RELATIONS BETWEEN DERIVATIONS ARISING FROM MODULAR FORMS

AARON POLLACK

Thesis submitted to Department of Mathematics
Duke University, Durham, NC 2009

Acknowledgments It is a pleasure to thank Professor R. Hain for his invaluable help on this project, and otherwise. I am deeply grateful to him for generously giving me such a fruitful problem, and for his countless hours of patient tutelage and guidance. Additionally, he has carefully read this manuscript, and made numerous useful suggestions to improve the exposition. Thanks also to Professor D. Kraines and Duke's PRUV program for supporting this research in Summers 2007 and 2008.

Abstract. Denote by $\mathbb{L}(a, b)$ the free complex Lie algebra on the two generators a and b. For each integer $m \geq 0$ there is a derivation ϵ_{2m} on $\mathbb{L}(a, b)$ that satisfies $\epsilon_{2m}([a, b]) = 0$ and $\epsilon_{2m}(a) = ad(a)^{2m}(b)$. In this paper we study the derivation subalgebra $\mathfrak{u}$ generated by the ϵ_{2m}. In particular, we study the relations between the ϵ_{2m} and find that these relations are related to the period polynomials of modular forms.

Contents

1. Introduction 3
2. Period polynomials of modular forms 8
3. Derivations and Trees 9
3.1. Rooted trees and $\mathbb{L}(H)$ 9
3.2. Non-rooted trees and derivations 10
3.3. A better understanding of $\mathcal{A}$ as derivations 12
3.4. Specialization to $\mathfrak{u}$ 15
4. Proofs of Main Theorems 17
4.1. The filtration Θ^k 17
4.2. Proof of the Main Result 19
5. Remarks, Examples, and Galois action computations 20
5.1. Remarks 21
5.2. Examples 21
5.3. Computations of $\tilde{z}_j$ actions 23
References 24

1. Introduction

Set $H = \mathbb{C}a \oplus \mathbb{C}b$ and denote by $\mathbb{L}(H)$ the free Lie algebra on H, or equivalently the free Lie algebra on the generators a and b. The Lie algebra of derivations $\mathrm{Der}\,\mathbb{L}(H)$ on $\mathbb{L}(H)$ contains a set of distinguished derivations ϵ_{2m}, m a nonnegative integer, that satisfy

$$\epsilon_{2m}([a,b]) = 0$$

and

$$\epsilon_{2m}(a) = ad(a)^{2m}(b).$$

When $m = 0$, ϵ_0 is defined by $\epsilon_0(a) = b$ and $\epsilon_0(b) = 0$. When $m > 0$

$$\epsilon_{2m}(a) = a^{2m} \cdot b \tag{1}$$

and

$$\epsilon_{2m}(b) = [b, a^{2m-1} \cdot b] + \sum_{1 \le j < m} (-1)^j [a^j \cdot b, a^{2m-1-j} \cdot b] \tag{2}$$

with the empty sum for $m = 1$ being interpreted as 0. Here and throughout, we will use the shorthand $(u_1 \dots u_r) \cdot x$ to denote $ad(u_1) \dots ad(u_r)(x) = [u_1, [\dots [u_r, x] \dots]]$.

Denote by $\mathfrak{u}$ the Lie subalgebra of $\mathrm{Der}\,\mathbb{L}(H)$ generated by the ϵ_{2m}. The problem attacked in this paper is

Problem (Rough Version)**.** Find a presentation for $\mathfrak{u}$, or equivalently all the Lie algebraic relations between the ϵ_{2m}.

We will define a bigrading $(*,*) = (\text{depth}, \text{weight})$ on $\mathbb{L}(H)$, which induces one on $\mathrm{Der}\,\mathbb{L}(H)$ and $\mathfrak{u}$. Give a depth 0, weight 1, and b depth 1, weight 1. Extend to $\mathbb{L}(H)$ by the Lie bracket. That is, a is defined to be of type $(0,1)$, b to be of type $(1,1)$, and if $u, v \in \mathbb{L}(H)$ are of types (d,w), (d',w'), then $[u,v]$ has type $(d+d', w+w')$. Thus one has the decomposition $\mathbb{L}(H) = \bigoplus_{d,w \ge 0} \mathbb{L}_{d,w}(H)$ as a direct sum of its depth d, weight w parts. The bigrading on $\mathbb{L}(H)$ induces one on $\mathrm{Der}\,\mathbb{L}(H)$ in the usual manner[1]:

$$\mathrm{Der}\,\mathbb{L}_{d,w}(H) := \{\delta \in \mathrm{Der}\,\mathbb{L}(H) : \delta(x) \in \mathbb{L}_{d+d',w+w'}(H) \text{ for every } x \in \mathbb{L}_{d',w'}(H)\}.$$

If $\delta \in \mathrm{Der}\,\mathbb{L}_{d,w}(H)$ and $\delta' \in \mathrm{Der}\,\mathbb{L}_{d',w'}(H)$, then $[\delta, \delta']$ is in $\mathrm{Der}\,\mathbb{L}_{d+d',w+w'}(H)$. The equations (1), (2) imply that the derivation ϵ_{2m} is of type $(1, 2m)$. As a consequence, this bigrading passes to $\mathfrak{u}$. A commutator of d $\epsilon's$ is thus of type $(d,*)$. So, the depth of an element of $\mathfrak{u}$ coincides with its commutator length. Since depth produces a grading on $\mathfrak{u}$, the relations are also graded by depth.

Equations (1) and (2) imply that ϵ_2 is the derivation $-ad([a,b])$. Since the ϵ_{2m} annihilate $[a,b]$, one has

$$[\epsilon_{2m}, \epsilon_2] = 0. \tag{3}$$

[1] We will frequently define gradings or filtrations on H or $\mathbb{L}(H)$ and then extend them to $\mathrm{Der}\,\mathbb{L}(H)$ (or a subalgebra) in this way. In the sequel we will just refer to this extension as "the usual manner".

This relation is atypical in that it is extremely simple and holds for all m. More typical examples are the depth 2 relations:

$$[\epsilon_{10}, \epsilon_4] - 3[\epsilon_8, \epsilon_6] = 0 \tag{4}$$

and

$$2[\epsilon_{14}, \epsilon_4] - 7[\epsilon_{12}, \epsilon_6] + 11[\epsilon_{10}, \epsilon_8] = 0. \tag{5}$$

The simplest depth 3 relation is

$$\begin{aligned} &80[\epsilon_{12}, [\epsilon_4, \epsilon_0]] + 16[\epsilon_4, [\epsilon_{12}, \epsilon_0]] - 250[\epsilon_{10}, [\epsilon_6, \epsilon_0]] - 125[\epsilon_6, [\epsilon_{10}, \epsilon_0]] + \\ &280[\epsilon_8, [\epsilon_8, \epsilon_0]] - 462[\epsilon_4, [\epsilon_4, \epsilon_8]] - 1725[\epsilon_6, [\epsilon_6, \epsilon_4]] = 0. \end{aligned} \tag{6}$$

These relations are of type $(2, 2m+2), (2, 14), (2, 18)$ and $(3, 16)$, respectively.

The coefficients appearing in the relations above have number theoretic significance; they are periods of modular forms for $SL_2(\mathbb{Z})$.[2] To fix notation, denote

$$r_f(X, Y) = \sum a_f(k) X^{n-2-k} Y^k$$

the period polynomial of the modular form f of weight n. Recall that when f is a cusp form

$$r_f(X, Y) = \int_0^{i\infty} f(\tau)(X - \tau Y)^{n-2} d\tau.$$

Zagier [6] has extended the definition of r_f to all modular forms. Denote by

$$r_f^+(X, Y) = \frac{1}{2}(r_f(X, Y) + r_f(X, -Y))$$

the even degree part of the period polynomial r_f and

$$r_f^- = \frac{1}{2}(r_f(X, Y) - r_f(X, -Y))$$

the odd degree part, so that $r_f = r_f^+ + r_f^-$. The main theorem of this paper is a result on the relations between the ϵ_{2m} of depth d for every $d \geq 2$, and the depth 2 case implies

Theorem 1. *For n a fixed positive even integer,*

$$\sum_{p+q=n+2} \beta(p, q)[\epsilon_p, \epsilon_q] = 0$$

if and only if

$$\sum \beta(p, q)(X^{p-2} Y^{q-2} - X^{q-2} Y^{p-2}) = r_f^+(X, Y)$$

for some modular form f of weight n.

---

[2] Every modular form from now on will be assumed to be for the full modular group $SL_2(\mathbb{Z})$.

For example, the relation (3) corresponds to the Eisenstein series of weight $2m$, as $r^+_{G_{2m}}$ is proportional to $X^{2m-2} - Y^{2m-2}$. A linear combination of G_{12} and Ramanujan's Δ gives a modular form f with

$$r^+_f(X,Y) = X^8Y^2 - X^2Y^8 - 3(X^6Y^4 - X^4Y^6),$$

implying (4). Similarly, there is a basis $\{G_{12}, g\}$ of the weight 16 modular forms with

$$r^+_g(X,Y) = 2(X^{12}Y^2 - X^2Y^{12}) - 7(X^{10}Y^4 - X^4Y^{10}) + 11(X^8Y^6 - X^6Y^8),$$

so that g corresponds to (5).

As is discussed in detail below, there is a notion of "indecomposable relation": the "indecomposable" part of a relation r is its class $[r]$ in the quotient space $\{\text{relations}\}/\{\text{consequences of relations}\}$. A consequence of a relation v, for example $v = [\epsilon_{10}, \epsilon_4] - 3[\epsilon_8, \epsilon_6]$, is a relation of the form $[u, v]$. Work of Richard Hain and Makoto Matsumoto on Mixed Elliptic Motives implies that each cusp form f of (modular form) weight n for the full modular group $SL_2(\mathbb{Z})$ determines an indecomposable relation of depth d and weight $n+2d-2$, for every $d \geq 2$. For example, the space of relations proportional to 36 times (3) with $2m = 12$ minus 691 times (4) comes from the 1 dimensional space of cusp forms of weight 12; similarly, (5) is essentially due to the 1 dimensional space of cusp forms of weight 16. The relation (6) corresponds to Δ.

