Thakur, D. S. (2009) "Relations Between Multizeta Values for $\mathbb{F}_q[t]$,"
International Mathematics Research Notices, Vol. 2009, No. 12, pp. 2318–2346
Advance Access publication March 17, 2009
doi:10.1093/imrn/rnp018

Relations Between Multizeta Values for $\mathbb{F}_q[t]$

Dinesh S. Thakur

University of Arizona, Tucson, AZ 85721

Correspondence to be sent to: thakur@math.arizona.edu

Despite the failure of naive analogs of the sum shuffle or integral shuffle relations, and despite the lack of understanding of analogs of many classical structures that exist in the corresponding theory in the number field case, the multizeta values defined by the author are proved (and conjectured) to satisfy many interesting and combinatorially involved identities. The connections of these multizeta values with iterated extensions of Carlitz–Tate t-motives, analogs of Ihara power series, and Deligne–Soulé cocycles, etc., make it an interesting challenge to understand all the identities and discover the other relevant underlying structures.

Introduction

The multizeta values introduced and studied originally by Euler have been pursued again recently with renewed interest because of their emergence in studies in mathematics and mathematical physics connecting diverse viewpoints. They occur naturally as coefficients of the Drinfeld associator, and thus have connections to quantum groups, knot invariants, and mathematical physics. They also occur in the Grothendieck–Ihara program to study the absolute Galois group through the fundamental group of the projective line minus three points and related studies of iterated extensions of Tate motives, Feynman path integral renormalizations, etc. We refer the reader to papers on this subject by Broadhurst, Cartier, Deligne, Drinfeld, Écalle, Furusho, Goncharov,

Received September 18, 2008; Revised January 23, 2009; Accepted January 29, 2009
Communicated by Prof. Barry Mazur

Hoffman, Kreimer, Racinet, Terasoma, Waldschmidt, Zagier, Zudilin to mention just a few names.

Having learned about these rich interconnections at the Arizona Winter school, the author, in 2002, defined and studied two types of multizeta values [11, Section 5.10] for function fields, one complex valued (generalizing special values of Artin–Weil zeta functions) and the other with values in Laurent series over finite fields (generalizing Carlitz zeta values). For the $\mathbb{F}_q[t]$ case, the first type was completely evaluated in [11] (see [9] for a study in the higher genus case), for both types some identities were established, and for the second type failure of the shuffle identities was noted. Because of the failure, some other variants of the second type were also investigated in [11]. Here, we deal only with the second type, and we restrict attention exclusively to $\mathbb{F}_q[t]$.

In [4], we introduced "degenerate multizeta values" to remedy the sum shuffle failure and gave period interpretations to multizeta values in terms of explicit iterated extensions of the Carlitz–Tate t-motives in the sense of Anderson's t-motives (related to Drinfeld's shtukas) [1], the degenerate case needing field of definition which is inseparable extension of the base.

In this paper, we give a strong evidence that the introduction of the degenerate multizeta values is not necessary, in the sense that the span of the originally defined multizeta values is also an algebra because of new kind of (combinatorially involved) shuffle-type identities (i.e., giving the product of multizeta values as the sum of multizeta values) involving digit expansions of the arguments. We see such "digit phenomena" in many other arithmetic aspects in function fields, such as (i) the Galois group study via Ihara power series [2, 6] (this represents the étale side at the meta-abelian level in contrast to the DeRham–Betti side at the nilpotent level giving the multizeta story in the classical fundamental group story mentioned above), (ii) exceptional orders of vanishing of zeta [11, 5.4], (iii) the description of the divided power series corresponding to the zeta measure [11, 5.7], (iv) the zero distribution of the zeta [11, 5.8], and (v) the functional equations for the gamma [11, 4.6]. (Only in the last case, we seem to have understood this digit influence well.)

In contrast to the classical division between the convergent versus the divergent (normalized) values, all the values are convergent in our case. In place of the sum or the integral shuffle relations, we have different kinds of relations: the shuffle-type relations with $\mathbb{F}_p$-coefficients and the Euler-type relations with $\mathbb{F}_p(t)$-coefficients. (Classically, of course, there is no such distinction, the rational number field being the prime field in that case.)

As an application of the formulas in this paper, we can derive many transcendence results for multizeta values. This will be treated elsewhere. Because of the recent advance [3, 10] in the relevant transcendence theory, we are in a curious situation: In contrast to the classical case, we know that all the identities are motivic, but again in contrast to the classical case, we do not know a simple description of all identities even conjecturally, while classically there is a precise conjecture of what all the identities are. We refer the reader to Cartier's *Séminaire Bourbabki,* no. 885 (2000), *Astérisque*, no. 282 (2002) for survey, several references till then and historical comments (in addition to the papers by the authors mentioned at the start for more recent developments).

We hope that eventually these relations and their connections with the relevant structures in the function field arithmetic will be understood better.

The first section of this paper introduces the objects of study. The second section states various results on the relations between multizeta values, while the third section develops tools and provides proofs. While the first three sections of this paper give definitions, theorems, and proofs, the last two sections provide observations, guesses from numerical calculations (in particular, complete conjectural recipe, when $q = 2$, to express the product of two zeta values as the sum of multizeta values), as well as mixture of heuristics and precise arguments.

1 Multiple Zeta Values Over $\mathbb{F}_q[t]$

1.1 Notation

$$\begin{aligned}
\mathbb{Z} &= \{\text{integers}\},\\
\mathbb{Z}_+ &= \{\text{positive integers}\},\\
q &= \text{a power of a prime } p,\\
A &= \mathbb{F}_q[t],\\
A_+ &= \text{monics in } A,\\
K &= \mathbb{F}_q(t),\\
K_\infty &= \mathbb{F}_q((1/t)) = \text{completion of } K \text{ at } \infty,\\
\mathbb{C}_\infty &= \text{completion of algebraic closure of } K_\infty,\\
[n] &= t^{q^n} - t,\\
D_n &= \textstyle\prod_{i=0}^{n-1} \left(t^{q^n} - t^{q^i}\right),\\
\ell_n &= \textstyle\prod_{i=1}^{n} \left(t - t^{q^i}\right) \;=\; (-1)^n L_n,\\
\tilde{\pi} &= \text{fundamental period of Carlitz module},\\
\text{“even”} &= \text{multiple of } q-1, \text{ and}\\
\deg &= \text{function assigning to } x \in K_\infty \text{ its degree in } t.
\end{aligned}$$

1.2 Carlitz zeta values and power sums

1.2.1 *Carlitz zeta*

For $s \in \mathbb{Z}_+$, put

$$\zeta(s) = \sum_{a \in A_+} \frac{1}{a^s} \in K_\infty.$$

These are the *Carlitz zeta values*. See [8, 11] and references therein for more on this and basic analogies between function field and number field situations.

1.2.2 *Power sums*

It is convenient to break the Carlitz zeta values into power sums grouped by degree: Given integers $s > 0$ and $d \geq 0$, put

$$S_d(s) = \sum_{\substack{a \in A_+ \\ \deg a = d}} \frac{1}{a^s}.$$

(This is $S_d(-s)$ in the notation of [11].)

1.3 Multiple zeta values

1.3.1 *Definition*

For $s_i \in \mathbb{Z}_+$, we define multizeta value $\zeta(s_1, \ldots, s_r)$ following [11, Section 5.10] (where it was denoted by ζ_d to stress the role of the degree) by

$$\zeta(s_1, \ldots, s_r) = \sum_{d_1 > \cdots > d_r \geq 0} S_{d_1}(s_1) \cdots S_{d_r}(s_r) = \sum \frac{1}{a_1^{s_1} \cdots a_r^{s_r}} \in K_\infty,$$

where the second sum is over $a_i \in A_+$ of degree d_i satisfying conditions as in the first sum.

We say that this multizeta value has depth r and weight $\sum s_i$.

1.3.2 *Remarks*

(1) In contrast to $\mathbb{Z}$, on which we have a natural total order, on A, we only have natural partial order according to the degree, or equivalently the size of

the norm. Many aspects such as interpolation [11, 5.5, 5.10], relations with motives [4, 5], work smoothly when we group according to the degree, so we define multizeta as above.

(2) We do not need $s_1 > 1$ condition for convergence as in the classical case.

(3) The depth and the weight are really associated with the tuple $(s_1, \ldots, s_r)$ and not to the corresponding multizeta value. A priori, the depth and the weight are not fully specified by the value. In the classical and our case though, conjecturally, two values with different weight should not be equal, whereas in the classical case, as the Euler identity recalled in 2.6 shows, the two values with different depth can be equal.

