

# ON THE ALGEBRAIC STRUCTURE OF ITERATED INTEGRALS OF QUASIMODULAR FORMS

NILS MATTHES

ABSTRACT. We study the algebra  $\mathcal{I}^{QM}$  of iterated integrals of quasimodular forms for  $\mathrm{SL}_2(\mathbb{Z})$ , which is the smallest extension of the algebra  $QM_*$  of quasimodular forms, which is closed under integration. We prove that  $\mathcal{I}^{QM}$  is a polynomial algebra in infinitely many variables, given by Lyndon words on certain monomials in Eisenstein series. We also prove an analogous result for the  $M_*$ -subalgebra  $\mathcal{I}^M$  of  $\mathcal{I}^{QM}$  of iterated integrals of modular forms.

## 1. INTRODUCTION

Quasimodular forms are generalizations of modular forms, which have first been introduced in [10], in a context motivated by mathematical physics. The  $\mathbb{C}$ -algebra  $QM_*$  of quasimodular forms for the full modular group  $\mathrm{SL}_2(\mathbb{Z})$  can be defined, in a slightly ad hoc fashion, as the polynomial ring  $\mathbb{C}[E_2, E_4, E_6]$ , where  $E_{2k}$  denotes the normalized Eisenstein series of weight  $2k$ :

$$E_{2k}(\tau) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} n^{2k-1} \frac{q^n}{1-q^n}, \quad q = e^{2\pi i \tau},$$

where  $B_{2k}$  are the Bernoulli numbers. In particular,  $QM_*$  contains the algebra of modular forms  $M_* \cong \mathbb{C}[E_4, E_6]$ .

The derivative of a quasimodular form (of weight  $k$ ) is again a quasimodular form (of weight  $k+2$ ); this was essentially already known to Ramanujan (cf. [19], Proposition 15). On the other hand, the integral of a quasimodular form is in general not quasimodular. For example, a primitive of  $E_2$  would have to be of weight zero, but every quasimodular form of weight zero is constant.

The goal of this paper is to study the smallest algebra extension of  $QM_*$ , which is closed under integration. For this, the idea is to iteratively adjoin primitives to  $QM_*$ , which eventually leads to adjoining all (indefinite) *iterated integrals*

$$(1.1) \quad I(f_1, \dots, f_n; \tau) = (2\pi i)^n \int_{\tau \leq \tau_1 \leq \dots \leq \tau_n \leq i\infty} \dots \int f_1(\tau_1) \dots f_n(\tau_n) d\tau_1 \dots d\tau_n,$$

where  $f_1, \dots, f_n$  are quasimodular forms (a precise definition will be given in Section 2.3). The integrals (1.1) have first been studied by Manin [13] and later by Brown [3] and Hain [9], in the case where all the  $f_i$  are modular forms.<sup>1</sup> In all of these treatments, the focus lies rather on arithmetic aspects of these iterated integrals, for example their special values at cusps of the upper half-plane. By contrast, we study them solely as holomorphic functions of  $\tau$ . It is also worth noting that even in the

2010 *Mathematics Subject Classification.* 11F11, (11F67).

*Key words and phrases.* Quasimodular forms, iterated integrals.

<sup>1</sup>More precisely, Manin only defined iterated integrals of cusp forms, and the extension to all modular forms is due to Brown.

modular case, the iterated integrals we study in the present paper are slightly more general than the ones introduced in [3, 9, 13]. For example, if  $f(\tau)$  is a modular form of weight  $k$ , then the integral  $\int_{\tau}^{i\infty} f(\tau_1)\tau_1^n d\tau_1$  is an iterated integral of modular forms in the sense of the present paper for every  $n \geq 0$ , while [3, 9, 13] also require  $n \leq k - 2$ .

Now let  $\mathcal{I}^{QM}$  be the  $QM_*$ -algebra generated by all the integrals (1.1), which is the smallest algebra extension of  $QM_*$ , closed under integration. It turns out that  $\mathcal{I}^{QM}$  is not finitely generated, but still has a manageable structure, which is captured by the notion of shuffle algebra (which is just the graded dual of the tensor algebra with a certain commutative multiplication, the so-called shuffle product) [17]. More precisely, let  $V = \mathbb{C} \cdot E_2 \oplus M_*$  be the  $\mathbb{C}$ -vector space spanned by all modular forms and the Eisenstein series  $E_2$ , and let  $\mathbb{C}\langle V \rangle$  be the shuffle algebra on  $V$ . Our main result is the following.

**Theorem** (Theorem 4.3 below). *The  $QM_*$ -linear morphism*

$$\begin{aligned} \varphi^{QM} : QM_* \otimes_{\mathbb{C}} \mathbb{C}\langle V \rangle &\rightarrow \mathcal{I}^{QM} \\ [f_1 | \dots | f_n] &\mapsto I(f_1, \dots, f_n; \tau) \end{aligned}$$

*is an isomorphism of  $QM_*$ -algebras.*

A similar result holds for the  $M_*$ -subalgebra  $\mathcal{I}^M$  of  $\mathcal{I}^{QM}$  of iterated integrals of modular forms (cf. Theorem 4.5).<sup>2</sup> The surjectivity of  $\varphi^{QM}$  can be reduced to the fact that every quasimodular form can be written uniquely as a polynomial in  $n$ -th derivatives of modular forms and the Eisenstein series  $E_2$  (cf. [19], Proposition 20). The proof of injectivity is more elaborate and amounts to showing that iterated integrals of modular forms and the Eisenstein series  $E_2$  are linearly independent over  $QM_*$ . It extends a result of [12] which dealt with iterated integrals of Eisenstein series. In both cases, the key is to use a general result on linear independence of iterated integrals [5]. It would be interesting to prove similar results for quasimodular forms for congruence subgroups.

The Milnor–Moore theorem [15] states that if  $k$  has characteristic zero, then  $k\langle V \rangle$  is isomorphic to a polynomial algebra (usually in infinitely many variables). Fixing a (totally ordered) basis  $\mathcal{B}$  of  $V$ , Radford [16] has given explicit generators of  $k\langle V \rangle$  in terms of Lyndon words on  $\mathcal{B}$  (cf. Section 4). Using this, we get the following theorem.

**Theorem** (Theorem 4.9 below). *Let  $\mathcal{B}$  be a basis of  $\mathbb{C} \cdot E_2 \oplus M_*$ . We have a natural isomorphism*

$$(1.2) \quad \mathcal{I}^{QM} \cong QM_*[\text{Lyn}(\mathcal{B}^*)],$$

*where the right hand side is the polynomial  $QM_*$ -algebra on the set  $\text{Lyn}(\mathcal{B}^*)$  of Lyndon words of  $\mathcal{B}$ .*

Again, a similar result holds for  $\mathcal{I}^M$ . Since  $QM_*$  has an explicit basis given by monomials in the Eisenstein series  $E_2$ ,  $E_4$  and  $E_6$ , the isomorphism (1.2) can be made completely explicit, and may be viewed as an analog of the isomorphism  $QM_* \cong \mathbb{C}[E_2, E_4, E_6]$  [10].

<sup>2</sup>After this paper has been submitted for publication, the author learned that, in the case of iterated integrals of modular forms, a very similar result has also been proved by Brown (cf. [4], Proposition 4.4), using a slightly different method.

Finally, we note that classically, integrals of modular forms play an important role in Eichler–Shimura theory, where they give rise to group-cocycles (say for  $\mathrm{SL}_2(\mathbb{Z})$  or more generally for some congruence subgroup thereof) with values in homogeneous polynomials. This has been generalized by Manin [13], and later by Brown [3] and Hain [9], who attach certain non-abelian cocycles to iterated integrals of modular forms. Although it is not the main focus of this article, in the appendix we show how one can attach cocycles to quasimodular forms (for  $\mathrm{SL}_2(\mathbb{Z})$ ), partly since we found no mention of this in the literature. On the other hand, we leave the definition and study of cocycles attached to iterated integrals of quasimodular forms for future investigation.

