DOUBLE ZETA VALUES AND MODULAR FORMS

HERBERT GANGL, MASANOBU KANEKO, AND DON ZAGIER

Dedicated to the memory of Tsuneo Arakawa

1. Introduction and main results

The *double zeta values*, which are defined for integers $r \geqslant 2$, $s \geqslant 1$, by

$$\zeta(r,s) \;=\; \sum_{m>n>0} \frac{1}{m^r n^s}\,, \tag{1}$$

are subject to numerous relations. Already Euler found that when the weight $k = r+s$ is odd the double zeta values can be reduced to products of usual zeta values. Furthermore, he gave the sum formula

$$\sum_{r=2}^{k-1} \zeta(r, k-r) \;=\; \zeta(k) \qquad (k>2). \tag{2}$$

The aims of the present paper are:

- to give other interesting relations among double zeta values,
- to show that the structure of the $\mathbb{Q}$-vector space of all relations among double zeta values of weight k is connected in many different ways with the structure of the space of modular forms M_k of weight k on the full modular group $\Gamma_1 = \mathrm{PSL}(2,\mathbb{Z})$, and
- to introduce and study both transcendental and combinatorial "double Eisenstein series" which explain the relation between double zeta values and modular forms and provide new realizations of the space of double zeta relations.

Double zeta values are a special case of *multiple zeta values*, defined by sums like (1) but with longer decreasing sequences of integers, which are known to satisfy a collection of relations called the *double shuffle relations* (cf., e.g., [3], [5], [12]). The specialization of these relations to the double zeta case is given by the following two sets of easily proved relations (see Section 2):

$$\begin{gathered}\zeta(r,s)+\zeta(s,r) = \zeta(r)\,\zeta(s) - \zeta(k) \qquad (r+s=k;\ r,s \geqslant 2)\,,\\ \sum_{r=2}^{k-1}\left[\binom{r-1}{j-1}+\binom{r-1}{k-j-1}\right]\zeta(r,k-r) = \zeta(j)\,\zeta(k-j) \qquad (2\leqslant j\leqslant \frac{k}{2})\,.\end{gathered} \tag{3}$$

We wish to study the relations which can be deduced from (3). Since we want to do this algebraically, it is useful to work, not with the double zeta values themselves, which for all we know may satisfy other relations than (3) (it is not even known that any $\zeta(r,s)/\pi^{r+s}$ is irrational), but with the *formal double zeta space* $\mathcal{D}_k$, generated by formal symbols $Z_{r,s}$, $P_{r,s}$ and Z_k subject to the relations (3), with $Z_{r,s}$, $P_{r,s}$ and Z_k taking the role of $\zeta(r,s)$, $\zeta(r)\zeta(s)$ and $\zeta(k)$, respectively, and where r and s are allowed to assume the value 1.

In $\mathcal{D}_k$ we can prove a number of explicit relations. In particular, Euler's result that all $Z_{r,s}$ are rational linear combinations of the $P_{r,s}$ when the weight k is odd holds in the formal double zeta space $\mathcal{D}_k$, so that we can (and usually will) assume that k is even. Similarly, the formal analogue of Euler's sum formula (2) holds in

$\mathcal{D}_k$, and in fact (for k even) has a refinement giving the sums of the even- and odd-argument double zeta values of weight k separately. Surprisingly, they are always in the ratio 3:1, independently of k:

Theorem 1. *For even $k > 2$, one has*

$$\sum_{\substack{r=2 \\ r \text{ even}}}^{k-1} Z_{r,k-r} = \frac{3}{4}\, Z_k\,, \qquad \sum_{\substack{r=2 \\ r \text{ odd}}}^{k-1} Z_{r,k-r} = \frac{1}{4}\, Z_k\,. \tag{4}$$

As an example of a more complicated identity, we show that, for $m, n \geqslant 1$ odd, $m + n = k > 2$,

$$2\sum_{\nu=0}^{n-1} \binom{-m}{\nu} B_\nu\, Z_{n-\nu,m+\nu} = \sum_{r+s=k} (-1)^{s-1} \lambda_{m,n}(r,s) P_{r,s}\,, \tag{5}$$

where B_ν is the νth Bernoulli number and

$$\lambda_{m,n}(r,s) = \sum_{\nu=0}^{n-1} \binom{m+\nu-1}{\nu} \binom{r-1}{n-\nu-1} B_\nu \tag{6}$$

(which despite appearances is symmetric in r and s). Since $B_\nu = 0$ for all odd ν except $\nu = 1$, this implies that any $Z_{\mathrm{ev,ev}}$ can be written in terms of $Z_{\mathrm{od,od}}$'s and $P_{r,s}$'s. But in fact only $Z_{\mathrm{od,od}}$'s are required:

Theorem 2. *Let $k > 2$ be even. Then the $Z_{r,k-r}$ with $0 < r < k$ odd are a basis of $\mathcal{D}_k$. There are explicit representations of the elements of various bases of $\mathcal{D}_k$ as linear combinations of the $Z_{\mathrm{od,od}}$'s.*

Theorem 2 will be proved in Section 4 by rewriting the defining relations (3) of $\mathcal{D}_k$ algebraically in terms of the action of the group ring $\mathbb{Z}[\Gamma_1]$ on a space of polynomials. This leads to both a simple proof of the first statement and to several concrete versions of the second. One of these, a variant of (5), is

$$Z_{m+1,k-m-1} + \frac{1}{2} Z_k \;=\; -\frac{2}{m} \sum_{\substack{r+s=k \\ r,\, s \geqslant 1 \text{ odd}}} \lambda^0_{m,k-m}(r,s)\big(Z_{r,s} + \frac{1}{2} Z_k\big) \tag{7}$$

for $m = 1, 3, \ldots, k-3$, where

$$\lambda^0_{m,n}(r,s) = \lambda_{m,n}(r,s) - \binom{s-1}{m-1} B_{s-m} = \sum_{\ell=0}^{r-2} \binom{k-2-\ell}{m-1} \binom{r-1}{\ell} B_{n-\ell-1}$$

(with $B_\nu = 0$ for $\nu < 0$). Since Z_k equals $4\sum_{r>1 \text{ odd}} Z_{r,k-r}$ by Theorem 1, this expresses all even-argument double zeta values in terms of odd-argument ones.

Theorem 2 is false for double zeta values. Instead we have the following result, which gives the first connection with modular forms:

Theorem 3. (Rough statement.) *The values $\zeta(\mathrm{od},\mathrm{od})$ of weight k satisfy at least $\dim S_k$ linearly independent relations, where S_k denotes the space of cusp forms of weight k on Γ_1.*

Example. For $k = 12$ and $k = 16$, the first weights for which there are non-zero cusp forms on Γ_1, we have the identities

$$28\,\zeta(9,3) + 150\,\zeta(7,5) + 168\,\zeta(5,7) \;=\; \frac{5197}{691}\,\zeta(12) \tag{8}$$

$$66\,\zeta(13,3) + 375\,\zeta(11,5) + 686\,\zeta(9,7) \;+\; 675\,\zeta(7,9) + 396\,\zeta(5,11) \;=\; \frac{78967}{3617}\,\zeta(16)\,,$$

which can be written in terms only of $\zeta(\mathrm{od},\mathrm{od})$'s using Theorem 1. Conjecturally (and numerically), these are the only relations over $\mathbb{Q}$ among odd-argument double zeta values up to weight $\leqslant 16$, and more generally we expect that there are no further relations among the $\zeta(\mathrm{od},\mathrm{od})$ except the ones predicted by Theorem 3.

Although Theorem 3 holds for the "true" double zeta world and is false in the formal one, it is in fact a consequence of a result in the formal space. In fact, it follows from two different—though complementary—results. Both of them involve period polynomials. We recall the definition of these polynomials. (A more detailed review will be given in Section 5.) For each even k we consider the space V_k of homogeneous polynomials of degree $k-2$ in two variables and the subspace $W_k \subset V_k$ of polynomials satisfying the relations $P(X,Y)+P(-Y,X)=0$, $P(X,Y)+P(X-Y,X)+P(Y,Y-X)=0$. It splits as the direct sum of subspaces W_k^+ and W_k^- of polynomials which are symmetric and antisymmetric with respect to $X \leftrightarrow Y$, with the former being odd and the latter even with respect to $X \mapsto -X$. The Eichler-Shimura-Manin theory tells us that there are canonical isomorphisms over $\mathbb{C}$ between S_k and W_k^+ and between M_k and W_k^-. The full statement of Theorem 3, given in Section 5, associates to any polynomial in W_k^-, in an injective way, an explicit relation among the numbers $Z_{\mathrm{od},\mathrm{od}}$ and $P_{\mathrm{ev},\mathrm{ev}}$ (and Z_k). For the above example (8), for instance, the polynomial $X^2Y^2(X^2-Y^2)^3$ in W_{12}^- leads to the relation

$$28\,Z_{9,3}+150\,Z_{7,5}+168\,Z_{5,7} \;=\; 28P_{4,8}+\frac{95}{3}P_{6,6}-\frac{167}{3}Z_{12}\,, \tag{9}$$

which by Euler's theorem agrees with (8) modulo $\mathbb{Q}\pi^{12}$, and similarly the complete version of the relation given above between odd double zeta values in weight 16 is

$$\begin{aligned}66\,Z_{13,3}&+375\,Z_{11,5}+686\,Z_{9,7}+675\,Z_{7,9}+396\,Z_{5,11}\\ &=\;66P_{4,12}+185P_{6,10}+\frac{364}{3}P_{8,8}-\frac{1081}{3}Z_{16}\,.\end{aligned}$$

The other result about formal double zeta values which implies Theorem 3 involves the space W_k^+ rather than W_k^-. More precisely, it involves a certain 1-dimensional extension $\widehat{W}_k^+ \subset V_k+\mathbb{C}\cdot\left(X^{k-1}Y^{-1}+X^{-1}Y^{k-1}\right)$ (see Section 6 for details) which is isomorphic to M_k rather than S_k:

Theorem 4. *If $\{Z_{r,s},\,P_{r,s},\,Z_k\}$ is a collection of numbers satisfying the double shuffle relations in weight k, then the polynomial*

$$\sum_{\substack{r+s=k\\ r,\ s\ \text{even}}} P_{r,s}\,X^{r-1}Y^{s-1}\;-\;\frac{Z_k}{2}\left(X^{k-1}Y^{-1}+X^{-1}Y^{k-1}\right)$$

belongs to $\widehat{W}_k^+$ (and to W_k^+ if $Z_k=0$). Every element of $\widehat{W}_k^+$ arises in this way.

From one point of view, this says that the subspace $\mathcal{P}_k^{\mathrm{ev}}$ of $\mathcal{D}_k$ spanned by the $P_{r,s}$ with r and s even is canonically dual to $\widehat{W}_k^+$. From another, it says that there are $k/6+\mathrm{O}(1)$ relations among the $P_{\mathrm{ev},\mathrm{ev}}$, these relations being the same as the relations satisfied by the coefficients of period polynomials in W_k^+. In fact, we will prove Theorem 4 in this form. It is this point of view which leads to the most direct connection with modular forms, because it is known (as a consequence of the so-called Rankin-Selberg or unfolding method) that the coefficients of (extended) symmetric period polynomials satisfy the same linear relations as the products $G_rG_s \in M_k$ $(r+s=k)$, where G_r denotes the Eisenstein series of weight r on $\mathrm{PSL}_2(\mathbb{Z})$. (When r or s is equal to 2, the product G_2G_{k-2} must be modified slightly by adding an appropriate multiple of G'_{k-2} to compensate for the non-modularity of G_2.) Thus the proof of Theorem 4, combined with the known facts that the products G_rG_s span M_k and, after dividing by π^k, have rational Fourier coefficients, also leads to the following, more intuitive, statement:

Theorem 5. *The space $\mathcal{P}_k^{\mathrm{ev}}$ is canonically isomorphic to $M_k^{\mathbb{Q}}$, by a map which sends $P_{r,s}$ to $(2\pi i)^{-k}G_rG_s$ (plus a multiple of G'_{k-2} if r or $s = 2$) and Z_k to $(2\pi i)^{-k}G_k$.*

Theorem 5 tells us that there is a realization of the symmetric ($\mathcal{P}$-) part of the double shuffle relations given by products of Eisenstein series. This implies by linear algebra that there must be a realization (and in fact, infinitely many realizations) of the full space $\mathcal{D}_k$ having these products as its symmetric part. It is then natural to ask whether there is a natural choice of such a realization. In the last part of the paper we show, following an idea already adumbrated in [15], that there are in fact two such choices. More precisely, we show that one can extend the map $\mathcal{P}_k^{\mathrm{ev}} \to M_k$ in two different ways to a map from $\mathcal{D}_k$ to a larger space of functions, by finding "double Eisenstein series" which are related to products of Eisenstein series in exactly the same way as double zeta values are related to products of Riemann zeta values. One of these ways is transcendental, in terms of holomorphic functions in the upper half plane, and the other combinatorial, in terms of formal power series in q with rational coefficients. Both ways are interesting, and they also turn out to be related: the Fourier expansion of the transcendental double Eisenstein series splits up into three terms, the most complicated of which is (a multiple of) the combinatorial double Eisenstein series. We now explain this in more detail.

