Multiple Polylogarithms: A Brief Survey

Douglas Bowman and David M. Bradley

Abstract. We survey various results and conjectures concerning multiple polylogarithms and the multiple zeta function. Among the results we announce our resolution of several conjectures on multiple zeta values. We also provide a new integral representation for the general multiple polylogarithm, and develop a q-analogue of the shuffle product.

1. Introduction

In recent years, nested harmonic sums have attracted increasing attention in both the mathematics and physics communities. The sums occur within the context of knot theory and quantum field theory, yet their rich structure offers much to fascinate theoreticians in such diverse areas as algebra, number theory, and combinatorics. Multiple polylogarithms generalize the aforementioned nested sums, as well as the Riemann zeta function and the classical polylogarithm, while still retaining many interesting properties. Their study has led to many beautiful yet unproven conjectures, including evaluations at arbitrary depth discovered with the use of recently developed integer relations-finding algorithms and high precision numerical computations in the thousands of digits.

Multiple polylogarithms [13, 46, 48] are multiply nested sums of the form

\begin{equation}
\text{Li}_{s_1,\ldots,s_k}(z_1,\ldots,z_k) := \sum_{n_1>\ldots>n_k>0} \prod_{j=1}^{k} n_j^{-s_j} z_j^{n_j},
\end{equation}

where \( s_1,\ldots,s_k \) and \( z_1,\ldots,z_k \) are complex numbers suitably restricted so that the sum (1.1) converges. Instances of multiple polylogarithms have occurred in several disparate fields, such as combinatorics (analysis of quad-trees [41, 59] and of lattice reduction algorithms [37]), knot theory [24, 26, 25, 60, 61, 77], perturbative quantum field theory [7, 19, 20, 22] and mirror symmetry [55].

1991 Mathematics Subject Classification. Primary 33E20; Secondary 11G55, 11M99, 40B05.

Key words and phrases. Euler sums, multiple zeta values, polylogarithms, multiple harmonic series, quantum field theory, knot theory, Riemann zeta function.

The first author was supported in part by NSF Grant DMS-9705782.

The second author was partially supported by the University of Maine summer faculty research fund.

\( \odot 2000 \) American Mathematical Society
sophisticated work relating polylogarithms and their generalizations to arithmetic and algebraic geometry, and to algebraic $K$-theory [8, 27, 28, 46, 47, 48, 79, 80, 81].

Figuring prominently are the nested sums (1.1) in which each $z_j = 1$. These latter are now commonly referred to as multiple zeta values [12, 13, 17, 68, 82] and are denoted by

\[(1.2) \quad \zeta(s_1, \ldots, s_k) := \sum_{n_1 > \cdots > n_k > 0} \prod_{j=1}^{k} n_j^{-s_j}.\]

The study of such sums goes back to Euler [39], who showed that

\[(1.3) \quad 2\zeta(m, 1) = m\zeta(m + 1) - \sum_{j=1}^{m-2} \zeta(m - j)\zeta(j + 1), \quad 2 \leq m \in \mathbb{Z}.\]

It can be shown that the sum (1.2) is absolutely convergent in the region

\[\{(s_1, \ldots, s_k) \in \mathbb{C}^k : \sum_{j=1}^{r} \Re(s_j) > r \text{ for } 1 \leq r \leq k\}.\]

(The condition given in Proposition 1 of [83] is insufficient to guarantee absolute convergence.)

Define the depth of the multiple polylogarithm (1.1) to be the number $k$ of nested summations. A good deal of work on multiple polylogarithms, and more specifically multiple zeta values, has been motivated by the problem of determining which sums can be expressed (say polynomially with rational coefficients) in terms of other sums of lesser depth. Solving this question in complete generality is currently beyond the reach of number theory. For example, proving the irrationality of expressions such as $\zeta(5, 3)/\zeta(5)\zeta(3)$ appears to be impossible with current techniques. Nevertheless, considerable progress has been made with regard to proving specific classes of reductions, even at arbitrary depth. The first nontrivial success at arbitrary depth was the settling [12, 13] of Zagier’s conjecture [82]

\[(1.4) \quad \zeta(3, 1, 3, 1, \ldots, 3, 1) = 4^{-n}\zeta(4, 4, \ldots, 4) = \frac{2\pi^{4n}}{(4n + 2)!} \quad 0 \leq n \in \mathbb{Z},\]

in which the $2n$ and $n$ beneath the underbraces in (1.4) denote the depth of the respective multiple zeta values. Subsequent work (see eg. [16, 17, 18]) has largely focused on developing a suitable framework for dealing with ultimately periodic argument strings in general, and additionally sums of multiple zeta values whose set of argument strings is fixed by the action of certain subgroups of the group of permutations.

It is instructive to trace the development of the subject and see for oneself how ad hoc techniques and considerations have in many cases evolved into more systematic methods of ongoing interest. In this connection, one might begin by citing the partial fractions technique of Euler [39] and Nielsen [67], subsequently employed by many others eg. [10, 15, 52, 56, 64, 66, 76], and which Ohno recently parlayed in his exceedingly clever proof of the cyclic sum formula [57, 69]. Techniques based on elementary integration formulæ and identities for special functions tailored to specific examples eg. [9, 10, 38, 66] have evolved [11, 13] into quite general, sophisticated, and powerful analytic methods [16, 18]. The naïve
approach of deriving elementary series transformation identities and solving the
resulting systems of linear equations [15, 72]; used to prove reducibility results of
depth three or less, has been largely superceded (eg, by methods based on contour
integration [40]) and supplanted by considerations of the shuffle and stuffle [13] multi-
uplications, and relatedly the harmonic algebra and the algebra of quasi-symmetric
functions [53, 54, 57].

Computational issues—both numeric and symbolic—have also come into play.
Relations satisfied by multiple polylogarithms, and multiple zeta values in partic-
ular, can be exploited by symbolic computer algebra systems to prove reductions
of small weight [65]. (Here the weight of the multiple zeta value (1.2) is simply the
sum of the arguments $s_1 + \cdots + s_k$.) Interest in high-precision, rapid computation
of multiple zeta values [35, 36] (see also [13, §7.2]) has been stimulated by the ability
to numerically hunt for or rule out identities (to a high degree of probability) with
the aid of recently developed integer relations finding algorithms [42, 50, 62].

In addition to the as yet unsolved problem of classifying all possible relations-
ships between multiple zeta values at positive integer arguments, one can also con-
sider (1.2) as a function of the complex numbers $s_1, \ldots, s_k$ and consider ques-
tions regarding analytic continuation, trivial zeros, and values at the non-positive
integers. The analytic continuation of (1.2) in the case $k = 2$ was established
by Atkinson [5] via the Poisson summation formula, and later by Apostol and
Vu [3], who used the Euler-Maclaurin summation formula. Subsequently, Arakawa
and Kaneko [4] proved that if $s_2, \ldots, s_k$ are fixed positive integers, then (1.2) can
be meromorphically continued as a function of $s_1$ to the whole complex $s_1$-plane.
The analytic continuation of (1.2) as a function defined on $\mathbb{C}^k$ was established by
An independent approach due to Zhao [83] uses properties of Gelfand and Shilov’s
generalized functions [44]. Zhao also attempts a discussion of trivial zeros for $k \leq 3$.
To our knowledge, no-one has yet determined the trivial zeros of (1.2) for general
$k$.

