Algebra & Number Theory

Volume 7

2013

No. 1

msp

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2013 is US $200/year for the electronic version, and $350/year (+$40, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.1

Powers of ideals and the cohomology of stalks and fibers of morphisms

Marc Chardin

À Jean-Pierre Jouanolou, avec admiration et amitié

We first provide here a very short proof of a refinement of a theorem of Kodiyalam and Cutkosky, Herzog and Trung on the regularity of powers of ideals. This result implies a conjecture of Hà and generalizes a result of Eisenbud and Harris concerning the case of ideals primary for the graded maximal ideal in a standard graded algebra over a field. It also implies a new result on the regularities of powers of ideal sheaves. We then compare the cohomology of the stalks and the cohomology of the fibers of a projective morphism to the effect of comparing the maximums over fibers and over stalks of the Castelnuovo–Mumford regularities of a family of projective schemes.

1. Introduction

An important result of Kodiyalam and Cutkosky, Herzog and Trung states that the Castelnuovo–Mumford regularity of the power I^t of an ideal over a standard graded algebra is eventually a linear function in t. The leading term of this function has been determined by Kodiyalam in his proof.

This result was first obtained for standard graded algebras over a field, and later extended by Trung and Wang to standard graded algebras over a Noetherian ring.

We first provide here a very short proof of a refinement of this result.

Theorem 1.1. *Let A be a positively graded Noetherian algebra, $M \neq 0$ be a finitely generated graded A-module, I be a graded A-ideal, and set*

$$d := \min\{\mu \mid \textit{there exists } p,\ (I_{\leq\mu})I^pM = I^{p+1}M\}.$$

Then

$$\lim_{t\to\infty}(\operatorname{end}(H^i_{A_+}(I^tM)) + i - td) \in \mathbb{Z}\cup\{-\infty\}$$

exists for any i, and is at least equal to the initial degree of M for some i.

MSC2000: primary 13D02; secondary 13A30, 13D45, 14A15.
Keywords: cohomology, stalks, Rees algebras, fibers of morphisms, powers of ideals, Castelnuovo–Mumford regularity.

The end of a graded module H is $\mathrm{end}(H) := \sup\{\mu \mid H_\mu \neq 0\}$ if $H \neq 0$ and $-\infty$ otherwise. Recall that for a graded A-module N, $\mathrm{reg}(N) = \max_i\{\mathrm{end}(H^i_{A_+}(N)) + i\}$.

Very interesting examples showing hectic behavior of the value of

$$a^i(t) := \mathrm{end}(H^i_{A_+}(I^t))$$

as t varies were given in [Cutkosky 2000]. These examples point out that the existence of the limit quoted above does not imply that all of the functions $a^i(t)$ are eventually linear functions of t. It only implies that at least one of them is eventually linear in t. For instance, in the examples given by Cutkosky, the limit in the theorem is $-\infty$ for all $i \neq 0$.

More recently, Eisenbud and Harris proved that in the case of a standard graded algebra A over a field, for a graded ideal that is A_+-primary and generated in a single degree, the constant term in the linear function is the maximum of the regularity of the fibers of the morphism defined by a set of minimal generators. In a recent preprint, Huy Tài Hà [2011, 1.3] generalized this result by proving that if an ideal is generated in a single degree d, a variant of the regularity (the a^*-invariant) satisfies $a^*(I^t) = dt + a$ for $t \gg 0$, where a can be expressed in terms of the maximum of the values of a^* on the stalks of the projection π from the closure of the graph of the map defined by the generators to its image. He conjectures that a similar result holds for the regularity.

In Theorem 5.3 we prove this conjecture of Hà. More precisely, we show that the limit in the theorem above is the maximum of the end degree of the i-th local cohomology of the stalks of π, for ideals generated in a single degree. This holds for graded ideals in a Noetherian positively graded algebra.

An interesting, and perhaps surprising, consequence of this result is the following result on the limit of the regularity of saturation of powers, or equivalently of powers of ideal sheaves, in a positively graded Noetherian algebra:

Corollary 1.2. *Let I be a graded ideal generated in a single degree d. Then,*

$$\lim_{t\to\infty}(\mathrm{reg}((I^t)^{\mathrm{sat}}) - dt)$$

exists and the following are equivalent:

(i) *the limit is nonnegative*,

(ii) *the limit is not $-\infty$*,

(iii) *the projection π from the closure of the graph of the function defined by minimal generators of I to its image admits a fiber of positive dimension.*

This can be applied to ideals generated in degree at most d, replacing I by $I_{\geq d}$.

It gives a simple geometric criterion for an ideal I generated in degree (at most) d to satisfy $\mathrm{reg}((I^t)^{\mathrm{sat}}) = dt + b$ for $t \gg 0$: This holds if and only if there exists a subvariety V of the closure of the graph that is contracted in its projection to the closure of the image (that is, $\dim(\pi(V)) < \dim V$). A very simple example is the following. In a polynomial ring in $n+1$ variables, any graded ideal generated by n forms of the same degree d satisfies $\operatorname{reg}((I^t)^{\mathrm{sat}}) = dt + b$ for $t \gg 0$, with $b \geq 0$. The same result holds if a reduction of the ideal is generated by at most n elements (in other words, if the analytic spread of I is at most n).

The result of Eisenbud and Harris is stated in terms of regularity of fibers. For a finite morphism, there is no difference between the regularity of stalks and the regularity of fibers. This follows from the following result that is likely part of folklore, but that we didn't find in several of the classical references in the field:

Lemma 1.3. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring, $S := R[X_1, \ldots, X_n]$ be a polynomial ring over R with $\deg X_i > 0$ and $\mathbb{M}$ be a finitely generated graded S-module. Set $d := \dim(\mathbb{M} \otimes_R k)$. Then $H^i_{S_+}(\mathbb{M}) = 0$ for $i > d$ and the natural graded map $H^d_{S_+}(\mathbb{M}) \otimes_R k \to H^d_{S_+}(\mathbb{M} \otimes_R k)$ is an isomorphism.*

For morphisms that are not finite or flat, the situation is more subtle — see Proposition 6.3. We show that for families of projective schemes that are close to being flat (if the Hilbert polynomial of any two fibers differ at most by a constant, in the standard graded situation), the maximum of the regularities of stalks and the maximum of the regularities of fibers agree. Also the maximum regularity of stalks bounds above the one for fibers under a weaker hypothesis. Putting this together provides a collection of results that covers the results obtained in [Eisenbud and Harris 2010; Hà 2011]. See Theorem 6.11.

To simplify the statements, we introduce the notion of regularity over a scheme, generalizing the usual notion of regularity with reference to a polynomial extension of a ring. This is natural in our situation: The family of schemes given by the closure of the graph over the parameter space given by the closure of the image of our map, considered as a projective scheme, is a key ingredient of this study.

2. Notation and general setup

Let R be a commutative ring and S a polynomial ring over R in finitely many variables.

If S is $\mathbb{Z}$-graded, $R \subset S_0$, and $X_1, \ldots, X_n$ are the variables with positive degrees, the Čech complex $\mathscr{C}^\bullet_{(S_+)}(M)$ with

$$\mathscr{C}^0_{(S_+)}(M) = M \quad \text{and} \quad \mathscr{C}^i_{(S_+)}(M) = \bigoplus_{j_1 < \cdots < j_i} M_{X_{j_1} \cdots X_{j_i}} \quad \text{for } i > 0$$

is graded whenever M is a graded S-module.

There is an isomorphism $H^i_{(S_+)}(M) \simeq H^i(\mathscr{C}^\bullet_{(S_+)}(M))$ for all i, which is graded if M is. One then defines two invariants attached to such a graded S-module M:

$$a^i(M) := \sup\{\mu \mid H^i_{(S_+)}(M)_\mu \neq 0\}$$

if $H^i_{(S_+)}(M) \neq 0$ and $a^i(M) := -\infty$ otherwise, and

$$b_j(M) := \sup\{\mu \mid \operatorname{Tor}^S_j(M, S/(S_+))_\mu \neq 0\}$$

if $\operatorname{Tor}^S_j(M, S/(S_+)) \neq 0$ and $b_j(M) := -\infty$ otherwise. Notice that $a^i(M) = -\infty$ for $i > n$ and $b_j(M) = -\infty$ for $j > n$. The Castelnuovo–Mumford regularity of a graded S-module M is then defined as

$$\operatorname{reg}(M) := \max_i\{a^i(M) + i\} = \max_j\{b_j(M) - j\} + n - \sigma$$

where σ is the sum of the degrees of the variables with positive degrees. Other options are possible, in particular when S is not standard graded (when $\sigma \neq n$). Another related invariant is

$$a^*(M) := \max_i\{a^i(M)\} = \max_j\{b_j(M)\} - \sigma.$$

The following classical result is usually stated for positive grading.

Theorem 2.1. *Let S be a finitely generated $\mathbb{Z}$-graded algebra over a Noetherian ring $R \subseteq S_0$ and M be a finitely generated graded S-module. Assume S is generated over R by elements of nonzero degree. Then, for any i,*

(i) $a^i(M) \in \{-\infty\} \cup \mathbb{Z}$,

(ii) *the R-module $H^i_{(S_+)}(M)_\mu$ is finitely generated for any $\mu \in \mathbb{Z}$.*

Proof. S is an epimorphic image of a polynomial ring S' over R by a graded morphism. Considering M as an S'-module, one has $H^i_{(S_+)}(M) \simeq H^i_{(S'_+)}(M)$ via the natural induced map, so that we may replace S by S' and assume that

$$S = R[Y_1, \dots, Y_m, X_1, \dots, X_n]$$

with $\deg Y_i \leq -1$ and $\deg X_j \geq 1$ for all i and j. We recall that $H^i_{(S_+)}(S) = 0$ for $i < n$ and $H^n_{(S_+)}(S) = (X_1 \cdots X_n)^{-1} R[Y_1, \dots, Y_m, X_1^{-1}, \dots, X_n^{-1}]$, and notice that $H^n_{(S_+)}(S)_\mu$ is a finitely generated free R-module for any μ.

Let $F_\bullet$ be a graded free S-resolution of M with F_i finitely generated. Both spectral sequences associated to the double complex $\mathscr{C}^\bullet_{(S_+)} F_\bullet$ degenerate at step 2 and provide graded isomorphisms

$$H^i_{(S_+)}(M) \simeq H_{n-i}(H^n_{(S_+)}(F_\bullet)),$$

which shows that $H^i_{(S_+)}(M)_\mu$ is a subquotient of $H^n_{(S_+)}(F_{n-i})_\mu$ and hence a finitely generated R-module that is zero in degrees greater than $-n+b_{n-i}$, where b_j is the highest degree of a basis element of F_j over S. $\square$

3. Regularity over a scheme

Local cohomology and the torsion functor commute with localization on the base R, providing natural graded isomorphisms for a graded S-module M:

$$H^i_{(S\otimes_R R_\mathfrak{p})_+}(M\otimes_R R_\mathfrak{p}) \simeq H^i_{S_+}(M)\otimes_R R_\mathfrak{p}$$

and

$$\mathrm{Tor}_i^{S\otimes_R R_\mathfrak{p}}(M\otimes_R R_\mathfrak{p}, R_\mathfrak{p}) \simeq \mathrm{Tor}_i^S(M, R)\otimes_R R_\mathfrak{p}.$$

Hence $a^i(M) = \sup_{\mathfrak{p}\in\mathrm{Spec}(R)} a^i(M\otimes_R R_\mathfrak{p})$ and $b_j(M) = \sup_{\mathfrak{p}\in\mathrm{Spec}(R)} b_j(M\otimes_R R_\mathfrak{p})$. It follows that the regularity is a local notion on R:

$$\mathrm{reg}(M) = \sup_{\mathfrak{p}\in\mathrm{Spec}(R)} \mathrm{reg}(M\otimes_R R_\mathfrak{p}).$$

These supremums are maximums whenever $\mathrm{reg}(M) < +\infty$, for instance if R is Noetherian and M is finitely generated. The same holds for $a^*(M)$.

In the following, this definition is extended in a natural way to the case where the base is a scheme.

Definition 3.1. Let Y be a scheme, $\mathscr{E}$ be a locally free $\mathscr{O}_Y$-module of finite rank, and $\mathscr{F}$ be a graded sheaf of $\mathrm{Sym}_Y(\mathscr{E})$-modules. Then

$$a^i(\mathscr{F}) := \sup_{y\in Y} a^i(\mathscr{F}\otimes_{\mathscr{O}_Y}\mathscr{O}_{Y,y}) \quad \text{and} \quad \mathrm{reg}(\mathscr{F}) := \max_i\{a^i(\mathscr{F})+i\}.$$

If $\mathscr{E}$ is free, $\mathrm{Sym}_Y(\mathscr{E}) = \mathscr{O}_Y[X_1,\dots,X_n]$, and the definition of regularity above makes sense for nonstandard grading.

A closed subscheme Z of $\mathrm{Proj}(\mathrm{Sym}_Y(\mathscr{E}))$ corresponds to $\mathscr{I}_Z$, a unique graded $\mathrm{Sym}_Y(\mathscr{E})$-ideal sheaf saturated with respect to $\mathrm{Sym}_Y(\mathscr{E})_+$. We set

$$a^i(Z) := \sup_{y\in Y} a^i(\mathscr{O}_{Y,y}[X_0,\dots,X_n]/(\mathscr{I}_Z\otimes_{\mathscr{O}_Y}\mathscr{O}_{Y,y}))$$

(notice that $a^0(Z) = -\infty$) and $\mathrm{reg}(Z) := \max_i\{a^i(Z)+i\}$.

The following proposition is immediate from the definition and the corresponding results over an affine scheme.

Proposition 3.2. *Assume Y is Noetherian, $\mathscr{E}$ is a locally free coherent sheaf on Y and $\mathscr{F}\neq 0$ is a coherent graded sheaf of* $\mathrm{Sym}_Y(\mathscr{E})$*-modules. Then* $\mathrm{reg}(\mathscr{F})\in\mathbb{Z}$. *If $Z\neq\varnothing$ is a closed subscheme of $\mathbb{P}^{n-1}_Y$, then* $\mathrm{reg}(Z)\geq 0$.

4. First result on cohomology of powers

We now prove the first statement of our text on cohomology of powers of ideals. It refines earlier results on the regularity of powers [Kodiyalam 2000; Cutkosky et al. 1999; Trung and Wang 2005]. The argument is based on Theorem 2.1 applied to a Rees algebra and a lemma due to Kodiyalam.

Theorem 4.1. *Let A be a positively graded Noetherian algebra, $M \neq 0$ be a finitely generated graded A-module, I be a graded A-ideal, and set*

$$d := \min\{\mu \mid \text{there exists } p,\ (I_{\leq\mu})I^pM = I^{p+1}M\}.$$

Then

$$\lim_{t\to\infty}(a^i(I^tM)+i-td) \in \mathbb{Z}\cup\{-\infty\}$$

exists for any i, and is at least equal to $\operatorname{indeg}(M)$ *for some i.*

Proof. Set $J := I_{\leq d}$ and write $J = (g_1, \dots, g_s)$ with $\deg g_i = d$ for $1 \leq i \leq m$ and $\deg g_i < d$ otherwise. Let

$$\mathcal{R}_J := \bigoplus_{t\geq 0} J(d)^t = \bigoplus_{t\geq 0} J^t(td) \quad \text{and} \quad \mathcal{R}_I := \bigoplus_{t\geq 0} I(d)^t = \bigoplus_{t\geq 0} I^t(td),$$

and $S_0 := A_0[T_1, \dots, T_m]$, $S := S_0[T_{m+1}, \dots, T_s, X_1, \dots, X_n]$, with $\deg(T_i) := \deg(g_i) - d$. Setting $\operatorname{bideg}(T_i) := (\deg(T_i), 1)$ and $\operatorname{bideg}(X_j) := (\deg(X_j), 0)$, one has $J_{\deg(g_i)} = (\mathcal{R}_J)_{\deg g_i - d, 1}$ and hence a bigraded onto map

$$S \to \mathcal{R}_J, \quad T_i \mapsto g_i.$$

As $M\mathcal{R}_I$ is finite over $\mathcal{R}_J$ according to the definition of d, the bigraded embedding $\mathcal{R}_J \to \mathcal{R}_I$ makes $M\mathcal{R}_I$ a finitely generated bigraded S-module.

The equality of graded A-modules $H^i_{(S_+)}(M\mathcal{R}_I)_{(*,t)} = H^i_{A_+}(M\mathcal{R}_I)_{(*,t)}$ shows that

$$H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu,t)} = H^i_{A_+}((M\mathcal{R}_I)_{(*,t)})_\mu = H^i_{A_+}(MI^t)_{\mu+td}.$$

By Theorem 2.1(i), $a^i(M\mathcal{R}_I) < +\infty$ and the equalities above show

$$a^i(MI^t) \leq td + a^i(M\mathcal{R}_I),$$

and that equality holds for some t.

Furthermore, Theorem 2.1(ii) shows that $K_{i,\mu} := H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu,*)}$ is a finitely generated graded S_0-module (for the standard grading $\deg(T_i) = 1$). It follows that $H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu,t)} = 0$ for $t \gg 0$ if and only if $K_{i,\mu}$ is annihilated by a power of $\mathfrak{n} := (T_1, \dots, T_m)$. Hence

$$\lim_{t\to+\infty}(a^i(MI^t) - td) = -\infty$$

if $K_{i,\mu}$ is annihilated by a power of $\mathfrak{n}$ for every $\mu \leq a^i(M\mathcal{R}_I)$, and otherwise

$$\lim_{t\to+\infty}(a^i(MI^t)-td)=\max\{\mu \mid K_{i,\mu}\neq H^0_{\mathfrak{n}}(K_{i,\mu})\}.$$

As $\operatorname{reg}(MI^t)\geq \operatorname{end}(MI^t/A_+MI^t)$, the last claim follows from the next lemma, due to Kodiyalam. □

Lemma 4.2. *With the hypotheses of Theorem 4.1,*

$$\operatorname{end}(MI^t/A_+MI^t)\geq \operatorname{indeg}(M)+td \quad \textit{for all } t.$$

Proof. The proof goes along the same lines as in the proof of [Kodiyalam 2000, Proposition 4]. The needed graded version of Nakayama's lemma does apply. □

5. Cohomology of powers and cohomology of stalks

The following result is a more elaborated, and more technical, version of Theorem 4.1 that essentially follows from its proof. It implies a conjecture of Hà on the regularity of powers of ideals, and refines the main result in [Hà 2011]. We will see later that, combined with a result on the regularity of stalks and fibers of a morphism, it also implies the result in [Eisenbud and Harris 2010].

Proposition 5.1. *Let A be a positively graded Noetherian algebra, M be a finitely generated graded A-module, I be a graded A-ideal and $J\subseteq I$ be a graded ideal such that $JI^pM=I^{p+1}M$ for some p.*

Assume that the ideal J is generated by r forms $f_1,\dots,f_r$ of respective degrees $d_1=\cdots=d_m>d_{m+1}\geq\cdots\geq d_r$. Set $d:=d_1$, $\deg(T_i):=\deg(f_i)-d$, $\operatorname{bideg}(T_i):=(\deg(T_i),1)$ and $\operatorname{bideg}(a):=(\deg(a),0)$ for $a\in A$. Consider the natural bigraded morphism of bigraded A_0-algebras

$$S:=A[T_1,\dots,T_r]\xrightarrow{\psi}\mathcal{R}_I:=\bigoplus_{t\geq 0}I(d)^t=\bigoplus_{t\geq 0}I^t(dt),$$

sending T_i to f_i, and the bigraded map of S-modules

$$M[T_1,\dots,T_r]\xrightarrow{1_M\otimes_A\psi}M\mathcal{R}_I:=\bigoplus_{t\geq 0}MI^t(dt).$$

Let $B:=A_0[T_1,\dots,T_m]$ and $B':=B/\operatorname{ann}_B(\ker(1_M\otimes_A\psi))$.

Then,

$$\lim_{t\to+\infty}(a^i(MI^t)-td)=\max_{\mathfrak{q}\in\operatorname{Proj}(B')}\{a^i(M\mathcal{R}_I\otimes_{B'}B'_{\mathfrak{q}})\}.$$

Proof. First remark that in the proof of Theorem 4.1 we only need the equality $JI^pM=I^{p+1}M$ for some p (as a consequence, for all p big enough). We have shown there that

$$\lim_{t\to+\infty}(a^i(MI^t)-td)=-\infty, \qquad (*)$$

if and only if the finitely generated B-module $H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu,*)}$ is supported in $V(T_1,\ldots,T_m)$ for any μ. As local cohomology commutes with flat base change and elements in B have degree 0,

$$H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu,*)}\otimes_{B'} B'_{\mathfrak{q}} = H^i_{(S_+)}(M\mathcal{R}_I\otimes_{B'} B'_{\mathfrak{q}})_{(\mu,*)};$$

hence $(*)$ holds if and only if $H^i_{(S_+)}(M\mathcal{R}_I\otimes_{B'} B'_{\mathfrak{q}})=0$ for any $\mathfrak{q}\in \mathrm{Proj}(B')$. On the other hand, if this does not hold, there exists μ_0 the maximum value such that $H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu_0,*)}$ is not supported in $V(T_1,\ldots,T_m)$, and choosing $\mathfrak{q}\in\mathrm{Proj}(B')\cap \mathrm{Supp}(H^i_{(S_+)}(M\mathcal{R}_I)_{(\mu_0,*)})$ shows that both members in the asserted equality are equal to μ_0. □

Remark 5.2. In the proposition above, as well as in other places in this text, we localize at homogeneous primes $\mathfrak{q}\in\mathrm{Proj}(C)$ for some standard graded algebra C, in other words, at graded prime ideals that do not contain C_+. We may as well replace these localizations by the degree zero part of the localization at such a prime ideal, usually denoted by $C_{(\mathfrak{q})}$: The multiplication by an element $\ell\in C_1\setminus\mathfrak{q}$ induces an isomorphism $(C_{\mathfrak{q}})_\mu\simeq (C_{\mathfrak{q}})_{\mu+1}$ for any μ. Hence, for any C-module M, $M\otimes_C C_{\mathfrak{q}}=0$ if and only if $M\otimes_C C_{(\mathfrak{q})}=0$.

In the equal degree case, the following corollary, which we state in a more geometric fashion, implies the conjecture of Hà [2011].

Theorem 5.3. *Let $A:=A_0[x_0,\ldots,x_n]$ be a positively graded Noetherian algebra and I be a graded A-ideal generated by $m+1$ forms of degree d. Set $Y:=\mathrm{Spec}(A_0)$ and $X:=\mathrm{Proj}(A/I)\subset\mathrm{Proj}(A)\subseteq\tilde{\mathbb{P}}^n_Y$. Let $\phi:\tilde{\mathbb{P}}^n_Y\setminus X\to\mathbb{P}^m_Y$ be the corresponding rational map, W be the closure of the image of ϕ, and*

$$\Gamma\subset\tilde{\mathbb{P}}^n_W\subseteq\tilde{\mathbb{P}}^n_{\mathbb{P}^m_Y}=\tilde{\mathbb{P}}^n_Y\times_Y\mathbb{P}^m_Y$$

be the closure of the graph of ϕ. Let $\pi:\Gamma\to W$ be the projection induced by the natural map $\tilde{\mathbb{P}}^n_{\mathbb{P}^m_Y}\to\mathbb{P}^m_Y$. Then

$$\lim_{t\to+\infty}(a^i(I^t)-dt)=a^i(\Gamma).$$

Proof. Choose $J:=I$ and $M:=A$ in Proposition 5.1. The equality

$$\lim_{t\to+\infty}(a^i(I^t)-dt)=a^i(\Gamma)$$

directly follows from the conclusion of Proposition 5.1 according the definition of $a^i(\Gamma)$ for $\Gamma\subset\tilde{\mathbb{P}}^n_W$ given in Definition 3.1. □

6. Cohomology of stalks and cohomology of fibers

We will now compare the cohomology of stalks and of fibers of a projective morphism, in order to compare their Castelnuovo–Mumford regularities. It will need results on the support of Tor modules. These are likely part of folklore. However, we include a proof as we did not find a reference that properly fits our exact need.

Lemma 6.1. *Let $R \to S$ be a homomorphism of Noetherian rings, $\mathbb{M}$ be a finitely generated S-module and N be a finitely generated R-module.*

Then the S-modules $\operatorname{Tor}_q^R(\mathbb{M}, N)$ are finitely generated over S and

(i) $\operatorname{Supp}_S(\operatorname{Tor}_q^R(\mathbb{M}, N)) \subseteq \operatorname{Supp}_S(\mathbb{M} \otimes_R N)$ *for any q,*

(ii) *if further $(R, \mathfrak{m})$ is local, $S = R[X_1, \dots, X_n]$, with $\deg X_i > 0$ and $\mathbb{M}$ is a graded S-module, then $\operatorname{Supp}_S(\operatorname{Tor}_q^R(\mathbb{M}, R/\mathfrak{m})) \subseteq \operatorname{Supp}_S(\operatorname{Tor}_1^R(\mathbb{M}, R/\mathfrak{m}))$ for any $q \geq 1$.*

Proof. First the modules $\operatorname{Tor}_q^R(\mathbb{M}, N)$ are finitely generated over S by [Bourbaki 1980, X §6 N°4 Corollaire]. Second,

$$\operatorname{Supp}_S(\mathbb{M} \otimes_R N) = \operatorname{Supp}_S(\mathbb{M}) \cap \varphi^{-1}(\operatorname{Supp}_R(N)),$$

where $\varphi : \operatorname{Spec}(S) \to \operatorname{Spec}(R)$ is the natural map induced by $R \to S$, by [Bourbaki 1985, II §4 N°4, Propositions 18 and 19], since $\mathbb{M} \otimes_R N = \mathbb{M} \otimes_S (N \otimes_R S)$. For $\mathfrak{P} \in \operatorname{Spec}(S)$, set $\mathfrak{p} := \varphi(\mathfrak{P})$. Then $\operatorname{Tor}_q^R(\mathbb{M}, N)_{\mathfrak{P}} = \operatorname{Tor}_q^{R_{\mathfrak{p}}}(\mathbb{M}_{\mathfrak{P}}, N_{\mathfrak{p}})$ vanishes if either $\mathbb{M}_{\mathfrak{P}} = 0$ or $N_{\mathfrak{p}} = 0$.

For (ii), we can reduce to the case of a local morphism by localizing S and $\mathbb{M}$ at $\mathfrak{m} + S_+$. In this local situation, $\operatorname{Tor}_1^R(\mathbb{M}, R/\mathfrak{m}) = 0$ if and only if $\mathbb{M}$ is A-flat by [André 1974, Lemme 58], which proves our claim by localization at primes $\mathfrak{P}$ such that $\varphi(\mathfrak{P}) = \mathfrak{m}$. □

Let R be a commutative ring, N be a R-module, $S := R[X_1, \dots, X_n]$ be a positively graded polynomial ring over R and $\mathbb{M}$ be a graded S-module. For a S-module $\mathbb{M}$, we will denote by $\operatorname{cd}_{S_+}(\mathbb{M})$ the cohomological dimension of $\mathbb{M}$ with respect to S_+, which is the maximal index i such that $H^i_{S_+}(\mathbb{M}) \neq 0$ (and $-\infty$ if all these local cohomology groups are 0). The following lemma is a natural way for comparing cohomology of stalks to cohomology of fibers.

Lemma 6.2. *There are two converging spectral sequences of graded S-modules with the same abutment $H^\bullet$ and with respective second terms*

$${}'_2E^p_q = H^p_{S_+}(\operatorname{Tor}_q^R(\mathbb{M}, N)) \Rightarrow H^{p-q} \quad \text{and} \quad {}''_2E^p_q = \operatorname{Tor}_q^R(H^p_{S_+}(\mathbb{M}), N) \Rightarrow H^{p-q}.$$

Let $d := \max\{i \mid H^i_{S_+}(\mathbb{M}\otimes_R N)\neq 0\}$. *If* R *is Noetherian*, N *is finitely generated over* R *and* $\mathbb{M}$ *is finitely generated over* S, *then*

$$H^d_{S_+}(\mathbb{M}\otimes_R N)\simeq H^d_{S_+}(\mathbb{M})\otimes_R N$$

and $\mathrm{Tor}^R_q(H^i_{S_+}(\mathbb{M}), N) = H^i_{S_+}(\mathrm{Tor}^R_q(\mathbb{M}, N)) = 0$ *for any* q *if* $i > d$.

Proof. Let $F_\bullet$ be a free R-resolution of N. Consider the double complex

$$\mathscr{C}^\bullet_{S_+}(\mathbb{M}\otimes_R F_\bullet) = \mathscr{C}^\bullet_{S_+}(\mathbb{M})\otimes_R F_\bullet,$$

totalizing to $T^\bullet$ with $T^i = \bigoplus_{p-q=i}\mathscr{C}^p_{S_+}(\mathbb{M})\otimes_R F_q$. It gives rise to two spectral sequences abutting to the homology $H^\bullet$ of $T^\bullet$.

One has first terms $\mathscr{C}^p_{S_+}(\mathrm{Tor}^R_q(\mathbb{M}, N))$ and second terms $H^p_{S_+}(\mathrm{Tor}^R_q(\mathbb{M}, N))$.

The other spectral sequence has first terms $H^p_{S_+}(\mathbb{M})\otimes_R F_q$ and second terms $\mathrm{Tor}^R_q(H^p_{S_+}(\mathbb{M}), N)$. It provides the quoted spectral sequences.

Recall that if P is a finitely presented S-module, one has $\mathrm{cd}_{S_+}(P')\le \mathrm{cd}_{S_+}(P)$ whenever $\mathrm{Supp}(P')\subseteq \mathrm{Supp}(P)$. This is proved in [Divaani-Aazar et al. 2002, 2.2] under the assumption that S is Noetherian and P' is finitely generated, which is enough for our purpose.

By Lemma 6.1(i), $\mathrm{Supp}(\mathrm{Tor}^R_q(\mathbb{M}, N))\subseteq \mathrm{Supp}(\mathbb{M}\otimes_R N)$ for any q, which implies that $H^i_{S_+}(\mathrm{Tor}^R_q(\mathbb{M}, N)) = 0$ for any q if $i > d$. It follows that $H^d = H^d_{S_+}(\mathbb{M}\otimes_R N)$ and $H^i = 0$ for $i > d$.

On the other hand, choose i maximal such that $H^i_{S_+}(\mathbb{M})\otimes_R N\neq 0$. Then $\mathrm{Tor}^R_q(H^p_{S_+}(\mathbb{M}), N) = 0$ for any q if $p > i$, because $H^p_{S_+}(\mathbb{M})_\mu$ is a finitely generated R-module for every μ, and hence $H^i = H^i_{S_+}(\mathbb{M})\otimes_R N\neq 0$ and $H^j = 0$ for $j > i$. The conclusion follows. □

The following statement extends a classical result on the cohomology of fibers in a flat family; see for instance [Hartshorne 1977, III 9.3]. The hypothesis on the cohomological dimension of Tor modules that appears in (ii) will be connected to the variation of the Hilbert polynomial of fibers in the corresponding family of sheaves in Lemma 6.6; it is a weakening of the flatness condition for this family.

Proposition 6.3. *Let* $(R, \mathfrak{m}, k)$ *be a Noetherian local ring*, $S := R[X_1, \dots, X_n]$ *be a polynomial ring over* R, *with* $\deg X_i > 0$ *for all* i, *and* $\mathbb{M}$ *be a finitely generated graded* S*-module. Set* $\mathrm{M} := \mathbb{M}\otimes_R k$ *and* $d := \dim \mathrm{M}$. *Then one has the following*:

(i) *The natural graded map* $H^d_{S_+}(\mathbb{M})\otimes_R k \to H^d_{S_+}(\mathrm{M})$ *is an isomorphism and* $d = \max\{i \mid H^i_{S_+}(\mathbb{M})\neq 0\}$. *In particular*,

$$a^d(\mathbb{M}) = a^d(\mathrm{M})\in\mathbb{Z}.$$

(ii) *For any integers* μ *and* ℓ, *if* $\mathrm{cd}_{S_+}(\mathrm{Tor}^R_1(\mathbb{M}, k))\le \ell+1$ *then*

$$\{H^i_{S_+}(\mathbb{M})_\mu = 0 \text{ for all } i\ge \ell\} \quad \text{implies} \quad \{H^i_{S_+}(\mathrm{M})_\mu = 0 \text{ for all } i\ge \ell\},$$

and both conditions are equivalent if $\mathrm{cd}_{S_+}(\mathrm{Tor}_1^R(\mathbb{M},k)) \leq \ell$. *In particular,* $\mathrm{reg}(\mathrm{M}) \leq \mathrm{reg}(\mathbb{M})$ *if* $\mathrm{cd}_{S_+}(\mathrm{Tor}_1^R(\mathbb{M},k)) \leq 1$ *and equality holds if* $\mathrm{depth}_{S_+}(\mathbb{M}) > 0$.

Proof. We consider the two spectral sequences in Lemma 6.2,

$$ {}_2'E_q^p = H^p_{S_+}(\mathrm{Tor}_q^R(\mathbb{M},k)) \Rightarrow H^{p-q} \quad \text{and} \quad {}_2''E_q^p = \mathrm{Tor}_q^R(H^p_{S_+}(\mathbb{M}),k) \Rightarrow H^{p-q}. $$

Let $B := k[X_1,\dots,X_n]$. The module $\mathrm{Tor}_q^R(\mathbb{M},k)$ is a $R[X_1,\dots,X_n]$-module of finite type, annihilated by $\mathfrak{m}$ and $\mathrm{ann}_S(\mathbb{M})$. Hence M is a graded B-module of finite type and $\mathrm{Tor}_q^R(\mathbb{M},k)$ is a graded $(B/\mathrm{ann}_B(\mathrm{M}))$-module of finite type, for any q.

Notice that $d = \mathrm{cd}_{S_+}(\mathrm{M}) = \mathrm{cd}_{B_+}(\mathrm{M})$. It follows that ${}_2'E_q^p = 0$ if $p > d$, and ${}_2'E_0^d \neq 0$.

By Lemma 6.2, ${}_2''E_q^p = 0$ for all q if $p > d$, in particular $H^p_{S_+}(\mathbb{M})_\mu \otimes_R k = 0$ for any μ if $p > d$. Hence $H^p_{S_+}(\mathbb{M})_\mu = 0$ for any μ if $p > d$. In other words, $H^p_{S_+}(\mathbb{M}) = 0$ for any $p > d$.

The same lemma shows that $H^d_{S_+}(\mathrm{M}) = H^d_{S_+}(\mathbb{M}) \otimes_R k$, and finishes the proof of (i).

For (ii), let μ be an integer. We prove the result by descending induction on ℓ from the case $\ell = d$, which we already proved.

Assume the results hold for $\ell + 1$. Recall that, for any p, the maps

$$ {}_r'd_{1-r}^{p-r} : {}_r'E_{1-r}^{p-r} \to {}_r'E_0^p \quad \text{and} \quad {}_r''d_0^p : {}_r''E_0^p \to {}_r''E_{-r}^{p+1-r} $$

are the zero map for $r \geq 2$ and $r \geq 1$, respectively.

If $H^i_{S_+}(\mathbb{M})_\mu = 0$, for all $i \geq \ell$, then $({}_2''E_q^p)_\mu = 0$ for $p \geq \ell$ and all q. As ${}_2''E_q^p = 0$ for $q < 0$, it follows that $({}_2''E_q^p)_\mu = 0$ if $p - q \geq \ell$.

If $\mathrm{cd}_{S_+}(\mathrm{Tor}_1^R(\mathbb{M},k)) \leq \ell+1$ then ${}_2'E_q^p = 0$ for $p \geq \ell+2$ and $q > 0$ by Lemma 6.1(ii), in particular the map

$$ ({}_r'd_0^\ell)_\mu : ({}_r'E_0^\ell)_\mu \to ({}_r'E_{r-1}^{\ell+r})_\mu $$

is the zero map for any $r \geq 2$, and hence $H^\ell_{S_+}(\mathrm{M})_\mu = ({}_2'E_0^\ell)_\mu = ({}_\infty'E_0^\ell)_\mu = 0$ as claimed.

For the reverse implication, the hypothesis implies that ${}_2'E_q^p = 0$ if $q \geq 1$ and $p \geq \ell + 1$ by Lemma 6.1(ii). Hence $({}_2'E_q^p)_\mu = 0$ for $p - q \geq \ell$ if $H^\ell_{S_+}(\mathrm{M})_\mu = 0$. By induction hypothesis, $H^p_{S_+}(\mathbb{M})_\mu \otimes_R k = 0$ for $p \geq \ell + 1$. Hence

$$ ({}_2'E_q^p)_\mu = \mathrm{Tor}_q^R(H^p_{S_+}(\mathbb{M})_\mu, k) = 0 $$

for $p \geq \ell + 1$ and all q. It implies that $H^\ell_{S_+}(\mathbb{M})_\mu \otimes_R k = ({}_\infty''E_0^\ell)_\mu = 0$, and proves the claimed equivalence.

Finally, recall that $H^i_{S_+}(\mathbb{M}) = 0$ for $i < \mathrm{depth}_{S_+}(\mathbb{M})$. □

Remark 6.4. Notice that $\mathrm{reg}(\mathrm{M}) \leq \mathrm{reg}(\mathbb{M})$ does not hold without the hypothesis $\mathrm{cd}_{S_+}(\mathrm{Tor}_1^R(\mathbb{M},k)) \leq 1$. To see this, consider generic polynomials of some given degrees $d_1, \dots, d_r$:

$$P_i := \sum_{|\alpha|=d_i} U_{i,\alpha} X^\alpha \in k[U_{i,\alpha}][X_1, \dots, X_n],$$

with $r \le n$ and a specialization map $\phi : k[U_{i,\alpha}] \to k$ to the field k with kernel $\mathfrak{m}$. Set $R := k[U_{i,\alpha}]_{\mathfrak{m}}$. As the P_i form a regular sequence in $k[U_{i,\alpha}][X_1, \dots, X_n]$, they also form one in $S := R[X_1, \dots, X_n]$ and show that $\mathbb{M} := S/(P_1, \dots, P_r)$ has regularity $d_1 + \cdots + d_r - r$. On the other hand, the regularity of

$$M = k[X_1, \dots, X_n]/(\phi(P_1), \dots, \phi(P_r)),$$

need not be bounded by $d_1 + \cdots + d_r - r$.

For instance, with $n = 4$ and $r = 3$, take

$$\phi(P_1) := X_1^{d-1} X_2 - X_3^{d-1} X_4, \quad \phi(P_2) := X_2^d \quad \text{and} \quad \phi(P_3) := X_4^d$$

(over any field). Then one has $\operatorname{reg}(M) = d^2 - 2$ for $d \ge 3$ (see [Chardin 2007, 1.13.6]), which is bigger than $\operatorname{reg}(\mathbb{M}) = 3d - 3$, and $\operatorname{cd}_{S_+}(\operatorname{Tor}_1^R(\mathbb{M}, k)) = 2$.

Remark 6.5. In the other direction, it may of course be that $\operatorname{reg}(\mathbb{M}) > \operatorname{reg}(M)$. If for instance (R, π, k) is a DVR, one may take $\mathbb{M} := R[X]/(\pi X^d)$, so that $\operatorname{reg}(\mathbb{M}) = d - 1$ and $\operatorname{reg}(M) = 0$, with $\operatorname{cd}_{S_+}(\operatorname{Tor}_1^R(\mathbb{M}, k)) = 1$.

More interesting is the example $R := \mathbb{Q}[a, b]$, $\mathfrak{m} := (a, b)$ and

$$\mathbb{M} := \operatorname{Sym}_R(\mathfrak{m}^3) = R[X_1, \dots, X_4]/(bX_1 - aX_2, bX_2 - aX_3, bX_3 - aX_4).$$

Then for any morphism from R to a field k, $\operatorname{reg}(\mathbb{M} \otimes_R k) = 0$, while $\operatorname{reg}(\mathbb{M}) = 1$.

Similar examples arises from the symmetric algebra of other ideals that are not generated by a proper sequence.

The characterization of flatness in terms of the constancy of the Hilbert polynomial of fibers extends as follows.

Lemma 6.6. *Let p be an integer. In the setting of Proposition 6.3, assume that R is reduced and S is standard graded. Then the following are equivalent*:

(i) $\dim(\operatorname{Tor}_1^R(\mathbb{M}, k)) \le p$.

(ii) *The Hilbert polynomials of $\mathbb{M} \otimes_R k$ and $\mathbb{M} \otimes_R (R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}})$ differ at most by a polynomial of degree $< p$, for any $\mathfrak{p} \in \operatorname{Spec}(R)$.*

Proof. We induct on p. The result is standard when $p = 0$; see for instance [Hartshorne 1977, III 9.9; Eisenbud 1995, Exercise 20.14].

Assume (i) and (ii) are equivalent for $p - 1 \ge 0$, for any Noetherian local reduced ring, standard graded polynomial ring over it and graded module of finite type.

Set $K := R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$, $M_K := \mathbb{M} \otimes_R K$, $B := k[X_1, \dots, X_n]$ and $C := K[X_1, \dots, X_n]$. Consider variables $U_1, \dots, U_n$ (of degree 0) and let $\ell := U_1 X_1 + \cdots + U_n X_n$. By the Dedekind–Mertens lemma,

(a) $\ker(\mathbb{M}[U] \xrightarrow{\times\ell} \mathbb{M}[U](1)) \subseteq H^0_{S_+}(\mathbb{M})[U]$,

(b) $\ker(M[U] \xrightarrow{\times\ell} M[U](1)) \subseteq H^0_{B_+}(M)[U]$,

(c) $\ker(M_K[U] \xrightarrow{\times\ell} M_K[U](1)) \subseteq H^0_{C_+}(M_K)[U]$, and

(d) $\ker(\mathrm{Tor}_1^R(\mathbb{M}, k)[U] \xrightarrow{\times\ell} \mathrm{Tor}_1^R(\mathbb{M}, k)[U](1)) \subseteq H^0_{B_+}(\mathrm{Tor}_1^R(\mathbb{M}, k))[U]$.

Let $R' := R(U)$ be obtained from $R[U]$ by inverting all polynomials whose coefficient ideal is the unit ideal, and denote by N' the extension of scalars from R to R' for the module N. Recall that $R(U)$ is local reduced with maximal ideal $\mathfrak{m}R(U)$, residue field $k' = k(U)$ and that $K' = K(U)$ — see for instance [Nagata 1962, page 17]. As the zero local cohomology modules above vanish in high degrees, (b) and (c) show that $\mathbb{M}'/\ell\mathbb{M}'$ satisfies condition (ii) of the lemma for $p-1$, R' and $R'[X_1, \ldots, X_n]$. Now (a) and (d) provide an exact sequence for $\mu \gg 0$:

$$0 \longrightarrow \mathrm{Tor}_1^R(\mathbb{M}', k')_{\mu-1} \xrightarrow{\times\ell} \mathrm{Tor}_1^{R'}(\mathbb{M}', k')_\mu \longrightarrow \mathrm{Tor}_1^{R'}(\mathbb{M}'/\ell\mathbb{M}', k')_\mu \longrightarrow 0,$$

which shows in particular that

$$\dim \mathrm{Tor}_1^{R'}(\mathbb{M}'/\ell\mathbb{M}', k') = \dim \mathrm{Tor}_1^{R'}(\mathbb{M}', k') - 1 = \dim \mathrm{Tor}_1^R(\mathbb{M}, k) - 1,$$

if $\dim \mathrm{Tor}_1^R(\mathbb{M}, k)$ is positive, and proves our claim by induction. □

Remark 6.7. If the grading is not standard, a quasipolynomial is attached to any finitely generated graded module, and in Lemma 6.6 property (ii) should be replaced by the following:

(ii) The difference between the quasipolynomials of $\mathbb{M} \otimes_R k$ and $\mathbb{M} \otimes_R (R_\mathfrak{p}/\mathfrak{p}R_\mathfrak{p})$ is a quasipolynomial of degree $< p$ for any $\mathfrak{p} \in \mathrm{Spec}(R)$.

The degree of a quasipolynomial is the highest degree of the polynomials that define it. The proof of [Hartshorne 1977, III 9.9] extends to this case when $p = 0$, and our proof extends after a slight modification: in the proof that (ii) implies (i), one should take $\ell := U_1 X_1^{w/w_1} + \cdots + U_n X_n^{w/w_n}$, where $w_i := \deg(X_i)$ and $w := \mathrm{lcm}(w_1, \ldots, w_n)$.

The local statement of Lemma 6.6 implies a global statement, by comparing Hilbert functions at generic points of the components and at closed points. We state it below in a ring theoretic form.

Proposition 6.8. *Let p be an integer, R be a reduced commutative ring, S be a Noetherian positively graded polynomial ring over R and $\mathbb{M}$ be a finitely generated graded S-module. Then the following are equivalent:*

(i) $H^i_{S_+}(\mathrm{Tor}_1^R(\mathbb{M}, R/\mathfrak{m})) = 0$ *for all $i > p$ and $\mathfrak{m}$ maximal in* $\mathrm{Spec}(R)$.

(ii) *For any two ideals* $\mathfrak{p} \subset \mathfrak{q}$ *in* $\operatorname{Spec}(R)$, *the quasipolynomials of* $\mathbb{M} \otimes_R R/\mathfrak{p}$ *and* $\mathbb{M} \otimes_R R/\mathfrak{q}$ *differ by a quasipolynomial of degree* $< p$.

(iii) *Over a connected component of* $\operatorname{Spec}(R)$, *the quasipolynomials of two fibers differ by a quasipolynomial of degree* $< p$.

In parallel to the definition of the regularity over a scheme, we define the fiber-regularity $\mathfrak{f}\mathrm{reg}$ as the maximum over the fibers of their regularity.

Definition 6.9. In the setting of Definition 3.1,

$$\tilde{a}_i(\mathscr{F}) := \sup_{y \in Y} a^i(\mathscr{F} \otimes_{\mathbb{O}_Y} k(y)), \quad \mathfrak{f}\mathrm{reg}(\mathscr{F}) := \max_i \{\tilde{a}_i(\mathscr{F}) + i\},$$

and $\mathfrak{f}\mathrm{reg}(Z) := \max_{i \geq 1} \{\tilde{a}_i(\operatorname{Sym}_Y(\mathscr{E})/\mathscr{I}_Z) + i\}$.

Notice that $\mathfrak{f}\mathrm{reg}(\mathscr{F})$ is finite if Y is covered by finitely many affine charts and $\mathscr{F}$ is coherent. This holds since the regularity of a graded module over a polynomial ring over a field is bounded in terms of the number of generators and the degrees of generators and relations; see for instance [Chardin et al. 2008, 3.5].

We now return to the problem of studying the ending degree of local cohomologies of powers of a graded ideal I in a positively graded Noetherian algebra A.

From the comparison of cohomology of stalks and cohomology of fibers, we get from Theorem 5.3 the following result. As in Theorem 5.3 we use geometric language and do not introduce a graded module (or a sheaf) to make the exposition more simple. In case a more general statement is needed, it can be easily derived by using Proposition 5.1 in place of Theorem 5.3. The six statements are not independent, but each of them answers a question that is quite natural to ask. Notice that (iv) is essentially equivalent to one of the main results of Eisenbud and Harris [2010, 2.2].

Remark 6.10. It follows from Theorem 5.3 that the dimension of any fiber of the projection π of the graph to its image (see Theorem 5.3 or below for the precise definition of π) is bounded above by the cohomological dimension of A/I with respect to A_+.

Theorem 6.11. *Let* $A := A_0[x_0, \ldots, x_n]$ *be a positively graded Noetherian algebra and* I *be a graded* A*-ideal generated by* $m+1$ *forms of degree* d. *Set* $Y := \operatorname{Spec}(A_0)$ *and* $X := \operatorname{Proj}(A/I) \subset \operatorname{Proj}(A) \subseteq \tilde{\mathbb{P}}^n_Y$. *Let* $\phi : \tilde{\mathbb{P}}^n_Y \setminus X \to \mathbb{P}^m_Y$ *be the corresponding rational map*, W *be the closure of the image of* ϕ, *and*

$$\Gamma \subset \tilde{\mathbb{P}}^n_W \subseteq \tilde{\mathbb{P}}^n_{\mathbb{P}^m_Y} = \tilde{\mathbb{P}}^n_Y \times_Y \mathbb{P}^m_Y$$

be the closure of the graph of ϕ. *Let* $\pi : \Gamma \to W$ *be the projection induced by the natural map* $\tilde{\mathbb{P}}^n_{\mathbb{P}^m_Y} \to \mathbb{P}^m_Y$. *Then we have the following*:

(i) $\lim_{t \to +\infty}(\operatorname{reg}((I^t)^{\mathrm{sat}}) - dt) = \max_{i \geq 2}\{a^i(\Gamma) + i\}$.

(ii) *If π admits a fiber $Z \subseteq \tilde{\mathbb{P}}^n_{\operatorname{Spec}\mathfrak{K}}$ of dimension $i-1$, then*

$$\lim_{t\to\infty}(a^i(I^t)+i-td) \geq a^i(Z)+i = \tilde{a}^i(Z)+i \geq 0.$$

(iii) *Let δ be the maximal dimension of a fiber of π. Then,*

$$a^{\delta+1}(I^t)-td = a^{\delta+1}(\Gamma) = \tilde{a}^{\delta+1}(\Gamma) \quad \textit{for all } t \gg 0.$$

(iv) *If π is finite, for instance if $X = \varnothing$, then*

$$\operatorname{reg}(I^t) = a^1(I^t)+1 = \mathfrak{f}\mathrm{reg}(\Gamma)+td \quad \textit{for all } t \gg 0$$

and $\lim_{t\to\infty}(a^i(I^t)-td) = -\infty$ *for $i \neq 1$.*

(v) *If π has fibers of dimension at most one, for instance if the canonical map $X \to Y$ is finite, then*

$$\operatorname{reg}(I^t)-td = \operatorname{reg}(\Gamma) \geq \mathfrak{f}\mathrm{reg}(\Gamma) \quad \textit{for all } t \gg 0,$$

and $\lim_{t\to\infty}(a^i(I^t)-td) = -\infty$ *for $i \geq 2$.*

If furthermore A is standard graded and reduced, π has fibers of dimension one, all of same degree, then $\mathfrak{f}\mathrm{reg}(\Gamma) = \operatorname{reg}(\Gamma)$,

$$\lim_{t\to\infty}(a^1(I^t)-td) \geq \tilde{a}^1(\Gamma)$$

and equality holds if $\operatorname{reg}(I^t) = a^1(I^t)+1$ *for $t \gg 0$.*

(vi) *If A is reduced and, for every connected component T of W, the Hilbert quasipolynomials of fibers of π over any two points in* $\operatorname{Spec}(T)$ *differ by a periodic function, then*

$$\operatorname{reg}(I^t) = \mathfrak{f}\mathrm{reg}(\Gamma)+td \quad \textit{for all } \mu \gg 0.$$

Proof. Part (i) is a direct corollary of Theorem 5.3. Statements (ii), (iii) and (iv) follow from Theorem 5.3 and Proposition 6.3(i).

Statements (v) and (vi) follow from Theorem 5.3, Proposition 6.3(ii) — notice that $\operatorname{depth}_{S_+}(\mathcal{R}_I) \geq 1$ — and the equivalence of (i) and (iii) in Proposition 6.8 applied on the affine charts covering $\pi(\Gamma)$. □

Remark 6.12. Cutkosky, Ein and Lazarsfeld proved in [Cutkosky et al. 2001] that the limit $s(I) := \lim_{t\to\infty}\operatorname{reg}((I^t)^{\mathrm{sat}})/t$ exists and is equal to the inverse of a Seshadri constant, when A_0 is a field and A is standard graded.

Using the existence of c such that $\operatorname{reg}(MI^t) \leq dt+c$ for all t when I is generated in degree at most d and M is finitely generated, one can easily derive the existence of this limit in our more general setting. Indeed, let

$$r_p := \operatorname{reg}((I^p)^{\mathrm{sat}}) \quad \text{and} \quad d_p := \min\{\mu \mid (I^p)^{\mathrm{sat}} = ((I^p)^{\mathrm{sat}}_{\leq\mu})^{\mathrm{sat}}\}.$$

One has $d_{p+q} \leq d_p + d_q$; hence $s := \lim_{p\to\infty}(d_p/p)$ exists. For any p there exists c_p such that

$$\operatorname{reg}(((I^p)^{\mathrm{sat}}_{\leq d_p})^t I^q) \leq td_p + c_p \quad \text{for all } t \geq 1 \text{ and } 0 \leq q < p.$$

The inequalities $d_{pt+q} \leq r_{pt+q} \leq td_p + c_p$ show that $\lim_{p\to\infty}(r_p/p) = s$ and that $d_p \geq ps$ for all p.

The same argument applies to any graded ideal J such that $\operatorname{Proj}(A/J) \to Y$ is finite (that is, $\operatorname{cd}_{A_+}(A/J) \leq 1$). Setting $r_p^J := \operatorname{reg}(I^p :_A J^\infty) \leq \operatorname{reg}(I^p)$ and defining d_p^J similarly to the above,

$$d_p^J := \min\{\mu \mid ((I^p : J^\infty)_{\leq\mu}) : J^\infty = I^p : J^\infty\},$$

the limits of r_p^J/p and d_p^J/p exist and are equal. For example, if X is a scheme with isolated nonlocally complete intersection points, then $\lim_{p\to\infty}\operatorname{reg}(I^{(p)}/p)$ exists, where $I^{(p)}$ denotes the p-th symbolic power of I.

On the other hand, when A/J has cohomological dimension 2 it may be that $\operatorname{reg}(I : J^\infty) > \operatorname{reg}(I)$ for J an embedded prime of I. This shows that the argument above is not directly applicable for symbolic powers in general. It however implies that $s^J := \lim_{p\to\infty}(d_p^J/p)$ exists for any J and is equal to $\lim_{p\to\infty}(\rho_p^J/p)$, where

$$\rho_p^J := \min\{\operatorname{reg}(K) \mid K \subseteq (I^p : J^\infty),\, K : J^\infty = I^p : J^\infty\}.$$

Remark 6.13. If I is generated in degree at most d, Theorem 6.11 implies that $s(I) < d$ if and only if the morphism π corresponding to the ideal (I_d) is finite. More precisely, by Remark 6.12, π is finite if and only if $\operatorname{Proj}(A/I^t)$ is defined by equations of degree $< dt$ for some t, and if not, $\operatorname{reg}((I^t)^{\mathrm{sat}}) - td$ is a nonnegative constant for $t \gg 0$.

This has been remarked in [Niu 2013], using the definition of $s(I)$ as (the inverse of) a Seshadri constant.

Theorem 6.11 also has a consequence on the dimension of the fibers. Assume for simplicity that A_0 is a field. Set $X := \operatorname{Proj}(A/I)$, with I generated in degree at most d and let $0 \leq i \leq \dim X$.

Part (ii) in Theorem 6.11 then shows that the morphism π associated to (I_d) has no fiber of dimension greater than i if there exists $p \geq 1$ and an ideal K, generated in degree less than pd, such that $\operatorname{Proj}(A/I^p)$ and $\operatorname{Proj}(A/K)$ coincide locally at each point $x \in \mathbb{P}^n$ of dimension at least i. Indeed if this happens, then

$$H^j_{A_+}(A/I^{ps}) \simeq H^j_{A_+}(A/K^s) \quad \text{for all } j > i,\ s \geq 1,$$

and therefore there exists c_p such that $a^j(I^{ps}) \leq (pd-1)s + c_p$ for all s and $j \geq i$, showing that $\lim_{t\to\infty}(a^j(I^t) - td) = -\infty$ for $j \geq i$.

Acknowledgments

This work was inspired by results of Huy Tài Hà [2011] and of David Eisenbud and Joe Harris [2010]. Bernd Ulrich made remarks on a very early version of some of these results and motivated my study of the difference between the regularity of stalks and the regularity of fibers, and Joseph Oesterlé provided references concerning Lemma 6.1. It is my pleasure to thank them for their contribution.

References

[André 1974] M. André, *Homologie des algèbres commutatives*, Die Grundlehren der mathematischen Wissenschaften **206**, Springer, Berlin, 1974. MR 50 #4707 Zbl 0284.18009

[Bourbaki 1980] N. Bourbaki, *Éléments de mathématique: Algèbre, Chapitre 10: Algèbre homologique*, Masson, Paris, 1980. MR 82j:18022 Zbl 0455.18010

[Bourbaki 1985] N. Bourbaki, *Éléments de mathématique: Algèbre, Chapitres 1 á 4*, Masson, Paris, 1985. MR 86k:13001a Zbl 0547.13001

[Chardin 2007] M. Chardin, "Some results and questions on Castelnuovo–Mumford regularity", pp. 1–40 in *Syzygies and Hilbert functions*, Lect. Notes Pure Appl. Math. **254**, Chapman & Hall/CRC, Boca Raton, FL, 2007. MR 2008c:13023 Zbl 1127.13014

[Chardin et al. 2008] M. Chardin, A. L. Fall, and U. Nagel, "Bounds for the Castelnuovo–Mumford regularity of modules", *Math. Z.* **258**:1 (2008), 69–80. MR 2009a:13023 Zbl 1138.13007

[Cutkosky 2000] S. D. Cutkosky, "Irrational asymptotic behaviour of Castelnuovo–Mumford regularity", *J. Reine Angew. Math.* **522** (2000), 93–103. MR 2002c:14004 Zbl 0951.14001

[Cutkosky et al. 1999] S. D. Cutkosky, J. Herzog, and N. V. Trung, "Asymptotic behaviour of the Castelnuovo–Mumford regularity", *Compositio Math.* **118**:3 (1999), 243–261. MR 2000f:13037 Zbl 0974.13015

[Cutkosky et al. 2001] S. D. Cutkosky, L. Ein, and R. Lazarsfeld, "Positivity and complexity of ideal sheaves", *Math. Ann.* **321**:2 (2001), 213–234. MR 2002j:14020 Zbl 1029.14022

[Divaani-Aazar et al. 2002] K. Divaani-Aazar, R. Naghipour, and M. Tousi, "Cohomological dimension of certain algebraic varieties", *Proc. Amer. Math. Soc.* **130** (2002), 3537–3544. MR 2003j:13023 Zbl 0998.13007

[Eisenbud 1995] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, New York, 1995. MR 97a:13001 Zbl 0819.13001

[Eisenbud and Harris 2010] D. Eisenbud and J. Harris, "Powers of ideals and fibers of morphisms", *Math. Res. Lett.* **17**:2 (2010), 267–273. MR 2011e:14003 Zbl 1226.13012

[Hà 2011] H. T. Hà, "Asymptotic linearity of regularity and a^*-invariant of powers of ideals", *Math. Res. Lett.* **18**:1 (2011), 1–9. MR 2012c:13040 Zbl 1239.13027

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Kodiyalam 2000] V. Kodiyalam, "Asymptotic behaviour of Castelnuovo–Mumford regularity", *Proc. Amer. Math. Soc.* **128**:2 (2000), 407–411. MR 2000c:13027 Zbl 0929.13004

[Nagata 1962] M. Nagata, *Local rings*, Interscience Tracts in Pure and Applied Mathematics **13**, Wiley, New York-London, 1962. MR 27 #5790 Zbl 0123.03402

[Niu 2013] W. Niu, "Some results on asymptotic regularity of ideal sheaves", *J. Algebra* **377** (2013), 157–172. MR 3008900

[Trung and Wang 2005] N. V. Trung and H.-J. Wang, "On the asymptotic linearity of Castelnuovo–Mumford regularity", *J. Pure Appl. Algebra* **201** (2005), 42–48. MR 2006k:13039 Zbl 1100.13024

Communicated by Craig Huneke
Received 2010-06-22 Revised 2012-01-10 Accepted 2012-02-07

chardin@math.jussieu.fr *Institut de Mathématiques de Jussieu, CNRS & UPMC, 4, place Jussieu, F-75005 Paris, France*

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.19

Graphs of Hecke operators

Oliver Lorscheid

Let X be a curve over $\mathbb{F}_q$ with function field F. In this paper, we define a graph for each Hecke operator with fixed ramification. A priori, these graphs can be seen as a convenient language to organize formulas for the action of Hecke operators on automorphic forms. However, they will prove to be a powerful tool for explicit calculations and proofs of finite dimensionality results.

We develop a structure theory for certain graphs $\mathcal{G}_x$ of unramified Hecke operators, which is of a similar vein to Serre's theory of quotients of Bruhat–Tits trees. To be precise, $\mathcal{G}_x$ is locally a quotient of a Bruhat–Tits tree and has finitely many components. An interpretation of $\mathcal{G}_x$ in terms of rank 2 bundles on X and methods from reduction theory show that $\mathcal{G}_x$ is the union of finitely many cusps, which are infinite subgraphs of a simple nature, and a nucleus, which is a finite subgraph that depends heavily on the arithmetic of F.

We describe how one recovers unramified automorphic forms as functions on the graphs $\mathcal{G}_x$. In the exemplary cases of the cuspidal and the toroidal condition, we show how a linear condition on functions on $\mathcal{G}_x$ leads to a finite dimensionality result. In particular, we reobtain the finite-dimensionality of the space of unramified cusp forms and the space of unramified toroidal automorphic forms.

In an appendix, we calculate a variety of examples of graphs over rational function fields.

MSC2010: primary 11F41; secondary 05C75, 11G20, 14H60, 20C08.
Keywords: curve over a finite field, vector bundles, automorphic forms, Hecke operator, Bruhat–Tits tree.

Introduction

Hecke operators play a central role in the theory of automorphic forms, and for classical modular forms, they are also computationally well understood. The theory of arithmetic quotients of the Bruhat–Tits tree as studied in [Serre 2003] allowed the study of Hecke operators over p-adic fields by geometric methods. In this paper, we consider how to compute with Hecke operators for automorphic forms on PGL_2 over a global function field. Our theory can be understood as a global counterpart to Serre's viewpoint over p-adic fields.

There are a few applications of Serre's theory to automorphic forms over global fields, which, however, mainly concentrate on rational function fields; see [Gekeler 1995; 1997; Gekeler and Nonnengardt 1995]. The key ingredient of this application is the strong approximation property of SL_2, as we will explain below. We begin with reminding the reader of the definition of a Bruhat–Tits tree. Though this paper is independent from Serre's book [2003], we review some aspects of it since the global theory (as developed in this paper) and the local approach (as in Serre's book) go hand in hand. In later parts of the paper, we make a few remarks pointing out the connections with and the differences to Serre's theory.

Let F be a global function field and x be a fixed place. We denote by F_x the completion of F at x, by $\mathcal{O}_x$ its integers, by $\pi_x \in \mathcal{O}_x$ a uniformizer and by q_x the cardinality of the residue field $\mathcal{O}_x/(\pi_x) \simeq \mathbb{F}_{q_x}$. The Bruhat–Tits tree $\mathcal{T}_x$ of F_x is a graph with vertex set $\mathrm{PGL}_2(F_x)/\mathrm{PGL}_2(\mathcal{O}_x)$. There is an edge between two cosets $[g]$ and $[g']$ if and only if $[g']$ contains $g\left(\begin{smallmatrix}1 & \\ & \pi_x\end{smallmatrix}\right)$ or $g\left(\begin{smallmatrix}\pi_x & b \\ & 1\end{smallmatrix}\right)$ for some $b \in \mathbb{F}_{q_x}$. Note that this condition is symmetric in g and g', so $\mathcal{T}_x$ is a geometric graph. In fact, $\mathcal{T}_x$ is a (q_x+1)-regular tree.

Every subgroup of $\mathrm{PGL}_2(F_x)$ acts on $\mathcal{T}_x$ by multiplication from the left. We shall be interested in the following case. Let $\mathcal{O}_F^x \subset F$ be the Dedekind ring of all elements $a \in F$ with $\|a\|_y \le 1$ for all places $y \ne x$. Put $\Gamma = \mathrm{PGL}_2(\mathcal{O}_F^x)$. Serre [2003] investigates the quotient graph $\Gamma \backslash \mathcal{T}_x$. It is the union of a finite connected graph with a finite number of cusps. A cusp is an infinite graph of the form

and each cusp corresponds to an element of the class group of $\mathcal{O}_F^x$.

An unramified automorphic form over F_x can be interpreted as a function f on the vertices of $\Gamma \backslash \mathcal{T}_x$ such that the space of functions generated by $\{T_x^i(f)\}_{i\ge 0}$ is finite-dimensional, where the Hecke operator T_x is defined by the formula

$$T_x(f)([g]) = \sum_{\substack{\text{edges } e \text{ with origin } [g] \\ \text{and terminus } [g']}} [\mathrm{Stab}_\Gamma([g]) : \mathrm{Stab}_\Gamma(e)] \cdot f([g'])$$

for each coset $[g] \in \mathrm{PGL}_2(F_x)/\,\mathrm{PGL}_2(\mathcal{O}_x)$.

The inclusion of $\mathrm{PGL}_2(F_x)$ as x-component into $\mathrm{PGL}_2(\mathbb{A})$ induces a map

$$\Gamma \setminus \mathrm{PGL}_2(F_x)/\,\mathrm{PGL}_2(\mathcal{O}_x) \to \mathrm{PGL}_2(F) \setminus \mathrm{PGL}_2(\mathbb{A})/\,\mathrm{PGL}_2(\mathcal{O}_\mathbb{A}),$$

where $\mathcal{O}_\mathbb{A}$ is the maximal compact subring of the adeles $\mathbb{A}$ of F. In the case that F is a rational function field (as in [Gekeler 1995; 1997; Gekeler and Nonnengardt 1995], or, more generally, a function field with odd class number, and x is a place of odd degree, this map is a bijection as a consequence of the strong approximation property of SL_2 (more detail will be given in Section 3). The double coset space on the right hand side is the domain of automorphic forms over F, and the bijection is equivariant with respect to the Hecke operator T_x and its global equivalent Φ_x.

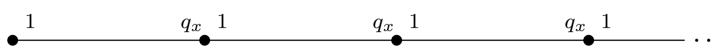
In this sense, it is possible to approximate automorphic forms in this case and use the theory from Serre's book. However, the method of approximation breaks down if the function field has even class number or if the Hecke operator of interest is attached to a place of even degree. For automorphic forms over any function field (with possibly even class number) or for the investigation of Hecke operators at any place of a given function field, respectively, a simultaneous description of all Hecke operators, the method of strong approximation is thus insufficient, and we see the need of a global analogue, which is the starting point of this paper.

The applications of this theory are primarily in explicit computations with automorphic forms. For instance, Lorscheid [2012] uses graphs of Hecke operators to calculate the dimensions of spaces of cusp forms and toroidal automorphic forms. From a more conceptual viewpoint, it might be fruitful to explore the connections between graphs of Hecke operators and Drinfeld modules; in particular, it might contribute to the Langlands program since there is a generalization of graphs of Hecke operator to all reductive groups via adelic Bruhat–Tits buildings, which we forgo explaining here.

We give an overview of the content of this paper. In Section 1, we introduce the graph of a Hecke operator as a graph with weighted edges that encodes the action of a Hecke operator on automorphic forms. This definition applies to every Hecke operator of $\mathrm{PGL}_2(\mathbb{A})$ over a global field. We collect first properties of these graphs and describe how the algebraic structure of the Hecke algebra is reflected in dependencies between the graphs. In Section 2, we describe the graph $\mathcal{G}_x$ of the unramified Hecke operators Φ_x (which correspond to the local Hecke operators T_x as introduced above) in terms of coset representatives. In Section 3, we make the connection to Bruhat–Tits trees precise: Each component of $\mathcal{G}_x$ is a quotient of $\mathcal{T}_x$ by a certain subgroup of $\mathrm{PGL}_2(F_x)$, and the components of $\mathcal{G}_x$ are counted by the 2-torsion of the class group of $\mathcal{O}_F^x$. In Section 4, we associate to each vertex of $\mathcal{G}_x$ a coset in $\mathrm{Cl}\,F/2\,\mathrm{Cl}\,F$ where $\mathrm{Cl}\,F$ is the divisor class group of F. We describe how these labels are distributed in $\mathcal{G}_x$ in dependence of x.

In Section 5, we give the vertices and edges of $\mathcal{G}_x$ a geometric meaning following ideas connected to the geometric Langlands program. Namely, the vertices correspond to the isomorphism classes of $\mathbb{P}^1$-bundles on the smooth projective curve X with function field F, and the edges correspond to certain exact sequences of sheaves on X. In Section 6, we distinguish three classes of rank 2 bundles: those that decompose into a sum of two line bundles, those that are the trace of a line bundle over the quadratic constant extension X' of X and those that are geometrically indecomposable. This divides the vertices of $\mathcal{G}_x$ into three subclasses $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X$, $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}} X$ and $\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}} X$. The former two sets of vertices have a simple description in terms of the divisor class groups of X and X'.

In Section 7, we introduce the integer valued invariant δ on the set of vertices, which is closely connected to reduction theory of rank 2 bundles. This helps us to refine our view on the vertices: $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}} X$ and $\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}} X$ are contained in the finite set of vertices v with $\delta(v) \leq 2g_X - 2$, where g_X is the genus of X. In Section 8, we describe the edges between vertices: $\mathcal{G}_x$ decomposes into a finite graph, which depends heavily on the arithmetic of F, and class-number-many cusps, which are infinite weighted subgraphs of the form



We conclude with a summary of results on $\mathcal{G}_x$ and illustrate them in Figure 8a.

In Section 9, we explain how abstract properties of unramified automorphic forms — namely, the compact support of cusp forms and eigenvalue equations for Eisenstein series — lead to an explicit description of them as functions on the vertices of the graphs $\mathcal{G}_x$. In Section 10, we show that the spaces of functions on $\mathrm{Vert}\,\mathcal{G}_x$ that satisfy the cuspidal or toroidal conditions, respectively, are finite dimensional. In particular, these spaces of functions contain only automorphic forms.

In the appendix, we will give a series of examples for a rational function field: $\mathcal{G}_x$ for $\deg x \leq 5$, the graphs of Φ_x^2 and Φ_x^3 for $\deg x = 1$ and the graphs of two ramified Hecke operators. We give short explanations on how to calculate these examples.

1. Definitions

In this section, we set up our notation and introduce the notion of a graph of a Hecke operator. We collect first properties of these graphs and describe how the algebraic structure of the Hecke algebra is reflected in dependencies between the graphs of different Hecke operators.

1.1. Let q be a prime power and F be the function field of a smooth projective curve X over $\mathbb{F}_q$. Let $\|X\|$ be the set of closed points of X, which we identify with the set of places of F. We denote by F_x the completion of F at $x \in \|X\|$ and by $\mathcal{O}_x$ the integers of F_x. We choose a uniformizer $\pi_x \in F$ for every place x. Let $\kappa_x = \mathcal{O}_x/(\pi_x)$ be the residue field. Let $\deg x$ be the degree of x and let $q_x = q^{\deg x}$ be the cardinality of κ_x. We denote by $\|\cdot\|_x$ the absolute value on F_x and F, respectively, such that $\|\pi_x\|_x = q_x^{-1}$.

Let $\mathbb{A}$ be the adèle ring of F and $\mathbb{A}^\times$ the idèle group. Put $\mathcal{O}_\mathbb{A} = \prod \mathcal{O}_x$, where the product is taken over all places x of F. The idèle norm is the quasicharacter $\|\cdot\| : \mathbb{A}^\times \to \mathbb{C}^\times$ that sends an idèle $(a_x) \in \mathbb{A}^\times$ to the product $\prod \|a_x\|_x$ over all local norms. By the product formula, this defines a quasicharacter on the idèle class group $\mathbb{A}^\times/F^\times$.

We think of F_x being embedded into the adèle ring $\mathbb{A}$ by sending an element a of F_x to the adèle (a_y) with $a_x = a$ and $a_y = 0$ for $y \neq x$. It being not quite compatible with this embedding, we think of the unit group $F_x^\times$ as a subgroup of the idèle group $\mathbb{A}^\times$ by sending an element b of $F_x^\times$ to the idèle (b_y) with $b_x = b$ and $b_y = 1$ for $y \neq x$. We will explain, in case of ambiguity, which of these embeddings we use.

Let $G = \mathrm{PGL}_2$. Following the habit of literature about automorphic forms, we will often write $G_\mathbb{A}$ instead of $G(\mathbb{A})$ for the group of adelic points and G_F instead of $G(F)$ for the group of F-valued points, et cetera. Note that $G_\mathbb{A}$ comes together with an adelic topology that turns $G_\mathbb{A}$ into a locally compact group. Let $K = G_{\mathcal{O}_\mathbb{A}}$ be the standard maximal compact open subgroup of $G_\mathbb{A}$. We fix the Haar measure on $G_\mathbb{A}$ for which $\operatorname{vol} K = 1$.

The *Hecke algebra $\mathcal{H}$ for $G_\mathbb{A}$* is the complex vector space of all compactly supported locally constant functions $\Phi : G_\mathbb{A} \to \mathbb{C}$ together with the convolution product

$$\Phi_1 * \Phi_2 : g \mapsto \int_{G_\mathbb{A}} \Phi_1(gh^{-1})\Phi_2(h)\, dh.$$

A Hecke operator $\Phi \in \mathcal{H}$ acts on the space $\mathcal{V} = C^0(G_F \backslash G_\mathbb{A})$ of continuous functions $f : G_F \backslash G_\mathbb{A} \to \mathbb{C}$ by the formula

$$\Phi(f)(g) = \int_{G_\mathbb{A}} \Phi(h) f(gh)\, dh.$$

Let K' be a compact open subgroup of $G_{\mathbb{A}}$. Then we denote by $\mathcal{H}_{K'}$ the subalgebra of $\mathcal{H}$ that consists of all bi-K'-invariant functions. The action above restricts to an action of $\mathcal{H}_{K'}$ on $\mathcal{V}^{K'}$, the space of right K'-invariant functions.

Lemma 1.2. *For every K' and every $\Phi \in \mathcal{H}_{K'}$, there are $h_1, \dots, h_r \in G_{\mathbb{A}}$ and $m_1, \dots, m_r \in \mathbb{C}$ for some integer r such that for all $g \in G_{\mathbb{A}}$ and all $f \in \mathcal{V}^{K'}$,*

$$\Phi(f)(g) = \sum_{i=1}^{r} m_i \cdot f(gh_i).$$

Proof. Since Φ is K'-biinvariant and compactly supported, it is a finite linear combination of characteristic functions on double cosets of the form $K'hK'$ with $h \in G_{\mathbb{A}}$. So we may reduce the proof to the case $\Phi = \mathrm{char}_{K'hK'}$. Again, since $K'hK'$ is compact, it equals the union of a finite number of pairwise distinct cosets $h_1K', \dots, h_rK'$, and thus, for arbitrary $g \in G_{\mathbb{A}}$,

$$\begin{aligned} \int_{G_{\mathbb{A}}} \mathrm{char}_{K'hK'}(h') f(gh')\, dh' &= \sum_{i=1}^{r} \int_{G_{\mathbb{A}}} \mathrm{char}_{h_iK'}(h') f(gh')\, dh \\ &= \sum_{i=1}^{r} \mathrm{vol}(K') f(gh_i). \end{aligned}$$ □

We will write $[g] \in G_F \setminus G_{\mathbb{A}}/K'$ for the class that is represented by $g \in G_{\mathbb{A}}$. Other cosets will also occur in this paper, but it will be clear from the context what kind of class the square brackets relate to.

Proposition 1.3. *For all $\Phi \in \mathcal{H}_{K'}$ and $[g] \in G_F \setminus G_{\mathbb{A}}/K'$, there is a unique set of pairwise distinct classes $[g_i] \in G_F \setminus G_{\mathbb{A}}/K'$ and numbers $m_i \in \mathbb{C}^\times$, for $1 \le i \le r$, such that for all $f \in \mathcal{V}^{K'}$,*

$$\Phi(f)(g) = \sum_{i=1}^{r} m_i f(g_i).$$

Proof. Uniqueness is clear, and existence follows from Lemma 1.2 after we have taken care of putting together values of f in same classes of $G_F \setminus G_{\mathbb{A}}/K'$ and excluding the zero terms. □

Definition 1.4. With the notation of the preceding proposition we define

$$\mathcal{U}_{\Phi,K'}([g]) = \{([g], [g_i], m_i)\}_{i=1,\dots,r}.$$

The classes $[g_i]$ are called the *Φ-neighbors of* $[g]$ (*relative to* K'), and the m_i are called their *weights*.

The *graph* $\mathcal{G}_{\Phi,K'}$ *of* Φ (*relative to* K') consists of vertices

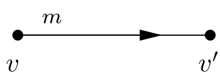
$$\mathrm{Vert}\, \mathcal{G}_{\Phi,K'} = G_F \setminus G_{\mathbb{A}}/K'$$

and oriented weighted edges

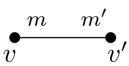
$$\operatorname{Edge} \mathscr{G}_{\Phi,K'} = \bigcup_{v \in \operatorname{Vert} \mathscr{G}_{\Phi,K'}} \mathscr{U}_{\Phi,K'}(v).$$

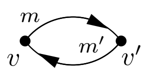
Remark 1.5. The usual notation for an edge in a graph with weighted edges consists of pairs that code the origin and the terminus, and an additional function on the set of edges that gives the weight. For our purposes, it is more convenient to replace the set of edges by the graph of the weight function and to call the resulting triples that consist of origin, terminus and the weight the edges of $\mathscr{G}_{\Phi,K'}$.

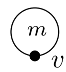
1.6. We make the following drawing conventions to illustrate the graph of a Hecke operator: vertices are represented by labeled dots, and an edge (v, v', m) together with its origin v and its terminus v' is drawn as

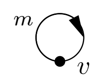


If there is precisely one edge from v to v' and precisely one from v' to v, which we call the inverse edge, we draw



in place of


and


in place of


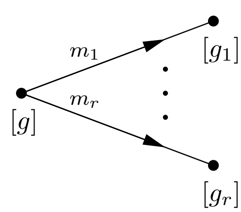
.

There are various examples for rational function fields in the appendix, and in [Lorscheid 2012], one finds graphs of Hecke operators for elliptic function fields.

1.7. We collect some properties that follow immediately from the definition of a graph of a Hecke operator Φ. For $f \in \mathscr{V}^{K'}$ and $[g] \in G_F \setminus G_{\mathbb{A}}/K'$, we have

$$\Phi(f)(g) = \sum_{\substack{([g],[g'],m') \\ \in \operatorname{Edge} \mathscr{G}_{\Phi,K'}}} m' f(g').$$

Hence one can read off the action of a Hecke operator on $f \in \mathscr{V}^{K'}$ from the illustration of the graph



Since $\mathscr{H} = \bigcup \mathscr{H}_{K'}$, with K' running over all compact opens in $G_{\mathbb{A}}$, the notion of the graph of a Hecke operator applies to any $\Phi \in \mathscr{H}$. The set of vertices of the graph of a Hecke operator $\Phi \in \mathcal{H}_{K'}$ only depends on K', and only the edges depend on the particular chosen Φ. There is at most one edge for each pair of vertices and each direction, and the weight of an edge is always nonzero. Each vertex is connected with only finitely many other vertices.

The algebra structure of $\mathcal{H}_{K'}$ has the following implications on the structure of the set of edges (with the convention that the empty sum is defined as 0). For the zero element $0 \in \mathcal{H}_{K'}$, the multiplicative unit $1 \in \mathcal{H}_{K'}$, and arbitrary $\Phi_1, \Phi_2 \in \mathcal{H}_{K'}$ and $r \in \mathbb{C}^\times$, we obtain

$$\begin{aligned}
\operatorname{Edge}\mathcal{G}_{0,K'} &= \varnothing, \\
\operatorname{Edge}\mathcal{G}_{1,K'} &= \{(v, v, 1)\}_{v \in \operatorname{Vert}\mathcal{G}_{1,K'}}, \\
\operatorname{Edge}\mathcal{G}_{\Phi_1+\Phi_2,K'} &= \Big\{(v, v', m) \,\Big|\, m = \sum_{\substack{(v,v',m') \\ \in \operatorname{Edge}\mathcal{G}_{\Phi_1,K'}}} m' + \sum_{\substack{(v,v',m'') \in \\ \operatorname{Edge}\mathcal{G}_{\Phi_2,K'}}} m'' \neq 0\Big\}, \\
\operatorname{Edge}\mathcal{G}_{r\Phi_1,K'} &= \big\{(v, v', rm) \,\big|\, (v, v', m) \in \operatorname{Edge}\mathcal{G}_{\Phi_1,K'}\big\}, \\
\operatorname{Edge}\mathcal{G}_{\Phi_1*\Phi_2,K'} &= \Big\{(v, v', m) \,\Big|\, m = \sum_{\substack{(v,v'',m') \in \operatorname{Edge}\mathcal{G}_{\Phi_1,K'}, \\ (v'',v',m'') \in \operatorname{Edge}\mathcal{G}_{\Phi_2,K'}}} m' \cdot m'' \neq 0\Big\}.
\end{aligned}$$

If $K'' < K'$ and $\Phi \in \mathcal{H}_{K'}$, then also $\Phi \in \mathcal{H}_{K''}$. This implies that we have a canonical map $P : \mathcal{G}_{\Phi,K''} \to \mathcal{G}_{\Phi,K'}$, which is given by

$$\begin{aligned}
\operatorname{Vert}\mathcal{G}_{\Phi,K''} = G_F \backslash G_{\mathbb{A}}/K'' &\xrightarrow{P} G_F \backslash G_{\mathbb{A}}/K' = \operatorname{Vert}\mathcal{G}_{\Phi,K'}, \\
\operatorname{Edge}\mathcal{G}_{\Phi,K''} &\xrightarrow{P} \operatorname{Edge}\mathcal{G}_{\Phi,K'}, \quad (v, v', m') \mapsto (P(v), P(v'), m').
\end{aligned}$$

1.8. One can also collect the data of $\mathcal{G}_{\Phi,K'}$ in an infinite-dimensional matrix $M_{\Phi,K'}$, which we call *the matrix associated with* $\mathcal{G}_{\Phi,K'}$, by putting $(M_{\Phi,K'})_{v',v} = m$ if $(v, v', m) \in \operatorname{Edge}\mathcal{G}_{\Phi,K'}$, and $(M_{\Phi,K'})_{v',v} = 0$ otherwise. Thus each row and each column has only finitely many nonvanishing entries.

The properties of the last paragraph imply the following:

$$\begin{aligned}
&M_{0,K'} = 0, \text{ the zero matrix}, && M_{\Phi_1+\Phi_2,K'} = M_{\Phi_1,K'} + M_{\Phi_2,K'}, \\
&M_{1,K'} = 1, \text{ the identity matrix}, && M_{r\Phi_1,K'} = r M_{\Phi_1,K'}, \\
& && M_{\Phi_1*\Phi_2,K'} = M_{\Phi_2,K'} M_{\Phi_1,K'}.
\end{aligned}$$

Let $\mathcal{J}(K') \subset \mathcal{H}_{K'}$ be the ideal of operators that act trivially on $\mathcal{V}^{K'}$. Then we may regard $\mathcal{H}_{K'}/\mathcal{J}(K')$ as a subalgebra of the algebra of $\mathbb{C}$-linear maps

$$\bigoplus_{G_F \backslash G_{\mathbb{A}}/K'} \mathbb{C} \to \bigoplus_{G_F \backslash G_{\mathbb{A}}/K'} \mathbb{C}.$$

2. Unramified Hecke operators

From now on we will restrict ourselves to unramified Hecke operators, which means elements in $\mathcal{H}_K$. In particular, we will investigate the graphs $\mathcal{G}_x$ of certain generators Φ_x of $\mathcal{H}_K$ in more detail.

2.1. Consider the uniformizers $\pi_x \in F$ as idèles via the embedding $F^\times \subset F_x^\times \subset \mathbb{A}^\times$ and define for every place x the unramified Hecke operator Φ_x as the characteristic function of $K\left(\begin{smallmatrix}\pi_x & \\ & 1\end{smallmatrix}\right)K$. It is well known that $\mathcal{H}_K \simeq \mathbb{C}[\Phi_x]_{x\in\|X\|}$ as an algebra, which means, in particular, that $\mathcal{H}_K$ is commutative. By the relations from Section 1.7, it is enough to know the graphs of generators to determine all graphs of unramified Hecke operators. We use the shorthand notation $\mathcal{G}_x$ for the graph $\mathcal{G}_{\Phi_x,K}$, and $\mathcal{U}_x(v)$ for the Φ_x-neighbors $\mathcal{U}_{\Phi_x,K}(v)$ of v.

We introduce the *lower x convention* that says that a lower index x on an algebraic group defined over the adèles of F will consist of only the component at x of the adelic points, for example, $G_x = G_{F_x}$. Analogously, we put $K_x = G_{\mathcal{O}_x}$.

The *upper x convention* means that an upper index x on an algebraic group defined over the adèles of F will consist of all components except for the one at x. In particular, we first define $\mathbb{A}^x = \prod'_{y\neq x} F_y$, the restricted product relative to $\mathcal{O}^x = \prod_{y\neq x}\mathcal{O}_y$ over all places y that do not equal x. Another example is $G^x = G_{\mathbb{A}^x}$. We put $K^x = G_{\mathcal{O}^x}$.

2.2. We embed κ_x via $\kappa_x \subset F_x \subset \mathbb{A}$; thus an element $b \in \kappa_x$ will be considered as the adèle whose component at x is b and whose other components are 0. Let $\mathbb{P}^1$ be the projective line. Define, for $w \in \mathbb{P}^1(\kappa_x)$,

$$\xi_w = \begin{pmatrix}\pi_x & b\\ & 1\end{pmatrix} \quad \text{if } w=[1:b] \quad \text{and} \quad \xi_w = \begin{pmatrix}1 & \\ & \pi_x\end{pmatrix} \quad \text{if } w=[0:1].$$

It is well known (see [Gelbart 1975, Lemma 3.7]) that the domain of Φ_x can be described as

$$K\begin{pmatrix}\pi_x & \\ & 1\end{pmatrix}K = \coprod_{w\in\mathbb{P}^1(\kappa_x)} \xi_w K.$$

Consequently the weights of edges in $\mathcal{G}_x$ are positive integers (recall that $\operatorname{vol} K = 1$). We shall also refer to the weights as the *multiplicity* of a Φ_x-neighbor. The above implies the following.

Proposition 2.3. *The Φ_x-neighbors of $[g]$ are the classes $[g\xi_w]$ with ξ_w as in the previous paragraph, and the multiplicity of an edge from $[g]$ to $[g']$ equals the number of $w \in \mathbb{P}^1(\kappa_x)$ such that $[g\xi_w]=[g']$. The multiplicities of the edges originating in $[g]$ sum up to $\#\,\mathbb{P}^1(\kappa_x) = q_x + 1$.*

3. Connection with Bruhat–Tits trees

Fix a place x. In this section we construct maps from Bruhat–Tits trees to $\mathcal{G}_x$. This will enable us to determine the components of $\mathcal{G}_x$.

Definition 3.1. The *Bruhat–Tits tree* $\mathcal{T}_x$ for F_x is the (unweighted) graph with vertices $\operatorname{Vert}\mathcal{T}_x = G_x/K_x$ and edges

$$\operatorname{Edge}\mathcal{T}_x = \{([g],[g']) \mid \exists w \in \mathbb{P}^1(\kappa_x),\ g \equiv g'\xi_w \pmod{K_x}\}.$$

3.2. Consider G_x to be embedded in $G_{\mathbb{A}}$ as the component at x. For each $h \in G_{\mathbb{A}}$, we define a map $\Psi_{x,h} : \mathcal{T}_x \to \mathcal{G}_x$ by

$$\begin{aligned} \operatorname{Vert}\mathcal{T}_x = G_x/K_x \to G_F \backslash G_{\mathbb{A}}/K = \operatorname{Vert}\mathcal{G}_x, \qquad & \operatorname{Edge}\mathcal{T}_x \to \operatorname{Edge}\mathcal{G}_x, \\ [g] \mapsto [hg], \qquad & ([g],[g']) \mapsto ([hg],[hg'],m), \end{aligned}$$

with m being the number of vertices $[g'']$ that are adjacent to $[g]$ in $\mathcal{T}_x$ such that $\Psi_{x,h}([g'']) = \Psi_{x,h}([g'])$.

By Proposition 2.3 and the definition of a Bruhat–Tits tree, $\Psi_{x,h}$ is well-defined and *locally surjective*, that is, it is locally surjective as a map between the associated simplicial complexes of $\mathcal{T}_x$ and $\mathcal{G}_x$ with suppressed weights.

Since Bruhat–Tits trees are indeed trees [Serre 2003, II.1, Theorem 1], hence in particular connected, the image of each $\Psi_{x,h}$ is precisely one component of $\mathcal{G}_x$, that is, a subgraph that corresponds to a connected component of the associated simplicial complex.

Every edge of the Bruhat–Tits tree has an inverse edge, which implies the analogous statement for the graphs $\mathcal{G}_x$. Namely, if $(v, v', m) \in \operatorname{Edge}\mathcal{G}_x$, then there is an $m' \in \mathbb{C}^\times$ such that $(v', v, m') \in \operatorname{Edge}\mathcal{G}_x$.

Remark 3.3. This symmetry of edges is a property that is particular to unramified Hecke operators for $G = \mathrm{PGL}_2$. In case of ramification, the symmetry is broken; see Example A.7.

3.4. The algebraic group SL_2 has the *strong approximation property*, that is, for every place x, $\mathrm{SL}_2 F$ is a dense subset of $\mathrm{SL}_2 \mathbb{A}^x$ with respect to the adelic topology. See [Bourbaki 1965, §2, nombre 4; Kneser 1966; Moore 1968, Chapter IV, Lemma 13.1; Margulis 1977; Prasad 1977] for the development of the strong approximation results and their generalizations to all simple groups. See also [Laumon 1997, Theorem E.2.1] for a proof. We explain what implication this has on PGL_2. More detail for the outline in this paragraph can be found in [van der Put and Reversat 1997, (2.1.3)].

Let x be a place of degree d. In accordance to the upper x convention, let $\mathcal{O}^x = \prod_{y \neq x} \mathcal{O}_y$. As a consequence of the strong approximation property of SL_n, the determinant map on GL_2 induces a bijection on double cosets:

$$\mathrm{GL}_2(F) \setminus \mathrm{GL}_2(\mathbb{A}^x)/ \mathrm{GL}_2(\mathcal{O}^x) \xrightarrow{\det} F^\times \setminus (\mathbb{A}^x)^\times/(\mathcal{O}^x)^\times.$$

The quotient group $F^\times \setminus (\mathbb{A}^x)^\times/(\mathcal{O}^x)^\times$ is nothing else but the ideal class group $\mathrm{Cl}\,\mathcal{O}_F^x$ of the *integers* $\mathcal{O}_F^x = \mathcal{O}^x \cap F$ *coprime to* x. Let $\mathrm{Cl}\,F = F^\times \setminus \mathbb{A}^\times/\mathcal{O}_\mathbb{A}^\times$ be the divisor class group of F and $\mathrm{Cl}^0\,F = \{[a] \in \mathrm{Cl}\,F \mid \deg a = 0\}$ be the ideal class group. Then we have bijections

$$\mathrm{GL}_2(F) \setminus \mathrm{GL}_2(\mathbb{A}^x)/ \mathrm{GL}_2(\mathcal{O}^x) \simeq F^\times \setminus (\mathbb{A}^x)^\times/(\mathcal{O}^x)^\times \simeq \mathrm{Cl}\,\mathcal{O}_F^x \simeq \mathrm{Cl}^0 F \times \mathbb{Z}/d\mathbb{Z}.$$

Let $S \subset \mathrm{GL}_2(\mathbb{A}^x)$ be a set of representatives for $\mathrm{GL}_2(F)\setminus\mathrm{GL}_2(\mathbb{A}^x)/\mathrm{GL}_2(\mathcal{O}^x)$. Then for every $g = g^x g_x \in \mathrm{GL}_2(\mathbb{A})$ (with $g^x \in \mathrm{GL}_2(\mathbb{A}^x)$ and $g_x \in \mathrm{GL}_2(F_x)$), there are $s \in S$, $\gamma \in \mathrm{GL}_2(F)$ and $k \in \mathrm{GL}_2(\mathcal{O}^x)$ such that $g = \gamma s k \tilde{g}_x$, where $\gamma s k$ equals g in all components $z \neq x$ and $\tilde{g}_x = \gamma^{-1} g_x$. The condition $[\det s] = [\det g^x]$ as cosets in $F^\times \setminus (\mathbb{A}^x)^\times/(\mathcal{O}^x)^\times$ implies that $s \in S$ is uniquely determined by g^x. Let Z be the center of GL_2. Then

$$\begin{aligned}\mathrm{GL}_2(\mathbb{A})/ \mathrm{GL}_2(\mathcal{O}_\mathbb{A}) Z_x &= \mathrm{GL}_2(\mathbb{A}^x)/ \mathrm{GL}_2(\mathcal{O}^x) \times G_x/K_x\\ &= \mathrm{GL}_2(\mathbb{A}^x)/ \mathrm{GL}_2(\mathcal{O}^x) \times \mathrm{Vert}\,\mathcal{T}_x.\end{aligned}$$

Define $\Gamma_s = \mathrm{GL}_2(F) \cap s\,\mathrm{GL}_2(\mathcal{O}^x) s^{-1}$. Then we obtain the following; see [van der Put and Reversat 1997, (2.1.3)].

Proposition 3.5. *The decomposition* $g = \gamma s k \tilde{g}_x$ *induces a bijective map*

$$\mathrm{GL}_2(F) \setminus \mathrm{GL}_2(\mathbb{A}^x)/ \mathrm{GL}_2(\mathcal{O}_\mathbb{A}) Z_x \to \coprod_{s \in S} \Gamma_s \setminus \mathrm{Vert}\,\mathcal{T}_x, \quad [g] \mapsto (s, [\tilde{g}_x]).$$

Its inverse is obtained by joining the components $s \in \mathrm{GL}_2(\mathbb{A}^x)$ *and* $\tilde{g}_x \in G_x$.

Remark 3.6. On the right side of the bijection in Proposition 3.5, we have a finite union of quotients of the form $\Gamma_s \setminus \mathrm{Vert}\,\mathcal{T}_x$. If s is the identity element e, then $\Gamma = \Gamma_e = \mathrm{GL}_2(\mathcal{O}_F^x)$ is an arithmetic group of the form considered in [Serre 2003, II.2.3]. For general s, I am not aware of any results about $\Gamma_s \setminus \mathrm{Vert}\,\mathcal{T}_x$.

3.7. So far, we have only divided out the action of the x-component Z_x of the center. We still have to consider the action of Z^x. The image of Z^x under the determinant $\det : \mathrm{GL}_2(\mathbb{A}^x) \to \mathrm{Cl}\,\mathcal{O}_F^x$ is $2\,\mathrm{Cl}\,\mathcal{O}_F^x$. Thus we obtain a bijection

$$Z^x\,\mathrm{GL}_2(F) \setminus \mathrm{GL}_2(\mathbb{A}^x)/ \mathrm{GL}_2(\mathcal{O}^x) \xrightarrow{\det} \mathrm{Cl}\,\mathcal{O}_F^x/2\,\mathrm{Cl}\,\mathcal{O}_F^x.$$

The double quotient on the left side can be identified with $G_F \setminus G^x/K^x$. Let $J = \{z \in Z^x \mid \det z = 0 \in \mathrm{Cl}\,\mathcal{O}_F^x\}$ be the kernel of the restriction $\det : Z^x \to \mathrm{Cl}\,\mathcal{O}_F^x$ and define $\tilde{\Gamma}_s = \mathrm{GL}_2(F) \cap J s\,\mathrm{GL}_2(\mathcal{O}^x) s^{-1}$. If we let $S' \subset S$ be a set of representatives for $\mathrm{Cl}\,\mathbb{O}_F^x/2\,\mathrm{Cl}\,\mathbb{O}_F^x$ (with respect to the determinant map), and $h_2 = \#(\mathrm{Cl}\,F)[2]$ the cardinality of the 2-torsion, then we obtain:

Proposition 3.8. *The decomposition $g = \gamma s k \tilde{g}_x$ induces a bijective map*

$$G_F \setminus G_{\mathbb{A}}/K \to \coprod_{s\in S'} \tilde{\Gamma}_s \setminus \mathrm{Vert}\,\mathcal{T}_x.$$

The inverse maps an element $(s, [\tilde{g}_x])$ to the class of the adelic matrix with components $s \in G^x$ and $\tilde{g}_x \in G_x$. The number of components of $\mathcal{G}_x$ equals

$$\#(\mathrm{Cl}\,\mathbb{O}_F^x/2\,\mathrm{Cl}\,\mathbb{O}_F^x) = \#(\mathrm{Cl}\,O_F^x)[2] = \begin{cases} h_2 & \textit{if } \deg x \textit{ is odd,} \\ 2h_2 & \textit{if } \deg x \textit{ is even.} \end{cases}$$

Proof. Everything follows from Proposition 3.5 and Section 3.7 except for the two equalities in the last line. The former equality follows from the general fact that one has $\#\ker f = \#(G/\operatorname{im} f)$ for a homomorphism f acting on a finite group G (in our case f is the multiplication by 2). The latter equality follows immediately from the observation $\mathrm{Cl}\,\mathbb{O}_F^x \simeq \mathrm{Cl}^0\,F \times \mathbb{Z}/d\mathbb{Z}$, where $d = \deg x$. □

4. A vertex labeling

In this section, we associate to each vertex of $\mathcal{G}_x$ an element of $\mathrm{Cl}\,F/2\,\mathrm{Cl}\,F$ and determine how these labels are distributed over the components of $\mathcal{G}_x$.

4.1. Let $\mathfrak{Q}_{\mathbb{A}} = \langle a^2 \mid a \in \mathbb{A}^\times \rangle$ be the subgroup of squares. We look once more at the determinant map

$$\mathrm{Vert}\,\mathcal{G}_x = G_F \setminus G_{\mathbb{A}}/K \xrightarrow{\det} F^\times \setminus \mathbb{A}^\times/\mathbb{O}_{\mathbb{A}}^\times \mathfrak{Q}_{\mathbb{A}} \simeq \mathrm{Cl}\,F/2\,\mathrm{Cl}\,F.$$

This map assigns to every vertex in $\mathcal{G}_x$ a label in $\mathrm{Cl}\,F/2\,\mathrm{Cl}\,F$, which has $2h_2$ elements, where $h_2 = \#(\mathrm{Cl}\,F)[2]$, for the same reason as used in the proof of Proposition 3.8.

Proposition 4.2. *If the prime divisor x is a square in the divisor class group, then all vertices in the same component of $\mathcal{G}_x$ have the same label, and there are $2h_2$ components, each of which has a different label. Otherwise, the vertices of each component have one of two labels that differ by x in $\mathrm{Cl}\,F/2\,\mathrm{Cl}\,F$, and two adjacent vertices have different labels, so each connected component is bipartite.*

Proof. First of all, observe that each label is realized, since if we represent a label by some idèle a, then the vertex represented by $\begin{pmatrix} a & \\ & 1 \end{pmatrix}$ has this label.

Let $\mathfrak{Q}_x = \langle b^2 \mid b \in F_x^\times \rangle$ and $\mathrm{Cl}\,F_x = F_x^\times/\mathbb{O}_x^\times$, a group isomorphic to $\mathbb{Z}$. For the Bruhat–Tits tree $\mathcal{T}_x$, the determinant map

$$\mathrm{Vert}\,\mathcal{T}_x = G_x/K_x \xrightarrow{\det} F_x^\times/\mathbb{O}_x^\times \mathfrak{Q}_x \simeq \mathrm{Cl}\,F_x/2\,\mathrm{Cl}\,F_x \simeq \mathbb{Z}/2\mathbb{Z}$$

defines a labeling of the vertices, and the two classes of $F_x^\times/\mathcal{O}_x^\times\mathcal{Q}_x$ are represented by 1 and π_x. Two adjacent vertices have the different labels since for $g \in G_x$ and ξ_w as in Definition 3.1, $\det(g\xi_w) = \pi_x \det g$ represents a class different from $\det g$ in $\operatorname{Vert}\mathcal{T}_x$.

Define for $a \in \mathbb{A}^\times$ a map $\psi_{x,a} : F_x^\times/\mathcal{O}_x^\times\mathcal{Q}_x \to F^\times\backslash\mathbb{A}^\times/\mathcal{O}_\mathbb{A}^\times\mathcal{Q}_\mathbb{A}$ by $\psi_{x,a}([b]) = [ab]$, where b is viewed as the idèle concentrated in x. For every $h \in G_\mathbb{A}$ we obtain a commutative diagram

$$\begin{array}{ccccccc} \operatorname{Vert}\mathcal{T}_x & = & G_x/K_x & \xrightarrow{\Psi_{x,h}} & G_F \backslash G_\mathbb{A}/K & = & \operatorname{Vert}\mathcal{G}_x \\ \downarrow & & \downarrow{\scriptstyle\det} & & \downarrow{\scriptstyle\det} & & \downarrow \\ \operatorname{Cl} F_x/2\operatorname{Cl} F_x & \simeq & F_x^\times/\mathcal{O}_x^\times\mathcal{Q}_x & \xrightarrow{\psi_{x,\det h}} & F^\times \backslash \mathbb{A}^\times/\mathcal{O}_\mathbb{A}^\times\mathcal{Q}_\mathbb{A} & \simeq & \operatorname{Cl} F/2\operatorname{Cl} F. \end{array}$$

This means that vertices with equal labels map to vertices with equal labels.

Each component of $\mathcal{G}_x$ lies in the image of a suitable $\Psi_{x,h}$, and thus has at most two labels. On the other hand, the two labels of $\mathcal{T}_x$ map to $\psi_{x,\det h}([1]) = [a]$ and $\psi_{x,\det h}([\pi_x]) = [a\pi_x]$, where $a = \det h$. The divisor classes of $[a]$ and $[a\pi_x]$ differ by the class of the prime divisor x, and are equal if and only if x is a square in the divisor class group. If so, according to Proposition 3.8, there must be $2h_2$ components, so that the $2h_2$ labels are spread over all components. If x is not a square, then by the local surjectivity of $\Psi_{x,h}$ on edges two adjacent vertices of $\mathcal{G}_x$ also have different labels. □

5. Geometric interpretation of unramified Hecke operators

A fundamental observation in the geometric Langlands program (for PGL_2, in this case) is that the domain of automorphic forms (with a certain ramification level) corresponds to the isomorphism classes of $\mathbb{P}^1$-bundles (with a corresponding level structure). The action of Hecke operators can be given a geometric meaning, which makes it possible to let algebraic geometry enter the field. We will use this geometric view point for a closer examination of the graphs of unramified Hecke operators. We begin with recalling the geometric interpretation of unramified Hecke operators. For more reference, see [Gaitsgory 2003].

5.1. Let $\mathcal{O}_X$ be the structure sheaf of the smooth projective curve X and η the generic point. We can identify the stalks $\mathcal{O}_{X,x}$ of the structure sheaf $\mathcal{O}_X$ at closed points $x \in \|X\|$ and their embeddings into the generic stalk $\mathcal{O}_{X,\eta}$ with

$$\mathcal{O}_{X,x} \simeq \mathcal{O}_x \cap F \hookrightarrow F \simeq \mathcal{O}_{X,\eta}.$$

We identify vector bundles on X with the corresponding locally free sheaf [Hartshorne 1977, Exercise II.5.18]. We denote by $\mathrm{Bun}_n X$ the set of isomorphism classes of *rank n bundles* over X and by Pic X the *Picard group*. For $\mathcal{L}_1, \mathcal{L}_2 \in \operatorname{Pic} X$, we use the shorthand notation $\mathcal{L}_1\mathcal{L}_2$ for $\mathcal{L}_1 \otimes \mathcal{L}_2$. The group Pic X acts on $\operatorname{Bun}_n X$ by tensor products. Let $\mathbb{P}\operatorname{Bun}_n X$ be the orbit set $\operatorname{Bun}_n X/\operatorname{Pic} X$, which is nothing but the set of isomorphism classes of $\mathbb{P}^{n-1}$-bundles over X [ibid., Ex. II.7.10].

We will call the elements of $\mathbb{P}\operatorname{Bun}_2 X$ *projective line bundles*. If we regard the total space of a projective line bundle as a scheme, then we obtain a ruled surface; see [ibid., Proposition V.2.2]. Thus $\mathbb{P}\operatorname{Bun}_2 X$ may also be seen as the set of isomorphism classes of ruled surfaces over X.

If two vector bundles $\mathcal{M}_1$ and $\mathcal{M}_2$ are in the same orbit of the action of Pic X, we write $\mathcal{M}_1 \sim \mathcal{M}_2$ and say that $\mathcal{M}_1$ and $\mathcal{M}_2$ are *projectively equivalent*. When we say $[\mathcal{M}] \in \mathbb{P}\operatorname{Bun}_2 X$, we mean the class that is represented by the rank 2 bundle $\mathcal{M}$.

Let $\operatorname{Cl} X = \operatorname{Cl} F$ be the divisor group of X. Every divisor $D \in \operatorname{Cl} X$ defines the *associated line bundle* $\mathcal{L}_D$, which defines an isomorphism $\operatorname{Cl} X \to \operatorname{Pic} X$ of groups [ibid., Proposition II.6.15]. The degree $\deg \mathcal{M}$ of a vector bundle $\mathcal{M}$ with $\det \mathcal{M} \simeq \mathcal{L}_D$ is defined as $\deg D$; see [ibid., Ex. II.6.12]. For a torsion sheaf $\mathcal{F}$, the degree is defined by $\deg \mathcal{F} = \sum_{x\in\|X\|} \dim_{\mathbb{F}_q}(\mathcal{F}_x)$. The degree is additive in short exact sequences.

Remark 5.2. Note that if $D = x$ is a prime divisor, the notation for the associated line bundle $\mathcal{L}_x$ coincides with the notation for the stalk of $\mathcal{L}$ at x. In order to avoid confusion, we will reserve the notation $\mathcal{L}_x$ strictly for the associated line bundle. In case we have to consider the stalk of a line bundle, we will use a symbol different from $\mathcal{L}$ for the line bundle.