The transcendental version of the double Eisenstein series $G_{r,s}(\tau)$ is defined, in complete analogy with (1), as

$$G_{r,s}(\tau) = \sum_{\substack{\mathbf{m},\,\mathbf{n}\in\mathbb{Z}\tau+\mathbb{Z}\\ \mathbf{m}\succ\mathbf{n}\succ 0}} \frac{1}{\mathbf{m}^r\mathbf{n}^s} \qquad (\tau\in\mathfrak{H} = \text{upper half-plane}),$$

where $\mathbf{n}\succ 0$ means $\mathbf{n} = n\tau + b$ with $n > 0$ or $n = 0, b > 0$ and $\mathbf{m}\succ\mathbf{n}$ means $\mathbf{m}-\mathbf{n}\succ 0$. The series converges absolutely for $r\geqslant 3$, $s\geqslant 2$, and also makes sense for $s = 1$ if the sum over $\mathbf{n}$ (for $\mathbf{m}$ fixed) is interpreted as a Cauchy principal value. The same combinatorial proof that establishes (3) shows that, at least in the convergent cases, the corresponding equations still hold with $\zeta(r,s)$ replaced by $G_{r,s}(\tau)$ and with (each) $\zeta(k)$ replaced by the function

$$G_k(\tau) = \sum_{\substack{\mathbf{m}\in\mathbb{Z}\tau+\mathbb{Z}\\ \mathbf{m}\succ 0}} \frac{1}{\mathbf{m}^k} \qquad (k\geqslant 2)$$

(again to be interpreted as a Cauchy principal value if $k = 2$), which equals the previously mentioned Eisenstein series if k is even. In other words, at least for the cases of absolute convergence, we have a realization of the double shuffle relations on the space of holomorphic functions in $\mathfrak{H}$ given by

$$Z_{r,s}\mapsto G_{r,s}(\tau), \qquad P_{r,s}\mapsto G_r(\tau)G_s(\tau), \qquad Z_k\mapsto G_k(\tau)\,.$$

The combinatorial/arithmetic aspect emerges when we study the Fourier expansions of the single and double Eisenstein series. The former are given by the well-known formula

$$(2\pi i)^{-k}G_k(\tau) \;=\; \widetilde{\zeta}(k) \;+\; g_k(q)\,, \tag{10}$$

where $q = e^{2\pi i\tau}$, $\widetilde{\zeta}(k) = (2\pi i)^{-k}\zeta(k)$, and

$$g_k(q) \;=\; \frac{(-1)^k}{(k-1)!}\sum_{u,n>0} u^{k-1}\,q^{un} \qquad (k\geqslant 2)\,. \tag{11}$$

The corresponding result for $G_{r,s}(\tau)$ is given by

Theorem 6. *The Fourier expansion of $G_{r,s}(\tau)$ for $r\geqslant 3,\ s\geqslant 2$ is given by*

$$(2\pi i)^{-r-s}\,G_{r,s}(\tau) \;=\; \widetilde{\zeta}(r,s) \;+\; \sum_{\substack{h+p=r+s\\ h,\,p>1}} C^p_{r,s}\,g_h(q)\,\widetilde{\zeta}(p) \;+\; g_{r,s}(q) \tag{12}$$

with $q = e^{2\pi i\tau}$, $\widetilde{\zeta}(r,s) = (2\pi i)^{-r-s}\zeta(r,s)$,

$$C^p_{r,s} \;=\; \delta_{s,p} \;+\; (-1)^s\binom{p-1}{s-1} \;+\; (-1)^{p-r}\binom{p-1}{r-1} \quad \in\; \mathbb{Z} \tag{13}$$

and

$$g_{r,s}(q) \;=\; \frac{(-1)^{r+s}}{(r-1)!(s-1)!}\sum_{\substack{m>n>0\\ u,\,v>0}} u^{r-1}v^{s-1}q^{um+vn} \qquad \in\; \mathbb{Q}[[q]]\,. \tag{14}$$

We can reinterpret this theorem in the light of the following considerations. If k is even, the case when $G_k(\tau)$ is modular (or quasi-modular if $k=2$), then by Euler's theorem the number $\widetilde{\zeta}(k)$ occurring on the right-hand side of (10) is the rational number $-B_k/2k!$, which we denote by β_k. Hence this right-hand side can be replaced by the expression

$$Z_k(q) \;=\; g_k(q) \;+\; \beta_k \qquad (k\geqslant 2)\,, \tag{15}$$

which we call the *combinatorial Eisenstein series* because it is purely combinatorially defined as an element of $\mathbb{Q}[[q]]$ and is proportional to the usual Eisenstein series (and hence modular) when k is even and $\geqslant 4$. In the same way, we define

$$\beta_{r,s}(q) = \sum_{h+p=r+s} C^p_{r,s}\,\beta_p\,g_h(q) \qquad (r,\,s\geqslant 2) \tag{16}$$

and set

$$Z_{r,s}(q) \;=\; g_{r,s}(q) \;+\; \beta_{r,s}(q) \qquad (r\geqslant 3,\ s\geqslant 2)\,, \tag{17}$$

the *combinatorial double Eisenstein series.* Then the right-hand side of (12) can be rewritten as

$$\widetilde{\zeta}(r,s) \;+\; \sum_{\substack{h+p=r+s\\ h,\,p>1,\,p\ \text{odd}}} C^p_{r,s}\,\widetilde{\zeta}(p)\,g_h(q) \;+\; Z_{r,s}(q)\,. \tag{18}$$

The three pieces in (18) lie in three non-intersecting $\mathbb{Q}$-subspaces of $\mathbb{C}[[q]]$: the first term is in $\mathbb{C}$ (more precisely, in $\mathbb{R}$ or $i\mathbb{R}$ depending on the parity of $r+s$), the second term in $i\mathbb{R}[[q]]^0$, and the third in $\mathbb{Q}[[q]]^0$, where $A[[q]]^0 = qA[[q]]$ denotes the space of power series without constant term with coefficients in a vector space A. The first term is our familiar double zeta value realization of the double shuffle relations. The second also fulfils the double shuffle relations, independently of the arithmetic natures of $\widetilde{\zeta}(p)$ and $g_h(q)$, because by a simple result which will be proved in Section 2 (Corollary 2) the numbers $C^p_{r,s}$ for any odd value of p less than $r+s$ already satisfy these relations. The following theorem, which we will prove in Section 7, says that the combinatorial double Eisenstein series, suitably extended to the missing cases $r=1,\,2$ and $s=1$, also satisfies the double shuffle relations.

Theorem 7. (Rough statement.) *There is a realization of the double shuffle relations in $\mathbb{Q}[[q]]^0$ which in the region corresponding to absolute convergence agrees with* (17) *and* (15) *and sends $P_{r,s}$ to $Z_r(q)Z_s(q) - \beta_r\beta_s$ for $r,\,s>2$.*

If we now use (18) with the extended definition of $Z_{r,s}(q)$ to *define* the double Eisenstein series $G_{r,s}(\tau)$ in the previously undefined cases $r=1$, $r=2$, and $s=1$, then we find that there is also a realization $\{Z_{r,s}, P_{r,s}, Z_k\}$ of the double shuffle relations in the space of holomorphic functions in the upper half-plane which maps $Z_{r,s}$ to $G_{r,s}(\tau)$ for $r\geqslant 3$, $s\geqslant 2$, $P_{r,s}$ to $G_r(\tau)G_s(\tau)$ for $r,\, s>2$ and maps Z_k to $G_k(\tau)$ for all $k>2$.

Remark. Some of the ideas developed in this paper were already mentioned, in a very preliminary form, in [15] and [16]. The discovery that there are unexpected relations among $\zeta(\mathrm{od},\mathrm{od})$'s starting in weight 12 originated with a question posed by T. Terasoma about the linear independence, modulo π^{12}, of $\zeta(r,s)$ with $r,\, s>1$ odd, $r+s=12$. We also mention that there is a related phenomenon for the "stable derivation algebra" of Y. Ihara [4] inside the Lie algebra of derivations of the free Lie algebra on two generators. The recent paper of L. Schneps [14] should have a close connection to our present work. Also related are several results of A.B. Goncharov, who defined a coproduct structure on (formal) multiple zeta values in [6] and described relations between double zeta values and the cohomology of $\mathrm{PSL}_2(\mathbb{Z})$ in [5].

Acknowledgements. H. G. and M. K. gratefully acknowledge financial support by the Collège de France (Paris) and the Max-Planck-Institut (Bonn). M. K. was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 15340014, 2003–2005.

2. The formal double zeta space

We begin by discussing the double shuffle relations (3). The first follows from the obvious decomposition of lattice points in $\mathbb{N}\times\mathbb{N}$ into the three disjoint subsets $\{(m,n)\mid m>n\}$, $\{(m,n)\mid m<n\}$ and $\{(m,n)\mid m=n\}$, giving the identity

$$\Big(\sum_{m>n}+\sum_{m<n}+\sum_{m=n}\Big)\frac{1}{m^r n^s}=\sum_{m\geqslant 1}\frac{1}{m^r}\sum_{n\geqslant 1}\frac{1}{n^s}\,,$$

which is precisely the first equation in (3). For the second, we can use the partial fraction expansion

$$\frac{1}{m^i n^j}=\sum_{r+s=k}\left[\frac{\binom{r-1}{i-1}}{(m+n)^r n^s}+\frac{\binom{r-1}{j-1}}{(m+n)^r m^s}\right]\qquad (i+j=k)\,. \tag{19}$$

(*Proof:* Compute the poles of both sides as rational functions of n, with m fixed.)

In the formal setting, it is convenient to extend the set of generators and relations in (3) slightly by including the case $r=1$ (in the case of double zeta values, this would give a non-convergent series): we introduce formal variables $Z_{r,s}$, $P_{r,s}$ and Z_k and impose the relations

$$\begin{aligned} Z_{r,s}\;+\;Z_{s,r}\;&=\;P_{r,s}-Z_k \qquad (r+s=k)\,,\\ \sum_{r+s=k}\left[\binom{r-1}{i-1}\;+\;\binom{r-1}{j-1}\right]Z_{r,s}\;&=\;P_{i,j}\qquad (i+j=k)\,. \end{aligned}\tag{20}$$

(From now on, whenever we write $r+s=k$ or $i+j=k$ without comment, it is assumed that the variables are integers $\geqslant 1$.)

The *formal double zeta space* is now defined as the $\mathbb{Q}$-vector space

$$\mathcal{D}_k=\frac{\{\mathbb{Q}\text{-linear combinations of formal symbols } Z_{r,s},P_{r,s},Z_k\}}{\big\langle\text{relations (20)}\big\rangle}\,.$$

Alternatively, since Eqs. (20) express the $P_{i,j}$ in terms of the $Z_{r,s}$, we can define $\mathcal{D}_k$ as

$$\mathcal{D}_k = \frac{\{\mathbb{Q}\text{-linear combinations of formal symbols } Z_{r,s}, Z_k\}}{\langle \text{relation (22)} \rangle}, \tag{21}$$

where relation (22) is given by taking the difference of Eqs. (20):

$$\sum_{r+s=k} \left[\binom{r-1}{i-1} + \binom{r-1}{j-1} \right] Z_{r,s} = Z_{i,j} + Z_{j,i} + Z_k \qquad (i+j=k). \tag{22}$$

Of course, since both sides of (22) are symmetric in i and j, it is enough to take (22) for $i \leqslant j$. We thus have (for k even) k generators and $k/2$ relations, so

$$\dim \mathcal{D}_k \geqslant \frac{k}{2} \qquad (k \text{ even}). \tag{23}$$

(We will see below that in fact equality holds.) Finally, we define the *A-valued points* $\mathcal{D}_k(A)$ of $\mathcal{D}_k$ for any $\mathbb{Q}$-vector space A by

$$\mathcal{D}_k(A) = \mathrm{Hom}_{\mathbb{Q}}(\mathcal{D}_k, A) = \{(Z_{r,s}, Z_k)_{r+s=k} \in A^k\,,\ \text{satisfying (22)}\}\,;$$

this can also be represented as the set of $(2k-1)$-tuples $(Z_{r,s}, P_{r,s}, Z_k)$ satisfying (20), and we will use both forms. An element of $\mathcal{D}_k(A)$ will be called a *realization* of the double zeta space in A. For example, with $A = \mathbb{R}$ and any $\kappa \in \mathbb{R}$ we have an $\mathbb{R}$-realization of $\mathcal{D}_k$ (for $k > 2$) given by

$$\begin{aligned}
Z_{r,s} &\mapsto \begin{cases} \zeta(r,s), & \text{if } r > 1, \\ \kappa, & \text{if } r = 1, \end{cases} \\
P_{r,s} &\mapsto \begin{cases} \zeta(r)\zeta(s), & \text{if } r,s > 1, \\ \kappa + \zeta(k-1,1) + \zeta(k), & \text{if } r = 1 \text{ or } s = 1, \end{cases} \\
Z_k &\mapsto \zeta(k)\,.
\end{aligned} \tag{24}$$

Here we could also treat κ as a variable and consider this as a realization in $\mathbb{R} + \mathbb{Q}\cdot\kappa$ or $\mathbb{R}[\kappa]$.

We now introduce two convenient ways to work with $\mathcal{D}_k$. The first is by generating functions. Let $(Z_{r,s}, P_{r,s}, Z_k)_{r+s=k} \in \mathcal{D}_k(A)$ be a realization of $\mathcal{D}_k$ in A. Then we can see easily that the identities (20) are equivalent to the relations

$$\begin{aligned}
\mathfrak{Z}_k(X,Y) + \mathfrak{Z}_k(Y,X) &= \mathfrak{P}_k(X,Y) - Z_k \cdot \frac{X^{k-1} - Y^{k-1}}{X - Y}, \\
\mathfrak{Z}_k(X+Y,Y) + \mathfrak{Z}_k(X+Y,X) &= \mathfrak{P}_k(X,Y)
\end{aligned} \tag{25}$$

for the generating functions

$$\mathfrak{Z}_k(X,Y) = \sum_{r+s=k} Z_{r,s} X^{r-1} Y^{s-1}\,, \qquad \mathfrak{P}_k(X,Y) = \sum_{r+s=k} P_{r,s} X^{r-1} Y^{s-1}$$

of the $Z_{r,s}$ and $P_{r,s}$, respectively, in $A[X,Y]$. (Equations (20) just express the equality of the coefficient of $X^{r-1}Y^{s-1}$ in (25).) Similarly, (22) is equivalent to the single relation

$$\begin{aligned}
\mathfrak{Z}_k(X+Y,Y) + \mathfrak{Z}_k(X+Y,X) - \mathfrak{Z}_k(X,Y) - \mathfrak{Z}_k(Y,X)& \\
= Z_k \cdot \frac{X^{k-1} - Y^{k-1}}{X - Y}&
\end{aligned} \tag{26}$$

for the polynomial $\mathfrak{Z}_k$.