The issue of values of (1.2) at the non-positive integers is subtle, since for $k > 1$
the result will in general depend on the order in which the respective limits are
taken. Thus, for example, if $n$ is a non-negative integer, $s(k, j)$ and $S(k, j)$ denote
the Stirling numbers of the first and second kind, respectively, and $B_j$ denotes the
$j$th Bernoulli number, then [2]

$$\lim_{s_1 \to -n} \lim_{s_2 \to 0} \cdots \lim_{s_k \to 0} \zeta(s_1, \ldots, s_k) = \frac{(-1)^{n+1}}{n+1} \sum_{j=1}^{n+1} \frac{(-1)^{k+j} j! S(n+1, j)}{k+j},$$

whereas

$$\lim_{s_k \to 0} \cdots \lim_{s_2 \to 0} \lim_{s_1 \to -n} \zeta(s_1, \ldots, s_k) = (-1)^k \delta_{n, 0} - \frac{1}{(k-1)!} \sum_{j=1}^{k} \frac{s(k, j) B_{n+j}}{n+j},$$

where $\delta_{n, 0} = 0$ if $n > 0$ and $\delta_{0, 0} = 1$. In particular (cf also [83])

$$\lim_{s_1 \to 0} \lim_{s_2 \to 0} \zeta(s_1, s_2) = \frac{5}{12}, \quad \text{but} \quad \lim_{s_2 \to 0} \lim_{s_1 \to 0} \zeta(s_1, s_2) = \frac{5}{12}.$$

We are unaware of any systematic treatment in the case of arbitrary non-positive
integer arguments.
1.1. Notation. Let $\sigma_1, \ldots, \sigma_k \in \{-1, 1\}$. We will have occasion to discuss the particular sums of the form

\begin{equation}
\text{Li}_{\sigma_1, \ldots, \sigma_k}(x) = \sum_{n_1 \ldots n_k \geq 0} x^{n_1} \prod_{j=1}^{k} n_j^{-\sigma_j} s_j^{n_j},
\end{equation}

in which $0 \leq x \leq 1$ is real and $s_1, \ldots, s_k$ are positive integers with $x = s_1 = \sigma_1 = 1$ excluded for convergence. Accepted practice dictates that (1.5) may be abbreviated by $\zeta_x(s_1, \ldots, s_k)$ with a bar placed over $s_j$ if and only if $\sigma_j = -1$. When $x = 1$, these are called Euler sums. Thus a multiple zeta value is an Euler sum with no alternations. We adopt the convention that $\zeta_0() := 1$ when no arguments are present ($k = 0$). We also drop the subscript $x$ when $x = 1$ since $\zeta_1(s_1, \ldots, s_k)$ agrees with (1.2) when each $s_j$ is bar-free. For example,

$$\zeta(\mathbb{Z}, 1) = \sum_{n=1}^{\infty} (-1)^n \prod_{m=1}^{n-1} \frac{1}{m}.$$ 

It will be convenient to abbreviate strings of repeated arguments by using exponents to denote concatenation powers. Then the first two members of (1.4) may be written $\zeta(\{3, 1\}^n) = 4^{-n}\zeta(\{4\}^n)$.

Finally, as customary the Gaussian hypergeometric function and the logarithmic derivative of the Euler gamma function are denoted by

$$F(a, b; c; x) = \sum_{n=0}^{\infty} x^n \prod_{j=0}^{n-1} \frac{(a + j)(b + j)}{(1 + j)(c + j)}$$ and $\psi = \frac{\Gamma'}{\Gamma},$

respectively. We also abbreviate the set of the first $k$ positive integers $\{1, 2, \ldots, k\}$ by $N_k$.

2. Stuffles

For the sake of brevity and simplicity, we shall restrict the discussion in this section to multiple zeta values. For a discussion of the more general polylogarithmic case, see [13].

As Hoffman [57] observed, one can view multiple zeta values as values of a homomorphism on a commutative $\mathbb{Q}$-algebra in two ways; the $\mathbb{Q}$-algebra multiplications have been referred to elsewhere [13] as “shuffle” and “stuffle.” It is conjectured that all relations between multiple zeta values are a consequence of the collision of the two multiplications, provided one admits the divergent sums (1.2) with $s_1 = 1$ (suitably renormalized) into the model. However, there seems little hope of proving this conjecture in the near future, and at present a wide variety of analytic, algebraic, and combinatorial techniques are used to prove identities for multiple zeta values.

Stuffle relations, or more simply stuffles—see §3.1 and §5 below for a discussion of shuffles—arise when one multiplies two nested series of the form (1.2) and expands the product distributively. Thus if $u$ and $v$ are (ordered) lists of positive integers, then

$$\zeta(u)\zeta(v) = \sum_{w \subseteq u \ast v} \zeta(w),$$
where \( u \ast v \) is the multiset defined by the recursion
\[
(su \ast tv) = s(u \ast tv) \cup t(su \ast v) \cup (s + t)(s \ast t), \quad 1 \leq s, t \in \mathbb{Z}.
\]
In (2.1) it is to be understood that if \( M \) is a multiset of lists and \( a \) is an integer, then \( aM \) denotes the multiset of lists obtained by placing \( a \) at the front of each list in \( M \). For example \((s, t) \ast u = \{(s, t, u), (s, t + u), (s, u, t), (s + u, t), (u, s, t)\} \) and correspondingly \( \zeta(s, t)\zeta(u) = \zeta(s, t, u) + \zeta(s, t + u) + \zeta(s, u, t) + \zeta(s + u, t) + \zeta(u, s, t) \).

Let \( f(|u|, |v|) \) denote the number of lists in \( u \ast v \). The recursive decomposition (2.1) shows that the generating function
\[
F(x, y) := \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f(m, n)x^m y^n
\]
satisfies the functional equation \( F(x, y) = 1 + xF(x, y) + yF(x, y) + xyF(x, y) \). It follows that \( F(x, y) = (1 - x - y - xy)^{-1} \) and hence that
\[
f(m, n) = \sum_{k=0}^{m} \binom{m}{k} \binom{n+k}{m} = \sum_{k=0}^{\min(m, n)} \binom{n}{k} \binom{m}{k} 2^k.
\]
One can also give a direct, combinatorial proof of (2.2) by considering how the indices interlace in the product of two nested series of the form (1.2).

There are interesting connections between stuffles, polyominoes, and codes which we briefly indicate. To begin, note that a stuffle counted by \( f(m, n) \) can be viewed as a pair \((\phi, \psi)\) of order-preserving injections
\[
\phi : N_m \rightarrow N_r, \quad \psi : N_n \rightarrow N_r
\]
where \( r \) is chosen so that \( \max(m, n) \leq r \leq m + n \) and the union of the images of \( \phi \) and \( \psi \) is all of \( N_r \). One can associate to such a pair a sequence of integers \( b_1, \ldots, b_m \) by defining \( a_1 = \phi(1) - 1 \) and \( a_j = \phi(j) - \phi(j-1) - 1 \) for \( 2 \leq j \leq m \) and then letting
\[
b_j = \begin{cases} -a_j & \text{if } \phi(j) \text{ is in the image of } \psi, \\ a_j & \text{otherwise} \end{cases}
\]
for each \( j \in N_m \). Since \( \phi \) is order-preserving, \( a_j \geq 0 \) for each \( j \in N_m \) and \( \sum_{j=1}^{m} |b_j| = \phi(m) - m \leq n \). Conversely, given a sequence of integers \( b_1, \ldots, b_m \) satisfying \( \sum_{j=1}^{m} |b_j| \leq n \), the pair \((\phi, \psi)\) is uniquely determined. Let \( p = \{|j : b_j < 0|\} \). We have \( r = m + n - p, \phi(1) = |b_1| + 1 \) and \( \phi(j) = \phi(j-1) + |b_j| + 1 \) for \( 2 \leq j \leq m \).

Put \( \psi(j) = \phi(j) \) if \( b_j < 0 \). The remaining values of \( \psi \) are determined by the requirement that it be an order-preserving injective map of \( N_n \) to \( N_r \). Thus, there is a one-to-one correspondence between the stuffles counted by \( f(n, m) \) and the sets of integer lattice points whose cardinalities satisfy
\[
(b_1, \ldots, b_m) \in \mathbb{Z}^m : \sum_{j=1}^{m} |b_j| \leq n \quad \text{and} \quad \left( b_1, \ldots, b_n \right) \in \mathbb{Z}^n : \sum_{j=1}^{n} |b_j| \leq m,
\]
the identity (2.3) holding in view of the obvious symmetry \( f(m, n) = f(n, m) \).

