Algebra & Number Theory

Volume 10

2016

No. 3

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2016 is US $/year for the electronic version, and $/year (+$, 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 MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.451

Group schemes and local densities of ramified hermitian lattices in residue characteristic 2 Part I

Sungmun Cho

The obstruction to the local-global principle for a hermitian lattice (L, H) can be quantified by computing the mass of (L, H). The mass formula expresses the mass of (L, H) as a product of local factors, called the local densities of (L, H). The local density formula is known except in the case of a ramified hermitian lattice of residue characteristic 2.

Let F be a finite unramified field extension of $\mathbb{Q}_2$. Ramified quadratic extensions E/F fall into two cases that we call *Case 1* and *Case 2*. In this paper, we obtain the local density formula for a ramified hermitian lattice in *Case 1*, by constructing a smooth integral group scheme model for an appropriate unitary group. Consequently, this paper, combined with the paper of W. T. Gan and J.-K. Yu (*Duke Math. J.* **105** (2000), 497–524), allows the computation of the mass formula for a hermitian lattice (L, H) in *Case 1*.

1. Introduction

1A. *Introduction.* The subject of this paper is old and has intrigued many mathematicians. If (V, H) and (V', H') are two hermitian k'-spaces (or quadratic k-spaces), where k is a number field and k' is a quadratic field extension of k, then it

MSC2010: primary 11E41; secondary 11E95, 14L15, 20G25, 11E39, 11E57.
Keywords: local density, mass formula, group scheme, smooth integral model.

is well known that they are isometric if and only if for all places v, the localizations (V_v, H_v) and (V'_v, H'_v) are isometric. That is, the local-global principle holds for hermitian spaces and quadratic spaces. It is natural to ask whether the local-global principle holds for a hermitian R'-lattice or quadratic R-lattice (L, H), where R' and R are the rings of integers of k' and k, respectively. In general, the answer to this question is no. However, there is a way, namely, the mass of (L, H), to quantify the obstruction to the local-global principle. An essential tool for computing the mass of a quadratic or hermitian lattice is the mass formula. The mass formula expresses the mass of (L, H) as a product of local factors, called the local densities of (L, H).

Therefore, it suffices to find the explicit local density formula in order to obtain the mass formula and thus quantify the obstruction to the local-global principle.

For a quadratic lattice, the local density formula was first computed by G. Pall [1965] (for $p \neq 2$) and G. L. Watson [1976] (for $p = 2$). For an expository sketch of their approach, see [Kitaoka 1993]. There is another proof of Y. Hironaka and F. Sato [2000] computing the local density when $p \neq 2$. They treat an arbitrary pair of lattices, not just a single lattice, over $\mathbb{Z}_p$ (for $p \neq 2$). J. H. Conway and J. A. Sloane [1988] further developed the formula for any p and gave a heuristic explanation for it. Later, W. T. Gan and J.-K. Yu [2000] (for $p \neq 2$) and S. Cho [2015a] (for $p = 2$) provided a simple and conceptual proof of Conway and Sloane's formula by explicitly constructing a smooth affine group scheme $\underline{G}$ over $\mathbb{Z}_2$ with generic fiber $\mathrm{Aut}_{\mathbb{Q}_2}(L, H)$, which satisfies $\underline{G}(\mathbb{Z}_2) = \mathrm{Aut}_{\mathbb{Z}_2}(L, H)$.

There has not been as much work done in computing local density formulas for hermitian lattices as in the case of quadratic lattices. Although the local density formula for a quadratic lattice with $p = 2$ was first proved in the author's paper [2015a], the formula was proposed in Conway and Sloane's paper [1988]. However, the local density formula for a ramified hermitian lattice with $p = 2$ has not been proposed yet and therefore, the mass formula, when the ideal (2) is ramified in k'/k, is not known.

Hironaka [1998; 1999] obtained the local density formula for an unramified hermitian lattice. In addition, M. Mischler [2000] computed the formula for a ramified hermitian lattice ($p \neq 2$) under restricted conditions. Later, Gan and Yu [2000] found a conceptual and elegant proof of the local density formula for an unramified hermitian lattice without any restriction on p, and for a ramified hermitian lattice with the restriction $p \neq 2$, by explicitly constructing certain smooth affine group schemes (called smooth integral models) of a unitary group.

As discussed further on p. 456, we distinguish two cases for a ramified quadratic extension E/F, where F is an unramified finite extension of $\mathbb{Q}_2$, depending on the lower ramification groups G_i of the Galois group $\mathrm{Gal}(E/F)$. The division is as follows:

$$\begin{cases} \textit{Case 1}: & G_{-1} = G_0 = G_1, \;\; G_2 = 0; \\ \textit{Case 2}: & G_{-1} = G_0 = G_1 = G_2, \;\; G_3 = 0. \end{cases}$$

These two cases should be handled independently because of technical difficulty and complexity. The methodologies of the two cases are basically the same, but *Case 2* is much more difficult than *Case 1*.

The main contribution of this paper is to get an explicit formula for the local density of a hermitian B-lattice (L, h) in *Case 1*, by explicitly constructing a certain smooth group scheme associated to it that serves as an integral model for the unitary group associated to $(L \otimes_A F, h \otimes_A F)$ and by investigating its special fiber, where B is a ramified quadratic extension of A and A is an unramified finite extension of $\mathbb{Z}_2$ with F as the quotient field of A. The local density formula in *Case 2* is handled in [Cho 2015b].

In conclusion, this paper, combined with [Gan and Yu 2000] and [Cho 2015a], allows the computation of the mass formula for a hermitian R'-lattice (L, H) when $k_v/\mathbb{Q}_2$ is unramified, and $k'_{v'}/k_v$ satisfies *Case 1* or is unramified. Here, $k'_{v'}$ (resp. k_v) is the completion of k' (resp. k) at the place v' (resp. v), where v' lies over v and v lies over the ideal (2). As the simplest case, we can compute the mass formula for an arbitrary hermitian lattice explicitly when k is $\mathbb{Q}$ and k' is any quadratic field extension of $\mathbb{Q}$ such that the completion of k' at any place lying over the ideal (2) satisfies *Case 1* or is unramified over $\mathbb{Q}_2$.

Let us briefly comment on the proofs. A key input into the local density formula is

$$\lim_{N\to\infty} f^{-N\dim G}\#\underline{G}'(A/\pi^N A), \tag{1-1}$$

where f is the cardinality of the residue field of A, π is a uniformizer in A, and $\underline{G}'$ is the naive integral model for the unitary group G associated to $(L \otimes_A F, h \otimes_A F)$, which represents the functor $R \mapsto \mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$.

Now if we are lucky enough that $\underline{G}'$ is smooth, then the limit in (1-1) would stabilize at $N = 1$, which would reduce us to simply finding $\underline{G}'(\kappa)$, where κ denotes the residue field of A. A key observation of Gan and Yu is that, even when $\underline{G}'$ is not smooth, one can employ a certain smooth group scheme $\underline{G}$ lurking in the background, which is a smooth integral model of G that satisfies $\underline{G}(R) = \underline{G}'(R)$ for every étale A-algebra R. The existence and uniqueness of such a $\underline{G}$ is guaranteed by the general theory of group smoothening. Then the problem essentially reduces to constructing $\underline{G}$ explicitly, so that one can compute the cardinality of the group $\underline{G}(\kappa)$ of κ-points of its special fiber. This tells us what the analog of (1-1) for $\underline{G}$ is, and further, it so turns out that one can deduce the expression (1-1) from its analog for $\underline{G}$. For a detailed explanation about this, see Section 3 of [Gan and Yu 2000].

Let us now describe, therefore, how we construct $\underline{G}$ and study its special fiber. As $\underline{G}'$ fails to be smooth, one must impose more equations than merely the ones related to the preservation of $(L \otimes_A R, h \otimes_A R)$. Towards this, note that there exist several sublattices L' of L such that any element of $\mathrm{Aut}_B(L, h)$ automatically also preserves L' (and such that, for any étale A-algebra R, any element of

$\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$ automatically also preserves $L'\otimes_A R$). For instance, the sublattice L' of elements $x\in L$ such that $h(x,L)$ belongs to a given ideal of B necessarily satisfies this property. This gives us additional equations to impose — these equations leave the group of R-points for any étale A-algebra R untouched, while taking us closer to smoothness. It so happens that taking sufficiently many sublattices L' into consideration, and imposing further restrictions arising from the behavior of an element of $\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$ on some of their quotients, do leave us with enough equations to ensure that the group scheme $\underline{G}$ defined by them is smooth. This step already turns out to be much harder for $p=2$ than for odd p, since in this case there are many more isomorphism classes of hermitian lattices. Another source of complications is the fact that the equations involve quadratic forms over the residue field κ of A that arise as quotients of some of the lattices L' mentioned above (the theory of quadratic forms over finite fields is more complicated in characteristic 2 than in other characteristics).

Now let us describe some of the ideas involved in the computation of the special fiber $\widetilde{G}$ of $\underline{G}$. Since the quotients of some pairs of lattices of the form L' alluded to in the previous paragraph naturally support symplectic or quadratic forms, it is not hard to construct a map φ from $\widetilde{G}$ to a suitable product of symplectic and orthogonal groups. This step occurs in [Gan and Yu 2000], too. However, p being even for us poses at least two new difficulties. Firstly, although this product of symplectic and orthogonal groups contains the identity component of the maximal reductive quotient of $\widetilde{G}$, this fact seems to be difficult to prove directly. Rather, we prove this fact indirectly, by explicitly computing the dimension of the kernel of φ. Secondly, φ does not quite define the maximal reductive quotient of $\widetilde{G}$: this maximal reductive quotient is built up from φ together with a few additional homomorphisms $\widetilde{G}\to\mathbb{Z}/2\mathbb{Z}$.

Our construction of these homomorphisms $\widetilde{G}\to\mathbb{Z}/2\mathbb{Z}$ is quite indirect. A typical homomorphism is constructed in the following manner. We define a certain new hermitian lattice, say (L'',h''), starting from (L,h). This lattice naturally gives us a homomorphism $\widetilde{G}\to\widetilde{G}''$, where $\widetilde{G}''$ is the special fiber of the smooth integral model obtained by applying our construction to (L'',h'') in place of (L,h). The analog φ'' of φ defines a map from $\widetilde{G}''$ to (a product of symplectic and orthogonal groups, and in particular) an orthogonal group, and, by composing with the Dickson invariant, one gets a homomorphism $\widetilde{G}''\to\mathbb{Z}/2\mathbb{Z}$. Precomposing this with the homomorphism $\widetilde{G}\to\widetilde{G}''$ yields a homomorphism $\widetilde{G}\to\mathbb{Z}/2\mathbb{Z}$. All our homomorphisms $\widetilde{G}\to\mathbb{Z}/2\mathbb{Z}$ are constructed in this way.

To show that the candidate for the maximal reductive quotient of $\widetilde{G}$ obtained from φ and the morphisms $\widetilde{G}\to\mathbb{Z}/2\mathbb{Z}$ is indeed the maximal reductive quotient, one shows that its kernel is isomorphic, as an affine variety, to an affine space over κ. This implies by a theorem of Lazard that the kernel of our candidate for maximal reductive quotient is indeed a connected unipotent group scheme, as desired.

Our main results are Theorem 3.8, Theorem 4.12 and Theorem 5.2. Theorem 3.8 shows that the group scheme $\underline{G}$ we construct is indeed the sought after smooth group scheme over A, Theorem 4.12 gives the maximal reductive quotient of $\widetilde{G}$, and Theorem 5.2 (supplemented by Remark 5.3) gives us the final local density formulas as follows. The local density of (L, h) is

$$\beta_L = f^N \cdot f^{-\dim G} \# \widetilde{G}(\kappa).$$

Here, N is a certain integer which can be found in Theorem 5.2 and $\#\widetilde{G}(\kappa)$ can be computed explicitly based on Remark 5.3(1) and Theorem 4.12.

Appendix B is devoted to illustrating our method with a simple example: the case where $L = B \cdot e$ is of rank one and h is defined by $h(le, l'e) = \sigma(l)l'$, σ being the unique nontrivial element of $\mathrm{Gal}(E/F)$. Section B.1 describes how the usual approach that works when $p \neq 2$ (and yields the obvious integral model for the "norm one" torus associated to B/A) fails when $p = 2$, and how one may fix this from "first principles", without using any of our techniques. We hope this helps clarify some of the issues involved. Section B.2 illustrates how our construction specializes to this case; we hope that the simplicity of this case may better motivate our general construction. Some readers may therefore prefer to look at Appendix B before perusing the general constructions of Sections 3 and 4 and Appendix A.

This paper is organized as follows. We first state a structure theorem for integral hermitian forms in Section 2. We then give an explicit construction of $\underline{G}$ (in Section 3) and study its special fiber (in Section 4) in *Case 1*. Finally, we obtain an explicit formula for the local density in Section 5 in *Case 1*. In Appendix B, we provide an example to describe the smooth integral model and its special fiber and to compute the local density for a unimodular lattice of rank 1.

The reader might want to skip to Appendix B and at least go to Section B.1 to get a first glimpse into why the case of $p = 2$ is really different. Some of the ideas behind our construction can be seen in the simple example illustrated in Section B.2.

The construction of smooth integral models and the investigation of their special fibers in this paper basically follow the arguments in [Gan and Yu 2000] and [Cho 2015a]. As in [Gan and Yu 2000], the smooth group schemes constructed in this paper should be of independent interest.

2. Structure theorem for hermitian lattices and notations

2A. ***Notation.*** Notation and definitions in this section are taken from [Cho 2015a; Gan and Yu 2000; Jacobowitz 1962].

- Let F be an unramified finite extension of $\mathbb{Q}_2$ with A its ring of integers and κ its residue field.
- Let E be a ramified quadratic field extension of F with B its ring of integers.

- Let σ be the nontrivial element of the Galois group $\mathrm{Gal}(E/F)$.

- The lower ramification groups G_i of the Galois group $\mathrm{Gal}(E/F)$ satisfy one of the following:

$$\begin{cases} \textit{Case 1}: & G_{-1}=G_0=G_1,\ G_2=0; \\ \textit{Case 2}: & G_{-1}=G_0=G_1=G_2,\ G_3=0. \end{cases}$$

We explain the above briefly. Based on Section 6 and Section 9 of [Jacobowitz 1962], we can select a suitable choice of a uniformizer π of B in the following way. In *Case 1*, $E=F(\sqrt{1+2u})$ for some unit u of A and $\pi=1+\sqrt{1+2u}$. Then $\sigma(\pi)=\epsilon\pi$, where $\epsilon\equiv 1 \bmod \pi$ and $\frac{\epsilon-1}{\pi}$ is a unit in B. So we have that $\sigma(\pi)+\pi, \sigma(\pi)\cdot\pi\in(2)\backslash(4)$. In *Case 2*, $E=F(\pi)$. Here, $\pi=\sqrt{2\delta}$, where $\delta\in A$ and $\delta\equiv 1 \bmod 2$. Then $\sigma(\pi)=-\pi$.

From now on, a uniformizing element π of B, u, and δ are fixed as explained above throughout this paper. The constructions of smooth integral models associated to these two cases are different and we will treat them independently.

- We consider a B-lattice L with a hermitian form

$$h: L\times L\to B,$$

where $h(a\cdot v, b\cdot w)=\sigma(a)b\cdot h(v,w)$ and $h(w,v)=\sigma(h(v,w))$. Here, $a,b\in B$ and $v,w\in L$. We denote by a pair (L,h) a hermitian lattice. We assume that $V=L\otimes_A F$ is nondegenerate with respect to h.

- We denote by (ϵ) the B-lattice of rank 1 equipped with the hermitian form having Gram matrix (ϵ). We use the symbol $A(a,b,c)$ to denote the B-lattice $B\cdot e_1+B\cdot e_2$ with the hermitian form having Gram matrix $\begin{pmatrix} a & c\\ \sigma(c) & b\end{pmatrix}$. For each integer i, the lattice of rank 2 having Gram matrix $\begin{pmatrix} 0 & \pi^i\\ \sigma(\pi^i) & 0\end{pmatrix}$ is called the π^i-modular hyperbolic plane and denoted by $H(i)$.

- A hermitian lattice L is the orthogonal sum of sublattices L_1 and L_2, written $L=L_1\oplus L_2$, if $L_1\cap L_2=0$, L_1 is orthogonal to L_2 with respect to the hermitian form h, and L_1 and L_2 together span L.

- The ideal in B generated by $h(x,x)$ as x runs through L will be called the norm of L and written $n(L)$.

- By the scale $s(L)$ of L, we mean the ideal generated by the subset $h(L,L)$ of B.

- We define the dual lattice of L, denoted by $L^{\perp}$, as

$$L^{\perp}=\{x\in L\otimes_A F : h(x,L)\subset B\}.$$

Definition 2.1. Let L be a hermitian lattice. Then:

(a) For any nonzero scalar a, define $aL=\{ax \mid x\in L\}$. It is also a lattice in the space $L\otimes_A F$. Call a vector x of L maximal in L if x does not lie in πL.

(b) The lattice L will be called π^i-modular if the ideal generated by the subset $h(x, L)$ of E is $\pi^i B$ for every maximal vector x in L. Note that L is π^i-modular if and only if $L^\perp = \pi^{-i} L$. We can also see that $H(i)$ is π^i-modular.

(c) Assume that i is even. A π^i-modular lattice L is of *parity type I* if $n(L) = s(L)$, and of *parity type II* otherwise. The zero lattice is considered to be *of parity type II*. We caution that we do not assign a *parity type* to a π^i-modular lattice L with i odd.

2B. ***A structure theorem for integral hermitian forms.*** We state a structure theorem for π^i-modular lattices in this subsection. Note that if L is π^{2i}-modular (resp. π^{2i+1}-modular), then $\pi^{-i} L \subset L \otimes_A F$ is π^0-modular (resp. π^1-modular). We will emphasize this in Remark 2.3(a) again. Thus it is enough to provide a structure theorem for π^0-modular or π^1-modular lattices.

Theorem 2.2. *Let $i = 0$ or 1.*

(a) *Let L be a π^i-modular lattice of rank at least 3. Then $L = \bigoplus_\lambda H_\lambda \oplus K$, where K is π^i-modular of rank 1 or 2, and each $H_\lambda = H(i)$.*

(b) *We denote by (1) or (2) the ideal of B generated by the element 1 or 2, respectively. Assume that K is π^i-modular of rank 1 or 2. Then, depending on i, the rank of K, the case that E/F falls into, the parity type of L (when applicable), and $n(L)$ which is the norm of L, we may take K to be of the following form:*

Rank of K	i	E/F	Parity type of L	$n(L)$	Form for K
1	0	Case 1	I^*	$(1)^*$	(a), $a \in A$, $a \equiv 1 \bmod 2$
1	0	Case 2	I^*	$(1)^*$	(a), $a \in A$, $a \equiv 1 \bmod 2$
2	0	Case 1	I	$(1)^*$	$A(1, 2b, 1)$, $b \in A$
2	0	Case 2	I	$(1)^*$	$A(1, 2b, 1)$, $b \in A$
2	0	Case 1	II	$(2)^*$	$H(0)$
2	0	Case 2	II	$(2)^*$	$A(2\delta, 2b', 1)$, $b' \in A$
2	1	Case 1		$(2)^*$	$A(2, 2a, \pi)$, $a \in A$
2	1	Case 2		(2)	$A(4a, 2\delta, \pi)$, $a \in A$
2	1	Case 2		(4)	$H(1)$

Here, the superscript $*$ indicates the value in the table necessarily holds.

Proof. Part (a) is proved in Proposition 10.3 of [Jacobowitz 1962].

For part (b), when the rank of K is 1, it is clear that $K \cong (a')$ for a certain unit $a' \in A$ with a basis e. Since the residue field κ is perfect, there is a unit element a'' in A such that $a' \equiv a''^2 \bmod 2$. The reader can check that replacing e by $(1/a'')e$ realizes K in the manner dictated by the theorem.

From now on, we assume that the rank of K is 2. Suppose that $i = 0$. Then $n(K) = (1)$ or $n(K) = (2)$ since $n(K) \supseteq n(H(0)) = (2)$ (Proposition 9.1(a) and

Equation 9.1 with $k = 0$ in [Jacobowitz 1962]). If $n(K) = (1)$, then we can use Proposition 10.2 of [Jacobowitz 1962] to get $K \cong A(1, a, 1)$ with respect to a basis (e_1, e_2). Furthermore, the determinant $a - 1$ is a unit in A. To show this, we observe that K has an orthogonal basis, since $n(K) = s(K) = (1)$ (Proposition 4.4 in [Jacobowitz 1962]), and so the determinant should be a unit in order for K to be π^0-modular. Since the residue field κ is perfect, there is a unit element β in A such that $a - 1 \equiv \frac{1}{\beta^2} \bmod 2$. We now choose another basis $(e_1, (1-\beta)e_1 + \beta e_2)$. With this basis, it is easy to see that $K \cong A(1, 2b, 1)$ for a certain $b \in A$.

Now assume that $n(K) = (2)$ so that we cannot use Proposition 10.2 of [Jacobowitz 1962]. We choose a basis of K so that $K \cong A(x, y, 1)$ for some $x, y \in A$. Since $n(K) = (2)$, both x and y should be contained in the ideal (2). Thus $K \cong A(2a, 2b, 1)$ for some $a, b \in A$. Furthermore, in *Case 1*, if $n(K) = (2)$ then $K \cong H(0)$ by parts (a) and (b) of Proposition 9.2 of [Jacobowitz 1962].

The remaining case we need to prove when $i = 0$ is then that

$$K = A(2\delta, 2b', 1)$$

for certain $b' \in A$, in *Case 2* if $n(K) = (2)$. By Proposition 9.2(a) of [Jacobowitz 1962], if K is isotropic then $K \cong H(0)$ so that we can choose $b' = 0$. Furthermore, the lattice K with $n(K) = (2)$ is determined by its determinant up to isomorphism (Proposition 10.4 in [Jacobowitz 1962]). Since the determinant $d(K)$ of K is a unit and is well-defined modulo $N_{B/A}B^\times$, there are at most two cases of $d(K)$ because $|A^\times/N_{B/A}B^\times| = 2$. Here, $B^\times$ and $A^\times$ are the unit groups of B and A, respectively, and $N_{B/A}B^\times$ is the norm of $B^\times$. We observe that $d(A(2\delta, 0, 1))$ and $d(A(2\delta, 2d/\delta, 1))$, which are clearly π^0-modular, give different classes in $A^\times/N_{B/A}B^\times$, where d is as defined in Lemma 2.4. Thus, a lattice K with $n(K) = (2)$ in *Case 2* should be isomorphic to one of these two. In other words, such K is isomorphic to $K = A(2\delta, 2b', 1)$ with $b' = 0$ or $b' = d/\delta$.

We next suppose that $i = 1$. In *Case 1*, $n(H(1)) = (2)$ and so $n(K) = (2)$ since $s(K) \supseteq n(K) \supseteq n(H(1))$. Thus K is also determined by its determinant up to isomorphism (Proposition 10.4 in [Jacobowitz 1962]). This fact implies that there are at most two cases for K since the determinant of K divided by 2 is a unit in A and the cardinality of $A^\times/N_{B/A}B^\times$ is 2. By Lemma 2.5, $d(A(2, 0, \pi))$ and $d(A(2, 2ud, \pi))$, which are clearly π^1-modular, give different classes in $A^\times/N_{B/A}B^\times$, where u and d are as defined in Lemma 2.5. Thus, a lattice K with $n(K) = (2)$ in *Case 1* should be isomorphic to one of these two. In other words, such K is isomorphic to $d(A(2, 0, \pi))$ or $d(A(2, 2ud, \pi))$.

In *Case 2*, $n(H(1)) = (4)$ and so $n(K) = (2)$ or $n(K) = (4)$. If $n(K) = (2)$, then we can use Proposition 10.2(b) of [Jacobowitz 1962] (take $m = 1$) to get $K \cong A(2\delta, 4a, \pi)$ with basis (e_1, e_2). If we use a basis $(e_2, -e_1)$, then $K \cong$

$A(4a, 2\delta, \pi)$. If $n(K) = (4)$, then by Proposition 9.2(a–b) of [Jacobowitz 1962], $K \cong H(1) \cong A(0, 4\delta, \pi)$.

These complete the proof. □

Remark 2.3. (a) If L is π^i-modular, then $\pi^j L$ is π^{i+2j}-modular for any integer j. Thus, the above theorem implies its obvious generalization to the case where i is allowed to be any element of $\mathbb{Z}$.

(b) [Jacobowitz 1962, Section 4] For a general lattice L, we have a Jordan splitting, namely $L = \bigoplus_i L_i$ such that L_i is $\pi^{n(i)}$-modular and such that the sequence $\{n(i)\}_i$ increases. Two Jordan splittings $L = \bigoplus_{1 \leqq i \leqq t} L_i$ and $K = \bigoplus_{1 \leqq i \leqq T} K_i$ will be said to be of the same type if $t = T$ and, for $1 \leqq i \leqq T$, the following conditions are satisfied: $s(L_i) = s(K_i)$, rank L_i = rank K_i, and $n(L_i) = s(L_i)$ if and only if $n(K_i) = s(K_i)$. Jordan splitting is not unique but partially canonical in the sense that two Jordan splittings of isometric lattices are always of the same type.

(c) If we allow some of the L_i's to be zero, then we may assume that $n(i) = i$ for all i. In other words, for all $i \in \mathbb{N} \cup \{0\}$ we have $s(L_i) = (\pi^i)$, and, more precisely, L_i is π^i-modular. Then we can rephrase part (b) above as follows. Let $L = \bigoplus_i L_i$ be a Jordan splitting with $s(L_i) = (\pi^i)$ for all $i \geq 0$. Then the scale, rank and parity type of L_i depend only on L. We will deal exclusively with a Jordan splitting satisfying $s(L_i) = (\pi^i)$ from now on.

Lemma 2.4. *Assume that B/A satisfies Case 2. Then there is an element $d \in A^\times$ such that $1 - 4d$ and 1 give different classes in $A^\times / N_{B/A} B^\times$.*

Proof. Using our knowledge of the lower ramification groups G_i for Gal(E/F), we can compute the higher ramification groups G^i for the same extension:

$$G^{-1} = G^0 = G^1 = G^2 \quad \text{and} \quad G^3 = 0.$$

Let $U^i = 1 + (2)^i$ be the i-th higher unit group in F with $i \geq 1$. Then by *local class field theory*, the image of G^i under the isomorphism Gal$(E/F) \cong F^*/N_{E/F}E^*$ is $U^i/(U^i \cap N_{E/F}E^*)$. We apply this when i is 2. Then we can easily verify the existence of a d as stated in the lemma. □

Lemma 2.5. *Assume that B/A satisfies Case 1. Then there is an element $d \in A^\times$ such that $1 + 2d$ and 1 give different classes in $A^\times / N_{B/A} B^\times$.*

Proof. The proof of this lemma is similar to that of the above lemma. In this case the higher ramification groups are as follows:

$$G^{-1} = G^0 = G^1 \quad \text{and} \quad G^2 = 0.$$

Again we use *local class field theory* as explained in the proof of the above lemma but with $i = 1$. Then we can easily verify the existence of a d as stated in the lemma. □

2C. *Lattices.* In this subsection, we will define several lattices and associated notation. Fix a hermitian lattice (L, h). We denote by (π^l) the scale $s(L)$ of L.

(1) Define $A_i = \{x \in L \mid h(x, L) \in \pi^i B\}$.

(2) Define $X(L)$ to be the sublattice of L such that $X(L)/\pi L$ is the radical of the symmetric bilinear form $\frac{1}{\pi^l}h \bmod \pi$ on $L/\pi L$.

Let $l = 2m$ or $l = 2m - 1$. We consider the function defined over L by

$$\tfrac{1}{2^m}q : L \to A, \quad x \mapsto \tfrac{1}{2^m}h(x, x).$$

Then $\frac{1}{2^m}q \bmod 2$ defines a quadratic form $L/\pi L \to \kappa$. It can be easily checked that $\frac{1}{2^m}q \bmod 2$ on $L/\pi L$ is an additive polynomial if $l = 2m$, or if $l = 2m - 1$ and E/F satisfies *Case 2*. Otherwise, that is, if $l = 2m - 1$ and E/F satisfies *Case 1*, it is not additive. We define a lattice $B(L)$ as follows.

(3) If $\frac{1}{2^m}q \bmod 2$ on $L/\pi L$ is an additive polynomial, then $B(L)$ is defined to be the sublattice of L such that $B(L)/\pi L$ is the kernel of the additive polynomial $\frac{1}{2^m}q \bmod 2$ on $L/\pi L$. If $\frac{1}{2^m}q \bmod 2$ on $L/\pi L$ is not an additive polynomial, then $B(L) = L$.

To define a few more lattices, we need some preparation as follows. For the remainder of the paper, set

$$\xi := \pi \cdot \sigma(\pi).$$

Assume $B(L) \subsetneq L$ and l is even. Then the bilinear form $\xi^{-l/2}h \bmod \pi$ on the κ-vector space $L/X(L)$ is nonsingular symmetric and nonalternating. It is well known that there is a unique vector $e \in L/X(L)$ such that

$$\left(\xi^{-l/2}h(v, e)\right)^2 = \xi^{-l/2}h(v, v) \bmod \pi$$

for every vector $v \in L/X(L)$. Let $\langle e\rangle$ denote the 1-dimensional vector space spanned by the vector e and denote by $e^{\perp}$ the 1-codimensional subspace of $L/X(L)$ which is orthogonal to the vector e with respect to $\xi^{-l/2}h \bmod \pi$. Then

$$B(L)/X(L) = e^{\perp}.$$

If $B(L) = L$, the bilinear form $\xi^{-l/2}h \bmod \pi$ on the κ-vector space $L/X(L)$ is nonsingular symmetric and alternating. In this case, we put $e = 0 \in L/X(L)$ and note that it is characterized by the same identity.

The remaining lattices we need for our definition are:

(4) Define $W(L)$ to be the sublattice of L such that

$$\begin{cases} W(L)/X(L) = \langle e\rangle & \text{if } l \text{ is even;} \\ W(L) = X(L) & \text{if } l \text{ is odd.} \end{cases}$$

(5) Define $Y(L)$ to be the sublattice of L such that $Y(L)/\pi L$ is the radical of

$$\begin{cases} \text{the form } \frac{1}{2^m}h \bmod \pi \text{ on } B(L)/\pi L & \text{if } l=2m; \\ \text{the form } \frac{1}{\pi}\cdot\frac{1}{2^{m-1}}h \bmod \pi \text{ on } B(L)/\pi L & \text{if } l=2m-1 \text{ in } \textit{Case 2}. \end{cases}$$

Both forms are alternating and bilinear.

(6) Define $Z(L)$ to be the sublattice of L such that $Z(L)/\pi L$ in *Case 1* or $Z(L)/\pi B(L)$ in *Case 2* is the radical of

$$\begin{cases} \text{the form } \frac{1}{2^m}q \bmod 2 \text{ on } L/\pi L & \text{if } l=2m-1 \text{ in } \textit{Case 1}; \\ \text{the form } \frac{1}{2^{m+1}}q \bmod 2 \text{ on } B(L)/\pi B(L) & \text{if } l=2m \text{ in } \textit{Case 2}. \end{cases}$$

Both forms are quadratic.

See, e.g., page 813 of [Sah 1960] for the notion of the radical of a quadratic form on a vector space over a field of characteristic 2.

Remark 2.6. (a) We can associate the 5 lattices $(B(L), W(L), X(L), Y(L), Z(L))$ above with (A_i, h) in place of L. Let B_i, W_i, X_i, Y_i, Z_i denote the resulting lattices.

(b) As κ-vector spaces, the dimensions of A_i/B_i and W_i/X_i are at most 1.

Let $L=\bigoplus_i L_i$ be a Jordan splitting. We assign a type to each L_i as follows:

parity of i	type of L_i	condition
even	I	L_i is of parity type I
even	I^o	L_i is of parity type I and the rank of L_i is odd
even	I^e	L_i is of parity type I and the rank of L_i is even
even	II	L_i is of parity type II
odd	II	E/F satisfies *Case 1* or E/F satisfies *Case 2* with $A_i=B_i$
odd	I	E/F satisfies *Case 2* and $A_i \supsetneq B_i$

In addition, we assign a subtype to L_i in the following manner:

parity of i	subtype of L_i	condition
even	bound of type I	L_i is of type I and either L_{i-2} or L_{i+2} is of type I
even	bound of type II	L_i is of type II and either L_{i-1} or L_{i+1} is of type I
odd	bound	either L_{i-1} or L_{i+1} is of type I

In all other cases, L_i is called *free*.

Notice that the type of each L_i is determined canonically regardless of the choice of a Jordan splitting.

2D. ***Sharpened structure theorem for integral hermitian forms.*** While Theorem 2.2 lets us work with a restricted set of candidates for each L_i, further pruning is facilitated by the type of each L_i. For this, we need a series of lemmas.

Lemma 2.7 [Jacobowitz 1962, Proposition 9.2]. *Let L be a π^i-modular lattice of rank 2 with $n(L) = n(H(i))$. Then $L \cong H(i)$ in Case 2 with i odd and in Case 1 with i even.*

Lemma 2.8 [Jacobowitz 1962, Proposition 4.4]. *A π^i-modular lattice L has an orthogonal basis if $n(L) = s(L)$.*

Lemma 2.9. *Assume that E/F satisfies Case 2.*

(1) *Let $L = A(4a, 2\delta, \pi) \oplus (2c)$ with respect to a basis (e_1, e_2, e_3), where $c \equiv 1 \bmod 2$. Then $L \cong H(1) \oplus (2c')$ where $c' \equiv 1 \bmod 2$.*

(2) *Let $L = A(4a, 2\delta, \pi) \oplus (c)$ with respect to a basis (e_1, e_2, e_3), where $c \equiv 1 \bmod 2$. Then $L \cong H(1) \oplus (c')$ where $c' \equiv 1 \bmod 2$.*

Proof. For (1), we work with the basis $(e_1 - (2a\pi/\delta)e_2, e_2 + e_3, (c\pi/\delta)e_1 + e_3)$ of L. With respect to this basis, $L \cong A(-4a - 16a^2, 2(\delta + c), \pi(1 + 4a)) \oplus (2c(1 - 4ac/\delta))$. Moreover, $n(A(-4a - 16a^2, 2(\delta + c), \pi(1 + 4a))) = n(H(1)) = (4)$. Combined with the lemma above, this completes the proof.

For (2), we note that the sublattice of L spanned by $(e_1, e_2, \pi e_3)$ is isomorphic to $A(4a, 2\delta, \pi) \oplus (2c')$ where $c' \equiv 1 \bmod 2$. If we apply (1) to this sublattice by choosing a basis $(e_1 - (2a\pi/\delta)e_2, e_2 + \pi e_3, (c'\pi/\delta)e_1 + \pi e_3)$, then $A(4a, 2\delta, \pi) \oplus (2c')$ is isomorphic to $H(1) \oplus (2c'')$ where $c'' \equiv 1 \bmod 2$. Now the sublattice of L spanned by $(e_1 - (2a\pi/\delta)e_2, e_2 + \pi e_3, \frac{1}{\pi}((c'\pi/\delta)e_1 + \pi e_3))$, which is the same as L, is isomorphic to $H(1) \oplus (-c''/\delta)$. ☐

The above lemmas will contribute to the proof of Theorem 2.10 below in the following manner. For a given Jordan splitting $L = \bigoplus_i L_i$ in *Case 2*, assume that L_1 is bound of type I. Theorem 2.2 tells us that there are two different possibilities for L_1 as a hermitian lattice and if $L_1 = \bigoplus H(1)$ then the conclusion of the as yet unstated Theorem 2.10, for $i = 1$, will follow. If $L_1 = \bigoplus H(1) \oplus A(4a, 2\delta, \pi)$ and either L_0 or L_2 is of type I^o, then by Lemma 2.9 and the above paragraph, $L_0 \oplus L_1 \oplus L_2 = L_0' \oplus L_1' \oplus L_2'$ such that $L_1' = \bigoplus H(1)$ and the types of L_0 and L_2 are the same as those of L_0' and L_2', respectively. In case either L_0 or L_2 is of type I^e, say L_2 is of type I^e, $L_2 = (\bigoplus H(2)) \oplus (2a) \oplus (2b)$ where $a, b \equiv 1 \bmod 2$ by Lemma 2.8. Then we use Lemma 2.9 on $L_1 \oplus (2b)$ to get $L_1 \oplus (2b) = (\bigoplus H(1)) \oplus (2b')$ with $b' \equiv 1 \bmod 2$. Thus $L_1 \oplus L_2 = L_1' \oplus L_2'$ where $L_1' = \bigoplus H(1)$ and the type of $L_2' = (\bigoplus H(2)) \oplus (2a) \oplus (2b')$ is the same as that of L_2. We conclude that $L = L_0' \oplus L_1' \oplus L_2' \oplus (\bigoplus_i L_i)$ is another Jordan splitting of L and in this case, $L_1' = \bigoplus H(1)$. Therefore, if L_1 is bound of type I in *Case 2*, then L_1 can always be replaced by $\bigoplus H(1)$. Combined with Theorem 2.2, this yields the following structure theorem:

Theorem 2.10. *There exists a suitable choice of a Jordan splitting of the given lattice $L = \bigoplus_i L_i$ such that $L_i = \bigoplus_\lambda H_\lambda \oplus K$, where each $H_\lambda = H(i)$ and K is π^i-modular of rank 1 or 2, with the following descriptions. Let $i = 0$ or $i = 1$. Then*

(a) *In Case 1,*

$$K = \begin{cases} (a) \text{ where } a \equiv 1 \bmod 2 & \text{if } i = 0 \text{ and } L_0 \text{ is of type } I^o; \\ A(1, 2b, 1) & \text{if } i = 0 \text{ and } L_0 \text{ is of type } I^e; \\ H(0) & \text{if } i = 0 \text{ and } L_0 \text{ is of type } II; \\ A(2, 2b, \pi) & \text{if } i = 1. \end{cases}$$

(b) *In Case* 2,

$$K=\begin{cases}(a)\ \textit{where}\ a\equiv 1 \bmod 2 & \textit{if}\ i=0\ \textit{and}\ L_0\ \textit{is of type}\ I^o;\\ A(1,2b,1) & \textit{if}\ i=0\ \textit{and}\ L_0\ \textit{is of type}\ I^e;\\ A(2\delta,2b,1) & \textit{if}\ i=0\ \textit{and}\ L_0\ \textit{is of type}\ II;\\ A(4a,2\delta,\pi) & \textit{if}\ i=1\ \textit{and}\ L_1\ \textit{is free of type}\ I;\\ H(1) & \textit{if}\ i=1,\ \textit{and}\ L_1\ \textit{is bound of type}\ I\ \textit{or of type}\ II.\end{cases}$$

Here, $a, b \in A$ *and* δ, π *are explained in Section 2A.*

From now on, the pair (L, h) is fixed throughout this paper.

Remark 2.11. Working with a basis furnished by Theorem 2.10, we can describe our lattices A_i through Z_i more explicitly. We use the following conventions. Let $\mathcal{L}_i$ denote $\bigoplus_{j\neq i}\pi^{\max\{0,i-j\}}L_j$. Further, the $\bigoplus_\lambda H_\lambda$ will be denoted by $\mathcal{H}_i$. Theorem 2.10 involves a basis for a lattice K, which we will write as $\{e_1^{(i)}, e_2^{(i)}\}$ according to the ordering contained therein. For all cases, we have $A_i = \mathcal{L}_i \oplus L_i$ and $X_i = \mathcal{L}_i \oplus \pi L_i$.

In order to write W_i, we should first find the vector $e \in A_i/X_i$ explained in the paragraph right after the definition of $B(L)$ in Section 2C. In order to simplify notations, let us work with one example. Assume that (e_1, e_2, e_3, e_4) is a B-basis of L with respect to which h is represented by the matrix

$$\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&1&1\\0&0&1&2\end{pmatrix}.$$

So L $(= L_0 = A_0)$ is of type I^e and our basis is as explained in Theorem 2.10. Now, in order to find W_0, we should find the vector $e \in L/\pi L$ explained in Section 2C (after the definition of $B(L)$). If $v = (x, y, z, w)$ is a vector in $L/\pi L$, then $h(v, v) \bmod \pi = z^2$. On the other hand, if $e = (0, 0, 0, 1) \in L/\pi L$, then $(h(v, e))^2 \bmod \pi = z^2$. Therefore, by uniqueness of the vector e, $(0, 0, 0, 1) \in L/\pi L$ is the vector e we are looking for.

Since W_0 is the sublattice of L such that $W_0/X_0 = W_0/\pi L$ is the subspace of $L/\pi L$ spanned by the vector e, W_0 is spanned by $(\pi e_1, \pi e_2, \pi e_3, e_4)$, and it is easy to see that B_0 is spanned by $(e_1, e_2, \pi e_3, e_4)$.

To describe all lattices, it is good to start with the matrix of our fixed hermitian form h with respect to a basis furnished by Theorem 2.10.

Case 1, i even: For type I, $e = (0, \cdots, 0, 1) \in A_i/X_i$. The following table describes the lattices:

Type	B_i	W_i	Y_i
I^o	$\mathcal{L}_i \oplus \mathcal{H}_i \oplus (\pi)e_1^{(i)}$	$\mathcal{L}_i \oplus \pi\cdot\mathcal{H}_i \oplus Be_1^{(i)}$	X_i
I^e	$\mathcal{L}_i \oplus \mathcal{H}_i \oplus (\pi)e_1^{(i)} \oplus Be_2^{(i)}$	$\mathcal{L}_i \oplus \pi\cdot\mathcal{H}_i \oplus (\pi)e_1^{(i)} \oplus Be_2^{(i)}$	W_i
II	A_i	X_i	X_i

Case 1, i odd. We have $B_i = A_i$, $W_i = X_i$, and Y_i is not defined. Also, Z_i is a sublattice of A_i and so we should have congruence conditions for L_j. Namely,

$$Z_i = \bigoplus_{j \notin \{i\} \cup \mathcal{E}} \pi^{\max\{0,i-j\}} L_j \oplus \pi L_i \oplus \bigoplus_{j \in \mathcal{E}} \pi^{\max\{0,i-j\}} \big(\mathcal{H}_j \oplus B e_2^{(j)}\big)$$
$$\oplus \left\{ \sum_{j \in \mathcal{E}} \pi^{\max\{0,i-j\}} \cdot a_j e_1^{(j)} \;\Big|\; \text{for each } a_j \in B, \sum_{j \in \mathcal{E}} a_j \in (\pi) \right\}.$$

Here, $\mathcal{E} = \{j \in \{i-1, i+1\} \mid L_j \text{ is of type } I\}$ and the $e_2^{(j)}$ factor should be ignored for those $j \in \mathcal{E}$ such that L_j is of type I^o.

The following example would be helpful to have a better understanding of the notions of "bound" and "free" and of the notion of type when i is odd. Let $L = L_1 \oplus L_2 = A(0, 0, \pi) \oplus (2)$, so that L_1 is bound of type I (since $A_1 \neq B_1$) and L_2 is free of type I.

Case 2, i even. The B_i, W_i, and Y_i are exactly as in the table given for *Case 1*. The lattice Z_i is a little complicated. Note that when L_i is of type I or bound of type II, the dimension of Y_i/Z_i as a κ-vector space is 1. We describe it case by case below.

- Let $\mathcal{E}' = \{j \in \{i-2, i+2\} \mid L_j \text{ is of type } I\}$. If L_i is of type I so that L_{i-1} and L_{i+1} are bound,

$$Z_i = \bigoplus_{j \notin \{i, i \pm 2\}} \pi^{\max\{0,i-j\}} L_j \oplus \bigoplus_{j \in \{i \pm 2\}} \pi^{\max\{0,i-j\}} \big(\mathcal{H}_j \oplus B e_2^{(j)}\big) \oplus \pi \mathcal{H}_i$$
$$\oplus \left\{ \left(\sum_{j \in \mathcal{E}'} \pi^{\max\{0,i-j\}} \cdot a_j e_1^{(j)} \right) + \big(\pi \cdot a_i e_1^{(i)} + b \cdot b_i e_2^{(i)}\big) \;\Big| \right.$$
$$\left. \text{for each } a_j \in B, \left(\sum_{j \in \mathcal{E}'} a_j \right) + a_i + b \cdot b_i \in (\pi) \right\},$$

  where the $e_2^{(j)}$ (resp. $e_2^{(i)}$) factor should be ignored for those $j \in \{i \pm 2\}$ (resp. i) such that L_j (resp. L_i) is not of type I^e, and $b \in B$ is such that $L_i = \pi^{i/2} (\mathcal{H}_i \oplus A(1, 2b, 1))$ when L_i is of type I^e.

- If L_i is free of type II (so that all of $L_{i\pm 2}$ and $L_{i \pm 1}$ are of type II), then $Z_i = X_i$.
- If L_i is bound of type II, then with $\mathcal{E}_1 = \{j \in \{i-1, i+1\} \mid L_j \text{ is free of type } I\}$ and $\mathcal{E}_2 = \{j \in \{i-2, i+2\} \mid L_j \text{ is of type } I\}$, we have

$$Z_i = \bigoplus_{j \notin \{i, i\pm 1, i \pm 2\}} \pi^{\max\{0,i-j\}} L_j \oplus \pi L_i$$
$$\oplus \bigoplus_{j \in \{i \pm 1\}} \pi^{\max\{0,i-j\}} \big(\mathcal{H}_j \oplus B e_1^{(j)}\big) \oplus \bigoplus_{j \in \{i \pm 2\}} \pi^{\max\{0,i-j\}} \big(\mathcal{H}_j \oplus B e_2^{(j)}\big)$$
$$\oplus \left\{ \left(\sum_{j \in \mathcal{E}_1} \pi^{\max\{0,i-j\}} \cdot a_j e_2^{(j)} \right) + \left(\sum_{j \in \mathcal{E}_2} \pi^{\max\{0,i-j\}} \cdot a_j e_1^{(j)} \right) \;\Big| \right.$$
$$\left. \text{for each } a_j \in B, \left(\sum_{j \in \mathcal{E}_1 \cup \mathcal{E}_2} a_j \right) \in (\pi) \right\}.$$

  For example, if $i+1 \in \mathcal{E}_1$, then $i+2 \notin \mathcal{E}_2$. And if $i+2 \in \mathcal{E}_2$, then $i+1 \notin \mathcal{E}_1$.

Case 2, i odd. In this case, $W_i = X_i$ and Z_i is not defined.

Type	B_i	Y_i
free of type I	$\mathcal{L}_i \oplus \mathcal{H}_i \oplus Be_1^{(i)} \oplus (\pi)e_2^{(i)}$	$\mathcal{L}_i \oplus \pi\mathcal{H}_i \oplus Be_1^{(i)} \oplus (\pi)e_2^{(i)}$
bound of type I	see below	see below
type II	A_i	X_i

When L_i is bound of type I, the dimension of A_i/B_i as κ-spaces is 1.

$$B_i = \bigoplus_{j\notin\{i\}\cup\mathcal{E}} \pi^{\max\{0,i-j\}} L_j \oplus L_i \oplus \bigoplus_{j\in\mathcal{E}} \pi^{\max\{0,i-j\}}\big(\mathcal{H}_j \oplus Be_2^{(j)}\big)$$
$$\oplus \Big\{ \sum_{j\in\mathcal{E}} \pi^{\max\{0,i-j\}} \cdot a_j e_1^{(j)} \;\Big|\; \text{for each } a_j \in B, \sum_{j\in\mathcal{E}} a_j \in (\pi) \Big\},$$
$$Y_i = \bigoplus_{j\notin\{i\}\cup\mathcal{E}} \pi^{\max\{0,i-j\}} L_j \oplus \pi L_i \oplus \bigoplus_{j\in\mathcal{E}} \pi^{\max\{0,i-j\}}\big(\mathcal{H}_j \oplus Be_2^{(j)}\big)$$
$$\oplus \Big\{ \sum_{j\in\mathcal{E}} \pi^{\max\{0,i-j\}} \cdot a_j e_1^{(j)} \;\Big|\; \text{for each } a_j \in B, \sum_{j\in\mathcal{E}} a_j \in (\pi) \Big\}.$$

Here, $\mathcal{E} = \{j \in \{i-1, i+1\} \mid L_j \text{ is of type } I\}$.

3. The construction of the smooth model

Let $\underline{G}'$ be the naive integral model of the unitary group $U(V, h)$, where $V = L \otimes_A F$, such that for any commutative A-algebra R,

$$\underline{G}'(R) = \mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R).$$

The scheme $\underline{G}'$ is then an (possibly nonsmooth) affine group scheme over A with smooth generic fiber $U(V, h)$. Then by Proposition 3.7 in [Gan and Yu 2000], there exists a unique smooth integral model, denoted by $\underline{G}$, with generic fiber $U(V, h)$, characterized by

$$\underline{G}(R) = \underline{G}'(R)$$

for any étale A-algebra R. Note that every étale A-algebra is a finite product of finite unramified extensions of A. This section, Section 4 and Appendix A are devoted to gaining an explicit knowledge of the smooth integral model $\underline{G}$ in *Case 1*, which will be used in Section 5 to compute the local density of (L, h) (again, in *Case 1*). For a detailed exposition of the relation between the local density of (L, h) and $\underline{G}$, see [Gan and Yu 2000, Section 3].

In this section, we give an explicit construction of the smooth integral model $\underline{G}$ when E/F satisfies *Case 1*. The construction of $\underline{G}$ is based on that of Section 5 in [Gan and Yu 2000] and Section 3 in [Cho 2015a]. Since the functor $R \mapsto \underline{G}(R)$ restricted to étale A-algebras R determines $\underline{G}$, we first list out some properties that are satisfied by each element of $\underline{G}(R) = \underline{G}'(R)$.

We choose an element $g \in \underline{G}(R)$ for an étale A-algebra R. Then g is an element of $\mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$. Here we consider $\mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$ as a subgroup of $\mathrm{Res}_{E/F}\,\mathrm{GL}_E(V)(F \otimes_A R)$. To ease the notation, we say $g \in \mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$ stabilizes a lattice $M \subseteq V$ if $g(M \otimes_A R) = M \otimes_A R$.

3A. ***Main construction.*** Let R be an étale A-algebra. In this subsection, as mentioned above, we observe properties of elements of $\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$ and their matrix interpretations. We choose a Jordan splitting $L=\bigoplus_i L_i$ and a basis of L as explained in Theorem 2.10 and Remark 2.3(a). Let $n_i=\mathrm{rank}_B L_i$, and $n=\mathrm{rank}_B L=\sum n_i$. Assume that $n_i=0$ unless $0\leq i<N$. Let g be an element of $\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$. We always divide a matrix g of size $n\times n$ into N^2 blocks such that the block in position (i,j) is of size $n_i\times n_j$. For simplicity, the row and column numbering starts at 0 rather than 1.

(1) First of all, g stabilizes A_i for every integer i. In terms of matrices, this fact means that the (i,j)-block has entries in $\pi^{\max\{0,j-i\}}B\otimes_A R$. From now on, we write

$$g=\left(\pi^{\max\{0,j-i\}}g_{i,j}\right).$$

(2) The element g stabilizes A_i, B_i, W_i, X_i and induces the identity on A_i/B_i and W_i/X_i. We also interpret these facts in terms of matrices as described below:

(a) If i is odd or L_i is of type *II*, then $A_i=B_i$ and $W_i=X_i$ and so there is no contribution.

(b) If L_i is of type I^o, the diagonal (i,i)-block $g_{i,i}$ is of the form

$$\begin{pmatrix} s_i & \pi y_i \\ \pi v_i & 1+\pi z_i \end{pmatrix}\in \mathrm{GL}_{n_i}(B\otimes_A R),$$

where s_i is an $(n_i-1)\times(n_i-1)$-matrix, etc.

(c) If L_i is of type I^e, the diagonal (i,i)-block $g_{i,i}$ is of the form

$$\begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & 1+\pi x_i & \pi z_i \\ v_i & u_i & 1+\pi w_i \end{pmatrix}\in \mathrm{GL}_{n_i}(B\otimes_A R),$$

where s_i is an $(n_i-2)\times(n_i-2)$-matrix, etc.

3B. ***Construction of $\underline{M}$.*** We define a functor from the category of commutative flat A-algebras to the category of monoids as follows. For any commutative flat A-algebra R, let

$$\underline{M}(R)\subset\{m\in \mathrm{End}_{B\otimes_A R}(L\otimes_A R)\}$$

to be the set of $m\in \mathrm{End}_{B\otimes_A R}(L\otimes_A R)$ satisfying the following conditions:

(1) m stabilizes $A_i\otimes_A R$, $B_i\otimes_A R$, $W_i\otimes_A R$, $X_i\otimes_A R$ for all i.

(2) m induces the identity on $A_i\otimes_A R/B_i\otimes_A R$, $W_i\otimes_A R/X_i\otimes_A R$ for all i.

Remark 3.1. We give another description for the functor $\underline{M}$ and using this, we show that it is represented by a polynomial ring. Let us define a functor from the category of commutative flat A-algebras to the category of rings as follows:

For any commutative flat A-algebra R, define

$$\underline{M}'(R)\subset\{m\in \mathrm{End}_{B\otimes_A R}(L\otimes_A R)\}$$

to be the set of $m\in \mathrm{End}_{B\otimes_A R}(L\otimes_A R)$ satisfying the following conditions:

(1) m stabilizes $A_i\otimes_A R$, $B_i\otimes_A R$, $W_i\otimes_A R$, $X_i\otimes_A R$ for all i.