As an example of the use of these equations, we will prove the first two identities mentioned in the Introduction, namely the fact that all $Z_{r,s}$'s are combinations of $P_{r,s}$'s and of Z_k if k is odd, and the separate even and odd sum formulas as given in Theorem 1 if k is even. For the first, we can work with (20) with the right-hand sides both replaced by 0 (because we want to work *modulo* all $P_{r,s}$'s and Z_k). Then (25) become simply $\mathfrak{Z}_k(X,Y)+\mathfrak{Z}_k(Y,X)=0$ and $\mathfrak{Z}_k(X+Y,Y)+\mathfrak{Z}_k(X+Y,X)=0$. Rewriting the latter equation as $\mathfrak{Z}_k(X,Y)+\mathfrak{Z}_k(X,X-Y)=0$, we see that $\mathfrak{Z}_k$ is anti-invariant under the two involutions $\varepsilon:(X,Y)\mapsto(Y,X)$ and $\tau:(X,Y)\mapsto(X,X-Y)$. Since $(\varepsilon\tau)^3$ maps (X,Y) to $(-X,-Y)$ and $\mathfrak{Z}_k$ is homogeneous of degree $k-2$, these two relations imply $\mathfrak{Z}_k(X,Y)=(-1)^k\mathfrak{Z}_k(X,Y)$, so $\mathfrak{Z}_k=0$ if k is odd, proving the first identity. (One can refine this proof to give an explicit formula for $\mathfrak{Z}_k(X,Y)$ as $A(X,Y)-A(X,X-Y)+A(Y,Y-X)$, where $A(X,Y)=\sum_{2|r}P_{r,s}X^{r-1}Y^{s-1}-\frac{Z_k}{2}\frac{X^{k-1}-Y^{k-1}}{X-Y}$.) For Theorem 1, it suffices to apply (26) with $(X,Y)=(1,0)$ and $(1,-1)$. This gives (for even k)

$$\mathfrak{Z}_k(1,1)-\mathfrak{Z}_k(0,1)=Z_k\,,\qquad \mathfrak{Z}_k(1,-1)-\mathfrak{Z}_k(0,1)=-\frac{1}{2}Z_k\,,$$

and Theorem 1 follows by adding and subtracting the equations.

We remark that it is occasionally convenient to work with the infinite product $\mathcal{D}=\prod_k\mathcal{D}_k$ consisting of collections of numbers $\bigl\{\{Z_{r,s}\}_{r,s\geqslant 1},\{P_{r,s}\}_{r,s\geqslant 1},\{Z_k\}_{k\geqslant 1}\bigr\}$ satisfying (20) for all k. Then the corresponding generating functions $\mathfrak{Z}(X,Y)$, $\mathfrak{P}(X,Y)$ and $\mathfrak{z}(T)=\sum_{k\geqslant 1}Z_kT^{k-1}$ satisfy

$$\begin{aligned}\mathfrak{Z}(X,Y)+\mathfrak{Z}(Y,X)\ &=\ \mathfrak{P}(X,Y)\ -\ \frac{\mathfrak{z}(X)-\mathfrak{z}(Y)}{X-Y}\,,\\ \mathfrak{Z}(X+Y,Y)+\mathfrak{Z}(X+Y,Y)\ &=\ \mathfrak{P}(X,Y)\,,\end{aligned}\tag{27}$$

and similarly for (26). For example, the reader may want to verify that the function $\mathfrak{Z}(X,Y)-\mathfrak{Z}(0,Y)$ is equal to $\sum_{m>n>0}X/m(m-X)(n-Y)$ for the realization (24) and to use this to verify the k-less version of (26) directly for this generating function. (The calculation—which requires some work—gives the result only up to an additive constant, corresponding to the fact that (24) holds only for $k>2$.)

The following proposition, which will be used in Section 7, gives some easy solutions of relations (26) (with $Z_k=0$).

Proposition 1. *Let $A(X,Y)\in V_k$ be a polynomial which is even with respect to Y. Then the function*

$$\mathfrak{Z}_k(X,Y)=A(X,Y)-A(X,X-Y)+A(Y,Y-X)\tag{28}$$

gives a realization of Equation (26) *with $Z_k=0$.*

Proof. One checks by direct calculation that if $\mathfrak{Z}_k(X,Y)$ is defined by (28) then both $\mathfrak{Z}_k(X,Y)+\mathfrak{Z}_k(Y,X)$ and $\mathfrak{Z}_k(X+Y,Y)+\mathfrak{Z}_k(X+Y,X)$ equal $A(X,Y)+A(Y,X)$. Note that the assertion of the proposition also holds if $A(X,Y)=A(Y,-X)$ or if A is anti-symmetric (with $\mathfrak{P}_k\equiv 0$ in the latter case). $\square$

Corollary. *Let $0<p<k$ be two integers with p odd. Then the numbers $Z_{r,s}=C^p_{r,s}$ $(r+s=k)$ with $C^p_{r,s}$ defined by Equation* (13) *satisfy* (22) *with $Z_k=0$.*

Proof. This is simply Proposition 2 applied to $A(X,Y)=X^{k-p-1}Y^{p-1}$. The corresponding numbers $P_{r,s}$ in (20) are equal to $\delta_{r,p}+\delta_{s,p}$. $\square$

The second way of working with $\mathcal{D}_k$ is by studying the *relations* among the $Z_{r,s}$ (or $Z_{r,s}$, $P_{r,s}$ and Z_k). The following result gives a useful description of them. We introduce the notation

$$V_k=\bigl\langle X^{r-1}Y^{s-1}\,\big|\,r+s=k\bigr\rangle\,,\qquad V_k^*=\bigl\langle\frac{1}{m^rn^s}\,\big|\,r+s=k\bigr\rangle\,.$$

We define an isomorphism $V_k\to V_k^*$ by

$$F(X,Y)=\sum_{r+s=k}\binom{k-2}{r-1}f_{r,s}X^{r-1}Y^{s-1}\ \mapsto\ F^*(m,n)=\sum_{r+s=k}\frac{f_{r,s}}{m^rn^s}\,.$$

Then we have the following

Lemma. *Let F, G, $H \in V_k$ and F^*, G^*, H^* the corresponding elements of V_k^*. Then the following two statements are equivalent:*

(i) $H^*(m,n) = F^*(m+n,n) + G^*(m,m+n)$,

(ii) $F(X,Y) = H(X,X+Y)\,, \quad G(X,Y) = H(X+Y,Y)$.

Proof. Equation (19) implies that any element $h \in V_k^*$ can be decomposed as $f(m+n,n) + g(m,m+n)$ for some f and g in V_k^*, and this decomposition is obviously unique since $f(1,x)$ has poles only at $x=0$ and $g(1-x,1)$ only at $x=1$. If $f = F^*$ etc., then an inspection of (19) shows that the coefficients $f_{r,s}$ and $g_{r,s}$ of F and G are related to the coefficients $h_{r,s}$ of H by

$$f_{r,s} = \sum_{i+j=k} \binom{r-1}{i-1} h_{i,j}\,, \qquad g_{r,s} = \sum_{i+j=k} \binom{s-1}{j-1} h_{i,j}\,.$$

Using the binomial coefficient identity $\binom{k-2}{r-1}\binom{r-1}{i-1} = \binom{k-2}{j-1}\binom{j-1}{s-1}$ $(r+s=i+j=k)$, we find that these formulas are equivalent to (ii). $\square$

Proposition 2. *Let $a_{r,s}$ and λ be rational numbers. Then the following three statements are equivalent:*

(i) The relation

$$\sum_{r+s=k} a_{r,s} Z_{r,s} = \lambda Z_k \tag{29}$$

holds in $\mathcal{D}_k$.

(ii) The generating function

$$A(X,Y) = \sum_{r+s=k} \binom{k-2}{r-1} a_{r,s}\, X^{r-1} Y^{s-1} \quad \in V_k \tag{30}$$

can be written as $H(X,X+Y) - H(X,Y)$ for some symmetric homogeneous polynomial $H \in \mathbb{Q}[X,Y]$ of degree $k-2$, and

$$\lambda = \frac{k-1}{2} \int_0^1 H(t,1-t)dt\,. \tag{31}$$

(iii) The generating function

$$A^*(m,n) = \sum_{r+s=k} \frac{a_{r,s}}{m^r n^s} \quad \in V_k^* \tag{32}$$

can be written as $f(m,n) - f(m+n,m) - f(m+n,n)$ for some $f \in V_k^*$, and

$$\lambda = \frac{f(1,1) - A^*(1,1)}{2} = f(2,1)\,. \tag{33}$$

Proof. If we choose the symmetric polynomial $H(X,Y) = X^{i-1}Y^{j-1} + X^{j-1}Y^{i-1}$ and use the binomial theorem to compute the $a_{r,s}$ in (30) and the beta integral to compute $\lambda = (i-1)!(j-1)!/(k-2)!$, then we find that (29) reduces to (22). Since these H's span the space of symmetric polynomials in V_k, this proves the equivalence of the first two statements.

The equivalence of (ii) and (iii) follows by applying the lemma with $F = A + H$, $G(X,Y) = F(Y,X)$ and $f = F^*$, $g(m,n) = f(n,m)$. To check that the values of λ in (ii) and (iii) agree, we again use the beta integral $\int_0^1 t^{r-1}(1-t)^{s-1}dt = \frac{(r-1)!(s-1)!}{(k-1)!}$ to get $(k-1)\int_0^1 H(t,1-t)dt = \sum h_{r,s} = H^*(1,1)$. $\square$

Remark. We can also write (31) as $\lambda = \frac{1}{2}\sum h_{r,s}$, where $H = \sum \binom{k-2}{r-1} h_{r,s} X^{r-1}Y^{s-1}$.

The two approaches outlined above are equivalent by a duality which we will discuss below, but it is very convenient to have both. As an example of the use of the proposition, we give a second quick proof of Theorem 1 from the Introduction. Taking $H = X^{k-2} + Y^{k-2}$ in the proposition gives $a_{r,s} = 1$ $(r \neq 1)$, $a_{1,k-1} = 0$, $\lambda = 1$, while taking $H = (X-Y)^{k-2}$ gives $a_{r,s} = (-1)^r$ $(r \neq 1)$, $a_{1,k-1} = 0$, $\lambda = \frac{1}{2}$. Again adding and subtracting the two relations thus obtained gives (4).

As a second example, we observe that Eq. (22) contains no $Z_{1,k-1}$, so that $Z_{1,k-1}$ is a free variable (as we already saw in the realization (24)). Thus $a_{1,k-1}$ must vanish in any relation of the form (29), and we can also see this in the proposition by setting $X = 0$.

3. Using the action of $\mathrm{PGL}_2(\mathbb{Z})$

We have already repeatedly used the space V_k of homogeneous polynomials of degree $k-2$ in X and Y. We now make this approach more systematic by exploiting two further structures on V_k: the action of the group $\Gamma = \mathrm{PGL}_2(\mathbb{Z})$ and the Γ-invariant scalar product. The former is defined in the obvious way by $(F|\gamma)(X,Y) = F(aX+bY, cX+dY)$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ (we suppose throughout that k is even) and the latter by

$$\left\langle X^{r-1}Y^{s-1},\, X^{m-1}Y^{n-1} \right\rangle = \frac{(-1)^r}{\binom{k-2}{m-1}} \delta_{(r,s),(n,m)} \tag{34}$$

for $r,s,m,n \geqslant 1$, $r+s = m+n = k$. The invariance property $\langle F|\gamma,\, G|\gamma\rangle = \langle F, G\rangle$ is easily checked. We extend the action of Γ on V_k to an action of the group ring $R = \mathbb{Z}[\Gamma]$ by linearity. Then $\langle F|\xi, G\rangle = \langle F, G|\xi^*\rangle$, where $\xi \mapsto \xi^*$ is the anti-automorphism of R induced by $\gamma \mapsto \gamma^{-1}$. We occasionally work with the model of V_k consisting of polynomials $f(x)$ of one variable of degree $\leqslant k-2$, corresponding to the homogeneous model via $f(x) = F(x,1)$, $F(X,Y) = Y^{k-2}f(X/Y)$. The group operation in this version takes the form $(f|\gamma)(x) = (cx+d)^{k-2} f((ax+b)/(cx+d))$.

The group Γ contains distinguished elements. First there are the commuting involutions

$$\varepsilon = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \delta = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

sending $F(X,Y)$ to $F(Y,X)$ and $F(-X,Y)$, respectively. The (±1)-eigenspaces of ε will be denoted by $V_k^{\pm}$ and the (±1)-eigenspaces of δ by V_k^{ev} and V_k^{od}; we also write $V_k^{+,\mathrm{ev}}$ for the space of even symmetric polynomials and similarly for the other three double eigenspaces of dimension $k/4 + \mathrm{O}(1)$. In $\mathrm{PSL}_2(\mathbb{Z})$, we have the elements

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = US = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad T' = U^2S = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

with the relations $S^2 = U^3 = 1$, $S = \varepsilon\delta$ and

$$\varepsilon U \varepsilon = U^2, \qquad \varepsilon T \varepsilon = T', \qquad \delta T \delta = T^{-1}, \qquad STS = T'^{-1}.$$

We will also consider various special elements of the group ring $\mathbb{Z}[\Gamma]$. First, we have the projections

$$\pi^+ \;=\; \frac{1}{2}(\varepsilon+1), \quad \pi^{\mathrm{od}} = \frac{1}{2}(1-\delta), \quad \pi^{+,\mathrm{od}} \;=\; \pi^+\pi^{\mathrm{od}} = \pi^{\mathrm{od}}\pi^+, \text{ etc.}$$

onto V^+, V^{od}, $V^{+,\mathrm{od}}$ etc. Next, we have the element

$$\Delta \;=\; (T-1)(\varepsilon+1)$$

which by (26) essentially characterizes $\mathcal{D}_k(\mathbb{Q})$: the codimension 1 subspace $\mathcal{D}_k^0(\mathbb{Q})$ of realizations with $Z_k = 0$ is identified precisely with $\mathrm{Ker}(\Delta)$, and the full space

$\mathcal{D}_k(\mathbb{Q})$ corresponds to the space of $\mathfrak{Z} \in V_k$ such that $\mathfrak{Z}|\Delta \in \mathbb{Q} \cdot \frac{X^{k-1}-Y^{k-1}}{X-Y}$. We can now interpret part of Proposition 2 $\big($the equivalence of (i) and (ii) when $Z_k = 0\big)$ as the dual statement of this with respect to the non-degenerate scalar product (34): a relation (29) with $Z_k = 0$ is reformulated equivalently as $\langle A|S, \mathfrak{Z}\rangle = 0$ (compare equations (29) and (30); the extra "S" comes from the interchange of r and s and the sign $(-1)^r$ in (34)). So this holds for all $\mathfrak{Z} \in \mathrm{Ker}(\Delta)$ if and only if $A|S \in \mathrm{Im}(\Delta^*) = V_k^+|(T^{-1}-1)$, i.e., if and only if $A \in V_k^+|(T^{-1}-1)S = V_k^+|(T'-1)$ (for the last step, use $V_k^+ = V_k^+|S$ and $ST^{-1}S = T'$), and this is just (ii) of Proposition 2. Finally, we have the element

$$\Lambda \;=\; 1 - \varepsilon U + U^2 \quad \in \quad \mathbb{Z}[\Gamma]\,. \tag{35}$$

It is related to the above elements by

$$\Lambda\,\Delta \;=\; -4\,\pi^{\mathrm{od}}\,\pi^+ \quad \in \quad \mathbb{Z}[\Gamma]\,, \tag{36}$$

as we see by the calculation

$$\begin{aligned}\Lambda\,(T-1) \;&=\; (1-\varepsilon U+U^2)(US-1)\\ &=\; \big[1-U(\varepsilon-1)\big]\,S \,+\, U^2(\varepsilon-1) \,-\, 1\\ &=\; \delta - 1 \,+\, (\delta - US + U^2)(\varepsilon-1)\,,\end{aligned}$$

followed by multiplying both sides on the right by $\varepsilon + 1$. It follows from (36) that for any polynomial $A \in V_k$ which is even *or* antisymmetric *or* S-invariant, the coefficients of $A|\Lambda$ give a realization of $\mathcal{D}_k$ with $Z_k = 0$ by taking $\mathfrak{Z}_k = A|\Lambda$ and $\mathfrak{P}_k = 2\pi^+(A)$. This is equivalent to Proposition 2 and the remark in its proof.

