The Multiple Zeta Value data mine

J. Blümleina,∗, D.J. Broadhurstb, J.A.M. Vermaserena,c

a Deutsches Elektronen-Synchrotron, DESY, Platanenallee 6, D-15738 Zeuthen, Germany
b Physics and Astronomy Department, Open University, Milton Keynes, MK7 6AA, UK
c Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands

A R T I C L E   I N F O
Article history:
Received 15 July 2009
Received in revised form 9 November 2009
Accepted 14 November 2009

A B S T R A C T
We provide a data mine of proven results for Multiple Zeta Values (MZVs) of the form
\[ \zeta(s_1, s_2, \ldots, s_k) = \sum_{n_1 > n_2 > \ldots > n_k > 0} \{1/(n_1^{s_1} \cdots n_k^{s_k})\} \] with weight \( w = \sum_{i=1}^{k} s_i \) and depth \( k \) and for Euler sums of the form
\[ \sum_{n_1 > n_2 > \ldots > n_k > 0} (\epsilon_1^{n_1} \cdots \epsilon_k^{n_k})/(n_1^{s_1} \cdots n_k^{s_k}) \] with signs \( \epsilon_i = \pm 1 \). Notably, we achieve explicit proven reductions of all MZVs with weights \( w \leq 22 \), and all Euler sums with weights \( w \leq 12 \), to bases whose dimensions, bigraded by weight and depth, have sizes in precise agreement with the Broadhurst–Kreimer and Broadhurst conjectures. Moreover, we lend further support to these conjectures by studying even greater weights (\( w \leq 30 \)), using modular arithmetic. To obtain these results we derive a new type of relation for Euler sums, the Generalized Doubling Relations. We elucidate the "pushdown" mechanism, whereby the ornate enumeration of primitive MZVs, by weight and depth, is reconciled with the far simpler enumeration of primitive Euler sums. There is some evidence that this pushdown mechanism finds its origin in doubling relations. We hope that our data mine, obtained by exploiting the unique power of the computer algebra language \( \text{form} \), will enable the study of many more such consequences of the double-shuffle algebra of MZVs, and their Euler cousins, which are already the subject of keen interest, to practitioners of Quantum Field Theory, and to mathematicians alike.

Crown Copyright © 2009 Published by Elsevier B.V. All rights reserved.

1. Introduction

Multiple Zeta Values (MZVs) and Euler sums [1–3] have been of interest to mathematicians [1,4–7] and physicists [8] for a long time. One place in physics in which they are important is perturbative Quantum Field Theory. The interest became even larger when higher order calculations in Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD) started to need the multiple harmonic sums \( S_{\vec{c}}(N) \) [9–11]. Euler sums are obtained as the limit \( N \to \infty \) of the related multiple sums
\[ \zeta_{\vec{c}} = \sum_{k=1}^{\infty} \frac{(\sigma(b))^k}{k!} Z_{\vec{a}}(k-1), \] (1.1)
with \( \vec{c} = (b, \vec{a}), b, a_i \in \mathbb{Z} \) and
\[ Z_{b,\vec{a}}(N) = \sum_{k=1}^{N} \frac{(\sigma(b))^k}{k!} Z_{\vec{a}}(k-1), \] (1.2)
with \( \sigma(b) = \text{sign}(b) \). Euler sums for which all indices are positive are called Multiple Zeta Values. Euler sums and MZVs with the first index \( b = 1 \) diverge, but will be included symbolically in the following, for convenience. Their degree of divergence can be uniquely traced...
back to a polynomial in the single harmonic sum $S_1(\infty) = \sum_{N=\infty}^{\infty} \sum_{k=1}^{N} \frac{1}{k}$ shown later in the text. We call the number of indices of the Euler sums and MZVs their depth $d$ and

$$w = \sum_{k=1}^{d} |c_k|$$

(13)

their weight.

The number of Euler sums, resp. MZVs, up to a given weight $w$ grows rapidly and amounts to $2 \cdot 3^w - 1$ and $2^w - 1$, respectively. A central question thus concerns to find all the relations between the Euler sums, resp. MZVs for fixed weight and depth, and even more importantly, new relations between MZVs at the one hand and Euler sums on the other hand, and the corresponding bases. Besides weight and depth, another degree of freedom, being discussed later, the pushdown $p$, quantifies the relation between MZVs and Euler sums. The way to view MZVs, embedded into Euler sums, dates back to Broadhurst [12], who conjectured the counting of basis elements at fixed $(w,d)$. The corresponding conjecture for the MZVs is due to Broadhurst and Kreimer [13]. For the number of basis elements for MZVs of a given weight, without regard to depth, an upper bound has been proven in [14]. This coincides with the result obtained by summing the numbers conjectured in [13] over all depths at a fixed weight.

The relations between MZVs and Euler sums in Ref. [12] are conjectured using algorithms for integer relations as PSLQ [15] and LLL [16] which use representations based on a large number of digits.

It is well known that MZVs obey shuffle- and stuffle-relations. This is due to their representation in terms of Poincaré iterated integrals [17] at argument $x=1$, which are harmonic polylogarithms [18] on the one hand, and harmonic sums [9–11] on the other hand. The former quantities obey a shuffle- the latter a quasi-shuffle algebra, i.e. shuffling with “stuff” from polynomials of harmonic sums of lower weight. Currently no other independent relation is known between MZVs. The Euler sums are also related by both the shuffle- and stuffle-relations, where now also negative indices occur to indicate alternating sums. However, these relations are not sufficient to obtain the minimal set of basis elements as being conjectured in [12]. Starting with $w=8$ it requires the doubling relation and with $w=11$ generalized doubling relations derived in the present paper. Beginning with $w=12$ relations occur, which allow to express MZVs of a given depth in terms of Euler sums of a lesser depth. Part of these relations have been conjectured in the past using integer relations [12,19]. A main objective of the present paper is to prove these relations applying computer algebra methods and to find relations of this type in a more systematic way.

We investigate the Euler sums to $w=12$ completely, deriving basis-representations for all individual values in an explicit analytic calculation. For the MZVs the same analysis is being performed up to $w=22$. To $w=24$ we checked the conjectured size of the basis using modular arithmetic. Under the further conjecture that the basis elements can be chosen out of MZVs of depth $d \leq w/3$ we confirm the conjecture up to $w=26$. Furthermore, the following runs at limited depth, using modular arithmetic keeping the highest weight terms only, were performed: $d=7$, $w=27$; $d=6$, $w=28$; $d=7$, $w=29$; $d=6$, $w=30$. For the Euler sums complete results were obtained for $d \leq 3$, $w=29$; $d \leq 4$, $w=22$; $d \leq 5$, $w=17$ and for $d \leq 3$, $w=51$; $d \leq 4$, $w=30$; $d \leq 5$, $w=21$; $d \leq 6$, $w=17$ using modular arithmetic neglecting products of lower weight. The conjectures on the number of basis elements w.r.t. $(w,d)$ were verified in all these cases. The results of our analysis are made available in the Multiple Zeta Value Data Mine [20], to allow users to search for yet undiscovered relations.

The paper is organized as follows. In Section 2 we summarize basic notations and the well-known relations between Euler sums. A novel type of relations, the generalized doubling relations, is derived in Section 4. There we also discuss its impact in finding the basis elements at a given weight $w$ and depth $d$. In Section 5 an outline is given on the details of the computer algebra code, which allowed to derive the basis-representations of the MZVs and Euler sums. Details on the running for the different cases are reported in Section 6. The results are stored in the Multiple Zeta Value Data Mine, which is described in Section 7. To establish the solution of the problems dealt with in the current project required some new features of FORM [21] and TFORM [22], which are described in Section 8. In Section 9 we briefly review the status achieved by other groups and present first results of the analysis. In particular a series of conjectures made in the mathematical literature are confirmed within the range explored in the present study. Here we discuss also particular choices for the respective bases. An interesting aspect representing MZVs by Euler sums concerns the so-called pushdowns, i.e. the representation of a MZV of a given depth $d$ with Euler sums of depth $d'$ with $d' < d$. These are studied in Section 10 in which we also introduce a new kind of object, the $A_2$-functions. They play a key role in representing a class of Euler sums. Some more special Euler sums are studied in Section 11. Section 12 contains the conclusions and an outlook. In Appendices A–C we provide different basis representations and discuss the pushdowns in more detail.

2. Basic formalism

In the following we work with three types of objects, the finite nested harmonic $S_2$-sums, $S_3$-sums, both at argument $N \in \mathbb{N}$, and the harmonic polylogarithms $H_n$ at argument $x$, $0 \leq x \leq 1$. They all can be used to define the MZVs and the Euler sums in the limit $N \rightarrow \infty$ and $x=1$, respectively. We generally consider the case of colored objects corresponding to $n=2$, i.e. numerator weights with $|\pm 1|^k$, i.e. polylogarithms of square root of unity.

The harmonic $S$-sums are defined by

$$S_n(0) = 0,$$

$$S_n(N) = \sum_{k=1}^{N} \frac{(\sigma(b))^k}{k|b|},$$

\cite{2} Conjectures for fixed weight are due to Zagier [2] and probably also independently due to Drinfel'd, Goncharov and Kontsevich.

\cite{3} It goes without saying that also the Euler sums are covered here.
\[ S_{\vec{a},\vec{b}}(N) = \sum_{k=1}^{N} \frac{(\sigma(b))^k}{k^{\vec{b}}} S_{\vec{a}}(k). \]  

(2.1)

In this form these sums are usually used by physicists. In particular results in QCD [23–26] are expressed in terms of these objects.  

Next there are the Z-sums. They are defined in (1.2). These are of course very similar to the S-sums and it is straightforward to convert from one notation to the other. The Z-sums are mostly used by mathematicians. In the limit \( N \to \infty \) and when \( \sigma(b) = 1 \) for all \( b \) they define the Multiple Zeta Values (MZVs):  

\[ \zeta_{\vec{a}} = \lim_{N \to \infty} \ Z_{\vec{a}}(N). \]  

(2.2)

When we allow \( \sigma(b) \) to take the values +1 or −1 and we take the limit \( N \to \infty \) we speak of Euler sums.

Finally, there are the harmonic polylogarithms, which we will also call \( H \)-functions. We consider the alphabets  

\[ \mathfrak{h} = \{ 0, 1, -1 \} \quad \text{and} \quad \mathfrak{f} = \{ 1/x, 1/(1-x), 1/(1+x) \}, \]  

(2.3)

which define the elements of the index set of the harmonic polylogarithms and the functions in the iterated integrals, respectively. Let \( \vec{a} = \{ m_1, \ldots, m_k \} \), \( m_i, b \in \mathfrak{h}, k \geq 1 \), then  

\[ H_{\vec{b},\vec{a}}(x) = \int_0^x dz \ f_0(z) H_{\vec{a}}(z), \]

\[ f_0(z) = 1/z, \]

\[ f_1(z) = 1/(1-z), \]

\[ f_{-1}(z) = 1/(1+z), \]

\[ H_0(x) = \log(x), \]

\[ H_1(x) = -\log(1-x), \]

\[ H_{-1}(x) = \log(1+x). \]  

(2.4)

The sums to infinity and the \( H \)-functions at unity are all related and can be readily transformed into each other. For some applications it is most convenient to work with one set of objects and for others other objects may be more useful. For reasons being explained later our computer programs work mostly with \( H \)-functions at unity.

A first aspect to note is that the index fields of the sums and the functions are of a different nature. This can be seen by introducing the notation in which the index \( n \) in the sums can alternatively be written as \( n - 1 \) zeroes followed by a one and \( -n \) is written as \( n - 1 \) zeroes followed by a minus one. In the \( H \)-functions we can absorb alternatively the zeroes in the nonzero number to their right by raising its absolute value by one for each zero being absorbed. This leaves only the rightmost zeroes. Hence:

\[ S_{-3,4}(N) = S_{0,0,0,0,0,1}(N), \]

\[ Z_{2,-5}(N) = Z_{0,1,0,0,0,0,0,1}(N), \]

\[ H_{0,1,-1,0,0,-1,0,0}(x) = H_{2,-1,-1,0,0}(x). \]  

(2.5)

The notation in terms of the 0, ±1 we call the (iterated) integral notation. The natural notation of the sums we call the (nested) sum notation.

Reference to the alphabet \( \mathfrak{h} \) allows us to count the number of objects and to classify them. The number of indices in this integral notation is called the weight of the sum or the function. For a given weight \( w \) there are \( 2 \cdot 3^{w-1} \) sums and \( 3^w \) \( H \)-functions. When the sums are written in the original sum notation, the number of indices indicates the number of nested sums. This is also called the depth of the sum. When there are no trailing zeroes in the \( H \)-functions we can introduce the depth in the same way. Because of algebraic relations we can express the functions with trailing zeroes as products of powers of \( \log(x) \) and \( H \)-functions with fewer indices [18,37]. In that case the concept of depth can be used in a similar way as with the sums.

For any argument \( x \neq 1 \) the \( H \)-functions form a shuffle algebra:

\[ H_{\vec{p}}(x) H_{\vec{q}}(x) = \sum_{\vec{r} \in \mathfrak{f} \sqcup \mathfrak{h}} H_{\vec{r}}(x), \]  

(2.6)

where \( \vec{p} \sqcup \vec{q} \) denotes the shuffle product, cf. e.g. [37], and \( p_i, q_i \in \mathfrak{h} \). When \( x = 1 \) \( H \)-functions for which the first index is one are divergent. It is however possible to express them in terms of a single divergent object and other finite terms in a consistent way. The only thing that breaks down is that there are correction terms to the shuffle relations when both objects in the left-hand side are divergent, see also Ref. [18]. Because the number of non-zero indices remains the same during the shuffle operation, we call it depth preserving.

---

4 The class of Euler sums is known to be too small in general to represent all Feynman diagrams for no-scale processes in scalar field theories, but have to be extended in higher orders [27–30]. This will apply also for field theories as QCD and QED. Feynman-integrals are periods [31] if all ratios of Lorenz invariants and masses have rational values [32].

5 Special cases are the classical polylogarithms [33] and the Nielsen polylogarithms [34]. Generalizations of harmonic polylogarithms are found in [35,36].
For general argument \( N \) the sums form a stuffle algebra [37]. This is a general property of sums which we show here for a double sum:

\[
\sum_{i=1}^{N} \sum_{j=1}^{N} = \sum_{i=1}^{N} \sum_{j=1}^{N} + \sum_{j=1}^{N} \sum_{i=1}^{N} - \sum_{i=1}^{N} \sum_{j=1}^{N} - \sum_{j=1}^{N} \sum_{i=1}^{N} + \sum_{i=1}^{N} \sum_{j=1}^{N} + \sum_{j=1}^{N} \sum_{i=1}^{N} .
\] (2.7)

The diagonal terms give extra ‘stuff’ beyond the normal shuffling in the natural notation for the sums. Even though the diagonal terms add terms usually the stuffle relations have fewer terms because most of the time some of the indices will have an absolute value greater than one. We write in terms of \( S \)- or \( Z \)-notation:

\[
S_m(N)S_n(N) = S_m,n(N) + S_{m,n}(N), \quad S_m(N)S_{n,k}(N) = S_{m,n,k}(N) + S_{m,n,k}(N) - S_{m,n,k}(N) - S_{m,n,k}(N), \quad Z_m(N)Z_n(N) = Z_{m,n}(N) + Z_{n,m}(N) + Z_{m,n}(N) + Z_{m,n}(N). \]

Here the operator \( \& \) is defined by

\[
m \& n = \sigma (m) \sigma (n) (|m| + |n|) = \sigma_m + \sigma_n. \]

(2.12)

The above algebraic relations can be used to bring an expression with many harmonic polylogarithms or harmonic sums into a standard form. For evaluation, however, it is often useful to work it the other way and reduce the number of objects at the highest weight in favor of products of objects with a lower weight which are easier to evaluate. For this the theory of Lyndon words [38] applies, but especially with the stuffles the extra terms which have the same weight but a lower depth have to be taken along and make things considerably more involved than pure shuffles.

A \( k \)-ary Lyndon word of length \( n \) is an \( n \)-letter concatenation product over an alphabet of size \( k \), which, observing lexicographical ordering is smaller than all its suffixes. Equivalently, it is the unique minimal element in the lexicographical ordering of all its cyclic permutations. The uniqueness implies that a Lyndon word is aperiodic. So it differs from any of its non-trivial rotations. In our case we will usually replace minimal by maximal when we form Lyndon words of indices of MZVs or Euler sums. That is, we will put the larger indices to the left. One could also say that the concept of greater than is defined in a special way inside the alphabet. The practical advantage is that this guarantees that none of the MZVs of which the index string forms a Lyndon word is divergent.

When we use the stuffle relations to simplify the set of objects at a given weight, we can arrange that they are used in such a way that they never raise the value of the depth parameter. Some terms will have a lower value for the depth. Therefore we call the stuffles potentially depth lowering.

When we consider the sums to infinity there are two classes of extra relations worth mentioning. The first is the ‘rule of the triangle’ which is based on

\[
\lim_{N \to \infty} \sum_{i=1}^{N} \sum_{j=1}^{N} = \lim_{N \to \infty} \sum_{i=1}^{N} \sum_{j=1}^{N} + \lim_{N \to \infty} \sum_{j=1}^{N} \sum_{i=1}^{N} .
\] (2.13)

For most sums the second term will give a limit that goes to zero with at least one power of \( 1/N \), possibly multiplied by powers of \( \log(N) \). This system can be generalized to the product of any pair of sums and it can be proven that the limit of the second term vanishes when at least one of the sums in the left-hand side is finite [10]. When both are divergent it is possible to work out which extra terms are needed. Because the sums of the first term in the right-hand side can be worked out, even in the most general case, the above gives us an extra algebraic relation for the sums to infinity. These relations are depth preserving.

When we consider the \( H \)-functions at unity, it is easy to see that they can be written as nested sums to infinity of the same variety as the \( Z \)-sums or the \( S \)-sums. Hence they now obey also the stuffle algebra. And it can be shown that the ‘rule of the triangle’ is no more than the equivalent of the shuffle algebra for the \( H \)-functions, with the same restrictions for the double divergent terms.

The next set of relations is easy to see for finite sums:

\[
S_m(N) = \sum_{i=1}^{N} \frac{1}{i^m} = \sum_{i=1}^{2^m} \left( \frac{1}{(2i)^m} \right) = \sum_{i=1}^{2^m} \frac{2^{m-1}(1 + (-1)^i)}{i^m}
\]

\[= 2^{m-1} \left[ S_m(2N) + S_{\overline{m}}(2N) \right], \]

(2.14)

which generalizes into

\[
S_{n_1,\ldots,n_p}(N) = 2^{n_1+\cdots+n_p-p} \sum_{\sigma} S_{\pm n_1,\ldots,\pm n_p}(2N).
\]

(2.15)

Here the sum is over all \( 2^p \) plus/minus combinations. These relations are called the ‘doubling relations’. For finite sums with \( n_1 \neq 1 \) these relations can be used directly. In the case that divergent sums are involved there are again correction terms.

The equivalent formula for the \( H \)-functions is obtained by looking at \( H_d(x^2) \) and noticing that at \( x = 1 \) this is the same as \( H_d(x) \). In that case we have

\[
H_{1,0.1}(x^2) = 2 \left[ H_{1.0,1}(x) - H_{1.0,-1}(x) - H_{1.0,-1}(x) + H_{1.0,-1}(x) \right],
\]

(2.16)

which generalizes to any number of indices. The rule is that the factor is identical to \( 2^m \) in which \( m \) is the number of zeroes in the indices, and each one in the left-hand side gives a doubling of terms in the right-hand side: one term with a corresponding 1 and one
with a corresponding $-1$ and an extra overall minus sign. In the left-hand side one cannot have negative indices. Again one should be careful with the divergent functions.

Divergences are expressed in terms of the object $S_1(\infty)$. In most cases one can use this as a regular symbol and take it along in the equations and expressions. Unless we mention the problems explicitly, one can exchange limits and sums when this object is combined with finite sums. The reason is that our finite sums converge faster than that this object diverges. A problem occurs when we use the doubling formula on it. We find:

$$S_1(\infty) = S_1(2\infty) + S_{-1}(2\infty) = S_1(2\infty) - \log(2),$$

which just shows that the divergence of $S_1(\infty)$ is logarithmic, since

$$S_1(N) = \ln(N) + \gamma_E + \frac{1}{2N} + \frac{1}{12N^2} + O\left(\frac{1}{N^3}\right),$$

cf. [25]. One can however use the stuffle relations on these objects. This allows one in principle to express the divergent sums in terms of products of $S_1(\infty)$ and finite sums as in

$$S_1(N)S_{m,n}(N) = S_{1,m,n}(N) + S_{m,1,n}(N) + S_{m,n}(N) - S_{n,1,n}(N) - S_{m,n}(N).$$

If we assume $m \neq 1$ this allows us to express the divergent sum $S_{1,m,n}(\infty)$ the way we want it. Similarly one can now look at stuffles of $S_1$ to determine $S_{1,1}$ and then look at stuffles of $S_{1,1}(N)$ with finite sums. In the programs we give $S_1(\infty)$ the name $\text{SiInf}$ which, due to the above, can be treated as a regular symbol.