The plan of the paper is as follows. In Section 2, we collect the necessary background on quasimodular forms and their iterated integrals. In Section 3, we prove a linear independence result for iterated integrals of quasimodular forms. This result is then put to use in Section 4, where the main results are proved. In the appendix, we discuss the above-mentioned generalization of the classical Eichler–Shimura theory to quasimodular forms for  $\mathrm{SL}_2(\mathbb{Z})$ .

**Acknowledgments:** Very many thanks to Pierre Lochak for bringing the article [5] to the author’s attention. Also, many thanks to Francis Brown, Erik Panzer and the referees for corrections as well as very helpful suggestions and to Don Zagier for inspiring discussions on the appendix. The results of this paper were found while the author was a PhD student at Universität Hamburg under the supervision of Ulf Kühn.

## 2. PRELIMINARIES

Throughout the paper, all modular and quasimodular forms will be for  $\mathrm{SL}_2(\mathbb{Z})$ . We fix some notation. Let  $\mathfrak{H} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$  be the upper half-plane with canonical coordinate  $\tau$ . For every  $k \in \mathbb{Z}$ , we have a group action of  $\mathrm{SL}_2(\mathbb{Z})$  on the set of all functions  $f : \mathfrak{H} \rightarrow \mathbb{C}$  (not necessarily holomorphic), defined by  $(\gamma, f) \mapsto f|_k\gamma$ , where

$$(f|_k\gamma)(\tau) := (c\tau + d)^{-k} f\left(\frac{a\tau + b}{c\tau + d}\right).$$

For fixed  $\tau \in \mathfrak{H}$ , we also define a map  $X : \mathrm{SL}_2(\mathbb{Z}) \rightarrow \mathbb{C}$  by  $X(\gamma) = \frac{1}{2\pi i} \frac{c}{c\tau + d}$ . Note that  $X$  has infinite, and thus Zariski dense, image.

**2.1. Recap of modular forms.** Denote by  $M_k$  the space of modular forms of weight  $k \in \mathbb{Z}$ . By definition, these are the holomorphic functions  $f : \mathfrak{H} \rightarrow \mathbb{C}$ , which satisfy  $f|_k\gamma = f$  for all  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ , and which are “holomorphic at the cusp”. The latter condition means that in the Fourier expansion  $f(\tau) = \sum_{n \in \mathbb{Z}} a_n q^n$  (which exists since for  $\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ , the condition  $f|_k\gamma = f$  is just  $f(\tau + 1) = f(\tau)$  for all  $\tau$ ), all  $a_n = 0$  for  $n < 0$ . Examples of modular forms include the Eisenstein series

$$E_{2k}(\tau) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} n^{2k-1} \frac{q^n}{1 - q^n} = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \left( \sum_{d|n} d^{2k-1} \right) q^n,$$

which is a modular form of weight  $2k$ , for  $k \geq 2$  (the  $B_{2k}$  are Bernoulli numbers). The  $\mathbb{C}$ -vector space of all modular forms  $M_*$  is a graded (for the weight)  $\mathbb{C}$ -algebra  $M_* = \bigoplus_{k \in \mathbb{Z}} M_k$ , which is well-known to be isomorphic to the polynomial algebra

$\mathbb{C}[E_4, E_6]$ . Proofs of all these facts and much more on modular forms can be found for example in [19].

**2.2. Quasimodular forms.** Quasimodular forms are a generalization of modular forms, which have first been introduced in [10] (see also [1], §3 and [19], §5.3). The definition we give here is due to W. Nahm<sup>3</sup> and is also used for example in [14].

**Definition 2.1.** Let  $k, p \in \mathbb{Z}$  with  $p \geq 0$ . A *quasimodular form* of weight  $k$  and depth  $\leq p$  is a function  $f : \mathfrak{H} \rightarrow \mathbb{C}$  with the following property: there exist holomorphic functions  $f_r : \mathfrak{H} \rightarrow \mathbb{C}$ , for  $0 \leq r \leq p$ , which have Fourier expansions  $\sum_{n=0}^{\infty} a_n q^n$ , such that

$$(2.1) \quad (f|_k \gamma)(\tau) = \sum_{r=0}^p f_r(\tau) X(\gamma)^r, \quad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}).$$

We denote by  $QM_k^{\leq p}$  the  $\mathbb{C}$ -vector space of quasimodular forms of weight  $k$  and depth  $\leq p$  and set

$$QM_k := \bigcup_{p \geq 0} QM_k^{\leq p}, \quad QM_* := \bigoplus_{k \in \mathbb{Z}} QM_k.$$

**Remark 2.2.** (i) It is clear from the definition that, if  $f_1 \in QM_{k_1}^{\leq p_1}$ ,  $f_2 \in QM_{k_2}^{\leq p_2}$ , then  $f_1 f_2 \in QM_{k_1+k_2}^{\leq p_1+p_2}$ . In other words,  $QM_*$  is a graded (for the weight) and filtered (for the depth)  $\mathbb{C}$ -algebra.

(ii) Using that  $X$  is Zariski dense, it is easy to see that the functions  $f_r(\tau)$  are uniquely determined by  $f(\tau)$ . Also, applying (2.1) with  $\gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , we see that  $f_0(\tau) = f(\tau)$ . In particular, every quasimodular form is holomorphic on  $\mathfrak{H}$  and at the cusp.

Every modular form is a quasimodular form of depth zero, more precisely,  $M_k = QM_k^{\leq 0}$ . An example of a quasimodular form, which is not modular is the Eisenstein series of weight two  $E_2(\tau) = 1 - 24 \sum_{n=1}^{\infty} n \frac{q^n}{1-q^n}$ , which transforms as

$$(2.2) \quad (E_2|_2 \gamma)(\tau) = E_2(\tau) + 12X(\gamma) = E_2(\tau) - \frac{6i}{\pi} \frac{c}{c\tau + d},$$

for all  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ . In particular,  $E_2 \in QM_2^{\leq 1} \setminus M_2$ .

The following proposition recalls basic properties of  $QM_*$  that will be of use later.

**Proposition 2.3.** (i) *The  $\mathbb{C}$ -algebra  $QM_*$  is closed under the differential operator  $D := \frac{1}{2\pi i} \frac{d}{d\tau} = q \frac{d}{dq}$ . More precisely, for  $f$  quasimodular of weight  $k$  and depth  $\leq p$ , we have*

$$(D(f)|_{k+2} \gamma)(\tau) = \sum_{r=0}^{p+1} (D(f_r)(\tau) + (k-r+1)f_{r-1}(\tau))X(\gamma)^r.$$

*In particular,  $D(QM_k^{\leq p}) \subset QM_{k+2}^{\leq p+1}$  for all  $k, p \in \mathbb{Z}$ .*

(ii) *We have*

$$QM_k = \begin{cases} \{0\}, & \text{if } k < 0 \\ \mathbb{C} \cdot E_2, & \text{if } k = 2 \\ D(QM_{k-2}) \oplus M_k & \text{else.} \end{cases}$$

<sup>3</sup>Cf. [19], Section 5.3.

In particular,  $QM_* = \mathbb{C} \cdot E_2 \oplus D(QM_*) \oplus M_*$ , and

$$QM_* \cong \mathbb{C}[E_2, E_4, E_6]$$

as graded  $\mathbb{C}$ -algebras.