As an example for how to work with the structures just introduced, we prove Eq. (5) from the Introduction. To do this, we define

$$B_{m,n}(X,Y) \;=\; \binom{k-2}{m-1} Y^{k-2} B_{n-1}(X/Y) \qquad (m+n=k), \tag{37}$$

where $B_\nu(x) = \sum_{\mu=0}^{\nu} \binom{\nu}{\mu} B_\mu\, x^{\nu-\mu}$ denotes the νth Bernoulli polynomial. The numbers $\lambda_{m,n}(r,s)$ defined in (6) are the coefficients of the generating series

$$\sum_{r+s=k} \binom{k-2}{r-1}\, \lambda_{m,n}(r,s)\, X^{r-1}Y^{s-1} \;=\; B_{m,n}(X, X+Y)\,. \tag{38}$$

The symmetry $\lambda_{m,n}(r,s) = (-1)^{m-1}\lambda_{m,n}(s,r)$ mentioned for m odd in the Introduction follows from this formula together with the standard property $B_\nu(1-x) = (-1)^\nu B_\nu(x)$ of Bernoulli polynomials (cf., e.g., [2]). Set $a_{r,s} = (-1)^{s-1}\big[\binom{s-1}{m-1}B_{s-m} - \lambda_{m,n}(r,s)\big]$. Then we see that the polynomial (30) has the form $H(X, X+Y) - H(X,Y)$ with $H(X,Y) = B_{m,n}(X, X-Y)$. The symmetry property just mentioned implies that $H(X,Y) = H(Y,X)$, so we can apply Proposition 2 to get (29) with $\lambda = \frac{1}{2}\sum(-1)^{s-1}\lambda_{m,n}(r,s)$. This is (5).

4. Representing even double zeta values in terms of odd ones

In this section, we prove Theorem 2. Since we already know that $\dim \mathcal{D}_k \geqslant k/2$ (cf. (23)), we have only to show that any $Z_{\mathrm{ev,ev}}$ is a linear combination of $Z_{\mathrm{od,od}}$'s. This means that any collection of numbers $\{a_{r,s} \mid r+s=k;\ r,s \text{ even}\}$ can be completed to a collection $\{a_{r,s}\ (r+s=k),\ \lambda\}$ satisfying (29) in $\mathcal{D}_k$. By Proposition 2, this is equivalent to showing that any polynomial $F \in V_k^{\mathrm{od}}$ is the odd part of a polynomial of the form $H|(T-1)$ with $H \in V_k^+$ (since then $F|\varepsilon$ is the odd part of $H|(T'-1)$). Thus the result to be proved is:

Proposition 3. *The space V_k ($k>2$ even) has the decomposition*

$$V_k^{\mathrm{ev}} + V_k^{+}|(T-1) = V_k\,.$$

Proof. Here it is more convenient to use the 1-variable model. Let $V_k^{T\delta}\subset V_k$ be the fixed point set of the involution $g(x)\mapsto g(1-x)$. Then we have the following commutative diagram with the top row exact:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{Q}\cdot 1 & \longrightarrow & V_k^{T\delta} & \xrightarrow{T-1} & V_k^{\mathrm{od}} & \longrightarrow & 0\\
 & & & & \Big\downarrow{\scriptstyle 1+\varepsilon-\varepsilon U} & & \Big\uparrow{\scriptstyle \pi^{\mathrm{od}}} & & \\
 & & & & V_k^{+} & \xrightarrow{T-1} & V_k^{+}|(T-1) & &
\end{array}$$

To see the exactness, we first observe that if $g\in V_k^{T\delta}$ then the polynomial $f=g|(T-1)$ is odd because, from $T\delta T=\delta$ and $g|T\delta=g$, we deduce $g|T=g|\delta$. Conversely, an odd polynomial $f(x)\in V_k$ has degree $\leqslant k-3$ (since k is even) and hence can be written as $g(x+1)-g(x)$ for some $g\in V_k$. But then $g|(T\delta-1)(1-T)=g|(T-1)(1+\delta)=f|(1+\delta)=0$, so $g|(T\delta-1)$ is a constant and hence zero since it vanishes at $x=1/2$. (One can also argue that the map $V_k^{T\delta}/\mathbb{Q}\cdot 1\xrightarrow{T-1}V_k^{\mathrm{od}}$ which is obviously injective, must be an isomorphism because both sides have dimension $k/2-1$.) It is clear that the kernel of $V_k^{T\delta}\xrightarrow{T-1}V_k^{\mathrm{od}}$ is $\mathbb{Q}\cdot 1$. Next, we have to show that $h=g|(1+\varepsilon-\varepsilon U)$ is symmetric for $g\in V_k^{T\delta}$. This follows from $\varepsilon U\varepsilon=U^2=TSU$ and thus $g|\varepsilon U\varepsilon=g|TSU=g|\delta SU=g|\varepsilon U$. The commutativity of the square now follows from the calculation

$$(\varepsilon-\varepsilon U)(T-1)=(\varepsilon T-1)\varepsilon(1+\delta)+(1-T\delta)\varepsilon(1-U)T\,,$$

which implies that $(h-g)|(T-1)=g|(\varepsilon-\varepsilon U)(T-1)=g|(\varepsilon T-1)\varepsilon(1+\delta)$ which vanishes under $|(1-\delta)$.

It follows from the diagram that the map $\pi^{\mathrm{od}}: V_k^{+}|(T-1)\to V_k^{\mathrm{od}}$ is surjective which is equivalent to the statement of the theorem. $\square$

The proposition and its proof give us an explicit way to realize the asserted decomposition by starting with any basis of $V_k^{T\delta}$. To obtain a relation (29) with prescribed values $a_{r,s}=f_{r,s}$ for r and s even, we write the generating function $f|\varepsilon\in V_k^{\mathrm{od}}$ as $g|(T-1)$ with $g\in V_k^{T\delta}$, then $A=g|(1+\varepsilon-\varepsilon U)(T-1)\varepsilon=g|(1+\varepsilon-\varepsilon U)(T'-1)$ has odd part f and belongs to $V_k^{+}|(T'-1)$, so that Proposition 2 applies. To obtain explicit relations of this decomposition, we can choose any basis of the space of functions symmetric about $x=1/2$. In particular, from the three bases $g(x)=(2x-1)^{k-2-2\nu}$, $(x^2-x)^{\nu}$ and $B_{2\nu}(x)$, where $0\leqslant\nu\leqslant(k-2)/2$, we get three explicit collections of relations. For the first one, suitably normalized, we find that the coefficients of the associated relation (29) are

$$a_{r,s}=\begin{cases} 2^{s-1}\dbinom{r-1}{2\nu} & (r,s\text{ even})\\[2mm] -2^{r-1}\dbinom{s-1}{2\nu}+\displaystyle\sum_{\alpha+\beta=2\nu}(-1)^{\alpha}\binom{r-1}{\alpha}\binom{s-1}{\beta} & (r,s\text{ odd}),\end{cases}$$

$$\lambda=\frac{k-1}{2}\binom{k-2}{2\nu}\Big[\int_0^1(2-3t)^{k-2\nu-2}t^{2\nu}dt-\frac{1}{2(2\nu+1)}\Big]\,,$$

Table 1: Relations among double zeta values of weight 12 coming from Proposition 4. Each row of the table gives a relation among the (formal) double zeta values displayed in the top line. The $Z_{1,11}$ column has been omitted, since all of its entries would be zero.

$g(x)$	$Z_{2,10}$	$Z_{4,8}$	$Z_{6,6}$	$Z_{8,4}$	$Z_{10,2}$	Z_{12}	$Z_{3,9}$	$Z_{5,7}$	$Z_{7,5}$	$Z_{9,3}$	$Z_{11,1}$
$\frac{1}{2}\binom{10}{10}(2x-1)^{10}$	512	128	32	8	2	$-\frac{1355}{4}$	-3	-15	-63	-255	-1023
$\frac{1}{2}\binom{10}{8}(2x-1)^{8}$	0	384	320	168	72	$-\frac{21}{4}$	-99	-243	-387	-243	45
$\frac{1}{2}\binom{10}{6}(2x-1)^{6}$	0	0	160	280	252	$\frac{129}{2}$	-294	-238	-62	-14	210
$\frac{1}{2}\binom{10}{4}(2x-1)^{4}$	0	0	0	56	168	$\frac{51}{2}$	-126	-14	2	-14	210
$\frac{1}{2}\binom{10}{2}(2x-1)^{2}$	0	0	0	0	18	$-\frac{7}{4}$	9	-3	-3	13	45
$\binom{10}{10}(x^2-x)^5$	1	$\frac{1}{6}$	$\frac{1}{126}$	0	0	$-\frac{331}{504}$	0	0	$-\frac{1}{21}$	$-\frac{4}{9}$	-2
$\binom{10}{8}(x^2-x)^4$	0	3	$\frac{10}{7}$	0	0	$-\frac{1}{4}$	0	0	$-\frac{15}{7}$	-2	0
$\binom{10}{6}(x^2-x)^3$	0	0	5	$\frac{7}{2}$	0	$\frac{1}{2}$	0	-14	-10	0	0
$\binom{10}{4}(x^2-x)^2$	0	0	0	7	0	$-\frac{4}{3}$	0	28	30	$\frac{28}{3}$	0
$\binom{10}{2}(x^2-x)$	0	0	0	0	9	$\frac{19}{4}$	-18	-24	-24	-16	0
$\binom{10}{10}B_{10}(x)$	1	0	0	0	0	$-\frac{767}{1155}$	$\frac{5}{33}$	$-\frac{41}{165}$	$-\frac{31}{231}$	$-\frac{41}{165}$	$-\frac{61}{33}$
$\binom{10}{8}B_{8}(x)$	0	3	0	0	0	$\frac{1}{6}$	-3	5	2	$-\frac{11}{3}$	-3
$\binom{10}{6}B_{6}(x)$	0	0	5	0	0	$-\frac{4}{3}$	10	-18	-15	$\frac{16}{3}$	10
$\binom{10}{4}B_{4}(x)$	0	0	0	7	0	$\frac{13}{6}$	-14	14	16	$-\frac{14}{3}$	-14
$\binom{10}{2}B_{2}(x)$	0	0	0	0	9	1	-3	-9	-9	-1	15
$\frac{1}{2}$	0	0	0	0	0	$-\frac{1}{4}$	1	1	1	1	1

and for the second basis we find

$$\frac{1}{2}\binom{k-2}{r-1}\,a_{r,s} = \begin{cases} \binom{\nu}{s-\nu-1} & (r, s \text{ even}), \\ (-1)^{\nu}\binom{k-2-2\nu}{r-\nu-1} - \binom{\nu}{r-\nu-1} & (r, s \text{ odd}) \end{cases}$$

(we omit the value of λ in this case). In both cases, the coefficients for r, s even form a triangular matrix. The third family $g(x) = B_{2\nu}(x)$ yields Eq. (7), as the reader can check as an exercise by imitating the proof of Eq. (5) which was given at the end of Section 3. The following table gives these three collections of relations for the case $k = 12$.

5. Double zeta values and period polynomials

In this section we describe various connections between period polynomials and the (formal) double zeta space, and prove Theorems 3 and 4. Since both of them involve period polynomials, we begin by reviewing these. The definition of period polynomials was already given briefly in the Introduction. The motivation comes from the connection with modular forms, which we will review in the next section. Here we discuss only the algebraic properties.

The space W_k is defined as

$$W_k \;=\; \mathrm{Ker}(1+S) \cap \mathrm{Ker}(1+U+U^2) \quad \subset V_k\,, \tag{39}$$

i.e. as the intersection of the (-1)-eigenspace of the involution S and the sum of the $\left(\frac{-1\pm\sqrt{-3}}{2}\right)$-eigenspaces of the element U of order 3. Since $\varepsilon S\varepsilon = S$ and

$\varepsilon U\varepsilon = U^2$, the involution ε acts on W_k and splits it as the direct sum of subspaces $W_k^\pm = W_k \cap V_k^\pm$ of symmetric and antisymmetric polynomials. Since elements in W_k are also (-1)-eigenfunctions of S and since $S\varepsilon = \varepsilon S = \delta$, we also have $W_k^+ = W_k^{\mathrm{od}} \subset V_k^{+,\mathrm{od}}$ and $W_k^- = W_k^{\mathrm{ev}} \subset V_k^{-,\mathrm{ev}}$. Another important property of period polynomials is given by the following lemma.