To see how (6) relates to Δ, note that dividing (6) by 40 gives

$$\begin{aligned} &4\{\tfrac{[\epsilon_{12},[\epsilon_4,\epsilon_0]]}{2} + \tfrac{[\epsilon_4,[\epsilon_{12},\epsilon_0]]}{10}\} - 25\{\tfrac{[\epsilon_{10},[\epsilon_6,\epsilon_0]]}{4} + \tfrac{[\epsilon_6,[\epsilon_{10},\epsilon_0]]}{8}\} + 42\{\tfrac{[\epsilon_8,[\epsilon_8,\epsilon_0]]}{6}\} \\ &- \tfrac{231}{20}[\epsilon_4,[\epsilon_4,\epsilon_8]] - \tfrac{345}{8}[\epsilon_6,[\epsilon_6,\epsilon_4]] = 0. \end{aligned} \tag{7}$$

The *odd degree part* of the period polynomial of Δ is proportional to

$$4(X^9Y + XY^9) - 25(X^7Y^3 + X^3Y^7) + 42(X^5Y^5).$$

So, the terms in this relation that contain ϵ_0 correspond to terms of r^-_Δ.

Main Theorem (Rough Version)**.** *If v is a relation between the ϵ_{2m}'s of depth d and weight $n+2d-2$, then the coefficients of the terms in v with $(d-2)$ ϵ_0's correspond to terms of r^+_f (resp. r^-_f) when d is even (resp. odd).*

In the rest of the introduction we give a precise version of the problem and the Main Theorem.

First, we will describe an $\mathfrak{sl}_2$ action on $\operatorname{Der}\mathbb{L}(H)$. Identify $H = \mathbb{C}a \oplus \mathbb{C}b$ with the standard representation of $\mathfrak{sl}_2$, with basis a, b. So, the matrix

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

in $\mathfrak{sl}_2$ acts on H by taking a to $pa + rb$ and b to $qa + sb$. Thus $\mathfrak{sl}_2$ is contained in the weight 0 part of $\operatorname{Der}\mathbb{L}(H)$, so this action of $\mathfrak{sl}_2$ on H extends to an action on $\mathbb{L}(H)$ and also on $\operatorname{Der}\mathbb{L}(H)$.

In defining the $\mathfrak{sl}_2$ action on H as described above, we have chosen a Cartan under which a has $\mathfrak{sl}_2$-weight 1 and b has $\mathfrak{sl}_2$-weight -1. This $\mathfrak{sl}_2$-weight defines a grading

$H = H_1 \oplus H_{-1}$ on H. When we extend this grading to $\mathbb{L}(H)$ and $\operatorname{Der}\mathbb{L}(H)$ in the usual manner, ϵ_{2m} has $\mathfrak{sl}_2$-weight $2m-2$. Contrast the $\mathfrak{sl}_2$-weight to the weight defined earlier, which we now call the *motivic weight*[3], or just m-weight for short. In general, one has

$$\mathfrak{sl}_2\text{-weight} = m\text{-weight } - 2 \text{ depth.} \tag{8}$$

Under the $\mathfrak{sl}_2$ action, ϵ_{2m} is a highest weight vector when $m > 0$. That is, if r denotes the derivation $a\frac{\partial}{\partial b}$, i.e. $r(b) = a$ and $r(a) = 0$, then $[r, \epsilon_{2m}] = 0$. Also note that ϵ_0 is the derivation $b\frac{\partial}{\partial a}$, and thus $\mathfrak{u}$ is an $\mathfrak{sl}_2$ submodule of $\operatorname{Der}\mathbb{L}(H)$.

The derivations ϵ_{2m+2} and ϵ_0 generate a copy of $S^{2m}H$, the $2m^{\text{th}}$ symmetric power of H, in $\operatorname{Der}\mathbb{L}(H)$. Furthermore, if H is given the basis $\{x, y\}$, the standard basis $\{x^k y^{2m-k}\}$ of $S^{2m}H$ corresponds to the derivations

$$\mu_{2m}^k = \frac{1}{\binom{2m}{k}} \frac{{\epsilon_0}^{2m-k}}{(2m-k)!} \cdot \epsilon_{2m+2}.$$

The derivation μ_{2m}^k is of type $(2m-k+1, 2m+2)$. Define $P = \bigoplus_{m\geq 0} S^{2m}H$, and write $\{x_{2m}^k\}$ for the standard basis of the copy of $S^{2m}H$ inside P. Put a bigrading $(*,*)$ on $\mathbb{L}(P)$, the free Lie algebra on P, by defining x_{2m}^k to be of type $(2m-k+1, 2m+2)$. We now have an $\mathfrak{sl}_2$-equivariant Lie algebra homomorphism

$$\Phi : \mathbb{L}(P) \to \mathfrak{u}$$

defined by $\Phi(x_{2m}^k) = \mu_{2m}^k$ and extending by the Lie bracket. By the statement "finding all the relations between the ϵ_{2m}'s" we mean "determining $\ker\Phi$". Since Φ preserves the bigrading, $\ker\Phi$ is bigraded.

Since Φ is $\mathfrak{sl}_2$ equivariant, to describe $\ker\Phi$ it is enough to describe the space of highest weight vectors of $\ker\Phi$, which we denote by K. Since the operator $r = a\frac{\partial}{\partial b}$ maps $\mathbb{L}_{d,N}(P) \to \mathbb{L}_{d-1,N}(P)$, K is also bigraded $K = \bigoplus K_{d,N}$.

Previously we mentioned that indecomposable relations are relations modulo their consequences. Consider the free Lie algebra $\mathfrak{k} = \mathbb{L}(z_3, z_5, z_7, z_9, \ldots)$ where the generator z_j has depth j and m-weight $2j$. Hain and Matsumoto show [3] that the action of the absolute Galois group on the unipotent fundamental group of a smoothing of the nodal cubic induces a homomorphism $\mathfrak{k} \to \operatorname{Der}\mathbb{L}(H)$ that preserves depth and motivic weight. Furthermore, the image $\mathfrak{k} \to \operatorname{Der}\mathbb{L}(H)$ is $\mathfrak{sl}_2$ invariant and normalizes $\mathfrak{u}$. Denote the image of z_j by $\tilde{z}_j$. What we really wish to do is to determine the $\tilde{z}_j$ and quotient out the relations by their action. To this end, set

$$\tilde{P} = P \oplus \mathbb{C}z_3 \oplus \mathbb{C}z_5 \oplus \mathbb{C}z_7 \oplus \ldots.$$

The map Φ extends to $\tilde{\Phi} : L(\tilde{P}) \to \operatorname{Der}\mathbb{L}(H)$. Write $\tilde{K}$ for the space of highest weight vectors in $\ker\tilde{\Phi}$. Denote by $\mathcal{K}$ the quotient $\tilde{K}/[\mathbb{L}(\tilde{P}), \tilde{K}]$. The space $\mathcal{K}$ is again bigraded and is called the space of indecomposable relations. The results of Hain and Matsumoto imply that each modular form (resp. cusp form) f of weight

[3]The author is entirely ignorant of motives, and they will make no explicit appearance in this work, although they are lurking in the background.

n determines an element of $\mathcal{K}$ of depth d and motivic weight $n+2d-2$ for every $d>1$ when d is even (resp. odd).

Thus, the precise version of the problem studied in this paper is

Problem (Precise Version)**.** Determine the image of the elements $z_3, z_5, \dots$ in $\operatorname{Der}\mathbb{L}(H)$ and explicitly compute the space $\mathcal{K}_{d,N}$, for each $d>1$.

By no means is this problem solved in entirety. Determining the action of $\mathfrak{k}$ appears to be a difficult problem, and we only remark on it in the final section. In this paper we focus on determining $K_{d,N}$, and even then we only work modulo $\mathbb{L}^3(P) := [\mathbb{L}(P), [\mathbb{L}(P), \mathbb{L}(P)]]$ and a corresponding ideal $\Theta^3\mathfrak{u}$ of $\mathfrak{u}$ (defined below). As we show in §4 Proposition 4.1, the highest $\mathfrak{sl}_2$-weight vectors of $\mathbb{L}_{d,N}(P)/\mathbb{L}^3_{d,N}(P)$ are of the form

$$\sum \alpha(p,q) h^d_{p,q},$$

where

$$h^d_{p,q} = \sum_{i+j=d-2} (-1)^i \binom{d-2}{i} [x_p^{p-i}, x_q^{q-j}]$$

and $p+q=N-4$. Thus, if $v \in K_{d,N}$, then $v \equiv \sum \alpha(p,q) h^d_{p,q} \bmod \mathbb{L}^3(P)$ for some $\alpha(p,q)$. The main theorem of this paper says that the $\alpha(p,q)$ are periods of modular forms.

Theorem 2. *If $v \in K_{d,N}$, then*

$$v \equiv \sum_{p+q=N-4} a_f(p-d+2) h^d_{p,q} \ \mathit{mod}\ \mathbb{L}^3(P)$$

for some modular form f of weight $N-2d+2$.

In fact, we prove a stronger result. Define a filtration Θ^k on $\mathbb{L}(H)$ as follows. Set $\Theta^0\mathbb{L}(H) = \mathbb{L}(H)$, $\Theta^1\mathbb{L}(H) = \mathbb{L}(H)^2$, and if $k>1$ define

$$\Theta^k\mathbb{L}(H) = [\Theta^1\mathbb{L}(H), \Theta^{k-1}\mathbb{L}(H)].$$

The filtration Θ on $\mathbb{L}(H)$ induces one $\operatorname{Der}\mathbb{L}(H)$ and $\mathfrak{u}$ in the standard manner. If $\delta_1 \in \Theta^j\mathbb{L}(H)$ and $\delta_2 \in \Theta^k\mathbb{L}(H)$, then $[\delta_1, \delta_2] \in \Theta^{j+k}\mathbb{L}(H)$. In §4 we show that $\Phi(\mathbb{L}^k(P)) \subset \Theta^k\mathfrak{u}$. Thus, the quotient map

$$\bar{\Phi} : \mathbb{L}(P)/\mathbb{L}^3(P) \to \mathfrak{u}/\Theta^3\mathfrak{u}$$

is well-defined. Denote by $\bar{K}$ the space of highest weight vectors of $\ker\bar{\Phi}$. This space is again bigraded: $\bar{K} = \bigoplus \bar{K}_{d,N}$. The main theorem of this paper completely determines $\bar{K}_{d,N}$.

Theorem 3. *Identify $h^d_{p,q}$ with its image in $\mathbb{L}(P)/\mathbb{L}^3(P)$. Then $v \in \bar{K}_{d,N}$ if and only if*

$$v = \sum_{p+q=N-4} a_f(p-d+2) h^d_{p,q}$$

for some modular form f of weight $N-2d+2$.