Because we have two shuffle products — the shuffle-algebra is a quasi-shuffle algebra [39] — we can equate the result of the stuffle product of two objects with the result of the shuffle product of the same two objects. The resulting relation is called a double-shuffle relation and contains only objects of the same weight. These relations have been used in a number of calculations. For our type of calculations they are, however, not suitable. We will use the stuffle and the shuffle relations individually. This will allow a better optimization of the algorithms.

The concept of duality is very useful and allows us to roughly half the number of objects that need to be computed. The duality relation is defined in the integral notation using harmonic polylogarithms at one. It states that if we have a MZV and we reverse the order of its relations they are, however, not suitable. We will use the stuffle and the shuffle relations individually. This will allow a better optimization of the algorithms.

The concept of duality is very useful and allows us to roughly half the number of objects that need to be computed. The duality relation is defined in the integral notation using harmonic polylogarithms at one. It states that if we have a MZV and we reverse the order of its indices while at the same time transforming zeroes into ones and ones into zeroes the new object has the same value as the original. An example of this duality is the relation

$$H_{0,1,0,1,1,1,1,1} = H_{0,0,0,0,0,1,0,1},$$

In mathematics one traditionally considers this duality in sum notation. In that case, for a sequence

$$I = (p_1 + 1, \{1\}_{q_1-1}, p_2 + 1, \{1\}_{q_2-1}, \ldots, p_k + 1, \{1\}_{q_k-1})$$

there is a dual sequence

$$\tau(I) = (q_k + 1, \{1\}_{p_k-1}, q_{k-1} + 1, \{1\}_{p_{k-1}-1}, \ldots, q_1 + 1, \{1\}_{p_1-1}).$$

The duality theorem [2] states

$$\zeta_I = \zeta_{\tau(I)}.$$  

It was conjectured in [40] and is easily proven by the transformation $x \rightarrow 1 - t$ of the corresponding iterated integrals.

Because for even weights there are some elements that are self-dual this does not divide the number of terms exactly by two. Considering that we do not have to consider the divergent objects we have $2^{w-3}$ relevant objects when $w$ is odd and $2^{w-3} + 2^{w/2-2}$ relevant objects when $w$ is even.

For Euler sums the equivalent transformation is more complicated due to the three letter alphabet. It is obtained by studying the transformation

$$x \rightarrow \frac{1 - t}{1 + t},$$

in the integral representation. Its effect is that given the alphabet

$$A = 0 \leftarrow \frac{1}{x},$$

$$B = 1 \leftarrow \frac{1}{1 - x},$$

$$C = -1 \leftarrow \frac{1}{1 + x},$$

and a string of letters from this alphabet as indices of an Euler sum $H$, the 'dual expression' is obtained by reverting the string of letters and making the replacement

$$A \rightarrow B \oplus C,$$

$$B \rightarrow A \oplus C,$$

$$C \rightarrow C.$$
The addition and subtraction operators here mean that for each such transformation there will be a doubling of the number of terms, one with the first letter and the other with the other letter. The sign-operator $\oplus$ ($\ominus$) refers to the sign of the complete term. Because these relations can both raise and lower the depth of a term we call them depth mixing.

We have tested that this transformation does add something new beyond what the stuffles and the shuffles give us. In particular, when one derives equations for all sums at a given weight, they can be used to replace the doubling and the Generalized Doubling Relations (GDRs), see Section 4. We have tested this to weight $w = 12$. Unfortunately they cannot be used when the concept of depth of the sums is important and hence we have not used these equations in our programs.

A generalization of the Riemann $\zeta$-function is Hurwitz' $\zeta$-function [6,41]:

$$\zeta(n, a) = \sum_{k=1}^{\infty} \frac{(\text{sign}(n))^k}{(k+a)^n},$$

(2.27)

which can be extended to generalized Euler sums analogous to (1.1). Since $a$ is a real parameter, one may differentiate $\zeta(\ell, a)$ w.r.t. $a$ and seek for new relations. We investigated this possibility, but did not find new relations beyond those quoted above.

When we are discussing bases into which we write the MZVs and the Euler sums we recognize two types of basis:

**Definition.** A basis of a vector space of all Euler sums or MZVs at a given weight $w$ is called a Fibonacci basis.

**Definition.** A basis of the ring of all Euler sums or MZVs at a given weight $w$ is called a Lyndon basis if all its elements have an index field that forms a Lyndon word.

In a Fibonacci basis all basis elements are nested sums of the same weight. The name derives from the observation that the size of such a basis for the Euler sums seem to follow a Fibonacci rule [42]. Also the MZVs seem to follow the rule that the total number of their basis elements follow the Fibonacci-like Padovan numbers [43], see Appendix A.

In a Lyndon basis we write in the complete basis as many elements as possible as products of lower weight basis elements and what remains is the Lyndon basis. Simultaneously we require the index field to form a Lyndon word. Sometimes a Lyndon basis can be formed from a Fibonacci basis by just selecting the Lyndon words from it. The number of basis elements in the case of MZVs is counted by a Witt-type relation [44] based on the Perrin numbers [45]. In the case of the Euler sums the corresponding relation relies on the Lucas numbers [46], see Appendix A. Any other basis we will call a mixed basis.

We will usually try to arrange the Lyndon bases in such a way that they are ‘minimal depth’. This means that if an element can be expressed in terms of objects with a lower depth, it cannot be a member of the basis. Details on a variety of bases are given in Appendix A. The complete basis we actually selected for the MZVs is presented in Appendix B.

### 3. Conjectures on bases at fixed weight and depth

Broadhurst [12] and Broadhurst and Kreimer [13] formulated conjectures on the size of the basis for Euler sums and MZVs, respectively, which we summarize in the following.

Euler sums $\zeta_2$ at given weight and depth $w$, $d$ are called independent if there exists no relation between them, cf. Sections 2, 4. The elements of the basis through which all Euler sums can be represented in terms of polynomials are called primitive. The numbers of independent and primitive sums at a given weight are fixed, while different basis representations may be chosen.

Let $E_{w,d}$ be the number of independent Euler sums at weight $w > 2$ and depth $d$ that cannot be reduced to primitive Euler sums of lesser depth and their products. Thus we believe that $E_{1,1} = 1$, since there is no known relationship between $\zeta_3$, $\pi^2$ and $\ln(2)$. It is rather natural to guess that $E_{w,d}$ is given by a filtration of the coefficients of powers of $x$ and $y$ in the expansion of $1/(1 - xy - x^2)$, i.e. that

$$\prod_{w \geq 2} \prod_{d > 0} (1 - x^w y^d)^{E_{w,d}} \equiv \frac{1 - xy - x^2}{(1 - xy)(1 - x^2)} = 1 - \frac{x^3 y}{(1 - xy)(1 - x^2)},$$

(3.1)

It is then easy to obtain $E_{w,d}$ by Möbius transformation of the binomial coefficients in Pascal’s triangle. Let

$$T(a, b) = \frac{1}{a + b} \sum_{d|a, b} \mu(d) \frac{(a/d + b/d)!}{(a/d)! (b/d)!},$$

(3.2)

where the sum is over all positive integers $d$ that divide both $a$ and $b$ and the Möbius function is defined by

$$\mu(d) = \begin{cases} 
1 & \text{when } d = 1, \\
0 & \text{when } d \text{ is divisible by the square of a prime,} \\
(-1)^k & \text{when } d \text{ is the product of } k \text{ distinct primes.} 
\end{cases}$$

(3.3)

When $w$ and $d$ have the same parity, and $w > d$, one obtains from (3.1)

$$E_{w,d} = T\left(\frac{w - d}{2}, d\right).$$

(3.4)

With the exception of $\ln(2)$ and $\zeta_2$, which act as the seeds $xy$ and $x^2$, all elements of the basis are thereby conjecturally enumerated. In this paper we provide extensive evidence to support conjecture (3.1).

Now let $D_{w,d}$ be the number of independent MZVs at weight $w > 2$ and depth $d$ that cannot be reduced to primitive MZVs of lesser depth and their products. Thus we believe that $D_{3,2} = 1$, since there is no known relationship between the double sum $Z_{3,3} =$
\[ \sum_{m,n>0} \frac{1}{(m^2 n^2)} \] and single sums or their products. It is tempting to guess to that \( D_{w,d} \) is generated by filtration of the expansion of \( 1/(1-x^2-x^3 y) \), seeded by \( x^2 \) and \( \zeta_3 \). But this is not the case, since the solution of the double-shuffle algebra at weight \( w = 12 \) leaves one quadruple sum undetermined, while the obvious guess would leave none. The conjecture \[13\] in this case is rather ornate, cf. Table 16.

\[ \prod_{w>2d>0} (1-x^w y^d)^{D_{w,d}} = 1 - \frac{x^2 y}{1-x^2} + \frac{x^{12} y^2 (1-y^2)}{(1-x^4)(1-x^6)} \] (3.5)

with a correction term whose numerator, \( x^{12} y^2 (1-y^2) \), ensures that \( D_{12,4} = 1 \) and \( D_{12,2} = 1 \), in agreement with the solution of the double-shuffle algebra. The denominator \( (1-x^4)(1-x^6) \) is then chosen to give \( D_{2m,2} = \lfloor (m-1)/3 \rfloor \) for the number of primitive double sums with weight \( 2m \). Conjecture (3.5) is impressively supported by the data mine.

Furthermore,

\[ \prod_{w>2d>0} (1-x^w y^d)^{M_{w,d}} = 1 - \frac{x^2 - x^3 y}{1-x^2} \] (3.6)

is the conjectured generating function of the basis elements \( M_{w,d} \) of the MZVs when expressed as Euler sums in a minimal depth representation, see Table 17.

4. Generalized doubling relations

Up to \( w = 10 \) the shuffle-, stuffle-, and doubling relations were sufficient to express the alternating Euler sums over a basis whose size is in accordance with the conjecture in Ref. [12]. This is not the case from \( w = 11 \) onwards. Therefore one has to seek a new kind of relations, which we derive in the following. Of course, when we derive all relations at a given weight we could use the relations of (2.26).

The fact that they are depth mixing makes them useless for calculations in which the concept of depth plays a role. Hence we need our new (depth lowering) relations anyway. We first present the derivation of this class of relations which we call Generalized Doubling Relations (abbreviated to GDRs) and discuss then their effect on the number of basis elements representing the Euler sums.

4.1. Derivation of the Generalized Doubling Relations

The only relations we could find thus far adding something new to the system are the depth 2 relations of Ref. [12]. They are based on partial fractioning in two different ways. One way is:

\[ \frac{1}{(2i+j)(j)} = \frac{1}{(2i+2j)(2i+j)} + \frac{1}{(2i+2j)(j)} \] (4.1)

We can take out the factor two and in the first term the 2i is taken care of by changing the summation over \( i \) into a summation over the even numbers by including a factor \((1+(-1)^i)/2\), which introduces negative indices in some Euler sums. In the other way we use the more regular form

\[ \frac{1}{(2i+j)(j)} = \frac{1}{(2i)(j)} - \frac{1}{(2i)(2i+j)} \] (4.2)

Together these partial fractions produce new types of relations.

Here we will give the new set of relations and their derivation. We will work with the \( Z \)-sums. The reason is a particularly handy representation of these sums to infinity \[40,47\]:

\[ Z_{m_1,\ldots,m_p}(\infty) = \sum_{i_1>i_2>\cdots>i_p>0} \frac{\sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_p}}{m_1^{i_1} m_2^{i_2} \cdots m_p^{i_p}} \]

\[ = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \cdots \sum_{x_p=1}^{\infty} \frac{\sigma_{x_1+x_2+\cdots+x_p} \sigma_{x_2+x_3+\cdots+x_p} \cdots \sigma_{x_p}}{(x_1 + x_2 + \cdots + x_p)^{i_1} (x_2 + \cdots + x_p)^{i_2} \cdots (x_p)^{i_p}} \] (4.3)

in which we take \( n_i = |m_i| \) and \( \sigma_i \) to be the sign of \( m_i \).

Let us start with the re-derivation of the equation for depth \( d = 2 \). Actually we do not reproduce it exactly, but we obtain a similar equation. Here we write for brevity \( Z(a,b) = Z_{a,b}(\infty) \). Throughout this section we assume that \( a, b, c \) and \( d \) are positive integers. We consider the following combination of \( Z \)-sums:

\[ E(a,b) = \frac{1}{2} \left( Z(a,b) + Z(-a,-b) \right) \]

\[ = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \frac{1}{(x_1 + x_2)^a x_2^b} \frac{1 + (-1)^x_1}{2} = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \frac{1}{(2x_1 + x_2)^a x_2^b} \]

\[ = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \left[ \sum_{i=0}^{a} A_i^{(a,b)} \frac{1}{(2x_1 + x_2)^a + b - i} (2x_1 + x_2)^i \right] + \sum_{i=1}^{b} B_i^{(a,b)} \frac{1}{(2x_1 + x_2)^a + b - i} \]

\[ = \sum_{i=1}^{a} A_i^{(a,b)} 2^{-a-b} \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \frac{1}{(x_1 + x_2)^a + b - i} (2x_1 + x_2)^i \]
\[ + \sum_{i=1}^{b} B_{i}^{(a,b)} 2^{i-a-b} Z(a + b - i, i) \]
\[ = \sum_{i=1}^{b} B_{i}^{(a,b)} 2^{i-a-b} Z(a + b - i, i) + \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} \sum_{x_{1}=1}^{\infty} \sum_{x_{2}=x_{1}+1}^{\infty} \frac{1}{(x_{1} + x_{2})^{a+b-i}} \]
\[ = \sum_{i=1}^{b} B_{i}^{(a,b)} 2^{i-a-b} Z(a + b - i, i) + \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} Z(i, a + b - i) \]
\[ - \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} \sum_{x_{1}=1}^{\infty} \sum_{x_{2}=x_{1}+1}^{\infty} \frac{1}{(x_{1} + x_{2})^{a+b-i}} \]
\[ = \sum_{i=1}^{b} B_{i}^{(a,b)} 2^{i-a-b} Z(a + b - i, i) + \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} Z(i, a + b - i) \]
\[ - \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} \sum_{x_{1}=1}^{\infty} \sum_{x_{2}=x_{1}+1}^{\infty} \frac{1}{(x_{1} + x_{2})^{a+b-i}} \]
\[ = \sum_{i=1}^{b} B_{i}^{(a,b)} 2^{i-a-b} Z(a + b - i, i) + \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} Z(i, a + b - i) \]
\[ - \sum_{i=1}^{a} A_{i}^{(a,b)} \frac{1}{2} \left( Z(i, a + b - i) + Z(i, -(a + b - i)) \right) - \frac{(a + b - 1)!}{(a - 1)! b!} 2^{-a-b} Z(a + b). \tag{4.4} \]

with
\[ A_{i}^{a,b} = \frac{(a+b-i-1)!}{(a-i)!(b-1)!}, \tag{4.5} \]
\[ B_{i}^{a,b} = \frac{(a+b-i-1)!}{(b-i)!(a-1)!}. \tag{4.6} \]

Actually there is a slight problem with the above derivation. At two points we changed the summation range. Once from \( \infty \) to \( \infty/2 \) and once from \( \infty \) to \( 2\infty \). This causes no problems if the sum is finite, but for the divergent sums this needs a correction term. The second case is harmless as it concerns only an inner sum, the step in which \( -(1)^{x_{2}} \) is introduced. But the first case, in the very first step of the derivation, needs a correction term. Hence the full formula becomes:

\[ E(a, \sigma_{b}) = \frac{1}{2} Z(a, \sigma_{b}) + Z(-a, -\sigma_{b}) \]
\[ = \frac{1}{2} \delta(a - 1) Z(-1) Z(\sigma_{b}) - \frac{1}{2} \delta(a - 1) \delta(\sigma_{b} - 1) Z(-2) \]
\[ + \sum_{i=1}^{b} B_{i}^{(a,b)} 2^{i-a-b} Z(a + b - i, i, \sigma_{b}) + \sum_{i=1}^{a} A_{i}^{(a,b)} 2^{i-a-b} Z(\sigma_{b} i, a + b - i) \]
\[ - \sum_{i=1}^{a} A_{i}^{(a,b)} \frac{1}{2} \left( Z(\sigma_{b} i, \sigma_{b} (a + b - i)) + Z(\sigma_{b} i, -\sigma_{b} (a + b - i)) \right) \]
\[ - \frac{(a + b - 1)!}{(a - 1)! b!} 2^{-a-b} Z(a + b). \tag{4.7} \]

Here also the signs on the indices \( a \) and \( b \) are included which is only a very mild complication in the derivation. The function \( \delta(m) \) is one when \( m \) is zero and zero otherwise. The \( \sigma \)-variables have a value that is either +1 or -1 and indicate non-alternating and alternating.
sums. Due to the symmetry of the starting formula a sign on the first variable is not necessary. If we put it anyway in the form of \( \sigma_a, \sigma_b \) will have to be replaced by \( \sigma_a \sigma_b \) in the right-hand side.

It is quite relevant to take these \( \sigma \) factors along. Although they are usually not needed to get a complete coverage of depth \( d = 2 \) sums, in the case of greater depth sums they are necessary.

The above derivation shows basically all techniques we need for the derivation of the greater depth formulas. In the sequel we will only carry the \( \sigma \) factors that survive conditions posed during the derivation.

The derivation of the depth 3 formula follows a similar but slightly more complicated path. Again, we first omit the signs of the indices and the correction terms for divergent integrals when we double or half the summation range. Then we present the complete formula. In the derivation we will be a bit shorter this time as the techniques are all similar to what we have shown above.

\[
E(a, b, c) = \frac{1}{2} (Z(a, b, c) + Z(-a, -b, c))
\]

\[
= \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \sum_{x_3=1}^{\infty} \frac{1}{(x_1 + x_2 + x_3)^{a+b+1}} \frac{1}{(2x_1 + 2x_2 + 2x_3)^{a+b+1}}
\]

\[
= \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \sum_{x_3=1}^{\infty} \sum_{a=1}^{b} A_i^{(a,b)} 2^{2-a-b-i} Z(a + b - i, i, c)
\]

\[
= \sum_{a=1}^{b} A_i^{(a,b)} 2^{2-a-b-i} Z(a + b - i, i, c)
\]

\[
= \sum_{a=1}^{b} A_i^{(a,b)} 2^{2-a-b-i} K_{1}^{(1)}(a + b - i, i, c)
\]

\[
= \sum_{a=1}^{b} A_i^{(a,b)} 2^{2-a-b-i} K_{2}^{(1)}(a + b - i, i, c).
\]

(4.8)

with the \( K \) functions given below. The full formula becomes

\[
E(a, \sigma_b, \sigma_c) = \frac{1}{2} (Z(a, \sigma_b, \sigma_c) + Z(-a, -\sigma_b, \sigma_c))
\]

\[
= \frac{1}{2} Z(-1)Z(\sigma_b, \sigma_c)\delta(a-1)
\]

\[
- \frac{1}{2} Z(-2)Z(\sigma_c, \sigma_b)\delta(a-1)\delta(\sigma_b-1)
\]

\[
+ \frac{1}{2} Z(-3)\delta(a-1)\delta(\sigma_b-1)\delta(\sigma_c-1)
\]

\[
+ \sum_{a=1}^{b} B_i^{(a,b)} 2^{-a-b+i} Z(a + b - i, \sigma_b i, \sigma_c)
\]

\[
+ \sum_{a=1}^{b} A_i^{(a,b)} 2^{-a-b+i} Z(\sigma_b i, a + b - i, \sigma_c)
\]

\[
- \sum_{a=1}^{b} A_i^{(a,b)} 2^{-a-b+i} K_{1}^{(1)}(a + b - i, \sigma_b i, \sigma_c)
\]

\[
- \sum_{a=1}^{b} A_i^{(a,b)} 2^{-a-b+i} K_{2}^{(1)}(\sigma_b i, a + b - i, \sigma_c).
\]

(4.9)

The correction terms with the \( \delta \)-functions are due to the halving of the summation range in the first step. The \( K \)-functions are given by
The first index is one. Because the second index cannot be one in that case, we only need one correction term.