Lemma. *Let $k > 2$ be even. Then*

$$W_k \;=\; \mathrm{Ker}\bigl(1 - T - T',\, V_k\bigr)$$

and

$$W_k^\pm \;=\; \mathrm{Ker}\bigl(1 - T \mp T\varepsilon,\, V_k\bigr)\,.$$

Proof. It is equivalent for $f \in V_k$ to be in $\mathrm{Ker}(1-T-T')$ or to satisfy $f|(1+S) = f|(1+U+U^2)$, since $(1-T-T')S = (1+S)-(1+U+U^2)$. But a polynomial which is fixed by both S and U is fixed by the full modular group and thus vanishes. The second assertion of the lemma follows from the first, because if f is annihilated by $1-T-T' = 1-T-\varepsilon T\varepsilon$ and $f|\varepsilon = \pm f$ then f is also annihilated by $1-T\mp T\varepsilon$, and conversely if f is annihilated by $1-T\mp T\varepsilon$ then $f = f|T(1\pm\varepsilon) \in V_k^\pm$ and hence $f|(1-T-T') = f|(1-T-\varepsilon T\varepsilon) = 0$. $\square$

Remark. The operator $\mathcal{L} = 1-T-T'$ plays a key role in the discovery by J. Lewis that there are holomorphic functions annihilated by this operator which have the same relation to the so-called Maass wave forms as period polynomials have to holomorphic modular forms ([10], [11]). We call the equation $f|\mathcal{L} = 0$ the *Lewis equation.*

As in the Introduction, we denote by $\mathcal{P}_k$ the subspace of $\mathcal{D}_k$ spanned by the $P_{r,s}$ (and Z_k, but it can be omitted by virtue of Theorem 1), and by $\mathcal{P}_k^{\mathrm{ev}}$ the subspace spanned by the $P_{\mathrm{ev},\mathrm{ev}}$. Note that $\mathcal{P}_k^{\mathrm{ev}}$ corresponds to generating functions in V_k^{od} because of the shift by 1 in the exponents of X and Y.

Theorem 3. *The spaces $\mathcal{P}_k^{\mathrm{ev}}$ and W_k^- are canonically isomorphic to each other. More precisely, to each $p \in W_k^-$ we associate the coefficients $p_{r,s}$ and $q_{r,s}$ ($r+s=k$) which are defined by $p(X,Y) = \sum\binom{k-2}{r-1}p_{r,s}X^{r-1}Y^{s-1}$ and $p(X+Y,Y) = \sum\binom{k-2}{r-1}q_{r,s}X^{r-1}Y^{s-1}$. Then $q_{r,s} - q_{s,r} = p_{r,s}$ (in particular $q_{r,s} = q_{s,r}$ for r,s even) and*

$$\sum_{\substack{r+s=k\\ r,\,s\ \text{even}}} q_{r,s}Z_{r,s} \;\equiv\; 3\sum_{\substack{r+s=k\\ r,\,s\ \text{odd}}} q_{r,s}Z_{r,s} \quad (\mathrm{mod}\ Z_k)\,, \tag{40}$$

and conversely, an element $\sum_{r,\,s\ \text{odd}} c_{r,s}Z_{r,s} \in \mathcal{D}_k$ belongs to $\mathcal{P}_k^{\mathrm{ev}}$ if and only if $c_{r,s} = q_{r,s}$ arising in this way.

Remarks. 1. The equivalence of the first and last statements of the theorem follows from Theorem 2 : since the $Z_{\mathrm{od},\mathrm{od}}$ form a basis of $\mathcal{D}_k$, it is equivalent to speak of elements of $\mathcal{P}_k^{\mathrm{ev}}$ or of relations of the form $\sum(*)P_{\mathrm{ev},\mathrm{ev}} = \sum(*)Z_{\mathrm{od},\mathrm{od}}$.

2. Since the double zeta realizations $\zeta(r)\zeta(s)$ and $\zeta(k)$ of $P_{r,s}$ (r, s even) and Z_k are rational multiples of π^k, and since π^k is a $\mathbb{Q}$-linear combination of $Z_{\mathrm{od},\mathrm{od}}$'s by Theorem 1, Theorem 3 as stated here contains the "rough statement" given in the Introduction. (The number of relations drops from $\dim W_k^- = \dim M_k$ to $\dim S_k = \dim M_k - 1$ because one relation gets used up to eliminate $\zeta(k)$.)

Example 1. For every even $k > 2$, the space W_k^- contains the polynomial $p(x) = x^{k-2} - 1$ (in the inhomogeneous notation). Here $p(x+1) = \sum_{r\neq 1}\binom{k-2}{r-1}x^{r-1}$, i.e., $q_{1,k-1} = 0$ and all other $q_{r,s}$ are equal to 1, and Theorem 3 reduces to a weaker version of Theorem 1.

Table 2: The coefficients $p_{r,s}$ and $q_{r,s}$ for $p(x) = x^2(x^2-1)^3 \in W_{12}^-$.

r	1	2	3	4	5	6	7	8	9	10	11
s	11	10	9	8	7	6	5	4	3	2	1
$1260\, p_{r,s}$	0	0	-28	0	18	0	-18	0	28	0	0
$1260\, q_{r,s}$	0	$\underline{0}$	0	$\underline{84}$	168	$\underline{190}$	150	$\underline{84}$	28	$\underline{0}$	0

Example 2. The space W_{12}^- is 2-dimensional, spanned by the two polynomials $p(x) = x^{10} - 1$ and $x^2(x^2-1)^3$. For the latter, we have $p(x+1) = x^8 + 8x^7 + 25x^6 + 38x^5 + 28x^4 + 8x^3$, so the $p_{r,s}$ and $q_{r,s}$ of the theorem are given (after multiplication by 1260) by the table
The $q_{r,s}$ with r and s even (underlined) are symmetric and the relation (40), divided by 3, becomes

$$28Z_{8,4} + \frac{190}{3}Z_{6,6} + 28Z_{4,8} \equiv 28Z_{9,3} + 150Z_{7,5} + 168Z_{5,7} \qquad (\mathrm{mod}\ Z_{12}),$$

in agreement with Eq. (9) of the Introduction. The example for $k = 16$ given there arises in the same way from the polynomial $p(x) = x^2(x^2-1)^3(2x^4 - x^2 + 2) \in W_{16}^-$.

Proof. The function $q = p|T$ satisfies $q|(1-\varepsilon) = p|(T - T\varepsilon) = p|(T + \varepsilon T \varepsilon) = p$ because p is antisymmetric and satisfies the Lewis equation. This shows that $q_{r,s} - q_{s,r} = p_{r,s}$ and also means that if we decompose q in the obvious way as $q = q^{\mathrm{ev},+} + q^{\mathrm{ev},-} + q^{\mathrm{od},+} + q^{\mathrm{od},-}$, then $q^{\mathrm{ev},-} = \frac{1}{2}p$ and $q^{\mathrm{od},-} = 0$. Write $[a,b,c]$ to denote $aq^{\mathrm{ev},+} + bq^{\mathrm{ev},-} + cq^{\mathrm{od},+}$. Then

$$[0,\,2,\,0]\,\big|\,T' \;=\; p|T' \;=\; p|\varepsilon T \varepsilon \;=\; -p|T\varepsilon \;=\; -q|\varepsilon \;=\; [-1,\,1,\,-1]\,,$$
$$[1,\,-1,\,-1]\,\big|\,T' \;=\; q|ST' \;=\; p|TST' \;=\; p|S = -p \;=\; [0,\,-2,\,0]\,,$$

and hence

$$[2,\,0,\,-2]\,\big|\,(T'-1) \;=\; [-1,\,-3,\,-1] \,-\, [2,\,0,\,-2] \;=\; [-3,\,-3,\,1]\,.$$

This says that $q^{\mathrm{od}} - 3q^{\mathrm{ev}}$ is the image under $T' - 1$ of the symmetric polynomial $2\big(q^{\mathrm{ev},+} - q^{\mathrm{od},+}\big)$, so Proposition 2 implies Eq. (40). (The omitted coefficient of Z_k in (40) is easily determined, by computing the integral in (31), as $\sum(-1)^{r-1}q_{r,s}$.) This gives a map $W_k^- \longrightarrow \mathcal{P}_k^{\mathrm{ev}}$ which is obviously injective since the antisymmetrization of the coefficients on the right-hand side of (40) are the coefficients of p itself. We omit the proof of surjectivity since it will follow from Theorem 4 below that $\dim W_k^- = \dim \mathcal{P}_k^{\mathrm{ev}}$, so that injectivity suffices. $\square$

Theorem 3 tells us that, given *any* relation of the form (29) in $\mathcal{D}_k$ with $a_{r,s} = a_{s,r}$ for r and s even, there exists a unique element $p \in W_k^-$ such that the $a_{r,s}$ with r, s odd are equal to the numbers $q_{r,s}$ in the theorem. On the other hand, given an element of $\mathcal{P}_k^{\mathrm{ev}}$, its representation as a linear relation of the generators $P_{\mathrm{ev,ev}}$ is not unique, because these generators are not linearly independent. The next proposition, which is a first form of Theorem 4 of the Introduction, describes the relations among them, i.e., all relations of the form (29) with $a_{r,s} = a_{s,r}$ for *all* r and s. It turns out that in any such relation the odd-index $a_{r,s}$ all vanish (in accordance with the widely believed and numerically verified statement that there are no relations over $\mathbb{Q}$ among products of values of the Riemann zeta function at odd arguments), while the even-index $a_{r,s}$ are related to the space W_k^+.

Proposition 4. *Let $a_{r,s}$ and μ be numbers with $a_{s,r} = a_{r,s}$. Then the following three statements are equivalent:*

(i) The relation

$$\sum_{r+s=k} a_{r,s} P_{r,s} = \mu Z_k \tag{41}$$

holds in $\mathcal{D}_k$.

(ii) The generating function

$$A(X,Y) = \sum_{r+s=k} \binom{k-2}{r-1} a_{r,s} X^{r-1} Y^{s-1} \quad \in V_k \tag{42}$$

can be written as $A = H|(1-S)$ *for some* $H \in V_k^U \cap V_k^+$, *and*

$$\mu = \langle H, \frac{X^{k-1} - Y^{k-1}}{X - Y} \rangle \,. \tag{43}$$

(iii) The generating function

$$A^*(m,n) = \sum_{r+s=k} \frac{a_{r,s}}{m^r n^s} \quad \in V_k^* \tag{44}$$

can be written as $f(m,n) - f(m+n,n) - f(m,m+n)$ *for some symmetric* $f \in V_k^*$, *and*

$$\mu = f(1,1)\,. \tag{45}$$

If these statements hold, then $a_{r,s} = 0$ *for odd* r *and* s.

Proof. Except for the assertions about the value of the constant μ, which we will leave to the reader, each part of this proposition is equivalent to the corresponding part of Proposition 2 of Section 2 with the extra condition that $a_{r,s} = a_{s,r}$. For (i) this is obvious; for (iii) it follows because $f(m,n) - A^*(m,n) = f(m+n,m) + f(m+n,n)$ in Proposition 2 is always symmetric, so that A is symmetric if and only if f is; and for (ii) it follows because if $H \in V_k^+$, then the element $H|(T'-1)$ is symmetric if and only if $H = H|U$ (because $H|T'(\varepsilon - 1) = H|(1-U)T$), in which case $H|(T'-1) = H|(T-1) = H|(S-1)$. The last assertion of the proposition is then clear since $A \in V_k^+|(1-S) \subset V_k^{\mathrm{od}}$. $\square$

Proposition 5, without the statements concerning μ, says that the following are equivalent for symmetric A:

(i′) $\sum a_{r,s} P_{r,s} \equiv 0 \pmod{Z_k}$;
(ii′) $A \in V_k^{U,+}|(1-S)$;
(iii′) $A^* \in V_k^*|(1-T-T')$ (in the obvious notation).

Statement (ii′) in turn is equivalent to

(iv′) $A \in V_k^{\mathrm{od}}$ and $A \perp W_k^+$ with respect to the scalar product (34),

because for $v \in V$ we have:

$$\begin{aligned} & v \text{ is orthogonal to } (V_k^{U,+})|(1-S) = V_k|(1+U+U^2)(1+\varepsilon)(1-S) \\ \Leftrightarrow \quad & v|(1-S)(1+\varepsilon)(1+U+U^2) = 0 \\ \Leftrightarrow \quad & v|(1-S)(1+\varepsilon) \in \mathrm{Ker}(1+U+U^2) \cap \mathrm{Ker}(1+S) \cap \mathrm{Ker}(1-\varepsilon) = W_k^+ \\ \Leftrightarrow \quad & v \in W_k^+ + V_k^- + V_k^{\mathrm{ev}}\,. \end{aligned}$$

On the other hand, (i′) is equivalent to the condition that $\sum a_{r,s} P_{r,s} = 0$ for any realization $\{Z_{r,s},\, P_{r,s},\, Z_k\}$ of $\mathcal{D}_k$ with $Z_k = 0$, while (iv′) says that all $a_{\mathrm{od},\mathrm{od}}$ are zero and $\sum a_{r,s} p_{r,s} = 0$ for any $p = \sum p_{r,s} X^{r-1} Y^{s-1} \in W_k^+$. The equivalence of (i′) and (iv′) therefore says that *any* symmetric collection of numbers $\{P_{r,s}$ $(r,\, s \text{ odd})\}$ can be extended to a realization of $\mathcal{D}_k$ with $Z_k = 0$, while a symmetric collection of numbers $\{P_{r,s}$ $(r,\, s \text{ even})\}$ can be extended to a realization of $\mathcal{D}_k$ with $Z_k = 0$ if and only if the corresponding generating function belongs to W_k^+. This is precisely the statement of Theorem 4 as given in the Introduction in the case $Z_k = 0$. (For the full statement we need the extended period polynomial space $\widehat{W}_k^+$, which will be discussed in the next section in the context of modular forms.) Before doing that, we give a slight improvement of the result just stated. This result says that, if P is any polynomial belonging to $V_k^{+,\mathrm{ev}}$ or to $W_k^+ \subset V_k^{+,\mathrm{od}}$, then there exists a $Z \in V_k$ with

$$Z\,|\,(1+\varepsilon) \;=\; P\,, \qquad Z\,|\,T(1+\varepsilon) \;=\; P\,. \tag{46}$$

The following proposition makes this explicit:

Proposition 5. *(i) Let $P \in V_k^{+,\mathrm{ev}}$. Then $Z = \frac{1}{2}\,P|\Lambda$ with Λ as in* (35) *gives a solution of Eqs.* (46).