(2) m maps $A_i\otimes_A R$, $W_i\otimes_A R$ into $B_i\otimes_A R$, $X_i\otimes_A R$, respectively.

Then, by Lemma 3.1 of [Cho 2015a], $\underline{M}'$ is represented by a unique flat A-algebra $A(\underline{M}')$ which is a polynomial ring over A of $2n^2$ variables. Moreover, it is easy to see that $\underline{M}'$ has the structure of a scheme of rings since $\underline{M}'(R)$ is closed under addition and multiplication.

We consider a scheme $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$ such that the associated set to a commutative flat A-algebra R is $\mathrm{End}_{B\otimes_A R}(L\otimes_A R)$. Indeed, $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$ is a group scheme under addition. But at this moment, we consider it as a scheme of sets so as to embed $\underline{M}$ into this. Let us consider both $\underline{M}$ and $\underline{M}'$ as functors from the category of commutative flat A-algebras to the category of sets. Then they are subfunctors of $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$. Furthermore, the functor $\underline{M}$ (viewed as valued in sets) is the same as the functor $1+\underline{M}'$, where $(1+\underline{M}')(R)=\{1+m : m\in\underline{M}'(R)\}$. Here, the set $\mathrm{End}_{B\otimes_A R}(L\otimes_A R)$ has an obvious additive structure and the addition in the description of $(1+\underline{M}')(R)$ comes from this.

Therefore, $\underline{M}$ and $\underline{M}'$ are equivalent, as subfunctors of $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$. This fact induces that the functor $\underline{M}$ is also represented by a unique flat A-algebra $A[\underline{M}]$ which is a polynomial ring over A of $2n^2$ variables. Moreover, it is easy to see that $\underline{M}$ has the structure of a scheme of monoids since $\underline{M}(R)$ is closed under multiplication.

We can therefore now talk of $\underline{M}(R)$ for any (not necessarily flat) A-algebra R. However, for a general R, the above description for $\underline{M}(R)$ will no longer be true. For such R, we use our chosen basis of L to write each element of $\underline{M}(R)$ formally. We describe each element of $\underline{M}(R)$ as a formal matrix $\left(\pi^{\max\{0,j-i\}}m_{i,j}\right)$. Here, $m_{i,j}$, when $i\neq j$, is an $(n_i\times n_j)$-matrix with entries in $B\otimes_A R$ and

$$m_{i,i}=\begin{cases}\begin{pmatrix} s_i & \pi y_i\\ \pi v_i & 1+\pi z_i\end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^o;\\ \begin{pmatrix} s_i & r_i & \pi t_i\\ \pi y_i & 1+\pi x_i & \pi z_i\\ v_i & u_i & 1+\pi w_i\end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^e;\\ m_{i,i} & \text{otherwise, i.e., if } L_i \text{ is of type } II.\end{cases}$$

Here, s_i is an $(n_i-1\times n_i-1)$-matrix (resp. $(n_i-2\times n_i-2)$-matrix) with entries in $B\otimes_A R$ if L_i of type I^o (resp. of type I^e) and $y_i, v_i, z_i, r_i, t_i, y_i, x_i, u_i, w_i$ are matrices of suitable sizes with entries in $B\otimes_A R$. Similarly, if L_i is of type II, then $m_{i,i}$ is an $(n_i\times n_j)$-matrix with entries in $B\otimes_A R$. To simplify notation, each element

$$((m_{i,j})_{i\neq j},(m_{i,i})_{L_i\text{ of type II}},(s_i,y_i,v_i,z_i)_{L_i\text{ of type }I^o},(s_i,v_i,z_i,r_i,t_i,y_i,x_i,u_i,w_i)_{L_i\text{ of type }I^e}).$$

of $\underline{M}(R)$ is denoted by $(m_{i,j},s_i\cdots w_i)$.

In the next section, we need a description of an element of $\underline{M}(R)$ and its multiplication for a κ-algebra R. In order to prepare for this, we describe the multiplication explicitly only for a κ-algebra R. To multiply $(m_{i,j},s_i\cdots w_i)$ and $(m'_{i,j},s'_i\cdots w'_i)$, we form the matrices $m=\left(\pi^{\max\{0,j-i\}}m_{i,j}\right)$ and $m'=\left(\pi^{\max\{0,j-i\}}m'_{i,j}\right)$ with $s_i\cdots w_i$ and $s'_i\cdots w'_i$ and write the formal matrix product $\left(\pi^{\max\{0,j-i\}}m_{i,j}\right)\cdot\left(\pi^{\max\{0,j-i\}}m'_{i,j}\right)=\left(\pi^{\max\{0,j-i\}}\tilde{m}''_{i,j}\right)$ with

$$\tilde{m}''_{i,i}=\begin{cases}\begin{pmatrix} \tilde{s}''_i & \pi \tilde{y}''_i\\ \pi \tilde{v}''_i & 1+\pi \tilde{z}''_i\end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^o;\\ \begin{pmatrix} \tilde{s}''_i & \tilde{r}''_i & \pi \tilde{t}''_i\\ \pi \tilde{y}''_i & 1+\pi \tilde{x}''_i & \pi \tilde{z}''_i\\ \tilde{v}''_i & \tilde{u}''_i & 1+\pi \tilde{w}''_i\end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^e.\end{cases}$$

Let $(m''_{i,j}, s''_i \cdots w''_i)$ be formed by letting π^2 be zero in each entry of $(\tilde{m}''_{i,j}, \tilde{s}''_i \cdots \tilde{w}''_i)$. Then each matrix of $(m''_{i,j}, s''_i \cdots w''_i)$ has entries in $B \otimes_A R$ and so $(m''_{i,j}, s''_i \cdots w''_i)$ is an element of $\underline{M}(R)$ and is the product of $(m_{i,j}, s_i \cdots w_i)$ and $(m'_{i,j}, s'_i \cdots w'_i)$. More precisely,

(1) If $i \neq j$ or if $i = j$ and L_i is of type II,

$$m''_{i,j} = \sum_{k=1}^{N} \pi^{(\max\{0,k-i\}+\max\{0,j-k\}-\max\{0,j-i\})} m_{i,k} m'_{k,j};$$

(2) For L_i of type I^o, we write $m_{i,i-1}m'_{i-1,i} + m_{i,i+1}m'_{i+1,i} = \left(\begin{smallmatrix} a''_i & b''_i \\ c''_i & d''_i \end{smallmatrix}\right)$ and $m_{i,i-2}m'_{i-2,i} + m_{i,i+2}m'_{i+2,i} = \left(\begin{smallmatrix} \tilde{a}''_i & \tilde{b}''_i \\ \tilde{c}''_i & \tilde{d}''_i \end{smallmatrix}\right)$ where a''_i and $\tilde{a}''_i$ are $(n_i-1)\times(n_i-1)$-matrices, etc. Then

$$\begin{cases} s''_i = s_i s'_i + \pi a''_i; \\ y''_i = s_i y'_i + y_i + b''_i + \pi(y_i z'_i + \tilde{b}''_i); \\ v''_i = v_i s'_i + v'_i + c''_i + \pi(z_i v'_i + \tilde{c}''_i); \\ z''_i = z_i + z'_i + d''_i + \pi(z_i z'_i + v_i y'_i + \tilde{d}''_i). \end{cases}$$

(3) When L_i is of type I^e, we write

$$m_{i,i-1}m'_{i-1,i} + m_{i,i+1}m'_{i+1,i} = \begin{pmatrix} a''_i & b''_i & c''_i \\ d''_i & e''_i & f''_i \\ g''_i & h''_i & k''_i \end{pmatrix}$$

and

$$m_{i,i-2}m'_{i-2,i} + m_{i,i+2}m'_{i+2,i} = \begin{pmatrix} \tilde{a}''_i & \tilde{b}''_i & \tilde{c}''_i \\ \tilde{d}''_i & \tilde{e}''_i & \tilde{f}''_i \\ \tilde{g}''_i & \tilde{h}''_i & \tilde{k}''_i \end{pmatrix}$$

where a''_i and $\tilde{a}''_i$ are $(n_i-2)\times(n_i-2)$-matrices, etc. Then

$$\begin{cases} s''_i &= s_i s'_i + \pi(r_i y'_i + t_i v'_i + a''_i); \\ r''_i &= s_i r'_i + r_i + \pi(r_i x'_i + t_i u'_i + b''_i); \\ t''_i &= s_i t'_i + r_i z'_i + t_i + c''_i + \pi(t_i w'_i + \tilde{c}''_i); \\ y''_i &= y_i s'_i + y'_i + z_i v'_i + d''_i + \pi(x_i y'_i + \tilde{d}''_i); \\ x''_i &= x_i + x'_i + z_i u'_i + y_i r'_i + e''_i + \pi(x_i x'_i + \tilde{e}''_i); \\ z''_i &= z_i + z'_i + f''_i + \pi(y_i t'_i + x_i z'_i + z_i w'_i + \tilde{f}''_i); \\ v''_i &= v_i s'_i + v'_i + \pi(u_i y'_i + w_i v'_i + g''_i); \\ u''_i &= u_i + u'_i + v_i r'_i + \pi(u_i x'_i + w_i u'_i + h''_i); \\ w''_i &= w_i + w'_i + v_i t'_i + u_i z'_i + k''_i + \pi(w_i w'_i + \tilde{k}''_i). \end{cases}$$

Remark 3.2. We let d be the determinant homomorphism on the algebraic monoid $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$. We consider the inclusion

$$\iota : \underline{M} \longrightarrow \mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$$

between functors of sets on the category of commutative flat A-algebras. Note that this inclusion is a morphism of schemes by Yoneda's lemma since $\underline{M}$ is flat over A. It is not an immersion as schemes since the special fiber of $\underline{M}$ is no longer embedded into that of $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)$. For a commutative flat A-algebra R, the multiplication on $\underline{M}(R)$ is induced from that on $\mathrm{Res}_{B/A}\,\mathrm{End}_B(L)(R)$ under ι. Thus the morphism ι is a morphism of monoid schemes.

We consider d as the restriction of the determinant homomorphism under ι. Then $\mathrm{Spec}(A[\underline{M}]_d)$ is an open subscheme of $\underline{M}$, where $A[\underline{M}]_d$ is the localization of the ring $A[\underline{M}]$ at d. Note that $\mathrm{Spec}(A[\underline{M}]_d)(R)$, the set of R-points of $\mathrm{Spec}(A[\underline{M}]_d)$ for a commutative A-algebra R, is characterized by

$$\{m \in \underline{M}(R) : \text{there exists } \tilde{m}' \in \mathrm{End}_{B\otimes_A R}(L \otimes_A R) \text{ such that } \iota_R(m)\cdot\tilde{m}' = \tilde{m}'\cdot\iota_R(m) = 1\}.$$

Here, $\iota_R : \underline{M}(R) \to \mathrm{Res}_{B/A}\,\mathrm{End}_B(L)(R)$ is a morphism of monoids induced by ι. It is easy to see that the above set $\mathrm{Spec}(A[\underline{M}]_d)(R)$ is a monoid, and hence $\mathrm{Spec}(A[\underline{M}]_d)$ is a scheme of monoids.

We define a functor $\underline{M}^*$ from the category of commutative A-algebras to the category of groups as follows. For a commutative A-algebra R, set

$$\underline{M}^*(R) = \{m \in \underline{M}(R) : \text{there exists } m' \in \underline{M}(R) \text{ such that } m\cdot m' = m'\cdot m = 1\}.$$

We claim that $\underline{M}^*$ is representable by $\mathrm{Spec}(A[\underline{M}]_d)$. For any commutative A-algebra R, the inclusion $\underline{M}^*(R) \subseteq \mathrm{Spec}(A[\underline{M}]_d)(R)$ is obvious.

In order to show $\mathrm{Spec}(A[\underline{M}]_d)(R) \subseteq \underline{M}^*(R)$, we first prove that $\tilde{m}'$ $(\in \mathrm{End}_{B\otimes_A R}(L\otimes_A R))$ associated to $m \in \underline{M}(R)$ is an element of $\underline{M}(R)$ for every flat A-algebra R. To verify this statement, it suffices to show that $\tilde{m}'$ satisfies conditions (1) and (2) defining $\underline{M}$. This follows from the following fact: if L' is a sublattice of L and m is an element of $\mathrm{Spec}(A[\underline{M}]_d)(R)$ for a flat A-algebra R which stabilizes $L' \otimes_A R$, then $L' \otimes_A R$ is stabilized by $\tilde{m}'$ as well. This can be easily proved as in Lemma 3.2 of [Cho 2015a] and so we skip the proof. Thus $\underline{M}^*(R)$ is the same as $\mathrm{Spec}(A[\underline{M}]_d)(R)$ for a flat A-algebra R. In order to show $\underline{M}^*(R) = \mathrm{Spec}(A[\underline{M}]_d)(R)$ for any commutative A-algebra R, we consider the following well-defined map, for any flat A-algebra R:

$$\begin{aligned} \mathrm{Spec}(A[\underline{M}]_d)(R) &\to \mathrm{Spec}(A[\underline{M}]_d)(R) \times \mathrm{Spec}(A[\underline{M}]_d)(R) \\ m &\mapsto (m, \tilde{m}'). \end{aligned}$$

Since $\mathrm{Spec}(A[\underline{M}]_d)$ is flat, this map is represented by a morphism of schemes by Yoneda's lemma. On the other hand, since $\mathrm{Spec}(A[\underline{M}]_d)$ is a scheme of monoids, the map

$$\begin{aligned} \mathrm{Spec}(A[\underline{M}]_d)(R) \times \mathrm{Spec}(A[\underline{M}]_d)(R) &\to \mathrm{Spec}(A[\underline{M}]_d)(R) \\ (m, m') &\mapsto mm' \end{aligned}$$

is represented by a morphism of schemes. We consider the composite of these two morphisms. It is the constant map (at the identity) at least at the level of R-points, for a flat A-algebra R. To show that the composite is the constant morphism of schemes (at the identity), it suffices to show that it is uniquely determined at the level of R-points, for a flat A-algebra R. Note that $\mathrm{Spec}(A[\underline{M}]_d)$ is an irreducible smooth affine scheme. We consider the open subscheme of $\mathrm{Spec}(A[\underline{M}]_d)$ which is the complement of the closed subscheme of $\mathrm{Spec}(A[\underline{M}]_d)$ determined by the prime ideal (2). This open subscheme of $\mathrm{Spec}(A[\underline{M}]_d)$

is then nonempty and dense since $\mathrm{Spec}(A[\underline{M}]_d)$ is reduced and irreducible. Furthermore, all R-points of $\mathrm{Spec}(A[\underline{M}]_d)$, for a flat A-algebra R, factor through this open subscheme. Since a morphism of schemes is continuous, the above composite is uniquely determined at the level of R-points, for a flat A-algebra R.

Thus, the inverse of $m \in \mathrm{Spec}(A[\underline{M}]_d)(R)$, for any commutative A-algebra R, is also contained in $\mathrm{Spec}(A[\underline{M}]_d)(R) \subseteq \underline{M}(R)$. This fact implies $\underline{M}^*(R) \supseteq \mathrm{Spec}(A[\underline{M}]_d)(R)$. Consequently, for any commutative A-algebra R, we have

$$\underline{M}^*(R) = \mathrm{Spec}(A[\underline{M}]_d)(R).$$

Therefore, we conclude that $\underline{M}^*$ is an open subscheme of $\underline{M}$ (since $\underline{M}^* = \mathrm{Spec}(A[\underline{M}]_d)$, which is an open subscheme of $\underline{M}$), with generic fiber $M^* = \mathrm{Res}_{E/F}\,\mathrm{GL}_E(V)$, and that $\underline{M}^*$ is smooth over A. Moreover, $\underline{M}^*$ is a group scheme since $\underline{M}$ is a scheme in monoids.

3C. *Construction of $\underline{H}$.* Recall that the pair (L, h) is fixed throughout this paper and the lattices A_i, B_i, W_i, X_i only depend on the hermitian pair (L, h). For any flat A-algebra R, let $\underline{H}(R)$ be the set of hermitian forms f on $L \otimes_A R$ (with values in $B \otimes_A R$) such that f satisfies the following conditions:

(a) $f(L \otimes_A R, A_i \otimes_A R) \subset \pi^i B \otimes_A R$ for all i.

(b) $\xi^{-m} f(a_i, a_i) \bmod 2 = \xi^{-m} h(a_i, a_i) \bmod 2$, where $a_i \in A_i \otimes_A R$, and $i = 2m$.

(c) $\frac{1}{\pi^i} f(a_i, w_i) = \frac{1}{\pi^i} h(a_i, w_i) \bmod \pi$, where $a_i \in A_i \otimes_A R$ and $w_i \in W_i \otimes_A R$, and $i = 2m$.

We interpret the above conditions in terms of matrices. The matrix forms are taken with respect to the basis of L fixed in Theorem 2.10 and Remark 2.3(a). A matrix form of the given hermitian form h is described in Remark 3.3(1) below. We use σ to mean the automorphism of $B \otimes_A R$ given by $b \otimes r \mapsto \sigma(b) \otimes r$. For a flat A-algebra R, $\underline{H}(R)$ is the set of hermitian matrices

$$\left(\pi^{\max\{i,j\}} f_{i,j}\right)$$

of size $n \times n$ satisfying the following:

(1) $f_{i,j}$ is an $(n_i \times n_j)$-matrix with entries in $B \otimes_A R$.

(2) If i is even and L_i is of type I^o, then $\pi^i f_{i,i}$ is of the form

$$\xi^{i/2}\begin{pmatrix} a_i & \pi b_i \\ \sigma(\pi \cdot {}^t b_i) & 1+2c_i \end{pmatrix}.$$

Here, the diagonal entries of a_i are divisible by 2, where a_i is an $(n_i-1)\times(n_i-1)$-matrix with entries in $B \otimes_A R$, etc.

(3) If i is even and L_i is of type I^e, then $\pi^i f_{i,i}$ is of the form

$$\xi^{i/2}\begin{pmatrix} a_i & b_i & \pi e_i \\ \sigma({}^t b_i) & 1+2f_i & 1+\pi d_i \\ \sigma(\pi \cdot {}^t e_i) & \sigma(1+\pi d_i) & 2c_i \end{pmatrix}.$$

Here, the diagonal entries of a_i are divisible by 2, where a_i is an $(n_i-2)\times(n_i-2)$-matrix with entries in $B \otimes_A R$, etc.

(4) Assume that L_i is of type II. The diagonal entries of $f_{i,i}$ (resp. $\pi f_{i,i}$) are divisible by 2 if i is even (resp. odd).

(5) Since $\left(\pi^{\max\{i,j\}} f_{i,j}\right)$ is a hermitian matrix, its diagonal entries are fixed by the nontrivial Galois action over E/F and hence belong to R.

Let us consider the *hermitian functor* from the category of commutative flat A-algebras to the category of sets such that the associated set to R is the set of hermitian forms f on $L \otimes_A R$ (with values in $B \otimes_A R$). Indeed, this functor is represented by a commutative group scheme since it is closed under addition. Then $\underline{H}$ is a subfunctor of the hermitian functor. We consider another functor $\underline{H}'$ such that $\underline{H}'(R) = \{f - h : f \in \underline{H}(R)\}$. Note that h is the fixed hermitian form and the notion of $f - h$ follows from the additive structure of the hermitian functor. For a matrix interpretation of h, we refer to Remark 3.3(1) below.

Then by Lemma 3.1 of [Cho 2015a], $\underline{H}'$ is represented by a flat A-scheme which is isomorphic to an affine space. Since $\underline{H}$ and $\underline{H}'$ are equivalent as subfunctors of the hermitian functor, the functor $\underline{H}$ is also represented by a flat A-scheme which is isomorphic to an affine space.

To compute the dimension of $\underline{H}$, we see that each entry of the upper triangular matrix of an element of $\underline{H}(R)$, for a flat A-algebra R, gives two variables and each diagonal entry gives one variable. Furthermore, each lower triangular entry of the matrix representing an element of $\underline{H}(R)$ is completely determined by the corresponding upper triangular entry. Thus the dimension of $\underline{H}$ is $2 \cdot n(n-1)/2 + n = n^2$. This is also the same as $2n^2 - \dim U(V, h) = n^2$.

Now suppose that R is any (not necessarily flat) A-algebra. Recall that ϵ is a unit in B such that $\sigma(\pi) = \epsilon\pi$ and $(\epsilon - 1)/\pi$ is a unit in B. We also use ϵ to mean $\epsilon \otimes 1$ in $B \otimes_A R$. We again use σ to mean the automorphism of $B \otimes_A R$ given by $b \otimes r \mapsto \sigma(b) \otimes r$. By choosing a B-basis of L as explained in Theorem 2.10 and Remark 2.3(a), we describe each element of $\underline{H}(R)$ formally as a matrix $\left(\pi^{\max\{i,j\}} f_{i,j}\right)$ with the following:

(1) When $i \neq j$, $f_{i,j}$ is an $(n_i \times n_j)$-matrix with entries in $B \otimes_A R$ and $\epsilon^{\max\{i,j\}} \sigma({}^t f_{i,j}) = f_{j,i}$.

(2) Assume that $i = j$ is even. Then

$$\pi^i f_{i,i} = \begin{cases} \xi^{i/2} \begin{pmatrix} a_i & \pi b_i \\ \sigma(\pi \cdot {}^t b_i) & 1 + 2c_i \end{pmatrix} & \text{if } L_i \text{ is of type } I^o; \\ \xi^{i/2} \begin{pmatrix} a_i & b_i & \pi e_i \\ \sigma({}^t b_i) & 1 + 2f_i & 1 + \pi d_i \\ \sigma(\pi \cdot {}^t e_i) & \sigma(1 + \pi d_i) & 2c_i \end{pmatrix} & \text{if } L_i \text{ is of type } I^e; \\ \xi^{i/2} a_i & \text{if } L_i \text{ is of type } II. \end{cases}$$

Here, a_i is a formal $(n_i - 1 \times n_i - 1)$-matrix (resp. $(n_i - 2 \times n_i - 2)$-matrix or $(n_i \times n_i)$-matrix) when L_i is of type I^o (resp. of type I^e or of type II). Nondiagonal entries of a_i are in $B \otimes_A R$ and the j-th diagonal entry of a_i is of the form $2x_i^j$ with $x_i^j \in R$. In addition, for nondiagonal entries of a_i, we have the relation $\sigma({}^t a_i) = a_i$. And b_i, d_i, e_i are matrices of suitable sizes with entries in $B \otimes_A R$ and c_i, f_i are elements in R.

(3) Assume that $i = j$ is odd. Then

$$\pi^i f_{i,i} = \xi^{(i-1)/2} \pi a_i,$$

where a_i is a formal $(n_i \times n_i)$-matrix. Here, nondiagonal entries of a_i are in $B \otimes_A R$ and the j-th diagonal entry of a_i is of the form $\epsilon\pi x_i^j$ with $x_i^j \in R$. In addition, for nondiagonal entries of a_i, we have the relation $\epsilon \cdot \sigma({}^t a_i) = a_i$.

To simplify notation, each element

$$((f_{i,j})_{i<j}, (a_i, x_i^j)_{L_i \text{ of type } II}, (a_i, x_i^j, b_i, c_i)_{L_i \text{ of type } I^o}, (a_i, x_i^j, b_i, c_i, d_i, e_i, f_i)_{L_i \text{ of type } I^e})$$

of $\underline{H}(R)$ is denoted by $(f_{i,j}, a_i \cdots f_i)$.

Remark 3.3. (1) Note that the given hermitian form h is an element of $\underline{H}(A)$. We represent the given hermitian form h by a hermitian matrix $(\pi^i \cdot h_i)$ whose (i,i)-block is $\pi^i \cdot h_i$ for all i, and all of whose remaining blocks are 0. Then:

(a) If i is even and L_i is of type I^o, then $\pi^i \cdot h_i$ has the following form (with $\gamma_i \in A$):

$$\xi^{i/2}\begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \\ & & & 1+2\gamma_i \end{pmatrix}.$$

(b) If i is even and L_i is of type I^e, then $\pi^i \cdot h_i$ has the following form (with $\gamma_i \in A$):

$$\xi^{i/2}\begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \\ & & & \begin{pmatrix} 1 & 1 \\ 1 & 2\gamma_i \end{pmatrix} \end{pmatrix}.$$

(c) If i is even and L_i is of type II, then $\pi^i \cdot h_i$ has the following form:

$$\xi^{i/2}\begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix}.$$

(d) If i is odd, then $\pi^i \cdot h_i$ has the following form (with $\gamma_i \in A$):

$$\xi^{(i-1)/2}\begin{pmatrix} \begin{pmatrix} 0 & \pi \\ \sigma(\pi) & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & \pi \\ \sigma(\pi) & 0 \end{pmatrix} & \\ & & & \begin{pmatrix} 2 & \pi \\ \sigma(\pi) & 2\gamma_i \end{pmatrix} \end{pmatrix}.$$

(2) Let R be a κ-algebra. We also denote by h the element of $\underline{H}(R)$ which is the image of $h \in \underline{H}(A)$ under the natural map from $\underline{H}(A)$ to $\underline{H}(R)$. Recall that we denote each element of $\underline{H}(R)$ by $(f_{i,j}, a_i \cdots f_i)$. Then the tuple $(f_{i,j}, a_i \cdots f_i)$ denoting $h \in \underline{H}(R)$ is defined by the conditions:

(a) If $i \neq j$, then $f_{i,j} = 0$.

(b) If i is even, then

$$a_i = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix}, \quad \text{thus } x_i^j = 0,$$

$$b_i = 0, d_i = 0, e_i = 0, f_i = 0, c_i = \bar{\gamma}_i.$$

Here, $\bar{\gamma}_i \in \kappa$ is the reduction of $\gamma_i \mod 2$.

(c) If i is odd, then

$$a_i = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ \bar{\epsilon} & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & 1 \\ \bar{\epsilon} & 0 \end{pmatrix} & \\ & & & \begin{pmatrix} \pi \cdot \bar{\epsilon}\bar{\zeta} & 1 \\ \bar{\epsilon} & \pi \cdot \bar{\epsilon}\bar{\zeta}\bar{\gamma}_i \end{pmatrix} \end{pmatrix}.$$

Here, $\bar{\epsilon}$ is the reduction of $\epsilon \mod 2$, not mod π, so that $\bar{\epsilon}$ is an element of $B \otimes_A R$. In addition, $\bar{\zeta} \in \kappa$ is the reduction of $\zeta \mod 2$, where $\zeta \in A$ is the unit satisfying $2 = \xi \cdot \zeta$. Thus, $x_i^j = 0$ for all $1 \leq j \leq n_i - 2$ and $x_i^{n_i - 1} = \bar{\zeta}$ and $x_i^{n_i} = \bar{\zeta}\bar{\gamma}_i$.

3D. *The smooth affine group scheme $\underline{G}$.*

Theorem 3.4. *For any flat A-algebra R, the group $\underline{M}^*(R)$ acts on $\underline{H}(R)$ on the right by $f \circ m = \sigma({}^t m) \cdot f \cdot m$. This action is represented by an action morphism*

$$\underline{H} \times \underline{M}^* \longrightarrow \underline{H}.$$

Proof. We start with any $m \in \underline{M}^*(R)$ and $f \in \underline{H}(R)$. In order to show that $\underline{M}^*(R)$ acts on the right of $\underline{H}(R)$ by $f \circ m = \sigma({}^t m) \cdot f \cdot m$, it suffices to show that $f \circ m$ satisfies conditions (a) to (c) given in Section 3C. Since elements of $\underline{M}(R)$ preserve $L \otimes_A R$ and $A_i \otimes_A R$, $f \circ m$ satisfies condition (a). That $f \circ m$ satisfies condition (b) follows from the fact that m stabilizes A_i and B_i and induces the identity on A_i/B_i.

For condition (c), it suffices to show that $\frac{1}{\pi^i} f(ma_i, mw_i) \equiv \frac{1}{\pi^i} f(a_i, w_i) \mod \pi$. We denote $ma_i = a_i + b_i$ and $mw_i = w_i + x_i$, where $b_i \in B_i \otimes_A R$, $x_i \in X_i \otimes_A R$. Hence it suffices to show $\frac{1}{\pi^i} f(a_i + b_i, x_i) + \frac{1}{\pi^i} f(b_i, w_i) \mod \pi \equiv 0$. Firstly, $\frac{1}{\pi^i} f(a_i + b_i, x_i) \mod \pi \equiv 0$ due to the definition of the lattice X_i. Secondly, if $B_i \subsetneq A_i$, then $\frac{1}{\pi^i} f(b_i, w_i) \mod \pi \equiv 0$ because $\frac{1}{\pi^i} f(b_i, w_i) = \frac{1}{\pi^i} h(b_i, w_i) \mod \pi$ and $(\xi^{-m} h(b_i, e))^2 \equiv \xi^{-m} h(b_i, b_i) \equiv 0 \mod \pi$, where e is the unique vector chosen earlier. If $B_i = A_i$, then $W_i = X_i$ and thus $\frac{1}{\pi^i} f(b_i, w_i) \mod \pi \equiv 0$.

We now show that this action of the group $\underline{M}^*(R)$ on the right of $\underline{H}(R)$ is represented by an action morphism of schemes. We observe that the action map $\underline{H}(R) \times \underline{M}^*(R) \longrightarrow \underline{H}(R)$, $(f, m) \mapsto \sigma({}^t m) \cdot f \cdot m$ is given by polynomials over A. Thus it induces a ring homomorphism over A from the coordinate ring of $\underline{H}$ to the coordinate ring of $\underline{H} \times \underline{M}^*$, which accordingly induces a morphism from $\underline{H} \times \underline{M}^*$ to $\underline{H}$ such that the action map induced by this morphism at the level of R-points, for a flat A-algebra R, is the same as the action given in the theorem. □

Remark 3.5. Let R be a κ-algebra. We explain the above action morphism in terms of R-points. Choose an element $(m_{i,j}, s_i \cdots w_i)$ in $\underline{M}^*(R)$ as explained in Section 3B and express this element formally as a matrix $m = \left(\pi^{\max\{0, j-i\}} m_{i,j}\right)$. We also choose an element $(f_{i,j}, a_i \cdots f_i)$ of $\underline{H}(R)$ and express this element formally as a matrix $f = \left(\pi^{\max\{i,j\}} f_{i,j}\right)$ as explained in Section 3C.

We then compute the formal matrix product $\sigma({}^t m) \cdot f \cdot m$ and denote it by the formal matrix $\left(\pi^{\max\{i,j\}} \tilde{f}'_{i,j}\right)$ with $(\tilde{f}'_{i,j}, \tilde{a}'_i \cdots \tilde{f}'_i)$. Here, the description of the formal matrix $\left(\pi^{\max\{i,j\}} \tilde{f}'_{i,j}\right)$ with $(\tilde{f}'_{i,j}, \tilde{a}'_i \cdots \tilde{f}'_i)$ is as explained in Section 3C.

We now let π^2 be zero in each entry of the formal matrices $(\tilde{f}'_{i,j})_{i<j}$, $(\tilde{b}'_i)_{L_i \text{ of type } I^o}$, $(\tilde{b}'_i, \tilde{d}'_i, \tilde{e}'_i)_{L_i \text{ of type } I^e}$ and in each nondiagonal entry of the formal matrix $(\tilde{a}'_i)$. Then these entries are elements in $B \otimes_A R$. We also let π^2 be zero in $(\tilde{x}_i^j)'$, $(\tilde{c}'_i)_{L_i \text{ of type } I^o}$, $(\tilde{f}'_i, \tilde{c}'_i)_{L_i \text{ of type } I^e}$. Note that $(\tilde{x}_i^j)'$ is a diagonal entry of a formal matrix $\tilde{a}'_i$. Then these entries are elements in R.

Let $(f'_{i,j}, a'_i \cdots f'_i)$ be the reduction of $(\tilde{f}'_{i,j}, \tilde{a}'_i \cdots \tilde{f}'_i)$ as explained above, i.e., by letting π^2 be zero in the entries of formal matrices as described above. Then $(f'_{i,j}, a'_i \cdots f'_i)$ is an element of $\underline{H}(R)$ and the composition $(f_{i,j}, a_i \cdots f_i) \circ (m_{i,j}, s_i \cdots w_i)$ is $(f'_{i,j}, a'_i \cdots f'_i)$.

We can also write $(f'_{i,j}, a'_i \cdots f'_i)$ explicitly in terms of $(f_{i,j}, a_i \cdots f_i)$ and $(m_{i,j}, s_i \cdots w_i)$ like the product of $(m_{i,j}, s_i \cdots w_i)$ and $(m'_{i,j}, s'_i \cdots w'_i)$ explained in Section 3B. However, this is complicated and we do not use it in this generality. On the other hand, we explicitly calculate $(f_{i,j}, a_i \cdots f_i) \circ (m_{i,j}, s_i \cdots w_i)$ when $(f_{i,j}, a_i \cdots f_i)$ is the given hermitian form h and $(m_{i,j}, s_i \cdots w_i)$ satisfies certain conditions on each block. This explicit calculation will be done in Appendix A.

Theorem 3.6. *Let ρ be the morphism $\underline{M}^* \to \underline{H}$ defined by $\rho(m) = h \circ m$, which is induced by the action morphism of Theorem 3.4. Then ρ is smooth of relative dimension* $\dim U(V, h)$.

Proof. The theorem follows from Lemma 5.5.1 of [Gan and Yu 2000] and the following lemma. □

Lemma 3.7. *The morphism $\rho \otimes \kappa : \underline{M}^* \otimes \kappa \to \underline{H} \otimes \kappa$ is smooth of relative dimension* $\dim U(V, h)$.

Proof. The proof is based on Lemma 5.5.2 in [Gan and Yu 2000]. It is enough to check the statement over the algebraic closure $\bar{\kappa}$ of κ. By [Hartshorne 1977, Proposition III.10.4], it suffices to show that, for any $m \in \underline{M}^*(\bar{\kappa})$, the induced map on the Zariski tangent space $\rho_{*,m} : T_m \to T_{\rho(m)}$ is surjective.

We define the two functors from the category of commutative flat A-algebras to the category of abelian groups as follows:

$$T_1(R) = \{m - 1 : m \in \underline{M}(R)\},$$
$$T_2(R) = \{f - h : f \in \underline{H}(R)\}.$$

The functor T_1 (resp. T_2) is representable by a flat A-algebra which is a polynomial ring over A of $2n^2$ (resp. n^2) variables by Lemma 3.1 of [Cho 2015a]. Moreover, each of them is represented by a commutative group scheme since they are closed under addition. In fact, T_1 is the same as the functor $\underline{M}'$ in Remark 3.1 and T_2 is the same as the functor $\underline{H}'$ in Section 3C.

We still need to introduce another functor on flat A-algebras. Define $T_3(R)$ to be the set of all $(n \times n)$-matrices y over $B \otimes_A R$ satisfying the following conditions:

(a) The (i, j)-block of y has entries in $\pi^{\max\{i,j\}} B \otimes_A R$ so that

$$y = \left(\pi^{\max\{i,j\}} y_{i,j}\right).$$

Here, the size of $y_{i,j}$ is $n_i \times n_j$.

(b) If i is even and L_i is of type I^o, then $y_{i,i}$ is of the form

$$\begin{pmatrix} s_i & \pi y_i \\ \pi v_i & \pi z_i \end{pmatrix} \in M_{n_i}(B \otimes_A R)$$

where s_i is an $(n_i - 1) \times (n_i - 1)$-matrix, etc.

(c) If i is even and L_i is of type I^e, then $y_{i,i}$ is of the form

$$\begin{pmatrix} s_i & r_i & \pi t_i \\ y_i & x_i & \pi z_i \\ \pi v_i & \pi u_i & \pi w_i \end{pmatrix} \in M_{n_i}(B \otimes_A R)$$

where s_i is an $(n_i - 2) \times (n_i - 2)$-matrix, etc.

The functor T_3 is represented by a flat A-scheme which is isomorphic to an affine space by Lemma 3.1 of [Cho 2015a]. Moreover it is represented by a commutative group scheme since it is closed under addition. So far, we have defined three functors T_1, T_2, T_3 and these are represented by schemes. Therefore, we can talk about their $\bar{\kappa}$-points.

We now compute the map $\rho_{*,m}$ explicitly. We first describe an element of the tangent space T_m. Since $\underline{M}^*$ is an open subscheme of $\underline{M}$, the tangent space T_m may and shall be identified with the set of elements of $\underline{M}(\bar{\kappa}[\epsilon]/(\epsilon^2))$ whose reduction to $\underline{M}(\bar{\kappa})$ induced by the obvious map $\bar{\kappa}[\epsilon]/(\epsilon^2) \to \bar{\kappa}$ is m, by considering m as an element of $\underline{M}(\bar{\kappa})$. Recall from Remark 3.1 that we defined the functor $\underline{M}'$ such that $(1 + \underline{M}')(R) = \underline{M}(R)$ inside $\mathrm{End}_{B\otimes_A R}(L \otimes_A R)$ for a flat A-algebra R. Thus there is an isomorphism of schemes (as set valued functors)

$$1+ : \underline{M}' \longrightarrow \underline{M}.$$

Let m' be an element of $\underline{M}'(\bar{\kappa})$ which maps to m under the morphism $1+$ at the level of $\bar{\kappa}$-points. Then each element of the tangent space of $\underline{M}'$ at m' is of the form $m' + \epsilon X \in$

$\underline{M}'(\bar{\kappa}[\epsilon]/(\epsilon^2))$ for $X \in \underline{M}'(\bar{\kappa})$. We denote by $m+\epsilon X$ the image of $m'+\epsilon X$ under the morphism $1+$ at the level of $\bar{\kappa}[\epsilon]/(\epsilon^2)$-points. Thus we can express an element of T_m formally as $m+\epsilon X$ where $X \in \underline{M}'(\bar{\kappa})$. Similarly, an element of $T_{\rho(m)}$ can be expressed formally as $\rho(m)+\epsilon Y$ where $Y \in \underline{H}'(\bar{\kappa})$, by using an isomorphism of schemes (as set valued functors)

$$h+ : \underline{H}' \longrightarrow \underline{H}.$$

Here, $\underline{H}'$ is defined in Section 3C.

Before observing the image of $m+\epsilon X$ under the morphism ρ at the level of $\bar{\kappa}[\epsilon]/(\epsilon^2)$-points, we lift $m+\epsilon X$ to an element of $\underline{M}(R[\epsilon]/(\epsilon^2))$ as follows, where R is a local A-algebra whose residue field is $\bar{\kappa}$. Let $\tilde{m}' \in \underline{M}'(R)$ (resp. $\widetilde{X} \in \underline{M}'(R)$) be a lift of m' (resp. X) so that $\tilde{m}'+\epsilon\widetilde{X} \in \underline{M}'(R[\epsilon]/(\epsilon^2))$ is a lift of $m'+\epsilon X \in \underline{M}'(\bar{\kappa}[\epsilon]/(\epsilon^2))$. Let $\tilde{m} \in \underline{M}(R)$ be the image of $\tilde{m}'$ under the morphism $1+$. Then $\tilde{m}+\epsilon\widetilde{X}$ is an element of $\underline{M}(R[\epsilon]/(\epsilon^2))$ whose reduction to $\underline{M}(\bar{\kappa}[\epsilon]/(\epsilon^2))$ induced by the map $R[\epsilon]/(\epsilon^2) \to \bar{\kappa}[\epsilon]/(\epsilon^2)$ is $m+\epsilon X$. Here, the addition in $\tilde{m}+\epsilon\widetilde{X}$ is the addition inside $\mathrm{End}_{B\otimes_A R[\epsilon]/(\epsilon^2)}(L\otimes_A R[\epsilon]/(\epsilon^2))$ since $R[\epsilon]/(\epsilon^2)$ is flat over A (cf. Remark 3.1). This is illustrated in the following commutative diagrams:

$$\begin{array}{ccc}
\underline{M}'(R[\epsilon]/(\epsilon^2)) & \xrightarrow{1+} & \underline{M}(R[\epsilon]/(\epsilon^2)) \\
\downarrow & & \downarrow \\
\underline{M}'(\bar{\kappa}[\epsilon]/(\epsilon^2)) & \xrightarrow{1+} & \underline{M}(\bar{\kappa}[\epsilon]/(\epsilon^2))
\end{array}$$

$$\begin{array}{ccc}
\tilde{m}'+\epsilon\widetilde{X} & \longmapsto & \tilde{m}+\epsilon\widetilde{X} \\
\downarrow & & \downarrow \\
m'+\epsilon X & \longmapsto & m+\epsilon X
\end{array}$$

Note that the proof of Theorem 3.4 also gives the existence of the morphism $\underline{H}\times\underline{M} \to \underline{H}$, defined by $(f, m) \mapsto f\circ m = \sigma({}^t m)\cdot f\cdot m$, where $f \in \underline{H}(R)$ and $m \in \underline{M}(R)$ for a flat A-algebra R. This morphism induces the morphism $\underline{M} \to \underline{H}$ with $m \mapsto h\circ m$ whose reduction to $\underline{M}^*$ is the same as ρ. Thus the above morphism $\underline{M} \to \underline{H}$ can also be denoted by ρ. We can now talk about the image of $\tilde{m}+\epsilon\widetilde{X}$ under the morphism ρ at the level of $R[\epsilon]/(\epsilon^2)$-points. Since $R[\epsilon]/(\epsilon^2)$ is a flat A-algebra, the image of $\tilde{m}+\epsilon\widetilde{X}$ comes from a usual matrix product

$$\sigma(\tilde{m}+\epsilon\widetilde{X})^t\cdot h\cdot(\tilde{m}+\epsilon\widetilde{X}) = \sigma(\tilde{m})^t\cdot h\cdot\tilde{m}+\epsilon(\sigma(\tilde{m})^t\cdot h\cdot\widetilde{X}+\sigma(\widetilde{X})^t\cdot h\cdot\tilde{m}). \tag{3-1}$$

Thus the image of $m+\epsilon X$ under the morphism ρ at the level of $\bar{\kappa}[\epsilon]/(\epsilon^2)$-points is the reduction of $\sigma(\tilde{m})^t\cdot h\cdot\tilde{m}+\epsilon(\sigma(\tilde{m})^t\cdot h\cdot\widetilde{X}+\sigma(\widetilde{X})^t\cdot h\cdot\tilde{m})$ to $\underline{H}(\bar{\kappa}[\epsilon]/(\epsilon^2))$. It is obvious that $\rho(m)$ ($\in \underline{H}(\bar{\kappa})$) is the reduction of $\sigma(\tilde{m})^t\cdot h\cdot\tilde{m}$ ($\in \underline{H}(R)$) since $\tilde{m}$ is a lift of m and ρ is a morphism of schemes. To observe the reduction of $\sigma(\tilde{m})^t\cdot h\cdot\widetilde{X}+\sigma(\widetilde{X})^t\cdot h\cdot\tilde{m}$ ($\in \underline{H}'(R)$) to $\underline{H}'(\bar{\kappa})$, we consider a morphism $\underline{M}\times\underline{H}' \to \underline{H}'$ such that $(\tilde{m}, \widetilde{X})$ maps to $\sigma(\tilde{m})^t\cdot h\cdot\widetilde{X}+\sigma(\widetilde{X})^t\cdot h\cdot\tilde{m}$, where $(\tilde{m}, \widetilde{X}) \in \underline{M}(R)\times\underline{H}'(R)$ for a flat A-algebra R. To show that this map is well-defined, we need to show that $\sigma(\tilde{m})^t\cdot h\cdot\widetilde{X}+\sigma(\widetilde{X})^t\cdot h\cdot\tilde{m}$ is an element of $\underline{H}'(R)$. This can be easily shown by considering the morphism of tangent spaces induced from ρ at $\tilde{m} \in \underline{M}(R)$ (cf. Equation (3-1)). Since this morphism is representable, we can denote by $\sigma(m)^t \cdot h \cdot X + \sigma(X)^t \cdot h \cdot m$ ($\in \underline{H}'(\bar{\kappa})$) the reduction of $\sigma(\tilde{m})^t \cdot h \cdot \widetilde{X} + \sigma(\widetilde{X})^t \cdot h \cdot \tilde{m}$ ($\in \underline{H}'(R)$) to $\underline{H}'(\bar{\kappa})$. Then the image of $m + \epsilon X$ is a formal sum $\rho(m) + \epsilon(\sigma(m)^t \cdot h \cdot X + \sigma(X)^t \cdot h \cdot m)$ ($\in \underline{H}(\bar{\kappa}[\epsilon]/(\epsilon^2))$).

Thus if we identify T_m with $T_1(\bar{\kappa})$ and $T_{\rho(m)}$ with $T_2(\bar{\kappa})$, then

$$\begin{aligned}\rho_{*,m} : T_m &\to T_{\rho(m)}\\ X &\mapsto \sigma(m)^t \cdot h \cdot X + \sigma(X)^t \cdot h \cdot m.\end{aligned}$$

We explain how to compute $X \mapsto \sigma(m)^t \cdot h \cdot X + \sigma(X)^t \cdot h \cdot m$ explicitly. Recall that for a κ-algebra R, we denote an element m of $\underline{M}(R)$ by $(m_{i,j}, s_i \cdots w_i)$ with a formal matrix interpretation $m = (\pi^{\max\{0,j-i\}} m_{i,j})$ (cf. Section 3B) and we denote an element f of $\underline{H}(R)$ by $(f_{i,j}, a_i \cdots f_i)$ with a formal matrix interpretation $f = (\pi^{\max\{i,j\}} f_{i,j})$ (cf. Section 3C). Similarly, we can also denote an element X of $T_1(\bar{\kappa})$ by $(m'_{i,j}, s'_i \cdots w'_i)$ with a formal matrix interpretation $X = (\pi^{\max\{0,j-i\}} m'_{i,j})$ and an element Z of $T_2(\bar{\kappa})$ by $(f'_{i,j}, a'_i \cdots f'_i)$ with a formal matrix interpretation $Z = (\pi^{\max\{i,j\}} f'_{i,j})$. Then we formally compute $X \mapsto \sigma(m^t) \cdot h \cdot X + \sigma(X^t) \cdot h \cdot m$ and consider the reduction of the formal matrix $\sigma(m^t) \cdot h \cdot X + \sigma(X^t) \cdot h \cdot m$ in a manner similar to that of the reduction explained in Remark 3.5. We denote this reduction by $(f''_{i,j}, a''_i \cdots f''_i)$ with a formal matrix interpretation $(\pi^{\max\{i,j\}} f''_{i,j})$. This $(f''_{i,j}, a''_i \cdots f''_i)$ may and shall be identified with an element of $T_2(\bar{\kappa})$ in the manner just described. Then $\rho_{*,m}(X)$ is the element $Z = (f''_{i,j}, a''_i \cdots f''_i)$ of $T_2(\bar{\kappa})$.

To prove the surjectivity of $\rho_{*,m} : T_1(\bar{\kappa}) \to T_2(\bar{\kappa})$, it suffices to show the following three statements:

(1) $X \mapsto h \cdot X$ defines a bijection $T_1(\bar{\kappa}) \to T_3(\bar{\kappa})$;

(2) for any $m \in \underline{M}^*(\bar{\kappa})$, $Y \mapsto \sigma({}^t m) \cdot Y$ defines a bijection from $T_3(\bar{\kappa})$ to itself;

(3) $Y \mapsto \sigma({}^t Y) + Y$ defines a surjection $T_3(\bar{\kappa}) \to T_2(\bar{\kappa})$.

Here, all the above maps are interpreted as in Remark 3.5 (if they are well-defined). Then $\rho_{*,m}$ is the composite of these three. Condition (3) is direct from the construction of $T_3(\bar{\kappa})$. Hence we provide the proof of (1) and (2).

For (1), suppose that the two functors $T_1(R) \longrightarrow T_3(R)$, $X \mapsto h \cdot X (\in M_{n\times n}(B \otimes_A R))$ and $T_3(R) \longrightarrow T_1(R)$, $Y \mapsto h^{-1} \cdot Y (\in M_{n\times n}(B \otimes_A R))$ are well-defined for all flat A-algebras R. In other words, suppose that $h \cdot X \in T_3(R)$ and $h^{-1} \cdot Y \in T_1(R)$. These functors are then represented by morphisms of schemes by an argument similar to that used in the proof of Theorem 3.4, so we skip it. Thus they give maps at the level of κ-algebra points. Furthermore, the composition of these two maps at the level of κ-algebra points is the identity. To show this, it suffices to prove that the composition of two morphisms given by the actions of h and h^{-1} is uniquely determined at the level of R-points, for a flat A-algebra R. This is proved in Remark 3.2.

We now show that these two functors are well-defined for a flat A-algebra R. We represent h by a hermitian block matrix $(\pi^i \cdot h_i)$ with a matrix $(\pi^i \cdot h_i)$ for the (i,i)-block and 0 for the remaining blocks as in Remark 3.3(1).

For the first functor, it suffices to show that $h \cdot X$ satisfies the three conditions defining the functor T_3. Here, $X \in T_1(R)$ for a flat A-algebra R. We write

$$X = \left(\pi^{\max\{0, j-i\}} x_{i,j}\right).$$

Then

$$h \cdot X = \left(\pi^{\max(i,j)} y_{i,j}\right).$$

Here, $y_{i,i} = h_i \cdot x_{i,i}$. Therefore, it suffices to show that $y_{i,i} = h_i \cdot x_{i,i}$ satisfies conditions (b) and (c) in the description of $T_3(R)$ when L_i is of type I.

If L_i is of type I^o, then we express $x_{i,i}$ as a matrix $\left(\begin{smallmatrix} s_i & \pi y_i \\ \pi v_i & \pi z_i \end{smallmatrix}\right)$. The matrix form of h_i is

$$\epsilon^{i/2} \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \\ & & & 1+2\gamma_i \end{pmatrix}$$

as in Remark 3.3(1). Here, ϵ is a unit in B such that $\sigma(\pi) = \epsilon\pi$, as explained in Section 2A. To simplify our notation, write $h_i = \epsilon^{i/2}\left(\begin{smallmatrix} I_i & 0 \\ 0 & 1+2\gamma_i \end{smallmatrix}\right)$. Then we can see that

$$h_i \cdot x_{i,i} = \epsilon^{i/2} \begin{pmatrix} I_i & 0 \\ 0 & 1+2\gamma_i \end{pmatrix} \cdot \begin{pmatrix} s_i & \pi y_i \\ \pi v_i & \pi z_i \end{pmatrix} = \epsilon^{i/2} \begin{pmatrix} I_i s_i & \pi I_i y_i \\ \pi(1+2\gamma_i) v_i & \pi(1+2\gamma_i) z_i \end{pmatrix}.$$

Thus, $h_i \cdot x_{i,i}$ satisfies the congruence condition given in (b) of the description of $T_3(R)$.

If L_i is of type I^e, then we express $x_{i,i}$ as a matrix

$$\begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & \pi x_i & \pi z_i \\ v_i & u_i & \pi w_i \end{pmatrix}.$$

The matrix form of h_i is given as in Remark 3.3(1) and again, in order to simplify our notation, write

$$h_i = \epsilon^{i/2} \begin{pmatrix} I_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2\gamma_i \end{pmatrix}.$$

Then we can see that

$$\begin{aligned} h_i \cdot x_{i,i} &= \epsilon^{i/2} \begin{pmatrix} I_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2\gamma_i \end{pmatrix} \cdot \begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & \pi x_i & \pi z_i \\ v_i & u_i & \pi w_i \end{pmatrix} \\ &= \epsilon^{i/2} \begin{pmatrix} I_i s_i & I_i r_i & \pi I_i t_i \\ \pi y_i + v_i & \pi x_i + u_i & \pi(z_i + w_i) \\ \pi y_i + 2\gamma_i v_i & \pi x_i + 2\gamma_i u_i & \pi z_i + 2\gamma_i \pi w_i \end{pmatrix}. \end{aligned}$$

Thus, $h_i \cdot x_{i,i}$ satisfies the congruence condition given in c) of the description of $T_3(R)$ and our functor is well-defined.

For the second functor, we write $Y=\left(\pi^{\max(i,j)}y_{i,j}\right)$ and $h^{-1}=\left(\pi^{-i}\cdot h_i^{-1}\right)$. Then we have the following:

$$h^{-1}\cdot Y=\left(\pi^{\max\{0,j-i\}}x_{i,j}\right).$$

Here, $x_{i,i}=h_i^{-1}\cdot y_{i,i}$.

Then it suffices to show that $h^{-1}\cdot Y=\left(\pi^{\max\{0,j-i\}}x_{i,j}\right)$ satisfies the conditions defining $T_1(R)$ for a flat A-algebra R. Indeed, we do not describe the conditions defining $T_1(R)$ explicitly in this paper. However, these conditions can be read off from the conditions defining $\underline{M}(R)$ because of the definition of the functor T_1. The matrix form of an element of $\underline{M}(R)$ is described in Section 3B and based on this, it suffices to observe the diagonal blocks $x_{i,i}=h_i^{-1}\cdot y_{i,i}$ when L_i is of type I.