Define an \( n \)-dimensional polyomino formed by adding \( m \) coats to a single-celled polyomino, where a coat consists of just enough cells to cover each previously exposed \((n-1)\)-dimensional face. There is clearly a bijection between such polyominos and the second set of lattice points (2.3). The relationship is explored in greater detail in [45].
3. Integral Representations

3.1. The Drinfeld Integral. There is also a representation for multiple zeta values in terms of an “iterated” (Drinfeld) integral due to Kontsevich [82]. For real 0 ≤ x ≤ 1 and positive integers s₁, ..., sₖ with x = s₁ = 1 excluded for convergence, we have

\[ \zeta_{x}(s₁, \ldots, sₖ) = \int \prod_{j=1}^{k} \left( \frac{\partial \phi(j)}{\partial \phi(j^p)} \right) \frac{\partial \psi(j)}{\partial \psi(j^p)} \]  

where the integral is over the simplex

x > t^{(1)}₁ > \cdots > t^{(1)}ₙ > \cdots > t^{(k)}₁ > \cdots > t^{(k)}ₙ > 0,

and is abbreviated by

\[ \int_{0}^{x} \prod_{j=1}^{k} a^{s_j - 1} b, \quad a = dt/t, \quad b = dt/(1 - t). \]

Making the simultaneous change of variable t → 1 - t at each level of integration and then reversing the order of integration makes transparent the duality identity for multiple zeta values:

(3.3) \[ \zeta(s₁ + 2, \{1\}^{r₁}, \ldots, sₖ + 2, \{1\}^{rₖ}) = \zeta(rₖ + 2, \{1\}^{sₖ}, \ldots, r₁ + 2, \{1\}^{s₁}), \]

first conjectured in [52] and proved in [82].

A related integral representation enabled Ohno [68] to prove the following beautiful generalization of (3.3). Let

\[ S(p₁, \ldots, pₙ; m) := \sum_{c₁ + \cdots + cₙ = m} \zeta(p₁ + c₁, \ldots, pₙ + cₙ), \]

where the sum is over all non-negative integers c₁, ..., cₙ which sum to m. As in (3.3) define the dual argument lists

\[ p := (s₁ + 2, \{1\}^{r₁}, \ldots, sₖ + 2, \{1\}^{rₖ}) \]

and

\[ p' := (rₖ + 2, \{1\}^{sₖ}, \ldots, r₁ + 2, \{1\}^{s₁}). \]

Then [68] S(p; m) = S(p'; m). When m = 0, Ohno’s result reduces to (3.3). Another interesting specialization is obtained by taking p = (k + 1) and m = n - k - 1; one then deduces Granville’s theorem [49], originally conjectured independently by Courtney Moen [52] and Michael Schmidt [64]:

\[ \sum_{s₁ + \cdots + sₖ = n} \zeta(s₁, \ldots, sₖ) = \zeta(n), \]

where the sum is over all positive integers s₁, ..., sₖ which sum to n and s₁ > 1.

The iterated integral representation is also responsible for a second multiplication rule satisfied by multiple zeta values. Suppose that x, y ∈ ℝ and \( f_j : [y, x] \to ℝ \) are integrable functions for j = 1, 2, ..., n. It is customary to make the abbreviation

\[ \int_{0}^{x} \prod_{j=1}^{n} \alpha_j := \int_{x > t₁ > t₂ > \cdots > tₙ > y} \prod_{j=1}^{n} f_j(t_j) \ d t_j, \quad \alpha_j := f_j(t_j) \ d t_j, \]
with the convention that (3.4) is equal to 1 if $n = 0$ regardless of the values of $x$ and $y$. There is an alternative definition of iterated integrals which explains their name. For $j = 1, 2, \ldots, n$ again define the 1-forms $\alpha_j$ by $\alpha_j := f_j(t_j) \, dt_j$. Then put

$$
\int_y^x \alpha_1 \alpha_2 \cdots \alpha_n := \begin{cases} \int_y^x f_1(t_1) \, \int_y^{t_1} \, \alpha_2 \cdots \alpha_n \, dt_1 & \text{if } n > 0 \vspace{-0.1cm} \\
1 & \text{if } n = 0.
\end{cases}
$$

Expanding out this second definition, it is easy to see that it coincides with the definition as an integral over a simplex. Both definitions occur frequently in the literature.

Clearly the product of two iterated integrals of the form (3.4) consists of a sum of iterated integrals involving all possible interlacings of the variables. Therefore, if we denote the set of all $(m + n)!/m!n!$ permutations $\sigma$ of the indices $N_{m+n}$ satisfying $\sigma^{-1}(j) < \sigma^{-1}(k)$ for all $1 \leq j < k \leq m$ and $m + 1 \leq j < k \leq m + n$ by Shuff$(m,n)$, then we have the self-evident formula

$$
\left( \int_y^x \prod_{j=1}^m \alpha_j \right) \left( \int_y^x \prod_{j=m+1}^{m+n} \alpha_j \right) = \sum_{\sigma \in \text{Shuff}(m,n)} \int_y^x \prod_{j=1}^{m+n} \alpha_{\sigma(j)},
$$

and so define the shuffle product $\shuffle$ by

$$
\left( \prod_{j=1}^m \alpha_j \right) \shuffle \left( \prod_{j=m+1}^{m+n} \alpha_j \right) := \sum_{\sigma \in \text{Shuff}(m,n)} \prod_{j=1}^{m+n} \alpha_{\sigma(j)}.
$$

Thus, the sum is over all non-commutative products (counting multiplicity) of length $m + n$ in which the relative orders of the factors in the products $\alpha_1 \alpha_2 \cdots \alpha_m$ and $\alpha_{m+1} \alpha_{m+2} \cdots \alpha_{m+n}$ are preserved. The term “shuffle” is used because such permutations arise in riffle shuffling a deck of $m + n$ cards cut into one pile of $m$ cards and a second pile of $n$ cards.

The study of shuffles and iterated integrals was pioneered by Chen [30, 31] and subsequently formalized by Ree [74]. As with the case of shuffles, one can view an element of Shuff$(m,n)$ as a pair of order-preserving injections $(\phi, \psi)$ where now $\phi : N_m \to N_{m+n}$ and $\psi : N_n \to N_{m+n}$ have disjoint images. One can then define a vector $(a_1, \ldots, a_m)$ of non-negative integers by $a_1 = \phi(1) - 1$ and $a_j = \phi(j) - \phi(j-1) - 1$ for $2 \leq j \leq m$. Since $\phi$ is order-preserving, $a_j \geq 0$ for each $j \in N_m$ and $\sum_{j=1}^m a_j = \phi(m) - m \leq n$. Conversely, if we have such a vector of non-negative integers, then $\phi(1) = a_1 + 1$ and $\phi(j) = \phi(j-1) + a_j + 1$ for $2 \leq j \leq m$ defines an order-preserving injection $\phi : N_m \to N_{m+n}$, and hence a shuffle. Thus, there is a one-to-one correspondence between Shuff$(m,n)$ and the sets of non-negative integer lattice points whose cardinalities satisfy

$$
\left\{ (a_1, \ldots, a_m) \in \mathbb{Z}_{\geq 0}^m : \sum_{j=1}^m a_j \leq n \right\} = \left\{ (a_1, \ldots, a_n) \in \mathbb{Z}_{\geq 0}^n : \sum_{j=1}^n a_j \leq m \right\},
$$

the latter identity holding in light of the fact that Shuff$(m,n)$ is clearly symmetric in $m$ and $n$. A deeper study of the algebra and combinatorics of shuffles leads to an alternative proof of (1.4) and generalizations thereof; see §5.