*Proof:* For (i), simply apply  $D$  to both sides of (2.1). The first equality in (ii) follows from [19], Proposition 20.(iii), and the isomorphism  $QM_* \cong \mathbb{C}[E_2, E_4, E_6]$  is essentially a consequence of this, but can also be proved independently (cf. [1], Proposition 3.5.(ii)).  $\square$

**Remark 2.4.** Relaxing the condition in the definition of quasimodular forms that every  $f_r$  be a holomorphic function, one can define the notion of *weakly quasimodular form* of weight  $k$  and depth  $\leq p$  as a meromorphic function  $f : \mathfrak{H} \rightarrow \mathbb{C}$  satisfying (2.1), but where the functions  $f_r(\tau)$  are only required to be meromorphic on  $\mathfrak{H}$  and have Fourier series of the form  $\sum_{n=-M}^{\infty} a_n q^n$  ( $f_r$  is “meromorphic at the cusp”). As in the case of quasimodular forms, one shows easily that the functions  $f_r(\tau)$  are uniquely determined by  $f(\tau)$  (cf. Remark 2.2). Moreover, Proposition 2.3.(i) generalizes straightforwardly to weakly quasimodular forms.

We end this subsection with a short lemma, for which we couldn’t find a suitable reference. Denote by  $\Delta = \frac{1}{1728}(E_4^3 - E_6^2)$  Ramanujan’s cusp form of weight 12.

**Lemma 2.5.** *Let  $g \in QM_* \setminus \{0\}$  and  $\alpha \in \mathbb{C}$  such that*

$$(2.3) \quad D(g) = (\alpha E_2) \cdot g.$$

*Then  $\alpha$  is a non-negative integer, and  $g = \beta \Delta^\alpha$  for some  $\beta \in \mathbb{C} \setminus \{0\}$ .*

*Proof:* Let  $g = \sum_{n=0}^{\infty} a_n q^n$ , so that  $D(g) = \sum_{n=0}^{\infty} n a_n q^n$ . Comparing coefficients on both sides of (2.3) yields that  $\alpha$  equals the smallest integer  $m \geq 0$  such that  $a_m \neq 0$ . On the other hand,  $\frac{D(\Delta)}{\Delta} = E_2$  (cf. [19], proof of Proposition 7), and from the chain rule  $\frac{D(\Delta^\alpha)}{\Delta^\alpha} = \alpha E_2$ , which gives the result.  $\square$

**2.3. Iterated integrals on the upper half-plane.** Iterated integrals of modular forms have been considered first by Manin (for cusp forms) [13], and later by Brown (in general) [3]. They are generalizations of the classical Eichler integrals [6, 11]

$$(2.4) \quad \int_{\tau}^{i\infty} f(z) z^m dz, \quad m = 0, \dots, k-2$$

where  $f$  is a cusp form of weight  $k$ . Extending (2.4) to a general modular form poses the problem of logarithmic divergences, which arise from the constant term in the Fourier series of  $f$ . A procedure for regularizing such integrals is described in [3], and we borrow it to define iterated integrals of quasimodular forms. Since it is perhaps not so well-known, we give some details, for the convenience of the reader.

Let  $W \subset \mathcal{O}(\mathfrak{H})$  be the  $\mathbb{C}$ -subalgebra of holomorphic functions  $f : \mathfrak{H} \rightarrow \mathbb{C}$ , which have an everywhere convergent Fourier series  $f(\tau) = \sum_{n=0}^{\infty} a_n q^n$  with  $q = e^{2\pi i \tau}$ . Note that  $QM_* \subset W$ . For  $f(\tau) \in W$ , let  $f^\infty = a_0$ , and  $f^0(\tau) = f(\tau) - f^\infty = \sum_{n=1}^{\infty} a_n q^n$ . Let  $\mathbb{C}\langle W \rangle$  (sometimes denoted by  $T^c(W)$ ) be the shuffle algebra [17], i.e. the graded dual of the tensor algebra  $T(W) = \bigoplus_{k \geq 0} W^{\otimes k}$  on  $W$ , where the grading is by the length of tensors. Elements of  $(W^{\otimes n})^\vee$  will be written using bar notation  $[f_1 | f_2 | \dots | f_n]$ , and a general element of  $\mathbb{C}\langle W \rangle$  is a  $\mathbb{C}$ -linear combination of

those. The product on  $\mathbb{C}\langle W \rangle$  is the shuffle product  $\sqcup$ , which is defined on the basic elements by

$$(2.5) \quad [f_1 | \dots | f_r] \sqcup [f_{r+1} | \dots | f_{r+s}] = \sum_{\sigma \in \Sigma_{r,s}} [f_{\sigma(1)} | \dots | f_{\sigma(r+s)}],$$

where  $\Sigma_{r,s}$  denotes the set of all the permutations on the set  $\{1, \dots, r+s\}$  such that  $\sigma^{-1}(1) < \dots < \sigma^{-1}(r)$  and  $\sigma^{-1}(r+1) < \dots < \sigma^{-1}(r+s)$ .

Define a  $\mathbb{C}$ -linear map  $R : \mathbb{C}\langle W \rangle \rightarrow \mathbb{C}\langle W \rangle$  by the formula

$$R[f_1 | \dots | f_n] = \sum_{i=0}^n (-1)^{n-i} [f_1 | \dots | f_i] \sqcup [f_n^\infty | \dots | f_{i+1}^\infty].$$

Following [3], Section 4, we make the following definition.

**Definition 2.6.** For  $f_1, \dots, f_n \in W$ , define their regularized iterated integral

$$(2.6) \quad I(f_1, \dots, f_n; \tau) := (2\pi i)^n \sum_{i=0}^n (-1)^{n-i} \int_\tau^{i\infty} R[f_1 | \dots | f_i] \int_0^\tau [f_n^\infty | \dots | f_{i+1}^\infty],$$

where  $\int_a^b [f_1 | \dots | f_n] := \int_{0 \leq t_1 \leq \dots \leq t_n \leq 1} (\gamma_a^b)^*(f_1(\tau_1) d\tau_1) \dots (\gamma_a^b)^*(f_n(\tau_n) d\tau_n)$  denotes the usual iterated integral along the straight line path  $\gamma_a^b$  from  $a$  to  $b$ .

**Remark 2.7.** Using the change of variables  $\tau \mapsto q = e^{2\pi i \tau}$ , it is easy to see that  $I(f_1, \dots, f_n; \tau) \in W[\log(q)]$ , where  $\log(q) := 2\pi i \tau$ . By the same token, if all of the  $f_i$  have rational Fourier coefficients, then  $I(f_1, \dots, f_n; \tau)$  will also have rational coefficients, as a series in  $q$  and  $\log(q)$ .

**Proposition 2.8.** *The functions  $I(f_1, \dots, f_n; \tau)$  satisfy the following properties.*

(i) *The product of any two of them is given by the shuffle product:*

$$(2.7) \quad I(f_1, \dots, f_r; \tau) I(f_{r+1}, \dots, f_{r+s}; \tau) = \sum_{\sigma \in \Sigma_{r,s}} I(f_{\sigma(1)}, \dots, f_{\sigma(r+s)}; \tau).$$

(ii) *They satisfy the differential equation*

$$(2.8) \quad \frac{1}{2\pi i} \frac{d}{d\tau} \Big|_{\tau=\tau_0} I(f_1, \dots, f_n; \tau) = -f_1(\tau_0) I(f_2, \dots, f_n; \tau_0).$$

(iii) *We have the integration by parts formulas*

$$(2.9) \quad \begin{aligned} I(f_1, \dots, f_i, D(g), f_{i+1}, \dots, f_n; \tau) &= I(f_1, \dots, f_i, g f_{i+1}, \dots, f_n; \tau) \\ &\quad - I(f_1, \dots, f_i g, f_{i+1}, \dots, f_n; \tau), \end{aligned}$$

as well as

$$I(D(g), f_2, \dots, f_n; \tau) = I(g f_2, f_3, \dots, f_n; \tau) - g(\tau) I(f_2, \dots, f_n; \tau),$$

and

$$I(f_1, \dots, f_{n-1}, D(g); \tau) = g(i\infty) I(f_1, \dots, f_{n-1}; \tau) - I(f_1, \dots, f_{n-1} g; \tau).$$

*Proof:* Using the definition (2.6), all of these follow from the analogous properties for usual iterated integrals (cf. e.g. [7]).  $\square$

**2.4. A criterion for linear independence of iterated integrals.** Let  $\text{Frac}(W)$  be the field of fractions of the  $\mathbb{C}$ -algebra  $W$  introduced in the last subsection. By the quotient rule, it is easy to see that  $\text{Frac}(W)$  is closed under  $D = \frac{1}{2\pi i} \frac{d}{d\tau}$ .