A couple comments are now in order. First, Theorem 3 of course implies Theorem 2, but is stronger because it is an if and only if statement. Moreover, Theorem 3 completely determines $K(2,N)$, as $K_{2,N} = \bar{K}_{2,N}$. Thus, Theorem 3 implies Theorem 1. Additionally, it is psychologically important to have the stronger result Theorem 3, and not just the weaker Theorem 2. The reason is that if we had some $v = \sum a_f(p-d+2)h^d_{p,q}$ that did not vanish modulo $\Theta^3\mathfrak{u}$, it would be impossible for v to be extended to a bona fide element of K, as everything in $\mathbb{L}^3(P)$ maps to $\Theta^3\mathfrak{u}$.

The work of Hain-Matsumoto implies that the indecomposable relations in $K_{d,N}$ should only be parametrized by cusp forms, while the Eisenstein series G_n should determine the commutator $[z_{d-1}, \epsilon_n]$. Thus we believe that the element of $\bar{K}_{d,N}$ corresponding to the Eisenstein series will not extend to an element of K, but instead be the lowest depth term in $[z_{d-1}, \epsilon_n]$. The elements of $\bar{K}_{d,N}$ corresponding to cusp forms are expected to extend to actual elements of K. For more on these issues see the remarks in the final section.

Finally, other authors ([4],[5]), in particular Schneps [5], have found the appearance of period polynomials as giving congruences between special derivations of a free Lie algebra of rank 2. This free Lie algebra corresponds to the unipotent fundamental group of $\mathbb{P}^1 - \{0, 1, \infty\}$, whereas $\mathbb{L}(H)$ corresponds to the unipotent fundamental group of a punctured elliptic curve. Matsumoto has proved that the depth 2 relations found in this paper, i.e. Theorem 1, imply those found by Schneps, by specializing to the nodal cubic and observing that the ϵ_{2m+2} and her f_{2m+1} have identical images modulo a certain ideal. The results of this paper on higher depth relations are apparently new.

Recall that $\epsilon_{2m}([a,b]) = 0$. Denote by $\mathrm{Der}^0\,\mathbb{L}(H)$ the Lie subalgebra of $\mathrm{Der}\,\mathbb{L}(H)$ annihilating $[a,b]$:

$$\mathrm{Der}^0\,\mathbb{L}(H) = \{\delta \in \mathrm{Der}\,\mathbb{L}(H) : \delta([a,b]) = 0\}.$$

The subalgebra $\mathfrak{u}$ is contained in $\mathrm{Der}^0\,\mathbb{L}(H)$, which plays a critical role in this work. In fact, the main tool used to prove Theorem 3 is a representation of $\mathrm{Der}^0\,\mathbb{L}(H)$ as sums of uni-trivalent trees with leaves in H, which is due to Garoufalidis and Levine [2]. This is explained at length in §3. Period polynomials are briefly reviewed in §2, and the proof of the main theorem follows in §4. We conclude with some remarks and examples in §5.

2. Period polynomials of modular forms

To a cusp form f of weight n, one can associate the degree $n-2$ homogeneous polynomial $r_f(X,Y)$ defined in the introduction. There is an $SL_2(\mathbb{Z})$ action on homogeneous polynomials in the variables X and Y given by $X \mapsto aX + bY, Y \mapsto cX + dY$. If $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and p is a polynomial, we write this action as $p|\gamma$. (It is a right action.) Period polynomials are known to satisfy the following functional equation:

$$r_f - r_f|T - r_f|T' = 0, \tag{9}$$

where

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } T' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

As was remarked in the introduction, the polynomial r_f can be written $r_f = r_f^+ + r_f^-$, where

$$r_f^+(X,Y) = \frac{1}{2}(r_f(X,Y) + r_f(X,-Y)) \text{ and } r_f^-(X,Y) = \frac{1}{2}(r_f(X,Y) - r_f(X,-Y)).$$

The terms where the exponents of both X and Y are even are contained in r_f^+, and r_f^- contains the terms where both exponents are odd. Not only does r_f satisfy (9), but both r_f^+ and r_f^- do as well.

Define W_n to be the space of homogeneous polynomials in the variables X and Y of degree $n-2$ that satisfy (9). Define W_n^+ (resp. W_n^-) to be the subspace of W_n where the exponents of X and Y are both even (resp. odd). Denote by M_n the space of modular forms for $SL_2(\mathbb{Z})$ of weight n and by M_n^0 the subspace of cusp forms. The theory of Eichler-Shimura-Manin states that the map $r^- : M_n^0 \to W_n^-$ is an isomorphism, and that $r^+ : M_n^0 \to W_n^+$ is an isomorphism onto a codimension 1 subspace. From Zagier's extension, the Eisenstein series $G_n(\tau)$ of weight n yields an even degree period polynomial proportional to $X^{n-2} - Y^{n-2}$, so that r^+ extends to an isomorphism $M_n \to W_n^+$.

Remark 2.1. In the literature, one often sees functional equations with S and U, $S^2 = U^3 = -1$ used to define W_n, and not (9). That one can use T and T' instead of S and U is quite helpful for this paper, and was proved by Gangl-Kaneko-Zagier in [1].

3. Derivations and Trees

Define a symplectic inner product $(\ ,\ ) : \Lambda^2 H \to \mathbb{C}$ on H by $(a,b) = 1$. So, one has $(a,a) = (b,b) = 0$ and $(b,a) = -1$. It will be convenient for us to have a very simple notation for $[a,b]$. Write $\theta = [a,b]$. In this section we will describe a Lie algebra spanned by uni-trivalent trees with leaves in H, which was defined by Garoufalidis and Levine in [2]. This Lie algebra is known to be isomorphic to $\mathrm{Der}^0\,\mathbb{L}(H)$. The construction of the Lie algebra holds more generally over any vector space V over a field of characteristic 0 with a symplectic inner product. For simplicity we only discuss the case of H, which is what we will need. All the trees and subtrees mentioned in the sequel will be assumed to be connected.

3.1. **Rooted trees and $\mathbb{L}(H)$.** We will first describe how $\mathbb{L}(H)$ can be represented as linear combinations of special types of trees. The relevant set of trees is the set R of rooted, oriented, uni-trivalent trees with leaves in H. The uni-trivalent condition means that every vertex either has degree 1 or degree 3; the degree 1 vertices are the ones with leaves. The oriented condition means that there is a preferred cyclic order to the three edges meeting any degree 3 vertex. All this means is that the two trees in Figure 2 are distinct. One can represent the Lie algebra $\mathbb{L}(H)$ as linear combinations of trees in R, modulo a multilinearity, antisymmetry (AS), and Jacobi (IHX) condition, which express the corresponding properties of the Lie bracket. An example makes the construction clear. For instance, the Lie monomial $[[[P,Q],R],[S,T]]$ is represented by the tree in Figure 1. Denote by $\mathbf{R}$ the vector

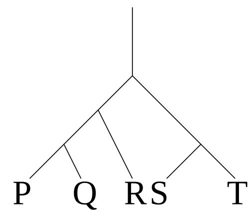

FIGURE 1. A tree for the Lie monomial $[[[P,Q],R],[S,T]]$.

space generated by R and $\mathcal{R}$ its quotient by these relations. We now describe these relations.

- **Multilinearity** Suppose $X, X', X'' \in R$, and that these trees are identical except at one leaf, where X' is labeled by u, X'' is labeled by v, and X is labeled by $u+v$. The multilinearity condition is simply $X = X' + X''$ in $\mathcal{R}$.
- **Antisymmetry** The antisymmetry relation is that the tree in Figure 2, i.e. the sum of the 2 trees seen there, is defined to be 0. In this relation, X and Y can be any subtree not containing the root of the tree, and $*$ represents the rest of the rooted tree. Thus, interchanging the two subtrees at a degree 3 vertex that do not contain the root of the tree multiplies the tree by -1.
- **Jacobi** The Jacobi or IHX relation is that the sum of the 3 trees seen in Figure 3 is defined to be 0. In this figure, X, Y, Z could be any subtrees not containing the root, and $*$ represents the rest of the rooted tree.

The way trees in R and thus $\mathbf{R}$ represent elements of $\mathbb{L}(H)$ can be defined inductively as follows. A tree $t \in R$ with only one leaf u represents the element $u \in H$, and if the tree X represents x, the tree Y represents y, then the first tree in Figure 2 represents $[x,y]$. By allowing linear combinations we get a surjection $\mathbf{R} \to \mathbb{L}(H)$, and this map passes to an isomorphism $\mathcal{R} \to \mathbb{L}(H)$. We will henceforth ignore the distinction between the trees in $\mathcal{R}$ and the elements of $\mathbb{L}(H)$ they represent.

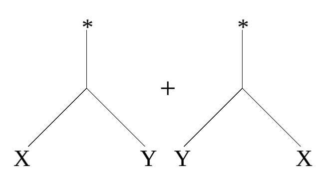

FIGURE 2. Anti-symmetry

3.2. **Non-rooted trees and derivations.** Now we will describe a Lie algebra of derivations on $\mathcal{R}$ that comes from special non-rooted trees. Denote by $\mathbf{A}$ the vector space spanned by the set of oriented, uni-trivalent trees (not-rooted) with leaves in H. If a tree T in $\mathbf{A}$ has n trivalent vertices, then it has $n+2$ univalent vertices and

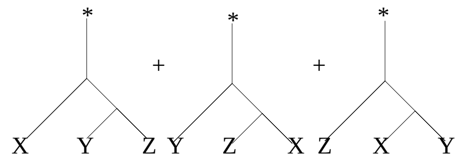

FIGURE 3. IHX for elements of $\mathbb{L}(H)$

$n + 2$ legs; such a tree is said to have degree n. Denote by $\mathbf{A}_n$ the span of all such trees, so $\mathbf{A} = \bigoplus \mathbf{A}_n$. Just as above, impose on trees in $\mathbf{A}$ a multilinearity, AS and IHX relation. The multilinearity condition is the obvious analogue of the one for $\mathcal{R}$. The AS relation is that the sum of the two trees seen in Figure 4 is defined to be 0, and the IHX condition is that the sum of the three trees seen in Figure 5 is defined to be 0. Here, W, X, Y, Z can be any subtree, equivalently any element of $\mathbb{L}(H)$. Denote the resulting quotient space by $\mathcal{A}$. The grading on $\mathbf{A}$ passes to the

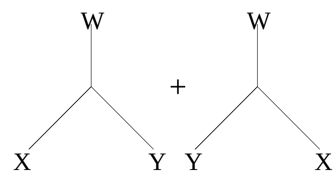

FIGURE 4. AS for elements of $\mathcal{A}$

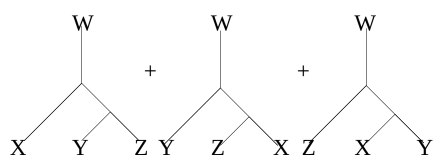

FIGURE 5. IHX for elements of $\mathcal{A}$

quotient, so we have $\mathcal{A} = \bigoplus \mathcal{A}_n$.