5.3. The correspondence between $\operatorname{Cl} X = F^\times \backslash \mathbb{A}^\times/\mathcal{O}_{\mathbb{A}}^\times$ and Pic X extends to higher rank. For more details on the following outline; see [Frenkel 2004, Lemma 3.1; Gaitsgory 2003, 2.1]. Let $\mathcal{M}$ be a rank 2 bundle. Then we can choose for every $x \in \|X\|$ a trivialization φ_x of $\mathcal{M}_x$ in a formal neighborhood of x, and a trivialization φ_η of the generic stalk $\mathcal{M}_\eta$. We define the matrix g_x as the base change matrix corresponding to

$$\mathcal{O}_{X,x}^2 \xrightarrow{\varphi_x} \mathcal{M}_x \hookrightarrow \mathcal{M}_\eta \xrightarrow{\varphi_\eta^{-1}} F^2$$

with respect to the standard bases of $\mathcal{O}_{X,x}^2$ and F^2. This yields an element $g = (g_x)$ of $\operatorname{GL}_2(\mathbb{A})$. A coordinate change of the stalks $\mathcal{M}_x$ corresponds to a matrix in $\operatorname{GL}_2(\mathcal{O}_{\mathbb{A}})$ and a coordinate change of $\mathcal{M}_\eta$ corresponds to a matrix in $\operatorname{GL}_2(F)$. Indeed, every double coset in $\operatorname{GL}_2(F) \backslash \operatorname{GL}_2(\mathbb{A})/\operatorname{GL}_2(\mathcal{O}_{\mathbb{A}})$ is obtained from a vector bundle in the described way, which yields a bijection

$$\begin{aligned} \operatorname{GL}_2(F) \backslash \operatorname{GL}_2(\mathbb{A})/\operatorname{GL}_2(\mathcal{O}_{\mathbb{A}}) &\xleftrightarrow{1:1} \operatorname{Bun}_2 X, \\ [g] &\longmapsto \mathcal{M}_g \end{aligned}$$

Furthermore, we have $\mathcal{M}_g \otimes \mathcal{L}_a = \mathcal{M}_{ag}$ for $a \in \mathbb{A}^\times$, and $\deg \mathcal{M}_g = \deg(\det g)$. Consequently, there is a bijection

$$G_F \setminus G_{\mathbb{A}}/K \overset{1:1}{\longleftrightarrow} \mathbb{P}\mathrm{Bun}_2\, X,$$

which allows us to identify the vertex set $\mathrm{Vert}\,\mathcal{G}_x = G_F \setminus G_{\mathbb{A}}/K$ with $\mathbb{P}\mathrm{Bun}_2\, X$.

5.4. The next task is to describe edges of $\mathcal{G}_x$ in geometric terms. We say that two exact sequences

$$0 \to \mathcal{F}_1 \to \mathcal{F} \to \mathcal{F}_1' \to 0 \quad \text{and} \quad 0 \to \mathcal{F}_2 \to \mathcal{F} \to \mathcal{F}_2' \to 0$$

of sheaves are *isomorphic with fixed* $\mathcal{F}$ if there are isomorphisms $\mathcal{F}_1 \to \mathcal{F}_2$ and $\mathcal{F}_1' \to \mathcal{F}_2'$ such that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{F}_1 & \longrightarrow & \mathcal{F} & \longrightarrow & \mathcal{F}_1' & \longrightarrow & 0 \\ & & \downarrow \simeq & & \| & & \downarrow \simeq & & \\ 0 & \longrightarrow & \mathcal{F}_2 & \longrightarrow & \mathcal{F} & \longrightarrow & \mathcal{F}_2' & \longrightarrow & 0 \end{array}$$

commutes.

Let $\mathcal{K}_x$ be the torsion sheaf that is supported at x and has stalk κ_x at x, where κ_x is the residue field at x. Fix a representative $\mathcal{M}$ of $[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2\, X$. Then we define $m_x([\mathcal{M}], [\mathcal{M}'])$ as the number of isomorphism classes of exact sequences

$$0 \to \mathcal{M}'' \to \mathcal{M} \to \mathcal{K}_x \to 0,$$

with fixed $\mathcal{M}$ and with $\mathcal{M}'' \sim \mathcal{M}'$. This number is independent of the choice of the representative $\mathcal{M}$ because for another choice, which would be a vector bundle of the form $\mathcal{M} \otimes \mathcal{L}$ for some $\mathcal{L} \in \mathrm{Pic}\, X$, we have the bijection

$$\begin{array}{ccc} \left\{ \begin{array}{c} \text{isomorphism classes} \\ 0 \to \mathcal{M}'' \to \mathcal{M} \to \mathcal{K}_x \to 0 \\ \text{with fixed } \mathcal{M} \end{array} \right\} & \to & \left\{ \begin{array}{c} \text{isomorphism classes} \\ 0 \to \mathcal{M}''' \to \mathcal{M} \otimes \mathcal{L} \to \mathcal{K}_x \to 0 \\ \text{with fixed } \mathcal{M} \otimes \mathcal{L} \end{array} \right\}, \\ (0 \to \mathcal{M}'' \to \mathcal{M} \to \mathcal{K}_x \to 0) & \mapsto & (0 \to \mathcal{M}'' \otimes \mathcal{L} \to \mathcal{M} \otimes \mathcal{L} \to \mathcal{K}_x \to 0). \end{array}$$

Definition 5.5. Let x be a place. For a projective line bundle $[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2\, X$ we define

$$\mathcal{U}_x([\mathcal{M}]) = \{([\mathcal{M}], [\mathcal{M}'], m) \mid m = m_x([\mathcal{M}], [\mathcal{M}']) \neq 0\},$$

and call the occurring $[\mathcal{M}']$ the Φ_x*-neighbors of* $[\mathcal{M}]$, and $m_x([\mathcal{M}], [\mathcal{M}'])$ their *multiplicity*.

5.6. We shall show that this concept of neighbors is the same as the one defined for classes in $G_F \setminus G_{\mathbb{A}}/K$ (Definition 1.4). Recall that in Proposition 2.3, we determined the Φ_x-neighbors of a class $[g] \in G_F \setminus G_{\mathbb{A}}/K$ to be of the form $[g\xi_w]$ for a $w \in \mathbb{P}^1(\kappa_x)$. The elements ξ_w define exact sequences

$$0 \to \prod_{y \in \|X\|} \mathcal{O}^2_{X,y} \xrightarrow{\xi_w} \prod_{y \in \|X\|} \mathcal{O}^2_{X,y} \to \kappa_x \to 0$$

of $\mathbb{F}_q$-modules and consequently an exact sequence $0 \to \mathcal{M}_{g\xi_w} \to \mathcal{M}_g \to \mathcal{K}_x \to 0$ of sheaves, where $\mathcal{M}_{g\xi_w}$ and $\mathcal{M}_g$ are the rank 2 bundles associated with $g\xi_w$ and g, respectively. This maps $w \in \mathbb{P}^1(\kappa_x)$ to the isomorphism class of $(0 \to \mathcal{M}_{g\xi_w} \to \mathcal{M}_g \to \mathcal{K}_x \to 0)$ with fixed $\mathcal{M}_g$. On the other hand, as we have chosen a basis for the stalk at x, each isomorphism class of sequences $(0 \to \mathcal{M}' \to \mathcal{M} \to \mathcal{K}_x \to 0)$ with fixed $\mathcal{M}$ defines an element in $\mathbb{P}(\mathcal{O}^2_{X,x}/(\pi_x\mathcal{O}_{X,x})^2) = \mathbb{P}^1(\kappa_x)$, which gives back w.

Thus for every $x \in \|X\|$, the map

$$\mathcal{U}_x([g]) \to \mathcal{U}_x([\mathcal{M}_g]), \quad ([g],[g'],m) \mapsto ([\mathcal{M}_g],[\mathcal{M}_{g'}],m)$$

is a well-defined bijection. We finally obtain the geometric description of the graph $\mathcal{G}_x$ of Φ_x.

Proposition 5.7. *Let $x \in \|X\|$. The graph $\mathcal{G}_x$ of Φ_x is described in geometric terms as*

$$\operatorname{Vert}\mathcal{G}_x = \mathbb{P}\mathrm{Bun}_2\, X \quad \textit{and} \quad \operatorname{Edge}\mathcal{G}_x = \coprod_{[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2 X} \mathcal{U}_x([\mathcal{M}]).$$

Remark 5.8. This interpretation shows that the graphs that we consider are a global version of the graphs of Serre [2003, Chapter II.2]. We are looking at all rank 2 bundles on X modulo the action of the Picard group of X while Serre considers rank 2 bundles that trivialize outside a given place x modulo line bundles that trivialize outside x. As already explained in Remark 3.6, we obtain a projection of the graph of Serre to the component of the trivial class c_0.

Serre describes his graphs as quotients of Bruhat–Tits trees by the action of the group $\Gamma = G_{\mathcal{O}_F^x}$ on both vertices and edges. This leads in general to multiple edges between vertices in the quotient graph; see for example [Serre 2003, 2.4.2c]. This does not happen with graphs of Hecke operators: There is at most one edge with given origin and terminus.

Relative to the action of Γ on Serre's graphs, one can define the weight of an edge as the order of the stabilizer of its origin in the stabilizer of the edge. The projection from Serre's graphs to graphs of Hecke operators identifies all the different edges between two vertices, adding up their weights to obtain the weight of the image edge.

6. Description of vertices

The aim of this section is to show that the set of isomorphism classes of projective line bundles over X can be separated into subspaces corresponding to certain quotients of the divisor class group of F, the divisor class group of $\mathbb{F}_{q^2}F$ and geometrically indecomposable projective line bundles. We recall a series of facts about vector bundles.

6.1. A vector bundle $\mathcal{M}$ is *indecomposable* if for every decomposition $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$ into two subbundles $\mathcal{M}_1$ and $\mathcal{M}_2$, one factor is trivial and the other is isomorphic to $\mathcal{M}$. The Krull–Schmidt theorem holds for the category of vector bundles over X, that is, every vector bundle $\mathcal{M}$ on X defined over $\mathbb{F}_q$ has, up to permutation of factors, a unique decomposition into a direct sum of indecomposable subbundles; see [Atiyah 1956, Theorem 2].

The map $p : X' = X \otimes \mathbb{F}_{q^i} \to X$ defines the *inverse image* or the *constant extension* of vector bundles

$$p^* : \mathrm{Bun}_n\, X \to \mathrm{Bun}_n\, X', \quad \mathcal{M} \mapsto p^*\mathcal{M}.$$

The isomorphism classes of rank n bundles that after extension of constants to $\mathbb{F}_{q^i}$ become isomorphic to $p^*\mathcal{M}$ are classified by $H^1(\mathrm{Gal}(\mathbb{F}_{q^i}/\mathbb{F}_q), \mathrm{Aut}(\mathcal{M} \otimes \mathbb{F}_{q^i}))$; see [Arason et al. 1992, Section 1]. The algebraic group $\mathrm{Aut}(\mathcal{M} \otimes \mathbb{F}_{q^i})$ is an open subvariety of the connected algebraic group $\mathrm{End}(\mathcal{M} \otimes \mathbb{F}_{q^i})$, and thus it is itself a connected algebraic group. As a consequence of Lang's theorem [1956, Corollary to Theorem 1], we have $H^1(\mathrm{Gal}(\mathbb{F}_{q^i}/\mathbb{F}_q), \mathrm{Aut}(\mathcal{M} \otimes \mathbb{F}_{q^i})) = 1$.

Thus p^* is injective. In particular, one can consider the constant extension to the geometric curve $\overline{X} = X \otimes \overline{\mathbb{F}}_q$ over an algebraic closure $\overline{\mathbb{F}}_q$ of $\mathbb{F}_q$. Then two vector bundles are isomorphic if and only if they are geometrically isomorphic, that is, if their constant extensions to $\overline{X}$ are isomorphic. We can therefore think of $\mathrm{Bun}_n\, X$ as a subset of $\mathrm{Bun}_n\, X'$ and $\mathrm{Bun}_n\, \overline{X}$.

On the other hand, $p : X' \to X$ defines the direct image or the *trace* of vector bundles

$$p_* : \mathrm{Bun}_n\, X' \to \mathrm{Bun}_{ni}\, X, \quad \mathcal{M} \mapsto p_*\mathcal{M}.$$

We have $p_*p^*\mathcal{M} \simeq \mathcal{M}^i$ for $\mathcal{M} \in \mathrm{Bun}_n\, X$ and $p^*p_*\mathcal{M} \simeq \bigoplus \mathcal{M}^\tau$ for $\mathcal{M} \in \mathrm{Bun}_n\, X'$, where τ ranges over $\mathrm{Gal}(\mathbb{F}_{q^i}/\mathbb{F}_q)$ and $\mathcal{M}^\tau$ is defined by the stalks $\mathcal{M}^\tau_x = \mathcal{M}_{\tau^{-1}(x)}$.

We call a vector bundle *geometrically indecomposable* if its extension to $\overline{X}$ is indecomposable. In [Arason et al. 1992, Theorem 1.8], it is shown that every indecomposable vector bundle over X is the trace of a geometrically indecomposable bundle over some constant extension X' of X.

There are certain compatibilities of the constant extension and the trace with tensor products. Namely, for a vector bundle $\mathcal{M}$ and a line bundle $\mathcal{L}$ over X, we have $p^*(\mathcal{M}\otimes\mathcal{L})\simeq p^*\mathcal{M}\otimes p^*\mathcal{L}$ and for a vector bundle $\mathcal{M}'$ over X',

$$p_*\mathcal{M}'\otimes\mathcal{L}\simeq p_*(\mathcal{M}'\otimes p^*\mathcal{L}).$$

Thus p^* induces a map, denoted by the same symbol,

$$p^*:\mathbb{P}\mathrm{Bun}_n\,X\to\mathbb{P}\mathrm{Bun}_n\,X',\quad [\mathcal{M}]\mapsto[p^*\mathcal{M}],$$

and p_* induces

$$p_*:\mathrm{Bun}_n\,X'/p^*\,\mathrm{Pic}\,X\to\mathbb{P}\mathrm{Bun}_{ni}\,X,\quad [\mathcal{M}]\mapsto[p_*\mathcal{M}].$$

6.2. We look at the situation for rank 2 bundles. Let σ be the nontrivial automorphism of $\mathbb{F}_{q^2}/\mathbb{F}_q$. The set $\mathbb{P}\mathrm{Bun}_2\,X$ is the disjoint union of the set of classes of decomposable rank 2 bundles, that is, rank 2 bundles that are isomorphic to the direct sum of two line bundles, and the set of classes of indecomposable bundles. We denote these sets by $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X$ and $\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}}\,X$, respectively. Let $\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}}\,X\subset\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}}\,X$ be the subset of classes of geometrically indecomposable bundles. Since the rank is 2, the complement $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X=\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}}\,X-\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}}\,X$ consists of classes of traces $p_*\mathcal{L}$ of certain line bundles $\mathcal{L}\in\mathrm{Pic}\,X'$ that are defined over the quadratic extension $X'=X\otimes\mathbb{F}_{q^2}$. More precisely, $p_*\mathcal{L}$ decomposes if and only if $\mathcal{L}\in p^*\,\mathrm{Pic}\,X$, and then $p_*\mathcal{L}\sim\mathcal{O}_X\oplus\mathcal{O}_X$. Thus, we have a disjoint union

$$\mathbb{P}\mathrm{Bun}_2\,X=\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X\sqcup\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X\sqcup\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}}\,X.$$

For $[D]\in\mathrm{Cl}\,X$, define

$$c_D=[\mathcal{L}_D\oplus\mathcal{O}_X]\in\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X,$$

and for a $[D]\in\mathrm{Cl}\,X'$, define

$$t_D=[p_*\mathcal{L}_D]\in\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X\cup\{c_0\}.$$

Note that σ acts on $\mathrm{Cl}\,X'$ in a way compatible with the identification $\mathrm{Cl}\,X'\simeq\mathrm{Pic}\,X'$. Since $p^*p_*(\mathcal{L})\simeq\mathcal{L}\oplus\mathcal{L}^\sigma\simeq p^*p_*(\mathcal{L}^\sigma)$ for $\mathcal{L}\in\mathrm{Pic}\,X'$, and isomorphism classes of vector bundles are stable under constant extensions, we have $t_D=t_{\sigma D}$.

We derive the following characterizations of $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X$ and $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X$:

Proposition 6.3. *The map* $\mathrm{Cl}\,X\to\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X$, $[D]\mapsto c_D$ *is surjective with fibers of the form* $\{[D],[-D]\}$.

Proof. Let $\mathcal{M}$ decompose into $\mathcal{L}_1\oplus\mathcal{L}_2$. Then

$$\mathcal{M}\simeq\mathcal{L}_1\oplus\mathcal{L}_2\sim(\mathcal{L}_1\oplus\mathcal{L}_2)\otimes\mathcal{L}_2^{-1}\simeq\mathcal{L}_1\mathcal{L}_2^{-1}\oplus\mathcal{O}_X,$$

thus surjectivity follows. Let $\mathcal{L}_{D'}\oplus\mathcal{O}_X$ represent the same projective line bundle as $\mathcal{L}_D\oplus\mathcal{O}_X$. Then, there is a line bundle $\mathcal{L}_0$ such that $\mathcal{L}_D\oplus\mathcal{O}_X\simeq(\mathcal{L}_{D'}\oplus\mathcal{O}_X)\otimes\mathcal{L}_0$, and thus either $\mathcal{L}_0 \simeq \mathcal{O}_X$ and $\mathcal{L}_D \simeq \mathcal{L}_{D'}$ or $\mathcal{L}_0 \simeq \mathcal{L}_D$ and $\mathcal{L}_{D'} \otimes \mathcal{L}_D \simeq \mathcal{O}_X$. Hence $[D']$ equals either $[D]$ or $[-D]$. □

Proposition 6.4. *The map* $\mathrm{Cl}\,X' / \mathrm{Cl}\,X \to \mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X \cup \{c_0\}$, $[D] \mapsto t_D$ *is surjective with fibers of the form* $\{[D], [-D]\}$.

Proof. From the previous considerations it is clear that this map is well-defined and surjective. Assume that $[D_1], [D_2] \in \mathrm{Cl}\,X'$ have the same image. Then there is an $\mathcal{L}_0 \in \mathrm{Pic}\,X$ such that $p_*\mathcal{L}_1 \simeq p_*\mathcal{L}_2 \otimes \mathcal{L}_0$, where we briefly wrote $\mathcal{L}_i$ for $\mathcal{L}_{D_i}$. Then in $\mathbb{P}\mathrm{Bun}_2\,X'$, we see that

$$\mathcal{L}_1 \oplus \mathcal{L}_1^\sigma \simeq p^* p_* \mathcal{L}_1 \simeq p^* p_* \mathcal{L}_2 \otimes p^* \mathcal{L}_0 \simeq (\mathcal{L}_2 \otimes p^* \mathcal{L}_0) \oplus (\mathcal{L}_2^\sigma \otimes p^* \mathcal{L}_0),$$

thus either $\mathcal{L}_1 \simeq \mathcal{L}_2 \otimes p^*\mathcal{L}_0$, which implies that D_1 and D_2 represent the same class in $\mathrm{Cl}\,X' / \mathrm{Cl}\,X$, or $\mathcal{L}_1 \simeq \mathcal{L}_2^\sigma \otimes p^*\mathcal{L}_0$, which means that D_1 represents the same class as σD_2. But in $\mathrm{Cl}\,X' / \mathrm{Cl}\,X$,

$$[\sigma D_2] = [\underbrace{\sigma D_2 + D_2}_{\in \mathrm{Cl}\,X} - D_2] = [-D_2]. \qquad \square$$

Lemma 6.5. *The constant extension restricts to an injective map*

$$p^* : \mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X \sqcup \mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X \hookrightarrow \mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X'.$$

Proof. Since $p^* p_*(\mathcal{L}) \simeq \mathcal{L} \oplus \mathcal{L}^\sigma$ for a line bundle $\mathcal{L}$ over X', it is clear that the image is contained in $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X'$. The images of $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X$ and $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X$ are disjoint since elements of the image of the latter set decompose into line bundles over X' that are not defined over X. If we denote taking the inverse elements by inv, then by Proposition 6.3, p^* is injective restricted to $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}}\,X$ because $(\mathrm{Cl}\,X/\,\mathrm{inv}) \to (\mathrm{Cl}\,X'/\,\mathrm{inv})$ is. Regarding $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}}\,X$, observe that

$$p^*(t_D) = p^* p_*(\mathcal{L}_D) \simeq \mathcal{L}_D \oplus \mathcal{L}_{\sigma D} \sim \mathcal{L}_{D - \sigma D} \oplus \mathcal{O}_{X'} = c_{D - \sigma D},$$

where by Proposition 6.4, D represents an element in $\big(\mathrm{Cl}\,X'/\,\mathrm{Cl}\,X\big)/\,\mathrm{inv}$, and by Proposition 6.3, $D - \sigma D$ represents an element in $\mathrm{Cl}\,X'/\,\mathrm{inv}$. If there are $[D_1]$, $[D_2] \in \mathrm{Cl}\,X'$ such that $(D_1 - \sigma D_1) = \pm(D_2 - \sigma D_2)$, then $D_1 \mp D_2 = \sigma(D_1 \mp D_2)$, and consequently $[D_1 \mp D_2] \in \mathrm{Cl}\,X$. □

Remark 6.6. The constant extension also restricts to a map

$$p^* : \mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}}\,X \to \mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}}\,X'.$$

But this restriction is in general not injective in contrast to the previous result. For a counterexample to injectivity, see [Lorscheid 2012, Remark 2.7].

7. Reduction theory for rank 2 bundles

In this section, we introduce reduction theory for rank 2 bundles, that is, an invariant δ closely related to the slope of a vector bundle and reduction theory. Namely, a rank 2 bundle $\mathcal{M}$ is (semi)stable if and only if $\delta(\mathcal{M})$ is negative (nonpositive). For the definition of the slope of a vector bundle and (semi)stable vector bundles, see [Harder and Narasimhan 1974/75]. The invariant δ is also defined for projective line bundles and will be help to determine the structure of the graphs $\mathcal{G}_x$.

7.1. In general, the cokernel of a sheaf morphism between two vector bundles might have nontrivial torsion. A *subbundle* of a vector bundle $\mathcal{M}$ is an injective morphism $\mathcal{M}' \to \mathcal{M}$ of vector bundles such that the cokernel is again a vector bundle. By a *line subbundle* $\mathcal{L} \to \mathcal{M}$ of a vector bundle $\mathcal{M}$, we mean a subbundle of $\mathcal{M}$ where $\mathcal{L}$ is a line bundle.

Every locally free subsheaf $\mathcal{L} \to \mathcal{M}$ of rank 1 extends to a uniquely determined line subbundle $\overline{\mathcal{L}} \to \mathcal{M}$, since $\overline{\mathcal{L}}$ is determined by the constraint $\mathcal{L} \subset \overline{\mathcal{L}}$ [Serre 2003, p. 100]. On the other hand, every rank 2 bundle has a line subbundle [Hartshorne 1977, Corollary V.2.7].

Two line subbundles $\mathcal{L} \to \mathcal{M}$ and $\mathcal{L}' \to \mathcal{M}$ are said to be the same if their images coincide, or, in other words, if there is an isomorphism $\mathcal{L} \simeq \mathcal{L}'$ that commutes with the inclusions into $\mathcal{M}$.

For a line subbundle $\mathcal{L} \to \mathcal{M}$ of a rank 2 bundle $\mathcal{M}$, we define

$$\delta(\mathcal{L}, \mathcal{M}) := \deg \mathcal{L} - \deg(\mathcal{M}/\mathcal{L}) = 2 \deg \mathcal{L} - \deg \mathcal{M},$$

$$\delta(\mathcal{M}) := \sup_{\substack{\mathcal{L} \to \mathcal{M} \\ \text{line subbundle}}} \delta(\mathcal{L}, \mathcal{M}).$$

If $\delta(\mathcal{M}) = \delta(\mathcal{L}, \mathcal{M})$, then we call $\mathcal{L}$ a *line subbundle of maximal degree*, or briefly, a *maximal subbundle*. Since $\delta(\mathcal{L} \otimes \mathcal{L}', \mathcal{M} \otimes \mathcal{L}') = \delta(\mathcal{L}, \mathcal{M})$ for a line bundle $\mathcal{L}'$, the invariant δ is well-defined on $\mathbb{P}\mathrm{Bun}_2 X$, and we put $\delta([\mathcal{M}]) = \delta(\mathcal{M})$.

Let g_X be the genus of X. Then the Riemann–Roch theorem and Serre duality imply:

Proposition 7.2 [Serre 2003, II.2.2, Propositions 6 and 7]. *Every rank* 2 *bundle* $\mathcal{M}$ *satisfies* $-2g_X \le \delta(\mathcal{M}) < \infty$. *If* $\mathcal{L} \to \mathcal{M}$ *is a line subbundle with* $\delta(\mathcal{L}, \mathcal{M}) > 2g_X - 2$, *then* $\mathcal{M} \simeq \mathcal{L} \oplus \mathcal{M}/\mathcal{L}$.

7.3. Every extension of a line bundle $\mathcal{L}'$ by a line bundle $\mathcal{L}$, that is, every exact sequence of the form $0 \to \mathcal{L} \to \mathcal{M} \to \mathcal{L}' \to 0$, determines a rank 2 bundle $\mathcal{M} \in \mathrm{Bun}_2 X$. This defines for all $\mathcal{L}, \mathcal{L}' \in \operatorname{Pic} X$ a map $\operatorname{Ext}^1(\mathcal{L}', \mathcal{L}) \to \mathrm{Bun}_2 X$, which maps the zero element to $\mathcal{L} \oplus \mathcal{L}'$. Since decomposable bundles may have line subbundles that differ from its given two factors, nontrivial elements can give rise to decomposable bundles.

Proposition 7.4. *The map*

$$\coprod_{\substack{-2g_X \le \deg \mathscr{L} \\ \le 2g_X - 2}} \operatorname{Ext}^1(\mathbb{O}_X, \mathscr{L}) \to \mathbb{P}\mathrm{Bun}_2 X$$

meets every element of $\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}} X$, *and the fiber of any* $[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2 X$ *is of the form*

$$\{0 \to \mathscr{L} \to \mathcal{M} \to \mathbb{O}_X \to 0 \mid \delta(\mathscr{L}, \mathcal{M}) \ge -2g_X\}.$$

Proof. We know that every $[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2 X$ has a reduction $0 \to \mathscr{L} \to \mathcal{M} \to \mathscr{L}' \to 0$ with $\delta(\mathscr{L}, \mathcal{M}) \ge -2g_X$, where we may assume that $\mathscr{L}' = \mathbb{O}_X$ by replacing $\mathcal{M}$ with $\mathcal{M} \otimes (\mathscr{L}')^{-1}$; hence $\delta(\mathscr{L}, \mathcal{M}) = \deg \mathscr{L}$. If $\deg \mathscr{L} > 2g_X - 2$, then $\mathcal{M}$ decomposes, that is, $\operatorname{Ext}^1(\mathbb{O}_X, \mathscr{L})$ is trivial (which is already clear from the proof [Serre 2003, II.2.2, Proposition 7]). This explains the form of the fibers and that $\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}} X$ is contained in the image. □

Corollary 7.5. *There are only finitely many isomorphism classes of indecomposable projective line bundles.*

Proof. This is clear since $\coprod_{-2g_X \le \deg \mathscr{L} \le 2g_X - 2} \operatorname{Ext}^1(\mathbb{O}_X, \mathscr{L})$ is a finite union of finite sets. □

Lemma 7.6. *If* $\mathscr{L} \to \mathcal{M}$ *is a maximal subbundle, then* $\delta(\mathscr{L}', \mathcal{M}) \le -\delta(\mathscr{L}, \mathcal{M})$ *for every line subbundle* $\mathscr{L}' \to \mathcal{M}$ *that is different from* $\mathscr{L} \to \mathcal{M}$. *Equality holds if and only if* $\mathcal{M} \simeq \mathscr{L} \oplus \mathscr{L}'$, *that is,* $\mathcal{M}$ *decomposes and* $\mathscr{L}'$ *is a complement of* $\mathscr{L}$ *in* $\mathcal{M}$.

Proof. Compare with [Schleich 1974, Lemma 3.1.1.]. Since $\mathscr{L}' \to \mathcal{M}$ is different from $\mathscr{L} \to \mathcal{M}$, there is no inclusion $\mathscr{L}' \to \mathscr{L}$ that commutes with the inclusions into $\mathcal{M}$. Hence the composed morphism $\mathscr{L}' \to \mathcal{M} \to \mathcal{M}/\mathscr{L}$ must be injective, and $\deg \mathscr{L}' \le \deg \mathcal{M}/\mathscr{L} = \deg \mathcal{M} - \deg \mathscr{L}$. This implies that

$$\delta(\mathscr{L}', \mathcal{M}) = 2 \deg \mathscr{L}' - \deg \mathcal{M} \le \deg \mathcal{M} - 2 \deg \mathscr{L} = -\delta(\mathscr{L}, \mathcal{M}).$$

Equality holds if and only if $\mathscr{L}' \to \mathcal{M}/\mathscr{L}$ is an isomorphism, and in this case, its inverse defines a section $\mathcal{M}/\mathscr{L} \simeq \mathscr{L}' \to \mathcal{M}$. □

Proposition 7.7.

(i) *A rank* 2 *bundle* $\mathcal{M}$ *has at most one line subbundle* $\mathscr{L} \to \mathcal{M}$ *such that* $\delta(\mathscr{L}, \mathcal{M}) \ge 1$.

(ii) *If* $\mathscr{L} \to \mathcal{M}$ *is a line subbundle with* $\delta(\mathscr{L}, \mathcal{M}) \ge 0$, *then* $\delta(\mathcal{M}) = \delta(\mathscr{L}, \mathcal{M})$.

(iii) *If* $\delta(\mathcal{M}) = 0$, *we distinguish three cases.*

(1) $\mathcal{M}$ *has only one maximal line bundle*; *this happens if and only if* $\mathcal{M}$ *is indecomposable.*

(2) $\mathcal{M}$ *has exactly two maximal subbundles* $\mathscr{L}_1 \to \mathcal{M}$ *and* $\mathscr{L}_2 \to \mathcal{M}$; *this happens if and only if* $\mathcal{M} \simeq \mathscr{L}_1 \oplus \mathscr{L}_2$ *and* $\deg \mathscr{L}_1 = \deg \mathscr{L}_2$, *but* $\mathscr{L}_1 \not\simeq \mathscr{L}_2$.

(3) $\mathcal{M}$ *has exactly* $q+1$ *maximal subbundles*; *this happens if and only if all maximal subbundles are of the same isomorphism type* $\mathcal{L}$ *and* $\mathcal{M} \simeq \mathcal{L} \oplus \mathcal{L}$.

(iv) $\delta(c_D) = \|\deg D\|$.

(v) $\delta(\mathcal{M})$ *is invariant under extension of constants for* $[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X$.

Proof. Everything follows from the preceding lemmas, except for the fact that $\mathcal{L} \oplus \mathcal{L}$ has precisely $q+1$ maximal subbundles in part (3), which needs some explanation.

If $\mathcal{M} = \mathcal{L} \oplus \mathcal{L}$ and $\mathcal{L}'$ is a third maximal subbundle of $\mathcal{M}$, then $\mathcal{M} \simeq \mathcal{L}' \oplus \mathcal{L}$ by Lemma 7.6, and thus there is an automorphism $\mathcal{M} \simeq \mathcal{L}' \oplus \mathcal{L} \to \mathcal{L} \oplus \mathcal{L} = \mathcal{M}$ that restricts to an isomorphism between $\mathcal{L}'$ and $\mathcal{L}$ by the Krull–Schmidt theorem; see [Atiyah 1956]. Thus the automorphism group $\mathrm{Aut}(\mathcal{M})$ of $\mathcal{M}$ acts transitively on the set of maximal line bundles of $\mathcal{M}$. Since $\mathrm{Aut}(\mathcal{M}) \simeq \mathrm{GL}_2(\mathbb{F}_q)$, the orbit of a maximal subbundle under $\mathrm{Aut}(\mathcal{M})$ is of cardinality $q+1$. □

Proposition 7.8. *Let* $p : X' = X \otimes \mathbb{F}_{q^2} \to X$ *and* $\mathcal{L} \in \mathrm{Pic}\, X'$, *then* $\delta(p_*\mathcal{L})$ *is an even nonpositive integer. It equals* 0 *if and only if* $\mathcal{L} \in p^* \mathrm{Pic}\, X$.

Proof. Over X', we have $p^* p_* \mathcal{L} \simeq \mathcal{L} \oplus \mathcal{L}^\sigma$ and $\deg \mathcal{L} = \deg \mathcal{L}^\sigma$. If $\mathcal{L}'$ is a line subbundle of $p_*\mathcal{L}$, then $p^*\mathcal{L}'$ is a subbundle of $\mathcal{L} \oplus \mathcal{L}^\sigma$. By the previous proposition, the degree of $p^*\mathcal{L}'$ (which is the same as the degree of $\mathcal{L}'$) equals the degree of $\mathcal{L}$ if and only if $p^*\mathcal{L}'$ is isomorphic to $\mathcal{L}$ or $\mathcal{L}^\sigma$, and it is smaller otherwise. In the former case, $\mathcal{L}$ is already defined over X; thus $p_*\mathcal{L} \simeq \mathcal{L}' \oplus \mathcal{L}'$ and $\delta(p_*\mathcal{L}) = 0$ if $\mathcal{L} \simeq p^*\mathcal{L}'$. In the latter case, that is, if $\mathcal{L}$ is not of the form $p^*\mathcal{L}'$ for a line bundle $\mathcal{L}'$ over X, we have $\delta(\mathcal{L}', p_*\mathcal{L}) < 0$ for every maximal subbundle $\mathcal{L}'$ of $p_*\mathcal{L}$. This shows that $\delta(p_*\mathcal{L})$ is nonpositive, and that it is 0 if and only if $\mathcal{L} \in p^* \mathrm{Pic}\, X$.

Finally note that by the very definition of $\delta(\mathcal{M})$ for rank 2 bundles $\mathcal{M}$, it follows that $\delta(\mathcal{M}) \equiv \deg \mathcal{M} \pmod 2$, and $\deg(p_*\mathcal{L}) = 2 \deg \mathcal{L}$ is even. □

Remark 7.9. We see that for $[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}} X$, the invariant $\delta(\mathcal{M})$ must get larger if we extend constants to $\mathbb{F}_{q^2}$, because $p^*(\mathcal{M})$ decomposes over X'. This stays in contrast to the result for classes in $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X$ (Proposition 7.7 (v)).

8. Nucleus and cusps

In this section, we will define certain subgraphs of $\mathcal{G}_x$ for a place x, namely, the cusp of a divisor class modulo x, which is an infinite subgraph of a simple nature, and the nucleus, which is a finite subgraph that depends heavily on the arithmetic of F. Finally, $\mathcal{G}_x$ can be described as the union of the nucleus with a finite number of cusps.

8.1. We use reduction theory to investigate sequences of the form

$$0 \to \mathcal{M}' \to \mathcal{M} \to \mathcal{K}_x \to 0,$$

which occur in the definition of $\mathcal{U}_x([\mathcal{M}])$. By additivity of the degree map (see Section 5.1), $\deg \mathcal{M}' = \deg \mathcal{M} - d_x$ where d_x is the degree of x.

Given an arbitrary inclusion $\mathcal{M}' \to \mathcal{M}$ of rank 2 bundles and a line subbundle $\mathcal{L} \to \mathcal{M}$, then we say that $\mathcal{L}$ lifts to $\mathcal{M}'$ if there exists a morphism $\mathcal{L} \to \mathcal{M}'$ such that the diagram

$$\begin{array}{ccc} & & \mathcal{L} \\ & \swarrow & \downarrow \\ \mathcal{M}' & \longrightarrow & \mathcal{M} \end{array}$$

commutes. In this case, $\mathcal{L} \to \mathcal{M}'$ is indeed a subbundle since otherwise it would extend nontrivially to a subbundle $\overline{\mathcal{L}} \to \mathcal{M}' \subset \mathcal{M}$ and would contradict the hypothesis that $\mathcal{L}$ is a subbundle of $\mathcal{M}$. In the case that $\mathcal{M}' \to \mathcal{M}$ is part of an exact sequence

$$0 \to \mathcal{M}' \to \mathcal{M} \to \mathcal{K}_x \to 0,$$

a line subbundle $\mathcal{L} \to \mathcal{M}$ lifts to $\mathcal{M}'$ if and only if the image of $\mathcal{L}$ in $\mathcal{K}_x$ is 0.

Let $\mathcal{I}_x \subset \mathcal{O}_X$ be the kernel of $\mathcal{O}_X \to \mathcal{K}_x$. This is also a line bundle, since $\mathcal{K}_x$ is a torsion sheaf. For every line bundle $\mathcal{L}$, we may think of $\mathcal{L}\mathcal{I}_x$ as a subsheaf of $\mathcal{L}$. In Pic X, the line bundle $\mathcal{I}_x$ represents the inverse of $\mathcal{L}_x$, the line bundle associated with the divisor x. In particular, $\deg \mathcal{I}_x = \deg \mathcal{L}_x^{-1} = -d_x$.

If $\mathcal{L} \to \mathcal{M}$ does not lift to a subbundle of $\mathcal{M}'$, we have that $\mathcal{L}\mathcal{I}_x \subset \mathcal{L} \to \mathcal{M}$ lifts to a subbundle of $\mathcal{M}'$:

$$\begin{array}{ccc} \mathcal{I}_x\mathcal{L} & \subset & \mathcal{L} \\ \downarrow & & \downarrow \\ \mathcal{M}' & \longrightarrow & \mathcal{M}. \end{array}$$

Note that every subbundle $\mathcal{L} \to \mathcal{M}'$ is a locally free subsheaf $\mathcal{L} \to \mathcal{M}$, which extends to a subbundle $\overline{\mathcal{L}} \to \mathcal{M}$. If thus $\mathcal{L} \to \mathcal{M}$ is a maximal subbundle that lifts to a subbundle $\mathcal{L} \to \mathcal{M}'$, then $\mathcal{L} \to \mathcal{M}'$ is a maximal subbundle. If, however, $\mathcal{L} \to \mathcal{M}$ is a maximal subbundle that does not lift to a subbundle $\mathcal{L} \to \mathcal{M}'$, then $\mathcal{L}\mathcal{I}_x \to \mathcal{M}'$ is a subbundle, which is not necessarily maximal. These considerations imply that

$$\delta(\mathcal{M}') \le 2 \deg \mathcal{L} - \deg \mathcal{M}' = 2 \deg \mathcal{L} - (\deg \mathcal{M} - d_x) = \delta(\mathcal{M}) + d_x,$$
$$\delta(\mathcal{M}') \ge 2 \deg \mathcal{I}_x\mathcal{L} - \deg \mathcal{M}' = 2 \deg \mathcal{L} - 2d_x - (\deg \mathcal{M} - d_x) = \delta(\mathcal{M}) - d_x.$$

Since $\delta(\mathcal{M}') \equiv \deg \mathcal{M}' = \deg \mathcal{M} - d_x \pmod 2$, we derive the following:

Lemma 8.2. *If* $0 \to \mathcal{M}' \to \mathcal{M} \to \mathcal{K}_x \to 0$ *is exact, then*

$$\delta(\mathcal{M}') \in \{\delta(\mathcal{M}) - d_x, \delta(\mathcal{M}) - d_x + 2, \dots, \delta(\mathcal{M}) + d_x\}.$$

8.3. Every line subbundle $\mathcal{L} \to \mathcal{M}$ defines a line $\mathcal{L}/\mathcal{L}\mathcal{I}_x$ in $\mathbb{P}\big(\mathcal{M}/(\mathcal{M}\otimes\mathcal{I}_x)\big)$. By the bijection

$$\left\{\begin{array}{c}\text{isomorphism classes of exact}\\ 0\to\mathcal{M}'\to\mathcal{M}\to\mathcal{K}_x\to 0\\ \text{with fixed }\mathcal{M}\end{array}\right\} \xrightarrow{1:1} \mathbb{P}(\mathcal{M}/(\mathcal{M}\otimes\mathcal{I}_x)),$$
$$(0\to\mathcal{M}'\to\mathcal{M}\to\mathcal{K}_x\to 0)\mapsto \mathcal{M}'/(\mathcal{M}\otimes\mathcal{I}_x)$$

(see Section 5.6), there is a unique $0\to\mathcal{M}'\to\mathcal{M}\to\mathcal{K}_x\to 0$ up to isomorphism with fixed $\mathcal{M}$, such that $\mathcal{M}'/(\mathcal{M}\otimes\mathcal{I}_x)=\mathcal{L}/\mathcal{L}\mathcal{I}_x$ in $\mathbb{P}(\mathcal{M}/(\mathcal{M}\otimes\mathcal{I}_x))$. This means that $\mathcal{L}$ is contained in the image of $\mathcal{M}'\to\mathcal{M}$ and that $\mathcal{L}\to\mathcal{M}$ lifts to a line subbundle $\mathcal{L}\to\mathcal{M}'$. We call $0\to\mathcal{M}'\to\mathcal{M}\to\mathcal{K}_x\to 0$ the *sequence associated with $\mathcal{L}\to\mathcal{M}$ relative to* Φ_x, or for short the *associated sequence*, and $[\mathcal{M}']$ the *associated Φ_x-neighbor*. It follows that $\delta(\mathcal{M}')\geq\delta(\mathcal{L},\mathcal{M})+d_x$.

We summarize this.

Lemma 8.4. *If $\mathcal{L}\to\mathcal{M}$ is a maximal subbundle, then the associated Φ_x-neighbor $[\mathcal{M}']$ has $\delta(\mathcal{M}')=\delta(\mathcal{M})+d_x$. Therefore,*

$$\sum_{\substack{([\mathcal{M}],[\mathcal{M}'],m)\in\mathcal{U}_x([\mathcal{M}])\\ \delta(\mathcal{M}')=\delta(\mathcal{M})+d_x}} m=\#\left\{\overline{\mathcal{L}}\in\mathbb{P}(\mathcal{M}/(\mathcal{M}\otimes\mathcal{I}_x))\;\middle|\;\begin{array}{c}\textit{there is a maximal submodule}\\ \mathcal{L}\to\mathcal{M}\textit{ with }\mathcal{L}\equiv\overline{\mathcal{L}}\pmod{\mathcal{M}\otimes\mathcal{I}_x}\end{array}\right\}.$$

Theorem 8.5. *Let x be a place and $[D]\in\operatorname{Cl}X$ be a divisor of nonnegative degree. The Φ_x-neighbors v of c_D with $\delta(v)=\deg D+d_x$ are given by the following list:*

$$\begin{aligned}(c_0,c_x,q+1)&\in\mathcal{U}_x(c_0),\\ (c_D,c_{D+x},2)&\in\mathcal{U}_x(c_D)\quad\textit{if }[D]\in(\operatorname{Cl}^0X)[2]-\{0\},\\ (c_D,c_{D+x},1),(c_D,c_{-D+x},1)&\in\mathcal{U}_x(c_D)\quad\textit{if }[D]\in\operatorname{Cl}^0X-(\operatorname{Cl}^0X)[2],\\ (c_D,c_{D+x},1)&\in\mathcal{U}_x(c_D)\quad\textit{if }\deg D\textit{ is positive.}\end{aligned}$$

For all Φ_x-neighbors v of c_D not occurring in this list, $\delta(v)<\delta(c_D)+d_x$. If furthermore $\deg D>d_x$, *then $\delta(v)=\deg D-d_x$, and if* $\deg D>m_X+d_x$ *where* $m_X=\max\{2g_X-2,0\}$, *then*

$$\mathcal{U}_x(c_D)=\{(c_D,c_{D-x},q_x),(c_D,c_{D+x},1)\}.$$

Proof. By Lemma 8.4, the Φ_x-neighbors v of c_D with $\delta(v)=\delta(c_D)+d_x$ counted with multiplicity correspond to the maximal subbundles of a rank 2 bundle $\mathcal{M}$ that represents c_D. Since $\delta(\mathcal{M})=\delta(c_D)\geq 0$, the list of all Φ_x-neighbors v of c_D with $\delta(v)=\deg D+d_x=\delta(c_D)+d_x$ follows from the different cases in Proposition 7.7 (i) and (iii). Be aware that $c_D=c_{-D}$ by Proposition 6.3; hence it makes a difference whether or not D is 2-torsion.

For the latter statements, write $\mathcal{M} = \mathcal{L}_D \oplus \mathcal{O}_X$ and let $\mathcal{M}'$ be a subsheaf of $\mathcal{M}$ with cokernel $\mathcal{K}_x$ such that $\delta(\mathcal{M}') < \delta(\mathcal{M}) + d_x$. Then $\mathcal{L}_D \to \mathcal{M}$ does not lift to $\mathcal{M}'$, but $\mathcal{L}_D\mathcal{I}_x \to \mathcal{M}'$ is a line subbundle and

$$\mathcal{M}'/\mathcal{L}_D\mathcal{I}_x \simeq (\det \mathcal{M}')(\mathcal{L}_D\mathcal{I}_x)^\vee \simeq (\det \mathcal{M})\mathcal{I}_x(\mathcal{L}_D\mathcal{I}_x)^\vee \simeq \mathcal{L}_D\mathcal{I}_x(\mathcal{L}_D\mathcal{I}_x)^\vee \simeq \mathcal{O}_X.$$

If $\deg D > d_x$, then

$$\delta(\mathcal{L}_D\mathcal{I}_x, \mathcal{M}') = \deg \mathcal{L}_D\mathcal{I}_x - \deg \mathcal{O}_X = \deg D - d_x > 0.$$

Proposition 7.7(i) implies that $\mathcal{L}_D \to \mathcal{M}$ is the unique maximal subbundle of $\mathcal{M}'$ and thus $\delta(\mathcal{M}') = \delta(\mathcal{M}) - d_x$.

If $\delta(\mathcal{M}) > m_X + d_x$, then $\delta(\mathcal{M}') > m_X \geq 2g_X - 2$; hence $\mathcal{M}'$ decomposes and represents c_{D-x}. Since the multiplicities of all Φ_x-neighbors of a vertex sum up to $q_x + 1$, this proves the last part of our assertions. □

Definition 8.6. Let x be a place. Let the divisor D represent a class

$$[D] \in \mathrm{Cl}\, \mathcal{O}_X^x = \mathrm{Cl}\, X/\langle x\rangle.$$

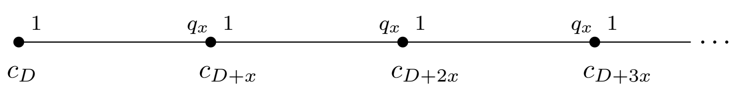
We define the *cusp* $\mathcal{C}_x(D)$ *(of D in* $\mathcal{G}_x$*)* as the full subgraph of $\mathcal{G}_x$ with vertices

$$\mathrm{Vert}\, \mathcal{C}_x(D) = \{c_{D'} \mid [D'] \equiv [D] \pmod{\langle x\rangle}, \text{ and } \deg D' > m_X\},$$

and the *nucleus* $\mathcal{N}_x$ *(of* $\mathcal{G}_x$*)* as the full subgraph of $\mathcal{G}_x$ with vertices

$$\mathrm{Vert}\, \mathcal{N}_x = \{[\mathcal{M}] \in \mathbb{P}\mathrm{Bun}_2\, X \mid \delta(\mathcal{M}) \leq m_X + d_x\}.$$

8.7. Theorem 8.5 determines all edges of a cusp $\mathcal{C}_x(D)$. If $m_X < \deg D \leq m_X + d_x$, the cusp can be illustrated as below. Note that a cusp is an infinite graph. It has a regular pattern that repeats periodically. In diagrams we draw the pattern and indicate its periodic continuation with dots.



We summarize the theory so far in the following theorem that describes the general structure of $\mathcal{G}_x$.

Theorem 8.8. *Let x be a place of degree d_x and $h_X = \#\,\mathrm{Cl}^0\, X$ be the class number.*

(i) *$\mathcal{G}_x$ has $h_X d_x$ cusps and*

$$\mathcal{G}_x = \mathcal{N}_x \cup \coprod_{[D] \in \mathrm{Cl}\, \mathcal{O}_F^x} \mathcal{C}_x(D),$$

where $\mathrm{Vert}\, \mathcal{N}_x \cap \mathrm{Vert}\, \mathcal{C}_x(D) = \{c_D\}$ *if $m_X < \deg D \leq m_X + d_x$. The union of the edges is disjoint. Different cusps are disjoint subgraphs.*

(ii) *$\mathcal{N}_x$ is finite and has $\#(\mathrm{Cl}\,\mathcal{O}_F^x\,/\,2\,\mathrm{Cl}\,\mathcal{O}_F^x)$ components. Each vertex of $\mathcal{N}_x$ is at distance $\leq (2g_X + m_X + d_x)/d_x$ from some cusp. The associated CW-complexes of $\mathcal{N}_x$ and $\mathcal{G}_x$ are homotopy equivalent.*

(iii) *If $[D] \in \mathrm{Cl}\,\mathcal{O}_F^x$, then $\mathrm{Vert}\,\mathcal{C}_x(D) \subset \mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X$. Furthermore,*

$$\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X \subset \{v \in \mathrm{Vert}\,\mathcal{G}_x \mid \delta(v) \geq 0\},$$
$$\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}} X \subset \{v \in \mathrm{Vert}\,\mathcal{G}_x \mid \delta(v) \leq 2g_X - 2\},$$
$$\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}} X \subset \{v \in \mathrm{Vert}\,\mathcal{G}_x \mid \delta(v) < 0 \textit{ and even}\}.$$

8.9. *Remark on Figure 8a.* Define $h = h_X,\ m = m_X,\ d = d_x$ and $q_x = q^{\deg x}$. Further let $D_1, \ldots, D_{hd}$ be representatives for $\mathrm{Cl}\,\mathcal{O}_F^x$ with $m < \deg D_i \leq m + d$ for $i = 1, \ldots, hd$. The cusps $\mathcal{C}_x(D_i)$ for $i = 1, \ldots, hd$ can be seen in Figure 8a as the subgraphs in the dashed regions that are open to the right. The nucleus $\mathcal{N}_x$ is contained in the dashed rectangle to the left. Since we have no further information about the nucleus, we leave the area in the rectangle open.

The δ-line on the bottom of the picture indicates the value $\delta(v)$ for the vertices v in the graph that lie vertically above $\delta(v)$.

The dotted regions refer to the sort of vertices, which are elements of either $\mathbb{P}\mathrm{Bun}_2^{\mathrm{gi}} X$, $\mathbb{P}\mathrm{Bun}_2^{\mathrm{tr}} X$, or $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X$. All lines are drawn with reference to the δ-line to reflect part (iii) of the theorem.

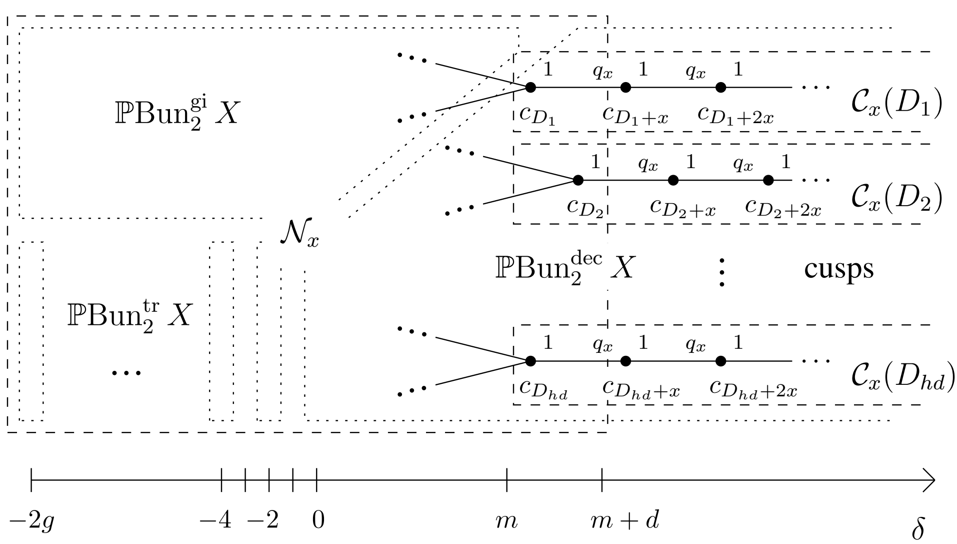


Figure 8a. General structure of $\mathcal{G}_x$.

Proof. The number of cusps is $\#\operatorname{Cl}\mathcal{O}_X^x = \#(\operatorname{Cl} X/\langle x\rangle) = \#\operatorname{Cl}^0 X \cdot \#(\mathbb{Z}/d_x\mathbb{Z}) = h_X d_x$. That the vertices of cusps are disjoint and only intersect in the given point with the nucleus is clear by definition. Regarding the edges, recall from Section 3.2 that if there is an edge from v to w in $\mathcal{G}_x$, then there is also an edge from w to v. But Theorem 8.5 implies that each vertex of a cusp that does not lie in the nucleus only connects to a vertex of the same cusp; hence every edge of $\mathcal{G}_x$ either lies in a cusp or in the nucleus. Different cusps are disjoint by definition. This shows (i).

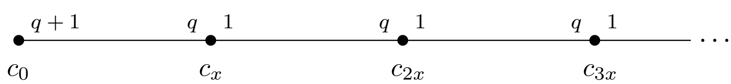
The nucleus is finite since $\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}} X$ is finite by Corollary 7.5 and since the intersection $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X \cap \operatorname{Vert}\mathcal{N}_x$ is finite by the definition of the nucleus and Proposition 6.3. Since the cusps are contractible as CW-complexes, $\mathcal{N}_x$ and $\mathcal{G}_x$ have the same homotopy type. Therefore $\mathcal{N}_x$ has $\#(\operatorname{Cl}\mathcal{O}_F^x \,/\, 2\mathcal{O}_F^x)$ components by Proposition 3.8. By Lemma 8.4, every vertex v has a Φ_x-neighbor w with $\delta(w) = \delta(v) + d_x$, which is the upper bound for the distance of vertices in the nucleus to one of the cusps. This proves (ii).

The four statements of (iii) follow from the definition of a cusp, Propositions 7.7(iv), 7.2 and 7.8, respectively. □

Example 8.10 (the projective line). Let X be the projective line over $\mathbb{F}_q$. Then $g_X = 0$, $h_X = 1$ and X has a closed point x of degree 1. This means that

$$\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X = \{c_{nx}\}_{n\geq 0}.$$

Since an indecomposable bundle $\mathcal{M}$ must satisfy both $\delta(\mathcal{M}) \geq 0$ and $\delta(\mathcal{M}) \leq -2$, which is impossible, all projective line bundles decompose. Theorem 8.5 together with the fact that the weights around each vertex sum to $q+1$ in the graph of Φ_x determines $\mathcal{G}_x$ completely, as illustrated here:



9. Application to automorphic forms

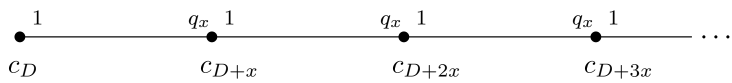
In this section, we explain how to recover automorphic forms as functions on the graph and indicate how unramified automorphic forms can be explicitly calculated as functions on the graph by solving a finite system of linear equations. We begin by recalling the definition of an automorphic form.

9.1. A function $f \in C^0(G_\mathbb{A})$ is called an *automorphic form* (*for* PGL_2 *over* F) if there is a compact open subgroup K' of $G_\mathbb{A}$ such that f is left G_F-invariant and right K'-invariant and if it generates a finite-dimensional $\mathcal{H}_{K'}$-subrepresentation $\mathcal{H}_{K'}(f)$ of $C^0(G_\mathbb{A})$. We denote the space of automorphic forms by $\mathcal{A}$ and note that the action of $\mathcal{H}$ on $C^0(G_\mathbb{A})$ restricts to $\mathcal{A}$. We denote the subspace of right K'-invariant automorphic forms by $\mathcal{A}^{K'}$, a space on which $\mathcal{H}_{K'}$ acts. We can reinterpret the elements in $\mathscr{A}^{K'}$ as functions on $G_F \setminus G_{\mathbb{A}}/K'$, which is the vertex set of the graph $\mathscr{G}_{\Phi,K'}$ of a Hecke operator $\Phi \in \mathscr{H}_{K'}$.

We shall investigate the space $\mathscr{A}^K$ of unramified automorphic forms in more detail. We write $f(v)$ or $f(\mathscr{M})$ for the value $f(g)$ if $v = [g]$ is the class of g in $G_F \setminus G_{\mathbb{A}}/K$ and $\mathscr{M} = \mathscr{M}_g$ is the rank 2 bundle that corresponds to g. In particular, we can see f also as a function on $\mathbb{P}\mathrm{Bun}_2\, X$.

The space of automorphic forms decomposes into a cuspidal part $\mathscr{A}_0$, a part $\mathscr{E}$ that is generated by Eisenstein series and their derivatives and a part $\mathscr{R}$ that is generated by residues of Eisenstein series and their derivatives (for complete definitions, see [Lorscheid 2010, Section 9.1]). The decomposition descends to unramified automorphic forms: $\mathscr{A}^K = \mathscr{A}_0^K \oplus \mathscr{E}^K \oplus \mathscr{R}^K$. We describe functions in these parts separately.

9.2. We start with some considerations for Φ_x-eigenfunctions as functions on a cusp $\mathscr{C}_x(D)$ where D is a divisor with $m_X < \deg D \leq m_X + d_x$:



Let $f \in \mathscr{A}^K$ satisfy the eigenvalue equation $\Phi_x f = \lambda f$, then we obtain for every $i \geq 1$,

$$f(c_{D+(i+1)x}) = \lambda f(c_{D+ix}) - q_x f(c_{D+(i-1)x}). \tag{9-1}$$

Thus the restriction of f to $\mathrm{Vert}\,\mathscr{C}_x(D)$ is determined by the eigenvalue λ once its values at c_D and c_{D+x} are given. The eigenvalue equation evaluated at c_D shows further that $f(c_{D+x})$ is a linear combination of values of f in vertices of the nucleus. This consideration justifies that we only have to evaluate the eigenvalue equation at vertices of the nucleus to determine the eigenfunctions of Φ_x.

9.3. The space $\mathscr{A}_0^K$ has a basis of $\mathscr{H}_K$-eigenfunctions and every unramified cusp form has a compact, that is, finite, support in $G_F \setminus G_{\mathbb{A}}/K$. By the eigenvalue (9-1) it follows that a Hecke eigenfunction $f \in \mathscr{A}_0^K$ must vanish on all vertices of a cusp in order to have compact support. Thus the support of a cusp form is contained in the finite set V of vertices v with $\delta(v) \leq m_X$, and $\mathscr{A}_0^K$ can be determined by considering a finite number of eigenvalue equations for Φ_x.

These eigenvalue equations can be described in terms of the matrix M_x associated with Φ_x; see Section 1.8. Namely, $\mathscr{A}_0^K$ is generated by the eigenfunctions of M_x whose support is contained in V. This problem can be rephrased into a question on the finite submatrix $M'_x = (a_{v,w})_{v \in V, w \in \mathrm{Vert}\,\mathscr{N}_x}$ of $M_x = (a_{v,w})_{v,w \in \mathrm{Vert}\,\mathscr{G}_x}$, which we forgo spelling out.

In [Moreno 1985] one finds a finite set S of places such that an $\mathscr{H}_K$-eigenfunction $f \in \mathscr{A}_0^K$ is already characterized (up to multiple) by its Φ_x-eigenvalues for $x \in S$.

This means that one finds the cuspidal $\mathcal{H}_K$-eigenfunctions by considering the eigenvalue equations for the finitely many vertices $v \in V$ and the finitely many Hecke operators Φ_x for $x \in S$.

9.4. We proceed with $\mathcal{E}^K \oplus \mathcal{R}^K$. This space decomposes into a direct sum of generalized (infinite-dimensional) Hecke eigenspaces $\mathcal{E}(\chi)$, where χ runs through all unramified Hecke characters, that is, continuous group homomorphisms

$$\chi : F^\times \backslash \mathbb{A}^\times / \mathcal{O}_\mathbb{A}^\times \to \mathbb{C}^\times,$$

modulo inversion; in particular, $\mathcal{E}(\chi) = \mathcal{E}(\chi^{-1})$. The generalized eigenspace $\mathcal{E}(\chi)$ is characterized by its unique Hecke eigenfunction $\tilde{E}(\,\cdot\,,\chi)$ (up to scalar multiple), which in turn is determined by its Φ_x-eigenvalues $\lambda_x(\chi) = q_x^{1/2}(\chi(\pi_x) + \chi^{-1}(\pi_x))$ for $x \in \|X\|$. We have $\mathcal{E}(\chi) \subset \mathcal{E}$ if and only if $\chi^2 \neq \|\cdot\|^{\pm 1}$, in which case $\tilde{E}(\,\cdot\,,\chi)$ is an Eisenstein series. For $\chi^2 = \|\cdot\|^{\pm 1}$, $\tilde{E}(\,\cdot\,,\chi)$ is a residue of an Eisenstein series. For details, see [Lorscheid 2010], in particular, Theorem 11.10.

We say that a subset $S \subset \|X\|$ generates $\mathrm{Cl}\, X$ if the classes of the prime divisors corresponding to the places in S generate $\mathrm{Cl}\, X$. Let S be a set of places that generates $\mathrm{Cl}\, X$ and satisfies that for every decomposition $S = S_+ \cup S_-$ either $2\,\mathrm{Cl}\, X = 2\langle S_+\rangle$ or $2\,\mathrm{Cl}\, X = 2\langle S_-\rangle$. This set can be chosen to be finite. Then the Hecke eigenfunction $\tilde{E}(\,\cdot\,,\chi)$ is uniquely determined (up to scalar multiples) by the Φ_x-eigenvalues $\lambda_x(\chi)$. For details, see [Lorscheid 2008, Theorem 3.7.6 and Section 3.7.10].

In order to describe an Eisenstein series or a residue of an Eisenstein series, one only needs to consider the finitely many eigenvalue equations for the vertices in the nuclei $\mathcal{N}_x$ of the finitely many Hecke operators Φ_x with $x \in S$. Derivatives of Eisenstein series or residues are similarly determined by generalized eigenvalue equations; see [Lorscheid 2010, Lemmas 11.2 and 11.7] for the explicit formulas.

In the case of a residue, that is, $\chi^2 = \|\cdot\|^{\pm 1}$, the function $f = \tilde{E}(\,\cdot\,,\chi)$ has a particular simple form. Namely, χ is of the form $\omega\|\cdot\|^{\pm 1/2}$ where $\omega^2 = 1$ and $\tilde{E}(g,\chi) = \omega \circ \det(g)$. This means that $f(g\xi_w) = \omega(\pi_x \det g) = \omega(\pi_x) f(g)$. Thus, as a function on $\mathrm{Vert}\,\mathcal{G}$, f satisfies $f(v) = \omega(\pi_x) f(w)$ for all adjacent vertices v and w.

Remark 9.5. The methods of this paragraph will be applied in [Lorscheid 2012] to determine the space of unramified cusp forms for an elliptic function field and to show that there are no unramified toroidal cusp forms in this case.

10. Finite-dimensionality results

In this section, we will show how the theory of the last sections can be used to show finite-dimensionality of subspaces of $C^0(G_\mathbb{A})^K$ whose elements f are defined by a condition of the form

$$\sum_{i=1}^{n} m_i \Phi(f)(g_i) = 0$$

for all $\Phi \in \mathcal{H}_K$ (with $m_i \in \mathbb{C}$ and $g_i \in G_{\mathbb{A}}$ being fixed). We will explain a general technique and apply it to show that the spaces of functions in $C^0(G_{\mathbb{A}})^K$ satisfying the cuspidal condition or the toroidal condition, respectively, are finite-dimensional. In particular, this implies that all functions satisfying one of these conditions are automorphic forms.

10.1. Write $\mathrm{Cl}^{\mathrm{pr}} X$ for the set of divisor classes that are represented by prime divisors and $\mathrm{Cl}^{\mathrm{eff}} X$ for the semigroup they generate, that is, for all classes that are represented by effective divisors. In particular, $\mathrm{Cl}^{\mathrm{eff}} X$ contains 0, the class of the zero divisor, and for all other $[D] \in \mathrm{Cl}^{\mathrm{eff}} X$, we have $\deg D > 0$. Denote by $\mathrm{Cl}^d X$ the set of divisor classes of degree d and by $\mathrm{Cl}^{\geq d} X$ the set of divisor classes of degree at least d. Let g_X be the genus of X.

Lemma 10.2. $\mathrm{Cl}^{\geq g_X} X \subset \mathrm{Cl}^{\mathrm{eff}} X.$.

Proof. Let C be a canonical divisor on X, which is of degree $2g_X - 2$. For a divisor D, define $l(D) = \dim_{\mathbb{F}_q} H^0(X, \mathcal{L}_D)$. We have $[D] \in \mathrm{Cl}^{\mathrm{eff}} X$ if and only if $l(D) > 0$; see [Hartshorne 1977, Section IV.1]. The Riemann–Roch theorem is

$$l(D) - l(D - C) = \deg D + 1 - g_X;$$

see [Hartshorne 1977, Theorem IV.1.3].

If now $[D] \in \mathrm{Cl}^{\geq g_X} X$, then $\deg D \geq g_X$ and the Riemann–Roch theorem implies that $l(D) \geq \deg D + 1 - g_X > 0$. □

10.3. Let D be an effective divisor. Then it can be written in a unique way up to permutation of terms as a sum of prime divisors $D = x_1 + \cdots + x_n$. We define Φ_0 as the identity operator and set $\Phi_D = \Phi_{x_1} \cdots \Phi_{x_n}$. Since $\mathcal{H}_K$ is commutative, Φ_D is well-defined. Further we briefly write $\mathcal{G}_D$ for the graph $\mathcal{G}_{\Phi_D,K}$ of Φ_D, and $\mathcal{U}_D(v)$ for $\mathcal{U}_{\Phi_D,K}(v)$.

Let $[D] \in \mathrm{Cl}\, X$. Recall from Section 5.1 that $\mathcal{L}_D$ denotes the associated line bundle and from Section 6.2 that c_D denotes the vertex that is represented by $\mathcal{L}_D \oplus \mathcal{O}_X$. Recall from Proposition 7.7(iv) that $\delta(c_D) = \|\deg D\|$, where δ is defined as in Section 7.1.

Lemma 10.4. *Let D be an effective divisor.*

(i) *Let $v, v' \in \mathrm{Vert}\, \mathcal{G}_D$. If v' is a Φ_D-neighbor of v, then* $\|\delta(v') - \delta(v)\| \leq \deg D$.

(ii) *Let $[\mathcal{M}] \in \mathrm{Vert}\, \mathcal{G}_D$. Every maximal subbundle $\mathcal{L} \to \mathcal{M}$ lifts to a maximal subbundle $\mathcal{L} \to \mathcal{M}'$ of a uniquely determined rank 2 bundle $\mathcal{M}'$ such that $[\mathcal{M}']$ is a Φ_D-neighbor of $[\mathcal{M}]$ with $\delta(\mathcal{M}') = \delta(\mathcal{M}) + \deg D$. Conversely, every* *maximal subbundle* $\mathcal{L} \to \mathcal{M}'$ *extends to a maximal subbundle* $\mathcal{L} \to \mathcal{M}$ *if* $[\mathcal{M}']$ *is a* Φ_D*-neighbor of* $[\mathcal{M}]$ *with* $\delta(\mathcal{M}') = \delta(\mathcal{M}) + \deg D$.

Proof. We do induction on the number of factors in $\Phi_D = \Phi_{x_1} \cdots \Phi_{x_n}$ with $x_1, \ldots, x_n$ being prime divisors. The lemma is trivial for the identity operator Φ_0.

If $n \geq 1$, write $x = x_n$ and $\Phi_D = \Phi_{D'}\Phi_x$ for the effective divisor

$$D' = x_1 + \cdots + x_{n-1},$$

which is of degree $\deg D' = \deg D - \deg x$. Assume that (i) and (ii) hold for D'. Let v' be a Φ_D-neighbor of v. Let m be the weight of the edge (v, v', m). As explained in Section 1.7, we have

$$\sum_{\substack{(v,v'',m')\in \mathrm{Edge}\,\mathcal{G}_{D'} \\ (v'',v',m'')\in \mathrm{Edge}\,\mathcal{G}_x}} m' \cdot m'' = m \neq 0,$$

which means that there is a v'' that is a $\Phi_{D'}$-neighbor of v and a Φ_x-neighbor of v'. Thus the inductive hypothesis and Lemma 8.2 imply

$$\|\delta(v') - \delta(v)\| \leq \|\delta(v') - \delta(v'')\| + \|\delta(v'') - \delta(v)\| \leq \deg D' + \deg x = \deg D.$$

This proves (i).

We proceed with (ii). Let $\mathcal{L} \to \mathcal{M}$ be a maximal subbundle. By the inductive hypothesis, there is a $\Phi_{D'}$-neighbor $\mathcal{M}''$ of $\mathcal{M}$ such that $\mathcal{L} \to \mathcal{M}$ lifts to a maximal subbundle of $\mathcal{M}''$ and such that $\delta(\mathcal{M}'') = \delta(\mathcal{M}) + \deg D'$. Let

$$0 \to \mathcal{M}' \to \mathcal{M}'' \to \mathcal{K}_x \to 0$$

be the sequence associated with $\mathcal{L} \to \mathcal{M}''$. This means that $\mathcal{L}$ lifts to a subbundle of $\mathcal{M}'$. As explained in Section 8.3, $\delta(\mathcal{L}, \mathcal{M}') = \delta(\mathcal{L}, \mathcal{M}'') + \deg x$, where $\delta(\mathcal{L}, \mathcal{M}'') = \delta(\mathcal{M}'')$ by the maximality of $\mathcal{L}$. By part (i) of the lemma, we have $\delta(\mathcal{M}') \leq \delta(\mathcal{M}'') + \deg x = \delta(\mathcal{L}, \mathcal{M}')$, which must be an equality in this case. Therefore $\mathcal{L} \to \mathcal{M}'$ is maximal and

$$\delta(\mathcal{M}') = \delta(\mathcal{M}'') + \deg x = \delta(\mathcal{M}) + \deg D' + \deg x = \delta(\mathcal{M}) + \deg D,$$

as desired.

Assume conversely that $\mathcal{M}'$ is a Φ_D-neighbor of $\mathcal{M}'$ with $\delta(\mathcal{M}') = \delta(\mathcal{M}) + \deg D$ and let $\mathcal{L} \to \mathcal{M}'$ be a maximal subbundle. As already explained in the proof of (i), there is an $\mathcal{M}''$, which is a $\Phi_{D'}$-neighbor of $\mathcal{M}'$ and a Φ_x-neighbor of $\mathcal{M}$. By (i), the difference of $\delta(\mathcal{M})$ and $\delta(\mathcal{M}')$ is maximal; therefore it must hold that $\delta(\mathcal{M}') = \delta(\mathcal{M}'') + \deg D'$ and $\delta(\mathcal{M}'') = \delta(\mathcal{M}) + \deg x$. By the inductive hypothesis, $\mathcal{L} \to \mathcal{M}''$ is a maximal subbundle, that is, $\delta(\mathcal{M}'') = \delta(\mathcal{L}, \mathcal{M}'')$. We derive

$$\delta(\mathcal{M}'') = \delta(\mathcal{M}) + \deg x \geq 2\deg \mathcal{L} - \deg \mathcal{M} + \deg x = 2\deg \mathcal{L} - \deg \mathcal{M}'' = \delta(\mathcal{M}'').$$

Consequently, all inequalities are equalities and $\mathscr{L} \to \mathscr{M}$ is a maximal subbundle, what was to be shown. □

10.5. We demonstrate how to use the lemma to show that the space $\mathscr{V}_0$ of all unramified functions on $G_F \setminus G_{\mathbb{A}}$ that satisfy the cuspidal condition is finite-dimensional. Namely, let $N \subset G$ be a unipotent subgroup. Then the cuspidal condition for $f \in C^0(G_F \setminus G_{\mathbb{A}})^K$ is that

$$\int_{N_F \setminus N_{\mathbb{A}}} \Phi(f)(n)\, dn = 0 \quad \text{for all } \Phi \in \mathscr{H}.$$

If f is an automorphic form, then this condition defines a cusp form. A posteriori it will be clear that $\mathscr{V}_0$ contains only automorphic forms and thus equals the space $\mathscr{A}_0^K$ of unramified cusp forms.

Theorem 10.6. *The dimension of $\mathscr{V}_0$ is finite and bounded by*

$$\dim \mathscr{V}_0 \leq \#\{[\mathscr{M}] \in \mathbb{P}\mathrm{Bun}_2 X \mid \delta(\mathscr{M}) \leq m_X\}.$$

Proof. Note that there are only finitely many projective line bundles $[\mathscr{M}]$ with $\delta(\mathscr{M}) \leq m_X$ since $\mathbb{P}\mathrm{Bun}_2^{\mathrm{indec}} X$ is finite and $\mathbb{P}\mathrm{Bun}_2^{\mathrm{dec}} X$ has only finitely many classes $[\mathscr{M}]$ with $\delta(\mathscr{M}) \leq m_X$. So the finite-dimensionality of $\mathscr{V}_0$ will follow from the inequality.

We proceed with the proof of the inequality. The geometric equivalent of the cuspidal condition is that

$$\sum_{\mathscr{M} \in \mathrm{Ext}^1(\mathscr{O}_X, \mathscr{O}_X)} \Phi(f)(\mathscr{M}) = 0 \quad \text{for all } \Phi \in \mathscr{H};$$

see [Gaitsgory 2003].

Since $\delta(\mathscr{O}_X, \mathscr{M}) = 0$ for $\mathscr{M} \in \mathrm{Ext}^1(\mathscr{O}_X, \mathscr{O}_X)$, we have that $\mathscr{O}_X \to \mathscr{M}$ is a maximal subbundle by Proposition 7.7(ii), and only in the case of the trivial extension $\mathscr{M} \simeq \mathscr{O}_X \oplus \mathscr{O}_X$ are there other maximal subbundles, namely, there exist $(q+1)$ different subbundles of the form $\mathscr{O}_X \to \mathscr{M}$. Note that in any case, $\delta(\mathscr{M}) = 0$.

Let D be a nontrivial effective divisor. In case $\mathscr{M}$ is the trivial extension $\mathscr{O}_X \oplus \mathscr{O}_X$, the vertex $c_0 = [\mathscr{M}]$ has the unique Φ_D-neighbor $v' = c_D$ with $\delta(v') = \deg D$, which is of multiplicity $q+1$, as follows from an easy induction using Theorem 8.5 and Lemma 10.4. In case $\mathscr{M}$ is a nontrivial extension of $\mathscr{O}_X$ by itself, the vertex $v = [\mathscr{M}]$ has a unique Φ_D-neighbor $v' = [\mathscr{M}']$ with $\delta(v') - \delta(v) = \deg D$, which has a unique maximal subbundle, namely, $\mathscr{O}_X \to \mathscr{M}'$.

Thus for every $\mathscr{M} \in \mathrm{Ext}^1(\mathscr{O}_X, \mathscr{O}_X)$ and every Φ_D-neighbor $[\mathscr{M}']$ of $[\mathscr{M}]$ with $\delta(\mathscr{M}') = \deg D$, the maximal subbundles of $\mathscr{M}'$ are of the form $\mathscr{O}_X \to \mathscr{M}'$. Thus if $\deg D > m_X$, then $\mathscr{M}' \simeq \mathscr{O}_X \oplus (\mathscr{M}'/\mathscr{O}_X)$ by Proposition 7.2. Since the determinant is multiplicative and $\det \mathscr{K}_x \simeq \mathscr{L}_x$ (see [Hartshorne 1977, Ex. 6.11]), a short exact sequence $0 \to \mathcal{M}_1 \to \mathcal{M}_2 \to \mathcal{K}_x \to 0$ yields $\det \mathcal{M}_2 \simeq \mathcal{L}_x \otimes \det \mathcal{M}_1$. An easy induction over the length of the prime decomposition $D = x_1 + \cdots + x_n$ shows that $\det \mathcal{M} \simeq \mathcal{L}_D \otimes \det \mathcal{M}'$. Therefore we have $\mathcal{M}'/\mathbb{O}_X \simeq \mathcal{L}_{-D}$, which shows that $[\mathcal{M}'] = c_D$.

We finish the proof of the theorem by showing that every $f \in \mathcal{V}_0$ is determined by its values in the vertices v with $\delta(v) \le m_X$. We make an induction on $d = \delta(c_D)$, where c_D varies through all vertices v with $\delta(v) > m_X$.

Let $d > m_X$. Assume that the values of f in all vertices v with $\delta(v) < d$ are given (which is the case when $d = m_X + 1$; thus the initial step). Let v be a vertex with $\delta(v) = d$. Then $v = c_D$ for an effective divisor D by Lemma 10.2 since $m_X = \max\{0, 2g_X - 2\} \ge g_X - 1$. For the Hecke operator Φ_D, the cuspidal condition reads by the previous argumentation and Lemma 10.4 as

$$(q + q^{e_1}) \cdot f(c_D) + \sum_{\delta(v')<d} a_{v'} f(v') = 0$$

for certain $a_{v'}$ and $e_1 = \dim \operatorname{Ext}^1(\mathbb{O}_X, \mathbb{O}_X)$. Thus $f(v)$ is determined by the values $f(v')$ in vertices v' with $\delta(v') < d$, which proves the theorem. □

10.7. While the finite-dimensionality of $\mathcal{V}_0$ can also be established without the techniques of this paper, we do not know any other method to prove the corresponding fact for toroidal functions. For more details on the following definitions, see [Lorscheid 2010].

Choose a basis of $\mathbb{F}_{q^2}$ over $\mathbb{F}_q$. This defines an embedding of $E = \mathbb{F}_{q^2} F$ into the algebra of 2×2-matrices with entries in F. The image of $E^\times$ is contained in $\mathrm{GL}_2(F)$ and defines a nonsplit torus T' of GL_2. The image of T' in $G = \mathrm{GL}_2 / Z$ defines a nonsplit torus T of G.

A function $f \in C^0(G_F \setminus G_{\mathbb{A}})^K$ is *E-toroidal* if for all $\Phi \in \mathcal{H}_K$,

$$\int_{T_F \setminus T_{\mathbb{A}}} \Phi(f)(t)\, dt = 0.$$

We denote the space of all E-toroidal functions $f \in C^0(G_F \setminus G_{\mathbb{A}})^K$ by $\mathcal{V}_{\mathrm{tor}}$. Note that in [Lorscheid 2010] one finds a toroidal condition that is stronger than E-toroidality. Namely, f has to be E'-toroidal for all separable quadratic algebra extensions E' of F. We forgo recalling complete definitions, but remark that the finite-dimensionality of the space of all toroidal $f \in C^0(G_F \setminus G_{\mathbb{A}})^K$ follows since it is a subspace of $\mathcal{V}_{\mathrm{tor}}$.

Let $p : X' \to X$ be the map of curves that corresponds to the field extension E/F.

Theorem 10.8. *Let $c_T = \mathrm{vol}(T_F \setminus T_{\mathbb{A}}) / \#\big(\operatorname{Pic} X' / p^*(\operatorname{Pic} X)\big)$. Then,*

$$\int_{T_F \setminus T_{\mathbb{A}}} f(t)\, dt = c_T \cdot \sum_{[\mathcal{L}] \in \operatorname{Pic} X'/p^*(\operatorname{Pic} X)} f([p_* \mathcal{L}]) \quad \textit{for all } f \in C^0(G_F \setminus G_{\mathbb{A}})^K.$$

Proof. Let $\mathbb{A}_E$ be the adèles of E. To avoid confusion, we write $\mathbb{A}_F$ for $\mathbb{A}$. We introduce the following notation. For an $x \in \|X\|$ that is inert in E/F, we define $\mathcal{O}_{E,x} := \mathcal{O}_{E,y}$, where y is the unique place that lies over x. For an $x \in \|X\|$ that is split in E/F, we define $\mathcal{O}_{E,x} := \mathcal{O}_{E,y_1} \oplus \mathcal{O}_{E,y_2}$, where y_1 and y_2 are the two places that lie over x. Note that there is no place that ramifies. Let $\mathcal{O}_{E_x}$ denote the completion of $\mathcal{O}_{E,x}$. Then $\mathcal{O}_{E_x}$ is a free module of rank 2 over $\mathcal{O}_{F_x} = \mathcal{O}_x$ for every $x \in \|X\|$.

Let $\Theta_E : \mathbb{A}_E^\times \to \mathrm{GL}_2(\mathbb{A}_F)$ be the base extension of the embedding $E^\times \to \mathrm{GL}_2(F)$ that defines T', which corresponds to the chosen basis of E over F that is contained in $\mathbb{F}_{q^2}$. This basis is also a basis of $\mathcal{O}_{E_x}$ over $\mathcal{O}_{F_x}$ for every $x \in \|X\|$. This shows that $\Theta_E^{-1}(\mathrm{GL}_2(\mathcal{O}_{\mathbb{A}_F})) = \mathcal{O}_{\mathbb{A}_E}^\times$ and that the diagram

$$\begin{array}{ccc} E^\times \backslash \mathbb{A}_E^\times / \mathcal{O}_{\mathbb{A}_E}^\times & \xrightarrow{1:1} & \mathrm{Pic}\, X' \\ \downarrow{\scriptstyle \Theta_E} & & \downarrow{\scriptstyle p_*} \\ \mathrm{GL}_2(F) \backslash \mathrm{GL}_2(\mathbb{A}_F) / \mathrm{GL}_2(\mathcal{O}_{\mathbb{A}_F}) & \xrightarrow{1:1} & \mathrm{Bun}_2\, X \end{array}$$

commutes, where the horizontal arrows are the bijections defined in Section 5.3.

The action of $\mathbb{A}_F$ on $E^\times \backslash \mathbb{A}_E^\times / \mathcal{O}_{\mathbb{A}_E}^\times$ and $\mathrm{GL}_2(F) \backslash \mathrm{GL}_2(\mathbb{A}_F) / \mathrm{GL}_2(\mathcal{O}_{\mathbb{A}_F})$ by scalar multiplication is compatible with the action of $\mathrm{Pic}\, X$ on $\mathrm{Pic}\, X'$ and $\mathrm{Bun}_2\, X$ by tensoring in the sense that all maps in the diagram above are equivariant if we identify $\mathrm{Pic}\, X$ with $F^\times \backslash \mathbb{A}_F^\times / \mathcal{O}_{\mathbb{A}_F}^\times$. Taking orbits under these compatible actions yields the commutative diagram

$$\begin{array}{ccc} E^\times \mathbb{A}_F^\times \backslash \mathbb{A}_E^\times / \mathcal{O}_{\mathbb{A}_E}^\times & \xrightarrow{1:1} & \mathrm{Pic}\, X' / p^* \,\mathrm{Pic}\, X \\ \downarrow{\scriptstyle \Theta_E} & & \downarrow{\scriptstyle p_*} \\ G_F \backslash G_{\mathbb{A}_F} / K & \xrightarrow{1:1} & \mathbb{P}\mathrm{Bun}_2 X. \end{array}$$

Since f is right K-invariant, we may take the quotient of the domain of integration by $T_{\mathbb{A}_F} \cap K$ from the right, which is the image of $\mathcal{O}_{\mathbb{A}_E}^\times$ in $G_{\mathbb{A}_F}$. We obtain the assertion of the theorem for some still undetermined value of c. The value of c is computed by plugging in a constant function for f. □

Theorem 10.9. *The space of unramified toroidal functions has finite dimension, bounded by*

$$\dim \mathcal{V}_{\mathrm{tor}} \le \#(\mathbb{P}\mathrm{Bun}_2\, X - \{c_D\}_{[D] \in \mathrm{Cl}^{\mathrm{eff}} X}).$$

Proof. Given the inequality in the theorem, finite-dimensionality follows since the right-hand set is finite. Indeed, by Lemma 10.2,

$$\mathbb{P}\mathrm{Bun}_2\, X - \{c_D\}_{[D] \in \mathrm{Cl}^{\mathrm{eff}} X} \subset \{v \in \mathbb{P}\mathrm{Bun}_2\, X \mid \delta(v) \le m_X\}$$

since $m_X \geq g_X - 1$, and the latter set is finite.

We now proceed with the proof of the inequality. Let $f \in \mathcal{V}_{\text{tor}}$. We will show by induction on $d = \deg D$ that the value of f at a vertex c_D with $[D] \in \mathrm{Cl}^{\text{eff}} X$ is uniquely determined by the values of f at the elements of $\mathbb{P}\mathrm{Bun}_2 X - \{c_D\}_{[D]\in \mathrm{Cl}^{\text{eff}} X}$. This will prove the theorem.

By Theorem 10.8, the condition for f to lie in $\mathcal{V}_{\text{tor}}$ reads

$$\sum_{[\mathcal{L}]\in(\mathrm{Pic}\, X'/p^*\,\mathrm{Pic}\, X)} \Phi(f)([p_*\mathcal{L}]) = 0 \quad \text{for all } \Phi \in \mathcal{H}.$$

If $d = 0$, take Φ as the identity element in $\mathcal{H}_K$. We know from Proposition 6.4 that $p_*(\mathrm{Pic}\, X'/p^*\,\mathrm{Pic}\, X) = \mathbb{P}\mathrm{Bun}_2^{\text{tr}} X \cup \{c_0\}$, so $f(c_0)$ equals a linear combination of values of f at vertices v in $\mathbb{P}\mathrm{Bun}_2^{\text{tr}} X$, which all satisfy $\delta(v) < 0$ by Proposition 7.8. Since the zero divisor class is the only class in $\mathrm{Cl}^{\text{eff}} X$ of degree 0, we have proven the case $d = 0$.

Next, let D be an effective divisor of degree $d > 0$ and put $\Phi = \Phi_D$. If v is a Φ_D-neighbor of w, then $\delta(v)$ and $\delta(w)$ can differ at most by d (Lemma 10.4(i)). Therefore all Φ_D-neighbors v of vertices in $\mathbb{P}\mathrm{Bun}_2^{\text{tr}} X$ have $\delta(v) < d$. The vertex c_D is the only Φ_D-neighbor v of c_0 with $\delta(v) = d$ (as already seen in the proof of Theorem 10.6). Thus

$$0 = \sum_{\mathcal{L}\in(\mathrm{Pic}\, X'/p^*\,\mathrm{Pic}\, X)} \Phi_D(f)([p_*\mathcal{L}]) = (q+1)f(c_D) + \sum_{\substack{\mathcal{L}\in(\mathrm{Pic}\, X'/p^*\,\mathrm{Pic}\, X),\\ ([p_*\mathcal{L}],v,\lambda)\in\mathcal{U}_D([p_*\mathcal{L}]),\\ \delta(v)<d}} \lambda f(v)$$

determines $f(c_D)$ as the linear combination of values of f at vertices v satisfying $\delta(v) < d$. By the inductive hypothesis, $f(c_D)$ is already determined by the values of f at vertices that are not contained in $\{c_D\}_{[D]\in \mathrm{Cl}^{\text{eff}} X}$. □

Example 10.10. If X is the projective line over $\mathbb{F}_q$, then all vertices v are of the form c_D for some effective divisor D (see Example 8.10). Thus $\mathcal{V}_{\text{tor}}$ is trivial. Since only $v = c_0$ satisfies $\delta(v) \leq m_X$, all values of $f \in \mathcal{V}_0$ are multiples of $f(c_0)$. However, $\mathrm{Ext}^1(\mathcal{O}_X, \mathcal{O}_X)$ is trivial, thus the cuspidal condition (applied to the trivial Hecke operator) is $f(c_0) = 0$. Thus also $\mathcal{V}_0$ is trivial. See [Lorscheid 2012] for the corresponding spaces in the case of an elliptic curve.

Appendix: Examples for rational function fields

We give examples of graphs of Hecke operators for a rational function field, which can be calculated by elementary matrix manipulations. We do not show all calculations, but hint on how to do them. The reader will find examples for elliptic function fields that are determined by geometric methods in [Lorscheid 2012].

Let F be $\mathbb{F}_q(T)$, the function field of the projective line over $\mathbb{F}_q$, which has $q+1$ $\mathbb{F}_q$-rational points and trivial class group. Fix a place x of degree 1.

A.1. Using strong approximation for SL_2 (see Proposition 3.8, where J is trivial in this case), we get a bijection by adding the identity matrix e at all places $y \neq x$:

$$\Gamma \setminus G_x / K_x \to G_F \setminus G_{\mathbb{A}} / K, \quad [g_x] \mapsto [(g_x, e)].$$

We introduce some notation. Elements of $\mathcal{O}_F^x = \mathcal{O}^x \cap F$ can be written in the form $\sum_{i=m}^{0} b_i \pi_x^i$ with $b_i \in \mathbb{F}_q$ for $i = m, \dots, 0$ for some integer $m \leq 0$. Let $\tilde{K}_x = \mathrm{GL}_2(\mathcal{O}_x)$, where we view $\mathcal{O}_x$ as the collection of all power series $\sum_{i \geq 0} b_i \pi_x^i$ with $b_i \in \mathbb{F}_q$ for $i \geq 0$. Let $\Gamma = \mathrm{GL}_2(\mathcal{O}_F^x)$ and let Z be the center of GL_2.

A.2. For better readability, we write π for the uniformizer π_x at x and g for a matrix in G_x. We say $g \sim g'$ if they represent the same class $[g] = [g']$ in $\Gamma \setminus G_x / K_x$, and indicate by subscripts to "$\sim$" how to alter one representative to another. The following changes of the representative g of a class $[g] \in \Gamma \setminus G_x / K_x$ provide an algorithm to determine a standard representative for the class of any matrix $g \in G_x$:

(i) By the Iwasawa decomposition, every class in $\Gamma \setminus G_x / K_x$ is represented by an upper triangular matrix, and

$$\begin{pmatrix} a & b \\ & d \end{pmatrix} \underset{/Z_x}{\sim} \begin{pmatrix} a & b \\ & d \end{pmatrix} \begin{pmatrix} d^{-1} & \\ & d^{-1} \end{pmatrix} = \begin{pmatrix} a/d & b/d \\ & 1 \end{pmatrix}.$$

(ii) Write $a/d = r\pi^n$ for some integer n and $r \in \mathcal{O}_x^\times$, then with $b' = b/d$, we have

$$\begin{pmatrix} r\pi^n & b' \\ & 1 \end{pmatrix} \underset{/\tilde{K}_x}{\sim} \begin{pmatrix} r\pi^n & b' \\ & 1 \end{pmatrix} \begin{pmatrix} r^{-1} & \\ & 1 \end{pmatrix} = \begin{pmatrix} \pi^n & b' \\ & 1 \end{pmatrix}.$$

(iii) If $b' = \sum_{i \geq m} b_i \pi^i$ for some integer m and coefficients $b_i \in \mathbb{F}_q$ for $i \geq m$, then

$$\begin{pmatrix} \pi^n & \sum_{i \geq m} b_i \pi^i \\ & 1 \end{pmatrix} \underset{/\tilde{K}_x}{\sim} \begin{pmatrix} \pi^n & \sum_{i \geq m} b_i \pi^i \\ & 1 \end{pmatrix} \begin{pmatrix} 1 & -\pi^{-n}(\sum_{i \geq n} b_i \pi^i) \\ & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \pi^n & b_m\pi + \dots + b_{n-1}\pi^{n-1} \\ & 1 \end{pmatrix}.$$

(iv) One can further perform the following step:

$$\begin{pmatrix} \pi^n & b_m\pi^m + \dots + b_{n-1}\pi^{n-1} \\ & 1 \end{pmatrix}$$
$$\underset{\Gamma\setminus}{\sim} \begin{pmatrix} 1 & -(b_m\pi^m + \dots + b_0\pi^0) \\ & 1 \end{pmatrix} \begin{pmatrix} \pi^n & b_m\pi^m + \dots + b_{n-1}\pi^{n-1} \\ & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \pi^n & b_1\pi + \dots + b_{n-1}\pi^{n-1} \\ & 1 \end{pmatrix}.$$

(v) If $b = b_1\pi + \cdots + b_{n-1}\pi^{n-1} \neq 0$, then $b = s\pi^k$ with $1 \leq k \leq n-1,\ s \in O_x^\times$ and

$$\begin{pmatrix} \pi^n & s\pi^k \\ & 1 \end{pmatrix} \underset{\Gamma\backslash/Z_x\tilde{K}_x}{\sim} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} \pi^n & s\pi^k \\ & 1 \end{pmatrix} \begin{pmatrix} s^{-1}\pi^{-k} & \\ & s^{-1}\pi^{-k} \end{pmatrix} \begin{pmatrix} & -s^2 \\ s\pi^{n-k} & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \pi^{n-2k} & s^{-1}\pi^{-k} \\ & 1 \end{pmatrix}.$$

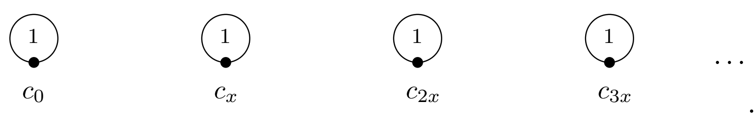
(vi) The last trick is

$$\begin{pmatrix} \pi^n & \\ & 1 \end{pmatrix} \underset{\Gamma\backslash/Z_x\tilde{K}_x}{\sim} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} \pi^n & \\ & 1 \end{pmatrix} \begin{pmatrix} \pi^{-n} & \\ & \pi^{-n} \end{pmatrix} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} \pi^{-n} & \\ & 1 \end{pmatrix}.$$

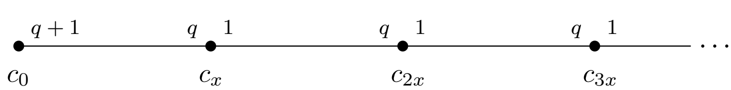
Executing these steps (possibly (iii)–(v) several times) will finally lead to a matrix of the form $p_n = \operatorname{diag}(\pi^{-n}, 1)$ for some $n \geq 0$. The matrix p_n represents the vertex c_{nx} in $\operatorname{Vert}\mathcal{G}_{\Phi,K} = \{c_{nx}\}_{n\geq 0}$ where Φ is any unramified Hecke operator (see Example 8.10). Thus we found a way to determine the vertex c_{nx} represented by an arbitrary matrix $g \in G_x \subset G_{\mathbb{A}}$.

Example A.3 (graph of 0 and 1). According to Section 1.7, the graph for the zero element 0 in $\mathcal{H}_K$ is

and the graph for the identity 1 in $\mathcal{H}_K$ is



Example A.4 (graph of Φ_x). By Proposition 2.3, the Φ_x-neighbors of p_i are of the form $p_i\xi_w$. With help of the reduction steps (i)–(vi) in A.2 one can determine easily the standard representative p_j of $p_i\xi_w$. We reobtain the graph of Φ_x as illustrated below (compare with Example 8.10).



Example A.5 (graph of Φ_y for $y \neq x$). If we want to determine the edges of $\mathcal{G}_y$ for a place y of degree d that differs from x, we have to find the standard representative p_j for each of the elements $p_i\xi_w$ where $w \in \mathbb{P}^1(\kappa_y)$, that is, ξ_w is an element of the form

$$\begin{pmatrix} \pi_y & b \\ & 1 \end{pmatrix} \quad \text{with } b \in \kappa_y, \qquad \text{or} \quad \begin{pmatrix} 1 & \\ & \pi_y \end{pmatrix}.$$

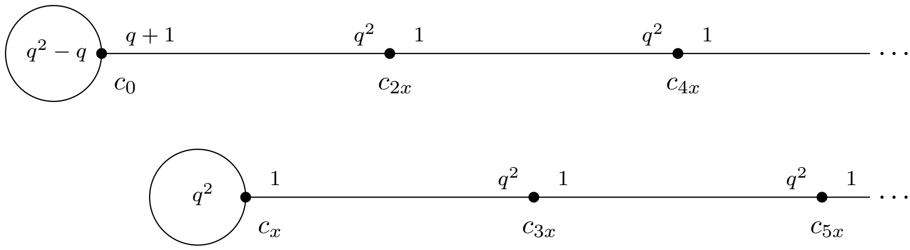


Figure A1. The graph of Φ_y for a place y of degree 2.

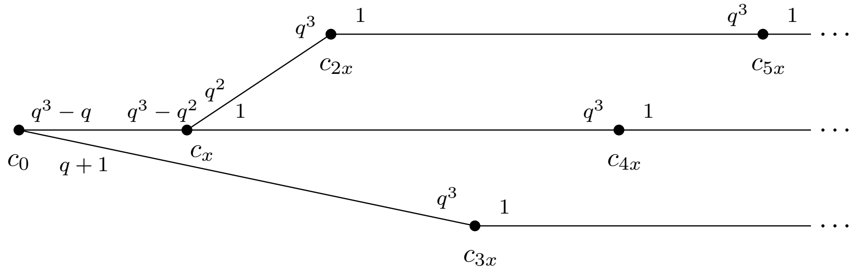


Figure A2. The graph of Φ_y for a place y of degree 3.

Since the class number of F is 1, the strong approximation property yields $G_F K^x = G^x_{\mathbb{A}}$ (see Proposition 3.8). This means that we find elements $\gamma \in G_F$ and $k \in K$ such that for all $z \neq x$, the adelic matrices ξ_w and γk have equal z-components $(\xi_w)_z = (\gamma k)_z$. Therefore, the only nontrivial component of the adelic matrix

$$\theta_w = \gamma^{-1} \xi_w k^{-1}$$

is its x-component. By an appropriate choice of k_x, we can normalize the x-component of θ_w to be equal to one of the matrices

$$\begin{pmatrix} \pi_x^d & b_0 + \cdots + b_{d-1}\pi_x^{d-1} \\ & 1 \end{pmatrix} \quad \text{with } b_i \in \kappa_x \text{ for } i = 0, \ldots, d-1, \quad \text{and} \quad \begin{pmatrix} 1 & \\ & \pi_x^d \end{pmatrix},$$

and for the different choices of $w \in \mathbb{P}^1(\kappa_y)$, each of these matrices occurs as the x-component of a (unique) θ_w. The reduction steps (i)–(vi) of A.2 tell us which classes p_j are represented by the matrices $\theta_w p_i = \gamma^{-1} p_i \xi_w k^{-1}$, and we are able to determine the edges similarly to the previous example. Thus we obtain that $\mathcal{G}_y$ only depends on the degree of y. Note that if y is of degree 1, then $\mathcal{G}_y$ equals $\mathcal{G}_x$. Figures A1, A2, A5, and A6 show the graphs for degrees 2, 3, 4 and 5, respectively.

Example A.6 (the graph of powers of Φ_x). It is interesting to compare the graph of Φ_y with $\deg y = d$ to the graph of Φ_x^d. The latter graph is easily deduced from $\mathcal{G}_x$ by

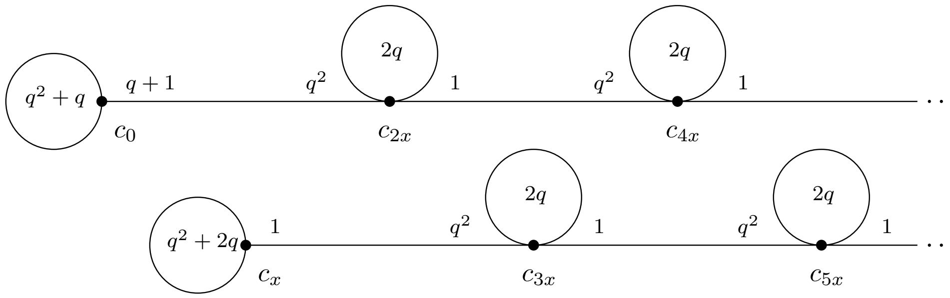


Figure A3. The graph of Φ_x^2.

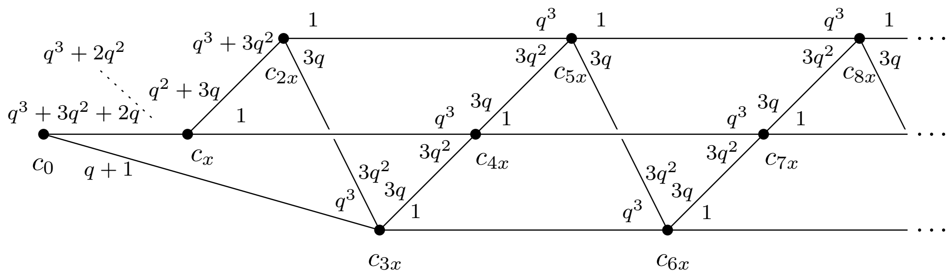


Figure A4. The graph of Φ_x^3.

means of Section 1.7. Namely, a vertex v' is a Φ_x^d-neighbor of a vertex v in $\mathcal{G}_{\Phi_x^d,K}$ if there is a path of length d from v to v' in $\mathcal{G}_x$, that is, a sequence $(v_0, v_1, \dots, v_d)$ of vertices in $\mathcal{G}_x$ with $v_0 = v$ and $v_d = v'$ such that for all $i = 1, \dots, d$, there is an edge (v_{i-1}, v_i, m_i) in $\mathcal{G}_x$. The weight of an edge from v to v' in the graph of $\mathcal{G}_x^d$ is obtained by taking the sum of the products $m_1 \cdots m_d$ over all paths of length d from v to v' in $\mathcal{G}_x$.

Figures A3 and A4 show the graphs of Φ_x^2 and Φ_x^3, respectively, and we see that for $\deg y = 2$, we have $\Phi_x^2 \equiv \Phi_y + 2q \cdot 1 \pmod{\mathcal{J}(K)}$ and for $\deg y = 3$, we have $\Phi_x^3 \equiv \Phi_y + 3q \cdot \Phi_x \pmod{\mathcal{J}(K)}$, where $\mathcal{J}(K)$ is the ideal of $\mathcal{H}_K$ of Hecke operators that operate trivially on $C^0(G_F \setminus G_{\mathbb{A}})$.

Example A.7 (the graphs of two ramified Hecke operators). It is also possible to determine examples for Hecke operators in $\mathcal{H}_{K'}$ by elementary matrix manipulations, when $K' < K$ is a subgroup of finite index. We will show two examples, which are illustrated in Figures A7 and A8. We omit the calculation, but only point out why the crucial differences between the two graphs occur.

For $K' = \left\{k \in K \;\middle|\; k_x \equiv \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \pmod{\pi_x}\right\}$, the fibers of the projection

$$P : G_F \setminus G_{\mathbb{A}}/K' \to G_F \setminus G_{\mathbb{A}}/K$$

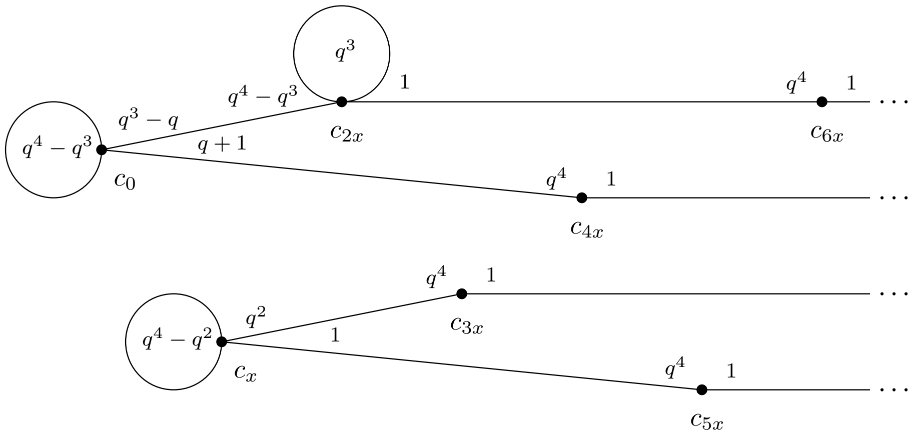


Figure A5. The graph of Φ_y for a place y of degree 4.

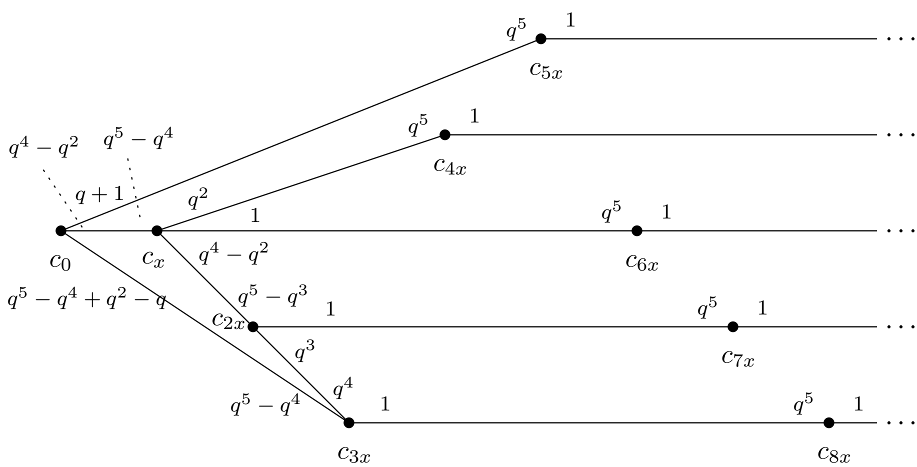


Figure A6. The graph of Φ_y for a place y of degree 5.

are given by $P^{-1}(c_0)=\{[p_0]\}$ and for positive n, by $P^{-1}(c_{nx})=\{[p_{nx}\vartheta_w]\}_{w\in\mathbb{P}^1(\kappa_x)}$ with $\vartheta_{[1:c]}=\left(\begin{smallmatrix}1&c\\&1\end{smallmatrix}\right)$ and $\vartheta_{[0:1]}=\left(\begin{smallmatrix}&1\\1&\end{smallmatrix}\right)$. The union of these fibers equals the set of vertices of a Hecke operator in $\mathcal{H}_{K'}$. We shall denote the vertices by $c'_0=[p_0]$ and $c'_{nx,w}=[p_{nx}\vartheta_w]$ for $n\geq 1$ and $w\in\mathbb{P}^1(\kappa_x)$. Note that $G_{\mathbb{F}_q}=G_{\kappa_x}$ acts on $\mathbb{P}^1(\kappa_x)$ from the right, so if $\gamma\in G_{\mathbb{F}_q}$, then $w\mapsto w\gamma$ permutes the elements of $\mathbb{P}^1(\kappa_x)$.