If L_i is of type I^o, then we express $y_{i,i}$ as a matrix $\left(\begin{smallmatrix} s_i & \pi y_i\\ \pi v_i & \pi z_i\end{smallmatrix}\right)$. The matrix form of h_i^{-1} is $h_i^{-1}=\epsilon^{-i/2}\left(\begin{smallmatrix} I_i & 0\\ 0 & 1+2\gamma_i'\end{smallmatrix}\right)$ for a certain $\gamma_i'\in A$. Then we can see that

$$h_i^{-1}\cdot y_{i,i}=\epsilon^{-i/2}\begin{pmatrix} I_i & 0\\ 0 & 1+2\gamma_i'\end{pmatrix}\cdot\begin{pmatrix} s_i & \pi y_i\\ \pi v_i & \pi z_i\end{pmatrix}=\epsilon^{i/2}\begin{pmatrix} I_i s_i & \pi I_i y_i\\ \pi(1+2\gamma_i')v_i & \pi(1+2\gamma_i')z_i\end{pmatrix}.$$

Thus, $h_i^{-1}\cdot y_{i,i}$ satisfies the relevant congruence condition in the definition of $T_1(R)$.

If L_i is of type I^e, then we express $y_{i,i}$ as a matrix

$$\begin{pmatrix} s_i & r_i & \pi t_i\\ y_i & x_i & \pi z_i\\ \pi v_i & \pi u_i & \pi w_i\end{pmatrix}.$$

The matrix form of h_i^{-1} is

$$h_i^{-1}=\epsilon^{-i/2}\begin{pmatrix} I_i & 0 & 0\\ 0 & 2\epsilon'\gamma_i & -\epsilon'\\ 0 & -\epsilon' & \epsilon'\end{pmatrix}.$$

Here, $\epsilon'=(2\gamma_i-1)^{-1}$ is a unit in A. Then we can see that $h_i^{-1}\cdot y_{i,i}$ is

$$\begin{aligned}&\epsilon^{-i/2}\begin{pmatrix} I_i & 0 & 0\\ 0 & 2\epsilon'\gamma_i & -\epsilon'\\ 0 & -\epsilon' & \epsilon'\end{pmatrix}\cdot\begin{pmatrix} s_i & r_i & \pi t_i\\ y_i & x_i & \pi z_i\\ \pi v_i & \pi u_i & \pi w_i\end{pmatrix}\\ &\qquad=\epsilon^{-i/2}\begin{pmatrix} I_i s_i & I_i r_i & \pi I_i t_i\\ \epsilon'(2\gamma_i y_i-\pi v_i) & \epsilon'(2\gamma_i x_i-\pi u_i) & \pi\epsilon'(2\gamma_i z_i-w_i)\\ \epsilon'(-y_i+\pi v_i) & \epsilon'(-x_i+\pi u_i) & \pi\epsilon'(-z_i+w_i)\end{pmatrix}.\end{aligned}$$

Thus, $h_i^{-1}\cdot y_{i,i}$ satisfies the relevant congruence condition in the definition of $T_1(R)$ and our functor is well-defined.

For (2), suppose that the functor

$$\underline{M}^*(R)\times T_3(R)\longrightarrow T_3(R),\ (m,Y)\mapsto\sigma({}^tm)\cdot Y,$$

for a flat A-algebra R, is well-defined. In other words, we suppose that $\sigma({}^tm)\cdot Y\in T_3(R)$. This functor is then represented by a morphism of schemes, a fact whose proof is similar to the argument used in the proof of Theorem 3.4, so we skip it. Thus it gives the map at the level of $\bar{\kappa}$-points

$$\underline{M}^*(\bar{\kappa}) \times T_3(\bar{\kappa}) \longrightarrow T_3(\bar{\kappa}),\ (m, Y) \mapsto \sigma({}^t m) \cdot Y.$$

This map implies that our map in (2) is well-defined. On the other hand, the inverse of our map in (2) is $Y \mapsto \sigma({}^t m)^{-1} \cdot Y$ and this map is well-defined as well since m^{-1} is also an element of $\underline{M}^*(\bar{\kappa})$. Therefore, the map in (2) is a bijection.

We now show that the above functor is well-defined. For a flat A-algebra, we choose an element $m \in \underline{M}^*(R)$ and $Y \in T_3(R)$ and we again express $m = \left(\pi^{\max\{0, j-i\}} m_{i,j}\right)$ and $Y = \left(\pi^{\max(i,j)} y_{i,j}\right)$. Then $\sigma({}^t m) \cdot Y$ obviously satisfies condition (a) in the definition of $T_3(R)$ and it suffices to show that $\sigma({}^t m_{i,i}) \cdot y_{i,i}$ satisfies conditions (b) and (c) when L_i is of type I.

If L_i is of type I^o, then we express $m_{i,i}$ as a matrix $\left(\begin{smallmatrix} s_i & \pi y_i \\ \pi v_i & 1+\pi z_i \end{smallmatrix}\right)$ and $y_{i,i}$ as a matrix $\left(\begin{smallmatrix} a_i & \pi b_i \\ \pi c_i & \pi d_i \end{smallmatrix}\right)$. Then

$$\sigma({}^t m_{i,i}) \cdot y_{i,i} = \begin{pmatrix} \sigma({}^t s_i) & \sigma(\pi \cdot {}^t v_i) \\ \sigma(\pi \cdot {}^t y_i) & 1+\sigma(\pi z_i) \end{pmatrix} \cdot \begin{pmatrix} a_i & \pi b_i \\ \pi c_i & \pi d_i \end{pmatrix}.$$

Then we can easily see that this matrix satisfies congruence condition (b) in the definition of $T_3(R)$.

If L_i is of type I^e, then we express $m_{i,i}$ and $y_{i,i}$ as matrices:

$$m_{i,i} = \begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & 1+\pi x_i & \pi z_i \\ v_i & u_i & 1+\pi w_i \end{pmatrix} \quad \text{and} \quad y_{i,i} = \begin{pmatrix} a_i & b_i & \pi c_i \\ d_i & e_i & \pi f_i \\ \pi g_i & \pi h_i & \pi k_i \end{pmatrix}.$$

Then

$$\sigma({}^t m_{i,i}) \cdot y_{i,i} = \begin{pmatrix} \sigma({}^t s_i) & \sigma(\pi \cdot {}^t y_i) & \sigma({}^t v_i) \\ \sigma({}^t r_i) & 1+\sigma(\pi x_i) & \sigma(u_i) \\ \sigma(\pi \cdot {}^t t_i) & \sigma(\pi z_i) & 1+\sigma(\pi w_i) \end{pmatrix} \cdot \begin{pmatrix} a_i & b_i & \pi c_i \\ d_i & e_i & \pi f_i \\ \pi g_i & \pi h_i & \pi k_i \end{pmatrix}.$$

Then we can easily see that this matrix satisfies congruence condition (c) in the definition of $T_3(R)$. □

Let $\underline{G}$ be the stabilizer of h in $\underline{M}^*$. It is an affine group subscheme of $\underline{M}^*$, defined over A. Thus we have the following theorem.

Theorem 3.8. *The group scheme $\underline{G}$ is smooth, and $\underline{G}(R) = \mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$ for any étale A-algebra R.*

Proof. Since $\underline{G}$ is the fiber of h along the smooth morphism $\rho : \underline{M}^* \to \underline{H}$, $\rho(m) = h \circ m$, the scheme $\underline{G}$ is smooth. Here, we use the fact that smoothness is stable under base change.

For the identity, we recall that each element of $\mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$, for an étale A-algebra R, satisfies all congruence conditions defining $\underline{M}$, which is explained in Section 3A. Since $\underline{G}(R)$ is the group of R-points of $\underline{M}^*$ stabilizing the given hermitian form h, we have the identity $\underline{G}(R) = \mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$ for any étale A-algebra R. □

Note that in the theorem, the equality holds only for an étale A-algebra R since we obtain conditions defining $\underline{M}$ by observing properties of elements of $\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$ for an étale A-algebra R (cf. Section 3A). For example, let (L, h) be the hermitian lattice of rank 1 as given in Appendix B. For simplicity, let $\pi+\sigma(\pi)=\pi^2=2$. As a set, $\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$ is the same as $\{(a,b): a,b\in R \text{ and } a^2+2ab+2b^2=1\}$ for a flat A-algebra R. Thus we cannot guarantee that $a-1$ is contained in the ideal (2), which should be necessary in order that (a,b) is an element of $\underline{G}(R)$.

4. The special fiber of the smooth integral model

In this section, we will determine the structure of the special fiber $\widetilde{G}$ of $\underline{G}$ by determining the maximal reductive quotient and the component group when E/F satisfies *Case 1*, by adapting the approach of Section 4 of [Cho 2015a]. From this section to the end, the identity matrix is denoted by id.

4A. The reductive quotient of the special fiber. Recall that Y_i is the sublattice of B_i such that $Y_i/\pi A_i$ is the radical of the alternating bilinear form $\xi^{-i/2}h \bmod \pi$ on $B_i/\pi A_i$ (when i is even) and that Z_i is the sublattice of A_i such that $Z_i/\pi A_i$ is the radical of the quadratic form $\frac{1}{2^m}q \bmod 2$ on $A_i/\pi A_i$, where $\frac{1}{2^m}q(x)=\frac{1}{2^m}h(x,x)$ (when $i=2m-1$ is odd).

Lemma 4.1. *Let i be odd. Consider the lattice $\pi A_{i-1}+A_{i+1}=\{x+y: x\in\pi A_{i-1},\ y\in A_{i+1}\}$. Then $\pi A_{i-1}+A_{i+1}=X_i$.*

Proof. Let $L=\bigoplus_i L_i$ be a Jordan splitting. We describe πA_{i-1}, A_{i+1}, X_i below:

$$\begin{aligned}\pi A_{i-1}&=\pi^i L_0\oplus\pi^{i-1}L_1\oplus\cdots\oplus\pi L_{i-1}\oplus\pi L_i\oplus\pi L_{i+1}\oplus\cdots;\\ A_{i+1}&=\pi^{i+1}L_0\oplus\pi^i L_1\oplus\cdots\oplus\pi^2 L_{i-1}\oplus\pi L_i\oplus L_{i+1}\oplus\cdots;\\ X_i&=\pi^i L_0\oplus\pi^{i-1}L_1\oplus\cdots\oplus\pi L_{i-1}\oplus\pi L_i\oplus L_{i+1}\oplus\cdots.\end{aligned}$$

Our claim follows directly from the above descriptions. □

Lemma 4.2. *Each element of $\underline{M}(R)$, for a flat A-algebra R, preserves $Y_i\otimes_A R$ (for i even) and $Z_i\otimes_A R$ (for i odd).*

Proof. The claim for Y_i follows from the fact that $Y_i=X_i$ or $Y_i=W_i$ according to the type of L_i as described in Remark 2.11.

To prove the claim for Z_i, use Lemma 4.1 to express a given arbitrary element of $Z_i\otimes_A R$ as $x+y$, where $x\in\pi A_{i-1}\otimes_A R$ and $y\in A_{i+1}\otimes_A R$. Let $g\in\underline{M}(R)$. Then $g(x+y)=g(x)+g(y)=(x+x')+(y+y')$, where $x'\in\pi B_{i-1}\otimes_A R$, $y'\in B_{i+1}\otimes_A R$ since g induces the identity on $(A_{i-1}\otimes_A R)/(B_{i-1}\otimes_A R)$ and on $(A_{i+1}\otimes_A R)/(B_{i+1}\otimes_A R)$. Since $\pi A_{i-1}\otimes_A R$ and $A_{i+1}\otimes_A R$ are contained in $W_i\otimes_A R$ and hence $\pi B_{i-1}\otimes_A R$ and $B_{i+1}\otimes_A R$ are contained in $Z_i\otimes_A R$, we have that $g(x+y)=(x+y)+x'+y'\in Z_i\otimes_A R$. □

Theorem 4.3. *Assume that i is even. Let h_i denote the nonsingular alternating bilinear form $\xi^{-i/2}h \bmod \pi$ on B_i/Y_i. Then there exists a unique morphism of algebraic groups*

$$\varphi_i:\widetilde{G}\longrightarrow \mathrm{Sp}(B_i/Y_i, h_i)$$

defined over κ such that for all étale local A-algebras R with residue field κ_R and every

$\tilde{m} \in \underline{G}(R)$ with reduction $m \in \widetilde{G}(\kappa_R)$, $\varphi_i(m) \in \mathrm{GL}(B_i \otimes_A R/Y_i \otimes_A R)$ is induced by the action of $\tilde{m}$ on $L \otimes_A R$ (which preserves $B_i \otimes_A R$ and $Y_i \otimes_A R$ by Lemma 4.2). Note that the dimension of B_i/Y_i, as a κ-vector space, is as follows:

$$\begin{cases} n_i & \text{if } L_i \text{ is of type } II; \\ n_i - 1 & \text{if } L_i \text{ is of type } I^o; \\ n_i - 2 & \text{if } L_i \text{ is of type } I^e. \end{cases}$$

Proof. Let R be an étale local A-algebra with κ_R as its residue field. Note that such an R is finite over A since any étale local algebra R over a henselian local ring is finite by Proposition 4 of Section 2.3 in [Bosch et al. 1990] and since A is henselian. For such a finite field extension κ_R of κ, R is uniquely determined up to isomorphism. Since $\underline{G}$ is smooth over A, the map $\underline{G}(R) \to \widetilde{G}(\kappa_R)$ is surjective by Hensel's lemma.

Now, we choose an element $m \in \widetilde{G}(\kappa_R)$ and a lift $\tilde{m} \in \underline{G}(R)$. Since the action of $\tilde{m}$ on $L \otimes_A R$ preserves $B_i \otimes_A R$ and $Y_i \otimes_A R$, $\tilde{m}$ determines an element of $\mathrm{GL}(B_i \otimes_A R/Y_i \otimes_A R)$. It is also easy to show that this element determined by $\tilde{m}$ fixes $h_i \otimes \kappa_R$ on $B_i/Y_i \otimes_\kappa \kappa_R$ ($= B_i \otimes_A R/Y_i \otimes_A R$). Thus $\tilde{m}$ determines an element of $\mathrm{Sp}(B_i/Y_i, h_i)(\kappa_R)$ and so we have a map from $\widetilde{G}(\kappa_R)$ to $\mathrm{Sp}(B_i/Y_i, h_i)(\kappa_R)$. Indeed, this map is well-defined, i.e., independent of a lift $\tilde{m}$ of m as will be explained later after describing a matrix interpretation of this map. In order to show that this map is well-defined and representable, we interpret it in terms of matrices. Recall that $\underline{G}$ is a closed subgroup scheme of $\underline{M}^*$ and $\widetilde{G}$ is a closed subgroup scheme of $\widetilde{M}$, where $\widetilde{M}$ is the special fiber of $\underline{M}^*$. Thus we may consider an element of $\widetilde{G}(\kappa_R)$ as an element of $\widetilde{M}(\kappa_R)$. Based on Section 3B, an element m of $\widetilde{G}(\kappa_R)$ may be written as, say, $(m_{i,j}, s_i \cdots w_i)$ and it has the following formal matrix description:

$$m = \left(\pi^{\max\{0, j-i\}} m_{i,j}\right).$$

Here, if i is even and L_i is of type I^o or of type I^e, then

$$m_{i,i} = \begin{pmatrix} s_i & \pi y_i \\ \pi v_i & 1 + \pi z_i \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & 1 + \pi x_i & \pi z_i \\ v_i & u_i & 1 + \pi w_i \end{pmatrix},$$

respectively, where $s_i \in M_{(n_i-1)\times(n_i-1)}(B \otimes_A \kappa_R)$ (resp. $s_i \in M_{(n_i-2)\times(n_i-2)}(B \otimes_A \kappa_R)$), etc., and s_i is invertible. For the remaining $m_{i,j}$'s except for the cases explained above, $m_{i,j} \in M_{n_i \times n_j}(B \otimes_A \kappa_R)$ and $m_{i,i}$ is invertible. Note that the description of the multiplication in $\widetilde{M}(\kappa_R)$ given in Section 3B forces s_i and $m_{i,i}$ to be invertible.

We can write $m_{i,i} = m_{i,i}^1 + \pi \cdot m_{i,i}^2$ when L_i is of type II and for each block of $m_{i,i}$ when L_i is of type I, $s_i = s_i^1 + \pi \cdot s_i^2$ and so on. Here, $m_{i,i}^1, m_{i,i}^2 \in M_{n_i \times n_i}(\kappa_R) \subset M_{n_i \times n_i}(B \otimes_A \kappa_R)$ when L_i is of type II and so on, and π stands for $\pi \otimes 1 \in B \otimes_A \kappa_R$. Then m maps to $m_{i,i}^1$ if L_i is of type II and s_i^1 if L_i is of type I. Since this map is independent of the choice of $m_{i,i}^2, s_i^2$ and so on, it is independent of the choice of $\tilde{m}$, i.e., this map is well-defined.

We note that this map is given by polynomials over A of degree at most 1 as well as a group homomorphism. Thus the above matrix interpretation induces a Hopf algebra homomorphism over A from the coordinate ring of $\mathrm{Sp}(B_i/Y_i, h_i)$ to the coordinate ring of $\widetilde{G}$, which accordingly induces an algebraic group homomorphism $\varphi_i : \widetilde{G} \to \mathrm{Sp}(B_i/Y_i, h_i)$

such that the group homomorphism induced by φ_i at the level of κ_R-points is the same as the map explained above.

Since $\widetilde{G}$ is smooth over κ and κ is perfect, the set of κ_R-points of $\widetilde{G}$ for all finite field extensions κ_R/κ is dense in $\widetilde{G}$ by [Bosch et al. 1990, Corollary 13 of Section 2.2]. Therefore, φ_i is uniquely determined by the map constructed above at the level of κ_R-points. $\square$

Theorem 4.4. *We next assume that $i = 2m-1$ is odd. Let $\bar{q}_i$ denote the nonsingular quadratic form $\frac{1}{2^m}q \bmod 2$ on A_i/Z_i. Then there exists a unique morphism of algebraic groups*

$$\varphi_i : \widetilde{G} \longrightarrow O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$$

defined over κ, where $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ is the reduced subgroup scheme of $O(A_i/Z_i, \bar{q}_i)$, such that for all étale local A-algebras R with residue field κ_R and every $\tilde{m} \in \underline{G}(R)$ with reduction $m \in \widetilde{G}(\kappa_R)$, $\varphi_i(m) \in \mathrm{GL}(A_i \otimes_A R/Z_i \otimes_A R)$ is induced by the action of $\tilde{m}$ on $L \otimes_A R$ (which preserves $A_i \otimes_A R$ and $Z_i \otimes_A R$ by Lemma 4.2).

Proof. The proof of this theorem is similar to that of Theorem 4.3 which deals with the case of even i. Thus we only provide the image of an element m of $\widetilde{G}(\kappa_R)$ in $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}(\kappa_R)$, where R is an étale local A-algebra with κ_R as its residue field. In this case, an element m of $\widetilde{G}(\kappa_R)$ maps to $m_{i,i}^1$ (if L_i is *free*) or to $\begin{pmatrix} m_{i,i}^1 & 0 \\ \delta_{i-1}e_{i-1}\cdot m_{i-1,i}^1+\delta_{i+1}e_{i+1}\cdot m_{i+1,i}^1 & 1 \end{pmatrix}$ (if L_i is *bound*). Here, $\delta_j = 1$ if L_j is of type I and $\delta_j = 0$ if L_j is of type II. Also, $e_j = (0, \cdots, 0, 1)$ (resp. $e_j = (0, \cdots, 0, 1, 0)$) of size $1 \times n_j$ if L_j is of type I^o (resp. of type I^e). $\square$

Notice that if the dimension of A_i/Z_i is even and positive, then $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ ($= O(A_i/Z_i, \bar{q}_i)$) is disconnected. If the dimension of A_i/Z_i is odd, then $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ ($= \mathrm{SO}(A_i/Z_i, \bar{q}_i)$) is connected. The dimension of A_i/Z_i, as a κ-vector space, is as follows:

$$\begin{cases} n_i & \text{if } L_i \text{ is free;} \\ n_i+1 & \text{if } L_i \text{ is bound.} \end{cases}$$

Note that the integer n_i, with i odd, is always even.

Theorem 4.5. *The morphism φ defined by*

$$\varphi = \prod_i \varphi_i : \widetilde{G} \longrightarrow \prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$$

is surjective.

Proof. Let us first prove the theorem under the assumption that

$$\dim \widetilde{G} = \dim \operatorname{Ker}\varphi + \sum_{i \text{ even}} (\dim \mathrm{Sp}(B_i/Y_i, h_i)) + \sum_{i \text{ odd}} (\dim O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}). \tag{4-1}$$

This equation will be proved in Appendix A. Thus $\operatorname{Im}\varphi$ contains the identity component of $\prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$. Here $\operatorname{Ker}\varphi$ denotes the kernel of φ and $\operatorname{Im}\varphi$ denotes the image of φ. Note that it is well known that the image of a homomorphism of algebraic groups is a closed subgroup.

Recall from Section 3B that a matrix form of an element of $\widetilde{G}(R)$ for a κ-algebra R is written $(m_{i,j}, s_i \cdots w_i)$ with the formal matrix interpretation

$$m = \left(\pi^{\max\{0, j-i\}} m_{i,j}\right).$$

We represent the given hermitian form h by a hermitian matrix $\left(\pi^i \cdot h_i\right)$ with $\pi^i \cdot h_i$ for the (i,i)-block and 0 for the remaining blocks, as in Remark 3.3(1).

Let $\mathcal{H}$ be the set of odd integers i such that $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ is disconnected. Notice that $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ is disconnected exactly when L_i with i odd is *free*. We first prove that φ_i, for such an odd integer i, is surjective. We prove this by a series of reductions, after which we will be able to assume that L is of rank two.

For such an odd integer i with a free lattice L_i, we define the closed scheme H_i of $\widetilde{G}$ by the equations $m_{j,k}=0$ if $j \neq k$, and $m_{j,j}=\mathrm{id}$ if $j \neq i$. An element of $H_i(R)$ for a κ-algebra R can be represented by a matrix of the form

$$\begin{pmatrix} \mathrm{id} & 0 & & \dots & & & 0 \\ 0 & \ddots & & & & & \\ & & \mathrm{id} & & & & \\ \vdots & & & m_{i,i} & & & \vdots \\ & & & & \mathrm{id} & & \\ & & & & & \ddots & 0 \\ 0 & & & \dots & & 0 & \mathrm{id} \end{pmatrix}.$$

Obviously, H_i has a group scheme structure. We claim that φ_i is surjective from H_i to $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ (recall that $Z_i = X_i$ since L_i is free). Note that equations defining H_i are induced by the formal matrix equation

$$\sigma({}^t m_{i,i})(\pi^i \cdot h_i) m_{i,i} = \pi^i \cdot h_i$$

which is interpreted as in Remark 3.5. We emphasize that, in this formal matrix equation, we work with $m_{i,i}$, not m, because of the description of H_i. Note that none of the congruence conditions mentioned in Section 3A involve any entry from $m_{i,i}$.

On the other hand, let us consider the hermitian lattice L_i independently as a π^i-modular lattice. Since there is only one nontrivial Jordan component for this lattice and i is odd, the smooth integral model associated to L_i is determined by the following formal matrix equation which is interpreted as in Remark 3.5:

$$\sigma({}^t m)(\pi^i \cdot h_i) m = \pi^i \cdot h_i,$$

where m is an $(n_i \times n_i)$-matrix and is not subject to any congruence condition.

We consider the map from H_i to the special fiber of the smooth integral model associated to the hermitian lattice L_i such that $m_{i,i}$ maps to m. Since $m_{i,i}$ and m are subject to the same set of equations, this map is an isomorphism as algebraic groups. In addition, this map induces compatibility between the morphism φ_i from H_i to $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ and the morphism from the special fiber of the smooth integral model associated to L_i to $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$. Thus, in order to show that φ_i is surjective from H_i to $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$, we may and do assume that $L = L_i$ and in this case $Z_i = X_i = \pi L_i$. For simplicity, we can also assume that $i = 1$.

Because of Equation (4-1) stated at the beginning of the proof, the dimension of the image of φ_i, as a κ-algebraic group, is the same as that of $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ $(= O(L_i/\pi L_i, \bar{q}_i))$. Therefore, the image of φ_i contains the identity component of $O(L_i/\pi L_i, \bar{q}_i)$, namely $\mathrm{SO}(L_i/\pi L_i, \bar{q}_i)$. Since $O(L_i/\pi L_i, \bar{q}_i)$ has two connected components, we only need to show the surjectivity of φ_i at the level of κ-points and it suffices to show that the image of $\varphi_i(\kappa)$ contains at least one element which is not contained in $\mathrm{SO}(L_i/\pi L_i, \bar{q}_i)(\kappa)$, where $\mathrm{SO}(L_i/\pi L_i, \bar{q}_i)(\kappa)$ is the group of κ-points of the algebraic group $\mathrm{SO}(L_i/\pi L_i, \bar{q}_i)$.

Recall that $L_i = \bigoplus_\lambda H_\lambda \oplus A(2, 2b, \pi)$ for a certain $b \in A$, cf. Theorem 2.10. We consider the orthogonal group associated to the quadratic κ-space $A(2, 2b, \pi)/\pi A(2, 2b, \pi)$ of dimension 2. Then this group is embedded into $O(L_i/\pi L_i, \bar{q}_i)(\kappa)$ as a closed subgroup and we denote the embedded group by $O(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$.

We express an element $m_{i,i} \in H_i(R)$, for a κ-algebra R, as $\left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right)$ such that $x = x^1 + \pi x^2$ and so on, where $x^1, x^2 \in M_{(n_i-2)\times(n_i-2)}(R) \subset M_{(n_i-2)\times(n_i-2)}(R \otimes_A B)$ and π stands for $1 \otimes \pi \in R \otimes_A B$. Consider the closed subscheme of H_i defined by the equations $x = \mathrm{id}$, $y = 0$, and $z = 0$. An argument similar to one used above to reduce to the case where $L = L_i$ shows that this subscheme is isomorphic to the special fiber of the smooth integral model associated to the hermitian lattice $A(2, 2b, \pi)$ of rank 2. Then under the map $\varphi_i(\kappa)$, an element of this subgroup maps to an element of $O(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$ of the form $\left(\begin{smallmatrix} \mathrm{id} & 0 \\ 0 & w^1 \end{smallmatrix}\right)$. Note that $O(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$ is not contained in $\mathrm{SO}(L_i/\pi L_i, \bar{q}_i)(\kappa)$. Thus it suffices to show that the restriction of $\varphi_i(\kappa)$ to the above subgroup of $H_i(\kappa)$, which is given by letting $x = \mathrm{id}$, $y = 0$, $z = 0$, is surjective onto $O(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$ and we may and do assume that $L = L_i = A(2, 2b, \pi)$ is of rank 2.

Let $m_{i,i} = \left(\begin{smallmatrix} r & s \\ t & v \end{smallmatrix}\right)$ be an element of $H_i(\kappa)$ such that $r = r_1 + \pi r_2$ and so on, where $r_1, r_2 \in R \subset R \otimes_A B$ and π stands for $1 \otimes \pi \in R \otimes_A B$. Recall that $\pi = 1 + \sqrt{1 + 2u}$ for a certain unit $u \in A$ so that $\pi + \sigma(\pi) = 2$, $\sigma(\pi) = \epsilon\pi$ with $\epsilon \equiv 1 \bmod \pi$, and $\pi^2 \equiv (\sigma(\pi))^2 \equiv \xi^- \equiv 2u \bmod 2\pi$ as mentioned in Section 2A. Let $\bar{u} \in \kappa$ be the reduction of u modulo π. Then the equations defining $H_i(\kappa)$ are

$$r_1^2 + r_1 t_1 + b t_1^2 = 1, \quad r_1 v_1 + t_1 s_1 = 1,$$
$$r_1 s_1 + b t_1 v_1 + \bar{u}(r_2 v_1 + r_1 v_2 + t_2 s_1 + t_1 s_2) = 0, \quad s_1^2 + s_1 v_1 + b v_1^2 = b.$$

Under the map $\varphi_i(\kappa)$, $m_{i,i}$ maps to $\left(\begin{smallmatrix} r_1 & s_1 \\ t_1 & v_1 \end{smallmatrix}\right)$. Note that the quadratic form $\bar{q}_i$ restricted to $A(2, 2b, \pi)/\pi A(2, 2b, \pi)$ is given by the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & b \end{smallmatrix}\right)$.

We now choose an element of $H_i(\kappa)$ by setting

$$r_1 = s_1 = v_1 = 1, \quad t_1 = 0, \quad 1 + \bar{u}(r_2 + v_2 + t_2) = 0.$$

Under the morphism $\varphi_i(\kappa)$, this element maps to $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in O(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$. The Dickson invariant of this element is nontrivial so that it is not contained in $\mathrm{SO}(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$.

Therefore, $\varphi_i(\kappa)$ induces a surjection from $H_i(\kappa)$ to $O(A(2, 2b, \pi)/\pi A(2, 2b, \pi), \bar{q}_i)(\kappa)$ for $i \in \mathcal{H}$.

We now prove that $\varphi = \prod_i \varphi_i$ is surjective. We consider the morphism

$$\prod_{i \in \mathcal{H}} H_i \to \widetilde{G}$$
$$(h_i)_{i \in \mathcal{H}} \mapsto \prod_{i \in \mathcal{H}} h_i.$$

By considering a formal matrix form of an element of $H_i(R)$ for a κ-algebra R as given above, it is easy to see the following two facts. Firstly, H_i and H_j commute with each other in the sense that $h_i \cdot h_j = h_j \cdot h_i$ for all $i \neq j$, where $h_i \in H_i(R)$ and $h_j \in H_j(R)$ for a κ-algebra R. Based on this, the above morphism becomes a group homomorphism. Secondly, $H_i \cap H_j = 0$ for all $i \neq j$. This fact implies that the morphism $H_i \times H_j \longrightarrow \widetilde{G}$, $(h_i, h_j) \mapsto h_i \cdot h_j$ is injective and so $H_i \times H_j$ is a closed subgroup scheme of $\widetilde{G}$. A matrix form of an element of $H_i(R)$ also implies that $(H_i \times H_j) \cap H_k = 0$ for all pairwise different three integers i, j, k and so the morphism $(H_i \times H_j) \times H_k \longrightarrow \widetilde{G}$, $(h_i, h_j, h_k) \mapsto h_i \cdot h_j \cdot h_k$ is injective. Thus $H_i \times H_j \times H_k$ is a closed subgroup scheme of $\widetilde{G}$. Therefore, by repeating this argument, the product $\prod_{i \in \mathcal{H}} H_i$ is embedded into $\widetilde{G}$ as a closed subgroup scheme. Since $\varphi_i|_{H_j}$ is trivial for $i \neq j$ with $i, j \in \mathcal{H}$, the morphism

$$\prod_{i \in \mathcal{H}} \varphi_i : \prod_{i \in \mathcal{H}} H_i \to \prod_{i \in \mathcal{H}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$$

is surjective. Therefore, φ is surjective. Now it suffices to prove Equation (4-1) made at the beginning of the proof, which is the next lemma. □

Lemma 4.6. Ker φ *is smooth and unipotent of dimension* l. *In addition, the number of connected components of* Ker φ *is* 2^{β}. *Here,*

- l *is such that*

$$l + \sum_{i \text{ even}} (\dim \mathrm{Sp}(B_i/Y_i, h_i)) + \sum_{i \text{ odd}} (\dim O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}) = \dim \widetilde{G}.$$

- β *is the number of even integers* j *such that* L_j *is of type* I *and* L_{j+2} *is of type* II.

Recall that the zero lattice is of type II. The proof is postponed to Appendix A.

Remark 4.7. We summarize the description of Im φ_i as follows.

type of lattice L_i	i	Im φ_i
II	even	$\mathrm{Sp}(n_i, h_i)$
I^o	even	$\mathrm{Sp}(n_i - 1, h_i)$
I^e	even	$\mathrm{Sp}(n_i - 2, h_i)$
free	odd	$O(n_i, \bar{q}_i)$
bound	odd	$\mathrm{SO}(n_i + 1, \bar{q}_i)$

Let i be odd and L_i be *free*. Then $A_i/Z_i = L_i/\pi L_i$ is a κ-vector space with even dimension. We now consider the question of whether the orthogonal group $O(A_i/Z_i, \bar{q}_i) = O(n_i, \bar{q}_i)$ is split or nonsplit.

By Theorem 2.10, we have that $L_i = \bigoplus_{\lambda} H_{\lambda} \oplus A(2, 2b_i, \pi)$ for certain $b_i \in A$. Thus the orthogonal group $O(A_i/Z_i, \bar{q}_i)$ $(= O(n_i, \bar{q}_i))$ is split if and only if the quadratic space $A(2, 2b_i, \pi)/\pi A(2, 2b_i, \pi)$ is isotropic. Recall that $\pi + \sigma(\pi) = 2$ and $\pi = 1 + \sqrt{1 + 2u}$ for a certain unit $u \in A$. Using this, the quadratic form on $A(2, 2b_i, \pi)/\pi A(2, 2b_i, \pi)$ is $q(x, y) = x^2 + xy + \bar{b}_i y^2$, where $\bar{b}_i$ is the reduction of b_i in κ.

We consider the identity $q(x, y) = x^2 + xy + \bar{b}_i y^2 = 0$. If $y = 0$, then $x = 0$. Assume that $y \neq 0$. Then we have that $\bar{b}_i = (x/y)^2 + x/y$.

Thus we can see that there exists a solution of the equation $z^2 + z = \bar{b}_i$ over κ if and only if $q(x, y)$ is isotropic if and only if $O(A_i/Z_i, \bar{q}_i)$ $(= O(n_i, \bar{q}_i))$ is split.

4B. ***The construction of component groups.*** The purpose of this subsection is to define a surjective morphism from $\widetilde{G}$ to $(\mathbb{Z}/2\mathbb{Z})^{\beta}$, where β is the number of even integers j such that L_j is of type I and L_{j+2} is of type II as defined in Lemma 4.6.

Definition 4.8. We set $L^0 = L$ and inductively define, for positive integers i,

$$L^i := \{x \in L^{i-1} \mid h(x, L^{i-1}) \subset (\pi^i)\}.$$

When $i = 2m$ is even,

$$L^{2m} = \pi^m(L_0 \oplus L_1) \oplus \pi^{m-1}(L_2 \oplus L_3) \oplus \cdots \oplus \pi(L_{2m-2} \oplus L_{2m-1}) \oplus \bigoplus_{i \geq 2m} L_i.$$

We choose a Jordan splitting for the hermitian lattice $(L^{2m}, \xi^{-m}h)$ as follows:

$$L^{2m} = \bigoplus_{i \geq 0} M_i,$$

where

$$\begin{aligned} M_0 &= \pi^m L_0 \oplus \pi^{m-1} L_2 \oplus \cdots \oplus \pi L_{2m-2} \oplus L_{2m}, \\ M_1 &= \pi^m L_1 \oplus \pi^{m-1} L_3 \oplus \cdots \oplus \pi L_{2m-1} \oplus L_{2m+1}, \\ M_k &= L_{2m+k} \text{ if } k \geq 2. \end{aligned}$$

Here, M_i is π^i-modular. We caution that the hermitian form we use on L^{2m} is not h, but its rescaled version $\xi^{-m}h$. Thus M_i is π^i-modular, not π^{2m+i}-modular.

Definition 4.9. We define $C(L)$ to be the sublattice of L such that

$$C(L) = \{x \in L \mid h(x, y) \in (\pi) \text{ for all } y \in B(L)\}.$$

We choose any even integer j such that L_j is of type I and L_{j+2} is of type II (possibly zero, by our convention), and consider the Jordan splitting $\bigoplus_{i \geq 0} M_i$ of L^j defined above. We stress that M_0 is nonzero and of type I, since it contains L_j as a direct summand so that $n(M_0) = s(M_0)$ (cf. Definition 2.1(c)), and $M_2 = L_{j+2}$ is of type II. Choose a basis $(\langle e_i \rangle, e)$ (resp. $(\langle e_i \rangle, a, e)$) for M_0 so that $M_0 = \bigoplus_{\lambda} H_{\lambda} \oplus K$ when the rank of M_0 is odd (resp. even). Here, we follow the notation from Theorem 2.10. Then $B(L^j)$ is spanned by

$$(\langle e_i \rangle, \pi e) \text{ (resp. } (\langle e_i \rangle, \pi a, e)) \quad \text{and} \quad M_1 \oplus \bigoplus_{i \geq 2} M_i$$

and $C(L^j)$ is spanned by

$$(\langle \pi e_i \rangle, e) \text{ (resp. } (\langle \pi e_i \rangle, \pi a, e)) \quad \text{and} \quad M_1 \oplus \bigoplus_{i \geq 2} M_i.$$

We now construct a morphism $\psi_j : \widetilde{G} \to \mathbb{Z}/2\mathbb{Z}$ as follows. (There are 2 cases depending on whether M_0 is of type I^e or of type I^o.)

(1) Firstly, we assume that M_0 is of type I^e. We choose a Jordan splitting for the hermitian lattice $(C(L^j), \xi^{-m}h)$ as follows:

$$C(L^j) = \bigoplus_{i \geq 1} M'_i,$$

where

$$M_1' = (\pi)a \oplus Be \oplus M_1, \quad M_2' = \Big(\bigoplus_i (\pi)e_i\Big) \oplus M_2, \quad \text{and} \quad M_k' = M_k \text{ if } k \geq 3.$$

Here, M_i' is π^i-modular and (π) is the ideal of B generated by a uniformizer π. Notice that M_2' is of type II, since both $\bigoplus_i (\pi)e_i$ and M_2 are of type II, so that M_1' is *free*.

If m is an element of the group of R-points of the naive integral model associated to the hermitian lattice L, for a flat A-algebra R, then m stabilizes the hermitian lattice $(C(L^j) \otimes_A R, \xi^{-m} h \otimes 1)$ as well. If we use this fact in the case of an F-algebra R, where F is the quotient field of A, then we obtain a morphism of algebraic groups from the unitary group associated to the hermitian space $L \otimes_A F$ to the unitary group associated to the hermitian space $(C(L^j) \otimes_A F, \xi^{-m} h)$ by Yoneda's lemma. Furthermore, if we use the above fact in the case of an étale A-algebra R, then the morphism between unitary groups is extended to give a map from the group of R-points of the naive integral model associated to the hermitian lattice L to that of the hermitian lattice $(C(L^j), \xi^{-m} h)$. Note that the naive integral model and the associated smooth integral model have the same generic fiber and are the same at the level of étale A-points. Thus by Proposition 2.3 of [Yu 2002], the morphism between unitary groups is uniquely extended to a morphism of group schemes from the smooth integral model associated to L to the smooth integral model associated to $(C(L^j), \xi^{-m} h)$ such that the map induced from it at the level of étale A-points is the same as that described above. Let G_j denote the special fiber of the latter smooth integral model. We now have a morphism from $\widetilde{G}$ to G_j. Moreover, since M_1' is *free* and nonzero, we have a morphism from G_j to the even orthogonal group associated to M_1' as explained in Section 4A. Thus, the Dickson invariant of this orthogonal group induces the morphism

$$\psi_j : \widetilde{G} \longrightarrow \mathbb{Z}/2\mathbb{Z}.$$

(2) We next assume that M_0 is of type I^o. We choose a Jordan splitting for the hermitian lattice $(C(L^j), \xi^{-m} h)$ as follows:

$$C(L^j) = \bigoplus_{i \geq 0} M_i',$$

where

$$M_0' = Be, \quad M_1' = M_1, \quad M_2' = \Big(\bigoplus_i (\pi)e_i\Big) \oplus M_2, \quad \text{and} \quad M_k' = M_k \text{ if } k \geq 3.$$

Here, M_i' is π^i-modular and (π) is the ideal of B generated by a uniformizer π. Notice that the rank of the π^0-modular lattice M_0' is 1 and the lattice M_2' is of type II. If G_j denotes the special fiber of the smooth integral model associated to the hermitian lattice $(C(L^j), \xi^{-m} h)$, then we have a morphism from $\widetilde{G}$ to G_j as in the above argument (1).

We now consider the new hermitian lattice $M_0' \oplus C(L^j)$. Then for a flat A-algebra R, there is a natural embedding from the group of R-points of the naive integral model associated to the hermitian lattice $(C(L^j), \xi^{-m} h)$ to that of the hermitian lattice $M_0' \oplus C(L^j)$ such that m maps to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & m \end{smallmatrix}\right)$, where m is an element of the former group. As in the previous argument (1), the above fact induces a closed immersion of algebraic groups from the unitary group associated to the hermitian space $(C(L^j)\otimes_A F, \xi^{-m}h)$ to the unitary group associated to the hermitian space $(M_0'\oplus C(L^j))\otimes_A F$ and its extension at the level of étale A-algebra points between the associated naive integral models. Thus by Proposition 2.3 of [Yu 2002], the morphism between unitary groups is uniquely extended to a morphism of group schemes from the smooth integral model associated to the hermitian lattice $(C(L^j), \xi^{-m}h)$ to the smooth integral model associated to the hermitian lattice $M_0'\oplus C(L^j)$ such that the map induced from it at the level of étale A-points is the same as that described above. In Remark 4.10, we describe this morphism explicitly in terms of matrices.

Thus we have a morphism from the special fiber G_j of the smooth integral model associated to $C(L^j)$ to the special fiber G_j' of the smooth integral model associated to $M_0'\oplus C(L^j)$. Note that $(M_0'\oplus M_0')\oplus\bigoplus_{i\geq 1} M_i'$ is a Jordan splitting of the hermitian lattice $M_0'\oplus C(L^j)$. Let G_j'' be the special fiber of the smooth integral model associated to $C\big((M_0'\oplus M_0')\oplus\bigoplus_{i\geq 1} M_i'\big)$. Since the π^0-modular lattice $M_0'\oplus M_0'$ is of type I^e, we have a morphism $G_j'\to\mathbb{Z}/2\mathbb{Z}$ obtained by factoring through G_j'' and the corresponding even orthogonal group with the Dickson invariant as constructed in argument (1). ψ_j is defined to be the composite

$$\psi_j:\widetilde{G}\to G_j\to G_j'\to\mathbb{Z}/2\mathbb{Z}.$$

Remark 4.10. In this remark, we describe the morphism from the smooth integral model $\underline{G}_j$ associated to the hermitian lattice $(C(L^j), \xi^{-m}h)$ to the smooth integral model $\underline{G}_j'$ associated to the hermitian lattice $M_0'\oplus C(L^j)$ as given in argument (2) above, in terms of matrices. Let R be a flat A-algebra. We choose an element in $\underline{G}_j(R)$ and express it as a matrix $m=\big(\pi^{\max\{0,j-i\}}m_{i,j}\big)$. Then $m_{0,0}=\big(1+\pi z_0\big)$ since M_0' is of type I with rank 1 so that we may and do write m as $m=\begin{pmatrix}1+\pi z_0 & m_1\\ m_2 & m_3\end{pmatrix}$. We consider a morphism from $\underline{G}_j$ to $\mathrm{Aut}_B(M_0'\oplus C(L^j))$ such that m maps to

$$T=\begin{pmatrix}1&0\\0&m\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1+\pi z_0&m_1\\0&m_2&m_3\end{pmatrix},$$

where the set of R-points of the group scheme $\mathrm{Aut}_B(M_0'\oplus C(L^j))$ is the automorphism group of $(M_0'\oplus C(L^j))\otimes_A R$ by ignoring the hermitian form. Then the image of this morphism is represented by an affine group scheme which is isomorphic to $\underline{G}_j$. Note that T preserves the hermitian form attached to the lattice $M_0'\oplus C(L^j)$.

We claim that $\begin{pmatrix}1&0\\0&m\end{pmatrix}$ is contained in $\underline{G}_j'(R)$. If this is true, then the above matrix description defines a morphism from $\underline{G}_j$ to $\underline{G}_j'$ by Yoneda's lemma since $\underline{G}_j$ is flat. Furthermore, this matrix description is the same as that of naive integral models explained in the above argument (2) when R is an F-algebra or an étale A-algebra, since the naive integral model and the associated smooth integral model have the same generic fiber and are the same at the level of étale A-points. Since the desired morphism is completely determined at the level of F-algebra points and étale A-algebra points by Proposition 2.3 of [Yu 2002], the morphism from $\underline{G}_j$ to $\underline{G}_j'$ obtained by the above matrix description is the morphism we want to describe.

We rewrite the hermitian lattice $M_0'\oplus C(L^j)$ as $(M_0'\oplus M_0')\oplus\left(\bigoplus_{i\geq 1}M_i'\right)$. Let (e_1,e_2) be a basis for $(M_0'\oplus M_0')$ so that the corresponding Gram matrix of $(M_0'\oplus M_0')$ is $\left(\begin{smallmatrix}a&0\\0&a\end{smallmatrix}\right)$, where $a\equiv 1 \bmod 2$. Then the hermitian lattice $(M_0'\oplus M_0')$ has Gram matrix $\left(\begin{smallmatrix}a&a\\a&2a\end{smallmatrix}\right)$ with respect to the basis (e_1,e_1+e_2). $(M_0'\oplus M_0')$ is *unimodular of type* I^e with rank 2. With this basis, T becomes

$$\widetilde{T}=\begin{pmatrix}1&-\pi z_0&-m_1\\0&1+\pi z_0&m_1\\0&m_2&m_3\end{pmatrix}.$$

On the other hand, an element of $\underline{G}'_j(R)$, with respect to a basis for $M_0'\oplus C(L^j)$ obtained by putting together the basis (e_1,e_1+e_2) for $(M_0'\oplus M_0')$ and a basis for $C(L^j)$, is given by an expression

$$\begin{pmatrix}1+\pi x_0'&-\pi z_0'&m_1'\\u_0'&1+\pi w_0'&m_1''\\m_2'&m_2''&m_3''\end{pmatrix},$$

cf. Section 3A. Then we can easily see that the congruence conditions on m_1,m_2,m_3 are the same as those of m_1'',m_2'',m_3'', respectively, and that the congruence conditions on m_1' are the same as those of m_1''. Thus $\widetilde{T}$ is an element of $\underline{M}^*_j(R)$, where $\underline{M}^*_j$ is the group scheme in Section 3B associated to $M_0'\oplus C(L^j)$ so that $\underline{G}'_j$ is defined as the closed subgroup scheme of $\underline{M}^*_j$ stabilizing the hermitian form on $M_0'\oplus C(L^j)$.

In conclusion, $\widetilde{T}$ preserves the hermitian form on $M_0'\oplus C(L^j)$. Therefore, it is an element of $\underline{G}'_j(R)$.

To summarize, if R is a nonflat A-algebra, then we can write an element of $\underline{G}_j(R)$ formally as $m=\left(\begin{smallmatrix}1+\pi z_0&m_1\\m_2&m_3\end{smallmatrix}\right)$. Then the image of m in $\underline{G}'_j(R)$ is $\widetilde{T}$ with respect to a basis as explained above.

(3) Combining all cases, the morphism

$$\psi=\prod_j\psi_j:\widetilde{G}\longrightarrow(\mathbb{Z}/2\mathbb{Z})^\beta,$$

where β is the number of even integers j such that L_j is of type I and L_{j+2} is of type II (possibly zero, by our convention).

Theorem 4.11. *The morphism*

$$\psi=\prod_j\psi_j:\widetilde{G}\longrightarrow(\mathbb{Z}/2\mathbb{Z})^\beta$$

is surjective. Moreover, the morphism

$$\varphi\times\psi:\widetilde{G}\longrightarrow\prod_{i\text{ even}}\mathrm{Sp}(B_i/Y_i,h_i)\times\prod_{i\text{ odd}}O(A_i/Z_i,\bar{q}_i)_{\mathrm{red}}\times(\mathbb{Z}/2\mathbb{Z})^\beta$$

is also surjective.

Proof. We first show that ψ_j is surjective. Recall that for such an even integer j, L_j is of type I and L_{j+2} is of type II (possibly zero by our convention). We define the closed subgroup scheme F_j of $\widetilde{G}$ defined by the following equations:

- $m_{i,k}=0$ *if* $i\neq k$;

- $m_{i,i} = \mathrm{id}$ *if* $i \neq j$;
- and for $m_{j,j}$,

$$\begin{cases} s_j = \mathrm{id},\ y_j = 0,\ v_j = 0 & \text{if } L_i \text{ is of type } I^o; \\ s_j = \mathrm{id},\ r_j = t_j = y_j = v_j = u_j = w_j = 0 & \text{if } L_i \text{ is of type } I^e. \end{cases}$$

A formal matrix form of an element of $F_j(R)$ for a κ-algebra R is then

$$\begin{pmatrix} \mathrm{id} & 0 & & \dots & & 0 \\ 0 & \ddots & & & & \\ & & \mathrm{id} & & & \\ \vdots & & & m_{j,j} & & \vdots \\ & & & & \mathrm{id} & \\ & & & & \ddots & 0 \\ 0 & & & \dots & 0 & \mathrm{id} \end{pmatrix}$$

such that

$$m_{j,j} = \begin{cases} \begin{pmatrix} \mathrm{id} & 0 \\ 0 & 1+\pi z_j \end{pmatrix} & \text{if } L_j \text{ is of type } I^o; \\ \begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & 1+\pi x_j & \pi z_j \\ 0 & 0 & 1 \end{pmatrix} & \text{if } L_j \text{ is of type } I^e. \end{cases}$$

In Lemma A.9, we will show that F_j is isomorphic to $\mathbb{A}^1 \times \mathbb{Z}/2\mathbb{Z}$ as a κ-variety so that it has exactly two connected components, by enumerating equations defining F_j as a closed subvariety of an affine space of dimension 2 (resp. 4) if L_j is of type I^o (resp. of type I^e). Here, $\mathbb{A}^1$ is an affine space of dimension 1. These equations are necessary in this theorem and thus we state them in Equation (4-2) below. We refer to Lemma A.9 for the proof. Let α be the unit in B such that $\epsilon = 1 + \alpha\pi$ as explained in Section 2A, and $\bar{\alpha}$ be the image of α in κ. We write $x_j = x_j^1 + \pi x_j^2$ and $z_j = z_j^1 + \pi z_j^2$, where $x_j^1, x_j^2, z_j^1, z_j^2 \in R \subset R \otimes_A B$ and π stands for $1 \otimes \pi \in R \otimes_A B$. Then the equations defining F_j as a closed subvariety of an affine space of dimension 2 (resp. 4) are

$$\begin{cases} (z_j^1/\bar{\alpha}) + (z_j^1/\bar{\alpha})^2 = 0 & \text{if } L_j \text{ is of type } I^o; \\ x_j^1 = z_j^1,\ (z_j^1/\bar{\alpha}) + (z_j^1/\bar{\alpha})^2 = 0,\ z_j^2 + x_j^2 + x_j^1 z_j^1 = 0 & \text{if } L_j \text{ is of type } I^e. \end{cases} \tag{4-2}$$

The proof of the surjectivity of ψ_j is given below. The main idea is to show that $\psi_j|_{F_j}$ is surjective. There are 4 cases according to the types of M_0 and L_j. Recall that $\bigoplus_{i\geq 0} M_i$ is a Jordan splitting of a rescaled hermitian lattice $(L^j, \frac{1}{\xi^{j/2}} h)$ and that $M_0 = \pi^{j/2} L_0 \oplus \pi^{j/2-1} L_2 \oplus \cdots \oplus \pi L_{j-2} \oplus L_j$.

(1) Assume that both M_0 and L_j are of type I^e. In this case and the next case, we will describe $\psi_j|_{F_j} : F_j \to \mathbb{Z}/2\mathbb{Z}$ explicitly in terms of a formal matrix. To do that, we will first describe a morphism from F_j to the special fiber of the smooth integral model associated to L^j and then to G_j. Recall that G_j is the special fiber of the smooth integral model associated to $C(L^j) = \bigoplus_{i\geq 1} M_i'$. Then we will describe a morphism from F_j to the even

orthogonal group associated to M_1' and compute the Dickson invariant of the image of an element of F_j in this orthogonal group.

We write $M_0 = N_0 \oplus L_j$, where N_0 is unimodular with even rank. Thus N_0 is either of type II or of type I^e. First we assume that N_0 is of type I^e. Then we can write $N_0 = \left(\bigoplus_{\lambda'} H_{\lambda'}\right) \oplus A(1, 2b, 1)$ and $L_j = \left(\bigoplus_{\lambda''} H_{\lambda''}\right) \oplus A(1, 2b', 1)$ by Theorem 2.10, where $H_{\lambda'} = H(0) = H_{\lambda''}$ and $b, b' \in A$. Thus we write $M_0 = \left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus A(1, 2b, 1) \oplus A(1, 2b', 1)$, where $H_\lambda = H(0)$. For this choice of a basis of $L^j = \bigoplus_{i \geq 0} M_i$, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & \begin{pmatrix} 1+\pi x_j & \pi z_j \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, id in the $(1, 1)$-block corresponds to the direct summand $\left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus A(1, 2b, 1)$ of M_0 and the diagonal block $\left(\begin{smallmatrix} 1+\pi x_j & \pi z_j \\ 0 & 1 \end{smallmatrix}\right)$ corresponds to the direct summand $A(1, 2b', 1)$ of M_0.

Let (e_1, e_2, e_3, e_4) be a basis for the direct summand $A(1, 2b, 1) \oplus A(1, 2b', 1)$ of M_0. Since this is *unimodular of type* I^e, we can choose another basis based on Theorem 2.10. With the basis $(-2be_1 + e_2, (2b'-1)e_1 + e_3 - e_4, e_3, e_2 + e_4)$, $A(1, 2b, 1) \oplus A(1, 2b', 1)$ becomes $A(2b(2b-1), 2b'(2b'-1), -(2b-1)(2b'-1)) \oplus A(1, 2(b+b'), 1)$. Since $A(2b(2b-1), 2b'(2b'-1), -(2b-1)(2b'-1))$ is *unimodular of type II*, it is isomorphic to $H(0)$ by Theorem 2.10. Thus we can write $M_0 = \left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus H(0) \oplus A(1, 2(b+b'), 1)$. For this basis, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix} * & *' & 0 \\ *'' & \begin{pmatrix} 1+\pi x_j & \pi z_j \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, the diagonal block $\left(\begin{smallmatrix} 1+\pi x_j & \pi z_j \\ 0 & 1 \end{smallmatrix}\right)$ corresponds to $A(1, 2(b+b'), 1)$ with basis $(e_3, e_2 + e_4)$ and the diagonal block $*$ corresponds to the direct summand $\left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus H(0)$ of M_0.