If one takes two elements X, Y of $\mathcal{R}$, and glues them together by "fusing" their roots, one will get an element of $\mathcal{A}$, as in Figure 6. Call this fusion a *connected sum.*

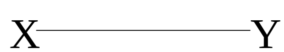

FIGURE 6. Write the above tree as $X\#Y$

Denote the resulting tree in $\mathcal{A}$ as $X\#Y$. Note that $X\#Y = Y\#X$. The connected sum is of course bilinear in each component:

$$(X_1 + X_2)\#Y = X_1\#Y + X_2\#Y.$$

A given tree D in $\mathcal{A}$ can be represented as the connected sum of elements in $\mathcal{R} = \mathbb{L}(H)$ in multiple ways. For example, the first tree in Figure 5 can be written in these four ways:

$$W\#[X,[Y,Z]], X\#[[Y,Z],W], Y\#[Z,[W,X]], Z\#[[W,X],Y].$$

As we said above, the space $\mathcal{A}$ actually acts as a space of derivations on $\mathcal{R}$. Let us now describe this action. So suppose $D \in \mathcal{A}$ is a tree with leaves $\{u_i\}$. Deleting the leaf u_i from D we obtain an element T_i of $\mathcal{R}$, which we call the *leaf complement* to u_i. Thus, for each i, we can write $D = T_i \# u_i$. If $r \in \mathcal{R}$ is a rooted tree, then similarly let $\{v_j\}$ be the leaves of r and let W_j be the doubly rooted tree obtained by deleting v_j from r. Then $D(r)$ is defined to be

$$\sum_{i,j} (u_i, v_j) T_i \# W_j,$$

where T_i is glued to W_j along the roots corresponding to the missing leaves u_i from D and v_j from r. Basically, one is performing a contraction between each leaf u_i of D and v_j of r. Due to the relations imposed on $\mathcal{A}$ and $\mathcal{R}$, this action is well-defined. Since elements of $\mathcal{A}$ act on trees in $\mathcal{R}$ one leaf at a time, it is clear that they act as derivations.

The commutator of derivations $T, T' \in \mathcal{A}$ also has a nice representation in this space of trees. Suppose u_i are the leaves of T, with corresponding complements T_i, and u'_j are the leaves of T', with corresponding complements T'_j. Then

$$[T, T'] = \sum_{i,j} (u_i, u'_j) T_i \# T'_j. \tag{10}$$

There is another way of describing the action of $\mathcal{A}$ on $\mathbb{L}(H)$. Note that since $\mathbb{L}(H)$ is free, a derivation δ is determined by its action on H. Therefore,

$$\operatorname{Der} \mathbb{L}(H) = \operatorname{Hom}(H, \mathbb{L}(H)).$$

But $\operatorname{Hom}(H, \mathbb{L}(H)) = H^* \otimes \mathbb{L}(H)$, and since H has the nondegenerate inner product $(\ ,\ )$, $H^* = H$. Hence $\operatorname{Der} \mathbb{L}(H) \cong H \otimes \mathbb{L}(H)$. In this identification, the element $u \otimes x$ acts as the derivation

$$t \mapsto -(t, u)x \text{ for } t \in H. \tag{11}$$

Now, if $D \in \mathcal{A}$, $\{u_i\}$ are the leaves of D, and T_i are their leaf complements, then D acts as the derivation $\sum_i u_i \otimes T_i$.

3.3. **A better understanding of $\mathcal{A}$ as derivations.** It turns out that every tree $T \in \mathcal{A}$ annihilates θ, and in fact $\mathcal{A} \cong \operatorname{Der}^0 \mathbb{L}(H)$. These facts are established in the literature, as are the facts contained in §§3.1 and 3.2. In this section we will prove these claims, among many other results, by what we believe to be novel methods. Although some of the results of this section are not essential for the proof of Theorem 3, it is hoped that they increase the conceptual understanding of the isomorphism $\mathcal{A} \cong \operatorname{Der}^0 \mathbb{L}(H)$.

Consider the space $\mathcal{D} = (\mathbb{L}(H) \otimes \mathbb{L}(H))/J$, where J is generated by

$$[u, v] \otimes w - u \otimes [v, w] + v \otimes [u, w]$$

for all $u, v, w \in \mathbb{L}(H)$.

Proposition 3.1. *$\mathcal{D} = H \otimes \mathbb{L}(H)$.*

Proof. The relation $[u,v]\otimes w = u\otimes[v,w]-v\otimes[u,w]$ in $\mathcal{D}$ allows one to find a unique representative in $H\otimes\mathbb{L}(H)$ for any $x\otimes y\in\mathcal{D}$. □

Proposition 3.1 implies $\mathcal{D}\cong \operatorname{Der}\mathbb{L}(H)$. We now make this isomorphism explicit. To do this, extend the inner product $(\ ,\ )$ to a map $T(H)^m\otimes T(H)^n\to T(H)^{m+n-2}$ defined as follows. If $x=Xu$ and $y=vY$, with $u,v\in H$, then $(x,y):=(u,v)XY$.

Proposition 3.2. *If $x\in\mathbb{L}(H)$ and δ represents the derivation $u\otimes v$, then $\delta(x)=-(x,u)\cdot v$.*

Proof. The above is true if $x,u\in H$. One proceeds by induction, writing $x=[x',x'']$ to show that the above expression is a derivation on $\mathbb{L}(H)$. Thus the formula must only be checked for $x\in H$. This again follows by induction, by writing $u=[u',u'']$ and using the definition of J. □

Now, the "inner product" $(\ ,\ )$ is no longer antisymmetric. Indeed, one can easily check that $(\theta,x)=(x,\theta)=-x$ for any $x\in\mathbb{L}(H)$. Hence we have the corollaries:

Corollary 3.3. *For any $v\in\mathbb{L}(H)$, $\theta\otimes v$ is the derivation $-ad(v)$.*

Corollary 3.4. *If δ represents the derivation $u\otimes v$, then $\delta(\theta)=[u,v]$. Hence*

$$\operatorname{Der}^0\mathbb{L}(H)=\ker\{[\ ,\ ]:(\mathbb{L}(H)\otimes\mathbb{L}(H))/J\to\mathbb{L}(H)\}.$$

The space $\mathcal{D}$ is intimately related to the representation of derivations as trees, as Proposition 3.6 shows. To see this connection, we must first briefly discuss *subtrees* of a tree $D\in\mathcal{A}$. Consider any edge of a tree D. If one cuts this edge at its midpoint, then D splits into two rooted trees X, Y with $D=X\#Y$. By a subtree we mean either X or Y. That is, a subtree of $D\in\mathcal{A}$ is an element $T\in\mathcal{R}$ such that there exists $T'\in\mathcal{R}$ with $D=T\#T'$.

For a subtree V of D, denote by $L(V)$ the subset of leaves of D contained in V. Then a family of subtrees $\{V_i\}$ of D is said to be *disjoint* if the corresponding sets of leaves $\{L(V_i)\}$ are pairwise disjoint, and is said to be *complete* if $\cup L(V_j)=L(D)$, i.e. every leaf of D is in some V_i. We are interested in complete, disjoint sets of subtrees of D. For example, the set of leaves $\{u_i\}$ of D is always a complete disjoint set, and no matter how big the subtrees W,X,Y,Z are, $\{W,X,Y,Z\}$ is a complete disjoint set of subtrees for the trees in Figure 5.

Here the simplest example of the use of such sets.

Proposition 3.5. *If $\{V_j\}$ are a complete disjoint set of subtrees of T' and W_j are their corresponding complements, then*

$$[T,T']=\sum_j T(V_j)\#W_j.$$

Proof. This follows straight away from (10) and the definition of how $\mathcal{A}$ acts on $\mathcal{R}$. □

Now, instead of writing $D=\sum_i u_i\otimes T_i$, let us replace the set of leaves $\{u_i\}$ with any other complete disjoint set of subtrees $\{V_j\}$. For each V_j, denote by W_j the rooted tree in $\mathcal{R}$ obtained by deleting V_j from D; that is, W_j is the *complement* of V_j. Then we have

Proposition 3.6. *If $D \in \mathcal{A}$ has a complete disjoint set of subtrees $\{V_j\}$ with corresponding complements W_j, then D acts as the derivation $\sum_j V_j \otimes W_j$ in $\mathcal{D}$.*

If the hypotheses of this Proposition are satisfied, call the equality $D = \sum V_j \otimes W_j$ a *tree representation* of the derivation D.

Proof. This follows by an easy induction on the size of the largest V_j, where by size we mean how deep each V_j is in the grading $\mathbb{L}(H) = \bigoplus \mathbb{L}_n(H)$, $\mathbb{L}_n(H)$ being the subspace of $\mathbb{L}(H)$ generated Lie monomials of length n. The base case is the case where each V_i is a leaf, and we know the claim is true in this case. If V_1 is the largest such V_i, and V_1 is not a leaf, then $V_1 = [A, B]$ for some A, B of size strictly smaller than the size of V_1. Hence

$$\begin{aligned}\sum_i V_i \otimes W_i &= [A,B] \otimes W_1 + \sum_{i>1} V_i \otimes W_i \\ &= A \otimes [B, W_1] + B \otimes [W_1, A] + \sum_{i>1} V_i \otimes W_i.\end{aligned}$$

But $\{A, B, V_2, \ldots\}$ is also a complete, disjoint set of subtrees, and the complements of A, B are $[B, W_1]$, $[W_1, A]$, respectively, so the result follows. □

Corollary 3.7. *If $D \in \mathcal{A}$, then $D(\theta) = 0$.*

Proof. Write D as a connected sum $X \# Y$. Then $\{X, Y\}$ is a complete disjoint set of subtrees of D, so D is the derivation $X \otimes Y + Y \otimes X$. Hence $D(\theta) = [X, Y] + [Y, X] = 0$. □

Hence the map $\mathcal{A} \to \operatorname{Der} \mathbb{L}(H)$ actually lands inside $\operatorname{Der}^0 \mathbb{L}(H)$. In fact, we will soon see the isomorphism $\mathcal{A} \cong \operatorname{Der}^0 \mathbb{L}(H)$.