The following theorem is a special case of the main result of [5].

**Theorem 2.9.** *Let  $\mathcal{F} = (f_i)_{i \in I}$  be a family of elements of  $W$ , and let  $\mathcal{C} \subset \text{Frac}(W)$  be a subfield, which is closed under  $D$  and contains  $\mathcal{F}$ . The following are equivalent:*

- (i) *The family of iterated integrals  $(I(f_1, \dots, f_n; \tau) \mid f_i \in I, n \geq 0)$  is linearly independent over  $\mathcal{C}$ .*
- (ii) *The family  $\mathcal{F}$  is linearly independent over  $\mathbb{C}$ , and we have*

$$D(\mathcal{C}) \cap \text{Span}_{\mathbb{C}}(\mathcal{F}) = \{0\}.$$

*Proof:* This is the special case of Theorem 2.1 in [5], with the notation of *loc.cit.*,  $k = \mathbb{C}$ ,  $(\mathcal{A}, d) = (\text{Frac}(\mathcal{O}(\mathfrak{H})), D)$ ,  $X = \{A_{f_i} \mid f_i \in \mathcal{F}\}$ ,  $M = -\sum_{i \in I} f_i A_{f_i}$  and  $S = \sum_{n \geq 0} \sum_{f_{i_1}, \dots, f_{i_n} \in S} I(f_1, \dots, f_n; \tau) \cdot A_{f_1} \dots A_{f_n}$ . Note that it follows from (2.8) that

$$D(S) = M \cdot S,$$

as required in Theorem 2.1 of [5]. □

**Remark 2.10.** Variants of Theorem 2.9 have been known before (cf. [2], Lemma 3.6).

### 3. LINEAR INDEPENDENCE OF ITERATED INTEGRALS OF QUASIMODULAR FORMS

In this section, we apply Theorem 2.9 to deduce linear independence of a large family of iterated integrals of quasimodular forms. More precisely, our main result is the following theorem.

**Theorem 3.1.** *Let  $\mathcal{B}$  be a  $\mathbb{C}$ -linearly independent family of elements of  $\mathbb{C} \cdot E_2 \oplus M_*$ . Then the family of iterated integrals*

$$(I(f_1, \dots, f_n; \tau) \mid f_i \in \mathcal{B})$$

*is linearly independent over  $\text{Frac}(QM_*) \cong \mathbb{C}(E_2, E_4, E_6)$ .*

**3.1. Two auxiliary lemmas.** For the proof of Theorem 3.1, we need two lemmas.

**Lemma 3.2.** *Let  $f, g \in \mathbb{C}[E_2, E_4, E_6]$  such that  $g \neq 0$  and such that  $f$  and  $g$  are coprime. Assume that  $D\left(\frac{f}{g}\right) \in \mathbb{C}[E_2, E_4, E_6]$ . Then  $g = \beta \Delta^\alpha$  for some  $\alpha \in \mathbb{Z}_{\geq 0}$  and some  $\beta \in \mathbb{C} \setminus \{0\}$ , where  $\Delta := \frac{1}{1728}(E_4^3 - E_6^2)$  is Ramanujan's cusp form of weight 12.*

*Proof:* By the quotient rule, we have

$$D\left(\frac{f}{g}\right) = \frac{D(f)g - fD(g)}{g^2} = \frac{D(f) - f\frac{D(g)}{g}}{g}.$$

The left hand side is contained in  $\mathbb{C}[E_2, E_4, E_6]$  by assumption, and since also  $D(f)$  and  $g$  are in  $\mathbb{C}[E_2, E_4, E_6]$ , we have  $f\frac{D(g)}{g} \in \mathbb{C}[E_2, E_4, E_6]$ . But then, as  $f$  and  $g$  have no common factor,  $g$  must divide  $D(g)$ , i.e. there exists  $h \in \mathbb{C}[E_2, E_4, E_6]$  such that

$$D(g) = gh.$$

Since the operator  $D : QM_* \rightarrow QM_*$  is homogeneous of weight 2 (cf. Proposition 2.3.(i)), we have  $h \in QM_2$ , i.e.  $h = \alpha E_2$  with  $\alpha \in \mathbb{C}$ . In other words,  $g$  solves the differential equation  $D(g) = (\alpha E_2) \cdot g$ . But by Lemma 2.5,  $\alpha$  must be a non-negative integer and  $g = \beta \Delta^\alpha$  for some  $\beta \in \mathbb{C} \setminus \{0\}$ .  $\square$

**Lemma 3.3.** *Let  $f$  be a weakly quasimodular form, such that its derivative  $D(f)$  is a quasimodular form. Then  $f$  is a quasimodular form.*

*Proof:* It is no loss of generality to assume that  $f$  is of weight  $k \in \mathbb{Z}$  and depth  $\leq p$ , where  $p \geq 0$ . By the definition of weakly quasimodular forms (cf. also Remark 2.2), there exist uniquely determined meromorphic functions  $f_r(\tau)$ , for  $0 \leq r \leq p$ , such that

$$(f|_k \gamma)(\tau) = \sum_{r=0}^p f_r(\tau) X(\gamma)^r,$$

for all  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ . Therefore, we only need to show that every  $f_r(\tau)$  is holomorphic, including at the cusp.

To this end, by Proposition 2.3.(i), we know that

$$(3.1) \quad (D(f)|_{k+2} \gamma)(\tau) = \sum_{r=0}^{p+1} (D(f_r)(\tau) + (k-r+1)f_{r-1}(\tau)) X(\gamma)^r,$$

and since  $D(f)$  is a quasimodular form by assumption, every coefficient of (3.1) is holomorphic, including at the cusp.

The constant term, with respect to  $X(\gamma)$ , in (3.1) equals  $D(f_0)(\tau)$ , which is holomorphic by assumption. But a meromorphic function whose derivative is holomorphic everywhere is itself holomorphic everywhere. An easy induction argument, using that the coefficients of (3.1) are holomorphic, now shows that in fact every  $f_r(\tau)$  is holomorphic.  $\square$

**3.2. Proof of Theorem 3.1.** We will use the criterion of Theorem 2.9 in the case where  $\mathcal{C} = \mathrm{Frac}(QM_*)$  and  $\mathcal{F} = \mathcal{B}$ . Since  $\mathcal{B}$  is linearly independent over  $\mathbb{C}$  by assumption, it is enough to prove that if  $h \in \mathrm{Frac}(QM_*)$  then

$$D(h) = \sum_{f \in \mathcal{B}} \alpha_f f, \alpha_f \in \mathbb{C} \quad \Rightarrow \quad \alpha_f = 0, \text{ for all } f \in \mathcal{B}.$$

Also, since  $\mathcal{B}$  spans a subspace of  $\mathbb{C} \cdot E_2 \oplus M_*$ , it clearly suffices to prove that  $D(h) \in \mathbb{C} \cdot E_2 \oplus M_*$  implies that  $D(h) = 0$ , or equivalently that  $h$  is constant. Thus, the following proposition completes the proof of Theorem 3.1.