### 3.2. A New Integral Representation

In light of the usefulness of the various integral representations, it may be of interest to give here a new integral representation for (1.1). The new representation appears to embody properties of
both the Drinfeld and partition integrals of [13], and therefore may be useful in proving certain results for multiple polylogarithms which have thus far withstood attacks based on traditional methods. The derivation employs MacMahon’s Omega operator, which discards terms with non-negative exponents from formal Laurent series in \( \lambda_1, \ldots, \lambda_k \). Thus, in view of (1.1), if \( 0 \leq x_1, \ldots, x_k \leq 1 \), we may write

\[
\operatorname{Li}_{s_1, \ldots, s_k}(x_1, \ldots, x_k) = \Omega \prod_{j=1}^{k} \sum_{n_j > 0} n_j^{-s_j} (x_j \lambda_j \lambda_{j-1}^{-1})^{n_j}, \quad \lambda_0 := 1
\]

\[
= \Omega \prod_{j=1}^{k} \operatorname{Li}_{s_j}(x_j \lambda_j \lambda_{j-1}^{-1})
\]

\[
= \Omega \prod_{j=1}^{k} \int_{u_1^{(j)} > \cdots > u_{s_j}^{(j)} > 0} \left( \prod_{r=1}^{s_j-1} \frac{du_r^{(j)}}{u_r^{(j)}} \right) \frac{x_j \lambda_j \lambda_{j-1}^{-1} du_s^{(j)}}{1 - x_j \lambda_j \lambda_{j-1}^{-1} u_s^{(j)}}
\]

\[
= \Omega \prod_{j=1}^{k} \int_{u_1^{(j)} > \cdots > u_{s_j}^{(j)} > 0} \left( \prod_{r=1}^{s_j-1} \frac{du_r^{(j)}}{u_r^{(j)}} \right) \sum_{m_j=1}^{\infty} (x_j \lambda_j \lambda_{j-1}^{-1})^{m_j} (u_{s_j}^{(j)})^{m_j-1} du_s^{(j)}
\]

\[
= \int_{\Delta(\mathfrak{s})} \left\{ \prod_{j=1}^{k} \left( \prod_{r=1}^{s_j-1} \frac{du_r^{(j)}}{u_r^{(j)}} \right) \right\} \sum_{m_1 > \cdots > m_k > 0} \prod_{j=1}^{k} (x_j u_{s_j}^{(j)})^{m_j} \frac{du_s^{(j)}}{u_s^{(j)}},
\]

where \( \Delta(\mathfrak{s}) \) denotes the set of all integration variables satisfying

\[
1 > u_1^{(j)} > u_2^{(j)} > \cdots > u_{s_j}^{(j)} > 0
\]

for \( j = 1, 2, \ldots, k \). Since \( 0 \leq y_j < 1 \) for each \( j = 1, 2, \ldots, k \) implies

\[
\sum_{m_1 > \cdots > m_k > 0} \prod_{j=1}^{k} y_j^{m_j}
\]

\[
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} y_1^{n_1+\cdots+n_k} y_2^{n_2+\cdots+n_k} \cdots y_k^{n_k}
\]

\[
= \frac{y_1}{1 - y_1} \cdot \frac{y_1 y_2}{1 - y_1 y_2} \cdots \frac{y_1 y_2 \cdots y_k}{1 - y_1 y_2 \cdots y_k},
\]

it follows that

\[
(3.8) \quad \operatorname{Li}_{s_1, \ldots, s_k}(x_1, \ldots, x_k) = \int_{\Delta(\mathfrak{s})} \prod_{j=1}^{k} \left\{ \tau \left( \prod_{m=1}^{r} x_m u_{s_m}^{(m)} \right) \prod_{r=1}^{s_j} \frac{du_r^{(j)}}{u_r^{(j)}} \right\},
\]

where \( \tau(x) := x/(1 - x) \).

4. Generating Functions

In many cases, generating functions provide the best means of stating reductions involving one or more parameters. A specific example of this which also illustrates how knowledge of the subject has progressed is given first. We then outline a systematic approach for tackling multiple zeta values with periodic argument lists, followed by additional examples to illustrate the richness of the theory.
4.1. Two-Parameter Symmetry. In connection with Euler’s result (1.3), Markett [64] derived
\[
\zeta(s,1,1) = \frac{1}{6} s(s+1)\zeta(s+2) - \frac{1}{2} (s-1)\zeta(2)\zeta(s) - \frac{s-4}{4} \sum_{n=0}^{s-4} \zeta(s-n-1)\zeta(n+3)
\]
\[+ \frac{1}{6} \sum_{n=0}^{s-4} \zeta(s-n-2) \sum_{m=0}^{n} \zeta(n-m+2)\zeta(m+2), \quad 3 \leq s \in \mathbb{Z},\]
via elementary but intricate series manipulations and partial fraction identities. An equivalent formula is proved in [10] using elementary facts about the dilogarithm, the polygamma function and the higher derivatives of the Euler beta function. For larger values of \(n\), the representation of \(\zeta(s,\{1\}^n)\) in terms of values of the Riemann zeta function becomes increasingly complicated. Nevertheless, there is an elegant generating function formulation which we restate here.

**Theorem 4.1 ([11]).** The bivariate formal power series identity
\[ (2) \quad \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} x^{m+1} y^{n+1} \zeta(m+2,\{1\}^n) \]
\[= 1 - \exp \left\{ \sum_{k=2}^{\infty} \frac{1}{k} (x^k + y^k - (x+y)^k) \zeta(k) \right\} \]
holds.

**Corollary 4.2.** Let \(n\) and \(s\) be non-negative integers with \(s \geq 2\). Then \(\zeta(s,\{1\}^n)\) lies in the polynomial ring \(\mathbb{Q}[\zeta(2),\zeta(3),\ldots,\zeta(s+n)]\).

By comparing coefficients of \(x^{s-1} y^{n+1}\) on both sides of (4.2), one sees that in fact, \(\zeta(s,\{1\}^n)\) is a rational linear combination of products of Riemann zeta values such that the sum of the arguments in each product is equal to \(s+n\). Moreover, Euler’s result (1.3) is an immediate consequence of comparing coefficients of \(x^{s-1} y^2\). Similarly, Markett’s formula (4.1) can be obtained most easily by comparing coefficients of \(x^{s-1} y^n\). Finally, as the right hand side of (4.2) is evidently symmetric in \(x\) and \(y\), the left hand side must also be. Thus Theorem 4.1 implies the special case
\[\zeta(m+2,\{1\}^n) = \zeta(n+2,\{1\}^m)\]
of the duality formula (3.3). It would be interesting to find a generating function formulation of duality at full strength.

4.2. Periodic Argument Lists. Results such as (1.4) and (4.2) suggest that one might profit from a more systematic study of multiple zeta values whose argument lists form an ultimately periodic sequence. This is indeed the case; such a study forms the basis of some of our current work in progress [18].

4.2.1. Period One. The case of all identical arguments is quite well understood. Nevertheless, there are a few items of interest worth recording here, in particular a connection to the problem of determining the number of unordered factorizations of an integer.

For \(\Re(s) > 1\), equation (1.2) implies
\[ (3) \quad \sum_{k=0}^{\infty} t^{ks} \zeta(\{s\}^k) = \prod_{j=1}^{\infty} \left( 1 + \frac{t^s}{j^s} \right). \]
If in \((4.3)\) we take \(s\) to be an even integer, say \(s = 2n\) where \(n\) is a positive integer, then we may rewrite \((4.3)\) in the form

\[
\sum_{k=0}^{\infty} (-1)^k 2^{kn} \zeta(\{2n\}^k) = \prod_{j=0}^{n-1} \sin(\pi t \rho^j),
\]

where \(\rho = e^{\pi i / n}\) and \(\sin x = \sin x / x\) for \(x \neq 0; \sin 0 := 1\). The identity \((4.4)\) is one of many possible generalizations of Euler’s formula for \(\zeta(2n)\), and moreover shows that \(\zeta(\{2n\}^k)\) is a rational multiple of \(\pi^{2kn}\).

Differentiating both sides of \((4.3)\) and equating coefficients yields the recurrence

\[
k \zeta(\{s\}^k) = \sum_{j=1}^{k} (-1)^{j+1} \zeta(\{s\}^k - j), \quad 0 \leq k \in \mathbb{Z}, \quad \Re(s) > 1,
\]

which is really just a special case of Newton’s formula

\[k \epsilon_k = \sum_{j=1}^{k} (-1)^{j+1} p_j \epsilon_{k-j}, \quad 0 \leq k \in \mathbb{Z},\]

relating the elementary symmetric functions and power sum symmetric functions

\[\epsilon_k = \sum_{j_1, \ldots, j_k \geq 0} \prod_{m=1}^{k} x_{j_m}, \quad p_k := \sum_{j \geq 0} x_j^k.\]

Substituting \(1/j^s\) for each indeterminate \(x_j\) yields \(\epsilon_k = \zeta(\{s\}^k)\) and \(p_k = \zeta(ks)\).

From \((4.5)\) it follows that if \(k\) is a positive integer and \(\Re(s) > 1\), then \(\zeta(\{s\}^k)\) lies in the polynomial ring \(\mathbb{Q}[\zeta(s), \zeta(2s), \ldots, \zeta(ks)]\). In fact, there is an explicit formula for \(\zeta(\{s\}^k)\) in terms of a sum over partitions of \(k\).