The first Hecke operator $\Phi'_{y,\gamma}\in\mathcal{H}_{K'}$ that we consider is $(\operatorname{vol}K/\operatorname{vol}K')$ times the characteristic function of $K'\left(\begin{smallmatrix}\pi_y&\\&1\end{smallmatrix}\right)\gamma K'$, where y is a degree one place different from x and $\gamma\in G_{\mathbb{A}}$ is a matrix whose only nontrivial component is $\gamma_x\in G_{\mathbb{F}_q}$. (The factor $(\operatorname{vol}K/\operatorname{vol}K')$ is included to obtain integer weights.) Since $K'\left(\begin{smallmatrix}\pi_y&\\&1\end{smallmatrix}\right)\gamma K'$ is

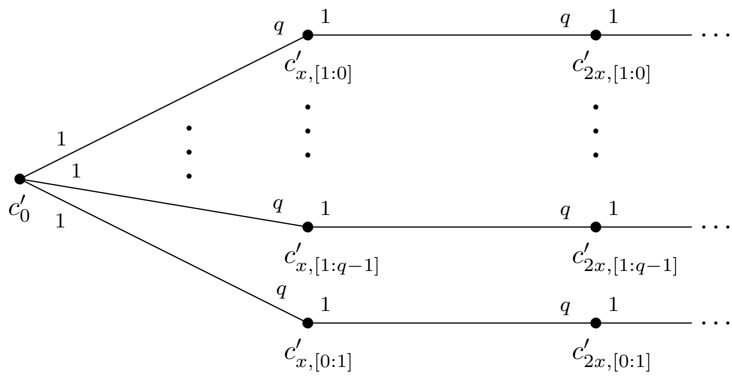


Figure A7. Graph of $\Phi'_{y,e}$ as defined in Example A.7.

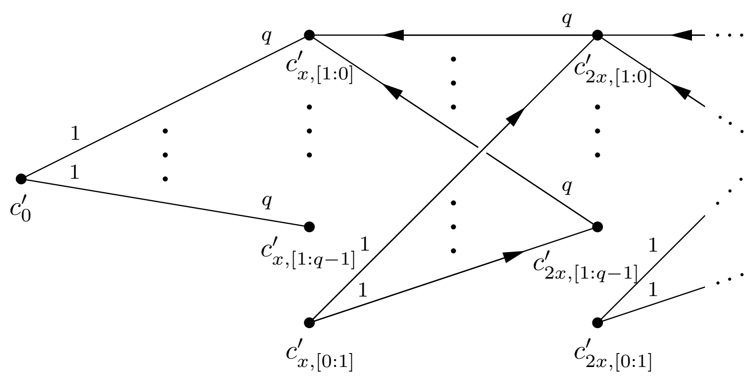


Figure A8. Graph of Φ'_x as defined in Example A.7.

contained in $K\begin{pmatrix}\pi_y & \\ & 1\end{pmatrix}\gamma K$, the graph of $\Phi'_{y,\gamma}$ relative to K' can have an edge from v to w only if $\mathcal{G}_y$ has an edge from $P(v)$ to $P(w)$. Because $K'_y = K_y$, we argue as for K that $K'\begin{pmatrix}\pi_y & \\ & 1\end{pmatrix}\gamma K' = \bigsqcup_{w\in\mathbb{P}^1(\kappa_y)} \xi_w \gamma K'$. Applying the same methods as in Example A.5, one obtains that

$$\mathcal{U}_{\Phi'_{y,\gamma},K'}(c'_0) \;=\; \{(c'_0, c'_{x,w}, 1)\}_{w\in\mathbb{P}^1(\kappa_x)}$$

and for every $n \geq 1$ and $w \in \mathbb{P}^1(\kappa_x)$ that

$$\mathcal{U}_{\Phi'_{y,\gamma},K'}(c'_{nx,w}) \;=\; \{(c'_{nx,w}, c'_{(n+1)x,w\gamma}, 1), (c'_{nx,w}, c'_{(n-1)x,w\gamma}, q)\}.$$

For the case that γ is equal to the identity matrix e, the graph is illustrated in Figure A7. Note that for general γ, an edge does not necessarily have an inverse edge since $w\gamma^2$ does not have to equal w.

The second Hecke operator $\Phi'_x \in \mathcal{H}_{K'}$ is $(\mathrm{vol}\, K / \mathrm{vol}\, K')$ times the characteristic function of $K'\left(\begin{smallmatrix}\pi_x & \\ & 1\end{smallmatrix}\right)K'$. This case behaves differently, since K'_x and K_x are not equal; in particular, we have $K'\left(\begin{smallmatrix}\pi_x & \\ & 1\end{smallmatrix}\right)K' = \coprod_{b\in\kappa_x} \left(\begin{smallmatrix}\pi_x & b\pi_x \\ & 1\end{smallmatrix}\right)K'$. This allows us to compute the edges as illustrated in Figure A8. Note that for $n \geq 1$, the vertices of the form $c'_{nx,[1:0]}$ and $c'_{nx,[0:1]}$ behave particularly.

Acknowledgements

This paper is extracted from my thesis [Lorscheid 2008]. First of all, I would like to thank Gunther Cornelissen for his advice during my graduate studies. I would like to thank Frits Beukers and Roelof Bruggeman for their numerous comments on a lecture series about my studies. I would like to thank the referee for his or her rich and detailed comments on the paper.

References

[Arason et al. 1992] J. K. Arason, R. Elman, and B. Jacob, "On indecomposable vector bundles", *Comm. Algebra* **20**:5 (1992), 1323–1351. MR 93e:14051 Zbl 0769.14004

[Atiyah 1956] M. Atiyah, "On the Krull–Schmidt theorem with application to sheaves", *Bull. Soc. Math. France* **84** (1956), 307–317. MR 19,172b Zbl 0072.18101

[Bourbaki 1965] N. Bourbaki, *Éléments de mathématique, Algèbre commutative, Chapitre 7: Diviseurs*, Actualités Scientifiques et Industrielles **1314**, Hermann, Paris, 1965. MR 41 #5339 Zbl 0141.03501

[Frenkel 2004] E. Frenkel, "Recent advances in the Langlands program", *Bull. Amer. Math. Soc.* (*N.S.*) **41**:2 (2004), 151–184. MR 2005e:11147 Zbl 1070.11051

[Gaitsgory 2003] D. Gaitsgory, "Informal introduction to geometric Langlands", pp. 269–281 in *An introduction to the Langlands program* (Jerusalem, 2001), edited by J. Bernstein and S. Gelbart, Birkhäuser, Boston, MA, 2003. MR 1990383 Zbl 1111.11308

[Gekeler 1995] E.-U. Gekeler, "Improper Eisenstein series on Bruhat–Tits trees", *Manuscripta Math.* **86**:3 (1995), 367–391. MR 95m:11043 Zbl 0884.11025

[Gekeler 1997] E.-U. Gekeler, "On the Drinfeld discriminant function", *Compositio Math.* **106**:2 (1997), 181–202. MR 98e:11071 Zbl 0930.11031

[Gekeler and Nonnengardt 1995] E.-U. Gekeler and U. Nonnengardt, "Fundamental domains of some arithmetic groups over function fields", *Internat. J. Math.* **6**:5 (1995), 689–708. MR 96i:11043 Zbl 0858.11025

[Gelbart 1975] S. S. Gelbart, *Automorphic forms on adèle groups*, Annals of Mathematics Studies **83**, Princeton University Press, 1975. MR 52 #280 Zbl 0329.10018

[Harder and Narasimhan 1974/75] G. Harder and M. S. Narasimhan, "On the cohomology groups of moduli spaces of vector bundles on curves", *Math. Ann.* **212** (1974/75), 215–248. MR 51 #509 Zbl 0324.14006

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Kneser 1966] M. Kneser, "Strong approximation", pp. 187–196 in *Algebraic Groups and Discontinuous Subgroups* (Boulder, CO, 1965), edited by A. Borel and G. D. Mostow, Amer. Math. Soc., Providence, R.I., 1966. MR 35 #4225 Zbl 0201.37904

[Lang 1956] S. Lang, "Algebraic groups over finite fields", *Amer. J. Math.* **78** (1956), 555–563. MR 19,174a Zbl 0073.37901

[Laumon 1997] G. Laumon, *Cohomology of Drinfeld modular varieties, II: Automorphic forms, trace formulas and Langlands correspondence*, Cambridge Studies in Advanced Mathematics **56**, Cambridge University Press, 1997. MR 98c:11045b Zbl 0870.14016

[Lorscheid 2008] O. Lorscheid, *Toroidal automorphic forms for function fields*, PhD thesis, University of Utrecht, 2008, available at http://igitur-archive.library.uu.nl.

[Lorscheid 2010] O. Lorscheid, "Toroidal automorphic forms for function fields", preprint, 2010. To appear in *Israel J. Math.* arXiv 1012.3223

[Lorscheid 2012] O. Lorscheid, "Automorphic forms for elliptic function fields", *Math. Z.* **272**:3-4 (2012), 885–911. MR 2995144

[Margulis 1977] G. A. Margulis, "Cobounded subgroups in algebraic groups over local fields", *Funkcional. Anal. i Priložen.* **11**:2 (1977), 45–57, 95. In Russian; translated in *Functional Anal. Appl.*, **11**:2 (1977), 119–122. MR 56 #495

[Moore 1968] C. C. Moore, "Group extensions of p-adic and adelic linear groups", *Inst. Hautes Études Sci. Publ. Math.* 35 (1968), 157–222. MR 39 #5575 Zbl 0159.03203

[Moreno 1985] C. J. Moreno, "Analytic proof of the strong multiplicity one theorem", *Amer. J. Math.* **107**:1 (1985), 163–206. MR 86m:22027 Zbl 0564.10035

[Prasad 1977] G. Prasad, "Strong approximation for semi-simple groups over function fields", *Ann. of Math.* (2) **105**:3 (1977), 553–572. MR 56 #2921 Zbl 0348.22006

[van der Put and Reversat 1997] M. van der Put and M. Reversat, "Automorphic forms and Drinfeld's reciprocity law", pp. 188–223 in *Drinfeld modules, modular schemes and applications* (Alden-Biesen, 1996), edited by E.-U. Gekeler et al., World Sci. Publ., River Edge, NJ, 1997. MR 99j:11059 Zbl 0924.11051

[Schleich 1974] T. Schleich, *Einige Bemerkungen zur Spektralzerlegung der Hecke-Algebra für die* PGL_2 *über Funktionenkörpern*, Bonner Mathematische Schriften **71**, Mathematisches Institut, Universität Bonn, Bonn, 1974. MR 56 #8759 Zbl 0345.10013

[Serre 2003] J.-P. Serre, *Trees*, Springer, Berlin, 2003. MR 2003m:20032 Zbl 1013.20001

Communicated by Edward Frenkel
Received 2011-04-11 Revised 2012-01-25 Accepted 2012-02-22

lorschei@impa.br *Instituto Nacional de Matemática Pura e Aplicada, IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, RJ, Brazil*
http://w3.impa.br/~lorschei/

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.63

Group actions of prime order on local normal rings

Franz Kiràly and Werner Lütkebohmert

Let B be a Noetherian normal local ring and $G \subset \operatorname{Aut}(B)$ be a cyclic group of local automorphisms of prime order. Let A be the subring of G-invariants of B and assume that A is Noetherian. We prove that B is a monogenous A-algebra if and only if the augmentation ideal of B is principal. If in particular B is regular, we prove that A is regular if the augmentation ideal of B is principal.

An important class of singularities is built by the famous Hirzebruch–Jung singularities. They arise by dividing out a finite cyclic group action on a smooth surface. Their resolution is well understood and has nice arithmetic properties related to continued fractions; see [Hirzebruch 1953; Jung 1908].

One can also look at such group actions from a purely algebraic point of view. So let B be a regular local ring and G a finite cyclic group of order n acting faithfully on B by local automorphisms. In the tame case, that is, the order of G is prime to the characteristic of the residue field k of B, there is a central result of J. P. Serre [1968] saying that the action is given by multiplying a suitable system of parameters $(y_1, \dots, y_d)$ by roots of unity $y_i \mapsto \zeta^{n_i} \cdot y_i$ for $i = 1, \dots, d$, where ζ is a primitive n-th root of unity. Moreover, the ring of invariants $A := B^G$ is regular if and only if $n_i \equiv 0 \bmod n$ for $d-1$ of the parameters. The latter is equivalent to the fact that $\operatorname{rk}((\sigma - \mathrm{id})|T) \le 1$ for the action of $\sigma \in G$ on the tangent space $T := \mathfrak{m}_B/\mathfrak{m}_B^2$. For more details see [Bourbaki 1981, Chapter 5, ex. 7].

Only very little is known in the case of a wild group action, that is, when $\gcd(n, \operatorname{char} k) > 1$. In this paper we will restrict ourselves to the case of p-cyclic group actions, that is, where $n = p$ is a prime number. We will present a sufficient condition for the ring of invariants A to be regular. Our result is also valid in the tame case, that is, where n is a prime different from $\operatorname{char} k$. As the method of Serre depends on an intrinsic formula for writing down the action explicitly, we provide also an explicit formula for presenting B as a free A-module if our condition is fulfilled.

MSC2010: primary 14L30; secondary 13A50.
Keywords: algebraic geometry, commutative algebra, group actions.

The interest in our problem arises from investigating the relationship between the regular and the stable R-model of a smooth projective curve X_K over the field of fractions K of a discrete valuation ring R. In general, the curve X_K admits a stable model X' over a finite Galois extension $R \hookrightarrow R'$. Then the Galois group $G = G(R'/R)$ acts on X'. Our result provides a means to construct a regular model over R by starting from the stable model X'. As a special case, we discuss in Section 4 the situation where X_K has good reduction after a Galois p-extension $R \hookrightarrow R'$. In this case there is a criterion for when the quotient of the smooth model is regular. We intend to work out more general situations in a further article.

1. The main result

In this paper we will study only local actions of a cyclic group G of prime order p on a normal local ring B. We fix a generator σ of G and obtain the *augmentation map*

$$I := I_\sigma := \sigma - \mathrm{id} : B \to B, \quad b \mapsto \sigma(b) - b.$$

We introduce the B-ideal

$$I_G := (I(b);\ b \in B) \subset B$$

which is generated by the image $I(B)$. This ideal is called *augmentation ideal*. If this ideal is generated by an element $I(y)$, we call y an *augmentation generator*. Note that this ideal does not depend on the chosen generator σ of G. Moreover, if y is an augmentation generator with respect to a generator σ of G, then y is also an augmentation generator for any other generator of G. Since B is local, the ideal I_G is generated by an augmentation generator if I_G is principal. Namely, $I_G/\mathfrak{m}_B I_G$ is a vector space over the residue field $k_B = B/\mathfrak{m}_B$ of B of dimension 1. So it is generated by the residue class of $I(y)$ for some $y \in B$, and hence, by Nakayama's lemma, I_G is generated by $I(y)$.

Definition 1. An action of a group G on a regular local ring B by local automorphisms is called a *pseudoreflection* if there exists a system of parameters $(y_1, \ldots, y_d)$ of B such that $y_2, \ldots, y_d$ are invariant under G.

Theorem 2. *Let B be a normal local ring with residue field $k_B := B/\mathfrak{m}_B$. Let p be a prime number and G a p-cyclic group of local automorphisms of B. Let I_G be the augmentation ideal. Let A be the ring of G-invariants of B. Consider the following conditions*:

(a) *$I_G := B \cdot I(B)$ is principal.*

(b) *B is a monogenous A-algebra.*

(c) *B is a free A-module.*

Then the following implications are true:

$$(a) \Longleftrightarrow (b) \Longrightarrow (c).$$

Assume, in addition, that B is regular. Consider the following conditions:

(d) *A is regular.*

(e) *G acts as a pseudoreflection.*

Then the condition (c) *is equivalent to* (d). *Moreover if, in addition, the canonical map $k_A \xrightarrow{\sim} k_B$ is an isomorphism, then condition* (a) *is equivalent to condition* (e).

We start the proof of the theorem with several preparations.

Remark 3. For $b_1, b_2, b \in B$, the following relations are true:

(i) $\quad I(b_1 \cdot b_2) = I(b_1) \cdot \sigma(b_2) + b_1 \cdot I(b_2).$

(ii) $$I(b^n) = \left(\sum_{i=1}^{n} \sigma(b)^{i-1} b^{n-i}\right) \cdot I(b).$$

(iii) $$I\left(\frac{b_1}{b_2}\right) = \frac{I(b_1)b_2 - b_1 I(b_2)}{b_2 \sigma(b_2)} \quad \text{if } b_2 \neq 0.$$

Proof. (i) follows by a direct calculation and (ii) by induction from (i).

As for (iii), the formula (i) holds for elements in the field of fractions as well. Therefore,

$$I(b_1) = I\left(\frac{b_1}{b_2} b_2\right) = I\left(\frac{b_1}{b_2}\right)\sigma(b_2) + \frac{b_1}{b_2} I(b_2),$$

and the formula follows. □

To prove that (a) implies (b) we need a technical lemma.

Lemma 4. *Let $y \in B$ be an augmentation generator. Then set, inductively*,

$$\begin{aligned} y_i^{(0)} &:= y^i && \text{for } i = 0, \dots, p-1, \\ y_i^{(1)} &:= I(y_i^{(0)})/I(y_1^{(0)}) && \text{for } i = 1, \dots, p-1, \\ y_i^{(n+1)} &:= I(y_i^{(n)})/I(y_{n+1}^{(n)}) && \text{for } i = n+1, \dots, p-1. \end{aligned}$$

Then

$$y_i^{(n)} = \sum_{0 \le k_1 \le \cdots \le k_{i-n} \le n} \ \prod_{j=1}^{i-n} \sigma^{k_j}(y) \quad \text{for } i = n, \dots, p-1,$$

and in particular,

$$y_n^{(n)} = 1, \qquad y_{n+1}^{(n)} = \sum_{j=1}^{n+1} \sigma^{j-1}(y), \qquad I(y_{n+1}^{(n)}) = \sigma^{n+1}(y) - y.$$

Furthermore, $y_{n+1}^{(n)}$ is again an augmentation generator for $n = 0, \dots, p-2$.

Proof. We proceed by induction on n. For $n = 0$ the formulas are obviously correct. For the convenience of the reader we also display the formulas for $n = 1$. Due to Remark 3 one has

$$y_i^{(1)} = \frac{I(y_i^{(0)})}{I(y_1^{(0)})} = \frac{I(y^i)}{I(y)} = \sum_{j=1}^{i} \sigma(y)^{j-1} y^{i-j} = \sum_{0 \le k_1 \le \cdots \le k_{i-1} \le 1} \prod_{\nu=1}^{i-1} \sigma^{k_\nu}(y),$$

since the last sum can be viewed as a sum over an index j where $i - j$ is the number of k_ν equal to 0. In particular, the formulas are correct for $y_1^{(1)}$ and $y_2^{(1)}$. Moreover

$$I(y_2^{(1)}) = I(\sigma(y) + y) = \sigma^2(y) - y.$$

Since σ^2 is generator of G for $2 < p$, the element $y_2^{(1)}$ is an augmentation generator as well.

Now assume that the formulas are correct for n. Since $y_{n+1}^{(n)}$ is an augmentation generator, $I(y_{n+1}^{(n)})$ divides $I(y_i^{(n)})$ for $i = n+1, \ldots, p-1$. Then it remains to show, upon substituting the expressions from the lemma for $y_i^{(n)}$ and $y_i^{(n+1)}$, that

$$I(y_i^{(n)}) = (\sigma^{n+1}(y) - y) \cdot y_i^{(n+1)} \quad \text{for } i = n+1, \ldots, p-1.$$

For the left hand side one computes

$$\begin{aligned} \text{LHS} &= I\left(\sum_{0 \le k_1 \le \cdots \le k_{i-n} \le n} \prod_{j=1}^{i-n} \sigma^{k_j}(y) \right) = \sum_{0 \le k_1 \le \cdots \le k_{i-n} \le n} I\left(\prod_{j=1}^{i-n} \sigma^{k_j}(y) \right) \\ &= \sum_{0 \le k_1 \le \cdots \le k_{i-n} \le n} \left(\prod_{j=1}^{i-n} \sigma^{k_j+1}(y) - \prod_{j=1}^{i-n} \sigma^{k_j}(y) \right) \\ &= \sum_{1 \le k_1 \le \cdots \le k_{i-n} \le n+1} \prod_{j=1}^{i-n} \sigma^{k_j}(y) - \sum_{0 \le k_1 \le \cdots \le k_{i-n} \le n} \prod_{j=1}^{i-n} \sigma^{k_j}(y). \end{aligned}$$

Now all terms occurring in both sums cancel. These are the terms with $k_{i-n} \le n$ in the first sum and $1 \le k_1$ in the second sum.
For the right hand side one computes

$$\begin{aligned} \text{RHS} &= (\sigma^{n+1}(y) - y) \cdot \sum_{0 \le k_1 \le \cdots \le k_{i-n-1} \le n+1} \prod_{j=1}^{i-n-1} \sigma^{k_j}(y) \\ &= \sum_{0 \le k_1 \le \cdots \le k_{i-n} = n+1} \prod_{j=1}^{i-n} \sigma^{k_j}(y) - \sum_{0 = k_1 \le \cdots \le k_{i-n} \le n+1} \prod_{j=1}^{i-n} \sigma^{k_j}(y). \end{aligned}$$

Both sides are seen to be equal. In particular we have

$$\begin{aligned} y_{n+1}^{(n+1)} &= 1, \\ y_{n+2}^{(n+1)} &= \sum_{0\le k_1\le n+1} \prod_{j=1}^{1} \sigma^{k_1}(y) = \sum_{j=1}^{n+2} \sigma^{j-1}(y), \\ I(y_{n+2}^{(n+1)}) &= \sigma^{n+2}(y) - y. \end{aligned}$$

So $y_{n+2}^{(n+1)}$ is an augmentation generator for $n+2 < p$, since σ^{n+2} generates G. This concludes the technical part. □

Proposition 5. *Assume that the augmentation ideal I_G is principal and let $y \in B$ be an augmentation generator. Then B decomposes into the direct sum*

$$B = A \cdot y^0 \oplus A \cdot y^1 \oplus \cdots \oplus A \cdot y^{p-1}.$$

Proof. Since $I(y) \neq 0$, the element y generates the field of fractions $Q(B)$ over $Q(A)$. Therefore

$$Q(B) = Q(A) \cdot y^0 \oplus Q(A) \cdot y^1 \oplus \cdots \oplus Q(A) \cdot y^{p-1}.$$

Then it suffices to show the following claim:

Let $a, a_0, \ldots, a_{p-1} \in A$. Assume that a divides

$$b = a_0 \cdot y^0 + a_1 \cdot y^1 + \cdots + a_{p-1} \cdot y^{p-1}.$$

Then a divides $a_0, a_1, \ldots, a_{p-1}$.

If $b = a \cdot \beta$, then $I(b) = a \cdot I(\beta)$. Since $I(\beta) = \beta_1 \cdot I(y)$, we get $I(b) = a\beta_1 \cdot I(y)$. So we see that a divides $I(b)/I(y) \in B$. Using the notation of Lemma 4, set

$$\begin{aligned} b^{(0)} &:= b = a_0 \cdot y^0 + a_1 \cdot y^1 + \cdots + a_{p-1} \cdot y^{p-1} \\ b^{(1)} &:= \frac{I(b^{(0)})}{I(y)} = a_1 + a_2 \frac{I(y^2)}{I(y)} + \cdots + a_{p-1} \frac{I(y^{p-1})}{I(y)} \\ &\quad = a_1 \cdot y_1^{(1)} + a_2 \cdot y_2^{(1)} + \cdots + a_{p-1} \cdot y_{p-1}^{(1)} \\ b^{(n)} &:= \frac{I(b^{(n-1)})}{I(y_n^{(n-1)})} = a_n \cdot y_n^{(n)} + a_{n+1} \cdot y_{n+1}^{(n)} + \cdots + a_{p-1} \cdot y_{p-1}^{(n)}. \end{aligned}$$

Due to the observation above, by induction a divides $b^{(0)}, b^{(1)}, \ldots, b^{(p-1)}$, since $y_{n+1}^{(n)}$ is an augmentation generator for $n = 1, \ldots, p-2$. So we obtain

$$a \mid b^{(p-1)} = a_{p-1} \cdot y_{p-1}^{(p-1)} = a_{p-1}.$$

Now proceeding downwards, one obtains

$$a \,\big|\, b^{(p-2)} = a_{p-2} + a_{p-1} \cdot y_{p-1}^{(p-2)}, \quad \text{hence } a \,\big|\, a_{p-2},$$
$$a \,\big|\, b^{(n)} = a_n + a_{n+1} \cdot y_{n+1}^{(n)} + \cdots + a_{p-1} \cdot y_{p-1}^{(n)}, \quad \text{hence } a \,\big|\, a_n$$

for $n = p-1, p-2, \ldots, 0$. □

Proof of the first part of Theorem 2. (a) $\Longrightarrow$ (b): This follows from Proposition 5.

(b) $\Longrightarrow$ (a): If $B = A[y]$ is monogenous, then $I_G = B \cdot I(y)$ is principal.

(b) $\Longrightarrow$ (c) is clear. Namely, if $B = A[y]$, the minimal polynomial of y over the field of fraction is of degree p and the coefficients of this polynomial belong to A. Then B has $y^0, y^1, \ldots, y^{p-1}$ as an A-basis. □

Next we do some preparations for proving the second part of the theorem where B is assumed to be regular.

Proposition 6. *Keep the assumption of the second part of Theorem 2, namely that B is regular and that the canonical morphism $k_A \xrightarrow{\sim} k_B$ is an isomorphism. Let $(y_1, \ldots, y_d)$ be a generating system of the maximal ideal $\mathfrak{m}_B$. Then the following assertions are true*:

(i) $I_G = B \cdot I(y_1) + \cdots + B \cdot I(y_d)$.

(ii) *If the ideal $I_G = B \cdot I(B)$ is principal, then there exists an index $i \in \{1, \ldots, d\}$ with $I_G = B \cdot I(y_i)$.*

Proof. (i) Recall that $A = B^G$ denotes the ring of invariants. Due to the assumption, we have $B = A + \mathfrak{m}_B$, and hence, $I(B) = I(\mathfrak{m}_B)$. Furthermore, we have

$$\mathfrak{m}_B = \mathfrak{m}_B^2 + \sum_{i=1}^{d} A \cdot y_i .$$

Since I is A-linear, we get

$$I(\mathfrak{m}_B) = I(\mathfrak{m}_B^2) + \sum_{i=1}^{d} A \cdot I(y_i).$$

Due to Remark 3, one knows $I(\mathfrak{m}_B^2) \subset \mathfrak{m}_B \cdot I(\mathfrak{m}_B)$. So, one obtains

$$I(\mathfrak{m}_B) \subset \mathfrak{m}_B \cdot I(\mathfrak{m}_B) + \sum_{i=1}^{d} B \cdot I(y_i).$$

Since B is local, Nakayama's lemma yields

$$I_G = B \cdot I(B) = B \cdot I(\mathfrak{m}_B) = \sum_{i=1}^{d} B \cdot I(y_i).$$

(ii) Since I_G is principal, $I_G/\mathfrak{m}_B I_G$ is generated by one of the $I(y_i)$, and hence, again by Nakayama's lemma, $I_G = B \cdot I(y_i)$ for a suitable $i \in \{1, \dots, d\}$. □

Proof of the second part of Theorem 2. (c) $\Longrightarrow$ (d) follows from [Matsumura 1980, Theorem 51]. Namely, B is noetherian due to the definition of a regular ring. Since $A \to B$ is faithfully flat, A is noetherian. Then one can apply [loc. cit.].

(d) $\Longrightarrow$ (c) follows from [Serre 1965, IV, Prop. 22].

(a) $\Longrightarrow$ (e): We assume that the canonical map $k_A \to k_B$ of the residue fields is an isomorphism. If I_G is principal, one can choose an augmentation generator $y \in \mathfrak{m}_B$ that is part of a system of parameters $(y, y_2, \dots, y_d)$ due to Proposition 6. Due to Proposition 5, we know that B decomposes into the direct sum

$$B = A \cdot y^0 \oplus A \cdot y^1 \oplus \cdots \oplus A \cdot y^{p-1}.$$

Now we can represent

$$y_j = \sum_{i=0}^{p-1} a_{i,j} \cdot y^i \quad \text{for } j = 2, \dots, d.$$

Then, set

$$\tilde{y}_j := y_j - \sum_{i=1}^{p-1} a_{i,j} y^i = a_{0,j} \in A \cap \mathfrak{m}_B = \mathfrak{m}_A \quad \text{for } j = 2, \dots, d.$$

So $(y, \tilde{y}_2, \dots, \tilde{y}_d)$ is a system of parameters of B as well. Thus G acts by a pseudoreflection.

(e) $\Longrightarrow$ (a): If G is a pseudoreflection, I_G is generated by $I(y)$ due to Proposition 6, where $y, x_2, \dots, x_p$ is a system of parameters with $x_i \in \mathfrak{m}_A$ for $i = 2, \dots, p$ if $k_A = k_B$. □

2. An example

If $k_A \to k_B$ is not an isomorphism, the implication (e) $\Longrightarrow$ (a) is false:

Example 7. Let k be a field of positive characteristic p and look at the polynomial ring $R := k[Z, Y, X_1, X_2]$ over k. We define a p-cyclic action of $G = \langle \sigma \rangle$ on R by

$$\sigma|k := \mathrm{id}_k, \quad \sigma(Z) = Z + X_1, \quad \sigma(Y) = Y + X_2, \quad \sigma(X_i) = X_i \quad \text{for } i = 1, 2.$$

This is a well-defined action of order p, since $p \cdot X_i = 0$ for $i = 1, 2$, and it leaves the ideal $\mathfrak{I} := (Y, X_1, X_2)$ invariant. Furthermore, for any $g \in k[Z] - \{0\}$ the image is given by $\sigma(g) = g + I(g)$ with $I(g) \in X_1 \cdot k[Z, X_1]$.

Then consider the polynomial ring $S := k(Z)[Y, X_1, X_2]$ over the field of fractions $k(Z)$ of the polynomial ring $k[Z]$. Then S has the maximal ideal $\mathfrak{m} = (Y, X_1, X_2)$.

Then set $B := S_{\mathfrak{m}} = k(Z)[Y, X_1, X_2]_{(Y,X_1,X_2)}$. We can regard all these rings as subrings of the field of fractions of R:

$$R \subset S \subset B \subset k(Z, Y, X_1, X_2).$$

Clearly, σ acts on R, and hence it induces an action on its field of fractions; denote this action by σ as well. Then we claim that the restriction of σ to B induces an action on B by local automorphisms. For this, it suffices to show that for any $g \in R - \mathfrak{I}$ the image $\sigma(g)$ does not belong to $\mathfrak{I}$. The latter is true, since $\sigma(g) = g + I(g)$ with $I(g) \in \mathfrak{I}$. The augmentation ideal $I_G = B \cdot X_1 + B \cdot X_2$ is not principal although G acts through a pseudoreflection.

3. A conjecture

Remark 8. In the tame case $p \neq \text{char}(k_B)$, the converse (d) $\Longrightarrow$ (a) is also true due to the theorem of Serre, as explained in the introduction.

In the case of a wild group action, that is, $p = \text{char}(k_B)$, it is not known whether the converse is true, but we conjecture it.

Conjecture 9. Let B be a regular local ring and let G be a p-cyclic group acting on B by local automorphisms. Then the following conditions are *conjectured* to be equivalent:

(1) I_G is principal.

(2) $A := B^G$ is regular.

The implication (1) $\Longrightarrow$ (2) was shown in Theorem 2. Of course the converse is true if $\dim A \leq 1$. In higher dimension, the converse (2) $\Longrightarrow$ (1) is uncertain, but it holds for small primes $p \leq 3$ as we explain now. Since A is regular, the ring B is a free A-module of rank p; see [Serre 1965, IV, Proposition 22]. So,

$$B/B\mathfrak{m}_A^n \text{ is a free } A/\mathfrak{m}_A^n\text{-module of rank } p \text{ for any } n \in \mathbb{N}. \qquad (*)$$

In the case $p = 2$, the rank of $\mathfrak{m}_B/B\mathfrak{m}_A$ is 0 or 1. In the first case, k_B is an extension of degree $[k_B : k_A] = 2$ over k_A and $\mathfrak{m}_B = B\mathfrak{m}_A$. So there exists an element $\beta \in B$ such that $B/B\mathfrak{m}_A$ is generated by the residue classes of 1 and β. Due to Nakayama's lemma, $B = A[\beta]$ is monogenous, and hence, I_G is principal. In the second case, where $k_A \to k_B$ is an isomorphism, there exists an element $\beta \in \mathfrak{m}_B$ such that $\mathfrak{m}_B = B\beta + B\mathfrak{m}_A$. Then G acts as a pseudoreflection, and hence, I_G is principal.

In the case $p = 3$ we claim that $B\mathfrak{m}_A \not\subset \mathfrak{m}_B^2$.

If we assume the contrary $B\mathfrak{m}_A \subset \mathfrak{m}_B^2$, then these ideals coincide; $B\mathfrak{m}_A = \mathfrak{m}_B^2$. Namely, the rank of $B/B\mathfrak{m}_A$ as $A/\mathfrak{m}_A$-module is 3 and the rank of $B/\mathfrak{m}_B^2$ is at least 3 due to $d := \dim B \geq 2$, so $B\mathfrak{m}_A = \mathfrak{m}_B^2$. Therefore the length of $B/B\mathfrak{m}_A^2 = B/\mathfrak{m}_B^4$

is 3 times the length of $A/\mathfrak{m}_A^2$, which is $3\cdot(\dim A+1)$. On the other hand the rank of $B/\mathfrak{m}_B^4$ is equal to

$$(1+\dim \mathfrak{m}_B/\mathfrak{m}_B^2)+\dim \mathfrak{m}_B^2/\mathfrak{m}_B^3+\dim \mathfrak{m}_B^3/\mathfrak{m}_B^4=\sum_{n=0}^{3}\binom{d+n-1}{d-1},$$

which is larger than $(1+\dim \mathfrak{m}_A/\mathfrak{m}_A^2)+(1+\dim \mathfrak{m}_A/\mathfrak{m}_A^2)+(1+\dim \mathfrak{m}_A/\mathfrak{m}_A^2)$, since for $d\geq 2$ both

$$\binom{d+1}{d-1}=\frac{(d+1)d}{2}\geq 1+d=1+\dim \mathfrak{m}_A/\mathfrak{m}_A^2$$

and

$$\binom{d+3-1}{d-1}=\frac{(d+2)(d+1)d}{2\cdot 3}>1+d$$

hold. Here we used the formula for the number $\lambda_{n,d}$ of monomials $T_1^{m_1}\cdots T_d^{m_d}$ in d variables of degree $n=m_1+\cdots+m_d$:

$$\lambda_{n,d}=\binom{d+n-1}{d-1}.$$

So, using only the condition $(*)$ and proceeding by induction on $\dim(A)$, we see that there exists a system of parameters $\alpha_1,\ldots,\alpha_d$ of A such that $\alpha_2,\ldots,\alpha_d$ is part of a system of parameters of B. In the case where $k_A\to k_B$ is an isomorphism, G acts as a pseudoreflection, and hence I_G is principal. If $k_A\to k_B$ is not an isomorphism, then we must have $\mathfrak{m}_B=B\mathfrak{m}_A$; otherwise the rank of $B/\mathfrak{m}_B$ is at least 4. Since $[k_B:k_A]\leq 3$, the field extension $k_A\to k_B$ is monogenous, and hence $A\to B$ is monogenous due to the lemma of Nakayama.

4. Relationship between the regular and the stable model of a smooth curve

As explained in the introduction, our incentive to study the invariant rings under a p-cyclic group action stems from the study of the relationship between the regular and the stable model of a smooth projective curve over the field of fractions K of a discrete valuation ring R. So let $R\hookrightarrow R'$ be a Galois extension of discrete valuation rings of prime order p and let π and π' be uniformizers of R and of R', respectively. Denote by K' the field of fractions of R' and let k and k' be the residue fields of R and R', respectively. Assume that $k=k'$ is algebraically closed and that $\mathrm{char}(k)=p$. Let G be the Galois group of R' over R.

In the tame case, the action can always be diagonalized and the invariant rings have the well-known Hirzebruch–Jung singularities. The tame case of higher dimension is also settled in [Edixhoven 1992, Proposition 3.5]. If the action of G is wild, this is in general not the case and the situation becomes quite capricious.

For example, consider an elliptic curve E over K having good reduction over K', and let X' be the corresponding proper smooth R'-model of $E \otimes_K K'$. Then G acts naturally on X', and hence one can consider the quotient $Y = X'/G$, which is a normal proper flat R-model of E. Assume that E has reduction of Kodaira type I_0^* over K; see [Silverman 1986, Theorem 15.2]. Curves of this type exist, since elliptic curves with Kodaira type I_0^* have integer j-invariant and thus potentially good reduction. Moreover, that a wild extension might be needed can be checked via Tate's algorithm [1975]. Let X be the minimal regular R-model of E. Then X happens to be a minimal blowing-up of Y and, in general, Y has singularities that are not of Hirzebruch–Jung type, since the special fiber of X contains components having three neighbors.

Our result now provides a tool to study the correspondence between X and the singularities of Y by looking at the group action G on X' and on R'-models Z', which are obtained by blowing-up G-invariant centers of X'. On these models, one can study the augmentation ideal and thereby obtain statements about which components have to occur in a desingularization of Y and in the regular model X, respectively. Since this analysis is beyond the scope of this article, we intend to explain this in greater detail in a further paper.

In the following we will look at Conjecture 9 in the case of relative curves.

Proposition 10. *Keep the situation of above. Let Y be an affine smooth relative curve over R' such that its closed fiber $Y \otimes_{R'} k'$ is irreducible. Assume that G acts on $Y \to \operatorname{Spec}(R')$ equivariantly. Let $B := \mathcal{O}_Y(Y)$ be the coordinate ring of Y. Then the following assertions are equivalent*:

(1) *The augmentation ideal I_G is locally principal.*

(2) *The ring $A := B^G$ of invariants is regular and $A/\mathfrak{p}$ is regular where $\mathfrak{p} = A \cap B\pi'$.*

Proof. (1) $\Longrightarrow$ (2). It follows from Theorem 2 that A is regular. It remains to show that the special fiber is regular. For showing this, it is enough to prove it after the π-adic completion, since the group action extends to the completion, taking invariants commutes with completion, and regularity of $A/\mathfrak{p}$ can be checked after π-adic completion. So we may assume that B is the coordinate ring of the associated formal completion of Y with respect to its special fiber. So set

$$\mathfrak{P} := B\pi' \quad \text{and} \quad \mathfrak{p} := A \cap \mathfrak{P}.$$

Then we obtain a finite extension of discrete valuation rings $A_{\mathfrak{p}} \hookrightarrow B_{\mathfrak{P}}$. Namely, the localization with respect to $A - \mathfrak{p}$ yields a finite flat extension $A_{\mathfrak{p}} \hookrightarrow B_{\mathfrak{p}}$. Since $\mathfrak{P}$ is the unique prime ideal of B lying above $\mathfrak{p}$, so $B_{\mathfrak{p}}$ is a local Dedekind ring, and hence we get $B_{\mathfrak{p}} = B_{\mathfrak{P}}$. Since A is regular, and hence locally factorial, the ideal $\mathfrak{p}$ is locally principal. The extended ideal $B\mathfrak{p}$ is locally principal and a power of $\mathfrak{P}$ and, hence, globally a power of $\mathfrak{P}$, that is, $\mathfrak{P}^e = B\mathfrak{p}$. The degree of the residue extension is denoted by $f := [Q(B/\mathfrak{P}) : Q(A/\mathfrak{p})]$. Moreover we have $p = e \cdot f$. In the case $f = p$ and $e = 1$ we have $\mathfrak{P} = B\mathfrak{p}$. Since $A \hookrightarrow B$ is faithfully flat, so $A/\mathfrak{p} \to B/\mathfrak{P}$ is faithfully flat as well. Then, due to [Matsumura 1980, Theorem 51], the ring $A/\mathfrak{p}$ is regular.

In the case $f = 1$, $e = p$, the ideal $\mathfrak{p}$ contains the uniformizer π of R. Since $\mathfrak{p}B = \mathfrak{P}^p$ due to $e = p$ and $\mathfrak{P} = B\pi'$ as Y is smooth over S, we obtain by faithfully flat descent $\mathfrak{p} = A\pi$. Therefore $A \otimes_R k$ is reduced and hence geometrically reduced. Then A is the set of all G-invariant functions f on Y that are bounded by 1 and also B consists of all functions on Y that are bounded by 1; see [Bosch et al. 1984, 6.4.3/4]. Moreover, it follows from [loc. cit.] that $A \otimes_R R'$ coincides with B. Thus we see that $A \otimes_R k = A \otimes_R R' \otimes_{R'} k' = B \otimes_{R'} k'$ is regular.

(2) $\Longrightarrow$ (1). For the converse implication, A is regular. Since B is regular as well, the extension $A \to B$ is faithfully flat; see [Serre 1965, IV, Proposition 22]. As above, we have the finite extension of discrete valuation rings $A_\mathfrak{p} \hookrightarrow B_\mathfrak{P}$ and its associated numbers e and f. In the case, $f = 1$ and $e = p$ the finite ring extension $A/\mathfrak{p} \to B/\mathfrak{P}$ is birational, and hence an isomorphism as $A/\mathfrak{p}$ is regular. So any local parameter of $A/\mathfrak{p}$ gives rise to a local parameter of $B/\mathfrak{P}$. Therefore, any maximal ideal of B is generated by a G-invariant element and π'. Therefore, $I_G = B \cdot I(\pi')$ is principal.

Now consider the case $f = p$ and $e = 1$. Since A is regular, the ideal $\mathfrak{p}$ is locally principal. So we may assume that $\mathfrak{p} = A\alpha$ is principal. Due to $e = 1$, we obtain $\mathfrak{P} = B\alpha$. Since $B/\mathfrak{P}$ is regular, any maximal ideal of B is generated by α and a lifting of a local parameter of $B/\mathfrak{P}$. Therefore, I_G is locally principal as it is generated by the $I(\beta)$, where β is a lifting of the local parameter $\bar{\beta}$ of $B/\mathfrak{P}$. □

Conjecture 11. In the case of an affine arithmetic surface, that is, Y is regular with irreducible special fiber, one conjectures that the following conditions are equivalent, where $\mathfrak{P} \subset B$ is the prime ideal whose locus is the special fiber and $\mathfrak{p} := A \cap \mathfrak{P}$:

(1) I_G is locally principal and $B/\mathfrak{P}$ is regular.

(2) A is regular and $A/\mathfrak{p}$ is regular.

The proof of the last proposition tells us that the implication (1) $\Longrightarrow$ (2) is true in the case $f = p$ and $e = 1$. In the case $f = 1$ and $e = p$, we used the fact that the formation of the ring of 1-bounded functions is compatible with base change; this is true when the multiplicity is 1. But it is not clear if one only knows that both models A and B have the same multiplicity in the special fiber over their base rings.

The implication (2) $\Longrightarrow$ (1) is true in the case $f = 1$ and $e = p$, as seen by the same arguments as given in Proposition 10. But the case $f = p$ and $e = 1$, is uncertain, although in this case the multiplicity behaves well.

Acknowledgements

S. Wewers [2010] has obtained partial versions of our results by different methods. It is our pleasure to thank the referee for a careful reading of the manuscript.

References

[Bosch et al. 1984] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis: A systematic approach to rigid analytic geometry*, Grundlehren der Mathematischen Wissenschaften **261**, Springer, Berlin, 1984. MR 86b:32031 Zbl 0539.14017

[Bourbaki 1981] N. Bourbaki, *Éléments de mathématique: Groupes et algèbres de Lie, Chapitres 4, 5 et 6*, Masson, Paris, 1981. MR 83g:17001 Zbl 0483.22001

[Edixhoven 1992] B. Edixhoven, "Néron models and tame ramification", *Compositio Math.* **81**:3 (1992), 291–306. MR 93a:14041 Zbl 0759.14033

[Hirzebruch 1953] F. Hirzebruch, "Über vierdimensionale Riemannsche Flächen mehrdeutiger analytischer Funktionen von zwei komplexen Veränderlichen", *Math. Ann.* **126** (1953), 1–22. MR 16,26d Zbl 0093.27605

[Jung 1908] H. W. E. Jung, "Darstellung der Funktionen eines algebraischen Körpers zweier unabhängiger Veränderlicher x, y in der Umgebung einer Stelle $x = a$, $y = b$", *Journal für die Reine und Angewandte Mathematik* **133** (1908), 289–314. JFM 39.0493.01

[Matsumura 1980] H. Matsumura, *Commutative algebra*, 2nd ed., Mathematics Lecture Note Series **56**, Benjamin/Cummings, Reading, MA, 1980. MR 82i:13003 Zbl 0441.13001

[Serre 1965] J.-P. Serre, *Algèbre locale: Multiplicités* (Cours au Collège de France, 1957–1958), 2nd ed., Lecture Notes in Mathematics **11**, Springer, Berlin, 1965. MR 34 #1352 Zbl 0142.28603

[Serre 1968] J.-P. Serre, "Groupes finis d'automorphismes d'anneaux locaux réguliers", pp. 11 in *Colloque d'Algèbre, Exp. 8* (Paris, 1967), Secrétariat mathématique, Paris, 1968. MR 38 #3267 Zbl 0200.00002

[Silverman 1986] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, 1986. MR 87g:11070 Zbl 0585.14026

[Tate 1975] J. Tate, "Algorithm for determining the type of a singular fiber in an elliptic pencil", pp. 33–52 in *Modular functions of one variable, IV* (Antwerp, 1972), edited by B. J. Birch and W. Kuyk, Lecture Notes in Math. **476**, Springer, Berlin, 1975. MR 52 #13850 Zbl 1214.14020

[Wewers 2010] S. Wewers, "Regularity of quotients by an automorphism of order p", preprint, 2010. arXiv 1001.0607

Communicated by David Eisenbud
Received 2011-04-14 Revised 2012-01-23 Accepted 2012-02-20

franz.j.kiraly@tu-berlin.de *Berlin Institute of Technology, Machine Learning Group, Marchstraße 23, 10587 Berlin, Germany*

Freie Universität Berlin, Discrete Geometry Group, Arnimallee 2, 14195 Berlin, Germany

Mathematisches Forschungsinstitut Oberwolfach, Schwarzwaldstraße 9-11, 77709 Oberwolfach, Germany

werner.luetkebohmert@uni-ulm.de
Dept. of Pure Mathematics, University of Ulm, Helmholtzstraße 18, 89069 Ulm, Germany

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.75

On the arithmetic and geometry of binary Hamiltonian forms

Jouni Parkkonen and Frédéric Paulin
Appendix by Vincent Emery

Given an indefinite binary quaternionic Hermitian form f with coefficients in a maximal order of a definite quaternion algebra over $\mathbb{Q}$, we give a precise asymptotic equivalent to the number of nonequivalent representations, satisfying some congruence properties, of the rational integers with absolute value at most s by f, as s tends to $+\infty$. We compute the volumes of hyperbolic 5-manifolds constructed by quaternions using Eisenstein series. In the appendix, V. Emery computes these volumes using Prasad's general formula. We use hyperbolic geometry in dimension 5 to describe the reduction theory of both definite and indefinite binary quaternionic Hermitian forms.

1. Introduction

Following [Weyl 1940; 1942], we will call a Hermitian form over Hamilton's real quaternion algebra with anti-involution the conjugation a *Hamiltonian form*.

Since Gauss, the reduction theory of the integral binary quadratic forms and the problem of representation of integers by them is quite completely understood. For binary Hermitian forms, these subjects have been well studied, starting with Hermite, Bianchi and especially Humbert, and much developed by Elstrodt, Grunewald and Mennicke; see for instance [Elstrodt et al. 1998]. In the recent paper [Parkkonen and Paulin 2011], we gave a precise asymptotic on the number of nonequivalent proper representations of rational integers with absolute value at most s by a given integral indefinite Hermitian form. Besides the general results on quadratic forms (see for instance [Weyl 1940; Cassels 1978]) and some special work (see for instance [Pronin 1967; Hashimoto and Ibukiyama 1980]), not much seemed to be precisely known on these questions for binary Hamiltonian forms.

The work in the appendix is supported by the Swiss National Science Foundation, project number PP00P2-128309/1.
MSC2010: primary 11E39, 11R52, 20G20; secondary 11N45, 15A21, 53A35, 11F06, 20H10.
Keywords: binary Hamiltonian form, representation of integers, group of automorphs, Hamilton–Bianchi group, hyperbolic volume, reduction theory.

In this paper, we use hyperbolic geometry in dimension 5 to study the asymptotic of the counting of representations of rational integers by integral binary Hamiltonian forms and to give a geometric description of the reduction theory of such forms. General formulas are known (by Siegel's mass formula; see for instance [Eskin et al. 1991]), but it does not seem to be easy (or even doable) to deduce our asymptotic formulas from them. There are numerous results on the counting of integer points with bounded norm on quadrics (or homogeneous varieties); see for instance the work of Duke, Eskin, McMullen, Oh, Rudnick, Sarnak and others. In this paper, we count appropriate orbits of integer points on which a fixed integral binary Hamiltonian form is constant, analogously to [Parkkonen and Paulin 2011].

Let $\mathbb{H}$ be Hamilton's quaternion algebra over $\mathbb{R}$, with $x \mapsto \overline{x}$ its conjugation, $\mathrm{n} : x \mapsto x\overline{x}$ its reduced norm and $\mathrm{tr} : x \mapsto x + \overline{x}$ its reduced trace. Let A be a quaternion algebra over $\mathbb{Q}$ that is definite ($A \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{H}$), with reduced discriminant D_A and class number h_A. Let $\mathcal{O}$ be a maximal order in A, and let $\mathfrak{m}$ be a (nonzero) left fractional ideal of $\mathcal{O}$, with reduced norm $\mathrm{n}(\mathfrak{m})$; see Section 2 for definitions.

Let $f : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ be a binary Hamiltonian form, with

$$f(u, v) = a\,\mathrm{n}(u) + \mathrm{tr}(\overline{u}\, b\, v) + c\,\mathrm{n}(v), \tag{1}$$

that is integral over $\mathcal{O}$ (its coefficients satisfy $a, c \in \mathbb{Z}$ and $b \in \mathcal{O}$) and indefinite (its discriminant $\Delta(f) = \mathrm{n}(b) - ac$ is positive); see Section 4. We denote by $\mathrm{SL}_2(\mathcal{O})$ the group of invertible 2×2 matrices with coefficients in $\mathcal{O}$; see Section 3. The group $\mathrm{SU}_f(\mathcal{O})$ of automorphs of f consists of those elements $g \in \mathrm{SL}_2(\mathcal{O})$ for which $f \circ g = f$. Given an arithmetic group Γ, such as $\mathrm{SL}_2(\mathcal{O})$ or $\mathrm{SU}_f(\mathcal{O})$, we will denote by $\mathrm{Covol}(\Gamma)$ the volume of the quotient by Γ of its associated symmetric space (assumed to be of noncompact type and normalized to have -1 as the maximum of its sectional curvature).

For every $s > 0$, we consider the integer

$$\psi_{f,\mathfrak{m}}(s) = \mathrm{Card}\ {}_{\mathrm{SU}_f(\mathcal{O})}\backslash \{(u, v) \in \mathfrak{m} \times \mathfrak{m} : \mathrm{n}(\mathfrak{m})^{-1} |f(u, v)| \leq s,\ \mathcal{O}u + \mathcal{O}v = \mathfrak{m}\},$$

which is the number of nonequivalent $\mathfrak{m}$-primitive representations by f of rational integers with absolute value at most s. The finiteness of $\psi_{f,\mathfrak{m}}(s)$ follows from general results on orbits of algebraic groups defined over number fields [Borel and Harish-Chandra 1962, Lemma 5.3].

Theorem 1. *As s tends to $+\infty$, we have the equivalence, with p ranging over positive rational primes,*

$$\psi_{f,\mathfrak{m}}(s) \sim \frac{45\, D_A\, \mathrm{Covol}(\mathrm{SU}_f(\mathcal{O}))}{2\pi^2 \zeta(3) \Delta(f)^2 \prod_{p | D_A} (p^3 - 1)}\, s^4$$

This result follows from the more general Theorem 13, which allows us in particular to count representations satisfying given congruence properties (see the end of Section 6).

Here is an example of our applications, concerning the asymptotic of the very useful real scalar product $(u, v) \mapsto \operatorname{tr}(\bar{u}\, v)$ on $\mathbb{H}$. See Section 6 for the proof and for further applications. Let

$$\mathrm{Sp}_1(\mathbb{O}) = \left\{ g \in \mathrm{SL}_2(\mathbb{O}) : {}^t\overline{g} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

Corollary 2. *As s tends to $+\infty$, we have the equivalence*

$$\mathrm{Card}\ {}_{\mathrm{Sp}_1(\mathbb{O})}\backslash\{(u, v) \in \mathbb{O}\times\mathbb{O} : |\operatorname{tr}(\bar{u}v)| \le s,\ \mathbb{O}u + \mathbb{O}v = \mathbb{O}\} \sim \frac{D_A}{48\,\zeta(3)} \prod_{p \mid D_A} \frac{p^2+1}{p^2+p+1}\, s^4.$$

To prove Theorem 1, applying a counting result of [Parkkonen and Paulin 2012] following from dynamical properties of the geodesic flow of real hyperbolic manifolds, we first prove that

$$\psi_{f,\mathfrak{m}}(s) \sim \frac{D_A \prod_{p \mid D_A}(p-1)\, \mathrm{Covol}(\mathrm{SU}_f(\mathbb{O}))}{512\pi^2 \Delta(f)^2\, \mathrm{Covol}(\mathrm{SL}_2(\mathbb{O}))}\, s^4.$$

The covolumes of the arithmetic groups $\mathrm{SL}_2(\mathbb{O})$ and $\mathrm{SU}_f(\mathbb{O})$ may be computed using the very general formula of [Prasad 1989]; see [Emery 2009] for an excellent exposition. Following the approach of [Rankin 1939a; 1939b; Selberg 1940], see also [Langlands 1966; Sarnak 1983] and others, we compute $\mathrm{Covol}(\mathrm{SL}_2(\mathbb{O}))$ in the main body of this paper (see Section 5) using Eisenstein series, whose analytic properties in the quaternion setting have been studied in [Krafft and Osenberg 1990]. We initially proved the case $h_A = 1$ of the following result, V. Emery proved the general case using Prasad's formula (see the appendix), and we afterwards managed to push the Eisenstein series approach to get the general result. The two proofs are completely different.

Theorem 3 (Emery; see the appendix). *We have*

$$\mathrm{Covol}(\mathrm{SL}_2(\mathbb{O})) = \frac{\zeta(3) \prod_{p \mid D_A}(p^3-1)(p-1)}{11520}.$$

In the final section, we give a geometric reduction theory of binary Hamiltonian forms using real hyperbolic geometry. The case of binary quadratic forms is well known, from either the arithmetic, geometric or algorithmic viewpoint; see for instance [Cassels 1978; Zagier 1981; Buchmann and Vollmer 2007]. We refer for instance to [Elstrodt et al. 1998] for the reduction theory of binary Hermitian forms. The case of binary Hamiltonian forms has been developed less; see for instance [Pronin 1967; Hashimoto and Ibukiyama 1980; 1981; 1983] for results in the positive definite case. We construct a natural map Ξ from the set $\mathcal{Q}(\mathcal{O}, \Delta)$ of binary Hamiltonian forms that are integral over $\mathcal{O}$ and have a fixed discriminant $\Delta \in \mathbb{Z} - \{0\}$ to the set of points or totally geodesic hyperplanes of the 5-dimensional real hyperbolic space $\mathbb{H}^5_{\mathbb{R}}$. For $\mathcal{F}_{\mathcal{O}}$ a Ford fundamental domain for the action of $\mathrm{SL}_2(\mathcal{O})$ on $\mathbb{H}^5_{\mathbb{R}}$, we say that $f \in \mathcal{Q}(\mathcal{O}, \Delta)$ is *reduced* if $\Xi(f)$ meets $\mathcal{F}_{\mathcal{O}}$. The finiteness of the number of orbits of $\mathrm{SL}_2(\mathcal{O})$ on $\mathcal{Q}(\mathcal{O}, \Delta)$, which can be deduced from general results of Borel and Harish-Chandra, then follows in an explicit way from the equivariance property of Ξ and the following result proved in Section 7.

Theorem 4. *There are only finitely many reduced integral binary Hamiltonian forms with a fixed nonzero discriminant.*

Answering the remark on page 257 of [Cassels 1978] that explicit sets of inequalities implying the reduction property were essentially only known for quadratic forms in dimension $n \leq 7$, we give an explicit such set in dimension 8 at the end of Section 7.

The knowledgeable reader may skip the background Sections 2 (except the new Lemma 6), 3 and 4 on respectively definite quaternion algebras over $\mathbb{Q}$, quaternionic homographies and real hyperbolic geometry in dimension 5, and binary Hamiltonian forms, though many references are made to them in the subsequent sections.

2. Background on definite quaternion algebras over $\mathbb{Q}$

A *quaternion algebra* over a field F is a four-dimensional central simple algebra over F. We refer for instance to [Vignéras 1980] for generalities on quaternion algebras.

A real quaternion algebra is isomorphic either to $\mathcal{M}_2(\mathbb{R})$ or to Hamilton's quaternion algebra $\mathbb{H}$ over $\mathbb{R}$, with basis elements $1, i, j, k$ as a $\mathbb{R}$-vector space, with unit element 1, satisfying $i^2 = j^2 = -1$ and $ij = -ji = k$. We define the *conjugate* of $x = x_0 + x_1 i + x_2 j + x_3 k$ in $\mathbb{H}$ by $\overline{x} = x_0 - x_1 i - x_2 j - x_3 k$, its *reduced trace* by $\mathrm{tr}(x) = x + \overline{x}$, and its *reduced norm* by $\mathrm{n}(x) = x\overline{x} = \overline{x}x$. Note that $\mathrm{n}(xy) = \mathrm{n}(x)\,\mathrm{n}(y)$, and $\mathrm{n}(x) \geq 0$ with equality if and only if $x = 0$; hence $\mathbb{H}$ is a division algebra. Furthermore, $\mathrm{tr}(\overline{x}) = \mathrm{tr}(x)$ and $\mathrm{tr}(xy) = \mathrm{tr}(yx)$. For every matrix $X = (x_{i,j})_{1 \leq i \leq p, 1 \leq j \leq q} \in \mathcal{M}_{p,q}(\mathbb{H})$, we denote by $X^* = (\overline{x_{j,i}})_{1 \leq i \leq q, 1 \leq j \leq p} \in \mathcal{M}_{q,p}(\mathbb{H})$ its adjoint matrix, which satisfies $(XY)^* = Y^* X^*$. The matrix X is *Hermitian* if $X = X^*$.

Let A be a quaternion algebra over $\mathbb{Q}$. We say that A is *definite* (or ramified over $\mathbb{R}$) if the real quaternion algebra $A \otimes_{\mathbb{Q}} \mathbb{R}$ is isomorphic to $\mathbb{H}$. In this paper, whenever we consider a definite quaternion algebra A over $\mathbb{Q}$, we will fix an identification between $A \otimes_{\mathbb{Q}} \mathbb{R}$ and $\mathbb{H}$, so that A is a $\mathbb{Q}$-subalgebra of $\mathbb{H}$.

The *reduced discriminant* D_A of A is the product of the primes $p \in \mathbb{N}$ such that the quaternion algebra $A \otimes_{\mathbb{Q}} \mathbb{Q}_p$ over $\mathbb{Q}_p$ is a division algebra, with $[\mathbb{H}^\times, \mathbb{H}^\times] = \mathrm{n}^{-1}(1)$. Two definite quaternion algebras over $\mathbb{Q}$ are isomorphic if and only if they have the same reduced discriminant, which can be any product of an odd number of primes; see [Vignéras 1980, page 74].

A $\mathbb{Z}$*-lattice* I in A is a finitely generated $\mathbb{Z}$-module generating A as a $\mathbb{Q}$-vector space. The intersection of finitely many $\mathbb{Z}$-lattices of A is again a $\mathbb{Z}$-lattice. An *order* in a quaternion algebra A over $\mathbb{Q}$ is a unitary subring $\mathbb{O}$ of A which is a $\mathbb{Z}$-lattice. In particular, $A = \mathbb{Q}\mathbb{O}$. Each order of A is contained in a maximal order. The *type number* $t_A \geq 1$ of A is the number of conjugacy (or equivalently isomorphism) classes of maximal orders in A (see for instance [Vignéras 1980, page 152] for a formula). For instance, $t_A = 1$ if $D_A = 2, 3, 5, 7, 13$ and $t_A = 2$ if $D_A = 11, 17$. If $\mathbb{O}$ is a maximal order in A, then the ring $\mathbb{O}$ has 2, 4 or 6 invertible elements except that $|\mathbb{O}^\times| = 24$ when $D_A = 2$, and $|\mathbb{O}^\times| = 12$ when $D_A = 3$. When $D_A = 2, 3, 5, 7, 13$, then (see [Eichler 1938, page 103])

$$|\mathbb{O}^\times| = \frac{24}{D_A - 1}. \tag{2}$$

Example 5 (See [Vignéras 1980, page 98]).

(1) The $\mathbb{Q}$-vector space $A = \mathbb{Q} + \mathbb{Q}i + \mathbb{Q}j + \mathbb{Q}k$ generated by $1, i, j, k$ in $\mathbb{H}$ is Hamilton's quaternion algebra over $\mathbb{Q}$. It is the unique definite quaternion algebra over $\mathbb{Q}$ (up to isomorphism) with discriminant $D_A = 2$. The *Hurwitz order* $\mathbb{O} = \mathbb{Z} + \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}(1 + i + j + k)/2$ is maximal, and it is unique up to conjugacy.

(2) Similarly, $A = \mathbb{Q} + \mathbb{Q}i + \mathbb{Q}\sqrt{p}\,j + \mathbb{Q}\sqrt{p}\,k$ is the unique (up to isomorphism) definite quaternion algebra over $\mathbb{Q}$ with discriminant $D_A = p$ for $p = 3, 7$, and $\mathbb{O} = \mathbb{Z} + \mathbb{Z}i + \mathbb{Z}(i + \sqrt{p}\,j)/2 + \mathbb{Z}(1 + \sqrt{p}\,k)/2$ is its unique (up to conjugacy) maximal order.

(3) Similarly, $A = \mathbb{Q} + \mathbb{Q}\sqrt{2}i + \mathbb{Q}\sqrt{p}\,j + \mathbb{Q}\sqrt{2p}\,k$ is the unique (up to isomorphism) definite quaternion algebra over $\mathbb{Q}$ with discriminant $D_A = p$ for $p = 5, 13$, and

$$\mathbb{O} = \mathbb{Z} + \mathbb{Z}\frac{1+\sqrt{2}i+\sqrt{p}\,j}{2} + \mathbb{Z}\frac{\sqrt{p}\,j}{2} + \mathbb{Z}\frac{2+\sqrt{2}i+\sqrt{2p}\,k}{2}$$

is its unique (up to conjugacy) maximal order.

Let $\mathbb{O}$ be an order in A. The reduced norm n and the reduced trace tr take integral values on $\mathbb{O}$. The invertible elements of $\mathbb{O}$ are its elements of reduced norm 1. Since $\bar{x} = \mathrm{tr}(x) - x$, any order is invariant under conjugation.

The *left order* $\mathbb{O}_\ell(I)$ of a $\mathbb{Z}$-lattice I is $\{x \in A : xI \subset I\}$; its *right order* $\mathbb{O}_r(I)$ is $\{x \in A : Ix \subset I\}$. A *left fractional ideal* of $\mathbb{O}$ is a $\mathbb{Z}$-lattice of A whose left order is $\mathbb{O}$. A *left ideal* of $\mathbb{O}$ is a left fractional ideal of $\mathbb{O}$ contained in $\mathbb{O}$. Right (fractional) ideals are defined analogously. The *inverse* of a right fractional ideal $\mathfrak{m}$ of $\mathbb{O}$ is

$\mathfrak{m}^{-1} = \{x \in A : \mathfrak{m}x\mathfrak{m} \subset \mathfrak{m}\}$. It is easy to check that for every $u, v \in \mathbb{O}$, if $uv \neq 0$, then

$$(u\mathbb{O} + v\mathbb{O})^{-1} = \mathbb{O}u^{-1} \cap \mathbb{O}v^{-1}. \tag{3}$$

If $\mathbb{O}$ is maximal, then $\mathfrak{m}^{-1}$ is a left fractional ideal of $\mathbb{O}$ and

$$\mathbb{O}_r(\mathfrak{m}^{-1}) = \mathbb{O}_\ell(\mathfrak{m}). \tag{4}$$

This formula follows from [Vignéras 1980, Lemma 4.3(3), page 21], which says that $\mathbb{O}_r(\mathfrak{m}^{-1})$ contains $\mathbb{O}_\ell(\mathfrak{m})$, since the maximality of $\mathbb{O}$ implies the maximality of $\mathbb{O}_\ell(\mathfrak{m})$, by [ibid., Exercice 4.1, page 28].

Two left fractional ideals $\mathfrak{m}$ and $\mathfrak{m}'$ of $\mathbb{O}$ are isomorphic as left $\mathbb{O}$-modules if and only if $\mathfrak{m}' = \mathfrak{m}c$ for some $c \in A^\times$. A (left) *ideal class* of $\mathbb{O}$ is an equivalence class of left fractional ideals of $\mathbb{O}$ for this equivalence relation. We will denote by ${}_{\mathbb{O}}\mathcal{I}$ the set of ideal classes of $\mathbb{O}$, and by $[\mathfrak{m}]$ the ideal class of a left fractional ideal $\mathfrak{m}$ of $\mathbb{O}$. The *class number* h_A of A is the number of ideal classes of a maximal order $\mathbb{O}$ of A. It is finite and independent of the maximal order $\mathbb{O}$; see for instance [ibid., pages 87–88]. See for instance [ibid., pages 152–155] for a formula for h_A, and for the fact that $h_A = 1$ if and only if $D_A = 2, 3, 5, 7, 13$. In particular D_A is prime if $h_A = 1$.

The *norm* $\mathrm{n}(\mathfrak{m})$ of a left (or right) ideal $\mathfrak{m}$ of $\mathbb{O}$ is the greatest common divisor of the norms of the nonzero elements of $\mathfrak{m}$. In particular, $\mathrm{n}(\mathbb{O}) = 1$. The *norm* of a left (or right) fractional ideal $\mathfrak{m}$ of $\mathbb{O}$ is $\mathrm{n}(c\mathfrak{m})/\mathrm{n}(c)$ for any $c \in \mathbb{N} - \{0\}$ such that $c\mathfrak{m} \subset \mathbb{O}$.

Note that a $\mathbb{Z}$-lattice Λ in A is a $\mathbb{Z}$-lattice in the Euclidean vector space $\mathbb{H}$ (with orthonormal basis $(1, i, j, k)$), and the volume $\mathrm{Vol}(\Lambda \setminus \mathbb{H})$ is finite. If $\mathbb{O}$ is maximal, we have (see for instance [Krafft and Osenberg 1990, Lemma 5.5])

$$\mathrm{Vol}(\mathbb{O}\backslash\mathbb{H}) = \frac{D_A}{4}\,. \tag{5}$$

The classical *zeta function* of A is

$$\zeta_A(s) = \sum_{\mathfrak{a}} \frac{1}{\mathrm{n}(\mathfrak{a})^{2s}}\,,$$

where the sum is over all left ideals $\mathfrak{a}$ in a maximal order $\mathbb{O}$ of A. It is independent of the choice of $\mathbb{O}$, it is holomorphic on $\{s \in \mathbb{C} : \mathrm{Re}\, s > 1\}$ and it satisfies by a theorem of Hey, with ζ the usual Riemann zeta function,

$$\zeta_A(s) = \zeta(2s)\zeta(2s-1) \prod_{p \mid D_A} (1 - p^{1-2s}), \tag{6}$$

where as usual the index p is prime; see [Schoeneberg 1939, page 88; Vignéras 1980, page 64]. Let $\mathfrak{m}$ be a left fractional ideal of a maximal order $\mathbb{O}$ in A. Define

$$\zeta(\mathfrak{m}, s) = \mathrm{n}(\mathfrak{m})^{2s} \sum_{x \in \mathfrak{m} - \{0\}} \frac{1}{\mathrm{n}(x)^{2s}},$$

which is also holomorphic on $\operatorname{Re} s > 1$ (and depends only on the ideal class of $\mathfrak{m}$), and

$$\zeta_{[\mathfrak{m}]}(s) = \sum \frac{1}{\mathrm{n}(\mathfrak{a})^{2s}}$$

where the sum is over all left ideals $\mathfrak{a}$ in $\mathbb{O}$ whose ideal class is $[\mathfrak{m}]$. The relations we will use in Section 5 between these zeta functions are the following ones, where $\operatorname{Re} s > 1$. The first one is obvious; see for instance respectively [Deuring 1968, page 134] and [Krafft and Osenberg 1990, page 436] for the other two:

$$\zeta_A(s) = \sum_{[\mathfrak{a}] \in_{\mathbb{O}} \mathcal{I}} \zeta_{[\mathfrak{a}]}(s), \tag{7}$$

$$\sum_{[\mathfrak{a}] \in_{\mathbb{O}} \mathcal{I}} \frac{1}{|\mathbb{O}_r(\mathfrak{a})^\times|} = \frac{1}{24} \prod_{p | D_A} (p-1), \tag{8}$$

$$\zeta(\mathfrak{m}, s) = |\mathbb{O}_r(\mathfrak{m})^\times| \, \zeta_{[\mathfrak{m}]}(s). \tag{9}$$

Note that when the class number h_A of A is 1, the formula (9) becomes

$$\zeta(\mathbb{O}, s) = |\mathbb{O}^\times| \, \zeta_A(s). \tag{10}$$

We end this section with the following lemma, which will be used in the proof of Theorem 13.

Lemma 6. *Let $\mathbb{O}$ be a maximal order in a definite quaternion algebra A over $\mathbb{Q}$, let $z \in A - \{0\}$ and let $\Lambda = \mathbb{O} \cap z\mathbb{O} \cap \mathbb{O}z \cap z\mathbb{O}z$. Then Λ is a $\mathbb{Z}$-sublattice of $\mathbb{O}$ such that*

$$[\mathbb{O} : \Lambda] \, \mathrm{n}(\mathbb{O}z^{-1} + \mathbb{O})^4 = 1.$$

Proof. This is a "prime by prime" type of proof, suggested by G. Chenevier. As an intersection of four $\mathbb{Z}$-lattices, Λ is a $\mathbb{Z}$-lattice, contained in $\mathbb{O}$. For every (positive rational) prime p, let v_p be the p-adic valuation on $\mathbb{Q}_p$; let us consider the quaternion algebra $A_p = A \otimes_{\mathbb{Q}} \mathbb{Q}_p$ over $\mathbb{Q}_p$, whose reduced norm is denoted by $\mathrm{n}_p : A_p \to \mathbb{Q}_p$; and for every $\mathbb{Z}$-lattice L of A, let $L_p = L \otimes_{\mathbb{Z}} \mathbb{Z}_p$. We embed A in A_p as usual by $x \mapsto x \otimes 1$. We then have the following properties (see for instance [Vignéras 1980, page 83-84]): L_p is a $\mathbb{Z}_p$-lattice of A_p; the map $L \mapsto L_p$ commutes with the inclusion, the sum and the intersection; if L and L' are $\mathbb{Z}$-lattices with $L \subset L'$, then

$$[L' : L] = \prod_p [L'_p : L_p];$$

if L is a left fractional ideal of $\mathcal{O}$, then L_p is a left fractional ideal of $\mathcal{O}_p$, and

$$\mathrm{n}(L)=\prod_p p^{\nu_p(\mathrm{n}_p(L_p))}.$$

Hence in order to prove Lemma 6, we only have to prove that for every prime p, if $z \in A_p^\times$ and $\Lambda_p=\mathcal{O}_p \cap z\mathcal{O}_p \cap \mathcal{O}_p\, z \cap z\mathcal{O}_p\, z$, we have

$$[\mathcal{O}_p:\Lambda_p]=p^{-4\nu_p(\mathrm{n}_p(\mathcal{O}_p z^{-1}+\mathcal{O}_p))}. \tag{11}$$

We distinguish two cases.

First assume that p does not divide D_A. Then we may assume that $A_p=\mathcal{M}_2(\mathbb{Q}_p)$ and $\mathcal{O}_p=\mathcal{M}_2(\mathbb{Z}_p)$ (by the uniqueness up to conjugacy of maximal orders). By Cartan's decomposition of $\mathrm{GL}_2(\mathbb{Q}_p)$ (see for instance [Bruhat and Tits 1972], or consider the action of $\mathrm{GL}_2(\mathbb{Q}_p)$ on its Bruhat–Tits tree as in [Serre 1977]), the element $z \in \mathrm{GL}_2(\mathbb{Q}_p)$ may be written

$$z=P\begin{pmatrix}p^a & 0\\ 0 & p^b\end{pmatrix}Q$$

with P, Q in the (good) maximal compact subgroup $\mathrm{GL}_2(\mathbb{Z}_p)$ and a, b in $\mathbb{Z}$. Since $\mathrm{GL}_2(\mathbb{Z}_p)$ preserves $\mathcal{O}_p=\mathcal{M}_2(\mathbb{Z}_p)$ by left and right multiplication, preserves the indices of $\mathbb{Z}$-lattices, and contains only elements of reduced norm (that is, of determinant) having valuation 0, we may assume that $P=Q=\mathrm{id}$. We hence have, by an easy matrix computation,

$$\begin{aligned}\Lambda_p&=\begin{pmatrix}\mathbb{Z}_p\cap p^a\,\mathbb{Z}_p\cap p^{2a}\,\mathbb{Z}_p & \mathbb{Z}_p\cap p^a\,\mathbb{Z}_p\cap p^b\,\mathbb{Z}_p\cap p^{a+b}\,\mathbb{Z}_p\\ \mathbb{Z}_p\cap p^a\,\mathbb{Z}_p\cap p^b\,\mathbb{Z}_p\cap p^{a+b}\,\mathbb{Z}_p & \mathbb{Z}_p\cap p^b\,\mathbb{Z}_p\cap p^{2b}\,\mathbb{Z}_p\end{pmatrix}\\ &=\begin{pmatrix}p^{2\max\{a,0\}}\,\mathbb{Z}_p & p^{\max\{a,0\}+\max\{b,0\}}\,\mathbb{Z}_p\\ p^{\max\{a,0\}+\max\{b,0\}}\,\mathbb{Z}_p & p^{2\max\{b,0\}}\,\mathbb{Z}_p\end{pmatrix}.\end{aligned}$$

Similarly, we have

$$\mathcal{O}_p\,z^{-1}+\mathcal{O}_p=\begin{pmatrix}p^{-a}\,\mathbb{Z}_p+\mathbb{Z}_p & p^{-b}\,\mathbb{Z}_p+\mathbb{Z}_p\\ p^{-a}\,\mathbb{Z}_p+\mathbb{Z}_p & p^{-b}\,\mathbb{Z}_p+\mathbb{Z}_p\end{pmatrix}=\mathcal{M}_2(\mathbb{Z}_p)\begin{pmatrix}p^{\min\{-a,0\}} & 0\\ 0 & p^{\min\{-b,0\}}\end{pmatrix}.$$

Therefore, since $\mathrm{n}_p(\mathcal{M}_2(\mathbb{Z}_p))=1$ and $\mathrm{n}_p=\det$ on $A_p=\mathcal{M}_2(\mathbb{Q}_p)$,

$$\begin{aligned}&[\mathcal{O}_p:\Lambda_p]\\ &\quad=\big|\mathbb{Z}_p/(p^{2\max\{a,0\}}\,\mathbb{Z}_p)\big|\,\big|\mathbb{Z}_p/(p^{\max\{a,0\}+\max\{b,0\}}\,\mathbb{Z}_p)\big|^2\,\big|\mathbb{Z}_p/(p^{2\max\{b,0\}}\,\mathbb{Z}_p)\big|\\ &\qquad=p^{4(\max\{a,0\}+\max\{b,0\})}=p^{-4(\min\{-a,0\}+\min\{-b,0\})}=p^{-4\,\nu_p(\mathrm{n}_p(\mathcal{O}_p z^{-1}+\mathcal{O}_p))},\end{aligned}$$

as wanted.

Now assume that p divides D_A, so that A_p is a division algebra. Let $\nu=\nu_p\circ\mathrm{n}_p$, which, by for instance [Vignéras 1980, page 34], is a discrete valuation on A_p, whose valuation ring is $\mathbb{O}_p$. The left ideals of $\mathbb{O}_p$ are two-sided ideals. Let π be a uniformizer of $\mathbb{O}_p$. Note that the residual field $\mathbb{O}_p/\pi\mathbb{O}_p$ has order p^2, and that $\mathrm{n}_p(\mathbb{O}_p)=1$ and $\mathrm{n}_p(\pi)=p$. We have

$$\Lambda_p=\mathbb{O}_p\pi^{2\max\{\nu(z),\,0\}}\quad\text{and}\quad \mathbb{O}_p z^{-1}+\mathbb{O}_p=\mathbb{O}_p\pi^{\min\{\nu(z^{-1}),0\}}.$$

Hence $[\mathbb{O}_p:\Lambda_p]=p^{4\max\{\nu(z),\,0\}}$ and $\nu_p(\mathrm{n}_p(\mathbb{O}_p\,z^{-1}+\mathbb{O}_p))=-\max\{\nu(z),0\}$, which is also as wanted. □

3. Background on Hamilton–Bianchi groups

The Dieudonné determinant (see [Dieudonné 1943; Aslaksen 1996]) Det is the group morphism from the group $\mathrm{GL}_2(\mathbb{H})$ of invertible 2×2 matrices with coefficients in $\mathbb{H}$ to $\mathbb{R}_+^*$, defined by

$$\begin{aligned}\left(\mathrm{Det}\begin{pmatrix}a&b\\c&d\end{pmatrix}\right)^2&=\mathrm{n}(a\,d)+\mathrm{n}(b\,c)-\mathrm{tr}(a\,\overline{c}\,d\,\overline{b})\\&=\begin{cases}\mathrm{n}(ad-aca^{-1}b)&\text{if } a\neq 0,\\ \mathrm{n}(cb-cac^{-1}d)&\text{if } c\neq 0,\\ \mathrm{n}(cb-db^{-1}ab)&\text{if } b\neq 0.\end{cases}\end{aligned}\tag{12}$$

It is invariant under the adjoint map $g\mapsto g^*$, by the properties of n and tr. We will denote by $\mathrm{SL}_2(\mathbb{H})$ the group of 2×2 matrices with coefficients in $\mathbb{H}$ with Dieudonné determinant 1, which equals the group of elements of (reduced) norm 1 in the central simple algebra $\mathcal{M}_2(\mathbb{H})$ over $\mathbb{R}$; see [Reiner 1975, Section 9a]. See [Kellerhals 2003] for more information on $\mathrm{SL}_2(\mathbb{H})$.

The group $\mathrm{SL}_2(\mathbb{H})$ acts linearly on the left on the right $\mathbb{H}$-module $\mathbb{H}\times\mathbb{H}$. Let $\mathbb{P}_r^1(\mathbb{H})=(\mathbb{H}\times\mathbb{H}-\{0\})/\mathbb{H}^\times$ be the right projective line of $\mathbb{H}$, identified as usual with the Alexandrov compactification $\mathbb{H}\cup\{\infty\}$ where $[1:0]=\infty$ and $[x:y]=xy^{-1}$ if $y\neq 0$. The projective action of $\mathrm{SL}_2(\mathbb{H})$ on $\mathbb{P}_r^1(\mathbb{H})$, induced by its linear action on $\mathbb{H}\times\mathbb{H}$, is then the action by homographies on $\mathbb{H}\cup\{\infty\}$ defined by

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\cdot z=\begin{cases}(az+b)(cz+d)^{-1}&\text{if } z\neq\infty,-c^{-1}d,\\ ac^{-1}&\text{if } z=\infty,\ c\neq 0,\\ \infty&\text{otherwise.}\end{cases}$$

This action induces a faithful left action of $\mathrm{PSL}_2(\mathbb{H})=\mathrm{SL}_2(\mathbb{H})/\{\pm\,\mathrm{id}\}$ on $\mathbb{H}\cup\{\infty\}$.

The group $\mathrm{PSL}_2(\mathbb{H})$ is very useful for studying 5-dimensional real hyperbolic geometry for the following reason. Let us endow $\mathbb{H}$ with its usual Euclidean metric $ds^2_{\mathbb{H}}$ (invariant under translations, with $(1,i,j,k)$ orthonormal). We will denote by $x=(z,r)$ a generic point in $\mathbb{H}\times\,]0,+\infty[$, and by $r:x\mapsto r(x)$ the second projection in this product. For the real hyperbolic space $\mathbb{H}^5_{\mathbb{R}}$ of dimension 5, we will use the upper halfspace model $\mathbb{H} \times]0, +\infty[$ with Riemannian metric $ds^2(x) = (ds^2_{\mathbb{H}}(z) + dr^2)/r^2$ at the point $x = (z, r)$, whose volume form is

$$d\,\mathrm{vol}_{\mathbb{H}^5_{\mathbb{R}}}(x) = \frac{d\,\mathrm{vol}_{\mathbb{H}}(z)\,dr}{r^5}. \tag{13}$$

The space at infinity $\partial_\infty \mathbb{H}^5_{\mathbb{R}}$ is hence $\mathbb{H} \cup \{\infty\}$.

By the Poincaré extension procedure (see for instance [Parkkonen and Paulin 2010, Lemma 6.6]), the action of $\mathrm{SL}_2(\mathbb{H})$ by homographies on $\partial_\infty \mathbb{H}^5_{\mathbb{R}}$ extends to a left action on $\mathbb{H}^5_{\mathbb{R}}$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot (z, r) = \left(\frac{(az+b)\overline{(cz+d)} + a\bar{c}r^2}{\mathrm{n}(cz+d) + r^2\,\mathrm{n}(c)}, \frac{r}{\mathrm{n}(cz+d) + r^2\,\mathrm{n}(c)} \right). \tag{14}$$

In this way, the group $\mathrm{PSL}_2(\mathbb{H})$ is identified with the group of orientation preserving isometries of $\mathbb{H}^5_{\mathbb{R}}$. Note that the isomorphism $\mathrm{PSL}_2(\mathbb{H}) \simeq \mathrm{SO}_0(1, 5)$ is one of the isomorphisms between connected simple real Lie groups of small dimensions in E. Cartan's classification.

Given an order $\mathbb{O}$ in a definite quaternion algebra A over $\mathbb{Q}$, define the *Hamilton–Bianchi group* as $\Gamma_{\mathbb{O}} = \mathrm{SL}_2(\mathbb{O}) = \mathrm{SL}_2(\mathbb{H}) \cap \mathcal{M}_2(\mathbb{O})$. Note that since the norm n takes integral values on $\mathbb{O}$, and since the Dieudonné determinant is a group morphism, we have $\mathrm{GL}_2(\mathbb{O}) = \mathrm{SL}_2(\mathbb{O})$. The Hamilton–Bianchi group $\Gamma_{\mathbb{O}}$ is a (nonuniform) arithmetic lattice in the connected real Lie group $\mathrm{SL}_2(\mathbb{H})$ (see for instance [Parkkonen and Paulin 2010, page 1104] for details). In particular, the quotient real hyperbolic orbifold $\Gamma_{\mathbb{O}} \backslash \mathbb{H}^5_{\mathbb{R}}$ has finite volume. The action by homographies of $\Gamma_{\mathbb{O}}$ preserves the right projective space $\mathbb{P}^1_r(\mathbb{O}) = A \cup \{\infty\}$, which is the set of fixed points of the parabolic elements of $\Gamma_{\mathbb{O}}$ acting on $\mathbb{H}^5_{\mathbb{R}} \cup \partial_\infty \mathbb{H}^5_{\mathbb{R}}$.

Remark 7. For every (u, v) in $\mathbb{O} \times \mathbb{O} - \{(0, 0)\}$, consider the two left ideals of $\mathbb{O}$

$$I_{u,v} = \mathbb{O}u + \mathbb{O}v \quad \text{and} \quad K_{u,v} = \begin{cases} \mathbb{O}u \cap \mathbb{O}v & \text{if } uv \neq 0, \\ \mathbb{O} & \text{otherwise.} \end{cases}$$

The map $\Gamma_{\mathbb{O}} \backslash \mathbb{P}^1_r(\mathbb{O}) \to ({}_{\mathbb{O}}\mathcal{I} \times {}_{\mathbb{O}}\mathcal{I})$ that associates to the orbit of $[u : v]$ in $\mathbb{P}^1_r(\mathbb{O})$ under $\Gamma_{\mathbb{O}}$ the couple of ideal classes $([I_{u,v}], [K_{u,v}])$ is a bijection. To see this, let $\ell_{u,v} : \mathbb{O} \times \mathbb{O} \to \mathbb{O}$ be the morphism of left $\mathbb{O}$-modules defined by $(o_1, o_2) \mapsto o_1 u + o_2 v$. The map $w \mapsto (wu^{-1}, -wv^{-1})$ is an isomorphism of left $\mathbb{O}$-modules from $\mathbb{O}u \cap \mathbb{O}v$ to the kernel of $\ell_{u,v}$ if $uv \neq 0$. The result then follows for instance from [Krafft and Osenberg 1990, Satz 2.1, 2.2], which says that the map $[u : v] \mapsto ([\mathrm{im}\,\ell_{u,v}], [\ker \ell_{u,v}])$ induces a bijection from $\Gamma_{\mathbb{O}} \backslash \mathbb{P}^1_r(\mathbb{O})$ into ${}_{\mathbb{O}}\mathcal{I} \times {}_{\mathbb{O}}\mathcal{I}$.

In particular, the number of cusps of $\Gamma_{\mathbb{O}}$ (or the number of ends of $\Gamma_{\mathbb{O}} \backslash \mathbb{H}^5_{\mathbb{R}}$) is the square of the class number h_A of A.

4. Background on binary Hamiltonian forms

With V the right $\mathbb{H}$-module $\mathbb{H}\times\mathbb{H}$, a *binary Hamiltonian form* $f : V \to \mathbb{R}$ is a map $X \mapsto \phi(X, X)$ where $\phi : V \times V \to \mathbb{H}$ is a Hermitian form on V with the conjugation as the anti-involution of the ring $\mathbb{H}$. That is, $\phi(X\lambda, Y) = \overline{\lambda}\phi(X, Y)$, $\phi(X + X', Y) = \phi(X, Y) + \phi(X', Y)$, $\phi(Y, X) = \overline{\phi(X, Y)}$ for $X, X', Y \in V$ and $\lambda \in \mathbb{H}$. Our convention of sesquilinearity on the left is the opposite of Bourbaki's unfortunate one in [Bourbaki 1959]. Equivalently, a binary Hamiltonian form f is a map $\mathbb{H}\times\mathbb{H} \to \mathbb{R}$ with

$$f(u, v) = a\,\mathrm{n}(u) + \mathrm{tr}(\overline{u}bv) + c\,\mathrm{n}(v),$$

whose *coefficients* $a = a(f)$ and $c = c(f)$ are real, and $b = b(f)$ lies in $\mathbb{H}$. Note that $f((u, v)\lambda) = \mathrm{n}(\lambda) f(u, v)$. The *matrix* $M(f)$ of f is the Hermitian matrix

$$\begin{pmatrix} a & b \\ \overline{b} & c \end{pmatrix},$$

so that

$$f(u, v) = \begin{pmatrix} u \\ v \end{pmatrix}^{*} \begin{pmatrix} a & b \\ \overline{b} & c \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

The *discriminant* of f is

$$\Delta = \Delta(f) = \mathrm{n}(b) - ac.$$

Note that the sign convention of the discriminant varies in the references. An easy computation shows that the Dieudonné determinant of $M(f)$ is equal to $|\Delta|$. If $a \neq 0$, then

$$f(u, v) = a\Big(\mathrm{n}\Big(u + \frac{bv}{a}\Big) - \frac{\Delta}{a^2}\,\mathrm{n}(v)\Big). \tag{15}$$

Hence the form f is *indefinite* (that is, f takes both positive and negative values) if and only if Δ is positive, and Δ is then equal to the Dieudonné determinant of $M(f)$. By (15), the form f is *positive definite* (that is, $f(x) \geq 0$ with equality if and only if $x = 0$) if and only if $a > 0$ and $\Delta < 0$.

The linear action on the left on $\mathbb{H}\times\mathbb{H}$ of the group $\mathrm{SL}_2(\mathbb{H})$ induces an action on the right on the set of binary Hermitian forms f by precomposition, that is, by $f \mapsto f \circ g$ for every $g \in \mathrm{SL}_2(\mathbb{H})$. The matrix of $f \circ g$ is $M(f \circ g) = g^* M(f) g$. Since the Dieudonné determinant is a group morphism, invariant under the adjoint map (and since $f \circ g$ is indefinite if and only if f is), we have, for every $g \in \mathrm{SL}_2(\mathbb{H})$,

$$\Delta(f \circ g) = \Delta(f). \tag{16}$$

Given an order $\mathbb{O}$ in a definite quaternion algebra over $\mathbb{Q}$, a binary Hamiltonian form f is *integral* over $\mathbb{O}$ if its coefficients belong to $\mathbb{O}$. Note that such a form f

takes integral values on $\mathbb{O} \times \mathbb{O}$. The lattice $\Gamma_{\mathbb{O}} = \mathrm{SL}_2(\mathbb{O})$ of $\mathrm{SL}_2(\mathbb{H})$ preserves the set of indefinite binary Hamiltonian forms f that are integral over $\mathbb{O}$. The stabilizer in $\Gamma_{\mathbb{O}}$ of such a form f is its *group of automorphs*

$$\mathrm{SU}_f(\mathbb{O}) = \{g \in \Gamma_{\mathbb{O}} : f \circ g = f\}.$$

For every indefinite binary Hamiltonian form f, with $a = a(f)$, $b = b(f)$ and $\Delta = \Delta(f)$, let

$$\mathscr{C}_\infty(f) = \{[u : v] \in \mathbb{P}^1_r(\mathbb{H}) : f(u, v) = 0\},$$
$$\mathscr{C}(f) = \{(z, r) \in \mathbb{H} \times \,]0, +\infty[\, : f(z, 1) + a\, r^2 = 0\}.$$

In $\mathbb{P}^1_r(\mathbb{H}) = \mathbb{H} \cup \{\infty\}$, the set $\mathscr{C}_\infty(f)$ is the 3-sphere of center $-b/a$ and radius $\sqrt{\Delta}/|a|$ if $a \neq 0$, and it is the union of $\{\infty\}$ with the real hyperplane

$$\{z \in \mathbb{H} : \mathrm{tr}(\bar{z} b) + c = 0\}$$

of $\mathbb{H}$ otherwise. The map $f \mapsto \mathscr{C}_\infty(f)$ induces a bijection between the set of indefinite binary Hamiltonian forms up to multiplication by a nonzero real factor and the set of 3-spheres and real hyperplanes in $\mathbb{H} \cup \{\infty\}$. The action of $\mathrm{SL}_2(\mathbb{H})$ by homographies on $\mathbb{H} \cup \{\infty\}$ preserves this set of 3-spheres and real hyperplanes, and the map $f \mapsto \mathscr{C}_\infty(f)$ is (anti)equivariant for the two actions of $\mathrm{SL}_2(\mathbb{H})$, in the sense that, for every $g \in \mathrm{SL}_2(\mathbb{H})$,

$$\mathscr{C}_\infty(f \circ g) = g^{-1}\, \mathscr{C}_\infty(f). \tag{17}$$

Given a finite index subgroup G of $\mathrm{SL}_2(\mathbb{O})$, an integral binary Hamiltonian form f is called *G-reciprocal* if there exists an element g in G such that $f \circ g = -f$. We define $R_G(f) = 2$ if f is G-reciprocal, and $R_G(f) = 1$ otherwise. The values of f are positive on one of the two components of $\mathbb{P}^1_r(\mathbb{H}) - \mathscr{C}_\infty(f)$ and negative on the other. As the signs are switched by precomposition by an element g as above, the G-reciprocity of the form f is equivalent to saying that there exists an element of G preserving $\mathscr{C}_\infty(f)$ and exchanging the two complementary components of $\mathscr{C}_\infty(f)$.

5. Using Eisenstein series to compute hyperbolic volumes

Let $\mathbb{O}$ be a maximal order in a definite quaternion algebra A over $\mathbb{Q}$.

In this section, we compute $\mathrm{Vol}(\mathrm{PSL}_2(\mathbb{O}) \backslash \mathbb{H}^5_{\mathbb{R}})$ using a method which goes back, in dimension 2, to Rankin and Selberg's method [Rankin 1939b; Selberg 1940] of integrating Eisenstein series on fundamental domains and "unfolding", generalized by [Langlands 1966] to the lattice of $\mathbb{Z}$-points of any connected split semisimple algebraic group over $\mathbb{Q}$. We follow the approach of [Sarnak 1983, pages 261–262] in dimension 3. See the appendix for a completely different proof of the same result by V. Emery.

Theorem 8. *Let $\mathbb{O}$ be a maximal order in a definite quaternion algebra A over $\mathbb{Q}$ with discriminant D_A. Then*

$$\operatorname{Vol}(\operatorname{PSL}_2(\mathbb{O})\backslash\mathbb{H}^5_{\mathbb{R}}) = \frac{\zeta(3)\prod_{p|D_A}(p^3-1)(p-1)}{11520}.$$

Proof. It is well known (see for instance [Parkkonen and Paulin 2010, Section 6.3, Example (3)]) that there exists $\underline{G}$, a connected semisimple linear algebraic group over $\mathbb{Q}$, such that $\underline{G}(\mathbb{R}) = \operatorname{SL}_2(\mathbb{H})$, $\underline{G}(\mathbb{Q}) = \operatorname{SL}_2(A)$ and $\underline{G}(\mathbb{Z}) = \operatorname{SL}_2(\mathbb{O})$. Let $\underline{P}$ be the parabolic subgroup of $\underline{G}$, defined over $\mathbb{Q}$, such that $\underline{P}(\mathbb{R})$ is the upper triangular subgroup of $\operatorname{SL}_2(\mathbb{H})$. By Borel's finiteness theorem [Borel 1966], the set $\operatorname{SL}_2(\mathbb{O})\backslash\operatorname{SL}_2(A)/\underline{P}(\mathbb{Q})$ is finite, and we will fix a subset $\mathcal{R}$ in $\operatorname{SL}_2(A)$ which is a system of representatives of this set of double cosets.

Let $\Gamma = \operatorname{SL}_2(\mathbb{O})$. For every $\alpha \in \mathcal{R}$, let $\Gamma_\alpha = \underline{P}(\mathbb{R}) \cap (\alpha^{-1}\Gamma\alpha)$ and let Γ'_α be its subgroup of unipotent elements. The group $\alpha\Gamma_\alpha\alpha^{-1}$ is the stabilizer of the parabolic fixed point $\alpha\infty$ in Γ. The action of Γ_α on $\mathbb{H}\cup\{\infty\}$ by homographies preserves ∞ and is cocompact on $\mathbb{H}$. If

$$\alpha^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix},$$

let $\mathfrak{u}_\alpha = c\mathbb{O} + d\mathbb{O}$, which is a right fractional ideal of $\mathbb{O}$, and $\mathfrak{v}_\alpha = \mathbb{O}\tilde{a} + \mathbb{O}\tilde{c}$, which is a left fractional ideal of $\mathbb{O}$.

For every $\alpha \in \mathcal{R}$, the *Eisenstein series* of the arithmetic group Γ for the cusp at infinity $\alpha\infty$ is the map $E_\alpha : \mathbb{H}^5_{\mathbb{R}} \times\,]4, +\infty[\, \to \mathbb{R}$ defined by

$$E_\alpha(x, s) = \sum_{\gamma\in(\alpha\Gamma_\alpha\alpha^{-1})\backslash\Gamma} r(\alpha^{-1}\gamma x)^s.$$

The summation does not depend on the choice of representatives of the left cosets in $(\alpha\Gamma_\alpha\alpha^{-1})\backslash\Gamma$ since Γ_α preserves ∞ and the Euclidean height r. The *Eisenstein series* of $\mathbb{O}$ is (for $x = (z, r) \in \mathbb{H}^5_{\mathbb{R}}$ and $s \in \mathbb{C}$ with $\operatorname{Re} s > 4$)

$$\hat{E}(x, s) = \sum_{(c,d)\in\mathbb{O}\times\mathbb{O}-\{0\}} \left(\frac{r}{\operatorname{n}(cz+d) + r^2\operatorname{n}(c)}\right)^s.$$

The next result concatenates results proven in [Krafft and Osenberg 1990].

Theorem 9 (Krafft and Osenberg). (i) *The Eisenstein series $E_\alpha(x, s)$ for $\alpha \in \mathcal{R}$ and $\hat{E}(x, s)$ converge absolutely and uniformly on compact subsets of* $\{s \in \mathbb{C} : \operatorname{Re} s > 4\}$, *uniformly on compact subsets of $x \in \mathbb{H}^5_{\mathbb{R}}$. They are invariant by the action of Γ on the first variable.*

(ii) *The map* $s \mapsto \hat{E}(x, s)$ *admits a meromorphic extension to* $\mathbb{C}$, *having only one pole*, *which is at* $s = 4$ *and is simple with residue*

$$\operatorname{Res}_{s=4} \hat{E}(x, s) = \frac{8\pi^4}{3\, {D_A}^2}\,. \tag{18}$$

Furthermore, *if* $c(\alpha, s) = \mathrm{n}(\mathrm{u}_\alpha)^s \zeta(\mathrm{u}_\alpha^{-1}, s/2)$ *for every* $\alpha \in \mathscr{R}$, *then*

$$\hat{E}(x, s) = \sum_{\alpha \in \mathscr{R}} c(\alpha, s) E_\alpha(x, s). \tag{19}$$

(iii) *For every* $\alpha, \beta \in \mathscr{R}$, *there exists a map* $s \mapsto \varphi_{\alpha,\beta}(s)$ *with* $(s-4)\varphi_{\alpha,\beta}(s)$ *bounded for* $s > 4$ *near* $s = 4$, *and a measurable map* $(x, s) \mapsto \Phi_{\alpha,\beta}(x, s)$ *such that* $(s-4)\Phi_{\alpha,\beta}(x, s)$ *is bounded by an integrable* (*for the hyperbolic volume*) *map*, *independent on* $s > 4$ *near* $s = 4$, *on* $x \in K \times [\epsilon, +\infty[$ *where* K *is a compact subset of* $\mathbb{H}$ *and* $\epsilon > 0$, *such that*

$$E_\alpha(\beta x, s) = \delta_{\alpha,\beta}\, r^s + \varphi_{\alpha,\beta}(s) r^{4-s} + \Phi_{\alpha,\beta}(x, s),$$

with $\delta_{\alpha,\beta} = 1$ *if* $\alpha = \beta$ *and* $\delta_{\alpha,\beta} = 0$ *otherwise.*

Proof. We are using Langlands' convention for the Eisenstein series; hence with Γ'_α the subgroup of unipotent elements of Γ_α, our Eisenstein series E_α is obtained from the one used in [Krafft and Osenberg 1990] by replacing α by α^{-1} and by multiplying by $1/[\Gamma_\alpha : \Gamma'_\alpha]$.

The part of claim (i) concerning the series $E_\alpha(x, s)$ for $\alpha \in \mathscr{R}$ is [Krafft and Osenberg 1990, Satz 3.2]. The rest follows from [ibid., Satz 4.2] with $M = \mathbb{O}$. The claim (ii) follows from [ibid., Korollar 5.6(a)] with $M = \mathbb{O}$, recalling that the reduced discriminant of any maximal order of A is equal to the reduced discriminant of A. The formula (19) follows from [ibid., Satz 4.3], recalling the above changes between our E_α and the one in [ibid.]. The claim (iii) follows from [ibid., Satz 3.3], again replacing β by β^{-1}, and using the second equation in [Magnus et al. 1966, page 85] to control the modified Bessel function. □

By a *fundamental domain* for a smooth action of a countable group G on a smooth manifold N, we mean a subset F of N such that F has negligible boundary, the interiors of the subsets gF for $g \in G$ are pairwise disjoint, and

$$N = \bigcup_{g \in G} gF.$$

Here is a construction of a fundamental domain $\mathscr{F}$ for Γ acting on $\mathbb{H}^5_{\mathbb{R}}$ that will be useful in this section (and is valid for any discrete subgroup of isometries of $\mathbb{H}^n_{\mathbb{R}}$ with finite covolume which is not cocompact). Let $\mathscr{P}$ be the set of parabolic fixed points of Γ. By the structure of the cusp neighborhoods, there exists a family $(\mathscr{H}_p)_{p \in \mathscr{P}}$ of pairwise disjoint closed horoballs in $\mathbb{H}^5_{\mathbb{R}}$, equivariant under Γ (that is,

$\gamma\mathscr{H}_p = \mathscr{H}_{\gamma p}$ for every $\gamma \in \Gamma$), with $\mathscr{H}_p$ centered at p. The *cut locus of the cusps* Σ is the piecewise hyperbolic polyhedral complex in $\mathbb{H}^5_{\mathbb{R}}$ consisting of the set of points outside the union of these horoballs that are equidistant to at least two of these horoballs (it is independent of the choice of this family when there is only one orbit of parabolic fixed points). Each connected component of the complement of Σ contains one and only one of these horoballs, is at bounded Hausdorff distance of it, is invariant under the stabilizer in Γ of its point at infinity, and is precisely invariant under the action of Γ. Recall that a subset A of a set endowed with an action of a group G is said to be *precisely invariant* under this group if for every $g \in G$, if $gA \cap A$ is nonempty, then $gA = A$.

For every $\beta \in \mathscr{R}$, let $\mathscr{D}_\beta$ be a compact fundamental domain for the action of Γ_β on $\mathbb{H}$, let $\widetilde{\mathscr{F}}_\beta$ be the closure of the component of the complement of Σ containing $\mathscr{H}_{\beta\infty}$, and define $\mathscr{F}_\beta = \widetilde{\mathscr{F}}_\beta \cap \beta(\mathscr{D}_\beta \times]0, +\infty[)$. Then $\mathscr{F}_\beta$ is a closed fundamental domain for the action of $\beta\Gamma_\beta\beta^{-1}$ on $\widetilde{\mathscr{F}}_\beta$, and there exists a continuous map $\sigma'_\beta : \mathscr{D}_\beta \to]0, +\infty[$, which hence has a positive lower bound, such that

$$\beta^{-1}\mathscr{F}_\beta = \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : z \in \mathscr{D}_\beta,\ r \geq \sigma'_\beta(z)\}. \tag{20}$$

Now define

$$\mathscr{F} = \bigcup_{\beta \in \mathscr{R}} \mathscr{F}_\beta. \tag{21}$$

Since $\mathscr{R}$ is a system of representatives of the cusps, $\mathscr{F}$ is a fundamental domain of Γ acting on $\mathbb{H}^5_{\mathbb{R}}$.

Note that, for every $\alpha \in \mathscr{R}$, there exists a continuous map $\sigma_\alpha : \mathscr{D}_\alpha \to [0, +\infty[$ (hence with a finite upper bound), with only finitely many zeros, such that, since $\alpha^{-1}\mathscr{F}$ is a fundamental domain for the action of $\alpha^{-1}\Gamma\alpha$ on $\mathbb{H}^5_{\mathbb{R}}$,

$$\bigcup_{\gamma \in (\alpha^{-1}\Gamma\alpha - \Gamma_\alpha)} \gamma\alpha^{-1}\mathscr{F} = \Gamma_\alpha\{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : z \in \mathscr{D}_\alpha, r < \sigma_\alpha(z)\}. \tag{22}$$

For every $\alpha \in \mathscr{R}$, let

$$b_\alpha(s) = \int_{\mathscr{F}} \big(E_\alpha(x, s) - r(\alpha^{-1}x)^s\big)\, d\,\mathrm{vol}_{\mathbb{H}^5_{\mathbb{R}}}(x).$$

When $s > 4$, we have

$$\begin{aligned} b_\alpha(s) &= \int_{\mathscr{F}} \Bigg(\sum_{\gamma \in (\alpha\Gamma_\alpha\alpha^{-1})\backslash\Gamma} r(\alpha^{-1}\gamma x)^s - r(\alpha^{-1}x)^s\Bigg)\, d\,\mathrm{vol}_{\mathbb{H}^5_{\mathbb{R}}}(x) \\ &= \int_{\mathscr{F}} \sum_{\substack{\gamma \in \\ \Gamma_\alpha\backslash(\alpha^{-1}\Gamma\alpha - \Gamma_\alpha)}} r(\gamma\alpha^{-1}x)^s\, d\,\mathrm{vol}_{\mathbb{H}^5_{\mathbb{R}}}(x) = \sum_{\substack{\gamma \in \\ \Gamma_\alpha\backslash(\alpha^{-1}\Gamma\alpha - \Gamma_\alpha)}} \int_{\mathscr{F}} r(\gamma\alpha^{-1}x)^s\, d\,\mathrm{vol}_{\mathbb{H}^5_{\mathbb{R}}}(x) \end{aligned}$$

$$= \sum_{\substack{\gamma\in\\ \Gamma_\alpha\backslash(\alpha^{-1}\Gamma\alpha-\Gamma_\alpha)}} \int_{\gamma\alpha^{-1}\mathscr{F}} r(x)^s\, d\,\mathrm{vol}_{\mathbb{H}^5_\mathbb{R}}(x) = \int_{\bigcup_{\gamma\in\Gamma_\alpha\backslash(\alpha^{-1}\Gamma\alpha-\Gamma_\alpha)}\gamma\alpha^{-1}\mathscr{F}} r(x)^s\, d\,\mathrm{vol}_{\mathbb{H}^5_\mathbb{R}}(x)$$
$$= \int_{z\in\mathscr{D}_\alpha}\int_0^{\sigma_\alpha(z)} r^{s-5}\,dr\,dz = \int_{z\in\mathscr{D}_\alpha} \frac{\sigma_\alpha(z)^{s-4}}{s-4}\,dz,$$

using for the succession of equations, respectively, the definition of E_α, the change of variables $\alpha^{-1}\gamma\alpha \to \gamma$, Fubini's theorem for positive functions, the invariance of the volume under the isometric change of variables $\gamma\alpha^{-1}x \to x$, the σ-additivity property, and the equations (22) and (13) and the invariance of the Euclidean height function r under Γ_α.