(4.10)

\[
K_1^{(1)}(\sigma_b, b, \sigma_c) = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \frac{\sigma_b^{2x_1 + x_2} \sigma_c^{x_2}}{(x_1 + x_2)^{a} (2x_1 + x_2)^{b} x_2^{c}}
\]

\[
= (-1)^b \sum_{i=1}^{a} A_i^{(a, b)} 2^{-a-i} Z(i, \sigma_b \sigma_c (a + b + c - i))
\]

\[
+ Z(\sigma_b \sigma_c (a + b + c - 1))
\]

\[
(4.11)
\]

\[
K_2^{(1)}(\sigma_d a, b, \sigma_c) = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \sum_{x_3=1}^{\infty} \frac{\sigma_d^{x_1 + 2x_2 + x_3} \sigma_c^{x_3}}{(x_1 + x_2 + x_3)^{a} (x_2 + x_3)^{b} x_3^{c}}
\]

\[
= (-1)^c 2^{b-1} \sum_{i=1}^{c} B_i^{(b, c)} (-1)^i Z(\sigma_d a, (b + c - i), \sigma_c i)
\]

\[
+ Z(\sigma_d a, -(b + c - i), -\sigma_c i)
\]

\[
- (-1)^{c-1} 2^{b-1} \sum_{i=1}^{b} A_i^{(b, c)} Z(\sigma_d a, \sigma_c (b + c - i), i)
\]

At depth 4 the relation becomes yet a bit more complicated but the derivation follows exactly the same path. We start with applying the non-trivial partial fractioning and then we have to try to rewrite the results in terms of \(Z\)-functions by percolating the factors two to the right. As there is one more sum this takes another step and we get two layers of \(K\)-functions:

\[
E(a, \sigma_b b, \sigma_c c, \sigma_d d) = \frac{1}{2} \left( Z(a, \sigma_b b, \sigma_c c, \sigma_d d) + Z(-a, -\sigma_b b, \sigma_c c, \sigma_d d) \right)
\]

\[
= \frac{1}{2} Z(-1) Z(\sigma_b b, \sigma_c c, \sigma_d d) \delta(a - 1)
\]

\[
- \frac{1}{2} Z(-2) Z(\sigma_c c, \sigma_d d) \delta(a - 1) \delta(\sigma_b b - 1)
\]

\[
+ \frac{1}{2} Z(-3) Z(\sigma_d d) \delta(a - 1) \delta(\sigma_b b - 1) \delta(\sigma_c c - 1)
\]

\[
- \frac{1}{2} Z(-4) \delta(a - 1) \delta(\sigma_b b - 1) \delta(\sigma_c c - 1) \delta(\sigma_d d - 1)
\]

\[
+ b \sum_{i=1}^{a} B_i^{(a, b)} 2^{-a-b+i} Z(a + b - i, \sigma_b i, \sigma_c c, \sigma_d d)
\]

\[
+ a \sum_{i=1}^{a} A_i^{(a, b)} 2^{-a-b+i} Z(\sigma_b i, a + b - i, \sigma_c c, \sigma_d d)
\]

\[
- a \sum_{i=1}^{a} A_i^{(a, b)} 2^{-a-b+i} K_1^{(1)}(a + b - i, \sigma_b i, \sigma_c c, \sigma_d d)
\]

\[
- a \sum_{i=1}^{a} A_i^{(a, b)} 2^{-a-b+i} K_2^{(1)}(a + b - i, \sigma_b i, \sigma_c c, \sigma_d d)
\]

(4.12)
\[ K_2^{(1)}(\sigma_a, b, \sigma_c, \sigma_d) = \sum_{x_1=1}^{\infty} \cdots \sum_{x_4=1}^{\infty} \frac{\sigma_a^{x_1+2x_2+x_3+4} \sigma_b^{x_2+x_3} \sigma_c^{x_1} \sigma_d^{x_4}}{(x_1 + 2x_2 + x_3 + 4)^d(x_2 + x_3 + 4)^b(x_3 + x_4)^c} x_4^d \]

\[ = (-1)^c 2^{b-1} \sum_{i=1}^{c} B_i^{(b,c)}(-1)^i \left( Z(\sigma_a, b + c - i, \sigma_i, \sigma_d) \right) + Z(\sigma_a, -(b + c - i), -\sigma_i, \sigma_d) \]

\[ + (-1)^c \sum_{i=1}^{b} A_i^{(b,c)} 2^{b-i} Z(\sigma_a, \sigma_c(b + c - i), \sigma_i, \sigma_d) \]

\[ - (1)^c \sum_{i=1}^{b} A_i^{(b,c)} 2^{b-i} K_1^{(2)}(\sigma_a, \sigma_c(b + c - i), i, \sigma_d) \]

\[ - (1)^c \sum_{i=1}^{b} A_i^{(b,c)} 2^{b-i} K_2^{(2)}(\sigma_a, \sigma_c(b + c - i), i, \sigma_d) \]

\[ K_2^{(2)}(\sigma_a, \sigma_b, c, \sigma_d) = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \sum_{x_3=1}^{\infty} \frac{\sigma_a^{x_1+x_2+x_3} \sigma_b^{2x_2+x_3} \sigma_d^{x_3}}{(x_1 + 2x_2 + x_3)^d(x_2 + x_3)^b(x_3 + x_4)^c} x_4^d \]

\[ = (-1)^d 2^{c-1} \sum_{i=1}^{d} B_i^{(c,d)}(-1)^i \left( Z(\sigma_a, b + c + d - i, \sigma_i, \sigma_d) \right) + Z(\sigma_a, -(b + c + d - i), -\sigma_i, \sigma_d) \]

\[ - (1)^d 2^{c-1} \sum_{i=1}^{c} A_i^{(c,d)} \left( Z(\sigma_a, \sigma_b \sigma_d(b + c + d - i), i) \right) + Z(\sigma_a, \sigma_b \sigma_d(b + c + d) \sigma_d(i), -(b + c + d)) \]

\[ - (1)^d 2^{c-1} \left( \frac{(c + d - 1)!}{(c - 1)!d!} \right) \left( Z(\sigma_a, (b + c + d)) \right) + Z(\sigma_a, -(b + c + d)), \]

\[ K_2^{(2)}(\sigma_a, \sigma_b, c, \sigma_d) = \sum_{x_1=1}^{\infty} \cdots \sum_{x_4=1}^{\infty} \frac{\sigma_a^{x_1+x_2+2x_3+x_4} \sigma_b^{2x_2+2x_3+x_4} \sigma_d^{x_4}}{(x_1 + 2x_2 + 2x_3 + x_4)^d(x_2 + 2x_3 + x_4)^b(x_3 + x_4)^c} x_4^d \]

\[ = (-1)^d 2^{c-1} \sum_{i=1}^{d} B_i^{(c,d)}(-1)^i \left( Z(\sigma_a, \sigma_b, (c + d - i), \sigma_d) \right) + Z(\sigma_a, \sigma_b, -(c + d - i), \sigma_d(i)) \]

\[ + (-1)^d \sum_{i=1}^{c} A_i^{(c,d)} 2^{c-i} Z(\sigma_a, \sigma_b, \sigma_d(c + d - i), i) \]

\[ - (1)^d 2^{c-1} \sum_{i=1}^{c} A_i^{(c,d)} \left( Z(\sigma_a, \sigma_b, \sigma_d(c + d - i), i) \right) + Z(\sigma_a, \sigma_b, \sigma_d(c + d - i), -(c + d - i)) \]

\[ - (1)^d 2^{c-1} \left( \frac{(c + d - 1)!}{(c - 1)!d!} \right) \left( Z(\sigma_a, \sigma_b, (b + c)) \right) + Z(\sigma_a, \sigma_b, -(b + c)). \]
When we do depth 5 we see that, like $K^{(1)}_1$, also the $K^{(1)}_2$ splits off two new functions. Hence to produce a generic routine for any depth we have to look at a few very general steps.

In the general case Eqs. (4.12), (4.13) and (4.15) stay more or less the same. They just get more indices to the right. The difference comes with the equations for $K^{(2)}$. We have to make a distinction whether there are still many indices to the right or whether we are terminating. The terminating equations are also more or less the same as the equations for $K^{(2)}$ above, but now with more indices to the left. This leaves the 'intermediary' objects:

$$K^{(i)}_1(M, \sigma_d a, b, \sigma_c c, (n_1, N))$$

$$= \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \frac{\sigma_d^{x_1+2x_1+x_2+x_N} \sigma_c^{x_2+x_N}}{(x_1 + 2x_1 + x_2 + x_N)^b (x_1 + x_2 + x_N)^b (x_2 + x_N)^b},$$

$$K^{(i)}_2(M, \sigma_d a, \sigma_d b, c, \sigma_d d, N)$$

$$= \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \frac{\sigma_d^{x_1+2x_1+x_2+x_N} \sigma_b^{x_1+x_2+x_N} \sigma_d^{x_2+x_N}}{(x_1 + 2x_1 + x_2 + x_N)^b (x_1 + x_2 + x_N)^b (x_2 + x_N)^b}.$$

In these formulas $M$ and $N$ indicate a range of indices. There are more sums and factors in the numerator and denominator, but we just omit them as they do not take part in the 'action'. We use the same techniques applied before to move the factor 2 that multiplies $x_1$ to $x_2$, to the right. When $N$ is empty we run into a termination condition and switch to the equations for $K^{(2)}$.

$$K^{(i)}_1((M), \sigma_d a, b, \sigma_c c, (n_1, N))$$

$$= (-1)^c \sum_{i=1}^{b} A^{(b,c)}_i 2^{b-i} Z(M, \sigma_d a, \sigma_c (b+c-i), i, n_1, N)$$

$$+ (-1)^c \sum_{i=1}^{c} B^{(b,c)}_i 2^{b-i} (-1)^i \left( Z(M, \sigma_d a, (b+c-i), \sigma_c i, n_1, N) + Z(M, \sigma_d a, -(b+c-i), -\sigma_c i, n_1, N) \right)$$

$$- (-1)^c \sum_{i=1}^{b} A^{(b,c)}_i 2^{b-i} K^{(i)}_1((M, \sigma_d a), \sigma_c (b+c-i), i, n_1, (N))$$

$$- (-1)^c \sum_{i=1}^{b} A^{(b,c)}_i 2^{b-i} K^{(i)}_2((M, \sigma_d a, \sigma_c (b+c-i), i, n_1, (N)).$$

$$K^{(i)}_2((M), \sigma_d a, \sigma_d b, c, \sigma_d d, (n_1, N))$$

$$= (-1)^d \sum_{i=1}^{c} A^{(c,d)}_i 2^{c-i} Z(M, \sigma_d a, \sigma_d b, \sigma_d (c+d-i), i, n_1, N)$$

$$+ (-1)^d \sum_{i=1}^{d} B^{(c,d)}_i 2^{c-i} (-1)^i \left( Z(M, \sigma_d a, \sigma_d (b+c+d-i), \sigma_d i, n_1, N) + Z(M, \sigma_d a, -\sigma_d (b+c+d-i), -\sigma_d i, n_1, N) \right)$$

$$- (-1)^d \sum_{i=1}^{c} A^{(c,d)}_i 2^{c-i} K^{(i)}_1((M, \sigma_d a), \sigma_d (b+c+d-i), i, n_1, (N))$$

$$- (-1)^d \sum_{i=1}^{c} A^{(c,d)}_i 2^{c-i} K^{(i)}_2((M, \sigma_d a, (b+c+d-i), \sigma_d i, n_1, (N)).$$

As one can see, each step of the iteration diminishes $N$ by one unit ($n_1$ is an index with its sign) and $M$ may or may not get one more index.

The above formulas can be programmed rather easily and compactly in a language like FORM. We have first programmed and tested the cases 2, 3, 4, 5 and after that we have made a generic routine that can handle any depth. Also this routine has been tested exhaustively. It can be found in the library.

4.2. The role of the generalized doubling relations

Let us start with a modification of the program for expressing Euler sums into a minimal set that was used for testing TFOM [22]. It was modified, so as to allow running only with sums/functions up to a given depth. We use the same relations, up to that depth, as in the complete program, i.e. we use the stuffles, the shuffles and the doubling relations, but not the GDRs. This should generate new information because one is often interested in sums of limited depth but large weight.
Table 1
Number of constants remaining when running at fixed depth for a given weight. With fixed depth we mean all depths up to the given value.

<table>
<thead>
<tr>
<th>Weight</th>
<th>Depth</th>
<th>Number</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>(d=2)</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>1</td>
<td>(d=2)</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1</td>
<td>(d=3)</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
<td>(d=2)</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
<td>(d=2, d=4)</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3</td>
<td>(3 \times (d=3))</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>2</td>
<td>(d=3, d=5)</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>(2 \times (d=2))</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>2</td>
<td>(2 \times (d=2))</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>6</td>
<td>(2 \times (d=2), 4 \times (d=4))</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>6</td>
<td>(2 \times (d=2), 4 \times (d=4))</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>3</td>
<td>(d=2, d=4, d=6)</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Number of shuffles and stuffles separated by depth contributing to Eq.(4.21).

<table>
<thead>
<tr>
<th>Depth</th>
<th>Shuffles</th>
<th>Stuffles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>52</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>41</td>
</tr>
</tbody>
</table>

When we compare the number of remaining variables with the conjectures \([12,13]\), we note that in many cases we have more variables left. However, if we increase the depth these remaining variables are eliminated after all. We set up the program in such a way that these objects may be recognized easily. In Table 1 we present how many of these constants are left and at which depth.

Table 1 indicates that there must be a significant ‘leaking’ of relations at greater depths that create nontrivial results at lower depth. As an example we derived the \(d=2\) relation at weight 6 without substituting the lower weight constants and keeping track of all products of lower weight objects that combined in shuffles and in stuffles. The relation we refer to is given as Eq. (27) in Ref.\([12]\):

\[
Z_{-4,-2} = -\frac{97}{420} \zeta^3 - \frac{3}{4} \zeta^2. \tag{4.21}
\]

The results are shown in Table 2.

We see that a total of 203 equations make contributions to the final result. Considering this, it should not come as a great surprise that attempts to derive this equation by hand using shuffle and stuffle relations have failed thus far.

It is of course possible to obtain this result by different means as was shown in Ref.\([26]\) where the finite harmonic sum \(S_{-4,-2}(N)\) was calculated in terms of the following one-dimensional integral representation:

\[
S_{-4,-2}(N) = -M \left[ \frac{4 \text{Li}_5(-x) - \ln(x) \text{Li}_4(-x)}{x-1} \right]_+ (N) + \frac{1}{2} \zeta_2 [S_4(N) - S_{-4}(N)] - \frac{3}{2} \zeta_3 S_3(N) + \frac{21}{8} \zeta_4 S_2(N) - \frac{15}{4} \zeta_5 S_1(N). \tag{4.22}
\]

where

\[M[f(x)](N) = \int_0^1 dx x^N f(x). \tag{4.23}\]

Since

\[
\int_0^1 \frac{4 \text{Li}_5(-x) + (15/16) \zeta_5 - \ln(x) \text{Li}_4(-x)}{x-1} = \frac{811}{840} \zeta^3 + \frac{3}{4} \zeta^2 \tag{4.24}\]

one obtains with

\[
Z_{-4,-2} = \lim_{N \to \infty} S_{-4,-2}(N) - \zeta_6 \tag{4.25}
\]

the above result. It should, however, be clear that if such methods are needed to replace the phenomenon of leakage, it will be a near impossibility to go to much greater values of the weight parameter.
Using the GDRs at depth $d = 2$ resolves the problem completely. Only the depth $d = 2$ shuffles and stuffles in combination with these GDRs give already the desired formula.

To study the problem at depth $d = 3$, we recreated an old program by one of us\(^6\) that only determines relations at leading depth for objects of which the index field is a Lyndon word. The FORM version of the program is rather fast when applied at depth $d = 3$, see Table 3.

We see a steady increase in the number of undetermined constants. In Tables 3, 4 we list under `expected` the number of undetermined constants according to conjecture [12]. The results for the weights 7 and 9 are in agreement with the numbers in Table 1.

To see whether we could improve the situation, we tried programming generalizations of the formulas $D_0$ and $D_1$ of Ref. [48]. They made no difference. Close inspection reveals that the formula $D_0$ is another form of the shuffle formulas with the combinatorics included properly. The formula $D_1$, or Markett formula [49], also does not add anything new. It seems to be a combination of shuffles and stuffles.

Next we applied the GDRs at depth $d = 3$ and these reduce the number of undetermined constants to their expected value. This means that if we include the GDRs we can run the program at maximum depth $d = 3$ and get a complete set of expressions for all depth $d = 1$, 2 and 3 objects. At the moment we have verified this for all weights up to $w = 51$. The run for the highest weight took about 20 hours of CPU time on a single Xeon processor at 2.33 GHz.

We have made a similar program for depth $d = 4$. This is of course much slower and hence we cannot go to such large values for the weight. The results are given in Table 4. Again we see an increase in the number of extra undetermined objects and again application of the GDRs resolved the issue.

The phenomenon of leakage is rather messy. Basically equations that are in nature of a greater depth have to combine first to eliminate most objects of this depth. After this a few equations remain between lower depth objects. Such leakage is impossible without the stuffle relations. The shuffle relations by themselves do not give terms with a lower depth and neither do the relations based on the doubling formula. But whether these extra relations come from the stuffles alone or materialize only after combining stuffles and shuffles, and maybe doublings, is currently not clear. What is clear is that they involve a very large number of equations. In all cases which we studied the leakage goes over at least two units of depth. This makes it very difficult to investigate. Fortunately the GDRs seem to resolve these problems. We formulate

\(^6\) The program had an error and hence gave rise to a wrong conjecture.
Table 5
Number of excess elements when no doubling relations (also no GDRs) are used, and when only no GDRs are used.

<table>
<thead>
<tr>
<th>Weight</th>
<th>No doubling</th>
<th>No GDRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Conjecture 1. The stuffle, shuffle, doubling and Generalized Doubling Relations are sufficient to reduce the Euler sums of a given weight and depth to a minimal set that is in agreement with the conjecture [12], both in weight and in depth.

Even if we could dispense with the GDRs up to weight $w = 10$, the whole situation changes at weight $w = 11$, see Table 5. Running only stuffles, shuffles and doubling relations leaves one variable in excess of the conjecture [12]. The GDRs provide the missing equation by which this variable is expressed in terms of the other remaining variables and agreement with conjecture [12] is reached. The same effect occurs at weight $w = 12$. Again there is one variable too many if the GDRs are not used. We cannot check this beyond weight $w = 12$, because leakage forces us to run all depths for a given weight if we exclude the GDRs. This becomes excessive in terms of current computer resources. Alternatively one could have used the relations of Eq. (2.26) to resolve this issue, but these relations do not help with the problem of running at a limited depth. Hence we have to add the GDRs anyway.

5. The computer program

We have combined the above relations into a new computer program to resolve all relations between MZVs and reduce them to a minimal set. In principle this is done by writing down all equations for the MZVs of a given weight and then solving the system. A few variables at the given weight may remain and there will be products of objects of lower weight.

Considering the size of the problem and its sparsity it did not look to us like a typical problem to solve by matrix techniques even though other people have done so [50,51]. Typically there would be many thousands of zeroes for each non-zero element. The advantage of computer algebra is that in a sparse polynomial representation those zeroes will not be present and need no attention. Hence we have selected a rather special method the essence of which has already been used in Refs. [10,18,22], although not described there in detail.

We select the FORM system, because it is by far the best suited for this kind of problems. Since we go to much greater weights than previously investigated, we take the opportunity to give here a better description of the completely renewed version of the program.

We start generating a master expression which contains one term for each sum that we want to compute. For the MZVs of weight $w = 4$ this expression looks in computer terms like

$$\text{FF} =$$

\[+E(0,0,0,1)*(H(0,0,0,1)) + E(0,0,1,1)*(H(0,0,1,1)) + E(0,1,0,1)*(H(0,1,0,1));\]

We have used already that we will only compute the finite elements and that there is a duality that allows us to eliminate all elements with a depth greater than half the weight. When the depth is exactly half the weight we choose from a sum and its dual the element that comes first lexicographically. We work in terms of the $H$-functions because for the Euler sums the basis of Ref. [12] turns out to be ideal. This basis consists of all Lyndon words of negative odd integers that add up in absolute value to the weight. For the MZVs these $H$-functions and the $Z$-functions are identical anyway and hence we could keep a single program for most procedures.

We pull the function $E$ outside brackets. The contents of a bracket is what we know about the object indicated by the indices of the function $E$. In the beginning this is all trivial knowledge.

Assume now that we generate the stuffle relation

$$H_{0,1}H_{0,1} = H_{0,0,0,1} + 2H_{0,1,0,1}. \quad (5.1)$$

The left-hand side can be substituted from the tables for the lower weight MZVs. Hence it becomes $\zeta_2^4$. In the program $\zeta_2$ is called $z_2$.

The right-hand side objects are replaced by the contents of the corresponding $E$ brackets in the master expression. These are for now trivial substitutions. From the result we generate the substitution

$$\text{id } H(0,1,0,1) = z_2^2/2 - H(0,0,0,1)/2;$$

which we apply to the master expression. Hence the master expression becomes

$$\text{FF} =$$

\[+E(0,0,0,1)*(H(0,0,0,1)) + E(0,0,1,1)*(H(0,0,1,1)) + E(0,1,0,1)*(z_2^2/2 - H(0,0,0,1)/2);\]

Let us now generate the corresponding shuffle relation:

$$H_{0,1}H_{0,1} = 4H_{0,0,1,1} + 2H_{0,1,0,1}. \quad (5.2)$$
and replace the right-hand side objects by the contents of the corresponding \( E \) brackets in the master expression. This gives

\[
\zeta_2^2 = 4H_{0,0,1.1} + \zeta_2^2 - H_{0,0,0.1}
\]

which leads to the substitution

\[
id H(0,0,1,1) = H(0,0,0,1)/4;
\]

and we obtain

\[
\begin{align*}
FF &= \\
&= +E(0,0,0,1)*(H(0,0,0,1)) \\
&+E(0,0,1,1)*(H(0,0,0,1)/4) \\
&+E(0,1,0,1)*(z2^2/2-H(0,0,0,1)/2);
\end{align*}
\]

We also need the divergent shuffles and stuffles. This is done by including the shuffles involving the basic divergent object and breaking down the multiple divergent sums with the stuffle relations as in:

\[
H_1 H_{0,0.1} = 2H_{0,0,1.1} + H_{0,1,0.1} + H_{1,0,0.1}
= -H_{0,0,0.1} + H_{0,0,1.1} + H_{0,1,0.1} + H_{1,0,0.1}.
\]

In the case we use \( H_1 \) as the only divergent object, this is equivalent to using Hoffmann’s \([52]\) relation. We can use any combination involving divergent objects, provided not both are divergent simultaneously. Substituting from the master expression we get the relation

\[
0 = -\frac{5}{4} H_{0,0,0.1} + \frac{1}{2} \zeta_2^2
\]

and hence the substitution

\[
id H(0,0,0,1) = z2^2/2/5;
\]

and finally the master expression becomes

\[
\begin{align*}
FF &= \\
&= +E(0,0,0,1)*(z2^2*2/5) \\
&+E(0,0,1,1)*(z2^2/10) \\
&+E(0,1,0,1)*(z2^2*3/10);
\end{align*}
\]

Now we can read off the values of all MZVs of weight 4 that we set out to compute. All other elements can be obtained from these by trivial operations that involve the use of one or two relations only.