Then the direct summand M_1' of $C(L^j) = \bigoplus_{i \geq 1} M_i'$ is $(\pi)e_3 \oplus B(e_2 + e_4) \oplus M_1$. The image of a fixed element of F_j in the special fiber of the smooth integral model associated to $C(L^j)$ is then

$$\begin{pmatrix} \begin{pmatrix} 1+\pi x_j & z_j \\ 0 & 1 \end{pmatrix} & 0 & *' \\ 0 & \mathrm{id} & *'' \\ *''' & *'''' & * \end{pmatrix}.$$

Here, the diagonal block $\left(\begin{smallmatrix} 1+\pi x_j & z_j \\ 0 & 1 \end{smallmatrix}\right)$ corresponds to $(\pi)e_3 \oplus B(e_2 + e_4)$ and the diagonal block id corresponds to the direct summand M_1 of M_1'.

Now, the image of a fixed element of F_j in the orthogonal group associated to $M_1'/\pi M_1'$ is

$$T_1=\begin{pmatrix}\begin{pmatrix}1 & z_j^1\\ 0 & 1\end{pmatrix} & 0\\ 0 & \mathrm{id}\end{pmatrix}.$$

Note that z_j^1 is in R such that $z_j=z_j^1+\pi z_j^2$ as explained in the paragraph before Equation (4-2). The Dickson invariant of T_1 is the same as that of $\left(\begin{smallmatrix}1 & z_j^1\\ 0 & 1\end{smallmatrix}\right)$. Here we consider $\left(\begin{smallmatrix}1 & z_j^1\\ 0 & 1\end{smallmatrix}\right)$ as an element of the orthogonal group associated to $((\pi)e_3\oplus B(e_2+e_4))/\pi((\pi)e_3\oplus B(e_2+e_4))$. In order to compute the Dickson invariant, we use the scheme-theoretic description of the Dickson invariant explained in Remark 4.4 of [Cho 2015a]. The Dickson invariant of an orthogonal group of the quadratic space with dimension 2 is explicitly given at the end of the proof of Lemma 4.5 in [Cho 2015a]. Based on this, the Dickson invariant of $\left(\begin{smallmatrix}1 & z_j^1\\ 0 & 1\end{smallmatrix}\right)$ is $z_j^1/\bar{\alpha}$. Note that $z_j^1/\bar{\alpha}$ is indeed an element of $\mathbb{Z}/2\mathbb{Z}$ by Equation (4-2).

In conclusion, $z_j^1/\bar{\alpha}$ is the image of a fixed element of F_j under the map ψ_j. Since $z_j^1/\bar{\alpha}$ can be either 0 or 1, $\psi_j|_{F_j}$ is surjective onto $\mathbb{Z}/2\mathbb{Z}$ and thus ψ_j is surjective.

If N_0 is of type II, then the proof of the surjectivity of ψ_j is similar to that of the above case and so we skip it.

(2) Assume that M_0 is of type I^e and L_j is of type I^o. We write $M_0=N_0\oplus L_j$, where N_0 is unimodular with odd rank so that it is of type I^o. Then we can write $N_0=\left(\bigoplus_{\lambda'}H_{\lambda'}\right)\oplus(a)$ and $L_j=\left(\bigoplus_{\lambda''}H_{\lambda''}\right)\oplus(a')$ by Theorem 2.10, where $H_{\lambda'}=H(0)=H_{\lambda''}$ and $a,a'\in A$ such that $a,a'\equiv 1 \bmod 2$. Thus we write $M_0=\left(\bigoplus_{\lambda}H_{\lambda}\right)\oplus(a)\oplus(a')$, where $H_\lambda=H(0)$. For this choice of a basis of $L^j=\bigoplus_{i\geq 0}M_i$, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix}\mathrm{id} & 0 & 0\\ 0 & \left(1+\pi z_j\right) & 0\\ 0 & 0 & \mathrm{id}\end{pmatrix}.$$

Here, id in the $(1,1)$-block corresponds to the direct summand $\left(\bigoplus_{\lambda}H_{\lambda}\right)\oplus(a)$ of M_0 and the diagonal block $\left(1+\pi z_j\right)$ corresponds to the direct summand (a') of M_0.

Let (e_1,e_2) be a basis for the direct summand $(a)\oplus(a')$ of M_0. Since this is *unimodular of type I^e*, we can choose another basis (e_1,e_1+e_2) such that the associated Gram matrix is $A(a,a+a',a)$, where $a+a'\in(2)$. For this basis, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix}\mathrm{id} & 0 & 0\\ 0 & \begin{pmatrix}1 & -\pi z_j\\ 0 & 1+\pi z_j\end{pmatrix} & 0\\ 0 & 0 & \mathrm{id}\end{pmatrix}.$$

Here, the diagonal block $\left(\begin{smallmatrix}1 & -\pi z_j\\ 0 & 1+\pi z_j\end{smallmatrix}\right)$ corresponds to $A(a,a+a',a)$ with a basis (e_1,e_1+e_2) and id in the $(1,1)$-block corresponds to the direct summand $\left(\bigoplus_{\lambda}H_{\lambda}\right)\oplus(a)$ of M_0.

Then the direct summand M_1' of $C(L^j)=\bigoplus_{i\geq 1}M_i'$ is $(\pi)e_1\oplus B(e_1+e_2)\oplus M_1$. The image of a fixed element of F_j in the special fiber of the smooth integral model associated to $C(L^j)$ is then

$$\begin{pmatrix} \begin{pmatrix} 1 & -z_j \\ 0 & 1+\pi z_j \end{pmatrix} & 0 & 0 \\ 0 & \mathrm{id} & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, the diagonal block $\left(\begin{smallmatrix} 1 & -z_j \\ 0 & 1+\pi z_j \end{smallmatrix}\right)$ corresponds to $(\pi)e_1 \oplus B(e_1+e_2)$ and id in the (2×2)-block corresponds to the direct summand M_1 of M_1'.

Now, the image of a fixed element of F_j in the orthogonal group associated to $M_1'/\pi M_1'$ is

$$T_1 = \begin{pmatrix} \begin{pmatrix} 1 & z_j^1 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & \mathrm{id} \end{pmatrix}.$$

Note that $z_j^1 \in R$ is such that $z_j = z_j^1 + \pi z_j^2$, as explained in the paragraph before Equation (4-2). The Dickson invariant of T_1 is the same as that of $\left(\begin{smallmatrix} 1 & z_j^1 \\ 0 & 1 \end{smallmatrix}\right)$. Here, we consider $\left(\begin{smallmatrix} 1 & z_j^1 \\ 0 & 1 \end{smallmatrix}\right)$ as an element of the orthogonal group associated to $((\pi)e_1 \oplus B(e_1+e_2))/\pi((\pi)e_1 \oplus B(e_1+e_2))$. Then as explained in the above case (1), the Dickson invariant of $\left(\begin{smallmatrix} 1 & z_j^1 \\ 0 & 1 \end{smallmatrix}\right)$ is $z_j^1/\bar{\alpha}$. Note that $z_j^1/\bar{\alpha}$ is indeed an element of $\mathbb{Z}/2\mathbb{Z}$ by Equation (4-2).

In conclusion, $z_j^1/\bar{\alpha}$ is the image of a fixed element of F_j under the map ψ_j. Since $z_j^1/\bar{\alpha}$ can be either 0 or 1, $\psi_j|_{F_j}$ is surjective onto $\mathbb{Z}/2\mathbb{Z}$ and thus ψ_j is surjective.

(3) Assume that both M_0 and L_j are of type I^o. In this case, we will describe $\psi_j|_{F_j} : F_j \to \mathbb{Z}/2\mathbb{Z}$ explicitly in terms of a formal matrix. To do that, we will first describe a morphism from F_j to the special fiber of the smooth integral model associated to L^j and then to G_j. Recall that G_j is the special fiber of the smooth integral model associated to $C(L^j) = \bigoplus_{i\geq 0} M_i'$. Then we will describe a morphism from F_j to the special fiber of the smooth integral model associated to $M_0' \oplus C(L^j)$ and to the special fiber of the smooth integral model associated to $C(M_0' \oplus C(L^j))$. Finally, we will describe a morphism from F_j to a certain even orthogonal group associated to $C(M_0' \oplus C(L^j))$ and compute the Dickson invariant of the image of an element of F_j in this orthogonal group.

We write $M_0 = N_0 \oplus L_j$, where N_0 is unimodular with even rank. Thus N_0 is either of type *II* or of type I^e. First we assume that N_0 is of type I^e. Then we can write $N_0 = \left(\bigoplus_{\lambda'} H_{\lambda'}\right) \oplus A(1, 2b, 1)$ and $L_j = \left(\bigoplus_{\lambda''} H_{\lambda''}\right) \oplus (a)$ by Theorem 2.10, where $H_{\lambda'} = H(0) = H_{\lambda''}$, $b \in A$, and $a\ (\in A) \equiv 1 \bmod 2$. Thus we write $M_0 = \left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus A(1, 2b, 1) \oplus (a)$, where $H_\lambda = H(0)$. For this choice of a basis of $L^j = \bigoplus_{i\geq 0} M_i$, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & \left(1+\pi z_j\right) & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, id in the $(1, 1)$-block corresponds to the direct summand $\left(\bigoplus_\lambda H_\lambda\right) \oplus A(1, 2b, 1)$ of M_0 and the diagonal block $\left(1+\pi z_j\right)$ corresponds to the direct summand (a) of M_0.

Let (e_1, e_2, e_3) be a basis for the direct summand $A(1, 2b, 1) \oplus (a)$ of M_0. Since this is *unimodular of type I^o*, we can choose another basis based on Theorem 2.10. Namely, if we choose $(-2be_1+e_2, -ae_1+e_3, e_2+e_3)$ as another basis, then $A(1,2b,1)\oplus(a)$ becomes $A(2b(2b-1), a(a+1), a(2b-1))\oplus(a+2b)$. Since $A(2b(2b-1), a(a+1), a(2b-1))$ is *unimodular of type II*, it is isomorphic to $H(0)$ by Theorem 2.10. Thus we can write $M_0=\left(\bigoplus_\lambda H_\lambda\right)\oplus H(0)\oplus(a+2b)$. For this basis, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix} * & *' & 0 \\ *'' & \left(1+\frac{a}{a+2b}\pi z_j\right) & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, the diagonal block $\left(1+\frac{a}{a+2b}\pi z_j\right)$ corresponds to $(a+2b)$ with a basis e_2+e_3 and the diagonal block $*$ corresponds to the direct summand $\left(\bigoplus_\lambda H_\lambda\right)\oplus H(0)$ of M_0.

Then the direct summand M_0' of $C(L^j)=\bigoplus_{i\geq 0}M_i'$ is $B(e_2+e_3)$ of rank 1. The image of a fixed element of F_j in the special fiber of the smooth integral model associated to $C(L^j)$ is then

$$\begin{pmatrix} \left(1+\frac{a}{a+2b}\pi z_j\right) & 0 & *' \\ 0 & \mathrm{id} & *'' \\ *''' & *'''' & * \end{pmatrix}.$$

Here, the diagonal block $\left(1+\frac{a}{a+2b}\pi z_j\right)$ corresponds to $M_0'=B(e_2+e_3)$ with a Gram matrix $(a+2b)$ and the diagonal block id corresponds to $M_1'=M_1$.

We now describe the image of the above in the special fiber of the smooth integral model associated to $M_0'\oplus C(L^j)=(M_0'\oplus M_0')\oplus\left(\bigoplus_{i\geq 1}M_i'\right)$. If (e_1', e_2') is a basis for $(M_0'\oplus M_0')$, then we choose another basis $(e_1', e_1'+e_2')$ for $(M_0'\oplus M_0')$. For this basis, based on the description of the morphism from the smooth integral model associated to $C(L^j)$ to the smooth integral model associated to $M_0'\oplus C(L^j)$ explained in Remark 4.10, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to $M_0'\oplus C(L^j)$ is

$$\begin{pmatrix} 1 & -\frac{a}{a+2b}\pi z_j & 0 & *' \\ 0 & 1+\frac{a}{a+2b}\pi z_j & 0 & *' \\ 0 & 0 & \mathrm{id} & *'' \\ 0 & *''' & *'''' & * \end{pmatrix}.$$

Here, the diagonal block $\left(\begin{smallmatrix} 1 & -\frac{a}{a+2b}\pi z_j \\ 0 & 1+\frac{a}{a+2b}\pi z_j \end{smallmatrix}\right)$ corresponds to $(M_0'\oplus M_0')$ with a basis $(e_1', e_1'+e_2')$ and the diagonal block id corresponds to $M_1'=M_1$.

We now follow step (1) with $M_0'\oplus C(L^j)=(M_0'\oplus M_0')\oplus\left(\bigoplus_{i\geq 1}M_i'\right)$. Namely,

$$\begin{aligned} C(M_0'\oplus C(L^j)) &= (\pi)e_1'\oplus B(e_1'+e_2')\oplus\left(\bigoplus_{i\geq 1}M_i'\right) \\ &= \left((\pi)e_1'\oplus B(e_1'+e_2')\oplus M_1'\right)\oplus\left(\bigoplus_{i\geq 2}M_i'\right). \end{aligned}$$

Here, $((\pi)e_1'\oplus B(e_1'+e_2')\oplus M_1')$ is π^1-modular and M_i' is π^i-modular with $i\geq 2$. Then the image of a fixed element of F_j in the special fiber of the smooth integral model associated to $C(M_0' \oplus C(L^j))$ is

$$\begin{pmatrix} 1 & -\frac{a}{a+2b}z_j & 0 & *' \\ 0 & 1+\frac{a}{a+2b}\pi z_j & 0 & *' \\ 0 & 0 & \mathrm{id} & *'' \\ 0 & *''' & *'''' & * \end{pmatrix}.$$

Here, the top left 3×3-matrix corresponds to $(\pi e_1' \oplus B(e_1' + e_2') \oplus M_1')$.

Now, the image of a fixed element of F_j in the orthogonal group associated to $(\pi e_1 \oplus B(e_1 + e_2) \oplus M_1')/\pi(\pi e_1 \oplus B(e_1 + e_2) \oplus M_1')$ is

$$T_1 = \begin{pmatrix} \begin{pmatrix} 1 & z_j^1 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & \mathrm{id} \end{pmatrix}$$

since mod 2 reduction of $\frac{a}{a+2b}$ is 1. Note that z_j^1 is in R such that $z_j = z_j^1 + \pi z_j^2$ as explained in the paragraph before Equation (4-2). Then, as explained in step (1), the Dickson invariant of this is $z_j^1/\bar{\alpha}$. Note that $z_j^1/\bar{\alpha}$ is indeed an element of $\mathbb{Z}/2\mathbb{Z}$ by Equation (4-2).

In conclusion, $z_j^1/\bar{\alpha}$ is the image of a fixed element of F_j under the map ψ_j. Since $z_j^1/\bar{\alpha}$ can be either 0 or 1, $\psi_j|_{F_j}$ is surjective onto $\mathbb{Z}/2\mathbb{Z}$ and thus ψ_j is surjective.

If N_0 is of type II, then the proof of the surjectivity of ψ_j is similar to that of the above case and so we skip it.

(4) Assume that M_0 is of type I^o and L_j is of type I^e. We write $M_0 = N_0 \oplus L_j$, where N_0 is unimodular with odd rank so that it is of type I^o. Then we can write $N_0 = \left(\bigoplus_{\lambda'} H_{\lambda'}\right) \oplus (a)$ and $L_j = \left(\bigoplus_{\lambda''} H_{\lambda''}\right) \oplus A(1, 2b, 1)$ by Theorem 2.10, where $H_{\lambda'} = H(0) = H_{\lambda''}$ and $a, b \in A$ such that $a \equiv 1 \mod 2$. We write $M_0 = \left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus (a) \oplus A(1, 2b, 1)$, where $H_{\lambda} = H(0)$. For this choice of a basis of $L^j = \bigoplus_{i \geq 0} M_i$, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & \begin{pmatrix} 1+\pi x_j & \pi z_j \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, id in the $(1, 1)$-block corresponds to the direct summand $\left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus (a)$ of M_0 and the diagonal block $\left(\begin{smallmatrix} 1+\pi x_j & \pi z_j \\ 0 & 1 \end{smallmatrix}\right)$ corresponds to the direct summand $A(1, 2b, 1)$ of M_0.

Let (e_1, e_2, e_3) be a basis for the direct summand $(a) \oplus A(1, 2b, 1)$ of M_0. Since this is *unimodular of type* I^o, we can choose another basis based on Theorem 2.10. Namely, if we choose $(-2be_2 + e_3, e_1 - ae_2, e_1 + e_3)$ as another basis, then $(a) \oplus A(1, 2b, 1)$ becomes $A(2b(2b-1), a(a+1), a(2b-1)) \oplus (a+2b)$. Since $A(2b(2b-1), a(a+1), a(2b-1))$ is *unimodular of type* II, it is isomorphic to $H(0)$ by Theorem 2.10. Thus we can write $M_0 = \left(\bigoplus_{\lambda} H_{\lambda}\right) \oplus H(0) \oplus (a+2b)$. For this basis, the image of a fixed element of F_j in the special fiber of the smooth integral model associated to L^j is

$$\begin{pmatrix} * & *' & 0 \\ *'' & \left(1+\frac{1}{a+2b}\pi z_j\right) & 0 \\ 0 & 0 & \mathrm{id} \end{pmatrix}.$$

Here, the diagonal block $\left(1+\frac{1}{a+2b}\pi z_j\right)$ corresponds to $(a+2b)$ with a basis (e_1+e_3) and the diagonal block $*$ corresponds to the direct summand $\left(\bigoplus_\lambda H_\lambda\right) \oplus H(0)$ of M_0.

Note that the reduction of $\frac{1}{a+2b}$ mod 2 is 1. The rest of the proof is similar to that of step (3) and so we skip it.

So far, we have proved that ψ_j is surjective. We now show that $\psi = \prod_j \psi_j$ is surjective. The proof is similar to the proof showing that $\prod_{i\in\mathcal{H}} H_i \to \widetilde{G}$ is a closed immersion in the last paragraph of the proof of Theorem 4.5.

We consider the morphism

$$F = \prod_j F_j \longrightarrow \widetilde{G}$$
$$(f_j) \mapsto \prod_j f_j.$$

By considering a matrix form of an element of $F_j(R)$ for a κ-algebra R as given at the beginning of the proof, it is easy to see the following two facts. Firstly, F_j and $F_{j'}$ commute with each other in the sense that $f_j \cdot f_{j'} = f_{j'} \cdot f_j$ for all even integers $j \neq j'$, where $f_j \in F_j(R)$ and $f_{j'} \in F_{j'}(R)$ for a κ-algebra R. Note that L_j and $L_{j'}$ (resp. L_{j+2} and $L_{j'+2}$) are of type I (resp. of type II). Based on this, the above morphism becomes a group homomorphism. Secondly, $F_j \cap F_{j'} = 0$ for all $j \neq j'$. This fact implies that the morphism $F_j \times F_{j'} \to \widetilde{G}$ with $(f_j, f_{j'}) \mapsto f_j \cdot f_{j'}$ is injective and so $F_j \times F_{j'}$ is a closed subgroup scheme of $\widetilde{G}$. A matrix form of an element of $F_j(R)$ also implies that $(F_j \times F_{j'}) \cap F_{j''} = 0$ for all pairwise different three integers j, j', j'' and so the morphism $(F_j \times F_{j'}) \times F_{j''} \to \widetilde{G}$ with $(f_j, f_{j'}, f_{j''}) \mapsto f_j \cdot f_{j'} \cdot f_{j''}$ is injective. Thus $F_j \times F_{j'} \times F_{j''}$ is a closed subgroup scheme of $\widetilde{G}$. Therefore, by repeating this argument, the product $F = \prod_j F_j$ is embedded into $\widetilde{G}$ as a closed subgroup scheme.

In addition, we claim that $\psi_j|_{F_{j'}}$ is trivial for all $j < j'$. The proof of our claim relies on the matrix interpretation of ψ_j. We first notice that $j' - j \geq 4$ since L_j is of type I and L_{j+2} is of type II. To obtain the morphism ψ_j, we observe that the lattice $C(L^j) = \bigoplus_{i\geq 1} M_i'$ (resp. $C(L^j) = \bigoplus_{i\geq 0} M_i'$) if M_0 is of type I^e (resp. of type I^o). In either case, $L_{j'}$ is a direct summand of $M_{j'-j}$ and the morphism ψ_j is attached to the Dickson invariant of the orthogonal group associated to M_1'. We should mention that if M_0 is of type I^o then we need a new hermitian lattice $M_0' \oplus C(L^j)$. In this case, the morphism ψ_j is also attached to the Dickson invariant of the orthogonal group associated to M_1' as a direct summand of $M_0' \oplus C(L^j)$. On the other hand, recall that G_j is the special fiber of the smooth integral model associated to $C(L^j)$. Then as a formal matrix, $F_{j'}$ maps to the block of G_j associated to $M_{j'-j}$. Therefore, since $j' - j$ is at least 4, the image of $F_{j'}$ under ψ_j is zero by observing the description of the orthogonal group associated to M_1' based on Section 4A.

We finally claim that the morphism ψ induces a surjective morphism from F to $(\mathbb{Z}/2\mathbb{Z})^\beta$ defined over κ. To show this, we express F as $F = F_{j_1} \times \cdots \times F_{j_\beta}$ and $(\mathbb{Z}/2\mathbb{Z})^\beta$ as $(\mathbb{Z}/2\mathbb{Z})^\beta = (\mathbb{Z}/2\mathbb{Z})_{j_1} \times \cdots \times (\mathbb{Z}/2\mathbb{Z})_{j_\beta}$, where $j_i < j_{i'}$ if $i < i'$. Choose an arbitrary element $(z_{j_1}, \cdots, z_{j_\beta})$ of $(\mathbb{Z}/2\mathbb{Z})_{j_1} \times \cdots \times (\mathbb{Z}/2\mathbb{Z})_{j_\beta}$ where each z_{j_i} is an element of $(\mathbb{Z}/2\mathbb{Z})_{j_i}$. We first choose $f_{j_1} \in F_{j_1}$ such that $\psi_{j_1}(f_{j_1}) = z_{j_1}$. Then choose $f_{j_2} \in F_{j_2}$ such that $\psi_{j_2}(f_{j_1} \cdot f_{j_2}) = z_{j_2}$. In this way, we choose $f_{j_t} \in F_{j_t}$ such that $\psi_{j_t}(f_{j_1} \cdot \cdots \cdot f_{j_t}) = z_{j_t}$. Note that $\psi_{j_t}(f_{j_{t'}}) = 0$ for all $t < t'$. Therefore, $\psi(f_{j_1} \cdots f_{j_\beta}) = \prod_t \psi_{j_t}(f_{j_1} \cdots f_{j_\beta}) = (z_{j_1}, \cdots, z_{j_\beta})$ and this shows the surjectivity of the morphism ψ.

For the surjectivity of $\varphi \times \psi$, we recall the following criterion ([Knus et al. 1998, Proposition 22.3]): the surjectivity of $\varphi \times \psi$ as algebraic groups is equivalent to the surjectivity of $\varphi \times \psi$ at the level of $\bar{\kappa}$-points since $\prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}} \times (\mathbb{Z}/2\mathbb{Z})^\beta$ is smooth.

Choose an element (x, y) in the group of $\bar{\kappa}$-points of

$$\prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}} \times (\mathbb{Z}/2\mathbb{Z})^\beta$$

such that $x \in \left(\prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}\right)(\bar{\kappa})$ and $y \in (\mathbb{Z}/2\mathbb{Z})^\beta(\bar{\kappa})$. Then there is an element $a \in \widetilde{G}(\bar{\kappa})$ such that $\varphi(a) = x$ since φ is surjective by Theorem 4.5. We choose an element $b \in F(\bar{\kappa})$ such that $\psi(ab) = y$. On the other hand, φ vanishes on F since the morphism φ_i vanishes on F_j for all i, j. Thus $\varphi(b) = 0$ and $(\varphi \times \psi)(ab) = (x, y)$. This completes the proof. □

4C. *The maximal reductive quotient.* We finally have the structure theorem for the algebraic group $\widetilde{G}$.

Theorem 4.12. *The morphism*

$$\varphi \times \psi : \widetilde{G} \longrightarrow \prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}} \times (\mathbb{Z}/2\mathbb{Z})^\beta$$

is surjective and the kernel is unipotent and connected. Consequently,

$$\prod_{i \text{ even}} \mathrm{Sp}(B_i/Y_i, h_i) \times \prod_{i \text{ odd}} O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}} \times (\mathbb{Z}/2\mathbb{Z})^\beta$$

is the maximal reductive quotient. Here, $\mathrm{Sp}(B_i/Y_i, h_i)$ *and* $O(A_i/Z_i, \bar{q}_i)_{\mathrm{red}}$ *are explained in Section 4A* (*especially Remark 4.7*) *and* β *is defined in Lemma 4.6.*

Proof. We only need to prove that the kernel is unipotent and connected. The kernel of φ is a closed subgroup scheme of the unipotent group $\widetilde{M}^+$ which is defined in Lemma A.2 and so it suffices to show that the kernel of $\varphi \times \psi$ is connected. Equivalently, it suffices to show that the kernel of the restricted morphism $\psi|_{\mathrm{Ker}\,\varphi}$ is connected. From Lemma 4.6, the number of connected components of $\mathrm{Ker}\,\varphi$ is 2^β. Since $\varphi|_F = 0$ so that $F = \prod_j F_j \subset \mathrm{Ker}\,\varphi$, the restricted morphism $\psi|_{\mathrm{Ker}\,\varphi}$ is surjective onto $(\mathbb{Z}/2\mathbb{Z})^\beta$. We complete the proof by counting the number of connected components. □

5. Comparison of volume forms and final formulas

This section is based on Section 7 of [Gan and Yu 2000] and Section 5 of [Cho 2015a]. Let H be the F-vector space of hermitian forms on $V = L \otimes_A F$. Let $M' = \mathrm{End}_B(L)$ and let $H' = \{f : f \text{ is a hermitian form on } L\}$. Regarding $\mathrm{End}_E V$ and H as varieties over F, let ω_M and ω_H be nonzero, translation-invariant forms on $\mathrm{End}_E V$ and H, respectively, with normalization

$$\int_{M'} |\omega_M| = 1 \quad \text{and} \quad \int_{H'} |\omega_H| = 1.$$

Let $M^* = \mathrm{Res}_{E/F}\,\mathrm{GL}_E(V)$. Define a map $\rho : M^* \to H$ by $\rho(m) = h \circ m$. Here $h \circ m$ is the hermitian form $(v, w) \mapsto h(mv, mw)$. Then the inverse of h under ρ is G, which is the unitary group associated to the hermitian space (V, h). It is also the generic fiber of $\underline{G}'$. Put $\omega^{\mathrm{ld}} = \omega_M/\rho^*\omega_H$. For a detailed explanation of what $\omega_M/\rho^*\omega_H$ means, we refer to Section 3.2 of [Gan and Yu 2000].

We choose two forms ω'_M and ω'_H as generators for the spaces of the top degree forms on $\underline{M}'$, which is identified with the Lie algebra of $\underline{M}^*$, and $\underline{H}'$, which is identified with the tangent space to $\underline{H}$ at h, respectively. Here $\underline{M}'$ is defined in Remark 3.1 and $\underline{H}'$ is defined in the paragraph following the matrix description of an element of $\underline{H}(R)$ for a flat A-algebra R in Section 3C. They are nonzero translation-invariant forms on $\mathrm{End}_E V$ and H, respectively, with normalization

$$\int_{\underline{M}(A)} |\omega'_M| = 1 \quad \text{and} \quad \int_{\underline{H}(A)} |\omega'_H| = 1.$$

By Theorem 3.6, we have an exact sequence of locally free sheaves on $\underline{M}^*$:

$$0 \longrightarrow \rho^*\Omega_{\underline{H}/A} \longrightarrow \Omega_{\underline{M}^*/A} \longrightarrow \Omega_{\underline{M}^*/\underline{H}} \longrightarrow 0.$$

Put $\omega^{\mathrm{can}} = \omega'_M/\rho^*\omega'_H$. For a detailed explanation of what $\omega'_M/\rho^*\omega'_H$ means, we refer to Section 3.2 of [Gan and Yu 2000]. It follows that ω^{can} is a differential of top degree on $\underline{G}$, which is invariant under the generic fiber of $\underline{G}$, and which has nonzero reduction on the special fiber.

Lemma 5.1. *We have*:

$$|\omega_M| = |2|^{N_M}|\omega'_M|, \qquad N_M = \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I}} (2n_i - 1) + \sum_{i<j} (j-i)\cdot n_i \cdot n_j,$$

$$|\omega_H| = |2|^{N_H}|\omega'_H|, \qquad N_H = \sum_{\substack{i \text{ even} \\ L_i \text{ of type I}}} (n_i - 1) + \sum_{i<j} j \cdot n_i \cdot n_j + \sum_{i \text{ even}} \frac{i+2}{2}\cdot n_i + \sum_{i \text{ odd}} \frac{i+1}{2}\cdot n_i + \sum_i d_i,$$

$$|\omega^{\mathrm{ld}}| = |2|^{N_M - N_H}|\omega^{\mathrm{can}}|.$$

Here, $d_i = i \cdot n_i \cdot (n_i - 1)/2$.

Proof. Note that both ω_M and ω'_M are volume forms on $\mathrm{End}_E V$ with different normalizations, so that they differ by a scalar. The "difference" between the Haar measures associated to these volume forms can be detected at the level of F-points of $\mathrm{End}_E V$, since $\mathrm{End}_E V$ is an affine space.

Since $\underline{M}(A) = 1 + \underline{M}'(A)$, where $\underline{M}'$ is defined in Remark 3.1, we have the identity $\int_{\underline{M}'(A)} |\omega'_M| = 1$. Note that $\underline{M}'(A)$ is a finitely generated free A-submodule of M' whose rank is the same as that of M'. Thus N_M is the "difference" between these two modules M' and $\underline{M}'(A)$. More precisely, N_M is the length of the finitely generated torsion A-module $M'/\underline{M}'(A)$. Note that 2 is a uniformizer of A.

Similarly, N_H is the length of the finitely generated torsion A-module $H'/\underline{H}'(A)$. Here, $\underline{H}'$ is defined in the paragraph following the matrix description of an element of $\underline{H}(R)$ for a flat A-algebra R in Section 3C.

Then the above formula for N_M (resp. N_H) can be read off from the matrix interpretation for $\underline{M}(A)$ (resp. $\underline{H}(A)$) given in Sections 3A and 3B (resp. Section 3C). □

Let f be the cardinality of κ. The local density is defined as

$$\beta_L = \frac{1}{[G : G^\circ]} \cdot \lim_{N\to\infty} f^{-N \dim G} \# \underline{G}'(A/\pi^N A).$$

Here, $\underline{G}'$ is the naive integral model described at the beginning of Section 3 and G is the generic fiber of $\underline{G}'$ and G° is the identity component of G. In our case, G is the unitary group $U(V, h)$, where $V = L \otimes_A F$. Since $U(V, h)$ is connected, G° is the same as G so that $[G : G^\circ] = 1$.

Then based on Lemma 3.4 and Section 3.9 of [Gan and Yu 2000], we finally have the following local density formula.

Theorem 5.2. *Let f be the cardinality of κ. The local density of (L, h) is*

$$\beta_L = f^N \cdot f^{-\dim U(V,h)} \# \widetilde{G}(\kappa),$$

where

$$N = N_H - N_M = \sum_{i<j} i \cdot n_i \cdot n_j + \sum_{i \text{ even}} \frac{i+2}{2} \cdot n_i + \sum_{i \text{ odd}} \frac{i+1}{2} \cdot n_i + \sum_i d_i - \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I}} n_i .$$

Here, $\#\widetilde{G}(\kappa)$ can be computed explicitly based on Remark 5.3(1) *below and Theorem 4.12.*

For convenience, we repeat the following remark from Remark 5.3 in [Cho 2015a].

Remark 5.3 [Cho 2015a, Remark 5.3]. (1) In the above local density formula, $\#\widetilde{G}(\kappa)$ is computed as follows. We denote by $R_u\widetilde{G}$ the unipotent radical of $\widetilde{G}$ so that the maximal reductive quotient of $\widetilde{G}$ is $\widetilde{G}/R_u\widetilde{G}$. That is, there is the following exact sequence of group schemes over κ:

$$1 \longrightarrow R_u\widetilde{G} \longrightarrow \widetilde{G} \longrightarrow \widetilde{G}/R_u\widetilde{G} \longrightarrow 1.$$

Furthermore, the following sequence of groups

$$1 \longrightarrow R_u\widetilde{G}(\kappa) \longrightarrow \widetilde{G}(\kappa) \longrightarrow (\widetilde{G}/R_u\widetilde{G})(\kappa) \longrightarrow 1$$

is also exact by Lemma A.1. Using Lemma A.1, one can see that $\#R_u\widetilde{G}(\kappa) = f^m$, where m is the dimension of $R_u\widetilde{G}$. Notice that the dimension of $R_u\widetilde{G}$ can be computed explicitly based on Theorem 4.12, since the dimension of $\widetilde{G}$ is n^2 with $n = \mathrm{rank}_B L$. In addition, the orders of orthogonal and symplectic groups defined over a finite field are well known. Thus, one can compute $\#(\widetilde{G}/R_u\widetilde{G})(\kappa)$ explicitly based on Theorem 4.12. Finally, the order of the group $\widetilde{G}(\kappa)$ is identified as follows:

$$\#\widetilde{G}(\kappa) = \#R_u\widetilde{G}(\kappa) \cdot \#(\widetilde{G}/R_u\widetilde{G})(\kappa).$$

(2) As in Remark 7.4 of [Gan and Yu 2000], although we have assumed that $n_i = 0$ for $i < 0$, it is easy to check that the formula in the preceding theorem remains true without this assumption.

Appendix A: The proof of Lemma 4.6

The proof of Lemma 4.6 is based on Proposition 6.3.1 in [Gan and Yu 2000]. We first state a theorem of Lazard which is repeatedly used in this paper. Let U be a group scheme of finite type over κ which is isomorphic to an affine space as an algebraic variety. Then U is connected smooth unipotent group (cf. IV, § 4, Theorem 4.1 and IV, § 2, Corollary 3.9 in [Demazure and Gabriel 1970]).

For preparation, we state several lemmas.

Lemma A.1 [Gan and Yu 2000, Lemma 6.3.3]. *Let $1 \to X \to Y \to Z \to 1$ be an exact sequence of group schemes that are locally of finite type over κ, where κ is a perfect field. Suppose that X is smooth, connected, and unipotent. Then $1 \to X(R) \to Y(R) \to Z(R) \to 1$ is exact for any κ-algebra R.*

Let $\widetilde{M}$ be the special fiber of $\underline{M}^*$ and let R be a κ-algebra. Recall that we have described an element and the multiplication of elements of $\underline{M}(R)$ in Section 3B. Based on these, an element of $\widetilde{M}(R)$ is

$$m = \left(\pi^{\max\{0,j-i\}} m_{i,j}\right).$$

Here, if i is even and L_i is of type I^o (resp. of type I^e), then

$$m_{i,i} = \begin{pmatrix} s_i & \pi y_i \\ \pi v_i & 1+\pi z_i \end{pmatrix} \quad (\text{resp. } \begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & 1+\pi x_i & \pi z_i \\ v_i & u_i & 1+\pi w_i \end{pmatrix}),$$

where $s_i \in M_{(n_i-1)\times(n_i-1)}(B\otimes_A R)$ (resp. $s_i \in M_{(n_i-2)\times(n_i-2)}(B\otimes_A R)$), etc., and $s_i \mod \pi\otimes 1$ is invertible. For the remaining $m_{i,j}$'s except for the cases explained above, $m_{i,j}$ is contained in $M_{n_i\times n_j}(B\otimes_A R)$ and $m_{i,i} \mod \pi\otimes 1$ is invertible.

Let

$$\widetilde{M}_i = \begin{cases} \mathrm{GL}_\kappa(B_i/Y_i) & \text{if } i \text{ is even;} \\ \mathrm{GL}_\kappa(A_i/X_i) & \text{if } i \text{ is odd.} \end{cases}$$

Let $s_i = m_{i,i}$ if L_i is of type *II* in the above description of $\widetilde{M}(R)$. Then $s_i \mod \pi\otimes 1$ is an element of $\widetilde{M}_i(R)$. Therefore, we have a surjective morphism of algebraic groups

$$r : \widetilde{M} \longrightarrow \prod \widetilde{M}_i,$$

defined over κ. We now have the following easy lemma:

Lemma A.2. *The kernel of r is the unipotent radical $\widetilde{M}^+$ of $\widetilde{M}$, and $\prod \widetilde{M}_i$ is the maximal reductive quotient of $\widetilde{M}$.*

Proof. Since $\prod \widetilde{M}_i$ is a reductive group, we only have to show that the kernel of r is a connected smooth unipotent group. Let R be a κ-algebra. By the description of the morphism r in terms of matrices explained above, an element of the kernel of r is

$$m = \left(\pi^{\max\{0,j-i\}} m_{i,j}\right)$$

satisfying the following. If i is even and L_i is of type I^o (resp. of type I^e), then

$$m_{i,i} = \begin{pmatrix} \mathrm{id} + \pi s_i' & \pi y_i \\ \pi v_i & 1 + \pi z_i \end{pmatrix} \quad \left(\text{resp. } \begin{pmatrix} \mathrm{id} + \pi s_i' & r_i & \pi t_i \\ \pi y_i & 1 + \pi x_i & \pi z_i \\ v_i & u_i & 1 + \pi w_i \end{pmatrix}\right),$$

where $\mathrm{id} + \pi \otimes 1 \cdot s_i' \in M_{(n_i-1)\times(n_i-1)}(B \otimes_A R)$ (resp. $\mathrm{id} + \pi \otimes 1 \cdot s_i' \in M_{(n_i-2)\times(n_i-2)}(B \otimes_A R)$), etc., such that s_i' has entries in $R \subset B \otimes_A R$. For the remaining $m_{i,j}$'s except for the cases explained above, $m_{i,j} \in M_{n_i \times n_j}(B \otimes_A R)$ and $m_{i,i} = \mathrm{id} + \pi \otimes 1 \cdot m_{i,i}'$ such that $m_{i,i}'$ has entries in $R \subset B \otimes_A R$. Note that there are no equations among the variables given above. Thus the kernel of r is isomorphic to an affine space as an algebraic variety over κ. Therefore, it is a connected smooth unipotent group by a theorem of Lazard which is stated at the beginning of Appendix A. □

Recall that we have defined the morphism φ in Section 4A. The morphism φ extends to an obvious morphism

$$\tilde{\varphi} : \widetilde{M} \longrightarrow \prod_{i \text{ even}} \mathrm{GL}_\kappa(B_i/Y_i) \times \prod_{i \text{ odd}} \mathrm{GL}_\kappa(A_i/Z_i)$$

such that $\tilde{\varphi}|_{\tilde{G}} = \varphi$. Note that $Y_i \otimes_A R$ and $Z_i \otimes_A R$ are preserved by an element of $\underline{M}(R)$ for a flat A-algebra R (cf. Lemma 4.2). By using this, the construction of $\tilde{\varphi}$ is similar to Theorems 4.3 and 4.4 and thus we skip it. Let R be a κ-algebra. Based on the description of the morphism φ_i explained in Section 4A, $\mathrm{Ker}\,\tilde{\varphi}(R)$ is the subgroup of $\widetilde{M}(R)$ defined by the following conditions:

(a) If i is even and L_i is of type I, $s_i = \mathrm{id} \mod \pi \otimes 1$.

(b) If i is even and L_i is of type II, $m_{i,i} = \mathrm{id} \mod \pi \otimes 1$.

(c) If i is odd, $m_{i,i} = \mathrm{id} \mod \pi \otimes 1$ and $\delta_{i-1} e_{i-1} \cdot m_{i-1,i} + \delta_{i+1} e_{i+1} \cdot m_{i+1,i} = 0 \mod \pi \otimes 1$. Here, $\delta_j = 1$ if L_j is of type I and $\delta_j = 0$ if L_j is of type II, and $e_j = (0, \cdots, 0, 1)$ (resp. $e_j = (0, \cdots, 0, 1, 0)$) of size $1 \times n_j$ if L_j is of type I^o (resp. of type I^e).

It is obvious that $\mathrm{Ker}\,\tilde{\varphi}$ is a closed subgroup scheme of $\widetilde{M}^+$ and is smooth and unipotent since it is isomorphic to an affine space as an algebraic variety over κ.

Recall from Remark 3.1 that we defined the functor $\underline{M}'$ such that $(1 + \underline{M}')(R) = \underline{M}(R)$ inside $\mathrm{End}_{B \otimes_A R}(L \otimes_A R)$ for a flat A-algebra R. Thus there is an isomorphism of set valued functors

$$1+ : \underline{M}' \to \underline{M}$$
$$m \mapsto 1 + m,$$

where $m \in \underline{M}'(R)$ for a flat A-algebra R. We define a new operation $\star$ on $\underline{M}'(R)$ such that $x \star y = x + y + xy$ for a flat A-algebra R. Since $\underline{M}'(R)$ is closed under addition and multiplication, it is also closed under the new operation $\star$. Moreover, it has 0 as an identity element with respect to $\star$. Thus $\underline{M}'$ may and shall be considered as a scheme of monoids with $\star$. We claim that the above morphism $1+$ is an isomorphism of monoid schemes. Namely, we claim the following commutative diagram of schemes:

$$\begin{tikzcd} \underline{M}' \times \underline{M}' \arrow[r, "(1+)\times(1+)"] \arrow[d, "\star"] & \underline{M} \times \underline{M} \arrow[d, "\text{multiplication}"] \\ \underline{M}' \arrow[r, "1+"] & \underline{M} \end{tikzcd}$$

Since all schemes are irreducible and smooth, it suffices to check the commutativity of the diagram at the level of flat A-points as explained in the third paragraph from below in Remark 3.2, and this is obvious.

Since $\underline{M}^*$ is an open subscheme of $\underline{M}$, $(1+)^{-1}(\underline{M}^*)$ is an open subscheme of $\underline{M}'$. The composite of the following three morphisms

$$(1+)^{-1}(\underline{M}^*) \xrightarrow{(1+)} \underline{M}^* \xrightarrow{\text{inverse}} \underline{M}^* \xrightarrow{(1+)^{-1}} (1+)^{-1}(\underline{M}^*)$$

defines the inverse morphism on the scheme of monoids $(1+)^{-1}(\underline{M}^*)$ with respect to the operation $\star$. Thus we can see that $(1+)^{-1}(\underline{M}^*)$ is a group scheme with respect to $\star$ and the morphism $1+$ is an isomorphism of group schemes between $(1+)^{-1}(\underline{M}^*)$ and $\underline{M}^*$.

Let R be a κ-algebra. Since the morphism $1+$ is an isomorphism of monoid schemes between $\underline{M}'$ and $\underline{M}$, we can write each element of $\underline{M}(R)$ as $1 + x$ with $x \in \underline{M}'(R)$. Here, $1 + x$ means the image of x under the morphism $1+$ at the level of R-points. Note that $\underline{M}'(R)$ is a $B \otimes_A R$-algebra for any A-algebra R with respect to the original multiplication on it, not the operation $\star$. In particular, $\underline{M}'(R)$ is a $(B/2B) \otimes_A R$-algebra for any κ-algebra R. Therefore, we consider the subfunctor $\underline{\pi M'} : R \mapsto (\pi \otimes 1)\underline{M}'(R)$ of $\underline{M}' \otimes \kappa$ and the subfunctor $\tilde{M}^1 : R \mapsto 1 + \underline{\pi M'}(R)$ of $\operatorname{Ker} \tilde{\varphi}$. Here, by $1 + \underline{\pi M'}(R)$, we mean the image of $\underline{\pi M'}(R)$ inside $\underline{M}(R)$ $(= \tilde{M}(R))$ under the morphism $1+$ at the level of R-points. That $1 + \underline{\pi M'}(R)$ is contained in $\operatorname{Ker} \tilde{\varphi}(R)$ can easily be checked by observing the construction of $\tilde{\varphi}$. The multiplication on $\tilde{M}^1$ is as follows: for two elements $1 + \pi x$ and $1 + \pi y$ in $\tilde{M}^1(R)$, based on the above commutative diagram, the product of $1 + \pi x$ and $1 + \pi y$ is

$$(1 + \pi x) \cdot (1 + \pi y) = 1 + \pi x \star \pi y = 1 + (\pi(x + y) + \pi^2(xy)) = 1 + \pi(x + y).$$

Here, π stands for $\pi \otimes 1 \in B \otimes_A R$. Then we have the following lemma.

Lemma A.3. (i) *The functor $\tilde{M}^1$ is representable by a smooth, connected, unipotent group scheme over κ. Moreover, $\tilde{M}^1$ is a closed normal subgroup of* $\operatorname{Ker} \tilde{\varphi}$.

(ii) *The quotient group scheme* $\operatorname{Ker} \tilde{\varphi} / \tilde{M}^1$ *represents the functor*

$$R \mapsto \operatorname{Ker} \tilde{\varphi}(R) / \tilde{M}^1(R)$$

by Lemma A.1 and is smooth, connected, and unipotent.

Proof. Let R be a κ-algebra. In the proof, π stands for $\pi \otimes 1 \in B \otimes_A R$. To show that $\widetilde{M}^1(R)$ is a subgroup of $\operatorname{Ker}\tilde{\varphi}(R)$, it suffices to show that the inverse $1+x'$ of $1+\pi x$ in $\operatorname{Ker}\tilde{\varphi}(R)$ is contained in $\widetilde{M}^1(R)$. From the identity

$$(1+x')(1+\pi x) = 1+x'\star\pi x = 1+(x'+\pi x+\pi x'x) = 1+0,$$

we see that x' is an element of $\pi\underline{M}'(R)$ so that $1+x'$ is an element of $\widetilde{M}^1(R)$, since $\underline{M}'(R)$ is closed under multiplication and addition which implies $x+x'x \in \underline{M}'(R)$.

Then the first sentence of (i) follows by a theorem of Lazard which is stated at the beginning of Appendix A since $\widetilde{M}^1$ is isomorphic to an affine space of dimension n^2 as an algebraic variety over κ.

To show that $\widetilde{M}^1(R)$ is a normal subgroup of $\operatorname{Ker}\tilde{\varphi}(R)$, we choose an element $1+\pi x \in \widetilde{M}^1(R)$ and $1+m \in \operatorname{Ker}\tilde{\varphi}(R)$ with $m \in \underline{M}'(R)$. Let $1+m'$ be the inverse of $1+m$ so that $(1+m')(1+m)=1$. Then we have the following identity:

$$(1+m')(1+\pi x)(1+m) = 1+m'\star\pi x\star m = 1+\pi(x+m'x+xm+m'xm).$$

Since $\underline{M}'(R)$ is closed under multiplication and addition, $x+m'x+xm+m'xm \in \underline{M}'(R)$ so that $(1+m')(1+\pi x)(1+m) \in \widetilde{M}^1(R)$.

For (ii), smoothness and connectedness are stable under quotienting by algebraic groups (Proposition 22.4 in [Knus et al. 1998]) and a quotient of a unipotent group is also a unipotent group by part (a) of the first corollary in Section 8.3 in [Waterhouse 1979]. □

This paragraph is a reproduction of [Gan and Yu 2000, 6.3.6]. Recall that there is a closed immersion $\widetilde{G} \to \widetilde{M}$. Notice that $\operatorname{Ker}\varphi$ is the kernel of the composition $\widetilde{G} \to \widetilde{M} \to \widetilde{M}/\operatorname{Ker}\tilde{\varphi}$. We define $\widetilde{G}^1$ as the kernel of the composition

$$\widetilde{G} \to \widetilde{M} \to \widetilde{M}/\widetilde{M}^1.$$

Then $\widetilde{G}^1$ is the kernel of the morphism $\operatorname{Ker}\varphi \to \operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ and, hence, is a closed normal subgroup of $\operatorname{Ker}\varphi$. The induced morphism $\operatorname{Ker}\varphi/\widetilde{G}^1 \to \operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ is a monomorphism, and thus $\operatorname{Ker}\varphi/\widetilde{G}^1$ is a closed subgroup scheme of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ by (Exposé $\mathrm{VI_B}$, Corollary 1.4.2 in [SGA 3_{I} 1970]).

Theorem A.4. *$\widetilde{G}^1$ is connected, smooth, and unipotent. Furthermore, the underlying algebraic variety of $\widetilde{G}^1$ over κ is an affine space of dimension*

$$\sum_{i<j} n_i n_j + \sum_{i \text{ odd}} \frac{n_i^2+n_i}{2} + \sum_{i \text{ even}} \frac{n_i^2-n_i}{2} + \#\{i : i \text{ is even and } L_i \text{ is of type } I\}.$$

Proof. We prove this theorem by writing out a set of equations completely defining $\widetilde{G}^1$ (after all there are so many different sets of equations defining $\widetilde{G}^1$). Let R be a κ-algebra. As explained in Remark 3.3(2), we consider the given hermitian form h as an element of $\underline{H}(R)$ and write it as a formal matrix $h = \big(\pi^i \cdot h_i\big)$ with $(\pi^i \cdot h_i)$ for the (i,i)-block and 0 for the remaining blocks. We also write h as $(f_{i,j}, a_i \cdots f_i)$. Recall that the notation $(f_{i,j}, a_i \cdots f_i)$ is defined and explained in Section 3C and explicit values of $(f_{i,j}, a_i \cdots f_i)$ for the h are given in Remark 3.3(2).

We choose an element $m = (m_{i,j}, s_i \cdots w_i) \in (\mathrm{Ker}\,\tilde{\varphi})(R)$ with a formal matrix interpretation $m = \big(\pi^{\max\{0,j-i\}} m_{i,j}\big)$, where the notation $(m_{i,j}, s_i \cdots w_i)$ is explained in Section 3B. Then $h \circ m$ is an element of $\underline{H}(R)$ and $(\mathrm{Ker}\,\varphi)(R)$ is the set of m such that $h \circ m = (f_{i,j}, a_i \cdots f_i)$. The action $h \circ m$ is explicitly described in Remark 3.5. Based on this, we need to write the matrix product $h \circ m = \sigma({}^t m) \cdot h \cdot m$ formally. To do that, we write each block of $\sigma({}^t m) \cdot h \cdot m$ as follows:

The diagonal (i,i)-block of the formal matrix product $\sigma({}^t m) \cdot h \cdot m$ is the following:

$$\pi^i \big(\sigma({}^t m_{i,i}) h_i m_{i,i} + \sigma(\pi) \cdot \sigma({}^t m_{i-1,i}) h_{i-1} m_{i-1,i} + \pi \cdot \sigma({}^t m_{i+1,i}) h_{i+1} m_{i+1,i}\big) \\ + \pi^i \big((\sigma\pi)^2 \cdot \sigma({}^t m_{i-2,i}) h_{i-2} m_{i-2,i} + \pi^2 \cdot \sigma({}^t m_{i+2,i}) h_{i+2} m_{i+2,i}\big), \quad \text{(A-1)}$$

where $0 \le i < N$.

The (i,j)-block of the formal matrix product $\sigma({}^t m) \cdot h \cdot m$, where $i < j$, is the following:

$$\pi^j \bigg(\sum_{i \le k \le j} \sigma({}^t m_{k,i}) h_k m_{k,j} + \sigma(\pi) \cdot \sigma({}^t m_{i-1,i}) h_{i-1} m_{i-1,j} + \pi \cdot \sigma({}^t m_{j+1,i}) h_{j+1} m_{j+1,j} \bigg), \quad \text{(A-2)}$$

where $0 \le i, j < N$.

Before studying $\widetilde{G}^1$, we describe the conditions for an element $m \in \widetilde{M}(R)$ as above to belong to the subgroup $\widetilde{M}^1(R)$.

(1) $m_{i,j} = \pi m'_{i,j}$ if $i \neq j$;

(2) $m_{i,i} = \mathrm{id} + \pi m'_{i,i}$ if L_i is of type II;

(3) $m_{i,i} = \begin{pmatrix} s_i & \pi y_i \\ \pi v_i & 1 + \pi z_i \end{pmatrix} = \begin{pmatrix} \mathrm{id} + \pi s'_i & \pi^2 y'_i \\ \pi^2 v'_i & 1 + \pi^2 z'_i \end{pmatrix}$ if i is even and L_i is of type I^o;

(4) $m_{i,i} = \begin{pmatrix} s_i & r_i & \pi t_i \\ \pi y_i & 1 + \pi x_i & \pi z_i \\ v_i & u_i & 1 + \pi w_i \end{pmatrix} = \begin{pmatrix} \mathrm{id} + \pi s'_i & \pi r'_i & \pi^2 t'_i \\ \pi^2 y'_i & 1 + \pi^2 x'_i & \pi^2 z'_i \\ \pi v'_i & \pi u'_i & 1 + \pi^2 w'_i \end{pmatrix}$ if i is even and L_i is of type I^e.