**Proposition 3.4.** *Let  $h \in \mathrm{Frac}(QM_*) \cong \mathbb{C}(E_2, E_4, E_6)$ , such that  $D(h) \in \mathbb{C} \cdot E_2 \oplus M_*$ . Then  $h$  is constant.*

*Proof:* Write  $h = \frac{f}{g}$  with  $f, g \in \mathbb{C}[E_2, E_4, E_6]$ ,  $g \neq 0$  and such that  $f$  and  $g$  are coprime. Writing  $f$  as a  $\mathbb{C}$ -linear combination of its homogeneous components, it is enough to show the proposition for  $f$  homogeneous of weight  $k_f$ .

First, we know from Lemma 3.2 that  $g = \beta \Delta^\alpha$  for some  $\alpha \in \mathbb{Z}_{\geq 0}$  and  $\beta \in \mathbb{C} \setminus \{0\}$ , where  $\Delta$  is Ramanujan's cusp form of weight 12. In particular,  $g$  is a cusp form of weight  $k_g = 12\alpha$ .



Since  $f$  is quasimodular of weight  $k_f$  and depth  $\leq p$ , there exist holomorphic (including at the cusp) functions  $f_r(\tau)$ , for  $0 \leq r \leq p$ , such that

$$(f|_{k_f}\gamma)(\tau) = \sum_{r=0}^p f_r(\tau)X(\gamma)^r,$$

for all  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ . Setting  $h_r(\tau) := \frac{f_r}{g}(\tau)$ , we also have, for  $k := k_f - k_g$

$$(h|_k\gamma)(\tau) = \sum_{r=0}^p h_r(\tau)X(\gamma)^r.$$

Moreover, the functions  $h_r(\tau)$  are meromorphic, thus,  $h$  is a weakly quasimodular form (of weight  $k$  and depth  $\leq p$ ). By assumption,  $D(h)$  is a quasimodular form (necessarily of weight  $k+2$  and depth  $\leq p+1$ ), and using Lemma 3.3, this implies that  $h \in QM_k^{\leq p}$ , therefore every  $h_r(\tau)$  is holomorphic, including at the cusp.

Summarizing, we have seen that  $h \in \mathrm{Frac}(QM_*)$  such that  $D(h) \in QM_*$  implies that  $h \in QM_*$ . But we even have  $D(h) \in \mathbb{C} \cdot E_2 \oplus M_*$  by assumption, and therefore Proposition 2.3.(ii) now implies that  $h$  is constant, as was to be shown.  $\square$

#### 4. ITERATED INTEGRALS OF QUASIMODULAR FORMS AND SHUFFLE ALGEBRAS

We describe the  $QM_*$ -algebra of iterated integrals of quasimodular forms, which is the smallest algebra, which contains  $QM_*$  and is closed under integration. Using the results of the last section, we show that it is canonically isomorphic to an explicit shuffle algebra. A similar result holds for the  $M_*$ -subalgebra of iterated integrals of modular forms.

##### 4.1. The algebra of iterated integrals of quasimodular forms.

**Definition 4.1.** Define  $\mathcal{I}^{QM}$  to be the  $QM_*$ -module generated by all iterated integrals of quasimodular forms:

$$\mathcal{I}^{QM} = \mathrm{Span}_{QM_*} \{I(f_1, \dots, f_n; \tau) \mid f_i \in QM_*\}.$$

We also denote by  $\mathcal{I}_n^{QM}$  the  $QM_*$ -linear submodule, which is spanned by all of the  $I(f_1, \dots, f_r; \tau)$  with  $r \leq n$ .

The subspaces  $\mathcal{I}_n^{QM}$  define an ascending filtration  $\mathcal{I}_\bullet^{QM}$  on  $\mathcal{I}^{QM}$ , called the length filtration (in analogy with the length filtration on iterated integrals [7]). It follows from (2.7) that  $\mathcal{I}^{QM}$  is a filtered  $QM_*$ -algebra. However, the length is not a grading, as shown by the next result.

**Proposition 4.2.** *Let  $f_1, \dots, f_n$  be quasimodular forms. Then*

$$I(f_1, \dots, f_{i-1}, D(f_i), f_{i+1}, \dots, f_n; \tau) \in \mathcal{I}_{n-1}^{QM}.$$

*Proof:* This is an immediate consequence of the integration by parts formula (2.9).  $\square$

**4.2.  $\mathcal{I}^{QM}$  as a shuffle algebra.** We let  $V$  be the  $\mathbb{C}$ -vector space  $\mathbb{C} \cdot E_2 \oplus M_*$ , and denote by  $\mathbb{C}\langle V \rangle$  the shuffle algebra on  $V$  (cf. Section 2.3). Recall that this is the graded dual of the tensor algebra  $T(V)$ , whose grading is given by the length of tensors. Elements of  $\mathbb{C}\langle V \rangle$  are  $\mathbb{C}$ -linear combination of the basic elements  $[f_1 | \dots | f_n]$ , and the product on  $\mathbb{C}\langle V \rangle$  is the shuffle product (2.5).

The following theorem is the main result of this paper.

**Theorem 4.3.** *The  $QM_*$ -linear map*

$$(4.1) \quad \begin{aligned} \varphi^{QM} : QM_* \otimes_{\mathbb{C}} \mathbb{C}\langle V \rangle &\rightarrow \mathcal{I}^{QM} \\ [f_1 | \dots | f_n] &\mapsto I(f_1, \dots, f_n; \tau) \end{aligned}$$

*is an isomorphism of  $QM_*$ -algebras.*

*Proof:* Let  $\mathcal{B}$  be a basis of  $V$ , so that the family  $([f_1 | \dots | f_n] \mid f_i \in \mathcal{B})$  is a basis of  $\mathbb{C}\langle V \rangle$ . The injectivity of  $\varphi^{QM}$  follows from the  $\text{Frac}(QM_*)$ -linear independence of the family

$$(4.2) \quad \mathcal{F} = (I(f_1, \dots, f_n; \tau) \mid f_i \in \mathcal{B}),$$

which is a consequence of Theorem 3.1.

In order to obtain the surjectivity, we need to prove that the family (4.2) generates  $\mathcal{I}^{QM}$ . To this end, we prove inductively that for every  $n \geq 0$ , we have  $\mathcal{I}_n^{QM} \subset \text{Span}_{QM_*} \mathcal{F}$ . The case  $n = 0$  is trivial. Now let  $n \geq 1$  and assume that for every  $r \leq n - 1$ , we have  $\mathcal{I}_r^{QM} \subset \text{Span}_{QM_*} \mathcal{F}$ . Given quasimodular forms  $f_1, \dots, f_n$ , we can write  $f_i = g_i + D(h_i)$ , where  $g_i \in \mathbb{C} \cdot E_2 \oplus M_*$  and  $h_i \in D(QM_*)$  by Proposition 2.3.(ii). Then by linearity

$$(4.3) \quad \begin{aligned} I(f_1, \dots, f_n; \tau) &= I(g_1, \dots, g_n; \tau) \\ &+ \sum_{i=1}^n I(g_1, \dots, g_{i-1}, D(h_i), g_{i+1}, \dots, g_n) + \dots, \end{aligned}$$

where the  $\dots$  above signifies iterated integrals, which have at least two  $D(h_i)$  as integrands. The first term on the right is contained in  $\text{Span}_{QM_*} \mathcal{F}$ , since  $g_i \in \mathbb{C} \cdot E_2 \oplus M_*$  for every  $i$  and  $\mathcal{B}$  is a basis. On the other hand, all other terms in the sum (4.3) are iterated integrals, which contain at least one  $D(h_i)$ . By Proposition 4.2, it thus follows that  $I(f_1, \dots, f_n; \tau) \equiv I(g_1, \dots, g_n; \tau) \pmod{\mathcal{I}_{n-1}^{QM}}$ , and we conclude using the induction hypothesis. Finally, it is clear that  $\varphi^{QM}$  is a homomorphism of algebras, since both sides of (4.1) are endowed with the shuffle product.  $\square$

**4.3. The algebra of iterated integrals of modular forms.** In this section, we study the subalgebra  $\mathcal{I}^M$  of  $\mathcal{I}^{QM}$ , generated by iterated integrals of modular forms.