Lemma 3.8. *If $\delta \in \operatorname{Der}^0 \mathbb{L}(H)$, $x, y \in T(H)$, and $(\ ,\ )$ is the extended inner product, then $\delta((x, y)) = (\delta(x), y) + (x, \delta(y))$.*

Proof. From the definition of the extended $(\ ,\ )$, it suffices to check the formula when $x, y \in H$, in which case one must show $(x, \delta(y)) + (\delta(x), y) = 0$. Start with $0 = -\delta(\theta) =$

$$-a\delta(b) + \delta(b)a - \delta(a)b + b\delta(a).$$

Contracting with a on the left and b on the right gives $(\delta(a), b) + (a, \delta(b)) = 0$. Other equalities are obtained similarly. □

Theorem 4. *The space of trees $\mathcal{A}$ is isomorphic to $\operatorname{Der}^0 \mathbb{L}(H)$. Additionally, if $\delta \in \operatorname{Der}^0_n \mathbb{L}(H)$ for some $n > 0$, and $\delta(a)$ is of depth d, then δ is represented by the tree $-\frac{1}{d+1} b \# \delta(a)$.*

Proof. First let us check injectivity of $\mathcal{A} \to \operatorname{Der}^0 \mathbb{L}(H)$. So, suppose $T \in \mathcal{A}$ is a tree that gives the zero derivation. Use multilinearity to express T as a tree whose leaves are just a, b, and suppose m of these leaves are b. Denote the leaf complements to these b's as $S_1, \ldots S_m$. Then $T = b \# S_i$ for each i, and thus $T = \frac{1}{m} \sum_{i=1}^m b \# S_i$ inside $\mathcal{A}$. But $T(a) = -\sum S_i = 0$, so linearity gives that T is 0 inside $\mathcal{A}$, proving injectivity.

For surjectivity we prove the second claim of the theorem. If $\eta = -\frac{1}{d+1} b\#\delta(a)$, then $\eta = -\frac{1}{d+1}(b \otimes \delta(a) + \delta(a) \otimes b)$. Clearly $-(b \otimes \delta(a))(a) = \delta(a)$ from Proposition 3.2. We have

$$\begin{aligned} -(\delta(a) \otimes b)(a) &= (a, \delta(a)) \cdot b \\ &= -(\delta(a), a) \cdot b \end{aligned}$$

from Lemma 3.8. Now, $\delta(a) = Ua + Vb$ for some U, V in the tensor algebra $T(H)$, so that $-(\delta(a), a) = V$. It is not hard to show via an inductive argument that $V \cdot b = d\delta(a)$. Hence $\eta \in \mathrm{Der}^0_n \mathbb{L}(H)$, and $\eta(a) = \delta(a)$, which implies $\eta = \delta$, as desired. □

So, for example, since $\epsilon_{2m}(a) = a^{2m} \cdot b$, we have $\epsilon_{2m} = -\frac{1}{2} b\#(a^{2m} \cdot b)$, which if $m > 0$ is equal to $\frac{1}{2}\theta\#(a^{2m-2} \cdot \theta)$.

As another consequence of Lemma 3.8, we have

Proposition 3.9. *If $\delta \in \mathrm{Der}^0 \mathbb{L}(H)$, then $[\delta, u \otimes v] = \delta(u) \otimes v + u \otimes \delta(v)$.*

Proof. This is an easy consequence of Proposition 3.2 and Lemma 3.8. □

3.4. **Specialization to $\mathfrak{u}$.** Equation (10) gives an expression for $[\epsilon_p, \epsilon_q]$ as a sum of $(p+2)(q+2)$ terms, although most of these terms are zero as $(a, a) = 0$. In fact, the commutator $[\epsilon_p, \epsilon_q]$ actually simplifies much further, and in the next section we will compute, as an easy example of one of the results, that $[\epsilon_p, \epsilon_q]$ is given by the single tree $\theta\#([a^{p-2} \cdot \theta, a^{q-2} \cdot \theta])$ when $p, q > 0$. We can use this fact to easily prove Theorem 1. Indeed, we have

$$[\epsilon_p, \epsilon_q] = \theta \otimes ([a^{p-2} \cdot \theta, a^{q-2} \cdot \theta]) + ([a^{p-2} \cdot \theta, a^{q-2} \cdot \theta]) \otimes \theta,$$

which using the relations from J can be written

$$\theta \otimes T_{p,q} + a \otimes S.$$

Here

$$T_{p,q} = [a^{p-2} \cdot \theta, a^{q-2} \cdot \theta] - a^{p-2} \cdot [\theta, a^{q-2} \cdot \theta] - a^{q-2} \cdot [a^{p-2} \cdot \theta, \theta]$$

and S is some not-so-important stuff. Hence $[\epsilon_p, \epsilon_q](a) = [a, T_{p,q}]$.

Since the Lie algebra $\mathbb{L}(H)$ is free, $[a, X] = 0$ implies X is a constant multiple of a. Hence if $\delta \in \mathrm{Der}^0_n \mathbb{L}(H)$ with $n > 0$, then $\delta(a) = 0$ implies $\delta(b) = 0$ as

$$[a, \delta(b)] = [b, \delta(a)].$$

Therefore, $\sum \beta(p, q)[\epsilon_p, \epsilon_q] = 0$ if and only if $\sum \beta(p, q)[a, T_{p,q}] = 0$, which holds in and only if $\sum \beta(p, q) T_{p,q} = 0$. But if we write $X^j Y^k$ for the element $(a^j \cdot \theta)(a^k \cdot \theta)$ of $T(H)$, then $T_{p,q}$ would be written

(12)

$$(X^{p-2}Y^{q-2} - X^{q-2}Y^{p-2}) - (X+Y)^{p-2}(Y^{q-2} - X^{q-2}) - (X+Y)^{q-2}(X^{p-2} - Y^{p-2}).$$

Rearranging the terms above, one sees that (12) is just

$$-(X^{p-2}Y^{q-2} - X^{q-2}Y^{p-2})|(-I + T + T').$$

Thus $\sum \beta(p,q)[\epsilon_p, \epsilon_q] = 0$ if and only if

$$\sum \beta(p,q)(X^{p-2}Y^{q-2} - X^{q-2}Y^{p-2})|(-I + T + T') = 0,$$

which proves the result.

Theorem 1 can be proved in ways simpler than the above, but we have presented this proof in a fashion somewhat parallel to that of Theorem 3, so that the proof of Theorem 3 is clear.

Let us now briefly discuss the trees for μ_{2m}^{2m-k} before moving on to §4. We have, for instance, that ϵ_6 is equal to $\frac{1}{2}$ the derivation in Figure 7, and the derivation μ_4^3 is $\frac{1}{4}$ the derivation in Figure 8.

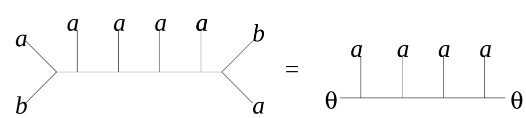

FIGURE 7. Two equivalent ways of writing the derivation $2\epsilon_6 = 2\mu_4^4$.

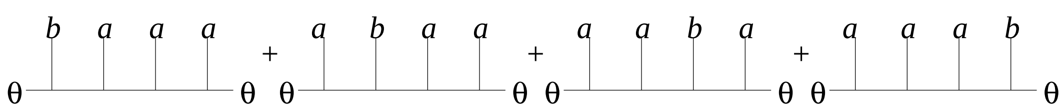

FIGURE 8. One-fourth the sum of the derivations above is the derivation μ_4^3.

For the general tree representation of μ_{2m}^k, set $s(j,k)$ to be the element of the tensor algebra $T(H)$ on H which is

$$\frac{1}{\binom{j+k}{k}} \sum u_1 \cdots u_{j+k},$$

where the sum is over every $u_1 \cdots u_{j+k}$ with j a's and k b's. For example, $s(3,1) = \frac{1}{4}(aaab + aaba + abaa + baaa)$. We have

Proposition 3.10. *The derivation μ_{2m}^{2m-k} corresponds to the tree*

$$\frac{1}{2}\theta\#(s(2m-k,k)\cdot\theta).$$

Proof. We have $\mu_{2m}^{2m-k} = \frac{1}{\binom{2m}{k}} \frac{{\epsilon_0}^k}{k!} \cdot (\mu_{2m}^{2m})$. The derivation μ_{2m}^{2m} comes from the tree $\frac{1}{2}\theta\#(a^{2m}\cdot\theta)$, and thus by Proposition 3.5,

$${\epsilon_0}^k(\theta\#(a^{2m}\cdot\theta)) = \theta\#({\epsilon_0}^k(a^{2m})\cdot\theta)$$

since $\epsilon_0(\theta) = 0$. Since $s(2m-k,k) = \frac{1}{\binom{2m}{k}} \frac{{\epsilon_0}^k}{k!}(a^{2m})$, the result follows. ∎

4. Proofs of Main Theorems

Recall that we are trying to find the space of highest $\mathfrak{sl}_2$−weight vectors part in $\ker \Phi$. Before we move on, let us prove a quick proposition characterizing highest weight vectors in $\mathbb{L}(P)$.

Proposition 4.1. *Denote by $\mathbb{L}_2(P)$ the subspace of $\mathbb{L}(P)$ spanned by commutators $[x_{2m}^j, x_{2n}^k]$. A basis for the highest weight vectors in $\mathbb{L}_{d,N}(P) \cap \mathbb{L}_2(P)$ is the set $\{h_{p,q}^d\}$, with*

$$h_{p,q}^d = \sum_{i+j=d-2} (-1)^i \binom{d-2}{i} [x_p^{p-i}, x_q^{q-j}],$$

$N = p+q+4$, and $p \geq q$ if d is odd and $p > q$ if d is even.