**Definition 4.3.** Let \(r\) be a non-negative integer and let \(\alpha = (\alpha_1, \alpha_2, \ldots)\) be a non-negative integer partition of \(r\). Let \(m_j = \# \{i : \alpha_i = j\}\) be the number of parts of size \(j\), and put \(c_\alpha = \prod_{j \geq 1} m_j \Gamma(-j)^{m_j}\). Furthermore, abbreviate \(r = \sum_{j \geq 1} \alpha_j\) by \(|\alpha|\) and \(\prod_{j \geq 1} p_{\alpha_j}\) by \(p_\alpha\).

In view of the generic relationship \([63]\)

\[
\sum_{k=0}^{\infty} \epsilon_k t^k = \exp \left\{ - \sum_{r=1}^{\infty} \frac{(-1)^r \rho^r}{r} t^r \right\} = \frac{1}{\prod_{\alpha} c_\alpha^{-1} p_\alpha},
\]

for \(\Re(s) > 1\) we therefore have

\[
\sum_{k=0}^{\infty} t^k \zeta(\{s\}^k) = \exp \left\{ - \sum_{r=1}^{\infty} \frac{(-1)^r \zeta(rs) t^r}{r} \right\} = \frac{1}{\prod_{\alpha} c_\alpha^{-1} \prod_{\alpha_j} \zeta(\alpha_j s)},
\]

i.e.

\[
\zeta(\{s\}^k) = (-1)^k \sum_{|\alpha| = k} c_\alpha^{-1} \prod_{\alpha_j > 0} \zeta(\alpha_j s).
\]

We note the following connection with factorisatio numerorum \([51]\). (See also \([29, 70, 78]\).) Let \(\alpha\) be as in Definition 4.3. Define the unrestricted divisor function associated with the partition \(\alpha\) by

\[d_\alpha(m) = \sum_{\prod_{j \geq 1} \delta_j^{\alpha_j m}} 1.\]
For example $d_{1,1}$ is the ordinary divisor function, and $d_2(m) = 1$ if $m$ is a perfect square and zero otherwise.

**Proposition 4.4.** Let $\tau_k(m)$ denote the number of unordered factorizations of $m$ into $k$ distinct factors. Then

$$\tau_k(m) = (-1)^k \sum_{|\alpha|=k} e^{-\lambda_1} d_{\alpha}(m).$$

**Proof.** Observe that for $\Re(s) > 1$,

$$\zeta\{s\}^k = \sum_{n_1, \ldots, n_k > 0} \prod_{j=1}^{k} n_j^{-s} = \sum_{m=1}^{\infty} \tau_k(m) m^{-s}.$$ 

Now compare coefficients of $m^{-s}$ in (4.6). □

**Example 4.5.** Since $\zeta\{s\}^2 = \frac{1}{2} \zeta^2(s) - \frac{1}{2} \zeta(2s)$, we get $\tau_2(m) = \frac{1}{2} d_{1,1}(m) - \frac{1}{2} d_2(m)$. In particular $\tau_2(12) = 3$.

4.2.2. Period Two and Beyond. In contrast with the situation in which all arguments are identical, much remains to be explored in the case of argument strings of period two and higher. In [13] and [16] differential equations were found to be a useful technique for analyzing the generating functions for period 2. We summarize here some results from [13] and [16] to indicate the richness and complexity of the resulting formulas arising from the solution of the associated fourth order differential equation.

**Definition 4.6.** For $0 \leq x \leq 1$ and $z \in \mathbb{C}$, let

$Y_1(x, z) := F(z, -z; 1; x)$,

$Y_2(x, z) := (1 - x) F(1 + z, 1 - z; 2; 1 - x)$,

$G(z) := \frac{1}{2} \{ \psi(1 + iz) + \psi(1 - iz) - \psi(1 + z) - \psi(1 - z) \}$.

**Theorem 4.7 ([13]).** Let $Y_1$ be as in Definition 4.6. Then for $0 \leq x \leq 1$ and $|z| < 1$,

$$\sum_{n=0}^{\infty} (-1)^n z^{4n} \zeta_x(\{3, 1\}^n) = Y_1(x, z) Y_1(x, iz).$$

**Theorem 4.8 ([16]).** Let $Y_1$, $Y_2$ and $G$ be as in Definition 4.6. Then for $0 \leq x \leq 1$ and $|z| < 1$,

$$\sum_{n=0}^{\infty} (-1)^n z^{n+2} 4^n \zeta_x(3, \{1, 3\}^n) = G(z) Y_1(x, z) Y_1(x, iz) - \frac{Y_1(x, iz) Y_2(x, z)}{4Y_1(1, iz)} + \frac{Y_1(x, z) Y_2(x, iz)}{4Y_1(1, iz)}.$$

Note that (4.7) proves (1.4). Similarly (4.8) proves

$$\zeta(3, \{1, 3\}^n) = 4^{-n} \sum_{k=0}^{n} \zeta(4k + 3) \zeta(\{4\}^{n-k})$$

$$= \sum_{k=0}^{n} \frac{2\pi 4^k}{(4k + 2)!} \left(-\frac{1}{4}\right)^{n-k} \zeta(4n - 4k + 3),$$
which escaped the extensive numerical and symbolic searches carried out in the preparation of [11, 12, 13]. Differentiation of (4.8) followed by a delicate analysis of the asymptotic behaviour of the requisite hypergeometric functions at their singular points proves [16] the reduction

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

conjectured in [11, 13].

The proof of Theorem 4.8 hinges on showing that both sides of (4.7) and (4.8) are annihilated by the same fourth order differential operator. In [13], computer algebra was used to establish this for (4.7). At the time, a conceptual proof was unavailable. Subsequently the present authors (see [16]) found a conceptual proof of the following more general result, which is perhaps best understood in the context of work going back to Orr [71] and Clausen [34] on differential equations satisfied by a product of hypergeometric series. The result is shown in [17] to be closely related to the combinatorial “shuffle” approach outlined in §5, and may be stated as follows.

**Lemma 4.9.** Let \( K \) be a differential field of characteristic not equal to 2 and let \( D \) be a derivation on \( K \). For each \( k \in K \), define a derivation \( D_k := kD \). Let \( t \) be a constant, and suppose that for some \( f, g, u, v \in K \) the differential equations \((D_f D_g + t)u = 0\) and \((D_f D_g - t)v = 0\) hold. Then \( uv \) is annihilated by the differential operator \((D_f^2 D_g^2 + 4t^2)\).

In particular, taking \( f(x) = 1 - x \), \( g(x) = x \) and \( t = z^2 \), given that \( Y_1 \) and \( Y_2 \) satisfy \((D_f D_g + z^2)y = 0\), Lemma 4.9 shows that each of the three linearly independent functions \( Y_1(x,z)Y_1(x,iz) \), \( Y_1(x,iz)Y_2(x,z) \), and \( Y_1(x,z)Y_2(x,iz) \) are annihilated by the operator \( D_f^2 D_g^2 + 4z^4 \). That \( L(x,z) \) and \( S(x,z) \) are annihilated by the same operator follows easily from the integral representation (3.1), whence Theorem 4.8 is proved.

Since \( D_f^2 D_g^2 + 4z^4 \) is a fourth order differential operator, one might legitimately ask in what context the fourth linearly independent solution \( Y_2(x,z)Y_2(x,iz) \) arises. It turns out that due to the double logarithmic singularity arising from the product of the underlying hypergeometric functions at \( x = 1 \), it is easier to ascribe a meaning to this solution in the case of alternating sums (1.5). Recalling the generating function

\[ A(z) := \sum_{n=0}^{\infty} z^n \zeta(\{T\}^n) = \prod_{j=1}^{\infty} \left( 1 + \frac{(-1)^j z}{j} \right) = \frac{\Gamma(1/2)}{\Gamma(1 + z/2)\Gamma(1/2 - z/2)} \]

from [11], we have

**Theorem 4.10 ([16]).** Let \( 0 \leq x \leq 1 \), and \( |t| < \infty \). Put \( z = (1 + i)t/2 \), \( s = (1 + x)/2 \), and let \( U(s,z) = Y_1(s,z) - zY_2(s,z) \), where \( Y_1 \) and \( Y_2 \) are as in Definition 4.6. Then,

\[ \sum_{n=0}^{\infty} \left[ t^{2n} \zeta_x(\{T,1\}^n) + t^{2n+1} \zeta_x(\{1,T\}^n) \right] = \frac{U(s,-z)U(s,iz)}{A(-z)A(iz)}. \]
Theorem 4.10 is a bivariate generalization of the conjecture [11, equation (14)] in the case \( x = 1 \), and may be viewed as an analytic extension of the purely combinatorial identity (5.4) below.