For any $\alpha \in \mathscr{R}$, the map ${\sigma_\alpha}^{s-4}$ converges pointwise, as $s \to 4^+$, to the map on $\mathscr{D}_\alpha$ with value 0 at the finitely many points where σ_α vanishes, and with value 1 otherwise. Since $\mathscr{D}_\alpha$ is compact and ${\sigma_\alpha}^{s-4}$ is uniformly bounded from above, Lebesgue's dominated convergence theorem gives

$$\lim_{s\to 4^+}(s-4)b_\alpha(s) = \mathrm{Vol}(\mathscr{D}_\alpha) = \mathrm{Vol}(\Gamma_\alpha\backslash\mathbb{H}).$$

Therefore by using (19), the map

$$s \mapsto b(s) = \int_{\mathscr{F}} \big(\hat{E}(x,s) - \sum_{\alpha\in\mathscr{R}} c(\alpha,s) r(\alpha^{-1}x)^s\big)\, d\,\mathrm{vol}_{\mathbb{H}^5_\mathbb{R}}(x) = \sum_{\alpha\in\mathscr{R}} c(\alpha,s) b_\alpha(s)$$

satisfies

$$\lim_{s\to 4^+}(s-4)b(s) = \sum_{\alpha\in\mathscr{R}} c(\alpha,4)\,\mathrm{Vol}(\Gamma_\alpha\backslash\mathbb{H}), \tag{23}$$

since $s \mapsto c(\alpha,s)$ is holomorphic for $\mathrm{Re}\, s > 2$.

On the other hand, let us prove that we may permute the limit as $s \to 4^+$ and the integral defining $(s-4)b(s)$. Using the equations (19) and (21), and an isometric, hence volume-preserving, change of variable, we have

$$b(s) = \sum_{\alpha,\beta\in\mathscr{R}} c(\alpha,s)\int_{\mathscr{F}_\beta}\big(E_\alpha(x,s) - r(\alpha^{-1}x)^s\big)\, d\,\mathrm{vol}_{\mathbb{H}^5_\mathbb{R}}(x)$$
$$= \sum_{\alpha,\beta\in\mathscr{R}} c(\alpha,s)\int_{\beta^{-1}\mathscr{F}_\beta}\big(E_\alpha(\beta x,s) - r(\alpha^{-1}\beta x)^s\big)\, d\,\mathrm{vol}_{\mathbb{H}^5_\mathbb{R}}(x).$$

If $x \in \beta^{-1}\mathscr{F}_\beta$, then $r(x)$ is bounded from below by a positive constant by the construction of $\mathscr{F}_\beta$; hence $r(x)^{4-s}$ is bounded from above for every $s \geq 4$. If $\alpha \neq \beta$ and $x \in \beta^{-1}\mathscr{F}_\beta$, then $r(\alpha^{-1}\beta x)^s$ is bounded from above for every $s \geq 0$, since $\alpha^{-1}\mathscr{F}_\beta$ is bounded in $\mathbb{H}\times\mathbb{R}$ by construction. Hence since $\beta^{-1}\mathscr{F}_\beta$ has finite hyperbolic volume, by Theorem 9(iii) separating the case $\alpha = \beta$ and the case $\alpha \neq \beta$, by Lebesgue's dominated convergence theorem, we may permute the limit as $s \to 4^+$ and the integral on $\beta^{-1}\mathscr{F}_\beta$ for the hyperbolic volume applied to $(s-4)\big(E_\alpha(\beta x, s) - r(\alpha^{-1}\beta x)^s\big)$. By a finite summation, we may indeed permute the limit as $s \to 4^+$ and the integral defining $(s-4)b(s)$.

Therefore, by (18),

$$\lim_{s\to 4^+}(s-4)b(s) = \frac{8\pi^4}{3{D_A}^2}\,\mathrm{Vol}(\mathrm{PSL}_2(\mathbb{O})\backslash \mathbb{H}^5_{\mathbb{R}}). \tag{24}$$

Finally, since for every $\rho \in A - \{0\}$ the element

$$\gamma_\rho = \begin{pmatrix} \rho & -1 \\ 1 & 0 \end{pmatrix}$$

of $\mathrm{SL}_2(A)$ maps ∞ to ρ, the element $\alpha \in \mathscr{R}$ may be chosen to be either id or γ_ρ for some $\rho \in A$. In the first case, $\mathfrak{u}_\alpha = \mathbb{O}$ and Γ'_α acts on $\mathbb{H}$ as the $\mathbb{Z}$-lattice $\mathbb{O}$, so that, by (5), since the subgroup $\{\pm\,\mathrm{id}\}$ of Γ_α is the kernel of its action on $\mathbb{H}$,

$$\mathrm{n}(\mathfrak{u}_\alpha)^4\,\mathrm{Vol}(\Gamma_\alpha\backslash\mathbb{H}) = \frac{2\,\mathrm{Vol}(\mathbb{O}\backslash\mathbb{H})}{[\Gamma_\alpha : \Gamma'_\alpha]} = \frac{D_A}{2[\Gamma_\alpha : \Gamma'_\alpha]}\,.$$

In the second case,

$$\alpha^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & \rho \end{pmatrix},$$

so that $\mathfrak{u}_\alpha = \mathbb{O}\rho + \mathbb{O}$ and Γ'_α acts on $\mathbb{H}$ as the $\mathbb{Z}$-lattice $\Lambda = \mathbb{O}\cap\rho^{-1}\mathbb{O}\cap\mathbb{O}\rho^{-1}\cap\rho^{-1}\mathbb{O}\rho^{-1}$ as we shall see in Lemma 15. By Lemma 6 applied with $z = \rho^{-1}$ and by (5), we hence have

$$\mathrm{n}(\mathfrak{u}_\alpha)^4\,\mathrm{Vol}(\Gamma_\alpha\backslash\mathbb{H}) = \mathrm{n}(\mathfrak{u}_\alpha)^4[\mathbb{O}:\Lambda]\,\frac{2\,\mathrm{Vol}(\mathbb{O}\backslash\mathbb{H})}{[\Gamma_\alpha : \Gamma'_\alpha]} = \frac{D_A}{2[\Gamma_\alpha : \Gamma'_\alpha]}\,.$$

Therefore, by the definition of $c(\alpha, s)$,

$$\sum_{\alpha\in\mathscr{R}} c(\alpha, 4)\,\mathrm{Vol}(\Gamma_\alpha\backslash\mathbb{H}) = \frac{D_A}{2}\sum_{\alpha\in\mathscr{R}}\frac{\zeta(\mathfrak{u}_\alpha^{-1}, 2)}{[\Gamma_\alpha : \Gamma'_\alpha]}\,. \tag{25}$$

Combining the equations (23), (24) and (25), we have

$$\mathrm{Vol}(\mathrm{PSL}_2(\mathbb{O})\backslash \mathbb{H}^5_{\mathbb{R}}) = \frac{3{D_A}^3}{16\,\pi^4}\sum_{\alpha\in\mathscr{R}}\frac{\zeta(\mathfrak{u}_\alpha^{-1}, 2)}{[\Gamma_\alpha : \Gamma'_\alpha]}\,. \tag{26}$$

Lemma 10. (1) *For every $\alpha \in \mathscr{R}$, we have $[\Gamma_\alpha : \Gamma'_\alpha] = |\mathbb{O}_r(\mathfrak{u}_\alpha^{-1})^\times||\mathbb{O}_r(\mathfrak{v}_\alpha)^\times|$.*

(2) *The map from $\mathscr{R}$ to ${}_{\mathbb{O}}\mathscr{I} \times {}_{\mathbb{O}}\mathscr{I}$ defined by $\alpha \mapsto ([\mathfrak{v}_\alpha], [\mathfrak{u}_\alpha^{-1}])$ is a bijection.*

Proof. (1) Let

$$\Gamma^+_\alpha = \big\{\gamma \in \Gamma_\alpha \;:\; (0\;1)\gamma = (0\;1)\big\} \quad\text{and}\quad \Gamma^-_\alpha = \big\{\gamma \in \Gamma_\alpha : \gamma\tbinom{1}{0} = \tbinom{1}{0}\big\},$$

which are normal subgroups of Γ_α, whose union generates Γ_α. By the top of page 434 in [Krafft and Osenberg 1990] (keeping in mind that our α is the inverse of the α in [ibid.]), we have $[\Gamma_\alpha^+ : \Gamma'_\alpha] = |\mathbb{O}_r(\mathfrak{v}_\alpha)^\times|$. Similarly, $[\Gamma_\alpha^- : \Gamma'_\alpha] = |\mathbb{O}_\ell(\mathfrak{u}_\alpha)^\times|$. Note that Γ'_α is a normal subgroup of Γ_α^-, Γ_α^+ and Γ_α, such that $\Gamma_\alpha^- \cap \Gamma_\alpha^+ = \Gamma'_\alpha$. Hence the product map from $\Gamma_\alpha^- \times \Gamma_\alpha^+$ to Γ'_α induces a bijection from $(\Gamma_\alpha^- / \Gamma'_\alpha) \times (\Gamma'_\alpha \backslash \Gamma_\alpha^+)$ to $\Gamma_\alpha / \Gamma'_\alpha$, since $\Gamma_\alpha / \Gamma'_\alpha$ is abelian. In particular, $[\Gamma_\alpha : \Gamma'_\alpha] = |\mathbb{O}_\ell(\mathfrak{u}_\alpha)^\times||\mathbb{O}_r(\mathfrak{v}_\alpha)^\times|$. Using (4), the result follows.

(2) Since these matrices act transitively on A by homographies, we may assume that every $\alpha \in \mathcal{R}$ either is the identity element id, or has the form

$$\begin{pmatrix} \rho_\alpha & -1 \\ 1 & 0 \end{pmatrix}$$

for some $\rho_\alpha \in A^\times$. Then α^{-1} is either id or

$$\begin{pmatrix} 0 & 1 \\ -1 & \rho_\alpha \end{pmatrix}.$$

Hence, $\mathfrak{u}_\alpha = \mathbb{O} + \rho_\alpha \mathbb{O}$ and $\mathfrak{v}_\alpha = \mathbb{O}\rho_\alpha + \mathbb{O}$, unless $\alpha = \mathrm{id}$, in which case $\mathfrak{u}_\alpha = \mathfrak{v}_\alpha = \mathbb{O}$. Since $\mathrm{SL}_2(A)$ acts (on the left) transitively by homographies on $\mathbb{P}^1_r(\mathbb{O})$ with stabilizer of $[1:0]$ equal to $\underline{P}(\mathbb{Q})$, the map from $\mathcal{R}$ to $\Gamma_{\mathbb{O}} \backslash \mathbb{P}^1_r(\mathbb{O})$ defined by $\alpha \mapsto \Gamma_{\mathbb{O}}\alpha[1:0]$ is a bijection. Note that $\alpha[1:0] = [\rho_\alpha : 1]$ if $\alpha \neq \mathrm{id}$. Using the notation of Remark 7, if $\alpha \neq \mathrm{id}$, we have $\mathfrak{v}_\alpha = I_{\rho_\alpha,1}$ and

$$[K_{\rho_\alpha,1}] = [\mathbb{O}\rho_\alpha \cap \mathbb{O}] = [\mathbb{O} \cap \mathbb{O}\rho_\alpha^{-1}] = [\mathfrak{u}_\alpha^{-1}]$$

by (3). The second assertion of this lemma then follows from Remark 7. □

Now, using respectively (9), Lemma 10(1), Lemma 10(2), the separation of variables and (7), (8), and (6) since $\zeta(4) = \pi^4/90$, we have

$$\begin{aligned}\sum_{\alpha\in\mathcal{R}} \frac{\zeta(\mathfrak{u}_\alpha^{-1},2)}{[\Gamma_\alpha : \Gamma'_\alpha]} &= \sum_{\alpha\in\mathcal{R}} \frac{|\mathbb{O}_r(\mathfrak{u}_\alpha^{-1})^\times|\,\zeta_{[\mathfrak{u}_\alpha^{-1}]}(2)}{[\Gamma_\alpha : \Gamma'_\alpha]} = \sum_{\alpha\in\mathcal{R}} \frac{\zeta_{[\mathfrak{u}_\alpha^{-1}]}(2)}{|\mathbb{O}_r(\mathfrak{v}_\alpha)^\times|} \\ &= \sum_{([I],[J])\in {}_{\mathbb{O}}\mathcal{I}\times {}_{\mathbb{O}}\mathcal{I}} \frac{\zeta_{[J]}(2)}{|\mathbb{O}_r(I)^\times|} = \zeta_A(2) \sum_{[I]\,\in\, {}_{\mathbb{O}}\mathcal{I}} \frac{1}{|\mathbb{O}_r(I)^\times|} \\ &= \frac{\zeta_A(2)}{24} \prod_{p|D_A} (p-1) = \frac{\zeta(3)\pi^4 \prod_{p|D_A}(1-p^{-3})(p-1)}{2160}. \end{aligned} \tag{27}$$

Theorem 8 follows from the equations (26) and (27). □

Corollary 11. *Let A be a definite quaternion algebra over $\mathbb{Q}$ with reduced discriminant D_A and class number* 1, *and let $\mathbb{O}$ be a maximal order in A. Then the* *hyperbolic volume of* $\mathrm{PSL}_2(\mathcal{O})\backslash\mathbb{H}^5_{\mathbb{R}}$ *is equal to*

$$\mathrm{Vol}(\mathrm{PSL}_2(\mathcal{O})\backslash\mathbb{H}^5_{\mathbb{R}}) = \frac{(D_A^3-1)(D_A-1)\zeta(3)}{11520}\,.$$

This is an immediate consequence of Theorem 8. But here is a proof directly from (26) that avoids using the technical Lemma 10 and the technical computation (27).

Proof. Since the number of cusps of $\mathrm{SL}_2(\mathcal{O})$ is the square of the class number h_A of A (see Remark 7), the set $\mathscr{R}$ has only one element, and we may choose $\mathscr{R} = \{\mathrm{id}\}$.

By definition of the Dieudonné determinant and since every element of $\mathcal{O}^\times$ has norm 1, the stabilizer Γ_∞ of ∞ in $\mathrm{SL}_2(\mathcal{O})$ is

$$\Gamma_\infty = \left\{\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} : a, d \in \mathcal{O}^\times, b \in \mathcal{O}\right\}.$$

The index in Γ_∞ of its unipotent subgroup is hence $|\mathcal{O}^\times|^2$. By the equations (10) and (6), Corollary 11 follows from (26), $\zeta(4) = \pi^4/90$, $|\mathcal{O}^\times| = 24/(D_A - 1)$ as seen in (2), and since D_A is prime when $h_A = 1$. □

Example 12. Let A be Hamilton's quaternion algebra over $\mathbb{Q}$, which satisfies $D_A = 2$ and $h_A = 1$. Let $\mathcal{O}$ be Hurwitz's maximal order in A. Applying Corollary 11, we get

$$\mathrm{Vol}(\mathrm{PSL}_2(\mathcal{O})\backslash\mathbb{H}^5_{\mathbb{R}}) = \frac{7\zeta(3)}{11520},$$

exactly four times the minimal volume of a cusped hyperbolic 5-orbifold, as we should because the Hurwitz modular group $\mathrm{PSL}_2(\mathcal{O})$ is a subgroup of index 4 in the group of the minimal volume cusped hyperbolic orbifold of dimension 5; see [Hild 2007, page 209; Johnson and Weiss 1999, page 186].

6. Representing integers by binary Hamiltonian forms

Let A be a definite quaternion algebra over $\mathbb{Q}$, and let $\mathcal{O}$ be a maximal order in A.

Let us introduce the general counting function we will study. For every indefinite integral binary Hamiltonian form f over $\mathcal{O}$, for every finite index subgroup G of $\mathrm{SL}_2(\mathcal{O})$, for every x, y in $\mathcal{O}$ not both zero, and for every $s > 0$, let

$$\psi_{f,G,x,y}(s) = \mathrm{Card}_{\,\mathrm{SU}_f(\mathcal{O})\cap G}\backslash \left\{(u, v) \in G(x, y) : \mathrm{n}(\mathcal{O}x + \mathcal{O}y)^{-1}|f(u, v)| \le s\right\}.$$

The counting function $\psi_{f,G,x,y}$ depends (besides on f, G) only on the G-orbit of $[x : y]$ in $\mathbb{P}^1_r(\mathcal{O})$.

Here is the notation for the statement of our main result which follows. Given $(x, y) \in \mathcal{O} \times \mathcal{O}$, let $\Gamma_{\mathcal{O},x,y}$ and $G_{x,y}$ be the stabilizers of (x, y) for the left linear actions of $\Gamma_{\mathcal{O}} = \mathrm{SL}_2(\mathcal{O})$ and G, respectively, and let $\mathfrak{u}_{xy^{-1}}$ be the right fractional ideal

$\mathbb{O}$ if $y=0$ and $\mathbb{O}+xy^{-1}\mathbb{O}$ otherwise. Let $\iota_G=1$ if $-\mathrm{id}\in G$, and $\iota_G=2$ otherwise. Note that the image of $\mathrm{SU}_f(\mathbb{O})\cap G$ in $\mathrm{PSL}_2(\mathbb{H})$ is again an arithmetic group.

Theorem 13. *Let f be an integral indefinite binary Hamiltonian form of discriminant $\Delta(f)$ over a maximal order $\mathbb{O}$ of a definite quaternion algebra A over $\mathbb{Q}$. Let x and y be elements in $\mathbb{O}$ not both zero, and let G be a finite index subgroup of $\Gamma_{\mathbb{O}}=\mathrm{SL}_2(\mathbb{O})$. Then, as s tends to $+\infty$, we have the equivalence*

$$\psi_{f,G,x,y}(s)\sim\frac{540\,\iota_G[\Gamma_{\mathbb{O},x,y}:G_{x,y}]\,\mathrm{Covol}(\mathrm{SU}_f(\mathbb{O})\cap G)}{\pi^2\zeta(3)|\mathbb{O}_\ell(\mathfrak{u}_{xy^{-1}})^\times|\Delta(f)^2[\Gamma_{\mathbb{O}}:G]\prod_{p|D_A}(p^3-1)(1-p^{-1})}s^4.$$

Proof. Let us first recall a geometric result from [Parkkonen and Paulin 2012] that will be used to prove this theorem.

Let $n\geq 2$ and let $\mathbb{H}^n_{\mathbb{R}}$ be the upper halfspace model of the real hyperbolic space of dimension n, with (constant) sectional curvature -1. Let F be a finite covolume discrete group of isometries of $\mathbb{H}^n_{\mathbb{R}}$. Let $1\leq k\leq n-1$ and let $\mathscr{C}$ be a real hyperbolic subspace of dimension k of $\mathbb{H}^n_{\mathbb{R}}$, whose stabilizer $F_{\mathscr{C}}$ in F has finite covolume. Let $\mathscr{H}$ be a horoball in $\mathbb{H}^n_{\mathbb{R}}$, which is precisely invariant under F, with stabilizer $F_{\mathscr{H}}$.

For every $\alpha,\beta\in F$, denote by $\delta_{\alpha,\beta}$ the common perpendicular geodesic arc between $\alpha\mathscr{C}$ and the horosphere $\beta\partial\mathscr{H}$ if it exists, and let $\ell(\delta_{\alpha,\beta})$ be its length, counted positively if $\delta_{\alpha,\beta}$ exits $\beta\mathscr{H}$ at its endpoint on $\beta\partial\mathscr{H}$, and negatively otherwise. Also define the multiplicity of $\delta_{\alpha,\beta}$ as $m(\alpha,\beta)=1/\mathrm{Card}(\alpha\,F_{\mathscr{C}}\alpha^{-1}\cap\beta\,F_{\mathscr{H}}\beta^{-1})$. Its denominator is finite, if the boundary at infinity of $\alpha\mathscr{C}$ does not contain the point at infinity of $\beta\mathscr{H}$, since then the subgroup $\alpha\,F_{\mathscr{C}}\alpha^{-1}\cap\beta\,F_{\mathscr{H}}\beta^{-1}$ that preserves both $\beta\mathscr{H}$ and $\alpha\mathscr{C}$ consists of elliptic elements. By convention, $\ell(\delta_{\alpha,\beta})=-\infty$ and $m(\alpha,\beta)=0$ if the boundary at infinity of $\alpha\mathscr{C}$ contains the point at infinity of $\beta\mathscr{H}$. In particular, there are only finitely many elements $[g]\in F_{\mathscr{C}}\backslash F/F_{\mathscr{H}}$ such that $m(g^{-1},\mathrm{id})$ is different from 1, or equivalently such that $g^{-1}F_{\mathscr{C}}g\cap F_{\mathscr{H}}\neq\{1\}$. For every $t\geq 0$, define $\mathscr{N}(t)=\mathscr{N}_{F,\mathscr{C},\mathscr{H}}(t)$ as the number, counted with multiplicity, of the orbits under F in the set of the common perpendicular arcs $\delta_{\alpha,\beta}$ for $\alpha,\beta\in F$ with length at most t:

$$\mathscr{N}(t)=\mathscr{N}_{F,\mathscr{C},\mathscr{H}}(t)=\sum_{\substack{(\alpha,\beta)\in F\backslash((F/F_{\mathscr{C}})\times(F/F_{\mathscr{H}}))\\ \ell(\delta_{\alpha,\beta})\leq t}}m(\alpha,\beta).$$

For every $m\in\mathbb{N}$, denoting by $\mathbb{S}_m$ the unit sphere of the Euclidean space $\mathbb{R}^{m+1}$ endowed with its induced Riemannian metric, we have the following result:

Theorem 14 [Parkkonen and Paulin 2012, Corollary 4.9]. *As $t\to+\infty$, we have*

$$\mathscr{N}(t)\sim\frac{\mathrm{Vol}(\mathbb{S}_{n-k-1})\,\mathrm{Vol}(F_{\mathscr{H}}\backslash\mathscr{H})\,\mathrm{Vol}(F_{\mathscr{C}}\backslash\mathscr{C})}{\mathrm{Vol}(\mathbb{S}_{n-1})\,\mathrm{Vol}(F\backslash\mathbb{H}^n_{\mathbb{R}})}e^{(n-1)t}.$$

Now, let A, $\mathcal{O}$, f, G, x and y be as in the statement of Theorem 13. We write f as in (1), and denote its discriminant by Δ. In order to apply Theorem 14, we first define the various objects n, k, F, $\mathscr{H}$, and $\mathscr{C}$ that appear in its statement.

Let $n = 5$ and $k = 4$, so that $\mathrm{Vol}(\mathbb{S}_{n-1}) = 8\pi^2/3$ and $\mathrm{Vol}(\mathbb{S}_{n-k-1}) = 2$. We use the description of $\mathbb{H}^5_{\mathbb{R}}$ given in Section 3.

For any subgroup S of $\mathrm{SL}_2(\mathbb{H})$, we denote by $\bar{S}$ its image in $\mathrm{PSL}_2(\mathbb{H})$, except that the image of $\mathrm{SU}_f(\mathcal{O})$ is denoted by $\mathrm{PSU}_f(\mathcal{O})$. We will apply Theorem 14 to $F = \overline{G}$.

Note that $\mathrm{Vol}(\overline{G} \setminus \mathbb{H}^5_{\mathbb{R}}) = [\overline{\Gamma_{\mathcal{O}}} : \overline{G}]\,\mathrm{Vol}(\overline{\Gamma_{\mathcal{O}}} \setminus \mathbb{H}^5_{\mathbb{R}})$ and $[\overline{\Gamma_{\mathcal{O}}} : \overline{G}] = (1/\iota_G)[\Gamma_{\mathcal{O}} : G]$ by the definition of ι_G. Thus, using Theorem 8 (or Theorem A.1), we have

$$\begin{aligned}\mathrm{Vol}(\overline{G}\backslash\mathbb{H}^5_{\mathbb{R}}) &= \frac{1}{\iota_G}[\Gamma_{\mathcal{O}} : G]\,\mathrm{Vol}(\overline{\Gamma_{\mathcal{O}}} \setminus \mathbb{H}^5_{\mathbb{R}}) = \frac{1}{\iota_G}[\Gamma_{\mathcal{O}} : G]\,\mathrm{Covol}(\Gamma_{\mathcal{O}}) \\ &= \frac{\zeta(3)[\Gamma_{\mathcal{O}} : G]}{11520\,\iota_G} \prod_{p|D_A} (p^3-1)(p-1). \end{aligned} \tag{28}$$

The point $\rho = xy^{-1} \in A \cup \{\infty\} \subset \partial_\infty \mathbb{H}^5_{\mathbb{R}}$ is a parabolic fixed point of $\overline{\Gamma_{\mathcal{O}}}$ and hence of $\overline{G}$. Let $\tau \in\,]0, 1]$ and $\mathscr{H}$ be the horoball in $\mathbb{H}^5_{\mathbb{R}}$ centered at ρ, with Euclidean height τ if $y \neq 0$, and consisting of the points of Euclidean height at least $1/\tau$ otherwise. Assume that τ is small enough so that $\mathscr{H}$ is precisely invariant under $\overline{\Gamma_{\mathcal{O}}}$ and hence under $\overline{G}$. Such a τ exists, as seen in the construction of the fundamental domain in Section 5. The stabilizer $\overline{\Gamma_{\mathcal{O},\rho}}$ in $\overline{\Gamma_{\mathcal{O}}}$ of the point at infinity ρ is equal to the stabilizer $(\overline{\Gamma_{\mathcal{O}}})_{\mathscr{H}}$ of the horoball $\mathscr{H}$.

Remark. If $\rho = \infty$ and $G = \Gamma_{\mathcal{O}}$, we may take $\tau = 1$ by [Kellerhals 2003, Proposition 5]. Then by an easy hyperbolic geometry computation, since the index in $(\overline{\Gamma_{\mathcal{O}}})_{\mathscr{H}}$ of the subgroup of translations by elements of $\mathcal{O}$ is $\frac{|\mathcal{O}^\times|^2}{2}$, and by using (5), we have

$$\mathrm{Vol}((\overline{\Gamma_{\mathcal{O}}})_{\mathscr{H}}\backslash\mathscr{H}) = \tfrac{1}{4}\,\mathrm{Vol}((\overline{\Gamma_{\mathcal{O}}})_{\mathscr{H}}\backslash\partial\mathscr{H}) = \frac{1}{2|\mathcal{O}^\times|^2}\,\mathrm{Vol}(\mathcal{O}\backslash\mathbb{H}) = \frac{D_A}{8|\mathcal{O}^\times|^2}.$$

The following lemma will allow us to generalize this formula.

Lemma 15. *Let* $\Lambda'_{\mathcal{O},\rho} = \mathcal{O} \cap \rho^{-1}\mathcal{O} \cap \mathcal{O}\rho^{-1} \cap \rho^{-1}\mathcal{O}\rho^{-1}$ *if* $x, y \neq 0$, *and* $\Lambda'_{\mathcal{O},\rho} = \mathcal{O}$ *otherwise. Then* $\Lambda'_{\mathcal{O},\rho}$ *is a* $\mathbb{Z}$*-lattice in* $\mathbb{H}$ *and we have*

$$\mathrm{Vol}(\overline{G}_{\mathscr{H}}\backslash\mathscr{H}) = \frac{\tau^4[(\overline{\Gamma_{\mathcal{O}}})_{\mathscr{H}} : \overline{G}_{\mathscr{H}}]}{4|\mathcal{O}_\ell(\mathfrak{u}_\rho)^\times|\,[(\overline{\Gamma_{\mathcal{O}}})_{\mathscr{H}} : \overline{\Gamma_{\mathcal{O},x,y}}]}\,\mathrm{Vol}(\Lambda'_{\mathcal{O},\rho}\backslash\mathbb{H}). \tag{29}$$

Proof. If $y = 0$, let $\gamma_\rho = \mathrm{id}$; otherwise let

$$\gamma_\rho = \begin{pmatrix} \rho & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{H}).$$

Note that γ_ρ^{-1} maps ρ to ∞ and $\mathscr{H}$ to the horoball $\mathscr{H}_\infty$ consisting of the points in $\mathbb{H}^5_\mathbb{R}$ with Euclidean height at least $1/\tau$.

Let

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \gamma' = \begin{pmatrix} 1 & b' \\ 0 & 1 \end{pmatrix}$$

be in $\mathrm{SL}_2(\mathbb{H})$. If $y = 0$, we have $\gamma_\rho^{-1}\gamma\gamma_\rho = \gamma'$ if and only if $a = 1$, $b = b'$, $c = 0$, $d = 1$. If $y \neq 0$, by an easy computation, we have $\gamma_\rho^{-1}\gamma\gamma_\rho = \gamma'$ (that is, $\gamma\gamma_\rho = \gamma_\rho\gamma'$) if and only if

$$c = -b', \quad a = 1 - \rho b', \quad d = 1 + b'\rho, \quad b = \rho b'\rho. \tag{30}$$

In particular, if $x,\ y \neq 0$, if $\gamma \in \mathrm{SL}_2(\mathbb{O})$ and $\gamma' = \gamma_\rho^{-1}\gamma\gamma_\rho \in \mathrm{SL}_2(A)$ is unipotent upper triangular, then these equations imply respectively that b' belongs to $\mathbb{O}$, $\rho^{-1}\mathbb{O}$, $\mathbb{O}\rho^{-1}$ and $\rho^{-1}\mathbb{O}\rho^{-1}$; therefore $b' \in \Lambda'_{\mathbb{O},\rho}$. If $x = 0$ or $y = 0$, we also have $b' \in \mathbb{O} = \Lambda'_{\mathbb{O},\rho}$

Conversely, if $b' \in \Lambda'_{\mathbb{O},\rho}$, then define a, b, c, d by the equations (30) if $y \neq 0$, and by $a = 1$, $b = b'$, $c = 0$, $d = 1$ otherwise, so that $a, b, c, d \in \mathbb{O}$. Let

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If $y \neq 0$, note that if $c = 0$, then $\gamma = \mathrm{id}$ and otherwise $cb - cac^{-1}d = -1$, so that $\gamma \in \mathrm{SL}_2(\mathbb{O})$ by (12). If $y \neq 0$, the equations (30) imply that $\gamma_\rho^{-1}\gamma\gamma_\rho$ is a unipotent upper triangular element of $\mathrm{SL}_2(\mathbb{O})$, and this is also the case if $y = 0$.

The abelian group $\Lambda'_{\mathbb{O},\rho}$ is a $\mathbb{Z}$-lattice in $\mathbb{H}$, as an intersection of at most four $\mathbb{Z}$-lattices in A. Since an isometry preserves the volume for the first equality, by an easy hyperbolic volume computation for the second one, and by the previous computation of the unipotent upper triangular subgroup Γ'_{γ_ρ} of $\gamma_\rho^{-1}\Gamma_{\mathbb{O},x,y}\gamma_\rho$ for the last one, we have

$$\begin{aligned}\mathrm{Vol}(\overline{\Gamma_{\mathbb{O},x,y}}\backslash\mathscr{H}) &= \mathrm{Vol}((\gamma_\rho^{-1}\overline{\Gamma_{\mathbb{O},x,y}}\gamma_\rho)\backslash\mathscr{H}_\infty) = \tfrac{1}{4}\mathrm{Vol}((\gamma_\rho^{-1}\overline{\Gamma_{\mathbb{O},x,y}}\gamma_\rho)\backslash\partial\mathscr{H}_\infty) \\ &= \frac{\tau^4}{4}\mathrm{Vol}\big((\gamma_\rho^{-1}\overline{\Gamma_{\mathbb{O},x,y}}\gamma_\rho)\backslash\mathbb{H}\big) = \frac{\tau^4}{4[\gamma_\rho^{-1}\Gamma_{\mathbb{O},x,y}\gamma_\rho : \Gamma'_{\gamma_\rho}]}\,\mathrm{Vol}(\Lambda'_{\mathbb{O},\rho}\backslash\mathbb{H}).\end{aligned}$$

With the notation of the proof of Lemma 10(1), we have $[\gamma_\rho^{-1}\Gamma_{\mathbb{O},x,y}\gamma_\rho : \Gamma'_{\gamma_\rho}] = |\mathbb{O}_\ell(\mathfrak{u}_\rho)^\times|$. Since covering arguments yield

$$\mathrm{Vol}(\,\overline{G}_\mathscr{H}\backslash\mathscr{H}) = [(\overline{\Gamma}_\mathbb{O})_\mathscr{H} : \overline{G}_\mathscr{H}]\,\mathrm{Vol}\big((\overline{\Gamma}_\mathbb{O})_\mathscr{H}\backslash\mathscr{H}\big) = \frac{[(\overline{\Gamma}_\mathbb{O})_\mathscr{H} : \overline{G}_\mathscr{H}]}{[(\overline{\Gamma}_\mathbb{O})_\mathscr{H} : \overline{\Gamma_{\mathbb{O},x,y}}]}\,\mathrm{Vol}(\overline{\Gamma_{\mathbb{O},x,y}}\backslash\mathscr{H}),$$

the result follows. □

Let us resume the proof of Theorem 13. Let $\mathscr{C} = \mathscr{C}(f)$, which is indeed a real hyperbolic hyperplane in $\mathbb{H}^5_{\mathbb{R}}$, whose set of points at infinity is $\mathscr{C}_\infty(f)$ (hence ∞ is a point at infinity of $\mathscr{C}(f)$ if and only if the first coefficient $a = a(f)$ of f is 0). Note that $\mathscr{C}$ is invariant under the group $\mathrm{SU}_f(\mathbb{O})$ by (17) (this equation implies that $\mathscr{C}(f \circ g) = g^{-1}\mathscr{C}(f)$ for every $g \in \mathrm{SL}_2(\mathbb{O})$). The arithmetic group $\mathrm{SU}_f(\mathbb{O})$ acts with finite covolume on $\mathscr{C}(f)$, its finite subgroup $\{\pm\,\mathrm{id}\}$ acting trivially. By definition,

$$\mathrm{Covol}(\mathrm{SU}_f(\mathbb{O}) \cap G) = \mathrm{Vol}\Big({}_{\mathrm{PSU}_f(\mathbb{O}) \cap \overline{G}}\backslash \mathscr{C}(f)\Big).$$

Note that $\mathrm{Covol}(\mathrm{SU}_f(\mathbb{O}) \cap G)$ depends only on the G-orbit of f, by (17) and since $\mathrm{SU}_{f\circ g}(\mathbb{O}) = g^{-1}\,\mathrm{SU}_f(\mathbb{O})g$ for every $g \in \mathrm{SL}_2(\mathbb{O})$. By its definition, $R_G(f)$ is the index of the subgroup $\mathrm{PSU}_f(\mathbb{O}) \cap \overline{G}$ in $\overline{G}_{\mathscr{C}}$; hence

$$\mathrm{Vol}(\overline{G}_{\mathscr{C}}\backslash \mathscr{C}) = \frac{1}{R_G(f)}\,\mathrm{Covol}(\mathrm{SU}_f(\mathbb{O}) \cap G). \tag{31}$$

The last step of the proof of Theorem 13 consists in relating the two counting functions $\psi_{f,G,x,y}$ and $\mathscr{N}_{\overline{G},\mathscr{C},\mathscr{H}}$, in order to apply Theorem 14.

For every $g \in \mathrm{SL}_2(\mathbb{H})$, let us compute the hyperbolic length of the common perpendicular geodesic arc $\delta_{g^{-1},\mathrm{id}}$ between the real hyperbolic hyperplane $g^{-1}\mathscr{C}$ and the horoball $\mathscr{H}$, assuming that they do not meet. We use the notation γ_ρ, $\mathscr{H}_\infty$ introduced in the proof of Lemma 15. Since γ_ρ^{-1} sends the horoball $\mathscr{H}$ to the horoball $\mathscr{H}_\infty$, it sends the common perpendicular geodesic arc between $g^{-1}\mathscr{C}$ and $\mathscr{H}$ to the (vertical) common perpendicular geodesic arc between $\gamma_\rho^{-1}g^{-1}\mathscr{C}$ and $\mathscr{H}_\infty$. Let r be the Euclidean radius of the 3-sphere $\mathscr{C}_\infty(f \circ g \circ \gamma_\rho)$, which is the image by γ_ρ^{-1} of the boundary at infinity of $g^{-1}\mathscr{C}$ by (17). Denoting by $a(f \circ g \circ \gamma_\rho)$ the coefficient of $\mathrm{n}(u)$ in $f \circ g \circ \gamma_\rho(u, v)$, we have, by (16),

$$\begin{aligned} r = \frac{\sqrt{\Delta(f \circ g \circ \gamma_\rho)}}{|a(f \circ g \circ \gamma_\rho)|} &= \frac{\sqrt{\Delta}}{|f \circ g \circ \gamma_\rho(1,0)|} \\ &= \frac{\sqrt{\Delta}}{|f \circ g(\rho, 1)|} = \frac{\mathrm{n}(y)\sqrt{\Delta}}{|f \circ g(x, y)|}, \end{aligned}$$

if $y \neq 0$ and $r = (\mathrm{n}(x)\sqrt{\Delta})/|f \circ g(x, y)|$ otherwise. An immediate computation gives

$$\ell(\delta_{g^{-1},\mathrm{id}}) = \ell(\gamma_\rho^{-1}\delta_{g^{-1},\mathrm{id}}) = \ln\frac{1}{\tau} - \ln r = \ln\frac{|f \circ g(x, y)|}{\tau\ \mathrm{n}(y)\sqrt{\Delta}}, \tag{32}$$

if $y \neq 0$ and

$$\ell(\delta_{g^{-1},\mathrm{id}}) = \ln\frac{|f \circ g(x, y)|}{\tau\ \mathrm{n}(x)\sqrt{\Delta}}$$

otherwise. With the conventions that we have taken, these formulas are also valid if $g^{-1}\mathscr{C}$ and $\mathscr{H}$ meet.

Recall that there are only finitely many elements $[g] \in \overline{G}_{\mathscr{C}} \backslash \overline{G} / \overline{G}_{\mathscr{H}}$ such that $g^{-1} \overline{G}_{\mathscr{C}}\, g \cap \overline{G}_{\mathscr{H}}$ is different from $\{1\}$ or such that the multiplicity $m(g^{-1}, \mathrm{id})$ is different from 1. If $y \neq 0$, using (32) for the third line below, [Parkkonen and Paulin 2011, Lemma 7] for the fourth one, and Theorem 14 applied to $F = \overline{G}$ for the sixth one, we hence have, as s tends to $+\infty$,

$$
\begin{aligned}
&\psi_{f,G,x,y}(s)\\
&\quad = \mathrm{Card}\big\{[g] \in (\mathrm{SU}_f(\mathbb{O}) \cap G) \backslash G / G_{x,y} : \mathrm{n}(\mathbb{O}x + \mathbb{O}y)^{-1} |f \circ g(x, y)| \leq s\big\}\\
&\quad = \mathrm{Card}\Big\{[g] \in (\mathrm{PSU}_f(\mathbb{O}) \cap \overline{G}) \backslash \overline{G} / \overline{G_{x,y}} : \ell(\delta_{g^{-1},\mathrm{id}}) \leq \ln \frac{s\, \mathrm{n}(\mathbb{O}x + \mathbb{O}y)}{\tau\, \mathrm{n}(y) \sqrt{\Delta}}\Big\}\\
&\quad \sim R_G(f) [\overline{G}_{\mathscr{H}} : \overline{G_{x,y}}]\, \mathrm{Card}\Big\{[g] \in \overline{G}_{\mathscr{C}} \backslash \overline{G} / \overline{G}_{\mathscr{H}} : \ell(\delta_{g^{-1},\mathrm{id}}) \leq \ln \frac{s\, \mathrm{n}(\mathbb{O}\rho + \mathbb{O})}{\tau \sqrt{\Delta}}\Big\}\\
&\quad \sim R_G(f) [\overline{G}_{\mathscr{H}} : \overline{G_{x,y}}]\, \mathcal{N}_{\overline{G},\mathscr{C},\mathscr{H}}\Big(\ln \frac{s\, \mathrm{n}(\mathbb{O}\rho + \mathbb{O})}{\tau \sqrt{\Delta}}\Big)\\
&\quad \sim R_G(f) [\overline{G}_{\mathscr{H}} : \overline{G_{x,y}}] \frac{6 \,\mathrm{Vol}(\overline{G}_{\mathscr{H}} \backslash \mathscr{H})\, \mathrm{Vol}(\overline{G}_{\mathscr{C}} \backslash \mathscr{C})}{8\pi^2\, \mathrm{Vol}(\overline{G} \backslash \mathbb{H}^n_{\mathbb{R}})} \Big(\frac{s\, \mathrm{n}(\mathbb{O}\rho + \mathbb{O})}{\tau \sqrt{\Delta}}\Big)^4.
\end{aligned}
$$

We replace the three volumes in the computation above by their expressions given in the equations (28), (29) and (31). We simplify the obtained expression using the following two remarks. Firstly,

$$
[\overline{G}_{\mathscr{H}} : \overline{G_{x,y}}] \frac{[(\overline{\Gamma_{\mathbb{O}}})_{\mathscr{H}} : \overline{G}_{\mathscr{H}}]}{[(\overline{\Gamma_{\mathbb{O}}})_{\mathscr{H}} : \overline{\Gamma_{\mathbb{O},x,y}}]} = \frac{[(\overline{\Gamma_{\mathbb{O}}})_{\mathscr{H}} : \overline{G_{x,y}}]}{[(\overline{\Gamma_{\mathbb{O}}})_{\mathscr{H}} : \overline{\Gamma_{\mathbb{O},x,y}}]} = [\overline{\Gamma_{\mathbb{O},x,y}} : \overline{G_{x,y}}] = [\Gamma_{\mathbb{O},x,y} : G_{x,y}].
$$

Secondly, we claim that

$$
\mathrm{Vol}(\Lambda'_{\mathbb{O},\rho} \backslash \mathbb{H})\, \mathrm{n}(\mathbb{O}\rho + \mathbb{O})^4 = \frac{D_A}{4}. \tag{33}
$$

If $x = 0$, then $\Lambda'_{\mathbb{O},\rho} = \mathbb{O}$; hence this claim is true, by (5) and since $\mathrm{n}(\mathbb{O}) = 1$. Otherwise, claim (33) follows from Lemma 6 with $z = \rho^{-1}$, since, by the definition of $\Lambda'_{\mathbb{O},\rho}$,

$$
\mathrm{Vol}(\Lambda'_{\mathbb{O},\rho} \backslash \mathbb{H})\, \mathrm{n}(\mathbb{O}\rho + \mathbb{O})^4 = \mathrm{Vol}(\Lambda \backslash \mathbb{H})\, \mathrm{n}(\mathbb{O}z^{-1} + \mathbb{O})^4 = \mathrm{Vol}(\mathbb{O} \backslash \mathbb{H}) [\mathbb{O} : \Lambda]\, \mathrm{n}(\mathbb{O}z^{-1} + \mathbb{O})^4,
$$

and by (5).

This concludes the proof of Theorem 13 if $y \neq 0$. The case $y = 0$ is similar to the case $x = 0$. □

Let us give a few corollaries of Theorem 13. The first one below follows by taking $G = \mathrm{SL}_2(\mathbb{O})$ in Theorem 13.

Corollary 16. *Let f be an integral indefinite binary Hamiltonian form of discriminant $\Delta(f)$ over a maximal order $\mathbb{O}$ of a definite quaternion algebra A over $\mathbb{Q}$.*

Let x and y be elements in $\mathcal{O}$ not both zero. Then, as s tends to $+\infty$, we have the equivalence

$$\psi_{f,\mathrm{SL}_2(\mathcal{O}),x,y}(s) \sim \frac{540\,\mathrm{Covol}(\mathrm{SU}_f(\mathcal{O}))}{\pi^2\zeta(3)|\mathcal{O}_\ell(\mathfrak{u}_{xy^{-1}})^\times|\Delta(f)^2\prod_{p|D_A}(p^3-1)(1-p^{-1})}\, s^4.$$

Remark 17. Recall that by Remark 7, the map from $\mathrm{SL}_2(\mathcal{O})\backslash\mathbb{P}^1_r(\mathcal{O})$ to ${}_{\mathcal{O}}\mathcal{I}\times{}_{\mathcal{O}}\mathcal{I}$ that associates, to the orbit of $[u:v]$ in $\mathbb{P}^1_r(\mathcal{O})$ under $\mathrm{SL}_2(\mathcal{O})$, the couple of ideal classes $([I_{u,v}],[K_{u,v}])$ is a bijection. The counting function $\psi_{f,\mathrm{SL}_2(\mathcal{O}),x,y}$ hence depends only on $([I_{x,y}],[K_{x,y}])$.

Given two left fractional ideals $\mathfrak{m}$ and $\mathfrak{m}'$ of $\mathcal{O}$, let $\psi_{f,\mathfrak{m},\mathfrak{m}'}(s)$ be the cardinality of the set

$$\mathrm{SU}_f(\mathcal{O})\Big\backslash\Big\{(u,v)\in\mathfrak{m}\times\mathfrak{m} : \frac{|f(u,v)|}{\mathrm{n}(\mathfrak{m})}\le s,\ I_{u,v}=\mathfrak{m},\ [K_{u,v}]=[\mathfrak{m}']\Big\}.$$

Note that this counting function depends only on the ideal classes of $\mathfrak{m}$ and $\mathfrak{m}'$.

Corollary 18. *Let f be an integral indefinite binary Hamiltonian form of discriminant $\Delta(f)$ over a maximal order $\mathcal{O}$ of a definite quaternion algebra A over $\mathbb{Q}$. Let $\mathfrak{m}$ and $\mathfrak{m}'$ be two left fractional ideals in $\mathcal{O}$. Then as s tends to $+\infty$, we have the equivalence*

$$\psi_{f,\mathfrak{m},\mathfrak{m}'}(s) \sim \frac{540\,\mathrm{Covol}(\mathrm{SU}_f(\mathcal{O}))}{\pi^2\zeta(3)|\mathcal{O}_r(\mathfrak{m}')^\times|\Delta(f)^2\prod_{p|D_A}(p^3-1)(1-p^{-1})} s^4.$$

Proof. By Remark 17, we have

$$\psi_{f,\mathfrak{m},\mathfrak{m}'} = \psi_{f,\mathrm{SL}_2(\mathcal{O}),x,y},$$

where (x,y) is any nonzero element of $\mathcal{O}\times\mathcal{O}$ such that $[I_{x,y}]=[\mathfrak{m}]$ and $[K_{x,y}]=[\mathfrak{m}']$. By the equations (4) and (3), if $xy\neq 0$, we have

$$|\mathcal{O}_\ell(\mathfrak{u}_{xy^{-1}})^\times| = |\mathcal{O}_r(\mathfrak{u}_{xy^{-1}}{}^{-1})^\times| = |\mathcal{O}_r(\mathcal{O}\cap\mathcal{O}yx^{-1})^\times| = |\mathcal{O}_r(K_{x,y})^\times|.$$

The first and last terms are also equal if $xy=0$. Hence the result follows from Corollary 16. □

Remark 19. With $\psi_{f,\mathfrak{m}}$ the counting function defined in the introduction, we have

$$\psi_{f,\mathfrak{m}} = \sum_{[\mathfrak{m}']\in{}_{\mathcal{O}}\mathcal{I}} \psi_{f,\mathfrak{m},\mathfrak{m}'}. \tag{34}$$

Therefore, since

$$\sum_{[\mathfrak{m}']\in{}_{\mathcal{O}}\mathcal{I}} \frac{1}{|\mathcal{O}_r(\mathfrak{m}')^\times|} = \frac{1}{24}\prod_{p|D_A}(p-1)$$

by (8), Theorem 1 in the introduction follows from Corollary 18.

We say $u, v \in \mathcal{O} \times \mathcal{O}$ are *relatively prime* if one of the following equivalent (by Remark 17) conditions is satisfied:

(i) There exists $g \in \mathrm{SL}_2(\mathcal{O})$ such that $g(1,0) = (u, v)$.

(ii) There exists $u', v' \in \mathcal{O}$ such that $\mathrm{n}(uv') + \mathrm{n}(u'v) - \mathrm{tr}(u\bar{v}v'\overline{u'}) = 1$.

(iii) The $\mathcal{O}$-modules $I_{u,v}$ and $K_{u,v}$ are isomorphic (as $\mathcal{O}$-modules) to $\mathcal{O}$.

We denote by $\mathcal{P}_{\mathcal{O}}$ the set of couples of relatively prime elements of $\mathcal{O}$.

Corollary 20. *Let f be an integral indefinite binary Hamiltonian form over a maximal order $\mathcal{O}$ in a definite quaternion algebra A over $\mathbb{Q}$, and let G be a finite index subgroup of $\Gamma_{\mathcal{O}} = \mathrm{SL}_2(\mathcal{O})$. Then, as s tends to $+\infty$, we have the equivalence*

$$\mathrm{Card}\ {}_{\mathrm{SU}_f(\mathcal{O}) \cap G}\backslash\big\{(u, v) \in \mathcal{P}_{\mathcal{O}} : |f(u, v)| \le s\big\}$$
$$\sim \frac{540\, \iota_G[\Gamma_{\mathcal{O},1,0} : G_{1,0}]\ \mathrm{Covol}(\mathrm{SU}_f(\mathcal{O}) \cap G)}{\pi^2 \zeta(3) |\mathcal{O}^\times| \Delta(f)^2 [\Gamma_{\mathcal{O}} : G] \prod_{d|D_A} (p^3 - 1)(1 - p^{-1})} s^4.$$

Proof. This follows from Theorem 13 by taking $x = 1$ and $y = 0$. □

Proof of Corollary 2 from the introduction. Consider the integral indefinite binary Hamiltonian form f over $\mathcal{O}$ defined by $f(u, v) = \mathrm{tr}(\bar{u}\, v)$, with matrix

$$M(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and discriminant $\Delta(f) = 1$. Its group of automorphs is

$$\mathrm{Sp}_1(\mathcal{O}) = \left\{ g \in \mathrm{SL}_2(\mathcal{O}) : g^* \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\},$$

which is an arithmetic lattice in the symplectic group over the quaternions $\mathrm{Sp}_1(\mathbb{H})$. We have

$$\mathcal{C}(f) = \{(z, r) \in \mathbb{H} \times]0, +\infty[: \mathrm{tr}(z) = 0\}.$$

The hyperbolic volume of the quotient of $\{(z, r) \in \mathbb{H} \times]0, +\infty[: \mathrm{tr}(z) = 0\}$ by $\mathrm{Sp}_1(\mathcal{O})$ has been computed as the main result of [Breulmann and Helmke 1996], yielding

$$\mathrm{Covol}(\mathrm{Sp}_1(\mathcal{O})) = \frac{\pi^2}{1080} \prod_{p|D_A} (p^2 + 1)(p - 1),$$

where p ranges over the primes dividing D_A.

Corollary 2 in the introduction then follows from Theorem 1 with $\mathfrak{m} = \mathcal{O}$. □

Remark 21. Theorem 13 and its Corollary 20 allow the asymptotic study of the counting of representations satisfying congruence properties. For instance, let $\mathcal{J}$ be a (nonzero) two-sided ideal in an order $\mathcal{O}$ in a definite quaternion algebra A over $\mathbb{Q}$. Let $\Gamma_{\mathcal{J}}$ be the kernel of the map $\mathrm{SL}_2(\mathcal{O}) \to \mathrm{GL}_2(\mathcal{O}/\mathcal{J})$ of reduction modulo

$\mathscr{I}$ of the coefficients, and $\Gamma_{\mathscr{I},0}$ the preimage of the upper triangular subgroup by this map. Then applying Corollary 20 with $G = \Gamma_{\mathscr{I}}$ and $G = \Gamma_{\mathscr{I},0}$ respectively, we get an asymptotic equivalence as $s \to +\infty$ of the number of relatively prime representations (u, v) of integers with absolute value at most s by a given integral binary Hamiltonian form, satisfying the additional congruence properties

$$\{u \equiv 1 \bmod \mathscr{I},\, v \equiv 0 \bmod \mathscr{I}\} \quad \text{or} \quad \{v \equiv 0 \bmod \mathscr{I}\}.$$

To give an even more precise result, the computation of the indices of $\Gamma_{\mathscr{I}}$ and $\Gamma_{\mathscr{I},0}$ in $\mathrm{SL}_2(\mathbb{O})$ would be needed.

7. Geometric reduction theory of binary Hamiltonian forms

Let $\mathbb{O}$ be a (not necessarily maximal) order in a definite quaternion algebra A over $\mathbb{Q}$.

Let $\mathscr{Q}$ be the 6-dimensional real vector space of binary Hamiltonian forms, $\mathscr{Q}^+$ the open cone of positive definite ones, $\mathscr{Q}^\pm$ the open cone of indefinite ones, $\mathscr{Q}(\mathbb{O})$ the discrete subset of the ones that are integral over $\mathbb{O}$, and

$$\mathscr{Q}^+(\mathbb{O}) = \mathscr{Q}^+ \cap \mathscr{Q}(\mathbb{O}), \quad \mathscr{Q}^\pm(\mathbb{O}) = \mathscr{Q}^\pm \cap \mathscr{Q}(\mathbb{O}).$$

For every $\Delta \in \mathbb{Z} - \{0\}$, let $\mathscr{Q}(\Delta) = \{f \in \mathscr{Q} : \Delta(f) = \Delta\}$, $\mathscr{Q}(\mathbb{O}, \Delta) = \mathscr{Q}(\Delta) \cap \mathscr{Q}(\mathbb{O})$ and

$$\mathscr{Q}^+(\mathbb{O}, \Delta) = \mathscr{Q}(\Delta) \cap \mathscr{Q}^+(\mathbb{O}), \quad \mathscr{Q}^\pm(\mathbb{O}, \Delta) = \mathscr{Q}(\Delta) \cap \mathscr{Q}^\pm(\mathbb{O}).$$

The group $\mathbb{R}_+^*$ acts on $\mathscr{Q}^+$ by multiplication; we will denote by $[f]$ the orbit of f and by $\overline{\mathscr{Q}}^+$ the quotient space $\mathscr{Q}^+/\mathbb{R}_+^*$. Similarly, the group $\mathbb{R}^*$ acts on $\mathscr{Q}^\pm$ by multiplication; we will denote by $[f]$ the orbit of f and by $\overline{\mathscr{Q}}^\pm$ the quotient space $\mathscr{Q}^\pm/\mathbb{R}^*$. The right action of $\mathrm{SL}_2(\mathbb{H})$ on $\mathscr{Q}$ preserves $\mathscr{Q}(\Delta)$, $\mathscr{Q}^+$ and $\mathscr{Q}^\pm$, commuting with the actions of $\mathbb{R}_+^*$ and $\mathbb{R}^*$ on these last two spaces. The subgroup $\mathrm{SL}_2(\mathbb{O})$ preserves $\mathscr{Q}(\mathbb{O})$, $\mathscr{Q}^+(\mathbb{O})$, $\mathscr{Q}^\pm(\mathbb{O})$, $\mathscr{Q}^+(\mathbb{O}, \Delta)$, $\mathscr{Q}^\pm(\mathbb{O}, \Delta)$.

Let $\mathscr{C}(\mathbb{H}^5_\mathbb{R})$ be the space of totally geodesic hyperplanes of $\mathbb{H}^5_\mathbb{R}$, with the Hausdorff distance on compact subsets.

Proposition 22. (1) *The map* $\Phi : \overline{\mathscr{Q}}^+ \to \mathbb{H}^5_\mathbb{R}$ *defined by*

$$[f] \mapsto \Big(-\frac{b(f)}{a(f)}, \frac{\sqrt{-\Delta(f)}}{a(f)}\Big)$$

is a homeomorphism, which is (anti)equivariant for the actions of $\mathrm{SL}_2(\mathbb{H})$: *For every* $g \in \mathrm{SL}_2(\mathbb{H})$, *we have* $\Phi([f \circ g]) = g^{-1}\Phi([f])$.

(2) *The map* $\Psi : \overline{\mathscr{Q}}^\pm \to \mathscr{C}(\mathbb{H}^5_\mathbb{R})$ *defined by* $[f] \mapsto \mathscr{C}(f)$ *is a homeomorphism, which is (anti)equivariant for the actions of* $\mathrm{SL}_2(\mathbb{H})$: *For every* $g \in \mathrm{SL}_2(\mathbb{H})$, *we have* $\Psi([f \circ g]) = g^{-1}\Psi([f])$.

Note that $\Phi([f])$ may be geometrically understood as the pair of the center and the imaginary radius of the imaginary sphere with equation $f(z, 1) = 0$, that is,

$$\mathrm{n}\Big(z + \frac{b(f)}{a(f)}\Big) = -\Big(\frac{\sqrt{-\Delta(f)}}{a(f)}\Big)^2.$$

Proof. (1) Since $a = a(f) > 0$ and $\Delta = \Delta(f) < 0$ when f is a positive definite binary Hamiltonian form, the map Φ is well-defined and continuous. Since the orbit by $\mathbb{R}_+^*$ of a positive definite binary Hamiltonian form has a unique element f such that $a(f) = 1$, and since $c(f)$ then is equal to $n(b(f)) - \Delta$, the map Φ is a bijection with continuous inverse $(z, r) \mapsto [f_{z,r}]$ where

$$f_{z,r} : (u, v) \mapsto \mathrm{n}(u) - \mathrm{tr}(\bar{u} z v) + (\mathrm{n}(z) + r^2)\, \mathrm{n}(v).$$

To prove the equivariance property of Φ, we could use (14) and the formula for the inverse of an element of $\mathrm{SL}_2(\mathbb{O})$ given for instance in [Kellerhals 2003], but the computations are quite technical and even longer than below. Hence we prefer to use the following lemma to decompose the computations.

Lemma 23. *The group* (*even the monoid*) $\mathrm{SL}_2(\mathbb{H})$ *is generated by the elements*

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad \text{with } \beta \in \mathbb{H}.$$

This is a consequence of a general fact about connected semisimple real Lie groups and their root groups, but the proof is short (and is one way to prove that the Dieudonné determinant of

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

is $\mathrm{n}(\gamma\beta - \gamma\alpha\gamma^{-1}\delta)$ if $\gamma \neq 0$).

Proof. This follows from the following facts, where $\alpha, \beta, \gamma, \delta \in \mathbb{H}$. If $\alpha \neq 0$, then

$$\begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} 1 & \alpha^{-1}\beta \\ 0 & 1 \end{pmatrix}, \quad \text{and}$$

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\alpha^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

If $\mathrm{n}(\alpha\delta) = 1$, there exist $u, v \in \mathbb{H}^\times$ such that $\alpha\delta = uvu^{-1}v^{-1}$, and

$$\begin{pmatrix} \alpha & 0 \\ 0 & \delta \end{pmatrix} = \begin{pmatrix} u & 0 \\ 0 & u^{-1} \end{pmatrix} \begin{pmatrix} v & 0 \\ 0 & v^{-1} \end{pmatrix} \begin{pmatrix} (vu)^{-1} & 0 \\ 0 & vu \end{pmatrix} \begin{pmatrix} \delta^{-1} & 0 \\ 0 & \delta \end{pmatrix}.$$

If $\gamma \neq 0$, then

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & \alpha\gamma^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma & 0 \\ 0 & -\beta + \alpha\gamma^{-1}\delta \end{pmatrix} \begin{pmatrix} 1 & \gamma^{-1}\delta \\ 0 & 1 \end{pmatrix}. \qquad \square$$

Now, to prove the equivariance property, one only has to prove it for the elements of the generating set of $\mathrm{SL}_2(\mathbb{H})$ given in the above lemma. Given $f \in \mathfrak{Q}^+$, let

$$M = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}$$

be the matrix of f and $\Delta = \Delta(f)$. Note that the matrix of $f \circ g$ is $g^* M g$.

If

$$g = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix},$$

we have $a(f \circ g) = a$ and $b(f \circ g) = a\beta + b$. Since

$$g^{-1} \cdot \Big(-\frac{b}{a}, \frac{\sqrt{-\Delta}}{a}\Big) = \Big(-\frac{b}{a} - \beta, \frac{\sqrt{-\Delta}}{a}\Big) = \Big(-\frac{b(f \circ g)}{a(f \circ g)}, \frac{\sqrt{-\Delta(f \circ g)}}{a(f \circ g)}\Big)$$

by (16), the result follows in this case.

If

$$g = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

then $a(f \circ g) = c$ and $b(f \circ g) = -\bar{b}$. By (14), for every $(z, h) \in \mathbb{H}^5_{\mathbb{R}}$, we have

$$g^{-1} \cdot (z, r) = \Big(\frac{-\bar{z}}{\mathrm{n}(z) + r^2}, \frac{r}{\mathrm{n}(z) + r^2}\Big).$$

Therefore, since $\Delta = \mathrm{n}(b) - ac$,

$$g^{-1} \cdot \Big(-\frac{b}{a}, \frac{\sqrt{-\Delta}}{a}\Big) = \left(\frac{-(-\frac{\bar{b}}{a})}{\frac{\mathrm{n}(b)}{a^2} + \frac{-\Delta}{a^2}}, \frac{\frac{\sqrt{-\Delta}}{a}}{\frac{c}{a}}\right) = \Big(-\frac{b(f \circ g)}{a(f \circ g)}, \frac{\sqrt{-\Delta(f \circ g)}}{a(f \circ g)}\Big).$$

The equivariance property of Φ follows.

(2) We have already seen that Ψ is a bijection. Its equivariance property follows from (17). Let $a = a(f)$, $b = b(f)$, $c = c(f)$ and $\Delta = \Delta(f)$. Since

$$\mathscr{C}(f) = \begin{cases} \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : \mathrm{n}(az + b) + a^2 r^2 = \Delta\} & \text{if } a \neq 0, \\ \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : \mathrm{tr}(\bar{z} b) + c = 0\} & \text{otherwise,} \end{cases}$$

the map Ψ is clearly a homeomorphism. □

In order to define a geometric notion of reduced binary Hamiltonian form, much less is needed than an actual fundamental domain for the group $\mathrm{SL}_2(\mathbb{O})$ acting on $\mathbb{H}^5_{\mathbb{R}}$. Though it might increase the number of reduced elements, this will make the verification that a given binary form is reduced much easier (see the end of this section). Indeed, due to the higher dimension, the number of inequalities is much larger than the one for $\mathrm{SL}_2(\mathbb{Z})$ or for $\mathrm{SL}_2(\mathbb{O}_K)$, where $\mathbb{O}_K$ is the ring of integers of an imaginary quadratic number field K; see for instance [Zagier 1981; Buchmann and Vollmer 2007; Elstrodt et al. 1998].

For $n \geq 2$, let us denote by $\|z\|$ the usual Euclidean norm on $\mathbb{R}^{n-1}$. Consider the upper halfspace model of the real hyperbolic n-space $\mathbb{H}^n_{\mathbb{R}}$, whose underlying manifold is $\mathbb{R}^{n-1} \times]0, +\infty[$, so that $\partial_\infty \mathbb{H}^n_{\mathbb{R}} = \mathbb{R}^{n-1} \cup \{\infty\}$. A *weak fundamental domain* for the action of a finite covolume discrete subgroup Γ of isometries of $\mathbb{H}^n_{\mathbb{R}}$ is a subset $\mathscr{F}$ of $\mathbb{H}^n_{\mathbb{R}}$ such that

(i) $\bigcup_{g \in \Gamma} g\mathscr{F} = \mathbb{H}^n_{\mathbb{R}}$,

(ii) there exists a compact subset K in $\mathbb{R}^{n-1}$ such that $\mathscr{F}$ is contained in $K \times]0, +\infty[$,

(iii) there exist $\kappa, \epsilon > 0$ and a finite set Z of parabolic fixed points of Γ such that $\mathscr{F} = \{(z, r) \in \mathscr{F} : r \geq \epsilon\} \cup (\bigcup_{s \in Z} \mathscr{E}_s)$, where $\mathscr{E}_s \subset \{(z, r) \in \mathscr{F} : \|z - s\| \leq \kappa r^2\}$.

Note that a weak fundamental domain for a finite index subgroup of Γ is a weak fundamental domain for Γ.

When ∞ is a parabolic fixed point of Γ, an example of a weak fundamental domain is any Ford fundamental domain of Γ, whose definition we now recall.

Given any isometry g of $\mathbb{H}^n_{\mathbb{R}}$ such that $g\infty \neq \infty$, the *isometric sphere* of g is the $(n-2)$-sphere S_g of $\mathbb{R}^{n-1}$ that consists of the points at which the tangent map of g is a Euclidean isometry. We then define S_g^+ as the set of points in $\mathbb{H}^n_{\mathbb{R}}$ that are in the closure of the unbounded component of the complement of the hyperbolic hyperplane whose boundary is S_g. For instance, if

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}_2(\mathbb{H}),$$

then $g\infty \neq \infty$ if and only if $\gamma \neq 0$ and its isometric sphere is then

$$S_g = \{z \in \mathbb{H} : \mathrm{n}(\gamma z + \delta) = 1\}, \quad \text{so that} \quad S_g^+ = \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : \mathrm{n}(\gamma z + \delta) + r^2 \geq 1\}.$$

Recall that since Γ has finite covolume, every parabolic fixed point ξ of Γ is *bounded*, that is, the quotient of $\partial_\infty \mathbb{H}^n_{\mathbb{R}} - \{\xi\}$ by the stabilizer of ξ in Γ is compact. Let $\mathscr{D}_\infty$ be a compact fundamental domain for the action of the stabilizer of ∞ in Γ on $\mathbb{R}^{n-1}$. Then the *Ford fundamental domain* $\mathscr{F}_\Gamma$ of Γ associated to $\mathscr{D}_\infty$ is

$$\mathscr{F}_\Gamma = \Big(\bigcap_{\substack{g \in \Gamma \\ g\infty \neq \infty}} S_g^+ \Big) \cap \big(\mathscr{D}_\infty \times]0, +\infty[\big).$$

It is well known (see for instance [Beardon 1983, page 239]) that $\mathscr{F}_\Gamma$ is a fundamental domain for Γ acting on $\mathbb{H}^n_{\mathbb{R}}$ (in particular, $\mathscr{F}_\Gamma$ satisfies condition (i) of a weak fundamental domain) and that the set of points at infinity of $\bigcap_{g \in \Gamma, g\infty \neq \infty} S_g^+$ is a locally finite set of parabolic fixed points in $\partial_\infty \mathbb{H}^n_{\mathbb{R}}$. Furthermore, since parabolic fixed points are bounded and have a precisely invariant horoball centered at them, and since the tangency of a circle and its tangent is quadratic, the condition (iii) is satisfied for every ϵ small enough, and κ large enough. Note that $\mathscr{F}_\Gamma$ satisfies condition (ii) with $K = \mathscr{D}_\infty$.

Let us fix a weak fundamental domain $\mathscr{F}$ for the action of $\mathrm{SL}_2(\mathbb{O})$ on $\mathbb{H}^5_\mathbb{R}$. A positive definite form $f \in \mathscr{Q}^+(\mathbb{O})$ is *reduced* if $\Phi([f]) \in \mathscr{F}$ and an indefinite form $f \in \mathscr{Q}^\pm(\mathbb{O})$ is *reduced* if $\Psi([f]) \cap \mathscr{F} \neq 0$. We say that a negative definite form $f \in -\mathscr{Q}^+(\mathbb{O})$ is *reduced* if $-f$ is reduced. The notion of being reduced does depend on the choice of a weak fundamental domain, which allows us to choose it appropriately when computing examples. Recall that $\mathscr{Q}(\Delta)$ is equal to $\mathscr{Q}^\pm(\Delta)$ if $\Delta > 0$ and to $\mathscr{Q}^+(\Delta) \cup -\mathscr{Q}^+(\Delta)$ if $\Delta < 0$.

Theorem 24. *For every $\Delta \in \mathbb{Z} - \{0\}$, the number of reduced elements of $\mathscr{Q}(\mathbb{O}, \Delta)$ is finite.*

This is a restatement of Theorem 4 in the introduction.

Proof. Note that the Euclidean norm on $\mathbb{H}$ is $\|z\| = \mathrm{n}(z)^{1/2}$.

Let us first prove that the number of reduced elements of $\mathscr{Q}^+(\mathbb{O}, \Delta)$ is finite.

For every $f \in \mathscr{Q}^+(\mathbb{O}, \Delta)$, let $a = a(f) > 0$, $b = b(f)$ and $c = c(f)$. We have $\mathrm{n}(b) - ac = \Delta < 0$; hence c is determined by a and b. The form f is reduced if and only if

$$\Phi([f]) = \left(-\frac{b}{a}, \frac{\sqrt{-\Delta}}{a}\right) \in \mathscr{F}.$$

By the condition (ii) and since K is compact, $\|b/a\|$ is bounded. Hence, if we have an upper bound on a, by the discreteness of $\mathbb{O}$, the elements a and b may take only finitely many values, and so does c, therefore the result follows.

Let κ, ϵ, Z be as in the condition (iii). If $\sqrt{-\Delta}/a \geq \epsilon$, then a is bounded from above, and we are done. Otherwise, by condition (iii), there exists s in the finite set Z such that $\Phi([f]) \in \mathscr{E}_s$. In particular,

$$\left\|-\frac{b}{a} - s\right\| \leq \kappa\left(\frac{\sqrt{-\Delta}}{a}\right)^2.$$

Since the set of parabolic elements of $\mathrm{SL}_2(\mathbb{O})$ is $A \cup \{\infty\}$, we may write $s = u/v$ with $u \in \mathbb{O}$ and $v \in \mathbb{N} - \{0\}$. The inequality above becomes

$$a\|bv + au\| \leq \kappa|\Delta|v.$$

The element $bv + au \in \mathbb{O}$ either is equal to 0 or has reduced norm, hence Euclidean norm, at least 1. In the second case, we have an upper bound on a, as wanted. In the first case, we have $b/a = -u/v$, that is $b = -au/v$. Hence

$$\Delta v^2 = (\mathrm{n}(b) - ac)v^2 = a(a\,\mathrm{n}(u) - cv^2).$$

Since $a\,\mathrm{n}(u) - cv^2 \in \mathbb{Z}$, the integer a divides the nonzero integer Δv^2; hence a is bounded, as wanted.

Let us now prove that the number of reduced elements of $\mathcal{Q}^{\pm}(\mathbb{O}, \Delta)$ is finite, which concludes the proof of Theorem 24.

We have $\Delta > 0$. With K a compact subset as in the condition (ii), let $\delta = \sup_{x \in K} \|x\|$. Let $f \in \mathcal{Q}^{\pm}(\mathbb{O}, \Delta)$ be reduced, and fix $(z, r) \in \mathscr{C}(f) \cap \mathscr{F}$. Let $a = a(f)$, $b = b(f)$ and $c = c(f)$.

Assume first that $a = 0$. Then $\mathrm{n}(b) = \Delta$; hence b takes only finitely many values, by the discreteness of $\mathbb{O}$. Recalling that $\mathscr{C}(f) = \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : \mathrm{tr}(\bar{z}b) + c = 0\}$, we have by the Cauchy–Schwarz inequality

$$|c| = |\mathrm{tr}(\bar{z}b)| \le 2\|z\|\,\|b\| \le 2\delta\sqrt{\Delta}.$$

Again by discreteness, c takes only finitely many values, and the result follows.

Assume that $a \neq 0$, and up to replacing f by $-f$ (which is reduced if f is), that $a > 0$. We have $\mathrm{n}(b) - ac = \Delta$, hence c is determined by a and b. Recalling that $\mathscr{C}(f) = \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : \mathrm{n}(az + b) + a^2 r^2 = \Delta\}$, we have by the triangular inequality

$$\left\|\frac{b}{a}\right\| \le \left\|z + \frac{b}{a}\right\| + \|z\| \le \sqrt{\Delta} + \delta.$$

Hence as in the positive definite case, if we have an upper bound on a, the result follows.

Let κ, ϵ, Z be as in the condition (iii). Note that $r \le \sqrt{\Delta}/a$. Hence if $r \ge \epsilon$, then we have an upper bound $a \le \sqrt{\Delta}/\epsilon$, as wanted. Therefore, we may assume that (z, r) belongs to $\mathscr{C}(f) \cap \mathscr{E}_s$ for some $s \in Z$. In particular,

$$\left\|z + \frac{b}{a}\right\| = \sqrt{\frac{\Delta}{a^2} - r^2} \quad \text{and} \quad \|z - s\| \le \kappa r^2.$$

First assume that $\|(b/a) + s\| \ge \sqrt{\Delta}/a$. Then by the inverse triangular inequality

$$\kappa r^2 \ge \|s - z\| \ge \left\|\frac{b}{a} + s\right\| - \left\|z + \frac{b}{a}\right\| \ge \frac{\sqrt{\Delta}}{a} - \sqrt{\frac{\Delta}{a^2} - r^2} \ge \frac{r^2}{2\sqrt{\Delta}/a}.$$

Therefore, we have an upper bound $a \le 2\kappa\sqrt{\Delta}$, as wanted.

Now assume that $\|(b/a) + s\| < \sqrt{\Delta}/a$. Write $s = u/v$ with $u \in \mathbb{O}$ and $v \in \mathbb{N} - \{0\}$. We have $\mathrm{n}(au + bv) < \Delta v^2$. The element $w = au + bv$, belonging to $\mathbb{O}$ and having reduced norm at most Δv^2, can take only finitely many values. The positive integer $v^2\Delta - \mathrm{n}(w)$ is equal to

$$\begin{aligned} v^2(\mathrm{n}(b) - ac) - \mathrm{n}(au + bv) &= -\,\mathrm{tr}(\overline{au}bv) - \mathrm{n}(au) - v^2 ac \\ &= -a(\mathrm{tr}(\bar{u}bv) + a\,\mathrm{n}(u) + v^2 c). \end{aligned}$$

Since $\mathrm{tr}(\bar{u}\,b\,v) + a\,\mathrm{n}(u) + v^2 c \in \mathbb{Z}$ by the properties of the reduced norm, the reduced trace and the conjugate of elements of $\mathbb{O}$, this implies that the integer a divides the nonzero integer $v^2\Delta - \mathrm{n}(w)$; hence a is bounded, as wanted.

This concludes the proof of Theorem 24. □

Corollary 25. *For every* $\Delta \in \mathbb{Z} - \{0\}$, *the number of orbits of* $\mathrm{SL}_2(\mathcal{O})$ *in* $\mathcal{Q}(\mathcal{O}, \Delta)$, *hence in* $\mathcal{Q}^+(\mathcal{O}, \Delta)$ *and in* $\mathcal{Q}^\pm(\mathcal{O}, \Delta)$, *is finite.*

Proof. This immediately follows from Theorem 24, by the equivariance properties in Proposition 22 and the assumption (i) on a weak fundamental domain (that was not used in the proof of Theorem 24). □

Example 26. Let A be Hamilton's quaternion algebra over $\mathbb{Q}$. Let $\mathcal{O}$ be Hurwitz's maximal order in A, and let $\mathcal{O}' = \mathbb{Z}+\mathbb{Z}i+\mathbb{Z}j+\mathbb{Z}k$ be the order of Lipschitz integral quaternions.

We identify $\mathbb{H}$ and $\mathbb{R}^4$ by the $\mathbb{R}$-linear map sending $(1, i, j, k)$ to the canonical basis of $\mathbb{R}^4$. Let $V \subset \mathcal{O}'$ denote the set of vertices of the 4-dimensional unit cube $[0, 1]^4$. We claim that the set

$$\mathcal{F} = \{(z, r) \in \mathbb{H}^5_{\mathbb{R}} : z \in [0, 1]^4,\ \mathrm{n}(z - s) + r^2 \geq 1 \text{ for all } s \in V\}$$

is a weak fundamental domain for $\mathrm{SL}_2(\mathcal{O}')$, and hence for $\mathrm{SL}_2(\mathcal{O})$. For every $s \in V$, the 3-sphere in $\mathbb{H}$ with equation $\mathrm{n}(z - s) = 1$ is the isometric sphere of

$$\begin{pmatrix} 0 & -1 \\ 1 & s \end{pmatrix} \in \mathrm{SL}_2(\mathcal{O}').$$

Since the diameter of the cube $[0, 1]^4$ is 2, the closed balls bounded by these spheres cover $[0, 1]^4$. This unit cube is a fundamental polytope of the subgroup of unipotent elements of $\mathrm{SL}_2(\mathcal{O}')$ fixing ∞. Thus, $\mathcal{F}$ contains a Ford fundamental domain of $\mathrm{SL}_2(\mathcal{O}')$, which implies property (i) of a weak fundamental domain. Property (ii) (with K the unit cube) is valid by the definition of $\mathcal{F}$. Property (iii) follows from the fact that the only point at infinity of $\mathcal{F}$ besides ∞ is the center point $(1+i+j+k)/2$ of the unit cube, which is the only point of this cube which does not belong to one of the open balls whose boundary is one of the isometric spheres used to define $\mathcal{F}$. Note that $(1 + i + j + k)/2 \in A$ is a parabolic fixed point of $\mathrm{SL}_2(\mathcal{O}')$.

Recall that a positive definite Hamiltonian form $f \in \mathcal{Q}^+(\mathcal{O}, \Delta)$ with coefficients $a = a(f)$, $b = b(f) = b_1 + b_2 i + b_3 j + b_4 k$ and $c = c(f)$ is reduced (for this choice of weak fundamental domain) if $(-b/a, \sqrt{-\Delta}/a) \in \mathcal{F}$. A straightforward manipulation of the defining inequalities of $\mathcal{F}$ shows that $f \in \mathcal{Q}^+(\mathcal{O}, \Delta)$ is reduced if and only if its coefficients satisfy the following set of 25 inequalities

$$a > 0, \quad 0 \leq -b_\ell \leq a, \quad a\Big(a - c - 2\sum_{m \in P} b_m\Big) \leq \mathrm{Card}(P) \tag{35}$$

for all $\ell \in \{1, 2, 3, 4\}$ and for all subsets $P \subset \{1, 2, 3, 4\}$. Theorem 24 implies that there are only a finite number of forms in $\mathcal{Q}^+(\mathcal{O}, \Delta)$ whose coefficients satisfy the inequalities (35).

Similarly, an indefinite Hamiltonian form $f \in \mathcal{Q}^{\pm}(\mathcal{O}, \Delta)$ with $a(f) = a > 0$, $b(f) = b_1 + b_2 i + b_3 j + b_4 k$ and $c(f) = c$ is reduced, that is, $\mathscr{C}(f)$ meets $\mathscr{F}$, if and only if the following system of 16 linear inequalities and one quadratic inequality in four real variables X_1, X_2, X_3, X_4 has a solution in the unit cube $[0, 1]^4$:

$$\sum_{\ell=1}^{4} 2X_\ell \frac{b_\ell}{a} + {X_\ell}^2 \le -\frac{c}{a}, \qquad \sum_{\ell=1}^{4} 2X_\ell \frac{b_\ell}{a} + \sum_{m \in P} 2X_m \le -1 - \frac{c}{a} + \mathrm{Card}(P),$$

for all subsets $P \subset \{1, 2, 3, 4\}$.

Appendix: The hyperbolic covolume of $\mathrm{SL}_2(\mathcal{O})$, by Vincent Emery

Let A be a definite quaternion algebra over $\mathbb{Q}$, with reduced discriminant D_A, and let $\mathcal{O}$ be a maximal order in A; see for instance [Vignéras 1980] and Section 2 for definitions and properties. Given a quaternion algebra A' over a field k, let $\mathrm{SL}_2(A') = \mathrm{SL}_1(\mathcal{M}_2(A'))$ be the group of elements of the central simple 2×2 matrix algebra $\mathcal{M}_2(A')$ having reduced norm 1. For any subring $\mathcal{O}'$ of A', let $\mathrm{SL}_2(\mathcal{O}') = \mathrm{SL}_2(A') \cap \mathcal{M}_2(\mathcal{O}')$ and $\mathrm{PSL}_2(\mathcal{O}') = \mathrm{SL}_2(\mathcal{O}')/\{\pm \mathrm{id}\}$. Fixing an identification between $A \otimes_{\mathbb{Q}} \mathbb{R}$ and Hamilton's real quaternion algebra $\mathbb{H}$ turns $\mathrm{SL}_2(\mathcal{O})$ into an arithmetic lattice in $\mathrm{SL}_2(\mathbb{H})$. Hence $\mathrm{SL}_2(\mathcal{O})$ acts by isometries with finite covolume on the real hyperbolic space $\mathbb{H}^5_{\mathbb{R}}$; see for instance Section 3 for generalities.

In this appendix, the following result is proved using Prasad's volume formula in [Prasad 1989]. See the main body of this paper for a proof using Eisenstein series.

Theorem A.1. *The hyperbolic covolume of* $\mathrm{SL}_2(\mathcal{O})$ *is*

$$\mathrm{Covol}(\mathrm{SL}_2(\mathcal{O})) = \frac{\zeta(3)}{11520} \prod_{p \mid D_A} (p^3 - 1)(p - 1),$$

where p ranges over the prime integers.

Proof. Let $\mathscr{P}$ be the set of positive primes in $\mathbb{Z}$. For every $p \in \mathscr{P}$, let $\mathcal{O}_p = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Z}_p$, which is a maximal order in the quaternion algebra $A_p = A \otimes_{\mathbb{Q}} \mathbb{Q}_p$ over $\mathbb{Q}_p$; see for instance [Vignéras 1980, page 84].

We refer for instance to [Tits 1966] for the classification of the semisimple connected algebraic groups over $\mathbb{Q}$. Let $\mathbf{G}$ be the (affine) algebraic group over $\mathbb{Q}$, having as its group of K-points, for each characteristic zero field K, the group

$$\mathbf{G}(K) = \mathrm{SL}_2(A \otimes_{\mathbb{Q}} K) = \mathrm{SL}_1(\mathcal{M}_2(A \otimes_{\mathbb{Q}} K)).$$

The group $\mathbf{G}$ is absolutely (quasi)simple and simply connected. Indeed, the $\mathbb{C}$-algebra $A \otimes_{\mathbb{Q}} \mathbb{C}$ is isomorphic to $\mathcal{M}_2(\mathbb{C})$ and thus the complex Lie group $\mathbf{G}(\mathbb{C})$ is isomorphic to $\mathrm{SL}_1(\mathcal{M}_4(\mathbb{C})) = \mathrm{SL}_4(\mathbb{C})$ (note that we are using the reduced norm and not the norm). Furthermore, $\mathbf{G}$ is an inner form of the split algebraic group $\mathscr{G} = \mathrm{SL}_4$ over $\mathbb{Q}$. The (absolute) rank of $\mathscr{G}$ and the exponents of $\mathscr{G}$ are given by

$$r = 3 \quad \text{and} \quad m_1 = 1,\ m_2 = 2,\ m_3 = 3; \tag{A1}$$

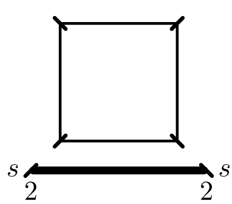
see for instance [Prasad 1989, page 96]. We consider the $\mathbb{Z}$-form of $\mathbf{G}$ such that $\mathbf{G}(\mathbb{Z}) = \mathrm{SL}_2(\mathcal{O})$ and $\mathbf{G}(\mathbb{Z}_p) = \mathrm{SL}_2(\mathcal{O}_p)$ for every $p \in \mathcal{P}$; see for instance [Parkkonen and Paulin 2010, page 382] for details.

Let $\mathcal{I}_{\mathbf{G},\mathbb{Q}_p}$ be the Bruhat–Tits building of $\mathbf{G}$ over $\mathbb{Q}_p$; see for instance [Tits 1979] for the necessary background on Bruhat–Tits theory. Recall that a subgroup of $\mathbf{G}(\mathbb{Q}_p)$ is *parahoric* if it is the stabilizer of a simplex of $\mathcal{I}_{\mathbf{G},\mathbb{Q}_p}$; a *coherent family of parahoric subgroups* of $\mathbf{G}$ is a family $(Y_p)_{p\in\mathcal{P}}$, where Y_p is a parahoric subgroup of $\mathbf{G}(\mathbb{Q}_p)$ and $Y_p = \mathbf{G}(\mathbb{Z}_p)$ for p big enough. The *principal lattice* associated with this family is the subgroup $\mathbf{G}(\mathbb{Q}) \cap \prod_p Y_p$ of $\mathbf{G}(\mathbb{Q})$ (diagonally contained in the group $\mathbf{G}(\mathbb{A}_f) = \prod'_p \mathbf{G}(\mathbb{Q}_p)$ of finite adèles of $\mathbf{G}$, where as usual $\prod'$ indicates the restricted product).

For every $p \in \mathcal{P}$, recall that by the definition of the discriminant D_A of A, if p does not divide D_A, then the algebra A_p is isomorphic to $\mathcal{M}_2(\mathbb{Q}_p)$, and otherwise A_p is a d^2-dimensional central division algebra with center $\mathbb{Q}_p$ with $d = 2$. Furthermore, for the discrete valuation $\nu = \nu_p \circ \mathrm{n}$, where ν_p is the discrete valuation of $\mathbb{Q}_p$ and n the reduced norm on A_p, the maximal order $\mathcal{O}_p$ is equal to the valuation ring of ν; see for instance [Vignéras 1980, page 34].

First assume that p does not divide D_A. Then $\mathbf{G}$ is isomorphic to $\mathscr{G} = \mathrm{SL}_4$ over $\mathbb{Q}_p$. The vertices of the building $\mathcal{I}_{\mathbf{G},\mathbb{Q}_p}$ are the homothety classes of $\mathbb{Z}_p$-lattices in ${\mathbb{Q}_p}^4$. In particular $\mathrm{SL}_2(\mathcal{O}_p) = \mathrm{SL}_4(\mathbb{Z}_p)$ is the stabilizer of the class of the standard $\mathbb{Z}_p$-lattice ${\mathbb{Z}_p}^4$ and hence is parahoric.

Now assume that p divides D_A. Then $\mathbf{G}(\mathbb{Q}_p) = \mathrm{SL}_m(A_p)$ with $m = 2$ and $\mathbf{G}(\mathbb{Q}_p)$ has local type ${}^dA_{md-1} = {}^2A_3$ in Tits' classification [1979, Section 4.4]. The corresponding local index is shown below:



Local index of type ${}^2A_3$.

The building $\mathcal{I}_{\mathbf{G},\mathbb{Q}_p}$ is a tree (see for instance [Serre 1977] for the construction of the Bruhat–Tits tree of $\mathrm{SL}_2(K)$ even when K is a noncommutative division algebra endowed with a discrete valuation). Its vertices are the homothety classes of

$\mathbb{O}_p$-lattices in the right A_p-vector space ${A_p}^2$. In particular $\mathrm{SL}_2(\mathbb{O}_p)$ is the stabilizer of the class of the standard $\mathbb{O}_p$-lattice ${\mathbb{O}_p}^2$, hence is parahoric.

Therefore, by definition, the family $(\mathrm{SL}_2(\mathbb{O}_p))_{p\in\mathcal{P}}$ is a coherent family of (maximal) parahoric subgroups of $\mathbf{G}$, and $\mathrm{SL}_2(\mathbb{O}) = \mathbf{G}(\mathbb{Z}) = \mathbf{G}(\mathbb{Q}) \cap \prod_{p\in\mathcal{P}} \mathbf{G}(\mathbb{Z}_p)$ is its associated principal lattice.

For every $p \in \mathcal{P}$, let $\overline{M}_p$ (respectively $\overline{\mathcal{M}}_p$) be the maximal reductive quotient, defined over the residual field $\mathbb{F}_p = \mathbb{Z}_p/p\mathbb{Z}_p$, of the identity component of the reduction modulo p of the smooth affine group scheme over $\mathbb{Z}_p$ associated with the vertex of $\mathcal{I}_{\mathbf{G},\mathbb{Q}_p}$ (respectively $\mathcal{I}_{\mathcal{G},\mathbb{Q}_p}$) stabilized by the parahoric subgroup $\mathrm{SL}_2(\mathbb{O}_p)$ (respectively $\mathrm{SL}_4(\mathbb{Z}_p)$); see for instance [Tits 1979, Section 3.5]. Note that $\overline{M}_p = \overline{\mathcal{M}}_p$ if p does not divide D_A, and that for every $p \in \mathcal{P}$ the algebraic group $\overline{\mathcal{M}}_p$ is isomorphic to SL_4 over $\mathbb{F}_p$. In particular $\overline{\mathcal{M}}_p(\mathbb{F}_p) = \mathrm{SL}_4(\mathbb{F}_p)$ and thus, for every $p \in \mathcal{P}$, the orders of finite groups of Lie type being listed for example in [Ono 1966, Table 1], we have

$$\dim \overline{\mathcal{M}}_p = 15 \quad \text{and} \quad |\overline{\mathcal{M}}_p(\mathbb{F}_p)| = p^6(p^2-1)(p^3-1)(p^4-1). \tag{A2}$$

If p divides D_A, by applying the theory in [Tits 1979, §3.5.2] on the local index ${}^2A_3$, we see that the semisimple part $\overline{M}_p^{ss}$ of $\overline{M}_p$ (given as the commutator algebraic group $[\overline{M}_p, \overline{M}_p]$) is of type ${}^2(A_1 \times A_1)$ and the radical $R(\overline{M}_p)$ of $\overline{M}_p$ must be a one-dimensional nonsplit torus over $\mathbb{F}_p$. In particular $|R(\overline{M}_p)(\mathbb{F}_p)| = p+1$ and $\overline{M}_p^{ss}(\mathbb{F}_p)$ has the same order as $\mathrm{SL}_2(\mathbb{F}_{p^2})$, that is, $p^2(p^4-1)$. Since the radical $R(\overline{M}_p)$ is central in $\overline{M}_p$ and the intersection $R(\overline{M}_p) \cap \overline{M}_p^{ss}$ is finite (see [Springer 1998, Proposition 7.3.1]), the product map

$$\overline{M}_p^{ss} \times R(\overline{M}_p) \to \overline{M}_p, \quad (x, y) \mapsto xy$$

is an isogeny (defined over $\mathbb{F}_p$) and using Lang's isogeny theorem (see for example [Platonov and Rapinchuk 1994, Proposition 6.3, page 290]), we obtain the order of $\overline{M}_p(\mathbb{F}_p)$ as the product $|\overline{M}_p^{ss}(\mathbb{F}_p)| \cdot |R(\overline{M}_p)(\mathbb{F}_p)|$.

Alternatively, the order of $\overline{M}_p(\mathbb{F}_p)$ can be deduced from the concrete structure of $\overline{M}_p$ given in [Bruhat and Tits 1984]. Namely, it follows from [ibid., Proposition 3.11 and Section 5.5] that $\overline{M}_p(\mathbb{F}_p)$ corresponds to the group of elements of reduced norm 1 in the $\mathbb{F}_p$-algebra $\mathcal{M}_2(\mathbb{F}_{p^2})$ (where $\mathbb{F}_{p^2}$ appears as the residue field of the division algebra A_p; see [Vignéras 1980, page 35]). The reduced norm (over $\mathbb{F}_p$) of an element $g \in \mathcal{M}_2(\mathbb{F}_{p^2})$ is $N_{\mathbb{F}_p^2|\mathbb{F}_p}(\det(g))$, where $N_{\mathbb{F}_{p^2}|\mathbb{F}_p}$ is the norm of the extension $\mathbb{F}_{p^2}|\mathbb{F}_p$. Thus $\overline{M}_p(\mathbb{F}_p)$ is the kernel of the surjective homomorphism $\mathrm{GL}_2(\mathbb{F}_{p^2}) \to \mathbb{F}_p^\times$ defined by $g \mapsto \det(g)^{p+1}$.

Therefore, from any of the two arguments above, we obtain that for every $p \in \mathcal{P}$ dividing D_A,

$$\dim \overline{M}_p = 7 \quad \text{and} \quad |\overline{M}_p(\mathbb{F}_p)| = p^2(p^4-1)(p+1). \tag{A3}$$

Let μ be the Haar measure on $\mathbf{G}(\mathbb{R}) = \mathrm{SL}_2(\mathbb{H})$ normalized as in [Prasad 1989]. That is, if w is the top degree exterior form on the real Lie algebra of $\mathbf{G}(\mathbb{R})$ whose associated invariant differential form on $\mathbf{G}(\mathbb{R})$ defines the measure μ and if $\mathbf{G}_u(\mathbb{R})$ is a compact real form of $\mathbf{G}(\mathbb{C})$, then the complexification $w_\mathbb{C}$ of w on the complex Lie algebra of $\mathbf{G}(\mathbb{C}) = \mathbf{G}_u(\mathbb{C})$ defines a top degree exterior form w_u on the real Lie algebra of $\mathbf{G}_u(\mathbb{R})$, whose associated invariant differential form on $\mathbf{G}_u(\mathbb{R})$ defines a measure μ_u, and we require that $\mu_u(\mathbf{G}_u(\mathbb{R})) = 1$.

Let μ' be the Haar measure on $\mathrm{PSL}_2(\mathbb{H}) = \mathrm{SO}_0(1,5)$ that disintegrates by the fibration $\mathrm{SO}_0(1,5) \to \mathrm{SO}_0(1,5)/\mathrm{SO}(5) = \mathbb{H}^5_\mathbb{R}$ with measures on the fibers of total mass one 1 and measure on the base the Riemannian measure $d\mathrm{vol}_{\mathbb{H}^5_\mathbb{R}}$ of the Riemannian metric of constant sectional curvature -1. Let $\tilde{\mu}'$ be the Haar measure on $\mathrm{SL}_2(\mathbb{H})$ such that the tangent map at the identity of the double cover of real Lie groups $\mathrm{SL}_2(\mathbb{H}) \to \mathrm{PSL}_2(\mathbb{R})$ preserves the top degree exterior forms defining the Haar measures. In particular, since $-\mathrm{id}$ belongs to $\mathrm{SL}_2(\mathbb{O})$,

$$\begin{aligned}\mathrm{Covol}(\mathrm{SL}_2(\mathbb{O})) &= \mathrm{Vol}(\mathrm{PSL}_2(\mathbb{O})\backslash\mathbb{H}^5_\mathbb{R})\\ &= \mu'(\mathrm{PSL}_2(\mathbb{O})\backslash \mathrm{PSL}_2(\mathbb{H})) = \tilde{\mu}'(\mathrm{SL}_2(\mathbb{O})\backslash \mathrm{SL}_2(\mathbb{H})).\end{aligned} \tag{A4}$$

Similarly, with $\mathbb{S}_5$ the 5-sphere endowed with its standard Riemannian metric of constant sectional curvature $+1$, let μ'_u be the Haar measure on $\mathrm{SO}(6)$ that disintegrates by the fibration $\mathrm{SO}(6) \to \mathrm{SO}(6)/\mathrm{SO}(5) = \mathbb{S}_5$ with measures on the fibers of total mass one 1 and measure on the base the Riemannian measure. In particular, $\mu'_u(\mathrm{SO}(6)) = \mathrm{Vol}(\mathbb{S}_5)$. Recall that

$$\mathrm{Vol}(\mathbb{S}_n) = \frac{2\pi^m}{(m-1)!} \quad \text{if } n = 2m-1 \geq 3.$$

It is well known (see for instance [Helgason 1978]) that the duality $G/K \mapsto G_u/K$ between irreducible symmetric spaces of noncompact type endowed with a left invariant Riemannian metric and the ones of compact type, where G_u is a compact form of the complexification of G, sends $\mathbb{H}^5_\mathbb{R}$ to $\mathbb{S}_5$, and hence μ' to μ'_u.

The maximal compact subgroup $\mathrm{SU}(4)$ of $\mathrm{SL}_4(\mathbb{C})$ is a covering of degree 2 of $\mathrm{SO}(6)$, which is the compact real form corresponding to $\mathrm{SO}_0(1,5)$. Hence we have (as first proved in [Emery 2009, Section 13.3])

$$\tilde{\mu}' = 2\,\mathrm{Vol}(\mathbb{S}_5)\mu = 2\pi^3\mu. \tag{A5}$$

By Prasad's volume formula [Prasad 1989, Theorem 3.7] (where with the notation of this theorem, $\ell = k = \mathbb{Q}$ (hence $D_k = D_\ell = 1$), $S = V_\infty = \{\infty\}$ and the Tamagawa number $\tau_\mathbb{Q}(\mathbf{G})$ is 1), we have, since $\overline{\mathcal{M}}_p = \overline{M}_p$ if p does not divide D_A and by (A1) for the second equality,

$$\mu(\mathrm{SL}_2(\mathbb{O})\backslash \mathrm{SL}_2(\mathbb{H})) = \prod_{i=1}^{r} \frac{(m_i)!}{(2\pi)^{m_i+1}} \prod_{p\in\mathcal{P}} \frac{p^{(\dim \overline{M}_p+\dim \overline{\mathcal{M}}_p)/2}}{|\overline{M}_p(\mathbb{F}_p)|}$$

$$= \frac{12}{(2\pi)^9} \prod_{p\in\mathcal{P}} \frac{p^{\dim \overline{\mathcal{M}}_p}}{|\overline{\mathcal{M}}_p(\mathbb{F}_p)|} \prod_{p|D_A} \frac{|\overline{\mathcal{M}}_p(\mathbb{F}_p)|}{|\overline{M}_p(\mathbb{F}_p)|} p^{(\dim \overline{M}_p-\dim \overline{\mathcal{M}}_p)/2}. \tag{A6}$$

Using Euler's product formula $\zeta(s) = \prod_{p\in\mathcal{P}} 1/(1-p^{-s})$ for Riemann's zeta function, we have by (A2), since $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$,

$$\prod_{p\in\mathcal{P}} \frac{p^{\dim \overline{\mathcal{M}}_p}}{|\overline{\mathcal{M}}_p(\mathbb{F}_p)|} = \zeta(2)\zeta(3)\zeta(4) = \frac{\pi^6\zeta(3)}{540}. \tag{A7}$$

Using the equations (A4), (A5), (A6), (A7), (A2) and (A3), the result follows. □

Acknowledgments

J. Parkkonen and F. Paulin thank P. Sarnak for his comments on the origin of volume computations using Eisenstein series, G. Chenevier for the proof of Lemma 6, Y. Benoist and F. Choucroun for discussions related to the appendix, and the referee for helpful comments, in particular for Lemma 23. F. Paulin thanks the University of Jyväskylä for the nice snow and its financial support. V. Emery thanks J. Parkkonen and F. Paulin for helpful discussions. He is particularly grateful to F. Paulin for his help concerning the Bruhat–Tits buildings appearing in the proof of Theorem A.1.

References

[Aslaksen 1996] H. Aslaksen, "Quaternionic determinants", *Math. Intelligencer* **18** (1996), 57–65. MR 97j:16028 Zbl 0881.15007

[Beardon 1983] A. F. Beardon, *The geometry of discrete groups*, Grad. Texts in Math. **91**, Springer, 1983. MR 85d:22026 Zbl 0528.30001

[Borel 1966] A. Borel, "Reduction theory for arithmetic groups", pp. 20–25 in *Algebraic groups and discontinuous subgroups* (Boulder, 1965), edited by A. Borel and G. D. Mostow, Proc. Sympos. Pure Math. **IX**, American Mathematical Society, 1966. MR 34 #4372 Zbl 0213.47201

[Borel and Harish-Chandra 1962] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups", *Ann. of Math.* **75** (1962), 485–535. MR 26 #5081 Zbl 0107.14804

[Bourbaki 1959] N. Bourbaki, *Algèbre Chapitre 9: Formes sesquilinéaires et formes quadratiques*, Hermann, 1959. MR 0107661 Zbl 0102.25503

[Breulmann and Helmke 1996] S. Breulmann and V. Helmke, "The covolume of quaternion groups on the four-dimensional hyperbolic space", *Acta Arith.* **77** (1996), 9–21. MR 97i:11051 Zbl 0848.11018

[Bruhat and Tits 1972] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local", *Publ. Math. Inst. Hautes Études Sci.* **41** (1972), 5–251. MR 48 #6265 Zbl 0254.14017

[Bruhat and Tits 1984] F. Bruhat and J. Tits, “Schémas en groupes et immeubles des groupes classiques sur un corps local”, *Bull. Soc. Math. France* **112** (1984), 259–301. MR 86i:20064 Zbl 0565.14028

[Buchmann and Vollmer 2007] J. Buchmann and U. Vollmer, *Binary quadratic forms: An algorithmic approach*, Algorithms Comput. Math. **20**, Springer, 2007. MR 2008b:11046 Zbl 1125.11028

[Cassels 1978] J. W. S. Cassels, *Rational quadratic forms*, London Math. Soc. Monogr. **13**, Academic Press, 1978. Zbl 0395.10029

[Deuring 1968] M. Deuring, *Algebren*, 2nd ed., Ergeb. Math. Grenzgeb. **41**, Springer, 1968. MR 37 #4106 Zbl 0159.04201

[Dieudonné 1943] J. Dieudonné, “Les déterminants sur un corps non commutatif”, *Bull. Soc. Math. France* **71** (1943), 27–45. MR 7,3a Zbl 0028.33904

[Eichler 1938] M. Eichler, “Über die Idealklassenzahl total definiter Quaternionenalgebren”, *Math. Z.* **43** (1938), 102–109. MR 1545717 JFM 3.0093.02

[Elstrodt et al. 1998] J. Elstrodt, F. Grunewald, and J. Mennicke, *Groups acting on hyperbolic space: Harmonic analysis and number theory*, Springer, 1998. MR 98g:11058 Zbl 0888.11001

[Emery 2009] V. Emery, *Du volume des quotients arithmétiques de l’espace hyperbolique*, thèse n^{o} 1648, Université de Fribourg (Suisse), 2009, available at http://www.unige.ch/math/folks/emery/Emery.pdf.