2 Relations Between Multizeta Values

Let us review first some simple relations, and comparison with well-known classical relations.

2.1 Relations coming from the p-power map

Since we are in characteristic p, definition immediately implies that the multizeta value at (ps_i) is the pth power of the corresponding multizeta value at (s_i).

2.2 Zeta at "even" s

Carlitz proved (see [11], 5.2.1) an analog of Euler's result that if s is "even" in the sense that $q-1$ divides s, then $\zeta(s)/\tilde{\pi}^s$ is in K. Here, $\tilde{\pi}$ is a fundamental period of the Carlitz module, and it being well defined up to multiplication by an element of $\mathbb{F}_q^*$, for "even" s, $\tilde{\pi}^s$ is well defined.

2.3 Low s relations

Using Carlitz $S_d(kp^n) = 1/\ell_d^{kp^n}$, for $0 < k < q$ [11, 5.9.1], it was noted in [11, 5.10.6] that nondegenerate multizeta values with weight not more than q satisfy the classical sum shuffle identities. More generally, the product over j of the multizeta values $\zeta(s_{ij})$ (where the j-th multizeta value has parameters s_{ij}) is a sum of multizeta values as in the classical sum shuffle identities, if for any i, $\sum_j s_{ij} \leq q$, or even more generally if each s_{ij} is of the form $k_{ij}p^{n_{ij}}$, with $k_{ij} < q$ and all the sums over j's of them are also of that form. For example, $\zeta(q-1)\zeta(p+1)$ satisfies the usual sum shuffle, if $q > p$.

So any classical sum shuffle relation with fixed s_{ij}'s works for q large enough! Of course, for our interests, q is fixed.

Apart from the sum shuffle, this implies other types of identities, for example, if $q+1=a+b=c+d$, then $\zeta(a)\zeta(b)-\zeta(a,b)-\zeta(b,a)$ is the same as $\zeta(c)\zeta(d)-\zeta(c,d)-\zeta(d,a)$, both being the same as $\sum 1/\ell_d^{q+1}$.

Remark. The look and consequences of these relations can be different, depending on q: By 2.3. we have $\zeta(1)\zeta(1)=\zeta(2)+2\zeta(1,1)$ for all $\mathbb{F}_q[t]$'s as q varies, but the right side reduces to $\zeta(2)$ when $p=2$, and the relation does not give any information on $\zeta(1,1)$. Also, by 2.1, when $p=2$, but not in general, the sum shuffle $\zeta(k)\zeta(k)=\zeta(2k)+2\zeta(k,k)$ works, but the right side reduces to $\zeta(2k)$ and the relation does not give any information on $\zeta(k,k)$. □

2.4 Failure of the naive sum shuffle

In [11, Theorem 5.10.12], it is shown that when $q=3$, the naive analog of classical sum shuffle $\zeta(2)^2=2\zeta(2,2)+\zeta(4)$ fails, by showing that $\zeta(2,2)/\tilde{\pi}^4$ is not in K and using the Carlitz result above for "even" s. (In fact, this follows much more simply as we will see below, but irrationality of the ratio is of independent interest.)

2.5 Failure of the naive integral shuffle

Classically, there are integral shuffle identities between multizeta values coming from their iterated integral expressions and thus connecting immediately to mixed motives. For example, in the usual iterated integral notation, we have

$$\int w_1w_0\int w_1w_0=2\int w_1w_0w_1w_0+4\int w_1w_1w_0w_0,$$

which, with $w_0=dz/z$ and $w_1=dz/(1-z)$, gives classically the identity $\zeta(2)^2=2\zeta(2,2)+4\zeta(3,1)$. In our case, the same identity does not work, for example, if $p=2$, because the right side then is zero, but the left side is nonzero and is, in fact, transcendental.

2.6 Failure of naive analog of the Euler identity

In contrast to Euler's $\zeta(3)=\zeta(2,1)$, we cannot have such identities for the simple reason that in the depth one, i.e., for the usual zeta value, its degree in t is zero, while for any higher depth multizeta value, the degree in t is less than zero as we can see just from the definitions.

In [12], we investigate degrees in t of various depth multizeta values. In the classical case, we have a generalization of the Euler identity that says the sum of all convergent multizeta values of given weight and a fixed length is independent of the length. In our case, this fails in general, for the same simple degree comparison reason.

2.7 Different identities with similar consequences

An important consequence of the shuffle identities is that a product of multizeta values can be expressed as a sum of multizeta values (all of the same weight which is the sum of the weights of multizeta values in the product), and thus the linear span of the multizeta values forms an algebra graded by weight. While we do not have the same shuffle identities, it seems that the previous statements are still true in our case.

Remark. While, classically, sums mean linear combinations with nonnegative integral coefficients, here $\mathbb{F}_p$-coefficients suffice, e.g., just zeros and ones when $p=2$.

We give several examples, postponing the proofs to the next section. □

2.7.1

For $q=2$, we showed [11] that the naive sum shuffle $\zeta(1)\zeta(2)=\zeta(3)+\zeta(2,1)+\zeta(1,2)$ does not work, as $\zeta(1,2)+\zeta(2,1)$ is not an algebraic multiple of $\tilde{\pi}^3$. Claim: The identity works if we drop the last term.

2.7.2

Claim. In our example of the failure of the naive sum shuffle above, we can just drop $2\zeta(2,2)$ to get $\zeta(2)^2=\zeta(4)$, when $q=3$. This works, by 2.1, for $p=2$ also, so even though the integral shuffle above fails, we get the correct consequence as above.

2.7.3

In 2.3, we mentioned the first part of the following theorem.

Theorem 1. We have the sum shuffle identity

$$\zeta(s_1)\zeta(s_2)=\zeta(s_1+s_2)+\zeta(s_1,s_2)+\zeta(s_2,s_1),$$

when (1) $s_1 + s_2 \leq q$,

or when (2) $s_1 = aq + b$, $s_2 = p$, if $0 < a < p$ and $0 < b < q - p$,

or when (3) $s_1 = q + p - 1$, $s_2 = q$, if $q > 2$,

or when (4) $s_1 = s_2 = q + p - 1$, if $q > p$,

or when (5) $s_1 = aq + b$, $s_2 = rp^2$, if $b + a + rp^2 < q$, and $p \leq a < 2p$ and $b_0 > 0$ and $b_0 + a_0 \leq p$, where a_0, b_0 denote the last base p digits of a and b, respectively. □

This is not an exhaustive list, but we have proved [12] that when $q = 2$, the identity of the theorem only works, at the degree level, by 2.1 and 2.3. Complete classification is not done yet in general.

The identity fails in general, but we have instead, for example, the theorem.

Theorem 2. (1) When $a, b \leq q$ and $a + b > q$, we have

$$\zeta(a)\zeta(b) = \zeta(a+b) + \zeta(a,b) + \zeta(b,a) + (a+b-q)\zeta(a+b-q+1, q-1).$$

(2) When $1 \leq b \leq q$, $q \neq 2$, we have

$$\zeta(b)\zeta(2q) = \zeta(2q+b) + \zeta(b,2q) + \zeta(2q,b) + b\zeta(q+b+1, q-1) + \binom{b+1}{2}\zeta(b+2, 2q-2).$$ □

2.7.4

Remark. Note in the special cases, such as when a or b is $q-1$ or when p divides $a+b$, three depth 2 multizeta values in part (1) can mix or disappear giving difference appearance to the identities. Also note that the multiple $a+b-q$ of the last multizeta value in (1) can also be written as $a+b$, as it only matters modulo p.

Claim. We have similar evaluations for $\zeta(2q+1)$, $\zeta(2q-2, 2)$ which can also be thought of as substitutes for the sum shuffles.

Claim. For $a, b < q$, $\zeta(a)\zeta(bq)$ is (explicitly given) as sum of multizeta values of the same weight and of depth at most 2.

2.7.5

Here is a situation of arbitrary weight, when $q = 2$.

Theorem 3. When we have $q = 2$,

$$\zeta(1)\zeta(a) = \zeta(1+a) + \sum_{i=1}^{a-1} \zeta(i, a+1-i). \qquad \square$$

2.7.6

If so far, we have restricted to a small q or a small weight compared to q (discounting 2.1 phenomena), here is an example without this restriction.