In recent work [18], the authors have greatly extended the differential equation approach. The authors have obtained results on more general generating functions which include not only multiple zeta values, but polylogarithmic and hyperlogarithmic [58] values in general. In fact, from the point of view of iterated integrals, arbitrary forms may occur in the iterated integrals studied. The differential equations are still present. The authors have classified various bases for the solutions of the differential equations, given matrices for change of basis, and found the explicit representations of the monodromy matrices of the associated differential equations. These results actually stand out with greater distinction in a more general setting. Taking arbitrary 1-forms on a manifold \( M \), an explicit homomorphism is obtained from \( \pi_1(M, x_0) \) into \( \text{GL}_n(\mathbb{C}) \). This gives rise to a transport between the manifold \( M \) and its principle bundle constructed from the representation into \( \text{GL}_n(\mathbb{C}) \). Finally these results can be cast yet more generally in the setting of differentiable spaces. Our homomorphism is similar to the celebrated homomorphism of K. T. Chen [30, 31, 32, 33] in that it is built out of a generating function of iterated integrals. The essential difference is that Chen’s homomorphism maps into a formal Lie group, while our homomorphism maps into \( \text{GL}_n(\mathbb{C}) \). Will our homomorphism give different information than Chen’s? We are currently investigating the geometric implications of our work in this area.

5. SHUFFLES AND CYCLIC INSERTION

As in [65] (cf. also [12, 74]) let \( A \) be a finite set and let \( A^* \) denote the free monoid generated by \( A \). We regard \( A \) as an alphabet, and the elements of \( A^* \) as words formed by concatenating any finite number of letters (repetitions permitted) from the alphabet \( A \). By linearly extending the concatenation product to the set \( \mathbb{Q}(A) \) of rational linear combinations of elements of \( A^* \), we obtain a noncommutative polynomial ring with the elements of \( A \) being indeterminates and with multiplicative identity 1 denoting the empty word.

The shuffle product (3.7) is alternatively defined first on words by the recursion

\[
\begin{align*}
&\forall w \in A^*, \quad 1 \shuffle w = w \shuffle 1 = w, \\
&\forall a, b \in A, \quad \forall u, v \in A^*, \quad au \shuffle bv = a(au \shuffle bv) + b(au \shuffle v),
\end{align*}
\]

and then extended linearly to \( \mathbb{Q}(A) \). One checks that the shuffle product so defined is associative and commutative, and thus \( \mathbb{Q}(A) \) equipped with the shuffle product becomes a commutative \( \mathbb{Q} \)-algebra, denoted \( \text{Sh}_\mathbb{Q}[A] \). Radford [73] has shown that \( \text{Sh}_\mathbb{Q}[A] \) is isomorphic to the polynomial algebra \( \mathbb{Q}[L] \) obtained by adjoining the transcendence basis \( L \) of Lyndon words to the field \( \mathbb{Q} \) of rational numbers.

The recursive definition (5.1) has its analytical motivation in the formula for integration by parts—equivalently, the product rule for differentiation. Thus, if we put \( a = f(t) dt, \quad b = g(t) dt \) and

\[
F(x) := \int_y^x (au \shuffle bv) = \left( \int_y^x f(t) \int_y^t u dt \right) \left( \int_y^x g(t) \int_y^t v dt \right),
\]
then writing \( F(x) = \int_y^x F'(s) \, ds \) and applying the product rule for differentiation yields

\[
F(x) = \int_y^x \left( f(s) \int_y^s u \left( \int_y^t v \, dt \right) \right) \, ds \\
+ \int_y^x g(s) \left( \int_y^s f(t) \int_y^t u \, dt \right) \, ds \\
= \int_y^x [a(u \shuffle bv) + b(au \shuffle v)].
\]

Alternatively, by viewing \( F \) as a function of \( y \), we see that the recursion (5.1) could equally well have been stated as

\[
\forall w \in A^*, \quad 1 \shuffle w = w \shuffle 1 = w, \\
\forall a, b \in A, \quad \forall u, v \in A^*, \quad ua \shuffle vb = (u \shuffle vb)a + (ua \shuffle v)b.
\]

Of course, both definitions are equivalent to (3.7).

The combinatorial proof [12] of Zagier’s conjecture (1.4) hinged on expressing the sum of the words comprising the shuffle product of \((ab)^m \shuffle (ab)^n\) as a linear combination of basis subwords. In [17] a more comprehensive study of the shuffle algebra \(\text{Sh}_Q[a, b]\) is undertaken, and as a consequence correspondingly deeper results for multiple zeta values are obtained. To highlight the most interesting of these results, we first recall the following

**Definition 5.1 ([12]).** For integers \( m \geq n \geq 0 \) let \( S_{m, n} \) denote the set of words occurring in the shuffle product \((ab)^m \shuffle (ab)^n\) in which the subword \( a^2 \) appears exactly \( n \) times, and let \( T_{m, n} \) be the sum of the \( m!/(2n)!(m-2n)! \) distinct words in \( S_{m, n} \). For all other integer pairs \( (m, n) \) it is convenient to define \( T_{m, n} := 0 \).

One then has

**Theorem 5.2 ([17]).** Let \( x \) and \( y \) be commuting indeterminates, and let \( m \) be a non-negative integer. In the commutative polynomial ring \((\text{Sh}_Q[a, b])[x, y]\) we have the shuffle convolution formula

\[
\sum_{k=0}^{m} x^k y^{m-k} \left[ (ab)^k \shuffle (ab)^{m-k} \right] = \sum_{n=0}^{\lfloor m/2 \rfloor} (4xy)^n(x+y)^{m-2n} T_{m, n}.
\]

A special case of Theorem 5.2 implies the intriguing shuffle factorization due to Broadhurst, and which in turn implies (1.4):

\[
A \left( \frac{z}{1-i} \right) \shuffle A \left( \frac{z}{1+i} \right) = M(z) \in (\text{Sh}_Q[a, b])[\langle z \rangle], \quad i^2 = -1,
\]

where

\[
A(z) := \sum_{n=0}^{\infty} (z^2ab)^n(1+za) \quad \text{and} \quad M(z) := \sum_{n=0}^{\infty} (z^4a^2b^2)^n(1+za+z^2a^2+z^3a^2b).
\]

The experts will recognize (4.9) as the analytic version of \( A(z) \) above, in which \( a = -dt/(1+t) \) and \( b = dt/(1-t) \). Similarly for \( M(z) \) and the left hand size of (4.10) when \( x = 1 \).
In addition, Theorem 5.2 plays a key role in a remarkable combinatorial generalization of (1.4) which we proceed to describe. Let \( S_{m,n} \) be as in Definition 5.1. Note that each word in \( S_{m,n} \) has a unique representation

\[(ab)^{m_0} \prod_{k=1}^{n} (a^2 b)(ab)^{m_{2k-1}} b(ab)^{m_{2k}},\]

in which \( m_0, m_1, \ldots, m_{2n} \) are non-negative integers with sum \( m - 2n \). Conversely, every ordered \((2n+1)\)-tuple \( (m_0, m_1, \ldots, m_{2n}) \) of non-negative integers with sum \( m - 2n \) gives rise to a unique word in \( S_{m,n} \) via (5.5). Thus, a bijective correspondence \( \varphi \) is established between the set \( S_{m,n} \) and the set \( C_{2n+1}(m - 2n) \) of ordered non-negative integer compositions of \( m - 2n \) with \( 2n + 1 \) parts. In view of the relationship (3.1) expressing multiple zeta values as iterated integrals, it therefore makes sense to define