**Definition 4.4.** Define  $\mathcal{I}^M$  to be the  $M_*$ -module generated by all iterated integrals of modular forms:

$$\mathcal{I}^M = \text{Span}_{M_*} \{I(f_1, \dots, f_n; \tau) \mid f_i \in M_*\}.$$

As in the case of  $\mathcal{I}^{QM}$ , the length of iterated integrals defines the length filtration  $\mathcal{I}_\bullet^M$  on  $\mathcal{I}^M$ , and  $\mathcal{I}^M$  is a filtered  $M_*$ -subalgebra of  $\mathcal{I}^{QM}$ . We let  $\mathbb{C}\langle M_* \rangle$  be the shuffle algebra on the  $\mathbb{C}$ -vector space  $M_*$ .

**Theorem 4.5.** *The  $M_*$ -linear map*

$$\begin{aligned} \varphi^M : M_* \otimes_{\mathbb{C}} \mathbb{C}\langle M_* \rangle &\rightarrow \mathcal{I}^M \\ [f_1 | \dots | f_n] &\mapsto I(f_1, \dots, f_n; \tau) \end{aligned}$$

*is an isomorphism of  $M_*$ -algebras.*

*Proof:* The morphism  $\varphi^M$  is surjective by definition. It is also injective, since for a basis  $\mathcal{B}_M$  of  $M_*$ , the iterated integrals  $I(f_1, \dots, f_n; \tau)$  with  $f_i \in \mathcal{B}_M$  are linearly independent over  $M_*$  by Theorem 3.1, as  $M_* \subset \text{Frac}(QM_*)$ .  $\square$

**4.4. A polynomial basis for  $\mathcal{I}^{QM}$ .** Recall from Proposition 2.3.(ii) that  $QM_*$  is isomorphic to the polynomial algebra  $\mathbb{C}[E_2, E_4, E_6]$ . A similar, but slightly more involved statement holds for the  $QM_*$ -algebra  $\mathcal{I}^{QM}$  of iterated integrals of quasimodular forms. Namely,  $\mathcal{I}^{QM}$  is a polynomial algebra over  $QM_*$  in infinitely many variables, which are given by certain Lyndon words.

In the following, if  $(S, <)$  is a totally ordered set, we will endow the free monoid  $S^*$  on  $S$  with the lexicographical order induced by  $<$ . Also, the *length* of  $w$  is simply the number of letters of  $w$ .

**Definition 4.6.** A *Lyndon word* on  $S^*$  is a non-trivial word,  $w \in S^* \setminus \{1\}$ , such that for all factorizations  $w = uv$  with  $u, v \neq 1$ , we have  $w < v$ . We denote by  $Lyn(S^*)$  the set of all Lyndon words on  $S^*$ .

**Example 4.7.** Let  $S = \{a, b\}$  with total order  $a < b$ . Then the Lyndon words on  $S^*$  of length at most four are

$$a, b, ab, aab, abb, aaab, aabb, abbb.$$

Now for a field  $k$  and any set  $S$ , define  $k\langle S \rangle$  to be the shuffle algebra on the free  $k$ -vector space generated by  $S$ . If  $k$  is of characteristic zero, then by the Milnor–Moore theorem [15],  $k\langle S \rangle$  is isomorphic to a polynomial algebra (in possibly infinitely many variables). The following refinement is due to Radford.

**Theorem 4.8** ([16]). *If  $k$  has characteristic zero, then  $k\langle S \rangle$  is freely generated, as a  $k$ -algebra, by the set of Lyndon words  $Lyn(S^*)$ . Equivalently,  $k\langle S \rangle \cong k[Lyn(S^*)]$ , the polynomial algebra on  $Lyn(S^*)$ .*

Returning to quasimodular forms, consider again the  $\mathbb{C}$ -vector space  $V = \mathbb{C} \cdot E_2 \oplus M_*$ , and let  $\mathcal{B} = \cup_{k \geq 0} \mathcal{B}_k$  be the homogeneous basis of  $V$ , given by  $\mathcal{B}_k = \{E_4^a E_6^b \mid 4a + 6b = k\}$  for  $k \neq 2$ , and  $\mathcal{B}_2 = \{E_2\}$ . The basis  $\mathcal{B}$  can be ordered for the lexicographical order as follows: if  $E_4^a E_6^b, E_4^{a'} E_6^{b'} \in \mathcal{B}_k$ , then

$$E_4^a E_6^b < E_4^{a'} E_6^{b'} \Leftrightarrow a < a', \text{ or } a = a', \text{ and } b < b',$$

and if  $f \in \mathcal{B}_k, g \in \mathcal{B}_{k'}$  with  $k < k'$ , then  $f < g$ .

Now, since for  $f_1, \dots, f_n \in \mathcal{B}$ , the iterated integrals  $I(f_1, \dots, f_n; \tau)$  are linearly independent over  $QM_*$  (by Theorem 3.1), we can canonically identify the set of all  $I(f_1, \dots, f_n; \tau)$  with the free monoid  $\mathcal{B}^*$ , and order  $\mathcal{B}^*$  for the lexicographical ordering induced from the order on  $\mathcal{B}$  above. The next result is a formal consequence of Theorems 4.3, 4.5 and 4.8.

**Theorem 4.9.** *The elements of  $Lyn(\mathcal{B}^*)$  are algebraically independent over  $QM_*$  and we have a natural isomorphism of  $QM_*$ -algebras*

$$QM_*[Lyn(\mathcal{B}^*)] \cong \mathcal{I}^{QM},$$

*which is filtered for the length, where the left hand side is the polynomial  $QM_*$ -algebra on  $Lyn(\mathcal{B}^*)$ . Explicitly, the isomorphism maps an element  $w = f_1 \dots f_n \in Lyn(\mathcal{B}^*)$  to the iterated integral  $I(f_1, \dots, f_n; \tau)$ . Similarly, we have a natural isomorphism of  $M_*$ -algebras*

$$M_*[Lyn(\mathcal{B}_M^*)] \cong \mathcal{I}^M,$$

where  $\mathcal{B}_M = \mathcal{B} \setminus \{E_2\}$ .

**Example 4.10.** The following table gives all elements of  $Lyn(\mathcal{B}^*)$  involving iterated integrals of length at most two of quasimodular forms of total weight at most 12. For ease of notation, we have dropped the  $\tau$  from  $I(f_1, \dots, f_n; \tau)$ .

Length \ Weight	0	1	2
0	—	$I(1)$	—
2	—	$I(E_2)$	—
4	—	$I(E_4)$	$I(1, E_4)$
6	—	$I(E_6)$	$I(1, E_6), I(E_2, E_4)$
8	—	$I(E_4^2)$	$I(1, E_4^2), I(E_2, E_6)$
10	—	$I(E_4 E_6)$	$I(1, E_4 E_6), I(E_2, E_4^2), I(E_4, E_6)$
12	—	$I(E_4^3), I(E_6^2)$	$I(1, E_4^3), I(1, E_6^2), I(E_2, E_4 E_6), I(E_4, E_4^2)$

Also, the list of all elements of  $Lyn(\mathcal{B}^*)$  consisting of iterated integrals of length at most three of quasimodular forms of total weight 12 is given by

$$\begin{aligned} & \{I(E_4^3), I(E_6^2), I(1, E_4^3), I(1, E_6^2), I(E_2, E_4 E_6), I(E_4, E_4^2), \\ & I(1, 1, E_4^3), I(1, 1, E_6^2), I(1, E_2, E_4 E_6), I(1, E_4, E_4^2), I(1, E_6, E_6), \\ & I(1, E_4^2, E_4), I(1, E_4 E_6, E_2), I(E_2, E_2, E_4^2), I(E_2, E_4, E_6), I(E_2, E_6, E_4)\}. \end{aligned}$$

#### APPENDIX A. EICHLER–SHIMURA FOR QUASIMODULAR FORMS

In this appendix, we show how one can attach one-cocycles to quasimodular forms. This extends the classical Eichler–Shimura theory of the cocycles attached to modular forms, and is probably well-known to the experts, but the author does not know of a suitable reference for the precise statements.