The method should be clear now: we generate the master expression that contains all nontrivial objects that we need to compute. Then we generate all known equations one by one, putting in the knowledge that is contained in the master expression. After that we incorporate the new knowledge in the master expression (provided the equation does not become trivial which will happen frequently, because we have more equations than variables).

With this method we do not need all equations to be in memory simultaneously. But there is a very important observation: the order in which the equations are generated will determine the size the master expression can have during the calculation. This intermediate expression swell should be controlled as much as possible, because it can make many orders of magnitude difference in the execution time and the space needed. And there is another problem: substituting a new equation in the master expression can be rather costly when this expression becomes rather big. To have to do this each time is wasteful because the master expression will have to be brought to normal order again. Therefore we have adopted a scheme in which we generate the equations in groups. Then we apply first a Gaussian elimination scheme among the equations in the group, eliminating both above and below the diagonal. If we have

\[
\text{normal order again. Therefore we have adopted a scheme in which we generate the equations in groups. Then we apply first a Gaussian elimination scheme among the equations in the group, eliminating both above and below the diagonal.}
\]

Now we can read off the values of all MZVs of weight 4 that we set out to compute. All other elements can be obtained from these by trivial operations that involve the use of one or two relations only.

The method should be clear now: we generate the master expression that contains all nontrivial objects that we need to compute. Then we generate all known equations one by one, putting in the knowledge that is contained in the master expression. After that we incorporate the new knowledge in the master expression (provided the equation does not become trivial which will happen frequently, because we have more equations than variables).

With this method we do not need all equations to be in memory simultaneously. But there is a very important observation: the order in which the equations are generated will determine the size the master expression can have during the calculation. This intermediate expression swell should be controlled as much as possible, because it can make many orders of magnitude difference in the execution time and the space needed. And there is another problem: substituting a new equation in the master expression can be rather costly when this expression becomes rather big. To have to do this each time is wasteful because the master expression will have to be brought to normal order again. Therefore we have adopted a scheme in which we generate the equations in groups. Then we apply first a Gaussian elimination scheme among the equations in the group, eliminating both above and below the diagonal. If we have \( G \) equations left we can substitute \( G \) variables in the master expression simultaneously. Again, this is not optimal yet as that would give \( G \) substitution statements and hence each term needs \( G \) pattern matchings. To improve upon this we enter these \( G \) objects in a temporary table and the substitution in the master expression is by a single table lookup. This is a binary search inside FORM and hence when we have grouped for instance 512 equations, the lookup takes only 9 compares, each of which is anyway much faster than a full pattern matching. The difference shows in a run we made on a machine with a single Opteron processor. When running the equations for MZVs one by one at weight 18, the run took 26761 sec, while with groups of 256 equations the same program ran in 2974 sec. Over the range in weights that we experimented in a run we made on a machine with a single Opteron processor. When running the equations for MZVs one by one at weight 18, the run took 26761 sec, while with groups of 256 equations the same program ran in 2974 sec. Over the range in weights that we experimented with, the optimal group size we found for the MZVs was close to 2

\[
\text{w}^{-1/2}.
\]

This is the value we use in the program. For the Euler sums the best value obeys a more involved relation because the number of variables goes with a power of three. We have measured the effect and it is shown in Table 6. From this table it looks like a decent value for the size of the groups is \( 2^{w/2-7} \) in which the exponent is rounded down to the nearest integer. We see, however, that the exact value is not very critical.

If it would be of great importance to improve over this scheme, one could set up a tree structure in the Gaussian scheme. This would change its quadratic (in the size of the groups) nature to a \( G \log(G) \) behavior. It would, however, make the code much more complicated and anyway, this is not where currently most computer time is used. As a consequence we decided to stay with the simple grouping.

This leaves determining a good order in which to generate the equations. It requires much trial and error and we are not claiming that we have the best scheme possible. The scheme for the stuffles is rather good, but for the shuffles it could probably be better. Once we could run what we wanted to run, we have stopped searching intensively. Anyway, the intermediate expression swell is rather moderate as is shown in the tables containing the results below.

Before we discuss the order of the equations, we make several observations:
Shuffles preserve depth.
Stuffles either preserve depth or lower it.
The number of indices that are one in sum notation is either preserved or lowered by stuffles.
The shuffle relations can contain many more terms than the stuffle relations.
The shuffles (which are executed in integral notation) can contain large combinatoric factors when there are long sequences of zeroes or ones. This lowers the number of terms in the equation.

Based on the above observations we start with the equations with the lowest depth, and then do the ones with the next depth, etc. In the case that we only look at the MZVs, we only need to go up to half the weight (rounded down), because the duality relation takes care of the other sums. In the case of the Euler sums we have to go ‘all the way’.

For each depth we do first the stuffles and then the shuffles. There are actually conjectures about that one does not need all stuffles but only a limited subset. We do not use these conjectures because they would make it necessary to apply more shuffle relations and those are more complicated than the stuffles that we would omit. We have verified experimentally that this would make the program significantly slower.

In the case of Euler sums we have two more categories of equations: the equations due to the doubling relation and the equations due to the GDRs. It looks like we do not need all equations from the latter category, but because they are not extremely costly, we have not been motivated enough to run many programs testing what can be done here. We just run them all and this way there is no risk that we omit something essential. They are, however, more costly than the shuffle equations and hence we put them after the shuffles. But more ordering within the group of (generalized) doubling equations is not relevant as there are only comparatively few substitutions generated by them.

To deal with the stuffles at a given weight and depth we generate an expression that contains one term for each stuffle relation that we will use. Then we apply several operations that multiply each term with a function with arguments based on the equation to be generated. The effect of this is that at the next sorting the equations will be ordered according to these arguments. This can be done in a rather flexible way. The ordering is in sum notation according to:

- The number of indices that are one.
- Next comes the number of indices that are two, then three, etc.
- The number of indices in the sum with the smallest depth.
- The largest first index in either of the two sums.

This relatively simple ordering is amazingly effective. When we compute the size of the basis, using arithmetic over a 31-bit prime number, it gives a nearly monotonically increasing size in the master expression, indicating that it will be very hard to improve upon it. Once we have this expression we use a feature of FORM that allows one to define a loop in which the loop variable takes a value which is (sequentially) each time a term from a given expression. This way we can now create expressions for each equation and each time we have enough equations to fill a group we call the routine that will expand the equations and process them. We do not consider stuffle equations that contain a divergent sum. Those are taken into consideration anyway when we have to extract the divergences in the shuffle equations, and for the Euler sums the GDRs.

For the shuffles things are more complicated. Again we generate an expression for all shuffles for the given depth. In this case we generate however only those objects that correspond to shuffles in which one of the objects is only of depth one. This seems to be sufficient. We have never run across a case where the other shuffles had any additional effect. It is actually possible to restrict the number of shuffle equations even more, although this is only based on conjectures and experimentation. A formal proof is missing. The ordering is now done according to

- The weight of the object of depth one.
- The number of indices that are one in sum notation.
- For each sum we compute the sum of the squares of the indices in sum notation. We order by the maximum of either of the two. The biggest comes first.
- We select which of the two sums has the smallest first index. The larger values for this number come first.
- We add the first indices of the two sums. The larger values come first.

The complicating factor here is that we have to keep divergent sums. We only keep those equations in which at most one object is divergent, and there is only a single divergence. Hence sums that have the first two indices equal to one are not considered.

According to observation the shuffle equations that fulfill all following requirements always reduce to trivial (0 = 0) equations:

- The combined depth is at least three.
- There is at least one index that is equal to one.
The depth one object has at least weight two.

If the depth one object has weight $w = 2$, there are at least two indices equal to one in the other object.

Harmonic sums with all the same index decompose algebraically into a polynomial of single harmonic sums. It is easily shown that the algebraic relations [37] always allow to write any harmonic sum in terms of polynomials of $S_1(N)$ and sums, which converge in the limit $N \rightarrow \infty$. All the above greatly reduces the number of shuffle equations that have to be evaluated. Because this evaluation is one of the expensive steps, it speeds up the program significantly. On the other hand, it is only an observation made in runs that do not involve the greatest weights. For the more critical runs we have left these equations active and spent the extra computer time.

The above describes the basic program. At this point we split it in several varieties. To first determine whether shuffles and stuffles are sufficient to reduce all MZVs to a basis of the conjectured size, we have made the simplifications:

- All products of lower weight objects are set to zero. This means we will only determine whether reduction to a Lyndon basis takes place.
- We work modulus a 31-bit prime.

We have also made runs over the rational numbers. This becomes only problematic for the very highest values of the weight.

For constructing tables of all sums at a given weight we run the full program. The performance of the program is shown in Fig. 1 for a complete run at weight $w = 21$ and a run to depth $d = 8$ at weight $w = 26$. We see that the stuffles give a steady growth of the master expression but that the shuffles cause intermediate expression swell which is worse when the depth is much less than half the weight. The result is that when we run the complete system most time is spent with the stuffle relations while for the limited depth runs by far most time is spent with the shuffle relations.

In the case of Euler sums the master expression is created with a three letter alphabet $(-1, 0, 1)$ rather than the two letter alphabet $(0, 1)$ for the MZVs. In addition there are many more equations to consider because the number of lower weight objects that we can multiply either by shuffles or stuffles is correspondingly greater. Of course also for the Euler sums it is possible to just study the basis.

In addition it is possible to study sums to a limited depth. This way we can go to much greater values of the weight. This is of course only possible if we use a basis in which the concept of depth is relevant, like the basis of the odd negative indices that form a Lyndon word. Without such a basis the calculations become much harder.

When we are constructing tables we cannot go quite as far in weight as when we are determining rank deficiency. When we use a Lyndon basis, the majority of terms consists of products of basis elements of lower weights. This means that we have many more terms to carry around. We observe, in addition, that the coefficients containing the most digits are in the terms with powers of $\zeta_2$. This is to be expected since $\zeta_2^N$ is our repository for all terms of the form $\zeta_2^m \zeta_2^n$ with $N = ma + nb$.

The representation we have selected, together with the modular arithmetic, makes for a very fast treatment of the terms. This is reflected in the number of terms that can be processed. In one run, which took more than 30 days the program generated a total of more than $7 \cdot 10^{12}$ terms. This seems to be a new record.

6. The running of the programs

We have used the programs of the previous section to obtain results to as high a weight and depth as possible, both for MZVs and Euler sums. Before we start discussing these results we show the parameters of these runs to give the reader an impression of what is available and why there are limitations to obtain more.

We start with the Euler sums. We have first run the complete system for the given weights, see Table 7. This means that for $w = 12$ there are expressions for all 236 196 Euler sums with that weight, all expressed in terms of the basis of Lyndon words of the negative odd integers, see Appendix A, which is the basis we use for all Euler sums, unless mentioned differently.

The columns marked ‘variables’ mentions how many variables there are at the start of the program. ‘Remaining’ tells how many basis elements remain in the end. Under ‘output’ we give the size of the output expression in text format. The column ‘size’ refers to the largest

---

7 With this we mean the programs that determine the size of the basis when using arithmetic over a prime number. Once we have established this, any further runs to for instance determine all values over the rational numbers, we can safely drop these equations.
Table 7
Runs on an 8-core Xeon computer at 3 GHz and with 32 Gbytes of memory. The column ‘Eqns’ gives the number of equations that was considered.

<table>
<thead>
<tr>
<th>w</th>
<th>Variables</th>
<th>Eqns</th>
<th>Remaining</th>
<th>Size</th>
<th>Output</th>
<th>Time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>36</td>
<td>57</td>
<td>1</td>
<td>4.3K</td>
<td>2.0K</td>
<td>0.06</td>
</tr>
<tr>
<td>5</td>
<td>108</td>
<td>192</td>
<td>2</td>
<td>21K</td>
<td>8.9K</td>
<td>0.12</td>
</tr>
<tr>
<td>6</td>
<td>324</td>
<td>665</td>
<td>2</td>
<td>98K</td>
<td>42K</td>
<td>0.37</td>
</tr>
<tr>
<td>7</td>
<td>972</td>
<td>2205</td>
<td>4</td>
<td>472K</td>
<td>219K</td>
<td>1.71</td>
</tr>
<tr>
<td>8</td>
<td>2916</td>
<td>7313</td>
<td>5</td>
<td>2.25M</td>
<td>1.15M</td>
<td>7.78</td>
</tr>
<tr>
<td>9</td>
<td>8748</td>
<td>23909</td>
<td>8</td>
<td>11M</td>
<td>6.3M</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>26244</td>
<td>77853</td>
<td>11</td>
<td>58M</td>
<td>36M</td>
<td>353</td>
</tr>
<tr>
<td>11</td>
<td>78732</td>
<td>251565</td>
<td>18</td>
<td>360M</td>
<td>213M</td>
<td>3266</td>
</tr>
<tr>
<td>12</td>
<td>236196</td>
<td>809177</td>
<td>25</td>
<td>3.1G</td>
<td>1.29G</td>
<td>47311</td>
</tr>
</tbody>
</table>

Table 8
Summary of the runs at \( d = 4 \). The runs were performed on a computer with 8 Xeons at 3 GHz, using (\( T \)FORM).

<table>
<thead>
<tr>
<th>Weight</th>
<th>Constants</th>
<th>Running time [sec]</th>
<th>Output [Mbyte]</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>956</td>
<td>7</td>
<td>0.26</td>
</tr>
<tr>
<td>10</td>
<td>1412</td>
<td>13</td>
<td>0.64</td>
</tr>
<tr>
<td>11</td>
<td>1996</td>
<td>24</td>
<td>1.25</td>
</tr>
<tr>
<td>12</td>
<td>2724</td>
<td>39</td>
<td>3.18</td>
</tr>
<tr>
<td>13</td>
<td>3612</td>
<td>68</td>
<td>5.04</td>
</tr>
<tr>
<td>14</td>
<td>4676</td>
<td>108</td>
<td>17.1</td>
</tr>
<tr>
<td>15</td>
<td>5932</td>
<td>199</td>
<td>17.1</td>
</tr>
<tr>
<td>16</td>
<td>7396</td>
<td>436</td>
<td>71.1</td>
</tr>
<tr>
<td>17</td>
<td>9084</td>
<td>602</td>
<td>54.9</td>
</tr>
<tr>
<td>18</td>
<td>11012</td>
<td>1323</td>
<td>275.9</td>
</tr>
<tr>
<td>19</td>
<td>13196</td>
<td>2761</td>
<td>1571</td>
</tr>
<tr>
<td>20</td>
<td>15652</td>
<td>5424</td>
<td>877</td>
</tr>
<tr>
<td>21</td>
<td>18396</td>
<td>14090</td>
<td>395</td>
</tr>
<tr>
<td>22</td>
<td>21444</td>
<td>21875</td>
<td>2559</td>
</tr>
</tbody>
</table>

Table 9
Summary of the runs at \( d = 5 \). Same computer as used in Table 8.

<table>
<thead>
<tr>
<th>Weight</th>
<th>Constants</th>
<th>Remaining</th>
<th>Running time [sec]</th>
<th>Output [Mbyte]</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>3394</td>
<td>7</td>
<td>27</td>
<td>1.15</td>
</tr>
<tr>
<td>10</td>
<td>5702</td>
<td>7</td>
<td>72</td>
<td>3.11</td>
</tr>
<tr>
<td>11</td>
<td>9042</td>
<td>13</td>
<td>172</td>
<td>8.5</td>
</tr>
<tr>
<td>12</td>
<td>13686</td>
<td>11</td>
<td>478</td>
<td>20.9</td>
</tr>
<tr>
<td>13</td>
<td>19938</td>
<td>22</td>
<td>7</td>
<td>68.9</td>
</tr>
<tr>
<td>14</td>
<td>28134</td>
<td>17</td>
<td>4306</td>
<td>133</td>
</tr>
<tr>
<td>15</td>
<td>38642</td>
<td>35</td>
<td>27607</td>
<td>473</td>
</tr>
<tr>
<td>16</td>
<td>51862</td>
<td>24</td>
<td>110336</td>
<td>688</td>
</tr>
<tr>
<td>17</td>
<td>68226</td>
<td>55</td>
<td>450462</td>
<td>2767</td>
</tr>
</tbody>
</table>

We have a nice example here of what happens when we change the order in which we deal with the shuffles and the stuffles. We reran the program of Table 9 for the weights \( w = 14 \) and \( w = 15 \) under these conditions, obtaining running times of 100973 and 493489 sec respectively. This is more than an order of magnitude slower than the order we select in the regular programs.

Because we like to compare results of the MZV runs with those of the Euler runs to as high a weight as possible we made also runs in which we do all calculus modulus a 31-bit prime number. The number we selected is 2147479273. We never ran into a case in which this seemed to cause problems. In the programs in which we used this modulus we also dropped all terms that are products of lower objects.

8 The first time we ran the \( w = 12 \) case on an 8-core Xeon machine at 2.33 GHz the run took two full weeks. It just shows how good a test case this problem is. Both (\( T \))FORM and the MZV program have been improved greatly during this project.
Table 10
Summary of the runs at \(d = 4\) in modular arithmetic, dropping all terms that are products of lower weight objects.

<table>
<thead>
<tr>
<th>Weight</th>
<th>Constants</th>
<th>Running time [sec]</th>
<th>Output [Mbyte]</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>4 676</td>
<td>35</td>
<td>1.3</td>
</tr>
<tr>
<td>16</td>
<td>7 396</td>
<td>105</td>
<td>2.9</td>
</tr>
<tr>
<td>18</td>
<td>11 012</td>
<td>323</td>
<td>6.0</td>
</tr>
<tr>
<td>20</td>
<td>15 652</td>
<td>939</td>
<td>11.3</td>
</tr>
<tr>
<td>22</td>
<td>21 444</td>
<td>2211</td>
<td>20.5</td>
</tr>
<tr>
<td>24</td>
<td>28 516</td>
<td>5335</td>
<td>35</td>
</tr>
<tr>
<td>26</td>
<td>36 996</td>
<td>13 127</td>
<td>57</td>
</tr>
<tr>
<td>28</td>
<td>47 012</td>
<td>47 056</td>
<td>89</td>
</tr>
<tr>
<td>30</td>
<td>58 692</td>
<td>100 813</td>
<td>137</td>
</tr>
</tbody>
</table>

Table 11
Summary of the runs at \(d = 5\) in modular arithmetic, dropping all terms that are products of lower weight objects.

<table>
<thead>
<tr>
<th>Weight</th>
<th>Constants</th>
<th>Running time [sec]</th>
<th>Output [Mbyte]</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>16 812</td>
<td>388</td>
<td>5.5</td>
</tr>
<tr>
<td>15</td>
<td>33 386</td>
<td>2832</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>60 044</td>
<td>18 836</td>
<td>53</td>
</tr>
<tr>
<td>19</td>
<td>100 236</td>
<td>118 874</td>
<td>131</td>
</tr>
<tr>
<td>21</td>
<td>157 932</td>
<td>554 870</td>
<td>299</td>
</tr>
</tbody>
</table>

Table 12
Summary of the runs at \(d = 6\) in modular arithmetic, dropping all terms that are products of lower weight objects. Times refer to an 8 Xeon-core machine at 3 GHz and 32 GBytes of memory.

<table>
<thead>
<tr>
<th>Weight</th>
<th>Constants</th>
<th>Remaining</th>
<th>Running time [sec]</th>
<th>Output [Mbyte]</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>56 940</td>
<td>22</td>
<td>2611</td>
<td>5.5</td>
</tr>
<tr>
<td>14</td>
<td>90 564</td>
<td>37</td>
<td>12 716</td>
<td>51</td>
</tr>
<tr>
<td>15</td>
<td>138 636</td>
<td>35</td>
<td>55 204</td>
<td>87</td>
</tr>
<tr>
<td>16</td>
<td>205 412</td>
<td>66</td>
<td>206 951</td>
<td>214</td>
</tr>
<tr>
<td>17</td>
<td>295 916</td>
<td>55</td>
<td>789 540</td>
<td>288</td>
</tr>
<tr>
<td>18</td>
<td>416 004</td>
<td>109</td>
<td>2 622 157</td>
<td>711</td>
</tr>
</tbody>
</table>

Table 13
Runs on an 8-core Xeon computer at 3 GHz and with 32 GBytes of memory. ‘Num’ indicates, for the final expressions, the maximum number of decimal digits in either a numerator or a denominator. ‘Eff.’ is the ratio of CPU time versus real time indicating how well the processors are used. The meaning of the column labeled ‘Rat.’ is explained in the text. The anomaly between size and output for \(w = 21\) is due to the fact that the output is in text and size is in FORM binary notation.