\[ Z(\vec{s}) := \int_0^1 \varphi(\vec{s}), \quad \vec{s} \in C_{2n+1}(m - 2n), \]

where as in (3.2), we now identify the abstract letters \( a \) and \( b \) with the differential 1-forms \( \frac{dt}{t} \) and \( \frac{dt}{(1 - t)} \), respectively. Thus, if \( \vec{s} = (m_0, m_1, \ldots, m_{2n}) \), then

\[ Z(\vec{s}) = \int_0^1 (ab)^{m_0} \prod_{k=1}^{n} (a^2 b)(ab)^{m_{2k-1}} b(ab)^{m_{2k}} \]

\[ = \zeta(\{2\}^{m_0}, \{2\}^{m_1}, \ldots, \{2\}^{m_{2n-1}}, 1, \{2\}^{m_n}), \]

in which the argument string consisting of \( m_j \) consecutive twos is inserted after the \( j \)-th element of the string \( \{3,1\}^n \) for each \( j = 0, 1, 2, \ldots, 2n \). It turns out [17] that

\[ Z(\vec{s}) = \sum_{\vec{s} \in C_{2n+1}(m-2n)} Z(\vec{s}) = \frac{2\pi^{2m}}{(2m+2)!} \left( \frac{m+1}{2n+1} \right), \]

for all non-negative integers \( m \) and \( n \) with \( m \geq 2n \). The proof uses Theorem 5.2 at essentially full strength combined with some tricky generating functionology. Observe that equation (1.4) is the special case of (5.6) in which \( m = 2n \), since \( Z(\{0\}^{2n+1}) = \zeta(\{3,1\}^n) \).

A more compelling formulation of (5.6) can be given as follows. Again, let \( \vec{s} = (m_0, m_1, \ldots, m_{2n}) \) and put

\[ C(\vec{s}) := Z(\vec{s}) + \sum_{j=1}^{2n} Z(m_j, m_{j+1}, \ldots, m_{2n}, m_0, \ldots, m_{j-1}). \]

In other words, sum over all cyclic permutations of the argument list \( \vec{s} \). Then [17]

\[ \sum_{\vec{s} \in C_{2n+1}(m-2n)} C(\vec{s}) = Z(m) \times |C_{2n+1}(m-2n)| = \frac{\pi^{2m}}{(2m+1)!} \left( \frac{m}{2n} \right) \]

is an equivalent formulation of (5.6). Here, we have used

\[ Z(m) = \zeta(\{2\}^m) = \frac{\pi^{2m}}{(2m+1)!}, \quad 0 \leq m \in \mathbb{Z}, \]

which follows from (4.4). The cyclic insertion conjecture [13] can be restated as the assertion that \( C(\vec{s}) = Z(m) \) for all \( \vec{s} \in C_{2n+1}(m-2n) \) and integers \( m \geq 2n \geq 0 \). Thus, (5.7) reduces the problem to that of establishing the invariance of \( C(\vec{s}) \) on
$C_{2n+1}(m - 2n)$. It is likely that this remaining step can be accomplished using only the shuffle property of multiple zeta values in conjunction with (4.4).

6. Dimension Conjectures

Broadhurst [21] has conjectures concerning the size of various bases (graded by weight and depth) for expressing multiple zeta values in terms of either irreducible multiple zeta values, or irreducible Euler sums, and also for expressing Euler sums in terms of irreducible Euler sums. The adjunction of additional differential forms appears to simplify the problem at each stage. Thus, if $D(n, k)$ denotes the number of multiple zeta values of weight $n$ and depth $k$ in a minimal $\mathbb{Q}$-basis for reducing all multiple zeta values to a $\mathbb{Q}$-linear combination of products of basis multiple zeta values, it is conjectured that

$$\prod_{n \geq 3} \prod_{k \geq 1} (1 - x^n y^k)^{D(n, k)} \equiv 1 - \frac{x^3 y}{1 - x^2} + \frac{x^{12} y^2 (1 - y^2)}{(1 - x^3)(1 - x^6)}.$$  

However, if we allow Euler sums into the basis, letting $M(n, k)$ denote the minimal number of Euler sums of weight $\eta$ and depth $k$ needed to reduce all multiple zeta values to basis Euler sums, then

$$\prod_{n \geq 3} \prod_{k \geq 1} (1 - x^n y^k)^{E(n, k)} \equiv 1 - \frac{x^3 y}{(1 - x^2)(1 - xy)}.$$  

One can also consider the problem of reducing Euler sums in terms of basis Euler sums. Let $E(\eta, k)$ denote the minimal number of Euler sums of weight $\eta$ and depth $k$ required to reduce all Euler sums to basis Euler sums. It is conjectured that

$$\prod_{n \geq 3} \prod_{k \geq 1} (1 - x^n y^k)^{E(\eta, k)} \equiv 1 - \frac{x^3 y}{1 - x^2}.$$  

Adjoining forms associated with sixth roots of unity to the set of possible differential forms yields the multiple Clausen values [14], and here it is conjectured that the number $F(n, k)$ of irreducible multiple Clausen values of weight $\eta$ and depth $k$ is generated by

$$\prod_{n \geq 1} \prod_{k > 0} (1 - x^n y^k)^{F(n, k)} \equiv 1 - \frac{x^2 y}{1 - x}.$$  

7. $q$-Shuffles

Here we consider a $q$-analogue of the shuffle algebra discussed in §5. Let $A$ be a set, not necessarily finite, and let $\eta : A \to A$ be bijective. Now form the free monoid generated by $A$ and call it $A^*$ as before. Extend the action of $\eta$ to $A^*$ in the obvious way so that $\eta$ becomes an automorphism of $A^*$. Again regard $A$ as an alphabet, and the elements of $A^*$ as words formed by concatenating any finite number of letters (repetitions permitted) from the alphabet $A$. By linearly extending the concatenation product to the set $\mathbb{Q}(A)$ of rational linear combinations of elements of $A^*$, we obtain a non-commutative polynomial ring with the elements of $A$ being indeterminates and with multiplicative identity $1$ denoting the empty word. It is clear that $\eta$ now extends to an automorphism of $\mathbb{Q}(A)$. 
A $q$-shuffle algebra is defined to be the ordered pair $(\mathbb{Q}(A), \shuffle_q)$ where $\shuffle_q$ is a commutative and associative bilinear operator on $\mathbb{Q}(A)$ satisfying the identity

\begin{equation}
\begin{cases}
\forall w \in A^*, & 1 \shuffle_q w = w \shuffle_q 1 = w, \\
\forall a, b \in A, \forall u, v \in A^*, & au \shuffle_q bv = a(u \shuffle_q bv) + b(\eta(au) \shuffle_q v).
\end{cases}
\end{equation}

We denote a $q$-shuffle algebra over $A$ by $\text{Sh}_{\mathbb{Q}}[A]$. It will be observed that a $q$-shuffle algebra is a commutative $\mathbb{Q}$-algebra.

Our definition implies that there may be more than one way of writing the $q$-shuffle product of two words. For example, letting $a, b \in A$, it is easy to see that in $\text{Sh}_{\mathbb{Q}}[A]$ one has $a \shuffle_q b = ab + b\eta a = ba + a\eta b$. As the length of the words being multiplied increases the number of different expressions also grows.

The motivation for our definition of $\text{Sh}_{\mathbb{Q}}[A]$ is not difficult to see. As the shuffle algebra is motivated by the property (3.6) of iterated integrals, one wants a similar identity to hold for iterated Jackson $q$-integrals [43]. Recall the definition of a Jackson $q$-integral. For $x > 0$, let $f: [0, x] \to \mathbb{R}$ be Riemann integrable. The Jackson $q$-integral of $f$ on $[0, x]$ is defined by

\begin{equation}
\int_0^x f(t) \, dq := \sum_{n \geq 0} f(xq^n) xq^n(1-q).
\end{equation}

Because for any $0 < q < 1$ the sum on the right hand side of (7.2) is a Riemann sum for $\int_0^x f(t) \, dt$, it follows that the Jackson $q$-integral tends to the ordinary Riemann integral in the limit as $q$ approaches 1.