Theorem 4.

$$\zeta(q^n - 1)\zeta((q-1)q^n) = \zeta(q^{n+1} - 1) + \zeta(q^n - 1, (q-1)q^n). \qquad \square$$

2.7.7

In place of the Euler identity of 2.6, we have the following theorem.

Theorem 5. When $q = 3$, we have $\zeta(1, 2) = \zeta(3)/\ell_1 = \zeta(3)/(t - t^3)$. More generally, for any q, we have

$$\zeta(m, m(q-1)) = \zeta(mq)/\ell_1^m, \quad m \leq q,$$
$$\zeta(1, q^2 - 1) = \zeta(q^2)(1/\ell_2 + 1/\ell_1). \qquad \square$$

2.7.8

Remark. When $q = 3$ (when "even" agrees with even), in comparison with the Euler identity we have an order switch. But, using the sum shuffle identity 2.3 for $\zeta(1)\zeta(2)$, we can express $\zeta(2, 1)$ in terms of $\zeta(3) = \zeta(1)^3$ and $\tilde{\pi}$.

Note the following strange evaluations (with direct transcendence applications from results of [10] on the nature of special values of logarithm).

Theorem 6. (1) We have

$$\zeta(1, q^3 - 1) = \zeta(q^3)\left(\frac{1}{\ell_3} + \frac{1}{\ell_2} + \frac{t}{\ell_2}\right) - \frac{1}{\ell_2}\left(\log\left(t^{1/q}\right)\right)^{q^3}.$$

(2) $\zeta(1, q^n - 1)$ is (explicit) $\zeta(q^n)$ times a rational plus linear combination of q-power powers of logarithms of q-power roots of polynomials.

(3) When $q = 2$,

$$\zeta(3,5) = \frac{1}{[2][1]^3}\Big(t\zeta(8) + \big(\log\big(t^{1/2}\big)\big)^8\Big). \qquad \square$$

Here is a series of identities of the type of Theorem 3.

Theorem 7. When $q = 2$, we have (1) if b is odd,

$$\zeta(2)\zeta(b) = \zeta(2+b) + \sum_{1 \le i \le (b-3)/2} \zeta(2i+1, 1+b-2i),$$

and (2) if b is even, then

$$\zeta(2)\zeta(b) = \zeta(2+b) + \sum_{1 \le i \le b/2-1} \zeta(2i, b+2-2i). \qquad \square$$

Finally, here are some examples in higher depth.

Theorem 8. When $q = 2$, we have

(1) $\zeta(1)\zeta(1,2) = \zeta(1,3) + \zeta(2,2)$,

(2) $\zeta(1)\zeta(2,1) = \zeta(3,1) + \zeta(1,2,1)$, and

(3) $\zeta(2)\zeta(1,1) = \zeta(3,1) + \zeta(1,3) + \zeta(2,2) + \zeta(2,1,1) + \zeta(1,1,2)$. $\square$

3 Proofs

3.1 Notation

Recall the notation $[n]$, ℓ_n, D_n from the first section, and note

$$\ell_0 = D_0 = 1, \quad (-1)^n L_n = \ell_n = -[n]\ell_{n-1}, \quad D_n = [n]D_{n-1}^q.$$

Put

$$\binom{x}{q^d} = \sum_{i=0}^{d} \frac{x^{q^i}}{D_i \ell_{d-i}^{q^i}},$$

$$S_{<d}(k) = \sum_{a \in A_+, \deg(a) < d} \frac{1}{a^k}.$$

3.2 Tools

Then we have the following results of Carlitz (see e.g., [11, 2.5, 5.6] or [7, 8]):

$$D_i = \prod_{a\in A_+, \deg(a)=i} a, \qquad L_i = \text{lcm of } a \in A_+ \text{ of degree } i,$$
$$\binom{x}{q^d} = \prod_{a\in A, \deg(a)<d} (x+a)/D_d.$$

Hence, $\binom{t^d}{q^d} = 1$ and thus

$$\binom{x+t^d}{q^d} = 1 + \binom{x}{q^d} = \prod_{a\in A_+, \deg(a)=d} (1+x/a).$$

Taking the logarithmic derivative, expanding by the geometric series, and using the fact that $\sum \theta^k$, where θ ranges over the elements of $\mathbb{F}_q^*$, is -1 or 0 according as to whether k is "even" or not, we see that

$$\frac{x}{\ell_d\left(1-\binom{x}{q^d}\right)} = \sum_{k=1}^{\infty} S_d(k)x^k,$$
$$\frac{x}{\ell_d\binom{x}{q^d}} = 1 + \sum_{k>0 \text{ ``even''}} S_{<d}(k)x^k.$$

3.3 Basic formulas

In other words, $S_d(s+1)$ is the coefficient of x^s in $(1+\binom{x}{q^d}+\binom{x}{q^d}^2+\cdots+\binom{x}{q^d}^s)/\ell_d$, or equivalently

$$S_d(k+1) = \frac{1}{\ell_d^{k+1}} \sum_{\sum_{i=0}^d k_i q^i = k, k_i\geq 0} \binom{\sum k_i}{k_i} \prod \frac{(\ell_d/\ell_{d-i})^{q^i k_i}}{D_i^{k_i}},$$
$$S_d(k+1) = \frac{1}{\ell_d^{k+1}} \sum_{k=k_d q^d+\cdots+k_0, k_i\geq 0} \binom{k_d+\cdots+k_0}{k_d,\ldots,k_0} \prod_{i=1}^{d} \left(\frac{(-1)^i([d]\cdots[d-i+1])^{q^i}}{D_i}\right)^{k_i}.$$

3.3.1 *Special cases*

Here are some special cases:

$$S_d(ap^n) = 1/\ell_d^{ap^n}, \quad \text{if } a \le q,$$
$$S_d(q+b) = \frac{1}{\ell_d^{q+b}}\left(1 - b\frac{[d]^q}{[1]}\right), \text{ if } 1 \le b < q,$$
$$S_d(aq+b) = \frac{1}{\ell_d^{aq+b}}\left(1 + \sum_{j=1}^{a}(-1)^j\binom{b+j-1}{j}\frac{[d]^{jq}}{[1]^j}\right), \quad \text{if } a, b < q,$$
$$S_d(q^2+1) = \frac{1}{\ell_d^{q^2+1}}\left(1 + \sum_{j=1}^{q}(-1)^j\frac{[d]^{jq}}{[1]^j} + \frac{[d]^{q^2}[d-1]^{q^2}}{[2][1]^q}\right).$$

3.3.2

Note that in the second generating function case, the powers k's that appear with nonzero coefficient are all "even," i.e., multiples of $q-1$, as they should be, because "even" powers kill all the signs that we consider here. Here are some special cases from the second generating function:

$$S_{<d}(m(q-1)) = \frac{[d]^m/[1]^m}{\ell_{d-1}^{m(q-1)}}, \quad m \le q,$$
$$S_{<d}(q^i - 1) = \frac{\ell_{d+i-1}}{\ell_i \ell_{d-1}^{q^i}}.$$

Remark. These can also be easily derived by induction (see, e.g., [11, 5.10.13]) from the corresponding S_d identities (or vice versa by just a subtraction) without using the generating function. In fact, we first proved and used it this way in special cases, before realizing that the generalization is available.

When $q = 2$, we have

$$S_{<d}(5) = \frac{[d+1]}{[2]}\frac{[d]^3}{[1]^3}\frac{1}{\ell_{d-1}^5}.$$ □

3.3.3

The following can be derived from the first generating function above after some manipulation, but follows from the $S_{<d}$ formula above just by subtraction:

$$S_d(q^i - 1) = \frac{\ell_{d+i-1}}{\ell_{i-1}\ell_d^{q^i}} = \frac{[d+i-1]\cdots[d+1]}{[i-1]\cdots[1]\ell_d^{q^i-1}}.$$

3.4 Proofs of the theorems

We will be using the tools and the special cases developed above without specific reference, the formula to be used being always the obvious one in the context from the list above.

Put, with $d_i = \deg(a_i)$ as usual,

$$S_d(s_1, s_2) = \sum_{d=d_1>d_2} \frac{1}{a_1^{s_1} a_2^{s_2}}.$$

We can similarly define higher-depth versions.

3.4.1 *Proof of Claim 2.7.2*

When $q = 3$, $\zeta(2)^2 = \zeta(4)$ follows easily from verified $S_d(2)S_d(2) = S_d(4) + S_d(2, 2)$. Another proof would be a straight calculation using Bernoulli polynomial calculation using 2.2.