<table>
<thead>
<tr>
<th>(w)</th>
<th>(d)</th>
<th>(C)</th>
<th>Size</th>
<th>Output</th>
<th>Num</th>
<th>CPU [sec]</th>
<th>Real [sec]</th>
<th>Eff.</th>
<th>Rat.</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>8</td>
<td>128</td>
<td>11M</td>
<td>7M</td>
<td>22</td>
<td>289</td>
<td>56</td>
<td>5.16</td>
<td>0.99</td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>256</td>
<td>30M</td>
<td>21M</td>
<td>19</td>
<td>677</td>
<td>129</td>
<td>5.25</td>
<td>0.96</td>
</tr>
<tr>
<td>18</td>
<td>9</td>
<td>256</td>
<td>88M</td>
<td>64M</td>
<td>29</td>
<td>3071</td>
<td>517</td>
<td>5.94</td>
<td>1.11</td>
</tr>
<tr>
<td>19</td>
<td>9</td>
<td>512</td>
<td>224M</td>
<td>182M</td>
<td>28</td>
<td>6848</td>
<td>1206</td>
<td>5.68</td>
<td>1.00</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>512</td>
<td>790M</td>
<td>558M</td>
<td>36</td>
<td>44 883</td>
<td>6834</td>
<td>6.57</td>
<td>1.42</td>
</tr>
<tr>
<td>21</td>
<td>10</td>
<td>1024</td>
<td>1766M</td>
<td>1821M</td>
<td>40</td>
<td>86318</td>
<td>13851</td>
<td>6.23</td>
<td>1.12</td>
</tr>
<tr>
<td>22</td>
<td>11</td>
<td>1024</td>
<td>8856M</td>
<td>5927M</td>
<td>46</td>
<td>1 572 605</td>
<td>208 972</td>
<td>7.53</td>
<td>3.18</td>
</tr>
</tbody>
</table>

weight objects. This means that in the end all sums are expressed into elements from the same-weight Lyndon part of the basis only. Such programs are much faster. This can be seen in Tables 10, 11 and 12 which are for depth \(d \leq 4\), depth \(d \leq 5\) and depth \(d \leq 6\), respectively.

The run at \(w = 18, d = 6\) deserves some special attention. It was our most costly run and during the running TFOM processed more than \(7 \cdot 10^{12}\) terms.

We come now to our runs for the Multiple Zeta Values. Those runs look more spectacular because there is much more literature on them. First we present the ‘complete’ runs in which all calculus is over the rational numbers and all terms are kept, cf. Table 13. ‘Rat.’ is the real time of this run divided by the real time of a run with a 31-bit prime number dropping also products of lower weight objects. Together with the numbers in the ‘Num’ column it shows that making several runs modulus a 31-bit prime and then using the Chinese remainder theorem [53], will not be efficient. We would need at least 12 runs for the \(w = 22\) case and even then we have to account for dropping the lower weight terms.

We indicate the maximum value of the depth which, due to the duality relation for MZVs, is sufficient to obtain all MZVs at the given weight.

The basis in which these results are presented is described in Appendix B. If we let the program select the basis, the outputs are shorter but from the viewpoint of basis elements selected there is less structure.

The next sequence of runs is performed using modular arithmetic in which we refer to the same 31-bit prime number as before. Again we run the full range of depths needed to obtain all sums. As usual in modular runs, we drop the products of lower weight objects. The results are given in Table 14.

The output of the run at \(w = 23\) gives the results for \(2^{20}\) MZVs expressed in terms of the 28 same-weight elements of a Lyndon basis selected by the program.
Lille group presented their studied the MZVs and the Euler sums to it was shown, that the basis has the expected length, but modulo powers of are not excessive, higher weights could not be investigated yet because of memory limitations. Since these methods are based on the Euler sums is mentioned by the Lille-group. However, the data-tables made available[58] only contain the relations to authors, among them D. Broadhurst, to[54], which gave a nearly complete list for the Euler sums of literature. The relations between MZVs were studied both by mathematicians and physicists. An early study is due to Gastmans and Troost to make them more accessible. In the higher runs some elements will be missing. In the runs on an 8-core Xeon computer at 3 GHz and with 32 Gbytes of memory. G is the size of the group used in the Gaussian elimination. Size is the maximum size of the master expression during the run, Output is the size of the master expression in the end, CPU is the total CPU time of all processors together in seconds. ‘Real’ denotes the elapsed time in seconds and ‘Eff.’ is the pseudo efficiency, defined by the CPU time divided by the real time.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>128</td>
<td>1.7M</td>
<td>1.2M</td>
<td>300</td>
<td>57</td>
<td>5.25</td>
</tr>
<tr>
<td>17</td>
<td>256</td>
<td>5.6M</td>
<td>3.2M</td>
<td>713</td>
<td>134</td>
<td>5.32</td>
</tr>
<tr>
<td>18</td>
<td>256</td>
<td>14.4M</td>
<td>7.2M</td>
<td>2706</td>
<td>465</td>
<td>5.82</td>
</tr>
<tr>
<td>19</td>
<td>512</td>
<td>39M</td>
<td>19M</td>
<td>6901</td>
<td>1206</td>
<td>5.72</td>
</tr>
<tr>
<td>20</td>
<td>512</td>
<td>104M</td>
<td>45M</td>
<td>30097</td>
<td>4819</td>
<td>6.25</td>
</tr>
<tr>
<td>21</td>
<td>1024</td>
<td>239M</td>
<td>114M</td>
<td>75302</td>
<td>12379</td>
<td>6.08</td>
</tr>
<tr>
<td>22</td>
<td>1024</td>
<td>767M</td>
<td>280M</td>
<td>449202</td>
<td>65644</td>
<td>6.84</td>
</tr>
<tr>
<td>23</td>
<td>2048</td>
<td>217G</td>
<td>734M</td>
<td>992431</td>
<td>151337</td>
<td>6.56</td>
</tr>
<tr>
<td>24</td>
<td>2048</td>
<td>8.04G</td>
<td>1.77G</td>
<td>9251325</td>
<td>1268247</td>
<td>7.29</td>
</tr>
</tbody>
</table>

Table 15

Runs on an 8-core Xeon computer at 3 GHz and with 32 Gbytes of memory. D indicates the maximum depth (see text). We reran at w = 23 and w = 24 to have information for extrapolation purposes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>7</td>
<td>2048</td>
<td>1.55G</td>
<td>89M</td>
<td>61447</td>
<td>9579</td>
<td>6.41</td>
</tr>
<tr>
<td>24</td>
<td>8</td>
<td>2048</td>
<td>673M</td>
<td>380M</td>
<td>536921</td>
<td>72991</td>
<td>7.36</td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>4096</td>
<td>6.37G</td>
<td>244M</td>
<td>369961</td>
<td>50197</td>
<td>7.37</td>
</tr>
<tr>
<td>26</td>
<td>8</td>
<td>4096</td>
<td>38.3G</td>
<td>1160M</td>
<td>4786841</td>
<td>651539</td>
<td>7.35</td>
</tr>
<tr>
<td>27</td>
<td>7</td>
<td>6144</td>
<td>12.7G</td>
<td>914M</td>
<td>2152321</td>
<td>277135</td>
<td>7.77</td>
</tr>
<tr>
<td>28</td>
<td>6</td>
<td>6144</td>
<td>2.88G</td>
<td>314M</td>
<td>235972</td>
<td>30960</td>
<td>7.62</td>
</tr>
<tr>
<td>29</td>
<td>7</td>
<td>6144</td>
<td>41.0G</td>
<td>3007M</td>
<td>8580364</td>
<td>1112836</td>
<td>7.71</td>
</tr>
<tr>
<td>30</td>
<td>6</td>
<td>6144</td>
<td>6.27G</td>
<td>658M</td>
<td>829701</td>
<td>106353</td>
<td>7.80</td>
</tr>
</tbody>
</table>

In Table 15 we give the statistics of runs to a more restricted depth. If the conjecture [13] is correct the runs at w = 25, 26 should still give us a complete basis. In the higher runs some elements will be missing.

We would have liked to have a run for depth d ≤ 9 at w = 27, but it would probably take more than a year with current technology. A run for depth d ≤ 8 at w = 28 will require a smaller CPU time. The reason why these runs are interesting is explained in Section 10 on pushdowns. They may give us a new type of basis element that would indicate a double pushdown.

The outputs of all of the above runs are collected in the data mine, together with some files in which the results have been processed to make them more accessible.

At the end of this section we would like to discuss the status of the general investigation of MZVs and Euler sums in the foregoing literature. The relations between MZVs were studied both by mathematicians and physicists. An early study is due to Gastmans and Troost [54], which gave a nearly complete list for the Euler sums of w = 4 and many relations for w = 5, supplemented in [11] later. Various authors, among them D. Broadhurst, to w = 9, and D. Zagier, performed precision numerical studies [55] using PARI [56] during the 1990s for MZVs, which were not published. A very far-reaching investigation concerned the study of some of the MZVs at w = 23 and depth d = 7 by Broadhurst by numerical techniques (PSLQ). Double sums were studied in [57] using the PSLQ method [15]. Vermaseren both studied the MZVs and the Euler sums to w = 9 [10] using a FORM program [21]. This was the situation around the year 2000, when the Lille group presented their w = 12 results for the MZVs and w = 7 results for the Euler sums [58]. In Ref. [59] the solution of w = 8 for the Euler sums is mentioned by the Lille-group. However, the data-tables made available [58] only contain the relations to w = 7. Moreover, the relations used in [59] do not cover the doubling relation, which is needed to reduce to the conjectured basis at this weight, as will be shown later. For the MZVs w = 10 had been solved in [60] and w = 13 in [61], cf. [62]. Vermaseren could extend the MZVs to w = 16 [63]. Studies for w = 16 were also performed at Lille [64] without making the results public. In the studies by Vermaseren also the divergent harmonic sums \( \zeta(1, \bar{1}) \) were included, as this is sometimes necessary for physics applications, cf. also [11].

The primary goal in this paper is to derive explicit representations of the MZVs over several bases suitable to the respective questions investigated. If one only wants to determine the size of the basis one may proceed differently, cf. [50]. Here for w = 19 in the MZV case it was shown, that the basis has the expected length, but modulo powers of \( \pi^2 \) at even weights. In [51] the case w = 20 was studied determining the size of the basis calculating the rank of the associated matrix modulo a 15-bit prime. Although the computation times are not excessive, higher weights could not be investigated yet because of memory limitations. Since these methods are based on the respective algebra only they can be extended to colored Multiple Zeta Values by extending the underlying alphabet.

7. The data mine

The results of our runs, together with a number of FORM programs to manipulate them and clarifying text, are available on the Internet in pages that we call the MZV data mine. It can be located as a link in the FORM home page [65]. Here we will describe the notations and how to use the programs.

The notations we use in the data mine are that the MZVs are represented either by a function \( Z \) of which the variables are its indices or by a single symbol that consists of a string of objects of which the first character is the letter  \( z \) and the remaining characters are decimal digits. Each of these strings refers to an index of the MZV. Let us give an example:

\[
z11z3z3 = Z(11, 3, 3)
\]
For the Euler sums we use mostly the function \( \zeta \). It can have positive and negative indices, the negative ones indicating alternating or Euler sums. When we use basis elements a compact notation is the letter \( \zeta \) followed by a number of alphabetic characters or digits. Each character stands for a negative index. The digits 1,...,9 stand for the indices \(-1,...,-9\) and the upper case characters \( A,...,Z \) stand for the indices \(-10,...,-35\). We had no need to go further in this notation. The next example should illustrate this:

\[
\text{hL33} = \zeta(-21,-3,-3)
\]

If there is ever any doubt about which variable indicates which object one can look in the corresponding library file (always included as a file with the extension .h in the directory in which the integrals reside) in the procedure ‘frombasis’.

For reasons of economy\(^8\) the \( \zeta \)-functions with a single negative index have a different notation. They are related to the constants \( \eta_k \) defined by

\[
\eta_k = \left(1 - \frac{1}{2k-1}\right)\zeta_k.
\] (7.1)

In the program we call these constants \( e3, e5, \ldots \).

In some cases we use a variable with a notation similar to the notation for the MZVs, except for that the character \( z \) is replaced by the character \( a \).

\[
aiajak = a(i,j,k)
\]

These are taken outside brackets, the contents of each bracket are what each \( \zeta \)-function \( z \) is multiplied by what this MZV is equal to. In the case that we fixed a basis this can be an expression that consists of symbols like we defined above. This holds also for long numbers. These are broken off by a backslash character and continued on the next line. The problem is usually that the number of digits in the fractions is somewhat smaller in \( \eta \)-notation than in \( \zeta \)-notation.

In exceptional cases we refer to \( \zeta \)-functions with negative indices. The most common notation for this in the literature is to put a bar over the number. This is however a notation that cannot be used in programs like FORM. Hence we use negative indices for the alternating sums there. For the symbolic variables we use the notation for the MZVs but with the character \( m \) between \( z \) and the number:

\[
zmlzm3z3 = z(-11,-3,3) = -\zeta(-11,3,3)
\]

The programs run in what we call integral notation. This means that the master expression has the index fields of the functions \( E, \zeta \) and \( \zeta \) in terms of the three letter alphabet \( \{0,1,-1\} \) for Euler sums and the two letter alphabet \( \{0,1\} \) for MZVs. This is then the way the outputs are presented. Actually, internally the whole string of indices is put together as one large ternary number for Euler sums and one large binary number for MZVs. This speeds up the calculation, but makes it virtually impossible to interpret intermediate results.

The outputs are presented in a method that one may consider unusual. In FORM it is often more efficient to have one big expression, rather than \( 2^{20} \) expressions as would be the case for the MZVs at \( w = 23 \). Hence the output contains functions \( \zeta \) with the indices of the corresponding MZV and each \( \zeta \) is multiplied by what this MZV is equal to. In the case of Euler sums a prime and only wanted to determine a basis, it will be an expression that consists of symbols like we defined above. In the case that we did the calculus modulus a prime and only wanted to determine a basis, this can be an expression that consists of symbols like we defined above. Often at the end of the program there is a list of the \( \zeta \) functions used. Because FORM will print the output in such a way that the functions \( \zeta \) are taken outside brackets, the contents of each bracket are what each \( \zeta \) function is equal to. With a decent editor it takes very few (< 4) edit commands to convert such output into the definition of \( 2^{20} \) table elements.

If this output should be used as input for other systems, this can be done, provided that the expressions do not cause memory problems. The format is in principle compatible with Pari/GP, Reduce and Maple. There may be a problem with large coefficients. FORM does not like to make output lines that are longer than a typical screen width. Hence they are usually broken up after some 75 characters. This holds also for long numbers. These are broken off by a backslash character and continued on the next line. The problem is usually that FORM places all white space at the beginning of the line and some programs may have problems with that. Hence one can use an editor to remove all white space (blanks and tabs) at the beginning of the lines.

The data mine consists of several parts. The main part is formed by the different data sets. The remainder files give information about how to use the data mine and links to other useful information and/or programs. The data are divided over a number of directories, each containing the results of one type of runs for a range of values of the weight. In each directory there are several types of files again. The log-files of the runs are stored. These contain the run time statistics and the output of the runs in text format. Then there are the table files. They are in text format and contain table definitions for FORM programs. Their extension is .prc. Some of these files have been split into several files because they become much to big to be handled conveniently. These tables can be read and compiled. Yet the case of the MZVs at \( w = 22 \) with its nearly 6 Gbytes can be too large for a system with ‘only’ 16 Gbytes. If one does not have a bigger machine to ones disposal, one should use either the binary .sav file or the .tbl file defined below.

The third type of files are the binary .sav files. They can be used to read in the complete tables without having to go through the compiler and without having to load the complete table as table elements (which needs also big compiler buffers). Finally we have created so-called tablebases which allow very fast access to individual elements. A tablebase is a type of database for large tables. They are particular to FORM and have been used with great success in a number of very large calculations. Their working is explained in Ref. \[66\] and the FORM manual. The tablebase files have traditionally the .tbl extension.

In each directory we have also the programs that were used to create the various files and in some cases some example programs.

There is another section in the data mine that contains pages in which it is explained how to manipulate the information in the files. Although many files are in text format it is not easy to manipulate a 4 Gbyte text file and hence it might become necessary to either

---

8. It turns out that the number of digits in the fractions is somewhat smaller in \( \eta \)-notation than in \( \zeta \)-notation.

9. The function \( \eta_k \) is the same as the function \( \zeta_k \). We need two different names because when we present the results the function \( \eta \) marks the brackets and the function \( \zeta \) marks the remaining basis elements.
use FORM and one of the binary files, or to use the STedi editor which has been used to manipulate these files on a computer with 16 Gbytes of main memory. Links are set to these programs. FORM programs are provided for the most common manipulations of the data. They contain much commentary. This should make it easy for the user to customize the programs should the need arise. The data mine is located at http://www.nikhef.nl/~form/datamine/datamine.html. Its structure is given in Fig. 2. In this figure we use the following names:

- **complete**: Complete expressions over the rational numbers.
- **modular**: Products of lower weight terms are dropped and the computation is performed modulus a large prime.
- **limited**: As modular but incomplete bases.
- **rational**: Complete expressions over the rational numbers.
- **other things**: Conventions, publications, help, links, etc.

The main problem with the data mine is its size. Many files are several Gbytes long. We have used bzip2 on most files, because it gives a better compression ratio than gzip, even though it is much slower, both in compressing and decompressing. But even with bzip2 the combined files are larger than 30 Gbytes.

All programs are FORM (or TFORM) codes. They will run with the latest versions of FORM (or TFORM). The executables of FORM can be obtained from the FORM web site: http://www.nikhef.nl/~form. Please remember the license condition: if you use FORM (or TFORM) for a publication, you should refer to Ref. [21].

8. FORM aspects

As mentioned the running of the programs posed great challenges for FORM and TFORM. This is not simply a matter of whether the system contains errors. It is much more a matter of whether the system deals with the problem in a sensible and efficient way. Where are the bottlenecks? What is inefficient? A clear example is the conversion between sum notation and integral notation. This can be programmed in one line:

```plaintext
repeat id H(?a,n?!{-1,0,1},?b) = H(?a,0,n-sig_(n),?b);
```

for going to integral notation and

```plaintext
repeat id H(?a,0,n?!{0,0},?b) = H(?a,n+sig_(n),?b);
```

for going to sum notation. It turns out that when one goes to large weights (for instance more than 20), this becomes very slow because it involves very much pattern matching. Considering also that the use of harmonic sums is becoming more and more popular it was decided to build two new commands in FORM for this transformation:

```plaintext
ArgImplode,H;
ArgExplode,H;
```

The first one converts $H$ to sum notation and the second one to integral notation. This made the program noticeably faster and easier to read.
Another addition to \textsc{form} concerns built-in shuffle and stuffle commands. One of the problems with shuffles is that the simple programming of it usually gives many identical terms. This means that the shuffle product of two MZVs can become very slow, which is illustrated by the following little program:

\begin{verbatim}
S n1,n2;
CF H,HH;
L F = H(3,5,3)*H(6,2,5);
ArgExplode,H;
Multiply HH;
repeat;
  id HH(a)*H(b)*H(c) = +HH(a,n1)*H(b)*H(n2,?c)
                   +HH(a,n2)*H(n1,?b)*H(c);
endrepeat;
id HH(a)*H(b)*H(c) = H(a,b,c);
.end
\end{verbatim}

Time = 37.38 sec Generated terms = 2496144
F Terms in output = 2146
Bytes used = 63176

By putting much combinatorics in the built-in shuffle statement we could solve most of these problems (although not all as the combinatorics can become very complicated). With the shuffle command the program becomes:

\begin{verbatim}
S n1,n2;
CF H,HH;
L F = H(3,5,3)*H(6,2,5);
ArgExplode,H;
Shuffle,H;
.end
\end{verbatim}

Time = 0.01 sec Generated terms = 5163
F Terms in output = 2146
Bytes used = 63176

This is a great improvement of course.

For the stuffle product things are much easier. There we have the complication that there are two definitions. One is the product used for the $Z$-sums and the other is the product used for the $S$-sums. We have resolved that by appending a $+$ for the $Z$-notation and a $-$ for the $S$-notation:

\begin{verbatim}
stuffle,Z+;
stuffle,S-;
\end{verbatim}

Not only did this make the program significantly faster, it also made it more readable.