Here, all matrices having $'$ in the superscript are considered as matrices with entries in R. When i is even and L_i is of type I, we formally write $m_{i,i} = \mathrm{id} + \pi m'_{i,i}$. Then $\widetilde{G}^1(R)$ is the set of $m \in \widetilde{M}^1(R)$ such that $h \circ m = h = (f_{i,j}, a_i \cdots f_i)$. Since $h \circ m$ is an element of $\underline{H}(R)$, we can write $h \circ m$ as $(f'_{i,j}, a'_i \cdots f'_i)$. In what follows, we will write $(f'_{i,j}, a'_i \cdots f'_i)$ in terms of $h = (f_{i,j}, a_i \cdots f_i)$ and m, and will compare $(f'_{i,j}, a'_i \cdots f'_i)$ with $(f_{i,j}, a_i \cdots f_i)$, in order to obtain a set of equations defining $\widetilde{G}^1$.

If we put all these (1)–(4) into (A-2), then we obtain

$$\pi^j \big(\sigma(1 + \pi \cdot {}^t m'_{i,i}) h_i \pi m'_{i,j} + \sigma(\pi \cdot {}^t m'_{j,i}) h_j (1 + \pi m'_{j,j})\big).$$

Therefore,

$$f'_{i,j} = \big(\sigma(1 + \pi \cdot {}^t m'_{i,i}) h_i \pi m'_{i,j} + \sigma(\pi \cdot {}^t m'_{j,i}) h_j (1 + \pi m'_{j,j})\big),$$

where this equation is considered in $B \otimes_A R$ and π stands for $\pi \otimes 1 \in B \otimes_A R$. Thus each term having π^2 as a factor is 0 and we have

$$f'_{i,j} = h_i \pi m'_{i,j} + \sigma(\pi \cdot {}^t m'_{j,i}) h_j, \quad \text{where } i < j. \tag{A-3}$$

This equation is of the form $f'_{i,j} = X + \pi Y$ since it is an equation in $B \otimes_A R$. By letting $f'_{i,j} = f_{i,j} = 0$, we obtain

$$\bar{h}_i m'_{i,j} + {}^t m'_{j,i} \bar{h}_j = 0, \quad \text{where } i < j, \tag{A-4}$$

where $\bar{h}_i$ (resp. $\bar{h}_j$) is obtained by letting each term in h_i (resp. h_j) having π as a factor be zero so that this equation is considered in R. Note that $\bar{h}_i$ and $\bar{h}_j$ are invertible as matrices with entries in R by Remark 3.3. Thus $m'_{i,j} = \bar{h}_i^{-1} \cdot {}^t m'_{j,i} \cdot \bar{h}_j$. This induces that each entry of $m'_{i,j}$ is expressed as a linear combination of the entries of $m'_{j,i}$. Thus there are exactly $n_i n_j$ independent linear equations among the entries of $m'_{i,j}, m'_{j,i}$.

Next, we put (1)–(4) into (A-1). Then we obtain

$$\pi^i \big(\sigma(1 + \pi \cdot {}^t m'_{i,i}) h_i (1 + \pi m'_{i,i}) \big). \tag{A-5}$$

We interpret this so as to obtain equations defining $\widetilde{G}^1$. There are 4 cases, indexed by (i), (ii), (iii), (iv), according to types of L_i.

(i) Assume that i is odd. Then $\pi^i h_i = \xi^{(i-1)/2} \pi a_i$ as explained in Section 3C and thus we have

$$a'_i = \sigma(1 + \pi \cdot {}^t m'_{i,i}) a_i (1 + \pi m'_{i,i}).$$

Here, the nondiagonal entries of this equation are considered in $B \otimes_A R$ and each diagonal entry of a'_i is of the form $\epsilon \pi x_i$ with $x_i \in R$.

Thus, we can cancel terms having π^2 as a factor and the above equation equals

$$a'_i = a_i + \sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i}.$$

By letting $a'_i = a_i$, we have the following equation

$$\sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i} = 0.$$

Since this is an equation in $B \otimes_A R$, it is of the form $X + \pi Y = 0$. Note that the reduction of $\epsilon \mod \pi$ is 1. We denote by $\bar{a}_i$ the reduction of $a_i \mod \pi$. Thus we have

$${}^t m'_{i,i} \bar{a}_i + \bar{a}_i m'_{i,i} = 0.$$

This is a matrix equation over R, in a usual sense, and $\bar{a}_i$ is symmetric and the diagonal entries of $\bar{a}_i$ are 0. More precisely,

$$\bar{a}_i = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix}.$$

Then we can see that there is no contribution coming from the diagonal entries of ${}^t m'_{i,i}\bar{a}_i + \bar{a}_i m'_{i,i} = 0$ and that there are exactly $(n_i^2 - n_i)/2$ independent linear equations. Thus $(n_i^2 + n_i)/2$ entries of $m'_{i,i}$ determine all entries of $m'_{i,i}$. Note that the conditions on $m'_{i,i}$, viewed as a matrix with entries in κ, are tantamount to this matrix belonging to the Lie algebra of a symplectic group associated to an obvious alternating form given by $\bar{a}_i$. Then $(n_i^2 + n_i)/2$ is the dimension of this symplectic group.

For example, let $m'_{i,i} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ and $\bar{a}_i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then

$$ {}^t m'_{i,i}\bar{a}_i + \bar{a}_i m'_{i,i} = \begin{pmatrix} 2z & x+w \\ x+w & 2y \end{pmatrix} = \begin{pmatrix} 0 & x+w \\ x+w & 0 \end{pmatrix}. $$

Thus there is one linear equation $x + w = 0$ and x, y, z determine all entries of $m'_{i,i}$.

(ii) Assume that i is even and L_i is of type II. This case is parallel to the previous case. Then $\pi^i h_i = \xi^{i/2} a_i$ as explained in Section 3C and we have

$$ a'_i = \sigma(1 + \pi \cdot {}^t m'_{i,i}) a_i (1 + \pi m'_{i,i}). $$

Here, the nondiagonal entries of this equation are considered in $B \otimes_A R$ and each diagonal entry of a'_i is of the form $2x_i$ with $x_i \in R$. Now, the nondiagonal entries of $\sigma(\pi \cdot {}^t m'_{i,i}) a_i (\pi m'_{i,i})$ are all 0 since they contain π^2 as a factor. The diagonal entries of $\sigma(\pi \cdot {}^t m'_{i,i}) a_i (\pi m'_{i,i})$ are also 0 since they contain π^4 as a factor. Thus, the above equation equals

$$ a'_i = a_i + \sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i}. $$

By letting $a'_i = a_i$, we have the following equation

$$ \sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i} = 0. $$

Based on (2) of the description of $\underline{H}(R)$ for a κ-algebra R, which is explained in Section 3C, in order to investigate this equation, we need to consider the nondiagonal entries of $\sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i}$ as elements of $B \otimes_A R$ and the diagonal entries of $\sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i}$ as of the form $2x_i$ with $x_i \in R$. Recall from Remark 3.3 that

$$ a_i = \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix}. $$

Then we can see that each diagonal entry as well as each nondiagonal (upper triangular) entry of $\sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i}$ produces a linear equation. Thus there are exactly $(n_i^2 + n_i)/2$ independent linear equations and $(n_i^2 - n_i)/2$ entries of $m'_{i,i}$ determine all entries of $m'_{i,i}$.

For example, let $m'_{i,i} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ and $\bar{a}_i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then

$$ \sigma(\pi) \cdot {}^t m'_{i,i} a_i + \pi \cdot a_i m'_{i,i} = \sigma(\pi) \begin{pmatrix} z & x \\ w & y \end{pmatrix} + \pi \begin{pmatrix} z & w \\ x & y \end{pmatrix} = \begin{pmatrix} (\sigma(\pi) + \pi) z & \sigma(\pi) x + \pi w \\ \sigma(\pi) w + \pi x & (\sigma(\pi) + \pi) y \end{pmatrix}. $$

Recall that $\sigma(\pi) = \epsilon\pi$ with $\epsilon \equiv 1 \mod \pi$ and $\sigma(\pi) + \pi = 2$, as explained at the beginning of Section 2A. Thus there are three linear equations $z = 0$, $x + w = 0$, $y = 0$ and x determines every other entry of $m'_{i,i}$.

(iii) Assume that i is even and L_i is of type I^o. Then $\pi^i h_i = \xi^{i/2}\begin{pmatrix} a_i & \pi b_i \\ \sigma(\pi\cdot {}^t b_i) & 1+2c_i \end{pmatrix}$ as explained in Section 3C and we have

$$\begin{pmatrix} a_i' & \pi b_i' \\ \sigma(\pi\cdot {}^t b_i') & 1+2c_i' \end{pmatrix} = \sigma(1+\pi\cdot {}^t m_{i,i}')\cdot \begin{pmatrix} a_i & \pi b_i \\ \sigma(\pi\cdot {}^t b_i) & 1+2c_i \end{pmatrix}\cdot(1+\pi m_{i,i}'). \tag{A-6}$$

Here, the nondiagonal entries of a_i' as well as the entries of b_i' are considered in $B\otimes_A R$, each diagonal entry of a_i' is of the form $2x_i$ with $x_i\in R$, and c_i' is in R. In addition, $b_i=0$, $c_i=\bar{\gamma}_i$ as explained in Remark 3.3(2) and a_i is the diagonal matrix with $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on the diagonal.

Note that in this case, $m_{i,i}' = \begin{pmatrix} s_i' & \pi y_i' \\ \pi v_i' & \pi z_i' \end{pmatrix}$. Compute $\sigma(\pi\cdot {}^t m_{i,i}')\cdot\begin{pmatrix} a_i & 0 \\ 0 & 1+2c_i \end{pmatrix}\cdot(\pi m_{i,i}')$ formally and this equals $\sigma(\pi)\pi\begin{pmatrix} {}^t s_i' a_i s_i' + \pi^2 X_i & \pi Y_i \\ \sigma(\pi\cdot {}^t Y_i) & \pi^2 Z_i \end{pmatrix}$ for certain matrices X_i, Y_i, Z_i with suitable sizes. Thus we can ignore the contribution from $\sigma(\pi\cdot {}^t m_{i,i}')\begin{pmatrix} a_i & 0 \\ 0 & 1+2c_i \end{pmatrix}(\pi m_{i,i}')$ in Equation (A-6) and so Equation (A-6) equals

$$\begin{pmatrix} a_i' & \pi b_i' \\ \sigma(\pi\cdot {}^t b_i') & 1+2c_i' \end{pmatrix} = \begin{pmatrix} a_i & 0 \\ 0 & 1+2c_i \end{pmatrix} + \sigma(\pi)\begin{pmatrix} {}^t s_i' & \sigma(\pi)\cdot {}^t v_i' \\ \sigma(\pi)\cdot {}^t y_i' & \sigma(\pi) z_i' \end{pmatrix}\begin{pmatrix} a_i & 0 \\ 0 & 1+2c_i \end{pmatrix} + \pi\begin{pmatrix} a_i & 0 \\ 0 & 1+2c_i \end{pmatrix}\begin{pmatrix} s_i' & \pi y_i' \\ \pi v_i' & \pi z_i' \end{pmatrix}.$$

We interpret each block of the above equation below:

(a) Firstly, we consider the $(1,1)$-block. The computation associated to this block is similar to that for the above case (ii). Hence there are exactly $((n_i-1)^2+(n_i-1))/2$ independent linear equations and $((n_i-1)^2-(n_i-1))/2$ entries of s_i' determine all entries of s_i'.

(b) Secondly, we consider the $(1,2)$-block. We can ignore the contribution from ${}^t v_i' c_i$ since it contains π^3 as a factor. Then the $(1,2)$-block is

$$\pi b_i' = \sigma(\pi)\pi\cdot(\epsilon\cdot {}^t v_i' + 1/\epsilon\cdot a_i y_i'). \tag{A-7}$$

By letting $b_i'=b_i=0$, we have

$$\sigma(\pi)\cdot(\epsilon\cdot {}^t v_i' + 1/\epsilon\cdot a_i y_i') = 0$$

as an equation in $B\otimes_A R$. Thus there are exactly (n_i-1) independent linear equations among the entries of v_i' and y_i' and the entries of v_i' determine all entries of y_i'.

(c) Finally, we consider the $(2,2)$-block. This is

$$1+2c_i' = 1+2c_i + (\pi^2+(\sigma(\pi))^2)z_i' + 2(\pi^2+(\sigma(\pi))^2)c_i z_i'. \tag{A-8}$$

Since $\pi^2+(\sigma(\pi))^2 = (\pi+\sigma(\pi))^2 - 2\sigma(\pi)\pi$, we see that $\pi^2+(\sigma(\pi))^2$ contains 4 as a factor. Thus by letting $c_i'=c_i$, this equation is trivial.

By combining the three cases (a)–(c), there are exactly $((n_i-1)^2+(n_i-1))/2+(n_i-1) = (n_i^2+n_i)/2-1$ independent linear equations and $(n_i^2-n_i)/2+1$ entries of $m_{i,i}'$ determine all entries of $m_{i,i}'$.

(iv) Assume that i is even and L_i is of type I^e. Then

$$\pi^i h_i = \xi^{i/2}\begin{pmatrix} a_i & b_i & \pi e_i \\ \sigma({}^t b_i) & 1+2f_i & 1+\pi d_i \\ \sigma(\pi\cdot{}^t e_i) & \sigma(1+\pi d_i) & 2c_i \end{pmatrix}$$

as explained in Section 3C and we have

$$\begin{pmatrix} a_i' & b_i' & \pi e_i' \\ \sigma({}^t b_i') & 1+2f_i' & 1+\pi d_i' \\ \sigma(\pi\cdot{}^t e_i') & \sigma(1+\pi d_i') & 2c_i' \end{pmatrix}$$
$$= \sigma(1+\pi\cdot{}^t m_{i,i}')\cdot\begin{pmatrix} a_i & b_i & \pi e_i \\ \sigma({}^t b_i) & 1+2f_i & 1+\pi d_i \\ \sigma(\pi\cdot{}^t e_i) & \sigma(1+\pi d_i) & 2c_i \end{pmatrix}\cdot(1+\pi m_{i,i}'). \quad \text{(A-9)}$$

Here, the nondiagonal entries of a_i' as well as the entries of b_i', e_i', d_i' are considered in $B\otimes_A R$, each diagonal entry of a_i' is of the form $2x_i$ with $x_i \in R$, and c_i', f_i' are in R. In addition, $b_i = 0$, $d_i = 0$, $e_i = 0$, $f_i = 0$, $c_i = \bar{\gamma}_i$ as explained in Remark 3.3(2) and a_i is the diagonal matrix with $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on the diagonal.

Notice that in this case,

$$m_{i,i}' = \begin{pmatrix} s_i' & r_i' & \pi t_i' \\ \pi y_i' & \pi x_i' & \pi z_i' \\ v_i' & u_i' & \pi w_i' \end{pmatrix}.$$

We compute

$$\sigma(\pi\cdot{}^t m_{i,i}')\cdot\begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i \end{pmatrix}\cdot(\pi m_{i,i}')$$

formally and this equals

$$\sigma(\pi)\pi\begin{pmatrix} {}^t s_i' a_i s_i' + \pi^2 X_i & Y_i & \pi Z_i \\ \sigma({}^t Y_i) & {}^t r_i' a_i r_i' + \pi^2 X_i' & \pi Y_i' \\ \sigma(\pi\cdot{}^t Z_i) & \sigma(\pi\cdot{}^t Y_i') & \pi^2 Z_i' \end{pmatrix}$$

for certain matrices $X_i, Y_i, Z_i, X_i', Y_i', Z_i'$ with suitable sizes. Thus we can ignore the contribution from this part in Equation (A-9) and so Equation (A-9) equals

$$\begin{pmatrix} a_i' & b_i' & \pi e_i' \\ \sigma({}^t b_i') & 1+2f_i' & 1+\pi d_i' \\ \sigma(\pi\cdot{}^t e_i') & \sigma(1+\pi d_i') & 2c_i' \end{pmatrix} = \begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i \end{pmatrix}$$
$$+\sigma(\pi)\begin{pmatrix} {}^t s_i' & \sigma(\pi)\cdot{}^t y_i' & {}^t v_i' \\ {}^t r_i' & \sigma(\pi) x_i' & u_i' \\ \sigma(\pi)\cdot{}^t t_i' & \sigma(\pi) z_i' & \sigma(\pi) w_i' \end{pmatrix}\begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i \end{pmatrix}$$
$$+\pi\begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i \end{pmatrix}\begin{pmatrix} s_i' & r_i' & \pi t_i' \\ \pi y_i' & \pi x_i' & \pi z_i' \\ v_i' & u_i' & \pi w_i' \end{pmatrix}.$$

We interpret each block of the above equation as follows:

(a) Let us consider the $(1,1)$-block. The computation associated to this block is similar to that for the previous case (ii). Hence there are exactly $((n_i-2)^2+(n_i-2))/2$ independent linear equations and $((n_i-2)^2-(n_i-2))/2$ entries of s_i' determine all entries of s_i'.

(b) We consider the $(1,2)$-block. This gives

$$b_i' = \pi(\epsilon^2\pi\cdot {}^t y_i' + \epsilon\cdot {}^t v_i' + a_i r_i'). \tag{A-10}$$

This is an equation in $B\otimes_A R$. By letting $b_i'=b_i=0$, there are exactly (n_i-2) independent linear equations among the entries of v_i', r_i'.

(c) The $(1,3)$-block is

$$\pi e_i' = \pi^2(\epsilon^2\cdot {}^t y_i' + (2/\pi)\cdot\epsilon\cdot {}^t v_i' c_i + a_i t_i').$$

By letting $e_i'=e_i=0$, we have

$$\begin{aligned} e_i' &= \pi(\epsilon^2\cdot {}^t y_i' + (2/\pi)\cdot\epsilon\cdot {}^t v_i' c_i + a_i t_i') \\ &= \pi(\epsilon^2\cdot {}^t y_i' + a_i t_i') = \pi({}^t y_i' + a_i t_i') = 0. \end{aligned} \tag{A-11}$$

This is an equation in $B\otimes_A R$. Thus there are exactly (n_i-2) independent linear equations among the entries of y_i', t_i'.

(d) The $(2,3)$-block is

$$1+\pi d_i' = 1+\sigma(\pi)(\sigma(\pi)x_i' + 2u_i' c_i) + \pi^2(z_i'+w_i').$$

By letting $d_i'=d_i=0$, we have

$$d_i' = \pi(\epsilon^2 x_i' + z_i' + w_i') = \pi(x_i'+z_i'+w_i') = 0. \tag{A-12}$$

This is an equation in $B\otimes_A R$. Thus there is exactly one independent linear equation among the entries of x_i', z_i', w_i'.

(e) The $(2,2)$-block is

$$\begin{aligned} 1+2f_i' &= 1+\sigma(\pi)(\sigma(\pi)x_i' + u_i') + \pi(\pi x_i' + u_i') \\ &= 1+2u_i' + ((\pi+\sigma(\pi))^2 - 2\pi\sigma(\pi))x_i'. \end{aligned}$$

By letting $f_i'=f_i=0$, we have

$$f_i' = u_i' + ((\pi+\sigma(\pi)) - \pi\sigma(\pi))x_i' = u_i' = 0.$$

This is an equation in R. Thus $u_i'=0$ is the only independent linear equation.

(f) The $(3,3)$-block is

$$\begin{aligned} 2c_i' &= 2c_i + \sigma(\pi)(\sigma(\pi)z_i' + 2\sigma(\pi)w_i' c_i) + \pi(\pi z_i' + 2\pi w_i' c_i) \\ &= 2c_i + ((\pi+\sigma(\pi))^2 - 2\pi\sigma(\pi))(z_i' + 2w_i' c_i). \end{aligned} \tag{A-13}$$

Since $((\pi+\sigma(\pi))^2 - 2\pi\sigma(\pi))$ contains 4 as a factor, by letting $c_i'=c_i$, this equation is trivial.

By combining the six cases (a)–(f), there are exactly $((n_i-2)^2+(n_i-2))/2+2(n_i-2)+2 = (n_i^2+n_i)/2-1$ independent linear equations and $(n_i^2-n_i)/2+1$ entries of $m_{i,i}'$ determine all entries of $m_{i,i}'$.

We now combine all the work done in this proof. Namely, we collect the above (i), (ii), (iii), (iv) which are the interpretations of Equation (A-5), together with Equation (A-4). Then there are exactly

$$\sum_{i<j} n_i n_j + \sum_{i \text{ odd}} \frac{n_i^2 - n_i}{2} + \sum_{i \text{ even}} \frac{n_i^2 + n_i}{2} - \#\{i : i \text{ is even and } L_i \text{ is of type } I\}$$

independent linear equations among the entries of m. Furthermore, all coefficients of these equations are in κ. Therefore, we consider $\widetilde{G}^1$ as a subvariety of $\widetilde{M}^1$ determined by these linear equations. Since $\widetilde{M}^1$ is an affine space of dimension n^2, the underlying algebraic variety of $\widetilde{G}^1$ over κ is an affine space of dimension

$$\sum_{i<j} n_i n_j + \sum_{i \text{ odd}} \frac{n_i^2 + n_i}{2} + \sum_{i \text{ even}} \frac{n_i^2 - n_i}{2} + \#\{i : i \text{ is even and } L_i \text{ is of type } I\}.$$

This completes the proof by using a theorem of Lazard which is stated at the beginning of Appendix A. □

Let R be a κ-algebra. We describe the functor of points of the scheme $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ by using points of the scheme $(\underline{M}' \otimes \kappa)/\pi\underline{M}'$, based on Lemma A.3. Recall from two paragraphs before Lemma A.3 that $(1+)^{-1}(\underline{M}^*)$, which is an open subscheme of $\underline{M}'$, is a group scheme with the operation $\star$. Let $\widetilde{M}'$ be the special fiber of $(1+)^{-1}(\underline{M}^*)$. Since $\widetilde{M}^1$ is a closed normal subgroup of $\widetilde{M}(=\underline{M}^* \otimes \kappa)$ (cf. Lemma A.3(i)), $\pi\underline{M}'$, which is the inverse image of $\widetilde{M}^1$ under the isomorphism $1+$, is a closed normal subgroup of $\widetilde{M}'$. Therefore, the morphism $1+$ induces the following isomorphism of group schemes, which is also denoted by $1+$,

$$1+ : \widetilde{M}'/\pi\underline{M}' \longrightarrow \widetilde{M}/\widetilde{M}^1.$$

Note that $\widetilde{M}'/\pi\underline{M}'(R) = \widetilde{M}'(R)/\pi\underline{M}'(R)$ by Lemma A.1. Each element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ is therefore uniquely written as $1+\bar{x}$, where $\bar{x} \in \widetilde{M}'(R)/\pi\underline{M}'(R)$. Here, by $1+\bar{x}$, we mean the image of $\bar{x}$ under the morphism $1+$ at the level of R-points.

We still need a better description of an element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ by using a point of the scheme $(\underline{M}' \otimes \kappa)/\pi\underline{M}'$. Note that $(\underline{M}' \otimes \kappa)/\pi\underline{M}'$ is a quotient of group schemes with respect to the addition, whereas $\widetilde{M}'/\pi\underline{M}'$ is a quotient of group schemes with respect to the operation $\star$.

We claim that the open immersion $\iota : \widetilde{M}' \to \underline{M}' \otimes \kappa$ with $x \mapsto x$ induces a monomorphism of schemes

$$\bar{\iota} : \widetilde{M}'/\pi\underline{M}' \to (\underline{M}' \otimes \kappa)/\pi\underline{M}'.$$

Choose $x \in \widetilde{M}'(R)$ and $\pi y \in \pi\underline{M}'(R)$ for a κ-algebra R. Since $x \star \pi y = x + \pi(y + xy)$, both x and $x \star \pi y$ give the same element in $((\underline{M}' \otimes \kappa)/\pi\underline{M}')(R)$. Thus the morphism $\bar{\iota}$ is well-defined.

In order to show that $\bar{\iota}$ is a monomorphism, choose $x, y \in \widetilde{M}'(R)$ such that $x = y + \pi z$ with $\pi z \in \pi\underline{M}'(R)$. Let y' $(\in \widetilde{M}'(R))$ be the inverse of y so that $y \star y' = y + y' + yy' = 0$. Then $\pi(z + y'z)$ is an element of $\pi\underline{M}'(R)$. We have the following identity:

$$x \star \pi(z + y'z) = (y + \pi z) \star \pi(z + y'z) = y + \pi(y + y' + yy')z = y.$$

Therefore, x and y give the same element in $(\widetilde{M}'/\pi\underline{M'})(R)$, which shows the injectivity of the above morphism.

Note that the operation $\star$ is closed in $\underline{M}'\otimes\kappa$ as mentioned in the third paragraph following Lemma A.2. We can also easily check that the operation $\star$ is well-defined on $(\underline{M}'\otimes\kappa)/\pi\underline{M}'$, which turns to be a scheme of monoids with respect to $\star$, and that the morphism $\bar{\iota}$ is a monomorphism of monoid schemes.

To summarize, the morphism $1+ : \widetilde{M}'/\pi\underline{M'} \longrightarrow \widetilde{M}/\widetilde{M}^1$ is an isomorphism of group schemes and the morphism $\bar{\iota} : \widetilde{M}'/\pi\underline{M}' \to (\underline{M}'\otimes\kappa)/\pi\underline{M}'$ is a monomorphism preserving the operation $\star$. Therefore, each element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ is uniquely written as $1+\bar{x}$, where $\bar{x}\in(\underline{M}'\otimes\kappa)(R)/\pi\underline{M}'(R)$. Here, by $1+\bar{x}$, we mean $(1+)\circ\bar{\iota}^{-1}(\bar{x})$. From now on to the end of this paper, we keep the notation $1+\bar{x}$ to express an element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ such that $\bar{x}$ is an element of $(\underline{M}'\otimes\kappa)(R)/\pi\underline{M}'(R)$ which is a quotient of R-valued points of group schemes with respect to addition. Then the product of two elements $1+\bar{x}$ and $1+\bar{y}$ is the same as $1+\bar{x}\star\bar{y}$ $(=1+(\bar{x}+\bar{y}+\bar{x}\bar{y}))$.

Remark A.5. By the above argument, we write an element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ formally as $m=\left(\pi^{\max\{0,j-i\}}m_{i,j}\right)$ with $s_i,\cdots,w_i$ as in Section 3B such that each entry of each of the matrices $(m_{i,j})_{i\neq j}, s_i,\cdots,w_i$ is in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)\cong R$. In particular, based on the description of $\operatorname{Ker}\tilde{\varphi}(R)$ given at the paragraph following Lemma A.2, we have the following conditions on m:

(1) Assume that i is even and L_i is of type I. Then $s_i=\mathrm{id}$.

(2) $m_{i,i}=\mathrm{id}$ if L_i is of type II.

(3) Assume that i is odd. Then $\delta_{i-1}e_{i-1}\cdot m_{i-1,i}+\delta_{i+1}e_{i+1}\cdot m_{i+1,i}=0$. Here, δ_j, e_j are as explained in the description of $\operatorname{Ker}\tilde{\varphi}(R)$.

Theorem A.6. *$\operatorname{Ker}\varphi/\widetilde{G}^1$ is isomorphic to $\mathbb{A}^{l'}\times(\mathbb{Z}/2\mathbb{Z})^\beta$ as a κ-variety, where $\mathbb{A}^{l'}$ is an affine space of dimension l'. Here,*

- *l' is such that $l'+\dim\widetilde{G}^1=l$. Notice that l is defined in Lemma 4.6 and that the dimension of $\widetilde{G}^1$ is given in Theorem A.4.*
- *β is the number of even integers j such that L_j is of type I and L_{j+2} is of type II.*

Proof. Lemma A.1 and Theorem A.4 imply that $\operatorname{Ker}\varphi/\widetilde{G}^1$ represents the functor $R\mapsto \operatorname{Ker}\varphi(R)/\widetilde{G}^1(R)$. Recall that $\operatorname{Ker}\varphi/\widetilde{G}^1$ is a closed subgroup scheme of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ as explained at the paragraph just before Theorem A.4. Let $m=\left(\pi^{\max\{0,j-i\}}m_{i,j}\right)$ be an element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ such that m belongs to $(\operatorname{Ker}\varphi/\widetilde{G}^1)(R)$. We want to find equations which m satisfies. Note that the entries of m involve $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ as explained in Remark A.5.

Recall that h is the fixed hermitian form and we consider it as an element in $\underline{H}(R)$ as explained in Remark 3.3(2). We write it as a formal matrix $h=\left(\pi^i\cdot h_i\right)$ with $(\pi^i\cdot h_i)$ for the (i,i)-block and 0 for the remaining blocks. We choose a representative $1+x\in\operatorname{Ker}\varphi(R)$ of m so that $h\circ(1+x)=h$. Any other representative of m in $\operatorname{Ker}\tilde{\varphi}(R)$ is of the form $(1+x)(1+\pi y)$ with $y\in\underline{M}'(R)$ and we have $h\circ(1+x)(1+\pi y)=h\circ(1+\pi y)$. Notice that $h\circ(1+\pi y)$ is an element of $\underline{H}(R)$ so we express it as $(f'_{i,j}, a'_i\cdots f'_i)$. We also let

$h=(f_{i,j},a_i\cdots f_i)$. Here, we follow notation from Section 3C, the paragraph just before Remark 3.3. Recall that $h=(f_{i,j},a_i\cdots f_i)$ is described explicitly in Remark 3.3(2). Now, $1+\pi y$ is an element of $\widetilde{M}^1(R)$ and so we can use our result (Equations (A-3), (A-7), (A-8), (A-10), (A-11), (A-12), (A-13)) stated in the proof of Theorem A.4 in order to compute $h\circ(1+\pi y)$. Based on this, we enumerate equations which m satisfies as follows:

(1) Assume $i<j$. By Equation (A-3) which involves an element of $\widetilde{M}^1(R)$, each entry of $f'_{i,j}$ has π as a factor so that $f'_{i,j}\equiv f_{i,j}(=0) \mod (\pi\otimes 1)(B\otimes_A R)$. In other words, the (i,j)-block of $h\circ(1+x)(1+\pi y)$ divided by $\pi^{\max\{i,j\}}$ is $f_{i,j}(=0)$ modulo $(\pi\otimes 1)(B\otimes_A R)$, which is independent of the choice of $1+\pi y$. Let $\tilde{m}\in \operatorname{Ker}\tilde{\varphi}(R)$ be a lift of m. Therefore, if we write the (i,j)-block of $\sigma({}^t\tilde{m})\cdot h\cdot\tilde{m}$ as $\pi^{\max\{i,j\}}\mathcal{X}_{i,j}(\tilde{m})$, where $\mathcal{X}_{i,j}(\tilde{m})\in M_{n_i\times n_j}(B\otimes_A R)$, then the image of $\mathcal{X}_{i,j}(\tilde{m})$ in $M_{n_i\times n_j}(B\otimes_A R)/(\pi\otimes 1)M_{n_i\times n_j}(B\otimes_A R)\cong M_{n_i\times n_j}(R)$ is independent of the choice of the lift $\tilde{m}$ of m. Therefore, we may denote this image by $\mathcal{X}_{i,j}(m)$. On the other hand, by Equation (A-2), we have the following identity:

$$\mathcal{X}_{i,j}(m)=\sum_{i\le k\le j}\sigma({}^t m_{k,i})\bar{h}_k m_{k,j}\ \text{ if }\ i<j. \tag{A-14}$$

We explain how to interpret the above equation. We know that $\mathcal{X}_{i,j}(m)$ and $m_{k,k'}$ (with $k\neq k'$) are matrices with entries in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$, whereas $m_{i,i}$ and $m_{j,j}$ are formal matrices as explained in Remark A.5. Thus we consider $\bar{h}_k$, $m_{i,i}$, and $m_{j,j}$ as matrices with entries in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ by letting π be zero in each entry of the formal matrices h_k, $m_{i,i}$, and $m_{j,j}$. Here we keep using $m_{i,i}$ and $m_{j,j}$ for matrices with entries in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ in the above equation in order to simplify notation. Later in Equation (A-23), they are denoted by $\bar{m}_{i,i}$ and $\bar{m}_{j,j}$. Then the right hand side is computed as a sum of products of matrices (involving the usual matrix addition and multiplication) with entries in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Thus, the assignment $m\mapsto\mathcal{X}_{i,j}(m)$ is polynomial in m. Furthermore, since m actually belongs to $\operatorname{Ker}\varphi(R)/\widetilde{G}^1(R)$, we have the following equation by the argument made at the beginning of this paragraph:

$$\mathcal{X}_{i,j}(m)=f_{i,j} \mod (\pi\otimes 1)(B\otimes_A R)=0.$$

Thus we get an $n_i\times n_j$ matrix $\mathcal{X}_{i,j}$ of polynomials on $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ defined by Equation (A-14), vanishing on the subscheme $\operatorname{Ker}\varphi/\widetilde{G}^1$.

Before moving to the following steps, we fix notation. Let m be an element in $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ and $\tilde{m}\in\operatorname{Ker}\tilde{\varphi}(R)$ be its lift. For any block x_i of m, $\tilde{x}_i$ is denoted by the corresponding block of $\tilde{m}$ whose reduction is x_i. Since x_i is a block of an element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$, it involves $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ as explained in Remark A.5, whereas $\tilde{x}_i$ involves $B\otimes_A R$. In addition, for a block a_i of h, $\bar{a}_i$ is denoted by the image of a_i in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$.

(2) Assume that i is even and L_i is of type I^o. By Equation (A-7) which involves an element of $\widetilde{M}^1(R)$, each entry of b'_i has π as a factor so that $b'_i\equiv b_i=0 \mod (\pi\otimes 1)(B\otimes_A R)$. Let $\tilde{m}\in\operatorname{Ker}\tilde{\varphi}(R)$ be a lift of m. By using an argument similar to the paragraph just before Equation (A-14) of step (1), if we write the $(1,2)$-block of the (i,i)-block of the formal matrix product $\sigma({}^t\tilde{m})\cdot h\cdot\tilde{m}$ as $\xi^{i/2}\cdot\pi\mathcal{X}_{i,1,2}(\tilde{m})$, where $\mathcal{X}_{i,1,2}(\tilde{m})\in M_{(n_i-1)\times 1}(B\otimes_A R)$, then the image of $\mathcal{X}_{i,1,2}(\tilde{m})$ in $M_{(n_i-1)\times 1}(B\otimes_A R)/(\pi\otimes 1)M_{(n_i-1)\times 1}(B\otimes_A R)$ is independent

of the choice of the lift $\tilde{m}$ of m. Therefore, we may denote this image by $\mathcal{X}_{i,1,2}(m)$. As for Equation (A-14) of step (1), we need to express $\mathcal{X}_{i,1,2}(m)$ as matrices. Recall that $\pi^i h_i = \xi^{i/2}\left(\begin{smallmatrix} a_i & 0 \\ 0 & 1+2c_i \end{smallmatrix}\right) = \pi^i \cdot \epsilon^{i/2}\left(\begin{smallmatrix} a_i & 0 \\ 0 & 1+2c_i \end{smallmatrix}\right)$ and $\epsilon \equiv 1 \mod \pi \otimes 1$. We write $m_{i,i}$ as $\left(\begin{smallmatrix} \mathrm{id} & \pi y_i \\ \pi v_i & 1+\pi z_i \end{smallmatrix}\right)$ and $\tilde{m}_{i,i}$ as $\left(\begin{smallmatrix} \tilde{s}_i & \pi \tilde{y}_i \\ \pi \tilde{v}_i & 1+\pi \tilde{z}_i \end{smallmatrix}\right)$ such that $\tilde{s}_i = \mathrm{id} \mod \pi \otimes 1$. Then

$$\sigma({}^t\tilde{m}_{i,i}) h_i \tilde{m}_{i,i} = \epsilon^{i/2} \begin{pmatrix} \sigma({}^t\tilde{s}_i) & \sigma(\pi \cdot {}^t\tilde{v}_i) \\ \sigma(\pi \cdot {}^t\tilde{y}_i) & 1+\sigma(\pi \tilde{z}_i) \end{pmatrix} \begin{pmatrix} a_i & 0 \\ 0 & 1+2c_i \end{pmatrix} \begin{pmatrix} \tilde{s}_i & \pi \tilde{y}_i \\ \pi \tilde{v}_i & 1+\pi \tilde{z}_i \end{pmatrix}. \tag{A-15}$$

Then the $(1,2)$-block of $\sigma({}^t\tilde{m}_{i,i}) h_i \tilde{m}_{i,i}$ is $\epsilon^{i/2}\pi(a_i \tilde{y}_i + \epsilon\sigma({}^t\tilde{v}_i)) + \pi^2(*)$ for a certain polynomial $(*)$. Therefore, by observing the $(1,2)$-block of Equation (A-1), we have

$$\mathcal{X}_{i,1,2}(m) = \bar{a}_i y_i + {}^t v_i + \mathcal{P}^i_{1,2}.$$

Here, $\mathcal{P}^i_{1,2}$ is a polynomial with variables in the entries of $m_{i-1,i}, m_{i+1,i}$. Note that this is an equation in $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R)$. Thus ϵ, which is appeared in the $(1,2)$-block of $\sigma({}^t\tilde{m}_{i,i}) h_i \tilde{m}_{i,i}$, has been ignored since $\epsilon \equiv 1 \mod \pi \otimes 1$. Furthermore, since m actually belongs to $\mathrm{Ker}\,\varphi(R)/\widetilde{G}^1(R)$, we have the following equation by the argument made at the beginning of this paragraph:

$$\mathcal{X}_{i,1,2}(m) = \bar{a}_i y_i + {}^t v_i + \mathcal{P}^i_{1,2} = \bar{b}_i = 0. \tag{A-16}$$

Thus we get polynomials $\mathcal{X}_{i,1,2}$ on $\mathrm{Ker}\,\tilde{\varphi}/\widetilde{M}^1$, vanishing on the subscheme $\mathrm{Ker}\,\varphi/\widetilde{G}^1$.

(3) Assume that i is even and L_i is of type I^e. The argument used in this step is similar to that of step (2) above. By Equations (A-10), (A-11) and (A-12), which involve an element of $\widetilde{M}^1(R)$, each entry of b_i', e_i', d_i' has π as a factor so that $b_i' \equiv b_i = 0$, $e_i' \equiv e_i = 0$, $d_i' \equiv d_i = 0 \mod (\pi \otimes 1)(B \otimes_A R)$. Let $\tilde{m} \in \mathrm{Ker}\,\tilde{\varphi}(R)$ be a lift of m. By using an argument similar to the paragraph just before Equation (A-14) of step (1), if we write the $(1,2)$, $(1,3)$, $(2,3)$-blocks of the (i,i)-block of the formal matrix product $\sigma({}^t\tilde{m}) \cdot h \cdot \tilde{m}$ as $\xi^{i/2} \cdot \pi \mathcal{X}_{i,1,2}(\tilde{m})$, $\xi^{i/2} \cdot \pi \mathcal{X}_{i,1,3}(\tilde{m})$, $\xi^{i/2} \cdot \pi \mathcal{X}_{i,2,3}(\tilde{m})$, respectively, where $\mathcal{X}_{i,1,2}(\tilde{m})$ and $\mathcal{X}_{i,1,3}(\tilde{m}) \in M_{(n_i-2)\times 1}(B \otimes_A R)$ and $\mathcal{X}_{i,2,3}(\tilde{m}) \in B \otimes_A R$, then the images of $\mathcal{X}_{i,1,2}(\tilde{m})$ and $\mathcal{X}_{i,1,3}(\tilde{m})$ in $M_{(n_i-2)\times 1}(B \otimes_A R)/(\pi \otimes 1)M_{(n_i-2)\times 1}(B \otimes_A R)$ and the image of $\mathcal{X}_{i,2,3}(\tilde{m})$ in $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R)$ are independent of the choice of the lift $\tilde{m}$ of m. Therefore, we may denote these images by $\mathcal{X}_{i,1,2}(m)$, $\mathcal{X}_{i,1,3}(m)$, and $\mathcal{X}_{i,2,3}(m)$, respectively. As for Equation (A-14) of step (1), we need to express $\mathcal{X}_{i,1,2}(m)$, $\mathcal{X}_{i,1,3}(m)$, and $\mathcal{X}_{i,2,3}(m)$ as matrices. Recall that

$$\pi^i h_i = \xi^{i/2} \begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i \end{pmatrix} = \pi^i \cdot \epsilon^{i/2} \begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i \end{pmatrix}$$

and $\epsilon \equiv 1 \mod \pi \otimes 1$. We write

$$m_{i,i} = \begin{pmatrix} id & r_i & \pi t_i \\ \pi y_i & 1+\pi x_i & \pi z_i \\ v_i & u_i & 1+\pi w_i \end{pmatrix} \quad \text{and} \quad \tilde{m}_{i,i} = \begin{pmatrix} \tilde{s}_i & \tilde{r}_i & \pi \tilde{t}_i \\ \pi \tilde{y}_i & 1+\pi \tilde{x}_i & \pi \tilde{z}_i \\ \tilde{v}_i & \tilde{u}_i & 1+\pi \tilde{w}_i \end{pmatrix}$$

such that $\tilde{s}_i = \mathrm{id} \mod \pi \otimes 1$. Then

$$\sigma({}^t\tilde{m}_{i,i})h_i\tilde{m}_{i,i} = \epsilon^{i/2}\begin{pmatrix} \sigma({}^t\tilde{s}_i) & \sigma(\pi\cdot{}^t\tilde{y}_i) & \sigma({}^t\tilde{v}_i) \\ \sigma({}^t\tilde{r}_i) & 1+\sigma(\pi\tilde{x}_i) & \sigma(\tilde{u}_i) \\ \sigma(\pi\cdot{}^t\tilde{t}_i) & \sigma(\pi\cdot{}^t\tilde{z}_i) & 1+\sigma(\pi\tilde{w}_i)\end{pmatrix} \begin{pmatrix} a_i & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2c_i\end{pmatrix}\begin{pmatrix} \tilde{s}_i & \tilde{r}_i & \pi\tilde{t}_i \\ \pi\tilde{y}_i & 1+\pi\tilde{x}_i & \pi\tilde{z}_i \\ \tilde{v}_i & \tilde{u}_i & 1+\pi\tilde{w}_i\end{pmatrix}. \qquad \text{(A-17)}$$

Then the $(1,2)$-block of $\sigma({}^t\tilde{m}_{i,i})h_i\tilde{m}_{i,i}$ is $\epsilon^{i/2}(a_i\tilde{r}_i + \sigma({}^t\tilde{v}_i)) + \pi(*)$, the $(1,3)$-block is $\epsilon^{i/2}\pi(a_i\tilde{t}_i + \epsilon\sigma({}^t\tilde{y}_i) + \sigma({}^t\tilde{v}_i)\tilde{z}_i) + \pi^2(**)$, and the $(2,3)$-block is $\epsilon^{i/2}(1+\pi(\sigma({}^t\tilde{r}_i)a_i\tilde{t}_i + \epsilon\sigma(\tilde{x}_i)+\tilde{z}_i+\tilde{w}_i+\sigma(\tilde{u}_i)\tilde{z}_i)+\pi^2(***))$ for certain polynomials $(*)$, $(**)$, $(***)$. Therefore, by considering the $(1,2)$, $(1,3)$, $(2,3)$-blocks of Equation (A-1) again, we have

$$\begin{cases} \mathcal{X}_{i,1,2}(m) = \bar{a}_i r_i + {}^t v_i; \\ \mathcal{X}_{i,1,3}(m) = \bar{a}_i t_i + {}^t y_i + {}^t v_i z_i + \mathcal{P}^i_{1,3}; \\ \mathcal{X}_{i,2,3}(m) = {}^t r_i\bar{a}_i t_i + x_i + z_i + w_i + u_i z_i + \mathcal{P}^i_{2,3}. \end{cases}$$

Here, $\mathcal{P}^i_{1,3}, \mathcal{P}^i_{2,3}$ are suitable polynomials with variables in the entries of $m_{i-1,i}, m_{i+1,i}$. These equations are considered in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Since m actually belongs to $\operatorname{Ker}\varphi(R)/\widetilde{G}^1(R)$, we have the following equation by the argument made at the beginning of this paragraph:

$$\begin{cases} \mathcal{X}_{i,1,2}(m) = \bar{a}_i r_i + {}^t v_i = \bar{b}_i = 0; \\ \mathcal{X}_{i,1,3}(m) = \bar{a}_i t_i + {}^t y_i + {}^t v_i z_i + \mathcal{P}^i_{1,3} = \bar{e}_i = 0; \\ \mathcal{X}_{i,2,3}(m) = {}^t r_i\bar{a}_i t_i + x_i + z_i + w_i + u_i z_i + \mathcal{P}^i_{2,3} = \bar{d}_i = 0. \end{cases} \qquad \text{(A-18)}$$

Thus we get polynomials $\mathcal{X}_{i,1,2}$, $\mathcal{X}_{i,1,3}$, $\mathcal{X}_{i,2,3}$ on $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$, vanishing on the subscheme $\operatorname{Ker}\varphi/\widetilde{G}^1$.

(4) Assume that i is even and L_i is of type I. By Equations (A-8) and (A-13) which involve an element of $\widetilde{M}^1(R)$, $c_i' \equiv c_i = 0 \mod (\pi\otimes 1)(B\otimes_A R)$. Let $\tilde{m} \in \operatorname{Ker}\tilde{\varphi}(R)$ be a lift of m. By using an argument similar to the paragraph just before Equation (A-14) of step (1), if we write the $(2,2)$-block (when L_i is of type I^o) or the $(3,3)$-block (when L_i is of type I^e) of the (i,i)-block of $h\circ\tilde{m} = \sigma({}^t\tilde{m})\cdot h\cdot\tilde{m}$ as $\xi^{i/2}\cdot(1+2\mathcal{X}_{i,i}(\tilde{m}))$ or $\xi^{i/2}\cdot(2\mathcal{X}_{i,i}(\tilde{m}))$ respectively, where $\mathcal{X}_{i,i}(\tilde{m}) \in B\otimes_A R$, then the image of $\mathcal{X}_{i,i}(\tilde{m})$ in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ is independent of the choice of the lift $\tilde{m}$ of m. Therefore, we may denote this image by $\mathcal{X}_{i,i}(m)$. As in Equation (A-14) of step (1), we need to express $\mathcal{X}_{i,i}(m)$ as matrices. By considering Equations (A-15) and (A-17), the $(2,2)$-block (when L_i is of type I^o) or the $(3,3)$-block (when L_i is of type I^e) of the formal matrix product $\sigma({}^t\tilde{m}_{i,i})h_i\tilde{m}_{i,i}$ is ($\epsilon^{i/2}$ if L_i is of type I^o) $+\epsilon^{i/2}(2c_i + (\pi+\sigma(\pi))\tilde{z}_i + \pi\sigma(\pi)\tilde{z}_i^2) + 4(*)$ for a certain polynomial $(*)$. Therefore, by considering the $(2,2)$-block (when L_i is of type I^o) or the $(3,3)$-block (when L_i is of type I^e) of Equation (A-1) again, we have

$$\begin{aligned}\mathcal{X}_{i,i}(\tilde{m}) = \frac{1}{\pi^2}\big((\pi+\sigma(\pi))\tilde{z}_i + \pi\sigma(\pi)\tilde{z}_i^2 + \sigma({}^t\tilde{m}'_{i-1,i})\cdot\sigma(\pi)h_{i-1}\cdot\tilde{m}'_{i-1,i} \\ + \sigma({}^t\tilde{m}'_{i+1,i})\cdot\pi h_{i+1}\cdot\tilde{m}'_{i+1,i} + \sigma({}^t\tilde{m}'_{i-2,i})\cdot\sigma(\pi)^2 h_{i-2}\cdot\tilde{m}'_{i-2,i} \\ + \sigma({}^t\tilde{m}'_{i+2,i})\cdot\pi^2 h_{i+2}\cdot\tilde{m}'_{i+2,i}\big).\end{aligned}$$

Here, $\tilde{m}'_{j,i}$ is the last column vector of the matrix $\tilde{m}_{j,i}$. Note that the right hand side is a formal polynomial with entries in $\tilde{m}$. This equation should be interpreted as follows. We formally compute the right hand side and then it is of the form $1/\pi^2(\pi^2X)$. The left hand side $\mathcal{X}_{i,i}(\tilde{m})$ is defined as the modified X by letting each term having π^2 as a factor in X be zero. It is a polynomial with entries in $B\otimes_A R$. Furthermore, $\mathcal{X}_{i,i}(m)$ is the image of $\mathcal{X}_{i,i}(\tilde{m})$ in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Let α be the unit in B such that $\epsilon=1+\alpha\pi$, as explained in Section 2A. Then $(\pi+\sigma(\pi))z_i+\pi\sigma(\pi)z_i^2=(2+\alpha\pi)\pi z_i+(1+\alpha\pi)\pi^2z_i^2$ and so $\mathcal{X}_{i,i}(\tilde{m})$ is written as follows:

$$\begin{aligned}\mathcal{X}_{i,i}(\tilde{m})=\frac{1}{\pi^2}\big(&\alpha\pi^2\tilde{z}_i+\pi^2\tilde{z}_i^2+\sigma({}^t\tilde{m}'_{i-1,i})\cdot\sigma(\pi)h_{i-1}\cdot\tilde{m}'_{i-1,i}+\sigma({}^t\tilde{m}'_{i+1,i})\cdot\pi h_{i+1}\cdot\tilde{m}'_{i+1,i}\\&+\sigma({}^t\tilde{m}'_{i-2,i})\cdot\sigma(\pi)^2h_{i-2}\cdot\tilde{m}'_{i-2,i}+\sigma({}^t\tilde{m}'_{i+2,i})\cdot\pi^2h_{i+2}\cdot\tilde{m}'_{i+2,i}\big).\end{aligned}$$

We can then write $\mathcal{X}_{i,i}(m)$ by using m and $\tilde{m}$ as follows:

$$\begin{aligned}\mathcal{X}_{i,i}(m)=\ &(\bar{\alpha}z_i+z_i^2+{}^tm'_{i-2,i}\cdot\bar{h}_{i-2}\cdot m'_{i-2,i}+m'_{i+2,i}\cdot\bar{h}_{i+2}\cdot m'_{i+2,i})\\&+\frac{1}{\pi^2}\big(\sigma({}^t\tilde{m}'_{i-1,i})\cdot\sigma(\pi)h_{i-1}\cdot\tilde{m}'_{i-1,i}+\sigma({}^t\tilde{m}'_{i+1,i})\cdot\pi h_{i+1}\cdot\tilde{m}'_{i+1,i}\big).\end{aligned}\tag{A-19}$$

Here, $\bar{\alpha}$ is the image of α in κ and $m'_{j,i}$ is the last column vector of the matrix $m_{j,i}$. Note that the σ-action on $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ is trivial and so we remove σ in the first line of the above equation. Here, the reason we do not express $\mathcal{X}_{i,i}(m)$ based only on the entries in m as in steps (1)–(3) is that two terms involving h_{i-1} and h_{i+1} have only π as a factor which makes the expression with m complicated notation wise. Thus, in the above expression of $\mathcal{X}_{i,i}(m)$, the first line is just a polynomial in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ and the second line is interpreted as explained above as a formal expression. Note that the second line is independent of the choice of lifts $\tilde{m}'_{i-1,i}$ and $\tilde{m}'_{i+1,i}$ of $m'_{i-1,i}$ and $m'_{i+1,i}$, respectively, as explained in the first paragraph of step (4). For example, let $\pi h_{i+1}=\begin{pmatrix}2&\pi\\\sigma(\pi)&2b\end{pmatrix}$ with $b\in A$ and let $\tilde{m}'_{i+1,i}=\begin{pmatrix}x_1+\pi x_2\\y_1+\pi y_2\end{pmatrix}$ such that $m'_{i+1,i}=\begin{pmatrix}x_1\\y_1\end{pmatrix}$. By Section 2A, we may assume that $\pi+\sigma(\pi)=2$ and $\pi\cdot\sigma(\pi)=\epsilon\pi^2=2u$ with $\epsilon\equiv 1 \bmod \pi$ and a unit $u\in A$. Then as a part of $\mathcal{X}_{i,i}(m)$, we can see that

$$\frac{1}{\pi^2}\sigma({}^t\tilde{m}'_{i+1,i})\cdot\pi h_{i+1}\cdot\tilde{m}'_{i+1,i}=\frac{1}{u}(x_1^2+x_1y_1+by_1^2).$$

Since m actually belongs to $\operatorname{Ker}\varphi(R)/\widetilde{G}^1(R)$, we have the following equation by the argument made at the beginning of this paragraph:

$$\begin{aligned}\mathcal{F}_i:\mathcal{X}_{i,i}(m)=\ &(\bar{\alpha}z_i+z_i^2+{}^tm'_{i-2,i}\cdot\bar{h}_{i-2}\cdot m'_{i-2,i}+m'_{i+2,i}\cdot\bar{h}_{i+2}\cdot m'_{i+2,i})\\&\frac{1}{\pi^2}\big(\sigma({}^t\tilde{m}'_{i-1,i})\cdot\sigma(\pi)h_{i-1}\cdot\tilde{m}'_{i-1,i}+\sigma({}^t\tilde{m}'_{i+1,i})\cdot\pi h_{i+1}\cdot\tilde{m}'_{i+1,i}\big)\\=\ &\bar{c}_i=0.\end{aligned}\tag{A-20}$$

Thus we get polynomials $\mathcal{X}_{i,i}$ on $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$, vanishing on the subscheme $\operatorname{Ker}\varphi/\widetilde{G}^1$.