[Eskin et al. 1991] A. Eskin, Z. Rudnick, and P. Sarnak, “A proof of Siegel’s weight formula”, *Internat. Math. Res. Notices* **1991**:5 (1991), 65–69. MR 92m:11040 Zbl 0743.11023

[Hashimoto and Ibukiyama 1980] K.-i. Hashimoto and T. Ibukiyama, “On class numbers of positive definite binary quaternion Hermitian forms, I”, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **27** (1980), 549–601. MR 82j:10038 Zbl 0452.10029

[Hashimoto and Ibukiyama 1981] K.-i. Hashimoto and T. Ibukiyama, “On class numbers of positive definite binary quaternion Hermitian forms, II”, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **28** (1981), 695–699. MR 83m:10029 Zbl 0493.10030

[Hashimoto and Ibukiyama 1983] K.-i. Hashimoto and T. Ibukiyama, “On class numbers of positive definite binary quaternion Hermitian forms, III”, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **30** (1983), 393–401. MR 85i:11030 Zbl 0533.10019

[Helgason 1978] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Math. **80**, Academic Press, 1978. MR 80k:53081 Zbl 0451.53038

[Hild 2007] T. Hild, “The cusped hyperbolic orbifolds of minimal volume in dimensions less than ten”, *J. Algebra* **313** (2007), 208–222. MR 2008g:57038 Zbl 1119.52011

[Johnson and Weiss 1999] N. W. Johnson and A. I. Weiss, “Quaternionic modular groups”, *Linear Algebra Appl.* **295** (1999), 159–189. MR 2000j:20096 Zbl 0960.20031

[Kellerhals 2003] R. Kellerhals, “Quaternions and some global properties of hyperbolic 5-manifolds”, *Canad. J. Math.* **55** (2003), 1080–1099. MR 2005b:57032 Zbl 1054.57019

[Krafft and Osenberg 1990] V. Krafft and D. Osenberg, “Eisensteinreihen für einige arithmetisch definierte Untergruppen von $SL_2(\mathbf{H})$”, *Math. Zeit.* **204** (1990), 425–449. MR 92f:11067 Zbl 0725.11024

[Langlands 1966] R. P. Langlands, “The volume of the fundamental domain for some arithmetical subgroups of Chevalley groups”, pp. 143–148 in *Algebraic groups and discontinuous subgroups* (Boulder, 1965), edited by A. Borel and G. D. Mostow, Proc. Sympos. Pure Math. **IX**, American Math. Soc., 1966. MR 35 #4226 Zbl 0218.20041

[Magnus et al. 1966] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and theorems for the special functions of mathematical physics*, 3rd ed., Grund. math. Wiss. **52**, Springer, 1966. MR 38 #1291 Zbl 0143.08502

[Ono 1966] T. Ono, “On algebraic groups and discontinuous groups”, *Nagoya Math. J.* **27** (1966), 279–322. MR 33 #7342 Zbl 0166.29802

[Parkkonen and Paulin 2010] J. Parkkonen and F. Paulin, “Prescribing the behaviour of geodesics in negative curvature”, *Geom. Topol.* **14** (2010), 277–392. MR 2011a:53060 Zbl 1191.53026

[Parkkonen and Paulin 2011] J. Parkkonen and F. Paulin, “On the representation of integers by indefinite binary Hermitian forms”, *Bull. Lond. Math. Soc.* **43** (2011), 1048–1058. MR 2861527 Zbl 05989570

[Parkkonen and Paulin 2012] J. Parkkonen and F. Paulin, “Équidistribution, comptage et approximation par irrationnels quadratiques”, *J. Mod. Dyn.* **6** (2012), 1–40. MR 2929128 Zbl 06049546

[Platonov and Rapinchuk 1994] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics **139**, Academic Press, 1994. MR 95b:11039 Zbl 0841.20046

[Prasad 1989] G. Prasad, “Volumes of S-arithmetic quotients of semi-simple groups”, *Publ. Math. Inst. Hautes Études Sci.* **69** (1989), 91–117. MR 91c:22023 Zbl 0695.22005

[Pronin 1967] L. N. Pronin, “Integral binary Hermitian forms over the skewfield of quaternions”, *Vestnik Har’kov. Gos. Univ.* **1967**:26 (1967), 27–41. In Russian. MR 40 #2606 Zbl 0258.10008

[Rankin 1939a] R. A. Rankin, “Contributions to the theory of Ramanujan’s function $\tau(n)$ and similar arithmetical functions, I: The zeros of the function $\sum_{n=1}^{\infty} \tau(n)/n^s$ on the line $\Re s = 13/2$”, *Proc. Cambridge Philos. Soc.* **35** (1939), 351–356. MR 1,69d Zbl 0021.39201

[Rankin 1939b] R. A. Rankin, “Contributions to the theory of Ramanujan’s function $\tau(n)$ and similar arithmetical functions, II: The order of the Fourier coefficients of integral modular forms”, *Proc. Cambridge Philos. Soc.* **35** (1939), 357–372. MR 1,69d Zbl 0021.39202

[Reiner 1975] I. Reiner, *Maximal orders*, London Math. Soc. Monogr. **5**, Academic Press, 1975. MR 52 #13910 Zbl 0305.16001

[Sarnak 1983] P. Sarnak, “The arithmetic and geometry of some hyperbolic three-manifolds”, *Acta Math.* **151** (1983), 253–295. MR 85d:11061

[Schoeneberg 1939] B. Schoeneberg, “Über die ζ-Funktion einfacher hyperkomplexer Systeme”, *Math. Ann.* **117** (1939), 85–88. MR 1,203e Zbl 0021.38802

[Selberg 1940] A. Selberg, “Bemerkungen über eine Dirichletsche Reihe, die mit der Theorie der Modulformen nahe verbunden ist”, *Arch. Math. Naturvid.* **43** (1940), 47–50. MR 2,88a JFM 66.0377.01

[Serre 1977] J.-P. Serre, *Arbres, amalgames,* SL_2, Astérisque **46**, Soc. Math. de France, 1977. MR 57 #16426 Zbl 0369.20013

[Springer 1998] T. A. Springer, *Linear algebraic groups*, 2nd ed., Prog. Math. **9**, Birkhäuser, Boston, MA, 1998. MR 99h:20075 Zbl 0927.20024

[Tits 1966] J. Tits, “Classification of algebraic semisimple groups”, pp. 33–62 in *Algebraic groups and discontinuous subgroups* (Boulder, 1965), edited by A. Borel and G. D. Mostow, Proc. Sympos. Pure Math. **IX**, American Mathematical Society, 1966. MR 37 #309 Zbl 0238.20052

[Tits 1979] J. Tits, “Reductive groups over local fields”, pp. 29–69 in *Automorphic forms, representations and L-functions, I* (Corvallis, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **XXXIII**, American Mathematical Society, 1979. MR 80h:20064

[Vignéras 1980] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Mathematics **800**, Springer, 1980. MR 82i:12016 Zbl 0422.12008

[Weyl 1940] H. Weyl, "Theory of reduction for arithmetical equivalence, I", *Trans. Amer. Math. Soc.* **48** (1940), 126–164. MR 2,35h Zbl 0024.14802

[Weyl 1942] H. Weyl, "Theory of reduction for arithmetical equivalence, II", *Trans. Amer. Math. Soc.* **51** (1942), 203–231. MR 2,35h Zbl 0028.01201

[Zagier 1981] D. B. Zagier, *Zetafunktionen und quadratische Körper: Eine Einführung in die höhere Zahlentheorie*, Springer, 1981. MR 82m:10002 Zbl 0459.10001

Communicated by Marie-France Vignéras
Received 2011-05-11 Revised 2011-12-14 Accepted 2012-01-30

parkkone@maths.jyu.fi *Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35, FI-40014 University of Jyväskylä, Finland*

frederic.paulin@math.u-psud.fr *Département de mathématique, UMR 8628 CNRS, Université Paris-Sud, Bât. 425, 91405 ORSAY Cedex, France*

vincent.emery@gmail.com *Section de mathématiques, 2-4 rue du Lièvre, Case postale 64, 1211 Genève 4, Switzerland*

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.117

L-functions and periods of adjoint motives

Michael Harris

The article studies the compatibility of the refined Gross–Prasad (or Ichino–Ikeda) conjecture for unitary groups, due to Neal Harris, with Deligne's conjecture on critical values of L-functions. When the automorphic representations are of motivic type, it is shown that the L-values that arise in the formula are critical in Deligne's sense, and their Deligne periods can be written explicitly as products of Petersson norms of arithmetically normalized coherent cohomology classes. In some cases this can be used to verify Deligne's conjecture for critical values of adjoint type (Asai) L-functions.

Introduction

The refined Gross–Prasad conjecture, or Ichino–Ikeda conjecture, is an explicit and exact expression for certain products of special values of automorphic L-functions in terms of automorphic periods. In the situation of the present article, π and π' are automorphic representations of unitary groups $U(W)$ and $U(W')$, respectively, where W is a hermitian space of dimension n over a CM field $\mathcal{K}$ and $W' \subset W$ is a nondegenerate hermitian subspace of codimension 1. We assume π and π' admit base change to automorphic representations $BC(\pi)$ and $BC(\pi')$ of $GL(n, \mathcal{K})$ and $GL(n-1, \mathcal{K})$, respectively. The original Ichino–Ikeda conjecture is stated for inclusions of special orthogonal groups; the version for unitary groups, due to Neal Harris [N. Harris 2011], gives a formula for the quotient

$$\frac{L\left(\frac{1}{2}, BC(\pi) \times BC(\pi')\right)}{L(1, \pi, \mathrm{Ad})L(1, \pi', \mathrm{Ad})} \tag{0.1}$$

in terms of global periods, local integrals, and some elementary terms (for details, see Section 2.1). Here the numerator is a Rankin–Selberg tensor product L-function for $GL(n) \times GL(n-1)$, and the L-functions attached to the adjoint

Institut des Mathématiques de Jussieu, U.M.R. 7586 du CNRS. Membre, Institut Universitaire de France. This work was partially supported by the Association of Members of the Institute for Advanced Study.
MSC2010: primary 11F67; secondary 11F70, 14G35, 11G09.
Keywords: adjoint L-functions, automorphic forms, motives, Ichino–Ikeda conjecture, periods.

representations of the L-groups of unitary groups can be identified with the *Asai L-functions* $L(s, \mathrm{BC}(\pi), \mathrm{As}^{\pm})$, $L(1, \mathrm{BC}(\pi'), \mathrm{As}^{\mp})$ of the conjugate self-dual representations $\mathrm{BC}(\pi)$, $\mathrm{BC}(\pi')$ as follows (see [N. Harris 2011, Remark 1.4; Gan et al. 2012a, Proposition 7.4]):

$$\begin{aligned} L(s, \pi, \mathrm{Ad}) &= L(s, \mathrm{BC}(\pi), \mathrm{As}^{(-1)^n}), \\ L(s, \pi', \mathrm{Ad}) &= L(s, \mathrm{BC}(\pi), \mathrm{As}^{(-1)^{n-1}}). \end{aligned} \tag{0.2}$$

In its formulation for special orthogonal groups, the Ichino–Ikeda conjecture is inspired by formulas for the central values of L-functions of GL(2), due to Waldspurger [1985] and others, and represents the culmination of several decades of work in connection with the Birch–Swinnerton-Dyer conjecture, including various attempts to generalize the Gross–Zagier formula. It is natural to focus on the central value in the numerator in the Ichino–Ikeda conjecture, and to view the L-values in the denominator as error terms. The present paper is instead primarily concerned with the denominator.

In what follows, when π is attached to a motive M of rank n over a number field, the value $L(1, \pi, \mathrm{Ad}) = L(s, \mathrm{BC}(\pi), \mathrm{As}^{(-1)^n})$ is critical in Deligne's sense [1979a], and is expected to be closely connected to the classification of p-adic deformations of the mod p Galois representations attached to M. For $n = 2$ this principle is well understood and there are very precise results due to Hida [1981], Diamond–Flach–Guo [2004], and Dimitrov [2009]. This is the first of a series of papers whose goal is to indicate a way to prove similar results for $n > 2$. The approach suggested here is heuristic and speculative, inasmuch as the Ichino–Ikeda conjecture has only been proved in special cases,[1] and a number of the steps rely on nonvanishing results for special values of L-functions, and ergodicity results for automorphic periods, that have yet to be studied seriously. Nevertheless, the Ichino–Ikeda conjecture, in conjunction with Deligne's conjecture on critical values of L-functions, indicates the existence of structural links between congruences among automorphic forms and the divisibility of the value $L(1, \pi, \mathrm{Ad})$, and these links seem worth exploring.

The function $L(s, \pi, \mathrm{Ad})$ is interpreted as the L-function of the Asai motive $\mathrm{As}^{(-1)^n}(M)$ attached to M. The present paper introduces the family of cohomological realizations that should be attached to the conjectural object $\mathrm{As}^{(-1)^n}(M)$ and explains how to relate them to automorphic forms. The main results interpret the Deligne period of $\mathrm{As}^{(-1)^n}(M)$ in terms of coherent cohomological automorphic forms, and show how the Ichino–Ikeda conjecture can be used to prove a version of Deligne's conjecture for the critical value $L(1, \pi, \mathrm{Ad}) = L(1, \mathrm{As}^{(-1)^n}(M))$,

[1] Added in proof: Since this paragraph was written, Wei Zhang has made remarkable progress on the conjecture, especially on the case considered in the final section of this paper. I will be returning to this question in forthcoming work with Harald Grobner.

assuming certain nonvanishing conjectures for twists of standard L-functions of unitary groups by finite order characters. Heuristic evidence for the nonvanishing conjectures is provided by the existence of p-adic L-functions: when π varies in a Hida family of ordinary automorphic representations with global root number $+1$, the p-adic L-function of the family is generically nonzero at the central critical point. Although the foundations are largely available for general CM fields, the main applications of the present article are limited to the case where $\mathcal{K}$ is a quadratic imaginary field and n is even; this provides for some simplification of the main formulas, while presenting the general picture. The author and L. Guerberoff hope to treat the general case in a subsequent article. Applications to congruence modules, in Hida's sense, will be treated in forthcoming joint work with C. Skinner.

The present paper can also be read as a confirmation of the compatibility between the Ichino–Ikeda conjecture and Deligne's conjecture for pairs of automorphic motives satisfying the inequalities (2.3.4), which correspond to period integrals on totally definite hermitian spaces W and W'. It appears that compatibility in general cannot be established by purely automorphic methods.

Notation and conventions

Throughout the article, we let $\mathcal{K}$ be a CM quadratic extension of a totally real field F, with $c \in \mathrm{Gal}(\mathcal{K}/F)$ complex conjugation. Let Σ_F denote the set of real places of F, and let Σ denote a CM type of $\mathcal{K}$, a set of extensions of Σ_F to $\mathcal{K}$, so that $\Sigma \coprod c \cdot \Sigma$ is the set of archimedean embeddings of $\mathcal{K}$. If $\sigma \in \Sigma_F$, we let $\sigma_{\mathcal{K}}$ denote its extension in Σ. We let $\eta_{\mathcal{K}/F} : \mathrm{Gal}(\bar{F}/F) \to \{\pm 1\}$ denote the Galois character attached to the quadratic extension $\mathcal{K}/F$.

Unless otherwise indicated, a discrete series representation of an algebraic group G over $\mathbb{R}$ will always be assumed to be *algebraic*, in the sense that its infinitesimal character is the same as that of a finite-dimensional representation. This is of course a condition on the central character.

Let E be a number field, and let $\alpha, \beta \in E \otimes_{\mathbb{Q}} \mathbb{C}$. Following Deligne, we write $\alpha \sim_E \beta$ if either $\beta \notin (E \otimes_{\mathbb{Q}} \mathbb{C})^\times$ or $\beta^{-1}\alpha \in E = E \otimes_{\mathbb{Q}} \mathbb{Q}$. In the situations that arise, if $\beta \notin (E \otimes_{\mathbb{Q}} \mathbb{C})^\times$ then we will assume $\beta = 0$.

Suppose $\mathcal{K}$ is a number field with a given embedding in $\mathbb{C}$. Then we write $\alpha \sim_{E,\mathcal{K}} \beta$ if either $\beta \notin (E \otimes_{\mathbb{Q}} \mathbb{C})^\times$ or $\beta^{-1}\alpha \in E \otimes_{\mathbb{Q}} \mathcal{K} \subset E \otimes_{\mathbb{Q}} \mathbb{C}$.

1. Deligne periods of polarized regular motives

1.1. ***Polarized regular motives over CM fields.*** Let Π be a cuspidal cohomological automorphic representation Π of $\mathrm{GL}(n, \mathcal{K})$ satisfying the polarization condition

$$\Pi^\vee \xrightarrow{\sim} \Pi^c. \tag{1.1.1}$$

Let $E = E(\Pi)$ denote a field of definition of Π_f.[2] This is a CM field [Blasius et al. 1994] and in what follows we will consider c-linear automorphisms of E-vector spaces. By the results of a number of people, collected in [Chenevier and Harris 2013], Π gives rise to a compatible system of λ-adic representations $\rho_{\Pi,\lambda} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathcal{K}) \to \mathrm{GL}(n, E_\lambda)$, where λ runs over places of E, with a nondegenerate pairing

$$\rho_{\Pi,\lambda} \otimes \rho^c_{\Pi,\lambda} \to E_\lambda(1-n). \tag{1.1.2}$$

To keep these Galois representations company, we postulate the existence of a pure motive $M = M_\Pi$ over $\mathcal{K}$ of rank n and weight $w = n-1$, with coefficients in E, whose λ-adic realization is $\rho_{\Pi,\lambda}$ and whose other realizations can be constructed using automorphic forms. For the present purposes, all we know of M is its family of realizations, together with compatibility isomorphisms. The relation between M and Π is encapsulated in the formula

$$L(s, M) = L(s + \tfrac{1}{2}(1-n), \Pi) = L(s, \Pi \otimes (|\cdot| \circ \det)^{(1-n)/2}) \tag{1.1.3}$$

Consider the motives $RM = R_{\mathcal{K}/F} M$ and $\mathcal{R}M = R_{\mathcal{K}/\mathbb{Q}} M$ over F and $\mathbb{Q}$, respectively. The base change $RM_{\mathcal{K}}$ of RM breaks up as $M \oplus M^c$, where the distinction between M and M^c depends on the choice of CM type Σ. Indeed, for each real embedding σ of F we can consider $RM_{B,\sigma}$, which can be interpreted as the topological cohomology $H^*(RM \times_{\sigma,\mathbb{C}} (\mathbb{C}), E)$; then

$$RM_{\mathcal{K},B,\sigma} = H^*(RM \times_{\sigma_{\mathcal{K}},\mathbb{C}} (\mathbb{C}), E) \oplus H^*(RM \times_{c\sigma_{\mathcal{K}},\mathbb{C}} (\mathbb{C}), E).$$

The polarization is a nondegenerate pairing

$$\langle \cdot, \cdot \rangle_B : M \otimes M^c \to E(1-n) \tag{1.1.4}$$

whereas F_∞ is just an isomorphism of Betti realizations that is linear with respect to the E-module structure:

$$F_\infty : M_B \xrightarrow{\sim} M^c_B. \tag{1.1.5}$$

We choose an E-basis $(e_1, \dots, e_n)$ of M_B and let $e^c_i = F_\infty(e_i)$ for $i = 1, \dots, n$. I refer to my paper [Harris 1997] for generalities about Deligne's conjectures [1979a] on special values of L-functions, as specialized to polarized regular motives. In that paper it is assumed $M \xrightarrow{\sim} M^c$, or equivalently that Π is a base change from F to $\mathcal{K}$, so that the superscripts ^{c} can be removed in (1.1.1) and (1.1.2). The

[2]To be completely accurate, although it is known that Π_f has a model over its field of rationality, it is not known that the motive we construct below has coefficients in the same field; for example, it has not been checked that the associated Galois representations can be realized over the λ-adic completions of $E(\Pi)$, because of the possibility of a nontrivial Brauer obstruction. So we will take $E(\Pi)$ to be a finite extension of the field of rationality of Π_f over which all the subsequent constructions are valid.

arguments in general are simple modifications of this self-dual case; however, there are roughly twice as many invariants in the general case. I follow [Harris et al. 2011], where these invariants are discussed in connection with automorphic forms on unitary groups.

The restriction of scalars $R_{\mathcal{K}/\mathbb{Q}}M_\Pi$ is naturally a motive of rank n over $\mathbb{Q}$ with coefficients in $E(\Pi)\otimes\mathcal{K}$. The de Rham realization of $R_{\mathcal{K}/\mathbb{Q}}M_\Pi$, denoted $M_{\mathcal{K}/\mathbb{Q},DR}(\Pi)$, is a free rank n module over $E(\Pi)\otimes\mathcal{K}$. The Hodge decomposition

$$M_{\mathcal{K}/\mathbb{Q},DR}(\Pi)\otimes\mathbb{C}\xrightarrow{\sim}\bigoplus_{p+q=n-1}M^{p,q}_{\mathcal{K}/\mathbb{Q}}(\Pi) \tag{1.1.6}$$

and the natural decomposition of $E(\Pi)\otimes\mathcal{K}\otimes\mathbb{C}$-modules

$$M_{\mathcal{K}/\mathbb{Q},DR}(\Pi)\otimes\mathbb{C}\xrightarrow{\sim}\bigoplus_{\sigma:E(\Pi)\otimes\mathcal{K}\to\mathbb{C}}M_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi) \tag{1.1.7}$$

are compatible with the $E(\Pi)\otimes\mathcal{K}$-action in the sense that complex conjugation c defines antilinear isomorphisms

$$c: M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi)\xrightarrow{\sim}M^{q,p}_{\mathcal{K}/\mathbb{Q},c\sigma}(\Pi) \tag{1.1.8}$$

such that

$$c(am)=c(a)c(m)\quad\text{for } a\in E(\Pi)\otimes\mathcal{K},\ m\in M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi). \tag{1.1.9}$$

Here

$$M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi)=M^{p,q}_{\mathcal{K}/\mathbb{Q}}(\Pi)\cap M_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi).$$

1.1.10 *Formal properties of polarized regular motives.* One expects the following properties to hold:

(a) For all p,q,σ, $\dim M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi)\le 1$.

(b) For all p,q, $\dim M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi)$ is independent of the restriction of σ to $E(\Pi)\otimes 1$.

(c) Let σ be as above and denote by $w\in\Sigma_{\mathcal{K}}$ its restriction to $1\otimes\mathcal{K}$, and $w^+\in\Sigma_F$ its restriction to F. Let $\mu(w)$ be the infinitesimal character of the finite-dimensional representation W_w defined in [Harris et al. 2011, Section 2.3] and let

$$p(w)=\mu(w)+\tfrac{n-1}{2}(1,1,\dots,1):=(p_1(w),p_2(w),\dots,p_n(w))$$

so that for all i, [Harris et al. 2011, (2.3.2)] implies that

$$p_i(w)+p_{n+1-i}(cw)=n-1.$$

Then $\dim M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi)=1$ if and only if $(p,q)=(p_i(w),p_{n+1-i}(cw))$ for some $i\in\boldsymbol{n}:=\{1,\dots,n\}$.

(d) The motive $R_{\mathcal{K}/\mathbb{Q}}M_\Pi$ has a nondegenerate polarization

$$\langle\,\cdot\,,\,\cdot\,\rangle : R_{\mathcal{K}/\mathbb{Q}}M_\Pi \otimes R_{\mathcal{K}/\mathbb{Q}}M_\Pi \to \mathbb{Q}(1-n)$$

that is alternating if n is even and symmetric if n is odd. The involution † on the coefficients $E(\Pi)\otimes\mathcal{K}$ induced by this polarization,

$$\langle ax, y\rangle = \langle x, a^\dagger y\rangle \quad \text{for } a \in E(\Pi)\otimes\mathcal{K} \text{ and } x, y \in R_{\mathcal{K}/\mathbb{Q}}M_\Pi,$$

coincides with complex conjugation. In particular, the polarization induces a nondegenerate hermitian pairing

$$\langle\,\cdot\,,\,\cdot\,\rangle_{i,w} : M^{p_i(w),p_{n+1-i}(cw)}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi) \otimes M^{p_i(cw),n-1-p_i(w)}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi) \to \mathbb{C}$$

for each pair (i, w).

Let $q_i(w) = n-1-p_i(w) = p_{n+1-i}(cw)$. For each pair $(i, w) \in \boldsymbol{n} \times \Sigma_{\mathcal{K}}$, we let $\omega_{i,w}(\Pi) \in M^{p_i(w),q_i(w)}_{\mathcal{K}/\mathbb{Q},\tau}(\Pi)$ be the nonzero image of some F-rational class in the appropriate stage of the Hodge filtration on $M_{\mathcal{K}/F,DR}(\Pi)$; see [Harris 1997, Section 1.4]. Via the comparison isomorphism

$$RM_B \otimes \mathbb{C} \xrightarrow{\sim} RM_{DR} \otimes \mathbb{C}$$

there is an action of F_∞ on RM_{DR}, linear with respect to the coefficients E, that exchanges M_{DR} with M^c_{DR}. Define the de Rham polarization $\langle\,\cdot\,,\,\cdot\,\rangle_{DR}$ by analogy with (1.1.4). It restricts to perfect pairings

$$M^{p_i(w),\,n-1-p_i(w)} \otimes M^{p_{n+1-i}(cw),\,n-1-p_{n+1-i}(cw)} \to E(1-n).$$

Let

$$Q_{i,w}(\Pi) = \langle \omega_{i,w}(\Pi), F_\infty(\omega_{i,w}(\Pi))\rangle_{DR} \in \mathbb{R}^\times. \tag{1.1.11}$$

Here F_∞ is complex conjugation on the Betti realization of $M_{\mathcal{K}/\mathbb{Q},DR}(\Pi)$; see [Harris 1997, (1.0.4)]. Then we may assume

$$F_\infty(\omega_{i,w}(\Pi)) = Q_{i,w}(\Pi)\cdot\omega_{n+1-i,cw}(\Pi). \tag{P}$$

For the rest of Section 1 we will assume $F = \mathbb{Q}$, since the main applications will be in this setting. We can thus choose an embedding $w : \mathcal{K} \hookrightarrow \mathbb{C}$ once and for all and drop the subscripts w in what follows, writing for example ω_i for $\omega_{i,w}$.

1.2. ***The determinant motive.*** The determinant $\det(M)$ is a rank one motive over $\mathcal{K}$ of weight $nw = n(n-1)$ with coefficients in E. Since its λ-adic realization is the Galois character $\xi_{\Pi,\lambda} = \det\rho_{\Pi,\lambda}$ we can write $\det(M) = M(\xi_\Pi)$ where

$$\xi_\Pi = \chi_\Pi \cdot \|\cdot\|^{-n(n-1)/2} \tag{1.2.1}$$

is the indicated shift of the central character χ_Π of Π, calculated using (1.1.3).

The polarization of M defines a polarization

$$M(\xi_\Pi) \otimes M(\xi_\Pi^c) \to E(n(1-n)), \tag{1.2.2}$$

which is obviously consistent with (1.2.1). Taking $\Omega_M = \bigwedge_{i=1}^n \omega_i$ as an E-rational basis of $\det(M)_{DR}$, and defining Ω_M^c analogously, relation (P) yields

$$F_\infty(\Omega_M) = Q_{\det(M)}\Omega_M^c, \quad Q_{\det(M)} = \prod_{i=1}^n Q_i. \tag{1.2.3}$$

On the other hand, letting e_M and e_M^c denote E-rational bases of $\det(M)_B$ and $\det(M^c)_B$ respectively, we can write

$$e_M = \delta(M)\Omega_M, \tag{1.2.4}$$

where following Deligne we let $\delta(M)$ denote the determinant of the comparison isomorphism $I_\infty : M_B \otimes \mathbb{C} \xrightarrow{\sim} M_{DR} \otimes \mathbb{C}$ calculated in E-rational bases; $\delta(M)$ is well-defined as an element of $(E \otimes \mathbb{C})^\times / E^\times$; see [Harris 1997, (1.2.2)].[3] The determinant of the dual map $(I_\infty^\vee)^{-1} : M_B^\vee \otimes \mathbb{C} \xrightarrow{\sim} M_{DR}^\vee \otimes \mathbb{C}$ equals $\delta(M)^{-1}$, up to a multiple in $E^\times$, but by the polarization we find that this is the determinant of

$$I^c(1-n)_\infty : M^c(n-1)_B \otimes \mathbb{C} \xrightarrow{\sim} M^c(n-1)_{DR} \otimes \mathbb{C}.$$

This in turn is $(2\pi i)^{n(n-1)}$ times the determinant of $I_\infty^c : M_B^c \otimes \mathbb{C} \xrightarrow{\sim} M_{DR}^c \otimes \mathbb{C}$; in other words,

$$\delta(M)^{-1} = (2\pi i)^{n(n-1)}\delta(M^c). \tag{1.2.5}$$

Or, with respect to the comparison isomorphism,

$$e_M^c = (2\pi i)^{n(1-n)}\delta(M)^{-1}\Omega_M^c. \tag{1.2.6}$$

Now by (1.2.3) and (1.2.4) we have

$$\Omega_M^c = Q_{\det(M)}^{-1} F_\infty(\Omega_M) = Q_{\det(M)}^{-1}\delta(M)^{-1}e_M^c,$$

which combined with (1.2.6) yields the following:

Lemma 1.2.7. *Under the hypotheses of Section 1.1, we have the relation*

$$\prod_{i=1}^n Q_i = Q_{\det(M)} = (2\pi i)^{n(1-n)}\delta(M)^{-2}$$

as elements of $(E \otimes \mathbb{C})^\times / E^\times$. In other words, there is an element $d(M) \in E^\times$ such that

$$\delta(M)^{-1} = d(M)^{1/2} \cdot (2\pi i)^{n(n-1)/2} \cdot Q_{\det(M)}^{1/2},$$

[3] Deligne's δ is the determinant of the period matrix of a motive over $\mathbb{Q}$; here the motive is over $\mathcal{K}$.

where the choice of square root $d(M)^{1/2}$ depends on the choice of square root of $Q_{\det(M)}$ in $(E \otimes \mathbb{C})^\times / E^\times$.

This is to be compared to [Harris 1997, Lemma 1.4.12]. There the independent definition of $\delta(M)$ determines a square root of $d(M) = [d_{DR}(M)/d_B(M)]$. Presumably $d(M)$ is again a ratio of discriminants of forms attached to the polarization, and its square root can therefore be given an independent definition in an appropriate quadratic extension of E.

1.3. ***Asai motives.*** We postulate that the adjoint motive $\mathrm{Ad}(M) = M \otimes M^\vee$ descends to a motive over F, denoted $\mathrm{As}(M)$ (for Asai). This is true for the ℓ-adic realizations, as explained in [Gan et al. 2012a], and we introduce the corresponding ad hoc descents of the de Rham and Betti realizations in order to define the Deligne periods.

More precisely, in the article [Gan et al. 2012a] of Gan, Gross and Prasad, there are two descents, denoted $\mathrm{As}(M)^+$ and $\mathrm{As}(M)^-$, that differ from one another by twist by the quadratic character $\eta_{\mathcal{K}/F}$, and are distinguished by the signature of F_∞, which is $n(n \pm 1)/2$ on $\mathrm{As}(M)^\pm$. Ours is the one denoted $\mathrm{As}(M)^{(-1)^n}$, as one sees by the definition of the F_∞ action below. Because the signs interfere with the notation for Deligne's periods, we write $\mathrm{As}(M)$ instead of $\mathrm{As}(M)^{(-1)^n}$ and $(\mathrm{As}(M)_B)^\pm$ with parentheses to designate the ± 1-eigenspaces of F_∞.

We denote by $\mathbb{Q}(\eta_{\mathcal{K}/F})$ the Artin motive of rank 1 over F attached to the character $\eta_{\mathcal{K}/F}$. Let e_η denote a basis vector for $\mathbb{Q}(\eta_{\mathcal{K}/F})_B$. The archimedean Frobenius F_∞ acts as -1 on $\mathbb{Q}(\eta_{\mathcal{K}/F})_B$. Let t be a rational basis of $\mathbb{Q}(1)_{DR} = \mathbb{Q}$ (see [Harris 1997, 1.1]), $t_B = 2\pi i t$ a rational basis of $\mathbb{Q}(1)_B = (2\pi i)\mathbb{Q}$; then $F_\infty(t_B) = -t_B$.

We identify $\mathrm{Ad}(M)^c \xrightarrow{\sim} \mathrm{Ad}(M)$ by composing

$$\begin{aligned} \mathrm{Ad}(M)^c = M^c \otimes M^{\vee,c} &\xrightarrow{\sim} M^\vee(1-n) \otimes (M^c(n-1))^c \\ &= M^\vee \otimes M \xrightarrow{\sim} M \otimes M^\vee = \mathrm{Ad}(M), \end{aligned}$$

where the last isomorphism is just exchanging the factors and the first is defined by the polarization. As a model for $\mathrm{As}(M)_B$ over F we take

$$\mathrm{As}(M)_B = M_B \otimes M_B^c(1-n) \otimes \mathbb{Q}(\eta_{\mathcal{K}/F})^{\otimes n}$$

with the action

$$\begin{aligned} F_\infty(e_i \otimes e_j^c \otimes t_B^{1-n} \otimes e_\eta^{\otimes n}) &= e_j \otimes e_i^c \otimes (-1)^{1-n} t_B^{1-n} \otimes (-1)^n e_\eta^{\otimes n} \\ &= -e_j \otimes e_i^c \otimes \otimes t_B^{1-n} e_\eta^{\otimes n}. \end{aligned}$$

Here we have exchanged the first two factors after applying complex conjugation. Thus the vectors

$$\{e_{ij}^+ = [e_i \otimes e_j^c - e_j \otimes e_i^c] \otimes t_B^{1-n} \otimes e_\eta^{\otimes n}, i < j\}$$

and

$$\{e_{ij}^- = [e_i \otimes e_j^c + e_j \otimes e_i^c] \otimes t_B^{1-n} \otimes e_\eta^{\otimes n}, i \le j\}$$

form bases for $(\mathrm{As}(M)_B)^+$ and $(\mathrm{As}(M)_B)^-$, respectively, in Deligne's notation (where we have added parentheses as explained above). In particular,

$$\dim(\mathrm{As}(M)_B)^+ = \tfrac{1}{2}n(n-1) \quad \text{and} \quad \dim(\mathrm{As}(M)_B)^- = \tfrac{1}{2}n(n+1). \tag{1.3.1}$$

But, in the applications we will be interested in the special value $L(1, \mathrm{As}(M)) = L(0, \mathrm{As}(M)(1))$. The action of F_∞ on the Tate twist

$$\mathrm{As}(M)(1)_B = M_B \otimes M_B^c(2-n) \otimes \mathbb{Q}(\eta_{\mathcal{K}/F})^{\otimes n}$$

is as above, with $(1-n)$ replaced by n. The motive $\mathrm{As}(M)(1)$ is pure of weight -2, and the dimension calculation shows that F_∞ acts as the scalar $+1$ on the space of $(-1,-1)$ classes; thus $\mathrm{As}(M)(1)$ is *critical* in Deligne's sense.[4] This implies in particular that the Hodge filtration of $\mathrm{As}(M)(1)_{DR}$ has two distinguished steps $F^\pm \mathrm{As}(M)(1)_{DR}$ (see [Harris 1997, Section 1.2]) uniquely determined by the equalities

$$\dim F^\pm \mathrm{As}(M)(1)_{DR} = \dim(\mathrm{As}(M)(1)_B)^\pm = \tfrac{1}{2}n(n \pm 1),$$

where the dimension calculation follows from (1.3.1), bearing in mind that F_∞ acts as -1 on $\mathbb{Q}(1)_B$. We can similarly define steps in the filtration of $\mathrm{As}(M)_{DR}$:

$$n^\pm := \dim F^\pm \mathrm{As}(M)_{DR} = \dim(\mathrm{As}(M)_B)^\pm = \tfrac{1}{2}n(n \mp 1). \tag{1.3.2}$$

Thus,

$$F^+ \mathrm{As}(M)_{DR} \subsetneq F^- \mathrm{As}(M)_{DR} \quad \text{and} \quad F^- \mathrm{As}(M)(1)_{DR} \subsetneq F^+ \mathrm{As}(M)(1)_{DR}.$$

With respect to the isomorphism $M^\vee \xrightarrow{\sim} M^c(n-1)$, we can take the differentials $\omega_j^c(n-1) = \omega_j^c \otimes t^{\otimes n-1}$ as a basis of $M_{DR}^\vee$. It follows from the dimension calculation above that the relevant step $F^+ \mathrm{As}(M)_{DR}$ in the Hodge filtration is spanned by the classes $\omega_{ij} = \omega_i \otimes \omega_j^c(n-1)$, of Hodge type

$$H_{ij}(\mathrm{As}(M)) := (p_i + p_j^c + 1 - n, n - 1 - p_i - p_j^c)$$

satisfying the condition

$$p_i + p_j^c > n - 1. \tag{$C(+)$}$$

This is equivalent to $p_i - p_{n+1-j} > 0$, and since the p_i are strictly decreasing, $(C(+))$ is true if and only if $i + j \le n + 1$. Similarly $F^- \mathrm{As}(M)_{DR}$ is spanned by

[4] Dick Gross has pointed out that this can be seen purely in terms of representation theory. The local L-factor at infinity $L_\infty(s, \mathrm{As}(M))$ has no pole at $s = 1$ because discrete series parameters are generic, and no pole at $s = 0$ because the corresponding representations are in the discrete series.

ω_{ij} satisfying

$$p_i + p_j^c \geq n - 1 \quad (n \text{ even}), \tag{$C(-)$}$$

which holds if and only if $i + j \leq n + 1$.

We define the motives $\bigwedge^2 M$ and $\mathrm{Sym}^2 M$ over $\mathcal{K}$ in the obvious way. Because we will need a uniform notation we write $S^+(M) = \mathrm{Sym}^2 M$ and $S^-(M) = \bigwedge^2 M$.

Write

$$\omega_j = \sum a_{ij} e_i \quad \text{and} \quad \omega_j^c = \sum a_{ij}^c e_i^c.$$

Then we have the relation $a_{i,n+1-j}^c = Q_j^{-1} a_{ij}$. Now let $\{e_{ik}^{\pm,*}\}$ denote the dual basis to the basis $\{e_{ik}^{\pm}\}$ of $(\mathrm{As}(M)_B)^{\pm}$ introduced above. It follows from the identity (P) that we have

$$\begin{aligned} e_{ik}^{\pm,*}(\omega_{j,n+1-\ell}) &\sim [a_{ij} a_{k,n+1-\ell}^c \pm a_{kj} a_{i,n+1-\ell}^c](2\pi i)^{1-n} \\ &\sim (2\pi i)^{1-n} Q_\ell^{-1}(a_{ij} a_{k,\ell} \pm a_{kj} a_{i,\ell}) \\ &\sim (2\pi i)^{1-n} Q_\ell^{-1} e_{ik}^{\pm,*}(\omega_j \otimes \omega_\ell), \end{aligned}$$

where $\sim$ means that the calculations are up to factors in the coefficient field. Now if $H_{j,n+1-\ell}(\mathrm{As}(M))$ satisfies $(C(+))$, then $j < \ell$. The arguments of [Harris 1997, Section 1.5] allow us to calculate the matrix for the Deligne period $c^+(\mathrm{As}(M)^\vee)$ of the *dual* of $\mathrm{As}(M)$. However, the self-duality of $\mathrm{Ad}(M)$ easily implies that $\mathrm{As}(M)$ is self-dual, so the calculation that follows gives an expression for $c^+(\mathrm{As}(M))$. The entries in the matrix are given by $e_{ik}^{+,*}(\omega_{j\ell})$ as (i,k) varies over pairs with $i \leq k$ and $j \leq \ell$ if n is odd, with strict inequalities if n is even.

Keep $n^{\pm}$ as in (1.3.2). Then the determinant of the period matrix calculating $c^{\pm}(\mathrm{As}(M))$ is equal to a certain product $Q^{\pm}(\mathrm{As}(M))$ of factors of the form Q_ℓ^{-1}, to be determined below, multiplied by the determinant Δ of the matrix

$$(e_i \otimes e_k - e_k \otimes e_i)^*(\omega_j \otimes \omega_\ell)$$

as (i,k) ranges over pairs with $i \leq k$ and (j,ℓ) ranges over pairs with $j < \ell$, the whole multiplied by $(2\pi i)^{(1-n)n^{\pm}}$. The determinant Δ is precisely the inverse of the determinant of the full period matrix of the motive $S^{\mp}(M)$ in the implicit bases, which Deligne denotes $\delta(S^{\mp}(M))$.

The factor $Q^{\pm}(\mathrm{As}(M))$ is determined as follows. For $1 \leq \ell \leq n$, let $m^+(\ell)$ and $m^-(\ell)$ denote the number of j such that $j \leq \ell$ and $j < \ell$, respectively. Then $m^+(\ell) = \ell$ and $m^-(\ell) = \ell - 1$. Let

$$Q^+(M) = \prod_\ell Q_\ell^{-m^+(\ell)} = \prod_\ell Q_\ell^{-\ell} \quad \text{and} \quad Q^-(M) = \prod_\ell Q_\ell^{-m^-(\ell)} = \prod_\ell Q_\ell^{1-\ell}.$$

It follows that:

Formula 1.3.3. $$Q^{\pm}(\mathrm{As}(M)) = Q^{\mp}(M).$$

This proves the first statement of the following proposition; the second statement is proved analogously.

Proposition 1.3.4. *Let M be a polarized motive satisfying the conditions of 1.1.10, and with the property that* $\mathrm{Ad}(M)$ *descends to* $F = \mathbb{Q}$. *Then*

$$c^+(\mathrm{As}(M)) = (2\pi i)^{(1-n)n^+} Q^-(M)\delta(S^-(M))^{-1},$$
$$c^-(\mathrm{As}(M)) = (2\pi i)^{(1-n)n^-} Q^+(M)\delta(S^+(M))^{-1}.$$

Applying [Deligne 1979a, formula (5.1.8)], with n^- as in (1.3.2), we have

$$c^+(\mathrm{As}(M)(1)) = c^-(\mathrm{As}(M))(2\pi i)^{n^-}.$$

One calculates easily that $\delta(S^\pm(M)) = \delta(\det(M)^{n\pm 1}) = \delta(M)^{n\pm 1}$, where the last equality follows from the considerations of Section 1.2.

Combining the formulas of this section with Lemma 1.2.7, we can therefore write the Deligne period for the motive of interest explicitly in terms of the Q_j and δ.

Corollary 1.3.5. *Under the above hypotheses, we have the following expression for* $c^+(\mathrm{As}(M)(1))$:

$$\begin{aligned} c^+(\mathrm{As}(M)(1)) &= (2\pi i)^{n^-}(2\pi i)^{(1-n)n^-} Q^+(M)\delta(S^+(M))^{-1} \\ &= d(M)^{1/2}(2\pi i)^{n(n+1)/2}[Q_{\det(M)}]^{(n-1)/2} \cdot \textstyle\prod_\ell Q_\ell^{1-\ell} \\ &= d(M)^{1/2}(2\pi i)^{n(n+1)/2} \textstyle\prod_\ell Q_\ell^{(n+1)/2-\ell} \end{aligned}$$

We see that $\delta(S^-(M))^{-1}$ is an odd power of $\delta(M)^{-1}$; therefore we need to include the factor $d(M)^{1/2}$ introduced in Lemma 1.2.7 along with the half-integral power of $Q_{\det(M)}$. The half-integral powers of the Q_ℓ that occur in the expression for even n are not meaningful individually, and have only been included for their suggestive similarity with the standard expression for the half-sum of positive roots.

Remark 1.3.6. If one defines Q_ℓ^c by analogy with the definition of Q_ℓ above, one sees easily that $Q_\ell^c = Q_{n+1-\ell}^{-1}$. It is obvious that the expression in Corollary 1.3.5 is invariant when M and M^c are exchanged, as it should be.

1.4. ***Tensor products.*** In subsequent sections we will explore the relations between the calculations of the previous section and the Ichino–Ikeda conjecture. Here we briefly explain how a similar calculation determines the Deligne period of the tensor product of two motives of the type considered in Section 1.

Suppose M and M' are two motives of dimension n and n', respectively, both of the type considered above. We let ω_a, ω_t^c, e_i, e_i^c, where $1 \le a, t, i \le n$, be the basis vectors defined for M above. For M' we use the notation η_b, η_u^c, f_j, f_j^c, with $1 \le b, u, j \le n'$. The Hodge types for M are $(p_i, n-1-p_i)$ and $(p_i^c, n-1-p_i^c)$

as before; for M' we write $(r_j, n'-1-r_j)$ and $(r_j^c, n'-1-r_j^c)$. The tensor product motive we consider is not $RM \otimes RM'$ but rather $R(M \otimes M') = R_{\mathcal{K}/F}(M \otimes M')$, whose Betti realization is $M_B \otimes M'_B \oplus M^c_B \otimes (M')^c_B$, and whose de Rham realization breaks up analogously. In particular, the differentials $\omega_a \otimes \eta_b$ and $\omega_t^c \otimes \eta_u^c$ form a basis for $R(M \otimes M')_{DR}$.

The motive $R(M \otimes M')$ is of dimension $2nn'$ over its coefficient field and of weight $w = n + n' - 2$. We will only need to consider the case when n and n' are of opposite parity; for example, when $n' = n - 1$, as in the original Gross–Prasad conjecture. Then w is odd and $R(M \otimes M')$ has no $(0, 0)$ classes; it follows that the value $(w + 1)/2 = (n + n' - 1)/2$ is a critical value of the L-function $L(s, R(M \otimes M'))$.

The basis for $R(M \otimes M')^{\pm}_B$ is then $e_i \otimes f_j \pm e_i^c \otimes f_j^c$ for $1 \leq i \leq n$ and $i \leq j \leq n'$. To determine the basis for $F^+R(M \otimes M')_{DR} = F^-R(M \otimes M')_{DR}$ we need to determine the sets $A(M, M')$ and $T(M, M')$ of pairs a, b and t, u such that $p_a + r_b \geq (w + 1)/2$ and $p_t^c + r_u^c \geq (w + 1)/2$, respectively. Bearing in mind Hodge duality, the cardinality

$$|A(M, M')| + |T(M, M')| = nn' = \dim F^+R(M \otimes M')_{DR}.$$

The set $\{\omega_a \otimes \eta_b \mid (a, b) \in A(M, M')\} \cup \{\omega_t^c \otimes \eta_u^c \mid (t, u) \in T(M, M')\}$ forms a basis for $F^+R(M \otimes M')_{DR}$. A calculation using the relation (P), as in Section 1.3, shows that:

Lemma 1.4.1.

$$\begin{aligned} c^+(R(M \otimes M')^\vee) &= \pm c^-(R(M \otimes M')^\vee) \\ &= \prod_{(t,u) \in T(M,M')} Q_{n+1-t}(M)^{-1} Q_{n'+1-u}(M')^{-1} \cdot \delta(M \otimes M')^{-1}, \end{aligned}$$

where δ is the determinant of the full period matrix for $M \otimes M'$, viewed as a motive over $\mathcal{K}$.

More precisely, letting (i, j) run over pairs of integers with $1 \leq i \leq n$ and $i \leq j \leq n'$, the Deligne period $c^+(R(M \otimes M'))$ is the determinant of the matrix whose first $|A(M, M')|$ columns, indexed by pairs $(a, b) \in A(M, M')$, are the vectors $(a_{ia} b_{jb})$, and whose last $|T(M, M')|$ columns, indexed by pairs $(t, u) \in T(M, M')$, are the vectors $(a_{it}^c b_{ju}^c)$. Here as above, we have written

$$\omega_a = \sum a_{ia} e_i, \quad \eta_b = \sum b_{jb} f_j, \quad \omega_t^c = \sum a_{it}^c e_i^c, \quad \eta_u^c = \sum b_{ju}^c f_j^c.$$

By identity (P) we have

$$\omega_t^c = Q_{n+1-t}(M)^{-1} \sum a_{it} e_i^c \quad \text{and} \quad \eta_u^c = Q_{n'+1-u}(M')^{-1} \sum b_{ju} f_j^c.$$

The formula for $c^+(R(M \otimes M')^\vee)$ then follows as in Section 1.3.

Because the Hodge types satisfy $p_t^c > p_{t+1}^c$ and $r_u^c > r_{u+1}^c$, we have this:

Lemma 1.4.2. *The set $T(M, M')$ is a tableau: if $(t, u) \in T(M, M')$, then for any $t' < t$ and $u' < u$, the pairs (t', u) and (t, u') are also in $T(M, M')$.*

We can represent $T(M, M')$ geometrically as a tableau in the rectangular grid of height n and width n', whose boxes are indexed by pairs with $1 \le t \le n$ and $1 \le u \le n'$. The box at position (t, u) is filled in if $(t, u) \in T(M, M')$. Then the lemma asserts that if a given box (t, u) is filled in, all boxes above it or to the left of it are also filled in.

In the notation of the introduction, the set $T(M, M')$ determines the pair of hermitian spaces $W' \subset W$ whose automorphic periods are expressed by the Ichino–Ikeda conjecture as the quotient of the central critical value of $L(s, R(M \otimes M'))$ by a product of critical values at $s = 1$ of Asai L-functions. The automorphic periods can be normalized as in [Harris 2012], where they are called *Gross–Prasad periods*. The relation between Gross–Prasad periods and motivic periods is in general not transparent, and it is therefore not clear how to establish compatibility between the Ichino–Ikeda and Deligne conjectures in general. We will return to this topic in a subsequent article. The remainder of the present article is devoted to studying a special case where compatibility of the two conjectures can be studied.

2. The Ichino–Ikeda conjecture for unitary groups

In the present section, W denotes an n-dimensional hermitian space over $\mathcal{K}$, relative to conjugation over F; until the end of Section 2.4, we allow F to be an arbitrary totally real field. If W_1 and W_2 are two such spaces, then for almost all finite primes v of F we have

$$U(W_1 \otimes F_v) \xrightarrow{\sim} U(W_2 \otimes F_v) \tag{2.0.1}$$

This allows us to consider automorphic representations of all unitary groups $\mathrm{U}(W)$ simultaneously, and to organize them into *near equivalence classes*: the automorphic representations π_1 of $\mathrm{U}(W_1)$ and π_2 of $\mathrm{U}(W_2)$ are nearly equivalent if, for all but finitely many v for which (2.0.1) holds, the local components $\pi_{1,v}$ and $\pi_{2,v}$ are equivalent.

The Gross–Prasad and Ichino–Ikeda conjectures concern special values of L-functions and local ε-factors for near equivalence classes of local and automorphic representations respectively. A given near equivalence class gives rise to a family of motives (or at least realizations) in the cohomology of the corresponding Shimura varieties; the details are recalled in Section 2.4.

All the automorphic representations in a near equivalence class are supposed to have a common base change, say Π, an automorphic representation of $\mathrm{GL}(n)_{\mathcal{K}}$ that satisfies the polarization condition (1.1.1). This has been proved in a great many cases (see [Labesse 2011; White 2010], for example) and will be taken as an axiom in what follows. The near equivalence class will sometimes be denoted $\Phi(\Pi)$ — convention actually dictates it should be $\Pi(\Phi)$, or even $\Pi(\Phi(\Pi))$, where Φ is supposed to suggest the Langlands parameter of Π, but since the letter Π is otherwise engaged this looks problematic.

2.1. ***Statement of the conjecture.*** Let $W' \subset W$ a codimension one subspace on which the restriction of the hermitian form is nondegenerate, so that $W = W' \oplus W_0$ with $W_0 = W^{\perp}$. The unitary groups of W, W' and W_0 are reductive algebraic groups over F; we write $G' = U(W')$, $G_0 = U(W_0)$ and $G = U(W)$.

Let π, π' and π_0 be tempered cuspidal automorphic representations of G, G' and G_0, respectively. Let

$$\chi_\pi : Z_G(\mathbf{A})/Z_G(F) \to \mathbb{C}^\times \quad \text{and} \quad \chi'_\pi : Z_{G'}(\mathbf{A})/Z_{G'}(F) \to \mathbb{C}^\times$$

denote their central characters — π_0 is itself a character — and assume that

$$\chi_\pi \cdot \chi'_\pi \otimes \pi_0 \big|_{Z_G(\mathbf{A})} = 1. \tag{2.1.1}$$

Fix factorizations

$$\pi \xrightarrow{\sim} \otimes'_v \pi_v, \quad \pi' \xrightarrow{\sim} \otimes'_v \pi'_v, \quad \pi^\vee \xrightarrow{\sim} \otimes'_v \pi^\vee_v, \quad \pi'^{,\vee} \xrightarrow{\sim} \otimes'_v \pi'^{,\vee}_v \tag{2.1.2}$$

and likewise for the contragredients $\pi^\vee$ and $\pi'^{,\vee}$. We assume the factorizations (2.1.2) are compatible with factorizations of pairings

$$\langle \cdot, \cdot \rangle_\pi = \prod_v \langle \cdot, \cdot \rangle_{\pi_v} \quad \text{and} \quad \langle \cdot, \cdot \rangle_{\pi'} = \prod_v \langle \cdot, \cdot \rangle_{\pi'_v},$$

where in each case the left hand side is the L_2 pairing on cusp forms and the right hand side is the product of canonical pairings between a representation and its contragredient. We define

$$\begin{aligned} P(f, f') &= \int_{G'(F)\backslash G'(\mathbf{A})} f(g') f'(g')\, dg', \\ P(f^\vee, f'^{,\vee}) &= \int_{G'(F)\backslash G'(\mathbf{A})} f^\vee(g') f'^{,\vee}(g')\, dg', \end{aligned} \tag{2.1.3}$$

$$\begin{aligned} Q(f, f^\vee) &= \int_{G(F)\backslash G(\mathbf{A})} f(g) f^\vee(g)\, dg, \\ Q(f', f'^{,\vee}) &= \int_{G'(F)\backslash G'(\mathbf{A})} f'(g') f'^{,\vee}(g')\, dg'. \end{aligned} \tag{2.1.4}$$

For any place v of F, write $G_v = G(F_v)$ and $G'_v = G'(F_v)$. Let dg and dg' denote Tamagawa measures on $G(\mathbf{A})$ and $G'(\mathbf{A})$, respectively. We choose factorizations $dg = \prod_v dg_v$, $dg' = \prod_v dg'_v$ over the places of v with these properties:

- For every finite v, the measures dg_v and dg'_v take rational values on open subsets of G_v and G'_v, respectively.
- For all v outside a finite set S, including all archimedean places and all places at which either π or π' is ramified,
$$\int_{K_v} dg_v = \int_{K'_v} dg'_v = 1,$$
where K_v and K'_v are hyperspecial maximal compact subgroups of G_v and G'_v respectively.

Assume $f \in \pi$, $f' \in \pi'$, $f^\vee \in \pi^\vee$, $f'^{,\vee} \in \pi'^{,\vee}$ are factorizable vectors, that is,
$$f = \bigotimes_v f_v, \qquad \text{where } f_v \in \pi_v, \qquad f' = \bigotimes_v f'_v, \qquad \text{etc.}$$
with respect to the isomorphisms (2.1.2). In what follows, we have:

(a) $|S(\pi, \pi')|$ is an integer measuring the size of the global L-packets of π and π'.

(b) Δ_G is the value at $s = 0$ of the L-function of the *Gross motive* of the group G; explicitly,
$$\Delta_G = \prod_{i=1}^{n} L(i, \eta^i_{\mathcal{K}/F}).$$

(c) For each finite v,
$$\begin{aligned} Z_v &= Z_v(f_v, f_v^\vee, f'_v, f_v'^{,\vee}) \\ &= \int_{G'_v} c_{f_v, f_v^\vee}(g'_v) c_{f'_v, f_v'^{,\vee}}(g'_v)\, dg'_v \cdot \frac{L(1, \pi_v, \mathrm{Ad}) L(1, \pi'_v, \mathrm{Ad})}{L(\frac{1}{2}, \mathrm{BC}(\pi_v) \times \mathrm{BC}(\pi'_v))}. \end{aligned}$$

(d) For each archimedean v,
$$Z_v = Z_v(f_v, f_v^\vee, f'_v, f_v'^{,\vee}) = \int_{G'_v} c_{f_v, f_v^\vee}(g'_v) c_{f'_v, f_v'^{,\vee}}(g'_v)\, dg'_v.$$

(e) In (c) and (d), the notation $c_{f_v, f_v^\vee}(g'_v)$ designates the local matrix coefficient $c_{f_v, f_v^\vee}(g_v) = (\pi(g_v) f_v, f_v^\vee)$ with respect to the canonical local pairing of representations (likewise for $c_{f'_v, f_v'^{,\vee}}$).

The Ichino–Ikeda conjecture is the assertion that
$$\frac{P(f, f') P(f^\vee, f'^{,\vee})}{Q(f, f^\vee) Q(f', f'^{,\vee})} = 2^{-|S(\pi, \pi')|} \Delta_G \prod_v Z_v \cdot \frac{L(\frac{1}{2}, \mathrm{BC}(\pi) \times \mathrm{BC}(\pi'))}{L(1, \pi, \mathrm{Ad}) L(1, \pi', \mathrm{Ad})} \quad (2.1.5)$$

Here the L-functions are defined in [N. Harris 2011] by Euler products over finite primes only. One of the main results of [Ichino and Ikeda 2010; N. Harris 2011] is that $Z_v = 1$ for all v outside a finite set S, including all archimedean places;

thus convergence of the product $\prod_v Z_v$ is not an issue. We can rewrite the right hand side

$$2^{|S(\pi,\pi')|}\Delta_G Z_{\mathrm{loc}} \cdot \frac{L(\frac{1}{2}, \mathrm{BC}(\pi) \times \mathrm{BC}(\pi'))}{L(1, \pi, \mathrm{Ad})L(1, \pi', \mathrm{Ad})}$$

with $Z_{\mathrm{loc}} = \prod_{v \in S} Z_v$.

2.2. ***Local vanishing and the Gross–Prasad conjecture.*** The map $P : \pi \otimes \pi' \to \mathbb{C}$ of (2.1.3) is invariant under $G'(\mathbf{A})$. Its nontriviality therefore implies that, for every v, there is a bilinear map

$$P_v : \pi_v \otimes \pi'_v \to \mathbb{C}, \tag{2.2.1}$$

invariant under the diagonal action of G'_v. (The integral Z_v defines a multilinear form on $(\pi_v \otimes \pi'_v) \otimes (\pi_v^\vee \otimes \pi_v'^{,\vee})$.)

The existence of G'_v-invariant maps like (2.2.1) is the subject of the Gross–Prasad conjecture [Gan et al. 2012a]. For the purposes of the present exposition, it will suffice to assume $\pi_v \otimes \pi'_v$ to be tempered. Assume that L-packets can be attached consistently to tempered Langlands parameters for the group $G_v \times G'_v$ and all its inner twists; see [Mœglin 2007]. Let $L(\pi_v, \pi'_v)$ denote the space of G'_v-invariant maps (2.2.1).

Conjecture 2.2.2 (local Gross–Prasad conjecture). *Let WD_{F_v} denote the Weil–Deligne group of F_v, and let*

$$\Phi_v \times \Phi'_v : WD_{F_v} \to {}^L(G_v \times G'_v)$$

denote a tempered Langlands parameter for the group $G_v \times G'_v$ and all its inner twists. Then

$$\sum_{W_v = W'_v \oplus W_{0,v}} \; \sum_{\substack{\pi_v \otimes \pi'_v \in \\ \Pi(\Phi_v \times \Phi'_v; U(W_v) \times U(W'_v))}} \dim L(\pi_v, \pi'_v) = 1.$$

Here the outer sum runs over isometry classes of pairs of hermitian spaces over F_v, as in Section 2.1, and the inner sum runs over the L-packet of the given inner form of $G_v \times G'_v$ attached to $\Phi_v \times \Phi'_v$.

The full Gross–Prasad conjecture treats more general inclusions of groups and gives a formula in terms of the Langlands parameter determining the unique pair $\pi_v \otimes \pi'_v$ in the L-packet for which $L(\pi_v, \pi'_v) \neq 0$. This has been proved for special orthogonal groups by Waldspurger in the tempered case and by Moeglin and Waldspurger in general; see [Moeglin and Waldspurger 2012]. Conjecture 2.2.2 for unitary groups is the subject of work in progress by R. Beuzart-Plessis.[5]

[5] Assuming standard conjectures on L-packets of unitary groups, Beuzart-Plessis has now proved Conjecture 2.2.2 together with its refinement.

Now let (π, π') be a pair of tempered cuspidal automorphic representations of G and G', as in Section 2.1. For each place v, Conjecture 2.2.2 asserts the existence of unique (strong) inner forms $G_{1,v}$ and $G'_{1,v}$ of G_v and G'_v, respectively, and unique representations $\pi_{1,v}$ and $\pi'_{1,v}$ of $G_{1,v}$ and $G'_{1,v}$ in the L-packets given by the local Langlands parameters of π_v and π'_v, such that $L(\pi_{1,v}, \pi'_{1,v}) \neq 0$. The following is a restatement of [Gan et al. 2012a, Conjecture 26.1] in the present situation.

Conjecture 2.2.3 (global Gross–Prasad conjecture). *With π and π' as above, the following are equivalent*:

(1) *There are unitary groups $G_1 \supset G'_1$ over F with local forms the given $G_{1,v}$ and $G'_{1,v}$, automorphic representations π_1 and π'_1 with the given local components, and forms $f_1 \in \pi_1$ and $f'_1 \in \pi'_1$, such that the period integral $P(f_1, f'_1)$ is not zero.*

(2) *The central value $L(\frac{1}{2}, \mathrm{BC}(\pi_1) \otimes \mathrm{BC}(\pi'_1)) = L(\frac{1}{2}, \mathrm{BC}(\pi) \otimes \mathrm{BC}(\pi')) \neq 0$.*

The Ichino–Ikeda conjecture (2.1.5) is a refinement of Conjecture 2.2.3.[6] As a part of their refinement of the global Gross–Prasad conjecture for special orthogonal groups, Ichino and Ikeda have proposed a refinement of the local conjecture as well. I state it here in the unitary case. (It seems not to have been stated in [N. Harris 2011], though it is certainly compatible with the global conjecture stated there.)

Conjecture 2.2.4 (of Ichino–Ikeda [2010, Conjecture 1.3]). *Under the hypotheses of Conjecture 2.2.2 — in particular, assuming π_v and π'_v belong to tempered L-packets — we have $L(\pi_v, \pi'_v) \neq 0$ if and only if the local integral Z_v defines a nonzero multilinear form on $(\pi_v \otimes \pi'_v) \otimes (\pi_v^\vee \otimes \pi_v'^{,\vee})$. In other words, the local zeta integral defines a basis vector in the one-dimensional vector space $L(\pi_v, \pi'_v) \otimes L(\pi_v^\vee, \pi_v'^{,\vee})$.*

If one admits these conjectures, the nonvanishing of the numerator of the quotient of L-functions on the right hand side of (2.1.5), together with the local nonvanishing Conjecture 2.2.3, picks out a unique global pair of hermitian spaces $W \supset W'$ and a unique pair of automorphic representations π, π' of the chosen inner forms $\mathrm{U}(W)$ and $\mathrm{U}(W')$, for which the left hand side and the product Z_v do not vanish. The arithmetic meaning of the local conditions at finite primes is not yet understood, but the local conditions at archimedean primes can be translated into simple conditions on the relative positions of the Hodge structures attached to the motives $M(\pi)$ and $M(\pi')$. The next two sections explain these conditions when W and W' are totally definite, and interprets the expressions on the left hand side of (2.1.5).

[6] Added in proof: Wei Zhang has now proved this under some local restrictions.

2.3. *Hodge structures in the definite case.* When v is a real place of F and π_v and π'_v are discrete series representations of G_v and G'_v, the dimension of $L(\pi_v, \pi'_v)$ is determined in [Gan et al. 2012b, Section 2] in terms of the local Langlands parameters. The relation with Hodge types is reduced there to a calculation of signs, which in general is rather elaborate.

The definite case is simpler. Let H denote the compact Lie group $\mathrm{U}(n)$, the symmetry group of the hermitian form $\sum_{i=1}^{n} z_i \bar{z}_i$. Let $H' = U(n-1) \times U(1)$, diagonally embedded in H, and fix an irreducible representation τ of H, with highest weight $a_1 \geq a_2 \geq \cdots \geq a_n$, where $a_i \in \mathbb{Z}$, in the standard normalization. The classic branching formula [Fulton and Harris 1991] determines the highest weights of the representations τ' that occur in the restriction of τ to H'.

Formula 2.3.1 (branching formula). Let τ' be the irreducible representation of H' with highest weight $(b_1, \ldots, b_{n-1}; b_n) \in \mathbb{Z}^n$, where $b_1 \geq \cdots \geq b_{n-1}$ is a highest weight for $\mathrm{U}(n-1)$ and b_n is the weight of a character of $\mathrm{U}(1)$. Then $L(\tau, \tau') \neq 0$ if and only if

- $\sum_{i=1}^{n} a_i = -\sum_{i=1}^{n} b_i$,
- $a_1 \geq -b_{n-1} \geq a_2 \geq -b_{n-2} \geq \cdots \geq a_{n-1} \geq -b_1 \geq a_n$.

Assume W is a totally definite hermitian space over $\mathcal{K}$, and let π and π' be automorphic representations of G and G', whose base changes to $\mathrm{GL}(n, \mathcal{K})$ and $\mathrm{GL}(n-1, \mathcal{K})$ are denoted Π and Π'. Choose a pair (w, cw) of conjugate complex embeddings of $\mathcal{K}$ over the real embedding w^+ of F, with $w \in \Sigma$, and extend w to a map $\sigma : E(\Pi) \otimes \mathcal{K} \to \mathbb{C}$ as in Section 1.1. Suppose $\pi_{w^+} = \tau$, $\pi'_{w^+} = \tau'$, with parameters as in Formula 2.3.1. The condition 1.1.10(c) determines the Hodge numbers of $R_{\mathcal{K}/\mathbb{Q}} M_{\Pi}$. Bearing in mind that Π is an automorphic representation whose local component Π_w has cohomology with coefficients in the *dual* representation $\tau^\vee$ of $\mathrm{GL}(n, \mathbb{C})$, we have

$$\dim M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi) = 1 \quad \text{if and only if, for some } i, \\ (p,q) = (p_i(w), q_i(w)) = (n - i - a_{n+1-i}, i - 1 + a_{n+1-i}). \tag{2.3.2}$$

Similarly,

$$\dim M^{p,q}_{\mathcal{K}/\mathbb{Q},\sigma}(\Pi') = 1 \quad \text{if and only if, for some } i, \\ (p,q) = (p'_i(w), q'_i(w)) = (n - 1 - i - b_{n-i}, i - 1 + b_{n-i}). \tag{2.3.3}$$

Comparing this to Formula 2.3.1(2), we find that

$$p_1(w) > p'_1(cw) \geq p_2(w) > p'_2(cw) \geq \cdots \\ > p_{n-1}(w) > p'_{n-1}(cw) \geq p_n(w) \tag{2.3.4}$$

2.4. ***Realizations of motives in unitary group Shimura varieties.*** The hermitian spaces W and W' are assumed definite at infinity, as in the previous section. Let Π be a cuspidal cohomological automorphic representation of $\mathrm{GL}(n)_{\mathcal{K}}$ satisfying (1.1.1). We consider the near equivalence class $\Phi(\Pi)$ of automorphic representations of varying $\mathrm{U}(W)$. The hermitian pairing $\langle\cdot,\cdot\rangle_W$ on W defines an involution $\tilde{c}$ on the algebra $\mathrm{End}_F(W)$ via $\langle a(v), v'\rangle_W = \langle v, a^{\tilde{c}}(v')\rangle_W$. For each such W, there is a Shimura variety $Sh(W)$ attached to the rational similitude group $\mathrm{GU}(W)$, defined as the functor on the category of $\mathbb{Q}$-algebras R by

$$\mathrm{GU}(W)(R) = \{g \in \mathrm{GL}(V \otimes_{\mathbb{Q}} R) \mid g \cdot \tilde{c}(g) = \nu(g) \text{ for some } \nu(g) \in R^{\times}\}.$$

For each automorphic representation $\pi \in \Phi(\Pi)$ of $\mathrm{U}(W)$, we choose an extension π^+ to an automorphic representation of $\mathrm{GU}(W)$; we can arrange that the central character χ_{π^+} of π^+ is independent of $\pi \in \Phi(\Pi)$. We summarize the discussions in [Harris 1997, Section 2] (for $F = \mathbb{Q}$) and [Harris et al. 2011, §3.2], and provide a few additional details.

For each W, we fix an irreducible admissible representation $\pi_f = \pi_{f,W}$ of $\mathrm{U}(W)(\mathbf{A}^f)$ such that $\pi_\infty \otimes \pi_f \in \Phi(\Pi)$ for some discrete series representation π_∞ of $\mathrm{U}(W_{\mathbb{R}}) := U(W \otimes_{\mathbb{Q}} \mathbb{R})$. For each place w of $\mathcal{K}$, let (r_w, s_w) denote the signature of the hermitian space W_w, and let $d_W = \sum_{v:F\hookrightarrow\mathbb{R}} r_w \cdot s_w$, where w is one of the two extensions of v to $\mathcal{K}$ and $r_w \cdot s_w$ does not depend on the choice. Define the Shimura variety $Sh(W)$ and the local system $\tilde{W}^+(\Pi)$ over $Sh(W)$ as in [Harris et al. 2011, Section 3.2]; here $\tilde{W}^+(\Pi)$ is attached to a finite-dimensional algebraic representation $W^+(\Pi)$ of $\mathrm{GU}(W)$. Then the motivic realization of Π on $Sh(W)$ is the motive

$$\begin{aligned} M(\pi_f^+) &= \mathrm{Hom}_{\mathrm{GU}(\mathbf{A}^f)}(\pi_f^+, H^{d_W}(Sh(W), \tilde{W}^+(\Pi))) \\ &= \mathrm{Hom}_{\mathrm{GU}(\mathbf{A}^f)}(\pi_f^+, H^{d_W}(Sh(W)^*, j_{!*}\tilde{W}^+(\Pi))), \end{aligned} \tag{2.4.1}$$

where $j : Sh(W) \hookrightarrow Sh(W)^*$ is the embedding of $Sh(W)$ in its Baily–Borel compactification.

Let M_Π be the rank n motive over $\mathcal{K}$ introduced in Section 1.1 and $M_{\mathcal{K}/\mathbb{Q}}(\Pi)$ for its restriction of scalars to $\mathbb{Q}$. As in [Harris et al. 2011, (3.2.4)], we have

$$M(\pi_f^+) \xrightarrow{\sim} \bigotimes_{w\in\Sigma} \textstyle\bigwedge^{s_w}(St) M_{F/\mathbb{Q}}(\Pi) \otimes (M(\chi_{\pi^+,W})(t_W)), \tag{2.4.2}$$

where $t_W = \frac{1}{2}\sum_{w\in\Sigma} s_w(s_w - 1)$.

All the motives $M(\pi_f^+)$ are assumed to have coefficients in a common field $E(\pi_f)$. Let $E(W)$ be the reflex field of $Sh(W)$; it is contained in the Galois closure of $\mathcal{K}$ over $\mathbb{Q}$, and of course it depends on the signatures of W at places of Σ. The de Rham realization $M_{\mathcal{K}/\mathbb{Q},DR}(\pi_f^+)$ is free over $E(W) \otimes E(\pi_f^+)$ of rank $\prod_w \binom{n}{s_w}$; the lowest nontrivial stage of its Hodge filtration $F^{\max}_{\mathcal{K}/\mathbb{Q},DR}(\pi_f^+)$ is a free rank one $E(W)\otimes E(\pi_f^+)$-submodule. Let $\Omega_W(\Pi)$ be any $E(W)\otimes E(\pi_f^+)$-basis of $F^{\max}_{\mathcal{K}/\mathbb{Q},DR}(\pi_f^+)$. By analogy with (1.1.11), we define

$$Q_W(\Pi)=\langle \omega_W(\Pi), F_\infty(\omega_W(\Pi))\rangle_{DR}\in (E(W)\otimes E(\pi_f^+)\otimes\mathbb{R})^\times. \tag{2.4.3}$$

We now simplify formulas by assuming $F=\mathbb{Q}$. The index W is in fact superfluous in the character $\chi_{\pi^+,W}$, given the presence of the twist t_W, but we will leave it in place. In [Harris 1997] there is a parameter denoted c in the highest weight of the representation $W^+(\Pi)$, corresponding to the restriction of the central character to the diagonal subgroup $\mathbb{G}_{m,\mathbb{Q}}\subset \mathrm{GU}(W)$. Dually, the central character χ_{π^+} of π^+ has the property that

$$\chi_{\pi^+}(t)=t^{-c}\quad \text{for } t\in\mathbb{R}^\times\subset Z_{\mathrm{GU}(W)}(\mathbb{R}). \tag{2.4.4}$$

Let $W(\Pi)$ denote the restriction of $W^+(\Pi)$ to $\mathrm{U}(W)$, and identify $W(\Pi)$ with the representation $\tau^\vee$ of Section 2.3, with parameters as in 2.3.1. Then $c\equiv\sum_i a_i \pmod 2$. To simplify the formulas, we assume $\sum_i a_i$ to be *even* and take $c=0$. Then $M(\chi_{\pi^+,W})$ is a motive of weight 0.

2.5. ***Automorphic forms on definite unitary groups.*** Let $G=\mathrm{U}(W)$, $G'=\mathrm{U}(W')$, as in Section 2.1, and assume W and W' are totally definite. We can define Shimura data $(G,x)\supset(G',x')$, where $x=x'$ is the point consisting of the trivial homomorphism from $R_{\mathbb{C}/\mathbb{R}}\mathbb{G}_{m,\mathbb{C}}$ to the group G'. This satisfies all the axioms of [Deligne 1979b, (2.1.1)] with the exception of (2.1.1.3), which is in fact unnecessary except for considerations having to do with strong approximation. All points of the corresponding Shimura varieties are defined over (the reflex field) $\mathbb{Q}$, but automorphic forms are rational over the fields of definition of their coefficients.

We can determine these fields of definition easily. Let (ρ,V) be an irreducible algebraic representation of G. An automorphic form on G of type ρ is a function $f:G(F)\backslash G(\mathbf{A})\to V(\mathbb{C})$, locally constant with respect to $G(\mathbf{A}^f)$, and satisfying

$$f(gg_\infty)=\rho^{-1}(g_\infty)f(g),\quad \text{for } g\in G(\mathbf{A}),\quad g_\infty\in G_\infty=G(F\otimes_\mathbb{Q}\mathbb{R}). \tag{2.5.1}$$

Let $\mathcal{A}(G,\rho)$ denote the space of automorphic forms of type ρ. It follows from (2.5.1) that the restriction map

$$R_f:\mathcal{A}(G,\rho)\to C^\infty(G(F)\backslash G(\mathbf{A}^f),V(\mathbb{C}))$$

is an isomorphism. If V is realized over the number field E_V, then

$$M_{DR}(S(G,x),V):=C^\infty(G(F)\backslash G(\mathbf{A}^f),V(E_V))$$

is an E_V-rational model for $\mathcal{A}(G,\rho)$, and for any $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, there is a canonical isomorphism

$$\sigma(M_{DR}(G,V))\xrightarrow{\sim} M_{DR}(G,\sigma(V)). \tag{2.5.2}$$

The same naturally holds for G'.

Let V_{triv} denote the trivial one-dimensional representation of G.

Lemma 2.5.3. *There is a perfect pairing*

$$M_{DR}(S(G,x),V)\otimes M_{DR}(S(G,x),V^{\vee})\to M_{DR}(S(G,x),V_{\text{triv}})(E_V)\to E_V,$$

where the first map is defined by the natural pairing on coefficients and the second map is integration with respect to Tamagawa measure. The pairings transform under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *by the action* (2.5.2) *on the coefficients* V.

Proof. The first map is obviously rational over E_V, and the second map is rational because the Tamagawa measure of $G(F)\backslash G(\mathbf{A})$ is a rational number. The pairing is perfect because it is essentially given by the L_2-pairing on automorphic forms; see [Harris 1997, Proposition 2.6.12]. □

Now suppose $V\to V'$ is a projection to an irreducible G'-invariant quotient, and let $(V')^{\vee}\to V^{\vee}$ denote the dual inclusion map. The following lemma is proved in the same way as Lemma 2.5.3.

Lemma 2.5.4. *Under these hypotheses, there is a natural* $E_{V,V'}=E_V\cdot E_{V'}$*-rational pairing*

$$M_{DR}(S(G,x),V)\otimes M_{DR}(S(G',x'),(V')^{\vee})\\ \to M_{DR}(S(G',x'),V_{\text{triv}})(E_{V,V'})\to E_{V,V'},$$

where the first map is defined by the natural pairing on coefficients and the second map is integration with respect to Tamagawa measure. The pairings transform under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ *by the action* (2.5.2) *on the coefficients* V, V'.

Corollary 2.5.5. *Let* E *be a number field containing* $E_{V,V'}$*, and suppose*

$$f\in M_{DR}(S(G,x),V)(E),\qquad f^{\vee}\in M_{DR}(S(G,x),V^{\vee})(E),$$
$$f'\in M_{DR}(S(G',x'),(V')^{\vee})(E),\quad f^{\prime,\vee}\in M_{DR}(S(G',x'),V')(E).$$

Define $P(f,f')$, $Q(f,f^{\vee})$, $P(f^{\vee},f^{\prime,\vee})$ *and* $Q(f',f^{\prime,\vee})$ *as in Section 2.1. Then the left hand side of* (2.1.5),

$$\frac{P(f,f')P(f^{\vee},f^{\prime,\vee})}{Q(f,f^{\vee})Q(f',f^{\prime,\vee})},$$

belongs to E *and for any* $\sigma\in\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$,

$$\sigma\left(\frac{P(f,f')P(f^{\vee},f^{\prime,\vee})}{Q(f,f^{\vee})Q(f',f^{\prime,\vee})}\right)=\frac{P(\sigma(f),\sigma(f'))P(\sigma(f^{\vee}),\sigma(f^{\prime,\vee}))}{Q(\sigma(f),\sigma(f^{\vee}))Q(\sigma(f'),\sigma(f^{\prime,\vee}))},$$

where $\sigma(f)\in M_{DR}(S(G,x),\sigma(V))(\sigma(E))$, *etc.*

In [Harris 1997, (2.6.11)] it is explained how to use the highest weight Λ of V, relative to a fixed maximal torus H, to identify $\mathscr{A}(G,\rho)$, and therefore $M_{DR}(S(G,x),V)$, with a subspace of the space $\mathscr{A}(G)$ of $\mathbb{C}$-valued automorphic forms on $G(F)\backslash G(\mathbf{A})$:

$$M_{DR}(S(G,x),V) \xrightarrow{\sim} \mathrm{Hom}_H(\mathbb{C}_{-\Lambda}, \mathscr{A}(G)_{V^\vee}), \tag{2.5.6}$$

where $\mathbb{C}_{-\Lambda}$ is the Λ^{-1}-eigenspace for H in $V^\vee$ and $\mathscr{A}(G)_{V^\vee}$ is the $V^\vee$-isotypic subspace for the action of G_∞ by right translation. The image under this identification naturally has a rational structure over the extension $E(V,\Lambda)\supset E(V)$ over which the Λ-eigenspace in V is rational, and as V and H vary the maps (2.5.6) are rational over $E(V,\Lambda)$ and transform naturally under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Lemma 2.5.7. *The map* (2.5.6) *takes the pairing of Lemma 2.5.3 to a rational multiple of the L_2-pairing on $\mathscr{A}(G)$.*

Proof. This is [Harris 1997, Proposition 2.6.12]. □

2.6. ***Fields of rationality of automorphic representations of unitary groups.*** In this section, F is a general totally real field. Let Π be a cohomological cuspidal automorphic representation of $\mathrm{GL}(n,\mathcal{K})$, and let $E(\Pi)$ be the field fixed by the subgroup of $\mathrm{Aut}(\mathbb{C})$ consisting of σ such that $\Pi_f^\sigma \xrightarrow{\sim} \Pi_f$. It is known [Clozel 1990] that $E(\Pi)$ is a number field and that Π_f has a rational model over $E(\Pi)$. Moreover, for any σ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ there is a (unique) cuspidal cohomological representation $\sigma(\Pi)$ with $\sigma(\Pi)_f \xrightarrow{\sim} \sigma(\Pi_f)$ — one obtains $\sigma(\Pi)_\infty$ from Π_∞ by letting σ permute the archimedean places of $\mathcal{K}$.

Suppose Π satisfies the polarization condition (1.1.1) and G is quasisplit at all finite places of v. Then Π descends to an L-packet $\{\pi_\alpha, \alpha\in A\}$ of G [Labesse 2011, Theorem 5.4]. We mean this in the following sense: let w be a finite place of $\mathcal{K}$ at which $\mathcal{K}/F$ and Π are unramified, and let v denote the restriction of w to F. If v splits in $\mathcal{K}$, we write $\Pi_v=\Pi_w\otimes\Pi_{cw}$; if v is inert, then $\Pi_v=\Pi_w$. Then for all α, $\pi_{\alpha,v}$ is spherical and the Satake parameters of Π_v are obtained from those of $\pi_{\alpha,v}$ by the stable base change map [Mínguez 2011, Theorem 4.1]. It then follows that π_∞ is the unique irreducible representation of the (compact) group G_∞ with the same infinitesimal character as Π_∞ [Labesse 2011, Theorem 5.5].

Proposition 2.6.1. *If Π is a cohomological cuspidal polarized representation of* $\mathrm{GL}(n)$ *that descends to an L-packet $\{\pi_\alpha\}$ of G, then the collection $\{\pi_{\alpha,f}\}$ is rational over $E(\Pi)$. Moreover, for any σ in* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, *the conjugate $\sigma(\Pi)$ descends to $\{\sigma(\pi)\}$.*

Proof. Let S be the set of finite primes v at which $\mathcal{K}/F$ and Π are unramified. We first note that for all $v\notin S$, the spherical representation $\pi_{\alpha,v}$ is defined over the field of definition of Π_v. Indeed, this is clear from the relation [Mínguez 2011, Theorem 4.1] of Satake parameters. Now let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. If σ fixes $E(\Pi)$, then $\sigma(\pi_{\alpha,v}) \xrightarrow{\sim} \pi_{\alpha,v}$ for all $v \notin S$. Thus by definition, the stable base change of $\sigma(\pi_f)$ is Π, so $\sigma(\pi_\alpha)$ is a $\pi_{\alpha'}$. The same argument implies the last assertion. □

3. Abelian representations of U(m)

3.1. *Existence of abelian representations.* In this section, the Weil group of a local or global field L is denoted W_L.

Let W' be an m-dimensional hermitian space over $\mathcal{K}$, and $\mathrm{U}(W')$ be the unitary group. Let μ be a Hecke character of $\mathcal{K}$ extending $\eta_{\mathcal{K}/F}$, that is, $\mu|_{\mathbf{A}_F^\times} = \eta_{\mathcal{K}/F}$. Let $H = U(1)^m$ and let $\xi_\mu : {}^L H \to {}^L U(W')$ be the L-homomorphism (in the Weil group form over F) considered by White [2010, Section 3]. On the dual group $\hat{H} = \mathrm{GL}(1, \mathbb{C})^m$, ξ_μ is just the diagonal embedding

$$(g_1, \dots, g_m) \mapsto \mathrm{diag}(g_1, \dots, g_m) \in \hat{U}(W') = \mathrm{GL}(m, \mathbb{C}).$$

The Hecke character μ defines a character $W_{\mathcal{K}} \to W_{\mathcal{K}}^{ab} \xrightarrow{\sim} \mathbf{A}_{\mathcal{K}}^\times/\mathcal{K}^\times \xrightarrow{\mu} \mathbb{C}$, also denoted μ. Set $\mu_m = \mu$ if m is even, $\mu_m = 1$ (the trivial character) if m is odd.

If $w \in W_{\mathcal{K}}$, we have

$$\begin{aligned}\xi_\mu(1, 1, \dots, 1) \times w = \mu_m(w) \cdot I_m \times w \in \mathrm{GL}(n, \mathbb{C}) \times W_{\mathcal{K}} \\ \subset \mathrm{GL}(m, \mathbb{C}) \ltimes W_F = {}^L U(W'). \quad (3.1.1)\end{aligned}$$

The map ξ_μ is characterized by these formulas and by its value on a single element of $(1 \times W_F) \setminus (1 \times W_{\mathcal{K}})$, as in [White 2010]; we omit the formula.

Let $\boldsymbol{\chi} = (\chi_1, \dots, \chi_m)$ be an m-tuple of Hecke characters of $\mathrm{U}(1)(\mathbf{A}_F)/U(1)(F)$; $\boldsymbol{\chi}$ is an automorphic representation of H, and we can consider its functorial transfer to $\mathrm{U}(W')$ via the L-homomorphism ξ_μ. Concretely, an automorphic representation $\pi(\boldsymbol{\chi})$ of $\mathrm{U}(W')$ is a functorial transfer of $\boldsymbol{\chi}$ if its formal base change $\Pi(\boldsymbol{\chi}) = \mathrm{BC}(\pi(\boldsymbol{\chi}))$ to $\mathrm{GL}(m)_{\mathcal{K}}$ is a (noncuspidal) automorphic representation with the property

$$L(s, \Pi(\boldsymbol{\chi})) = \prod_{i=1}^{m} L\big(s + \tfrac{1}{2}(m-1), \mathrm{BC}(\chi_i) \cdot \mu_m\big). \quad (3.1.2)$$

Here,

$$\mathrm{BC}(\chi)(z) = \chi(z/c(z)), z \in \mathbf{A}_{\mathcal{K}}^\times, \quad (3.1.3)$$

where c denotes Galois conjugation; this was denoted $\tilde{\chi}$ in [Harris 1997]. By definition, the functorial transfers of $\boldsymbol{\chi}$ to $\mathrm{U}(W')$ form a single L-packet $\boldsymbol{\pi}(\boldsymbol{\chi})$ such that, for each place v of F, π_v is a local functorial transfer of $\boldsymbol{\chi}_v$ for any $\pi \in \boldsymbol{\pi}(\boldsymbol{\chi})$.

An L-packet of the form $\boldsymbol{\pi}(\boldsymbol{\chi})$ will be called an *abelian L-packet* of $\mathrm{U}(W')$, and a member of $\boldsymbol{\pi}(\boldsymbol{\chi})$ that occurs with nonzero multiplicity in the automorphic spectrum of $\mathrm{U}(W')$ is called an *abelian representation*. The existence of abelian representations in this sense is considered in [White 2010], along with other cases of endoscopic transfer. More precisely, one can say that the local functorial transfers are the L-packets defined by Moeglin [2007] — we denote them $\boldsymbol{\pi}(\boldsymbol{\chi}_v)$ — and that if we choose one $\pi_v \in \boldsymbol{\pi}(\boldsymbol{\chi}_v)$ for each v, then we can ask for the multiplicity of $\bigotimes'_v \pi_v$ in the automorphic spectrum of $\mathrm{U}(W')$. These multiplicities are predicted by Arthur's conjectures. We return to this point in Section 4.3.

Let v be a real prime of F and suppose $\chi_{j,v}(e^{i\theta}) = e^{ik_j\theta}$, with $k_j \in \mathbb{Z}$. We say that k_j is the *weight* of χ_j at v (or of $\chi_{j,v}$). The Langlands parameter of $\chi_{j,v}$ is given by the homomorphism $\phi(\chi_{j,v}) : W_{\mathbb{R}} \to {}^L U(1) = \mathrm{GL}(1, \mathbb{C}) \rtimes \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ whose restriction to $\mathbb{C}^\times = W_{\mathbb{C}}$ is

$$W_{\mathbb{C}} \ni z \mapsto (z/\bar{z})^{k_j}.$$

Then $\mathrm{BC}_{\mathbb{C}/\mathbb{R}}(\Pi(\boldsymbol{\chi}_v))$ is the representation of $\mathrm{GL}(n, \mathbb{C})$ with Langlands parameter

$$\phi(\boldsymbol{\chi}_v) : W_{\mathbb{C}} \ni z \mapsto \mathrm{diag}((z/\bar{z})^{k_1} \cdot \mu_m(z), \dots, (z/\bar{z})^{k_m} \cdot \mu_m(z)) \in \mathrm{GL}(m, \mathbb{C}). \quad (3.1.4)$$

This descends to a discrete series L-packet of $\mathrm{U}(W')_v$, for any W', if and only if the k_j are all distinct [White 2010, Definition 5.3]; then the infinitesimal character of the discrete series L-packet coincides with the Langlands parameter, and we say $\boldsymbol{\chi}_v$ is *regular*.

On $\mathrm{U}(1) \subset \mathbb{C}^\times$ we write $\mu_m(e^{i\theta}) = e^{it_m\theta}$ for some $t_m \in \mathbb{Z}$. We order the k_i so that

$$k_i > k_{i+1} \quad (3.1.5)$$

with k_i defined by

$$(z/\bar{z})^{k_i} \cdot \mu_m(z) = (z/\bar{z})^{k_i + t_m/2}, \quad k_i + \tfrac{1}{2}t_m \in \mathbb{Z} + \tfrac{1}{2}(m-1).$$

The half-integrality of $k_i + \frac{1}{2}t_m$ follows from the parity of μ_m and is as it should be; see [Clozel 1990, Section 3.5].

We can immediately prove the following:

Lemma 3.1.6. *Suppose $\boldsymbol{\chi}_v$ is regular for all real primes v. Then the local Langlands parameter $\phi(\boldsymbol{\chi}_v)$ is relevant for all $\mathrm{U}(W')_v$ and for any W' the L-packet $\boldsymbol{\pi}(\boldsymbol{\chi})$ of $\mathrm{U}(W')$ is of discrete series type at all real places.*

The definite case. Suppose now $\mathrm{U}(W'_v)$ is the compact form of $\mathrm{U}(m)$. Then the L-packet $\boldsymbol{\pi}(\boldsymbol{\chi})$ is a singleton τ' with highest weight $(b_1 \geq b_2 \geq \cdots \geq b_m)$, in the notation of Section 2.3. The relation between b_i and k_i is given by

$$b_i = k_i - \tfrac{1}{2}(-t_m + m + 1 - 2i) \quad (3.1.7)$$

so that $b_i \geq b_{i+1}$, as required.

In what follows, we assume we are given a nontrivial abelian L-packet $\boldsymbol{\pi}(\boldsymbol{\chi})$ and apply it in the Ichino–Ikeda conjecture. Henceforward we specialize to the case $F=\mathbb{Q}$, $m=n-1$, with n even, so $\mu_m=1$ and $k_i=b_i+\frac{1}{2}n-i$. This will suffice to illustrate the general principles guiding this work. We hope to treat the general case in a subsequent paper.

3.2. ***Review of CM periods.*** We review the properties of the CM period invariants, as discussed in [Harris 1997, (1.10) and (3.6)]. Since the final results will only be stated when $F=\mathbb{Q}$, we only consider the CM periods attached to imaginary quadratic fields. Details of the more general CM periods have only been written up in the present language up to algebraic factors; most of the results of the present paper can be extended to general CM fields without going beyond the available literature, provided one is will to settle for rationality up to $\overline{\mathbb{Q}}^{\times}$.

Thus, $\mathcal{K}$ is an imaginary quadratic field, with chosen embedding $\mathcal{K}\to\mathbb{C}$, denoted $\mathbb{1}$. Let $\eta:\mathbf{A}_{\mathcal{K}}^{\times}/\mathcal{K}^{\times}\to\mathbb{C}^{\times}$ be a Hecke character whose archimedean part is algebraic: $\eta_\infty(z)=z^{-a_{\mathbb{1}}}\cdot(cz)^{-a_c}$ for $z\in\mathbb{C}^{\times}$, with the exponents in $\mathbb{Z}$. Let $E(\eta)\supset\mathcal{K}$ be the field generated by $\eta|_{\mathbf{A}_{\mathcal{K}}^{f,\times}}$, and let ${}^c\eta=\eta\circ c$. There are then two period invariants

$$p(\eta,\mathbb{1}),\ p(\eta,c)=p({}^c\eta,\mathbb{1})\in(E(\eta)\otimes\mathbb{C})^{\times}/E(\eta)^{\times}.$$

These invariants satisfy the multiplicative relations

$$p(\eta_1,\cdot\,)p(\eta_2,\cdot\,)\sim_{\tilde{E}(\eta_1,\eta_2)}p(\eta_1\eta_2,\cdot\,),\quad\text{where }\cdot=\mathbb{1},c,\tag{3.2.1}$$

and the normalization conditions (here $\|\cdot\|$ is the norm)

$$p(\|\cdot\|^a,\mathbb{1})=p(\|\cdot\|^a,c)=(2\pi i)^{-a}.\tag{3.2.2}$$

If η is the Hecke character attached to a Dirichlet character of conductor N (with archimedean component a power of the sign character) and $\psi:\mathbb{Z}/N\mathbb{Z}\to\mathbb{C}^{\times}$ is an additive character, then

$$p(\eta,\mathbb{1})=g(\eta,\psi)^{-1},\tag{3.2.3}$$

where $g(\eta,\psi)=\sum_{b\in(\mathbb{Z}/N\mathbb{Z})^{\times}}\eta(b)\psi(b)$ is the standard Gauss sum. If $(a_{\mathbb{1}},a_c)=(k,0)$, with $k>0$, then for all critical values m of the Hecke L-function $L(s,\eta)$, we have

$$L(m,\eta)=L(0,\eta\cdot\|\cdot\|^{-m})\sim_{E(\eta),\mathcal{K}}(2\pi i)^m p(\check{\eta},\mathbb{1})\tag{3.2.4}$$

where $\check{\eta}(z)=\eta^{-1}(cz)$. In particular, if χ is a character of the group $\mathrm{U}(1)$ as above, then $\mathrm{BC}(\chi)=\mathrm{BC}(\chi)\check{}$, so for critical values

$$\begin{aligned}L(m,\mathrm{BC}(\chi))&\sim_{E(\chi),\mathcal{K}}(2\pi i)^m p(\mathrm{BC}(\chi),\mathbb{1})\\&\sim_{E(\chi),\mathcal{K}}(2\pi i)^m p(\chi^+,\mathbb{1})p({}^c\chi^+,\mathbb{1})^{-1}\end{aligned}\tag{3.2.5}$$

for any extension χ^+ of χ to an algebraic Hecke character of $\mathcal{K}$.

3.3. ***Asai L-functions of abelian representations.*** Fix $\boldsymbol{\chi}$ as in the previous section, and let $\Pi = \Pi(\boldsymbol{\chi})$. The formula (3.1.2) gives an explicit expression for the motive $M_{\Pi(\boldsymbol{\chi})}$ over $\mathcal{K}$:

$$M_{\Pi(\boldsymbol{\chi})} = \bigoplus_{i=1}^{n-1} M_{\mathrm{BC}(\chi_i)}\Big(\frac{2-n}{2}\Big). \tag{3.3.1}$$

It then follows from the definitions that $L(s, \mathrm{As}(M_{\Pi(\boldsymbol{\chi})}))$, which is an L-function over F $(=\mathbb{Q})$, decomposes as

$$\begin{aligned} L(s, \mathrm{As}(M_{\Pi(\boldsymbol{\chi})})) &= \prod_{1\le i<j\le n-1} L(s, AI_{\mathcal{K}/F}\,\mathrm{BC}(\chi_j\cdot\chi_i^{-1}))L(s,\eta_{\mathcal{K}/F})^{n-1} \\ &= \prod_{1\le i<j\le n-1} L(s, AI_{\mathcal{K}/F}\,\mathrm{BC}(\chi_{ij}))L(s,\eta_{\mathcal{K}/F})^{n-1}, \end{aligned} \tag{3.3.2}$$

where $\chi_{ij} = \chi_j/\chi_i$. Indeed,

$$L(s, \mathrm{Ad}(M_{\Pi(\boldsymbol{\chi})})) = \prod_{1\le i\ne j\le n-1} L(s, \mathrm{BC}(\chi_j\cdot\chi_i^{-1}))\zeta_{\mathcal{K}}(s)^{n-1},$$

where $\zeta_{\mathcal{K}}$ is the Dedekind zeta function. The two descents $\mathrm{As}^{\pm}$ are distinguished by their L-functions over F; in addition to the one indicated in (3.3.2), there is the one obtained by twisting by $\eta_{\mathcal{K}/F}$, namely

$$\prod_{1\le i<j\le n-1} L(s, AI_{\mathcal{K}/F}\,\mathrm{BC}(\chi_j\cdot\chi_i^{-1}))\zeta_F(s)^{n-1}.$$

The condition on the signature of F_∞ guarantees that (3.3.2) is the right choice for $\mathrm{As}(M_{\Pi(\boldsymbol{\chi})})$.

We evaluate the values at $s=1$ of the factors of (3.3.2) using Blasius' result on special values of Hecke L-series (Damarell's formula in this case). As in Section 3.1, we assume χ_i is of weight k_i at the archimedean prime, so that χ_{ij} is of weight $-k_{ij}$, with $k_{ij} = k_i - k_j$. We assume the χ_i are ordered so that $k_{ij} > 0$ for $i<j$, as in Formula 2.3.1. This is the normalization used in [Harris 1997]. As in [ibid., Section 2.9], we define

$$\chi_{ij}^{(2)} = \chi_{ij}^2\cdot(\chi_{ij,0}\circ N_{\mathcal{K}/\mathbb{Q}})^{-1}, \quad \text{where } \chi_{ij,0} = \chi_{ij}|_{\mathbf{A}_{\mathbb{Q}}^{\times}}\cdot\|\cdot\|_{\mathbf{A}}^{-k_{ij}}. \tag{3.3.3}$$

Then (see [Harris 1997, (3.6.1), (3.6.3)]),

$$L(1, \mathrm{BC}(\chi_{ij})) = L(1+k_i-k_j, \chi_{ij}^{(2)}) \sim (2\pi i)^{1+k_i-k_j} p((\chi_{ij}^{(2)})^{\vee}, 1).$$

By using the formula $\chi_{ij}^{(2)} = \chi_j^{(2)}/\chi_i^{(2)}$ and the relations in Section 3.2, we find that the value at 1 of (3.3.2) is

$$[(2\pi i)g(\eta_{\mathcal{K}/F})]^{n-1}\cdot\prod_{i<j}(2\pi i)^{1+k_i-k_j}p((\chi_{ij}^{(2)})^\vee, 1)$$

$$\sim [(2\pi i)g(\eta_{\mathcal{K}/F})]^{n-1}\cdot(2\pi i)^{(n-2)(n-1)/2}\cdot\prod_{i=1}^{n-1}[(2\pi i)^{k_i}p((\chi_i^{(2)})^\vee, 1)]^{2i-n}$$

$$\sim g(\eta_{\mathcal{K}/F})^{n-1}\cdot(2\pi i)^{n(n-1)/2}\cdot\prod_{i=1}^{n-1}[(2\pi i)^{k_i}p((\chi_i^{(2)})^\vee, 1)]^{2i-n} \tag{3.3.4}$$

Comparing this formula with Corollary 1.3.5(i), it is reasonable to suppose that

$$Q_\ell = [(2\pi i)^{k_\ell}p((\chi_\ell^{(2)})^\vee, 1)]^{-2} \quad \text{for } \ell = 1, \dots, n-1, \tag{3.3.5}$$

so that $[(2\pi i)^{k_\ell}p((\chi_\ell^{(2)})^\vee, 1)]^{2\ell-n} = Q_\ell^{((n-1)+1)/2-\ell}$, as predicted. However, it will not be necessary to verify this formula, since the same expression reappears in the numerator of the Ichino–Ikeda formula in the applications.

4. The critical value of the Asai L-function

We continue to assume $F = \mathbb{Q}$ and n is even. Henceforward the groups G and G' are assumed to be definite. We let f, $f^\vee$, f', $f'^{,\vee}$ be automorphic forms as in the statement of the Ichino–Ikeda conjecture, and we assume they are all E-rational, as in the statement of Corollary 2.5.5.

We begin by studying the L-functions that occur on the right hand side of the Ichino–Ikeda conjecture for the pair π and π'. Starting in Section 4.2, we will assume $\pi' \in \boldsymbol{\pi}(\boldsymbol{\chi})$ for an appropriate $(n-1)$-tuple $\boldsymbol{\chi}$ of Hecke characters. The weights of $\boldsymbol{\chi}$ will be chosen so that the unitary groups that occur on the left hand side of (2.1.5), and in the zeta integrals on the right hand side, are necessarily definite, as in Section 2.3. The left hand side is then an algebraic number, as we have seen in Corollary 2.5.5. We conclude with an expression for the value $L(1, \pi, \mathrm{Ad})$, which we compare to the conjectured expression from Section 1.3.

4.1. ***Elementary and local terms in the Ichino–Ikeda formula for definite groups.*** The left hand side of the Ichino–Ikeda conjecture (2.1.5) was studied in Section 2.5. Corollary 2.5.5 demonstrates that it is an algebraic number that transforms as expected under Galois conjugation. Thus the Ichino–Ikeda conjecture implies that the right hand side is also algebraic, and determines how it transforms under Galois conjugation. In this section we study the algebraicity of the elementary and local terms.

4.1.1. The power of 2 that appears as the first term is, of course, rational.

4.1.2 *The normalizing factor.* The abelian normalizing factor Δ_G is a product of n abelian L-functions of $\mathbb{Q}$ — either $\zeta(s)$ or $L(s, \eta_{\mathcal{K}/\mathbb{Q}})$ depending on the parity — evaluated at integer points. Each of the integer points is well known to be critical, and the formulas for the special values can be written as follows:

$$\Delta_G \sim_{\mathcal{K}} \prod_{i=1}^{n} g(\eta^i_{\mathcal{K}/\mathbb{Q}}) \cdot (2\pi i)^i = (2\pi i)^{n(n+1)/2} g(\eta_{\mathcal{K}/\mathbb{Q}})^{n/2}.$$

Here $\sim_{\mathcal{K}}$ means that the left hand side is a $\mathcal{K}^\times$-multiple of the right hand side. By the Iwasawa main conjecture, the integral properties of $\Delta_G/(2\pi i)^{n(n+1)/2}$ are closely related to orders of class groups of cyclotomic fields.

4.1.3 *Factorization.* For the next section, we need to write f, $f^\vee$, f', $f'^{,\vee}$ as tensor products of vectors $f = \bigotimes_v f_v$, $f_v \in \pi_v$, and so on. Let $E(\pi) \supset E(V)$ and $E(\pi') \supset E(V')$ denote fields of definition of π and π', respectively. In particular, each factor π_v is defined over $E(\pi)$, and we can assume that the isomorphisms $\pi \xrightarrow{\sim} \bigotimes_v \pi_v$ and $\pi' \xrightarrow{\sim} \bigotimes_v \pi'_v$ (and the corresponding dual maps) are defined over $E(\pi)$ and $E(\pi')$, respectively. Our hypothesis is that the test vectors on the left hand side of (2.1.5) are all E-rational; thus f_v, f'_v, $f^\vee_v$, $f'^{,\vee}_v$ are also E-rational for all v.

Moreover, the canonical local pairings $\langle \cdot, \cdot \rangle_{\pi_v}$ and $\langle \cdot, \cdot \rangle_{\pi'_v}$ are tautologically $E(\pi)$- and $E(\pi')$-rational, respectively. It follows that the matrix coefficients $c_{f_v, f^\vee_v}(g_v)$ and $c_{f'_v, f'^{,\vee}_v}(g'_v)$ are E-rational. For finite v, this means that they are functions that take values in the indicated number fields. For $v = \infty$, an E-rational matrix coefficient of the algebraic representation π_∞ is an element of the affine algebra $E(G)$ of the algebraic group G; likewise for π'_∞.

4.1.4 *Measures and archimedean local terms.* We want to prove that the product Z_{loc} of local terms on the right hand side of (2.1.5) is an algebraic number that transforms appropriately under Galois conjugation. We begin by reconsidering the factorization $dg' = \prod_v dg'_v$ of Tamagawa measure. For the moment F is an arbitrary totally real field, and $G_\infty = \prod_{v|\infty} G_v$ is the product of definite unitary groups. For $v \notin S$, let $K'_v \subset G'_v$ be a hyperspecial maximal compact subgroup; we recall from Section 2.1 that $\int_{K'_v} dg'_v = 1$ for $v \notin S$.

Lemma 4.1.5. *For any sufficiently small open subgroup $\prod_{v \in S} K'_v \subset \prod_{v \in S} G'_v$, the open subgroup $G'_\infty \times \prod_{v \nmid \infty} K'_v \subset G(\mathbf{A})$ acts freely (on the right) on $G'(F)\backslash G'(\mathbf{A})$ with finitely many orbits. In particular, $\int_{G'_\infty \times \prod_{v\nmid\infty} K'_v} dg$ is a rational number.*

Proof. Let $\mathrm{U} = G'_\infty \times \prod_v K'_v$, and let $g \in G(\mathbf{A})$ be a fixed point of some $u \in U$. Thus $gu = \gamma g$ for some $\gamma \in G'(F)$, or $gug^{-1} \in gUg^{-1} \cap G'(F)$. It's well known that this intersection is trivial if U is sufficiently small; see the proof of [Clozel et al. 2008, Lemma 3.3.1]. Finiteness of the number of orbits is clear because U is open in $G'(\mathbf{A})$ and $G'(F)\backslash G'(\mathbf{A})$ is compact. The final assertion follows from the first because the Tamagawa number of G' is rational (in fact it equals 2). $\square$

Corollary 4.1.6. *The volume of G'_∞ with respect to $dg_\infty = \prod_{v|\infty} dg_v$ is rational.*

Proof. Indeed,

$$\int_{G'_\infty} dg_\infty = \frac{\int_{G'_\infty \times \prod_v K'_v} dg}{\int \prod_{v \nmid \infty} K'_v}.$$

The numerator is rational by the lemma, and the denominator is rational by conditions (1) and (2) of Section 2.1. $\square$

Now for simplicity we assume $F = \mathbb{Q}$, so that there is only one archimedean prime.

Corollary 4.1.7. *The archimedean local factor Z_∞ of Z_{loc} is an algebraic number.*

Proof. It follows from Lemma 2.5.7 that Z_∞ is a rational multiple of the integral of a product of E-rational matrix coefficients of two algebraic representations of G'_v with respect to the measure of total volume 1. By the orthogonality relations, this is an element of E. $\square$

4.1.8 *Nonarchimedean local factors.* Let $p \in S$ be a finite prime and let E be a number field over which both π_p and π'_p are defined. Then it makes sense to speak of E-rational matrix coefficients $c_{f_p, f_p^\vee}$ and $c_{f'_p, f_p'^{,\vee}}$ of π_p and π'_p, respectively. Recall that in Section 2.1 we have assumed that local measures at finite primes take rational values on compact open subsets.

Lemma 4.1.9. *Suppose π_p and π'_p are tempered. For any E-rational matrix coefficients $c_{f_p, f_p^\vee}$ and $c_{f'_p, f_p'^{,\vee}}$ as above, the local zeta integral has the property that*

$$Z_p(f_p, f_p^\vee, f'_p, f_p'^{,\vee}) \in E.$$

In [Ichino and Ikeda 2010; N. Harris 2011] it is proved that the integral defining $Z_p(f_p, f_p^\vee, f'_p, f_p'^{,\vee})$ converges absolutely when the two representations are tempered, but no information is given about the rationality of the integral. Using Casselman's results on asymptotics of matrix coefficients, Moeglin and Waldspurger [2012, Lemma 1.7] decompose the analogous integral for pairs of special orthogonal groups (even in the nontempered case) into a finite sum of terms that can easily be seen to be rational over E.

More precisely, we write G and G' for the local groups at p. Assume π and π' are constituents of representations induced from supercuspidal representations of the Levi components M and M' of parabolic subgroups $P \subset G$ and $P' \subset G'$, respectively, with M and M' respectively of (split) rank t and t'. Thus π and π' belong to complex families (components of the respective Bernstein centers) $C(\pi)$

and $C(\pi')$ of dimension t and t', parametrized by characters $X(M)$ of M and M', modulo the actions of the normalizers $W_M = N_G(M)/M$ and $W_{M'} = N_{G'}(M')/M'$:

$$C(\pi) = \operatorname{Spec}(\mathbb{C}[X(M)]^{W_M}), \quad C(\pi') = \operatorname{Spec}(\mathbb{C}[X(M')]^{W_{M'}}). \tag{4.1.10}$$

These complex families have rational structures over $\mathbb{Q}$ whose E-rational points are the E-rational orbits of W_M and $W_{M'}$ on the character groups. The functions f_p, $f_p^\vee$ and f'_p, $f_p^{\prime,\vee}$ can be extended to E-rational algebraic functions on $C(\pi)$ and $C(\pi')$. The lemma proved by Moeglin and Waldspurger (in the orthogonal case, but the argument works as well for unitary groups) is then:

Lemma 4.1.11 (Moeglin, Waldspurger). *There are polynomials*

$$D, L \in \mathbb{C}[X(M), X(M')],$$

depending on $f_p, f_p^\vee, f'_p, f_p^{\prime,\vee}$, *such that* $D \cdot Z_p(f_p, f_p^\vee, f'_p, f_p^{\prime,\vee}) = L$.

For the proof of the lemma, it is not assumed that π and π' are tempered. In the tempered case, the convergence proved in [Ichino and Ikeda 2010; N. Harris 2011] implies that D has no pole at the point corresponding to $\pi, \pi' \in C(\pi) \times C(\pi')$.