3.4.2 *Proof of Theorem 1*

All the claims follow directly by straight calculations (keeping in mind that there is vanishing modulo p often) using the special cases above to check $S_d(s_1)S_d(s_2) = S_d(s_1 + s_2)$ under each of the hypotheses.

3.4.3 *Proof of Theorem 2*

For $1 \le m \le q$, and with $a + b = q + m$, $a, b \le q$, we have

$$\begin{aligned} S_d(q+m) &= \frac{1}{\ell_d^{q+m}} - \frac{m[d]^q/[1]}{\ell_d^{q+m}} = S_d(a)S_d(b) - \frac{m[d]/[1]}{\ell_{d-1}^{q-1}\ell_d^{m+1}} \\ &= S_d(a)S_d(b) - mS_{<d}(q-1)S_d(m+1), \end{aligned}$$

(where if $m = q$, the formula for $S_d(m + 1)$ does not match, but the whole identity is saved by m in front, which is zero). Summing over $d \ge 0$, we get the first part of the theorem. If $a = b = q$, it reduces to the sum shuffle identity 2.3.

To get the second part, we sum

$$S_d(2q+b) = \frac{1}{\ell_d^{2q+b}} - b\frac{[d]^q/[1]}{\ell_d^{2q+b}} + \binom{b+1}{2}\frac{[d]^{2q}/[1]^2}{\ell_d^{2q+b}}$$

and verify by straightforward manipulations that the first term is $S_d(2q)S_d(b)$, the second term is $-bS_d(q+b-1, q-1) - b(b+1)S_d(b+2, 2q-2)$, and the third term is $\binom{b+1}{2}S_d(b+2, 2q-2)$, and noting that when $b+2 > q$, we are still saved by vanishing of appropriate terms.

3.4.4 *Proof of Theorem 3*

It is enough to prove

$$S_d(1)S_d(a) = S_d(a+1) + \sum_{i=2}^{a} S_d(i, a+1-i).$$

Proof. Coefficient of x^{a+1} in

$$\frac{x}{\ell_d\binom{x}{q^d}} \frac{x}{\ell_d\left(1-\binom{x}{q^d}\right)} = \frac{x^2}{\ell_d^2\left(\binom{x}{q^d}^2 - \binom{x}{q^d}\right)}$$

is $\sum_{i=1}^{a} S_d(i, a+1-i) + S_d(a+1)$. Now

$$\binom{x}{q^d}^q - \binom{x}{q^d} = [d+1]\binom{x}{q^{d+1}},$$

so the coefficient is also $1/\ell_d$ times $S_{<d+1}(a) = S_{<d}(a) + S_d(a)$.

Remark. Let θ be a generator of $\mathbb{F}_q^*$. A simple substitution of x/θ^i for x in the last but one formula in 3.2 shows that

$$\frac{x}{\ell_d\left(\theta^i - \binom{x}{q^d}\right)} = \sum_{k=1}^{\infty} \theta^{-ik} S_d(k) x^k.$$

Hence, the above proof generalizes for any q to prove that

$$\sum \theta^{k_1+2k_2+\cdots+(q-1)k_{q-1}} S_d(k_1)\cdots S_d(k_{q-1}) S_{<d}\left(a+q-1-\sum k_i\right)$$

is the same as $-S_d(q-1)S_d(a) - S_d(q-1, a)$ or 0 according as whether a is "even" or not, where we have used temporary convention that $S_{<d}(0) = 1$ and we have $k_i \geq 1$, $\sum k_i \leq a+q-1$. This will be used in a sequel to this paper. □

3.4.5 *Proof of Theorem 4*

It is enough to prove

$$S_d(q^n-1)S_d((q-1)q^n) = S_d\left(q^{n+1}-1\right) - S_d((q-1)q^n, q^n-1).$$

Proof follows by the formulas we have for $S_d(q^n-1)$, $S_{<d}(q^n-1)$ in 3.3.2 and 3.3.3 and for $S_d(q^n)$ and $S_d((q-1)q^n)$ in 2.1, 2.3. In fact, the equality reduces, after removing the common factors by using the recursion formula in 3.1 for ℓ_d, to $1 = [d+n]/[n] - [d]^{q^n}/[n]$ which is clear.

3.4.6 *Proof of Theorem 5*

We have $\zeta(m, m(q-1)) = \sum_{d=1}^{\infty} S_d(m, m(q-1))$. Now

$$S_d(m, m(q-1)) = S_d(m)S_{<d}(m(q-1)) = \frac{[d]^m/[1]^m}{\ell_d^m \ell_{d-1}^{m(q-1)}} = \frac{(-1)^m}{[1]^m \ell_{d-1}^{mq}} = \frac{1}{\ell_1^m} S_{d-1}(mq).$$

Hence, the first claim follows by summing over d.

We give two proofs for the second claim. The first follows by summing the following identity, which can be verified using formulas above for the relevant sums, over $d \geq 0$:

$$S_d(q^2)\left(\frac{1}{\ell_1} + \frac{1}{\ell_2}\right) = S_d(q^2-1)\frac{1}{\ell_{d+1}} + S_{<d+1}(q^2-1)\frac{1}{\ell_{d+2}}.$$

For the second proof, we sum over $d \geq 0$, the following identity:

$$S_d(1, q^2-1) = \frac{1}{\ell_d}\frac{\ell_{d+1}}{\ell_2 \ell_{d-1}^{q^2-1}} = \frac{t - t^{q^{d+1}}}{\ell_2 \ell_{d-1}^{q^2}}$$

to get

$$\zeta(1, q^2-1) = \frac{1}{\ell_2}((\log(t)^{q^2} - t(\log(1))^{q^2}) = \zeta(q^2)\left(\frac{1}{\ell_1} + \frac{1}{\ell_2}\right),$$

where the log is Carlitz logarithm $\log(z) = \sum z^{q^d}/\ell_d$, and for the last equality we use $\zeta(1) = \log(1)$, $\zeta(1)^{q^2} = \zeta(q^2)$, which are consequences of 2.1 and 2.3, and $\log(t) = t\log(1) - \log(1)$, which can be directly verified by the definition of logarithm above and recursion for ℓ_n above.

3.4.7 *Remark*

Unlike the proofs of the shuffle-type identities, note that the parts are borrowed from the dth and $d+1$-th levels on the right side to match the dth level on the left.

3.4.8 *Proof of Theorem 6*

We generalize the second proof in 3.4.6. At the moment we do not know whether we are missing a clearer zeta expression and the first type of proof.

The first identity in the last proof generalizes

$$S_d(q^i)\left(\frac{1}{\ell_{i-1}}+\frac{1}{\ell_{i-2}}\right)=S_d(q^i-1)\frac{1}{\ell_{d+i-1}}+S_{<d+1}(q^2-1)\frac{1}{\ell_{d+i}}.$$

But, to get $\zeta(1,q^3-1)$, we need to add, to the sum of the right side over d, the sum

$$\sum_{d=0}^{\infty}\frac{S_d(q^i-1)}{\ell_{d+1}}=\frac{1}{\ell_2}\sum\frac{t-t^{q^{d+2}}}{\ell_d^{q^3}}=\frac{1}{\ell_2}\left(t\zeta(q^3)-\log\left(t^{1/q}\right)^{q^3}\right)$$

as claimed in part (1). Parts (2) and (3) follow by the same method and we omit the proofs.

3.4.9 *Proof of Theorem 7*

To prove part (1), it is enough to prove that

$$S_d(2)S_d(b)=S_d(b+2)+\sum_{1}^{(b-1)/2}S_d(2i+1,2+b-(2i+1))+S_d(2,b).$$

Since $q=2$, by taking a derivative or by subtracting the square of the generating function, we can isolate the odd powers, so that

$$\sum_{i=0}^{\infty}S_d(2i+1)x^{2i+1}=\frac{x}{\ell_d\left(1+\binom{x}{q^d}\right)}+\frac{x^2}{\ell_d^2\left(1+\binom{x}{q^d}\right)^2}.$$

Hence, the coefficient of x^{2+b} in

$$\frac{x}{\ell_d\binom{x}{q^d}}\left(\frac{x}{\ell_d\left(1+\binom{x}{q^d}\right)}+\frac{x^2}{\ell_d^2\left(1+\binom{x}{q^d}\right)^2}\right)$$

is $S_d(2+b)+\sum S_d(2i+1,2+b-(2i+1))$, where i runs through 0 to $(b-1)/2$.

