i) Show that for any admissible index $k$ we have $d s(z_1, z_k) = z_1 \star_{\bar{k}} z_k - z_1 \ast z_k \in \bar{H}^0$.

ii) Prove Hoffman’s relation (Proposition 3.5) by using the extended double shuffle relations.

Exercise 11.

i) Show that $z_k^* = \tau(z_k)$ for any admissible index $k$.

ii) Give the proof of Proposition 3.17.

Exercise 12.

i) Prove for $m \geq 1$ Euler’s formula (4.14)

$$Z(2m) = - \frac{B_{2m}}{2(2m)!} (-24Z(2))^m,$$

by assuming the condition in i) of Theorem 4.13.

ii*) Assume the conditions in ii) of Theorem 4.13 hold. Try to find an explicit formula for $Z(2m)$ as a polynomial in $\partial^j Z(2)$ for $j \geq 0$, which generalizes Euler’s formula

(Part ii) is a bonus exercise.)

More exercises will come. Always check the homepage for an updated version of the lecture notes and exercise sheets!
MZVs and modular forms • Multiple Eisenstein series and q-analogues of MZV

References


MZVs and modular forms

• Multiple Eisenstein series and q-analogues of MZV


[HS] M. Hirose, N. Sato: *Iterated integrals on \( \mathbb{P}^1 \setminus \{0,1,\infty,z\} \) and a class of relations among multiple zeta values*, preprint, arXiv:1801.03807.


MZVs and modular forms • Multiple Eisenstein series and q-analogues of MZV


MZVs and modular forms • Multiple Eisenstein series and q-analogues of MZV