Throughout this appendix, we will freely use some elementary concepts from the cohomology of groups, for which we refer to [18], Ch. 6.

**A.1. Cocycles attached to modular forms.** We begin by briefly recalling how modular forms give rise to cocycles for  $SL_2(\mathbb{Z})$ . A standard reference is [11], Ch. VI.

For  $d \geq 0$ , let  $\mathbb{Q}[X, Y]_d$  be the  $\mathbb{Q}$ -vector space of homogeneous polynomials in  $X$  and  $Y$  of degree  $d$ . It is a right  $SL_2(\mathbb{Z})$ -module by defining

$$P(X, Y)|_\gamma = P(aX + bY, cX + dY), \quad \text{for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}), \quad P \in \mathbb{Q}[X, Y]_d.$$

With this action, given a modular form  $f$  of weight  $k \geq 2$ , it is straightforward to verify that the holomorphic differential one-form

$$\underline{f}(\tau) := (2\pi i)^{k-1} f(\tau)(X - \tau Y)^{k-2} d\tau \in \Omega^1(\mathfrak{H}) \otimes_{\mathbb{Q}} \mathbb{Q}[X, Y]_{k-2}$$

is  $SL_2(\mathbb{Z})$ -invariant, where  $SL_2(\mathbb{Z})$  acts on  $\mathfrak{H}$  in the usual way via fractional linear transformations. Fixing a base point  $\tau_0$  of  $\mathfrak{H}$  (possibly  $i\infty$ ), it follows from the  $SL_2(\mathbb{Z})$ -invariance that the function

$$r_{f, \tau_0} : SL_2(\mathbb{Z}) \rightarrow \mathbb{C}[X, Y]_{k-2}, \quad \gamma \mapsto \int_\tau^{\tau_0} \underline{f}(\tau) - \left( \int_{\gamma \cdot \tau}^{\tau_0} \underline{f}(\tau) \right) \Big|_\gamma,$$

(regularized as in Section 2.3, if  $\tau_0 = i\infty$ ) is a one-cocycle, i.e. it satisfies  $r_{f, \tau_0}(\gamma_1 \gamma_2) = r_{f, \tau_0}(\gamma_1)|_{\gamma_2} + r_{f, \tau_0}(\gamma_2)$  for all  $\gamma_1, \gamma_2 \in SL_2(\mathbb{Z})$ . Its cohomology class does not depend on  $\tau_0$ , and we denote this class simply by  $[r_f]$ .

The same construction can also be applied to the complex conjugate  $\overline{f(\tau)} := (-2\pi i)^{k-1} \overline{f(\tau)} (X - \bar{\tau}Y)^{k-2} d\bar{\tau}$  of the one-form  $\underline{f}(\tau)$ , and we denote by  $[r_{\bar{f}}]$  the resulting cohomology class.

**Theorem A.1** (Eichler–Shimura). *For every  $k \geq 2$ , the morphism*

$$\begin{aligned} M_k \oplus \overline{S}_k &\rightarrow H^1(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X, Y]_{k-2}) \otimes_{\mathbb{Q}} \mathbb{C}, \\ (f, \bar{g}) &\mapsto [r_f] + [r_{\bar{g}}], \end{aligned}$$

is an isomorphism of  $\mathbb{C}$ -vector spaces. Here,  $\overline{S}_k$  denotes the complex conjugate of the  $\mathbb{C}$ -vector space of cusp forms of weight  $k$ .

**A.2. Cocycles for the braid group.** The fact that  $r_f$  is a cocycle hinges on the modularity of  $f$ . In order to incorporate quasimodular forms into the picture, we need to consider instead of  $\mathrm{SL}_2(\mathbb{Z})$  the braid group  $B_3 = \langle \sigma_1, \sigma_2 : \sigma_1 \sigma_2 \sigma_1 = \sigma_2 \sigma_1 \sigma_2 \rangle$  on three strands. It is a central extension

$$(A.1) \quad 1 \longrightarrow \mathbb{Z} \longrightarrow B_3 \longrightarrow \mathrm{SL}_2(\mathbb{Z}) \longrightarrow 1,$$

and also the fundamental group of the quotient of  $\mathbb{C}^\times \times \mathfrak{H}$  by the  $\mathrm{SL}_2(\mathbb{Z})$ -action

$$\gamma \cdot (z, \tau) = ((c\tau + d)z, \gamma \cdot \tau), \quad \text{for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

where  $\mathrm{SL}_2(\mathbb{Z})$  acts on  $\mathfrak{H}$  as before. We refer to [8], §8, for more details and further equivalent descriptions of  $B_3$ .

Next, we compute the cohomology groups  $H^1(B_3, \mathbb{Q}[X, Y]_d)$ , where  $B_3$  acts on  $\mathbb{Q}[X, Y]_d$  via the projection  $B_3 \rightarrow \mathrm{SL}_2(\mathbb{Z})$ .

**Proposition A.2.** *We have canonical isomorphisms*

$$H^1(B_3, \mathbb{Q}[X, Y]_d) \cong \begin{cases} H^1(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X, Y]_d), & \text{for } d \geq 1, \\ \mathbb{Q}, & \text{for } d = 0. \end{cases}$$

*Proof:* The Hochschild–Serre spectral sequence ([18], Ch. 6.8.3) associated to the extension (A.1) yields an exact sequence

$$0 \rightarrow H^1(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X, Y]_d) \rightarrow H^1(B_3, \mathbb{Q}[X, Y]_d) \rightarrow H^1(\mathbb{Z}, \mathbb{Q}[X, Y]_d)^{\mathrm{SL}_2(\mathbb{Z})} \rightarrow 0,$$

where we have used that  $H^2(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X, Y]_d) = \{0\}$ , as  $\mathrm{SL}_2(\mathbb{Z})$  has virtual cohomological dimension equal to one. The proposition now follows easily from this.  $\square$

**A.3. Quasimodular forms and braid group cocycles.** In light of Theorem A.1, Proposition A.2 suggests to attach a one-cocycle  $B_3 \rightarrow \mathbb{C}$  to the Eisenstein series  $E_2$ . Indeed, this can be done as follows.

First, the modular transformation property of  $E_2$  (2.2) implies that the differential one-form

$$(A.2) \quad 2\pi i E_2(\tau) d\tau - 12 \frac{dz}{z} \in \Omega^1(\mathbb{C}^\times \times \mathfrak{H})$$

is  $\mathrm{SL}_2(\mathbb{Z})$ -invariant, i.e. it descends to the quotient  $\mathrm{SL}_2(\mathbb{Z}) \backslash (\mathbb{C}^\times \times \mathfrak{H})$ . Denote by

$$\underline{E}_2(\xi, \tau) := \varphi^* \left( 2\pi i E_2(\tau) d\tau - 12 \frac{dz}{z} \right) = 2\pi i E_2(\tau) d\tau - 12 d\xi \in \Omega^1(\mathbb{C} \times \mathfrak{H}),$$

the pull-back of (A.2) along the universal covering map  $\varphi : \mathbb{C} \times \mathfrak{H} \rightarrow \mathrm{SL}_2(\mathbb{Z}) \backslash (\mathbb{C}^\times \times \mathfrak{H})$ . Clearly,  $\underline{E}_2(\xi, \tau)$  is  $B_3$ -invariant and it follows that for any base point  $(\xi_0, \tau_0)$  (for example  $(\xi_0, \tau_0) = (0, i\infty)$ ), the function

$$r_{E_2, (\xi_0, \tau_0)} : B_3 \rightarrow \mathbb{C}$$

$$\gamma \mapsto \int_{(\xi, \tau)}^{(\xi_0, \tau_0)} \underline{E}_2(\xi, \tau) - \left( \int_{\gamma \cdot (\xi, \tau)}^{(\xi_0, \tau_0)} \underline{E}_2(\xi, \tau) \right) \Big|_\gamma$$

is a well-defined cocycle (again, regularization is needed if  $\tau_0 = i\infty$ ).