Now we use the same identity as in the proof of Theorem 3, so that after you cross-multiply, the first term is $x^2(\ell_d \ell_{d+1}\binom{x}{q^{d+1}})^{-1}$, and hence we get $1/\ell_d = S_d(1)$ times the coefficient of x^{1+b} in $\sum S_{<d+1}(k)x^k$. Hence, the contribution from the first term is $S_d(1, 1+b) + S_d(1)S_d(1+b)$.

Using the same identity, we rewrite the second term (after cross-multiplying) as $x^2(\ell_d \ell_{d+1}\binom{x}{q^{d+1}})^{-1} x(\ell_d(1+\binom{x}{q^d}))^{-1}$, so that the contribution is

$$S_d(1)[S_d(1+b) + \sum_1^b S_d(k, 1+b-k) + S_d(k)S_d(1+b-k)],$$

which is $S_d(1)[S_d(1+b) + S_d(1,b) + S_d(1+b) + S_d(1)S_d(b) + S_d(1+b)]$ (the last part of the sum cancels except for the middle term, since the terms occur twice and we are in characteristic two. This is the same as $S_d(1)S_d(1+b) + S_d(2,b) + S_d(2)S_d(b)$.

So the total contribution is $S_d(2)S_d(b) + S_d(1, 1+b) + S_d(2, b)$ proving the claim by matching the two expressions.

Part (2) follows by Theorem 3 combined with 2.1.

3.4.10 *Proof of Theorem 8*

Part (1). By shuffle and 2.3, the left side is

$$2\zeta(1,1,2) + \zeta(1,2,1) + \zeta(2,2) + \sum_{d>d_1} S_d(1)S_{d_1}(1)S_{d_1}(2),$$

so it is enough to show that

$$S_d(1)\sum_{d>d_1} S_{d_1}(1)S_{d_1}(2) = S_d(1,3) + S_d(1,2,1).$$

The left side is $(1/\ell_d)\sum 1/\ell_{d_1}^3$ and since

$$S_{d_1}(2,1) = \left(1/\ell_{d_1}^2\right)([d_1]/[1])/\ell_{d_1-1} = [d_1]^2/\left([1]\ell_{d_1}^3\right),$$

whereas $S_{d_1}(3) = (1 + [d_1]^2/[1])/\ell_{d_1}^3$, the claim follows.

Part (2). The same method reduces the proof to the identity

$$S_d(1)S_d(2)\sum_{d>d_1} S_{d_1}(1) = S_d(2,2) + S_d(3,1),$$

whose left side equals $(1/\ell_d^3)([d]/[1])/\ell_{d-1}$ and the right side equals

$$\left(1/\ell_d^2\right)([d]^2/[1]^2)/\ell_{d-1}^2 + (1+[d]^2/[1])/\ell_d^3 * ([d]/[1])/\ell_{d-1},$$

and thus both are easily seen to be equal.

Part (3). This reduces by the shuffle to proving that

$$S_d(2)S_d(1)\sum S_{d_1}(1) + S_d(1)\sum S_{d_1}(2)S_{d_1}(1)$$

is the same as $S_d(2,2)+S_d(3,1)$ plus $S_d(1,2,1)+S_d(1,3)$, which is exactly what we saw in the two calculations above.

3.4.11 *Proof of Claims in 2.7.4*

Claims in 2.7.4 follow as in the proof of Theorem 2, but with much more involved calculation which we omit.

4 Pattern Recognition Attempts by Numerical Calculations

In this section, we give guesses about identities, with some proofs, some numerical verification, and some pattern recognition (or failure of it!). We will explain in the last section, how each particular identity between specific multizeta values (rather than an identity scheme containing parameters as in our theorems, e.g., because of the parameter a in Theorem 3, the identity scheme that Theorem 3 proves contains infinitely many identities, one for each value of a), if true, should have a mechanical proof. We have indeed verified, a few, but not all of the following predictions.

Remark. These calculations were mostly first guessed and checked by hand calculations and then only verified more by maxima by substituting exact rational functions, without using automated program. We plan to do more extensive verification (using automated programs, power series calculations, etc.) and proofs soon and report in a sequel to this paper. □

The material in 4.1.1–4.1.3 is implied by 4.1.4, but it is still presented here, as it provides concrete examples to the reader without need of much notation, and it formed the basis on which 4.1.4 was formulated.

4.1 When $q=2$

4.1.1 *One index small*

Here is a prediction of how to write $\zeta(a)\zeta(b)$, when $q=2$, as a sum of multizeta values, when a is of form $a_0 2^n$ with $a_0 \leq 32$. We give it at the "motivic" level of S_d's rather than of zetas (for which the corresponding identities are easy to deduce by summing over d and using shuffle). We predict

$$S_d(a)S_d(b) = S_d(a+b) + \sum S_d(a_i, a+b-a_i),$$

where the list of the distinct a_i's that appear, given a and b, is described below.

Let us start with some observations.

In Theorems 3 and 7, we have seen that, if $a=1$, then a_i's are exactly all the integers between 2 and b, whereas if $a=2$, then if b is even, a_i's are all the even integers between b and 4 (as also follows from the $a=1$ case, together with 2.1), and if b is odd, a_i's are all the odd integers between b and 3 together with $a_i=2$. (In particular, the a_i's are between 2 and b. This should be a feature in general.)

In other words, given a_i's for $(1,b)$ (i.e., corresponding to $S_d(1)S_d(b)$), you obtain those for $(1,b+1)$ by just adding (i.e., appending) to the list of a_i's the top possible entry, i.e., $b+1$; and given a_i's for $(2,b)$, you obtain those for $(2,b+2)$ by adding the top possible entry (i.e., $b+2$). So the recursion length for pattern for $a=1,2$ is $1,2$, respectively. For $3,4$ it is 4.

In general, here is the recursive recipe, in four parts, to get a_i's:

(i) If m is the smallest integer such that $a \leq 2^m$, then the recursion length is 2^m.

(ii) For $a_0 = 1,3,5,7,9,11,13,15,17,19,21,23,25,27,29,31$, you add $1,2,4,2,8,4,4,2,16,8,8,4,8,4,4,2$ terms, respectively, to the list of a_i's at a recursive step.

(iii) If $n=0$, they are the top possible terms (i.e., b downward) (spread patterns by n otherwise) unless $a_0 = 11,19,21,23,27$, when you add, respectively,

top 2, gap of 2, next 2,

top 2, gap 2, next 2, gap 2, next 2, gap 2, next 2,

top 4, gap 4, next 4,

top 2, gap 6, next 2, and

top 2, gap 2, next 2.

(iv) Finally, a short table (we give here a shorter version for $a \leq 9$ rather than for the full range $a \leq 32$) takes care of the initial conditions from which this recursive recipe gives the full prediction. For a given $a \leq 9$, we will write down $(b|a_i's)$ for b up to the recursion length (though we could have further restricted, without loss of generality, to $b > a$, using symmetry in a and b, and with at least one of them odd, using 2.1). Let $c..d$ temporarily denote all the integers between c and d inclusive, and let $-$ denote the empty string of a_i's.

$a = 1$: $(1|-)$.
$a = 2$: $(1|2)$, $(2|-)$.
$a = 3$: $(1|3, 2)$, $(2|3, 2)$, $(3|-)$, $(4|4)$.
$a = 4$: $(1|4..2)$, $(2|4)$, $(3|4)$, $(4|-)$.
$a = 5$: $(1|5..2)$, $(2|5, 3, 2)$, $(3|5..2)$, $(4|5..2)$, $(5|-)$, $(6|6, 4)$, $(7|7, 6)$, $(8|8, 7, 6)$.
$a = 6$: $(1|6..2)$, $(2|6, 4)$, $(3|6, 5, 3, 2)$, $(4|6, 4)$, $(5|6, 4)$, $(6|-)$, $(7|7, 6)$, $(8|8)$.
$a = 7$: $(1|7..2)$, $(2|7, 5, 3, 2)$, $(3|7, 6)$, $(4|7, 4)$, $(5|7, 6)$, $(6|7, 6)$, $(7|-)$, $(8|8)$.
$a = 8$: $(1|8..2)$, $(2|8, 6, 4)$, $(3|8, 7, 4)$, $(4|8)$, $(5|8, 7, 6)$, $(6|8)$, $(7|8)$, $(8|-)$.
$a = 9$: $(1|9..2)$, $(2|9, 7, 5, 3, 2)$, $(3|9, 8, 5..2)$, $(4|9, 5..2)$, $(5|9..2)$, $(6|9, 7..2)$, $(7|9..2)$, $(8|9..2)$, $(9|-)$, $(10|10, 8, 6, 4)$, $(11|11, 10, 7, 6)$, $(12|12..10, 8..6)$, $(13|13..10)$, $(14|14..10, 8)$, $(15|15..10)$, $(16|16..10)$.