This way the stuffle product of two Euler sums in integral notation becomes in principle (assuming that we are in integral notation):

\begin{verbatim}
ArgImplode,H;
#call convertHtoZ(H,Z)
Stuffle,Z+;
#call convertZtoH(Z,H)
ArgExplode,H;
\end{verbatim}

except for that in the actual program we substituted the contents of the two conversion procedures. Of course for MZVs the conversions are not needed and we can use just:

\begin{verbatim}
ArgImplode,H;
Stuffle,H+;
ArgExplode,H;
\end{verbatim}

A third improvement concerns the parallelization. The original parallelization of \textsc{tforn} [22] assumed the treatment of a single large expression of which the terms are distributed over the workers and later gathered in by the master. During the phase in which we execute a Gaussian elimination inside a group of identities, this is very inefficient, because we deal with many small expressions, each giving a certain amount of overhead when they are distributed over the worker threads. Hence it was decided to create a new form of parallelization in which the user tells the program that there are many small expressions coming. The reaction of the master thread is now to divide the expressions over the workers. It only has to tell each worker which expression to do next, after which the worker is
The conjecture is in accordance with the upper bound for the size of the basis being derived in Refs. [14].

There may be 2 basis elements of depth 9 for weight 27 with our runs to depth 7, the Zagier conjecture holds also at weight 27. This at most one third of the weight we can even say that it holds to weight 26. If we combine the findings in the thesis of Racinet [68] that spurious zeroes, we can say that it holds to weight 24. With the additional assumption that all (Lyndon) basis elements have a depth of 9.1. Checking some conjectures with the data mine

The possibility to perform the calculus modulus a prime number has existed in FORM since its first version. Much of it remained untested because these facilities had not been used extensively. It turned out to be necessary to redesign parts of it and add a few new features.

Other aspects of TFORM performed amazingly well. We have seen the program running with eight workers who all eight had to enter the fourth stage of the sorting simultaneously. This is rather rare even for single threads and only happens for very large expressions. It gives a bit of a slow down due to the great amount of disk accesses, but it all worked without any problems. The most impressive single module result observed was

| Time = 15720.03 sec | Generated terms = 1202653196013 |
| FF Terms in output = 1508447974 |
| substitution(7-sh)-7621 Bytes used = 36215474400 |

The execution time is that of the master. Actually the master spent 1000 CPU sec on this step and the eight workers each almost 200000 CPU sec.

One may wonder about the probability that calculations, done with a system under development, give correct answers. We have several remarks concerning this topic:

- Whenever FORM failed, it was always in a very obvious way, like crashing because it couldn’t interpret something.
- The full all-depth outputs from the MZVs up to $w = 22$ and the Euler sums up to $w = 12$ have been tested numerically by completely independent programs, run under PARI-GP [56].
- Because of both TFORM and the MZV programs being under development many programs have been run at least several times with different configurations and/or different orders of solving the equations.
- TFORM operates in a rather non-deterministic fashion. Terms are rarely distributed twice in the same way over the workers because the master serves the workers when they have finished a task and this is usually not in the same order. In the case of errors this would lead to different results in different runs.
- There are effects that are expected on the basis of extrapolation, like the pushdowns and the construction of a basis. If anything goes wrong, such effects are absent.
- If for instance a term gets lost in a calculation over the rational numbers, usually the output would have terms with fractions that are abnormally much more complicated than the others. This is due to the fact that in intermediate stages the coefficients are usually much more complicated than at the end. Such terms are spotted relatively easily.

9. Results

Armed with the vast amount of information contained in the data mine we start with having a look at a number of conjectures in this field. They concern the number of basis elements, either just as a function of the weight or as a function of weight and depth. We first check some conjectures made in the literature using the data mine and then describe the selection of the basis to represent the Euler sums and MZVs in the data mine.

9.1. Checking some conjectures with the data mine

Zagier conjecture. (See [2].) The number of elements in a Lyndon basis for the MZVs at weight $w$ is given by Eq. (A.13).

As far as we can check, the Zagier conjecture holds to weight 22. Assuming that in the modular calculus no terms were lost due to spurious zeroes, we can say that it holds to weight 24. With the additional assumption that all (Lyndon) basis elements have a depth of at most one third of the weight we can even say that it holds to weight 26. If we combine the findings in the thesis of Racinet [68] that there may be 2 basis elements of depth 9 for weight 27 with our runs to depth 7, the Zagier conjecture holds also at weight 27. This conjecture is in accordance with the upper bound for the size of the basis being derived in Refs. [14].

Hoffman conjecture. (See [12].) A Fibonacci-basis for the MZVs at a given weight $w$ is formed out of MZVs the index set of which is formed out of all words over the alphabet $[2, 3]$.

We could test the basis conjectured by Hoffman up to weight $w = 22$. If we take the sub-variety in which we only look at the Lyndon words made from the indices 2 and 3, we can even verify this Lyndon basis to weight 24. Because this basis is not centered around the concept of depth, we cannot use the partial runs at larger weights and limited depths for further validation.

Broadhurst conjecture. (See [12].) The number of basis elements of the Euler sums at fixed weight $w$ and depth $d$ is given by Eq. (3.4).
The number of basis elements for MZVs when expressed in terms of Euler sums in a minimal depth representation is given by Eq. (3.6). Another conjecture by Hoffman.

than the first part, because for this we need the results for the corresponding Euler sums. 5 complete verification to weight 17 and modular verification to weight 21. For depth 4 these numbers are weight 22 and weight 30 respectively.

Broadhurst–Kreimer conjecture. (See [13].) The number of basis elements of the MZVs at fixed weight w and depth d is given by Eq. (3.5). The number of basis elements for MZVs when expressed in terms of Euler sums in a minimal depth representation is given by Eq. (3.6).

The runs for the MZVs confirm this conjecture over a large range, cf. Tables 16, 17. The second part of the conjecture is harder to check than the first part, because for this we need the results for the corresponding Euler sums.

Another conjecture by Hoffman. (See [3].)

\[ H_{2,1,2,2,2,2} - 2H_{2,2,2,2,2} = 0, \]
\[ H_{2,1,2,2,2,3} - H_{2,2,2,2,2,2} = 0, \]
\[ H_{2,1,2,2,2,2,3} - H_{2,2,2,2,2,3,3} = 0, \]
\[ H_{2,1,2,2,2,2,3} - H_{2,2,2,2,2,2,2} = 0, \]
\[ H_{2,1,2,2,2,2,3} - H_{2,2,2,2,2,2,3} = 0. \]

We verified these relations up to weight \( w = 22 \). At \( w = 24 \) we checked the weight-24 part, since we have only the modular representation at this level.

There are identities for special patterns of indices as

\[ 2\zeta_{m,1} = m\zeta_{m+1} - \sum_{k=1}^{m-2} \zeta_{m-k}\zeta_{k+1}, \quad 2 \leq m \in \mathbb{Z}, \]

(9.6)

cf. [1,4] or

\[ \zeta_{[3,1]_n} = \frac{1}{2n+1}\zeta_{[2]_n} = \frac{1}{4n}\zeta_{[4]_n} = \frac{2\pi^{4n}}{(4n+2)!}, \]

(9.7)
conjectured in [2] and proven in [47]. Another relation is

\[ \zeta_{2,1,3|n} = \frac{1}{4n} \sum_{k=0}^{n} (-1)^k \zeta(4k-3) \left( 4k+1 \right) \zeta(4k+2) - 4 \sum_{j=1}^{k} \zeta(4j-1) \zeta(4j+3) \]  

(9.8)

conjectured in [19] and proven in [70]. For the Euler sums one finds, [71].

\[ \zeta_{3|n} = 8^n \zeta_{(-2,1)|n}. \]  

(9.9)

In Ref. [19] conjectures were given for special cases based on PSLQ,

\[ \zeta_{4,1,1|2} = \frac{3\pi^4}{16} \left[ \zeta(6,2) - 4 \zeta_5 \zeta_3 \right] - \frac{41\pi^6}{5040} \left[ \zeta_5^2 - \frac{77023\pi^6}{14414400} \right] + \frac{397}{8} \zeta_9 \zeta_3 + \zeta_5^4, \]  

(9.10)

\[ \zeta_{2,1,2,3,2} = \frac{75\pi^2}{32} \left[ \zeta_8 - 2 \zeta_7 \zeta_3 + \frac{34}{225} \zeta_5^2 + \frac{4528801\pi^10}{61297236000} \right] - \frac{825}{8} \zeta_7 \zeta_5, \]  

(9.11)

which we verified. A series of special relations for the Euler sums were conjectured in [19] based on PSLQ:

\[ \zeta_{2,1,-2,-2} = \frac{39}{128} \zeta_4 \zeta_3 - \frac{193}{64} \zeta_5 \zeta_2 + \frac{593}{128} \zeta_7, \]  

(9.12)

\[ \zeta_{2,-2,1,2} = \frac{9}{128} \zeta_4 \zeta_3 + \frac{447}{128} \zeta_5 \zeta_2 - \frac{1537}{256} \zeta_7, \]  

(9.13)

\[ \zeta_{1,-3,1|2} = -7 \left[ \alpha_9 - \frac{39}{64} \zeta_5 + \frac{1}{8} \zeta_4 \ln(2) \right] \zeta_3 + \left[ 2\alpha_4 - \frac{1}{4} \zeta_4 \right]^2 + 2 \left[ \alpha_4 - \frac{15}{16} \zeta_4 + \frac{7}{8} \zeta_3 \ln(2) \right]^2 - \frac{1}{32} \zeta_8. \]  

(9.14)

Here

\[ \alpha_n = \text{Li}_n(1/2) + (-1)^n \left[ \frac{\ln(2)}{n!} - \frac{\zeta_2 \ln(n-2)(2)}{2(n-2)!} \right]. \]  

(9.15)

These relations are verified analytically as well by our database. Relations (9.10)-(9.13) were also obtained in [58].
In Ref. [12] a series of relations was conjectured for weight \( w = 8 \ldots 12 \) and \( d = 3,4 \) for Euler sums being related to values \( \zeta(-|a_1|, -|a_2|) \):

\[
\begin{align*}
\xi_{3,-3, -3} &= 6\zeta(5, -1, -3) + 6\zeta(3, -1, -5) - \frac{315}{32} \ln(2)\zeta_3 + 6\zeta(5, -1) \\
&= \frac{40005}{128} \zeta_2 \zeta_3 - \frac{39}{64} \zeta_3^2 + \frac{1993}{256} \zeta_3 \zeta_2 + \frac{8295}{128} \zeta_4 - \frac{226369}{384} \zeta_9. \\
\xi_{3,-5, -3} &= \frac{1059}{80} \zeta(5, 3, 3) + 15\zeta(7, -1, -3) + 15\zeta(5, -1, -7) + \frac{701}{69} \zeta(5, -3) \\
&= \frac{6615}{256} \ln(2)\zeta_3 \zeta_5 - \frac{11852967}{256} \zeta_1 + \frac{301599}{128} \zeta_2 \zeta_9 \\
&= \frac{124943}{5888} \zeta_2^2 + \frac{1753577}{35328} \zeta_8 + \frac{2960103}{5120} \zeta_4 \zeta_7 + \frac{3405}{32} \zeta_5 \zeta_6. \\
\xi_{3,-1, -3} &= \frac{61}{27} \xi(-3, -1, -1) - \frac{14}{3} \zeta(-5, -1, -1) - \frac{185}{27} \zeta(-5, -1, -2) \\
&= \frac{163499}{22356} \zeta(-3, -1, -1) + \frac{2051}{54} \zeta(-7, -1, -1) - \frac{28}{9} \ln(2)\zeta_5 + \frac{35}{96} \ln(2)\zeta_3^2 \\
&= \frac{581}{64} \ln(2)\zeta_3^2 - \frac{8735}{576} \ln(2)\zeta_2 \zeta_3 - \frac{903}{64} \ln(2)\zeta_3 \zeta_4 \\
&= \frac{1441}{288} \zeta_2 \zeta_3^2 + \frac{476928}{36916435} \zeta_3 \zeta_5 + \frac{1907712}{36916435} \zeta_8.
\end{align*}
\]

(9.16)

(9.17)

(9.18)

(9.19)

These relations were verified using the current database. Eq. (9.19) is particularly interesting since it implies a relation between MZVs mediated by one term of the kind \( \zeta(-|a_1|, -|a_2|) \).

There is a series of theorems proven on the MZVs, which can be verified using the database. We used already the duality theorem [2].

For the MZVs a large variety of relations has been proven, which can be verified for specific examples using the data mine.

The first of these general relations is the sum theorem, Refs. [1,72].

\[
\sum_{i_1 + \ldots + i_k = n, i_1 > 1} \zeta_{i_1, \ldots, i_k} = \zeta_n.
\]

(9.20)

The sum theorem was conjectured in [40], cf. [39]. For its derivation using the Euler connection formula for polylogarithms, cf. [73].

Further identities are given by the derivation theorem [40,52]. Let \( I = (i_1, \ldots, i_k) \) any sequence of positive integers with \( i_1 > 1 \). Its derivation \( D(I) \) is given by

\[
D(I) = (i_1 + 1, i_2, \ldots, i_k) + (i_1, i_2 + 1, \ldots, i_k) + \ldots + (i_1, i_2, \ldots, i_k + 1).
\]

(9.21)

The derivation theorem states

\[
D(I) = \zeta_I (D(\tau(I))).
\]

(9.22)

Here \( \tau \) denotes the duality-operation (2.23). We call an index-word \( w \) admissible, if its first letter is not 1. The words form the set \( \mathcal{S}^D \). \( |w| = w \) is the weight and \( d(w) \) the depth of \( w \). For the MZVs the words \( w \) are build in terms of concatenation products \( x_0^{i_0-1} x_1^{j_1} x_0^{i_2-1} x_1 x_0^{i_3-1} x_1 \). The height of a word, \( h(w) \), counts the number of (non-commutative) factors \( x_0 x_1^n \) of \( w \). The operator \( D \) and its dual \( \tilde{D} \) act as follows [7],

\[
D x_0 = 0, \quad D x_1 = x_0 x_1, \quad \tilde{D} x_0 = x_0 x_1, \quad \tilde{D} x_1 = 0.
\]

Define an anti-symmetric derivation

\[
\partial_t x_0 = x_0 (x_0 + x_1)^{n-1} x_1.
\]

A generalization of the Derivation Theorem was given in [52,74]: The identity

\[
\zeta(\partial_t w) = 0
\]

holds for any \( n \geq 1 \) and any word \( w \in \mathcal{S}^D \). Further theorems are the Le–Murakami theorem [75], the Ohno theorem [76], which generalizes the Sum- and Duality Theorems, the Ohno–Zagier Theorem [77], which covers the Le–Murakami Theorem and the Sum Theorem, and generalizes a theorem by Hoffman [39,40], and the Cyclic Sum Theorem [78].
Finally, we mention a main conjecture for the MZVs. Consider tuples \( k = (k_1, \ldots, k_r) \in \mathbb{N}^r \), \( k_1 \geq 1 \). One defines

\[
\begin{align*}
Z_0 & := \mathbb{Q}, \\
Z_1 & := [0], \\
Z_w & := \sum_{|k|=w} \mathbb{Q} : \zeta(k) \subset \mathbb{R}.
\end{align*}
\]  

(9.24)

If further

\[
\begin{align*}
Z^{Go} & := \sum_{w=0}^{\infty} Z_w \subset \mathbb{R} \quad \text{(Goncharov)}, \\
Z^{Ca} & := \bigoplus_{w=0}^{\infty} Z_w \quad \text{(Cartier)}
\end{align*}
\]

(9.25) (9.26)

the conjecture states

(a) \( Z^{Go} \cong Z^{Ca} \). There are no relations over \( \mathbb{Q} \) between the MZVs of different weight \( w \).

(b) \( \dim Z_w = d_w \), with \( d_0 = 1 \), \( d_1 = 0 \), \( d_2 = 1 \), \( d_w = d_{w-2} + d_{w-3} \).

(c) All relations between MZVs are given by the extended double-shuffle relations [79], cf. also [80]. If this conjecture turns out to be true all MZVs are irrational numbers.

### 9.2. Selection of a basis

Thus far we have not specified which basis we have been using for the MZVs. In first instance, we actually let the program select the basis. The result was the collection of remaining elements after elimination of as many elements as possible. The ordering in the elimination process was such that the remaining elements would be minimal in depth and maximal in their sum notation. Hence \( Z_{20,2,1,1} \) would be preferred over \( Z_{18,4,1,1} \). As it turned out, all remaining elements had an index field which formed a Lyndon word. This is not really surprising due to the ordering. Unfortunately there was not much systematics found in these elements.

Next came the idea that if the Euler sums have a basis made out of Lyndon words of only negative odd indices, maybe one should investigate to which extent one can write a basis for the MZVs in terms of Lyndon words with positive odd indices only. It turns out that a number of elements can be selected with odd-only indices, but it is not possible for the whole basis. A number of basis elements needs at least two even indices.

**Definition.** \( L_w \) is the set of Lyndon words made out of positive odd-integer indices, with no index \( i = 1 \) at given weight \( w \).

We observed that Table 17 can be reproduced by basis elements with indices in \( L_w \). As mentioned, this is not a basis for the MZVs, but if we write as many elements of the basis as possible as elements of the set \( L_w \), the remaining elements of the basis have a depth that is at least two greater than the elements that are remaining in the \( L_w \) set and need at least two even indices. Additionally, it looks like that they can be written as an extended version of these remaining elements by adding two indices 1 at the end and subtracting one from the first two indices as in

\[
Z_{7,5,3} \to Z_{6,4,3,1,1}.
\]  

(9.27)

We have been able to construct bases with these properties all the way up to weight \( w = 26 \). The complete (non-unique) recipe for such bases is:

1. Construct the set \( L_w \) of all Lyndon words of positive odd integers excluding one that add up to \( w \).
2. Starting at lowest depth, write as many basis elements of the basis as possible in terms of elements of \( L_w \). Call the remaining elements in \( L_w \) at this depth \( R_{w}^{(D)} \).
3. At the next depth, two units larger than the previous one, write again as many basis elements of the basis as possible in terms of elements of \( L_w \) and construct \( R_{w}^{(D+2)} \).
4. Write the elements of the basis with depth \( D + 2 \) that could not be written as elements of \( L_w \) as 1-fold extended elements of \( R_{w}^{(D)} \).
5. Write the elements of the basis with depth \( D + 2 \) that could not be written as elements of \( L_w \) or 1-fold extended elements of \( R_{w}^{(D)} \) as 2-fold extended elements of what remains of \( R_{w}^{(D-2)} \), etc.
6. If we are not done yet, raise \( D \) by two and go back to step 3.

The concept of \( n \)-fold extension is defined by subtracting one from the first \( 2n \) indices and adding \( 2n \) indices with the value one at the end of the index set.

To illustrate this we give two examples. First the basis at weight \( w = 12 \):

\[
\begin{align*}
L_{12} & : \quad H_{9,3} \quad H_{7,5} \\
P_{12} & : \quad H_{9,3} \quad H_{6,4,1,1}
\end{align*}
\]
Table 18
Number of basis elements for MZVs as a function of weight, depth and extension (or pushdown). If there are several numbers, separated by commas, the first indicates the number of elements that came from $L_w$, the second the number of 1-fold extensions from depth $d = 2$, the third the number of 2-fold extensions from depth $d = 4$, etc. A single number refers to the elements of $L_w$.

<table>
<thead>
<tr>
<th>w/d</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>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td></td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td></td>
<td>4</td>
<td></td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td></td>
<td>4.1</td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
<td>3.2</td>
<td>1.2</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td></td>
<td>6.1</td>
<td>6.3</td>
<td>1.3</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td></td>
<td>3</td>
<td>10.1</td>
<td>4.7</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td></td>
<td>8</td>
<td>11.4</td>
<td>11.12</td>
<td>0.1</td>
<td>0.1</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td></td>
<td>8</td>
<td>14.2</td>
<td>8.6</td>
<td>4.11</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td></td>
<td>10</td>
<td>18.5</td>
<td>25.23</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>4</td>
<td></td>
<td>11</td>
<td>29.7</td>
<td>31.14</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td></td>
<td>14</td>
<td>25.2</td>
<td>31.14</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td></td>
<td>14</td>
<td>42.8</td>
<td>25.23</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td></td>
<td>14</td>
<td>42.8</td>
<td>25.23</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td></td>
<td>14</td>
<td>42.8</td>
<td>25.23</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td></td>
<td>14</td>
<td>42.8</td>
<td>25.23</td>
<td>14.22</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and next the basis at weight $w = 18$:

$L_{18}: \quad H_{15,3} \quad H_{13,5} \quad H_{11,7} \quad H_{9,3,3,3} \quad H_{7,5,3,3} \quad H_{7,3,5,3} \quad H_{7,3,3,5} \quad H_{5,5,5,3}$

$P_{18}: \quad H_{15,3} \quad H_{13,5} \quad H_{10,6,1,1} \quad H_{9,3,3,3} \quad H_{6,4,3,3,1,1} \quad H_{7,3,5,3} \quad H_{7,3,3,5} \quad H_{5,5,5,3}$

From the basis at weight 18 it should be clear why we put so much effort in obtaining the results for the Euler sums at weight 18, depth 6.