Proof. Any element of $\mathbb{L}_{d,N}(P) \cap \mathbb{L}_2(P)$ must be a sum of elements of the form $\sum_{i+j=d-2} \alpha(i,j)[x_p^{p-i}, x_q^{q-j}]$ for some $\alpha(i,j)$, where $N = p+q+4$ and $p \geq q$ if d is odd and $p > q$ if q is even. Thus, the only question is what are the correct $\alpha(i,j)$.

Denote by $r = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ the element of $\mathfrak{sl}_2$ that determines highest weight vectors by its kernel. That is, if W is an $\mathfrak{sl}_2$ module, $w \neq 0 \in W$ is a highest weight vector if and only if $rw = 0$. Then $rx_p^k = (p-k)x_p^{k+1}$, so applying r to $\sum_{i+j=d-2} \alpha(i,j)[x_p^{p-i}, x_q^{q-j}]$ and setting equal to 0 we find the $\alpha(i,j)$ must satisfy

$$\alpha(i,j)i + \alpha(i-1,j+1)(j+1) = 0.$$

If we normalize $\alpha(d-2,0) = 1$, then it follows that $\alpha(i,j) = (-1)^i\binom{d-2}{i}$. □

We will now prove a few results about the filtration Θ^k defined in the introduction. The most important result in this section is Corollary 4.7. If $\delta \in \Theta^A \operatorname{Der}^0 \mathbb{L}(H)$ and $\delta' \in \Theta^B \operatorname{Der}^0 \mathbb{L}(H)$, and these derivations have nice tree representations, then Corollary 4.7 shows how to calculate the image of $[\delta, \delta']$ in $\Theta^{A+B} \operatorname{Der}^0 \mathbb{L}(H)$ modulo $\Theta^{A+B+1} \operatorname{Der}^0 \mathbb{L}(H)$ very easily using trees. Using the results of this section, Theorem 3 will follow easily after a short calculation.

4.1. The filtration Θ^k.

This first proposition relates the filtration Θ^k to the tree representation of $\operatorname{Der}^0 \mathbb{L}(H)$.

Proposition 4.2. *If $\delta \in \operatorname{Der}^0 \mathbb{L}(H)$ can be written as a tree with $k+1$ θ's, then $\delta \in \Theta^k \operatorname{Der}^0 \mathbb{L}(H)$.*

Proof. From the definition of how trees act as derivations, it is clear that $\delta(u) \in \Theta^k\mathbb{L}(H)$ for any $u \in H$. Using the fact that $\delta(\theta) = 0$, it follows that $\delta \in \Theta^k \operatorname{Der}^0 \mathbb{L}(H)$. □

Proposition 3.10 shows that μ_{2m}^k can be written as a tree with 2 θ's, and is thus in $\Theta^1 \operatorname{Der}^0 \mathbb{L}(H)$. Hence

Corollary 4.3. *$\Phi(\mathbb{L}(P)^k) \subset \Theta^k \operatorname{Der}^0 \mathbb{L}(H)$.*

To handle elements of $\mathbb{L}(H)$ more easily, we will represent them as polynomials in the indeterminates $x_1, x_2, \ldots$ in the following manner. To the element $a^p \cdot b$ of $\mathbb{L}(H)$ we associate the polynomial x_1^p, and to $[a^p \cdot b, a^q \cdot b] = (a^p \cdot b)(a^q \cdot b) - (a^q \cdot b)(a^p \cdot b)$ the polynomial $x_1^p x_2^q - x_1^q x_2^p$. In general, to $\sum_P \lambda_P a^{p_1} \cdot b \cdots a^{p_r} \cdot b$, associate the polynomial

$$f(x_1, x_2, \ldots, x_r) = \sum_P \lambda_P x_1^{p_1} \cdots x_r^{p_r}.$$

Call this map $F : \mathbb{L}(H) \to \mathbb{C}[X]$, where $X = \{x_1, x_2, \ldots\}$. One of the reasons this map F is useful is that it respects the filtration Θ in the following sense: If $u \in \Theta^k \mathbb{L}(H)$, then Fu is a sum of monomials divisible by k *distinct* x_i. That is, if $\mathfrak{m}_k = \langle x_{i_1} x_{i_2} \ldots x_{i_k} : i_\alpha \neq i_\beta \text{ when } \alpha \neq \beta \rangle$ is the ideal generated by products of k distinct x_{i_j}, then $F\Theta^k \mathbb{L}(H) \subset \mathfrak{m}_k$. One also has the converse.

Lemma 4.4. *If $\eta \in \mathbb{L}(H)$ satisfies $F\eta \in \mathfrak{m}_k$, then $\eta \in \Theta^k \mathbb{L}(H)$.*

Proof. All that $F\eta \in \mathfrak{m}_k$ means is that, when written as a sum of monomials in the tensor algebra $T(H)$, each term has k θ's in it. That is, η is a sum of terms of the form $U_1 \theta U_2 \theta U3 \ldots U_k \theta U_{k+1}$, with each $U_i \in T(H)$.

Denote by $P : T(H) \to \mathbb{L}(H)$ the linear map defined by setting $P(u_1 \ldots u_n)$ equal to $\frac{1}{n}(u_1 \cdots u_{n-1}) \cdot u_n$ whenever each u_i is in H. This map is a projection on to $\mathbb{L}(H)$. Since η is Lie, one thus has $\eta = P(\eta)$. But $P(U_1 \theta U_2 \theta U3 \ldots U_k \theta U_{k+1})$ is clearly in $\Theta^k \mathbb{L}(H)$, so η is in $\Theta^k \mathbb{L}(H)$. □

More importantly, we have the following.

Lemma 4.5. *If $\delta \in \operatorname{Der}^0 \mathbb{L}(H)$ and $\delta(a) \in \Theta^k \mathbb{L}(H)$, then $\delta \in \Theta^k \operatorname{Der}^0 \mathbb{L}(H)$.*

Proof. First we need to show that if $[a, f] \in \Theta^k \mathbb{L}(H)$ for some $f \in \mathbb{L}(H)$, then $f \in \Theta^k \mathbb{L}(H)$. So, assume $[a, f] \in \Theta^k \mathbb{L}(H)$. We have then that $F[a, f] \in \mathfrak{m}_k$, so for any distinct $i_1, \ldots i_{k-1}$, $F[a, f]$ is 0 when evaluated at $x_\alpha = 0, \alpha \neq i_j$. But $F[a, f] = (x_1 + \cdots + x_r) Ff(x_1, \ldots, x_r)$, so Ff is 0 when evaluated at $x_\alpha = 0, \alpha \neq i_j$. Hence $Ff \in \mathfrak{m}_k$ so $f \in \Theta^k \mathbb{L}(H)$ by Lemma 4.4.

Now, since $\delta \in \operatorname{Der}^0 \mathbb{L}(H)$, $[a, \delta(b)] = [b, \delta(a)]$. Thus by the remarks above $\delta(b)$ is also in $\Theta^k \mathbb{L}(H)$. Again using the fact that $\delta(\theta) = 0$, it is easy to see that $\delta \in \Theta^k \operatorname{Der}^0 \mathbb{L}(H)$. □

Note that in Proposition 3.5 there is an asymmetry in the expression for $[T, T']$. One can apply T to the tree parts of T', or by interchanging the roles of T and T', apply T' to the tree parts of T. One gets two expressions, that while representing the same derivation, look different. This next proposition will remove this asymmetry and take full advantage of the fact that the derivations in $\operatorname{Der}^0 \mathbb{L}(H)$ annihilate θ. In fact, this "θ trick" is the reason that using trees to compute with $\operatorname{Der}^0 \mathbb{L}(H)$ is beneficial. Recall the definition of tree representation from section §3. Then

Proposition 4.6 (The θ trick)**.** *Suppose $\delta_1, \delta_2 \in \operatorname{Der}^0 \mathbb{L}(H)$, and $\delta_1 = \sum \theta \otimes s_i + \sum u_j \otimes t_j$, $\delta_2 = \sum \theta \otimes e_k + \sum v_l \otimes f_l$ are tree representations of these derivations. Then $[\delta_1, \delta_2]$ is given by the tree $\sum \theta \# [s_i, e_k] + \sum (u_j, v_l) t_j \# f_l$.*

This Proposition says that, when computing $[\delta_1, \delta_2]$, one can do a contraction of the θ's of δ_1 with the θ's of δ_2 (which is very easy), and add that to the contraction of the "non-θ's" of δ_1 with the "non-θ's" of δ_2. The point is that there are no "cross terms" coming from contracting the θ's of δ_1 with the non-θ's of δ_2, and vice versa.

One easy and pertinent application of Proposition 4.6 is the case $\delta_1 = \mu_{2m}^{2m}$ and $\delta_2 = \mu_{2n}^{2n}$. A tree representation of μ_{2m}^{2m} is $\frac{1}{2}\theta \otimes a^{2m} \cdot \theta + \frac{1}{2}\theta \otimes a^{2m} \cdot \theta + a \otimes \text{Stuff}$, or simply

$$\theta \otimes a^{2m} \cdot \theta + a \otimes \text{Stuff}_m.$$

Hence $[\mu_{2m}^{2m}, \mu_{2n}^{2n}]$ is given by the tree $\theta\#[a^{2m} \cdot \theta, a^{2n} \cdot \theta] + (a,a)\text{Stuff}_m\#\text{Stuff}_n$, which is just $\theta\#[a^{2m} \cdot \theta, a^{2n} \cdot \theta]$ since $(a,a) = 0$.