For our purposes, the important point is that every step in the proof in [Moeglin and Waldspurger 2012] is rational over E. The main reduction step is the expression of the integral as a finite sum of terms indexed by rational parabolic subgroups of G or G', in which the matrix coefficients are replaced by corresponding expressions involving the nonnormalized Jacquet modules. Since the nonnormalized Jacquet functor preserves rationality over $\mathbb{Q}$, the proof of Lemma 4.1.11 actually yields Lemma 4.1.9.

4.1.12 *Conclusion.* Combining the results obtained above with Corollary 2.5.5, we find that

$$\frac{(2\pi i)^{n(n+1)/2} L(\frac{1}{2}, \mathrm{BC}(\pi) \times \mathrm{BC}(\pi'))}{L(1, \pi, \mathrm{Ad}) L(1, \pi', \mathrm{Ad})} \in \overline{\mathbb{Q}}. \tag{4.1.13}$$

For all $\sigma \in \operatorname{Gal}(\overline{\mathbb{Q}}/\mathcal{K})$,

$$\sigma\left[\frac{(2\pi i)^{n(n+1)/2} L(\frac{1}{2}, \mathrm{BC}(\pi) \times \mathrm{BC}(\pi'))}{L(1, \pi, \mathrm{Ad}) L(1, \pi', \mathrm{Ad})}\right] = \frac{(2\pi i)^{n(n+1)/2} L(\frac{1}{2}, \mathrm{BC}(\sigma(\pi)) \times \mathrm{BC}(\sigma(\pi')))}{L(1, \sigma(\pi), \mathrm{Ad}) L(1, \sigma(\pi'), \mathrm{Ad})}. \tag{4.1.14}$$

Including the Gauss sums that appear in 4.1.2 in the expression (4.1.13) would allow us to assert the modified version of (4.1.14) for all $\sigma \in \operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. However, the subsequent calculations are taken from [Harris 1997] and have are only been proved for conjugation by $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathcal{K})$.

4.2. ***Tensor products involving abelian representations.*** Let π and π' be automorphic representations of the definite unitary groups G and G', as in Section 2.3, with base changes Π and Π' to $\mathrm{GL}(n)_{\mathcal{K}}$ and $\mathrm{GL}(n-1)_{\mathcal{K}}$, respectively, and with central characters χ_π and $\chi_{\pi'}$. We assume $L(\tau, \tau') \neq 0$, with $\tau = \pi_\infty$ and $\tau' = \pi'_\infty$; thus the highest weights of τ and τ' satisfy the branching law 2.3.1. Our goal is to understand the special value $L(1, \pi, \mathrm{Ad})$. This is unchanged when π is twisted by a Hecke character, so we lose no generality if we assume the highest weight of $\tau = \pi_\infty$, with parameters as in Section 2.3, has the form $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$. It then follows from 2.3.1 that the k_j are all *negative*.

We assume $\pi' \in \mathbf{\Pi}(\boldsymbol{\chi})$. Then (since $\mu_{n-1} = 1$)

$$L(s, \Pi \times \Pi') = \prod_{i=1}^{n-1} L(s, \Pi \otimes \mathrm{BC}(\chi_j)) = \prod_{i=1}^{n-1} L(s, \pi \otimes \chi_i \circ \det, St). \qquad (4.2.1)$$

Here St is the standard L-function of the L-group of G in the unitary normalization, as in [Harris 1997]. In the motivic normalization (see [Harris 1997]), we then have

$$L(s, \Pi \times \Pi') = \prod_{i=1}^{n-1} L^{\mathrm{mot}}\big(s + \tfrac{1}{2}(n-1), \pi \otimes \chi_i \circ \det, St\big). \qquad (4.2.2)$$

Lemma 4.2.3. *The value $s_0 = n/2$ is critical in Deligne's sense for each of the factors $L^{\mathrm{mot}}(s, \pi \otimes \chi_i \circ \det, St)$.*

(*If n were odd, there would be a shift of $\frac{1}{2}$ to compensate the character μ.*)

Proof. The line $\mathrm{Re}(s) = s_0$ is the axis of symmetry for the functional equation, and the integral point on the axis of symmetry of the L-function of a motive is critical whenever the motive is of odd weight. The motive in question is $M(\Pi) \otimes M(\mathrm{BC}(\chi_i))$. Since $M(\Pi)$ is of weight $n-1$ and $M(\mathrm{BC}(\chi))$ is of weight 0 for any algebraic Hecke character χ, the lemma follows. □

Thus $L(s_0, \Pi \times \Pi')$ can be expressed in terms of automorphic periods using the formulas in [Harris 1997; Harris 2008].

Lemma 4.2.4. *In the terminology of* [Harris 1997, Section 1.7], *the character* $\mathrm{BC}(\chi_i)$ *belongs to the i-th critical interval for $M(\Pi)$, where $i = 1, \ldots, n-1$.*

Proof. Recall from [Harris 1997] that the i-th critical interval is the interval

$$\begin{aligned}[n - 2p_i, n - 2p_{i+1} - 2] &= [n - 2(n - i - a_{n+1-i}), n - 2(n - i - a_{n-i})] \\ &= [2a_{n+1-i} - n + 2i, 2a_{n-i} - n + 2i],\end{aligned}$$

where the first equality is (2.3.2). On the other hand, up to a twist by a power of the norm character $z\bar{z}$, $\mathrm{BC}(\chi_i)_\infty$ is of weight $-2k_i = -2b_i - n + 2i$ (according to the conventions of [Harris 1997, p. 92]), so the lemma follows from the inequalities Formula 2.3.1(2). □

Now suppose the following hypothesis is satisfied:

Hypothesis NE. For every inner form J of G_∞, there exists an inner form G_J of G with $G_{J,\infty} = J$ and a holomorphic automorphic representation π_J of G_J that is nearly equivalent to π; in other words, such that $\pi_{J,v} \xrightarrow{\sim} \pi_v$ for all but finitely many places v.

Then we can apply [Harris 2008, Theorem 4.3] and find that

$$L^{\mathrm{mot}}(\tfrac{1}{2}n, \pi \otimes \chi_i, St)$$
$$\sim_{E(\pi,\chi_i),\mathcal{K}} (2\pi i)^{n/2+k_i(2i-n)} g(\eta_{\mathcal{K}/F})^{n/2} P^{(n-i)}(\Pi) p((\chi_i^{(2)})^\vee, 1)^{2i-n}$$
$$\sim_{E(\pi,\chi_i),\mathcal{K}} (2\pi i)^{n/2} G(i, \boldsymbol{\chi}) P^{(n-i)}(\Pi),$$

where we have introduced the abbreviation

$$G(i, \boldsymbol{\chi}) = \left[(2\pi i)^{k_i} \cdot p((\chi_i^{(2)})^\vee, 1)\right]^{2i-n},$$

and we have chosen to ignore powers of $g(\eta_{\mathcal{K}/F})$.

The periods $P^{(s)}(\Pi)$ were defined in [Harris 1997, (2.8.2)], where they were denoted $P^{(s)}(\pi, V; \beta)$. Roughly speaking, $P^{(s)}(\pi, V; \beta)$ is the normalized Petersson square norm of a holomorphic automorphic form β on the Shimura variety attached to a unitary group $\mathrm{GU}(V)$ of a hermitian space V of signature (r, s); we assume β is rational over an appropriate coefficient field, and the period $P^{(s)}(\pi, V; \beta)$ is well-defined up to multiplication by a scalar in this coefficient field. In [Harris 1997, Corollary 3.5.12], it is proved under somewhat restrictive hypotheses that $P^{(s)}(\pi, V; \beta)$ depends only on the near equivalence class of π (and on the signature (r, s)), and therefore only on Π. The argument used to prove that corollary can be applied to the result of [Harris 2008, Theorem 4.3] to obtain the same statement under a much weaker hypothesis, namely when the L-functions $L^{\mathrm{mot}}(s, \pi \otimes \chi_i, St)$ have nonvanishing critical values for some χ_i in the corresponding critical interval for Π. Since this is a consequence of hypothesis (3) of Theorem 4.2.6, we will just assume this to be the case; thus it is legitimate to write $P^{(s)}(\Pi)$ as a function of the near-equivalence class.[7]

The statement of [Harris 2008, Theorem 4.3] is conditional on the possibility of representing the special value in question as an integral of a holomorphic automorphic form — hence the need for Hypothesis NE — against an Eisenstein series realized by means of the Siegel–Weil formula. That this is possible for the central value is proved in [Harris et al. 2011, Section 4.2].

[7] Under Hypotheses 4.1.4, 4.1.10, and 4.1.14 of [Harris 2007], Theorem 4.2.1 therein implies immediately that $P^{(s)}(\pi, V; \beta)$ depends only on the near equivalence class of π. The most important of these hypotheses is 4.1.10: Π is cohomological with nontrivial cohomology with coefficients in a representation of $\mathrm{GL}(n)$ of regular highest weight.

In other words,

$$L^{\mathrm{mot}}\big(\tfrac{1}{2}n, \pi \otimes \pi'\big) = \prod_{i=1}^{n-1} L^{\mathrm{mot}}\big(\tfrac{1}{2}n, \pi \otimes \chi_i, St\big)$$

$$\sim_{E(\pi,\{\chi_i\}),\mathcal{K}} (2\pi i)^{n(n-1)/2} \prod_{i=1}^{n-1} G(i, \boldsymbol{\chi}) \cdot P^{(n-i)}(\Pi).$$

Combining this with (3.3.4), and bearing in mind $L(s, \mathrm{As}(\pi')) = L(s, \mathrm{As}(M_{\Pi(\boldsymbol{\chi})}))$, we find

$$\begin{aligned}\frac{L^{\mathrm{mot}}(\frac{1}{2}n, \pi\otimes\pi')}{L(1, \mathrm{As}(\pi'))} &\sim_{E(\pi,\{\chi_i\}),\mathcal{K}} (2\pi i)^{n(n-1)/2} \frac{\prod_{i=1}^{n-1} G(i, \boldsymbol{\chi})\cdot P^{(n-i)}(\Pi)}{(2\pi i)^{n(n-1)/2}\cdot\prod_{i=1}^{n-1} G(i, \boldsymbol{\chi})} \\ &\sim_{E(\pi,\{\chi_i\}),\mathcal{K}} \prod_{i=1}^{n-1} P^{(n-i)}(\Pi)\end{aligned} \tag{4.2.5}$$

The next theorem then follows immediately from (4.2.5) and 4.1.12.

Theorem 4.2.6. *We admit the Ichino–Ikeda conjecture* (2.1.5). *Fix a representation τ of G_∞, and an automorphic representation π of G of infinity type τ. Suppose π satisfies Hypothesis NE, and suppose there exists an* $(n-1)$*-tuple $\boldsymbol{\chi}$ satisfying the following*:

(1) *The L-packet $\Pi(\boldsymbol{\chi})$ on G' is nontrivial.*

(2) *Let τ' denote the common archimedean component of all elements of* $\Pi(\boldsymbol{\chi})$. *Then τ' satisfies the inequalities of Formula 2.3.1(2) relative to τ, that is, $L(\tau, \tau') \neq 0$.*

(3) *For each χ_i, the central value $L^{\mathrm{mot}}(\frac{1}{2}n, \pi \otimes \chi_i, St) = L(\frac{1}{2}, \pi \otimes \chi_i, St) \neq 0$.*

Then

$$L(1, \pi, \mathrm{Ad}) \sim_{E(\pi),\mathcal{K}} (2\pi i)^{n(n+1)/2} \prod_{i=1}^{n-1} P^{(n-i)}(\Pi).$$

Remark 4.2.7. (a) It is legitimate to replace $E(\pi, \{\chi_i\})$ by $E(\pi)$ because we can let the χ_i vary over their Galois conjugates; only π remains on the two sides.

(b) Hypotheses (1) and (3) imply that the central value $L(\frac{1}{2}, \Pi \times BC(\Pi(\chi)))$, which is another expression for the numerator of the left-hand side of (4.2.5), does not vanish. The Ichino–Ikeda conjecture, together with the Gross–Prasad conjecture, then picks out a pair (G_1, G_1') of inner forms of G and G', respectively, and automorphic representations π_1 and π_1' on G_1 and G_1', with $\mathrm{BC}(\pi_1) = \Pi$, $\mathrm{BC}(\pi_1') = \mathrm{BC}(\Pi(\chi))$, such that the left hand side of the identity (2.1.5) does not vanish for some choice of data $f, f', f^\vee, f'^{,\vee}$. In particular, $L(\pi_{1,v}, \pi'_{1,v}) \otimes L(\pi^\vee_{1,v}, \pi'^{,\vee}_{1,v}) \neq 0$ for all places v. Moreover, the quadruple $(G_1, G_1', \pi_1, \pi_1')$ is unique. It follows from hypothesis (2) that $G_{1,\infty} = G_\infty$ and $G'_{1,\infty} = G'_\infty$ are compact. Since $n-1$ is odd, this implies that G'_1 and G' are isomorphic. On the other hand, G_1 may well be different from G at finite places, but since $L(1, \pi_1, \mathrm{Ad}) = L(1, \pi, \mathrm{Ad})$, we need not refer to π_1 in the statement of Theorem 4.2.6.

4.3. *Verification of the hypotheses of Theorem 4.2.6.*

4.3.1. The existence of L-packets $\Pi(\chi)$ satisfying hypotheses (1) and (2) is predicted in most cases by the Langlands functoriality conjectures. Proofs of endoscopic functoriality in related situations are based on the stable Arthur-Selberg trace formula. In the situation at hand, where G' is definite at archimedean places, White has some results to this effect in his thesis [2010, Theorems 5.12 and Theorem 5.15]. Complete results for endoscopic transfer can be found in recent papers of C. P. Mok when the target group G' is quasisplit. There may be obstructions at finite places at which G' is not quasi-split; this should be settled by additional work on the stable trace formula.

4.3.2 *The nonvanishing hypothesis (3) of Theorem 4.2.6.* This hypothesis is not accessible at present. One can conjecture that it is always true, given the freedom one has in choosing $\boldsymbol{\chi}$ in the proof of 4.3.1. For each i one needs to find χ_i of the appropriate weight such that $L(\frac{1}{2}, \pi \otimes \chi_i, St) \neq 0$; equivalently, with χ_i fixed, one needs to find χ'_i of finite order, with trivial restriction to the idèles of $\mathbb{Q}$, such that $L(\frac{1}{2}, \pi \otimes \chi_i \cdot \chi'_i, St) \neq 0$.

The first condition is to find χ'_i such that the sign of the functional equation of $L(\frac{1}{2}, \pi \otimes \chi_i \cdot \chi'_i, St)$ is $+1$. This is a local problem and can always be solved. As explained in [Harris et al. 2011], the local signs $\varepsilon(1/2, \pi_v \otimes \chi_{i,v} \cdot \cdot \chi'_i) \in \{\pm 1\}$ determine a certain Siegel–Weil Eisenstein series on a quasisplit unitary group $\mathrm{U}(n, n)$, and the vanishing of the central value $L(\frac{1}{2}, \pi \otimes \chi_i \cdot \chi'_i, St)$ corresponds to the triviality of the pairing of this Eisenstein series with vectors in

$$(\pi \otimes \chi_i \cdot \chi'_i) \otimes (\pi \otimes \chi_i \cdot \chi'_i)^\vee$$

in the doubling method. However, the Eisenstein series itself is nontrivial, so there are certainly representations π for which $L(\frac{1}{2}, \pi \otimes \chi_i \cdot \chi'_i, St) \neq 0$!

One would like to say that the L-function does not vanish for most π in a family of representations. For the families typically considered by analytic number theorists this also seems to be an inaccessible problem. On the other hand, one can prove such a generic nonvanishing result for p-adic families of automorphic representations, provided one has well-behaved p-adic L-functions for these families. This will be explained in more detail in forthcoming work of the author with Eischen, Li, and Skinner.

4.4. ***Comparison of Theorem 4.2.6 with Deligne's conjecture.*** It remains to compare the expression

$$(2\pi i)^{n(n+1)/2}\prod_{i=1}^{n-1} P^{(n-i)}(\Pi)$$

of Theorem 4.2.6 with the expression

$$d(M)^{1/2}(2\pi i)^{n(n+1)/2}[Q_{\det(M)}]^{(n-1)/2}\cdot\prod_{\ell} Q_{\ell}^{1-\ell}$$

predicted by Deligne's conjecture as expressed in Corollary 1.3.5; in other words, we wish to justify a comparison

$$\prod_{i=1}^{n-1} P^{(n-i)}(\Pi) \sim_{\mathcal{K}} d(M)^{1/2}[Q_{\det(M)}]^{(n-1)/2}\cdot\prod_{\ell} Q_{\ell}^{1-\ell}. \tag{4.4.1}$$

The comparison can only be heuristic, because the invariants Q_ℓ are defined in terms of a hypothetical polarized regular motive, whereas the $P^{(n-i)}(\Pi)$ are normalized Petersson square norms of arithmetic holomorphic automorphic forms on Shimura varieties. We reason as in [Harris 1997, Section 3.7], deriving a version of (4.4.1) from the Tate conjecture. Briefly, we stipulate that the Q_ℓ are defined for a motive $M(\Pi)$ with λ-adic realizations $\rho_{\Pi,\lambda}$, as in Section 1.1, while the $P^{(s)}(\Pi)$ are periods of a motive, say $M^{(s)}(\Pi)$, whose λ-adic realization is isomorphic to an explicit abelian twist of $\bigwedge^{n-s} M(\Pi)^\vee$; see [Harris 1997, 2.7.6.1, 2.7.7, 3.7.9] and the subsequent discussion. More precisely, in view of the Tate conjecture, the relation of L-functions asserted as [ibid., Conjecture 2.7.7] motivates the following version of [ibid., Hypothesis 3.7.9]:[8]

$$\begin{aligned} M^{(s)}(\Pi) &\xrightarrow{\sim} \bigwedge^{r} M(\Pi)^\vee \otimes M(\chi_{\pi^+})(\tfrac{1}{2}r(r-1)),\\ &\xrightarrow{\sim} \Big(\bigwedge^{s} M(\Pi)\Big)\otimes M(\chi_{\pi^+})^{-1}(\tfrac{1}{2}r(r-1)), \end{aligned} \tag{4.4.2}$$

where $r = n - s$ and χ_{π^+} is the central character of any of the representations π^+ of one of the groups $\mathrm{GU}(W) \supset U(W) = G$, the base change of whose restriction to G is Π. With χ_Π as in Section 1.2, we thus have

$$\chi_\Pi = \chi_{\pi^+}/\chi^c_{\pi^+}. \tag{4.4.3}$$

[8] Thanks to progress on the stable trace formula, especially the proof of the Fundamental Lemma, Langlands' Conjecture 2.7.7 on the cohomology of Shimura varieties attached to unitary groups is much closer to being established now than when [Harris 1997] was published. The conjecture has been proved in a number of cases, under simplifying hypotheses, the corresponding relations of automorphic representations are the subject of [Clozel et al. 2011].

To be completely accurate, the restriction of π^+ to G may have several irreducible components π, but they all have the same base change to $\mathrm{GL}(n)$. Note that the relation (4.4.3) is insensitive to the choice of extension of the central character of one such π to the center of $\mathrm{GU}(W)$, which is isomorphic to $\mathrm{GL}(1)_{\mathcal{K}}$. We have made the simplifying hypothesis that the parameter c of (2.4.4) equals 0, so we may assume the restriction of χ_{π^+} to the idèles of $\mathbb{Q}$ is a Hecke character of finite order, in other words a Dirichlet character χ_0.

As in [Harris 1997], (4.4.2) motivates the following relations:

$$P^{(n-i)}(\Pi) \sim_{\mathcal{K}} \prod_{\ell=1}^{n-i} Q_\ell \cdot Q(\chi_{\pi^+})^{-1}.$$

Here $Q(\chi_{\pi^+})$ is defined by analogy with $Q_{\det M}$.

The Tate twist is invisible at this stage because the periods $P^{(s)}$ and Q_ℓ are defined with respect to the de Rham pairing, and $\mathbb{Q}(1)_{DR} = \mathbb{Q}$. Then the left hand side of (4.4.1) is

$$\sim_{\mathcal{K}} \left[\prod_{i=1}^{n-1}\prod_{\ell=1}^{n-i} Q_\ell\right] \cdot Q(\chi_{\pi^+})^{1-n} \sim_{\mathcal{K}} \left[\prod_{\ell=1}^{n-1} Q_\ell^{n-\ell}\right] \cdot Q(\chi_{\pi^+})^{1-n}.$$

Thus the relation (4.4.1) follows from

$$\begin{aligned} Q(\chi_{\pi^+}) \sim_{\mathcal{K}} d(M)^{1/2} Q_{\det M(\Pi)}^{1/2} &\sim_{\mathcal{K}} d(M)^{1/2} Q(\xi_\Pi)^{1/2} \\ &= d(M)^{1/2} Q(\chi_\Pi)^{1/2}, \end{aligned} \tag{4.4.4}$$

where the last relation is (1.2.1), bearing in mind that the Tate twist does not contribute to this calculation, so that $Q(\xi_\Pi) = Q(\chi_\Pi)$. By (4.4.3), the relation (4.4.4) is equivalent to

$$Q(\chi_{\pi^+}) \sim_{\mathcal{K}} d(M)^{1/2} Q(\chi_{\pi^+}/\chi_{\pi^+}^c)^{1/2}. \tag{4.4.5}$$

But $\chi_\pi^c = \chi_\pi^{-1}$ (since it is a character of $\mathrm{U}(1)$), so $\chi_{\pi^+} \cdot \chi_{\pi^+}^c$ factors through the norm from $\mathcal{K}$ to $\mathbb{Q}$.

We hope to provide a hypothetical interpretation of $d(M)$ in a subsequent paper with Guerberoff. In the meantime, we may as well square the two sides of (4.4.5), which reduces the question to

$$Q(\chi_{\pi^+} \cdot \chi_{\pi^+}^c) \sim_{\mathcal{K}} Q(\chi_0 \circ N_{\mathcal{K}/\mathbb{Q}}) \sim_{\mathcal{K}} 1, \tag{4.4.6}$$

with χ_0 as above. Finally, if we are willing to accept the analogue of the relation (3.3.5) (with $k_\ell = 0$), namely,

$$Q(\chi_0 \circ N_{\mathcal{K}/\mathbb{Q}}) = p(([\chi_0 \circ N_{\mathcal{K}/\mathbb{Q}}]^{(2)})^\vee, 1)^{-2},$$

then we are done, because the definition implies that $\chi^{(2)}$ is trivial for any Dirichlet character χ composed with the norm.

Acknowledgments

This article was written in the stimulating environment of the Institute for Advanced Study in Princeton, during the special year on Galois Representations and Automorphic Forms organized by Richard Taylor. I thank the Institute and the organizer for the invitation that made my stay possible. I thank D. Blasius and B. Gross for helpful discussions, Lucio Guerberoff for pointing out several inconsistencies that allowed me to correct my calculations of Deligne periods, and Daniel Barrera Salazar for pointing out a sign mistake in the proof of Lemma 4.2.4. For the statements of the Ichino–Ikeda conjecture for unitary groups I thank Wee Teck Gan and Neal Harris. The anonymous referee deserves special thanks for a very careful reading that led to substantial improvements and clarifications. Finally, I thank the Association of Members of the IAS for their generous support.

References

[Blasius et al. 1994] D. Blasius, M. Harris, and D. Ramakrishnan, "Coherent cohomology, limits of discrete series, and Galois conjugation", *Duke Math. J.* **73**:3 (1994), 647–685. MR 95b:11054 Zbl 0811.11034

[Chenevier and Harris 2013] G. Chenevier and M. Harris, "Construction of automorphic Galois representations, II", 2013. To appear in *Cambridge J. Math.*

[Clozel 1990] L. Clozel, "Motifs et formes automorphes: applications du principe de fonctorialité", pp. 77–159 in *Automorphic forms, Shimura varieties, and L-functions* (Ann Arbor, MI, 1988), vol. 1, edited by L. Clozel and J. S. Milne, Perspect. Math. **10**, Academic Press, Boston, MA, 1990. MR 91k:11042 Zbl 0705.11029

[Clozel et al. 2008] L. Clozel, M. Harris, and R. Taylor, "Automorphy for some l-adic lifts of automorphic mod l Galois representations", *Publ. Math. Inst. Hautes Études Sci.* 108 (2008), 1–181. MR 2010j:11082 Zbl 1169.11020

[Clozel et al. 2011] L. Clozel, M. Harris, and J.-P. Labesse, "Construction of automorphic Galois representations, I", pp. 497–527 in *Stabilization of the trace formula, Shimura varieties, and arithmetic applications, I: On the stabilization of the trace formula*, edited by L. Clozel et al., International Press, 2011. MR 2856383

[Deligne 1979a] P. Deligne, "Valeurs de fonctions L et périodes d'intégrales", pp. 313–346 in *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), vol. 2, edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, R.I., 1979. MR 81d:12009 Zbl 0449.10022

[Deligne 1979b] P. Deligne, "Variétés de Shimura: Interprétation modulaire, et techniques de construction de modèles canoniques", pp. 247–289 in *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), vol. 2, edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math., Amer. Math. Soc., Providence, R.I., 1979. MR 81i:10032 Zbl 0437.14012

[Diamond et al. 2004] F. Diamond, M. Flach, and L. Guo, "The Tamagawa number conjecture of adjoint motives of modular forms", *Ann. Sci. École Norm. Sup.* (4) **37**:5 (2004), 663–727. MR 2006e:11089 Zbl 1121.11045

[Dimitrov 2009] M. Dimitrov, "On Ihara's lemma for Hilbert modular varieties", *Compos. Math.* **145**:5 (2009), 1114–1146. MR 2011d:11117 Zbl 05625824

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics **129**, Springer, New York, 1991. MR 93a:20069 Zbl 0744.22001

[Gan et al. 2012a] W. T. Gan, B. Gross, and D. Prasad, "Symplectic local root numbers, central critical L-values, and restriction problems in the representation theory of classical groups", pp. 1–110 in *Sur les conjectures de Gross et Prasad*, Astérisque **346**, Société Mathématique de France, Paris, 2012.

[Gan et al. 2012b] W. T. Gan, B. Gross, and D. Prasad, "Symplectic local root numbers, central critical L-values and restriction problems in the representation theory of classical groups", pp. 111—170 in *Sur les conjectures de Gross et Prasad*, Astérisque **346**, Société Mathématique de France, Paris, 2012.

[Harris 1997] M. Harris, "L-functions and periods of polarized regular motives", *J. Reine Angew. Math.* **483** (1997), 75–161. MR 98b:11070 Zbl 0859.11032

[Harris 2007] M. Harris, "Cohomological automorphic forms on unitary groups, II: Period relations and values of L-functions", pp. 89–149 in *Harmonic analysis, group representations, automorphic forms and invariant theory*, edited by J.-S. Li et al., Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **12**, World Sci. Publ., Hackensack, NJ, 2007. MR 2009i:11065

[Harris 2008] M. Harris, "A simple proof of rationality of Siegel–Weil Eisenstein series", pp. 149–185 in *Eisenstein series and applications*, edited by W. T. Gan et al., Progr. Math. **258**, Birkhäuser, Boston, MA, 2008. MR 2009g:11061 Zbl 1225.11069

[Harris 2012] M. Harris, "Beilinson–Bernstein localization over $\mathbb{Q}$ and periods of automorphic forms", *Int. Math. Res. Notices* (2012).

[Harris et al. 2011] M. Harris, J.-S. Li, and B. Sun, "Theta correspondences for close unitary groups", pp. 265–307 in *Arithmetic geometry and automorphic forms*, edited by J. Cogdell et al., Adv. Lect. Math. (ALM) **19**, International Press, 2011. MR 2906912

[Hida 1981] H. Hida, "Congruence of cusp forms and special values of their zeta functions", *Invent. Math.* **63**:2 (1981), 225–261. MR 82g:10044 Zbl 0459.10018

[Ichino and Ikeda 2010] A. Ichino and T. Ikeda, "On the periods of automorphic forms on special orthogonal groups and the Gross–Prasad conjecture", *Geom. Funct. Anal.* **19**:5 (2010), 1378–1425. MR 2011a:11100 Zbl 1216.11057

[Labesse 2011] J.-P. Labesse, "Changement de base CM et séries discrètes", pp. 429–470 in *Stabilization of the trace formula, Shimura varieties, and arithmetic applications, I: On the stabilization of the trace formula*, edited by L. Clozel et al., International Press, 2011. MR 2856380

[Mínguez 2011] A. Mínguez, "Unramified representations of unitary groups", pp. 389–410 in *Stabilization of the trace formula, Shimura varieties, and arithmetic applications, I: On the stabilization of the trace formula*, edited by L. Clozel et al., International Press, 2011. MR 2856377

[Mœglin 2007] C. Mœglin, "Classification et changement de base pour les séries discrètes des groupes unitaires p-adiques", *Pacific J. of Mathematics* **233**:1 (2007), 159–204. MR 2009d:22022 Zbl 1157.22010

[Moeglin and Waldspurger 2012] C. Moeglin and J.-L. Waldspurger, "La conjecture locale de Gross–Prasad pour les groupes spéciaux orthogonaux : le cas général", in *Sur les conjectures de Gross et Prasad, II*, Astérisque **347**, Société Mathématique de France, Paris, 2012.

[N. Harris 2011] N. Harris, *A refined Gross–Prasad conjecture for unitary groups*, thesis, University of California, San Diego, 2011.

[Waldspurger 1985] J.-L. Waldspurger, "Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie", *Compositio Math.* **54**:2 (1985), 173–242. MR 87g:11061b Zbl 0567.10021

[White 2010] P.-J. White, *Le produit tensoriel automorphe et l'endoscopie sur le groupe unitaire*, thesis, Université Paris Diderot- Paris 7, 2010.

Communicated by Richard Taylor
Received 2011-07-10 Revised 2011-10-12 Accepted 2012-02-20

harris@math.jussieu.fr *Tour 15-25, 4ème étage, bureau 420, Institut de MathÃ©matiques de Jussieu, 4, place Jussieu, 75252 Paris* CEDEX *05, France*
http://people.math.jussieu.fr/~harris/

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.157

Galois module structure of local unit groups

Romyar Sharifi

We study the groups U_i in the unit filtration of a finite abelian extension K of $\mathbb{Q}_p$ for an odd prime p. We determine explicit generators of the U_i as modules over the $\mathbb{Z}_p$-group ring of $\mathrm{Gal}(K/\mathbb{Q}_p)$. We work in eigenspaces for powers of the Teichmüller character, first at the level of the field of norms for the extension of K by p-power roots of unity and then at the level of K.

1. Introduction

Fix an odd prime p and a finite unramified extension E of $\mathbb{Q}_p$. We use F_n to denote the field obtained from E by adjoining to E the p^nth roots of unity in an algebraic closure of $\mathbb{Q}_p$. The ith unit group in the unit filtration of F_n will be denoted by $U_{n,i}$. The object of this paper is to describe generators of the groups $U_{n,i}$ as modules over the $\mathbb{Z}_p$-group ring of $G_n = \mathrm{Gal}(F_n/\mathbb{Q}_p)$. We express these generators in terms of generators of the pro-p completion D_n of $F_n^{\times}$ as a Galois module. In fact, one consequence of our work is a rather elementary proof of an explicit presentation of D_n as such a module, as was proven by Greither [1996] using Coleman theory.

Instead of working with all of D_n at once, we find it easier to work with certain eigenspaces of it. For this and several other purposes, it will be useful to think of the Galois group G_n as a direct product of cyclic subgroups

$$G_n = \Delta \times \Gamma_n \times \Phi,$$

where $\Delta \times \Gamma_n = \mathrm{Gal}(F_n/E)$ with $|\Delta| = p-1$ and $|\Gamma_n| = p^{n-1}$, and Φ is isomorphic to $\mathrm{Gal}(E/\mathbb{Q}_p)$. We then decompose D_n into a direct sum of $p-1$ eigenspaces for powers of the Teichmüller character $\omega\colon \Delta \to \mathbb{Z}_p^{\times}$. For any integer r, the ω^r-eigenspace $D_n^{(r)}$ of D_n is the subgroup of elements upon which $\sigma \in \Delta$ acts by left multiplication by $\omega(\sigma)^r$. This definition depends only on r modulo $p-1$, so we fix r with $2 \le r \le p$. Note that $D_n^{(r)}$ is a module over the group ring $A_n = \mathbb{Z}_p[\Gamma_n \times \Phi]$.

Supported in part by an NSF Postdoctoral Research Fellowship, an NSERC Discovery Grant, the Canada Research Chairs program and NSF award DMS-0901526.
MSC2010: 11SXX.
Keywords: Galois module structure, unit filtration, local field.

In fact, as we shall see in Section 3.1, the A_n-module $D_n^{(r)}$ has a generating set with just one element if $r \le p-2$, three elements if $r = p-1$, and two elements if $r = p$.

We will be interested in the A_n-module structure of the groups $V_{n,i}^{(r)} = D_n^{(r)} \cap U_{n,i}$. It turns out that

$$V_{n,i}^{(r)} \supsetneq V_{n,i+1}^{(r)} = V_{n,i+2}^{(r)} = \cdots = V_{n,i+p-1}^{(r)}$$

for all $i \equiv r \bmod p-1$ (see Lemma 2.1), so we will consider only such i and set $V_{n,i} = V_{n,i}^{(r)}$.

Our main results, Theorems 4.3.1 and 4.3.3, provide a small set of at most $n+1$ generators of $V_{n,i}$ as an A_n-module and state that any proper generating subset of it has cocardinality 1. The elements of this set are written down explicitly as A_n-linear combinations of elements of the generators of $D_n^{(r)}$. In Section 4.2, elements of a special form are constructed so as to lie as deep in the unit filtration as possible. In Section 4.3, these are refined to elements of the same form that instead lie just deep enough to be in $V_{n,i}$, which are in turn the generators that we use.

It is convenient to work first in the field of norms F of Fontaine–Wintenberger for the tower of extensions F_n of E. This is a field of characteristic p, the multiplicative group of which is the inverse limit of the $F_n^\times$. We prove analogues of all of the above-mentioned results first at this infinite level, prior to applying them in descending to the level of F_n. The fact that the pth power map is an automorphism of $F^\times$ simplifies some of the computations. Moreover, the structure of the eigenspaces of the pro-p completion of $F^\times$, which we study in Section 3.1, is somewhat simpler than that of the $D_n^{(r)}$. We construct special elements in the eigenspaces of the groups in the unit filtration in Section 3.2, refine them in Section 3.3, and prove generation and a minimality result in Section 3.4.

We see a number of interesting potential applications for the results of this paper. To mention just one, it appears to make possible the computation of the conductors of all degree p^n Kummer extensions of F_n in terms of the Kummer generator of the extension. The problem of making this computation, which was approached by the author in three much earlier papers, has until now seemed beyond close reach in this sort of generality.

2. Preliminaries

We maintain the notation of the introduction and introduce some more. Recall from [Wintenberger 1983] that the field of norms F for the extension $F_\infty = \bigcup_n F_n$ of E is a local field of characteristic p with multiplicative group

$$F^\times = \varprojlim F_n^\times,$$

the inverse limit being taken with respect to norm maps.

Let $\zeta = (\zeta_{p^n})_n$ be a norm compatible sequence of p-power roots of unity, with ζ_{p^n} a primitive p^nth root of unity in F_n. Then $\lambda = 1 - \zeta = (1 - \zeta_{p^n})_n$ is a prime element of F.

For $m \geq n$, let $N_{m,n} \colon F_m \to F_n$ be the norm map. Recall that the addition on F is given by

$$(\alpha + \beta)_n = \lim_{m \to \infty} N_{m,n}(\alpha_m + \beta_m)$$

for $\alpha = (\alpha_n)_n$ and $\beta = (\beta_n)_n$ in F. We fix an isomorphism of the residue field of E (and thereby each F_n) with $\mathbb{F}_q$, with q the order of the residue field. Using this, the field $\mathbb{F}_q$ is identified with a subfield of F via the map that takes $\xi \in \mathbb{F}_q^\times$ to $(\tilde{\xi}^{p^{-n}})_n \in F^\times$, where $\tilde{\xi}$ is the $(q-1)$st root of unity in E lifting ξ. The field F may then be identified with the field of Laurent series $\mathbb{F}_q((\lambda))$.

If F_∞ is the union of the F_n, then $G = \mathrm{Gal}(F_\infty/\mathbb{Q}_p)$ acts as automorphisms on the field F. As with G_n, we may decompose $G = \mathrm{Gal}(F_\infty/\mathbb{Q}_p)$ into a direct product of procyclic subgroups

$$G = \Delta \times \Gamma \times \Phi,$$

where $\mathrm{Gal}(F_\infty/E) = \Delta \times \Gamma$, the group Δ has order $p-1$, the group Γ is isomorphic to $\mathbb{Z}_p$, and Φ is isomorphic to $\mathrm{Gal}(E/\mathbb{Q}_p)$. Let γ denote the topological generator of Γ such that $\gamma(\zeta_{p^n}) = \zeta_{p^n}^{1+p}$ for all n.

The pro-p completion D of $F^\times$ decomposes into a direct sum of eigenspaces for the powers of the Teichmüller character ω on Δ. For an integer r, we let $D^{(r)} = D^{\varepsilon_r}$, where ε_r is the idempotent

$$\varepsilon_r = \frac{1}{p-1} \sum_{\delta \in \Delta} \omega(\delta)^{-r} \delta \in \mathbb{Z}_p[\Delta].$$

For $i \geq 1$, let U_i denote the ith group in the unit filtration of F. We then set

$$V_i^{(r)} = U_i \cap D^{(r)} \quad \text{and} \quad (V_i^{(r)})' = V_i^{(r)} - V_{i+1}^{(r)}.$$

The following is [Sharifi 2002, Lemma 2.3] (with F_n replaced by F).

Lemma 2.1. *We have $V_i^{(r)}/V_{i+p-1}^{(r)} \cong \mathbb{F}_q$ for every $i \geq 1$, and $(V_i^{(r)})' \neq \varnothing$ if and only if $i \equiv r \bmod p-1$.*

From now on, we set $V_i = V_i^{(r)}$ and $V_i' = (V_i^{(r)})'$ if $i \equiv r \bmod p-1$. As a consequence of Lemma 2.1, an element $z \in V_i$ is determined modulo λ^{i+p-1} by its expansion

$$z \equiv 1 + \xi\lambda^i \bmod \lambda^{i+1} \tag{2.1}$$

with $\xi \in \mathbb{F}_q$.

The following is [Sharifi 2002, Lemma 2.4] (with F_n replaced by F).

Lemma 2.2. *Let $z \in V_i'$. If $p \nmid i$, then $z^{\gamma-1} \in V_{i+p-1}'$. Otherwise, $z^{\gamma-1} \in V_{i+2(p-1)}$.*

We identify $\Lambda = \mathbb{Z}_p[\![\Gamma]\!]$ with the power series ring $\mathbb{Z}_p[\![T]\!]$ via the continuous, $\mathbb{Z}_p$-linear isomorphism that takes $\gamma - 1$ to T, and we use additive notation to describe the action of $\mathbb{Z}_p[\![T]\!]$ on D. Ramification theory would already have told us that $T \cdot V_i \subseteq V_{i+p-1}$ for all i. On the other hand, explicit calculation will yield the following two lemmas and proposition, which provide more precise information on how powers of T move elements of V_i.

For $\xi \in \mathbb{F}_q^\times$, we let $V_i(\xi)$ denote the set of $z \in V_i$ for which z has an expansion of the form in (2.1). We use $[k]$ to denote the smallest nonnegative integer congruent to $k \in \mathbb{Z}$ modulo p.

Lemma 2.3. *Let $z \in V_i(\xi)$ for some i. Then, for $0 \le j \le [i]$, we have*

$$T^j z \in V_{i+j(p-1)}\Big(\frac{[i]!}{([i]-j)!} \cdot \xi\Big).$$

Proof. Note that

$$\lambda^\gamma = 1 - \zeta^{1+p} = 1 - (1-\lambda)(1-\lambda^p) = \lambda + \lambda^p - \lambda^{p+1}. \tag{2.2}$$

Using this, we see, for any $i \ge 1$, that

$$(1+\xi\lambda^i)^{\gamma-1} \equiv 1 + i\xi\lambda^{i+p-1}\frac{1-\lambda}{1+\xi\lambda^i} \bmod \lambda^{i+2p-2}. \tag{2.3}$$

Hence,

$$(1+\xi\lambda^i)^{\gamma-1} \equiv 1 + i\xi\lambda^{i+p-1} \bmod \lambda^{i+p}. \tag{2.4}$$

Applying (2.4) recursively, we obtain the result. □

Lemma 2.4. *Let $z \in V_{pi-p+1}(\xi)$ for some $i \ge 2$. If j is a nonnegative multiple of $p-1$, then $T^{j+1}z \in V_{p(i+j)}(\xi)$.*

Proof. Let us begin by proving slightly finer versions of (2.3) in two congruence classes of exponents modulo p. For any $t \ge 1$, we have

$$\begin{aligned}
(1+\xi\lambda^{pt})^{\gamma-1} &= \frac{1+\xi\lambda^{pt}(1+\lambda^{p(p-1)}-\lambda^{p^2})^t}{1+\xi\lambda^{pt}} \equiv 1 \bmod \lambda^{p(t+p-1)},\\
(1+\xi\lambda^{pt+1})^{\gamma-1} &= 1+\xi\lambda^{pt+1}\frac{\sum_{m=1}^{pt+1}\binom{pt+1}{m}(\lambda^{p-1}-\lambda^p)^m}{1+\xi\lambda^{pt+1}}\\
&\equiv 1+\xi(\lambda^{p(t+1)}-\lambda^{p(t+1)+1}) \bmod (\lambda^{p(t+p-1)+1}, \lambda^{p(2t+1)+1}),
\end{aligned}$$

the latter congruence following from the fact that $p \mid \binom{pt+1}{m}$ for $2 \le m < p$. Via some obvious inequalities, we conclude that

$$(1+\xi\lambda^{pt})^{\gamma-1} \equiv 1 \bmod \lambda^{p(t+2)}, \tag{2.5}$$

$$(1+\xi\lambda^{pt+1})^{\gamma-1} \equiv (1+\xi\lambda^{p(t+1)})(1-\xi\lambda^{p(t+1)+1}) \bmod \lambda^{p(t+2)}. \tag{2.6}$$

Let $x = 1 + \xi\lambda^{pi-p+1}$. Recursively applying (2.5) and (2.6), we see that

$$x^{(\gamma-1)^{k+1}} \equiv (1 + (-1)^k \xi\lambda^{p(i+k)})(1 + (-1)^{k+1}\xi\lambda^{p(i+k)+1}) \bmod \lambda^{p(i+k+1)},$$

for any positive integer k, as (2.4) implies that $U_{p(i+k)}^{\gamma-1} \subseteq U_{p(i+k+1)}$. The result now follows by application of ε_i, since

$$z^{-1}x^{\varepsilon_i} \in V_{pi}, \quad T^{j+1}x^{\varepsilon_i} \in V_{p(i+j)}(\xi), \quad \text{and} \quad T^{j+1}V_{pi} \subset V_{p(i+j+1)-1},$$

the latter by Lemma 2.2. □

Let us use $\{k\}$ to denote the smallest nonnegative integer congruent to $k \in \mathbb{Z}$ modulo $p-1$. For $i \geq 1$ with $p \nmid i$, we define a monotonically increasing function $\phi^{(i)}\colon \mathbb{Z}_{\geq 0} \to \mathbb{Z}$ by $\phi^{(i)}(0) = i$ and

$$\phi^{(i)}(a) = pa + (i - [i]) + \{[i] - a\} \quad \text{for } a \geq 1. \tag{2.7}$$

Proposition 2.5. *Let $z \in V_i(\xi)$ for some $i \geq 2$ with $p \nmid i$. Then, for $j \geq 1$, we have*

$$T^j z \in V_{\phi^{(i)}(j)}\Big(\frac{[i]!}{\{[i]-j\}!}\xi\Big).$$

Proof. Lemma 2.3 implies that

$$T^{[i]-1}z \in V_{\phi^{(i)}([i]-1)}([i]! \cdot \xi),$$

and note that $\phi^{(i)}([i]-1) \equiv 1 \bmod p$. Set $k = \{[i]-j\}$. Since $j+k-[i]$ is divisible by $p-1$, Lemma 2.4 then implies that

$$T^{j+k}z \in V_{\phi^{(i)}(j+k)}([i]! \cdot \xi). \tag{2.8}$$

It follows from (2.7) that

$$\phi^{(i)}(j+k) - i = p(j+k-[i]) + (p-1)[i],$$

and so, given (2.8), Lemma 2.2 forces $T^l z \in V'_{\phi^{(i)}(l)}$ for all $l \leq j+k$. In particular, applying Lemma 2.3 with j replaced by k and z replaced by $T^j z$, we see that for (2.8) to hold, $T^j z$ must have the stated form. □

Remark 2.6. The obvious analogues of the results of this section all hold at the level of F_n for $n \geq 2$, with λ replaced by $\lambda_n = 1 - \zeta_{p^n}$. In fact, Lemmas 2.1 and 2.2 were originally proven in that setting in [Sharifi 2002]. That the other results hold breaks down to the fact that p is a unit times $\lambda_n^{p^{n-1}(p-1)}$ in F_n, which in particular tells us that (2.2) can be replaced by $\lambda_n^\gamma \equiv \lambda_n + \lambda_n^p - \lambda_n^{p+1} \bmod \lambda_n^{p(p-1)+1}$.

3. The infinite level

3.1. ***Structure of the eigenspaces.*** In this subsection, we fix choices of certain elements that will be used throughout the paper. From now on, we let ξ denote an element of $\mathbb{F}_q$ with $\mathrm{Tr}_\Phi\, \xi = 1$, the conjugates of which form a normal basis of $\mathbb{F}_q$ over $\mathbb{F}_p$. Let $\varphi \in \Phi$ denote the Frobenius element. Let $N_\Phi \in \mathbb{Z}_p[\Phi]$ denote the norm element. Let $\zeta = (\zeta_{p^n})_n$ be a norm-compatible system of primitive p^nth roots of unity as before.

Let r be an integer satisfying $2 \le r \le p$. If $2 \le r \le p-2$, we simply fix an element $u_r \in V_r(\xi)$. In the case that $r = p-1$, generation of $D^{(p-1)}$ requires one additional element $\pi \in D^{(p-1)}$, a non-unit, chosen along with $u_{p-1} \in V_{p-1}(\xi)$ in the lemma which follows. The case of $r = p$ shall require more work, but we will fix elements $w \in V_1(-\xi)$ and $u_p \in V_p(\xi)$ as in Proposition 3.1.3 below.

Lemma 3.1.1. *There exist elements $\pi \in D^{(p-1)}$ and $u_{p-1} \in V_{p-1}(\xi)$ such that $\pi^\varphi = \pi$ and $\pi^{\gamma-1} = u_{p-1}^{N_\Phi}$.*

Proof. Set $\pi = \lambda^{\varepsilon_{p-1}}$, which satisfies $\pi^\varphi = \pi$ and $\pi^{\gamma-1} \in V_{p-1}(1)$. Since every unit is a norm in an unramified extension, there exists $u'_{p-1} \in D^{(p-1)}$ such that $(u'_{p-1})^{N_\Phi} = \pi^{\gamma-1}$, and such an element must lie in $V_{p-1}(\xi')$ for some $\xi' \in \mathbb{F}_q$ with $\mathrm{Tr}_\Phi\, \xi' = 1$. Hilbert's Theorem 90 tells us that $\xi' = \xi + (\varphi - 1)\eta$ for some $\eta \in \mathbb{F}_q$. Let $z \in V_{p-1}(\eta)$, and set $u_{p-1} = u'_{p-1} z^{1-\varphi}$. □

In fact, one could have chosen $u_{p-1} \in V_p(\xi)$ arbitrarily and then taken π to satisfy the relations, as can be seen using the results of the following section.

Lemma 3.1.2. *There exist elements $w \in V_1(-\xi)$ and $u_p \in V_p(\xi)$ with $w^{N_\Phi} = \zeta$ and $u_p^{\varphi-1} = w^{\gamma-1-p}$.*

Proof. First, local class field theory yields the existence of an element $w' \in D^{(p)}$ with $(w')^{N_\Phi} = \zeta$. Since $\zeta \in V_1(-1)$, we must have $w' \in V_1(-\xi')$ for some $\xi' \in \mathbb{F}_q$ with $\mathrm{Tr}_\Phi\, \xi' = 1$. Since $\xi' = \xi + (\varphi - 1)\eta$ for some $\eta \in \mathbb{F}_q$, we choose any $y \in V_1(\eta)$, and then $w = w' y^{1-\varphi} \in V_1(-\xi)$ satisfies $w^{N_\Phi} = \zeta$ as well.

Next, note that $(w^{\gamma-1-p})^{N_\Phi} = 1$, and so Hilbert's Theorem 90 allows us to choose an element $u'_p \in D^{(p)}$ with $(u'_p)^{\varphi-1} = w^{\gamma-1-p}$. A simple computation using (2.4) tells us that $w^{\gamma-1-p} \in V_p(\xi^p - \xi)$, and therefore $u'_p \in V_p(\xi + a)$ for some $a \in \mathbb{F}_p$. We may then choose $z \in V_p(a)$ with $z^\varphi = z$ and take $u_p = u'_p z^{-1} \in V_p(\xi)$. □

We need slightly finer information on the relationship between w and u_p inside the unit filtration, as found in the following proposition.

Proposition 3.1.3. *There exist elements $w \in V_1(-\xi)$ and $u_p \in V_p(\xi)$ with $w^{N_\Phi} = \zeta$ and $u_p^{\varphi-1} = w^{\gamma-1-p}$ such that the element $y = u_p w^{p\varphi^{-1}}$ lies in $V_{2p-1}(-\xi)$.*

Proof. For now, fix any choices of u_p and w as in Lemma 3.1.2. We must have $u_p = (1+\xi\lambda^p)^{\varepsilon_1}\alpha$ with $\alpha \in V_{2p-1}$ and $w = (1-\xi\lambda)^{\varepsilon_1}\beta$ with $\beta \in V_p$. Note that

$$(1+\xi\lambda^p)^{\varphi-1} \equiv 1+(\xi^p-\xi)\lambda^p \mod \lambda^{2p},$$
$$(1-\xi\lambda)^{\gamma-1-p} = \frac{1-\xi(\lambda+\zeta\lambda^p)}{(1-\xi\lambda)(1-\xi^p\lambda^p)} \equiv 1+\Big(\xi^p-\xi\frac{1-\lambda}{1-\xi\lambda}\Big)\lambda^p \mod \lambda^{2p}.$$

We then have

$$\frac{(1-\xi\lambda)^{\gamma-1-p}}{(1+\xi\lambda^p)^{\varphi-1}} \equiv 1+\frac{\xi(1-\xi)}{1-\xi\lambda}\lambda^{p+1} \mod \lambda^{2p}. \tag{3.1.1}$$

We denote the quantity on the right side of (3.1.1) by θ. By Lemma 2.2, we have $\beta^{\gamma-1-p} \in V_{3p-2}$, from which it follows that $\alpha^{\varphi-1}\theta^{-\varepsilon_1} \in V_{3p-2}$. On the other hand, by Lemma 2.1, we have

$$y\alpha^{-1} = (1+\xi\lambda^p)^{\varepsilon_1}(1-\xi\lambda^p)^{\varepsilon_1}\beta^{p\varphi^{-1}} \in V_{3p-2},$$

so in fact we have $y^{\varphi-1}\theta^{-\varepsilon_1} \in V_{3p-2}$. If we can show that $\theta^{\varepsilon_1} \in V_{2p-1}(\xi-\xi^p)$, we will then have $y \in V_{2p-1}(-\xi+a)$ for some $a \in \mathbb{F}_p$. As in the proof of Lemma 3.1.2, we can then choose an element $z \in V_{2p-1}(a)$ with $z^\varphi = z$ and replace u_p by $u_p z^{-1}$ to obtain the result.

By Proposition 2.5, we see that to show that $\theta^{\varepsilon_1} \in V_{2p-1}(\xi-\xi^p)$, it suffices to show that $\theta^{\varepsilon_1(\gamma-1)^{p-1}} \in V_{p^2}(\xi^p-\xi)$. Since $p^2 \equiv 1 \mod p-1$, for this, it suffices to show that

$$\theta^{(\gamma-1)^{p-1}} \equiv 1+(\xi^p-\xi)\lambda^{p^2} \mod \lambda^{p^2+1}.$$

This is a simple consequence of Lemma 3.1.4, which follows. That is, in the notation of said lemma, Fermat's little theorem and the binomial theorem tell us that $d_{p-1,k} = -1$ for all positive integers $k \le p-1$. □

Lemma 3.1.4. *For each positive integer $j \le p-1$, one has*

$$\Big(1+\frac{\xi(1-\xi)}{1-\xi\lambda}\lambda^{p+1}\Big)^{(\gamma-1)^j} \equiv 1+\Big(\sum_{k=1}^{j} d_{j,k}\xi^k(1-\xi)\Big)\lambda^{(j+1)p} \mod \lambda^{(j+1)p+1},$$

where

$$d_{j,k} = \sum_{h=1}^{k}(-1)^{j+h}\binom{k}{h}h^j \in \mathbb{F}_p$$

for positive integers $k \le j$.

Proof. We make the expansion

$$\theta = 1+\frac{\xi(1-\xi)}{1-\xi\lambda}\lambda^{p+1} \equiv \prod_{k=1}^{p-1}(1+\xi^k(1-\xi)\lambda^{p+k}) \mod \lambda^{2p}.$$

Since $U_s^{\gamma-1} \subseteq U_{s+p-1}$ for all s, as follows from (2.4), to compute $\theta^{(\gamma-1)^j}$ modulo $\lambda^{(j+1)p+1}$, it suffices to compute $(1+\xi^k(1-\xi)\lambda^{p+k})^{(\gamma-1)^j}$ modulo $\lambda^{(j+1)p+1}$.

Fix a positive integer $k \le p-1$. We claim that the coefficient of $\lambda^{(j+1)p}$ in the expansion of $(1+\xi^k(1-\xi)\lambda^{p+k})^{(\gamma-1)^j}$ as a power series in $\mathbb{F}_q[\![\lambda]\!]$ is 0 if $j<k$ and $\xi^k(1-\xi)d_{j,k}$ if $j \ge k$. As a consequence of (2.3), one sees that

$$(1+\xi\lambda^t)^{\gamma-1} \equiv (1+t\xi\lambda^{t+p-1})(1-t\xi\lambda^{t+p}) \bmod \lambda^{t+2p-2} \quad \text{for any } t \ge p-1.$$

Using this and the finer congruence (2.6) when possible, an induction yields that the expansion in question is determined by

$$\prod_{m=0}^{\min(j,k)} \prod_{(a_i)\in P_{j,k,m}} \left(1+\xi^k(1-\xi)\frac{k!}{(k-m)!}\prod_{i=1}^{j-m} a_i \cdot \lambda^{(j+1)p+k-m}\right)^{(-1)^{j-m}} \bmod \lambda^{(j+1)p+k+1},$$

where

$$P_{j,k,m} = \{(a_1,a_2,\dots,a_{j-m}) \in \mathbb{Z}^{j-m} \mid k-m \le a_1 \le a_2 \le \dots \le a_{j-m} \le k\}$$

if $j>m$ and $P_{j,k,j}=\{0\}$, and we consider the empty product to be 1. In particular, the coefficient in question is indeed 0 for $j<k$ and is $\xi^k(1-\xi)c_{j,k}$ for $j \ge k$, where

$$c_{j,k} = (-1)^{j-k}k! \sum_{(a_i)\in P_{j,k,k}} \prod_{i=1}^{j-k} a_i.$$

It remains to verify that $c_{j,k}=d_{j,k}$ for $j \ge k$.

Let D denote the differential operator $x\frac{d}{dx}$ on $\mathbb{F}_p[x]$. By the binomial theorem, we have

$$D^j((1-x)^k)|_{x=1} = \sum_{h=1}^{k}(-1)^h\binom{k}{h}h^j x^h\Big|_{x=1} = (-1)^j d_{j,k}.$$

On the other hand, repeated application of the product formula for the derivative yields

$$\begin{aligned} D^j((1-x)^k)|_{x=1} &= (-1)^k \sum_{h=1}^{\min(j,k)} \frac{k!}{(k-h)!} \sum_{(a_i)\in P_{j,h,h}} \prod_{i=1}^{j-h} a_i \cdot (x-1)^{k-h}x^h\Big|_{x=1} \\ &= (-1)^j c_{j,k} \end{aligned}$$

for all $j \ge k$ and hence the result. □

In the next section, we will obtain the following very slight refinement of what is essentially a result of [Greither 1996, Sections 2 and 3]; see also [Sharifi 2002, Corollary 2.2].

Theorem 3.1.5. *For $r \leq p-2$, the A-module $D^{(r)}$ is freely generated by any $u_r \in V_r(\xi)$. The A-module $D^{(p-1)}$ has a presentation*

$$D^{(p-1)} = \langle \pi, u_{p-1} \mid \pi^{\varphi} = \pi,\ u_{p-1}^{N_\Phi} = \pi^{\gamma-1} \rangle,$$

for some $u_{p-1} \in V_{p-1}(\xi)$ and $\pi \in D^{(p-1)}$. The A-module $D^{(p)}$ has a presentation

$$D^{(p)} = \langle u_p, w \mid w^{\gamma-1-p} = u_p^{\varphi-1} \rangle,$$

for some $u_p \in V_p(\xi)$ and $w \in V_1(-\xi)$ such that $w^{N_\Phi} = \zeta$.

3.2. ***Special elements.*** Fix r such that $2 \leq r \leq p$, and define $\phi \colon \mathbb{Z}_{\geq 0} \to \mathbb{Z}$ by $\phi(a) = \phi^{(r)}(a)$ for $a \geq 1$. Set

$$\delta = \begin{cases} 0 & \text{if } 2 \leq r \leq p-1, \\ 1 & \text{if } r = p. \end{cases}$$

For all $a \geq 1$, we have $\phi(a) = p(a+\delta) + \{r - \delta - a\}$, so $\phi(a)$ is the smallest integer that is at least $p(a+\delta)$ and congruent to r modulo $p-1$.

From now on, i will be used solely to denote a positive integer congruent to r modulo $p-1$. We will write $\alpha \sim \beta$ to denote that both α and β lie in $V_i(\xi)$ for some i and $\xi \in \mathbb{F}_q^\times$. We use additive notation for the action of $A = \mathbb{Z}_p[\Phi][\![T]\!]$ on $D^{(r)}$. We begin with the following useful lemma.

Lemma 3.2.1. *Let j be a positive integer.*

a. *We have*

$$T^j u_r \in V_{\phi(j)}\Big(\frac{[r]!}{\{r-\delta-j\}!}\xi\Big).$$

b. *If $j \equiv r - \delta \bmod (p-1)$ so that $T^j u_r \sim pz$ for some $z \in D^{(r)}$, then*

$$T^j u_r - pz \in V_{\phi(j)+p-1}\left(-[r]!\,\xi\right).$$

Proof. For $r < p$, part a is a direct consequence of Proposition 2.5 and the fact that $u_r \in V_r(\xi)$. For $r = p$, Proposition 2.5 and the fact that $\phi = \phi^{(2p-1)}$ on positive integers would tell us more directly that $T^j y \in V_{\phi(j)}(\frac{1}{\{-j\}!}\,\xi)$ for $j \geq 1$ and y as in Proposition 3.1.3. Note, however, that $Tu_p = Ty - p\varphi^{-1}Tw \sim Ty$, since $pTw \in V_{p^2}$. This is also the key point of part b. That is, we have

$$T(T^j u_r - pz) \sim T^{j+1} u_r$$

as $pTz \in V_{p\phi(j)}$ and

$$\phi(j+1) \leq \phi(j) + 2(p-1) < p\phi(j).$$

Since $T^{j+1}u_r \in V_{\phi(j)+2(p-1)}([r]!\,\xi)$, a final application of Proposition 2.5 tells us that $T^j u_r - pz$ had to be in the stated group. □

For a nonnegative integer m, let us define $\phi_m : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by $\phi_m = p^m(\phi+1)-1$. We remark that

$$p\phi_m = \phi \circ (\phi_m - \delta). \tag{3.2.1}$$

From now on, we set $\rho = p\varphi^{-1}$ for brevity of notation. We define special elements in the unit filtration of $D^{(r)}$.

Theorem 3.2.2. *Let m and j be nonnegative integers. Define*

$$\alpha_{m,j} = \frac{1}{[r]!}\left(\{r-\delta-j\}!\rho^m T^j - \sum_{k=1}^{m} \rho^{m-k} T^{\phi_{k-1}(j)-\delta}\right)u_r,$$

unless $j = 0$ and $r = p-1$, in which case we replace $\{r-\delta-j\}!$ with -1 in the formula. Then $\alpha_{m,j} \in V_{\phi_m(j)}(\xi)$. Furthermore, $(p^m bT^j + c)u_r \notin V_{\phi_m(j)+p-1}$ for all $b \in \mathbb{Z}_p[\Phi]$ with $b \not\equiv 0 \bmod p$ and $c \in T^{j+1}A$.

Proof. We work by induction, the case of $m = 0$ being Lemma 3.2.1a, aside from the case $j = 0$, in which case it is simply the definition of u_r. Assume we have proven the first statement for m. Then $p\alpha_{m,j} \in V_{p\phi_m(j)}(\xi^p)$ and, using Lemma 3.2.1a and (3.2.1), we have

$$T^{\phi_m(j)-\delta}u_r \in V_{p\phi_m(j)}\big([r]!\,\xi\big).$$

Lemma 3.2.1b then tells us that

$$\alpha_{m+1,j} = \rho\alpha_{m,j} - \frac{1}{[r]!}T^{\phi_m(j)-\delta}u_r \in V_{p\phi_m(j)+p-1}(\xi).$$

Now assume the second statement is true for m. (For $m = 0$, this is a consequence of the fact that the conjugates of ξ are $\mathbb{F}_p$-linearly independent.) Suppose that

$$\alpha = (p^{m+1}bT^j + c)u_r \in V'_i$$

with $i \geq \phi_{m+1}(j)$, $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$ and $c \in T^{j+1}A$. We write $c = (pc' + T^h\nu)u_r$ for some $c', \nu \in A$ with $\nu \not\equiv 0 \bmod (p, T)$ and $h \geq j+1$. By induction, we have

$$(p^m bT^j + c')u_r \notin V_{\phi_m(j)+p-1}.$$

Since $\phi_{m+1}(j) = p\phi_m(j) + p - 1$ and $\alpha \in V_{\phi_{m+1}(j)}$ by assumption, this forces

$$p(p^m bT^j + c')u_r \sim -T^h\nu u_r,$$

which tells us by Lemma 3.2.1a that $\phi(h) \leq p\phi_m(j)$. On the other hand, it follows from Lemma 3.2.1b that $\alpha \in V'_{\phi(h)+p-1}$, which forces $i = \phi_{m+1}(j)$. □

The second statement of Theorem 3.2.2 insures, in particular, that $au_r \neq 0$ for all nonzero $a \in A$. We therefore have the following corollary.

Corollary 3.2.3. *The A-submodule of $D^{(r)}$ generated by u_r is free.*

In the exceptional case that $r = p$, we require additional elements. First, we modify the function ϕ_m for this r. For nonnegative integers m and j, we set $\phi'_m(j) = \phi_m(j)$ unless $r = p$ and $j = p^l - 1$ for some $l \geq 0$, in which case we set

$$\phi'_m(p^l - 1) = p^{m+l+1} + p^{m+1} - 1 = \phi_m(p^l - 1) + p^m(p-1).$$

Theorem 3.2.4. *Let l and m be nonnegative integers. Define*

$$\beta_{m,l} = \left(\rho^m T^{p^l-1} + \sum_{k=1}^{m} \rho^{m-k} T^{\phi'_{k-1}(p^l-1)-1}\right) u_p + \rho^{m+l+1} w.$$

Then $\beta_{m,l} \in V_{\phi'_m(p^l-1)}(-\xi)$. Moreover, for any $j \geq 0$, we have

$$(p^m b T^j + c) u_p + d w \notin V_{\phi'_m(j)+p-1}$$

for all $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1}A$ and $d \in \mathbb{Z}_p[\Phi]$.

Proof. The proof is similar to that of Theorem 3.2.2. Since Lemma 3.2.1a and the definition of w tell us that

$$T^{p^l-1} u_p \in V_{p^{l+1}}(\xi) \quad \text{and} \quad \rho^{l+1} w \in V_{p^{l+1}}(-\xi),$$

Lemma 3.2.1b yields $\beta_{0,l} = \rho^{l+1} w + T^{p^l-1} u_p \in V_{p^{l+1}+p-1}(-\xi)$. For any $m \geq 0$, we have

$$\beta_{m+1,l} = \rho \beta_{m,l} + T^{\phi'_m(p^l-1)-1} u_p.$$

By induction and Lemma 3.2.1a, we have

$$\rho \beta_{m,l} \in V_{p\phi'_m(p^l-1)}(-\xi) \quad \text{and} \quad T^{\phi'_m(p^l-1)-1} u_p \in V_{p\phi'_m(p^l-1)}(\xi).$$

Since $\phi'_{m+1}(p^l - 1) = p\phi'_m(p^l - 1) + p - 1$, that $\beta_{m+1,l} \in V_{\phi'_{m+1}(p^l-1)}(-\xi)$ is just another application of Lemma 3.2.1b.

Let $j \geq 0$, $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1}A$, and $d \in \mathbb{Z}_p[\Phi]$. First, suppose that $\alpha = (bT^j + c)u_p + dw \in V'_i$ for some $i \geq \phi'_0(j)$. Note that $(bT^j + c)u_p \in V'_{\phi(j)}$ by Lemma 3.2.1a, while $dw \in V'_{p^l}$ for some $l \geq 0$. Since $\alpha \in V_{\phi'_0(j)}$, we must have $p^l \geq \phi(j)$. We then have $\alpha \in V'_{\phi(j)}$ unless $\phi(j) = p^l$. This occurs if and only if $l \geq 1$ and $j = p^{l-1} - 1$, in which case $\phi'_0(j) = p^l + p - 1$. For this to hold, we must have $(bT^j + c)u_p \sim -dw$. Lemma 3.2.1b then implies that $\alpha \in V'_{p^l+p-1}$, so $i = \phi'_0(j)$ in all cases.

Suppose now that $\alpha = (p^{m+1} b T^j + c)u_p + dw \in V'_i$ for some $i \geq \phi'_{m+1}(j)$. Rewrite c as $pc' + T^h \nu$ for some $h \geq j + 1$ and $c', \nu \in A$ with $\nu \not\equiv 0 \bmod (p, T)$. If we are to have $\alpha \in V_p$, we may also write $d = pd'$ for some $d' \in \mathbb{Z}_p[\Phi]$. By induction, we have

$$(p^m b T^j + c')u_p + d' w \notin V_{\phi'_m(j)+p-1},$$

and so in order that $\alpha \in V_{\phi'_{m+1}(j)}$, we must have

$$(p^{m+1}bT^j + pc')u_p + dw \sim -T^h \nu u_p,$$

which tells us using Lemma 3.2.1a that $\phi(h) \le p\phi'_m(j)$. On the other hand, Lemma 3.2.1b tells us that $\alpha \in V'_{\phi(h)+p-1}$, so we must have $i = \phi'_{m+1}(j)$. □

Theorem 3.1.5 may now be proven as a consequence of the description of the elements above and their place in the unit filtration.

Proof of Theorem 3.1.5. For $r \le p-1$, the union of the disjoint images of the functions ϕ_m is exactly the set of positive integers congruent to r modulo $p-1$. Therefore, Theorem 3.2.2 implies that there exists an element of the A-module generated by u_r in $V_i(\xi)$ for each $i \equiv r \mod p-1$. In particular, u_r therefore clearly generates V_r as an A-module, which equals $D^{(r)}$ for $r \le p-2$, and it is free by Corollary 3.2.3. Every element of $D^{(p-1)}$ may then be written in the form $\pi^m u_{p-1}^a$ with $m \in \mathbb{Z}_p$ and $a \in A$, and such an element can clearly only be trivial if m is, and therefore a is as well. Noting that our choices of π and u_{p-1} as in Lemma 3.1.1 satisfy the desired relations, the presentation for $r = p-1$ is as stated.

For $r = p$, the union of $\{1\}$ and the images of the functions ϕ_m and ϕ'_m is the set of positive integers that are congruent to 1 modulo $p-1$. Theorem 3.2.2 and Theorem 3.2.4 imply that there exists an element of the A-module generated by u_p and w in $V_i(\xi)$ for each $i \equiv 1 \mod p-1$. Thus, this A-module is $D^{(p)}$. Our choices of u_p and w satisfy the relations of Lemma 3.1.2, and it follows from the second statement of Theorem 3.2.4 that if either $c \in A$ or $d \in \mathbb{Z}_p[\Phi]$ is nonzero, then so is $cu_p + dw$. □

3.3. ***Refined elements.*** In this section, we provide refinements of the elements constructed in Theorem 3.2.2 and Theorem 3.2.4. We maintain the notation of Section 3.2. We begin by constructing certain one-sided inverses to the monotonically increasing functions ϕ and ϕ_m.

For any nonnegative integer a and positive integer t, let us set

$$\langle a \rangle_t = \max(a + \{t-a\}, t).$$

Therefore, $\langle a \rangle_t$ is the smallest integer greater than or equal to t and a and congruent to t modulo $p-1$. Define $\psi \colon \mathbb{Z}_{\ge 0} \to \mathbb{Z}_{\ge 0}$ by

$$\psi(a) = \left\lfloor \frac{\langle a \rangle_r + 1}{p} \right\rfloor - \delta$$

except for $r = p-1$ and $a \le p-1$, in which case we set $\psi(a) = 0$. For $m \ge 0$, define $\psi_m \colon \mathbb{Z}_{\ge 0} \to \mathbb{Z}_{\ge 0}$ by

$$\psi_m(a) = \psi\left(\left\lceil \frac{a+1}{p^m} \right\rceil - 1\right).$$

Note that $\psi_0 = \psi$.

Lemma 3.3.1. *We have $\psi_m(\phi_m(j)) = j$ for all nonnegative integers j. Moreover, for all such j and positive integers a, we have $\phi_m(j) \geq a$ if and only if $j \geq \psi_m(a)$.*

Proof. First, note that $\phi(j)$ is congruent to r modulo $p-1$, so we have

$$\psi(\phi(j)) = \left\lfloor \frac{\phi(j)+1}{p} \right\rfloor - \delta = \left\lfloor \frac{p(j+\delta)+\{r-\delta-j\}+1}{p} \right\rfloor - \delta = j,$$

unless $r \geq p-1$ and $j = 0$, but one checks immediately that $\psi(\phi(0)) = \psi(r) = 0$ if $r \geq p-1$ as well. It follows that we have

$$\psi_m(\phi_m(j)) = \psi\left(\left\lceil \frac{p^m(\phi(j)+1)}{p^m} \right\rceil - 1\right) = \psi(\phi(j)) = j.$$

Therefore, if $\phi_m(j) \geq a$, then $j = \psi_m(\phi_m(j)) \geq \psi_m(a)$, since ψ_m is nondecreasing.

To finish the proof, we need only show that $\phi_m(\psi_m(a)) \geq a$, since ϕ_m is nondecreasing (in fact, strictly increasing). First, note that the definition of ψ is such that $\psi(a) = \psi(\langle a \rangle_r)$. For $i \equiv r \bmod p-1$ with $i \neq 1, p-1$, the value $\phi(\psi(i))$ is the unique integer between $p\lfloor (i+1)/p \rfloor$ and $p\lfloor (i+1)/p \rfloor + p - 2$ that is congruent to $r \bmod p-1$. This implies that

$$\phi(\psi(a)) = \begin{cases} \langle a \rangle_r & \text{if } \langle a \rangle_r \not\equiv -1 \bmod p, \text{ or } a \leq r = p-1, \\ \langle a \rangle_r + p - 1 & \text{otherwise,} \end{cases} \tag{3.3.1}$$

which is, in particular, at least a. By definition of ϕ_m and ψ_m, we then have

$$\phi_m(\psi_m(a)) = p^m(\phi(\psi_m(a)) + 1) - 1 \geq p^m \left\lceil \frac{a+1}{p^m} \right\rceil - 1 \geq a. \qquad \square$$

We actually need a version of Lemma 3.3.1 with ϕ_m replaced by ϕ'_m and ψ_m replaced by an appropriate function $\psi'_m \colon \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$, which we now define. Set $\psi'_m = \psi_m$ if $r \leq p-1$ and, if $r = p$, let

$$\psi'_m(a) = \begin{cases} \psi_m(a) - 1 & \text{if } p^{m+l+1} + p^m \leq a \leq p^{m+l+1} + p^{m+1} - 1 \text{ for some } l \geq 0, \\ \psi_m(a) & \text{otherwise.} \end{cases}$$

Note that $\psi'_m(a) = \psi_m(a) - 1$ for $r = p$ if and only if $\phi_m(p^l - 1) < a \leq \phi'_m(p^l - 1)$ for some $l \geq 0$, in which case $\psi'_m(a) = p^l - 1$. One then easily checks the following:

Corollary 3.3.2. *We have $\psi'_m(\phi'_m(j)) = j$ for all nonnegative integers j. Moreover, for all such j and positive integers a, we have $\phi'_m(j) \geq a$ if and only if $j \geq \psi'_m(a)$.*

For the rest of this section, we fix a positive integer i with $i \equiv r \bmod p-1$.

Remark 3.3.3. Lemma 3.3.1 and Theorem 3.2.2 tell us that each $\alpha_{m,\psi_m(i)}$ lies in V_i. Corollary 3.3.2 and Theorem 3.2.4 tell us that each $\beta_{m,l}$ with $\psi'_m(i) = p^l - 1$ lies in V_i. These elements have the form $(p^m bT^j + c)u_r + dw$ for $j = \psi'_m(i)$, where

$b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1}A$, and $d \in \mathbb{Z}[\Phi]$, with $d = 0$ if $r \neq p$. The same results also show that no such element with $j < \psi'_m(i)$ can lie in V_i.

For any $m \geq 0$, define $\theta_m \colon \mathbb{Z}_{\geq 1} \to \mathbb{Z}_{\geq 0}$ by

$$\theta_m(a) = \psi\Big(\Big\lceil \frac{\langle a \rangle_r}{p^m} \Big\rceil\Big).$$

By Lemma 3.2.1a and Lemma 3.3.1, the value $\theta_m(i)$ for $i \equiv r \bmod p-1$ is the minimal integer j such that $p^m T^j u_r \in V_i$. In particular, $\theta_m(i) \geq \psi_m(i)$ for all i.

Lemma 3.3.4. *For all positive integers m and k with $k \leq m$, we have*

$$\phi'_{k-1}(\psi'_m(i)) - \delta \geq \theta_{m-k}(i) - 1,$$

with equality if and only if

$$p^{m-k+1}\phi'_{k-1}(\psi'_m(i)) < i. \tag{3.3.2}$$

Moreover, we have $\psi'_m(i) \geq \theta_m(i) - 1$, with equality if and only if the equivalent conditions above hold for $k = 1$.

Proof. Let us check the case that $r = p$ and $\psi'_m(i) = p^l - 1$ for some $l \geq 0$ separately. First, suppose that $p^{m+l+1} < i < p^{m+l+1} + p^{m+1}$. In this case, we have $\psi'_m(i) = \theta_m(i) - 1$. We also have

$$\phi'_{k-1}(\psi'_m(i)) = \phi'_{k-1}(p^l - 1) = p^{k+l} + p^k - 1 = \psi(p^{k+l+1} + p^{k+1}) \geq \theta_{m-k}(i),$$

with equality if and only if

$$p^{m-k+1}\phi'_{k-1}(\psi'_m(i)) = p^{m+l+1} + p^{m+1} - p^{m-k+1} < i. \tag{3.3.3}$$

Moreover, in the case that $p^{m+l+1} - p^{m+1} + 2p^m < i \leq p^{m+l+1}$, the values $\phi'_{k-1}(\psi'_m(i))$ and $p^{m-k+1}\phi'_{k-1}(\psi'_m(i))$ are the same as in the previous case, while $\theta_{m-k}(i) - 1$ and i are smaller. So, we may assume from this point forward that r and i are such that $\psi'_m(i) = \psi_m(i)$ and $\phi'_{k-1}(\psi'_m(i)) = \phi_{k-1}(\psi_m(i))$ for all k.

We claim that $\rho^{m-k}T^{\phi_{k-1}(\psi_m(i))+1-\delta}u_r$ lies in V_{i+p-1} for all positive $k \leq m$ and that $\rho^{m-k}T\alpha_{k,\psi_m(i)}$ lies in V_{i+p-1} for all nonnegative $k \leq m$. Note that $T\alpha_{m,\psi_m(i)}$ lies in V_{i+p-1} as a consequence of Theorem 3.2.2. Suppose that $\rho^{m-k}T\alpha_{k,\psi_m(i)} \in V_{i+p-1}$ for some positive $k \leq m$. We then have

$$\rho^{m-k}T^{\phi_{k-1}(\psi_m(i))+1-\delta}u_r \sim -[r]!\rho^{m-k}T\alpha_{k,\psi_m(i)} \in V_{i+p-1},$$

which also forces $\rho^{m-k+1}T\alpha_{k-1,\psi_m(i)} \in V_{i+p-1}$, since

$$\rho^{m-k+1}T\alpha_{k-1,\psi_m(i)} = \rho^{m-k}T\alpha_{k,\psi_m(i)} + \frac{1}{[r]!}\rho^{m-k}T^{\phi_{k-1}(\psi_m(i))+1-\delta}u_r,$$

proving the claim. In particular, since $\rho^m T\alpha_{0,\psi_m(i)} \in V_{i+p-1}$, we have

$$\rho^m T^{\psi_m(i)+1} u_r \in V_{i+p-1}$$

as well. The definition of $\theta_{m-k}(i)$ now yields the desired inequalities.

Now (3.3.2) holds for a given k if and only if $\rho^{m-k+1}\alpha_{k-1,\psi_m(i)} \notin V_i$. Since

$$[r]!\,\rho\alpha_{k-1,\psi_m(i)} \sim T^{\phi_{k-1}(\psi_m(i))-\delta} u_r,$$

this occurs if and only if $\rho^{m-k}T^{\phi_{k-1}(\psi_m(i))-\delta}u_r \notin V_i$ and, therefore, if and only if

$$\phi_{k-1}(\psi_m(i)) - \delta \le \theta_{m-k}(i) - 1,$$

which must then be an equality. Also, $\psi_m(i) < \theta_m(i)$ if and only if $\rho^m\alpha_{0,\psi_m(i)} \notin V_i$, which holds by Lemma 3.2.1a if and only if $p^m\phi(\psi_m(i)) < i$, the same condition as (3.3.2) for $k = 1$. □

From now on, we set $i_m = \lceil \frac{i}{p^m} \rceil$ for all $m \ge 0$.

Lemma 3.3.5. *For any pair of positive integers m and k with $k \le m$, we have $\phi'_{k-1}(\psi'_m(i)) - \delta \ge \theta_{m-k}(i) - 1$, with equality if and only if*

(1) $i_{m+\epsilon} \not\equiv 0 \bmod p$, *or* $r = p-1$ *and* $i_m = p$,

(2) $i_{m+\epsilon} \equiv r+1 \bmod p-1$, *but not* $r = p-1$ *and* $i_m = 1$, *and*

(3) $i \equiv -j \bmod p^{m+\epsilon}$ *for some* $0 < j < p^{m+1-k}$,

where $\epsilon = 0$ unless $r = p$ and $i_{m+1} = p^l + 1$ for some $l \ge 0$, in which case we set $\epsilon = 1$. Moreover, we have $\psi'_m(i) \ge \theta_m(i) - 1$, with equality if and only if the above conditions hold with $k = 1$.

Proof. The case that $r = p$ and $\psi'_m(i) = p^l - 1$ for some $l \ge 0$ follows from the proof of Lemma 3.3.4, noting that if $i_{m+1} = p^l + 1$, then it is both nonzero modulo p and congruent to $p+1$ modulo $p-1$, and the third condition of the lemma holds exactly when (3.3.3) does. On the other hand, for the remaining i with $\psi'_m(i) = p^l - 1$, we have $i_{m+1} = p^l$, and the fact that the inequality is strict was shown in the proof of Lemma 3.3.4. So, we again assume that $r \ne p$ or i is such that $\psi'_m(i) \ne p^l - 1$ for all $l \ge 0$.

By Lemma 3.3.4, it suffices to determine the precise conditions under which (3.3.2) holds. Let us set $a = (i+1)_m$. It follows from (3.3.1) that we have

$$p^{m-k+1}\phi_{k-1}(\psi_m(i)) = \begin{cases} p^m\langle a\rangle_{r+1} - p^{m-k+1} & \text{if } p \nmid \langle a\rangle_{r+1}, \\ p^m\langle a\rangle_{r+1} + p^m(p-1) - p^{m-k+1} & \text{otherwise,} \end{cases} \tag{3.3.4}$$

unless $r = p-1$ and $\langle a\rangle_{r+1} = p$, in which case

$$p^{m-k+1}\phi_{k-1}(\psi_m(i)) = p^{m+1} - p^{m-k+1}.$$

Aside from this exceptional case, (3.3.4) implies that p cannot divide $\langle a\rangle_{r+1}$ if (3.3.2) is to hold. Moreover, if $\langle a\rangle_{r+1} > a$, then again (3.3.2) cannot hold, so for it to hold, we must have $a \equiv r+1 \mod p-1$, but not $r = p-1$ and $a = 1$. Assuming that these necessary conditions hold, the condition that

$$p^{m-k+1}\phi_{k-1}(\psi_m(i)) = p^m a - p^{m-k+1} < i$$

is exactly that $i \equiv -j \mod p^m$ with $0 < j < p^{m-k+1}$. □

For $m \geq 0$, we will define new elements $\kappa_{m,i}$ of V_i that involve fewer terms and easier-to-compute exponents of powers of T than the expressions for $\alpha_{m,\psi_m(i)}$ and $\beta_{m,l}$. In preparation, set $\sigma(m,i) = \lfloor \log_p(p^m i_m - i)\rfloor$ for any $m \geq 0$ such that $p^m \nmid i$. Note that $0 \leq \sigma(m,i) \leq m-1$ when it is defined and $\sigma(m+1,i)$ is defined and greater than or equal to $\sigma(m,i)$ whenever $\sigma(m,i)$ is defined.

First, supposing either that $r \leq p-1$ or that $r = p$ and $i_{m+1}-1$ is not a power of p, we set

$$\kappa_{m,i} = \rho^m T^{\theta_m(i)} u_r \tag{3.3.5}$$

if $i_m \not\equiv r+1 \mod p-1$, $p \mid i_m$, $i < p^m$, or $p^m \mid i$, unless $r = p-1$ and $i_m = p$, and

$$\kappa_{m,i} = \left(\rho^m T^{\theta_m(i)-1} - a_{m,i}\sum_{k=\sigma(m,i)}^{m-1} \rho^k T^{\theta_k(i)-1}\right) u_r \tag{3.3.6}$$

otherwise, where $a_{m,i}$ denotes the least positive residue of $(\{r+1-\delta-\theta_m(i)\}!)^{-1}$ modulo p unless $r = p-1$ and $\theta_m(i) = 1$, in which case we take $a_{m,i} = -1$. In the remaining case that $r = p$ and $i_{m+1}-1$ is a power of p, we set

$$\kappa_{m,i} = \left(\rho^m T^{\theta_m(i)-1} + \sum_{k=\sigma(m+1,i)}^{m-1} \rho^k T^{\theta_k(i)-1}\right) u_r + \rho^{m+\log_p(i_{m+1}-1)+1} w. \tag{3.3.7}$$

For consistency, we let $a_{m,i} = -1$ for such m. Note that Lemma 3.3.5 tells us that each $\kappa_{m,i}$ has the form $(\rho^m T^{\psi'_m(i)} + c)u_r + dw$ for some $c \in T^{\psi'_m(i)+1}A$ and $d \in \mathbb{Z}_p[\Phi]$, with d taken to be zero if $r \leq p-1$.

We give two examples for $p = 5$ and particular values of i.

Example 3.3.6. Suppose that $p = 5$, $r = 3$, and $i = 11899$. Then we have

$$\kappa_{0,i} = T^{2380}u_3, \qquad \kappa_{1,i} = \rho T^{476}u_3,$$
$$\kappa_{2,i} = (\rho^2 T^{95} - \rho T^{475} - T^{2379})u_3, \qquad \kappa_{3,i} = (\rho^3 T^{19} - \rho^2 T^{95})u_3,$$
$$\kappa_{4,i} = \rho^4 T^4 u_3, \qquad \kappa_{5,i} = (\rho^5 - \rho^4 T^3 - \rho^3 T^{19})u_3.$$

Example 3.3.7. Suppose that $p = 5$, $r = 5$, and $i = 92729$. Then we have

$$\kappa_{0,i} = T^{18545}u_5, \qquad \kappa_{1,i} = (\rho T^{3708} - T^{18544})u_5,$$
$$\kappa_{2,i} = \rho^2 T^{741}u_5, \qquad \kappa_{3,i} = (\rho^3 T^{147} - \rho^2 T^{740} - \rho T^{3708})u_5,$$
$$\kappa_{4,i} = \rho^4 T^{29}u_5, \qquad \kappa_{5,i} = (\rho^5 T^4 + \rho^4 T^{28})u_5 + \rho^7 w,$$
$$\kappa_{6,i} = \rho^6 u_5 + \rho^7 w.$$

Remark 3.3.8. It is not hard to see from the definition of $\sigma(m,i)$ that $\sigma(m,i) \geq k$ for $k < m$ if and only if $p^{m-k} \nmid i_k$. Moreover, if for a given k there exists $m > k$ such that $\sigma(m,i)$ is less than k or not defined, then $p \mid i_k$ so $\kappa_{k,i} = \rho^k T^{\theta_k(i)} u_r$ unless $r = p$ and $i_{k+1} - 1$ is a power of p or $r = p-1$ and $i_k = p$. The previous examples illustrate some of this.

Let us show that the $\kappa_{m,i}$ are actually elements of V_i. In the process, we see how they compare to the elements $\alpha_{m,\psi_m(i)}$ and $\beta_{m,l}$ previously defined.

Proposition 3.3.9. *The elements $\kappa_{m,i}$ lie in V_i for all nonnegative integers m.*

Proof. Suppose first that $r \neq p$ or i does not satisfy $i_m = p^l + 1$ for any $l \geq 0$ (and omitting the case $r = p-1$ and $\psi_m(i) = 0$, for which one should take the fractions in the following two equations to be 1). If $\psi_m(i) = \theta_m(i)$, then we have

$$\kappa_{m,i} = \frac{[r]!}{\{r - \delta - \psi_m(i)\}!}\, \rho^m \alpha_{0,\psi_m(i)},$$

and this lies in V_i by the definition of $\theta_m(i)$. If $\psi_m(i) = \theta_m(i) - 1$, we claim that

$$\kappa_{m,i} \sim \frac{[r]!}{\{r - \delta - \psi_m(i)\}!}\, \rho^{\sigma(m,i)} \alpha_{m-\sigma(m,i),\psi_m(i)}. \tag{3.3.8}$$

To see this, note that

$$\kappa_{m,i} = \rho^{\sigma(m,i)}\left(\rho^{m-\sigma(m,i)} T^{\psi_m(i)} - a_{m,i} \sum_{k=1}^{m-\sigma(m,i)} \rho^{m-\sigma(m,i)-k} T^{\theta_{m-k}(i)-1}\right) u_r.$$

It follows from Lemma 3.3.5 that $\theta_{m-k}(i) - 1 = \phi_{k-1}(\psi_m(i)) - \delta$ if and only if $p^{m-k+1} > p^m i_m - i$, and therefore if $k \leq m - \sigma(m,i)$, proving the claim. (Note that we the reason we do not have actual equality in (3.3.8) is simply that we took $a_{m,i}$ to be an inverse to $\{r - \delta - \psi_m(i)\}!$ modulo p, not in $\mathbb{Z}_p^\times$.) Moreover, we have by Theorem 3.2.2 that $\kappa_{m,i} \in V_t$ with $t = p^{\sigma(m,i)} \phi_{m-\sigma(m,i)}(\psi_m(i))$. Since $p^{\sigma(m,i)} \leq p^m i_m - i$, Lemma 3.3.5 implies that

$$\phi_{m-\sigma(m,i)}(\psi_m(i)) - \delta \geq \theta_{\sigma(m,i)-1}(i),$$

and Lemma 3.3.4 then states that $t \geq i$.

Finally, if $r = p$ and $i_{m+1} = p^l + 1$ for some $l \geq 0$, then Lemma 3.3.5 similarly implies that $\kappa_{m,i} = \rho^{\sigma(m+1,i)} \beta_{m-\sigma(m+1,i),l}$. By Theorem 3.2.4, we have in this case that $\kappa_{m,i} \in V_t$ with

$$t = p^{\sigma(m+1,i)} \phi'_{m-\sigma(m+1,i)}(p^l - 1) \geq i,$$

the inequality again following from Lemmas 3.3.4 and 3.3.5. □

3.4. *Generating sets.* In this subsection, we give explicit minimal generating sets of all of the A-modules V_i in terms of the elements $\kappa_{m,i}$ of the previous section. We begin with generation. Recall that $\delta \in \{0, 1\}$ is 1 if and only if $r = p$.

Theorem 3.4.1. *We let $S_i = \{\kappa_{m,i} \mid 0 \leq m \leq s\}$ for*

$$s = \left\lceil \log_p \left(\frac{i+1}{r+1+\delta(p-1)} \right) \right\rceil.$$

If $2 \leq r \leq p-1$, then S_i generates V_i as an A-module, while if $r = p$, then $S_i \cup \{p^{\lceil \log_p(i) \rceil} w\}$ generates V_i as an A-module.

Proof. Let $t = (i+1)_m - 1$. In the case that $2 \leq r \leq p-1$, we have $\psi_m(i) = \psi(t)$ and $\psi(t) > 0$ if and only if $\frac{i+1}{p^m} > r+1$, or $m < \log_p(\frac{i+1}{r+1})$. The smallest m such that $\psi_m(i) = 0$ is therefore s. If $r = p$, then $\psi'_m(i) = \psi(t) - \epsilon_t$, where $\epsilon_t \in \{0, 1\}$ is 1 if and only if $p^{l+1} + 1 \leq t \leq p^{l+1} + p - 1$ for some $l \geq 0$. In particular, we have $\psi(t) > \epsilon_t$ if and only if $t \geq 2p$, so the smallest m such that $\psi'_m(i) = 0$ is again s.

It suffices to show that the images of our elements generate V_i / V_{i+p-1}. Suppose that $\alpha = (\rho^k b T^j + c) u_r + d w \in V_i$ for some nonnegative integers j, k, $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1} A$, and $d \in \mathbb{Z}_p[\Phi]$ (with $d = 0$ if $r \neq p$). Let $m = \min(k, s)$. Then $j \geq \psi'_m(i)$ by Theorems 3.2.2 and 3.2.4 and Corollary 3.3.2 (and the fact that $\psi'_s(i) = 0$), and we set

$$\alpha' = \alpha - \rho^{k-m} b T^{j - \psi'_m(i)} \kappa_{m,i} \in V_i \cap (A(T^{j+1} u_r, w)).$$

If $r \leq p-1$, we may repeat this process recursively until we obtain an element of V_{i+p-1}. If $r = p$, either $\kappa_{m,i} \in A u_p$ or $\kappa_{m,i} \in \rho^{m+l+1} w + A u_p$ for some $l \geq 0$ with $i < p^{m+l+1} + p^{m+1}$. Since $(T, p) p^{m+l+1} w \subseteq V_{p^{m+l+2}}$, there exists an element

$$\alpha'' \in V_i \cap (T^{j+1} A u_r + \mathbb{Z}_p[\Phi] w)$$

with $\alpha'' - \alpha' \in V_{i+p-1}$, and again we may repeat the process until we obtain an element of V_{i+p-1} plus an element of $V_i \cap \mathbb{Z}_p[\Phi] w = \mathbb{Z}_p[\Phi] p^{\lceil \log_p(i) \rceil} w$. □

Lemma 3.4.2. *If $m \geq 1$ is such that $\theta_m(i) \geq 1$, then $\theta_{m-1}(i) \geq \theta_m(i) + 2$.*

Proof. First, suppose that $\theta_m(i) \geq 1$, and note that $i_{m-1} \geq p(i_m - 1) + 1$. Therefore,

$$\theta_{m-1}(i) \geq \psi(p(i_m - 1) + 1) = i_m - 1 + \left\lfloor \frac{2 + \{r - i_m\}}{p} \right\rfloor - \delta. \tag{3.4.1}$$

On the other hand,

$$\theta_m(i) = \left\lfloor \frac{i_m+1+\{r-i_m\}}{p} \right\rfloor - \delta. \tag{3.4.2}$$

In particular, $\theta_m(i) = 1$ exactly when $r+1 \le i_m \le r+(\delta+1)(p-1)$. In this case,

$$\theta_{m-1}(i) \ge r+1-\delta \ge 3 = \theta_m(i)+2.$$

In general, (3.4.1) and (3.4.2) tell us that

$$\theta_{m-1}(i) \ge i_m - 1 - \delta \quad \text{and} \quad \frac{i_m}{p} + 1 - \delta \ge \theta_m(i),$$

and we have

$$i_m - 1 - \delta \ge \frac{i_m}{p} + 3 - \delta$$

if and only if $i_m \ge \frac{4p}{p-1}$, which holds for $i_m \ge r+p$ unless $i_m = 5$, $r = 2$, and $p = 3$, in which case $\theta_{m-1}(i) \ge 5$ and $\theta_m(i) = 2$. □

For each $m \ge 0$, let us set $\epsilon_m(i) = \theta_m(i) - \psi'_m(i)$, which lies in $\{0, 1\}$ by Lemma 3.3.4 and the remark before it. The following corollary is useful in understanding the form of our special elements.

Corollary 3.4.3. *For every $m \ge 0$, we have $\psi'_m(i) \ge \psi'_{m+1}(i)$, with equality if and only if $\psi'_m(i) = 0$.*

Proof. If $\theta_{m+1}(i) \ge 1$, Lemma 3.4.2 and the fact that $\epsilon_k(i) \in \{0, 1\}$ for all k imply that $\psi'_m(i) > \psi'_{m+1}(i)$. Otherwise, $\psi'_{m+1}(i) = 0$, and the inequality holds automatically, with equality exactly if $\psi'_m(i) = 0$. □

We next show that the sets given in Theorem 3.4.1 are minimal unless $r = p$. It is in the proof of this result that the refined elements $\kappa_{m,i}$ first hold an advantage of ease of use over the elements of Section 3.2.

Theorem 3.4.4. *For $r \le p-1$, no proper subset of S_i generates V_i as an A-module. For $r = p$, every proper subset of $S_i \cup \{p^{\lceil \log_p(i) \rceil} w\}$ that generates V_i as an A-module must contain S_i.*

Proof. Assume first that $2 \le r \le p-1$. Suppose that

$$\sum_{m=0}^{s} c_m \kappa_{m,i} = 0, \tag{3.4.3}$$

where $c_m \in A$ for $m \le s$. We must show that no c_m is a unit. We prove the somewhat stronger claim that $c_m \in (p, T^{\epsilon_m(i)+1})$ for each m.

Fix a nonnegative integer $m \le s$. If $\epsilon_m(i) = 0$, then $\kappa_{m,i} = \rho^m T^{\theta_m(i)} u_r$ by (3.3.5). If $\epsilon_m(i) = 1$, then (3.3.6) tells us that

$$\kappa_{m,i} \equiv \rho^m T^{\theta_m(i)-1} u_r \bmod A T^{\theta_m(i)+1} u_r,$$

noting Lemma 3.4.2. Set

$$X_m = \{k \in \mathbb{Z} \mid m < k \leq s,\ \epsilon_k(i) = 1,\ \sigma(k, i) \leq m\}, \tag{3.4.4}$$

which is actually a set of cardinality at most one, though we do not need this fact. Let $k \leq s$. If $k \in X_m$, then (3.3.6) and Lemma 3.4.2 together imply that

$$\kappa_{k,i} \equiv -a_{k,i}\rho^m T^{\theta_m(i)-1}u_r \bmod (p^{m+1}, T^{\theta_m(i)+1})u_r,$$

and if $k \notin X_m$, they and (3.3.5) similarly imply that $\kappa_{k,i} \in (p^{m+1}, T^{\theta_m(i)+1})u_r$. Thus, (3.4.3) yields the congruence

$$c_m\rho^m T^{\psi_m(i)} \equiv \sum_{k \in X_m} c_k a_{k,i}\rho^m T^{\theta_m(i)-1} \bmod (p^{m+1}, T^{\theta_m(i)+1}). \tag{3.4.5}$$

If the claim holds for all $k > m$, then we have $c_k \in (p, T^2)$ for each $k \in X_m$, so $c_m \in (p, T^{\epsilon_m(i)+1})$, as desired.

If $r = p$, a completely analogous argument shows that at most $p^{\lceil \log_p(i) \rceil}w$ is unnecessary for generation, if one works modulo $Aw = \mathbb{Z}_p[\Phi]w + A(\varphi - 1)u_p$ throughout. Here, one should replace X_m by

$$X'_m = \{k \in \mathbb{Z} \mid m < k \leq s, \epsilon_k(i) = 1, \sigma'(k, i) \leq m\}, \tag{3.4.6}$$

where we set $\sigma'(k, i) = \sigma(k, i)$ unless $i_{k+1} = p^l + 1$ for some $l \geq 0$, in which case we set $\sigma'(k, i) = \sigma(k + 1, i)$. □

For the purpose of completeness, we also give the precise condition on i under which no proper subset of $S_i \cup \{p^{\lceil \log_p(i) \rceil}w\}$ generates V_i in the case that $r = p$.

Proposition 3.4.5. *For $r = p$, the set S_i generates V_i if and only if $i_s = p + 1$.*

Proof. To determine whether $p^{\lceil \log_p(i) \rceil}w$ is or is not necessary, we work in distinct ranges of i separately. Note that the definition of s forces $2p^s < i < 2p^{s+1}$.

Case 1: $2p^s < i \leq p^{s+1}$. In this case, all of the elements $\kappa_{m,i}$ lie in Au_p, and therefore $p^{s+1}w$ is necessary.

Case 2: $p^{s+1} < i \leq p^{s+1} + p^s - p^{s-1}$. In this range, we have

$$\kappa_{s,i} = \rho^s u_p + \rho^{s+1}w \quad \text{and} \quad \kappa_{s-1,i} = \rho^{s-1}T^{p-1}u_p + \rho^{s+1}w.$$

Note that $(T - p)\kappa_{s,i} = \rho^s(T - p)u_p + \rho^{s+1}(\varphi - 1)u_p = \rho^s(T - \rho)u_p$, so

$$\rho^s T u_p \equiv \rho^{s+1}u_p \bmod A\kappa_{s,i} \quad \text{and} \quad \rho^s u_p \equiv -\rho^{s+1}w \bmod A\kappa_{s,i}. \tag{3.4.7}$$

Applying these to $\rho\kappa_{s-1,i}$, we obtain

$$\rho\kappa_{s-1,i} \equiv \rho^{s+p-1}u_p + \rho^{s+2}w \equiv (-\rho^{s+p} + \rho^{s+2})w \bmod A\kappa_{s,i},$$

which in particular tells us that $p^{s+2}w \in A(\kappa_{s-1,i}, \kappa_{s,i})$.

Case 3. $p^{s+1}+p^s-p^{s-1}<i\le p^{s+1}+p^s$. In this range, we have

$$\kappa_{s,i}=\rho^s u_p+\rho^{s+1}w \quad \text{and} \quad \kappa_{s-1,i}=\Bigl(\rho^{s-1}T^{p-1}+\sum_{k=\sigma(s,i)}^{s-2}\rho^k T^{\theta_k(i)-1}\Bigr)u_p+\rho^{s+1}w$$

with $\sigma(s,i)\le s-2$. Moreover, $\kappa_{m,i}\in Au_r$ for $m\le s-2$.

Set $\nu_{k,i}=\rho^k T^{\theta_k(i)}u_r$ for all nonnegative k. We note that $\nu_{m,i}\in A(\kappa_{0,i},\ldots,\kappa_{m,i})$ for $m\le s-2$: If $\kappa_{m,i}\ne\nu_{m,i}$, which is to say $\epsilon_m(i)=1$, then

$$\nu_{m,i}=T\kappa_{m,i}+a_{m,i}\sum_{k=\sigma(m,i)}^{m-1}\nu_{k,i}.$$

Let $j=\theta_{s-2}(i)-p$, and note that $j\ge p^2-1\ge 2$. Since

$$T^j\kappa_{s-1,i}\equiv\rho^{s-1}T^{\theta_{s-2}(i)-1}u_p+\rho^{s+1}T^j w \mod A(\nu_{\sigma(s,i),i},\ldots,\nu_{s-2,i})$$

and $\rho^{k+1}T^{\theta_k(i)-1}u_p\in A\nu_{k+1,i}$ for all k with $\sigma(s,i)\le k\le s-3$, we therefore have

$$(\rho-T^j)\kappa_{s-1,i}\equiv\rho^sT^{p-1}u_p+\rho^{s+1}(\rho-T^j)w \mod A(\nu_{\sigma(s,i),i},\ldots,\nu_{s-2,i}). \quad (3.4.8)$$

Using (3.4.7) to reduce (3.4.8), we see that

$$(\rho-T^j)\kappa_{s-1,i}\equiv\rho^{s+2}(1-\rho^{p-2}-\rho^{j-1})w \mod A(\nu_{\sigma(s,i),i},\ldots,\nu_{s-2,i},\kappa_{s,i}),$$

which implies that $p^{s+2}w\in A(\kappa_{0,i},\ldots,\kappa_{s,i})$.

Case 4: $p^{s+1}+p^s<i<2p^{s+1}$. In this case, all of the $\kappa_{m,i}$ with $m\le s-1$ lie in AT^pu_p, and so for $p^{s+2}w$ to be unnecessary, there would have to exist $c\in A$ such that

$$c\kappa_{s,i}\equiv p^{s+2}w \mod AT^2u_p. \quad (3.4.9)$$

Note that $c\kappa_{s,i}\equiv c(\rho^su_p+\rho^{s+1}w) \mod AT^2u_p$, which forces $c\equiv T^2c' \mod (\varphi-1)$ for some $c'\in A$. This means that

$$c\kappa_{s,i}\equiv c'p^{s+3}w \mod A(u_p,(\varphi-1)w),$$

but $p^{s+2}w\notin A(p^{s+3}w,(\varphi-1)w,u_p)$, so (3.4.9) cannot hold. □

4. The finite level

4.1. ***Norms and eigenspace structure.*** In this section, we explore the consequences of the results of Section 3 for unit groups of actual abelian local fields of characteristic 0. Fix a positive integer n. Recall from the introduction that F_n is the field obtained from E by adjoining the p^nth roots of unity and that $U_{n,t}$ denotes the tth unit group of F_n for $t\ge 1$. As before, we set $\Gamma_n=\mathrm{Gal}(F_n/F_1)$.

For positive integers $m \geq n$, let $N_{m,n}$ and $\mathrm{Tr}_{m,n}$ denote, respectively, the norm and trace from F_m to F_n. We also let N_n denote the restriction map $N_n \colon F^\times \to F_n^\times$ on norm compatible sequences. Recall that $\lambda_n = N_n(\lambda) = 1 - \zeta_{p^n}$, where $\zeta_{p^n} = N_n(\zeta)$ is a primitive p^nth root of unity. We require a few preliminary lemmas.

Lemma 4.1.1. *One has*

$$\mathrm{Tr}_{n+1,n}(\lambda_{n+1}^{pk-\epsilon}) \equiv p\lambda_n^{k-\epsilon} \bmod p^3$$

for all $k \geq 1$ *and* $\epsilon \in \{0, 1\}$.

Proof. An easy calculation shows that

$$\mathrm{Tr}_{n+1,n}(\lambda_{n+1}^t) = p \sum_{j=0}^{\lfloor \frac{t}{p} \rfloor} \binom{t}{pj} (-\zeta_{p^n})^j$$

for every $t \geq 0$. The result follows since

$$\binom{pk-\epsilon}{pj} = \binom{k-\epsilon}{j} \prod_{\substack{s=1 \\ p\nmid s}}^{p(k-j)} \Bigl(1 + \frac{pj}{s}\Bigr) \equiv \binom{k-\epsilon}{j} \bmod p^2 \quad \text{for any } j \geq 0. \qquad \square$$

Let $e_n = p^{n-1}(p-1)$ denote the ramification index of E. In applying Lemma 4.1.1, it is useful to make note of the fact that

$$p \equiv -\lambda_n^{e_n} \bmod \lambda_n^{p^n}. \tag{4.1.1}$$

Lemma 4.1.2. *For* $t \geq 1$ *and any unit* η *in* E, *one has*

$$N_{n+1,n}(1 + \eta\lambda_{n+1}^t) \equiv \begin{cases} 1 + \eta^p \lambda_n^t \bmod \lambda_n^{t+1} & \text{if } t < p^n - 1, \\ 1 + (\eta^p - \eta)\lambda_n^{p^n-\epsilon} \bmod \lambda_n^{p^n+1-\epsilon} & \text{if } t = p^n - \epsilon,\ \epsilon \in \{0, 1\}, \\ 1 - \eta\lambda_n^{e_n+k-\epsilon} \bmod \lambda_n^{e_n+k+1-\epsilon} & \text{if } t = pk - \epsilon > p^n,\ \epsilon \in \{0, 1\}. \end{cases}$$

Moreover, we have

$$N_{n+1,n}(1 + \eta\lambda_{n+1}^t) \equiv 1 \bmod \lambda_n^{e_n + \lfloor t/p \rfloor}$$

for all $t > p^n$.