At the suggestion of the referee, we add a detailed example to make this complicated combinatorial recipe clearer.

Example. Let $a = 19$, so that by (i), the recursion length is 32 and by (ii) we add 8 new terms (which are described in (iii)) at each recursive step. This implies that, with the notation as above (so that a_i's correspond to $S_d(19)S_d(b)$),

$$\begin{aligned} S_d(19)S_d(b+32) = {} & S_d(19+b+32) + \sum S_d(a_i, b+51-a_i) \\ & + S_d(b+32, 19) + S_d(b+31, 20) + S_d(b+28, 23) + S_d(b+27, 24) \\ & + S_d(b+24, 27) + S_d(b+23, 28) + S_d(b+20, 31) + S_d(b+19, 32). \end{aligned}$$

To start this recursion, we need a_i's corresponding to $b \leq 32$. By reversing the role of a and b, the short table in (iv) gives them for $b \leq 9$. Also, e.g., since $S_d(19)S_d(19) = S_d(38)$ there are no a_i's in this case, so that a_i's for $S_d(19)S_d(51)$ are (51, 50, 47, 46, 43, 42, 39, 38) by the recipe. In 4.1.4, we give the full recipe in general, and it would be a good exercise for the interested reader to use it to fill in the full table for $a = 19$ and $b \leq 32$ completing the detailed recipe in the special case $a = 19$. □

4.1.2 *Special large indices*

You add 2 top terms for $a = 2^n - 1$, and 2^n top terms for $a = 2^n + 1$.

4.1.3 *Both indices large*

(The first prediction below follows from Theorem 4)

$$S_d(2^n - 1)S_d(2^n) = S_d\left(2^{n+1} - 1\right) + S_d(2^n, 2^n - 1),$$

$$S_d(2^n)S_d(2^n + 1) = S_d\left(2^{n+1} + 1\right) + \sum_{i=2}^{2^n+1} S_d\left(i, 2^{n+1} + 1 - i\right),$$

$$S_d(2^n - 1)S_d(2^n + 1) = S_d\left(2^{n+1}\right) + \sum_{i=2}^{2^n+1} S_d\left(i, 2^{n+1} - i\right),$$

$$S_d\left(2^{n-1}\right)S_d(2^n + 1) = S_d\left(3 * 2^{n-1} + 1\right) + S_d\left(2^n + 1, 2^{n-1}\right) + \sum_{i=2}^{2^{n-1}+1} S_d\left(i, 3 * 2^{n-1} + 1 - i\right).$$

When $k < n - 1$, we have

$$S_d(2^n - 2^k - 1)S_d(2^n - 2^k) = S_d\left(2^{n+1} - 2^{k+1} - 1\right) + S_d(2^n - 2^k, 2^n - 2^k - 1) + S_d\left(2^n - 2^{k+1}, 2^n - 1\right).$$

4.1.4 *General predictions when $q = 2$*

Now, we give a full conjectural recipe ((1–3) below) to write down a product of two zeta values as a sum of multizeta values, when $q = 2$, guessed (and verified to some extent) from the calculations mentioned above.

We start with some definitions to be used only in this subsection:

(i) Let r_a be the smallest power of 2 not less than a, i.e., $r_a = 2^{\lceil \log_2(a) \rceil}$.

(ii) For sets A, B we put $A \oplus B = A \cup B - (A \cap B)$ (this operation corresponds below to the addition at multizeta level, as we are in characteristic two).

Here is the conjectural recipe/hypothesis:

(1) $S_d(a)S_d(b) = S_d(a + b) + \sum S_d(a_i, a + b - a_i)$, where the set, denoted by $S(a, b)$, of a_i's is independent of d, so that in fact we have

$$\zeta(a)\zeta(b) = \zeta(a + b) + \zeta(a, b) + \zeta(b, a) + \sum \zeta(a_i, a + b - a_i).$$

We will denote the size of $S(a, b)$ by $s(a, b)$.

Remark. Then $S(a,a)$ is empty and $S(a,b) = S(b,a)$, so that we can assume without loss of generality, when needed, that $a < b$. □

Though $S(a,b) = S(b,a)$, we will give a nonsymmetric description (we would love to have a symmetric and/or nonrecursive direct description!) by fixing a and describing in (2) recursion of length r_a in b, where at each recursion step one adds new entries to the set in the manner described in (2). The initial conditions of the recursion follow from the "half-recursion" formulated in (3).

(2) $S(a,b) = S(a, b - r_a) \oplus T(a,b)$, where we now describe recipe for $T(a,b)$.

Example. Theorem 3 shows that $S(1,b) = \{2, \ldots, b\}$ and $s(1,b) = b - 1$, $T(1,b) = \{b\}$.

First consider odd a. Write $a = \sum_1^n 2^{e_i} - \sum_1^n 2^{\overline{e_i}} + 2^{e_0}$, with $e_0 = 0$, $e_n > \overline{e_n} > e_{n-1} > \cdots \overline{e_1} > e_0 = 0$. (In other words, the base 2 expansion of $a-1$ is decomposed in consecutive bunches of 1's and 0's: more precisely, it consists of consecutive 1's between the places $e_i - 1$ to $\overline{e_i}$, both inclusive, and 0's elsewhere.) We give recipe inductively on n ($n = 0$ being taken care of by the example above): If $a' = a + 2^{e_{n+1}} - 2^{\overline{e_{n+1}}}$, then

$$T(a',b) = T(a,b) \oplus T(a, b - 2^{e_n}) \oplus \cdots \oplus T\big(a, b - \big(2^{\overline{e_{n+1}} - e_n} - 1\big)2^{e_n}\big).$$

We have

$$T(2^m a, b) = \{b, b - 2^m b_1, \ldots, b - 2^m b_k\}, \text{ if } T(a,b) = \{b, b - b_1, \ldots, b - b_k\}.$$ □

Example. We leave it to the reader to verify that $T(a,b) = \{b, b-1\}$ as claimed in 4.1.1, and observe that our recipe then implies $T(19,b) = \{b, b-1, b-4, b-5, b-8, b-9, b-12, b-13\}$, as claimed in 4.1.1, since $a = 3 = 2^2 - 2^1 + 2^0$, $a' = 19 = 3 + 2^5 - 2^4$, so that $e_n = 2$ and $\overline{e_{n+1}} = 4$. □

Remark. (a) It follows from the recipe that the size $t(a,b)$ of the set $T(a,b)$ is independent of b and equals t_a defined as 2 raised to the number of zeros in the base 2 expansion of $a - 1$: For odd a, this follows from the interpretation of e_i and $\overline{e_i}$ above, since in (2), if $t_a = 2^{\mu_a}$, then we see that $t_{a'} = 2^{\mu_a + \overline{e_{n+1}} - e_n}$. In general, we see this from the last displayed formula and the odd case, by a straight calculation.

(b) Note the trivial evaluations $r_a = 2^{e_n}$ and $r_{a'} = 2^{e_{n+1}}$.

(c) In (2), we could have replaced $\oplus$ by $\cup$ as the unions can be seen to be disjoint.

(d) We see by a straight induction on n that $S(a,b) \subset \{2, 3, \ldots, b\}$, and that if we write $T(a,b) = \{b, b - b_1, \ldots, b - b_k\}$, then $b_i \leq r_a/2$ and that b_i are independent of b, so that original definition of $T(a,b)$, which makes sense only for $b > r_a$, can be extended by this equality to $b > r_a/2$ still leading to the set of positive integers. □

(3) To describe the initial values $S(a,b)$, for $b \le r_a$ of the recursion we have the following inductive recipe:

$$S(a,b) \oplus S(a, b + r_a/2) = T(a, b + r_a/2) \oplus S(b, \overline{a}),$$

where $a \equiv \overline{a} \mod r_a/2$, $0 < \overline{a} \le r_a/2$, and $0 < b \le r_a/2$. □

Example. If $a = 2^m + 1$, then $S(a,b) \oplus S(a, b + r_a/2) = \{b + r_a/2, b + r_a/2 - 1, \ldots, 2\}$. So the two sets on the left are sort of complementary. □

This finishes the description. But, let us see in more detail how it takes care of the initial values for the recursion: We do induction on a and apply this to $b \le r_a/2$. Since $b < a$ and $\overline{a} \le r_b$, we know (using remark (d) above) the right side of the displayed formula. We also know the set $S(a,b) = S(b,a)$, by using (2) and induction, since $b < a$. So the formula gives $S(a, b + r_a/2)$, which finishes the job.