(ii) Let $P \in W_k^+$. Then $Z = \frac{1}{3}\,P|(T^{-1}+1)$ gives a solution of Eqs. (46).

Proof. (i) The identities $\Lambda(1+\varepsilon) \;=\; 1+\varepsilon$ and $\Lambda T \;=\; S + T(1-\varepsilon)$ together with $P|\varepsilon = P|\delta = P$ immediately imply (46).

(ii) By Lemma 5, P satisfies the Lewis equation $P|(1-T-T') = 0$. Hence, using $P|\varepsilon = P$ and $P|\delta = -P$, we find

$$P|(1+T^{-1})(1+\varepsilon) = 2P + P|(T^{-1} - \delta T^{-1}\varepsilon) = 3P + P|(1-T-T')T^{-1} = 3P$$

and $P|(T^{-1}+1)T(1+\varepsilon) = 2P \,+\, P|(T+T') = 3P\,.$ $\square$

6. Double zeta values and modular forms

In this section we reinterpret the results of Section 5 from the modular point of view. To do this, we begin by reviewing the theory of period polynomials of modular forms on $\mathrm{PSL}_2(\mathbb{Z})$, including various supplementary results which are less well-known and which are needed here.

The *period polynomial associated to a cusp form* $f \in S_k$ can be defined by

$$P_f(X,Y) \;=\; \int_0^\infty (X - Y\tau)^{k-2}\, f(\tau)\, d\tau\,. \tag{47}$$

The identity $(aX+bY) - (cX+dY)\tau = (a-c\tau)(X - \gamma^{-1}(\tau)Y)$ together with the modularity of f shows that

$$P_f\,\big|\gamma(X,Y) \;=\; \int_{\gamma^{-1}(0)}^{\gamma^{-1}(\infty)} (X - Y\tau)^{k-2}\, f(\tau)\, d\tau$$

for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1\,.$ In particular,

$$P_f|(1-T-T') = \bigg(\int_0^\infty - \int_{-1}^\infty - \int_0^{-1}\bigg)(X - Y\tau)^{k-2}\, f(\tau)\, d\tau \;=\; 0\,,$$

so $P_f \in W_k$ by Lemma 5. (One can also verify that $P_f|(1+S) = P_f|(1+U+U^2) = 0$ directly by a similar calculation.) This gives the basic connection between cusp forms and period polynomials. The complete Eichler-Shimura-Manin theory, of which summaries can be found in many places (e.g. [9]), tells us that the maps assigning to f the symmetric (odd) and antisymmetric (even) parts P_f^+ and P_f^- of P_f give isomorphisms from the space S_k of cusp forms of weight k on Γ_1 onto W_k^+ and a codimension 1 subspace of W_k^-. (The latter was determined in [8].) The theory is also "defined over $\mathbb{Q}$" in the sense that the even and odd period polynomials of a normalized Hecke eigenform f are proportional to polynomials with coefficients in the number field generated by the Fourier coefficients of f, and transform properly under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. For instance, the even and odd period polynomials of the modular form $\Delta \in S_{12}$, in the inhomogeneous version, are multiples of $\frac{36}{691}(x^{10}-1) - x^2(x^2-1)^3$ and $x(x^2-1)^2(x^2-4)(4x^2-1)$, respectively.

For $f \in M_k$ with $f(\infty) = a_0 \neq 0$ the integral (47) diverges, but the modified definition [17]

$$\widehat{P}_f(X,Y) \;=\; \int_{\tau_0}^{\infty} (X-Y\tau)^{k-2}\bigl(f(\tau)-a_0\bigr)d\tau \;+\; \int_0^{\tau_0} (X-Y\tau)^{k-2}\Bigl(f(\tau)-\frac{a_0}{\tau^k}\Bigr)d\tau$$
$$+\;\frac{a_0}{k-1}\left(\frac{1}{Y}\;-\;\frac{\tau_0^{1-k}}{X}\right)(X-Y\tau_0)^{k-1} \qquad (\text{any } \tau_0 \in \mathfrak{H})$$

makes sense, since the integrals converge and the derivative of the right-hand side with respect to τ_0 vanishes. We call this function the *extended period polynomial* of f. For instance, the extended period polynomial of the Eisenstein series G_k is

$$\frac{2\pi i\,\zeta(k-1)}{2k-2}(X^{k-2}-Y^{k-2}) \;-\; \frac{(2\pi i)^k}{2k-2}\sum_{r=0}^{k}\frac{B_r}{r!}\,\frac{B_{k-r}}{(k-r)!}\,X^{r-1}Y^{k-r-1}\,,$$

where B_r denotes the rth Bernoulli number.

The function $\widehat{P}_f(X,Y)$ no longer lies in V_k, but in the larger space $\widehat{V}_k = \bigoplus_{\substack{r+s=k\\ r,s\geqslant 0}} \mathbb{C}\cdot X^{r-1}Y^{s-1}$. On the other hand, the same calculation as before shows that it is again annihilated by $1+S$, $1+U+U^2$ and $1-T-T'$. (The group Γ_1 does not act on $\widehat{V}_k$, but it acts on the larger space $\mathbb{C}(X,Y)$ of rational functions in two variables, so this statement makes sense.) In other words, $\widehat{P}_f$ belongs to the space

$$\widehat{W}_k \;:=\; \mathrm{Ker}(1+S,\,\widehat{V}_k)\cap\mathrm{Ker}(1+U+U^2,\,\widehat{V}_k) \;=\; \mathrm{Ker}(1-T-T',\,\widehat{V}_k)\,,$$

which we call the space of *extended period polynomials*. This space fits into a short exact sequence $0 \longrightarrow W_k \longrightarrow \widehat{W}_k \xrightarrow{\lambda} \mathbb{C} \longrightarrow 0$, where λ associates to $\widehat{P} \in \widehat{W}_k$ the coefficient of X^{k-1}/Y in $\widehat{P}$. By writing $\widehat{P} = P + \lambda(\widehat{P})(X^{k-1}/Y + Y^{k-1}/X)$ we can identify $\widehat{W}_k$ with the space

$$\widehat{\widehat{W}}_k \;=\; \bigl\{(P,\lambda) \in V_k\times\mathbb{C} \mid P|(1+S) = P|(1+U+U^2)+\lambda\Phi_k = 0\bigr\}\,,$$

where $\Phi_k \in V_k$ is the polynomial

$$\Phi_k(X,Y) \;=\; \Bigl(\frac{X^{k-1}}{Y}+\frac{Y^{k-1}}{X}\Bigr)\Big|\,(1+U+U^2)$$
$$=\;\Bigl(\frac{X^{k-1}-(X-Y)^{k-1}}{Y}+\frac{Y^{k-1}+(X-Y)^{k-1}}{X}+\frac{X^{k-1}-Y^{k-1}}{X-Y}\Bigr)\,.$$

Corresponding to the canonical splitting $M_k = S_k \oplus \mathbb{C}\cdot G_k$ of modular forms into cusp forms and Eisenstein series, we have the splittings $\widehat{W}_k = W_k \oplus \mathbb{C}\cdot\widehat{\mathcal{E}}_k$, $\widehat{\widehat{W}}_k = W_k \oplus \mathbb{C}\cdot(\mathcal{E}_k, B_k)$, where

$$\widehat{\mathcal{E}}_k(X,Y) = \sum_{\substack{r+s=k\\ r,\,s\geqslant 0}}\binom{k}{r}B_rB_sX^{r-1}Y^{s-1} = \mathcal{E}_k(X,Y) + B_k\Bigl(\frac{Y^{k-1}}{X}+\frac{X^{k-1}}{Y}\Bigr)\,. \tag{48}$$

The space $\widehat{W}_k$ also splits into symmetric and antisymmetric parts $\widehat{W}_k^{\pm}$. Since $\widehat{\mathcal{E}}_k$ is symmetric, we have $\widehat{W}_k^{+} = W_k^{+}\oplus\mathbb{C}\cdot\widehat{\mathcal{E}}_k$ and $\widehat{W}_k^{-} = W_k^{-}$.

If f and g are two cusp forms, then the Petersson scalar product (f,g) is proportional to the pairing $\langle P_f^+|(T-T^{-1}),\,P_g^-\rangle$ ([7], [8]). (It is essential that odd period polynomials are always paired with even ones, because if f is a normalized Hecke eigenform then the coefficients of $P_f^{\pm}$ belong to $\omega_{\pm}(f)\,\overline{\mathbb{Q}}[X,Y]$ with some constants $\omega_{\pm}(f)$ whose product is essentially (f,f). But the "straight" pairing $\langle P_f^+,\,P_g^-\rangle$ would vanish since P_f^+ and P_g^- have opposite symmetry properties and also opposite parity. The effect of $|(T-T^{-1})$ is to change odd polynomials to even ones and vice versa.) This pairing is Hecke invariant, but we do not explain this here since we have not discussed the action of Hecke operators on period polynomials. It extends in a natural way to a pairing $\widehat{W}_k^+ \times W_k^- \to \mathbb{C}$ in such a way that the codimension 1 subspace of W_k^- corresponding to period polynomials of cusp forms is precisely the space of polynomials whose pairing with $\widehat{\mathcal{E}}_k$ vanishes (cf. [8]). The result (Theorem 9 of [8]) is

$$\{P_f^- \mid f \in S_k\} = \Big\{ \sum_{r+s=k} \tbinom{k-2}{r-1} a_{r,s} X^{r-1}Y^{s-1} \,\Big|\, \sum_{r+s=k} (-1)^{\frac{r-1}{2}} \kappa_{r,s} a_{r,s} = 0 \Big\},$$

where

$$\begin{aligned} \kappa_{r,s} \;=\; -\kappa_{s,r} \;=\; 2 \sum_{\substack{0<j\leqslant k \\ j \text{ even}}} \binom{j-1}{r-1}\binom{k}{j} B_j B_{k-j} \;-\; \binom{k}{r} B_r B_s \\ +\; \left[(-1)^{r-1} + \binom{k-1}{r-1} - \binom{k-1}{r} \right] B_k\,. \end{aligned} \tag{49}$$

Equation (49) says that $\kappa_{r,s}$ is essentially the coefficient of $X^{r-1}Y^{s-1}$ in $\mathcal{E}_k|T$ with $\mathcal{E}_k$ as in (48).

We can now easily extend Proposition 5 (and hence the preliminary version of Theorem 4 mentioned in the previous section) to a statement concerning $\widehat{W}_k^+$ rather than W_k^+. Given any symmetric extended period polynomial

$$\widehat{P} = \sum_{\substack{r+s=k \\ r,s\geqslant 0 \\ r,\,s \text{ even}}} p_{r,s} X^{r-1}Y^{s-1} \;\in\; \widehat{W}_k^+$$

(so that $\lambda(\widehat{P}) = p_{k,0} = p_{0,k}$), there is a realization of $\mathcal{P}_k$ with

$$P_{r,s} \mapsto \begin{cases} p_{r,s} & (r,\ s \text{ even}), \\ 0 & (r,\ s \text{ odd}), \end{cases} \qquad Z_k \mapsto -2\,p_{k,0}\,.$$

To see this, instead of going through the whole proof of Proposition 5, keeping careful track of the constant μ, it is sufficient (since W_k^+ has codimension 1 in $\widehat{W}_k^+$) to check this for one single extended "polynomial" $\widehat{P}$ which belongs to $\widehat{W}_k^+$ but not to W_k^+, and this is easy: the function $\widehat{\mathcal{E}}_k$ defined in (48) has coefficients $P_{r,s} = \binom{k}{r} B_r B_s = 4k!\,\beta_r\beta_s$, where $\beta_r = \zeta(r)/(2\pi i)^r = -B_r/2r!$ ($r \geqslant 0$ even), and on the other hand the original double zeta realization of $\mathcal{D}_k$ has $P_{r,s} = \zeta(r)\zeta(s) = (2\pi i)^k \beta_r\beta_s$ and $Z_k = \zeta(k) = (2\pi i)^k \beta_k$. This completes the discussion of extended period polynomials and the proof of Theorem 4.

We also mention the corresponding extension of Proposition 5:

Supplement to Proposition 5: *If $\widehat{P} = P + \lambda(X^{k-1}Y^{-1} + X^{-1}Y^{k-1})$, then we have a solution of* (25) *with $\mathfrak{P}_k = P$ and $Z_k = -2\lambda$ given by*

$$\begin{aligned} \mathfrak{Z}_k \;&=\; \frac{1}{3} P\big|(T^{-1}+1) \;+\; \frac{\lambda}{6}\frac{X^{k-1}}{Y}\,\bigg|\, U^2(1+\varepsilon)(5-3U+U\varepsilon) \\ &=\; \frac{1}{3} P\big|(T^{-1}+1) \;+\; \frac{\lambda}{6}\frac{X^{k-1}-Y^{k-1}}{X-Y}\,\bigg|\,(5-3U+U\varepsilon)\,. \end{aligned}$$

Proof. Just apply the calculation of the proof of part (ii) of Proposition 5 to $\widehat{P}$ with $\widehat{\mathfrak{Z}}_k = \frac{1}{3}\widehat{P}|(1+T^{-1})$ since $\widehat{P}$ satisfies the same relations as P. $\square$

If we apply this result to the special case $\widehat{P} = \frac{1}{4k!}\widehat{\mathcal{E}}_k$ then we find that, as well as the "Euler realization" of $\mathcal{D}_k$ in $\mathbb{R}$ with $Z_{r,s}^E = \zeta(r,s)$, $P_{r,s}^E = \zeta(r)\zeta(s)$ and $Z_k^E = \zeta(k)$, we also have a "Bernoulli realization" with $Z_k^B = \beta_k$, $P_{r,s}^B = \beta_r\beta_s$ (so that $P^B_{r,s} = 0$ if r and s are odd and $k > 2$), and with $Z^B_{r,s}$ equal to the "double Bernoulli number"

$$\begin{aligned} Z^B_{r,s} &= \frac{1}{3}\sum_{m+n=k}\binom{m-1}{r-1}\beta_m\beta_n + \frac{\beta_r\beta_s}{3} - \frac{\beta_k}{12}\left[5+3\binom{k-1}{r-1}-\binom{k-1}{r}\right] \\ &= \frac{1}{3}\sum_{\substack{m+n=k\\ m,n\geqslant 0}}\binom{m-1}{r-1}\beta_m\beta_n - \frac{\beta_k}{12}\left[1+\binom{k-1}{r-1}-\binom{k-1}{r}\right] + \frac{1}{3}(\beta_r\beta_s-\beta_k)\,. \end{aligned}$$

These are almost exactly the same as the numbers (49) occurring in the result on the image of cusp forms in W_k^- cited above! Also note that the number $Z^B_{r,s}$ is essentially the double zeta value $\zeta(1-r,1-s)/(r-1)!(s-1)!$ at negative arguments, which has been studied in [1].