Proof of Proposition. Let $A = \sum \theta \otimes s_i$ denote the θ terms of δ_1 and $B = \sum u_j \otimes t_j$ denote the non-θ terms. Similarly, denote the θ and non-θ terms of δ_2 by C and D, respectively. To compute $[\delta_1, \delta_2]$, we must contract each of A and B with each of C and D. Write this as $(A+B)\times(C+D)$. We have $(A+B)\times(C+D) = A\times(C+D) + (A+B)\times C - A\times C + B\times D$. But since δ_2 annihilates θ, $A\times(C+D) = 0$. Similarly since δ_1 annihilates θ $(A+B)\times C = 0$. Hence $(A+B)\times(C+D) = -A\times C + B\times D$. It is not difficult to see that the contraction $-A\times C$ of the θ's of A with the θ's of B is just $\sum \theta\#[s_i, e_k]$. The term $B \times D$ is of course $\sum (u_j, v_l) t_j \# f_l$. □

Corollary 4.7. *Suppose $\delta_1 \in \Theta^A \operatorname{Der}^0 \mathbb{L}(H)$ and $\delta_2 \in \Theta^B \operatorname{Der}^0 \mathbb{L}(H)$. Further, suppose δ_1 has a tree representation $\delta_1 = \sum \theta \otimes s_i + \sum u_j \otimes t_j$ where $s_i \in \Theta^A \mathbb{L}(H)$ and $t_j \in \Theta^{A+1}\mathbb{L}(H)$. Similarly suppose δ_2 has a tree representation $\delta_2 = \sum \theta \otimes e_k + \sum v_l \otimes f_l$ with $e_k \in \Theta^B \mathbb{L}(H)$ and $f_l \in \Theta^{B+1}\mathbb{L}(H)$. Then modulo Θ^{A+B+1}, $[\delta_1, \delta_2]$ is represented by the tree $\sum \theta\#[s_i, e_k]$.*

Proof. The trees $t_j\#f_l$ all have $A + B + 2$ θ's, and are thus in $\Theta^{A+B+1} \operatorname{Der}^0 \mathbb{L}(H)$, i.e. are 0 modulo Θ^{A+B+1}. □

4.2. **Proof of the Main Result.**

Set $r_{2m}^k = s(k, 2m-k) \cdot \theta$. Then $\mu_{2m}^k = \Phi(x_{2m}^k)$ is represented by the tree $\frac{1}{2}\theta\#r_{2m}^k$, which, modulo Θ^2, is given by $\theta \otimes r_m^{m-k}$. Thus we have

Corollary 4.8. *$\bar{\Phi}([x_p^{p-i}, x_q^{q-j}])$ is represented by the tree $\theta\#[r_p^{p-i}, r_q^{q-j}]$.*

Proof of main theorem. We will compute $\Phi(h_{p,q}^d)(a)$ modulo $\Theta^3\mathbb{L}(H)$. The equalities below are equalities in the space $\mathbb{L}(H)/\Theta^3\mathbb{L}(H)$.

From Corollary 4.8 and Proposition 3.3, one has that $\Phi([x_p^{p-i}, x_q^{q-j}])(a) = [a, T_{p,q}^{i,j}]$ where $T_{p,q}^{i,j} =$

$$-[s(p-i,i)\cdot\theta, s(q-j,j)\cdot\theta] + s(p-i,i)\cdot[\theta, s(q-j,j)\cdot\theta] + s(q-j,j)\cdot[s(p-i,i)\cdot\theta, \theta].$$

Setting, for ease of notation $p' = p-d+2$, $q' = q-d+2$, $d' = d-2$, and $\theta_U^V = a^V b^U \cdot \theta$, we have that

$$\begin{aligned} T_{p,q}^{i,j} &= -[a^{p'} \cdot \theta_i^j, a^{q'} \cdot \theta_j^i] + a^{p'} a^j b^i \cdot [\theta, a^{q'} \theta_j^i] + a^{q'} a^i b^j \cdot [a^{p'} \theta_i^j, \theta] \\ &= -[a^{p'} \cdot \theta_i^j, a^{q'} \cdot \theta_j^i] + \sum_{l,m} \binom{i}{l} \binom{j}{m} \Big\{ a^{p'} \cdot \Big([\theta_l^m, a^{q'} \cdot \theta_{d'-l}^{d'-m}] \Big) \\ &\quad + a^{q'} \cdot \Big([a^{p'} \cdot \theta_{d'-m}^{d'-l}, \theta_m^l] \Big) \Big\}. \end{aligned}$$

Now $\Phi(h_{p,q}^d)(a) = [a, T_{p,q}]$, where $T_{p,q} = \sum_{i=0}^{d'} \binom{d'}{i} (-1)^i T_{p,q}^{i,j}$. Fixing l, m, the sum

$$\sum_{i,j} \binom{d'}{i} (-1)^i \binom{i}{l} \binom{j}{m}$$

is easily seen to be $(-1)^l \binom{d'}{l} \delta_{m,d'-l}$, i.e. is 0 unless $d' = l + m$, in which case it is $(-1)^l \binom{d'}{l}$. Thus, we have that $T_{p,q}$

$$\begin{aligned} &= \sum_{l+m=d'} \binom{d'}{l} (-1)^l \Big\{ -[a^{p'} \cdot \theta_l^m, a^{q'} \cdot \theta_m^l] + a^{p'} \cdot \Big([\theta_l^m, a^{q'} \cdot \theta_m^l] \Big) \\ &\quad + a^{q'} \cdot \Big([a^{p'} \cdot \theta_l^m, \theta_m^l] \Big) \Big\} \\ &= \sum_{l+m=d'} \binom{d'}{l} (-1)^l \Big\{ - \Big(a^{p'} \cdot \theta_l^m a^{q'} \cdot \theta_m^l - (-1)^{d'} a^{q'} \cdot \theta_l^m a^{p'} \cdot \theta_m^l \Big) + \\ &\quad a^{p'} \cdot \Big(\theta_l^m a^{q'} \cdot \theta_m^l - (-1)^{d'} a^{q'} \cdot \theta_l^m \theta_m^l \Big) + \\ &\quad a^{q'} \cdot \Big(a^{p'} \cdot \theta_l^m \theta_m^l - (-1)^{d'} \theta_l^m a^{p'} \cdot \theta_m^l \Big) \Big\} \end{aligned}$$

The point of the above expression is that the first part is a $-I$ term, the second part a T term, and the third part a T' term. Thus if $v = \sum a_f(p-d+2) h_{p,q}^d$ for some modular form f of weight $N-2d+2$, then summing over p, q while keeping l, m fixed, we see that each term of the sum above vanishes, and thus that $\Phi(v)(a) = [a, 0] = 0$ modulo $\Theta^3\mathbb{L}(H)$. Hence by Lemma 4.5, $\Phi(v) \in \Theta^3\mathfrak{u}$, so $v \in \bar{K}(d, N)$.

For the converse, expand the above in the Lazard basis $\{a^k \cdot b\}$ and just consider the terms of the form $a^j \cdot ba^k \cdot bb^{d-2}$ with $j \geq 1$. For the sake of clarity, call such a term "good". The good terms are outside the span of the rest of the terms, which are of the form bG, for some G in $T(H)$. Hence if the entire sum $\sum \alpha(p,q) T_{p,q}$ vanishes, so must the good terms. But only the parts of the sum above with $l = 0$ contribute to the good terms, and the good terms also form a "$-I + T + T'$" expression, so the $\alpha(p,q)$ must be of the form $a_f(p-d+2)$ for some modular form f of weight $N - 2d + 2$, as desired. □

5. Remarks, Examples, and Galois action computations

In the first part of this section we will make a few remarks regarding Theorem 3. Part 2 of this section includes specific examples of relations between the ϵ_{2m}'s of depths 2, 3, and 4, i.e. elements of $K_{d,N}$ for $d = 2, 3, 4$. In part 3 we will present trees for $\tilde{z}_3$ and $\tilde{z}_5$ and expressions for $[\tilde{z}_3, \epsilon_n]$ for a few small n. In this part we also remark about two interesting congruences observed between the action of $\tilde{z}_3$ and the depth 4 relations, which exist modulo numerators of Bernoulli numbers. The first congruence is between $[\tilde{z}_3, \epsilon_{12}]$ and the depth 4 relation corresponding to Δ, and is modulo 691. The second congruence is between $[\tilde{z}_3, \epsilon_{16}]$ and the depth 4 relation corresponding to the weight 16 cusp form; this congruence is modulo 3617.

5.1. **Remarks.** Note that since $K \subset \mathbb{L}^2(P)$, $[\mathbb{L}(P), K] \subset \mathbb{L}^3(P)$. Therefore, the $\mathbb{L}(P)$-consequences of the relations are 0 in the quotient space $\mathbb{L}(P)/\mathbb{L}^3(P)$. Thus, if $v \neq 0$ in $\bar{K}_{d,N}$ lifts to a true element of K, then v is not an $\mathbb{L}(P)$-consequence of some other relation. Hence, the results of this paper suggest that there is a map

$$\text{indecomposable relations} \to \text{modular forms}$$

given by $v \mapsto f$ if v mod $\mathbb{L}^3(P)$ comes from the period polynomial of f.

Now, for d even, Theorem 3 implies that $h^d_{d-2,n+d-4}$ is in $\bar{K}_{d,n+2d-2}$, since the polynomial $X^{n-2} - Y^{n-2}$ is in W_n^+, corresponding to the Eisenstein series of weight n. For clarity, set $E^d_n = h^d_{d-2,n+d-4}$, the "E" being for "Eisenstein." We conjecture that E^d_n never comes from an actual element of $K_{d,N}$, while the ones coming from cusp forms actually do give honest relations.

It appears however that this spurious element E^d_n is actually related to the Galois action. Indeed, note that the element $[\tilde{z}_{d-1}, \epsilon_n]$ is of type $(d, n + 2d - 2)$. Hain has shown that the relation $[z_{d-1}, \epsilon_n] = u$ for some $u \in \mathfrak{u}$ should correspond to the Eisenstein series G_n in the same way that relations of type $(d, n + 2d - 2)$ between the ϵ_{2m} correspond to cusp forms. As we said above, the results of this paper suggest that the way to determine how indecomposable relations in $\mathfrak{u}$ correspond to cusp forms is by looking modulo $\mathbb{L}^3(P)$. Thus, it should be the case that there is some $v \in \mathbb{L}(P)$ such that $\Phi(v) = [\tilde{z}_{d-1}, \epsilon_n]$ and v is proportional to E^d_n modulo $\mathbb{L}^3(P)$. This is indeed the case in all the examples computed in part 3 of this section.

5.2. **Examples.** Here are some examples of the interesting relations holding between the ϵ_{2m}, for depths 2,3, and 4. No relations of depth 5 or higher have as yet been computed. The coefficients in these relations rapidly get very large, so we only list the first two relations for each depth. We will use the x^k_{2m} notation to express the relations, as this notation makes the correspondence with period polynomials more apparent. With each relation we also list the corresponding *cusp form* weight (not $\mathfrak{sl}_2$ weight or m-weight) for easy reference.

5.2.1. *Depth 2.* The coefficients appearing in the depth 2 relations are precisely the even degree parts of the period polynomials.