Proof. The jump in the ramification filtration of $\mathrm{Gal}(F_{n+1}/F_n)$ occurs at $p^n - 1$. By [Serre 1979, Lemmas V.4 and V.5], we have

$$N_{n+1,n}(1 + \eta\lambda_{n+1}^t) \equiv 1 + \eta \, \mathrm{Tr}_{n+1,n}(\lambda_{n+1}^t) + \eta^p \lambda_n^t \bmod \lambda_n^{e_n + \lfloor 2t/p \rfloor},$$
$$\mathrm{Tr}_{n+1,n}(\lambda_{n+1}^t) \equiv 0 \bmod \lambda_n^{e_n + \lfloor t/p \rfloor}.$$

The result is then a corollary of Lemma 4.1.1, upon applying (4.1.1). $\square$

Let D_n be the pro-p completion of $F_n^\times$, and let $D_n^{(r)} = D_n^{\varepsilon_r}$ for any $r \in \mathbb{Z}$. As before, we fix r with $2 \le r \le p$, and i will always denote a positive integer with $i \equiv r \bmod p-1$. Let $V_{n,i} = U_{n,i}^{\varepsilon_r} = U_{n,i} \cap D_n^{(r)}$ for any such i. These $V_{n,i}$ are all modules over $A_n = \mathbb{Z}_p[\Gamma_n \times \Phi]$. As in Lemma 2.1, we have isomorphisms

$$V_{n,i}/V_{n,i+p-1} \xrightarrow{\sim} \mathbb{F}_q$$

that send $1 + x\lambda_n^i$ for some x in the valuation ring of F_n to the element $\bar{x}$ of $\mathbb{F}_q$ that is identified with the image of x in the residue field of F_n under the isomorphism fixed in Section 2. We may then set $V'_{n,i} = V_{n,i} - V_{n,i+p-1}$ and define $V_{n,i}(\eta)$ for $\eta \in \mathbb{F}_q^\times$ as the set of elements $1 + x\lambda_n^i$ with $\bar{x} = \eta$.

We have the following consequence of Lemma 4.1.2.

Lemma 4.1.3. *For any $t \ge -1$, we have $N_{n+1,n}(V_{n+1,p^n+t}) \subseteq V_{n,p^n+t-(p-1)\lfloor (t+1)/p \rfloor}$, with equality for $t \ge 0$.*

Proof. Note that Lemma 4.1.2 yields $N_{n+1,n}(U_{n+1,p^n+pk-\epsilon}) = U_{n,p^n+k-\epsilon}$ for all $k \ge 0$ and $\epsilon \in \{0, 1\}$ with $k \ge \epsilon$, since every element in $U_{n,p^n+k-\epsilon}$ can be written as a product of elements of the form $1 + \eta_t \lambda_n^{p^n+t}$ with $t \ge k - \epsilon$ and $\eta_t \in \mathbb{F}_q$. (For $k = 0$ and $\epsilon = -1$, it tells us just that any element of U_{n+1,p^n-1} has a norm in U_{n,p^n-1}.)

Note that

$$U^{\varepsilon_r}_{n,p^n+k-\epsilon} = V_{n,p^n+k-\epsilon+\{r-k+\epsilon-1\}} \quad \text{and} \quad U^{\varepsilon_r}_{n+1,p^n+pk-\epsilon} = V_{n,p^n+pk-\epsilon+\{r-k+\epsilon-1\}}.$$

For any $t \ge 0$, we may write $t = pk - \epsilon + \{r - k + \epsilon - 1\}$ for some k, ϵ, and r, and we have

$$t - (p-1)\left\lfloor \frac{t+1}{p} \right\rfloor = k - \epsilon + \{r - k + \epsilon - 1\} \qquad \square.$$

The next corollary is almost immediate from Lemmas 4.1.2 and 4.1.3, so we leave it to the reader.

Corollary 4.1.4. *For any unit η in E, one has*

$$N_{n+1,n}(V_{n+1,i}(\eta)) \subseteq \begin{cases} V_{n,i}(\eta^p) & \text{if } i < p^n - 1, \\ V_{n,i}(\eta^p - \eta) & \text{if } i = p^n - 1, \\ V_{n,p^n+k-1}(-\eta) & \text{if } i = p^n + pk - 1 \text{ for some } k > 0, \end{cases}$$

with equality if $r \neq p-1$ or $i > p^n$.

As for the p-power map, we have a well-known and easy-to-prove fact:

Lemma 4.1.5. *Suppose that $i > p^{n-1}$. Then the pth power map induces an isomorphism $V_{n,i} \xrightarrow{\sim} V_{n,i+e_n}$, and we have $V_{n,i}(\eta)^p = V_{n,i+e_n}(-\eta)$ for all $\eta \in \mathbb{F}_q^\times$.*

Next, we discuss the restriction map from the field of norms to the finite level.

Proposition 4.1.6. *The map N_n induces maps $N_n\colon D^{(r)} \to D_n^{(r)}$ that are surjections for $r \neq p-1$ and which have procyclic cokernel for $r = p-1$. For $\eta \in \mathbb{F}_q^\times$, we have*

$$N_n(V_i(\eta)) \subseteq \begin{cases} V_{n,i}(\eta^{p^{-n}}) & \text{if } i \le p^n - 2, \\ V_{n,p^n-1}(\eta^{p^{-n}} - \eta^{p^{-n-1}}) & \text{if } i = p^n - 1, \\ V_{n,p^n+k-1}(-\eta^{p^{-n-1}}) & \text{if } i = p^n + pk - 1 \text{ for some } 0 < k < e_n. \end{cases}$$

Moreover, we have induced maps $V_i/V_{i+1} \to V_{n,i}/V_{n,i+1}$ for all $i < p^n$, and these are isomorphisms for $i \neq p^n - 1$. For $i \le p^n$, we have $V_{n,i} = N_n V_i$ if $r \neq p-1$, and $V_{n,i}/N_n V_i$ is procyclic if $r = p-1$.

Proof. That the cokernel of N_n is trivial if $r \neq p-1$ and procyclic if $r = p-1$ follows easily from local class field theory, but it is also a consequence of the argument that follows. The first jump in the ramification filtration of $\mathrm{Gal}(F_\infty/F_{n+1})$ is at $p^{n+1}-1$. In particular, for t less than this value, repeated application of Lemma 4.1.2 tells us

$$N_{n+1}(1+\eta\lambda^t) = \lim_{m\to\infty} N_{m,n+1}(1+\eta^{p^{-m}}\lambda_m^t) \equiv 1+\eta^{p^{-n-1}}\lambda_{n+1}^t \bmod \lambda_{n+1}^{t+1}.$$

Moreover, repeated application of Corollary 4.1.4 followed by two applications of Lemma 4.1.3 tells us that $N_n(V_{p^{n+1}-1+\{r\}}) \subseteq V_{n,p^n+e_n-1+\{r\}}$. An application of Corollary 4.1.4 then yields the stated containments.

Since $\eta^{p^{-n}}$ and $-\eta^{p^{-n-1}}$ run through all elements of $\mathbb{F}_q$ as $\eta \in \mathbb{F}_q$ varies, we obtain $V_{n,i} = N_n V_i + V_{n,i+p-1}$ for all $i \le p^n$ but $p^n - 1$. Noting Lemma 4.1.5, this implies

$$V_{n,i+ke_n} = N_n V_{p^k i} + V_{n,i+ke_n+p-1}$$

for $p^{n-1} < i \le p^n$ with $i \neq p^n - 1$ and $k \ge 0$. Note that every element of every $V_{n,i}$ may be written as an infinite product over $j \ge 0$ of one element from each of a fixed set of representatives of the $V_{n,i+j(p-1)}/V_{n,i+(j+1)(p-1)}$. Thus, we have $N_n V_i = V_{n,i}$ so long as $r \neq p-1$.

If $r = p-1$, we can choose an element z_n of $V_{n,p^n-1}(\xi)$ that is not a norm. By the formula proven above for $N_n(1+\eta\lambda^{p^n-1})$ modulo $\lambda_n^{p^n}$, we have

$$V_{n,p^n+ke_n-1} = N_n V_{p^k(p^n-1)} + V_{n,p^n+ke_n+p-2} + \langle z_n^{p^k} \rangle$$

for $k = 0$, and then for all $k \ge 0$ by taking powers. Therefore, $V_{n,i}/N_n V_i$ is generated by z_n for all $i < p^n$ with $i \equiv 0 \bmod p-1$. □

The following structural result is again essentially found in [Greither 1996], without the stated congruences. Here, we derive it from more basic principles.

Theorem 4.1.7. *For $r \le p-2$, the A_n-module $D_n^{(r)}$ is freely generated as an A_n-module by an element $u_{n,r} \in V_{n,r}(\xi)$. The A_n-module $D_n^{(p-1)}$ has a presentation*

$$D_n^{(p-1)} = \langle \pi_n, u_{n,p-1}, v_n \mid \pi_n^{\varphi} = \pi_n, \pi_n^{\gamma-1} = u_{n,p-1}^{N_\Phi}, v_n^{\gamma} = v_n, u_{n,p-1}^{N_{\Gamma_n}} = v_n^{1-\varphi} \rangle,$$

where $v = v_n^{\varphi^{2-n}} \equiv 1 + p\xi \mod p^2$ *is independent of* n *and* $u_{n,p-1} \in V_{n,p-1}(\xi)$ *for* $n \geq 2$, *while* $u_{1,p-1} \in V_{n,p-1}(\xi - \xi^{p^{-1}})$. *The* A_n*-module* $D_n^{(p)}$ *has a presentation*

$$D_n^{(p)} = \langle u_{n,p}, w_n \mid w_n^{\gamma-1-p} = u_{n,p}^{\varphi-1} \rangle$$

with $u_{n,p} \in V_{n,p}(\xi)$ *and* $w_n \in V_{n,1}(-\xi)$ *such that* $w_n^{N_\Phi} = \zeta_{p^n}$.

Proof. We set $u_{n,r} = (N_n u_r)^{\varphi^n}$, $\pi_n = N_n \pi$, and $w_n = (N_n w)^{\varphi^n}$ with u_r, π, and w as in Theorem 3.1.5. It follows from the surjectivity of N_n for $r \neq p-1$ in Proposition 4.1.6 that the element $u_{n,r}$ generates $D_n^{(r)}$ for $r \leq p-2$, while the elements w_n and $u_{n,r}$ generate $D_n^{(p)}$. By Hilbert's Theorem 90, the kernel of N_n consists exactly of elements of the form $\alpha^{\gamma^{p^{n-1}}-1}$ with $\alpha \in D$, and therefore it follows that $D_n^{(r)}$ is free of rank 1 on $u_{n,r}$ over A_n for $r \leq p-2$ and that $D_n^{(p)}$ has the stated presentation. (That $u_{1,p} \in V_{1,p}(\xi)$ requires a simple check using Propositions 3.1.3 and 4.1.6.)

The elements π_n and $u_{n,p-1}$ automatically satisfy the first two relations in the desired presentation of $D_n^{(p-1)}$. In particular,

$$u_{n,p-1}^{N_{\Gamma_n \times \Phi}} = \pi_n^{(\gamma-1)N_{\Gamma_n}} = 1,$$

so Hilbert's Theorem 90 tells us that $u_{n,p-1}^{N_{\Gamma_n}} = v_n^{1-\varphi}$ for some v_n in the pro-p completion of $E^\times$. By Proposition 4.1.6, we have

$$u_{n,p-1}^{N_{\Gamma_n}} \equiv 1 + (\xi^{p^{n-1}} - \xi^{p^{n-2}})\lambda_1^{p-1} \mod \lambda_1^p.$$

Noting (4.1.1), we may in fact choose $v_n \equiv 1 + p\varphi^{n-2}(\xi) \mod p^2$ with $v = v_n^{\varphi^{2-n}}$ independent of n.

Hilbert's Theorem 90 and Theorem 3.1.5 tell us that the A_n-module generated by $u_{n,p-1}$ is isomorphic to $A_n/(N_{\Gamma_n \times \Phi})$. By Proposition 4.1.6, the cokernel of N_n on $D^{(p-1)}$ is isomorphic to $\mathbb{Z}_p$. We claim that the image of v topologically generates this cokernel. If this is the case, then clearly $D_n^{(p-1)}$ is generated by π_n, $u_{n,p-1}$, and v, and any solution with $b, d \in \mathbb{Z}_p$ and $c \in A_n$ to $\pi_n^b u_{n,p-1}^c v^d = 1$ must satisfy $b = d = 0$ and $c \in \mathbb{Z}_p N_{\Gamma_n \times \Phi}$.

It remains only to demonstrate the claim. Suppose by way of contradiction that there exists $a \in A_n$ such that $x = v u_{n,p-1}^a$ is a pth power in $D_n^{(p-1)}$. This implies that $x^{\gamma-1} = u_{n,p-1}^{a(\gamma-1)}$ is a pth power in the A_n-module generated by $u_{n,p-1}$. It follows that $a(\gamma-1) \in A_n(p, N_{\Gamma_n \times \Phi})$, which forces $a(\gamma-1) \equiv 0 \mod p$, so $a \in A_n(p, N_{\Gamma_n})$. It then suffices to show that

$$v u_{n,p-1}^{bN_{\Gamma_n}} = v^{1+b\varphi^{n-2}(1-\varphi)}$$

is not a pth power in F_n for any $b \in \mathbb{Z}_p[\Phi]$. If it were for some b, then v^{N_Φ} and hence $1+p$ would be a pth power in F_n as well, but this is clearly not the case. □

4.2. *Special elements.* We assume for the rest of the paper that $n \geq 2$, the case that $n = 1$ being slightly exceptional but also completely straightforward. In this subsection, we construct special elements in the groups in the unit filtration of $F_n^\times$. Aside from the case that $r = p - 1$, these arise as restrictions of the elements introduced in Section 3.2.

Note that $\mathbb{Z}_p[\Gamma_n] \cong \mathbb{Z}_p[T]/(f_n)$, where $f_n = (T+1)^{p^{n-1}} - 1$. Of course, we can then speak of the action of T on an element of $D_n^{(r)}$. Once again reverting to additive notation, the following is now an immediate corollary of Theorem 3.2.2 and Proposition 4.1.6.

Proposition 4.2.1. *Let m and j be nonnegative integers with $\phi_m(j) < p^n - 1$. Define*

$$\alpha_{n,m,j} = \frac{1}{[r]!}\left(\{r-\delta-j\}!\,\rho^m T^j - \sum_{k=1}^{m} \rho^{m-k} T^{\phi_{k-1}(j)-\delta}\right) u_{n,r},$$

unless $j = 0$ and $r = p-1$, in which case we replace $\{r-\delta-j\}!$ with -1 in the formula. Then $\alpha_{n,m,j} \in V_{n,\phi_m(j)}(\xi)$. Furthermore, $(p^m b T^j + c)u_{n,r} \notin V_{n,\phi_m(j)+p-1}$ for all $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$ and $c \in T^{j+1}A_n$.

For nonnegative $m \leq n-2$, define $\phi'_{n,m} : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by $\phi'_{n,m}(j) = \phi'_m(j)$ unless $r = p-1$ and $j = e_{n-m-1}$, in which case we set

$$\phi'_{n,m}(e_{n-m-1}) = e_n + p^{m+1} - 1 = \phi_m(e_{n-m-1}) + p^m(p-1).$$

For nonnegative k, define $\vartheta_{2,k} = 1 + \varphi^{-1} + \cdots + \varphi^{-k}$ and $\vartheta_{j,k} = 1$ for $j > 2$. Note that $\vartheta_{2,k} \in p\mathbb{Z}_p[\Phi]$ if and only if $k \equiv -1 \bmod p|\Phi|$.

By Theorem 4.1.7, every element of $V_{n,p-1}$ may be written as $cu_{n,p-1} + dv$ with $c \in A_n$ and $d \in \mathbb{Z}_p$, and this representation is unique up to the choice of c modulo $N_{\Gamma_n \times \Phi}$. For $a, b \in D_n^{(r)}$, we again write $a \sim b$ if $a, b \in V_{n,i}(\eta)$ for some i and $\eta \in \mathbb{F}_q^\times$.

Theorem 4.2.2. *Let $m \leq n-2$ be a nonnegative integer, and define*

$$\omega_{n,m} = \sum_{k=0}^{m} \rho^{m-k} \vartheta_{n-m,k} T^{p^{n-m+k-2}(p-1)+p^k-1} u_{n,p-1} - v.$$

Then we have $\omega_{n,m} \in V_{n,e_n+p^{m+1}-1}(\xi)$. Furthermore, if $j \geq 0$ with $\phi_m(j) < p^n$, then

$$(p^m T^j b + c)u_{n,p-1} + dv \notin V_{\phi'_{n,m}(j)+p-1}$$

for all $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1}A_n$, and $d \in \mathbb{Z}_p$.

Proof. Let l be a nonnegative integer with $l \leq m$. We define

$$\omega_{n,m,l} = \sum_{k=0}^{l} \rho^{m-k} \vartheta_{n-m,k} T^{p^{n-m+k-2}(p-1)+p^k-1} u_{n,p-1} - v. \tag{4.2.1}$$

We claim not only that $\omega_{n,m} = \omega_{n,m,m} \in V_{n,e_n+p^{m+1}-1}(\xi)$, but that, for $l < m$,

$$\omega_{n,m,l} \in \begin{cases} V_{n,e_n+p^{m+1}-p^{m-l}}(\vartheta_{n-m,l+1}\xi) & \text{if } p \nmid \vartheta_{n-m,l+1}, \\ V_{n,e_n+p^{m+1}-p^{m-l-1}}(\xi) & \text{if } p \mid \vartheta_{n-m,l+1}. \end{cases}$$

We note, to begin with, that $\omega_{n,m,l} \in V_{n,e_n+p-1}$, since Lemma 3.2.1a implies

$$\omega_{n,m,l} + v \sim \rho^m T^{e_{n-m-1}} u_{n,p-1} \in V_{n,e_n}(-\xi).$$

For a given i, we take $V_{n,i}(0)$ to mean $V_{n,i+p-1}$ in what follows.

If $p \nmid \vartheta_{n-m,l}$, then Lemmas 3.2.1a and 4.1.5 imply that

$$T\omega_{n,m,l} \sim \rho^{m-l} \vartheta_{n-m,l} T^{p^{n-m+l-2}(p-1)+p^l} u_{n,p-1}$$

if $l < m$, $m = 0$, or $m < n-2$, and we have

$$T\omega_{n,m,l} \in \begin{cases} V_{n,e_n+p^{m+1}+p^{m-l}(p-2)}(-\xi) & \text{if } m < n-2 \text{ or } m = 0, \\ V_{n,p^n+p^{m-l-1}(p-2)}(\vartheta_{2,l}\varphi^{-1}\xi) & \text{if } l < m = n-2. \end{cases} \tag{4.2.2}$$

On the other hand, if $p \mid \vartheta_{n-m,l}$, then we have $T\omega_{n,m,l} \sim T\omega_{n,m,l-1}$, so we can still apply (4.2.2). Moreover, since $\vartheta_{2,n-3}\varphi^{-1} - \vartheta_{2,n-2} = -1$ for $n \geq 3$, we have

$$T\omega_{n,n-2} \sim \rho\vartheta_{2,n-3} T^{p^{n-3}} u_{n,p-1} + \vartheta_{2,n-2} T^{p^{n-2}} u_{n,p-1} \in V_{n,p^n+p-2}(-\xi).$$

We prove our claim by induction on m. In the case that $m = 0$, we have that $T\omega_{n,0} \in V_{n,e_n+2p-2}(-\xi)$ by (4.2.2), and we have seen that $\omega_{n,0} \in V_{n,e_n+p-1}$, so Proposition 2.5 forces $\omega_{n,0} \in V_{n,e_n+p-1}(\xi)$. For $m \geq 1$, that $\omega_{n,m} \in V_{n,e_n+p^{m+1}-1}$ follows from the claim for $l = m-1$ and the fact that

$$\omega_{n,m} - \omega_{n,m,m-1} = \vartheta_{n-m,m} T^{e_{n-1}+p^m-1} u_{n,p-1}$$

is an element of $V_{n,e_n+p^{m+1}-p}(-\vartheta_{n-m,m}\xi)$. Since $T\omega_{n,m} \in V_{n,e_n+p^{m+1}+p-2}(-\xi)$, an application of Proposition 2.5 would then yield that $\omega_{n,m} \in V_{n,e_n+p^{m+1}-1}(\xi)$. So, to perform the inductive step for $l < m$, we assume that either $p \nmid \vartheta_{n-m,l+1}$ or $l = m-1$, since otherwise $\omega_{n,m,l} \sim \omega_{n,m,l+1}$ and $l+1 < m$.

By Lemma 4.1.5 and induction, we have

$$\begin{aligned} &N_{n,n-1}(\omega_{n,m,l}) \\ &\quad = p\omega_{n-1,m-1,l} \in \begin{cases} V_{n-1,2e_{n-1}+p^m-p^{m-l-1}}(-\vartheta_{n-m,l+1}\xi) & \text{if } l < m-1, \\ V_{n-1,2e_{n-1}+p^m-p^{m-l-1}}(-\xi) & \text{if } l = m-1. \end{cases} \end{aligned} \tag{4.2.3}$$

Let i be such that $\omega_{n,m,l} \in V'_{n,i}$, and set $t = e_n + p^{m+1} - p^{m-l}$. By Lemma 4.1.3, we have both that $i \leq t+p-1$ and that there exists $x \in V'_{n,t}$ with $N_{n,n-1}(x) = p\omega_{n-1,m-1,l}$. Hilbert's Theorem 90 implies that $x - \omega_{n,m,l} \in A_n f_{n-1} u_{n,p-1}$. Note that

$$pf_{n-1}u_{n,p-1} \sim pT^{p^{n-2}} u_{n,p-1} \in V_{n,p^n+p-2},$$

while $\omega_{n,m,l} \notin V_{n,p^n+p-2}$. It follows that

$$x \sim \omega_{n,m,l} + bT^g u_{n,p-1}, \tag{4.2.4}$$

for some $b \in A_n$ with $b \notin (p, T)$ and $g \geq p^{n-2}$. Since $bT^{g+1}u_{n,p-1} \in V'_{n,\phi(g+1)}$ by Lemma 3.2.1a and both Tx and $T\omega_{n,m,l}$ lie in $V_{n,t+p-1}$, the latter by (4.2.2), we have $\phi(g+1) > t$ and hence $\phi(g) \geq t$. Therefore, we have $bT^g u_{n,p-1} \in V_{n,t}$, and (4.2.4) now forces $i \geq t$, which means that $i \in \{t, t+p-1\}$.

If $l < n-3$, then Lemma 2.2 forces $i = t$ in order for (4.2.2) to hold. If $l = n-3$ and $i = t+p-1$, then Proposition 2.5 and (4.2.2) force $\omega_{n,n-2,n-3}$ to be in $V_{n,p^n-1}(-\vartheta_{2,n-3}\varphi^{-1}\xi)$. By Corollary 4.1.4, this implies that

$$N_{n,n-1}(\omega_{n,n-2,n-3}) \in V_{n-1,2e_{n-1}+p^{n-2}-1}(\vartheta_{2,n-3}\varphi^{-1}\xi),$$

and then (4.2.3) tells us that $p \mid \vartheta_{2,n-2}$ and $\omega_{n,n-2,n-3} \in V_{n,p^n-1}(\xi)$.

If $i = t$, then Lemma 3.2.1a implies that

$$\omega_{n,m,l} \sim -dT^{e_{n-1}+p^m-p^{m-l-1}}u_{n,p-1} \tag{4.2.5}$$

for some $d \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$. Set

$$z = \omega_{n,m,l} + dT^{e_{n-1}+p^m-p^{m-l-1}}u_{n,p-1} \in V_{n,t+p-1}.$$

By (4.2.2) and Lemma 3.2.1a, we have $Tz \in V_{n,t+2(p-1)}(-d'\xi)$, where $d' = d$ if $l < n-3$ and $d' = d - \vartheta_{2,n-3}\varphi^{-1}$ if $l = n-3$. We therefore have $z \in V_{n,t+p-1}(d'\xi)$, and then

$$N_{n,n-1}(z) \in V_{n-1,2e_{n-1}+p^m-p^{m-l-1}}(-d'\xi)$$

by Corollary 4.1.4. On the other hand, we have

$$N_{n,n-1}(T^{e_{n-1}+p^m-p^{m-l-1}}u_{n,p-1}) = \varphi T^{e_{n-1}+p^m-p^{m-l-1}}u_{n-1,p-1} \in V_{n-1,t},$$

so we have $N_{n,n-1}(z) \sim N_{n,n-1}(\omega_{n,m,l})$. By (4.2.3), we then have $d = \vartheta_{n-m,l+1}$. If $p \mid \vartheta_{n-m,l+1}$, then $l = n-3$ by assumption, and this contradicts our assumption on i and implies the claim for $\omega_{n,n-2,n-3}$. Otherwise, we have already shown that $i = t$, and (4.2.5) and Lemma 3.2.1a yield the claim.

Suppose now that $j \geq 0$, $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1}A_n$, and $d \in \mathbb{Z}_p$ are such that $\phi_m(j) < p^n$ and

$$\omega = (p^m bT^j + c)u_{n,p-1} + dv \in V'_{n,i}$$

for some $i \geq \phi'_{n,m}(j)$. We suppose that $\phi_m(j) \geq e_n$, as the result otherwise reduces to Proposition 4.2.1. For $m = 0$, if $(bT^j + c)u_{n,p-1} \not\sim -dv$, then $i = e_n$ or $i = \phi(j) \leq \phi'_{n,0}(j)$. Otherwise, we must have $j = e_{n-1}$, and since $T\omega \sim bT^{j+1}u_{n,p-1}$, the argument of Lemma 3.2.1b tells us that $i = e_n + p - 1$.

For $m \geq 1$, we rewrite c as $pc' + T^h\nu$ for some $h \geq j+1$ and $c', \nu \in A_n$ with $\nu \notin (p, T)$. Note that $\phi'_{n,m}(j) = p\phi'_{n-1,m-1}(j) + p - 1$. By induction, we have

$$(p^{m-1}bT^j + c')\varphi u_{n-1,p-1} + dv \notin V_{n-1,\phi'_{n-1,m-1}(j)+p-1}.$$

The pth power of this element is the norm from F_n of

$$\omega' = \omega - T^h \nu u_{n,p-1} = (p^m b T^j + pc')u_{n,p-1} + d\nu,$$

and $\omega' \notin V_{n,\phi'_{n,m}(j)+p-1}$ by Lemma 4.1.3. If $\omega' \in V_{n,\phi'_{n,m}(j)}$, then the fact that $\phi'_{n,m}(j)$ is -1 modulo p and therefore not a value of ϕ implies that $\omega' \not\sim -T^h \nu u_{n,p-1}$, so we have $i = \phi'_{n,m}(j)$.

So, assume that $\omega' \notin V_{n,\phi'_{n,m}(j)}$. Then $\omega' \sim -T^h \nu u_{n,p-1}$, and Lemma 3.2.1a implies that $\phi(h) < \phi'_{n,m}(j) \le i$. If $\omega \notin V_{n,\phi(h+1)}$, then we must have

$$i = \phi(h) + p - 1 = \phi'_{n,m}(j).$$

So, we assume moreover that $\omega \in V_{n,\phi(h+1)}$, in which case $T\omega' \sim -T^{h+1}\nu u_{n,p-1}$. Since $T\omega'$ is a power of p, either $\phi(h+1)$ is divisible by p and less than p^n, or $\phi(h+1) > p^n$. In the former case, unless $\phi(h+2) > p^n$, we would have $T^2\omega' \in V_{n,\phi(h+1)+p(p-1)}$ and then $T^2\omega \in V'_{n,\phi(h+2)}$, contradicting $\omega \in V_{n,\phi(h+1)}$. We therefore have $\phi(h+2) > p^n$ in both cases, so $T\omega' \in V_{n,p^n-p}$. By Proposition 4.2.1 and the fact that $\phi_m(p^{n-m-1}) > p^n$, this forces $j = p^{n-m-1} - 1$. If $m < n-2$, then

$$p^n - p \le \phi(h) + p - 1 \le \phi'_{n,m}(j) = \phi_m(j) = p^n - p^{m+1} + p^m - 1,$$

which is a contradiction. We therefore have $m = n-2$ and $j = p-1$, so

$$p^n - 1 = \phi'_{n,n-2}(p-1) > \phi_{n-2}(p-1),$$

which, noting Proposition 4.2.1, implies that $p \nmid d$ and then, noting Theorem 4.1.7, that $\omega \notin pD_n^{(p-1)}$. In particular, $\omega \notin V_{n,p^n+p-2}$, so $i = p^n - 1$. □

Remark 4.2.3. Note that $\phi_{n-1}(0) = p^n - 1 < p^n$ as well, but in this case, the element $u_{n,p-1}^{N_{\Gamma_n \times \Phi}} = 1$ has the form $(p^{n-1}b + c)u_{n,p-1}$ with $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$ and $c \in TA$.

For $r = p$, the following is a consequence of Theorem 3.2.4 and Proposition 4.1.6.

Proposition 4.2.4. *Let m and l be nonnegative integers with $\phi_m(p^l - 1) \le p^n$. Let*

$$\beta_{n,m,l} = \left(\rho^m T^{p^l-1} + \sum_{k=1}^{m} \rho^{m-k} T^{\phi'_{k-1}(p^l-1)-1}\right) u_{n,p} + \rho^{m+l+1} w_n.$$

Then $\beta_{n,m,l} \in V_{n,\phi'_m(p^l-1)}(-\xi)$ unless $l = n-1$ and $m = 0$, in which case $\beta_{n,0,n-1} \in V_{n,p^n}(\xi^{p^{-1}})$. Furthermore, for any $j \ge 0$ with $\phi_m(j) \le p^n$, we have

$$(p^m b T^j + c)u_{n,p} + d w_n \notin V_{n,\min(\phi'_m(j)+p-1,\, p^n+p-1)}$$

for all $b \in \mathbb{Z}_p[\Phi] - p\mathbb{Z}_p[\Phi]$, $c \in T^{j+1}A_n$ and $d \in \mathbb{Z}_p[\Phi]$.

4.3. ***Generating sets.*** In this final subsection, we turn to the task of finding small generating sets for the groups $V_{n,i}$ as A_n-modules. First, we define the refined elements that will be used in forming these sets.

Suppose that $i \le p^n$ and

$$0 \le m \le \left\lceil \log_p\Bigl(\frac{i+1}{r+1+\delta(p-1)}\Bigr)\right\rceil.$$

Aside from the case that $r = p-1$ and $p^m < i - e_n < p^{m+1}$, we set $\kappa_{n,m,i} = \varphi^n N_n \kappa_{m,i}$, which can be written down explicitly as in the formulas (3.3.5), (3.3.6), and (3.3.7), but now with u_r replaced by $u_{n,r}$ and w replaced by w_n. By Propositions 3.3.9, 4.1.6 and 4.2.4, we have $\kappa_{n,m,i} \in V_{n,i}$.

If $r = p-1$ and $p^m < i - e_n < p^{m+1}$, then we set $\kappa_{n,m,i} = \omega_{n,m,m-\sigma(m+1,i)}$ with $\omega_{n,m,l}$ for $l \ge 0$ defined as in (4.2.1). Then $\kappa_{n,m,i} \in V_i$ by the claim in the proof of Theorem 4.2.2. Moreover, we have

$$\kappa_{n,m,i} = \sum_{k=\sigma(m+1,i)}^{m} \rho^k \vartheta_{n-m,m-k} T^{\theta_k(i)-1} u_{n,p-1} - v, \tag{4.3.1}$$

since $\theta_k(i) = p^{n-k-2}(p-1) + p^{m-k}$ if $k \ge \sigma(m+1,i)$.

Our next result is the analogue of Theorem 3.4.1 at the finite level.

Theorem 4.3.1. *Let μ be the smallest nonnegative integer for which $i \le \mu e_n + p^n$. Let*

$$S_{n,i} = \{p^{\mu}\kappa_{n,m,i-\mu e_n} \mid 0 \le m \le s\}, \quad \textit{where } s = \left\lceil \log_p\Bigl(\frac{i-\mu e_n+1}{r+1+\delta(p-1)}\Bigr)\right\rceil.$$

If $2 \le r \le p-2$, then the A_n-module $V_{n,i}$ is generated by $S_{n,i}$. If $r = p-1$, it is generated by $S_{n,i} \cup \{p^{\mu} v\}$ if $i \le (\mu+1)e_n$ and $S_{n,i}$ otherwise, and if $r = p$, it is generated by $S_{n,i} \cup \{p^{\mu+\lceil \log_p(i-\mu e_n)\rceil} w_n\}$.

Proof. Suppose first that $i \le p^n$. If $r \ne p-1$, then $V_{n,i} = N_n V_i$ by Proposition 4.1.6. For such i, the generation then follows immediately from Theorem 3.4.1.

Similarly, if $r = p-1$, then $v \in V_{n,e_n}(-\xi)$ generates the cokernel of N_n. If $i \le e_n$, then $S_{n,i} \cup \{v\}$ generates $V_{n,i}$ by a similar argument to that given in Theorem 3.4.1 (or by Proposition 4.1.6 and Theorem 3.4.1 itself). If $e_n < i < p^n$, then similarly $S_{n,i} \cup \{pv\}$ generates $V_{n,i}$, but we now claim that pv is in the A_n-submodule generated by $S_{n,i}$. To see this, suppose that $m \le n-2$ is such that $p^m < i - e_n < p^{m+1}$. Note that $A_n S_{n,i}$ contains $\nu_{n,k,i} = \rho^k T^{\theta_k(i)} u_{n,p-1}$ for each $0 \le k \le n-1$. (If $\kappa_{n,k,i}$ is not this element, one can multiply it by T and subtract off multiples of the $\nu_{n,h,i}$ for $h < k$ to reduce it to this form.) Noting (4.3.1), we have

$$\rho v = -\rho\kappa_{n,m,i} + \sum_{k=\sigma(m+1,i)}^{m} \vartheta_{n-m,m-k} T^{\theta_k(i)-\theta_{k+1}(i)-1} \nu_{n,k+1,i} \in A_n S_{n,i}.$$

In the case of arbitrary r and i, Lemma 4.1.5 tells us that $V_{n,i} = p^{\mu} V_{n,i-\mu e_n}$, and we again have the desired generation. □

Remark 4.3.2. For $i \le p^n$, the integer s in Theorem 4.3.1 is unique such that i lies in the half-open interval $[(r+1)p^{s-1}, (r+1)p^s)$ if $r \le p-1$ and $[2p^s, 2p^{s+1})$ if $r = p$. Since $S_{n,i}$ has $s+1$ elements, the generating set $S'_{n,i}$ provided in Theorem 4.3.1 has at most $n+1$ elements. Since $S'_{n,i} = p^{\mu} S'_{n,i-\mu e_n}$, the latter statement holds for all i. In fact, for $i > p^{n-1}$, the set $S'_{n,i}$ has either n or $n+1$ elements, depending for each r on which of two ranges i lies in modulo e_n.

Finally, we prove a slightly weaker minimality statement than Theorem 3.4.4, since in the finite case there are many values of i for which the analogous statement to Theorem 3.4.4 is simply not true, so long as $r \le p-1$.

Theorem 4.3.3. *Every generating subset of the generating set for $V_{n,i}$ of Theorem 4.3.1 is of cocardinality at most one.*

Proof. We maintain the notation of Theorem 4.3.1. By Lemma 4.1.5, the p^{μ}th power map defines an isomorphism $V_{i-\mu e_n} \xrightarrow{\sim} V_i$, and $S_{n,i} = p^{\mu} S_{n,i-\mu e_n}$. We therefore assume that $i \le p^n$ for the rest of the proof. Note that we have

$$\theta_k(i) \le p^{n-k-1} \tag{4.3.2}$$

for all $0 \le k \le n-1$, and we have $\theta_n(i) = 0$.

Case $r \le p-2$. In this case, N_n induces an isomorphism $D^{(r)}/f_n D^{(r)} \xrightarrow{\sim} D_n^{(r)}$, so Proposition 4.1.6 tells us that $V_{n,i} \cong V_i/(V_i \cap f_n D^{(r)})$. In other words, a subset Y_n of $S_{n,i}$ will generate $V_{n,i}$ if and only if the subset Y of S_i lifting it has the property that $Y \cup \{f_n u_r\}$ generates $V_i + f_n D^{(r)}$.

Recall that

$$f_n \equiv \sum_{k=0}^{n-1} p^k T^{p^{n-k-1}} \bmod (p^{n-1}T^2, p^{n-2}T^{2p}, \ldots, T^{2p^{n-1}}).$$

Noting (4.3.2), we have

$$f_n \equiv p^m T^{p^{n-m-1}} \bmod (p^{m+1}, T^{\theta_m(i)+1}) \tag{4.3.3}$$

for each $0 \le m \le s$. Let us set $I = (p, T, \varphi - 1)$ and $I_m = (p, T^{1+\epsilon_m(i)}, \varphi - 1)$ for the remainder of the proof.

The analogue of (3.4.3) in our current setting is

$$\sum_{m=0}^{s} c_m \kappa_{m,i} = b f_n u_r \tag{4.3.4}$$

for some $c_m \in A$ and $b \in A$. Given a solution to (4.3.4), we claim that there exist $q_k \in \mathbb{Z}_p$ for $k \le s$, independent of the solution, such that

$$c_k \equiv q_k b T^{\epsilon_k(i)} \bmod I_k. \tag{4.3.5}$$

Of course, only those $\kappa_{n,k,i}$ for k such that $p \nmid q_k$ and $\epsilon_k(i) = 0$ can possibly be A_n-linear combinations of the others. If k is such a value and we suppose that $c_k = 0$, then these congruences force $b \in I$ and therefore $c_m \in I$ for every other $m \le s$, proving the result.

We turn to the proof of the claim. In our current setting, (3.4.5) becomes

$$c_n \rho^n \equiv 0 \bmod (p^{n+1}, T)$$

for $m = n$ (if $s = n$, since $\theta_n(i) = 0$) and

$$c_m \rho^m T^{\psi_m(i)} \equiv \sum_{k \in X_m} c_k a_{k,i} \rho^m T^{\theta_m(i)-1} + b p^m T^{p^{n-m-1}} \bmod (p^{m+1}, T^{\theta_m(i)+1}) \tag{4.3.6}$$

for $m \le n-1$, with X_m as in (3.4.4). In the case that $s = n$, the claim for $k = n$ is then immediate. Moreover, supposing that we know the claim for k with $m+1 \le k \le s$, the congruence (4.3.6) implies that

$$c_m \equiv \sum_{k \in X_m} q_k a_{k,i} b T^{\epsilon_m(i)} + b T^{p^{n-m-1} - \psi_m(i)} \bmod I_m$$

upon application of (4.3.5) for $k \in X_m$. As $\epsilon_m(i) \le p^{n-m-1} - \psi_m(i)$ by (4.3.2), we have the claim for $k = m$ as well.

We remark that if $\theta_m(i) < p^{n-m-1}$ for all $m \le n-1$, which is to say that $i \le p^{n-1} r$, then we obtain recursively that $p \mid q_m$ for all $m \le s$. In other words, $S_{n,i}$ has no proper generating subset for such i. This is useful in the following case.

Case $r = p$. In the case $r = p$, we have $\theta_m(i) < p^{n-m-1}$ for all $m \le n-1$ and all $i \le p^n$ (since $\delta = 1$), and the analogous argument working modulo Aw and using the set X'_m of (3.4.6) shows that any subset of $S_{n,i} \cup \{p^{\lceil \log_p(i) \rceil} w_n\}$ that generates $V_{n,i}$ must contain $S_{n,i}$.

Case $r = p-1$. Finally, we consider the more subtle case that $r = p-1$. In this case, $s \le n-1$. Recall from Theorem 4.1.7 that

$$V_{p-1}/f_n V_{p-1} \cong A_n u_{n,p-1} \cong A_n/(N_{\Gamma_n \times \Phi})$$

and $A_n v = \mathbb{Z}_p v + \mathbb{Z}_p[\Phi] N_{\Gamma_n} u_{n,p-1}$. Note that N_{Γ_n} lifts to $T^{-1} f_n$ in A. As in (4.3.3), we have

$$T^{-1} f_n \equiv p^m T^{p^{n-m-1}-1} \bmod (p^{m+1}, T^{\theta_m(i)+1})$$

for $0 \le m \le n-2$ and

$$T^{-1} f_n \equiv p^{n-1}(1 - \tfrac{1}{2}T) \bmod (p^n, T^2). \tag{4.3.7}$$

Range $i \le e_n$: In this range, every $\kappa_{n,m,i}$ lies in $A_n u_r$, so v is in particular necessary to generate $V_{n,i}$. We also have $\theta_{n-1}(i) = 0$ and $\theta_m(i) \le e_{n-m-1}$ for all $m \le n-2$. Consider the following analogue of (3.4.3):

$$\sum_{m=0}^{s} c_m \kappa_{m,i} = bT^{-1} f_n u_{p-1}. \tag{4.3.8}$$

As before, we claim that there exist $q_k \in \mathbb{Z}_p$ for $k \le s$, independent of the solution to (4.3.8), such that (4.3.5) holds, from which the result follows in this range.

The analogue of (3.4.5) for $m \le s$ in the current setting is

$$c_m \rho^m T^{\psi_m(i)} \equiv \sum_{k \in X_m} c_k a_{k,i} \rho^m T^{\theta_m(i)-1} + bp^m T^{p^{n-m-1}-1} \bmod (p^{m+1}, T^{\theta_m(i)+1}). \tag{4.3.9}$$

If $s = n-1$, we then obtain $c_{n-1} \equiv b \bmod I$. If $s \ge n-2$, we have

$$c_{n-2} \equiv bT^{p-1-\theta_{n-2}(i)+\epsilon_{n-2}(i)} \bmod I_{n-2},$$

and hence the claim for $k = n-2$. For $m \le n-3$, we have $\theta_m(i) \le p^{n-m-1} - 2$, and assuming the claim for $m+1 \le k \le s$, we see recursively using (4.3.9) that

$$c_m \equiv \sum_{k \in X_m} q_k a_{k,i} bT^{\epsilon_m(i)} \bmod I_m.$$

Range $e_n < i < p^n$. In this range, $s = n-1$, $\theta_{n-1}(i) = 1$, and $\theta_{n-2}(i) = p$. Let $l \le n-2$ be such that $p^l < i - e_n < p^{l+1}$, so $\kappa_{n,l,i}$ is the lone element of $S_{n,i}$ that does not lie in $A_n u_{n,p-1}$. Thus, if we were to have

$$\sum_{m=0}^{n-1} d_m \kappa_{n,m,i} = 0 \tag{4.3.10}$$

for some $d_m \in A_n$, then we would have to have $d_l \in A_n(T, \varphi - 1)$ in order that $d_l \kappa_{n,l,i} \in A_n u_{n,p-1}$. Let

$$\kappa'_{l,i} = \sum_{j=\sigma(l+1,i)}^{l} \rho^j \vartheta_{n-l,l-j} T^{\theta_j(i)} u_{p-1}$$

so that $T\kappa_{n,l,i} = \varphi^n N_n \kappa'_{l,i}$. Let $\kappa'_{m,i} = \kappa_{m,i}$ for $m \le n-1$ with $m \ne l$.

Now (4.3.10) implies that

$$\sum_{m=0}^{n-1} c_m \kappa'_{m,i} \equiv bT^{-1} f_n u_{p-1} \bmod A(\varphi - 1)u_{p-1} \tag{4.3.11}$$

for some $b \in A$ and where $c_m \in A$ reduces to d_m for $m \neq l$ and $c_l \in A$ is such that Tc_l reduces to d_l modulo $A_n(\varphi - 1)$. Similarly to before, we claim that there exist $q_m \in \mathbb{Z}_p$ for $m \leq n-2$, independent of the solution to (4.3.11), such that (4.3.5) holds, and that $b \in I$ if and only if $c_{n-1} \in I$. From this, it follows that a solution to (4.3.11) with $c_k = 0$ for some k has $c_m \in I$ for every other $m \leq n-1$.

Note that $\epsilon_l(i) = 0$, and let τ_m be $\vartheta_{n-l,l-m}$ if $\sigma(l+1, i) \leq m < l$ and 0 otherwise. Equations (4.3.7) and (4.3.11) yield

$$c_{n-1} \equiv b(1 - \tfrac{1}{2}T) \bmod (p, T^2, \varphi - 1), \tag{4.3.12}$$

and, for arbitrary $m \leq n-2$, we have

$$c_m T^{\psi_m(i)} \equiv \sum_{k \in X_m} c_k a_{k,i} T^{\theta_m(i)-1} - \tau_m c_l T^{\theta_m(i)} + bT^{p^{n-m-1}-1} \bmod (p, T^{\theta_m(i)+1}, \varphi - 1). \tag{4.3.13}$$

For $m = n-2$, note that (4.3.12), (4.3.13), and $a_{n-1,i} = -1$ imply that

$$c_{n-2} T^{1-\epsilon_{n-2}(i)} \equiv b - c_{n-1} \equiv \tfrac{1}{2} bT \bmod (p, T^2, \varphi - 1), \tag{4.3.14}$$

so (4.3.5) holds with $q_{n-2} = \frac{1}{2}$. For m with $\sigma(n-1, i) \leq m \leq n-3$ (which exists only if $l = n-2$), we have $X_m = \{n-1\}$ and $\theta_m(i) = p^{n-m-1}$, and we obtain from (4.3.13) and (4.3.14) that

$$\begin{aligned} c_m T^{1-\epsilon_m(i)} &\equiv -c_{n-1} - c_{n-2}\tau_m T + b \\ &\equiv \tfrac{1}{2}(1 - \vartheta_{2,n-m-2}) bT \bmod (p, T^2, \varphi - 1), \end{aligned} \tag{4.3.15}$$

so (4.3.5) holds with $q_m = -\frac{1}{2}(n-m-2)$. For $m < \sigma(n-1, i)$, we have $\theta_m(i) < p^{n-m-1}$, and (4.3.13) and (4.3.14) yield recursively that

$$c_m \equiv \sum_{k \in X_m} q_k a_{k,i} bT^{\epsilon_m(i)} - q_l \tau_m bT^{\epsilon_m(i)} + bT^{p^{n-m-1}-1-\psi_m(i)} \bmod I_m,$$

verifying (4.3.5) for $k = m$. □

Acknowledgments

The idea for this paper originated with my 1999 PhD thesis, and initial computations were performed on an evening in June 2001 during a visit to the University of Nottingham. I thank Ivan Fesenko for his hospitality. I wrote a short draft of the paper in 2002 and made additions to it in August 2006. The paper tripled in size as I brought it to near final form in the summer of 2011. I also thank Richard Gottesman for his interest in this work, which inspired me to finish the paper.

References

[Greither 1996] C. Greither, "On Chinburg's second conjecture for abelian fields", *J. Reine Angew. Math.* **479** (1996), 1–37. MR 97e:11139 Zbl 0856.11051

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics **67**, Springer, New York, 1979. MR 82e:12016 Zbl 0423.12016

[Sharifi 2002] R. T. Sharifi, "Determination of conductors from Galois module structure", *Math. Z.* **241**:2 (2002), 227–245. MR 2003h:11147 Zbl 1017.11058

[Wintenberger 1983] J.-P. Wintenberger, "Le corps des normes de certaines extensions infinies de corps locaux; applications", *Ann. Sci. École Norm. Sup.* (4) **16**:1 (1983), 59–89. MR 85e:11098 Zbl 0516.12015

Communicated by John H. Coates
Received 2011-08-20 Revised 2011-11-29 Accepted 2012-02-20

sharifi@math.arizona.edu
Department of Mathematics, University of Arizona, 617 N. Santa Rita Ave, PO Box 210089, Tucson AZ 85721-0089, United States
http://math.arizona.edu/~sharifi

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.193

On the invariant theory for tame tilted algebras

Calin Chindris

We show that a tilted algebra A is tame if and only if for each generic root $\boldsymbol{d}$ of A and each indecomposable irreducible component C of $\operatorname{mod}(A, \boldsymbol{d})$, the field of rational invariants $k(C)^{\mathrm{GL}(\boldsymbol{d})}$ is isomorphic to k or $k(x)$. Next, we show that the tame tilted algebras are precisely those tilted algebras A with the property that for each generic root $\boldsymbol{d}$ of A and each indecomposable irreducible component $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, the moduli space $\mathcal{M}(C)^{ss}_{\theta}$ is either a point or just $\mathbb{P}^1$ whenever θ is an integral weight for which $C^s_{\theta} \neq \varnothing$. We furthermore show that the tameness of a tilted algebra is equivalent to the moduli space $\mathcal{M}(C)^{ss}_{\theta}$ being smooth for each generic root $\boldsymbol{d}$ of A, each indecomposable irreducible component $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, and each integral weight θ for which $C^s_{\theta} \neq \varnothing$. As a consequence of this latter description, we show that the smoothness of the various moduli spaces of modules for a strongly simply connected algebra A implies the tameness of A.

Along the way, we explain how moduli spaces of modules for finite-dimensional algebras behave with respect to tilting functors, and to theta-stable decompositions.

1. Introduction

Throughout this paper, we work over an algebraically closed field k of characteristic zero. All algebras (associative and with identity) are assumed to be finite-dimensional over k, and all modules are assumed to be finite-dimensional left modules.

One of the fundamental problems in the representation theory of algebras is that of classifying the indecomposable modules. Based on the complexity of the indecomposable modules, one distinguishes the class of tame algebras and that of wild algebras. According to the remarkable Tame-Wild Dichotomy Theorem of Drozd [1979], these two classes of algebras are disjoint and they cover the whole class of

The author was partially supported by NSF grant DMS-1101383.
MSC2010: primary 16G10; secondary 16R30, 16G60, 16G20.
Keywords: exceptional sequences, moduli spaces, rational invariants, tame and wild algebras, tilting.

algebras. Since the representation theory of a wild algebra is at least as complicated as that of a free algebra in two variables, and since the latter theory is known to be undecidable, one can hope to meaningfully classify the indecomposable modules only for tame algebras. For more precise definitions, see [Simson and Skowroński 2007, Chapter XIX] and the reference therein.

An interesting task in the representation theory of algebras is to study the geometry of affine varieties of modules of fixed dimension vectors and the actions of the corresponding products of general linear groups associated to a given finite-dimensional algebra A over k. In particular, it would be interesting to find characterizations of prominent classes of tame algebras via geometric properties of their module varieties. This research direction has attracted much attention during the last two decades; see for example [Bobiński 2008; Bobiński and Skowroński 1999a; 1999b; 2002; Geiss and Schröer 2003; Riedtmann 2004; Riedtmann and Zwara 2004; 2008; Skowroński and Weyman 2000].

In this paper, we seek for characterizations of tame algebras in terms of invariant theory. A first result in this direction was obtained by Skowroński and Weyman [2000, Theorem 1], who showed that a finite-dimensional algebra of global dimension one is tame if and only if all of its algebras of semiinvariants are complete intersections. Unfortunately, this result does not extend to algebras of higher global dimension (not even of global dimension two), as shown by Kraśkiewicz [2001]. As was suggested by Weyman, in order to characterize the tameness of an algebra via invariant theory, one should impose geometric conditions on the various moduli spaces of semistable modules rather than on the entire algebras of semiinvariants.

In [Chindris 2011], the author has found a description of the tameness of path algebras and of canonical algebras in terms of the invariant theory of the algebras in question; see also [Domokos 2011]. In this paper, we continue this line of inquiry for the class of tilted algebras. Recall that a tilted algebra is an algebra of the form $\mathrm{End}_H(T)$, where H is a connected finite-dimensional hereditary algebra and T is a multiplicity-free tilting H-module, that is, $\mathrm{Ext}^1_H(T, T) = 0$ and T is the direct sum of n pairwise nonisomorphic indecomposable modules with n the rank of the Grothendieck group $K_0(H)$ of H. It has been proved by Kerner [1989, Theorem 6.2] that a tilted algebra A is tame if and only if its Tits quadratic form q_A is weakly nonnegative (takes nonnegative values on nonnegative vectors).

Theorem 1.1. *Let A be a tilted algebra. Then the following conditions are equivalent*:

(1) *A is tame*;

(2) *for each generic root $\mathbf{d}$ of A and each indecomposable irreducible component C of* $\mathrm{mod}(A, \boldsymbol{d})$, *we have $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k$ or $k(x)$*;

(3) *for each generic root* $\boldsymbol{d}$ *of* A *and each indecomposable irreducible component* $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, *the moduli space* $\mathcal{M}(C)^{ss}_{\theta}$ *is either a point or* $\mathbb{P}^1$ *whenever* θ *is an integral weight of* A *for which* $C^s_{\theta} \neq \varnothing$;

(4) *for each generic root* $\boldsymbol{d}$ *of* A *and each indecomposable irreducible component* $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, *the moduli space* $\mathcal{M}(C)^{ss}_{\theta}$ *is smooth whenever* θ *is an integral weight of* A *for which* $C^s_{\theta} \neq \varnothing$.

Following [Skowroński 1993], a triangular algebra A is called strongly simply connected if the first Hochschild cohomology space $HH^1(C)$ of any convex subcategory C of A vanishes. It has been recently proved by Brüstle, de la Peña, and Skowroński [Brüstle et al. 2011, Main Theorem] that a strongly simply connected algebra A is tame if and only if its Tits form q_A is weakly nonnegative. As a consequence of Theorem 1.1 and another tameness criterion from [ibid., Corollary 1], we derive the following sufficient geometric criterion for the tameness of a strongly simply connected algebra:

Proposition 1.2. *Let* A *be a strongly simply connected algebra. Assume for each generic root* $\boldsymbol{d}$ *of* A, *each indecomposable irreducible component* $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, *and each integral weight* θ *for which* $C^s_{\theta} \neq \varnothing$, *that* $\mathcal{M}(C)^{ss}_{\theta}$ *is a smooth variety. Then*, A *is a tame algebra.*

We would like to point out that the equivalence of (1) and (3) in Theorem 1.1 settles in the affirmative a conjecture of Weyman for the class of tilted algebras, while Proposition 1.2 proves one implication of Weyman's conjecture for the class of strongly simply connected algebras (for more details, see Remark 4).

Our next theorem, which is key in proving Theorem 1.1 and Proposition 1.2, identifies integral weights of an algebra for which the corresponding moduli spaces of semistable modules are preserved under titling. Our next theorem generalizes [Domokos and Lenzing 2000, Theorem 6.3] to arbitrary bound quiver algebras. (The details of our notation can be found in Section 3B.)

Theorem 1.3. *Let* $A = kQ/I$ *be a bound quiver algebra*, T *a basic tilting* A*-module*, *and* θ *an integral weight of* A *that is well positioned with respect to* T. *Let* F *be either the functor* $\operatorname{Hom}_A(T, \_)$, *in case there are nonzero* θ*-semistable torsion* A*-modules*, *or the functor* $\operatorname{Ext}^1_A(T, \_)$, *in case there are nonzero* θ*-semistable torsion-free* A*-modules. Denote the algebra* $\operatorname{End}_A(T)^{op}$ *by* B *and let* $u : K_0(A) \to K_0(B)$ *be the isometry induced by the tilting module* T. *Then*,

(a) *the functor* F *defines an equivalence of categories between* $\operatorname{mod}(A)^{ss}_{\theta}$ *and* $\operatorname{mod}(B)^{ss}_{\theta'}$, *where* $\theta' = |\theta \circ u^{-1}|$; *and*

(b) *the bijective map* $f : \mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta} \to \mathcal{M}(B, \boldsymbol{d}')^{ss}_{\theta'}$ *induced by* F *is an isomorphism of algebraic varieties*, *where* $\boldsymbol{d}$ *is a* θ*-semistable dimension vector of* A *and* $\boldsymbol{d}' = u(\boldsymbol{d})$.

In particular, this theorem allows us to transfer much of the geometry of A over to that of B; see for example Proposition 4.1.

It is natural to ask if the description of the fields of rational invariants and of the moduli spaces in Theorem 1.1 can be extended to irreducible components that are not necessarily indecomposable. To answer this question, we rely on two general reduction results. The first such result has been recently proved in [Chindris 2011, Proposition 4.7] and allows one to compute fields of rational invariants on irreducible components by reducing the considerations to the case where the irreducible components involved are indecomposable. For the second general reduction result, the starting point is Derksen and Weyman's notion [2011] of θ-stable decomposition of representation spaces for quivers without oriented cycles. Here, we first extend their notion to irreducible components of module varieties, and then explain how to extend [Derksen and Weyman 2011, Theorem 3.20] to arbitrary bound quiver algebras:

Theorem 1.4. *Let $A = kQ/I$ be a bound quiver algebra and let $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$ be a θ-well-behaved irreducible component, where θ is an integral weight of A. Let*

$$C = m_1 \cdot C_1 \dotplus \cdots \dotplus m_n \cdot C_n$$

be the θ-stable decomposition of C, where $C_i \subseteq \operatorname{mod}(A, \boldsymbol{d}_i)$ with $1 \leq i \leq n$ are θ-stable irreducible components, and $\boldsymbol{d}_i \neq \boldsymbol{d}_j$ for all $1 \leq i \neq j \leq n$. Assume that

(1) *C contains the image of $X := C_1^{m_1} \times \cdots \times C_n^{m_n}$ through the natural (diagonal) embedding $\mathcal{V} := \operatorname{mod}(Q, \boldsymbol{d}_1)^{m_1} \times \cdots \times \operatorname{mod}(Q, \boldsymbol{d}_n)^{m_n} \hookrightarrow \operatorname{mod}(Q, \boldsymbol{d})$; and*

(2) *C is a normal variety.*

Then $\mathcal{M}(C)^{ss}_\theta \cong S^{m_1}(\mathcal{M}(C_1)^{ss}_\theta) \times \cdots \times S^{m_n}(\mathcal{M}(C_n)^{ss}_\theta)$.

Note that this reduction result allows us to "break" a moduli space of modules into smaller ones that are typically easier to handle; see Section 3C.

Recall that a quasitilted algebra is a basic and connected finite-dimensional algebra of the form $\operatorname{End}_{\mathcal{H}}(T)^{\mathrm{op}}$, where $\mathcal{H}$ is a hereditary category and $T \in \mathcal{H}$ is a tilting object. In [Happel et al. 1996, Theorem 2.3], Happel, Reiten, and Smalø proved that an algebra A is quasitilted if and only if A is of global dimension at most two and every indecomposable finite-dimensional A-module X has projective dimension or injective dimension at most one. It was shown by Skowroński [1998, Theorem A] that a quasitilted algebra A is tame if and only if its Tits form q_A is weakly nonnegative.

Using our results described above, we can prove this:

Proposition 1.5. *Let $A = kQ/I$ be a tame quasitilted algebra, $\boldsymbol{d}$ a dimension vector of A, and C an irreducible component of $\operatorname{mod}(A, \boldsymbol{d})$.*

(1) *The field of rational invariants satisfies* $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k(x_1, \ldots, x_N)$, *where* N *is the sum of the multiplicities of the isotropic imaginary roots that occur in the generic decomposition of* $\boldsymbol{d}$ *in* C.

(2) *If* $\boldsymbol{d}$ *is an isotropic root of* A, *then the moduli spaces* $\mathcal{M}(C)^{ss}_{\theta}$ *for* $\theta \in \mathbb{Z}^{Q_0}$ *are products of projective spaces.*

Our proof of Proposition 1.5(1) provides another approach to proving [Domokos and Lenzing 2002, Corollary 7.4].

The layout of this paper is as follows. In Section 2, we recall some background material on irreducible components of module varieties and their rational invariants. In Section 3, we first review King's construction of moduli spaces of modules for algebras, and then prove Theorem 1.3 in Section 3B. In Section 3C, we first explain how to extend Derksen and Weyman's notion [2011] of θ-stable decomposition to quivers with relations, and then prove Theorem 1.4. We prove Theorem 1.1 and Proposition 1.5 in Section 4.

2. Background on module varieties

Let $Q = (Q_0, Q_1, t, h)$ be a finite quiver with vertex set Q_0 and arrow set Q_1. The two functions $t, h : Q_1 \to Q_0$ assign to each arrow $a \in Q_1$ its tail ta and head ha, respectively.

A representation V of Q over k is a collection $(V(i), V(a))_{i \in Q_0, a \in Q_1}$ of finite-dimensional k-vector spaces $V(i)$, $i \in Q_0$, and k-linear maps

$$V(a) \in \mathrm{Hom}_k(V(ta), V(ha)) \quad \text{for } a \in Q_1.$$

The dimension vector of a representation V of Q is the function $\mathbf{dim}\, V : Q_0 \to \mathbb{Z}$ defined by $(\mathbf{dim}\, V)(i) = \dim_k V(i)$ for $i \in Q_0$. Let S_i be the one-dimensional representation of Q at vertex $i \in Q_0$, and let us denote by $\boldsymbol{e}_i$ its dimension vector. By a dimension vector of Q, we simply mean a function $\boldsymbol{d} \in \mathbb{Z}^{Q_0}_{\geq 0}$.

Given two representations V and W of Q, we define a morphism $\varphi : V \to W$ to be a collection $(\varphi(i))_{i \in Q_0}$ of k-linear maps with $\varphi(i) \in \mathrm{Hom}_k(V(i), W(i))$ for each $i \in Q_0$, and such that $\varphi(ha)V(a) = W(a)\varphi(ta)$ for each $a \in Q_1$. We denote by $\mathrm{Hom}_Q(V, W)$ the k-vector space of all morphisms from V to W. Let V and W be two representations of Q. We say that V is a subrepresentation of W if $V(i)$ is a subspace of $W(i)$ for each $i \in Q_0$ and $V(a)$ is the restriction of $W(a)$ to $V(ta)$ for each $a \in Q_1$. In this way, we obtain the abelian category $\mathrm{rep}(Q)$ of all representations of Q.

Given a quiver Q, its path algebra kQ has a k-basis consisting of all paths (including the trivial ones), and the multiplication in kQ is given by concatenation of paths. It is easy to see that any kQ-module defines a representation of Q, and vice-versa. Furthermore, the category $\mathrm{mod}(kQ)$ of kQ-modules is equivalent to the category rep(Q). In what follows, we identify mod(kQ) and rep(Q), and use the same notation for a module and the corresponding representation.

A two-sided ideal I of kQ is said to be *admissible* if there exists an integer $L \geq 2$ such that $R_Q^L \subseteq I \subseteq R_Q^2$. Here, R_Q denotes the two-sided ideal of kQ generated by all arrows of Q.

If I is an admissible ideal of KQ, the pair (Q, I) is called a *bound quiver* and the quotient algebra kQ/I is called the *bound quiver algebra* of (Q, I). Any admissible ideal is generated by finitely many admissible relations, and any bound quiver algebra is finite-dimensional and basic. Moreover, a bound quiver algebra kQ/I is connected if and only if (the underlying graph of) Q is connected; see for example [Assem et al. 2006].

It is well known that any basic algebra A is isomorphic to the bound quiver algebra of a bound quiver (Q_A, I), where Q_A is the Gabriel quiver of A; see [Assem et al. 2006]. (Note that the ideal of relations I is not uniquely determined by A.) We say that A is a *triangular* algebra if its Gabriel quiver has no oriented cycles.

Fix a bound quiver (Q, I) and let $A = kQ/I$ be its bound quiver algebra. We denote by e_i the primitive idempotent corresponding to the vertex $i \in Q_0$. A representation M of a A (or (Q, I)) is just a representation M of Q such that $M(r) = 0$ for all $r \in I$. The category mod(A) of finite-dimensional left A-modules is equivalent to the category rep(A) of representations of A. As before, we identify mod(A) and rep(A), and make no distinction between A-modules and representations of A.

Assume from now on that A has finite global dimension; this happens, for example, when Q has no oriented cycles. The Ringel form of A is the bilinear form $\langle \cdot, \cdot \rangle_A : \mathbb{Z}^{Q_0} \times \mathbb{Z}^{Q_0} \to \mathbb{Z}$ defined by

$$\langle \boldsymbol{d}, \boldsymbol{e} \rangle_A = \sum_{l \geq 0} (-1)^l \sum_{i,j \in Q_0} \dim_k \operatorname{Ext}_A^l(S_i, S_j) \boldsymbol{d}(i) \boldsymbol{e}(j).$$

Note that if M is a $\boldsymbol{d}$-dimensional A-module and N is an $\boldsymbol{e}$-dimensional A-module, then

$$\langle \boldsymbol{d}, \boldsymbol{e} \rangle_A = \sum_{l \geq 0} (-1)^l \dim_k \operatorname{Ext}_A^l(M, N).$$

The quadratic form induced by $\langle \cdot, \cdot \rangle_A$ is denoted by χ_A.

The *Tits form* of A is the integral quadratic form $q_A : \mathbb{Z}^{Q_0} \to \mathbb{Z}$ defined by

$$q_A(\boldsymbol{d}) := \sum_{i \in Q_0} \boldsymbol{d}^2(i) - \sum_{i,j \in Q_0} \dim_k \operatorname{Ext}_A^1(S_i, S_j) \boldsymbol{d}(i) \boldsymbol{d}(j) + \sum_{i,j \in Q_0} \dim_k \operatorname{Ext}_A^2(S_i, S_j) \boldsymbol{d}(i) \boldsymbol{d}(j).$$

If A is triangular, then $r(i, j) := |R \cap e_j \langle R \rangle e_i|$ is precisely $\dim_k \operatorname{Ext}^2_A(S_i, S_j)$, for all $i, j \in Q_0$, as shown by Bongartz [1983]. So, in the triangular case, we can write

$$q_A(\boldsymbol{d}) = \sum_{i \in Q_0} \boldsymbol{d}^2(i) - \sum_{a \in Q_1} \boldsymbol{d}(ta)\boldsymbol{d}(ha) + \sum_{i,j \in Q_0} r(i, j)\boldsymbol{d}(i)\boldsymbol{d}(j).$$

2A. ***The generic decomposition for irreducible components.*** Let $\boldsymbol{d}$ be a dimension vector of A (or equivalently, of Q). The variety of $\boldsymbol{d}$-dimensional A-modules is the affine variety

$$\operatorname{mod}(A, \boldsymbol{d}) = \Big\{ M \in \prod_{a \in Q_1} \operatorname{Mat}_{\boldsymbol{d}(ha) \times \boldsymbol{d}(ta)}(k) \;\Big|\; M(r) = 0 \text{ for all } r \in I \Big\}.$$

It is clear that $\operatorname{mod}(A, \boldsymbol{d})$ is a $\operatorname{GL}(\boldsymbol{d})$-invariant closed subset of the affine space $\operatorname{mod}(Q, \boldsymbol{d}) := \prod_{a \in Q_1} \operatorname{Mat}_{\boldsymbol{d}(ha) \times \boldsymbol{d}(ta)}(k)$. Note that $\operatorname{mod}(A, \boldsymbol{d})$ does not have to be irreducible. We call $\operatorname{mod}(A, \boldsymbol{d})$ the *module variety* of $\boldsymbol{d}$-dimensional A-modules. We also denote by $\operatorname{ind}(A, \boldsymbol{d})$ the (possibly empty) constructible subset of all indecomposable modules in $\operatorname{mod}(A, \boldsymbol{d})$.

Let C be an irreducible component of $\operatorname{mod}(A, \boldsymbol{d})$. We say that C is *indecomposable* if C has a nonempty open subset of indecomposable modules. We call C a *Schur* irreducible component if C contains a Schur A-module. (Recall that a Schur A-module is just an A-module M such that $\operatorname{End}_A(M) \simeq k$.) Note that a Schur irreducible component is always indecomposable. The converse is always true for path algebras of quivers without oriented cycles. Finally, we say that $\boldsymbol{d}$ is a *generic root* of A if $\operatorname{mod}(A, \boldsymbol{d})$ has an indecomposable irreducible component.

Let us consider a decomposition $\boldsymbol{d} = \boldsymbol{d}_1 + \cdots + \boldsymbol{d}_t$, where $\boldsymbol{d}_i \in \mathbb{Z}^{Q_0}_{\geq 0}$ for $1 \leq i \leq t$. If C_i is a $\operatorname{GL}(\boldsymbol{d}_i)$-invariant subset of $\operatorname{mod}(A, \boldsymbol{d}_i)$ for $1 \leq i \leq t$, we denote by $C_1 \oplus \cdots \oplus C_t$ the constructible subset of $\operatorname{mod}(A, \boldsymbol{d})$ consisting of all modules isomorphic to direct sums of the form $\bigoplus_{i=1}^t X_i$ with $X_i \in C_i$ for all $1 \leq i \leq t$.

As shown by de la Peña [1991, Section 1.3] and Crawley-Boevey and Schröer [2002, Theorem 1.1], if C is an irreducible component of $\operatorname{mod}(A, \boldsymbol{d})$, then there are unique generic roots $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_t$ of A such that $\boldsymbol{d} = \boldsymbol{d}_1 + \cdots + \boldsymbol{d}_t$ and

$$C = \overline{C_1 \oplus \cdots \oplus C_t}$$

for some indecomposable irreducible components C_i of $\operatorname{mod}(A, \boldsymbol{d}_i)$ for $1 \leq i \leq t$. Also, the indecomposable irreducible components C_i for $1 \leq i \leq t$ are uniquely determined by this property. We call $\boldsymbol{d} = \boldsymbol{d}_1 \oplus \cdots \oplus \boldsymbol{d}_t$ the generic decomposition of $\boldsymbol{d}$ in C, and $C = \overline{C_1 \oplus \cdots \oplus C_t}$ the generic decomposition of C.

Recall that for an irreducible component $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, the field of rational $\operatorname{GL}(\boldsymbol{d})$-invariants on C is

$$k(C)^{\operatorname{GL}(\boldsymbol{d})} = \{\phi \in k(C) \mid g \cdot \phi = \phi \text{ for all } g \in \operatorname{GL}(\boldsymbol{d})\}.$$

In what follows, if R is an integral domain, we denote its field of fractions by $\mathrm{Quot}(R)$. Moreover, if K/k is a field extension and m is a positive integer, we define $S^m(K/k)$ to be the field $(\mathrm{Quot}(K^{\otimes m}))^{S_m}$, which is in fact the same as $\mathrm{Quot}((K^{\otimes m})^{S_m})$, since S_m is a finite group.

Proposition 2.1 [Chindris 2011, Proposition 4.7]. *Assume that the generic decomposition of C is of the form*

$$C = \overline{C_1^{\oplus m_1} \oplus \cdots \oplus C_n^{\oplus m_n}},$$

where $C_i \subseteq \mathrm{mod}(A, \boldsymbol{d}_i)$ for $1 \le i \le n$ are indecomposable irreducible components, $m_1, \dots, m_n$ are positive integers, and $\boldsymbol{d}_i \neq \boldsymbol{d}_j$ for all $1 \le i \neq j \le n$. Then

$$k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq \mathrm{Quot}\Big(\bigotimes_{i=1}^{n} S^{m_i}(k(C_i)^{\mathrm{GL}(\boldsymbol{d}_i)}/k)\Big).$$

In the next section, we present a homological method for studying fields of rational invariants on indecomposable irreducible components in module varieties.

2B. ***Exceptional sequences and rational invariants.*** Recall that a sequence $\mathscr{E} = (E_1, \dots, E_t)$ of A-modules is called an *orthogonal exceptional sequence* if the following conditions are satisfied:

(1) E_i is an exceptional A-module, that is, $\mathrm{End}_A(E_i) = k$ and $\mathrm{Ext}^l_A(E_i, E_i) = 0$ for all $l \ge 1$ and $1 \le i \le t$.

(2) $\mathrm{Ext}^l_A(E_i, E_j) = 0$ for all $l \ge 0$ and $1 \le i < j \le t$.

(3) $\mathrm{Hom}_A(E_j, E_i) = 0$ for all $1 \le i < j \le t$.

Given an orthogonal exceptional sequence $\mathscr{E}$, consider the full subcategory $\mathrm{filt}_{\mathscr{E}}$ of $\mathrm{mod}(A)$ whose objects M have a finite filtration $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_s = M$ of submodules such that each factor M_j/M_{j-1} is isomorphic to one of the $E_1, \dots, E_t$. For a dimension vector $\boldsymbol{d}$ of A, we define

$$\mathrm{filt}_{\mathscr{E}}(\boldsymbol{d}) = \{M \in \mathrm{mod}(A, \boldsymbol{d}) \mid M \text{ is isomorphic to a module in } \mathrm{filt}_{\mathscr{E}}\}.$$

We will be especially interested in short orthogonal exceptional sequences. As a first step in proving the rationality of fields of rational invariants for A, we will use the following direct consequence of the reduction theorem [Chindris 2011, Theorem 1.2]:

Proposition 2.2. *Let $\boldsymbol{d}$ be a generic root of A and let $C \subseteq \mathrm{mod}(A, \boldsymbol{d})$ be an indecomposable irreducible component. Assume that there exists an orthogonal exceptional sequence $\mathscr{E} = (E_1, E_2)$ of A-modules such that $\boldsymbol{d} = \mathbf{dim}\, E_1 + \mathbf{dim}\, E_2$, $\mathrm{filt}_{\mathscr{E}}(\boldsymbol{d}) \cap C \neq \varnothing$, and $\dim \mathrm{Ext}^2_A(E_2, E_1) = 0$. Then $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k(x_1, \dots, x_{n-1})$ where $n = \dim_k \mathrm{Ext}^1_A(E_2, E_1)$.*

Proof. The triangular algebra $A_{\mathscr{C}}$ that arises from the (minimal) A_∞-algebra structure of the Yoneda algebra $\mathrm{Ext}^\bullet_A(E_1 \oplus E_2, E_1 \oplus E_2)$ is precisely the path algebra of the generalized Kronecker quiver, K_n, with two vertices and n arrows, all pointing in the same direction. It now follows from [Chindris 2011, Theorem 1.2] that $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k(\mathrm{mod}(K_n, (1, 1)))^{\mathrm{GL}((1,1))} \simeq k(x_1, x_2, \dots x_{n-1})$. □

3. Moduli spaces of modules

Let $A = kQ/I$ be a bound quiver algebra and let $\boldsymbol{d} \in \mathbb{Z}^{Q_0}_{\geq 0}$ be a dimension vector of A. We denote $\mathrm{GL}(\boldsymbol{d})/T_1$ by $\mathrm{PGL}(\boldsymbol{d})$, where $T_1 = \{(\lambda\, \mathrm{Id}_{\boldsymbol{d}(i)})_{i \in Q_0} \mid \lambda \in k^*\} \leq \mathrm{GL}(\boldsymbol{d})$. Note that there is a well-defined action of $\mathrm{PGL}(\boldsymbol{d})$ on $\mathrm{mod}(A, \boldsymbol{d})$ since T_1 acts trivially on $\mathrm{mod}(A, \boldsymbol{d})$.

We always identify $K_0(A)$ with the lattice $\mathbb{Z}^{Q_0}$, which, in turn, we identify with $\mathrm{Hom}_{\mathbb{Z}}(K_0(A), \mathbb{Z})$ via $\theta(\boldsymbol{d}) = \sum_{i \in Q_0} \theta(i)\boldsymbol{d}(i)$ for all $\theta \in \mathbb{Z}^{Q_0}$ and $\boldsymbol{d} \in \mathbb{Z}^{Q_0}$. Note that when A is triangular, any integral weight $\theta \in \mathbb{Z}^{Q_0}$ can be written as $\langle \boldsymbol{d}, \cdot \rangle_A$ for a unique vector $\boldsymbol{d} \in \mathbb{Z}^{Q_0}$. Similarly, θ can be written as $\langle \cdot, \boldsymbol{e} \rangle_A$ for a unique vector $\boldsymbol{e} \in \mathbb{Z}^{Q_0}$.

Note that any $\theta \in \mathbb{Z}^{Q_0}$ defines a rational character $\chi_\theta : \mathrm{GL}(\boldsymbol{d}) \to k^*$ by

$$\chi_\theta\big((g(i))_{i \in Q_0}\big) = \prod_{i \in Q_0} (\det g(i))^{\theta(i)}.$$

In this way, we can identify $\mathbb{Z}^{Q_0}$ with the group $X^\star(\mathrm{GL}(\boldsymbol{d}))$ of rational characters of $\mathrm{GL}(\boldsymbol{d})$, assuming that $\boldsymbol{d}$ is a sincere dimension vector. In general, we have only the natural epimorphism $\mathbb{Z}^{Q_0} \to X^*(\mathrm{GL}(\boldsymbol{d}))$.

Now, let $\theta \in \mathbb{Z}^{Q_0}$ be an integral weight of A. Following King [1994], an A-module M is said to be *θ-semistable* if $\theta(\mathbf{dim}\, M) = 0$ and $\theta(\mathbf{dim}\, M') \leq 0$ for all submodules $M' \leq M$. We say that M is *θ-stable* if M is nonzero, $\theta(\mathbf{dim}\, M) = 0$, and $\theta(\mathbf{dim}\, M') < 0$ for all submodules $\{0\} \neq M' < M$. Now, consider the (possibly empty) open subsets

$$\mathrm{mod}(A, \boldsymbol{d})^{ss}_\theta = \{M \in \mathrm{mod}(A, \boldsymbol{d}) \mid M \text{ is } \theta\text{-semistable}\}$$

and

$$\mathrm{mod}(A, \boldsymbol{d})^{s}_\theta = \{M \in \mathrm{mod}(A, \boldsymbol{d}) \mid M \text{ is } \theta\text{-stable}\}$$

of $\boldsymbol{d}$-dimensional θ(-semi)-stable A-modules.

The weight space of semiinvariants on $\mathrm{mod}(A, \boldsymbol{d})$ of weight $n\theta \in \mathbb{Z}^{Q_0}$, where $n \in \mathbb{Z}_{\geq 0}$, is

$$\mathrm{SI}(A, \boldsymbol{d})_{n\theta} := \{f \in k[\mathrm{mod}(A, \boldsymbol{d})] \mid g \cdot f = (n\theta)(g) f \text{ for all } g \in \mathrm{GL}(\boldsymbol{d})\}.$$

Using methods from GIT, King [1994] showed that the projective variety

$$\mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta} := \operatorname{Proj}\Big(\bigoplus_{n\geq 0} \operatorname{SI}(A, \boldsymbol{d})_{n\theta}\Big)$$

is a GIT-quotient of $\operatorname{mod}(A, \boldsymbol{d})^{ss}_{\theta}$ by the action of $\operatorname{PGL}(\boldsymbol{d})$. We say that $\boldsymbol{d}$ is a *θ-semistable dimension vector* if $\operatorname{mod}(A, \boldsymbol{d})^{ss}_{\theta} \neq \varnothing$.

For an irreducible component $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$, we similarly define C^{ss}_{θ}, C^{s}_{θ}, $\operatorname{SI}(C)_{n\theta}$, and $\mathcal{M}(C)^{ss}_{\theta}$.

3A. *Families of A-modules.* Let us denote by $\operatorname{mod}(A)^{ss}_{\theta}$ the full subcategory of $\operatorname{mod}(A)$ consisting of the θ-semistable modules. It is easy to see that $\operatorname{mod}(A)^{ss}_{\theta}$ is a full exact abelian subcategory of $\operatorname{mod}(A)$ that is closed under extensions and whose simple objects are precisely the θ-stable modules. Moreover, $\operatorname{mod}(A)^{ss}_{\theta}$ is Artinian and Noetherian, and hence every θ-semistable A-module has a Jordan–Hölder filtration in $\operatorname{mod}(A)^{ss}_{\theta}$.

Two θ-semistable A-modules are said to be *S-equivalent* if they have the same composition factors in $\operatorname{mod}(A)^{ss}_{\theta}$. It was proved in [King 1994, Proposition 4.2] that the points of $\mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta}$ are in one-to-one correspondence with the S-equivalence classes of $\boldsymbol{d}$-dimensional θ-semistable A-modules.

We now recall the definition of a *family of A-modules* over a variety that was introduced in this context by King [1994]. Let Z be a (reduced) variety and let $(V_z)_{z\in Z}$ be a collection of A-modules parametrized by Z. Following the presentation in [Domokos and Lenzing 2000, Section 6], we call $(V_z)_{z\in Z}$ a *family of A-modules* if the following two conditions are satisfied:

(i) $(V_z)_{z\in Z}$ is an algebraic vector bundle over Z; in particular, the vector spaces V_z for $z \in Z$ have the same dimension.

(ii) For each $a \in A$, the map $z \to a \cdot \operatorname{Id}_{V_z}$ $(z \in Z)$ is a section of the endomorphism bundle $(\operatorname{End}_k(V_z))_{z\in Z}$; in other words, the A-module structure on V_z varies algebraically with $z \in Z$.

King showed that $\mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta}$ is a *coarse moduli space* for families of $\boldsymbol{d}$-dimensional θ-semistable A-modules; see [King 1994, Proposition 5.2]. This essentially says that if $(V_z)_{z\in Z}$ is a family of $\boldsymbol{d}$-dimensional θ-semistable A-modules and ϕ is the (unique) set-theoretic map $Z \to \mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta}$ that sends each $z \in Z$ to the point representing the S-equivalence class of V_z, then ϕ is a morphism of varieties.

Lemma 3.1. *Let A and B be two bound quiver algebras, T an A-B-bimodule, Z a variety, and n a positive integer.*

(1) *Let $(V_z)_{z\in Z}$ be a family of A-modules parametrized by Z. Assume that for each $0 \leq l \leq n$, there exists an integer m_l such that* $\dim_k \operatorname{Ext}^l_A(T, V_z) = m_l$ *for all $z \in Z$. Then* $(\operatorname{Ext}^n_A(T, V_z))_{z\in Z}$ *is a family of B-modules.*

(2) *Let* $(W_z)_{z\in Z}$ *be a family of* B*-modules parametrized by* Z*. Assume that for each* $0 \le l \le n$, *there exists an integer* t_l *such that* $\dim_k \operatorname{Tor}^l_B(T, W_z) = t_l$ *for all* $z \in Z$*. Then* $(\operatorname{Tor}^n_B(T, W_z))_{z\in Z}$ *is a family of* A*-modules.*

Remark 1. For $n = 1$, this lemma was proved by Domokos and Lenzing [2000, Lemma 6.3]. Here, we explain how to prove the general case by working with Hochschild complexes.

Proof. In what follows, for a given integer $l \ge 0$, we write A^l and B^l for

$$\underbrace{A \otimes_k \cdots \otimes_k A}_{l} \quad \text{and} \quad \underbrace{B \otimes_k \cdots \otimes_k B}_{l}.$$

(1) For each $z \in Z$, we consider the Hochschild complex

$$K^*_z:\ 0 \to \operatorname{Hom}_k(T, V_z) \xrightarrow{d^0_z} \operatorname{Hom}_k(A \otimes_k T, V_z) \xrightarrow{d^1_z} \operatorname{Hom}_k(A^2 \otimes_k T, V_z) \longrightarrow \cdots,$$

where

$$\begin{aligned} d^l_z(\phi_l)(a_1 \otimes \cdots \otimes a_{l+1} \otimes t) &= a_1\phi_l(a_2 \otimes \cdots \otimes a_{l+1} \otimes t) + \sum_{i=1}^{l} (-1)^i \phi_l\big(a_1 \otimes \cdots \otimes (a_i a_{i+1}) \otimes \cdots \otimes t\big) \\ &\qquad + (-1)^{l+1}\phi_l\big(a_1 \otimes \cdots \otimes a_l \otimes (a_{l+1}t)\big). \end{aligned}$$

As k is a commutative field, we know that $H^l(K^*_z) \simeq \operatorname{Ext}^l_A(T, V_z)$ for all $l \ge 0$; see, for example, [Weibel 1994, Theorem 8.7.10 and Lemma 9.1.9].

It is now easy to see that $(d^l_z)_{z\in Z}$ is a morphism of vector bundles for each integer $l \ge 0$. Also, for each $0 \le l \le n$, the maps d^l_z for $z \in Z$, have constant rank, and hence the kernel and the image of $(d^l_z)_{z\in Z}$ are subbundles of $(\operatorname{Hom}_k(A^l \otimes_k T, V_z))_{z\in Z}$ and $(\operatorname{Hom}_k(A^{l+1} \otimes_k T, V_z))_{z\in Z}$, respectively [Le Potier 1997, Proposition 1.7.2]. Since these subbundles are clearly families of B-modules, $(\operatorname{Ext}^n_A(T, V_z))_{z\in Z}$ is indeed a family of B-modules.

(2) For this part, we work with the homology of the following complex (see for example [Weibel 1994, Section 8.7.1]):

$$K^*_z:\ 0 \longleftarrow T \otimes_k W_z \xleftarrow{(d_0)_z} T \otimes_k B \otimes_k W_z \xleftarrow{(d_1)_z} T \otimes_k B^2 \otimes_k W_z \longleftarrow \cdots.$$

As before, the differentials of this complex give rise to morphisms of vector bundles whose kernels and images are families of A-modules. From this, one immediately derives the desired claim. $\square$

3B. ***Moduli spaces and tilting.*** We now explain how moduli spaces of semistable A-modules behave under tilting. This was already discussed by Domokos and Lenzing [2000] in the context of moduli spaces of modules over canonical algebras.

Let T be a basic tilting A-module and denote $\mathrm{End}_A(T)^{\mathrm{op}}$ by B. The torsion pairs $(\mathcal{T}(T), \mathcal{F}(T))$ in $\mathrm{mod}(A)$ induced by T and $(\mathcal{X}(T), \mathcal{Y}(T))$ in $\mathrm{mod}(B)$ induced by $D(T) := \mathrm{Hom}_k(T, k)$ are

- $\mathcal{T}({}_AT) = \{M \in \mathrm{mod}(A) \mid \mathrm{Ext}^1_A(T, M) = 0\}$;
- $\mathcal{F}({}_AT) = \{M \in \mathrm{mod}(A) \mid \mathrm{Hom}_A(T, M) = 0\}$;
- $\mathcal{X}(T_B) = \{N \in \mathrm{mod}(B) \mid \mathrm{Hom}_B(N, D(T)) = 0\}$
  $= \{N \in \mathrm{mod}(B) \mid T \otimes_B N = 0\}$; and
- $\mathcal{Y}(T_B) = \{N \in \mathrm{mod}(B) \mid \mathrm{Ext}^1_B(N, D(T)) = 0\}$.
  $= \{N \in \mathrm{mod}(B) \mid \mathrm{Tor}^1_B(T, N) = 0\}$

The Brenner–Butler tilting theorem (see for example [Assem et al. 2006]) tells us that the tilting functor $\mathrm{Hom}_A(T, \_) : \mathrm{mod}(A) \to \mathrm{mod}(B)$ induces an equivalence of categories between $\mathcal{T}(T)$ and $\mathcal{Y}(T)$ with quasiinverse $T \otimes_B \_$. Furthermore, the functor $\mathrm{Ext}^1_A(T, \_) : \mathrm{mod}(B) \to \mathrm{mod}(A)$ induces an equivalence of categories between $\mathcal{F}(T)$ and $\mathcal{X}(T)$ with quasiinverse $Tor^B_1(T, \_)$.

We also have the isometry $u : K_0(A) \to K_0(B)$ defined by

$$u(\mathbf{dim}\, M) = \mathbf{dim}\, \mathrm{Hom}_A(T, M) - \mathbf{dim}\, \mathrm{Ext}^1_A(T, M)$$

for any A-module M.

Definition 3.2. We say an integral weight $\theta \in \mathrm{Hom}_{\mathbb{Z}}(K_0(A), \mathbb{Z})$ is *well-positioned with respect to T* if either

(1) there are nonzero θ-semistable A-modules, $\mathrm{mod}(A)^{ss}_\theta \subseteq \mathcal{T}(T)$, and

$$\theta(\mathbf{dim}\, M) < 0$$

for all nonzero modules M in $\mathcal{F}(T)$; or

(2) there are nonzero θ-semistable A-modules, $\mathrm{mod}(A)^{ss}_\theta \subseteq \mathcal{F}(T)$, and

$$\theta(\mathbf{dim}\, M) > 0$$

for all nonzero modules M in $\mathcal{T}(T)$.

Let θ be an integral weight of A that is well positioned with respect to T. We define $|\theta \circ u^{-1}|$ to be $\theta \circ u^{-1}$ if condition (1) above is satisfied; if condition (2) is satisfied, $|\theta \circ u^{-1}|$ is defined to be $-\theta \circ u^{-1}$.

Now we are ready to prove Theorem 1.3:

Proof of Theorem 1.3. (a) *Case 1*: $\operatorname{mod}(A)^{ss}_\theta \subseteq \mathcal{T}(T)$ and $\theta(\mathbf{dim}\, M) < 0$ for all nonzero modules M in $\mathcal{F}(T)$. In this case, $\theta' = \theta \circ u^{-1}$ and $F = \operatorname{Hom}_A(T, \_)$.

Let M be a θ-semistable A-module. We show that $N = F(M)$ is θ'-semistable. As M is a θ-semistable module lying in $\mathcal{T}(T)$, we deduce that $\theta'(\mathbf{dim}\, N) = 0$. Now, let N' be a submodule of N and let $M' \in \mathcal{T}(T)$ be such that $F(M') \simeq N'$. In particular, we get that $\theta'(\mathbf{dim}\, N') = \theta'(u(\mathbf{dim}\, M') = \theta(\mathbf{dim}\, M')$. If $\phi \in \operatorname{Hom}_A(M', M)$ is the morphism corresponding to the inclusion $N' \hookrightarrow N$, then $\ker(\phi) \in \mathcal{F}(T)$ as F is left exact. Using our assumption on θ, it is now clear that $\theta'(\mathbf{dim}\, N') \leq 0$. This shows that N is θ'-semistable.

Now, let $\widetilde{N}$ be a θ'-semistable B-module. First, we claim that $\widetilde{N} \in \mathcal{Y}(T)$. Indeed, let us consider the canonical sequence of $\widetilde{N}$ with respect to $(\mathcal{X}(T), \mathcal{Y}(T))$:

$$0 \to \operatorname{Ext}^1_A(T, \operatorname{Tor}^1_B(T, \widetilde{N})) \longrightarrow \widetilde{N} \longrightarrow \operatorname{Hom}_A(T, T \otimes_B \widetilde{N}) \to 0.$$

If $\widetilde{N}'$ denotes the B-module $\operatorname{Ext}^1_A(T, \operatorname{Tor}^1_B(T, \widetilde{N}))$, then

$$\mathbf{dim}\, \widetilde{N}' = -u(\mathbf{dim} \operatorname{Tor}^1_B(T, \widetilde{N})),$$

and so $\theta'(\mathbf{dim}\, \widetilde{N}') = -\theta(\mathbf{dim} \operatorname{Tor}^1_B(T, \widetilde{N}))$. Using again our assumption on θ, we have that $\theta'(\mathbf{dim}\, \widetilde{N}')$ is strictly positive unless $\operatorname{Tor}^1_B(T, \widetilde{N}) = \{0\}$. But since $\widetilde{N}$ is θ'-semistable, we must have $\operatorname{Tor}^1_B(T, \widetilde{N}) = \{0\}$, and hence $\widetilde{N} \simeq F(\widetilde{M})$, where $\widetilde{M} := T \otimes_B \widetilde{N} \in \mathcal{T}(T)$.

Next, we show that $\widetilde{M}$ is θ-semistable. It is clear that $\theta(\mathbf{dim}\, \widetilde{M}) = 0$. Now, let $\widetilde{M}'$ be a submodule of $\widetilde{M}$ and note that $\operatorname{coker} F(\pi) \in \mathcal{X}(T)$, where $\pi : \widetilde{M} \to \widetilde{M}/\widetilde{M}'$ is the canonical projection. So, there exists an A-module $\widetilde{M}''$ in $\mathcal{F}(T)$ such that $\mathbf{dim} \operatorname{coker}(F(\pi)) = \mathbf{dim} \operatorname{Ext}^1_A(T, \widetilde{M}'') = -u(\mathbf{dim}\, \widetilde{M}'')$. In particular, we get that $\theta'(\mathbf{dim} \operatorname{coker} F(\pi)) = -\theta(\mathbf{dim}\, \widetilde{M}'') \geq 0$, and from this we see that $\theta'(\mathbf{dim}\, F(\widetilde{M}/\widetilde{M}')) \geq 0$. But since $\theta'(\mathbf{dim}\, F(\widetilde{M}/\widetilde{M}')) = \theta(\mathbf{dim}\, \widetilde{M}/\widetilde{M}')$, we conclude that $\theta(\mathbf{dim}\, \widetilde{M}') \leq 0$. This proves part (a) in Case 1.

Case 2: $\operatorname{mod}(A)^{ss}_\theta \subseteq \mathcal{F}(T)$ and $\theta(\mathbf{dim}\, M) > 0$ for all nonzero modules M in $\mathcal{T}(T)$. In this case, $\theta' = -\theta \circ u^{-1}$ and $F = \operatorname{Ext}^1_A(T, \_)$. The proof in this case is essentially dual to the one above; one simply uses the existence of long exact sequences in (co)homology along with the fact that the projective dimension of T is at most one.

(b) For this part, we follow closely the arguments in [Domokos and Lenzing 2000, Section 6]. First, let us consider the canonical family $(V_M)_{M \in \operatorname{mod}(A, \boldsymbol{d})^{ss}_\theta}$ of $\boldsymbol{d}$-dimensional θ-semistable A-modules. By this we simply mean the trivial vector bundle $\operatorname{mod}(A, \boldsymbol{d})^{ss}_\theta \times V$, where $V = \bigoplus_{i \in Q_0} k^{\boldsymbol{d}(i)}$ and, for each $M \in \operatorname{mod}(A, \boldsymbol{d})^{ss}_\theta$, V is equipped with the A-module structure corresponding to M. Now, it follows from part (a) that for each $M \in \operatorname{mod}(A, \boldsymbol{d})^{ss}_\theta$, $F(V_M)$ is a $\boldsymbol{d}'$-dimensional θ'-semistable B-module. Consequently, we can apply Lemma 3.1 to conclude that $(F(V_M))_{M\in \mathrm{mod}(A,\boldsymbol{d})^{ss}_\theta}$ is actually a family of $\boldsymbol{d}'$-dimensional θ'-semistable B-modules. Hence, we get the morphism of varieties $\phi : \mathrm{mod}(A,\boldsymbol{d})^{ss}_\theta \to \mathcal{M}(B,\boldsymbol{d}')^{ss}_{\theta'}$ that sends $M \in \mathrm{mod}(A,\boldsymbol{d})^{ss}_\theta$ to the point of $\mathcal{M}(B,\boldsymbol{d}')^{ss}_{\theta'}$ corresponding to the S-equivalence class of $F(V_M)$. It is clear that ϕ is a $\mathrm{PGL}(\boldsymbol{d})$-invariant morphism. From the universal property of the GIT-quotient $\mathcal{M}(A,\boldsymbol{d})^{ss}_\theta$, we obtain the morphism of algebraic varieties $f : \mathcal{M}(A,\boldsymbol{d})^{ss}_\theta \to \mathcal{M}(B,\boldsymbol{d}')_{\theta'}$ for which $f \circ \pi = \phi$, where $\pi : \mathrm{mod}(A,\boldsymbol{d})^{ss}_\theta \to \mathcal{M}(A,\boldsymbol{d})^{ss}_\theta$ is the quotient morphism. To construct the inverse morphism of f, one follows the same arguments as above, with the functor F replaced by its quasiinverse. □

3C. *The theta-stable decomposition for irreducible components.* Derksen and Weyman [2011] introduced the notion of θ-stable decomposition for spaces of representations of quivers without relations. In this section, we explain how to extend [Derksen and Weyman 2011, Theorem 3.20] to quivers with relations.

Let $A = kQ/I$ be a bound quiver algebra, $\boldsymbol{d} \in \mathbb{Z}^{Q_0}_{\geq 0}$ a dimension vector of A, $C \subseteq \mathrm{mod}(A,\boldsymbol{d})$ an irreducible component, and $\theta \in \mathbb{Z}^{Q_0}$ an integral weight of A. We say that C is a θ(-semi)-stable irreducible component if C contains a θ(-semi)-stable A-module. A θ-semistable irreducible component $C \subseteq \mathrm{mod}(A,\boldsymbol{d})$ is said to be *θ-well-behaved* if $\mathrm{mod}(A,\boldsymbol{d}')$ has a unique θ-stable irreducible component whenever $\boldsymbol{d}'$ is the dimension vector of a factor of a Jordan–Hölder filtration in $\mathrm{mod}(A)^{ss}_\theta$ of a generic A-module in C.

Example 3.3. If A is a tame quasitilted algebra, then any θ-semistable irreducible component is θ-well-behaved. This is because for any generic root $\boldsymbol{d}$ of A, as shown by Bobiński and Skowroński [1999b], $\mathrm{mod}(A,\boldsymbol{d})$ has a unique indecomposable irreducible component.

Let C be a θ-well-behaved irreducible component of $\mathrm{mod}(A,\boldsymbol{d})$. We say that

$$C = C_1 \dot{+} \cdots \dot{+} C_l$$

is *the θ-stable decomposition of* C if

- the $C_i \subseteq \mathrm{mod}(A,\boldsymbol{d}_i)$ for $1 \leq i \leq l$ are θ-stable irreducible components; and
- the generic A-module M in C has a finite filtration $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_l = M$ of submodules such that each factor M_j/M_{j-1} for $1 \leq j \leq l$ is isomorphic to a θ-stable module in one of the $C_1,\dots,C_l$, and the sequence $(\mathbf{dim}\, M_1/M_0,\dots,\mathbf{dim}\, M/M_{l-1})$ is the same as $(\boldsymbol{d}_1,\dots,\boldsymbol{d}_l)$ up to permutation.

To prove the existence and uniqueness of the θ-stable decomposition of C, first note that the irreducible variety C^{ss}_θ is a disjoint union of sets of the form $\mathcal{F}_{(C_i)_{1\leq i\leq l}}$, where each $\mathcal{F}_{(C_i)_{1\leq i\leq l}}$ consists of those modules $M \in C$ that have a finite filtration $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_l = M$ of submodules with each factor

M_j/M_{j-1} isomorphic to a θ-stable module in one of the C_i for $1 \le i \le l$. (Note that the θ-well-behavedness of C is used to ensure that the union above is indeed disjoint.) Next, it is not difficult to show that each $\mathcal{F}_{(C_i)_{1\le i\le l}}$ is constructible; see for example [Crawley-Boevey and Schröer 2002, Section 3]. Hence, there is a unique (up to permutation) sequence $(C_i)_{1\le i\le l}$ of θ-stable irreducible components for which $\mathcal{F}_{(C_i)_{1\le i\le l}}$ contains an open and dense subset of C_θ^{ss} (or C).

Remark 2. Let us mention that the notion of θ-stable decomposition of a dimension vector in an irreducible component of a module variety was introduced in [Chindris 2011, Section 6.2]. It serves as a useful tool for finding convenient orthogonal exceptional sequences. But in order to understand how weight spaces of semiinvariants behave with respect to such a decomposition, one also needs to be able to keep track of the various θ-stable irreducible components that arise in the decomposition in question. This issue is now addressed in the notion above of θ-stable decomposition of a well-behaved irreducible component.

Next, we recall the following useful fact from invariant theory. Let G and G_1 be linearly reductive groups with $G_1 \le G$, let V be a finite-dimensional rational representation of G, and let V_1 be a vector subspace of V invariant under the action of G_1. The G_1-equivariant inclusion $\tau : V_1 \hookrightarrow V$ descends to a morphism

$$\psi : V_1/\!/G_1 \to V/\!/G$$

such that $\psi \circ \pi_1 = \pi \circ \tau$, where $\pi : V \twoheadrightarrow V/\!/G$ and $\pi_1 : V_1 \twoheadrightarrow V_1/\!/G_1$ are the categorical affine quotient morphisms. We denote the image of the zero vector of V through the two quotient morphisms by the same symbol 0. Consider the Hilbert's nullcones $\mathcal{N}_G(V) := \pi^{-1}(0)$ and $\mathcal{N}_{G_1}(V_1) := \pi_1^{-1}(0)$.

Lemma 3.4. *Keep the same notation as above. If $\psi^{-1}(0) = \{0\}$, then ψ is a finite morphism.*

Proof. Let I be the ideal of $K[V]$ generated by all homogeneous G-invariants of positive degree. By choosing homogeneous invariants $f_1, \dots, f_n \in K[V]^G$ such that $I = (f_1, \dots, f_n)$, Hilbert proved that $K[V]^G = K[f_1, \dots, f_n]$; see for example [Derksen and Kemper 2002, Theorem 2.2.10].

Now, if $\mathfrak{m}$ denotes the ideal of $K[V]^G$ generated by $f_1, \dots f_n$, then the zero set of $\mathfrak{m}$ in $V/\!/G$ is precisely $\{0\}$. From this fact and the assumption that $\psi^{-1}(0) = \{0\}$, we immediately deduce that the zero set of $\psi^*(f_1), \dots, \psi^*(f_n)$ in V_1 is precisely the nullcone $\mathcal{N}_{G_1}(V_1)$. Hence, $K[V_1]^{G_1}$ is a finite module over

$$K[\psi^*(f_1), \dots, \psi^*(f_n)];$$

see for example [Derksen and Kemper 2002, Lemma 2.4.5]. The proof follows. □

With the right definition of θ-stable decomposition, the proof of Theorem 1.4 is essentially the same as that of [Derksen and Weyman 2011, Theorem 3.20]. Nonetheless, we provide below a detailed proof for completeness. In what follows, if C' is a θ-stable irreducible component that occurs in the θ-stable decomposition of C with multiplicity m, we denote $\underbrace{C' \dotplus C' \dotplus \cdots \dotplus C'}_{m}$ by $m \cdot C'$.

Proof of Theorem 1.4. Without loss of generality, we assume that θ is indivisible, the induced character $\chi_\theta \in X^*(\mathrm{GL}(\boldsymbol{d}))$ is not trivial, and Q is connected.

We view $\mathcal{V}$ as a vector subspace of $\mathrm{mod}(Q, \boldsymbol{d})$ and denote by G the stabilizer of $\mathcal{V} \subseteq \mathrm{mod}(Q, \boldsymbol{d})$ in G_θ. It easy to see that G is isomorphic to the intersection of G_θ with

$$(S_{m_1} \ltimes \mathrm{GL}(\boldsymbol{d}_1)^{m_1}) \times \cdots \times (S_{m_n} \ltimes \mathrm{GL}(\boldsymbol{d}_n)^{m_n}).$$

(Here, S_m denotes the symmetric group on m elements.) Let

$$\psi : \mathcal{V}/\!/G \to \mathrm{mod}(Q, \boldsymbol{d})/\!/G_\theta$$

be the morphism induced by the G-equivariant inclusion $\tau : \mathcal{V} \hookrightarrow \mathrm{mod}(Q, \boldsymbol{d})$. Since X embeds G-equivariantly into C, ψ descends to a morphism

$$\widetilde{\psi} : X/\!/G \to C/\!/G_\theta$$

such that $\widetilde{\psi} \circ \pi_X = \pi_C \circ \tau|_X$, where $\pi_X : X \to X/\!/G$ and $\pi_C : C \to C/\!/G_\theta$ are the categorical quotient morphisms. Note that

$$K[C/\!/G_\theta] = \bigoplus_{m \geq 0} \mathrm{SI}(C)_{m\theta}, \quad \text{and} \quad K[X/\!/G] = \bigoplus_{m \geq 0} \bigotimes_{i=1}^{n} S^{m_i}(\mathrm{SI}(C_i)_{m\theta}),$$

and moreover, the pullback map $\widetilde{\psi}^*$ respects the gradings of the coordinate rings above. In what follows we show that $\widetilde{\psi}^*$ is an isomorphism.

Note that if $M \in \mathcal{V}$, then M is G-semistable, meaning that $0 \in \overline{GM}$ if and only if the direct summands of M are θ-semistable. This implies that $\psi^{-1}(0) = \{0\}$, and so ψ is a finite morphism by Lemma 3.4. But since $\widetilde{\psi}$ is the restriction of ψ to $X/\!/G$, we can immediately see that $\widetilde{\psi}$ is a finite morphism too.

Next, let $M \in C_\theta^{ss}$ be a module that has a filtration of the form

$$0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_l = M,$$

where the factors M_i/M_{i-1} for $1 \leq i \leq l$ are θ-stable and the sequence

$$(\mathbf{dim}\, M_1, \ldots, \mathbf{dim}\, M/M_{l-1})$$

is the same as $(\boldsymbol{d}_1^{m_1}, \ldots, \boldsymbol{d}_n^{m_n})$ up to permutation. Here, $l := m_1 + \cdots + m_n$. Now, let $\widetilde{M} \in X$ be a module isomorphic to $\bigoplus_{i=1}^{l} M_i/M_{i-1}$. Then, we have

$$\widetilde{\psi}(\pi_X(\widetilde{M})) = \pi_C(M),$$

and hence $\tilde{\Psi}$ is dominant. Denote by X^0 the nonempty open subset

$$(C^s_{1,\theta})^{m_1} \times \cdots \times (C^s_{n,\theta})^{m_n}$$

of X, and note that any point of X^0 has its G_θ-orbit closed in C. This implies that π_C is injective on X^0, and so the morphism $\tilde{\Psi}$ is injective on $\pi_X(X^0)$; in particular, $\tilde{\Psi}$ is injective on an open and dense subset of $X//G$. It is now clear that $\tilde{\Psi}$ has to be a birational morphism.

Finally, we know from geometric invariant theory that the affine quotient variety $C//G_\theta$ is normal, since C is assumed to be a normal variety. It now follows that $\tilde{\Psi}$ is an isomorphism, and this finishes the proof. □

Remark 3. Keep the same assumptions as in Theorem 1.4. If we further assume that A is tame, then for each $1 \le i \le n$, the moduli space $\mathcal{M}(C_i)^{ss}_\theta$ is of dimension $\dim C_i - \dim \mathrm{GL}(\boldsymbol{d}_i) + 1 \le 1$. More precisely, $\mathcal{M}(C_i)^{ss}_\theta$ is a curve if, for example, $q_A(\boldsymbol{d}_i) = 0$; see [de la Peña 1991, Proposition 1.2].