(5) We now choose an even integer j such that L_j is of type I and L_{j+2} is of type II (possibly zero, by our convention). For each such j, there is a nonnegative integer m_j such that L_{j-2l} is of type I for every l with $0\le l\le m_j$ and $L_{j-2(m_j+1)}$ is of type II. Then we

claim that the sum of equations

$$\sum_{l=0}^{m_j} \frac{1}{\bar{\alpha}^2} \mathcal{F}_{j-2l}$$

is the same as

$$\sum_{l=0}^{m_j}\Big(\frac{z_{j-2l}}{\bar{\alpha}}+\Big(\frac{z_{j-2l}}{\bar{\alpha}}\Big)^2\Big)=\Big(\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\bar{\alpha}}\Big)\Big(\sum_{l=0}^{m_j}(\frac{z_{j-2l}}{\bar{\alpha}})+1\Big)=0. \tag{A-21}$$

Here, $\bar{\alpha}$ is the image of α in κ and we consider this equation in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. We postpone the proof of this claim to Lemma A.7.

Let $G^\ddagger$ be the subfunctor of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ consisting of those m satisfying Equations (A-14), (A-16), (A-18) and (A-20). Note that such m also satisfy Equation (A-21). In Lemma A.8 below, we will prove that $G^\ddagger$ is represented by a smooth closed subscheme of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ and is isomorphic to $\mathbb{A}^{l'}\times(\mathbb{Z}/2\mathbb{Z})^\beta$ as a κ-variety, where $\mathbb{A}^{l'}$ is an affine space of dimension

$$l'=\sum_{i<j}n_in_j-\sum_{\substack{i \text{ odd}\\ L_i \text{ bound}}}n_i+\sum_{\substack{i \text{ even}\\ L_i \text{ of type } I^o}}(n_i-1)+\sum_{\substack{i \text{ even}\\ L_i \text{ of type } I^e}}(2n_i-2).$$

For ease of notation, let $G^\dagger=\operatorname{Ker}\varphi/\widetilde{G}^1$. Since $G^\dagger$ and $G^\ddagger$ are both closed subschemes of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ and $G^\dagger(\bar{\kappa})\subset G^\ddagger(\bar{\kappa})$, $(G^\dagger)_{\mathrm{red}}$ is a closed subscheme of $(G^\ddagger)_{\mathrm{red}}=G^\ddagger$. It is easy to check that $\dim G^\dagger=\dim G^\ddagger$ since $\dim G^\dagger=\dim\operatorname{Ker}\varphi-\dim\widetilde{G}^1=l-\dim\widetilde{G}^1$ and $\dim G^\ddagger=l'=l-\dim\widetilde{G}^1$. Here, $\dim\operatorname{Ker}\varphi=l$ is given in Lemma 4.6 and $\dim\widetilde{G}^1$ is given in Theorem A.4.

We claim that $(G^\dagger)_{\mathrm{red}}$ contains at least one (closed) point of each connected component of $G^\ddagger$. Choose an even integer j such that L_j is of type I and L_{j+2} is of type II (possibly zero, by our convention). Consider the closed subgroup scheme F_j of $\widetilde{G}$ defined by the following equations:

- $m_{i,k}=0$ if $i\neq k$;
- $m_{i,i}=\mathrm{id}$ if $i\neq j$;
- and for $m_{j,j}$,

$$\begin{cases} s_j=\mathrm{id},\ y_j=0,\ v_j=0 & \text{if } L_i \text{ is of type } I^o;\\ s_j=\mathrm{id},\ r_j=t_j=y_j=v_j=u_j=w_j=0 & \text{if } L_i \text{ is of type } I^e.\end{cases}$$

We will prove in Lemma A.9 below that each element of $F_j(R)$ for a κ-algebra R satisfies $(z_j^1/\bar{\alpha})+(z_j^1/\bar{\alpha})^2=0$, where $z_j=z_j^1+\pi z_j^2$, and that F_j is isomorphic to $\mathbb{A}^1\times\mathbb{Z}/2\mathbb{Z}$ as a κ-variety, where $\mathbb{A}^1$ is an affine space of dimension 1.

Notice that F_j and $F_{j'}$ commute with each other for all even integers $j\neq j'$, in the sense that $f_j\cdot f_{j'}=f_{j'}\cdot f_j$, where $f_j\in F_j$ and $f_{j'}\in F_{j'}$. Let $F=\prod_j F_j$. Then F is smooth and is a closed subgroup scheme of $\operatorname{Ker}\varphi$ as mentioned in the proof of Theorem 4.11. If $F^\dagger$ is the image of F in $G^\dagger$, then it is smooth and thus a closed subscheme of $(G^\dagger)_{\mathrm{red}}$. By observing Equation (A-21) and $(z_j^1/\bar{\alpha})+(z_j^1/\bar{\alpha})^2=0$ above, we can easily see that $F^\dagger$ contains at least one (closed) point of each connected component of $G^\ddagger$ and this proves our claim.

Combining this fact with $\dim G^\dagger=\dim G^\ddagger$, we conclude that $(G^\dagger)_{\mathrm{red}}\simeq G^\ddagger$, and hence, $G^\dagger=G^\ddagger$ because $G^\dagger$ is a subfunctor of $G^\ddagger$. This completes the proof. □

Lemma A.7. *Choose an even integer j such that L_j is of type I and L_{j+2} is of type II (possibly zero, by our convention). For such j, there is a nonnegative integer m_j such that L_{j-2l} is of type I for every l with $0 \leq l \leq m_j$ and $L_{j-2(m_j+1)}$ is of type II. Then the sum of the equations*

$$\sum_{l=0}^{m_j} \frac{1}{\bar{\alpha}^2} \mathcal{F}_{j-2l}$$

equals

$$\sum_{l=0}^{m_j} \Big(\frac{z_{j-2l}}{\bar{\alpha}} + \Big(\frac{z_{j-2l}}{\bar{\alpha}}\Big)^2 \Big) = \Big(\sum_{l=0}^{m_j} \frac{z_{j-2l}}{\bar{\alpha}} \Big) \Big(\sum_{l=0}^{m_j} (\frac{z_{j-2l}}{\bar{\alpha}}) + 1 \Big) = 0.$$

Proof. Our strategy to prove this lemma is the following. We will first prove that for each odd integer i, the terms containing an h_i add to zero in the sum $\sum_{l=0}^{m_j} \frac{1}{\alpha^2} \mathcal{F}_{j-2l}$. Then we will show that for each even integer i, the terms containing an $\bar{h}_i$ add to zero in the sum $\sum_{l=0}^{m_j} \frac{1}{\alpha^2} \mathcal{F}_{j-2l}$, so that only the terms containing the z_i remain.

We recall the notations used in the theorem. Let m be an element in $(\mathrm{Ker}\, \tilde{\varphi}/\widetilde{M}^1)(R)$ and $\tilde{m} \in \mathrm{Ker}\, \tilde{\varphi}(R)$ be its lift. For any block x_i of m, $\tilde{x}_i$ denotes the corresponding block of $\tilde{m}$ whose reduction is x_i. Since x_i is a block of an element of $(\mathrm{Ker}\, \tilde{\varphi}/\widetilde{M}^1)(R)$, its entries are elements of $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R) \cong R$ as explained in Remark A.5, whereas entries of $\tilde{x}_i$ are elements of $B \otimes_A R$. In addition, for a block a_i of h, where we consider h as an element of $\underline{H}(R)$ as explained in Remark 3.3(2), $\bar{a}_i$ denotes the image of a_i in $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R)$, which is mentioned in step (i) of the proof of Theorem A.4. If we write h as a formal matrix $h = \big(\pi^i \cdot h_i\big)$ with $(\pi^i \cdot h_i)$ for the (i, i)-block and 0 for the remaining blocks, then recall from the paragraph following Equation (A-14) that $\bar{h}_k$ is the matrix with entries in $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R) \cong R$ by letting π be zero in each entry of the formal matrix h_k. To help our computation, we write $\bar{h}_i$. Note that $\epsilon\ (\in B) \equiv 1 \mod \pi$.

$$\bar{h}_i = \begin{cases} \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \\ & & & 1 \end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^o; \\ \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \\ & & & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^e; \\ \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} & \text{if } i \text{ is odd or if } i \text{ is even and } L_i \text{ is of type } II. \end{cases} \tag{A-22}$$

We recall that $m_{i,i}$ is a formal matrix as described in Remark A.5, not a matrix in $M_{n_i\times n_i}((B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R))$, whereas $m_{i,j}$ for $i\neq j$ is a matrix with entries in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Thus we need to modify $m_{i,i}$ into a matrix with entries in $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ in order to use Equation (A-14) as explained in the paragraph following Equation (A-14). We define $\bar{m}_{i,i}\,(\in M_{n_i\times n_i}((B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)))$ to be obtained from $m_{i,i}$ by letting π be zero in each entry of the formal matrix $m_{i,i}$. The matrix $\bar{m}_{i,i}$ is described as follows.

$$\bar{m}_{i,i}=\begin{cases}\begin{pmatrix}\mathrm{id} & 0\\ 0 & 1\end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^o;\\ \begin{pmatrix}\mathrm{id} & r_i & 0\\ 0 & 1 & 0\\ v_i & u_i & 1\end{pmatrix} & \text{if } i \text{ is even and } L_i \text{ is of type } I^e;\\ \mathrm{id} & \text{if } i \text{ is even and } L_i \text{ is of type } \mathit{II};\\ \mathrm{id} & \text{if } i \text{ is odd.}\end{cases} \tag{A-23}$$

In addition, if i is odd, then we have

$$\delta_{i-1}e_{i-1}\cdot m_{i-1,i}+\delta_{i+1}e_{i+1}\cdot m_{i+1,i}=0. \tag{A-24}$$

Here, δ_j, e_j are as explained in the description of $\mathrm{Ker}\,\tilde{\varphi}(R)$, the paragraph following Lemma A.2.

We choose an even integer k (assuming $m_j>0$) such that $j-2(m_j-1)\leq k\leq j$ so that both L_k and L_{k-2} are of type I. We observe $\sigma({}^t\tilde{m}'_{k-1,k})\cdot\sigma(\pi)h_{k-1}\cdot\tilde{m}'_{k-1,k}$ in $\mathcal{F}_k$ and $\sigma({}^t\tilde{m}'_{k-1,k-2})\cdot\sigma(\pi)h_{k-1}\cdot\tilde{m}'_{k-1,k-2}$ in $\mathcal{F}_{k-2}$ (cf. Equation (A-20)). We claim that

$$\frac{1}{\pi^2}\big(\sigma({}^t\tilde{m}'_{k-1,k})\cdot\sigma(\pi)h_{k-1}\cdot\tilde{m}'_{k-1,k}+\sigma({}^t\tilde{m}'_{k-1,k-2})\cdot\sigma(\pi)h_{k-1}\cdot\tilde{m}'_{k-1,k-2}\big)=0. \tag{A-25}$$

Note that this equation is interpreted as explained in the paragraph following Equation (A-19).

We use Equation (A-14) for $i=k-1$ and $j=k$ so that we have

$$ {}^t\bar{m}_{k-1,k-1}\bar{h}_{k-1}m_{k-1,k}={}^tm_{k,k-1}\bar{h}_k\bar{m}_{k,k}. \tag{A-26}$$

Note that this equation is over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Indeed, there is a σ-action in Equation (A-14) but it is trivial over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Recall that $m'_{k-1,k}$ is the last column vector of $m_{k-1,k}$. Let $e_{k-1}=(0,\cdots,0,1)$ be of size $1\times n_k$. Then $m'_{k-1,k}=m_{k-1,k}\cdot{}^te_{k-1}$. We multiply both sides of the above equation by ${}^te_{k-1}$ on the right. Then the left hand side is ${}^t\bar{m}_{k-1,k-1}\bar{h}_{k-1}m_{k-1,k}\cdot{}^te_{k-1}=\bar{h}_{k-1}m'_{k-1,k}$ since ${}^t\bar{m}_{k-1,k-1}=\mathrm{id}$. The right hand side is ${}^tm_{k,k-1}\bar{h}_k\bar{m}_{k,k}\cdot{}^te_{k-1}$. Since $\bar{m}_{k,k}\cdot{}^te_{k-1}$ is the last column vector of $\bar{m}_{k,k}$, $\bar{m}_{k,k}\cdot{}^te_{k-1}={}^te_{k-1}$ by Equation (A-23) so that ${}^tm_{k,k-1}\bar{h}_k\bar{m}_{k,k}\cdot{}^te_{k-1}={}^tm_{k,k-1}\bar{h}_k\cdot{}^te_{k-1}$. Furthermore, $\bar{h}_k$ is symmetric over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$ and so ${}^tm_{k,k-1}\bar{h}_k\cdot{}^te_{k-1}={}^t(e_{k-1}\cdot\bar{h}_k m_{k,k-1})$. Then based on the matrix form of $\bar{h}_k$ in Equation (A-22), we have that $e_{k-1}\cdot\bar{h}_k$ is the same as e_k, where e_k is defined in the paragraph following Lemma A.2. (There, e_j is defined when j is even and L_j is of type I.) In conclusion, Equation (A-26) induces the equation

$$\bar{h}_{k-1}m'_{k-1,k}={}^t(e_k\cdot m_{k,k-1}) \tag{A-27}$$

over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$.

We again use Equation (A-14) for $i=k-2$ and $j=k-1$ so that we have

$$ {}^t\bar{m}_{k-2,k-2}\bar{h}_{k-2}m_{k-2,k-1} = {}^t m_{k-1,k-2}\bar{h}_{k-1}\bar{m}_{k-1,k-1}. \tag{A-28}$$

Note that this equation is over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Since $k-1$ is odd, $\bar{m}_{k-1,k-1}=\mathrm{id}$ by Equation (A-23). Recall that $m'_{k-1,k-2}$ is the last column vector of $m_{k-1,k-2}$. Let $e'_{k-1}=(0,\cdots,0,1)$ of size $1\times n_{k-2}$. Then $m'_{k-1,k-2}=m_{k-1,k-2}\cdot {}^t e'_{k-1}$. We multiply both sides of the above equation by e'_{k-1} on the left. Then the right hand side is $e'_{k-1}\cdot {}^t m_{k-1,k-2}\bar{h}_{k-1}\bar{m}_{k-1,k-1} = {}^t m'_{k-1,k-2}\bar{h}_{k-1}$. Note that $\bar{m}_{k-2,k-2}\cdot {}^t e'_{k-1} = {}^t e'_{k-1}$ by Equation (A-23) since this is the last column vector of $\bar{m}_{k-2,k-2}$. Thus in the left hand side, $e'_{k-1}\cdot {}^t\bar{m}_{k-2,k-2}\bar{h}_{k-2}m_{k-2,k-1} = e'_{k-1}\cdot\bar{h}_{k-2}m_{k-2,k-1}$. Based on the matrix form of $\bar{h}_k$ for an even integer k in Equation (A-22), $e'_{k-1}\cdot\bar{h}_{k-2}$ is the same as e_{k-2}, where e_k is defined in the paragraph following Lemma A.2. In conclusion, Equation (A-28) induces the equation

$$ {}^t m'_{k-1,k-2}\bar{h}_{k-1} = e_{k-2}\cdot m_{k-2,k-1} \tag{A-29}$$

over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$.

Now we use Equations (A-27) and (A-29). Based on the matrix form of $\bar{h}_{k-1}$ for an odd integer $k-1$ in Equation (A-22), we have that $\bar{h}_{k-1}\cdot\bar{h}_{k-1}=\mathrm{id}$ and $\bar{h}_{k-1}$ is symmetric. Thus, by multiplying Equations (A-27) and (A-29) by $\bar{h}_{k-1}$, we obtain $m'_{k-1,k}={}^t(e_k\cdot m_{k,k-1}\cdot\bar{h}_{k-1})$ and ${}^t m'_{k-1,k-2}=e_{k-2}\cdot m_{k-2,k-1}\cdot\bar{h}_{k-1}$, respectively, as equations over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$.

On the other hand, we observe that $k-1$ is odd and both L_{k-2} and L_k are of type I. Thus $e_{k-2}\cdot m_{k-2,k-1}=e_k\cdot m_{k,k-1}$ by Equation (A-24). We multiply this equation by $\bar{h}_{k-1}$ and so obtain

$$ e_{k-2}\cdot m_{k-2,k-1}\cdot\bar{h}_{k-1} = e_k\cdot m_{k,k-1}\cdot\bar{h}_{k-1}$$

as an equation over $(B\otimes_A R)/(\pi\otimes 1)(B\otimes_A R)$. Therefore, we have the equation

$$ m'_{k-1,k}=m'_{k-1,k-2}.$$

As mentioned at the paragraph following Equation (A-19), Equation (A-25) is independent of the choice of a lift of $m'_{k-1,k}$ and $m'_{k-1,k-2}$. Therefore, two terms in Equation (A-25) are same and this verifies our claim.

In the case of $\mathcal{F}_j$, following the proof of Equation (A-29), we have ${}^t m'_{j+1,j}\cdot\bar{h}_{j+1} = e_j\cdot m_{j,j+1}$. Since L_{j+2} is of type II (possibly zero, by our convention), $e_j\cdot m_{j,j+1}=0$ by Equation (A-24). Thus, the term involving h_{j+1} in $\mathcal{F}_j$ is zero. In the case of $j-2m_j$, where $m_j\geq 0$, the term involving h_{j-2m_j-1} in $\mathcal{F}_{j-2m_j}$ is zero in a manner similar to that of the above case of $\mathcal{F}_j$.

To summarize, for each odd integer i, the terms containing an h_i add to zero in $\sum_{l=0}^{m_j}\frac{1}{\alpha^2}\mathcal{F}_{j-2l}$.

We now prove that for each even integer i, the terms containing an $\bar{h}_i$ add to zero in $\sum_{l=0}^{m_j}\frac{1}{\alpha^2}\mathcal{F}_{j-2l}$. We again choose an even integer k (assuming $m_j>0$) such that $j-2(m_j-1)\leq k\leq j$ so that both L_k and L_{k-2} are of type I. We observe ${}^t m'_{k-2,k}\cdot\bar{h}_{k-2}\cdot m'_{k-2,k}$

in $\mathcal{F}_k$ and ${}^t m'_{k,k-2} \cdot \bar{h}_k \cdot m'_{k,k-2}$ in $\mathcal{F}_{k-2}$, and we claim that

$$ {}^t m'_{k-2,k} \cdot \bar{h}_{k-2} \cdot m'_{k-2,k} + {}^t m'_{k,k-2} \cdot \bar{h}_k \cdot m'_{k,k-2} = 0, \tag{A-30} $$

as an equation over $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R)$. Let $\widehat{m}'_{k-2,k}$ be the $(n_{k-2} \times n_k)$-th entry (resp. $((n_{k-2}-1) \times n_k)$-th entry) of $m_{k-2,k}$ when L_{k-2} is of type I^o (resp. I^e). We can also define $\widehat{m}'_{k,k-2}$ as the $(n_k \times n_{k-2})$-th entry (resp. $((n_k-1) \times n_{k-2})$-th entry) of $m_{k,k-2}$ when L_k is of type I^o (resp. I^e). Then the above Equation (A-30) is the same as

$$ (\widehat{m}'_{k-2,k})^2 + (\widehat{m}'_{k,k-2})^2 = 0. \tag{A-31} $$

We use Equation (A-14) for $i = k-2$ and $j = k$ so that we have

$$ {}^t \bar{m}_{k-2,k-2} \bar{h}_{k-2} m_{k-2,k} + {}^t m_{k-1,k-2} \bar{h}_{k-1} m_{k-1,k} + {}^t m_{k,k-2} \bar{h}_k \bar{m}_{k,k} = 0. \tag{A-32} $$

Let $\widetilde{e}_k = (0, \cdots, 0, 1)$ of size $1 \times n_k$ and $\widetilde{e}_{k-2} = (0, \cdots, 0, 1)$ of size $1 \times n_{k-2}$. Then we have

$$ \widetilde{e}_{k-2} \cdot {}^t \bar{m}_{k-2,k-2} \bar{h}_{k-2} m_{k-2,k} \cdot {}^t \widetilde{e}_k = \widehat{m}'_{k-2,k} \tag{A-33} $$

since $m_{k-2,k} \cdot {}^t \widetilde{e}_k = m'_{k-2,k}$ and $\bar{m}_{k-2,k-2} \cdot {}^t \widetilde{e}_{k-2} = {}^t \widetilde{e}_{k-2}$. We also have

$$ \widetilde{e}_{k-2} \cdot {}^t m_{k,k-2} \bar{h}_k \bar{m}_{k,k} \cdot {}^t \widetilde{e}_k = \widehat{m}'_{k,k-2} \tag{A-34} $$

since $\bar{m}_{k,k} \cdot {}^t \widetilde{e}_k = {}^t \widetilde{e}_k$ and $m_{k,k-2} \cdot {}^t \widetilde{e}_{k-2} = m'_{k,k-2}$. Note that we use Equations (A-22) and (A-23) for our matrix computation. On the other hand, due to the fact that $\bar{h}_{k-1} \cdot \bar{h}_{k-1} = \mathrm{id}$ and $\bar{h}_{k-1}$ is symmetric, we have

$$ \widetilde{e}_{k-2} \cdot {}^t m_{k-1,k-2} \bar{h}_{k-1} m_{k-1,k} \cdot {}^t \widetilde{e}_k = (\widetilde{e}_{k-2} \cdot {}^t m_{k-1,k-2} \bar{h}_{k-1}) \cdot \bar{h}_{k-1} \cdot (\bar{h}_{k-1} m_{k-1,k} \cdot {}^t \widetilde{e}_k). \tag{A-35} $$

Now, $\widetilde{e}_{k-2} \cdot {}^t m_{k-1,k-2} \bar{h}_{k-1} = {}^t m'_{k-1,k-2} \bar{h}_{k-1} = e_{k-2} \cdot m_{k-2,k-1}$ by Equation (A-29) and $\bar{h}_{k-1} m_{k-1,k} \cdot {}^t \widetilde{e}_k = \bar{h}_{k-1} m'_{k-1,k} = {}^t (e_k \cdot m_{k,k-1})$ by Equation (A-27). Since $e_{k-2} \cdot m_{k-2,k-1} = e_k \cdot m_{k,k-1}$ by Equation (A-24), Equation (A-35) equals

$$ (\widetilde{e}_{k-2} \cdot {}^t m_{k-1,k-2} \bar{h}_{k-1}) \cdot \bar{h}_{k-1} \cdot (\bar{h}_{k-1} m_{k-1,k} \cdot \widetilde{e}_k) = (e_k \cdot m_{k,k-1}) \cdot \bar{h}_{k-1} \cdot {}^t (e_k \cdot m_{k,k-1}) = 0. \tag{A-36} $$

We now combine Equations (A-33), (A-34), and (A-36). Namely, if we multiply $\widetilde{e}_{k-2}$ to the left of each side in Equation (A-32) and we multiply ${}^t \widetilde{e}_k$ to the right of each side in Equation (A-32), then we have

$$ \widehat{m}'_{k-2,k} + 0 + \widehat{m}'_{k,k-2} = 0 \tag{A-37} $$

and so Equations (A-31) and (A-30) are proved.

In the case of $\mathcal{F}_j$, the term ${}^t m'_{j+2,j} \cdot \bar{h}_{j+2} \cdot m'_{j+2,j} = 0$ since L_{j+2} is of type II (possibly zero, by our convention). Similarly, the term ${}^t m'_{j-2m_j-2,j-2m_j} \cdot \bar{h}_{j-2m_j-2} \cdot m'_{j-2m_j-2,j-2m_j}$ of $\mathcal{F}_{j-2m_j}$, where $m_j \geq 0$, is 0 since L_{j-2m_j-2} is of type II. Here, we use Equation (A-22) for our matrix multiplication.

To summarize, for each even integer i, the terms containing an h_i add to zero in $\sum_{l=0}^{m_j} \frac{1}{\alpha^2} \mathcal{F}_{j-2l}$.

Therefore, the sum of equations $\sum_{l=0}^{m_j}\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2l}$ equals

$$\sum_{l=0}^{m_j}\frac{1}{\bar{\alpha}^2}(\bar{\alpha}\bar{z}_{j-2l}+\bar{z}_{j-2l}^2)=0.$$

This is the same as

$$\sum_{l=0}^{m_j}\Big(\frac{\bar{z}_{j-2l}}{\bar{\alpha}}+\Big(\frac{\bar{z}_{j-2l}}{\bar{\alpha}}\Big)^2\Big)=\Big(\sum_{l=0}^{m_j}\frac{\bar{z}_{j-2l}}{\bar{\alpha}}\Big)\Big(\sum_{l=0}^{m_j}\Big(\frac{\bar{z}_{j-2l}}{\bar{\alpha}}\Big)+1\Big)=0. \tag{A-38}$$

This completes the proof of the lemma. □

Lemma A.8. *Let $G^\ddagger$ be the subfunctor of* $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ *consisting of those m satisfying Equations (A-14), (A-16), (A-18), and (A-20). Note that such m then satisfies Equation (A-21) as well. Then $G^\ddagger$ is represented by a smooth closed subscheme of* $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ *and is isomorphic to $\mathbb{A}^{l'}\times(\mathbb{Z}/2\mathbb{Z})^\beta$ as a κ-variety, where $\mathbb{A}^{l'}$ is an affine space of dimension l'. Here,*

$$l'=\sum_{i<j}n_in_j-\sum_{\substack{i \text{ odd}\\ L_i \text{ bound}}}n_i+\sum_{\substack{i \text{ even}\\ L_i \text{ of type } I^o}}(n_i-1)+\sum_{\substack{i \text{ even}\\ L_i \text{ of type } I^e}}(2n_i-2).$$

Proof. Let $\mathcal{J}$ be the set of even integers j such that L_j is of type I and L_{j+2} is of type II (possibly empty, by our convention). Note that Equation (A-20) implies Equation (A-21) by Lemma A.7. Equation (A-21) implies that $G^\ddagger$ is disconnected with at least 2^β connected components (Exercise 2.19 of [Hartshorne 1977]). Here, $\beta=\#\mathcal{J}$. Let $\mathcal{J}_1$ and $\mathcal{J}_2$ be a pair of two (possibly empty) subsets of $\mathcal{J}$ such that $\mathcal{J}$ is the disjoint union of $\mathcal{J}_1$ and $\mathcal{J}_2$. Let $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$ be the subfunctor of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ consisting of those m satisfying Equations (A-14), (A-16), (A-18), and (A-20), the equations $\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\bar{\alpha}}=0$ for any $j\in\mathcal{J}_1$, and the equations $\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\bar{\alpha}}=1$ for any $j\in\mathcal{J}_2$. Here m_j is the integer associated to j defined in Lemma A.7. We claim that $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$ is represented by a smooth closed subscheme of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ and is isomorphic to $\mathbb{A}^{l'}$. Since the scheme $G^\ddagger$ is a direct product of $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$'s for any such pair of $\mathcal{J}_1,\mathcal{J}_2$ by Exercise 2.19 of [Hartshorne 1977], the lemma follows from this claim.

It is obvious that $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$ is represented by a closed subscheme of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ since the equations defining $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$ as a subfunctor of $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ are all polynomials. Thus it suffices to show that $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$ is isomorphic to an affine space $\mathbb{A}^{l'}$. Our strategy to show this is that the coordinate ring of $\widetilde{G}^\ddagger_{\mathcal{J}_1,\mathcal{J}_2}$ is isomorphic to a polynomial ring. To do that, we use the following trick over and over. We consider the polynomial ring $\kappa[x_1,\cdots,x_n]$ and it quotient ring $\kappa[x_1,\cdots,x_n]/(x_1+P(x_2,\cdots,x_n))$. Then the quotient ring $\kappa[x_1,\cdots,x_n]/(x_1+P(x_2,\cdots,x_n))$ is isomorphic to $\kappa[x_2,\cdots,x_n]$ and in this case we say that x_1 *can be eliminated by* $x_2,\cdots,x_n$.

By the description of an element of $(\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1)(R)$ in Remark A.5, we see that $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ is isomorphic to an affine space of dimension

$$2\sum_{i<j}n_in_j-\sum_{\substack{i \text{ odd}\\ L_i \text{ bound}}}n_i+\sum_{\substack{i \text{ even}\\ L_i \text{ of type } I^o}}(2n_i-1)+\sum_{\substack{i \text{ even}\\ L_i \text{ of type } I^e}}(4n_i-4)$$

with variables

$$(m_{i,j})_{i\neq j},\quad (y_i,v_i,z_i)_{\substack{i\text{ even}\\ L_i\text{ of type }I^o}},\quad (r_i,t_i,y_i,v_i,x_i,z_i,u_i,w_i)_{\substack{i\text{ even}\\ L_i\text{ of type }I^e}}$$

such that $\delta_{i-1}e_{i-1}\cdot m_{i-1,i}+\delta_{i+1}e_{i+1}\cdot m_{i+1,i}=0$ with i odd. Here, δ_j, e_j are as explained in the description of $\operatorname{Ker}\tilde{\varphi}(R)$, the paragraph right after Lemma A.2.

From now on, we eliminate suitable variables based on Equations (A-14), (A-16), (A-18), and (A-20), the equations $\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\alpha}=0$ for all $j\in\mathcal{J}_1$, and the equations $\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\alpha}=1$ for all $j\in\mathcal{J}_2$.

(1) We first consider Equation (A-14). For two integers i, j with $i<j$, we have

$${}^tm_{j,i}\bar{h}_j\bar{m}_{j,j}=\sum_{i\leq k\leq j-1}{}^tm_{k,i}\bar{h}_km_{k,j}\ \text{ over }\ (B\otimes_AR)/(\pi\otimes1)(B\otimes_AR).$$

By Equation (A-23), $\bar{m}_{j,j}=\text{id}$ if L_j is not of type I^e. Thus the above equation equals ${}^tm_{j,i}\bar{h}_j=\sum_{i\leq k\leq j-1}{}^tm_{k,i}\bar{h}_km_{k,j}$ over $(B\otimes_AR)/(\pi\otimes1)(B\otimes_AR)$ if L_j is not of type I^e. Since $\bar{h}_j$ is a nonsingular matrix by Equation (A-22), $m_{j,i}$ can be eliminated by the right hand side. If L_j is of type I^e, we have

$$\bar{m}_{j,j}=\begin{pmatrix}\text{id}&r_j&0\\0&1&0\\v_j&u_j&1\end{pmatrix}\quad\text{and}\quad\bar{h}_j=\begin{pmatrix}a_j&0&0\\0&1&1\\0&1&0\end{pmatrix}$$

by Equations (A-23) and (A-22), respectively. Then

$$\bar{h}_j\bar{m}_{j,j}=\begin{pmatrix}a_j&a_jr_j&0\\v_j&1+u_j&1\\0&1&0\end{pmatrix}.$$

To compute ${}^tm_{j,i}\bar{h}_j\bar{m}_{j,j}$, we write ${}^tm_{j,i}=\begin{pmatrix}A_j&B_j&C_j\end{pmatrix}$ so that

$${}^tm_{j,i}\bar{h}_j\bar{m}_{j,j}=\begin{pmatrix}A_ja_j+B_jv_j&A_ja_jr_j+B_j(1+u_j)+C_j&B_j\end{pmatrix}.$$

By first considering the $(1,3)$-block of the matrix ${}^tm_{j,i}\bar{h}_j\bar{m}_{j,j}$, B_j can be eliminated by $\sum_{i\leq k\leq j-1}{}^tm_{k,i}\bar{h}_km_{k,j}$. Then we consider the $(1,1)$-block of ${}^tm_{j,i}\bar{h}_j\bar{m}_{j,j}$. Since a_j is a nonsingular matrix, we see that A_j can be eliminated by $\sum_{i\leq k\leq j-1}{}^tm_{k,i}\bar{h}_km_{k,j}$ with B_jv_j. By considering the $(1,2)$-block of ${}^tm_{j,i}\bar{h}_j\bar{m}_{j,j}$, C_j can be eliminated by $\sum_{i\leq k\leq j-1}{}^tm_{k,i}\bar{h}_km_{k,j}$ with $A_ja_jr_j+B_j(1+u_j)$. Therefore, all lower triangular blocks $m_{j,i}$ (with $j>i$) can be eliminated by upper triangular blocks $m_{i,j}$ together with r_j,v_j,u_j (resp. r_i,v_i,u_i) if L_j (resp. L_i) is of type I^e. Here r_i,v_i,u_i are nontrivial blocks of $\bar{m}_{i,i}$, if L_i is of type I^e, which appeared in the right hand side of the above equation.

On the other hand, the equation $\delta_{i-1}e_{i-1}\cdot m_{i-1,i}+\delta_{i+1}e_{i+1}\cdot m_{i+1,i}=0$ for an odd integer i, which is one equation defining $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ (cf. Remark A.5(3)), should be rewritten in terms of upper triangular blocks. To do that, we use Equation (A-27) with $i=k-1$. Note that the only assumption needed in Equation (A-27) is that L_k is of type I. Thus the above equation is the same as

$$\delta_{i-1}e_{i-1}\cdot m_{i-1,i}+\delta_{i+1}{}^t(\bar{h}_im'_{i,i+1})=0.$$

(2) We secondly consider Equation (A-16). If L_i is of type I^o, then v_i can be eliminated by y_i and $m_{i-1,i}, m_{i,i+1}$.

(3) Next, we consider Equation (A-18). By $\mathcal{X}_{i,1,2}$, v_i can be eliminated by r_i. By $\mathcal{X}_{i,1,3}$, y_i can be eliminated by t_i, v_i, z_i and entries from $m_{i-1,i}, m_{i,i+1}$. By $\mathcal{X}_{i,2,3}$, x_i can be eliminated by r_i, t_i, z_i, w_i, u_i and entries from $m_{i-1,i}, m_{i,i+1}$.

(4) Finally, we consider $\frac{1}{\bar{\alpha}^2}\mathcal{F}_i$, instead of $\mathcal{F}_i$ (Equation (A-20)), together with equations $\sum_{l=0}^{m_j} \frac{z_{j-2l}}{\alpha} = 0$ with $j \in \mathcal{J}_1$ and equations $\sum_{l=0}^{m_j} \frac{z_{j-2l}}{\alpha} = 1$ with $j \in \mathcal{J}_2$. Note that $\frac{1}{\bar{\alpha}^2}\mathcal{F}_i$ is equivalent to $\mathcal{F}_i$ since α is a unit in B. For each $j \in \mathcal{J}$, there is a nonnegative integer m_j such that L_{j-2l} is of type I for every l with $0 \le l \le m_j$ and $L_{j-2(m_j+1)}$ is of type II (cf. Lemma A.7).

To analyze these equations, we investigate $\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2l}$ for a fixed $j \in \mathcal{J}$. First assume that $m_j \ge 1$. Since we have eliminated all lower triangular blocks in step (1), we need to replace lower triangular blocks appeared in $\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2l}$ by suitable upper triangular blocks. If $m_j \ge 2$, then we choose an integer l such that $0 < l < m_j$. By definition, $\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2l}$ is

$$\begin{aligned}
&\frac{1}{\bar{\alpha}^2 \cdot \pi^2}\big(\sigma({}^t\tilde{m}'_{j-2l-1,j-2l}) \cdot \sigma(\pi) h_{j-2l-1} \cdot \tilde{m}'_{j-2l-1,j-2l} \\
&\qquad\qquad\qquad\qquad + \sigma({}^t\tilde{m}'_{j-2l+1,j-2l}) \cdot \pi h_{j-2l+1} \cdot \tilde{m}'_{j-2l+1,j-2l}\big) \\
&\qquad\qquad + \frac{z_{j-2l}}{\bar{\alpha}} + \Big(\frac{z_{j-2l}}{\bar{\alpha}}\Big)^2 + \frac{{}^t m'_{j-2l-2,j-2l} \cdot \bar{h}_{j-2l-2} \cdot m'_{j-2l-2,j-2l}}{\bar{\alpha}^2} \\
&\qquad\qquad + \frac{{}^t m'_{j-2l+2,j-2l} \cdot \bar{h}_{j-2l+2} \cdot m'_{j-2l+2,j-2l}}{\bar{\alpha}^2} \\
&= 0.
\end{aligned}$$

The first two lines are interpreted as explained in the paragraph following Equation (A-19) and the third and fourth line is a polynomial in $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R)$. We claim that the equation $\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2l}$ is the same as the following:

$$\begin{aligned}
&\frac{1}{\bar{\alpha}^2 \cdot \pi^2}\big(\sigma({}^t\tilde{m}'_{j-2l-1,j-2l}) \cdot \sigma(\pi) h_{j-2l-1} \cdot \tilde{m}'_{j-2l-1,j-2l} \\
&\qquad + (e_{j-2l} \cdot \sigma(\tilde{m}_{j-2l,j-2l+1})) \cdot \pi h^3_{j-2l+1} \cdot {}^t(e_{j-2l} \cdot \tilde{m}_{j-2l,j-2l+1})\big) \\
&\qquad\qquad + \frac{z_{j-2l}}{\bar{\alpha}} + \Big(\frac{z_{j-2l}}{\bar{\alpha}}\Big)^2 + \Big(\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}\Big)^2 + \Big(\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}\Big)^2 \\
&= 0. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(A-39)}
\end{aligned}$$

The third line easily follows from the definition of $\widehat{m}'_{k-2,k}$ and $\widehat{m}'_{k,k-2}$ (given in the paragraph following Equation (A-30)) combined with Equation (A-37). For the first two lines, we consider Equation (A-29) with $k-2 = j-2l$ which gives the identity $m'_{j-2l+1,j-2l} = \bar{h}_{j-2l+1} \cdot {}^t(e_{j-2l} \cdot m_{j-2l,j-2l+1})$ over $(B \otimes_A R)/(\pi \otimes 1)(B \otimes_A R)$. Note that the only assumption needed in Equation (A-29) is that L_{k-2} is of type I. Then $h_{j-2l+1} \cdot {}^t(e_{j-2l} \cdot \tilde{m}_{j-2l,j-2l+1})$ is a lift of $\bar{h}_{j-2l+1} \cdot {}^t(e_{j-2l} \cdot m_{j-2l,j-2l+1})$. The first line is independent of the choice of a lift $\tilde{m}'_{j-2l+1,j-2l}$ of $m'_{j-2l+1,j-2l}$ as explained at the paragraph following Equation (A-19). This fact completes our claim. The above equation is equivalent to

$$\frac{1}{\bar{\alpha}^2\cdot\pi^2}\big(\sigma({}^t\tilde{m}'_{j-2l-1,j-2l})\cdot\sigma(\pi)h_{j-2l-1}\cdot\tilde{m}'_{j-2l-1,j-2l} + (e_{j-2l}\cdot\sigma(\tilde{m}_{j-2l,j-2l+1}))\cdot\pi h^3_{j-2l+1}\cdot{}^t(e_{j-2l}\cdot\tilde{m}_{j-2l,j-2l+1})\big) + \Big(\frac{z_{j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}\Big)+\Big(\frac{z_{j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}\Big)^2 = \Big(\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}\Big) \tag{A-40}$$

by adding $\big(\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}\big)$ to both sides.

For $\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2m_j}$, we observe that L_{j-2m_j-2} is of type II. By Equation (A-27) with $k=j-2m_j$, we have ${}^tm'_{j-2m_j-1,j-2m_j}=e_{j-2m_j}\cdot m_{j-2m_j,j-2m_j-1}\bar{h}_{j-2m_j-1}$. Here we use the fact that $\bar{h}^2_{j-2m_j-1}=\mathrm{id}$ (cf. Equation (A-22)). Note that the only assumption needed in Equation (A-27) is that L_k is of type I. On the other hand, the equation in Remark A.5(3), when $i=j-2m_j-1$, is $e_{j-2m_j}\cdot m_{j-2m_j,j-2m_j-1}=0$ since L_{j-2m_j-2} is of type II. Thus ${}^tm'_{j-2m_j-1,j-2m_j}=0$. Therefore, $\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2m_j}$ is

$$\frac{1}{\bar{\alpha}^2\cdot\pi^2}\big((e_{j-2m_j}\cdot\sigma(\tilde{m}_{j-2m_j,j-2m_j+1}))\cdot\pi h^3_{j-2m_j+1}\cdot{}^t(e_{j-2m_j}\cdot\tilde{m}_{j-2m_j,j-2m_j+1})\big) +\frac{z_{j-2m_j}}{\bar{\alpha}}+\Big(\frac{z_{j-2m_j}}{\bar{\alpha}}\Big)^2+\Big(\frac{\widehat{m}'_{j-2m_j,j-2m_j+2}}{\bar{\alpha}}\Big)^2 = 0. \tag{A-41}$$

This equation is equivalent to

$$\frac{1}{\bar{\alpha}^2\cdot\pi^2}\Big((e_{j-2m_j}\cdot\sigma(\tilde{m}_{j-2m_j,j-2m_j+1}))\cdot\pi h^3_{j-2m_j+1}\cdot{}^t(e_{j-2m_j}\cdot\tilde{m}_{j-2m_j,j-2m_j+1})\Big) +\Big(\frac{z_{j-2m_j}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2m_j,j-2m_j+2}}{\bar{\alpha}}\Big)+\Big(\frac{z_{j-2m_j}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2m_j,j-2m_j+2}}{\bar{\alpha}}\Big)^2 =\frac{\widehat{m}'_{j-2m_j,j-2m_j+2}}{\bar{\alpha}} \tag{A-42}$$

by adding $\frac{\widehat{m}'_{j-2m_j,j-2m_j+2}}{\bar{\alpha}}$ to both sides.

We emphasize that it is unnecessary to investigate $\mathcal{F}_j$ since the equation $\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\bar{\alpha}}=0$ (resp. $\sum_{l=0}^{m_j}\frac{z_{j-2l}}{\bar{\alpha}}=1$) if $j\in\mathcal{J}_1$ (resp. if $j\in\mathcal{J}_2$) already implies Equation (A-21) so that $\sum_{l=0}^{m_j}\frac{1}{\bar{\alpha}^2}\mathcal{F}_{j-2l}=0$.

We now observe Equations (A-40) and (A-42). We introduce a new variable

$$z'_{j-2l}=\begin{cases}\dfrac{z_{j-2l}}{\bar{\alpha}}+\dfrac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}+\dfrac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}} & \text{if } 0<l<m_j;\\[2ex] \dfrac{z_{j-2m_j}}{\bar{\alpha}}+\dfrac{\widehat{m}'_{j-2m_j,j-2m_j+2}}{\bar{\alpha}} & \text{if } l=m_j.\end{cases}$$

Then z_{j-2l} can be eliminated by z'_{j-2l}, $\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}$, $\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}$. In addition, by using Equations (A-40) and (A-42), the term $\frac{\widehat{m}'_{j-2l-2,j-2l}}{\bar{\alpha}}+\frac{\widehat{m}'_{j-2l,j-2l+2}}{\bar{\alpha}}$ can be eliminated by

z'_{j-2l} and $m'_{j-2l-1,j-2l}, m_{j-2l,j-2l+1}$. Furthermore, the equation $\sum_{l=0}^{m_j} \frac{z_{j-2l}}{\bar{\alpha}} = 0$ (resp. $\sum_{l=0}^{m_j} \frac{z_{j-2l}}{\bar{\alpha}} = 1$) if $j \in \mathcal{J}_1$ (resp. if $j \in \mathcal{J}_2$) implies that z_j can be eliminated by z'_{j-2l}, $m'_{j-2l-1,j-2l}, m_{j-2l,j-2l+1}$ with $0 < l \leq m_j$.

If $m_j = 0$, then we can show that the equation $\frac{1}{\bar{\alpha}^2}\mathcal{F}_j$ is the same as

$$\frac{z_j}{\bar{\alpha}} + \left(\frac{z_j}{\bar{\alpha}}\right)^2 = 0$$

by using an argument similar to that used in the proof of Equation (A-41). Then the equation $\frac{z_j}{\bar{\alpha}} = 0$ (resp. $\frac{z_j}{\bar{\alpha}} = 1$) if $j \in \mathcal{J}_1$ (resp. if $j \in \mathcal{J}_2$) implies that z_j can be eliminated.

We now combine all cases (1)–(4) observed above.

(a) By (1), we eliminate $\sum_{i<j} n_i n_j$ variables.

(b) By (2), we eliminate $\sum_{i \text{ even and } L_i \text{ of type } I^o} (n_i - 1)$ variables.

(c) By (3), we eliminate $\sum_{i \text{ even and } L_i \text{ of type } I^e} (2(n_i - 2) + 1)$ variables.

(d) By (4), we eliminate $\#\{i : i \text{ is even and } L_i \text{ is of type } I\}$ variables.

Recall from the third paragraph of the proof that $\operatorname{Ker}\tilde{\varphi}/\widetilde{M}^1$ is isomorphic to an affine space of dimension

$$2\sum_{i<j} n_i n_j - \sum_{\substack{i \text{ odd} \\ L_i \text{ bound}}} n_i + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^o}} (2n_i - 1) + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^e}} (4n_i - 4).$$

Thus, $\widetilde{G}^{\ddagger}_{\mathcal{J}_1,\mathcal{J}_2}$ is isomorphic to an affine space of dimension

$$\Bigg(2\sum_{i<j} n_i n_j - \sum_{\substack{i \text{ odd} \\ L_i \text{ bound}}} n_i + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^o}} (2n_i - 1) + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^e}} (4n_i - 4)\Bigg) \tag{A-43}$$

$$-\Bigg(\sum_{i<j} n_i n_j + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^o}} (n_i - 1) + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^e}} (2(n_i - 2) + 1) + \#\{i : i \text{ is even and } L_i \text{ is of type } I\}\Bigg).$$

Therefore, the dimension of $\widetilde{G}^{\ddagger}_{\mathcal{J}_1,\mathcal{J}_2}$ is

$$\sum_{i<j} n_i n_j - \sum_{\substack{i \text{ odd} \\ L_i \text{ bound}}} n_i + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^o}} (n_i - 1) + \sum_{\substack{i \text{ even} \\ L_i \text{ of type } I^o}} (2n_i - 2), \tag{A-44}$$

which finishes the proof. □

Lemma A.9. *Let F_j be the closed subgroup scheme of $\widetilde{G}$ defined by the following equations*:

- $m_{i,k} = 0$ *if* $i \neq k$;
- $m_{i,i} = \mathrm{id}$ *if* $i \neq j$;
- *and for* $m_{j,j}$,

$$\begin{cases} s_j = \mathrm{id},\ y_j = 0,\ v_j = 0 & \text{if } L_i \text{ is of type } I^o; \\ s_j = \mathrm{id},\ r_j = t_j = y_j = v_j = u_j = w_j = 0 & \text{if } L_i \text{ is of type } I^e. \end{cases}$$

Then F_j is isomorphic to $\mathbb{A}^1 \times \mathbb{Z}/2\mathbb{Z}$ as a κ-variety, where $\mathbb{A}^1$ is an affine space of dimension 1, *and has exactly two connected components.*

Proof. A matrix form of an element m of $F_j(R)$ for a κ-algebra R is

$$\begin{pmatrix} \mathrm{id} & 0 & & \dots & & & 0 \\ 0 & \ddots & & & & & \\ & & \mathrm{id} & & & & \\ \vdots & & & m_{j,j} & & & \vdots \\ & & & & \mathrm{id} & & \\ & & & & & \ddots & 0 \\ 0 & & & \dots & & 0 & \mathrm{id} \end{pmatrix}$$

such that

$$m_{j,j} = \begin{cases} \begin{pmatrix} \mathrm{id} & 0 \\ 0 & 1+\pi z_j \end{pmatrix} & \text{if } L_j \text{ is of type } I^o; \\ \begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & 1+\pi x_j & \pi z_j \\ 0 & 0 & 1 \end{pmatrix} & \text{if } L_j \text{ is of type } I^e. \end{cases}$$

To prove the lemma, we consider the matrix equation $\sigma({}^t m) \cdot h \cdot m = h$. Recall that h, as an element of $\underline{H}(R)$, is as explained in Remark 3.3(2). Based on Equations (A-1) and (A-2), the diagonal (i,i)-blocks of $\sigma({}^t m) \cdot h \cdot m = h$ with $i \neq j$ are trivial and the nondiagonal blocks of $\sigma({}^t m) \cdot h \cdot m = h$ are also trivial. The (j,j)-block of $\sigma({}^t m) \cdot h \cdot m$ is

$$\begin{cases} \pi^j \cdot \begin{pmatrix} a_j & 0 \\ 0 & (1+\sigma(\pi z_j)) \cdot (1+2\bar{\gamma}_j) \cdot (1+\pi z_j) \end{pmatrix} & \text{if } L_j \text{ is of type } I^o; \\ \pi^j \cdot \begin{pmatrix} a_j & 0 & 0 \\ 0 & (1+\sigma(\pi x_j))(1+\pi x_j) & (1+\sigma(\pi x_j))(1+\pi z_j) \\ 0 & (1+\sigma(\pi z_j))(1+\pi x_j) & (1+\pi z_j)\sigma(\pi z_j)+\pi z_j+2\bar{\gamma}_j \end{pmatrix} & \text{if } L_j \text{ is of type } I^e. \end{cases}$$

We write $x_j = x_j^1 + \pi x_j^2$ and $z_j = z_j^1 + \pi z_j^2$, where $x_j^1, x_j^2, z_j^1, z_j^2 \in R \subset R \otimes_A B$ and π stands for $1 \otimes \pi \in R \otimes_A B$. When L_j is of type I^o, by considering the $(2,2)$-block of the matrix above, we obtain the equation

$$\bar{\alpha}(z_j^1) + (z_j^1)^2 = 0.$$

Recall that α is the unit in B such that $\epsilon = 1 + \alpha\pi$ as explained in Section 2A, and $\bar{\alpha}$ is the image of α in κ.

Then this equation is equivalent to

$$(z_j^1/\bar{\alpha}) + (z_j^1/\bar{\alpha})^2 = 0$$

by dividing by $\bar{\alpha}^2$ in both sides. Therefore, in this case, F_j is isomorphic to $\mathbb{A}^1 \times \mathbb{Z}/2\mathbb{Z}$ as a κ-variety.

When L_j is of type I^e, by considering the $(2,2)$-block of the matrix above, we obtain the equation

$$\bar{\alpha}(x_j^1) + (x_j^1)^2 = 0.$$

We also consider the (2, 3)-block of the matrix above, and we obtain two equations

$$x_j^1 + z_j^1 = 0, \quad \bar{\alpha} x_j^1 + x_j^2 + z_j^2 + \bar{\alpha} x_j^1 z_j^1 = 0.$$

By considering the (3, 3)-block of the matrix above, we obtain the equation

$$\bar{\alpha}(z_j^1) + (z_j^1)^2 = 0.$$

By combining all these, we see that F_j is isomorphic to $\mathbb{A}^1 \times \mathbb{Z}/2\mathbb{Z}$ as a κ-variety. □

We introduce the final lemma in order to prove Lemma 4.6 below. This lemma is about the number of connected components in a short exact sequence of algebraic groups.

Lemma A.10. *Assume that there is a short exact sequence*

$$1 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 1$$

of linear algebraic groups over κ. Let $\pi_0(B)$ be the component group of B which is defined as the spectrum of the largest separable subalgebra $\pi_0(\kappa[B])$ of $\kappa[B]$, where $\kappa[B]$ is the coordinate ring of B. Let $\#(\pi_0(B))$ be the order of $\pi_0(B)$, which is defined as the dimension of $\pi_0(\kappa[B])$ as a κ-vector space. Note that B is connected if and only if $\pi_0(B)$ is trivial if and only if $\#(\pi_0(B)) = 1$. Thus $\#(\pi_0(B))$ is the number of connected components of $B \otimes_\kappa \bar{\kappa}$. Then

$$\#(\pi_0(B)) \leq \#(\pi_0(A)) \cdot \#(\pi_0(C)).$$

Moreover, the equality holds if A is connected and in this case, $\pi_0(B) = \pi_0(C)$.

Proof. By definition of a component group, there exists a surjective morphism $\pi : B \longrightarrow \pi_0(B)$ whose kernel is connected. Let A' ($\subseteq \pi_0(B)$) be the image of A under the morphism π. Notice that A' is a normal subgroup of $\pi_0(B)$ and that $\#(A') \leq \#(\pi_0(A))$. Then the morphism π induces a surjective morphism from C to $\pi_0(B)/A'$ and so $\#(\pi_0(B)/A') \leq \#(\pi_0(C))$. Therefore, $\#(\pi_0(B)) \leq \#(\pi_0(A)) \cdot \#(\pi_0(C))$.

It is clear that $\#(\pi_0(C)) \leq \#(\pi_0(B))$. Thus, if A is connected, then $\#(\pi_0(C)) = \#(\pi_0(B))$. In this case, since there exists a surjective morphism from B to $\pi_0(C)$ (through C), there exists a surjective morphism from $\pi_0(B)$ to $\pi_0(C)$. Since $\#(\pi_0(C)) = \#(\pi_0(B))$, we can conclude that $\pi_0(B) = \pi_0(C)$. □

We finally prove Lemma 4.6.