Remark. (a) Here, we cannot replace $\oplus$ by $\cup$, as one can verify by data in 4.1.1, or from the example below.

(b) From the predictions above, we see that $s(a,b)$ can be approximated by $t_a b/r_a$, for $b >> a$, with bounded error as b approaches infinity. It seems that $s(2^n, b) + 1$ is 2^n divided by 2 raised to the number of 1's in the base 2 expansions of $b - 1$. We do not know a simple symmetric description of $s(a,b)$ in general. □

Example. We continue with the example $a = 19$ from 4.1.1. In this case, the only new part of the recipe is the inductive description of the initial conditions in (3). So let us evaluate $S(19, 20)$, for example. With $r_{19} = 32$, putting $b = 4$ in the recipe (3), we get $S(19,4) \oplus S(19,20) = T(19,20) \oplus S(4,3)$. By induction, or in our case by more explicit 4.1.1, the only unknown entry here is $S(19,20)$, as we know by 4.1.1 table that $S(4,3) = S(3,4) = \{4\}$, and so $S(19,4) = S(4,19) = \{4, 7, 11, 15, 19\}$ by the recursion. We also know that $T(19,b) = \{b, b-1, b-4, b-5, b-8, b-9, b-12, b-13\}$ so that $T(19,20) = \{20, 19, 16, 15, 12, 11, 8, 7\}$. Hence, we conclude that $S(19,20) = \{20, 16, 12, 8\}$. □

Remark. We can recover details of the recursion part of the recipe from another prediction from which, in fact, they were guessed,

$$S(b-1, b) = S(b - 2^j, b).$$

Here is the start: Let $j < n$. Then we have $\{2, \ldots, 2^n + 1\} = S(1, 2^n + 1) = S(2^n + 1 - 2^j, 2^n + 1)$, but the right side then equals $\{2, \ldots, 2^n - 2^j + 1\} \cup T(2^n + 1 - 2^j, 2^n + 1)$, since for $a = 2^n + 1 - 2^j$, we will have $r_a = 2^n$ and $S(2^n + 1 - 2^j, 1) = S(1, 2^n + 1 - 2^j) = \{2, \ldots, 2^n - 2^j + 1\}$. Hence, $T(2^n + 1 - 2^j, b)$ should consist of the top 2^j terms and $t_{2^n+1-2^j} = 2^j$. Varying n between 2 to 5, it takes care of all a's the recipe except for exactly the exceptional list 11, 19, 21, 23, 27. We explained 19 case above. Let us check one more. Consider e.g., $a = 27$. Now we know $T(3, b) = \{b, b-1\}$ and $r_3 = 4$. So $s(3, 35) = s(27, 35) = s(27, 3) + t_{27}$, which gives $t_{27} = 4$ and $T(27, b) = \{b, b-1, b-4, b-5\}$. It is a fun exercise to derive full 4.1.1 from 4.1.4. □

4.2 When $q = 3$

We have

$$S_d(1)S_d(k) = S_d(k+1) + S_d(k-1, 2) + S_d(k-3, 4) + \cdots$$

ending at $S_d(2, k-1)$ or $S_d(3, k-2)$ depending on whether k is odd or even, respectively.

$S_d(2)S_d(2k+1)$ has a_i description starting with $2k+1$ and going by gaps of $4, 2, 4, 2, \ldots$ and ending at least of 3 and 5, with signs alternating each time starting with negative.

Example.

$$S_d(2)S_d(17) - S_d(19) = -S_d(17, 2) + S_d(13, 6) - S_d(11, 8) + S_d(7, 12) - S_d(5, 14).$$

$S_d(2)S_d(2k)$ is also with alternating signs, starting with $-2k$ and with gaps of $4, 2, 4, 2\ldots$, except if you have $a_i = 6$ after gap of 2, end there, and if you reach ± 4, end with 2 with same sign ignoring the gap pattern rule.

In analogy with $q = 2$ case above, we can describe these by recursion:

For $a = 1$, the recursion length is 2, and to get $S_d(1)S_d(b)$ you add $a_i = b - 1$ to the corresponding list for $S_d(1)S_d(b-2)$.

For $a = 2$, the recursion length is 6, and to get $S_d(2)S_d(b)$ you add $-S_d(b, 2) + S_d(b - 4, 6)$ to the corresponding list of a_i's from $S_d(2)S_d(b-6)$ (keeping the same coefficients).

For $a = 3$, the recursion length is 6, and you add $a_i = b - 3$ to the corresponding list.

For $a = 4$, the recursion length is 18, and to get $S_d(4)S_d(b)$ from $S_d(4)S_d(b-18)$ you add the following to the corresponding sum (with weights adjusted as usual):

$$-S_d(b, 4) - S_d(b-2, 6) + S_d(b-4, 8) - S_d(b-10, 14) + S_d(b-12, 16) + S_d(b-14, 18).$$ □

4.3 For general q

4.3.1

We hope to address this in a sequel, but content ourselves here with predictions of the simplest features: In general, there should be recursions of the type above with the recursion length corresponding to a being $(q-1)q^n$, where q^n is the smallest power of q not less than a. The simplest case prediction is

$$S_d(1)S_d(a) = S_d(a+1) + \sum_{i<a,\ \text{“even”}} S_d(a+1-i, i).$$

When a is “even,” this is proved exactly as in the proof of Theorem 3: To pick up the exponents $a+1$ with a “even,” we just look at $\sum_{\theta \in \mathbb{F}_q^*} f(\theta x)$, where the left side is $xf(x)$.

4.3.2

For any a, b, we should have $2 \leq a_i \leq b$, if $a \leq b$ (which can be assumed by renaming). Now, of course, one needs to predict signs in $\mathbb{F}_q^*$ in front of the multizeta values as well as a_i's.

4.3.3

For weight, depth, and parity restrictions on possible relations, see the discussion in the next section.

4.3.4

Ongoing computer calculations by Javier Diaz–Vargas and Alejandro Lara–Rodriguez at the University of Arizona suggest that many features of the recipe in 4.1 generalize nicely for a general q, and in particular, that the recursion length prediction in 4.3.1 can even be improved from $(q-1)q^n$, mentioned there, to $(q-1)p^n$.

5 General Restrictions on Motivic Identities

We discuss some observations, heuristics, and proofs giving some general restrictions on possible relations.

5.1 Motivic identities

In [4], the multizeta values are expressed as periods of some explicit Anderson's t-motives (these "mixed Carlitz–Tate t-motives" are analogs of iterated extensions of Tate motives that occur in the classical case, but analogs of several structures in the classical theory, such as connection with the fundamental group of projective line minus three points, are missing or unclear right now). Hence, the algebraic relations between them should come from relations between the motives. A precise version of this general expectation (in the classical case, this is called the Grothendieck conjecture) has been proved in [3, 10]. But we will work here at the more naive level.

From the formulas for $S_d(k)$ in 3.3 (see [5, 3.7.4] for full details, but with a slightly different notation), we see that $S_d(k)$ is of the form $h_d(k)/\ell_d^k$, where $h_d(k)$ is the specialization of a polynomial $H_k(T, t)$ (in T with coefficients rational functions in t) at the graph of dth power Frobenius, namely at $T = t^{q^d}$. Let us call any function with this property an F-function. See [5] for the generating function for H_k and for the description of how they enter motivic (or rather t-motivic) picture at the zeta level (i.e., depth one). See [4] for the general multizeta values case.

So, the multizeta relations should come from the relations between these polynomials and their generalizations for higher depths (e.g., the corresponding polynomials for the $S_{<d}$'s that we have given above). The classical sum shuffle-type identity corresponds to a relation between H_k's alone, whereas other identities also need iterated polynomials.