The Bernoulli realization has the same even-index $p_{r,s}$ as the Euler realization (up to a factor $(2\pi i)^k$), but 0 instead of the presumably transcendental values $\zeta(\mathrm{od})\zeta(\mathrm{od})/(2\pi i)^k$; the fact that both the original $P_{r,s}$ and the new ones can be realized in $\mathcal{D}_k$ is an illustration of the fact that the numbers $P_{\mathrm{od},\mathrm{od}}$ in $\mathcal{P}_k$ are completely unconstrained (this corresponds to the vanishing of $a_{\mathrm{od},\mathrm{od}}$ in Proposition 5).

The results of Section 5 were formulated purely algebraically, but we can now easily relate them to the theory of modular forms. A result of Rankin ([13]; see also [8]) says that, for a normalized Hecke eigenform $f \in S_k$, one has

$$\sum_{\substack{r+s=k\\ r,\,s \text{ even}}} (f,\, G_rG_s)\, X^{r-1}Y^{s-1} = c_f\, P_f^+(X,Y)\,,$$

where c_f is essentially $P_f^-(1,0)$.

In particular, the cusp forms $G_rG_s - \frac{\beta_r\beta_s}{\beta_k}G_k$ satisfy the same linear relations as elements of W_k^+, so they form the coefficients of an element

$$\mathfrak{G}_k(\tau;X,Y) = \sum_{\substack{r,\,s\geqslant 1\\ r+s=k}}\Bigl(G_r(\tau)G_s(\tau) - \frac{\beta_r\beta_s}{\beta_k}G_k(\tau)\Bigr)X^{r-1}Y^{s-1} \;\in\; W_k^+\otimes S_k\,.$$

But then setting $\widehat{\mathfrak{G}}_k = \mathfrak{G}_k - \frac{1}{2\,B_k}G_k(\tau)\widehat{\mathcal{E}}_k$ gives the much simpler statement that the polynomial

$$\widehat{\mathfrak{G}}_k(\tau;X,Y) = \sum_{\substack{r,\,s\geqslant 0\\ r+s=k}} G_r(\tau)G_s(\tau)X^{r-1}Y^{s-1}$$

belongs to $\widehat{W}_k^+\otimes M_k$. Theorem 5 follows immediately.

7. Double Eisenstein series

At the end of the last section, we applied Rankin's result relating products of Eisenstein series to period polynomials of cusp forms to show that there is a realization of the double shuffle space $\mathcal{D}_k$ sending $P_{r,s}$ to $G_r(\tau)G_s(\tau)$ and Z_k to $G_k(\tau)$ for all r, s and k. This proof, however, is very indirect, and in view of the simplicity of the final statement one would expect that there should be a simpler and more natural argument. This is indeed the case, as we now explain. This alternative approach also leads directly to the double Eisenstein series which are the final topic of this paper.

We present the argument in a more abstract form than we will use here. Let A be a discrete subgroup of $\mathbb{C}$ (or possibly some more general commutative topological field) which is totally ordered, i.e. can be decomposed into a disjoint union $A^+ \cup \{0\} \cup (-A^+)$ with A^+ closed under addition. For $\mathbf{m}$, $\mathbf{n} \in A$ we write $\mathbf{n} \succ 0$ to mean $\mathbf{n} \in A^+$ and $\mathbf{m} \succ \mathbf{n}$ to mean $\mathbf{m} - \mathbf{n} \succ 0$. Then we claim that, at least in the range for which the sums $Z(r)=\sum_{\mathbf{m}\succ 0}\mathbf{m}^{-r}$ converge, the numbers $P_{r,s}=Z(r)Z(s)$ and $Z_k=Z(k)$ give a realization of $\mathcal{P}_k$. Indeed, by (iii) of Proposition 2 we can write $A^*(\mathbf{m},\mathbf{n})=f(\mathbf{m},\mathbf{n})-f(\mathbf{m}+\mathbf{n},\mathbf{n})-f(\mathbf{m},\mathbf{m}+\mathbf{n})$ for some $f\in V_k^*\otimes\mathbb{C}$ and

$$\sum_{r+s=k}a_{r,s}Z(r)Z(s) \;=\; \sum_{\mathbf{m},\,\mathbf{n}\succ 0}A^*(\mathbf{m},\mathbf{n}) \;=\; \Biggl(\sum_{\mathbf{m},\,\mathbf{n}\succ 0}-\sum_{\mathbf{m}\succ\mathbf{n}\succ 0}-\sum_{\mathbf{n}\succ\mathbf{m}\succ 0}\Biggr)f(\mathbf{m},\mathbf{n})$$
$$=\;\sum_{\mathbf{m}=\mathbf{n}\succ 0}A^*(\mathbf{m},\mathbf{n}) \;=\; f(1,1)\,Z(k)\,.$$

Applying this concept to the special case where $A=\mathbb{Z}\tau+\mathbb{Z}$ for some number τ in the upper half plane $\mathfrak{H}$ with the ordering described in the Introduction, we find that indeed the functions $\{G_r(\tau)\,G_s(\tau),\,G_k(\tau)\}$ give a realization of the $\{P_{r,s},Z_k\}$-part of $\mathcal{D}_k$. The corresponding realization of the $\{Z_{r,s}\}$-part is given by the double Eisenstein series $G_{r,s}$ as defined in the Introduction. In this section, we study these in more detail.

A simple estimate shows that the series defining G_k converges absolutely if and only if $k>2$ and the series defining $G_{r,s}$ if and only if $s>1$ and $r>2$. Our first object is to compute the Fourier expansion of $G_{r,s}$ (Theorem 6). We begin by recalling the corresponding computation for G_k, which is of course well-known. We define power series $\varphi_k^0(q)$ and $\varphi_k(q)$ in $\mathbb{Q}[[q]]$ — both actually polynomials of degree k in $q/(1-q)$ — by

$$\varphi_k^0(q) \;=\; \frac{(-1)^k}{(k-1)!}\sum_{u=1}^{\infty}u^{k-1}\,q^u\,,\quad \varphi_k(q)\;=\;-\frac{1}{2}\delta_{k,1}+\varphi_k^0(q)\qquad (k\geqslant 1).$$

The Lipschitz formula says that

$$\sum_{a\in\mathbb{Z}}\frac{1}{(\tau+a)^k}\;=\;(2\pi i)^k\varphi_k(q) \tag{50}$$

for $\tau\in\mathfrak{H}$ and all $k\geqslant 1$, where $q=e^{2\pi i\tau}$ as usual and where the sum on the left-hand side has to be interpreted as a Cauchy principal value if $k=1$. Applying this to $G_k(\tau)$ with $k\geqslant 2$, where the summation in the non-absolutely convergent case $k=2$ is to be carried out in the order $\succ$, we find

$$G_k(\tau)\;=\;\sum_{a>0}\frac{1}{a^k}\;+\;\sum_{m>0}\sum_{a\in\mathbb{Z}}\frac{1}{(m\tau+a)^k}\;=\;\zeta(k)\;+\;(2\pi i)^k\,g_k(q),$$

where

$$g_k(q)\;=\;\sum_{m>0}\varphi_k^0(q^m)\;=\;-\sum_{m,\,u>0}\frac{(-u)^{k-1}}{(k-1)!}\,q^{mu}\,. \tag{51}$$

The statement of Theorem 6, which we repeat here for convenience, was that the Fourier expansion of $G_{r,s}(\tau)$ in the convergent case is given by the analogous formula

$$G_{r,s}(\tau)\;=\;\zeta(r,s)\;+\sum_{h+p=k}C_{r,s}^p\,(2\pi i)^h g_h(q)\zeta(p)\;+\;(2\pi i)^k g_{r,s}(q)\,, \tag{52}$$

where $k=r+s$, $C_{r,s}^p$ is a simple numerical coefficient given by (13), and

$$g_{r,s}(q)=\sum_{m>n>0}\varphi_r^0(q^m)\varphi_s^0(q^n)=\sum_{\substack{m>n>0\\ u,\,v>0}}\frac{(-u)^{r-1}}{(r-1)!}\frac{(-v)^{s-1}}{(s-1)!}\,q^{mu+nv}. \tag{53}$$

(The condition $h>1$ and $p>1$ in (12) can be dropped, even though the definitions of $\widetilde{\zeta}(p)$ and $g_h(q)$ are problematic in these cases, because $C_{r,s}^p$ vanishes when $p=1$ or $p=r+s-1$ unless $r=1$.)

Proof of Theorem 6. We divide the sum of the defining series of

$$G_{r,s}(\tau)=\sum_{m\tau+a\succ n\tau+b\succ 0}\frac{1}{(m\tau+a)^r(n\tau+b)^s}$$

into four terms, according as $m=n=0$, $m>n=0$, $m=n>0$, or $m>n>0$. It is obvious that the terms of the first type give the double zeta value $\zeta(r,s)$ and that those of the second type give $(G_r(\tau)-\zeta(r))\zeta(s)=(2\pi i)^r g_r(q)\zeta(s)$, while those of the fourth type, again by virtue of the Lipschitz formula (50), give $(2\pi i)^k\sum_{m>n>0}\varphi_r(m\tau)\varphi_s(n\tau)$. (Note that it does not matter here whether we write $\varphi_r\varphi_s$ or $\varphi_r^0\varphi_s^0$ since we are assuming that both r and s are greater than 1.) Finally, the sum of the terms of the third type can be written as $\sum_{m>0}\Psi_{r,s}(m\tau)$, where

$$\Psi_{r,s}(\tau)=\sum_{a>b}\frac{1}{(\tau+a)^r(\tau+b)^s}\,.$$

This sum converges absolutely because we are assuming that $r,\ s\geqslant 2$, and is obviously periodic. To calculate its Fourier development, we use the partial fraction decomposition

$$\frac{1}{(\tau+a)^r(\tau+b)^s}\;=\;\sum_{h+p=r+s}\left[\frac{(-1)^s\binom{p-1}{s-1}}{c^p\,(\tau+a)^h}+\frac{(-1)^{p-r}\binom{p-1}{r-1}}{c^p\,(\tau+b)^h}\right]$$

(compare (19)), where $c=a-b>0$ and where we use our usual convention that the condition "$h+p=k$" tacitly includes "$h\geqslant 1,\ p\geqslant 1$." Using (50) yet again, we obtain

$$\Psi_{r,s}(\tau)\;=\;\sum_{h+p=r+s}\left[(-1)^s\binom{p-1}{s-1}\;+\;(-1)^{p-r}\binom{p-1}{r-1}\right]\zeta(p)\,(2\pi i)^h\varphi_h(\tau)\,,$$

where the implied interchange of order of summation is justified because the expression in square brackets vanishes if $p=1$ (because r and s are >1) or $h=1$ (because the binomial coefficient $\binom{r+s-2}{r-1}$ is symmetric in r and s). Replacing τ by $m\tau$ and summing over $m\geqslant 1$ replaces φ_h by g_h in this expression, and combining with the terms already computed, we obtain the desired formula (52). □

As explained in the Introduction, we want to do two things: find the "right" definition of the double Eisenstein series $G_{r,s}(\tau)$ in the cases when the original series defining it does not converge absolutely (i.e., if $r=1$ or 2 or if $s=1$), and give a purely combinatorial proof that the extended function satisfy the double shuffle relations. As also already explained, for the latter purpose we can ignore the term $\zeta(r,s)$ and the terms with p odd in the middle sum in (52), because they individually satisfy the double shuffle relations (the latter because of the corollary to Proposition 2 in Section 2). If we remove these terms, then what is left, after division by $(2\pi i)^k$ (where $k=r+s$ is the total weight as usual) is the power series $Z_{r,s}(q)$ defined in (17). We now extend this definition to all values of r and s by setting

$$Z_{r,s}(q)\;=\;g_{r,s}(q)\;+\;\beta_{r,s}(q)\;+\;\frac{1}{2}\,\varepsilon_{r,s}(q)\qquad(r,\ s\geqslant 1),\tag{54}$$

where $g_{r,s}(q)$ is defined by (53), $\beta_{r,s}(q)$ by

$$\beta_{r,s}(q)\;=\;\sum_{h+p=k}C_{r,s}^p\,\beta_p\,g_h(q)\qquad(r,\ s\geqslant 1,\ \ r+s=k)\tag{55}$$

with $\beta_p=-\dfrac{B_p}{2\,p!}$ (B_p = pth Bernoulli number), and

$$\varepsilon_{r,s}(q)=\delta_{r,2}\,g_s^*(q)-\delta_{r,1}\,g_{s-1}^*(q)+\delta_{s,1}\big(g_{r-1}^*(q)+g_r(q)\big)+\delta_{r,1}\delta_{s,1}g_2(q)\,,\tag{56}$$

$$g_k^*(q) = -\sum_{m>0} m\varphi_{k+1}^0(q^m) = \frac{(-1)^k}{k!}\sum_{m,\,u>0} mu^k q^{um} = \frac{1}{k}\,q\,\frac{d}{dq}\,g_k(q)\,. \tag{57}$$

Notice that β_p was previously used only for p even, where it was defined by $\beta_p = (2\pi i)^{-p}\zeta(p)$ and hence equal to $-B_p/2p!$ by Euler's theorem; now we take the latter formula as the definition for all values of p. This does not affect the definition when r and s are larger than 1, so that (55) agrees with (16) in these cases, because β_p is 0 for odd $p > 1$ and $C^1_{r,s} = 0$ for $r,\, s > 1$ (as we already used in the above proof). Since the further correction terms (56) are non-zero only if $r \le 2$ or $s = 1$, they do not occur in the region of convergence of $G_{r,s}$, so that (54) agrees with the earlier definition (17) in the cases where it was applicable. The reason for the various terms in (56) will become clear in the course of the proof of Theorem 7, but one term has a clear explanation which we can mention now: in the above derivation, the only place where it mattered that s was strictly greater than 1 was for the absolute convergence of the inner sum in $G_{r,s}(\tau) = \sum_{m,a}(m\tau+a)^{-r}\sum_n\sum_b(n\tau+b)^{-s}$, and if we interpret this sum as a Cauchy principal value and use (50) here too, then we see that the only effect on the final calculation is to replace the factor $\varphi_s^0(q)$ in (53) by $\varphi_s(q)$. Since they differ only by the constant $-1/2$ when $s = 1$, and not at all otherwise, this adds $-\frac{1}{2}\delta_{s,1}\sum_{m>0}(m-1)\varphi_r^0(q^m) = \frac{1}{2}\delta_{s,1}\bigl(g^*_{r-1}(q)+g_r(q)\bigr)$ to $g_{r,s}(q)$, and this accounts for the third term in the definition of $\varepsilon_{r,s}(q)$. Finally, we mention that (56) apparently contains the term $g_0^*(q)$ when $r = 1$ or $s = 1$, and this is not defined by the last formula in (57) (although the other two formulas do still make sense and lead to the definition $g_0^*(q) = g_2(q)$), but this is not important because the two terms in (56) that potentially involve $g_0^*(q)$ occur only when $r = s = 1$ and then cancel.