Because the construction does not tell which elements of $L_w$ to select the results are not unique. In fact quite a few selections are not possible because of dependencies between the elements of $L_w$. Hence the whole procedure requires a certain amount of experimenting before a good basis is found. In Appendix B we have tried to find a basis in which the elements that are taken from $L_w$ have the highest values when their index set is seen as a multi-digit number. Because of reasons being explained in the next section we call these bases ‘pushdown bases’.

We do not have complete runs for the weights $w = 27$ and $w = 28$. In these cases the elements with the greatest depth are missing. But we can go through the construction as far as possible and make predictions about the missing elements. It turns out that for both these weights a 2-fold extension is needed. For weight $w = 27$ this would be for depth 5 to depth 9 and for weight $w = 28$ for depth 4 to depth 8. This concept was not taken into account in the conjectures in Ref. [13]. Hence we formulate a new conjecture that not only specifies the number of elements for each weight and depth but also how many elements need how many extensions.

**Conjecture 2.** The number of basis elements $D(w, d, p)$ of MZVs with weight $w$, depth $d$, and pushdown $p$ is obtained from the generating function

$$
\prod_{w=3}^{\infty} \prod_{d=1}^{\infty} \prod_{p=0}^{\infty} (1 - x^w y^d z^p) D(w, d, p) = 1 - \frac{x^3 y}{1 - x^2} + \frac{x^{12} y^2 (1 - y^2 z)}{(1 - x^4)(1 - x^6)}
$$

solving for the coefficients of the monomials $x^w y^d z^p$.

This formula predicts the first $n$-fold extension ($n > 1$) at weight $w = 12n + 3$ and it will be to depth $d = 4n + 1$. The exception is the first extension at weight 12. We show this in Table 18.

It is a great pity that with the resources that were at our disposal we just could not get direct access to a double extension or pushdown. Extrapolating from the numbers in Table 15 indicates computer times of the order of half a year (for weight 28, depth 8) to more than a year (for weight 27, depth 9).
10. Pushdowns

As mentioned in the previous section, there are elements that as MZVs can only be written with a certain depth, while, when written in terms of Euler sums, can be written with a smaller depth. This phenomenon is called pushdown. The simplest example occurs at weight \( w = 12 \) and can be looked up in the tables for the Euler sums. It is

\[
Z_{6,4,1,1} = \frac{2107648}{15825} H_{-11,-1} + \frac{50048}{9495} H_{-9,-3} - \frac{117568}{237375} H_{-7,-5} + \frac{100352}{1583} \eta_2 H_{-9,-1} - \frac{3854}{1583} \eta_2^2 H_{-7,-3} + \frac{320}{57} \eta_2^2 H_{-7,-1} - \frac{64}{171} \eta_2^2 H_{-5,-1} - \frac{2535128220786914}{481025690578125} \eta_2^6
\]

in which we remind the reader that \( \eta_2 = H_{-2} \). The next equation is at weight \( w = 15 \) and is already considerably longer. The rhs of \( Z_{6,4,3,1,1} \) contains 49 terms when written in this form and some of the fractions consist of more than 100 decimal digits. The phenomenon of these pushdowns seems to be intimately connected with the doubling and generalized doubling relations. We have investigated this at the weight \( w = 12 \) system. This is the only system over which we have complete control, because we have the full results for all depths for all Euler sums up to this weight. If we run this system without the use of the doubling and generalized doubling relations there are three more elements left in the 'basis', see Table 5. Two are of depth 4 and one is of depth 2. And additionally there is no pushdown. We have investigated this at the weight \( w = 12 \) system. This is the only system over which we have complete control, because we have the full results for all depths for all Euler sums up to this weight. If we run this system without the use of the doubling and generalized doubling relations there are three more elements left in the 'basis', see Table 5. Two are of depth 4 and one is of depth 2. And additionally there is no pushdown. The element \( Z_{6,4,1,1} = H_{6,4,1,1} \) needs one of these extra elements at depth 4. If we use the doubling relations, but we do not use the GDRs, there is only one extra element of depth 4, but the pushdown does take place. If we use only the GDRs, there are no remaining elements beyond the regular basis and the pushdown takes place.

Unfortunately we cannot run this test for other weights. Not using the GDRs means that we cannot run at restricted depth, due to the phenomenon of leakage. Of course it is rather adventurous to make the statement that doubling is at the origin of the pushdowns, when we have only a single case, but there is more supporting evidence as we will see below.

The way we have presented the pushdown in (10.1), although correct, is not its most transparent form. One can rewrite it as many possible and obtain a much simpler representation. One can, for instance, write

\[
H_{-9,-3} = \frac{1055}{1024} \bigg[ -Z_{9,3} - \frac{185874}{5275} \eta_3 \xi_5 + \frac{37332}{1055} \eta_3 \xi_3 \bigg] + \frac{1024}{26375} H_{-7,-5} + \frac{18792}{5275} H_{-11,-1} - \frac{92649159488}{23101203125} \xi_6^2.
\]

(10.2)

Additionally, we introduce a new function \( A \) as

\[
A_{H_{-7,5,3}} = H_{7,5,3} - H_{-7,5,3} - H_{7,-5,3} + H_{-7,-5,3}.
\]

(10.3)

Notice that the last index is always positive. In terms of the \( Z \)-notation the function \( A \) is the sum over all \( Z \)-sums with an even number of negative indices, but the absolute values of the indices are identical to the indices of the \( A \)-function. We rewrite then

\[
H_{-7,-5} = -\frac{25}{3} A_{H_{7,5}} + \frac{1295}{2304} Z_{9,3} + \frac{461399}{15360} \eta_3 \xi_3 + \frac{3213}{128} \eta_3 \xi_3
\]

and finally the result for the pushdown becomes:

\[
Z_{6,4,1,1} = \frac{64}{27} A_{H_{7,5}} - \frac{7967}{1944} Z_{9,3} + \frac{1}{12} \xi_4 + \frac{11431}{1296} \eta_2 \xi_5
\]

\[
- \frac{799}{12} \eta_2 \xi_3 + 3 \xi_2 Z_{7,3} + \frac{7}{2} \xi_2^2 \eta_3 + 10 \xi_2 \xi_2 \xi_3
\]

\[
+ \frac{3}{5} \xi_2^2 Z_{5,3} - \frac{1}{5} \xi_2^2 \xi_3 - \frac{18}{35} \xi_2^2 \xi_3 - \frac{3213}{6081075} \eta_2 \xi_2.
\]

(10.6)

which is much simpler than Eq. (10.1). We see the same happening in the expression for \( Z_{6,4,3,1,1} \).

\[
Z_{6,4,3,1,1} = \frac{1408}{81} A_{H_{7,5,3}} + \frac{16663}{11664} Z_{9,3,3} + \frac{150481}{68040} Z_{7,3,5} + 10 \xi_3 Z_{6,4,1,1}
\]

\[
+ \frac{162823}{3888} \xi_3 Z_{9,3} - \frac{17}{20} \xi_3^4 - \frac{101437}{38880} \xi_3 Z_{7,3} - \frac{1520827}{38880} \xi_3
\]

\[
+ \frac{1408}{81} A_{H_{7,5,3}} + \frac{16663}{11664} Z_{9,3,3} + \frac{150481}{68040} Z_{7,3,5} + 10 \xi_3 Z_{6,4,1,1}
\]

\[
+ \frac{162823}{3888} \xi_3 Z_{9,3} - \frac{17}{20} \xi_3^4 - \frac{101437}{38880} \xi_3 Z_{7,3} - \frac{1520827}{38880} \xi_3
\]
there are 1365 finite
with at least one negative index that can be rewritten completely in terms of MZVs is given in Table 19. In Table 20 we show how the
second case can occur, it will for a large fraction of the
remaining element(s) that get(s) pushed down. But for all cases that we could check there is a single function
that are in modular arithmetic and without the products of lower weight
11. Special Euler sums
The discovery of the \( A \)-functions brings up a new point. Which Euler sums can be written as a linear combination of MZVs only? This
is of course a perfect question for a system like the data mine in which exhaustive searches are relatively cheap. At the same time we ask of course the question which \( A \)-functions can be written in terms of MZVs only. We should distinguish two cases:

- The object can be written in terms of MZVs that have at most the same depth as the object.
- The object must be a negative index. This occurs when there is already an \( A \)-function that is used in a pushdown. In that case many other \( A \)-functions may be rewritten in terms of this \( A \)-function and MZVs of the same depth or lower depth.

We find that whenever the second case can occur, it will for a large fraction of the \( A \)-functions of that depth. The number of \( H \)-functions
with at least one negative index that can be rewritten completely in terms of MZVs is given in Table 19. In Table 20 we show the same for the
\( A \)-functions. Here there are clearly many more. Actually a sizable fraction of the \( A \)-functions can be rewritten like this. For example, there are 1365 finite \( A \)-functions of \( w = 17 \), \( d = 5 \) of which 449 can be rewritten in terms of MZVs only.

Considering that a number of the Euler sums can be rewritten in terms of MZVs only, one may raise the question whether the
pushdowns can be rewritten in such a way that they do not have the \( A \)-functions, but rather have a single Euler sum in their right-hand
side. This turned out to be a difficult question to answer, because the pushdown at \( w = 21 \), \( d = 7 \) was very time consuming and took several days for each trial. At first the number of candidates was rather large. We could make a list of candidates in a way, similar to that of Table 19 for \( w = 21 \), \( d = 5 \) and see which Euler sums could be expressed in terms of MZVs and \( A_{7,5,3,3,3} \) which is the object that was used in the pushdown.\(^{11}\)

Unfortunately the results for \( w = 21 \), \( d = 5 \) are in modular arithmetic and without the products of lower weight

\(^{11}\) Originally we worked with \( A_{9,3,3,3,3} \) and it was only at a very late stage that we converted to \( A_{7,5,3,3,3} \). Hence a number of the ‘raw’ results still refer to \( A_{9,3,3,3,3} \).
objects. Trying several elements of the list gave negative results indicating that many objects that give only MZVs for the terms with the same weight may have terms that are products of Euler sums of a lower weight. Then, after constructing Table 19 we looked for patterns and we noticed that the only eligible elements for $w = 13$, $w = 15$, $w = 17$ are

$$Z_{3,-2,3} = H_{3,-2,3,3},$$
$$Z_{3,-4,3} = H_{3,-4,3,3},$$
$$Z_{3,-2,3,-2} = H_{3,-2,-2,3,3},$$
$$Z_{3,-6,3} = H_{3,-6,3,3},$$
$$Z_{3,-4,3,3} = H_{3,-4,3,3,3},$$
$$Z_{3,-2,3,-6,3} = H_{3,-2,-3,3,3}. \tag{11.1}$$

Trying to rewrite $Z_{3,-6,3,-6,3}$ in terms of $A_{7,5,3,3}$ by means of LLL (a 130 elements search) gave the desired result. Hence by now all pushdowns have been obtained as well in terms of MZVs as in terms of one single Euler sum only. Unfortunately the index field of these Euler sums seems to be completely unrelated to the index fields of our basis elements.

### 12. Outlook

The data mine has given us already much information and it may yield more yet. But the current results leave also many new questions. To name a few:

- Can the GDRs be derived and/or written in a simpler way?
- Why can the GDRs resolve the problem of 'leakage'?
- Why do we need the doubling relations at all?
- What is the relation between the doubling formula and the pushdowns?
- Is it possible to see which $A$-functions can be used for pushdowns without needing the Euler sums of the data mine?
- Can a pushdown basis be constructed without needing the MZVs of the data mine?

In addition there is some 'unfinished business'. We did not get more than partial evidence for the double pushdowns at weight 27 and weight 28. Although we can guess the basis at weight 27, an LLL search for the complete formula would involve more than 800 elements and probably more than 10 times the number of digits than what our current searches needed. Considering the asymptotic behavior of the LLL algorithm, this would mean at least $10^7$ times the computer time we needed for the current determinations. The data mine approach is also not very attractive. There we would need the Euler sums to weight 27, depth 9. This might need even more extra orders of magnitude in resources than for the LLL algorithm. What would be very welcome is an algorithm by which we can determine a (small) subset of the Euler sums that includes the $A$-functions and combine this subset with the MZVs. For the MZV part of these double...
The data mine will be extended whenever new and relevant results are obtained, there is a history page that shows additions and corrections. If others have interesting contributions, they should contact one of the authors.

Acknowledgements

The work has been supported in part by the research program of the Dutch Foundation for Fundamental Research of Matter (FOM), by DFG Sonderforschungsbereich Transregio 9, Computergestützte Theoretische Teilchenphysik, the European Commission MRTN HEPTOOLS under Contract No. MRTN-CT-2006-035505, J.V. would also like to thank the Humboldt foundation for its generous support, and DESY, Zeuthen and the University of Karlsruhe for its hospitality during this work. The runs for creating the data mine were done on the computer system of the Theoretical Particle Physics group (TTP) at the University of Karlsruhe and computers at DESY and Nikhef.

Appendix A. Fibonacci and Lyndon bases at fixed weight

In the past several bases have been considered for both the MZVs and the Euler sums. In some of these the concept of depth is not relevant and hence for the counting rules we should sum over the depth. We will discuss those bases in this appendix. For a number of these bases conjectures are formulated in the literature, which cannot be broken down fixing the depth. We hope, that the empirical discoveries we made in this paper for harmonic sums up to $w = 30$ will stimulate mathematical research and eventually lead to proofs of more far reaching theorems in the future. Here we regard the consideration of the embedding of the MZVs into the Euler sums of importance. Likewise one may consider ‘colored MZVs’ with even higher roots of unity [81] in the future, which have not been the objective of this paper.

The first 30 Padovan numbers.

<table>
<thead>
<tr>
<th>$w$</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>
</tr>
</thead>
<tbody>
<tr>
<td>$P_w$</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>
</tr>
<tr>
<td>$w$</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>$P_w$</td>
<td>12</td>
<td>16</td>
<td>21</td>
<td>28</td>
<td>37</td>
<td>49</td>
<td>65</td>
<td>86</td>
<td>114</td>
<td>151</td>
</tr>
<tr>
<td>$w$</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>$P_w$</td>
<td>200</td>
<td>265</td>
<td>351</td>
<td>465</td>
<td>616</td>
<td>816</td>
<td>1081</td>
<td>1432</td>
<td>1897</td>
<td>2513</td>
</tr>
</tbody>
</table>

The number of these basis elements is counted by the Padovan numbers, but start from the initial values $P_1 = P_2 = P_3 = 1$. Their generating function is

$$G_P(x) = \frac{1 + x}{1 - x^2 - x^3} = \sum_{k=0}^{\infty} x^k P_k.$$  (A.11)

They also obey a Binet-like formula. The first values are given in Table 21.

The above basis is of the Fibonacci type. Another basis of the Fibonacci type is the Hoffman basis [69] which consists of all elements of which the index field is made up from 2’s and 3’s only. If one uses the following construction it is easy to see that the number of basis elements follows the Padovan sequence:

$$w = 1 \ \emptyset$$
$$w = 2 \ \{2\}$$
$$w = 3 \ \{3\}.$$  (A.11)
Let us now turn to Lyndon bases for the MZVs. Using a Witt-type relation \[44\] the size of the basis is conjectured to be given by

\[ l(w) = \frac{1}{w} \sum_{d \mid w} \mu \left( \frac{w}{d} \right) P_d. \]

Here the sum runs over the divisors \( d \) of the weight \( w \) and \( P_d \) denotes the Perrin numbers \([45,46]\). They are given by the Binet-like formula

\[ P_n = \alpha^n + \beta^n + \gamma^n, \quad \text{with } \alpha, \beta, \gamma \text{ the roots of} \]

\[ x^3 - x - 1 = 0 \]

and can be derived from the generating function

\[ G(P_k, x) = \frac{3 - x^2}{1 - x^2 - x^3} = \sum_{k=0}^{\infty} x^k P_k. \]

The first values are given in Table 22.

For the basis different choices are possible, which yield equivalent representations. Here we choose the basis in terms of\( \zeta \)-values, with an index field which forms a Lyndon word (see Table 23). Our first choice consists of indices, which contain as widely as possible odd integers. In case of even weights in a series of cases also indices with only even numbers occur from\( w = 12 \) onwards, as e.g. for

\[ w = 18: \quad \zeta_{15.3}, \zeta_{13.5}, \zeta_{9.3.3.3}, \zeta_{7.5.3.3}, \zeta_{5.5.5.3}, \zeta_{3.2.2.2.2}, \zeta_{12.2.2.2}. \]

A second natural choice is to take the afore mentioned Hoffman basis and select from it only those elements of which the index field forms a Lyndon word. Because the algebraic relations for the product of basis elements of lower weight do not give objects that are closely related to the basis elements at the higher weight, this basis is not used very frequently.

As an example we consider the case \( w = 30 \) and calculate the size of the bases using the Witt formula \((A.13)\) resp. the number of Lyndon words made up by the letters 2 and 3 only with 2 an index field which forms a Lyndon word (see Table 23). We now calculate the number of Lyndon words for each of these contributions, with \( m_i = k_i + l_i \).

One obtains

\[ L_{(2,3)}(30) = \frac{1}{14} \left[ \frac{14!}{12!2!} - \frac{7!}{6!} \right] + \frac{1}{13} \left[ \frac{13!}{9!4!} \right] + \frac{1}{12} \left[ \frac{12!}{6!^2} - \frac{6!}{3!^2} - \frac{4!}{2!^2} + \frac{2!}{1!^2} \right] + \frac{1}{11} \left[ \frac{11!}{8!3!} \right] = 151. \]

Using \((A.13)\) the result is
A basis up to weight \( w = 17 \) for the MZVs was also constructed in \([82]\).

For the Euler sums the Fibonacci basis is counted by the Fibonacci numbers. When we consider also all divergent Multiple Zeta Values the Fibonacci sequence is merely shifted. It is easily shown that the divergent Euler sums can be represented by the convergent sums and the element \( \sigma_0 \). As in the MZV case we may span the vector space of the Euler sums by forming a basis, which includes products of lower weight basis elements contained in a Lyndon basis. One basis of this type, used in the \textit{summer} program \([10]\) reads

\[
\begin{align*}
w &= 1 \quad \ln(2) \quad (A.21) \\
w &= 2 \quad \zeta_2, \ln^2(2) \quad (A.22) \\
w &= 3 \quad \zeta_3, \zeta_2 \ln(2), \ln^3(2) \quad (A.23) \\
w &= 4 \quad \text{Li}_4(1/2), \zeta_3 \ln(2), \zeta_2^2, \zeta_2 \ln^2(2), \ln^4(2) \quad (A.24) \\
w &= 5 \quad \text{Li}_5(1/2), \zeta_5, \text{Li}_4(1/2) \ln(2), \zeta_3 \zeta_2, \zeta_3 \ln^2(2), \zeta_2 \ln^3(2), \zeta_2^2 \ln(2), \ln^5(2) \quad (A.25) \\
w &= 6 \quad \text{Li}_6(1/2), \zeta_5, -1, \text{Li}_5(1/2) \ln(2), \zeta_5 \ln(2), \text{Li}_4(1/2) \zeta_2, \text{Li}_4(1/2) \ln^2(2), \zeta_5^2, \zeta_3 \zeta_2 \ln(2), \zeta_3 \ln^3(2), \zeta_2^2, \zeta_2 \ln^2(2), \zeta_2 \ln^4(2), \ln^6(2), \text{etc.} \quad (A.26)
\end{align*}
\]

These bases are counted by the Fibonacci numbers \([42,83]\), \( f_{w+1} \), which obey the same recursion relation as the Lucas numbers, but with the initial conditions \( f_0 = 0, f_1 = 1 \). They are represented by the formula given by J.P.M. Binet (1843)\(^{12}\)

\[
f_d = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^d - \left( \frac{1 - \sqrt{5}}{2} \right)^d \right).
\]

result from the generating function

\[
G(f_k, x) = \frac{x}{1 - x - x^3} = \sum_{k=0}^{\infty} x^k f_k.
\]

The first values are given in Table 24.

Table 24

<table>
<thead>
<tr>
<th>( w )</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>
</tr>
</thead>
<tbody>
<tr>
<td>( f_w )</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>13</td>
<td>21</td>
<td>34</td>
<td>55</td>
</tr>
<tr>
<td>( w )</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>( f_w )</td>
<td>89</td>
<td>144</td>
<td>233</td>
<td>377</td>
<td>610</td>
<td>987</td>
<td>1597</td>
<td>2584</td>
<td>4181</td>
<td>6765</td>
</tr>
</tbody>
</table>

\[
l(30) = \frac{1}{30} [P_{30} - P_{15} - P_{10} - P_6 + f_5 + P_3 + P_2 - P_0] \\
= \frac{1}{30} [4610 - 68 - 17 - 5 + 5 + 3 + 2 - 0] = 151. \quad (A.20)
\]

The first words with indices out of \( I_w \) are \( H_{-1}(1) \) and \( H_{0,-1}(1) = H_{-2}(1) \) are chosen as basis elements.

**Conjecture 4.** With the above starting conditions, consider the index words at weight \( w \) to be

\[
I_w = \bigcup_{|a| = (w-1)} (-1, I_a) \cup \bigcup_{|b| = (w-2)} (-2, I_b).
\]

The basis elements for the Euler sums are then given by the \( \zeta \)-values with indices out of \( I_w \). The elements of which the index sets are a Lyndon word form a Lyndon basis.