**Remark A.3.** The integral  $I(E_2; \tau)$  introduced in Section 2.3 is actually equal to  $\int_\tau^{i\infty} \underline{E}_2(\xi, \tau)$ , where we embed  $\mathfrak{H}$  into  $\mathbb{C} \times \mathfrak{H}$  by  $\tau \mapsto (0, \tau)$ . However, that embedding is not  $B_3$ -equivariant, and indeed the integral  $I(E_2; \tau)$  does not give rise to a cocycle for  $B_3$ ; for this, one really needs to lift the form  $2\pi i E_2(\tau) d\tau$  to the form  $\underline{E}_2(\xi, \tau)$ .

Now since the cocycle  $r_{E_2, (\xi_0, \tau_0)}$  is non-zero, its cohomology class (which is again independent of the choice of base point  $(\xi_0, \tau_0)$ ) is non-trivial. The Eichler–Shimura theorem (Theorem A.1) together with Proposition A.2 then implies the next result.

**Corollary A.4.** *For every  $k \geq 2$ , the morphism*

$$V_k \oplus \overline{\mathcal{S}}_k \rightarrow H^1(B_3, \mathbb{Q}[X, Y]_{k-2}) \otimes_{\mathbb{Q}} \mathbb{C},$$

$$(f, \overline{g}) \mapsto [r_f] + [r_{\overline{g}}],$$

where  $V := M_* \oplus \mathbb{C} \cdot E_2$ , is an isomorphism of  $\mathbb{C}$ -vector spaces.

One can also attach a cocycle  $r_{f, \tau_0}$  to a general quasimodular form  $f \in QM_k$  of weight  $k$  as follows. By Proposition 2.3.(ii), we know that  $f$  can be written uniquely as a  $\mathbb{C}$ -linear combination of derivatives of modular forms and of derivatives of  $E_2$ . Thus, we can write

$$f = \sum \lambda_g \cdot D^{p_g}(g), \quad \lambda_g \in \mathbb{C}, \quad p_g \geq 0,$$

where  $g$  is either a modular form of weight  $k - 2p_g$  or  $g = E_2$ . Therefore, we may define  $r_{f, \tau_0} : B_3 \rightarrow \mathbb{C}[X, Y]_{\leq k-2} := \bigoplus_{0 \leq d \leq k-2} \mathbb{C}[X, Y]_d$  by

$$r_{f, \tau_0} := \sum \lambda_g \cdot r_{g, \tau_0}.$$

Using this definition, one sees in particular that the cocycles of quasimodular forms can be expressed in terms of the cocycles attached to modular forms and to  $E_2$ . This is of course in line with Corollary A.4.

**Remark A.5.** In [3, 9, 13], certain non-abelian  $\mathrm{SL}_2(\mathbb{Z})$ -cocycles given in terms of iterated integrals of modular forms are studied. It would be natural to try and extend this theory to non-abelian  $B_3$ -cocycles attached to iterated integrals of quasimodular forms (perhaps along the lines suggested in [9], §14), but this is beyond the scope of the present paper.

## REFERENCES

- [1] S. Bloch and A. Okounkov. The character of the infinite wedge representation. *Adv. Math.*, 149(1):1–60, 2000.
- [2] F. Brown. Multiple zeta values and periods of moduli spaces  $\overline{\mathcal{M}}_{0,n}$ . *Ann. Sci. Éc. Norm. Supér. (4)*, 42(3):371–489, 2009.
- [3] F. Brown. Multiple modular values and the relative completion of the fundamental group of  $\mathcal{M}_{1,1}$ . arXiv:1407.5167v3, 2016.

- [4] F. Brown. A class of non-holomorphic modular forms II: Equivariant iterated eisenstein integrals. arXiv:1708.03354, 2017.
- [5] M. Deneufchâtel, G. H. E. Duchamp, V. H. N. Minh, and A. I. Solomon. Independence of hyperlogarithms over function fields via algebraic combinatorics. In *Algebraic informatics*, volume 6742 of *Lecture Notes in Comput. Sci.*, pages 127–139. Springer, Heidelberg, 2011.
- [6] M. Eichler. Eine Verallgemeinerung der Abelschen Integrale. *Math. Z.*, 67:267–298, 1957.
- [7] R. Hain. The geometry of the mixed Hodge structure on the fundamental group. In *Algebraic geometry, Bowdoin, 1985 (Brunswick, Maine, 1985)*, volume 46 of *Proc. Sympos. Pure Math.*, pages 247–282. Amer. Math. Soc., Providence, RI, 1987.
- [8] R. Hain. Lectures on moduli spaces of elliptic curves. In *Transformation groups and moduli spaces of curves*, volume 16 of *Adv. Lect. Math. (ALM)*, pages 95–166. Int. Press, Somerville, MA, 2011.
- [9] R. Hain. The Hodge–de Rham theory of modular groups. In *Recent advances in Hodge theory*, volume 427 of *London Math. Soc. Lecture Note Ser.*, pages 422–514. Cambridge Univ. Press, Cambridge, 2016.
- [10] M. Kaneko and D. Zagier. A generalized Jacobi theta function and quasimodular forms. In *The moduli space of curves (Texel Island, 1994)*, volume 129 of *Progr. Math.*, pages 165–172. Birkhäuser Boston, Boston, MA, 1995.
- [11] S. Lang. *Introduction to modular forms*. Springer-Verlag, Berlin-New York, 1976. Grundlehren der mathematischen Wissenschaften, No. 222.
- [12] P. Lochak, N. Matthes, and L. Schneps. Elliptic multiple zeta values and the elliptic double shuffle relations. arXiv:1703.09410, 2017.
- [13] Y. I. Manin. Iterated integrals of modular forms and noncommutative modular symbols. In *Algebraic geometry and number theory*, volume 253 of *Progr. Math.*, pages 565–597. Birkhäuser Boston, Boston, MA, 2006.
- [14] F. Martin and E. Royer. Formes modulaires et périodes. In *Formes modulaires et transcendance*, volume 12 of *Sémin. Congr.*, pages 1–117. Soc. Math. France, Paris, 2005.
- [15] J. W. Milnor and J. C. Moore. On the structure of Hopf algebras. *Ann. of Math. (2)*, 81:211–264, 1965.
- [16] D. E. Radford. A natural ring basis for the shuffle algebra and an application to group schemes. *J. Algebra*, 58(2):432–454, 1979.
- [17] C. Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.
- [18] C. A. Weibel. *An introduction to homological algebra*, volume 38 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1994.
- [19] D. Zagier. Elliptic modular forms and their applications. In *The 1-2-3 of modular forms*, Universitext, pages 1–103. Springer, Berlin, 2008.

MAX-PLANCK-INSTITUT FÜR MATHEMATIK, VIVATSGASSE 7, D-53111, BONN, GERMANY  
E-mail address: nilsmath@mpim-bonn.mpg.de