After these long preliminaries we can finally state and prove the full version of Theorem 7 from the Introduction.

Theorem 7. *There is a realization in $\mathbb{Q}[[q]]^0$ of the double shuffle relations* (20) *for all weights with $Z_{r,s}(q)$ defined by* (54),

$$P_{r,s}(q) = g_r(q)g_s(q) + \beta_r g_s(q) + \beta_s g_r(q) + \frac{1}{2}(\delta_{r,2}\, g_s^*(q) + \delta_{s,2}\, g_r^*(q)) \tag{58}$$

for all $r,\, s \geqslant 1$, and with $Z_k(q) = g_k(q)$ for $k > 2$, $Z_2(q) = 0$.

Proof. The proof will be shorter than the discussion leading up to the statement. Of course we use generating functions. We drop the "(q)" in the names of elements of $\mathbb{Q}[[q]]$ and systematically write $\gamma(X)$ and $\gamma(X,Y)$ for the generating functions $\sum_{k\geqslant 1}\gamma_k X^{k-1}$ and $\sum_{r,\,s\geqslant 1}\gamma_{r,s}X^{r-1}Y^{s-1}$ associated to sequences $\{\gamma_k\}$ or $\{\gamma_{r,s}\}$ indexed by one or two integers, respectively. Then from the definitions (51), (57), (53), (55) and (56) we have

$$\begin{aligned}
\beta(X) \;&=\; \sum_{k\geqslant 1}\beta_k\,X^{k-1} \;=\; \frac{1}{2}\left(\frac{1}{X} - \frac{1}{e^X-1}\right),\\
g(X) \;&=\; \sum_{k\geqslant 1} g_k\,X^{k-1} \;=\; -\sum_{u>0} e^{-uX}\,\frac{q^u}{1-q^u}\,,\\
g^*(X) \;&=\; \sum_{k\geqslant 1} g_k^*\,X^{k-1} \;=\; \frac{1}{X}\left(\sum_{u>0} e^{-uX}\,\frac{q^u}{(1-q^u)^2} - g_2\right),\\
g(X,Y) \;&=\; \sum_{r,\,s\geqslant 1} g_{r,s}\,X^{r-1}Y^{s-1} \;=\; \sum_{\substack{m>n>0\\ u,\,v>0}} e^{-uX-vY}\,q^{mu+nv}\\
&=\; \sum_{u,\,v>0} e^{-uX-vY}\,\frac{q^u}{1-q^u}\,\frac{q^{u+v}}{1-q^{u+v}}\,,
\end{aligned}$$

$$\begin{aligned}
\beta(X,Y) \;&=\; \sum_{r,\,s\geqslant 1} \beta_{r,s}\, X^{r-1}Y^{s-1} \;=\; \sum_{h,\,p\geqslant 1} \beta_p\, g_h \Bigl(\sum_{r+s=k} C^p_{r,s} X^{r-1}Y^{s-1}\Bigr)\\
&= \sum_{h,\,p\geqslant 1} \beta_p\, g_h \bigl(Y^{p-1}X^{h-1} - (X-Y)^{p-1}(X^{h-1}-Y^{h-1})\bigr)\\
&= \beta(Y)\, g(X) \,-\, \beta(X-Y)\,\bigl(g(X)-g(Y)\bigr)\,,\\
\varepsilon(X,Y) \;&=\; \sum_{r,\,s\geqslant 1} \varepsilon_{r,s}\, X^{r-1}Y^{s-1} = (X-Y)g^*(Y) + Xg^*(X) + g(X) + g_2,
\end{aligned}$$

and we want to show that the generating functions

$$\begin{aligned}
\mathfrak{Z}(X,Y) \;&=\; g(X,Y) \,+\, \beta(X,Y) \,+\, \frac{1}{2}\,\varepsilon(X,Y)\,,\\
\mathfrak{P}(X,Y) \;&=\; g(X)g(Y) \,+\, \beta(X)g(Y) \,+\, \beta(Y)g(X) \,+\, \frac{1}{2}\bigl(Xg^*(Y)+Yg^*(X)\bigr)\,,\\
\mathfrak{z}(X) \;&=\; g(X) \,-\, g_2
\end{aligned}$$

satisfy (27). So we must calculate $Z(X,Y)+Z(Y,X)$ and $Z(X+Y,Y)+Z(X+Y,X)$ for each of the three pieces $Z=g$, β, and ε constituting $\mathfrak{Z}$.

From the above formulas for the generating functions we find

$$\begin{aligned}
g(X,Y) \,+\, g(Y,X) \;&=\; \sum_{u,\,v>0} e^{-uX-vY} \left(\frac{q^u}{1-q^u} \,+\, \frac{q^v}{1-q^v}\right) \frac{q^{u+v}}{1-q^{u+v}}\\
&=\; \sum_{u,\,v>0} e^{-uX-vY} \left(\frac{q^u}{1-q^u}\,\frac{q^v}{1-q^v} \,-\, \frac{q^{u+v}}{1-q^{u+v}}\right)\\
&=\; g(X)g(Y) \,-\, \sum_{w>0} \frac{e^{(1-w)Y}-e^{(1-w)X}}{e^X-e^Y}\,\frac{q^w}{1-q^w}\\
&=\; g(X)g(Y) \,+\, \frac{e^Y}{e^X-e^Y}\, g(Y) \,-\, \frac{e^X}{e^X-e^Y}\, g(X)\\
&=\; g(X)g(Y) \,-\, \frac{g(X)+g(Y)}{2} \,-\, \coth\bigl(\frac{X-Y}{2}\bigr)\,\frac{g(X)-g(Y)}{2}\,,\\
\beta(X,Y) \,+\, \beta(Y,X) \;&=\; \beta(Y)g(X) \,+\, \beta(X)g(Y)\\
&\qquad -\, \frac{g(X)-g(Y)}{X-Y} \,+\, \coth\bigl(\frac{X-Y}{2}\bigr)\,\frac{g(X)-g(Y)}{2}\\
\varepsilon(X,Y) \,+\, \varepsilon(Y,X) \;&=\; Xg^*(Y) \,+\, Yg^*(X) \,+\, g(X) \,+\, g(Y) + 2g_2\,,
\end{aligned}$$

and adding up these three equations (the last with a coefficient $1/2$) we obtain the first of equations (27) for $\mathfrak{Z}$, $\mathfrak{P}$ and $\mathfrak{z}$ as defined above. Similarly, we have

$$\begin{aligned}
g(X+Y,Y) \,+\, g(X+Y,X) \;&=\; \Bigl(\sum_{v>u>0} \,+\, \sum_{u>v>0}\Bigr)\, e^{-uX-vY}\, \frac{q^u}{1-q^u}\,\frac{q^v}{1-q^v}\\
&=\; \Bigl(\sum_{u,\,v>0} \,-\, \sum_{u=v>0}\Bigr)\, e^{-uX-vY}\, \frac{q^u}{1-q^u}\,\frac{q^v}{1-q^v}\\
&=\; g(X)g(Y) \,-\, \sum_{u>0} e^{-u(X+Y)} \left(\frac{q^u}{(1-q^u)^2} \,-\, \frac{q^u}{1-q^u}\right)\\
&=\; g(X)g(Y) \,-\, (X+Y)g^*(X+Y) \,-\, g_2 \,-\, g(X+Y)\,,\\
\beta(X+Y,Y) \,+\, \beta(X+Y,X) \;&=\; \beta(X)g(Y) \,+\, \beta(Y)g(X)\,,\\
\varepsilon(X+Y,Y) \,+\, \varepsilon(X+Y,X) \;&=\; Xg^*(Y) + Yg^*(X) + 2(X+Y)g^*(X+Y)\\
&\qquad +\, 2g(X+Y) + 2g_2\,,
\end{aligned}$$

and combining these gives the second of equations (27). $\square$

Remarks. 1. It is notable that the power series defined by (58), with the constant term $\beta_r\beta_s$ added, is a modular form of weight $r+s$ for all even $r, s > 0$, the correction terms $\delta_{r,2}g_s^*(q)$ and $\delta_{s,2}g_r^*(q)$ being just what is needed to compensate for the non-modularity of $(\beta_r + g_r(q))(\beta_s + g_s(q))$ when r or s is equal to 2.

2. One can also ask whether it is possible to lift Theorem 7 from $\mathbb{Q}[[q]]^0$ to all of $\mathbb{Q}[[q]]$ by adding a suitable constant term to $Z_{r,s}(q)$ in such a way that the relations (20) still hold when we add $\beta_r\beta_s$ to $P_{r,s}(q)$ in order to make it modular for r and s even. This is equivalent to finding a realization of $\mathcal{D}_k$ in $\mathbb{Q}$ with $P_{r,s} = \beta_r\beta_s$ for r and s even. One such realization is provided by the "Bernoulli realization" given in Section 6, but there may be other ones which are more naturally related to the combinatorial double Eisenstein series $Z_{r,s}(q)$.

References

[1] **Akiyama, S., Tanigawa, Y.** *Multiple zeta values at non-positive integers.* Ramanujan J. 5, no. 4 (2001), 327–351.

[2] **Arakawa, T., Ibukiyama, T., and Kaneko, M.** *Bernoulli numbers and zeta functions* (in Japanese; English translation in preparation). Makino Shoten, 2001.

[3] **Ihara, K., Kaneko, M., and Zagier, D.** *Derivation and double shuffle relations for multiple zeta values.* To appear in Compositio Math.

[4] **Ihara, Y.** *Some arithmetic aspects of Galois actions in the pro-p fundamental group of* $\mathbb{P}^1 - \{0, 1, \infty\}$. Arithmetic fundamental groups and noncommutative algebra (Berkeley, CA, 1999), 247–273, Proc. Sympos. Pure Math., 70, Amer. Math. Soc., Providence, RI, 2002.

[5] **Goncharov, A. B.** *The dihedral Lie algebras and Galois symmetries of* $\pi_1^{(l)}(\mathbb{P}^1 - (\{0, \infty\} \cup \mu_N))$. Duke Math. J. 110 (2001), no. 3, 397–487.

[6] **Goncharov, A. B.** *Galois symmetries of fundamental groupoids and noncommutative geometry.* Duke Math. J. 128, no. 2 (2005), 209–284

[7] **Haberland, K.** *Perioden von Modulformen einer Variablen und Gruppencohomologie.* Math. Nachr. 112 (1983), 245–282.

[8] **Kohnen, W., Zagier, D.** *Modular forms with rational periods.* Modular forms (Durham, 1983), 197–249, Ellis Horwood Ser. Math. Appl.: Statist. Oper. Res., Horwood, Chichester, 1984.

[9] **Lang, S.** *Introduction to modular forms.* Grundlehren der mathematischen Wissenschaften, No. 222. Springer-Verlag, Berlin-New York, 1976.

[10] **Lewis, J. B.** *Spaces of holomorphic functions equivalent to the even Maass cusp forms.* Invent. Math. 127 (1997), no. 2, 271–306.

[11] **Lewis, J., Zagier, D.** *Period functions for Maass wave forms. I.* Ann. of Math. (2) 153 (2001), no. 1, 191–258.

[12] **Racinet, G.** *Doubles mélanges des polylogarithmes multiples aux racines de l'unité.* Publ. Math. Inst. Hautes Etudes Sci. No. 95 (2002), 185–231.

[13] **Rankin, R. A.** *The scalar product of modular forms.* Proc. London Math. Soc. (3) 2, (1952). 198–217.

[14] **Schneps, L.** *On the Poisson Bracket on the Free Lie Algebra in two Generators.* Journal of Lie Theory 16, no. 1 (2006), 19–37.

[15] **Zagier, D.** *Periods of modular forms, traces of Hecke operators, and multiple zeta values.* Research into automorphic forms and L-functions (Kyoto, 1992). RIMS Kokyuroku No. 843, (1993), 162–170.

[16] **Zagier, D.** *Quelques conséquences surprenantes de la cohomologie de* $SL(2, \mathbb{Z})$. Leçons de Mathématiques d'aujourd'hui, Cassini, Paris (2000), 99-123.

[17] **Zagier, D.** *Periods of modular forms and Jacobi theta functions.* Invent. Math. 104 (1991), no. 3, 449–465.

[18] **Zagier, D.** *Values of zeta functions and their applications.* First European Congress of Mathematics, Vol. II (Paris, 1992), 497–512, Progr. Math., 120, Birkhäuser, Basel, 1994.