One defines iterated Jackson $q$-integrals in exactly the same way that ordinary iterated integrals are defined by (3.5). To this end, for $j = 1, 2, \ldots, n$ let $f_j: [0, x] \to \mathbb{R}$ and $\omega_j := f_j(t_j) \, dt_j$. Then put

\begin{equation}
\int_0^x \omega_1 \omega_2 \cdots \omega_n := \prod_{j=1}^n \int_0^{t_{j-1}} f_j(t_j) \, dt_j, \quad t_0 := x
\end{equation}

\begin{equation}
= \begin{cases}
\int_0^x f_1(t_1) \int_0^{t_1} \omega_2 \cdots \omega_n \, dt_1 & \text{if } n > 0 \\
1 & \text{if } n = 0.
\end{cases}
\end{equation}

Here the fact that the 1-forms on the right hand side of (7.3) are $q$-difference 1-forms implies that the integral on the left hand side of (7.3) is a $q$-iterated integral and not an ordinary iterated integral.

Iterating the definition of the $q$-iterated integral, one finds that

\begin{equation}
\int_0^x \omega_1 \cdots \omega_k = \sum_{n_1, \ldots, n_k \geq 0} f_1(xq^{n_1}) f_2(xq^{n_1+n_2-1}) \cdots f_k(xq^{n_1+\cdots+n_k-1}) \times q^{kn_k+(k-1)n_{k-1}+\cdots+2n_2+n_1}(1-q)^k a^k,
\end{equation}

but this is not a very convenient expression. To simplify (7.4) it helps to make the following change of indices:

\[ l_i := \sum_{j=k-i+1}^k n_j, \quad 1 \leq i \leq k. \]
Then (7.4) reduces to

\[
\int_0^x \omega_1 \cdots \omega_k = \sum_{0 \leq t_1 \leq t_2 \leq \cdots \leq t_k} (1 - q)^k \cdot xq^{t_1} f(xq^{t_1})xq^{t_2} f(xq^{t_2}) \cdots xq^{t_k} f(xq^{t_k}).
\]

This in turn motivates the definition of \(q\)-difference 1-forms by the equation

\[
\omega_i = \omega_i(t) = f_i(t) \, dt := f_i(t)(1 - q).
\]

Notice that with this definition, the \(q\)-integral reduces to a summation operator on these 1-forms:

\[
\int_0^x \omega = \sum_{j \geq 0} \omega(t_0q^j),
\]

which agrees with the original definition by virtue of \(t_0 = x\).

Now fix \(0 < q < 1\) and define the Rogers [75] \(q\)-difference operator \(\eta\) acting on continuous functions \(f : [0, \infty) \to \mathbb{R}\) by the equation \((\eta f)(x) := f(xq)\). Also define the \(q\)-derivative in the usual way:

\[
(D_q f)(x) := \frac{f(x) - f(xq)}{x(1 - q)}.
\]

The relevant fact about the \(q\)-derivative is the following well-known property:

\[
(D_q f)(x) = \frac{f(x) - f(xq)}{x(1 - q)}.
\]

We are ready to give the motivation for the \(q\)-shuffle product. The alphabet \(A\) now consists of \(q\)-difference 1-forms, and the automorphism \(\eta\) is Rogers’ \(q\)-difference operator. The action of \(\eta\) is extended to forms by the equation \(\eta \omega = \omega(qt) = f(qt)tq(1 - q)\). This of course defines a new form \(\omega' = \eta \omega\) by \(\omega' = g(t)t(1 - q)\), where \(g(t) = qf(tq)\). The alphabet \(A\) needs to be infinite to account for all the forms \(\eta^j \omega\) for \(j \in \mathbb{Z}\). The action of \(\eta\) is now extended to \(Q(A)\) in the obvious way. It clearly forms an automorphism of this algebra. We wish to define the \(q\)-shuffle product so that for \(u, v \in A^+\) the following equation is true:

\[
\int_0^x u \, \underline{\underline{\eta}} \nu = \left( \int_0^x u \right) \left( \int_0^x \nu \right).
\]

To accomplish this, one applies the \(q\)-analogue of the argument given in §5 for deriving the recursive definition of the shuffle product. Take \(a, b \in A\) and \(u, v \in A^+\). Put \(a = f(t) \, dt\), \(b = g(t) \, dt\) and

\[
F(x) := \int_0^x (au \, \underline{\underline{\eta}} \nu b) = \left( \int_0^x f(t) \, dt \right) \left( \int_0^x u \, dt \right) \left( \int_0^x g(t) \, dt \right) \left( \int_0^x v \, dt \right).
\]
Writing $F(x) = \int_0^x F'(s) \, ds$ where $F'(s) = (D_q F)(s)$, and applying the $q$-product rule for $q$-differentiation ($(D_q f g) = (D_q f) g + (\eta f)(D_q g)$) yields

$$F(x) = \int_0^x \left( f(s) \int_0^s u \left( \int_0^t v \, dq \right) \, dt \right) \, ds$$

$$+ \int_0^x g(s) \left( \int_0^s f(t) \, dq \right) \, ds$$

$$= \int_0^x \left[ a(u \cdot_\eta b) + b(\eta a \cdot_\eta v) \right],$$

where the first equality follows from (7.7) and the product rule for $D_q$. Hence the inductive definition of the $q$-shuffle product results. Notice that commutativity and associativity of $\cdot_\eta$ follow immediately from (7.8).

We will conclude with a few examples of the $q$-shuffle product illustrating how several equivalent sums can arise from this product. Taking $\omega_1 \cdot_\eta \omega_2 \cdot_\eta \omega_3$ using the inductive definition gives (among several possibilities):

$$\omega_1 \cdot_\eta \omega_2 \cdot_\eta \omega_3 = \omega_1 \omega_2 \omega_3 + \omega_2 (\eta \omega_1) \omega_3 + \omega_3 (\eta \omega_2) \omega_1$$

$$= \omega_1 \omega_2 \omega_3 + \omega_2 (\eta \omega_1 \omega_3) + \omega_3 (\eta \omega_2 \omega_1).$$

Writing $\omega_i = \omega_i(x)$, these equations translate into the easily verifiable generic series identities:

$$\sum_{0 \leq l_1} \omega_1(x^{l_1}) \sum_{0 \leq l_2 \leq l_3} \omega_2(x^{l_2}) \omega_3(x^{l_3})$$

$$= \sum_{0 \leq l_1, l_2 \leq l_3} \omega_1(x^{l_1}) \omega_2(x^{l_2}) \omega_3(x^{l_3}) + \sum_{0 \leq l_2 \leq l_3} \omega_2(x^{l_2}) \omega_1(x^{l_1+1}) \omega_3(x^{l_3})$$

$$+ \sum_{0 \leq l_2 \leq l_1 \leq l_3} \omega_2(x^{l_2}) \omega_3(x^{l_1}) \omega_1(x^{l_3+2})$$

$$= \sum_{0 \leq l_1 \leq l_2 \leq l_3} \omega_1(x^{l_1}) \omega_2(x^{l_2}) \omega_3(x^{l_3}) + \sum_{0 \leq l_2 \leq l_3} \omega_2(x^{l_2}) \omega_1(x^{l_1+1}) \omega_3(x^{l_3+1})$$

$$+ \sum_{0 \leq l_2 \leq l_1 \leq l_3} \omega_2(x^{l_2}) \omega_3(x^{l_1}) \omega_1(x^{l_3+1}).$$

Taking the $q$-shuffle product as acting on the non-commutative polynomials in forms, it follows that all these different expressions for the $q$-shuffle product tend to the ordinary shuffle product in the limit as $q$ approaches 1. Further results about $q$-shuffle algebras and their combinatorics will be given elsewhere.

References


[38] P. J. De Doelder, *On some series containing $\psi(x) - \psi(y)$ and $(\psi(x) - \psi(y))^2$ for certain values of $x$ and $y$*, J. Comput. Appl. Math. 37 (1991), 124–141.


University of Illinois at Urbana-Champaign, Department of Mathematics, 273 Altgeld Hall, 1408 W. Green St., Urbana, IL 61801 U.S.A.

E-mail address: bowman@math.uiuc.edu

Department of Mathematics and Statistics, University of Maine, 5752 Neville Hall, Orono, ME 04469-5752 U.S.A.

E-mail address: dbraden@math.ams.org