Hence, the "building blocks" $\mathcal{M}(C_1)^{ss}_\theta, \ldots, \mathcal{M}(C_n)^{ss}_\theta$ that make up the moduli space $\mathcal{M}(C)^{ss}_\theta$ are either points or projective curves in the tame case.

4. Tilted algebras

Recall that a quasitilted algebra is a basic and connected finite-dimensional algebra of the form $\mathrm{End}_{\mathcal{H}}(T)^{\mathrm{op}}$, where $\mathcal{H}$ is a hereditary category and $T \in \mathcal{H}$ is a tilting object.

4A. ***Singular moduli spaces of modules for wild tilted algebras.*** Let

$$B = \mathrm{End}_A(T)^{\mathrm{op}}$$

be a wild tilted algebra, where $A = kQ$ with Q a wild connected quiver and T is a basic tilting A-module. Our goal here is to show that B has a singular moduli space of modules. We achieve this by reducing the considerations to the case of wild hereditary algebras via Theorem 1.3.

Proposition 4.1. *If B is a wild tilted algebra, then there exist a generic root $\boldsymbol{d}$ of B, an indecomposable irreducible component C of* $\mathrm{mod}(B, \boldsymbol{d})$, *and an integral weight θ of B such that $C^s_\theta \neq \varnothing$ and the moduli space $\mathcal{M}(C)^{ss}_\theta$ is singular.*

Proof. First of all, we know from the main results in [Kerner 1989; 1997] and [Strauss 1991] that any wild tilted algebra contains a convex subcategory that is wild concealed (the titling module involved is either preprojective or preinjective). Consequently, we can assume that $B = \mathrm{End}_A(T)^{\mathrm{op}}$, where $A = kQ$ with Q a connected wild quiver and T is a basic preprojective tilting A-module. (The case when T is preinjective is dual.) Then, we know that the indecomposable A-modules in

$\mathcal{F}(T)$ are all preprojective and any regular or preinjective A-module belongs to $\mathcal{T}(T)$; see for example [Assem et al. 2006].

To construct a weight θ with the desired properties, we begin by choosing a regular A-module X_0 with the property that all $\tau_A^m X$ for $m \geq 0$ are sincere regular Schur A-modules and $\mathbf{dim}\, X_0$ is an imaginary, nonisotropic root of A; see [Kerner 1996, Proposition 10.2]. Denote the dimension vector of X_0 by $\boldsymbol{d}_0$ and let θ_0 be the weight $\langle \boldsymbol{d}_0, \cdot \rangle_A - \langle \cdot, \boldsymbol{d}_0 \rangle_A$. Then $n\boldsymbol{d}_0$ is θ_0-stable for all integers $n \in \mathbb{Z}_{>0}$ by [Schofield 1992, Theorem 6.1] and [Derksen and Weyman 2011, Proposition 3.16].

Next, we show that θ_0 is well positioned with respect to T, which is equivalent to showing that $\theta_0(\mathbf{dim}\, M) < 0$ for every preprojective A-module M. Assume to the contrary that there exists a preprojective A-module M such that $\langle \mathbf{dim}\, X, \mathbf{dim}\, M \rangle \geq \langle \mathbf{dim}\, M, \mathbf{dim}\, X \rangle$. But this is equivalent to

$$-\dim_k \operatorname{Ext}_A^1(X, M) \geq \dim_k \operatorname{Hom}_A(M, X),$$

and so $\dim_k \operatorname{Ext}_A^1(X, M) = 0$. Writing $M = \tau_A^{-m} P_i$ for uniquely determined $m \in \mathbb{Z}_{\geq 0}$ and $i \in Q_0$, we get that $\tau_A^{m+1} X(i) = \{0\}$, which contradicts that $\tau_A^{m+1} X$ is sincere. So, we conclude that θ_0 is well positioned with respect to T.

Let $u : K_0(A) \to K_0(B)$ be the isometry induced by T and let $\theta = \theta_0 \circ u^{-1}$. We claim that $C := \overline{\operatorname{mod}(B, \boldsymbol{d})_\theta^{ss}}$ is an irreducible component of $\operatorname{mod}(B, \boldsymbol{d})$, where $\boldsymbol{d} := u(n\boldsymbol{d}_0)$ and $n \in \mathbb{Z}_{>0}$. Indeed, it follows from the proof of Theorem 1.3(a) that the θ-semistable B-modules all lie in $\mathcal{Y}(T)$, and hence their projective dimension is at most one, as A is hereditary. Consequently, the subset $\operatorname{mod}_{\mathcal{P}}(B, \boldsymbol{d})$ of $\operatorname{mod}(B, \boldsymbol{d})$ consisting of all modules of projective dimension at most one is nonempty, and this implies that $\operatorname{mod}_{\mathcal{P}}(B, \boldsymbol{d})$ is an irreducible open subset of $\operatorname{mod}(B, \boldsymbol{d})$; see [Barot and Schröer 2001, Proposition 3.1]. This immediately implies our claim. Furthermore, as $n\boldsymbol{d}_0$ is θ_0-stable, we deduce from the proof of Theorem 1.3(a) that $\boldsymbol{d}$ is θ-stable, that is, $C_\theta^s \neq \varnothing$.

Using Theorem 1.3(b) again, we get that $\mathcal{M}(C)_\theta^{ss} \simeq \mathcal{M}(A, n\boldsymbol{d}_0)_{\theta_0}^{ss}$, which is known to be singular for $n = 3$; see for example [Domokos 2011]. □

Proof of Proposition 1.2. Assuming to the contrary that A is wild, it follows from [Brüstle et al. 2011, Corollary 1] that A contains a convex hypercritical algebra B. Then Proposition 4.1 provides us with a singular moduli space of B-modules, which contradicts our assumption on the moduli spaces of modules for A. □

Remark 4. In [Brüstle et al. 2011], Brüstle, de la Peña, and Skowroński proved that for a tame strongly simply connected algebra A, the convex hull of any indecomposable A-module inside A is a tame tilted algebra, or a coil algebra, or a $\mathbb{D}$-algebra; see [Brüstle et al. 2011, Corollary 5]. Hence, to prove the analogue of Theorem 1.1 for strongly simply connected algebras, which was conjectured to hold true by Weyman, it remains to study the geometry of modules over coil algebras and $\mathbb{D}$-algebras. We plan to address these issues in future work.

4B. ***Rational and GIT quotient varieties of modules for tame quasitilted algebras.*** In what follows, we review some important facts about the geometry of modules over quasitilted algebras, which are due to Bobiński and Skowroński.

By a *root* of a quasitilted algebra A, we simply mean the dimension vector of an indecomposable A-module. We say that a root $\boldsymbol{d}$ of A is *real* if $q_A(\boldsymbol{d}) = 1$. We call a root $\boldsymbol{d}$ of A *isotropic* if $q_A(\boldsymbol{d}) = 0$. If $\boldsymbol{d}$ is an isotropic generic root of A, we call the indecomposable irreducible components of $\operatorname{mod}(A, \boldsymbol{d})$ isotropic, too.

Now, we can state the following important result; see [Bobiński and Skowroński 1999b, Corollaries 3 and 2.5 and Proposition 2.3].

Theorem 4.2. *Let A be a tame quasitilted algebra and let $\boldsymbol{d}$ be a generic root of A. Then $\boldsymbol{d}$ is a Schur root with $q_A(\boldsymbol{d}) \in \{0, 1\}$. More precisely*:

(1) *If $q_A(\boldsymbol{d}) = 1$, there exists a unique, up to isomorphism, $\boldsymbol{d}$-dimensional indecomposable A-module M that is, in fact, exceptional; if this is the case, then $\overline{\mathrm{GL}(\boldsymbol{d})M}$ is the unique indecomposable irreducible component of* $\operatorname{mod}(A, \boldsymbol{d})$.

(2) *If $q_A(\boldsymbol{d}) = 0$, the support of $\boldsymbol{d}$ is a tame concealed or a tubular convex subcategory of A. Furthermore,* $\operatorname{mod}(A, \boldsymbol{d})$ *is a normal variety.*

Proposition 4.3 [Chindris 2011]. *Let A be a tame concealed or a tubular algebra, and $\boldsymbol{d}$ an isotropic Schur root of A. Then there exists a short orthogonal exceptional sequence $\mathscr{E} = (E_1, E_2)$ with* $\dim_k \operatorname{Ext}^1_A(E_2, E_1) = 2$ *and* $\operatorname{Ext}^2_A(E_2, E_1) = 0$, *and such that the generic module M in* $\operatorname{mod}(A, \boldsymbol{d})$ *fits into a short exact sequence of the form*

$$0 \longrightarrow E_1 \longrightarrow M \longrightarrow E_2 \longrightarrow 0.$$

Remark 5. This proposition has been proved for tame canonical algebras in [Chindris 2011, Proposition 6.7], but the exact same arguments work for arbitrary tame concealed algebras and for tubular algebras; see for example [Chindris 2012].

Proposition 4.4. *Let A be a quasitilted algebra.*

(1) *The following conditions are equivalent*:

  (a) *A is tame*;

  (b) *for each generic root $\boldsymbol{d}$ of A and each indecomposable irreducible component C of* $\operatorname{mod}(A, \boldsymbol{d})$, *either $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k$ or $k(x)$.*

(2) *Assume A is tame and let $\boldsymbol{d}$ be an isotropic root of A. Then $\mathcal{M}(\operatorname{mod}(A, \boldsymbol{d}))^{ss}_{\theta}$ is a product of projective spaces for every integral weight θ of A.*

Proof. (1) The implication (b) $\Longrightarrow$ (a) has been already proved in [Chindris 2011, Proposition 4.6].

Now, let us assume that A is tame and let $\boldsymbol{d}$ be a generic root of A. We know from Theorem 4.2 that $\boldsymbol{d}$ is a Schur root and $\operatorname{mod}(A, \boldsymbol{d})$ has a unique indecomposable irreducible component; call it C.

If $q_A(\boldsymbol{d}) = 1$, then $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k$ since C is just the closure of the $\mathrm{GL}(\boldsymbol{d})$-orbit of the $\boldsymbol{d}$-dimensional exceptional A-module.

It remains to look into the case when $\boldsymbol{d}$ is an isotropic Schur root of A. In this case, we simply use Proposition 4.3 and Proposition 2.2 to conclude that $k(C)^{\mathrm{GL}(\boldsymbol{d})} \simeq k(x)$.

(2) We know that $\operatorname{mod}(A, \boldsymbol{d})$ is normal by Corollary 3 in [Bobiński and Skowroński 1999b]. Now, let θ be an integral weight for which $\mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta} \neq \varnothing$, and note that $\operatorname{mod}(A, \boldsymbol{d})$ is θ-well-behaved by Theorem 4.2. Let $C_1, \dots, C_n$ be the pairwise distinct isotropic indecomposable irreducible components that occur in the θ-stable decomposition of $\operatorname{mod}(A, \boldsymbol{d})$, and denote by $m_1, \dots, m_n$ their multiplicities. It now follows from Theorem 1.4 that

$$\mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta} \cong \prod_{i=1}^{n} S^{m_i}(\mathcal{M}(C_i)^{ss}_{\theta}).$$

But, for each $1 \leq i \leq n$, $\mathcal{M}(C_i)^{ss}_{\theta}$ is a projective curve that is, first, normal, as C_i is normal by Theorem 4.2(2) and, second, rational, as proved in part (1). So, $\mathcal{M}(C_i)^{ss}_{\theta} \simeq \mathbb{P}^1$ for all $1 \leq i \leq n$, and hence $\mathcal{M}(A, \boldsymbol{d})^{ss}_{\theta} \cong \prod_{i=1}^{n} \mathbb{P}^{m_i}$. □

Remark 6. Let A be a tame quasitilted algebra, $\boldsymbol{d}$ a root of A, $C \subseteq \operatorname{mod}(A, \boldsymbol{d})$ an irreducible component, and θ an integral weight of A such that $C^s_{\theta} \neq \varnothing$. Then the proposition above tells us that $\mathcal{M}(C)^{ss}_{\theta}$ is either a point or just $\mathbb{P}^1$.

Proof of Theorem 1.1. The implications (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) were proved in Proposition 4.4. The implication (4) $\Longrightarrow$ (1) follows from Proposition 4.1. □

Proof of Proposition 1.5. We know from Theorem 1.1 that if C is an indecomposable irreducible component of $\operatorname{mod}(A, \boldsymbol{d})$, then $S^m(k(C)^{\mathrm{GL}(\boldsymbol{d})})$ is isomorphic to either k, in case $\boldsymbol{d}$ is a real Schur root, or $k(t_1, \dots, t_m)$, in case $\boldsymbol{d}$ is isotropic. The proof now follows from Proposition 2.1 and Proposition 4.4. □

Remark 7. In view of [Happel 2001], to prove the implication (4) $\Longrightarrow$ (1) of Theorem 1.1 for quasitilted algebras, one possible path is to prove first the analogue of Theorem 1.3 for tilting complexes, and then that of Proposition 4.1 for wild canonical algebras. We plan to explore this approach in a sequel to this work.

Acknowledgments

I am grateful to Otto Kerner for clarifying conversations on wild tilted algebras. I would also like to thank the referee for detailed suggestions that helped improve the exposition of the paper.

References

[Assem et al. 2006] I. Assem, D. Simson, and A. Skowroński, *Elements of the representation theory of associative algebras, 1: Techniques of representation theory*, London Mathematical Society Student Texts **65**, Cambridge University Press, 2006. MR 2006j:16020 Zbl 1092.16001

[Barot and Schröer 2001] M. Barot and J. Schröer, "Module varieties over canonical algebras", *J. Algebra* **246**:1 (2001), 175–192. MR 2003e:16013 Zbl 1036.16010

[Bobiński 2008] G. Bobiński, "On the zero set of semi-invariants for regular modules over tame canonical algebras", *J. Pure Appl. Algebra* **212**:6 (2008), 1457–1471. MR 2009c:16042 Zbl 1180.16011

[Bobiński and Skowroński 1999a] G. Bobiński and A. Skowroński, "Geometry of directing modules over tame algebras", *J. Algebra* **215**:2 (1999), 603–643. MR 2000f:16017 Zbl 0965.16009

[Bobiński and Skowroński 1999b] G. Bobiński and A. Skowroński, "Geometry of modules over tame quasi-tilted algebras", *Colloq. Math.* **79**:1 (1999), 85–118. MR 2000i:14067 Zbl 0994.16009

[Bobiński and Skowroński 2002] G. Bobiński and A. Skowroński, "Geometry of periodic modules over tame concealed and tubular algebras", *Algebr. Represent. Theory* **5**:2 (2002), 187–200. MR 2003d:16021 Zbl 1013.16009

[Bongartz 1983] K. Bongartz, "Algebras and quadratic forms", *J. London Math. Soc.* (2) **28**:3 (1983), 461–469. MR 85i:16036 Zbl 0532.16020

[Brüstle et al. 2011] T. Brüstle, J. A. de la Peña, and A. Skowroński, "Tame algebras and Tits quadratic forms", *Adv. Math.* **226**:1 (2011), 887–951. MR 2011m:16023 Zbl 05825430

[Chindris 2011] C. Chindris, "Geometric characterizations of the representation type of hereditary algebras and of canonical algebras", *Adv. Math.* **228**:3 (2011), 1405–1434. MR 2012h:16033 Zbl 05949109

[Chindris 2012] C. Chindris, "On the geometry of orbit closures for representation-infinite algebras", *Glasg. Math. J.* **54**:3 (2012), 629–636.

[Crawley-Boevey and Schröer 2002] W. Crawley-Boevey and J. Schröer, "Irreducible components of varieties of modules", *J. Reine Angew. Math.* **553** (2002), 201–220. MR 2004a:16020 Zbl 1062.16019

[Derksen and Kemper 2002] H. Derksen and G. Kemper, *Computational invariant theory*, Encyclopaedia of Mathematical Sciences **130**, Springer, Berlin, 2002. MR 2003g:13004 Zbl 1011.13003

[Derksen and Weyman 2011] H. Derksen and J. Weyman, "The combinatorics of quiver representations", *Ann. Inst. Fourier* (*Grenoble*) **61**:3 (2011), 1061–1131. MR 2918725 Zbl 06002991

[Domokos 2011] M. Domokos, "On singularities of quiver moduli", *Glasg. Math. J.* **53**:1 (2011), 131–139. MR 2012a:16029 Zbl 1241.16010

[Domokos and Lenzing 2000] M. Domokos and H. Lenzing, "Invariant theory of canonical algebras", *J. Algebra* **228**:2 (2000), 738–762. MR 2001h:16016 Zbl 0955.16015

[Domokos and Lenzing 2002] M. Domokos and H. Lenzing, "Moduli spaces for representations of concealed-canonical algebras", *J. Algebra* **251**:1 (2002), 371–394. MR 2003d:16016 Zbl 1013.16006

[Drozd 1979] J. A. Drozd, "Tame and wild matrix problems", pp. 39–74 in *Representations and quadratic forms*, edited by Y. A. Mitropol'skiĭ, Akad. Nauk Ukrain. SSR Inst. Mat., Kiev, 1979. In Russian; translated in *Amer. Math. Soc. Transl. Ser. 2* **128** (1986), 31–55. MR 82m:16028 Zbl 0454.16014

[Geiss and Schröer 2003] C. Geiss and J. Schröer, "Varieties of modules over tubular algebras", *Colloq. Math.* **95**:2 (2003), 163–183. MR 2004d:16026 Zbl 1033.16004

[Happel 2001] D. Happel, "A characterization of hereditary categories with tilting object", *Invent. Math.* **144**:2 (2001), 381–398. MR 2002a:18014 Zbl 1015.18006

[Happel et al. 1996] D. Happel, I. Reiten, and S. O., *Tilting in abelian categories and quasitilted algebras*, Mem. Amer. Math. Soc. **120**, 1996. MR 97j:16009 Zbl 0849.16011

[Kerner 1989] O. Kerner, "Tilting wild algebras", *J. London Math. Soc.* (2) **39**:1 (1989), 29–47. MR 90d:16025 Zbl 0675.16013

[Kerner 1996] O. Kerner, "Representations of wild quivers", pp. 65–107 in *Representation theory of algebras and related topics* (Mexico City, 1994), edited by R. Bautista et al., CMS Conf. Proc. **19**, Amer. Mathematical Society, Providence, RI, 1996. MR 97e:16028 Zbl 0863.16010

[Kerner 1997] O. Kerner, "Wild tilted algebras revisited", *Colloq. Math.* **73**:1 (1997), 67–81. MR 98b:16011 Zbl 0879.16006

[King 1994] A. D. King, "Moduli of representations of finite-dimensional algebras", *Quart. J. Math. Oxford Ser.* (2) **45**:180 (1994), 515–530. MR 96a:16009 Zbl 0837.16005

[Kraśkiewicz 2001] W. Kraśkiewicz, "On semi-invariants of tilted algebras of type A_n", *Colloq. Math.* **90**:2 (2001), 253–267. MR 2002m:16015 Zbl 0993.16010

[Le Potier 1997] J. Le Potier, *Lectures on vector bundles*, Cambridge Studies in Advanced Mathematics **54**, Cambridge University Press, 1997. MR 98a:14019 Zbl 0872.14003

[de la Peña 1991] J. A. de la Peña, "On the dimension of the module-varieties of tame and wild algebras", *Comm. Algebra* **19**:6 (1991), 1795–1807. MR 92i:16016 Zbl 0818.16013

[Riedtmann 2004] C. Riedtmann, "Tame quivers, semi-invariants, and complete intersections", *J. Algebra* **279**:1 (2004), 362–382. MR 2005j:16015 Zbl 1076.16013

[Riedtmann and Zwara 2004] C. Riedtmann and G. Zwara, "On the zero set of semi-invariants for tame quivers", *Comment. Math. Helv.* **79**:2 (2004), 350–361. MR 2005g:16024 Zbl 1063.14052

[Riedtmann and Zwara 2008] C. Riedtmann and G. Zwara, "The zero set of semi-invariants for extended Dynkin quivers", *Trans. Amer. Math. Soc.* **360**:12 (2008), 6251–6267. MR 2009i:14064 Zbl 1159.14306

[Schofield 1992] A. Schofield, "General representations of quivers", *Proc. London Math. Soc.* (3) **65**:1 (1992), 46–64. MR 93d:16014 Zbl 0795.16008

[Simson and Skowroński 2007] D. Simson and A. Skowroński, *Elements of the representation theory of associative algebras, 3: Representation-infinite tilted algebras*, London Mathematical Society Student Texts **72**, Cambridge University Press, 2007. MR 2008m:16001 Zbl 1131.16001

[Skowroński 1993] A. Skowroński, "Simply connected algebras and Hochschild cohomologies [MR1206961 (94e:16016)]", pp. 431–447 in *Representations of algebras* (Ottawa, 1992), CMS Conf. Proc. **14**, American Mathematical Society, Providence, RI, 1993. MR 1265301 Zbl 0806.16012

[Skowroński 1998] A. Skowroński, "Tame quasi-tilted algebras", *J. Algebra* **203**:2 (1998), 470–490. MR 99b:16019 Zbl 0908.16013

[Skowroński and Weyman 2000] A. Skowroński and J. Weyman, "The algebras of semi-invariants of quivers", *Transform. Groups* **5**:4 (2000), 361–402. MR 2001m:16017 Zbl 0986.16004

[Strauss 1991] H. Strauss, "On the perpendicular category of a partial tilting module", *J. Algebra* **144**:1 (1991), 43–66. MR 92m:16013 Zbl 0746.16009

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics **38**, Cambridge University Press, 1994. MR 95f:18001 Zbl 0797.18001

Communicated by J. Toby Stafford
Received 2011-09-17 Revised 2012-01-02 Accepted 2012-02-20

chindrisc@missouri.edu *Department of Mathematics, University of Missouri, Columbia, MO 65211, United States*
http://www.math.missouri.edu/~chindrisc

ALGEBRA AND NUMBER THEORY 7:1 (2013)
dx.doi.org/10.2140/ant.2013.7.215

Period functions and cotangent sums

Sandro Bettin and Brian Conrey

We investigate the period function of $\sum_{n=1}^{\infty} \sigma_a(n)\,\mathrm{e}(nz)$, showing it can be analytically continued to $|\arg z| < \pi$ and studying its Taylor series. We use these results to give a simple proof of the Voronoi formula and to prove an exact formula for the second moments of the Riemann zeta function. Moreover, we introduce a family of cotangent sums, functions defined over the rationals, that generalize the Dedekind sum and share with it the property of satisfying a reciprocity formula.

1. Introduction

In the well-known theory of period polynomials one constructs a vector space of polynomials associated with a vector space of modular forms. The Hecke operators act on each space and have the same eigenvalues. Thus, either vector space produces the usual degree 2 L-series associated with holomorphic modular forms. Lewis and Zagier [2001] extended this theory and defined spaces of period functions associated to nonholomorphic modular forms, that is, to Maass forms and real analytic Eisenstein series. Period functions are real analytic functions $\psi(x)$ that satisfy three-term relations

$$\psi(x) = \psi(x+1) + (x+1)^{-2s}\psi\left(\frac{x}{1+x}\right), \tag{1}$$

where $s = 1/2 + it$. The period functions for Maass forms are characterized by (1) together with the growth conditions $\psi(x) = o(1/x)$ as $x \to 0^+$ and $\psi(x) = o(1)$ as $x \to \infty$; for these, $s = 1/2 + ir$, where $1/4 + r^2$ is the eigenvalue of the Laplacian associated with a Maass form. For Eisenstein series, the o's in the growth conditions above are replaced by O's if $t \neq 0$ and by $O(1/(x|\log x|))$ and $O(\log x)$ if $t = 0$. They show that ψ, which is initially defined only in the upper half plane, actually has an analytic continuation to all of $\mathbb{C}$ apart from the negative real axis.

To each period function is also associated a periodic and holomorphic function f on the upper half plane,

$$f(z) = \psi(z) + z^{-2s}\psi(-1/z).$$

MSC2010: primary 11M06; secondary 11M41, 11L99.
Keywords: period functions, moments, mean values, Riemann zeta function, Eisenstein series, Voronoi formula, cotangent sums, Vasyunin sum, Dedekind sum.

In this paper we focus on the case of real analytic Eisenstein series. For these, the periodic function f turns out to be essentially

$$\sum_{n=1}^{\infty} \sigma_{2s-1}(n)\mathrm{e}(nz),$$

where, as usual, $\sigma_a(n) := \sum_{d|n} d^a$ indicates the sum of the a-th power of the divisors of n and $\mathrm{e}(z) := e^{2\pi i z}$. We interpret Lewis and Zagier's results directly in terms of this function, obtaining a better understanding of the Taylor series of the associated period function. It turns out that the case $s = 1/2$, that is, $t = 0$, is especially useful. In this case the arithmetic part of the n-th Fourier coefficient is $d(n)$, the number of divisors of n.

There are several nice applications that are consequences of the analytic continuation of the associated period function, that is, they are consequences of the surprising fact that the function

$$\sum_{n=1}^{\infty} d(n)e(nz) - \frac{1}{z}\sum_{n=1}^{\infty} d(n)e(-n/z),$$

which apparently only makes sense when the imaginary part of z is positive, actually has an analytic continuation to $\mathbb{C}'$ the slit complex plane (the complex with the negative real axis removed). First, we obtain a new formula for the weighted mean square of the Riemann zeta function on the critical line:

$$\int_0^{\infty} |\zeta(1/2+it)|^2 e^{-\delta t}\, dt.$$

Previously, the best formula for this quantity was a main term plus an asymptotic, but not convergent, series of powers of δ, each term an order of magnitude better than the previous as $\delta \to 0^+$. Our formula gives an asymptotic series that is also convergent. The situation is somewhat analogous to the situation of the partition function $p(n)$. Hardy and Ramanujan found an asymptotic series for $p(n)$ and subsequently Rademacher gave a series that was both asymptotic and convergent. In both the partition case and our case, the exact formula allows for the computation of the sought quantity to *any* desired degree of precision, whereas an asymptotic series has limits to its precision. Of course, an extra feature of $p(n)$, which is not present in our situation, is that since $p(n)$ is an integer it is known exactly once it is known to a precision of 0.5. However, our formula does have the extra surprising feature that the time required to calculate our desired mean square is basically independent of δ, apart from the intrinsic difficulty of the extra work required just to write down a high precision number δ.

A second application proves a surprising reciprocity formula for the Vasyunin sum, which is a cotangent sum that appears in the Nyman–Beurling criterion for the

Riemann hypothesis. Specifically, the Vasyunin sum appears as part of the exact formula for the twisted mean-square of the Riemann zeta function on the critical line:

$$\int_0^{\infty} |\zeta(1/2+it)|^2 (h/k)^{it} \frac{dt}{\frac{1}{4}+t^2}.$$

The fact that there is a reciprocity formula for the Vasyunin sum is a nonobvious symmetry relating this integral for h/k and the integral for $\overline{h}/k$ where $h\overline{h} \equiv 1 \bmod k$. It is not apparent from this integral that there should be such a relationship; our formula reveals a hidden structure.

The reciprocity formula is most simply stated in terms of the function

$$c_0(h/k) = -\sum_{m=1}^{k-1} \frac{m}{k} \cot \frac{\pi m h}{k},$$

defined initially for nonzero rational numbers h/k where h and k are integers with $(h,k)=1$ and $k>0$. The reciprocity formula can be simply stated as, “The function

$$c_0\left(\frac{h}{k}\right) + \frac{k}{h} c_0\left(\frac{k}{h}\right) - \frac{1}{\pi h}$$

extends from its initial definition on rationals $x = h/k$ to an (explicit) analytic function on the complex plane with the negative real axis deleted.” This is nearly an example of what Zagier calls a “quantum modular form” [Zagier 2010]. We proved this reciprocity formula in [Bettin and Conrey 2011]; in this paper, we generalize it to a family of “cotangent sums”, containing both c_0 and the Dedekind sum.

These (imperfect) quantum modular forms are analogous to the “quantum Maass forms” studied by Bruggeman [2007], the former being associated to Eisenstein series and the latter to Maass forms. The main difference between these two classes of quantum forms comes from the fact that the L-functions associated to Maass forms are entire, while for Eisenstein series the associated L-functions are not, since they are products of two shifted Riemann zeta functions. This translates into quantum Maass forms being quantum modular forms in the strict sense, whereas the reciprocity formulas for the cotangent sums contain a nonsmooth correction term.

As a third application, we give a generalization of the classical Voronoi summation formula, which is a formula for $\sum_{n=1}^{\infty} d(n) f(n)$, where $f(n)$ is a smooth rapidly decaying function. The usual formula proceeds from

$$\sum_{n=1}^{\infty} d(n) f(n) = \frac{1}{2\pi i} \int_{(2)} \zeta(s)^2 \tilde{f}(s)\, ds, \quad \text{where } \tilde{f}(s) = \int_0^{\infty} f(x) x^{-s}\, dx.$$

One obtains the formula by moving the path of integration to the left to $\operatorname{Re} s = -1$, say, and then using the functional equation

$$\zeta(s) = \chi(s)\zeta(1-s)$$

of $\zeta(s)$. Here, as usual,

$$\chi(s) = 2(2\pi)^{s-1}\Gamma(1-s).$$

In this way one obtains a leading term

$$\int_0^\infty f(u)(\log u + 2\gamma)\,du,$$

from the pole of $\zeta(s)$ at $s = 1$, plus another term

$$\sum_{n=1}^\infty d(n)\hat{f}(n),$$

where $\hat{f}(u)$ is a kind of Fourier–Bessel transform of f; specifically,

$$\hat{f}(u) = \frac{1}{2\pi i}\int_{(-1)} \chi(s)^2 u^{s-1}\tilde{f}(s)\,ds = \int_0^\infty f(t)C(2\pi\sqrt{tu})\,dt$$

with $C(z) = 4K_0(2z) - 2\pi Y_0(2z)$, where K and Y are the usual Bessel functions. By contrast, the period relation implies, for example, that for $0 < \delta < \pi$ and $z = 1 - e^{-i\delta}$,

$$\sum_{n=1}^\infty d(n)\mathrm{e}(nz) = \frac{1}{4} + 2\frac{\log(-2\pi i z) - \gamma}{2\pi i z} + \frac{1}{z}\sum_{n=1}^\infty d(n)\mathrm{e}\Big(\frac{-n}{z}\Big) + \sum_{n=1}^\infty c_n e^{-in\delta}, \quad (2)$$

where $c_n \ll e^{-2\sqrt{\pi n}}$. This is a useful formula that cannot be readily extracted from the Voronoi formula. In fact, the Voronoi formula is actually an easy consequence of the formula (2). In Section 4 we give some other applications of this extended Voronoi formula.

The theory and applications described above are for the period function associated with the Eisenstein series with $s = 1/2$. In this paper we work in a slightly more general setting with $s = a$, an arbitrary complex number. The circle of ideas presented here have other applications and further generalizations, for example to exact formulas for averages of Dirichlet L-functions, which will be explored in future work.

2. Statement of results

For $a \in \mathbb{C}$ and $\mathrm{Im}(z) > 0$, consider

$$\mathcal{S}_a(z) := \sum_{n=1}^{\infty} \sigma_a(n) \mathrm{e}(nz).$$

For $a = 2k+1$ with $k \in \mathbb{Z}_{\geq 1}$, the series $\mathcal{S}_a(z)$ is essentially the Eisenstein series of weight $2k+2$:

$$E_{a+1}(z) = 1 + \frac{2}{\zeta(-a)} \mathcal{S}_a(z),$$

for which the well-known modularity property

$$E_{2k}(z) - \frac{1}{z^{2k}} E_{2k}\Big(-\frac{1}{z}\Big) = 0$$

holds when $k \geq 2$. For other values of a this equality is no longer true, but the period function

$$\psi_a(z) := E_{a+1}(z) - \frac{1}{z^{a+1}} E_{a+1}\Big(-\frac{1}{z}\Big) \tag{3}$$

still has some remarkable properties.

Theorem 1. *Let* $\mathrm{Im}(z) > 0$ *and* $a \in \mathbb{C}$. *Then* $\psi_a(z)$ *satisfies the three-term relation*

$$\psi_a(z) - \psi_a(z+1) = \frac{1}{(z+1)^{1+a}} \psi_a\Big(\frac{z}{z+1}\Big) \tag{4}$$

and extends to an analytic function on $\mathbb{C}' := \mathbb{C} \setminus \mathbb{R}_{\leq 0}$ *via the representation*

$$\psi_a(z) = \frac{i}{\pi z} \frac{\zeta(1-a)}{\zeta(-a)} - i \frac{1}{z^{1+a}} \cot \frac{\pi a}{2} + i \frac{g_a(z)}{\zeta(-a)},$$

where

$$g_a(z) := -2 \sum_{1 \leq n \leq M} (-1)^n \frac{B_{2n}}{(2n)!} \zeta(1-2n-a)(2\pi z)^{2n-1} + \frac{1}{\pi i} \int_{(-\frac{1}{2}-2M)} \zeta(s)\zeta(s-a)\Gamma(s) \frac{\cos \pi a/2}{\sin \pi(s-a)/2} (2\pi z)^{-s} \, ds, \tag{5}$$

and M *is any integer greater or equal to* $-\frac{1}{2}\min(0, \mathrm{Re}(a))$.

Here and throughout the paper equalities are to be interpreted as identities between meromorphic functions in a. In particular, taking the limit $a \to 0^+$, we have

$$\psi_0(z) = -2 \frac{\log 2\pi z - \gamma}{\pi i z} - 2i g_0(z),$$

$$g_0(z) = \frac{1}{\pi i} \int_{(-\frac{1}{2})} \zeta(s)^2 \frac{\Gamma(s)}{\sin(\pi s/2)} (2\pi z)^{-s} \, ds = \frac{1}{\pi i} \int_{(-\frac{1}{2})} \frac{\zeta(s)\zeta(1-s)}{\sin \pi s} z^{-s} \, ds.$$

Theorem 1 is essentially a reformulation of Lewis and Zagier's results [2001] for the noncuspidal case and can be seen as a starting point for their theory of period functions.

For ease of reference, note that (3) can be rewritten in terms of $\mathcal{S}_a$ and g_a as

$$\mathcal{S}_a(z) - \frac{1}{z^{a+1}}\mathcal{S}_a\Big(-\frac{1}{z}\Big) = i\frac{\zeta(1-a)}{2\pi z} - \frac{\zeta(-a)}{2} + \frac{e^{\pi i(a+1)/2}\zeta(a+1)\Gamma(a+1)}{(2\pi z)^{a+1}} + \frac{i}{2}g_a(z). \quad (6)$$

Another important feature of the function $\psi_a(z)$ comes from the properties of its Taylor series. For example, in the case $a = 0$ one has

$$\frac{\pi i}{2}(1+z)\psi_0(1+z) = -1 - \frac{z}{2} + \sum_{m=2}^{\infty} a_m(-z)^m,$$

with

$$a_m := \frac{1}{n(n+1)} + 2b_n + 2\sum_{j=0}^{n-2}\binom{n-1}{j}b_{j+2} \quad \text{and} \quad b_n := \frac{\zeta(n)B_n}{n}$$

and where B_{2n} denotes the $2n$-th Bernoulli number. In particular, the values a_m are rational polynomials in π^2. The terms involved in the definition of a_m are extremely large, since

$$b_{2n} \sim \frac{B_{2n}}{2n} \sim (-1)^{n+1} 2\sqrt{\frac{\pi}{n}}\Big(\frac{n}{\pi e}\Big)^{2n}$$

as $n \to \infty$, though there is a lot of cancellation; for example, for $m = 20$ one has

$$\begin{aligned} a_m = \frac{1}{420} + \frac{\pi^2}{36} - \frac{19\pi^4}{600} + \frac{646\pi^6}{19845} - \frac{323\pi^8}{1500} + \frac{4199\pi^{10}}{343035} \\ - \frac{154226363\pi^{12}}{36569373750} + \frac{1292\pi^{14}}{1403325} - \frac{248571091\pi^{16}}{2170943775000} \\ + \frac{1924313689\pi^{18}}{288905366499750} - \frac{30489001321\pi^{20}}{252669361772953125} \\ = 0.0499998087\ldots \end{aligned}$$

Notice how close this number is to $\frac{1}{20}$; this observation can be made for all m and in fact in [Bettin and Conrey 2011] we proved that

$$a_m - \frac{1}{m} = 2^{5/4}\pi^{3/4}\frac{e^{-2\sqrt{\pi m}}}{m^{3/4}}\Big(\sin(2\sqrt{\pi m} + \tfrac{3}{8}\pi) + O\Big(\frac{1}{\sqrt{m}}\Big)\Big).$$

In this paper we show that similar results hold for the Taylor series at any point τ in the half plane $\mathrm{Re}(\tau) > 0$ and for any $a \in \mathbb{C}$. We give a proof in the following theorem, using g_a instead of ψ_a to simplify slightly the resulting formulas.

Theorem 2. *Let* $\mathrm{Re}(\tau) > 0$ *and for* $|z| < |\tau|$, *let*

$$g_a(\tau+z) := \sum_{m=0}^{\infty} \frac{g_a^{(m)}(\tau)}{m!} z^m$$

be the Taylor series of $g_a(z)$ *around* τ. *Then*

$$\begin{aligned}\frac{g_a^{(m)}(1)}{m!} = -\sum_{\substack{2n-1+k=m,\\ n,k\geq 1}} &(-1)^{n+m} B_{2n}\zeta(1-2n-a)\frac{\Gamma(2n+a+k)}{\Gamma(2n+a)k!(2n)!}2(2\pi)^{2n-1}\\ &+(-1)^m \cot\frac{\pi a}{2}\zeta(-a)\frac{\Gamma(1+a+m)}{\Gamma(1+a)m!}\\ &+(-1)^m\Big(\frac{\Gamma(1+a+m)}{\Gamma(a)(m+1)!}-1\Big)\frac{\zeta(1-a)}{\pi},\end{aligned} \quad (7)$$

and in particular if $a \in \mathbb{Z}_{\leq 0}$ *and* $(a, m) \neq (0, 0)$, *then* $\pi g_a^{(m)}(1)$ *is a rational polynomial in* π^2. *Moreover,*

$$\begin{aligned}\frac{g_a^{(m)}(\tau)}{m!} &= \cos\Big(\frac{\pi a}{2}\Big)\frac{2^{7/4-a/2}}{\pi^{3/4+a/2}}\frac{e^{-2\sqrt{\pi\tau m}}}{m^{1/4-a/2}\tau^{m+3/4+a/2}}\\ &\times\Big(\cos\big(2\sqrt{\pi\tau m}-\tfrac{1}{8}\pi(2a-1)+(\tau+m)\pi\big)+O_{\tau,a}\Big(\frac{1}{\sqrt{m}}\Big)\Big),\end{aligned} \quad (8)$$

as $m \to \infty$.

Some of the ideas used in the proofs of Theorems 1 and 2 can be easily generalized to a more general setting. For example, let $F(s)$ be a meromorphic function on $1-\omega \leq \mathrm{Re}(s) \leq \omega$ for some $1 < \omega < 2$ with no poles on the boundary and assume $|F(\sigma+it)| \ll_\sigma e^{(\pi/2-\eta)|t|}$ for some $\eta > 0$. Let

$$\begin{aligned}W_+(z) &:= \frac{1}{2\pi i}\int_{(\omega)} F(s)\Gamma(s)(-2\pi i z)^{-s}\,ds,\\ W_-(z) &:= \frac{1}{2\pi i}\int_{(\omega)} F(1-s)\Gamma(s)(-2\pi i z)^{-s}\,ds,\end{aligned} \quad (9)$$

for $\frac{\pi}{2}-\eta < \arg z < \frac{\pi}{2}+\eta$. (Notice that these functions are essentially convolutions of the exponential function and the Mellin transform of $F(s)$.) Then we have

$$\sum_{n=1}^{\infty} d(n)W_+(nz) - \frac{1}{z}\sum_{n=1}^{\infty} d(n)W_-\Big(-\frac{n}{z}\Big) = R(z)+k(z), \quad (10)$$

where $R(z)$ is the sum of the residues of $F(s)\Gamma(s)\zeta(s)^2(-2\pi i z)^{-s}$ between $1-\omega$ and ω, and

$$k(z) := \frac{1}{2\pi}\int_{(1-\omega)} F(s)\frac{\zeta(s)\zeta(1-s)}{\sin\pi s} z^{-s}\, ds$$

is holomorphic on $|\arg(z)| < \frac{\pi}{2}+\eta$. Moreover, if we assume that $F(s)$ is holomorphic on $\mathrm{Re}(s) < 1-\omega$, then it follows that the Taylor series of $k(z)$ converges very fast, that is,

$$\frac{k^{(n)}(\tau)}{n!} \ll n^{-B}|\tau|^{-n}$$

for any $B > 0$ and τ such that $|\arg\tau| < \eta$. Also, $W_-(z)$ decays faster than any power of z at infinity and so the second sum in (10) is rapidly convergent and is very small if we let z go to zero in $|\arg z| < \eta$. In Section 4 we will give an explicit example; a subsequent paper will elaborate on this.

The Voronoi summation formula is an important tool in analytic number theory; in its simplest form, it states that, if $f(u)$ is a smooth function of compact support, then

$$\sum_{n=1}^{\infty} d(n)f(n) = \sum_{n=1}^{\infty} d(n)\hat{f}(n) + \int_0^{\infty} f(t)(\log t + 2\gamma)\, dt + \tfrac{1}{4}f(0), \tag{11}$$

where

$$\hat{f}(x) := 4\int_0^{\infty} f(t)(K_0(4\pi\sqrt{tx}) - \tfrac{1}{2}\pi Y_0(4\pi\sqrt{tx}))\, dt.$$

This formula can be deduced from (10) (or also directly from (6)) as a very easy corollary. Actually, Voronoi's formula can be interpreted as a version of (6) confined to the positive real axis. If we get rid of this limitation and we use directly the period formula (6), we are able to obtain interesting results also for weight functions of the shape $f(u) = e^{-\delta u}$, for which the Voronoi summation formula fails to give a useful formula. (Try it!) Thus, we have a generalization of Voronoi's formula.

The use of a weight function of the shape $e^{-\delta u}$ is fundamental to investigate the smoothly weighted second moment of the Riemann zeta function,

$$L_{2k}(\delta) := \int_0^{\infty} \left|\zeta(\tfrac{1}{2}+it)\right|^{2k} e^{-\delta t}\, dt,$$

in the case $k = 1$. These integrals play a major role in the theory of the Riemann zeta function and getting good upper bounds on their growth as $\delta \to 0^+$ would imply the Lindelöf hypothesis. Unfortunately, the only two value of k for which the asymptotics are known are $k = 1$ [Hardy and Littlewood 1916] and $k = 2$ [Ingham 1927]. For other values we have just conjectures; see [Conrey and Ghosh 1998; Conrey and Gonek 2001; Keating and Snaith 2000]. For $k = 1$, it is easy to see that the smooth moment is strictly related to the sum $\mathcal{S}_0(-e^{-i\delta})$ and, from this, it is easy to deduce an asymptotic expansion for $L_{2k}(\delta)$. This classical asymptotic series is not convergent. Here we replace the series by two series, each of which are absolutely convergent asymptotic series. (See also [Motohashi 1997].) The following theorem provides a new exact formula for $L_1(\delta)$, by applying Theorem 1 and 2 to $\mathcal{S}_0(-e^{-i\delta})$.

Theorem 3. *For* $0 < \mathrm{Re}(\delta) < \pi$, *we have*

$$L_1(\delta) = \frac{\gamma - \log 2\pi\delta}{2\sin\delta/2} + \frac{\pi i}{\sin\delta/2}\mathcal{S}_0\Big(\frac{-1}{1-e^{-i\delta}}\Big) + h(\delta) + k(\delta),$$

where $k(\delta)$ *is analytic in* $|\mathrm{Re}(\delta)| < \pi$ *and* $h(\delta)$ *is* C^∞ *in* $\mathbb{R}$ *and holomorphic in*

$$\mathbb{C}'' := \mathbb{C} \setminus \{x + iy \in \mathbb{C} \mid x \in 2\pi\mathbb{Z},\ y \geq 0\}.$$

Moreover, $h(0) = 0$ *and*, *if* $\mathrm{Im}(\delta) \leq 0$,

$$h(\delta) = i\sum_{n\geq 0} h_n e^{-i(n+1/2)\delta},$$

with

$$h_n = 2^{7/4}\pi^{1/4}\frac{e^{-2\sqrt{\pi n}}}{n^{1/4}}\sin(2\sqrt{\pi n} + \tfrac{5}{8}\pi) + O\Big(\frac{e^{-2\sqrt{\pi n}}}{n^{3/4}}\Big),$$

as $n \to \infty$.

The most remarkable aspect of this theorem lies in the fact that the arithmetic sum $\mathcal{S}_0(-1/(1-e^{-i\delta}))$ decays exponentially fast for $\delta \to 0^+$, while the Fourier series $h(\delta)$ is very rapidly convergent. Moreover, Theorem 3 implies that $L_1(\delta)$ can be evaluated to any given precision in a time that is independent of δ.

For a rational number h/k, with $(h, k) = 1$ and $k > 0$, define

$$c_0\Big(\frac{h}{k}\Big) = -\sum_{m=1}^{k-1}\frac{m}{k}\cot\Big(\frac{\pi m h}{k}\Big).$$

The value of $c_0(h/k)$ is an algebraic number, that is, $c : \mathbb{Q} \to \overline{\mathbb{Q}}$, and, more precisely, $c_\ell(h/k)$ is contained in the maximal real subfield of the cyclotomic field of k-th roots of unity. Moreover, c_0 is odd and is periodic of period 1. See Figure 1 and Figure 2.

The cotangent sum $c_0(h/k)$ arises in analytic number theory in the value

$$D(0, h/k) = \frac{1}{4} + \frac{i}{2}c_0\Big(\frac{h}{k}\Big) \tag{12}$$

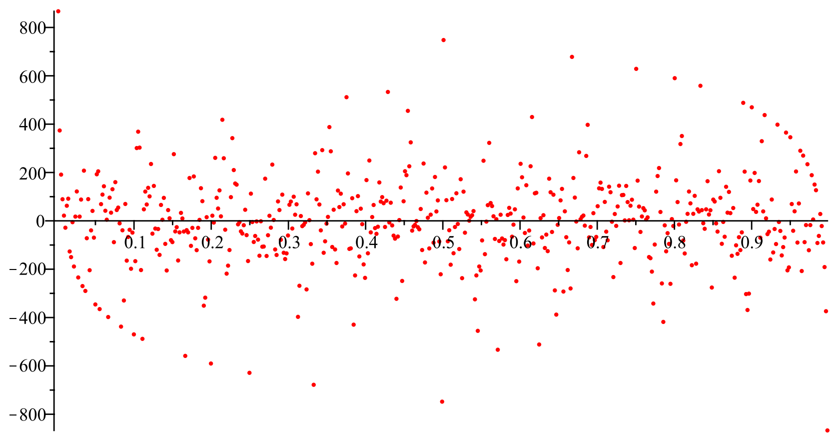


Figure 1. Graph of $c_0(h/k)$ for $1 \le h < k = 541$.

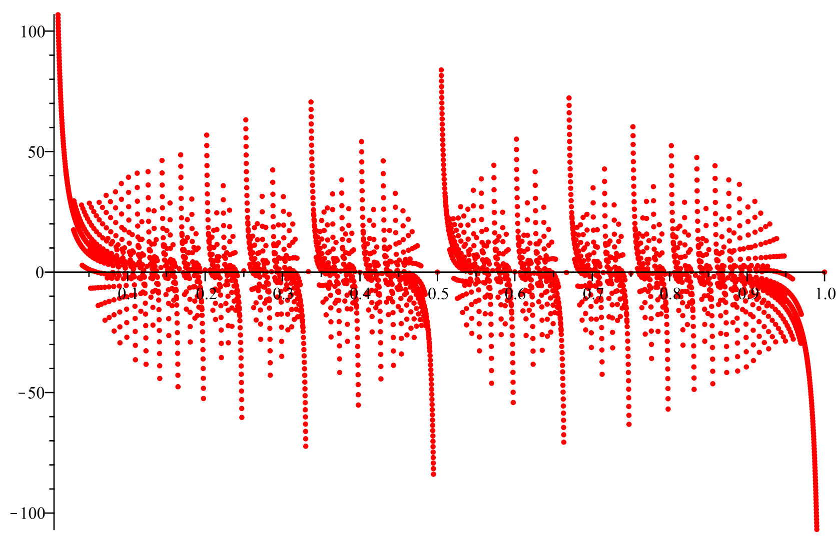


Figure 2. Graph of $c_0(h/k)$ for $1 \le h \le k \le 100$, with $(h,k) = 1$.

at $s = 0$ of the Estermann function, defined for $\mathrm{Re}(s) > 1$ by

$$D(s, h/k) := \sum_{n=1}^{\infty} \frac{d(n) \mathrm{e}(nh/k)}{n^s}.$$

The Estermann function extends analytically to $\mathbb{C} \setminus \{1\}$ and satisfies a functional equation; these properties are useful in studying the asymptotics of the mean square

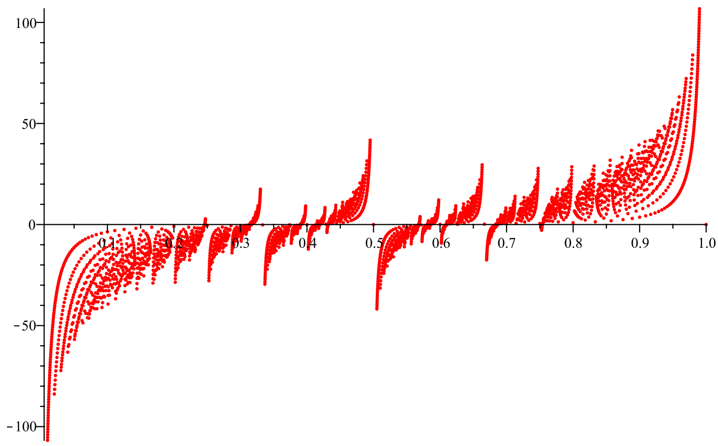


Figure 3. Graph of $V(h/k)$ for $1 \le h, k \le 100$ and $(h,k) = 1$.

of the Riemann zeta function multiplied by a Dirichlet polynomial (see [Balasubramanian et al. 1985]), which are needed, for example, for theorems that give a lower bound for the portion of zeros of $\zeta(s)$ on the critical line. See also [Conrey 1989; Iwaniec 1980]. The sum

$$V\Big(\frac{h}{k}\Big) := \sum_{m=1}^{k-1} \Big\{\frac{mh}{k}\Big\} \cot\Big(\frac{\pi m}{k}\Big) = -c_0(\bar{h}/k),$$

known as the Vasyunin sum (see Figure 3), arises in the study of the Riemann zeta function by virtue of the formula

$$\begin{aligned} \nu(h/k) &:= \frac{1}{2\pi\sqrt{hk}} \int_{-\infty}^{\infty} \big|\zeta(\tfrac{1}{2}+it)\big|^2 \Big(\frac{h}{k}\Big)^{it} \frac{dt}{\frac{1}{4}+t^2} \\ &= \frac{\log 2\pi - \gamma}{2}\Big(\frac{1}{h}+\frac{1}{k}\Big) + \frac{k-h}{2hk}\log\frac{h}{k} - \frac{\pi}{2hk}\Big(V\Big(\frac{h}{k}\Big)+V\Big(\frac{k}{h}\Big)\Big); \end{aligned} \tag{13}$$

see Figure 4.

This formula is relevant to the approach of Nyman, Beurling, Báez-Duarte and Vasyunin to the Riemann hypothesis, which asserts that the Riemann hypothesis is true if and only if $\lim_{N\to\infty} d_N = 0$, where

$$d_N^2 = \inf_{A_N} \frac{1}{2\pi} \int_{-\infty}^{\infty} \big|1 - \zeta A_N(\tfrac{1}{2}+it)\big|^2 \frac{dt}{\frac{1}{4}+t^2}$$

and the infimum is over all the Dirichlet polynomial $A_N(s) = \sum_{n=1}^{N} a_n/n^s$ of length N; see [Bagchi 2006] for a nice account of the Nyman–Beurling approach to the

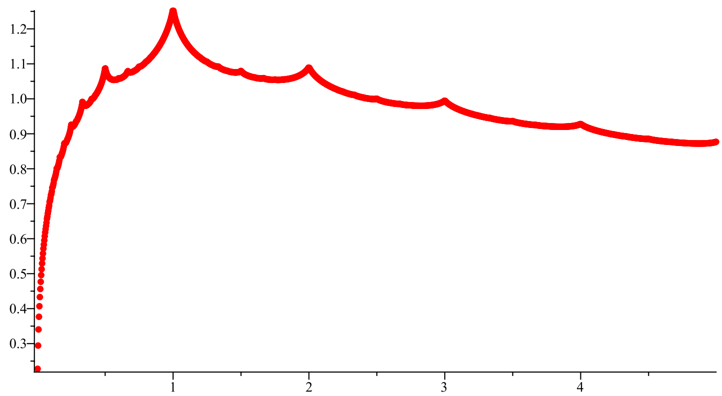


Figure 4. Graph of $\sqrt{hk}\,\nu(h/k)$ for $1 \le h \le 5k$, $k = 307$, and $(h,k) = 1$.

Riemann hypothesis with Báez-Duarte's significant contribution and see [Báez-Duarte et al. 2005; Landreau and Richard 2002] for information about the Vasyunin sums, as well as interesting numerical experiments about d_N and the minimizing polynomials A_N. Thus d_N^2 is a quadratic expression in the unknown quantities a_m in terms of the Vasyunin sums.

In [Bettin and Conrey 2011] we showed that $c_0(h/k)$ satisfies the reciprocity formula

$$c_0\Big(\frac{h}{k}\Big) + \frac{k}{h}c_0\Big(\frac{k}{h}\Big) - \frac{1}{\pi h} = \frac{i}{2}\psi_0\Big(\frac{h}{k}\Big) \tag{14}$$

(and in particular that $c_0(h/k)$ can be computed to within a prescribed accuracy in a time that is polynomial in $\log k$). See Figure 5.

This behavior is analogous to that of the Dedekind sum,

$$s\Big(\frac{h}{k}\Big) = -\frac{1}{4k}\sum_{m=1}^{k-1}\cot\Big(\frac{\pi m}{k}\Big)\cot\Big(\frac{\pi m h}{k}\Big),$$

which satisfies the well-known reciprocity formula

$$s\Big(\frac{h}{k}\Big) + s\Big(\frac{k}{h}\Big) - \frac{1}{12hk} = \frac{1}{12}\Big(\frac{h}{k} + \frac{k}{h} - 3\Big). \tag{15}$$

In this paper we prove that these results can be generalized to the sums

$$c_a\Big(\frac{h}{k}\Big) := k^a\sum_{m=1}^{k-1}\cot\Big(\frac{\pi m h}{k}\Big)\zeta\Big(-a, \frac{m}{k}\Big),$$

where $\zeta(s,x)$ is the Hurwitz zeta function (note that at $a = -1$ the poles of $\zeta(-a, m/k)$ cancel).

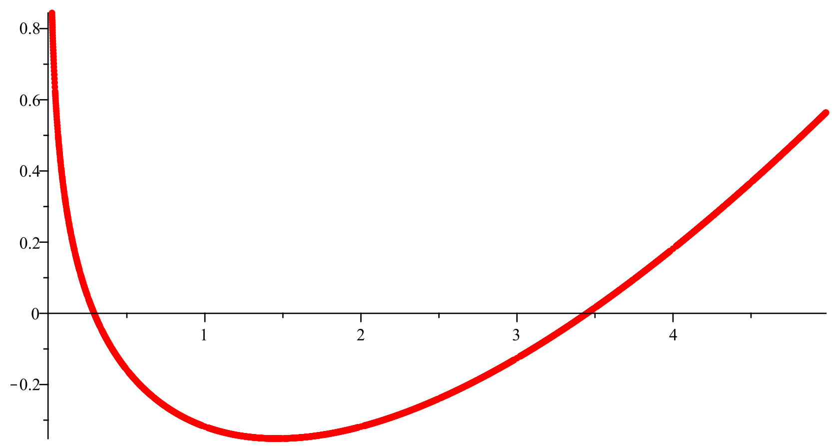


Figure 5. Graph of $c_0(h/k) + (k/h)c_0(k/h) - 1/\pi h$ for $h \le 5k$, $k \le 50$ and $(h, k) = 1$.

Notice that, for all a, $c_a(h/k)$ is odd and periodic in $x = h/k$ with period 1 and, for nonnegative integers a, it takes values in the maximal real subfield of the cyclotomic field of k-th roots of unity.

At the nonnegative integers, $a = n \ge 0$, these cotangent sums can be expressed in terms of the Bernoulli polynomials:

$$c_n\Big(\frac{h}{k}\Big) = -k^n \sum_{m=1}^{k-1} \cot\Big(\frac{\pi m h}{k}\Big) \frac{B_{n+1}(m/k)}{n+1},$$

which is most interesting when n is even, since $c_n \equiv 0$ for positive odd n.

If $a = -n$ is a negative integer one can write c_a as

$$c_{-n}\Big(\frac{h}{k}\Big) = \frac{(-1)^n}{k^n(n-1)!} \sum_{m=1}^{k-1} \cot\Big(\frac{\pi m h}{k}\Big) \Psi\Big(n-1, \frac{m}{k}\Big),$$

where

$$\Psi(m, z) := \frac{d^{m+1}}{dz^{m+1}} \log \Gamma(z)$$

is the polygamma function.

By the reflection formula for the polygamma function,

$$\Psi(m, 1-z) + (-1)^{m+1}\Psi(m, z) = (-1)^m \pi \frac{d^m}{dz^m} \cot(\pi z),$$

for a positive odd integer n we can write c_{-n} as

$$c_{-n}\Big(\frac{h}{k}\Big) = -\frac{\pi}{2k^n(n-1)!}\sum_{m=1}^{k-1}\cot\Big(\frac{\pi mh}{k}\Big)\frac{d^{n-1}}{dz^{n-1}}\cot(\pi z)\Big|_{z=m/k}$$

and, in particular,

$$c_{-1}\Big(\frac{h}{k}\Big) = 2\pi s\Big(\frac{h}{k}\Big).$$

Like the case $a=0$, these cotangent sums appear in the value

$$D\Big(0,a,\frac{h}{k}\Big) = -\frac{1}{2}\zeta(-a) + \frac{i}{2}c_a\Big(\frac{h}{k}\Big), \tag{16}$$

at $s=0$ of the function $D(s,a,h/k)$, defined for $\mathrm{Re}(s)>1$ by

$$D\Big(s,a,\frac{h}{k}\Big) := \sum_{n=1}^{\infty}\frac{\sigma_a(n)\mathrm{e}(nh/k)}{n^s}.$$

Moreover, the cotangent sums c_a appear also in a shifted version of Vasyunin's formula (13) (see Theorem 5 at the end of the paper for a new analytic proof).

Theorem 4. *Let $h,k\geq 1$, with $(h,k)=1$. Then*

$$c_a\Big(\frac{h}{k}\Big) - \Big(\frac{k}{h}\Big)^{1+a}c_a\Big(\frac{-k}{h}\Big) + k^a\frac{a\zeta(1-a)}{\pi h} = -i\zeta(-a)\psi_a\Big(\frac{h}{k}\Big). \tag{17}$$

(Note that, since $g_{-1}(z)$ is identically zero, for $a=-1$ the reciprocity formula reduces to (15).) In particular, $c_a(h/k)$ gives an example of an "imperfect" quantum modular form of weight $1+a$.

New formulas can be obtained by differentiating (17); for example, if we write

$$c_{-1}^*\Big(\frac{h}{k}\Big) := \frac{1}{k}\sum_{m=1}^{k-1}\cot\Big(\frac{\pi mh}{k}\Big)\gamma_1\Big(\frac{m}{k}\Big),$$

where $\gamma_1(x)$ is the first generalized Stieltjes constant defined by

$$\zeta(s,x) = \frac{1}{s-1} + \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\gamma_n(x)(s-1)^n,$$

then, taking the derivative at -1 of (17) multiplied by k^{-a}, we get the formula

$$c_{-1}^*\Big(\frac{h}{k}\Big) - c_{-1}^*\Big(\frac{-k}{h}\Big) + \frac{\zeta'(2)+\pi^2/6}{\pi kh} + \pi\log k\Big(\frac{1}{6}\frac{k}{h}-\frac{1}{2}\Big) = q\Big(\frac{h}{k}\Big),$$

where

$$q(z) := -\frac{1}{\pi z}\zeta'(2) + \frac{\pi}{2}(\log z+\gamma) + g_{-1}'(z)$$

is holomorphic in $\mathbb{C}'$.

3. The period function

In this section we give a proof of Theorems 1 and 2.

Proof of Theorem 1. Firstly, observe that the three-term relation (4) follows easily from the periodicity in z of $E(a,z)$.

$\mathcal{S}_a(z)$ can be written as

$$\begin{aligned}\mathcal{S}_a(z) &= \sum_{n=1}^{\infty} \sigma_a(n) \frac{1}{2\pi i} \int_{(2+\max(0,\mathrm{Re}(a)))} \Gamma(s)(-2\pi i n z)^{-s}\, ds \\ &= \frac{1}{2\pi i} \int_{(2+\max(0,\mathrm{Re}(a)))} \zeta(s)\zeta(s-a)\Gamma(s) e^{\pi i s/2} (2\pi z)^{-s}\, ds \\ &= \frac{1}{2\pi i} \int_{(-\frac{1}{2}-2M)} \zeta(s)\zeta(s-a)\Gamma(s) e^{\pi i s/2} (2\pi z)^{-s}\, ds + r_{a,M}(z),\end{aligned} \tag{18}$$

where M is any integer greater or equal to $-\frac{1}{2}\min(0,\mathrm{Re}(a))$ and

$$\begin{aligned}r_{a,M}(z) := -\tfrac{1}{2}\zeta(-a) + i\frac{\zeta(1-a)}{2\pi z} + i\frac{\zeta(1+a)\Gamma(1+a)e^{\pi i a/2}}{(2\pi z)^{1+a}} \\ - \sum_{1\le n\le M} i(-1)^n \frac{B_{2n}}{(2n)!}\zeta(1-2n-a)(2\pi z)^{2n-1}\end{aligned}$$

is the sum of the residues encountered moving the integral (and has to be interpreted in the limit sense if some of the terms have a pole). Now, consider

$$\begin{aligned}\frac{1}{z^{1+a}}\mathcal{S}_a\Big(-\frac{1}{z}\Big) &= \frac{1}{z^{1+a}}\frac{1}{2\pi i}\int_{(2+\max(0,\mathrm{Re}(a)))} \zeta(s)\zeta(s-a)\Gamma(s)e^{\pi i s/2}\Big(2\pi\frac{-1}{z}\Big)^{-s} ds \\ &= \frac{1}{2\pi i}\int_{(2+\max(0,\mathrm{Re}(a)))} \zeta(s)\zeta(s-a)\Gamma(s)e^{-\pi i s/2}(2\pi)^{-s}z^{s-1-a}\, ds,\end{aligned}$$

since in this context $0<\arg z<\pi$ and $0<\arg -1/z<\pi$, so $\arg -1/z = \pi - \arg z$. Applying the functional equation to both $\zeta(s)$ and $\zeta(s-a)$ we get, after the change of variable $s \to 1-s+a$,

$$\begin{aligned}&\frac{1}{z^{1+a}}\mathcal{S}_a\Big(-\frac{1}{z}\Big) \\ &\quad = -\frac{1}{2\pi}\int_{(-1+\min(0,\mathrm{Re}(a)))} \zeta(s-a)\zeta(s)\Gamma(s)\frac{e^{\pi i(s-a)/2}\cos\frac{\pi s}{2}}{\sin\frac{\pi(s-a)}{2}}(2\pi z)^{-s}\, ds \\ &\quad = -\frac{1}{2\pi}\int_{(-1/2-M)} \zeta(s-a)\zeta(s)\Gamma(s)\frac{e^{\pi i(s-a)/2}\cos\frac{\pi s}{2}}{\sin\frac{\pi(s-a)}{2}}(2\pi z)^{-s}\, ds,\end{aligned} \tag{19}$$

since the integrand doesn't have any pole on the left of $-1+\min(0,\mathrm{Re}(a))$. The theorem then follows summing (18) and (19) and using the identity

$$e^{\pi i s/2}+i\frac{e^{\pi i(s-a)/2}\cos\frac{\pi s}{2}}{\sin\frac{\pi(s-a)}{2}}=i\frac{\cos\frac{\pi a}{2}}{\sin\frac{\pi(s-a)}{2}}.\qquad\square$$

We remark that for $a=2k+1$, with $k\geq 1$, Theorem 1 reduces to

$$E_{2k}(z)-\frac{1}{z^{2k}}E_{2k}\Big(-\frac{1}{z}\Big)=0,$$

while, for $a=1$, the theorem reduces to the well-known identity

$$E_2(z)-\frac{1}{z^2}E_2\Big(\frac{-1}{z}\Big)=-\frac{12}{2\pi i z}.$$

To prove Theorem 2 we need the following lemma.

Lemma 1. *For fixed complex numbers A and α we have, as $n\to\infty$*

$$\begin{aligned}J_n &:= \int_0^\infty u^{n+\alpha}e^{-A\sqrt{u}}e^{-u}\frac{du}{u}\\ &= \sqrt{2\pi}e^{A^2/8}e^{-A\sqrt{n}}e^{-n}n^{n+\alpha-1/2}\Big(1-\frac{C}{\sqrt{n}}+O\Big(\frac{1}{n}\Big)\Big),\end{aligned}$$

where

$$C=\frac{4\alpha-1}{8}A+\frac{A^3}{96}.$$

Proof. After the change of variable $u=nx^2$, we have

$$\begin{aligned}J_n &= 2n^{n+\alpha}\int_0^\infty x^{2\alpha-1}e^{-A\sqrt{n}x-n(x^2-2\log x)}dx\\ &= 2n^{n+\alpha}e^{-A\sqrt{n}}\int_{-1}^\infty (x+1)^{2\alpha-1}e^{-A\sqrt{n}x-n((x+1)^2-2\log(x+1))}dx\\ &= 2n^{n+\alpha}e^{-A\sqrt{n}}e^{-n}(1+O(e^{-n\delta^2/2}))\\ &\qquad\times\int_{-\delta}^{\delta}(x+1)^{2\alpha-1}e^{-A\sqrt{n}x-2nx^2}(1+\tfrac{2}{3}nx^3+O(nx^4))dx\end{aligned}$$

for any small $\delta>0$. We can then approximate the binomial and extend the integral to $\mathbb{R}$ at a negligible cost, getting

$$J_n=2n^{n+\alpha}e^{-A\sqrt{n}}e^{-n}\int_{-\infty}^{\infty}(1+(2\alpha-1)x+\tfrac{2}{3}nx^3+O(x^2+nx^4))\times e^{-A\sqrt{n}x-2nx^2}dx.$$

Evaluating the integrals, the lemma follows. $\square$

Proof of Theorem 2. The three-term relation (4) implies that

$$g_a(z+1) = \frac{1}{(z+1)^{1+a}}\cot\Big(\frac{\pi a}{2}\Big)\zeta(-a) - \frac{1}{\pi z(z+1)^a}\zeta(1-a) + \frac{1}{\pi z(z+1)}\zeta(1-a) + g_a(z) - \frac{1}{(z+1)^{1+a}}g_a\Big(\frac{z}{z+1}\Big).$$

Now, from the definition (5) of $g_a(z)$, it follows that

$$g_a(z) = 2\sum_{1\le n\le M}(-1)^n\frac{B_{2n}}{(2n)!}\zeta(1-2n-a)(2\pi z)^{2n-1} + O(|z|^{2M+1/2})$$

for any $M \ge 1$. Thus

$$\begin{aligned} g_a(z) - \frac{g_a(z/(z+1))}{(z+1)^{1+a}} &= 2\sum_{1\le n\le M}(-1)^n\frac{B_{2n}}{(2n)!}\zeta(1-2n-a)(2\pi z)^{2n-1}\Big(1-\frac{1}{(z+1)^{2n+a}}\Big)+O(|z|^{2M+1/2}) \\ &= -2\sum_{m=1}^{2M}\Bigg(\sum_{\substack{2n-1+k=m,\\ n,k\ge 1}}(-1)^{n+m}B_{2n}\zeta(1-2n-a)\frac{\Gamma(2n+a+k)}{\Gamma(2n+a)k!(2n)!}(2\pi)^{2n-1}\Bigg)z^m \\ &\quad + O(|z|^{2M+\frac{1}{2}}). \end{aligned}$$

Therefore,

$$g_a(z+1) = \sum_{m=0}^{2M} b_m z^m + O(|z|^{2M+1/2}),$$

where

$$\begin{aligned} b_m := -2\sum_{\substack{2n-1+k=m,\\ n,k\ge 1}}&(-1)^{n+k}B_{2n}\zeta(1-2n-a)\frac{\Gamma(2n+a+k)}{\Gamma(2n+a)k!(2n)!}(2\pi)^{2n-1} \\ &+(-1)^m\cot\Big(\frac{\pi a}{2}\Big)\zeta(-a)\frac{\Gamma(1+a+m)}{\Gamma(1+a)m!} \\ &+(-1)^m\Big(\frac{\Gamma(1+a+m)}{\Gamma(a)(m+1)!}-1\Big)\frac{\zeta(1-a)}{\pi}, \end{aligned}$$

and, since $g_a(z)$ is holomorphic at 1, b_m must coincide with the m-th coefficient of the Taylor series of $g_a(z)$ at 1.

Now, let's prove the asymptotic (8). Fix any $M \ge -\frac{1}{2}\min(0, \mathrm{Re}(a))$ and assume $m \ge 2M+1$ and $\mathrm{Re}(\tau) > 0$. By the functional equation for ζ and basic properties of $\Gamma(s)$, we have

$$\begin{aligned}
&\frac{(2\pi)^a\tau^m}{\cos\frac{\pi a}{2}}g_a^{(m)}(\tau)\\
&\quad=\frac{(-1)^m}{\pi i}\int_{(-\frac12-2M)}\Gamma(s)\frac{\zeta(s)\zeta(s-a)}{\sin\frac{\pi(s-a)}{2}}s(s+1)\cdots(s+m-1)(2\pi)^{-s+a}\tau^{-s}\,ds\\
&\quad=\frac{(-1)^m}{\pi i}\int_{(-\frac12-2M)}\frac{\zeta(s)\zeta(s-a)}{\sin\frac{\pi(s-a)}{2}}\Gamma(s+m)(2\pi)^{-s+a}\tau^{-s}\,ds\\
&\quad=\frac{(-1)^m}{\pi^3 i}\int_{(-\frac12-2M)}\zeta(1-s)\zeta(1-s+a)\\
&\qquad\qquad\times\Gamma(1-s)\Gamma(1-s+a)\Gamma(s+m)\sin\Big(\frac{\pi s}{2}\Big)\Big(\frac{2\pi}{\tau}\Big)^s\,ds.
\end{aligned}$$

We can see immediately that $g_a^{(m)}(\tau)\ll_a m^{-B}|\tau|^{-m}m!$ for any fixed $B>0$, just by moving the path of integration to the line $\mathrm{Re}(s)=-B$ and using trivial estimates for Γ. To get a formula asymptotic as $m\to\infty$, we expand $\zeta(1-s)\zeta(1-s+a)$ into a Dirichlet series and integrate term-by-term; the main term arises from the first term of the sum. We have

$$g_a^{(m)}(\tau)=2\frac{(-\tau)^{-m}\cos\frac{\pi a}{2}}{\pi^2(2\pi)^a}\sum_{\ell=1}^{\infty}\frac{\sigma_{-a}(\ell)}{\ell}I_{m,a}\Big(\frac{\ell}{\tau}\Big),$$

where

$$I_{m,a}(x):=\frac{1}{2\pi i}\int_{(-\frac12-2M)}\Gamma(1-s)\Gamma(1-s+a)\Gamma(s+m)\sin\Big(\frac{\pi s}{2}\Big)(2\pi x)^s\,ds.$$

We reexpress this integral as a convolution integral. Recall that for $|\arg x|<\pi$ we have

$$\frac{1}{2\pi i}\int_{(\frac32+2M)}\Gamma(s)\Gamma(s+a)u^{-s}\,ds=2u^{a/2}K_a(2\sqrt{u}),$$

where K_a denotes the K-Bessel function of order a. Also,

$$\frac{1}{2\pi i}\int_{(-\frac12-2M)}\Gamma(s+m)u^{-s}\,ds=u^m e^{-u}.$$

Thus,

$$I_{m,a}(x)=I_{m,a}^{+}(x)+I_{m,a}^{-}(x),$$

where

$$I_{m,a}^{\pm}(x)=(2\pi x)^{1+a/2}e^{\pm\pi i a/4}\int_0^\infty u^{m+a/2}K_a(2e^{\pm\pi i/4}\sqrt{2\pi xu})e^{-u}\,du.$$

Now, for $|\arg z| < \frac{3}{2}\pi$

$$K_a(z) = \sqrt{\frac{\pi}{2z}}e^{-z}\Big(1 + \frac{4a^2-1}{8z} + O_a\Big(\frac{1}{|z|^2}\Big)\Big),$$

as $z \to \infty$, and

$$K_{-a}(z) = K_a(z) \sim \begin{cases} 2^{a-1}\Gamma(a)z^{-a} & \text{if } \mathrm{Re}(a) \geq 0,\ a \neq 0, \\ -\log(x/2) - \gamma & \text{if } a = 0, \end{cases}$$

as $z \to 0$. Therefore, by Lemma 1,

$$\begin{aligned} I^{\pm}_{m,a}(x) = (2\pi x)^{1+\frac{a}{2}} \frac{\pi^{\frac{1}{4}} e^{\pm\pi i(a-\frac{1}{2})/4}}{2^{\frac{5}{4}} x^{\frac{1}{4}}} \int_0^\infty & u^{m+\frac{a}{2}-\frac{1}{4}} e^{-u-2(1\pm i)\sqrt{\pi x u}} \\ & \times \Big(1 + \frac{4a^2-1}{2^{\frac{9}{2}}\pi^{\frac{1}{2}} e^{\pm\frac{\pi i}{4}}\sqrt{xu}} + O_a\Big(\frac{1}{u}\Big)\Big)\, du \\ \sim 2^{\frac{1}{4}+\frac{a}{2}} \pi^{\frac{7}{4}+\frac{a}{2}} e^{\pm\pi i(a-\frac{1}{2})/4} x^{\frac{3}{4}+\frac{a}{2}} & e^{\pm i\pi x} e^{-2(1\pm i)\sqrt{\pi x n}} e^{-m} m^{m+\frac{1}{4}+\frac{a}{2}} \\ & \times \Big(1 + \frac{\xi^{\pm}}{\sqrt{m}} + O\Big(\frac{1}{m}\Big)\Big), \end{aligned}$$

where

$$\xi^{\pm} = -\frac{(1\pm i)\sqrt{\pi x}(1+a)}{2} + \frac{(1\mp i)(\pi x)^{\frac{3}{2}}}{6} + \frac{(4a^2-1)(1\mp i)}{32\pi^{\frac{1}{2}}\sqrt{x}},$$

and (8) follows. □

4. An extension of Voronoi's formula

Formula (10) can be proved with the same techniques used to prove Theorems 1 and 2. In this section we give an application of this formula and we discuss a similar formula for convolutions of the exponential function. We conclude the section showing how these results can be used to prove Voronoi's formula.

Applying formula (10) to $F(s) = \Gamma(s/2)/2\Gamma(s)$ we get, for $\frac{1}{4}\pi < \arg(z) < \frac{3}{4}\pi$,

$$\sum_{n=1}^{\infty} d(n) e^{(2\pi n z)^2} = \frac{1}{z}\sum_{n=1}^{\infty} d(n) T(4\pi n z) + R(z) + k(z), \tag{20}$$

where, in the same range of $\arg(z)$,

$$T(z) := \frac{1}{\sqrt{\pi} i}\int_{(2)} \frac{\Gamma(s)}{\Gamma(1-s/2)}(-iz)^{-s}\, ds = \sum_{n=0}^{\infty} \frac{(iz)^n}{n!\,\Gamma(1+n/2)}$$

and

$$R(z) := \frac{1}{4} + \frac{2\log(-4\pi i z) - 3\gamma}{8\sqrt{\pi} i z},$$
$$k(z) := \frac{1}{4\pi^2} \int_{(-\frac{1}{2})} \Gamma(s/2)\Gamma(1-s)\zeta(s)\zeta(1-s)z^{-s}\, ds.$$

Notice that we have $T(z) \ll |z|^{-B}$ for all fixed $B > 0$; moreover, $k(z)$ is holomorphic in $|\arg(z)| < \frac{3}{4}\pi$ and, if $|\arg(\tau)| < \frac{1}{4}\pi$,

$$c_\tau(m) := \frac{k^{(m)}(\tau)}{m!} \ll |\tau|^{-m} m^{-B}$$

for all $B > 0$. In particular, if we set $z = i\delta$ with $0 < \delta \le 1$, taking the real part of (20) we get

$$\sum_{n=1}^{\infty} d(n) e^{-(2\pi n\delta)^2} = \frac{1}{4} + \frac{-2\log(4\pi\delta) - 3\gamma}{4\sqrt{\pi}\delta} + \mathrm{Re} \sum_{m=0}^{\infty} c_m \Big(\frac{\sqrt{3}}{2} + i\Big(\frac{1}{2} - \delta\Big)\Big)^m \quad (21)$$

with

$$c_m := c_{(\sqrt{3}+i)/2}(m) \ll m^{-B}$$

for all $B > 0$.

We now state a similar formula for convolutions of the exponential function and a function that is compactly supported on $\mathbb{R}_{>0}$.

Let $g(x)$ be a compactly supported function on $\mathbb{R}_{>0}$ and let

$$W_+(z) := \int_0^\infty f(1/x)\mathrm{e}(zx)\,\frac{dx}{x} \quad \text{and} \quad W_-(z) := \int_0^\infty f(x)\mathrm{e}(zx)\, dx.$$

If we denote the Mellin transform of $f(x)$ with $F(s)$, then it follows that $F(s)$ is entire and that $W_+(x)$ and $W_-(x)$ can be written as in (9). In particular, since

$$F(0) = \int_0^\infty f(x)\,\frac{dx}{x}, \quad F(1) = \int_0^\infty f(x)\, dx, \quad F'(1) = \int_0^\infty f(x)\log x\, dx,$$

formula (10) can be written as

$$\sum_{n=1}^{\infty} d(n) W_+(nz) - \frac{1}{z}\sum_{n=1}^{\infty} d(n) W_-(-n/z)$$
$$= \int_0^\infty f(x)\Big(\frac{1}{4x} - \frac{1}{4z} - \frac{\gamma - \log(2\pi z/x)}{2\pi i z}\Big)\, dx + k(z)$$
$$+ \int_0^\infty f(x) \int_{(-\frac{1}{2})} \frac{\zeta(s)\zeta(1-s)}{\sin \pi s}\Big(\frac{z}{x}\Big)^{-s} ds\, \frac{dx}{2\pi x} \quad (22)$$

for $\mathrm{Im}(z) > 0$.

Proof of Voronoi's formula. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be a smooth function that decays faster than any power of x and let

$$\tilde{f}(x) := 2 \int_0^\infty f(y) \cos(2\pi xy)\, dy$$

be the cosine transform of $f(x)$. Then, $\tilde{f}(x)$ is smooth and, by partial integration, $\tilde{f}^{(m)}(x) \ll 1/x^{2+m}$ for all $m \geq 0$. For $0 < \mathrm{Re}(s) < 2$, we can define the Mellin transform of $\tilde{f}$,

$$F(s) := \int_0^\infty \tilde{f}(x) x^{s-1}\, dx.$$

By partial integration we see that $F(s)$ extends to a meromorphic function on $\mathrm{Re}(s) < 2$ with simple poles at most at the nonpositive integers. Also, $F(s)$ decays rapidly on vertical strips. Moreover, by Parseval's formula, for $0 < \mathrm{Re}(s) < 1$ we have

$$\begin{aligned} F(s) &= \frac{2}{s} \int_0^\infty f(y)(2\pi y)^{-s} \Gamma(s+1) \cos\Big(\frac{\pi s}{2}\Big)\, dy \\ &= \frac{2}{s} \int_0^\infty f(y)\, dy - 2 \int_0^\infty f(y)(\log(2\pi y) + \gamma)\, dy + O(|s|) \\ &= \frac{F_{-1}}{s} + F_0 + O(|s|), \end{aligned}$$

say. For $\mathrm{Im}(z) \geq 0$ we can define

$$\begin{aligned} W_+(z) &:= \frac{1}{2\pi i} \int_{(\frac{3}{2})} F(s)\Gamma(s)(-2\pi i z)^{-s}\, ds = \int_0^\infty \tilde{f}\Big(\frac{1}{x}\Big) \mathrm{e}(zx)\, \frac{dx}{x}, \\ W_-(z) &:= \frac{1}{2\pi i} \int_{(\frac{3}{2})} F(1-s)\Gamma(s)(-2\pi i z)^{-s}\, ds \\ &= \int_0^\infty (\tilde{f}(x) - \mathrm{Res}_{s=0} F(s)) \mathrm{e}(zx)\, dx, \end{aligned} \tag{23}$$

with the second representation of $W_-(z)$ defined only on $\mathrm{Im}(z) > 0$. Since $F(s)$ is rapidly decaying at infinity, (10) holds for $\mathrm{Im}(z) \geq 0$ and so we can apply that formula for $z = 1$ and take the real part. By the definition of $\tilde{f}$, we have

$$\begin{aligned} \mathrm{Re}(W_+(n)) &= 2 \int_0^\infty f(y) \int_0^\infty \cos\Big(\frac{2\pi y}{x}\Big) \cos(nx)\, \frac{dx}{x}\, dy \\ &= \int_0^\infty f(y)(2K_0(4\pi\sqrt{ny}) - \pi Y_0(4\pi\sqrt{ny}))\, dy \end{aligned}$$

and

$$\begin{aligned}\mathrm{Re}(W_-(-n)) &= \lim_{\substack{z\to 1,\\ \mathrm{Im}(z)>0}} \mathrm{Re}(W_-(-nz))\\ &= \lim_{\substack{z\to 1,\\ \mathrm{Im}(z)>0}} \mathrm{Re}\int_0^\infty \tilde{f}(x)\mathrm{e}(-nzx)\,dx - \lim_{\substack{z\to 1,\\ \mathrm{Im}(z)>0}} \mathrm{Re}\,\frac{\mathrm{Res}_{s=0} F(s)}{-2\pi i n z}\\ &= \tfrac{1}{2} f(n),\end{aligned}$$

since $\mathrm{Res}_{s=0} F(s)$ is real. Moreover, $(2\pi)^{-1}\int_{(-1/2)} F(s)\frac{\zeta(s)\zeta(1-s)}{\sin \pi s} z^{-s}\,ds$ is purely imaginary on the real line, so we just need to compute

$$\begin{aligned}&\mathrm{Re}\big(\mathrm{Res}_{s=0,1} F(s)\Gamma(s)\zeta(s)^2(-2\pi i)^{-s}\big)\\ &\quad= \mathrm{Re}\bigg(\frac{F(1)(\gamma-\log(-2\pi i))+F'(1)}{-2\pi i}\\ &\qquad\qquad + \frac{-F_{-1}(\log(-2\pi i)+\gamma-2\log 2\pi)+F_0}{4}\bigg)\\ &\quad= -\frac{f(0)}{8} - \frac{1}{2}\int_0^\infty f(y)(\log y + 2\gamma)\,dy,\end{aligned}$$

since $F(1) = f(0)/2$ and $F'(1)$ is real. This completes the proof of the theorem. □

5. An exact formula for the second moment of $\zeta(s)$

In this section we prove the exact formula for the second moment of the Riemann zeta function.

Proof of Theorem 3. Firstly, observe that

$$L_2(\delta) = -ie^{-i\delta/2}\int_{\frac{1}{2}}^{\frac{1}{2}+i\infty} \zeta(s)\zeta(1-s)e^{i\delta s}\,ds.$$

The functional equation for $\zeta(s)$,

$$\zeta(1-s) = \chi(1-s)\zeta(s),$$

where

$$\chi(1-s) = (2\pi)^{-s}\Gamma(s)(e^{\pi i s/2} + e^{-\pi i s/2}),$$

allows us to split $L_2(\delta)$ as

$$L_2(\delta) = -ie^{-i\delta/2}\int_{\frac{1}{2}}^{\frac{1}{2}+i\infty} \chi(1-s)\zeta(s)^2 e^{i\delta s}\,ds = -ie^{-i\delta/2}(L^+(\delta)+L^-(\delta)),$$

where

$$L^{\pm}(\delta) = \int_{\frac{1}{2}}^{\frac{1}{2}+i\infty} (2\pi)^{-s}\Gamma(s)e^{\pm\pi is/2}\zeta(s)^2 e^{i\delta s}\, ds.$$

By Stirling's formula, $L^+(\delta)$ is analytic for $\mathrm{Re}(\delta) > -\pi$. Moreover, by contour integration,

$$L^-(\delta) = \int_{(2)} (2\pi)^{-s}\Gamma(s)e^{-\pi is/2}\zeta(s)^2 e^{i\delta s}\, ds - G(\delta) = J(\delta) - G(\delta),$$

say, where

$$G(\delta) := \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}} (2\pi)^{-s}\Gamma(s)e^{-\pi is/2}\zeta(s)^2 e^{i\delta s}\, ds + 2\pi i \,\mathrm{Res}_{s=1}\big((2\pi)^{-s}\Gamma(s)e^{-\pi is/2}\zeta(s)^2 e^{i\delta s}\big)$$

is analytic for $\mathrm{Re}(\delta) < \pi$. Now, expanding $\zeta(s)^2$ into its Dirichlet series, for $\mathrm{Re}(\delta) > 0$ we have

$$\begin{aligned} J(\delta) &= \sum_{n=1}^{\infty} d(n) \int_{2-i\infty}^{2+i\infty} \Gamma(s)(2\pi i n e^{-i\delta})^{-s}\, ds \\ &= 2\pi i \mathscr{S}_0(-e^{-i\delta}) = 2\pi i \mathscr{S}_0(1 - e^{-i\delta}). \end{aligned} \tag{24}$$

By Theorem 1, we can write this as

$$J(\delta) = \frac{\log 2\pi\delta - \gamma}{1 - e^{-i\delta}} - \pi g_0(1 - e^{-i\delta}) + \frac{2\pi i}{1 - e^{-i\delta}}\mathscr{S}_0\left(\frac{-1}{1 - e^{-i\delta}}\right) + ie^{i\delta}\omega(\delta),$$

where

$$\omega(\delta) = -\frac{\log((1 - e^{-i\delta})/\delta) - \frac{1}{2}\pi i}{2\sin(\delta/2)}$$

is holomorphic in $|\mathrm{Re}(\delta)| < \pi$. Summing up, we have

$$L_2(\delta) = \frac{\gamma - \log 2\pi\delta}{2\sin(\delta/2)} + \frac{\pi i}{\sin(\delta/2)}\mathscr{S}_0\left(\frac{-1}{1 - e^{-i\delta}}\right) + i\pi e^{-i\delta/2} g_0(1 - e^{-i\delta}) + \omega(\delta) - ie^{-i\delta/2}(L^+(\delta) - G(\delta)). \tag{25}$$

The theorem then follows after writing

$$h(\delta) := i\pi e^{-i\delta/2} g_0(1 - e^{-i\delta})$$

and applying Theorems 1 and 2. □

6. Cotangent sums

We start by recalling the basic properties of $D(s,a,h/k)$.

Lemma 2. *For* $(h,k)=1$, $k>0$ *and* $a\in\mathbb{C}$,

$$D\Big(s,a,\frac{h}{k}\Big)-k^{1+a-2s}\zeta(s-a)\zeta(s)$$

is an entire function of s. *Moreover*, $D(s,a,h/k)$ *satisfies a functional equation*,

$$\begin{aligned}D\Big(s,a,\frac{h}{k}\Big)\\ &=-\frac{2}{k}\Big(\frac{k}{2\pi}\Big)^{2-2s+a}\Gamma(1-s+a)\Gamma(1-s)\\ &\quad\times\Big(\cos\Big(\frac{\pi}{2}(2s-a)\Big)D\Big(1-s,-a,-\frac{\overline{h}}{k}\Big)-\cos\frac{\pi a}{2}\,D(1-s,-a,\frac{\overline{h}}{k})\Big),\end{aligned}\tag{26}$$

and

$$D\Big(0,a,\frac{h}{k}\Big)=\frac{i}{2}c_a\Big(\frac{h}{k}\Big)-\frac{1}{2}\zeta(-a).$$

Proof. The analytic continuation and the functional equation for $D(s,a,h/k)$ can be proved easily using the analogous properties for the Hurwitz zeta function and the observation that

$$D\Big(s,a,\frac{h}{k}\Big)=\frac{1}{k^{2s-a}}\sum_{m,n=1}^{k}\mathrm{e}\Big(\frac{mnh}{k}\Big)\zeta\Big(s-a,\frac{m}{k}\Big)\zeta\Big(s,\frac{n}{k}\Big).$$

Moreover, applying this equality at 0, we see that

$$\begin{aligned}D\Big(0,a,\frac{h}{k}\Big)&=-k^a\sum_{m,n=1}^{k-1}\mathrm{e}\Big(\frac{mnh}{k}\Big)\zeta\Big(-a,\frac{m}{k}\Big)B_1\Big(\frac{n}{k}\Big)-\frac{\zeta(-a)}{2}\\&=\frac{i}{2}c_a\Big(\frac{h}{k}\Big)-\frac{\zeta(-a)}{2},\end{aligned}$$

where we used

$$\sum_{n=1}^{k-1}B_1\Big(\frac{n}{k}\Big)\Big(\mathrm{e}\Big(\frac{mh}{k}\Big)\Big)^n=-\frac{1}{2}\frac{1+\mathrm{e}(\frac{mh}{k})}{1-\mathrm{e}(\frac{mh}{k})}=-\frac{i}{2}\cot\Big(\frac{\pi mh}{k}\Big),$$

which can be easily obtained from the equality

$$B_1(x)=\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{te^{xt}}{e^t-1}\Big)\Big|_{t=0}.$$

□

Proof of Theorem 4. First observe that we can assume $0\neq|a|<1$, since the result extends to all a by analytic continuation. Now, taking $z=\frac{h}{k}(1+i\delta)$, with $\delta>0$, we have

$$\begin{aligned}\mathcal{S}_a(z) &= \sum_{n\geq 1}\sigma_a(n)\mathrm{e}\Big(n\frac{h}{k}\Big)e^{-2\pi n(h/k)\delta}\\ &= \frac{1}{2\pi i}\int_{(2)}\Gamma(s)D\Big(s,a,\frac{h}{k}\Big)\Big(2\pi\frac{h}{k}\delta\Big)^{-s}ds.\end{aligned}$$

Therefore, moving the integral to $\sigma = -\frac{1}{2}$,

$$\mathcal{S}_a(z) = \frac{k^a}{2\pi h\delta}\zeta(1-a) + \frac{1}{(2\pi h\delta)^{1+a}}\zeta(1+a)\Gamma(1+a) + D\Big(0,a,\frac{h}{k}\Big) + O(\delta^{1/2}).$$

Similarly,

$$\begin{aligned}\frac{1}{z^{1+a}}\mathcal{S}_a\Big(\frac{-1}{z}\Big) &= \frac{1}{z^{1+a}}\sum_{n\geq 1}\sigma_a(n)\mathrm{e}\Big(-n\frac{k}{h}\Big)e^{-2\pi n(k/h)\delta/(1+i\delta)}\\ &= \frac{k^a}{2\pi\delta h}\zeta(1-a) + \frac{1}{(2\pi\delta h)^{1+a}}\zeta(1+a)\Gamma(1+a)\\ &\qquad - ia\frac{k^a}{2\pi h}\zeta(1-a) + \Big(\frac{k}{h(1+i\delta)}\Big)^{1+a}D\Big(0,a,-\frac{k}{h}\Big) + O(\delta^{1/2}).\end{aligned}$$

In particular, as δ goes to 0, we have

$$\mathcal{S}_a(z) - \frac{1}{z^{1+a}}\mathcal{S}_a\Big(\frac{-1}{z}\Big) \longrightarrow D\Big(0,a,\frac{h}{k}\Big) - \Big(\frac{k}{h}\Big)^{1+a}D\Big(0,a,-\frac{k}{h}\Big) + ia\frac{k^a}{2\pi h}\zeta(1-a).$$

Applying Theorem 1, it follows that

$$\begin{aligned}&D\Big(0,a,\frac{h}{k}\Big) - \Big(\frac{k}{h}\Big)^{1+a}D\Big(0,a,-\frac{k}{h}\Big) + ia\frac{k^a}{2\pi h}\zeta(1-a)\\ &\qquad = \frac{\zeta(-a)}{2}\Big(\Big(\frac{k}{h}\Big)^{1+a} - 1 + \psi_a\Big(\frac{h}{k}\Big)\Big),\end{aligned}$$

which is equivalent to (17). □

We conclude the paper by giving a new proof of Vasyunin's formula (with a shift).

Theorem 5. *Let* $(h,k)=1$, *with* $h,k\geq 1$. *Let* $|\mathrm{Re}(a)|<1$. *Then*

$$\begin{aligned}&\frac{1+a}{2\pi}\int_{-\infty}^{\infty}\zeta\Big(\frac{1}{2}+\frac{a}{2}+it\Big)\zeta\Big(\frac{1}{2}+\frac{a}{2}-it\Big)\Big(\frac{h}{k}\Big)^{-it}\frac{dt}{(\frac{1}{2}+\frac{a}{2}+it)(\frac{1}{2}+\frac{a}{2}-it)}\\ &\quad = -\frac{\zeta(1+a)}{2}\Big(\Big(\frac{k}{h}\Big)^{\frac{1}{2}+\frac{a}{2}} + \Big(\frac{h}{k}\Big)^{\frac{1}{2}+\frac{a}{2}}\Big) - \frac{\zeta(a)}{a}\Big(\Big(\frac{k}{h}\Big)^{\frac{1}{2}-\frac{a}{2}} + \Big(\frac{h}{k}\Big)^{\frac{1}{2}-\frac{a}{2}}\Big)\\ &\qquad - \Big(\frac{1}{hk}\Big)^{\frac{1}{2}+\frac{a}{2}}(2\pi)^a\Gamma(-a)\sin\frac{\pi a}{2}\Big(c_a\Big(\frac{\overline{h}}{k}\Big) + c_a\Big(\frac{\overline{k}}{h}\Big)\Big).\end{aligned}$$

Proof. We need to evaluate

$$\frac{1+a}{2\pi(hk)^{\frac{1}{2}+\frac{a}{2}}}\int_{-\infty}^{\infty}\zeta\Big(\frac{1}{2}+\frac{a}{2}+it\Big)\zeta\Big(\frac{1}{2}+\frac{a}{2}-it\Big)\Big(\frac{h}{k}\Big)^{it}\frac{dt}{(\frac{1}{2}+\frac{a}{2}+it)(\frac{1}{2}+\frac{a}{2}-it)}$$
$$=\frac{1+a}{2\pi i}\int_{(\frac{1}{2}-\frac{\mathrm{Re}(a)}{2})}\frac{\zeta(s+a)\zeta(1-s)}{h^{s+a}k^{1-s}}\frac{ds}{(s+a)(1-s)}.$$

We rewrite this as

$$\frac{1+a}{2\pi i}\int_{(\frac{1}{2}-\frac{\mathrm{Re}(a)}{2})}\frac{\zeta(s+a)\zeta(1-s)}{h^{s+a}k^{1-s}}\frac{ds}{(s+a)(1-s)}$$
$$=\frac{1}{2\pi i}\int_{(\frac{1}{2}-\frac{\mathrm{Re}(a)}{2})}\frac{\zeta(s+a)\zeta(1-s)}{h^{s+a}k^{1-s}}\frac{ds}{1-s}+\frac{1}{2\pi i}\int_{(\frac{1}{2}-\frac{\mathrm{Re}(a)}{2})}\frac{\zeta(s+a)\zeta(1-s)}{h^{s+a}k^{1-s}}\frac{ds}{s+a}$$
$$=I_a\Big(\frac{h}{k}\Big)+I_a\Big(\frac{k}{h}\Big),$$

where

$$I_a\Big(\frac{h}{k}\Big):=\frac{1}{2\pi i}\int_{(\frac{1}{2}-\frac{\mathrm{Re}(a)}{2})}\frac{\zeta(s+a)\zeta(1-s)}{h^{s+a}k^{1-s}}\frac{ds}{1-s}.$$

The integral is not absolutely convergent, so some care is needed. One could introduce a convergence factor $e^{\delta s^2}$ and let $\delta\to 0^+$ at the end of the argument, or one could work with the understanding that the integrals are to be interpreted as $\lim_{T\to\infty}\int_{c-iT}^{c+iT}$. We opt for the latter. Recall that $\zeta(s)=\chi(s)\zeta(1-s)$, where

$$\chi(1-s)=((-2\pi i)^{-s}+(2\pi i)^{-s})\Gamma(s).$$

This leads to

$$\frac{1}{2\pi i}\int_{(2)}\frac{\chi(1-s)}{1-s}u^{-s}\,ds=\frac{-1}{2\pi i}\int_{(2)}((-2\pi i)^{-s}+(2\pi i)^{-s})\frac{\Gamma(s)}{s-1}u^{-s}\,ds$$
$$=\frac{-1}{2\pi iu}\int_{(1)}((-2\pi i)^{-s-1}+(2\pi i)^{-s-1})\Gamma(s)u^{-s}\,ds=\frac{\sin 2\pi u}{\pi u}.$$

Using Cauchy's theorem, the functional equation for $\zeta(s)$, and the Dirichlet series for $\zeta(s+a)\zeta(s)$, we have

$$I_a\Big(\frac{h}{k}\Big)=-\mathrm{Res}_{s=1}\frac{\chi(1-s)\zeta(s+a)\zeta(s)}{h^{s+a}k^{1-s}(1-s)}-\mathrm{Res}_{s=1-a}\frac{\chi(1-s)\zeta(s+a)\zeta(s)}{h^{s+a}k^{1-s}(1-s)}$$
$$+\frac{1}{\pi h^{1+a}}\sum_{n=1}^{\infty}\frac{\sigma_{-a}(n)\sin(2\pi n\frac{h}{k})}{n}$$
$$=-\frac{\zeta(1+a)}{2h^{1+a}}-\frac{\zeta(a)}{ahk^a}+\frac{1}{\pi h^{1+a}}\sum_{n=1}^{\infty}\frac{\sigma_{-a}(n)\sin(2\pi n\frac{h}{k})}{n}.$$

By the functional equation for D we see that

$$\frac{D(s,-a,\frac{h}{k})-D(s,-a,-\frac{h}{k})}{2i}=\frac{2}{k}\Big(\frac{k}{2\pi}\Big)^{2-2s-a}\Gamma(1-s-a)\Gamma(1-s)$$
$$\times\Big(\cos\Big(\frac{\pi}{2}(2s+a)\Big)+\cos\frac{\pi a}{2}\Big)\Big(D\Big(1-s,a,\frac{\overline{h}}{k}\Big)-D\Big(1-s,a,-\frac{\overline{h}}{k}\Big)\Big),$$

so that, defining

$$S\Big(s,-a,\frac{h}{k}\Big):=\sum_{n=1}^{\infty}\frac{\sigma_{-a}(n)\sin(2\pi n\frac{h}{k})}{n^s},$$

we have

$$S\Big(s,-a,\frac{h}{k}\Big)=\frac{2}{k}\Big(\frac{k}{2\pi}\Big)^{2-2s-a}\Gamma(1-s-a)\Gamma(1-s)$$
$$\times\Big(\cos(\frac{\pi}{2}(2s+a))+\cos\frac{\pi a}{2}\Big)S\Big(1-s,a,\frac{\overline{h}}{k}\Big). \quad (27)$$

In particular, $S(s,-a,\frac{h}{k})$ is regular at $s=1$. Noting that

$$\lim_{s\to 1}\Gamma(1-s-a)\Gamma(1-s)\Big(\cos\Big(\frac{\pi}{2}(2s+a)\Big)+\cos\frac{\pi a}{2}\Big)=-\pi\Gamma(-a)\sin\frac{\pi a}{2}$$

and

$$S(0,a,\overline{h}/k)=\tfrac{1}{2}c_a(\overline{h}/k),$$

we obtain, by letting $s\to 1$ in (27), the identity

$$S\Big(1,-a,\frac{h}{k}\Big)=2^a\Big(\frac{\pi}{k}\Big)^{1+a}\Gamma(-a)\sin\frac{\pi a}{2}c_a\Big(\frac{\overline{h}}{k}\Big),$$

whence

$$\sum_{n=1}^{\infty}\frac{\sigma_{-a}(n)\sin(2\pi n\frac{h}{k})}{\pi n h^{1+a}}=-\Big(\frac{1}{hk}\Big)^{1+a}(2\pi)^a\Gamma(-a)\sin\frac{\pi a}{2}c_a\Big(\frac{\overline{h}}{k}\Big).$$

Thus,

$$I_a\Big(\frac{h}{k}\Big)=-\frac{\zeta(1+a)}{2h^{1+a}}-\frac{\zeta(a)}{ahk^a}-\Big(\frac{1}{hk}\Big)^{1+a}(2\pi)^a\Gamma(-a)\sin\frac{\pi a}{2}c_a\Big(\frac{\overline{h}}{k}\Big)$$

and the theorem follows. □

References

[Báez-Duarte et al. 2005] L. Báez-Duarte, M. Balazard, B. Landreau, and E. Saias, "Étude de l'autocorrélation multiplicative de la fonction 'partie fractionnaire'", *Ramanujan J.* **9**:1-2 (2005), 215–240. MR 2006i:11096 Zbl 1173.11343

[Bagchi 2006] B. Bagchi, "On Nyman, Beurling and Baez–Duarte's Hilbert space reformulation of the Riemann hypothesis", *Proc. Indian Acad. Sci. Math. Sci.* **116**:2 (2006), 137–146. MR 2007b:11126 Zbl 1125.11049

[Balasubramanian et al. 1985] R. Balasubramanian, J. B. Conrey, and D. R. Heath-Brown, "Asymptotic mean square of the product of the Riemann zeta-function and a Dirichlet polynomial", *J. Reine Angew. Math.* **357** (1985), 161–181. MR 87f:11061 Zbl 0549.10030

[Bettin and Conrey 2011] S. Bettin and J. Conrey, "A reciprocity formula for a cotangent sum", preprint, 2011.

[Bruggeman 2007] R. Bruggeman, "Quantum Maass forms", pp. 1–15 in *The conference on L-functions* (Fukuoka, 2006), edited by L. Weng and M. Kaneko, World Scientific Publishing, Hackensack, NJ, 2007. MR 2008g:11076 Zbl 1130.11019

[Conrey 1989] J. B. Conrey, "More than two fifths of the zeros of the Riemann zeta function are on the critical line", *J. Reine Angew. Math.* **399** (1989), 1–26. MR 90g:11120 Zbl 0668.10044

[Conrey and Ghosh 1998] J. B. Conrey and A. Ghosh, "A conjecture for the sixth power moment of the Riemann zeta-function", *Internat. Math. Res. Notices* **1998**:15 (1998), 775–780. MR 99h:11096 Zbl 0920.11060

[Conrey and Gonek 2001] J. B. Conrey and S. M. Gonek, "High moments of the Riemann zeta-function", *Duke Math. J.* **107**:3 (2001), 577–604. MR 2002b:11112 Zbl 1006.11048

[Hardy and Littlewood 1916] G. H. Hardy and J. E. Littlewood, "Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes", *Acta Math.* **41**:1 (1916), 119–196. MR 1555148 JFM 46.0498.01

[Ingham 1927] A. E. Ingham, "Mean-value theorems in the theory of the Riemann zeta-function", *Proc. London Math. Soc. (2)* **27**:1 (1927), 273. MR 1575391 JFM 53.0313.01

[Iwaniec 1980] H. Iwaniec, "On mean values for Dirichlet's polynomials and the Riemann zeta function", *J. London Math. Soc.* (2) **22**:1 (1980), 39–45. MR 81m:10076 Zbl 0439.10026

[Keating and Snaith 2000] J. P. Keating and N. C. Snaith, "Random matrix theory and $\zeta(1/2+it)$", *Comm. Math. Phys.* **214**:1 (2000), 57–89. MR 2002c:11107 Zbl 1051.11048

[Landreau and Richard 2002] B. Landreau and F. Richard, "Le critère de Beurling et Nyman pour l'hypothèse de Riemann: aspects numériques", *Experiment. Math.* **11**:3 (2002), 349–360. MR 2004c:11159 Zbl 1117.11305

[Lewis and Zagier 2001] J. Lewis and D. Zagier, "Period functions for Maass wave forms, I", *Ann. of Math.* (2) **153**:1 (2001), 191–258. MR 2003d:11068 Zbl 1061.11021

[Motohashi 1997] Y. Motohashi, *Spectral theory of the Riemann zeta-function*, Cambridge Tracts in Mathematics **127**, Cambridge University Press, 1997. MR 99f:11109 Zbl 0878.11001

[Zagier 2010] D. Zagier, "Quantum modular forms", pp. 659–675 in *Quanta of maths* (Paris, 2007), edited by E. Blanchard et al., Clay Math. Proc. **11**, American Mathematical Society, Providence, RI, 2010. MR 2012a:11066 Zbl 05902011

Communicated by Barry Mazur
Received 2011-12-01 Revised 2012-01-15 Accepted 2012-02-20

sandro.bettin@gmail.com *School of Mathematics, University of Bristol, Howard House, Queens Avenue, Bristol BS82NF, United Kingdom*
http://www.maths.bris.ac.uk/~maxsb/

conrey@aimath.org *American Institute of Mathematics, 360 Portage Avenue, Palo Alto, CA 94306, United States*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in ANT are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 7 No. 1 2013

1937-0652(2013)7:1;1-C