Proof. We start with the following short exact sequence

$$1 \longrightarrow \widetilde{G}^1 \longrightarrow \operatorname{Ker}\varphi \longrightarrow \operatorname{Ker}\varphi/\widetilde{G}^1 \longrightarrow 1.$$

It is obvious that $\operatorname{Ker}\varphi$ is smooth by Theorems A.4 and A.6. $\operatorname{Ker}\varphi$ is also unipotent since it is a subgroup of a unipotent group $\widetilde{M}^+$. Since $\widetilde{G}^1$ is connected by Theorem A.4, the component group of $\operatorname{Ker}\varphi$ is the same as that of $\operatorname{Ker}\varphi/\widetilde{G}^1$ by Lemma A.10. Moreover, the dimension of $\operatorname{Ker}\varphi$ is the sum of the dimension of $\widetilde{G}^1$ and the dimension of $\operatorname{Ker}\varphi/\widetilde{G}^1$. This completes the proof. □

Appendix B: Examples

In this appendix, we provide an example with a unimodular lattice (L, h) of rank 1. Let L be Be, a rank 1 hermitian lattice with hermitian form $h(le, l'e) = \sigma(l)l'$. With this lattice, we construct the smooth integral model and its special fiber and compute the local density.

B.1: ***Naive construction (without using our technique).*** We first construct the smooth integral model and its special fiber, without using any techniques introduced in this paper. If we write an element of L as $x + \pi y$ where $x, y \in A$, then it is easy to see that a naive integral model $\underline{G}'$ is Spec $A[x, y]/(x^2 + (\pi + \sigma(\pi))xy + \pi\sigma(\pi)y^2 - 1)$. As mentioned in Section 2A, we may assume that $\pi + \sigma(\pi) = 2$ and $\pi\sigma(\pi) = 2u$ for a unit $u \in A$. We remark that $\underline{G}'$ is smooth if $p \neq 2$, and in this case its special fiber is Spec $\kappa[x, y]/(x^2 - 1) = \mathbb{A}^1 \times \mu_2$ as a κ-variety. However, if $p = 2$, then its special fiber is no longer smooth since $\kappa[x, y]/(x^2 - 1) = \kappa[x, y]/(x - 1)^2$ is nonreduced. Some of the difficulty in the case $p = 2$ arises from this. The associated smooth integral model is obtained by a finite sequence of *dilatations* (at least once) of $\underline{G}'$ (cf. [Bosch et al. 1990]).

On the other hand, the difficulty can also be explained in terms of quadratic forms. Namely, the smoothness of any scheme over A should be closely related to the smoothness of its special fiber. If we define a function $q : L \longrightarrow A$ by $l \mapsto h(l, l)$, then $q \bmod 2$ is a quadratic form over κ. Therefore, the associated smooth integral model should contain information about this quadratic form, which is more subtle than quadratic forms over a field of characteristic not equal 2.

To construct the smooth integral model, we observe the characterization of $\underline{G}$ that $\underline{G}(R) = \underline{G}'(R)$ for an étale A-algebra R. Thus any element of $\underline{G}(R)$ is of the form $x + \pi y$ such that $x^2 + 2xy + 2uy^2 = 1$. Therefore, $(x - 1)^2$ is contained in the ideal (2) of R so that we can rewrite $x = 1 + 2x'$ since R is étale over A. With this, any element of $\underline{G}(R)$ is of the form $1 + 2x' + \pi y$ such that $y + uy^2 + 2(x' + (x')^2 + x'y) = 0$. We consider the affine scheme Spec $A[x, y]/(y + uy^2 + 2(x + x^2 + xy))$. Its special fiber is then reduced and smooth. Thus, this affine scheme is the desired smooth integral model $\underline{G}$. Furthermore, its special fiber Spec $\kappa[x, y]/(y + uy^2)$ is isomorphic to $\mathbb{A}^1 \times \mathbb{Z}/2\mathbb{Z}$ as a κ-variety so that the number of rational points is $2f$, where f is the cardinality of κ.

B.2: ***Construction following our technique.*** Define the map $q : L \to A$ via

$$l \mapsto h(l, l).$$

If we write $l = x + \pi y$ such that $x, y \in A$, then $q(l) = h(x + \pi y, x + \pi y) = x^2 + (\pi + \sigma(\pi))xy + \pi \cdot \sigma(\pi)y^2$. Thus $q \bmod 2$ is an additive polynomial over κ. Let $B(L)$ be the sublattice of L such that $B(L)/\pi L$ is the kernel of the additive polynomial $q \bmod 2$ on $L/\pi L$. In this case, $B(L) = \pi L$.

For an étale A-algebra R with $g \in \mathrm{Aut}_{B\otimes_A R}(L \otimes_A R, h \otimes_A R)$, it is easy to see that g induces the identity on $L/B(L) = L/\pi L$. Based on this, we construct the following functor from the category of commutative flat A-algebras to the category of monoids as follows. For any commutative flat A-algebra R, set

$$\underline{M}(R) = \{m \in \mathrm{End}_{B\otimes_A R}(L \otimes_A R)\} \mid m \text{ induces the identity on } L \otimes_A R/B(L) \otimes_A R\}.$$

This functor $\underline{M}$ is then representable by a polynomial ring and has the structure of a scheme of monoids. Let $\underline{M}^*(R)$ be the set of invertible elements in $\underline{M}(R)$ for any commutative A-algebra R. Then $\underline{M}^*$ is representable by a group scheme which is an open subscheme of $\underline{M}$ (Section 3B). Thus $\underline{M}^*$ is smooth. As a matrix, each element of $\underline{M}^*(R)$ for a flat A-algebra R can be written as $\big(1+\pi z\big)$.

We define another functor from the category of commutative flat A-algebras to the category of sets as follows. For any commutative flat A-algebra R, let $\underline{H}(R)$ be the set of hermitian forms f on $L\otimes_A R$ (with values in $B\otimes_A R$) such that $f(a,a) \bmod 2 = h(a,a) \bmod 2$, where $a\in L\otimes_A R$. As a matrix, each element of $\underline{M}^*(R)$ for a flat A-algebra R is $\big(1+2c\big)$.

Then for any flat A-algebra R, the group $\underline{M}^*(R)$ acts on the right of $\underline{H}(R)$ by $f\circ m = \sigma({}^t m)\cdot f\cdot m$ and this action is represented by an action morphism (Theorem 3.4)

$$\underline{H}\times\underline{M}^*\longrightarrow\underline{H}.$$

Let ρ be the morphism $\underline{M}^*\to\underline{H}$ defined by $\rho(m)=h\circ m$, which is obtained from the above action morphism. As a matrix, for a flat A-algebra R,

$$\rho(m)=\rho\big(\big(1+\pi z\big)\big)=\big(1+\pi z+\sigma(\pi z)+\pi\sigma(\pi)\cdot z\sigma(z)\big).$$

Then ρ is smooth of relative dimension 1 (Theorem 3.6). Let $\underline{G}$ be the stabilizer of h in $\underline{M}^*$. The group scheme $\underline{G}$ is smooth, and $\underline{G}(R)=\mathrm{Aut}_{B\otimes_A R}(L\otimes_A R, h\otimes_A R)$ for any étale A-algebra R (Theorem 3.8).

We now describe the structure of the special fiber $\widetilde{G}$ of $\underline{G}$. For a κ-algebra R, each element of $\underline{M}(R)$ (resp. $\underline{H}(R)$) can be written as a formal matrix $m=\big(1+\pi z\big)$ (resp. $f=\big(1+2c\big)$). Firstly, it is easy to see that $B_0=Y_0=\pi L$ so that the morphism φ in Section 4A is trivial.

For the component groups, as explained in Theorem 4.11, there is a surjective morphism from $\widetilde{G}$ to $\mathbb{Z}/2\mathbb{Z}$. Let us describe this morphism explicitly below. It is easy to see that $L^0=M_0=L$ and $C(L^0)=M_0'=L$. Here, we follow notation of Section 4B. Since $M_0=L$ is of type I^o, there exists a morphism from the special fiber $\widetilde{G}$ $(=G_0)$ to the special fiber of the smooth integral model associated to $M_0'\oplus C(L^0)=L\oplus L$ of type I^e as explained in the argument 2 just before Remark 4.10. Remark 4.10 tells us how to describe this morphism as formal matrices. Let (e_1,e_2) be a basis for $L\oplus L$ so that the associated Gram matrix of the hermitian lattice $L\oplus L$ with respect to this basis is $\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$. Then we consider the basis (e_1,e_1+e_2), with respect to which the morphism described in Remark 4.10 is given as

$$\big(1+\pi z\big)\mapsto\begin{pmatrix}1&-\pi z\\0&1+\pi z\end{pmatrix}.$$

We now construct a morphism from the special fiber of the smooth integral model associated to $M_0'\oplus C(L^0)=L\oplus L$ to $\mathbb{Z}/2\mathbb{Z}$ and describe the image of $\left(\begin{smallmatrix}1&-\pi z\\0&1+\pi z\end{smallmatrix}\right)$ in $\mathbb{Z}/2\mathbb{Z}$.

Let R be a κ-algebra. The Gram matrix for the hermitian lattice $L\oplus L$ with respect to the basis (e_1,e_1+e_2) is $\left(\begin{smallmatrix}1&1\\1&2\end{smallmatrix}\right)$. Since $L\oplus L$ is *unimodular of type* I^e, an R-point of the special fiber associated to $L\oplus L$ with respect to this basis is expressed as the formal matrix $\left(\begin{smallmatrix}1+\pi x'&\pi z'\\u'&1+\pi w'\end{smallmatrix}\right)$, as explained in Section 3B. Based on argument (1) following

Definition 4.9, the morphism mapping to $\mathbb{Z}/2\mathbb{Z}$ factors through the special fiber associated to $C(L\oplus L)$, composed with the Dickson invariant associated to the corresponding orthogonal group. $C(L\oplus L)$ is then generated by $(\pi e_1, e_1+e_2)$ and is π^1-modular. Thus there is no congruence condition on an element of the smooth integral model associated to $C(L\oplus L)$ as explained in Section 3B. Write $x'=x_1'+\pi x_2'$, $y'=y_1'+\pi y_2'$, and $z'=z_1'+\pi z_2'$. The image of $\begin{pmatrix} 1+\pi x' & \pi z' \\ u' & 1+\pi w' \end{pmatrix}$ in the special fiber associated to $C(L\oplus L)$ is $\begin{pmatrix} 1+\pi x_1' & z_1'+\pi z_2' \\ \pi u_1' & 1+\pi w_1' \end{pmatrix}$. Since $C(L\oplus L)$ is π^1-modular with rank 2, there is a morphism from the special fiber associated to $C(L\oplus L)$ to the orthogonal group associated to $C(L\oplus L)/\pi C(L\oplus L)$, as described in Theorem 4.4 or Remark 4.7. Then the image of $\begin{pmatrix} 1+\pi x_1' & z_1'+\pi z_2' \\ \pi u_1' & 1+\pi w_1' \end{pmatrix}$ in this orthogonal group is $\begin{pmatrix} 1 & z_1' \\ 0 & 1 \end{pmatrix}$. The Dickson invariant of $\begin{pmatrix} 1 & z_1' \\ 0 & 1 \end{pmatrix}$ is $z_1'/\bar{\alpha}$ as mentioned in step (1) of the proof of Theorem 4.11. Here, α is the unit in B such that $\epsilon=1+\alpha\pi$ as explained in Section 2A, and $\bar{\alpha}$ is the image of α in κ.

In conclusion, the image of $(1+\pi z)$, which is an element of $\widetilde{G}(R)$ for a κ-algebra R, in $\mathbb{Z}/2\mathbb{Z}$ is $z_1/\bar{\alpha}$, where we write $z=z_1+\pi z_2$. On the other hand, the equation defining $\widetilde{G}$ is $\bar{\alpha}z_1+z_1^2=0$ which is equivalent to $\frac{z_1}{\bar{\alpha}}+(\frac{z_1}{\bar{\alpha}})^2=0$. Thus, the morphism from $\widetilde{G}$ to $\mathbb{Z}/2\mathbb{Z}$ is surjective. Therefore the maximal reductive quotient of $\widetilde{G}$ is $\mathbb{Z}/2\mathbb{Z}$ and using Remark 5.3,

$$\#(\widetilde{G}(\kappa)) = \#(\mathbb{Z}/2\mathbb{Z})\cdot\#(\mathbb{A}^1) = 2f,$$

where f is the cardinality of κ. Based on Theorem 5.2, the local density is

$$\beta_L = f^0\cdot 2f = 2f.$$

Acknowledgements

The author greatly thanks the referee for putting incredible time and effort into reading this paper, and for providing a lot of valuable feedback. The author owes a special debt to Professor Brian Conrad for his immense patience and copious suggestions, which helped make this paper substantially more pleasant to read. This paper originated from the paper [Gan and Yu 2000] and the author's Ph.D. dissertation, and the author would like to express his deep appreciation to Professor Wee Teck Gan and Professor Jiu-Kang Yu. The author would like to thank his Ph.D. thesis advisor Jiu-Kang Yu for many valuable comments. In addition, the author would like to thank Professor Wai Kiu Chan, Professor Benedict H. Gross, and Professor Gopal Prasad for their interest in this project and their encouragement. The author would like to thank Radhika Ganapathy, Bogume Jang, Manish Mishra, Marco Rainho and Sandeep Varma for carefully reading a draft of this paper to help reduce the typographical errors and improve the presentation of this paper.

References

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Results in Mathematics and Related Areas (3) **21**, Springer, Berlin, 1990. MR 91i:14034 Zbl 0705.14001

[Cho 2015a] S. Cho, "Group schemes and local densities of quadratic lattices in residue characteristic 2", *Compos. Math.* **151**:5 (2015), 793–827. MR 3347991

[Cho 2015b] S. Cho, "Group schemes and local densities of ramified hermitian lattices in residue characteristic 2 Part II", preprint, 2015, Available at https://sites.google.com/site/sungmuncho12/.

[Conway and Sloane 1988] J. H. Conway and N. J. A. Sloane, "Low-dimensional lattices. IV. The mass formula", *Proc. Roy. Soc. London Ser. A* **419**:1857 (1988), 259–286. MR 90a:11074 Zbl 0655.10023

[Demazure and Gabriel 1970] M. Demazure and P. Gabriel, *Groupes algébriques. Tome I: Géométrie algébrique, généralités, groupes commutatifs*, Masson, Paris, 1970. MR 46 #1800 Zbl 0203.23401

[Gan and Yu 2000] W. T. Gan and J.-K. Yu, "Group schemes and local densities", *Duke Math. J.* **105**:3 (2000), 497–524. MR 2001m:11060 Zbl 1048.11028

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

[Hironaka 1998] Y. Hironaka, "Local zeta functions on Hermitian forms and its application to local densities", *J. Number Theory* **71**:1 (1998), 40–64. MR 99e:11045 Zbl 0932.11072

[Hironaka 1999] Y. Hironaka, "Spherical functions and local densities on Hermitian forms", *J. Math. Soc. Japan* **51**:3 (1999), 553–581. MR 2000c:11064 Zbl 0936.11024

[Jacobowitz 1962] R. Jacobowitz, "Hermitian forms over local fields", *Amer. J. Math.* **84** (1962), 441–465. MR 27 #131 Zbl 0118.01901

[Kitaoka 1993] Y. Kitaoka, *Arithmetic of quadratic forms*, Cambridge Tracts in Mathematics **106**, Cambridge University Press, Cambridge, 1993. MR 95c:11044 Zbl 0785.11021

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, American Mathematical Society Colloquium Publications **44**, Amer. Math. Soc., Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[Mischler 2000] M. Mischler, "Local densities of Hermitian forms", pp. 201–208 in *Quadratic forms and their applications* (Dublin, 1999), edited by E. Bayer-Fluckiger et al., Contemp. Math. **272**, Amer. Math. Soc., Providence, RI, 2000. MR 2001m:11055 Zbl 1028.11023

[Pall 1965] G. Pall, "The weight of a genus of positive n-ary quadratic forms", pp. 95–105 in *Theory of Numbers, Proc. Sympos. Pure Math., Vol. VIII*, edited by A. L. Whiteman, Amer. Math. Soc., Providence, R.I., 1965. MR 31 #3386 Zbl 0136.03101

[Sah 1960] C.-h. Sah, "Quadratic forms over fields of characteristic 2", *Amer. J. Math.* **82** (1960), 812–830. MR 22 #10954 Zbl 0100.25308

[Sato and Hironaka 2000] F. Sato and Y. Hironaka, "Local densities of representations of quadratic forms over p-adic integers (the non-dyadic case)", *J. Number Theory* **83**:1 (2000), 106–136. MR 2001e:11035 Zbl 0949.11023

[SGA 3_{I} 1970] M. Demazure and A. Grothendieck, *Schémas en groupes, Tome I: Propriétés générales des schémas en groupes, Exposés I–VII* (Séminaire de Géométrie Algébrique du Bois Marie 1962–1964), Lecture Notes in Math. **151**, Springer, Berlin, 1970. MR 43 #223a Zbl 0207.51401

[Waterhouse 1979] W. C. Waterhouse, *Introduction to affine group schemes*, Graduate Texts in Mathematics **66**, Springer, New York, 1979. MR 82e:14003 Zbl 0442.14017

[Watson 1976] G. L. Watson, "The 2-adic density of a quadratic form", *Mathematika* **23**:1 (1976), 94–106. MR 57 #16195 Zbl 0326.10022

[Yu 2002] J.-K. Yu, "Smooth models associated to concave functions in Bruhat–Tits theory", preprint, 2002, Available at http://www2.ims.nus.edu.sg/preprints/2002-20.pdf.

Communicated by Brian Conrad
Received 2013-08-30 Revised 2015-09-15 Accepted 2015-10-25

sungmuncho12@gmail.com *Department of Mathematics, University of Toronto, 40 St. George St., Room 6290, Toronto ON M5S 2E4, Canada*

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.533

Presentation of affine Kac–Moody groups over rings

Daniel Allcock

Tits has defined Steinberg groups and Kac–Moody groups for any root system and any commutative ring R. We establish a Curtis–Tits-style presentation for the Steinberg group $\mathfrak{St}$ of any irreducible affine root system with rank ≥ 3, for any R. Namely, $\mathfrak{St}$ is the direct limit of the Steinberg groups coming from the 1- and 2-node subdiagrams of the Dynkin diagram. In fact, we give a completely explicit presentation. Using this we show that $\mathfrak{St}$ is finitely presented if the rank is ≥ 4 and R is finitely generated as a ring, or if the rank is 3 and R is finitely generated as a module over a subring generated by finitely many units. Similar results hold for the corresponding Kac–Moody groups when R is a Dedekind domain of arithmetic type.

1. Introduction

Suppose R is a commutative ring and A is one of the *ABCDEFG* Dynkin diagrams, or equivalently its Cartan matrix. Steinberg [1968] defined what is now called the Steinberg group $\mathfrak{St}_A(R)$, by generators and relations. It plays a central role in K-theory and some aspects of Lie theory.

Kac–Moody algebras are infinite-dimensional generalizations of the semisimple Lie algebras. When $R = \mathbb{R}$ and A is an affine Dynkin diagram, the corresponding Kac–Moody group is a central extension of the loop group of a finite-dimensional Lie group. For a general ring R and any generalized Cartan matrix A, the definition of a Kac–Moody group is due to Tits [1987]. A difficulty in tracing the story is that Tits began by defining a "Steinberg group" which unfortunately differs from Steinberg's original group when A has an A_1 component. This was resolved by Morita and Rehmann [1990] by adding extra relations to Tits' definition. So there are two definitions of the Steinberg group. Increasing the chance of confusion, they agree for most A of interest, including the irreducible affine diagrams of rank ≥ 3. We follow Morita and Rehmann, so the Steinberg group $\mathfrak{St}_A(R)$ reduces to

Supported by NSF grant DMS-1101566.
MSC2010: primary 20G44; secondary 14L15, 22E67, 19C99.
Keywords: affine Kac–Moody group, Steinberg group, Curtis–Tits presentation.

Steinberg's original group when this is defined. See Section 3 for further background on $\mathfrak{St}$.

Tits then defined another functor $R \mapsto \tilde{\mathfrak{G}}_A(R)$ as a quotient of his version of the Steinberg group. In this paper we will omit the tilde and refer to $\mathfrak{G}_A(R)$ as the Kac–Moody group of type A over R. The relations added by Morita and Rehmann to the definition of $\mathfrak{St}_A(R)$ are among the relations that Tits imposed in his definition of $\mathfrak{G}_A(R)$. Therefore, we may regard $\mathfrak{G}_A(R)$ as a quotient of $\mathfrak{St}_A(R)$, just as Tits did, even though our $\mathfrak{St}_A(R)$ is not quite the same as his. (Tits actually defined $\tilde{\mathfrak{G}}_D(R)$ where D is a root datum; by $\mathfrak{G}_A(R)$ we refer to the root datum whose generalized Cartan matrix is A and which is "simply connected in the strong sense" [Tits 1987, p. 551]. The general case differs from this one by enlarging or shrinking the center of $\tilde{\mathfrak{G}}_D(R)$.) See Section 3 for further background on $\mathfrak{G}$.

The meaning of "Kac–Moody group" is far from standardized. Tits [1987] wrote down axioms (KMG1)–(KMG9) that one could demand of a functor from rings to groups before calling it a Kac–Moody functor. He showed in [loc. cit., Theorem 1′] that any such functor admits a natural homomorphism from $\mathfrak{G}_A$, which is an isomorphism at every field. So Kac–Moody groups over fields are well-defined, and over general rings $\mathfrak{G}_A$ approximates the yet unknown ultimate definition. This is why we refer to $\mathfrak{G}_A$ as the Kac–Moody group. But $\mathfrak{G}_A$ does not quite satisfy Tits' axioms, so ultimately some other language may be better. See Section 6 for more on this.

The purpose of this paper is to simplify Tits' presentations of $\mathfrak{St}_A(R)$ and $\mathfrak{G}_A(R)$ when A is an affine Dynkin diagram of rank (number of nodes) ≥ 3. We will always take affine diagrams to be irreducible. We will show that $\mathfrak{St}_A(R)$ and $\mathfrak{G}_A(R)$ are finitely presented under quite weak hypotheses on R. This is surprising because there is no obvious reason for an infinite-dimensional group over (say) $\mathbb{Z}$ to be finitely presented, and Tits' presentations are "very" infinite. His generators are indexed by all pairs (root, ring element), and his relations specify the commutators of many pairs of these generators. Subtle implicitly defined coefficients appear throughout his relations.

The main step in proving our finite presentation results is to first establish smaller, and more explicit, presentations for $\mathfrak{St}_A(R)$ and $\mathfrak{G}_A(R)$. These presentations are not necessarily finite, but they do apply to all R. In [Allcock 2015] we wrote down a presentation for a group functor we called the pre-Steinberg group $\mathfrak{PSt}_A$. We have reproduced it in Section 2, for any generalized Cartan matrix A. The generators are S_i and $X_i(t)$ with i varying over the nodes of the Dynkin diagram and t varying over R. The relations are (2-1)–(2-28), but (2-27)–(2-28) may be omitted when A is 2-spherical (it has no edges labeled "∞") and has no A_1 components. This case includes all affine diagrams of rank ≥ 3. The only way the presentation fails to be finite is that the $X_i(t)$ are parameterized by elements of R, and each Chevalley relation is parameterized by pairs of elements of R.

The name "pre-Steinberg group" reflects the fact that there is a natural map from $\mathfrak{PSt}_A(R)$ to the Steinberg group $\mathfrak{St}_A(R)$. In Section 3 we will describe this in a conceptual manner. But in terms of presentations it suffices to say that our $X_i(t)$ and S_i map to the group elements $x_{\alpha_i}(t)$ and $\hat{w}_{\alpha_i}(1)$ in Morita and Rehmann's definition of $\mathfrak{St}_A(R)$ [1990, §2]. Our general philosophy is that $\mathfrak{PSt}_A(R)$ is interesting only as a means of approaching $\mathfrak{St}_A(R)$, as in the following theorem, which is our main result.

Theorem 1.1 (presentation of affine Steinberg and Kac–Moody groups). *Suppose A is an affine Dynkin diagram of rank ≥ 3 and R is a commutative ring. Then the natural map from the pre-Steinberg group $\mathfrak{PSt}_A(R)$ to the Steinberg group $\mathfrak{St}_A(R)$ is an isomorphism. In particular, $\mathfrak{St}_A(R)$ has a presentation with generators S_i and $X_i(t)$, with i varying over the simple roots and t over R, and relations* (2-1)–(2-26).

One obtains Tits' Kac–Moody group $\mathfrak{G}_A(R)$ by adjoining the relations

$$\tilde{h}_i(u)\tilde{h}_i(v) = \tilde{h}_i(uv) \tag{1-1}$$

for all simple roots i and all units u, v of R, where

$$\tilde{h}_i(u) := \tilde{s}_i(u)\tilde{s}_i(-1), \quad \tilde{s}_i(u) := X_i(u)S_iX_i(1/u)S_i^{-1}X_i(u).$$

We remark that if A is a spherical diagram (that is, its Weyl group is finite) then it follows immediately from an alternate description of $\mathfrak{PSt}_A$ that $\mathfrak{PSt}_A \to \mathfrak{St}_A$ is an isomorphism; see Section 3 or [Allcock 2015, §7]. So Theorem 1.1 extends the isomorphism $\mathfrak{PSt}_A \cong \mathfrak{St}_A$ from the spherical case to the affine case, except for the two affine diagrams of rank 2. See [Allcock and Carbone 2016] for a further extension, to the simply laced hyperbolic case, and [Allcock 2015] for generalizations beyond the hyperbolic case.

For a moment we return to the case where A is an arbitrary generalized Cartan matrix. If $B_1 \subseteq B_2$ are two subdiagrams of A then there is a natural homomorphism $\mathfrak{PSt}_{B_1}(R) \to \mathfrak{PSt}_{B_2}(R)$. This is because the generators and relations of $\mathfrak{PSt}_{B_1}(R)$ are among those of $\mathfrak{PSt}_{B_2}(R)$, by the fact that our presentations of these groups are defined in terms of the nodes and edges of these subdiagrams of A. Using these maps, we consider the directed system of groups $\mathfrak{PSt}_B(R)$, where B varies over the subdiagrams of A of rank ≤ 2. It is a formality that the direct limit is $\mathfrak{PSt}_A(R)$; this is just an abstract way of saying that each generator or relation of $\mathfrak{PSt}_A(R)$ already appears in the presentation of some $\mathfrak{PSt}_B(R)$ with B of rank ≤ 2.

When A is affine of rank ≥ 3, $\mathfrak{PSt}_A(R) \to \mathfrak{St}_A(R)$ is an isomorphism by Theorem 1.1. And $\mathfrak{PSt}_B(R) \to \mathfrak{St}_B(R)$ is an isomorphism for every proper subdiagram B of A, since such subdiagrams are spherical. It follows that we may replace $\mathfrak{PSt}$ by $\mathfrak{St}$ throughout the preceding paragraph, proving the following result. The point is that affine Steinberg groups of rank ≥ 3 are built up from the classical Steinberg groups of types A_1, A_1^2, A_2, B_2 and G_2.

Corollary 1.2 (Curtis–Tits presentation). *Suppose A is an affine Dynkin diagram of rank ≥ 3 and R is a commutative ring. Then $\mathfrak{St}_A(R)$ is the direct limit of the groups $\mathfrak{St}_B(R)$, where B varies over the subdiagrams of A of rank ≤ 2, and the maps between these groups are as specified above. The same result also holds with $\mathfrak{St}$ replaced by $\mathfrak{G}$ throughout.* □

An informal way to restate Corollary 1.2 is that a presentation for $\mathfrak{St}_A(R)$ can be got by amalgamating one's favorite presentations for the $\mathfrak{St}_B(R)$. Splitthoff [1986] discovered quite weak sufficient conditions for the latter groups to be finitely presented. When these hold, one would therefore expect $\mathfrak{St}_A(R)$ also to be finitely presented. The next theorem expresses this idea precisely. Claim (ii) is part of [Allcock 2015, Theorem 1.4]. See Section 6 for the proof of claim (i).

Theorem 1.3 (finite presentability). *Suppose A is an affine Dynkin diagram and R is any commutative ring. Then the Steinberg group $\mathfrak{St}_A(R)$ is finitely presented as a group if either*

(i) $\operatorname{rk} A > 3$ *and R is finitely generated as a ring*, *or*

(ii) $\operatorname{rk} A = 3$ *and R is finitely generated as a module over a subring generated by finitely many units.*

In either case, if the unit group of R is finitely generated as an abelian group, then Tits' Kac–Moody group $\mathfrak{G}_A(R)$ is finitely presented as a group.

One of the main motivations for Splitthoff's work was to understand when the Chevalley–Demazure groups, over Dedekind domains of interest in number theory, are finitely presented. This was finally settled by Behr [1967; 1998], capping a long series of works by many authors. The following analogue of these results follows immediately from Theorem 1.3. How close the analogy is depends on how well $\mathfrak{G}_A$ approximates whatever plays the role of the Chevalley–Demazure group scheme in the setting of Kac–Moody theory.

Corollary 1.4 (finite presentation in arithmetic contexts). *Suppose K is a global field, meaning a finite extension of $\mathbb{Q}$ or $\mathbb{F}_q(t)$. Suppose S is a nonempty finite set of places of K, including all infinite places in the number field case. Let R be the ring of S-integers in K.*

Suppose A is an affine Dynkin diagram. Then Tits' Kac–Moody group $\mathfrak{G}_A(R)$ is finitely presented if $\operatorname{rk} A \geq 3$, *unless K is a function field and $|S| = 1$, when* $\operatorname{rk} A > 3$ *suffices.* □

We remark that if R is a field then the $\mathfrak{G}_A$ case of Corollary 1.2 is due to Abramenko and Mühlherr [1997] (see also [Devillers and Mühlherr 2007]). Namely, suppose A is any generalized Cartan matrix which is 2-spherical, and that R is a field (but not $\mathbb{F}_2$ if A has a double bond, and neither $\mathbb{F}_2$ nor $\mathbb{F}_3$ if A has a triple bond). Then $\mathfrak{G}_A(R)$ is the direct limit of the groups $\mathfrak{G}_B(R)$. Abramenko and Mühlherr

[1997, p. 702] state that if A is affine then one can remove the restrictions $R \neq \mathbb{F}_2, \mathbb{F}_3$.

One of our goals is to bring Kac–Moody groups into the world of geometric and combinatorial group theory, which mostly addresses finitely presented groups. For example, which Kac–Moody groups admit classifying spaces with finitely many cells below some chosen dimension? What other finiteness properties do they have? Do they have Kazhdan's property T? What isoperimetric inequalities do they satisfy in various dimensions? Are there (nonsplit) Kac–Moody groups over local fields whose uniform lattices (suitably defined) are word hyperbolic? Are some Kac–Moody groups (or classes of them) quasi-isometrically rigid? We find the last question very attractive, since the corresponding answer for lattices in Lie groups is deep (see [Eskin and Farb 1997; Farb and Schwartz 1996; Kleiner and Leeb 1997; Schwartz 1995]).

Regarding property T we would like to mention work of Hartnick and Köhl [2015] who showed that many Kac–Moody groups over local fields have property T when equipped with the Kac–Peterson topology. Also, Shalom [1999] and Neuhauser [2003] respectively showed that the loop groups of (i.e., the spaces of continuous maps from S^1 to) $\mathrm{SL}_n(\mathbb{C})$ and $\mathrm{Sp}_{2n}(\mathbb{C})$ have property T.

2. Presentation of the pre-Steinberg group $\mathfrak{PSt}_A(R)$

Suppose R is any commutative ring and A is any generalized Cartan matrix. Write I for the set of A's nodes, and for $i, j \in I$ write m_{ij} for the order of the product of the corresponding generators of the Weyl group. Following [Allcock 2015, §7], the pre-Steinberg group $\mathfrak{PSt}_A(R)$ is defined by the following presentation. The generators are S_i and $X_i(t)$ with $t \in R$. The relations are (2-1)–(2-28) below, in which i, j vary over I and t, u vary over R. We use the notation $Y \rightleftarrows Z$ to say that Y and Z commute.

If A has no A_1 components and is 2-spherical (all m_{ij} are finite), then the last two relations (2-27)–(2-28) follow from the others and may be omitted [Allcock 2015, Remark 7.13]. If A is affine of rank ≥ 3 then it satisfies this condition, and our main result (Theorem 1.1) is that the presentation equally well defines the Steinberg group $\mathfrak{St}_A(R)$.

For every $i \in I$ we impose the relations

$$X_i(t)X_i(u) = X_i(t+u), \tag{2-1}$$

$$S_i = X_i(1)S_iX_i(1)S_i^{-1}X_i(1). \tag{2-2}$$

For all i, j we impose the relations

$$S_i^2 S_j S_i^{-2} = S_j^{\varepsilon}, \tag{2-3}$$

$$S_i^2 X_j(t) S_i^{-2} = X_j(\varepsilon t), \tag{2-4}$$

where $\varepsilon = (-1)^{A_{ij}}$.

Whenever $m_{ij}=2$ we impose the relations

$$S_i S_j = S_j S_i, \tag{2-5}$$
$$S_i \rightleftarrows X_j(t), \tag{2-6}$$
$$X_i(t) \rightleftarrows X_j(u). \tag{2-7}$$

Whenever $m_{ij}=3$ we impose the relations

$$S_i S_j S_i = S_j S_i S_j, \tag{2-8}$$
$$S_j S_i X_j(t) = X_i(t) S_j S_i, \tag{2-9}$$
$$X_i(t) \rightleftarrows S_i X_j(u) S_i^{-1}, \tag{2-10}$$
$$[X_i(t), X_j(u)] = S_i X_j(tu) S_i^{-1}. \tag{2-11}$$

Whenever $m_{ij}=4$ we impose the following relations; in (2-14)–(2-17), s (resp. l) refers to whichever of i and j is the shorter (resp. longer) root:

$$S_i S_j S_i S_j = S_j S_i S_j S_i, \tag{2-12}$$
$$S_i S_j S_i \rightleftarrows X_j(t), \tag{2-13}$$
$$S_s X_l(t) S_s^{-1} \rightleftarrows S_l X_s(u) S_l^{-1}, \tag{2-14}$$
$$X_l(t) \rightleftarrows S_s X_l(u) S_s^{-1}, \tag{2-15}$$
$$[X_s(t), S_l X_s(u) S_l^{-1}] = S_s X_l(-2tu) S_s^{-1}, \tag{2-16}$$
$$[X_s(t), X_l(u)] = S_l X_s(-tu) S_l^{-1} \cdot S_s X_l(t^2u) S_s^{-1}. \tag{2-17}$$

Whenever $m_{ij}=6$ we impose the following relations; s and l have the same meaning they had in the previous paragraph:

$$S_i S_j S_i S_j S_i S_j = S_j S_i S_j S_i S_j S_i, \tag{2-18}$$
$$S_i S_j S_i S_j S_i \rightleftarrows X_j(t), \tag{2-19}$$
$$X_l(t) \rightleftarrows S_l S_s X_l(u) S_s^{-1} S_l^{-1}, \tag{2-20}$$
$$S_s S_l X_s(t) S_l^{-1} S_s^{-1} \rightleftarrows S_l S_s X_l(u) S_s^{-1} S_l^{-1}, \tag{2-21}$$
$$S_s X_l(t) S_s^{-1} \rightleftarrows S_l X_s(u) S_l^{-1}, \tag{2-22}$$
$$[X_l(t), S_s X_l(u) S_s^{-1}] = S_l S_s X_l(tu) S_s^{-1} S_l^{-1}, \tag{2-23}$$
$$[X_s(t), S_s S_l X_s(u) S_l^{-1} S_s^{-1}] = S_s X_l(3tu) S_s^{-1}, \tag{2-24}$$
$$\begin{aligned}[X_s(t), S_l X_s(u) S_l^{-1}] = S_s S_l X_s(-2tu) S_l^{-1} S_s^{-1} \cdot S_s X_l(-3t^2u) S_s^{-1} \\ \cdot S_l S_s X_l(-3tu^2) S_s^{-1} S_l^{-1},\end{aligned} \tag{2-25}$$
$$\begin{aligned}[X_s(t), X_l(u)] = S_s S_l X_s(t^2u) S_l^{-1} S_s^{-1} \cdot S_l X_s(-tu) S_l^{-1} \\ \cdot S_s X_l(t^3u) S_s^{-1} \cdot S_l S_s X_l(-t^3u^2) S_s^{-1} S_l^{-1}.\end{aligned} \tag{2-26}$$

This paper	[Allcock 2015]
(2-1)	(7.4)
(2-2)	(7.26)
(2-3)	(7.2)–(7.3)
(2-4)	(7.5)
(2-5) ∪ (2-8) ∪ (2-12) ∪ (2-18)	(7.1)
(2-6)	(7.6)
(2-7)	(7.10), the A_1^2 Chevalley relation
(2-9)	(7.7)
(2-10)–(2-11)	(7.11)–(7.12), the A_2 Chevalley relations
(2-13)	(7.8)
(2-14)–(2-17)	(7.13)–(7.16), the B_2 Chevalley relations
(2-19)	(7.9)
(2-20)–(2-26)	(7.17)–(7.23), the G_2 Chevalley relations
(2-27)	(7.24)
(2-28)	(7.25)

Table 1. Correspondence between our relations and those of [Allcock 2015].

Officially, the next two relations are part of the presentation of $\mathfrak{PSt}_A(R)$. But as mentioned above, they may be omitted if A is 2-spherical without A_1 components. We let r vary over the units of R and impose the relations

$$\tilde{h}_i(r)X_j(t)\tilde{h}_i(r)^{-1} = X_j(r^{A_{ij}}t), \tag{2-27}$$

$$\tilde{h}_i(r)\, S_j X_j(t) S_j^{-1}\, \tilde{h}_i(r)^{-1} = S_j X_j(r^{-A_{ij}}t) S_j^{-1}, \tag{2-28}$$

where $\tilde{h}_i(r)$ is defined in Theorem 1.1.

Because we have organized the relations differently than we did in [Allcock 2015], we will state the correspondence explicitly. See Table 1.

3. Steinberg and pre-Steinberg groups

Our goal in this section is to describe the Steinberg group and to give a second description of the pre-Steinberg group. This description makes visible the latter group's natural map to the Steinberg group, and is the form we will use for our calculations in Section 5.

We work in the setting of [Tits 1987] and [Allcock 2015], so R is a commutative ring and A is a generalized Cartan matrix. This matrix determines a complex Lie algebra $\mathfrak{g}$ called the Kac–Moody algebra, and we write Φ for the set of real roots of $\mathfrak{g}$. For each real root α, its root space $\mathfrak{g}_\alpha$ comes with a distinguished pair of (complex vector space) generators, each the negative of the other. We write $\mathfrak{g}_{\alpha,\mathbb{Z}}$ for their integral span, and we define the root group $\mathfrak{U}_\alpha$ as $\mathfrak{g}_{\alpha,\mathbb{Z}} \otimes R \cong R$. Tits' definition of the Steinberg group begins with the free product $*_{\alpha\in\Phi}\, \mathfrak{U}_\alpha$.

We emphasize that there is no natural way to choose an isomorphism $R \to \mathfrak{U}_\alpha$. If $\pm e$ are the two distinguished generators for $\mathfrak{g}_\alpha$, then there are two natural choices for the parameterization of $\mathfrak{U}_\alpha$, namely $t \mapsto (\pm e) \otimes t$. Often we will choose one of these and call it X_α; we speak of this as a "sign choice". Making such a choice makes computations more concrete, but breaks the symmetry.

In Tits' definition of $\mathfrak{St}_A(R)$, the relations have the following form. He calls a pair $\alpha, \beta \in \Phi$ *prenilpotent* if some element of the Weyl group W sends both α, β to positive roots, and some other element of W sends both to negative roots. A consequence of this condition is that every root in $\mathbb{N}\alpha + \mathbb{N}\beta$ is real, which enables Tits to write down Chevalley-style relators for α, β. That is, for every prenilpotent pair α, β he imposes relations of the form

$$[\text{element of } \mathfrak{U}_\alpha, \text{element of } \mathfrak{U}_\beta] = \prod_{\gamma \in \theta(\alpha,\beta) - \{\alpha,\beta\}} (\text{element of } \mathfrak{U}_\gamma), \tag{3-1}$$

where $\theta(\alpha, \beta) := (\mathbb{N}\alpha + \mathbb{N}\beta) \cap \Phi$ and $\mathbb{N} = \{0, 1, 2, \dots\}$. The exact relations are given in a rather implicit form in [Tits 1987, §3.6]. Writing them down explicitly requires choosing parameterizations of $\mathfrak{U}_\alpha$, $\mathfrak{U}_\beta$ and each $\mathfrak{U}_\gamma$. We suppose this has been done as above, with the parameterizations being X_α, X_β and the various X_γ. Then the relations take the form

$$[X_\alpha(t), X_\beta(u)] = \prod_{\substack{\text{roots } \gamma = m\alpha + n\beta \\ \text{with } m,n \geq 1}} X_\gamma\big(N_{\alpha\beta\gamma} t^m u^n\big), \tag{3-2}$$

where the $N_{\alpha\beta\gamma}$ are integers determined by the structure constants of $\mathfrak{g}$, the sign choices made in parameterizing the root groups, and the ordering of the terms on the right side. See [Tits 1987, §§3.4–3.6] for details, or Section 5 for the cases we will need. Morita [1987; 1988] showed that the right side has at most 1 term except when $(\mathbb{Q}\alpha \oplus \mathbb{Q}\beta) \cap \Phi$ has type B_2 or G_2, and found simple formulas for the constants (up to sign).

For Tits, this is the end of the definition of the Steinberg group. We called this group $\mathfrak{St}_A^{\text{Tits}}(R)$ in [Allcock 2015] to avoid confusion with $\mathfrak{St}_A(R)$ itself, which we take to also satisfy the Morita–Rehmann relations. These extra relations play the role of making the "maximal torus" and "Weyl group" in $\mathfrak{St}_A^{\text{Tits}}(R)$ act in the expected way on root spaces. These relations follow from the Chevalley relations when A is 2-spherical without A_1 components, so the reader could skip down to the definition of the Kac–Moody group $\mathfrak{G}_A(R)$.

Here is a terse description of the Morita–Rehmann relations; see [Morita and Rehmann 1990, Relations (B′)] or [Allcock 2015, §6] for details. For each simple root $\alpha \in \Phi$ and each of the two choices e for a generator of $\mathfrak{g}_{\alpha,\mathbb{Z}}$, we impose relations as follows. By a standard construction, the choice of e distinguishes a generator f for $\mathfrak{g}_{-\alpha,\mathbb{Z}}$. Using e and f as above, we obtain parameterizations of $\mathfrak{U}_\alpha$ and $\mathfrak{U}_{-\alpha}$

which we will call X_e and X_f. For $r \in R^*$ we define $\tilde{s}_e(r) = X_e(r)X_f(1/r)X_e(r)$ and $\tilde{h}_e(r) = \tilde{s}_e(r)\tilde{s}_e(-1)$. Morita and Rehmann impose relations that describe the actions of $\tilde{s}_e(1)$ and $\tilde{h}_e(r)$ on every $\mathfrak{U}_\beta$, where β varies over Φ. First, conjugation by $\tilde{s}_e(1)$ sends $\mathfrak{U}_\beta$ to $\mathfrak{U}_{s_\alpha(\beta)}$ in the same way that $s_e^* := (\exp \operatorname{ad}_e)(\exp \operatorname{ad}_f)(\exp \operatorname{ad}_e) \in \operatorname{Aut} \mathfrak{g}$ does. (Here s_α is the reflection in α, and for the relation to make sense one must check that s_e^* sends $\mathfrak{g}_{\beta,\mathbb{Z}}$ to $\mathfrak{g}_{s_\alpha(\beta),\mathbb{Z}}$.) Second, every $\tilde{h}_e(r)$ acts on $\mathfrak{U}_\beta \cong R$ by scaling by $r^{\langle \alpha^\vee, \beta \rangle}$, where $\alpha^\vee$ is the coroot associated to α.

The quotient of $\mathfrak{St}_A^{\mathrm{Tits}}(R)$ by all these relations is the definition of the Steinberg group $\mathfrak{St}_A(R)$ and agrees with Steinberg's original group when A is spherical. We remark that we let e vary over both possible choices of generator for $\mathfrak{g}_{\alpha,\mathbb{Z}}$ just to avoid choosing one. But one could choose one without harm, because it turns out that the relations imposed for e are the same as those imposed for $-e$. Also, Morita and Rehmann write $\hat{w}_\alpha$ rather than $\tilde{s}_e$, and their definition of it uses $X_f(-1/r)$ rather than $X_f(1/r)$. This sign just reflects the fact that they use a different sign on f than Tits does, in the "standard" basis e, f, h for $\mathfrak{sl}_2$.

The Kac–Moody group $\mathfrak{G}_A(R)$ is defined as the quotient of $\mathfrak{St}_A(R)$ by the relations (1-1).

In Section 2 we defined the pre-Steinberg group $\mathfrak{PSt}_A(R)$ in terms of generators and relations. It also has an "intrinsic" definition: the same as $\mathfrak{St}_A(R)$, except that Tits' Chevalley relations are imposed only for *classically nilpotent* pairs α, β. This means $(\mathbb{Q}\alpha + \mathbb{Q}\beta) \cap \Phi$ is finite and $\alpha + \beta$ is nonzero, which is equivalent to α, β satisfying $\alpha + \beta \neq 0$ and lying in some A_1, A_1^2, A_2, B_2 or G_2 root system. As the name suggests, such a pair is prenilpotent. So $\mathfrak{PSt}_A(R)$ is defined the same way as $\mathfrak{St}_A(R)$, just omitting the Chevalley relations for prenilpotent pairs that are not classically prenilpotent. In particular, $\mathfrak{St}_A(R)$ is a quotient of $\mathfrak{PSt}_A(R)$, hence the prefix "pre-".

In [Allcock 2015] we defined $\mathfrak{PSt}_A(R)$ this way and then showed that it has the presentation in Section 2. In this paper, for ease of exposition we defined $\mathfrak{PSt}_A(R)$ by this presentation. But we will use the above "intrinsic" description in the proof of Theorem 1.1. So equality between the two versions of $\mathfrak{PSt}_A(R)$ is essential for our work. We proved it in [loc. cit., Theorem 1.2] and restate it now:

Theorem 3.1 (the two models of $\mathfrak{PSt}_A(R)$). *Let A be a generalized Cartan matrix and R a commutative ring. For each simple root α_i, choose one of the two distinguished parameterizations $X_{e_i} : R \to \mathfrak{U}_{\alpha_i}$. Then the pre-Steinberg group as defined in Section 2 is isomorphic to the pre-Steinberg group as defined above, by $S_i \mapsto \tilde{s}_{e_i}(1)$ and $X_i(t) \mapsto X_{e_i}(t)$.* □

4. Nomenclature for affine root systems

Our proof of Theorem 1.1, appearing in the next section, refers to the root system as a whole, with the simple roots playing no special role. It is natural in this setting to

	[Moody and Pianzola 1995]	[Kac 1990]	condition
$\widetilde{A}_n$	$A_n^{(1)}$	$A_n^{(1)}$	$n \geq 1$
$\widetilde{B}_n$	$B_n^{(1)}$	$B_n^{(1)}$	$n \geq 2$
$\widetilde{C}_n$	$C_n^{(1)}$	$C_n^{(1)}$	$n \geq 2$
$\widetilde{D}_n$	$D_n^{(1)}$	$D_n^{(1)}$	$n \geq 3$
$\widetilde{E}_n$	$E_n^{(1)}$	$E_n^{(1)}$	$n = 6, 7, 8$
$\widetilde{F}_4$	$F_4^{(1)}$	$F_4^{(1)}$	
$\widetilde{G}_2$	$G_2^{(1)}$	$G_2^{(1)}$	
$\widetilde{B}_n^{\text{even}}$	$B_n^{(2)}$	$D_{n+1}^{(2)}$	$n \geq 2$
$\widetilde{C}_n^{\text{even}}$	$C_n^{(2)}$	$A_{2n-1}^{(2)}$	$n \geq 2$
$\widetilde{BC}_n^{\text{odd}}$	$BC_n^{(2)}$	$A_{2n}^{(2)}$	$n \geq 1$
$\widetilde{F}_4^{\text{even}}$	$F_4^{(2)}$	$E_6^{(2)}$	
$\widetilde{G}_2^{0 \bmod 3}$	$G_2^{(3)}$	$D_4^{(3)}$	

Table 2. Our and others' names for affine root systems; see Section 4.

use nomenclature for the affine root systems that emphasizes this global perspective. Our notation $\widetilde{X}_n^{\cdots}$ in Table 2 is close to that in [Moody and Pianzola 1995, §3.5]. The differences are that our superscripts describe the construction of the root systems, and that we use a tilde to indicate affineness. For the affine root systems obtained by "folding", Kac's nomenclature [1990, pp. 54–55] emphasizes not the affine root system itself but rather the one being folded.

It is very easy to describe the set Φ of real roots in the root system $\widetilde{X}_n^{\cdots}$. Let $\bar{\Phi}$ be a root system of type X_n, let $\bar{\Lambda}$ be its root lattice, and let Λ be $\bar{\Lambda} \oplus \mathbb{Z}$. Then $\Phi \subseteq \Lambda$ is the set of pairs (root of X_n, $m \in \mathbb{Z}$) satisfying the condition that if the root is long then m has the property "$\cdots$" indicated in the superscript, if any.

A set of simple roots can be described as follows. We begin with a set of simple roots for the root system $\Phi_0 \subseteq \Phi$ consisting of roots of the form $(\bar{\alpha}, 0)$. This is an X_n root system except for $\widetilde{BC}_n^{\text{odd}}$, when it has type B_n. The affinizing simple root is $(\bar{\alpha}, 1)$, where $\bar{\alpha}$ is the lowest root of Φ_0 in the absence of a superscript, or twice the lowest short root for $\widetilde{BC}_n^{\text{odd}}$, or the lowest short root in all other cases. This can be used to verify the correspondences between our nomenclature and those of [Kac 1990] and [Moody and Pianzola 1995].

The condition on n in Table 2 is the weakest condition for which the definition of $\widetilde{X}_n^{\cdots}$ makes sense. If one wishes to avoid duplication, so that each isomorphism class of affine root system appears exactly once, then one should omit one of $\widetilde{A}_3 \cong \widetilde{D}_3$, one of $\widetilde{B}_2 \cong \widetilde{C}_2$ and one of $\widetilde{B}_2^{\text{even}} \cong \widetilde{C}_2^{\text{even}}$. Both [Kac 1990] and [Moody and Pianzola 1995] omit $\widetilde{D}_3$, $\widetilde{B}_2$ and $\widetilde{C}_2^{\text{even}}$. Also, [Moody and Pianzola 1995] gives $A_1^{(2)}$ as an alternate name for $BC_1^{(2)}$.

5. The isomorphism $\mathfrak{PSt}_A(R) \to \mathfrak{St}_A(R)$

This section is devoted to proving Theorem 1.1, whose hypotheses we assume throughout. In light of Theorem 3.1, our goal is to show that the Chevalley relations for the classically prenilpotent pairs imply those of the remaining prenilpotent pairs. We will begin by saying which pairs of real roots are prenilpotent and which are classically prenilpotent. Then we will analyze the pairs that are prenilpotent but not classically prenilpotent.

We fix the affine Dynkin diagram A, write Φ, $\bar{\Phi}$, Λ, $\bar{\Lambda}$ as in Section 4, and use an overbar to indicate projections of roots from Φ to $\bar{\Phi}$. It is easy to see that $\alpha, \beta \in \Phi$ are classically prenilpotent if and only if $\alpha = \beta$ or $\bar{\alpha}, \bar{\beta} \in \bar{\Phi}$ are linearly independent. The following lemma describes which pairs of roots are prenilpotent but not classically prenilpotent, and what their Chevalley relations are (except for one special case discussed later).

Lemma 5.1. *The following are equivalent*:

(i) *α, β are prenilpotent but not classically prenilpotent.*

(ii) *$\alpha \neq \beta$ are not equal and $\bar{\alpha}, \bar{\beta}$ differ by a positive scalar factor.*

(iii) *$\alpha \neq \beta$, and either $\bar{\alpha} = \bar{\beta}$ are equal or else one is twice the other and $\Phi = \widetilde{BC}_n^{\text{odd}}$.*

When these equivalent conditions hold, the Chevalley relations between $\mathfrak{U}_\alpha, \mathfrak{U}_\beta$ are $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$, unless $\Phi = \widetilde{BC}_n^{\text{odd}}$, $\bar{\alpha}$ and $\bar{\beta}$ are the same short root of $\bar{\Phi} = BC_n$, and $\alpha + \beta \in \Phi$.

Proof. We think of the Weyl group W acting on affine space in the usual way, with each root corresponding to an open halfspace. A root is positive if its halfspace contains the fundamental chamber, or negative if not. Recall that two roots $\alpha, \beta \in \Phi$ form a prenilpotent pair if some element w_+ of W sends both to positive roots, and some $w_- \in W$ sends both to negative roots. The existence of both $w_\pm$ is equivalent to saying that some chamber lies in the halfspaces of both α and β, and some other chamber lies in neither of them. (Proof: apply $w_\pm$ to the fundamental chamber rather than to $\{\alpha, \beta\}$.) By Euclidean geometry, this happens only if either their bounding hyperplanes are nonparallel or their bounding hyperplanes are parallel and one halfspace contains the other. In the first case, $\bar{\alpha}$ and $\bar{\beta}$ are linearly independent, so α, β are classically prenilpotent. In the second case, $\bar{\alpha}$ and $\bar{\beta}$ differ by a positive scalar. If α and β are equal then they form a classically prenilpotent pair. Otherwise they do not, because $(\mathbb{Q}\alpha \oplus \mathbb{Q}\beta) \cap \Phi$ is infinite. This proves the equivalence of (i) and (ii).