Weight 12:
$h^2_{2,8} - 3h^2_{4,6}$

Weight 16:
$2h^2_{2,12} - 7h^2_{4,10} + 11h^2_{6,8}$

5.2.2. *Depth 3.* Write $[p, q, r]$ for the element $[x^p_p, [x^q_q, x^r_r]]$. The parts of the expressions below of the form $h^3_{p,q}$ are the parts determined by the theorems of this paper. They come from the odd degree part of the period polynomial of the corresponding cusp form.

Weight 12:
$160h^3_{2,10} - 1000h^3_{4,8} + 840h^3_{6,6} + 1725[4, 4, 2] - 462[2, 6, 2]$

Weight 16:
$22680h^3_{2,14} - 154350h^3_{4,12} + 339570h^3_{6,10} - 207900h^3_{8,8} + 32910[2, 10, 2]$
$+ 650[4, 8, 2] - 241164[6, 6, 2] + 500025[8, 4, 2]$

Note that if v_2 represents a depth 2 relation, then $[x^{2m}_{2m}, v_2]$ is a depth 3 relation for any $m > 0$. Thus, coefficients in depth 3 relations are only canonical up to multiples of expressions of the form $[x^{2m}_{2m}, v_2]$. So, for example, $[x^4_4, [x^2_2, x^8_8]] - 3[x^4_4, [x^4_4, x^6_6]]$ is a relation of modular weight 16. Thus the coefficients in the weight 16 relation above are only canonical up to multiples of this expression.

5.2.3. *Depth 4.* Write $d(p; q, r) = [x^p_p, h^3_{q,r}]$ and $[p, q, r, s] = [x^p_p, [x^q_q, [x^r_r, x^s_s]]]$. Again, only the coefficients of the terms of the form $h^4_{p,q}$ are determined by this paper. Here it is again the even degree part of the period polynomials that occurs as the coefficients multiplying these terms. Additionally, not all the coefficients below are canonically related to the corresponding cusp form, since adding multiples of consequences of relations to the relations below again gives a relation.

Weight 12:
$19958400h^4_{12,2} - 383090400h^4_{10,4} + 1149271200h^4_{8,6} - 1134826056d(2; 8, 2)$
$+ 94270176d(2; 6, 4) + 3552068520d(4; 6, 2) - 691691000d(4; 4, 4)$
$- 2708723160d(6; 4, 2) + 674053380d(8; 2, 2) + 808632825[2, 2, 4, 2]$

This relation is proportional to

$$\frac{36}{691}h^4_{12,2} - h^4_{10,4} + 3h^4_{8,6} + \dots,$$

as expected.

Weight 16:
$95351256000h^4_{16,2} - 1916030516400h^4_{14,4} + 6706106807400h^4_{12,6}$
$- 10538167840200h^4_{10,8} - 3151348328700d(2; 12, 2) - 7088135254200d(2; 10, 4)$
$+ 3907872076680d(2; 8, 6) + 8435253298500d(4; 10, 2) + 46246901337000d(4; 8, 4)$
$- 35901886645170d(4; 6, 6) - 23678235604296d(6; 8, 2) + 249393119976d(6; 6, 4)$
$+ 24337606022730d(8; 6, 2) - 6139661277750d(8; 4, 4) - 11064415962600d(10; 4, 2)$

$+\,2501333850015d(12;2,2) + 6429569526380[2,2,8,2] + 40400609689440[2,6,4,2]$
$-\,82392213676500[4,4,4,2]$

This relation is proportional to

$$-\frac{360}{3617}h^4_{16,2} + 2h^4_{14,4} - 7h^4_{12,6} + 11h^4_{10,8} + \dots,$$

again as expected.

5.3. **Computations of $\tilde{z}_j$ actions.** In this section we give formulas for $\tilde{z}_3$ and $\tilde{z}_5$ as trees, show the results of some of the computations of $[\tilde{z}_3, \epsilon_n]$ for a few small n, and discuss some intriguing congruences between the commutators $[\tilde{z}_3, \epsilon_n]$ and the depth 4 relations.

Recall that $\tilde{z}_j$, j odd > 1 is of type $(j, 2j)$, is $\mathfrak{sl}_2$ invariant, and normalizes $\mathfrak{u}$. This last condition is quite strong and can be used to compute the $\tilde{z}_j$ (up to scalar multiples, of course).

Just by considering the weight, depth, and $\mathfrak{sl}_2$ properties of $\tilde{z}_3$, Hain determined that $\tilde{z}_3$ must be given by the tree $\theta\#((b\theta a)\cdot\theta)$. A tree for $\tilde{z}_5$ is given by $\theta\#(v\cdot\theta)$, where v in the tensor algebra $T(H)$ is given by

$$\begin{aligned} v = &-11b^2a^4b^2 - 29ab^3aba^2 + 16ab^3a^3b + 101ab^2ab^2a^2 - 364ab^2aba^2b \\ &+ 148ab^2a^2bab - 81abab^3a^2 + 205abab^2a^2b + 15a^2babab^2. \end{aligned} \tag{13}$$

Note that there are multiple nonzero $k \in T(H)$ with $\theta\#(k\cdot\theta) = 0$, so this v is far from unique.

Equation (13) was obtained as follows. First, note that every $\delta \in \mathrm{Der}^0_n \mathbb{L}(H)$ with $n \geq 2$ can be written as a tree in the form $\theta\#(u\cdot\theta)$ for some u in $T(H)$. A spanning set for $\mathrm{Der}^0 \mathbb{L}_{5,10}(H)$ is thus the set of $\delta = \theta\#(u\cdot\theta)$ with u a monomial in $T(H)$ with 4 a's and 4 b's, i.e. of depth 4 and weight 8. This spanning set can be reduced to a basis for $\mathrm{Der}^0 \mathbb{L}_{5,10}(H)$ by eliminating derivations in the span of other ones. Proceeding in this way, one finds that the derivations $M_i = \theta\#(m_i\cdot\theta)$ for $\{m_i\}$ the set of monomial terms in Equation (13) gives such a basis. To find $\tilde{z}_5$, it is enough to impose the condition $[\tilde{z}_5, \epsilon_4] \in \mathfrak{u}$. One imposes this constraint by finding a basis $\{R_j\}$ of $\mathfrak{u}_{6,14}$, which is where $[\tilde{z}_5, \epsilon_4]$ must live, and looking for linear relations between the M_i and R_j. There is 1 such relation, and hence $\tilde{z}_5$ is determined. As a check, one can verify (not by hand!) that the resulting derivation is $\mathfrak{sl}_2$ invariant by computing $[\epsilon_0, \tilde{z}_5] = 0$. With more labor, this procedure can be extended to find $\tilde{z}_7$, although we have not attempted this computation.

Below are the results of some of the computations $[\tilde{z}_3, \epsilon_n]$. We express the commutator $[\tilde{z}_3, \epsilon_n]$ as an element $v \in \mathbb{L}_{4,6+2n}(P)$ with $\Phi(v) = [\tilde{z}_3, \epsilon_n]$. The choice of v is only unique up to elements of $K_{4,6+n}$, thus the v picked for $n = 12$ was not the only such choice. The notation is as in the Examples section above, except that we have replaced E^d_n for $h^d_{d-2,n+d-4}$.

$[\tilde{z}_3, \epsilon_4]$:
$\frac{2}{7}E_4^4$

$[\tilde{z}_3, \epsilon_6]$:
$\frac{3}{10}E_6^4 + \frac{63}{100}d(2;2,2)$

$[\tilde{z}_3, \epsilon_8]$:
$\frac{10}{33}E_8^4 + \frac{3}{7}d(2;4,2) + \frac{3}{4}d(4;2,2)$

$[\tilde{z}_3, \epsilon_{10}]$:
$\frac{691}{2275}E_{10}^4 + \frac{99}{125}d(6;2,2) + \frac{44}{125}d(2;6,2) - \frac{22}{315}d(2;4,4) + \frac{22}{35}d(4;4,2)$

$[\tilde{z}_3, \epsilon_{12}]$:
$\frac{210}{691}E_{12}^4 + \frac{30979}{114015}d(2;8,2) - \frac{468}{3455}d(2;6,4) + \frac{429}{691}d(4;6,2) + \frac{429}{691}d(6;4,2) + \frac{1742}{2073}d(8;2,2) - \frac{819}{5528}[2,2,4,2]$

It appears as if the coefficient of E_n^4 in $[\tilde{z}_3, \epsilon_n]$ is always $\frac{6B_{n+2}}{\binom{n+2}{2}B_n}$ where B_k is the Bernoulli number.

Let us move on to the congruences hinted at above. Note that all the denominators in $[\tilde{z}_3, \epsilon_{12}]$ are divisible by 691 but not 691^2. Hence multiplying this expression by 691 one obtains a element v which has well-defined coefficients in $\mathbb{Z}/691\mathbb{Z}$. Denote by R_{12} the depth 4 relation corresponding to Δ listed above. It turns out that there is some $\eta \in (\mathbb{Z}/691\mathbb{Z})^*$ with $\eta v \cong R_{12}$ mod 691. Although we have not written down $[\tilde{z}_3, \epsilon_{16}]$, there is again a congruence with the depth 4, weight 16 relations above, this time modulo 3617. These congruences appear to be explained by Hain's result that the $[z_j, \epsilon_n]$ action corresponds to Eisenstein series, as their is a congruence $G_{12} \equiv \Delta$ mod 691 in their q-expansions and similarly one between G_{16} and the weight 16 cusp form mod 3617.

References

[1] Gangl, H., Kaneko, M., and Zagier, D., *Double zeta values and modular forms*, Automorphic forms and zeta functions, Proceedings of the conference in memory of Tsuneo Arakawa, World Scientific, 2006.

[2] Garoufalidis, S., and Levine, J., *Tree-level invariants of 3 manifolds, Massey products and the Jonhson homomorphism*, Graphs and Patterns in Mathematics and Theoretical Physics, Proceedings Symp. Pure Math. 73 (2005) 173–205.

[3] Hain, R., Matsumoto, M. *Universal mixed elliptic motives*, in preparation, 2009.

[4] Luque, J.-G., Novelli, J.-C., and Thibon, J.-Y, *Period Polynomials and Ihara Brackets*, Journal of Lie Theory, 17 (2007), 229–239.

[5] Schneps, L. *On the Poisson bracket on the free Lie algebra in two generators*, Journal of Lie Theory, 16 (2006), 019-037.

[6] Zagier, D. *Periods of modular forms and Jacobi theta functions*, Inventiones Mathematicae 104, 449-465, 1991.