The Fibonacci version of this basis seems to have been discovered independently by S. Zlobin, see Ref. \([71]\).

This construction is analogous to that by Hoffman in the case of MZVs. It also uses a 2-letter alphabet. The different decomposition of the weight \( w \), however, leads to a basis of different length. Again we may derive the length of the basis using the Witt formula \((A.52)\) or counting the basis elements as Lyndon words of the index set \((A.30)\). Let us give an example for \( w = 20 \).

\(^{12}\) The relation was known to Euler and Moivre.
20 = k_1 * 1 + l_1 * 2 = 18 * 1 + 1 * 2 = 16 * 1 + 2 * 2 = 14 * 1 + 3 * 2
= 12 * 1 + 4 * 2 = 10 * 1 + 5 * 2 = 8 * 1 + 6 * 2 = 6 * 1 + 7 * 2
= 4 * 1 + 8 * 2 = 2 * 1 + 9 * 2.

(A.31)

Similar to the non-alternating case one obtains

\[ L_{(-1, -2)}(20) = \frac{1}{19} \binom{19}{18} \binom{18}{16} \binom{16}{14} \binom{14}{12} \binom{12}{10} \binom{10}{8} \binom{8}{6} \binom{6}{4} \binom{4}{2} \binom{2}{0} \]
\[ + \frac{1}{15} \binom{15}{13} \binom{13}{11} \binom{11}{9} \binom{9}{7} \binom{7}{5} \binom{5}{3} \binom{3}{1} \]
\[ + \frac{1}{12} \binom{12}{10} \binom{10}{8} \binom{8}{6} \binom{6}{4} \binom{4}{2} \binom{2}{0} \]
\[ = \frac{1}{20} (15! - 123 - 7 + 3) = 750. \]

(A.32)

Likewise the Witt formula (A.52) yields

\[ l(20) = \frac{1}{20} [l_{20} - l_{10} - l_2]. \]

(A.33)

The above basis suffers from the same shortcoming as the Hoffman basis in that the concept of depth lacks relevance. Hence we did not use it.

In a similar way we can construct yet another Fibonacci basis:

**Conjecture 5.** With the starting conditions of (A.29), consider the index words at weight \( w \) to be

\[ I_w = \bigcup_{|a|=(w-1)} (-1, l_a) \cup \bigcup_{|b|=(w-2)} (0, 0, l_b). \]

(A.34)

The basis elements for the Euler sums are then given by the \( \zeta \)-values of indices \( I_w \). The elements of which the index fields are a Lyndon word and all indices are odd valued if \( w > 2 \) form a Lyndon basis.

The Lyndon basis of this construction happens to be the basis proposed in Ref. [12]. We can divide \( I_w \)

\[ I_w^{\text{odd}} \oplus I_w^{\text{odd}}, \]

(A.35)

with the indices in \( I_w^{\text{odd}} \) are all odd and the last index of \( I_w^{\text{odd}} \) even, all others odd. The Lyndon words of \( I_w^{\text{odd}} \), \( I_w^{\text{odd}} \), form the basis elements at weight \( w \) and they are counted by (A.52). Note, that the basis element at \( w = 2 \) is not odd, which is an exception.

As an illustration we consider the case \( w = 6 \). The following words are generated, where we assume the ordering \( 0 < 1 \) and let the digit 1 play the role of \(-1.\)

\[
\{000001, 000011, 001001, 001101, 011111\};
\{100001, 100101, 101011, 110001, 110101, 111011, 111111\}. \]

(A.36)

The Lyndon words are

\[
(000011) \equiv (-5, -1); \quad (001111) \equiv (-3, -1, -1, -1);
\]
\[
(000001) \equiv (-6); \quad (001101) \equiv (-3, -1, -2). \]

(A.37)

The Lyndon words with odd indices taken as index of an Euler sum are basis elements, which we express through the harmonic polylogarithms at argument \( x = 1, H_{-5, -1}(1) \) and \( H_{-3, -1, -1, -1}(1) \). On the other hand,

\[
H_{-6} = \frac{62}{35} H_{-2}^3, \]

(A.38)

\[
H_{-3, -1, -2} = H_{-5, -1} + H_{-2} H_{-3, -1} + \frac{452}{105} H_{-2}^2 = \frac{55}{18} H_{-3}^2 \]

(A.39)

do not belong to the basis.

The last Lyndon basis is the one we actually use in the programs. It is depth oriented and no element can be written as a linear combination of elements of lower depth or products of elements with lower weight. To weight \( w = 12 \) the complete basis for the finite elements is given by

\[
w = 1 \quad H_{-1}; \]
\[
w = 2 \quad H_{-2}; \]
\[
w = 3 \quad H_{-3}; \]
\[
w = 4 \quad H_{-3, -1}; \]
\[
w = 5 \quad H_{-5, -3, -1, -1}; \]

(A.40) (A.41) (A.42) (A.43) (A.44)
The first 20 Lucas numbers.

Table 25

<table>
<thead>
<tr>
<th>w</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>
</tr>
</thead>
<tbody>
<tr>
<td>lw</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>11</td>
<td>18</td>
<td>29</td>
<td>47</td>
<td>76</td>
<td>123</td>
</tr>
</tbody>
</table>

For the Lyndon basis the conjectured length is \[12\] (A.45)

\[w = 6 \quad H_{-5,-1}, H_{-3,-1,-1,-1};\]

\[w = 7 \quad H_{-7}, H_{-5,-1,-1}, H_{-3,-3,-1}, H_{-3,-1,-1,-1};\]

\[w = 8 \quad H_{-7,-1}, H_{-5,-3}, H_{-5,-1,-1}, H_{-3,-3,-3}, H_{-3,-1,-1,-1};\]

\[H_{-3,-1,-1,-1,-1};\]

\[w = 9 \quad H_{-9}, H_{-7,-1,-1}, H_{-5,-3,-1}, H_{-3,-3,-1}, H_{-3,-1,-1,-1}, H_{-3,-1,-1,-1,-1};\]

\[H_{-3,-3,-3,-1,-1,-1};\]

\[w = 10 \quad H_{-9,-1}, H_{-7,-3}, H_{-7,-1,-1}, H_{-5,-3,-3}, H_{-5,-1,-1,-1}, H_{-3,-3,-3,-3}, H_{-3,-1,-1,-1,-1};\]

\[H_{-3,-3,-1,-1,-1,-1};\]

\[w = 11 \quad H_{-11}, H_{-9,-1,-1}, H_{-7,-3,-1}, H_{-7,-1,-3}, H_{-5,-5,-1}, H_{-5,-3,-3,-3}, H_{-3,-3,-3,-1,-1}, H_{-5,-1,-1,-1,-1,-1}, H_{-5,-1,-1,-1,-1,-1};\]

\[H_{-3,-3,-1,-1,-1,-1,-1};\]

\[w = 12 \quad H_{-7,-5}, H_{-9,-3}, H_{-11,-1}, H_{-5,-1,-1,-1}, H_{-5,-3,-1,-1,-3};\]

\[H_{-5,-3,-3,-1,-1}, H_{-5,-5,-1,-1}, H_{-7,-1,-1,-3}, H_{-7,-1,-3,-1};\]

\[H_{-7,-3,-1,-1}, H_{-9,-1,-1,-1}, H_{-3,-3,-1,-1,-1};\]

\[H_{-3,-3,-1,-1,-1,-1}, H_{-5,-1,-1,-1,-1,-1};\]

\[H_{-5,-1,-1,-1,-1,-1,-1};\]

For the Lyndon basis the conjectured length is \[12\] (A.46)

(A.47)

(A.48)

(A.49)

(A.50)

(A.51)

The first values are given in Table 25. The case \(w = 1\) is special as two elements contribute. See also Table 26.

Table 26

Number of basis elements of the Lyndon basis for the Euler sums for fixed weight \(w\).

<table>
<thead>
<tr>
<th>w</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>
</tr>
</thead>
<tbody>
<tr>
<td>lw</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

\[l_w = 18, 25, 40, 58, 90, 135, 210, 316, 492, 750.\]
Appendix B. Pushdown bases

We have tried to select a basis in which the elements of the set $L_w$ are maximal and the extended elements are minimal. At the same time the extended elements should be Lyndon words. This means for instance that an element like $H_{5,5,3}$ cannot be extended and hence has to be part of the basis, even though it is the minimal element at weight 18. One could of course reverse the criteria. For the construction of the bases this does not really diminish the amount of work. In both cases there are elements that should be skipped because of linear dependencies. We call the basis below the ‘minimal pushdown basis’. In addition we have used the requirement that for the extended elements the corresponding $A$-function should be usable for a pushdown. This requirement we could enforce up to weight $w = 22$. For higher weights we do not have the information in the data mine, and hence we do not know whether this requirement can be achieved.

\begin{align}
P_2 &= H_2; \quad \text{(B.1)} \\
P_3 &= H_3; \quad \text{(B.2)} \\
P_5 &= H_5; \quad \text{(B.3)} \\
P_7 &= H_7; \quad \text{(B.4)} \\
P_8 &= H_{5,3}; \quad \text{(B.5)} \\
P_9 &= H_9; \quad \text{(B.6)} \\
P_{10} &= H_{7,3}; \quad \text{(B.7)} \\
P_{11} &= H_{11,3}, H_{5,3,3}; \quad \text{(B.8)} \\
P_{12} &= H_{9,3}, H_{6,4,1,1}; \quad \text{(B.9)} \\
P_{13} &= H_{13,3}, H_{7,3,3}, H_{5,5,3}; \quad \text{(B.10)} \\
P_{14} &= H_{11,3}, H_9, H_{5,3,3,3}; \quad \text{(B.11)} \\
P_{15} &= H_{15,3}, H_{7,3,5}, H_{9,3,3}, H_{6,4,3,1,1}; \quad \text{(B.12)} \\
P_{16} &= H_{11,5}, H_{13,3}, H_{5,5,3,3}, H_{7,3,3,3}, H_{8,6,1,1}; \quad \text{(B.13)} \\
P_{17} &= H_{17,3}, H_{7,3,5}, H_{9,3,5}, H_{9,5,3}, H_{11,3,3}, H_{5,3,3,3,3}, H_{6,6,3,1,1}; \quad \text{(B.14)} \\
P_{18} &= H_{13,5}, H_{15,3}, H_{5,5,3,3}, H_{7,3,3,5}, H_{7,3,5,3}, H_{9,3,3,3,3}, H_{10,6,1,1}, H_{6,4,3,1,1}; \quad \text{(B.15)} \\
P_{19} &= H_{19,3}, H_{9,3,7}, H_{9,5,5}, H_{11,3,5}, H_{11,5,3}, H_{13,3,3}, H_{5,3,5,3,3}, H_{7,3,3,3,3}, H_{6,6,5,1,1}, H_{8,6,3,1,1}; \quad \text{(B.16)} \\
P_{20} &= H_{13,7}, H_{15,5}, H_{17,3}, H_{7,3,5,5}, H_{7,5,3,3}, H_{7,7,3,3}, H_{9,3,3,5}; \quad \text{(B.17)} \\
P_{21} &= H_{21}, H_{9,5,7}, H_{9,9,3}, H_{11,3,7}, H_{13,3,5}, H_{13,5,3}, H_{15,3,3}, H_{6,5,3,3,3}, H_{5,5,3,3,3}, H_{7,3,3,3,3}, H_{7,3,3,5,3}, H_{9,3,3,3,3}, H_{6,6,5,1,1}, H_{10,4,5,1,1,1}; \quad \text{(B.18)} \\
P_{22} &= H_{15,7}, H_{17,5}, H_{19,3}, H_{7,5,7,3}, H_{7,7,3,5}, H_{9,3,5,5}, H_{9,3,7,3}; \quad \text{(B.19)} \\
P_{23} &= H_{23}, H_{11,7,5}, H_{11,9,3}, H_{13,3,7}, H_{13,5,5}, H_{13,7,3}, H_{15,3,5}, H_{15,5,3}; \quad \text{(B.20)} \\
P_{24} &= H_{17,7}, H_{19,5}, H_{21,3}, H_{7,7,7,3}, H_{9,7,3,5}, H_{9,7,5,3}, H_{9,9,3,3}; \quad \text{(B.21)}
for these two elements the numbers 6 and 4 seem to play a special role.

\[ P_{25} = H_{25}, H_{11.11.3}, H_{13.3.5.7}, H_{13.3.7.5}, H_{13.9.3}, H_{15.3.7}, H_{15.5.5}, H_{15.7.3}, \\
H_{17.3.5}, H_{17.3.3}, H_{19.3.3}, H_{7.3.7.5}, H_{7.5.3.7}, H_{7.7.3.3}, \\
H_{9.3.3.3.3}, H_{8.3.3.5.5}, H_{9.3.3.7.3}, H_{9.3.5.3.5}, H_{9.3.5.5.3}, H_{9.3.7.3.3}, \\
H_{10.5.3.3.5}, H_{9.5.3.3.3}, H_{9.7.3.7.3}, H_{11.3.3.3.5}, H_{11.3.3.5.3}, \\
H_{11.3.5.3.3}, H_{11.5.3.3.3}, H_{13.3.3.3.3}, H_{8.8.7.1.1}, H_{10.4.9.1.1}, \\
H_{10.6.7.1.1}, H_{10.8.5.1.1}, H_{12.2.9.1.1}, H_{5.3.3.5.3.3}, H_{5.5.3.3.3.3}, \\
H_{5.5.3.3.3.3}, H_{7.3.3.3.3.3}, H_{6.2.5.5.5.1.1}, H_{6.4.3.5.5.1.1}, H_{6.6.5.3.5.1.1}, \\
H_{6.6.5.3.5.1.1}, H_{6.6.5.3.5.1.1}, H_{6.6.5.3.5.1.1}; \]

\[ P_{26} = H_{17.9}, H_{19.7}, H_{21.5}, H_{23.3}, H_{7.7.7.5}, H_{9.5.9.3}, H_{11.3.9.3}, H_{11.5.3.7}, \\
H_{11.5.5.5}, H_{11.5.7.3}, H_{11.7.3.5}, H_{11.7.5.3}, H_{11.9.3.3}, H_{13.3.3.7}, H_{13.3.5.5}, \\
H_{13.3.7.3}, H_{13.5.3.5}, H_{13.5.5.3}, H_{13.7.3.3}, H_{15.3.3.5}, H_{15.3.5.3}, H_{15.5.3.3}, \\
H_{17.3.3.3}, H_{14.10.1.1}, H_{5.5.5.3.5.3}, H_{5.5.5.3.5.3}, H_{7.3.3.5.5.3}, H_{7.3.5.3.5.3}, \\
H_{7.3.5.5.3.3}, H_{7.3.7.3.3.3}, H_{7.5.3.3.5.3}, H_{7.5.3.5.3.3}, H_{7.5.5.3.3.3}, \\
H_{7.7.3.3.3.3}, H_{9.3.3.3.5.3}, H_{9.3.3.5.5.3}, H_{9.3.5.3.3.3}, H_{9.3.5.3.3.3}, \\
H_{9.5.3.3.3.3}, H_{11.3.3.3.3.3}, H_{8.2.7.7.1.1}, H_{8.4.5.7.1.1}, H_{8.4.7.5.1.1}, \\
H_{8.6.3.7.1.1}, H_{8.6.5.5.1.1}, H_{8.6.7.3.1.1}, H_{8.8.3.5.1.1}, H_{8.8.5.3.1.1}, \\
H_{10.2.5.1.1}, H_{10.2.5.1.1}, H_{5.3.3.3.3.3}, H_{6.2.3.3.5.1.1}, H_{6.2.5.3.5.3.1.1}, H_{6.4.3.3.3.3.1}, H_{6.4.3.3.3.3.1}; \]

The above bases are complete. For the following basis we miss the two elements at depth 9 due to limited computer resources. Yet the construction based on $L_{27}$ allows us to predict the last two elements:

\[ P_{27} = H_{27}, H_{11.7.9}, H_{13.11.3}, H_{13.5.3.9}, H_{15.5.7}, H_{15.7.5}, H_{15.9.3}, H_{17.5.5}, H_{17.7.3}, \\
H_{19.3.5}, H_{19.5.3}, H_{21.3.3}, H_{7.5.5.7.3}, H_{7.5.7.3.5}, H_{7.7.3.7.3}, H_{7.7.7.3.3}, \\
H_{9.3.9.3.3}, H_{9.5.3.5.5}, H_{9.5.3.7.3}, H_{9.5.5.3.5}, H_{9.5.5.5.3}, H_{9.5.7.3.3}, H_{9.7.3.3.5}, \\
H_{9.7.3.5.3}, H_{9.7.5.3.3}, H_{9.9.3.3.3}, H_{11.3.3.3.7}, H_{11.3.3.5.5}, H_{11.3.3.7.3}, H_{11.3.5.3.5}, \\
H_{11.3.5.5.3}, H_{11.3.7.3.3}, H_{11.5.3.3.5}, H_{11.5.3.5.3}, H_{11.7.3.3.3}, H_{11.3.3.3.5}, \\
H_{11.3.3.3.5}, H_{11.3.3.3.7}, H_{11.3.3.3.9}, H_{10.8.7.1.1}, H_{10.10.5.1.1}, \\
H_{12.2.1.1.1}, H_{12.4.9.1.1}, H_{12.6.7.1.1}, H_{12.8.5.1.1}, H_{16.2.7.1.1}, \\
H_{5.3.5.3.3.3}, H_{5.5.3.3.3.3}, H_{5.5.3.5.3.3}, H_{5.5.3.5.3.3}, H_{5.5.3.5.3.3}, \\
H_{7.3.3.3.3.5}, H_{7.3.3.3.5.3}, H_{7.3.3.3.5.3}, H_{7.3.3.5.3.3}, H_{7.3.5.3.3.3}, \\
H_{9.3.3.3.3.3}, H_{9.4.5.5.5.1.1}, H_{9.6.3.5.1.1}, H_{9.6.6.3.5.1.1}, H_{9.6.6.5.3.1.1}, \\
H_{9.8.2.3.7.5.1.1}, H_{9.8.2.5.3.7.1.1}, H_{9.8.2.5.5.1.1}, H_{9.8.2.5.7.1.1}, \\
H_{9.8.2.7.5.1.1}, H_{9.8.4.3.3.7.1.1}, \\
H_{7.5.7.5.3} \rightarrow H_{6.4.6.4.3.1.1.1.1}, H_{7.5.3.3.3.3} \rightarrow H_{6.4.3.3.3.3.1.1}. \]

We have selected the last two elements for the necessary extension on the basis of the appendix in the thesis by Racinet [68] in which for these two elements the numbers 6 and 4 seem to play a special role.

Although we have also results for $P_{28}$ in which the leading depth is missing, there are too many elements missing to give a reliable list of the basis elements. It should be remarked though that also for $P_{28}$ we expect a 2-fold pushdown from depth 8 to depth 4.

Appendix C. Explicit pushdowns

Below we list all pushdowns up to $w = 21$ and one at $w = 22$ with the mixing with terms of equal weight and depth in the left-hand side and all remaining Euler sums in the right-hand side. The function $A$ is defined in (10.3).

We only list that part of the pushdowns that we consider particularly interesting. The complete formulas can be found in the data mine in the programs part. The name of the file is pushdowns.h.
The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the PARI algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.

The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the LLL algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.

The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the LLL algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.

The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the LLL algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.

The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the LLL algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.

The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the LLL algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.

The above results used the available resources to their limit. The formula in (C.15) needed 45 hours of running time using the LLL algorithm. Seven of them could be derived with the data mine. Unfortunately for depth $d = 5$ objects we have only exact results up to weight $w = 17$ and for depth $d = 4$ we have only exact results up to weight $w = 22$.
It is also possible to express each pushdown in terms of a single Euler sum rather than an $A$-function. In a sense this is less telling. After all the $A$-function contains half of the terms of the doubling relation and the doubling relations seem to be at the origin of the pushdowns. Also we could not find much structure concerning which Euler sum(s) to select. There are often many possibilities. In the case of the $A$-functions one can make a unique selection: the $A$-function should have the same index field as the element of the set $L_w$ that represents the pushdown. Anyway, for completeness we give here a single Euler sum for each of the pushdowns. We have dropped all factors and terms which have MZVs of the same weight or products of MZVs with lower weight.

\begin{align}
A_{7,3,5} &= + \frac{7649}{143360} Z_{7,3,5} - \frac{7089}{143360} \zeta_5 Z_{7,3} - \frac{2097}{71680} Z^3 \zeta_5 Z_{3,5} \\
&\quad - \frac{116396017}{1083797} \zeta_8 Z_{7,3,5} + \frac{40960}{81059} \zeta_{13} Z_{7,3,5} + \frac{81059}{17680} Z^2 \zeta_{11}.
\end{align}

(C.17)

Of course more complete results can be found in the data mine.

References


J.A. Lappo-Danilevsky, Mémoires sur la Théorie des Systèmes Differentiels Linéaires, Chelsea, New York, 1953;


D. Zagier, Multiple zeta values, preprint.