Note that given an identity of the shuffle type, it is mechanical to prove it, assuming it is "motivic" in this naive sense that it works at a degree level. (Given an identity scheme (see the start of Section 4), of course, it is no longer mechanical to verify it.) Thus, we can effectively determine all the motivic identities for a given weight and a given q.

5.2 Weight preservation

Out of the $\mathbb{F}_p$ and $\mathbb{F}_p(t)$-coefficient identities that we have seen, the first work at the degree level and the second almost seem to work at that level, except some degrees need to be combined (see 3.4.7). If we assume this to be a feature of the motivic identities (this is almost justified above with reference to [4]), then we see that the motivic identities preserve weight: This is because they reduce to the identities between F-functions (defined in 5.1) in the numerator when we make a common denominator ℓ_d^w, for the common weight w, and if we mix different weights, and bring a common denominator ℓ_d^w, for any w, then ℓ_d^m's which are not F-functions will be involved in numerators. (In fact, if it represented by a two-variable function H in the manner above, then $H(T^q, t)/H(T, t) = (T - t)^m$, solution to which (denoted by Ω^m and studied in [3–5]) is not F-function. (Following simple degree calculation proves a special case: For large enough d, the degree of F-function would be $m_1 q^d + m_2$ for some integers m_i, whereas ℓ_d^m has degree $m(q^{d+1} - q)/(q - 1)$, giving a contradiction at least when $q - 1$ does not divide m.)

5.3 Depth and "even" restriction

Apart from the weight preservation, we expect that the depth filtration, though not the depth itself, is preserved in the multizeta value identities and that all the iterated indices are "even" (when $q = 2$, all indices are automatically "even") at S_d level, whereas at the ζ level, the non-"even" indices can appear via the sum shuffle. Here are some heuristics reasons for this:

(1) In defining a zeta value or a power sum, we fix a sign. The polynomials of degree at most d form a vector space, and the monic polynomials of degree d make an affine space, but the collection of all the monic polynomials of degree at most d, which appears in the iterated sums, is not so nice. Only "even" powers get rid of this sign problem.

(2) We saw in our generating function for $S_{<d}$ that only the "even" powers occurred.

(3) If $\sum_{i=1}^{d} S_i(k) = g_d(k)/\ell_d^k$ for an F-function $g_d(k)$, then

$$h_d(k)/\ell_d^k = S_d(k) = g_d(k)/\ell_d^k - g_{d-1}(k)/\ell_{d-1}^k.$$

Suppose $k \leq q$, then $h_d(k) = 1$ by 2.3 and the two-variable function G corresponding to $g_d(k)$ satisfies $G^{(1)} - (T - t)^k G = 1$. This equation has solution only if k is "even," as follows by the comparison of T-degree of both sides. More generally, if $\deg(h_d(k)) < kq^d$ (this happens often, but not always, see [12] and 5.6 of [11] for formulas for this degree), then we see that the top T-degrees in $G^{(1)}$ and $(T - t)^k G$ have to be the same to get cancellation, and hence k has to be "even." It is possible that this argument generalizes to all k's without size or degree restriction. We have checked it for $k \leq q^2$, by the formulas for $S_d(aq + b)$ above, by checking that in this case, degrees of $G^{(1)}$ and $(T - t)^k G$ have to be the same, leading to k being "even."

5.4 Only one shuffle

We have found in many cases one relation expressing the product of multizeta values as the sum of multizeta values, while classically there are two such relations. We claim that in our case, there is only one. It is enough to prove that a nontrivial $\mathbb{F}_p$-linear combination of the multizeta values is nonzero. Since we assume that such identities can only be motivic, we restrict to the multizeta values of the same weight. By calculations of degree in t, we can easily check the following:

(1) It works in weight not more than 3. For example, the multizeta values of weight 3, namely $\zeta(3)$, $\zeta(1,2)$, $\zeta(2,1)$, $\zeta(1,1,1)$ have degrees $0, -q, -2q$, and less than $-2q$, respectively, and hence they are linearly independent over $\mathbb{F}_p$.

(2) In weight 4, if $q=2$, out of 8 multizeta values, the only degree clashes are for $\zeta(3,1)$ and $\zeta(2,2)$ of degree -4, but their sum being of degree -6 does not cancel.

(3) If we consider the product of two zeta values as sum of multizeta values of depth at most 2, then for weight $k \le q+1$ the degree of $\zeta(a, k-a)$ being $-aq$, there is only one relation as there are no degree clashes.

(4) If we assume relations are motivic and keep depth filtration, then here is a proof in the case of product of zeta values: It is enough to prove that depth two multizeta values are linearly independent over $\mathbb{F}_p$, and by the degree bounds, which we know [12] in depth two for any q, it is thus enough to check the lowest terms. In other words, it is enough to show that $\mathbb{F}_p$-span of $S_1(k)$'s as k ranges between 1 and m is m-dimensional. Now $S_1(k+1)$ is the coefficient of x^k in $(1+\sum+\sum^2+\cdots+\sum^k)/\ell_1$, where $\sum = (x-x^q)/\ell_1$. Hence, $S_1(k)$ is $1/\ell_1^k$ plus smaller powers of $1/\ell_1$, and hence the space is the same as that generated by $1/\ell_1^k$'s, and hence it is m-dimensional.

5.5 Questions

It is straightforward to generalize the definition of the multizeta values from the $\mathbb{F}_q[t]$ case that we focused on here to more general A's in the Drinfeld module context (these are coordinate rings of a complete, nonsingular curve over a finite field minus a point) in either relative or absolute context, using the recipes in the zeta case (5.1 in [11], [8]). We have shown using a different kind of proofs that some of our theorems proved for $\mathbb{F}_q[t]$ here are in fact universal, some universal for a given q, and some need modification in the general case. In a sequel to this paper, we hope to address these issues and our continuing work on some of the following interesting areas: Determination of all identities, situation in higher depths, counting dimensions of $\mathbb{F}_p(t)$-spans (dimension of $\mathbb{F}_p$-span would be the number of all multizeta under consideration, if the phenomena explained above hold in general), transcendence degrees for a given weight (and q), algebraic independence results, interpolated values [11, 5.10] and negative integer arguments, Hopf-algebra, and other relevant underlying structures.

Acknowledgments

I thank Pierre Deligne and Greg Anderson for their encouragement and for interesting conversations about structures related to multizeta. I thank the von Neumann Fund and the Ellentuck Fund for their support during my stay at the Institute for Advanced Study, Princeton, where this paper was mostly written. I also thank the referee for suggestions to improve the exposition. The author was supported by National Security Agency grant H98230-08-1-0049.

References

[1] Anderson, G. W. "t-motives." *Duke Mathematical Journal* 53 (1985): 457–502.

[2] Anderson, G. W. "Digit patterns and the formal additive group." *Israel Journal of Mathematics* 161 (2007): 125–39.

[3] Anderson, G. W., W. D. Brownawell, and M. A. Papanikolas. "Determination of the algebraic relations among special Γ-values in positive characteristic." *Annals of Mathematics* 160, no. 2 (2004): 237–313.

[4] Anderson, G. W., and D. S. Thakur. "Multizeta values for $F_q[t]$, their period interpretation, and relations between them." *International Mathematics Research Notices* 2009 (forthcoming).

[5] Anderson, G. W., and D. S. Thakur. "Tensor powers of the Carlitz module and zeta values." *Annals of Mathematics* 132, no. 2 (1990): 159–91.

[6] Anderson, G. W., and D. S. Thakur. "Ihara power series for $\mathbb{F}_q[t]$." (forthcoming).

[7] Carlitz, L. "On certain functions connected with polynomials in a Galois field." *Duke Mathematical Journal* 1 (1935): 139–58.

[8] Goss, D. *Basic Structures of Function Field Arithmetic*. Ergebnisse der Mathematik und ihrer Grenzgebiete 35. Berlin: Springer, 1996.

[9] Masri, R. "Multiple zeta values over global function fields." *Proceedings of the Symposium in Pure Mathematics* 75 (2006): 157–75.

[10] Papanikolas, M. "Tannakian duality for Anderson–Drinfeld motives and algebraic independence of Carlitz logarithms." *Inventiones Mathematicae* 171 (2008): 123–74.

[11] Thakur, D. S. *Function Field Arithmetic*. Hackensack, NJ: World Scientific, 2004.

[12] Thakur, D. S. "Power sums with applications to multizeta and zeta zero distribution for $\mathbb{F}_q[t]$." (forthcoming).