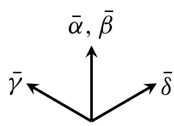
To see the equivalence of (ii) and (iii) we refer to the fact that $\bar{\Phi}$ is a reduced root system (i.e, the only positive multiple of a root that can be a root is that root itself) except in the case $\Phi = \widetilde{BC}_n^{\text{odd}}$. In this last case, the only way one root of $\bar{\Phi} = BC_n$ can be a positive multiple of a different root is that the long roots are got by doubling the short roots.

The proof of the final claim is similar. Except in the excluded case, we have $\bar{\Phi}\cap(\mathbb{N}\bar{\alpha}+\mathbb{N}\bar{\beta})=\{\bar{\alpha},\bar{\beta}\}$. The corresponding claim for Φ follows, so $\theta(\alpha,\beta)-\{\alpha,\beta\}$ is empty and the right-hand side of (3-2) is the identity. That is, the Chevalley relations for α, β read $[\mathfrak{U}_\alpha,\mathfrak{U}_\beta]=1$. (In the excluded case, $\Phi\cap(\mathbb{N}\alpha+\mathbb{N}\beta)=\{\alpha,\beta,\alpha+\beta\}$. So the Chevalley relations set the commutators of elements of $\mathfrak{U}_\alpha$ with elements of $\mathfrak{U}_\beta$ equal to certain elements of $\mathfrak{U}_{\alpha+\beta}$. See Case 6 below.) □

Recall from Theorem 3.1 that $\mathfrak{St}_A(R)$ may be got from $\mathfrak{PSt}_A(R)$ by adjoining the Chevalley relations for every prenilpotent pair α, β that is not classically prenilpotent. So to prove Theorem 1.1 it suffices to show that these relations already hold in $\mathfrak{PSt}:=\mathfrak{PSt}_A(R)$. In light of Lemma 5.1, the proof falls into seven cases, according to Φ and the relative position of $\bar{\alpha}$ and $\bar{\beta}$. Conceptually, they are organized as follows; see below for their exact hypotheses. Case 1 applies if $\bar{\alpha}=\bar{\beta}$ is a long root of some A_2 root system in $\bar{\Phi}$. Case 2 (resp. 3) applies if $\bar{\alpha}=\bar{\beta}$ is a long (resp. short) root of some B_2 root system in $\bar{\Phi}$. Case 4 applies if $\bar{\alpha}=\bar{\beta}$ is a short root of $\bar{\Phi}=G_2$. The rest of the cases are specific to $\Phi=\widetilde{BC}_n^{\mathrm{odd}}$. Case 5 applies if $\bar{\beta}=2\bar{\alpha}$. Case 6 or 7 applies if $\bar{\alpha}=\bar{\beta}$ is a short root of BC_n. There are two cases because $\alpha+\beta$ may or may not be a root.

In every case but one we must establish $[\mathfrak{U}_\alpha,\mathfrak{U}_\beta]=1$. Each case begins by choosing two roots in Φ, of which β is a specified linear combination, and whose projections to $\bar{\Phi}$ are specified. Given the global description of Φ from Section 4, this is always easy. Then we use the Chevalley relations for various classically prenilpotent pairs to deduce the Chevalley relations for α, β.

Case 1 of Theorem 1.1. Assume $\bar{\alpha}=\bar{\beta}$ is a root of $\bar{\Phi}=A_{n\geq 2}$, $\bar{\Phi}=D_n$ or $\bar{\Phi}=E_n$, or a long root of $\bar{\Phi}=G_2$. Choose $\bar{\gamma},\bar{\delta}\in\bar{\Phi}$ as shown, and choose lifts $\gamma,\delta\in\Phi$ summing to β. (Choose any $\gamma\in\Phi$ lying over $\bar{\gamma}$, define $\delta=\beta-\gamma$, and use the global description of Φ to check that $\delta\in\Phi$. This is trivial except in the case $\Phi=\widetilde{G}_2^{0\bmod 3}$, when it is easy.)



Because $\bar{\alpha}+\bar{\gamma},\bar{\alpha}+\bar{\delta}\notin\bar{\Phi}$, it follows that $\alpha+\gamma,\alpha+\delta\notin\Phi$. So the Chevalley relations $[\mathfrak{U}_\alpha,\mathfrak{U}_\gamma]=[\mathfrak{U}_\alpha,\mathfrak{U}_\delta]=1$ hold. The Chevalley relations for γ, δ imply $[\mathfrak{U}_\gamma,\mathfrak{U}_\delta]=\mathfrak{U}_{\gamma+\delta}=\mathfrak{U}_\beta$. (These relations are (2-23) in the G_2 case and (2-11) in the others. One can write them as $[X_\gamma(t),X_\delta(u)]=X_{\gamma+\delta}(tu)$ in the notation of the next paragraph.) Since $\mathfrak{U}_\alpha$ commutes with $\mathfrak{U}_\gamma$ and $\mathfrak{U}_\delta$, it commutes with the group they generate, hence $\mathfrak{U}_\beta$. □

The other cases use the same strategy: express an element of $\mathfrak{U}_\beta$ in terms of other root groups, and then evaluate its commutator with an element of $\mathfrak{U}_\alpha$. But the calculations are more delicate. We will work with explicit elements $X_\gamma(t) \in \mathfrak{U}_\gamma$ for various roots $\gamma \in \Phi$. Here t varies over R, and the definition of $X_\gamma(t)$ depends on choosing a basis vector e_γ for $\mathfrak{g}_{\gamma,\mathbb{Z}} \subseteq \mathfrak{g}$, as explained in Section 3. For each γ there are two possibilities for e_γ. The point of making these sign choices is to write down the relations explicitly.

For example, if $s, l \in I$ are the short and long roots of a B_2 subdiagram of A, then we copy their relations from (2-17): for all $t, u \in R$,

$$[X_s(t), X_l(u)] = S_l X_s(-tu) S_l^{-1} \cdot S_s X_l(t^2 u) S_s^{-1}. \tag{5-1}$$

The reason for writing the right side this way is to avoid making choices: to write down the relation, one only needs to specify generators e_s and e_l for $\mathfrak{g}_{s,\mathbb{Z}}$ and $\mathfrak{g}_{l,\mathbb{Z}}$, not the other root spaces involved. But for explicit computation one must choose generators for these other root spaces. Because S_s and S_l permute the root spaces in the same way the reflections in s and l do, the terms on the right in (5-1) lie in $\mathfrak{U}_{l+s}$ and $\mathfrak{U}_{l+2s}$. Therefore, after choosing suitable generators e_{l+s} and e_{l+2s} for $\mathfrak{g}_{l+s,\mathbb{Z}}$ and $\mathfrak{g}_{l+2s,\mathbb{Z}}$, we may rewrite (5-1) as

$$[X_s(t), X_l(u)] = X_{l+s}(-tu) \cdot X_{l+2s}(t^2 u). \tag{5-2}$$

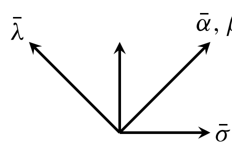
Now, if σ and λ are short and long simple roots for *any* copy of B_2 in Φ, then some element w of the Weyl group sends some pair of simple roots to them. Taking s and l to be this pair, and defining X_σ, X_λ, $X_{\lambda+\sigma}$ and $X_{\lambda+2\sigma}$ as the w-conjugates of X_s, X_l, X_{l+s} and X_{l+2s}, we can write the Chevalley relation for σ and λ by applying the substitution $s \mapsto \sigma$ and $l \mapsto \lambda$ to (5-2):

$$[X_\sigma(t), X_\lambda(u)] = X_{\lambda+\sigma}(-tu) \cdot X_{\lambda+2\sigma}(t^2 u). \tag{5-3}$$

In this way we can obtain the Chevalley relations we will need, for any classically prenilpotent pair, from the ones listed explicitly in Section 2. One could also refer to other standard references, for example, [Carter 1972, §5.2].

The root system $\widetilde{BC}_{n\geq 2}^{\text{odd}}$ appears as a possibility in several cases, including the next one. We will use "short", "middling" and "long" to refer to its three root lengths.

Case 2 of Theorem 1.1. Assume $\bar\alpha = \bar\beta$ is a long root of $\bar\Phi = B_{n\geq 2}$, $\bar\Phi = C_{n\geq 2}$, $\bar\Phi = BC_{n\geq 2}$ or $\bar\Phi = F_4$. Our first step is to choose roots $\bar\lambda, \bar\sigma \in \bar\Phi$ as pictured:



This is easily done using any standard description of $\bar\Phi$. (Note: although $\bar\lambda$ stands for "long" and $\bar\sigma$ for "short", $\bar\sigma$ is actually a middling root in the case $\bar\Phi = BC_n$.)

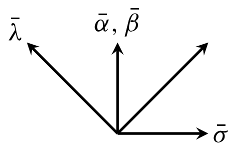
Our second step is to choose lifts $\lambda, \sigma \in \Phi$ with $\beta = \lambda + 2\sigma$. If Φ equals $\widetilde{B}_n$, $\widetilde{C}_n$ or $\widetilde{F}_4$ then one chooses any lift σ of $\bar{\sigma}$ and defines λ as $\beta - 2\sigma$. This works since every element of Λ lying over a root of $\bar{\Phi}$ is a root of Φ. If Φ equals $\widetilde{B}_n^{\text{even}}$, $\widetilde{C}_n^{\text{even}}$, $\widetilde{F}_4^{\text{even}}$ or $\widetilde{BC}_n^{\text{odd}}$ then this argument might fail since Φ is "missing" some long roots. Instead, one chooses any $\lambda \in \Phi$ lying over $\bar{\lambda}$ and defines σ as $(\beta - \lambda)/2$. Now, $\beta - \lambda = (\bar{\beta} - \bar{\lambda}, m)$ with m being even by the meaning of the superscript "even" or "odd". Also, $\bar{\beta} - \bar{\lambda}$ is divisible by 2 in $\bar{\Lambda}$ by the figure above. It follows that $\sigma \in \Lambda$. Then, as an element of Λ lying over a short (or middling) root of $\bar{\Phi}$, σ lies in Φ.

Because σ, λ are simple roots for a B_2 root system inside Φ, their Chevalley relation (5-3) holds in $\mathfrak{PSt}$. This shows that any element of $\mathfrak{U}_\beta = \mathfrak{U}_{\lambda+2\sigma}$ can be written in the form

$$(\text{some } x_{\lambda+\sigma} \in \mathfrak{U}_{\lambda+\sigma}) \cdot \big[(\text{some } x_\sigma \in \mathfrak{U}_\sigma), (\text{some } x_\lambda \in \mathfrak{U}_\lambda)\big]. \tag{5-4}$$

Referring to the picture of $\bar{\Phi}$ shows that $\alpha + \lambda + \sigma \notin \Phi$. Therefore, the Chevalley relations in $\mathfrak{PSt}$ include $[\mathfrak{U}_\alpha, \mathfrak{U}_{\lambda+\sigma}] = 1$. In particular, $\mathfrak{U}_\alpha$ commutes with the first term of (5-4). The same argument shows that $\mathfrak{U}_\alpha$ also commutes with the other terms, hence with any element of $\mathfrak{U}_\beta$. This shows that the Chevalley relations present in $\mathfrak{PSt}$ imply $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$, as desired. □

Case 3 of Theorem 1.1. Assume $\bar{\alpha} = \bar{\beta}$ is a short root of $\bar{\Phi} = B_{n\geq 2}$, $\bar{\Phi} = C_{n\geq 2}$ or $\bar{\Phi} = F_4$, or a middling root of $\bar{\Phi} = BC_{n\geq 2}$. We may choose $\lambda, \sigma \in \Phi$ with sum β and the following projections to $\bar{\Phi}$ (by a simpler argument than in the previous case):



The Chevalley relations for σ, λ are (5-3), showing that any element of $\mathfrak{U}_\beta = \mathfrak{U}_{\sigma+\lambda}$ can be written in the form

$$\big[(\text{some } x_\sigma \in \mathfrak{U}_\sigma), (\text{some } x_\lambda \in \mathfrak{U}_\lambda)\big] \cdot (\text{some } x_{\lambda+2\sigma} \in \mathfrak{U}_{\lambda+2\sigma}). \tag{5-5}$$

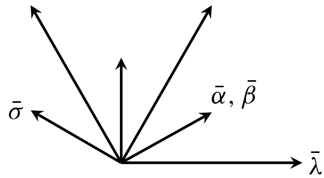
As in the previous case, we will conjugate this by an arbitrary element of $\mathfrak{U}_\alpha$. This requires the following Chevalley relations. We have $[\mathfrak{U}_\alpha, \mathfrak{U}_\lambda] = 1$ and $[\mathfrak{U}_\alpha, \mathfrak{U}_{\lambda+2\sigma}] = 1$ by the same argument as before. What is new is that the Chevalley relations for α, σ depend on whether $\alpha + \sigma$ is a root. If it is not, then $\mathfrak{U}_\alpha$ commutes with $\mathfrak{U}_\sigma$ and therefore with (5-5). That is, $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ as desired. If $\alpha + \sigma$ is a root then $[\mathfrak{U}_\alpha, \mathfrak{U}_\sigma] \subseteq \mathfrak{U}_{\alpha+\sigma}$. Then conjugating (5-5) by an element of $\mathfrak{U}_\alpha$ yields

$$\big[x_\sigma \cdot (\text{some } x_{\alpha+\sigma} \in \mathfrak{U}_{\alpha+\sigma}), x_\lambda\big] \cdot x_{\lambda+2\sigma},$$

which we can simplify by further use of Chevalley relations. Namely, neither $\lambda + \alpha + \sigma$ nor $\alpha + 2\sigma$ is a root, so $\mathfrak{U}_{\alpha+\sigma}$ centralizes $\mathfrak{U}_\lambda$ and $\mathfrak{U}_\sigma$. Therefore, $x_{\alpha+\sigma}$

centralizes the other terms in the commutator, and hence drops out, leaving (5-5). This shows that conjugation by any element of $\mathfrak{U}_\alpha$ leaves invariant every element of $\mathfrak{U}_\beta$. That is, $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$. □

Case 4 of Theorem 1.1. Assume $\bar{\alpha} = \bar{\beta}$ is a short root of $\bar{\Phi} = G_2$. This is the hardest case by far. Begin by choosing roots $\bar{\sigma}, \bar{\lambda} \in \bar{\Phi}$ as shown, with lifts $\sigma, \lambda \in \Phi$ summing to β:



Many different root groups appear in the argument, so we choose a generator e_γ of γ's root space, for each $\gamma \in \Phi$ that is a nonnegative linear combination of α, σ, λ.

Next we write down the G_2 Chevalley relations in $\mathfrak{PSt}$ that we will need, derived from (2-20)–(2-26). We will record them in the $\Phi = \widetilde{G}_2$ case and then comment on the simplifications that occur if $\Phi = \widetilde{G}_2^{0 \bmod 3}$. After negating some of the e_γ, for γ involving σ and λ but not α, we may suppose that the Chevalley relations (2-26) for σ, λ read

$$[X_\sigma(t), X_\lambda(u)] = X_{2\sigma+\lambda}(t^2u) X_{\sigma+\lambda}(-tu) X_{3\sigma+\lambda}(t^3u) X_{3\sigma+2\lambda}(-t^3u^2). \tag{5-6}$$

Then we may negate $e_{\alpha+2\sigma+\lambda}$, if necessary, to suppose the Chevalley relations (2-24) for α, $2\sigma+\lambda$ read

$$[X_\alpha(t), X_{2\sigma+\lambda}(u)] = X_{\alpha+2\sigma+\lambda}(3tu). \tag{5-7}$$

After negating some of the e_γ, for γ involving α and σ but not λ, we may suppose that the Chevalley relations (2-25) for σ and α read

$$[X_\sigma(t), X_\alpha(u)] = X_{\alpha+\sigma}(-2tu) X_{\alpha+2\sigma}(-3t^2u) X_{2\alpha+\sigma}(-3tu^2). \tag{5-8}$$

We know the Chevalley relations (2-24) for σ and $\alpha+\sigma$ have the form

$$[X_\sigma(t), X_{\alpha+\sigma}(u)] = X_{\alpha+2\sigma}(3\varepsilon tu), \tag{5-9}$$

where $\varepsilon = \pm 1$. We cannot choose the sign because we've already used our freedom to negate $e_{\alpha+2\sigma}$ in order to get (5-8). Similarly, we know that the Chevalley relations (2-23) for λ and $\alpha+2\sigma$ are

$$[X_\lambda(t), X_{\alpha+2\sigma}(u)] = X_{\alpha+2\sigma+\lambda}(\varepsilon' tu) \tag{5-10}$$

for some $\varepsilon' = \pm 1$. (We will see at the very end that $\varepsilon = \varepsilon' = 1$.)

We were able to write down these relations because we could work out the roots in the positive span of any two given roots. This used the assumption $\Phi = \widetilde{G}_2$, but now suppose $\Phi = \widetilde{G}_2^{0 \bmod 3}$. It may happen that some of the vectors appearing in the previous paragraph, projecting to long roots of $\bar{\Phi} = G_2$, are not roots of Φ. One can check that if $\alpha - \beta$ is divisible by 3 in Λ then there is no change. On the other hand, if $\alpha - \beta \not\equiv 0 \pmod 3$ then $\alpha + 2\sigma + \lambda$, $\alpha + 2\sigma$ and $2\alpha + \sigma$ are not roots. Because $(\mathbb{Q}\alpha \oplus \mathbb{Q}(2\sigma + \lambda)) \cap \Phi$ now has type A_2 rather than G_2, (5-7) is replaced by $[\mathfrak{U}_\alpha, \mathfrak{U}_{2\sigma+\lambda}] = 1$, from (2-10). And $(\mathbb{Q}\alpha \oplus \mathbb{Q}\sigma) \cap \Phi$ also has type A_2 now, so (5-8) is replaced by $[X_\sigma(t), X_\alpha(t)] = X_{\alpha+\sigma}(tu)$, obtained from (2-11), and (5-9) is replaced by $[\mathfrak{U}_\sigma, \mathfrak{U}_{\alpha+\sigma}] = 1$, from (2-10). Finally, there is no relation (5-10) because there is no longer a root group $\mathfrak{U}_{\alpha+2\sigma}$. The calculations below use the relations (5-6)–(5-10). To complete the proof, one must also carry out a similar calculation using (5-6) and the altered versions of (5-7)–(5-9). This calculation is so much easier that we omit it.

The long roots $3\sigma + 2\lambda$, $\alpha + 2\sigma + \lambda$ and $2\alpha + \sigma$ all lie over $3\bar{\sigma} + 2\bar{\lambda}$. These root groups commute with all others that will appear, by the Chevalley relations in $\mathfrak{PSt}$, and they commute with each other by Case 1 above. We will use this without specific mention.

Since $\beta = \sigma + \lambda$, we may take (5-6) with $t = 1$ and rearrange, to express any element of $\mathfrak{U}_\beta$ as

$$X_\beta(u) = X_{3\sigma+\lambda}(u) X_{3\sigma+2\lambda}(-u^2) \big[X_\lambda(u),\, X_\sigma(1)\big] X_{2\sigma+\lambda}(u). \tag{5-11}$$

We use this to express the commutators generating $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta]$:

$$\begin{aligned} [X_\alpha(t),\, X_\beta(u)] = {} & X_\alpha(t) X_{3\sigma+\lambda}(u) X_\alpha(t)^{-1} \cdot X_\alpha(t) X_{3\sigma+2\lambda}(-u^2) X_\alpha(t)^{-1} \\ & \cdot \big[X_\alpha(t) X_\lambda(u) X_\alpha(t)^{-1},\, X_\alpha(t) X_\sigma(1) X_\alpha(t)^{-1}\big] \\ & \cdot X_\alpha(t) X_{2\sigma+\lambda}(u) X_\alpha(t)^{-1} \\ & \cdot X_{2\sigma+\lambda}(-u) \big[X_\sigma(1),\, X_\lambda(u)\big] X_{3\sigma+2\lambda}(u^2) X_{3\sigma+\lambda}(-u). \end{aligned} \tag{5-12}$$

Because $\mathfrak{U}_\alpha$ centralizes $\mathfrak{U}_{3\sigma+\lambda}$, $\mathfrak{U}_{3\sigma+2\lambda}$ and $\mathfrak{U}_\lambda$, we may cancel all the $X_\alpha(t)$ in the first two terms, and in the first term of the first commutator. Because $\mathfrak{U}_{3\sigma+2\lambda}$ centralizes all terms present, we may cancel the terms $X_{3\sigma+2\lambda}(\pm u^2)$. The terms between the commutators assemble themselves into $[X_\alpha(t), X_{2\sigma+\lambda}(u)]$, which equals $X_{\alpha+2\sigma+\lambda}(3tu)$ by (5-7). Because $\mathfrak{U}_{\alpha+2\sigma+\lambda}$ centralizes all terms present, we may move this term to the very beginning. Finally, from (5-8) one can rewrite the second term of the first commutator as

$$X_\alpha(t) X_\sigma(1) X_\alpha(t)^{-1} = X_{2\alpha+\sigma}(3t^2) X_{\alpha+2\sigma}(3t) X_{\alpha+\sigma}(2t) X_\sigma(1).$$

After all these simplifications, (5-12) reduces to

$$\begin{aligned} [X_\alpha(t),\, X_\beta(u)] = {} & X_{\alpha+2\sigma+\lambda}(3tu) X_{3\sigma+\lambda}(u) \\ & \cdot [X_\lambda(u),\, X_{2\alpha+\sigma}(3t^2) X_{\alpha+2\sigma}(3t) X_{\alpha+\sigma}(2t) X_\sigma(1)][X_\sigma(1),\, X_\lambda(u)] X_{3\sigma+\lambda}(-u). \end{aligned} \tag{5-13}$$

Now we focus on the first commutator $[\cdots,\cdots]$ on the right side. All its terms commute with $\mathfrak{U}_{2\alpha+\sigma}$, so we may drop the $X_{2\alpha+\sigma}(3t^2)$ term. Writing out what remains gives

$$\begin{aligned}[\cdots,\cdots] = X_\lambda(u)X_{\alpha+2\sigma}(3t)X_{\alpha+\sigma}(2t)X_\sigma(1)\\ \cdot X_\lambda(-u)X_\sigma(-1)X_{\alpha+\sigma}(-2t)X_{\alpha+2\sigma}(-3t).\end{aligned}$$

By repeatedly using (5-9)–(5-10) and the commutativity of various pairs of root groups, we move all the X_λ and X_σ terms to the far right. A page-long computation yields

$$[\cdots,\cdots] = X_{\alpha+2\sigma+\lambda}(3\varepsilon' tu - 6\varepsilon\varepsilon' tu)\big[X_\lambda(u), X_\sigma(1)\big].$$

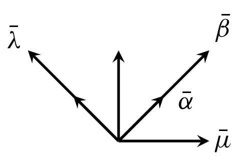
Plugging this into (5-13), and canceling the commutators and the $X_{3\sigma+\lambda}(\pm u)$ terms, yields

$$[X_\alpha(t), X_\beta(u)] = X_{\alpha+2\sigma+\lambda}(3tu + 3\varepsilon' tu - 6\varepsilon\varepsilon' tu) = X_{\alpha+2\sigma+\lambda}(Ctu),$$

where C equals 0, ± 6 or 12 depending on $\varepsilon, \varepsilon' \in \{\pm 1\}$.

If $C = 0$ (i.e., $\varepsilon = \varepsilon' = 1$) then we have established the desired Chevalley relation $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ and the proof is complete. Otherwise we pass to the quotient $\mathfrak{St}$ of $\mathfrak{PSt}$. Here $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$ commute, so we derive the relation $X_{\alpha+2\sigma+\lambda}(Ct) = 1$ in $\mathfrak{St}$. Since this identity holds universally, it holds for $R = \mathbb{C}$, so the image of $\mathfrak{U}_{\alpha+2\sigma+\lambda}(\mathbb{C})$ in $\mathfrak{St}(\mathbb{C})$ is the trivial group. This is a contradiction, since $\mathfrak{St}(\mathbb{C})$ acts on the Kac–Moody algebra $\mathfrak{g}$, with $X_{\alpha+2\sigma+\lambda}(t)$ acting (nontrivially for $t \neq 0$) by $\exp \operatorname{ad}(t e_{\alpha+2\sigma+\lambda})$. Since $C \neq 0$ leads to a contradiction, we must have $C = 0$ and so the Chevalley relation $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ holds in $\mathfrak{PSt}$. □

Case 5 of Theorem 1.1. Assume $\bar\beta = 2\bar\alpha$ in $\bar\Phi = BC_{n\geq 2}$. Choose $\bar\mu, \bar\lambda \in \bar\Phi$ as shown, and lift them to $\mu, \lambda \in \Phi$ with $2\mu + \lambda = \beta$. (Mnemonic: μ is middling and λ is long.)

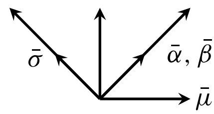


As in Case 2 (when $\bar\alpha$ and $\bar\beta$ were the same long root of $\bar\Phi = B_n$), we can express any element of $\mathfrak{U}_\beta$ in the form

$$(\text{some } x_{\mu+\lambda} \in \mathfrak{U}_{\mu+\lambda}) \cdot \big[(\text{some } x_\lambda \in \mathfrak{U}_\lambda), (\text{some } x_\mu \in \mathfrak{U}_\mu)\big].$$

The Chevalley relations in $\mathfrak{PSt}$ include the commutativity of $\mathfrak{U}_\alpha$ with $\mathfrak{U}_\lambda$, $\mathfrak{U}_\mu$ and $\mathfrak{U}_{\mu+\lambda}$. So $\mathfrak{U}_\alpha$ also centralizes $\mathfrak{U}_\beta$. □

Case 6 of Theorem 1.1. Assume $\bar\alpha = \bar\beta$ is a short root of $\bar\Phi = BC_{n\geq 2}$ and $\alpha + \beta$ is a root. This is the exceptional case of Lemma 5.1, and the Chevalley relation we must establish is not $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$. We will determine the correct relation during the proof. We begin by choosing $\bar{\mu}, \bar{\sigma} \in \bar{\Phi}$ as shown and lifting them to $\mu, \sigma \in \Phi$ with $\mu + \sigma = \beta$, so that σ, μ generate a B_2 root system:



We choose a generator e_γ for the root space of each nonnegative linear combination $\gamma \in \Phi$ of α, σ, μ. By changing the signs of $e_{\sigma+\mu}$ and $e_{2\sigma+\mu}$ if necessary, we may suppose that the Chevalley relations (2-17) for σ, μ are

$$[X_\sigma(t), X_\mu(u)] = X_{\sigma+\mu}(-tu) X_{2\sigma+\mu}(t^2 u). \tag{5-14}$$

Since $\sigma + \mu = \beta$ we may take $t = 1$ in (5-14) to express any element of $\mathfrak{U}_\beta$:

$$X_\beta(u) = X_{2\sigma+\mu}(u)[X_\mu(u), X_\sigma(1)]. \tag{5-15}$$

Using this one can express any generator for $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta]$:

$$\begin{aligned}[X_\alpha(t), X_\beta(u)] = X_\alpha(t) X_{2\sigma+\mu}(u) X_\alpha(t)^{-1} \\ \cdot [X_\alpha(t) X_\mu(u) X_\alpha(t)^{-1}, X_\alpha(t) X_\sigma(1) X_\alpha(t)^{-1}] \\ \cdot [X_\sigma(1), X_\mu(u)] \cdot X_{2\sigma+\mu}(-u).\end{aligned} \tag{5-16}$$

By the Chevalley relations $[\mathfrak{U}_\alpha, \mathfrak{U}_{2\sigma+\mu}] = [\mathfrak{U}_\alpha, \mathfrak{U}_\mu] = 1$, the $X_\alpha(t)^{\pm 1}$ cancel in the first term on the right side and in the first term of the first commutator.

Now we consider the Chevalley relations of α and σ. Since $\bar{\alpha} + \bar{\sigma}$ is a middling root of $\bar{\Phi}$, and Φ contains every element of Λ lying over every such root, we see that $\alpha + \sigma$ is a root of Φ. In particular, $(\mathbb{Q}\alpha \oplus \mathbb{Q}\sigma) \cap \Phi$ is a B_2 root system, in which α and σ are orthogonal short roots. The Chevalley relations (2-16) for α, σ are therefore

$$[X_\alpha(t), X_\sigma(u)] = X_{\alpha+\sigma}(-2tu), \tag{5-17}$$

after changing the sign of $e_{\alpha+\sigma}$ if necessary.

Next, $\mu + \sigma + \alpha = \alpha + \beta$ is a root by hypothesis. We choose $e_{\mu+\sigma+\alpha}$ so that the Chevalley relations (2-16) for μ, $\alpha + \sigma$ are

$$[X_\mu(t), X_{\alpha+\sigma}(u)] = X_{\mu+\alpha+\sigma}(-2tu). \tag{5-18}$$

Now we rewrite (5-16), applying the cancellations mentioned above and rewriting the second term in the first commutator using (5-17):

$$\begin{aligned}&[X_\alpha(t), X_\beta(u)] \\ &= X_{2\sigma+\mu}(u) \cdot \bigl[X_\mu(u), X_{\alpha+\sigma}(-2t) X_\sigma(1)\bigr] \cdot \bigl[X_\sigma(1), X_\mu(u)\bigr] \cdot X_{2\sigma+\mu}(-u).\end{aligned} \tag{5-19}$$

Now we restrict attention to the first commutator on the right side and use the Chevalley relations $[\mathfrak{U}_{\alpha+\sigma}, \mathfrak{U}_\sigma] = 1$ and (5-18) to obtain

$$\begin{aligned}[X_\mu(u), X_{\alpha+\sigma}(-2t)X_\sigma(1)] &= X_\mu(u)X_{\alpha+\sigma}(-2t)\cdot X_\sigma(1)\cdot X_\mu(-u)X_\sigma(-1)X_{\alpha+\sigma}(2t)\\ &= X_{\mu+\alpha+\sigma}(4tu)X_{\alpha+\sigma}(-2t)X_\mu(u)\cdot X_\sigma(1)\\ &\qquad\cdot X_{\mu+\alpha+\sigma}(4tu)X_{\alpha+\sigma}(2t)X_\mu(-u)X_\sigma(-1).\end{aligned}$$

The projections to $\bar\Phi$ of any two roots occurring as subscripts are linearly independent. Therefore, any two of them are classically prenilpotent, so their Chevalley relations are present in $\mathfrak{PSt}$. In particular, $\mathfrak{U}_{\mu+\alpha+\sigma}$ centralizes all the other terms; we gather the $X_{\mu+\alpha+\sigma}(4tu)$ terms at the beginning. Next, $[\mathfrak{U}_\sigma, \mathfrak{U}_{\alpha+\sigma}] = 1$, so we may move $X_\sigma(1)$ to the right across $X_{\alpha+\sigma}(2t)$. Then we can use (5-18) again to move $X_\mu(u)$ rightward across $X_{\alpha+\sigma}(2t)$. The result is

$$[X_\mu(u),\, X_{\alpha+\sigma}(-2t)X_\sigma(1)] = X_{\mu+\alpha+\sigma}(4tu)[X_\mu(u),\, X_\sigma(1)].$$

Plugging this into (5-19) and canceling the commutators gives

$$[X_\alpha(t),\, X_\beta(u)] = X_{2\sigma+\mu}(u)X_{\mu+\alpha+\sigma}(4tu)X_{2\sigma+\mu}(-u) = X_{\alpha+\beta}(4tu).$$

Tits' Chevalley relation in his definition of $\mathfrak{St}$ has the same form, with the factor 4 replaced by some integer C. (Although we don't need it, we remark that $C = \pm 4$ by the second displayed equation in [Tits 1987, §3.5], or from [Morita 1988, Theorem 2(2)]. This is related to the fact that $(\mathbb{Q}\alpha \oplus \mathbb{Q}\beta)\cap\Phi$ is a rank 1 affine root system, of type $\widetilde{BC}_1^{\,\mathrm{odd}}$.) If $C \neq 4$ then in $\mathfrak{St}$ we deduce $X_{\alpha+\beta}((C-4)tu) = 1$ for all $t, u \in R$ and all rings R, leading to the same contradiction we found in Case 4. Therefore, $C = 4$ and we have established that Tits' relation already holds in $\mathfrak{PSt}$. □

Case 7 of Theorem 1.1. Assume $\bar\alpha = \bar\beta$ is a short root of $\bar\Phi = BC_{n\geq 2}$ and $\alpha+\beta$ is not a root. This is similar to the previous case but much easier. We choose μ, σ and the e_γ in the same way, except that $\mu+\sigma+\alpha$ is no longer a root, so the Chevalley relation (5-18) is replaced by $[\mathfrak{U}_\mu, \mathfrak{U}_{\alpha+\sigma}] = 1$. We expand $X_\beta(u)$ as in (5-15) and obtain (5-19) as before. But this time the $X_{\alpha+\sigma}(-2t)$ term centralizes both $\mathfrak{U}_\mu$ and $\mathfrak{U}_\sigma$, so it vanishes from the commutator. The right side of (5-19) then collapses to 1 and we have proven $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ in $\mathfrak{PSt}$. □

6. Finite presentations

In this section we prove Theorem 1.3, that various Steinberg and Kac–Moody groups are finitely presented. At the end we make several remarks about possible variations on the definition of Kac–Moody groups.

Proof of Theorem 1.3. We must show that $\mathfrak{St}_A(R)$ is finitely presented under either of the two stated hypotheses. By Theorem 1.1 it suffices to prove this with $\mathfrak{PSt}$ in place of $\mathfrak{St}$.

(ii) We are assuming that $\operatorname{rk} A = 3$ and that R is finitely generated as a module over a subring generated by finitely many units. Theorem 1.4(ii) of [Allcock 2015] shows that if R satisfies this hypothesis and A is 2-spherical, then $\mathfrak{PSt}_A(R)$ is finitely presented. This proves (ii).

(i) Now we are assuming that $\operatorname{rk} A > 3$ and that R is finitely generated as a ring. Theorem 1.4(iii) of [Allcock 2015] gives the finite presentability of $\mathfrak{PSt}_A(R)$ if every pair of nodes of the Dynkin diagram lies in some irreducible spherical diagram of rank ≥ 3. (This use of a covering of A by spherical diagrams was also used by Capdeboscq [2013].) By inspecting the list of affine Dynkin diagrams of rank > 3, one checks that this treats all cases of (i) except

$A =$ 

(with some orientations of the double edges). In this case, no irreducible spherical diagram contains α and δ. Note that $\beta \neq \gamma$ since $\operatorname{rk} A > 3$.

For this case we use a variation on the proof of Theorem 1.4(iii) of [Allcock 2015]. Consider the direct limit G of the groups $\mathfrak{St}_B(R)$ as B varies over all irreducible spherical diagrams of rank ≥ 2. If $\operatorname{rk} B \geq 3$ then $\mathfrak{St}_B(R)$ is finitely presented by Theorem I of [Splitthoff 1986]. If $\operatorname{rk} B = 2$ then $\mathfrak{St}_B(R)$ is finitely generated by [Allcock 2015, Lemma 12.2]. Since every irreducible rank 2 diagram lies in one of rank > 2, it follows that G is finitely presented. Now, G satisfies all the relations of $\mathfrak{St}_A(R)$ except for the commutativity of $\mathfrak{St}_{\{\alpha\}}$ with $\mathfrak{St}_{\{\delta\}}$. Because these groups may not be finitely generated, we might need infinitely many additional relations to impose commutativity in the obvious way.

So we proceed indirectly. Let Y_α be a finite subset of $\mathfrak{St}_{\{\alpha\}}$ which together with $\mathfrak{St}_{\{\beta\}}$ generates $\mathfrak{St}_{\{\alpha,\beta\}}$. This is possible since $\mathfrak{St}_{\{\alpha,\beta\}}$ is finitely generated. We define Y_δ similarly, with γ in place of β. We define H as the quotient of G by the finitely many relations $[Y_\alpha, Y_\delta] = 1$, and we claim that the images in H of $\mathfrak{St}_{\{\alpha\}}$ and $\mathfrak{St}_{\{\delta\}}$ commute.

A computation in H establishes this: First, every element of Y_δ centralizes $\mathfrak{St}_{\{\beta\}}$ by the definition of G, and every element of Y_α by that of H. Therefore, it centralizes $\mathfrak{St}_{\{\alpha,\beta\}}$, hence $\mathfrak{St}_{\{\alpha\}}$. We've shown that $\mathfrak{St}_{\{\alpha\}}$ centralizes Y_δ, and it centralizes $\mathfrak{St}_{\{\gamma\}}$ by the definition of G. Therefore, it centralizes $\mathfrak{St}_{\{\gamma,\delta\}}$, hence $\mathfrak{St}_{\{\delta\}}$.

H has the same generators as $\mathfrak{PSt}_A(R)$, and its defining relations are among those defining $\mathfrak{PSt}_A(R)$. On the other hand, we have shown that the generators of H satisfy all the relations in $\mathfrak{PSt}_A(R)$. So $H \cong \mathfrak{PSt}_A(R)$. In particular, $\mathfrak{PSt}_A(R)$ is finitely presented.

It remains to prove the finite presentability of $\mathfrak{G}_A(R)$ under the extra hypothesis that the unit group of R is finitely generated as an abelian group. This follows from [Allcock 2015, Lemma 12.4], which says that the quotient of $\mathfrak{PSt}_A(R)$ by all the relations (1-1) is equally well-defined by finitely many of them. Choosing finitely many such relations, and imposing them on the quotient $\mathfrak{St}_A(R)$ of $\mathfrak{PSt}_A(R)$, gives all the relations (1-1). The quotient of $\mathfrak{St}_A(R)$ by these is the definition of $\mathfrak{G}_A(R)$, proving its finite presentation. □

Remark 6.1 (completions). We have worked with the "minimal" or "algebraic" forms of Kac–Moody groups. One can consider various completions, such as those surveyed in [Tits 1985]. None of these completions can possibly be finitely presented, so no analogue of Theorem 1.3 exists. But it is reasonable to hope for an analogue of Corollary 1.2.

Remark 6.2 (Chevalley–Demazure group schemes). If A is spherical then we write $\mathfrak{CD}_A$ for the simply connected version of the associated Chevalley–Demazure group scheme. This is the unique most natural (in a certain technical sense) algebraic group over $\mathbb{Z}$ of type A. If R is a Dedekind domain of arithmetic type, then the question of whether $\mathfrak{CD}_A(R)$ is finitely presented was settled by Behr [1967; 1998]. We emphasize that our Theorem 1.3 does not give a new proof of his results, because $\mathfrak{CD}_A(R)$ may be a proper quotient of $\mathfrak{G}_A(R)$. The kernel of $\mathfrak{St}_A(R) \to \mathfrak{CD}_A(R)$ is called $K_2(A; R)$ and contains the relators (1-1). It can be extremely complicated.

For a nonspherical Dynkin diagram A, the functor $\mathfrak{CD}_A$ is not defined. The question of whether there is a good definition, and what it would be, seems to be completely open. Only when R is a field is there known to be a unique "best" definition of a Kac–Moody group [Tits 1987, Theorem 1′, p. 553]. The main problem is what extra relations to impose on $\mathfrak{G}_A(R)$. The remarks below discuss the possible forms of some additional relations.

Remark 6.3 (Kac–Moody groups over integral domains). If R is an integral domain with fraction field k, then it is open whether $\mathfrak{G}_A(R) \to \mathfrak{G}_A(k)$ is injective. If $\mathfrak{G}_A$ satisfies Tits' axioms then this would follow from (KMG4), but Tits does not assert that $\mathfrak{G}_A$ satisfies his axioms. If $\mathfrak{G}_A(R) \to \mathfrak{G}_A(k)$ is not injective, then the image seems a better candidate than $\mathfrak{G}_A(R)$ itself for the role of "the" Kac–Moody group.

Remark 6.4 (Kac–Moody groups via representations). Fix a root datum D and a commutative ring R. By using Kostant's $\mathbb{Z}$-form of the universal enveloping algebra of $\mathfrak{g}$, one can construct a $\mathbb{Z}$-form $V^\lambda_{\mathbb{Z}}$ of any integrable highest-weight module V^λ of $\mathfrak{g}$. Then one defines V^λ_R as $V^\lambda_{\mathbb{Z}} \otimes R$. For each real root α, one can exponentiate $\mathfrak{g}_{\alpha,\mathbb{Z}} \otimes R \cong R$ to get an action of $\mathfrak{U}_\alpha \cong R$ on V^λ_R. One can define the action of the torus $(R^*)^n$ directly. Then one can take the group $\mathfrak{G}^\lambda_D(R)$ generated by these transformations and call it a Kac–Moody group. This approach is extremely natural and not yet fully worked out. The first such work for Kac–Moody groups over rings is Garland's landmark paper [1980] treating affine groups; see also Tits' survey [1985, §5], its references, and the recent articles [Bao and Carbone 2015] and [Carbone and Garland 2012].

Tits [1987, p. 554] asserts that this construction allows one to build a Kac–Moody functor satisfying all his axioms (KMG1)–(KMG9). We imagine that he reasoned as follows. First, show that each $\mathfrak{G}_D^\lambda$ is a Kac–Moody functor and therefore by Tits' theorem admits a canonical functorial homomorphism from $\mathfrak{G}_A$, where A is the generalized Cartan matrix of D. One cannot directly apply Tits' theorem, because $\mathfrak{G}_D^\lambda(R)$ only comes equipped with the homomorphisms $\mathrm{SL}_2(R) \to \mathfrak{G}_D^\lambda(R)$ required by Tits when $\mathrm{SL}_2(R)$ is generated by its subgroups $\left(\begin{smallmatrix}1 & *\\ 0 & 1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1 & 0\\ * & 1\end{smallmatrix}\right)$. Presumably this difficulty can be overcome. Second, define I as the intersection of the kernels of all the homomorphisms $\mathfrak{G}_A \to \mathfrak{G}_D^\lambda$, and then define the desired Kac–Moody functor as $\mathfrak{G}_A/I$.

Remark 6.5 (Kac–Moody groups as amalgams of Chevalley–Demazure groups). The difficulty in the previous remark, that $\mathrm{SL}_2(R)$ is not always generated by unipotent elements, might be resolved as follows. One can consider the spherical subdiagrams B of A, construct the corresponding Chevalley–Demazure groups $\mathfrak{CD}_B(R)$, and amalgamate these as in Corollary 1.2, rather than amalgamating Steinberg groups. Our results here and in [Allcock 2015] show that this amalgam satisfies the Chevalley relations of all of the prenilpotent pairs that are not classically prenilpotent. (For nonaffine diagrams this requires A to be 3-spherical; 2-sphericity will do if A is simply laced or R has no tiny quotients.) And it is the smallest extension of Tits' construction that recovers $\mathfrak{CD}_A(R)$ when A is spherical. We propose this amalgam, possibly with extra relations, as a reasonable candidate for the definition of Kac–Moody groups.

Remark 6.6 (loop groups). Suppose X is one of the *ABCDEFG* diagrams, $\widetilde{X}$ is its affine extension as in Section 4, and R is a commutative ring. The well-known description of affine Kac–Moody algebras and loop groups makes it natural to expect that $\mathfrak{G}_{\widetilde{X}}(R)$ is a central extension of $\mathfrak{G}_X(R[t^{\pm 1}])$ by R^*. The most general results along these lines that I know of are Theorems 10.1 and B.1 in [Garland 1980], although they concern slightly different groups. Instead, one might simply *define* the loop group $G_{\widetilde{X}}(R)$ as a central extension of $\mathfrak{CD}_X(R[t^{\pm 1}])$ by R^*, where the 2-cocycle defining the extension would have to be made explicit. Then one could try to show that $G_{\widetilde{X}}$ satisfies Tits' axioms.

It is natural to ask whether such a group $G_{\widetilde{X}}(R)$ would be finitely presented if R is finitely generated. If R^* is finitely generated then this is equivalent to the finite presentation of the quotient $\mathfrak{CD}_X\big(R[t^{\pm 1}]\big)$. If $\operatorname{rk} X \geq 3$ then $\mathfrak{St}_X\big(R[t^{\pm 1}]\big)$ is finitely presented by Theorem I of [Splitthoff 1986]. Then, as explained in Section 7 of [loc. cit.], the finite presentability of $\mathfrak{CD}_X(R[t^{\pm 1}])$ boils down to properties of $K_1(X, R[t^{\pm 1}])$ and $K_2(X, R[t^{\pm 1}])$.

Acknowledgements

The author is very grateful to the Japan Society for the Promotion of Science and to Kyoto University for their support and hospitality. He would also like to thank Lisa Carbone and the referees for very helpful comments on earlier versions of the paper.

References

[Abramenko and Mühlherr 1997] P. Abramenko and B. Mühlherr, "Présentations de certaines *BN*-paires jumelées comme sommes amalgamées", *C. R. Acad. Sci. Paris Sér. I Math.* **325**:7 (1997), 701–706. MR 1483702 Zbl 0934.20024

[Allcock 2015] D. Allcock, "Steinberg groups as amalgams", preprint, 2015. arXiv 1307.2689

[Allcock and Carbone 2016] D. Allcock and L. Carbone, "Presentation of hyperbolic Kac–Moody groups over rings", *J. Algebra* **445** (2016), 232–243. MR 3418056 Zbl 06509595

[Bao and Carbone 2015] L. Bao and L. Carbone, "Integral forms of Kac–Moody groups and Eisenstein series in low dimensional supergravity theories", preprint, 2015. arXiv 1308.6194

[Behr 1967] H. Behr, "Über die endliche Definierbarkeit verallgemeinerter Einheitengruppen, II", *Invent. Math.* **4** (1967), 265–274. MR 0249437 Zbl 0157.36702

[Behr 1998] H. Behr, "Arithmetic groups over function fields, I: A complete characterization of finitely generated and finitely presented arithmetic subgroups of reductive algebraic groups", *J. Reine Angew. Math.* **495** (1998), 79–118. MR 1603845 Zbl 0923.20038

[Capdeboscq 2013] I. Capdeboscq, "Bounded presentations of Kac–Moody groups", *J. Group Theory* **16**:6 (2013), 899–905. MR 3198724 Zbl 1286.20064

[Carbone and Garland 2012] L. Carbone and H. Garland, "Infinite dimensional Chevalley groups and Kac–Moody groups over $\mathbb{Z}$", preprint, 2012, available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.297.5555&rep=rep1&type=pdf.

[Carter 1972] R. W. Carter, *Simple groups of Lie type*, Pure and Applied Mathematics **28**, Wiley, London, 1972. MR 0407163 Zbl 0248.20015

[Devillers and Mühlherr 2007] A. Devillers and B. Mühlherr, "On the simple connectedness of certain subsets of buildings", *Forum Math.* **19**:6 (2007), 955–970. MR 2367950 Zbl 1188.51003

[Eskin and Farb 1997] A. Eskin and B. Farb, "Quasi-flats and rigidity in higher rank symmetric spaces", *J. Amer. Math. Soc.* **10**:3 (1997), 653–692. MR 1434399 Zbl 0893.22004

[Farb and Schwartz 1996] B. Farb and R. Schwartz, "The large-scale geometry of Hilbert modular groups", *J. Differential Geom.* **44**:3 (1996), 435–478. MR 1431001 Zbl 0871.11035

[Garland 1980] H. Garland, "The arithmetic theory of loop groups", *Inst. Hautes Études Sci. Publ. Math.* **52** (1980), 5–136. MR 601519 Zbl 0475.17004

[Hartnick and Köhl 2015] T. Hartnick and R. Köhl, "Two-spherical topological Kac–Moody groups are Kazhdan", *J. Group Theory* **18**:4 (2015), 649–654. MR 3365821 Zbl 06455987

[Kac 1990] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, 1990. MR 1104219 Zbl 0716.17022

[Kleiner and Leeb 1997] B. Kleiner and B. Leeb, "Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings", *Inst. Hautes Études Sci. Publ. Math.* **86** (1997), 115–197. MR 1608566 Zbl 0910.53035

[Moody and Pianzola 1995] R. V. Moody and A. Pianzola, *Lie algebras with triangular decompositions*, John Wiley & Sons, New York, 1995. MR 1323858 Zbl 0874.17026

[Morita 1987] J. Morita, "Commutator relations in Kac–Moody groups", *Proc. Japan Acad. Ser. A Math. Sci.* **63**:1 (1987), 21–22. MR 892949 Zbl 0626.22014

[Morita 1988] J. Morita, "Root strings with three or four real roots in Kac–Moody root systems", *Tohoku Math. J.* (2) **40**:4 (1988), 645–650. MR 972252 Zbl 0656.17012

[Morita and Rehmann 1990] J. Morita and U. Rehmann, "A Matsumoto-type theorem for Kac–Moody groups", *Tohoku Math. J.* (2) **42**:4 (1990), 537–560. MR 1076175 Zbl 0701.19001

[Neuhauser 2003] M. Neuhauser, "Kazhdan's property T for the symplectic group over a ring", *Bull. Belg. Math. Soc. Simon Stevin* **10**:4 (2003), 537–550. MR 2040529 Zbl 1063.22021

[Schwartz 1995] R. E. Schwartz, "The quasi-isometry classification of rank one lattices", *Inst. Hautes Études Sci. Publ. Math.* **82** (1995), 133–168. MR 1383215 Zbl 0852.22010

[Shalom 1999] Y. Shalom, "Bounded generation and Kazhdan's property (T)", *Inst. Hautes Études Sci. Publ. Math.* **90** (1999), 145–168. MR 1813225 Zbl 0980.22017

[Splitthoff 1986] S. Splitthoff, "Finite presentability of Steinberg groups and related Chevalley groups", pp. 635–687 in *Applications of algebraic K-theory to algebraic geometry and number theory* (Boulder, CO, 1983), vol. II, edited by S. J. Bloch et al., Contemp. Math. **55**, American Mathematical Society, Providence, RI, 1986. MR 862658 Zbl 0596.20034

[Steinberg 1968] R. Steinberg, *Lectures on Chevalley groups*, Yale University, New Haven, CT, 1968. MR 0466335 Zbl 1196.22001

[Tits 1985] J. Tits, "Groups and group functors attached to Kac–Moody data", pp. 193–223 in *Workshop Bonn 1984* (Bonn, 1984), edited by F. Hirzebruch et al., Lecture Notes in Math. **1111**, Springer, Berlin, 1985. MR 797422 Zbl 0572.17010

[Tits 1987] J. Tits, "Uniqueness and presentation of Kac–Moody groups over fields", *J. Algebra* **105**:2 (1987), 542–573. MR 873684 Zbl 0626.22013

Communicated by Edward Frenkel
Received 2014-09-23 Revised 2015-06-21 Accepted 2015-10-15

allcock@math.utexas.edu
Department of Mathematics, University of Texas at Austin, RLM 8.100, 2515 Speedway Stop C1200, Austin, TX 78712, United States

msp

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.557

Discriminant formulas and applications

Kenneth Chan, Alexander A. Young and James J. Zhang

The discriminant is a classical invariant associated to algebras which are finite over their centers. It was shown recently by several authors that if the discriminant of A is "sufficiently nontrivial" then it can be used to answer some difficult questions about A. Two such questions are: What is the automorphism group of A? Is A Zariski cancellative?

We use the discriminant to study these questions for a class of (generalized) quantum Weyl algebras. Along the way, we give criteria for when such an algebra is finite over its center and prove two conjectures of Ceken, Wang, Palmieri and Zhang.

Introduction

In algebraic number theory, the discriminant takes on a familiar form: let L be a Galois extension of the field $\mathbb{Q}$ and write $\mathcal{O}_L = \mathbb{Z}[\alpha] \cong \mathbb{Z}[x]/(f)$, where f is the minimal polynomial (or the characteristic polynomial) of α. Then we have

$$\Delta_{L/\mathbb{Q}} = \prod_{i \neq j} (r_i - r_j),$$

where $r_1, \ldots, r_n$ are the roots of f. In noncommutative algebra, the discriminant has long been used to study orders and lattices in a central simple algebra [Reiner 1975]. Recently, it has been shown that the discriminant plays a remarkable role in solving some classical and notoriously difficult questions:

(1) *Automorphism problem*: determining the full automorphism groups of noncommutative Artin–Schelter regular algebras [CPWZ 2015a; 2016].

(2) *Zariski cancellation problem*: concerning the cancellative property of noncommutative algebras such as skew polynomial rings [Bell and Zhang 2016].

(3) *Isomorphism problem*: finding a criterion for when two algebras are isomorphic, within certain classes of noncommutative algebras [CPWZ 2015b].

MSC2010: 16W20.

Keywords: discriminant, automorphism group, cancellation problem, quantum algebra, Clifford algebra, rings and algebras.

Despite the usefulness of the discriminant in algebraic number theory, algebraic geometry and noncommutative algebra, it is extremely hard to compute, especially in high dimensional and high rank cases. In [CPWZ 2015a; 2016], the authors made two conjectures on discriminant formulas for some classes of noncommutative algebras. Our main aim is to prove these two conjectures.

Let k be a base commutative domain and let $k^\times$ be the set of invertible elements in k. The discriminant of a noncommutative algebra A over a central subalgebra $Z \subseteq A$, denoted by $d(A/Z)$, will be reviewed in Section 1. Let $q \in k^\times$ be an invertible element in k and let A_q be the q-quantum Weyl algebra generated by x and y and subject to the relation $yx = qxy + 1$. Our first result is:

Theorem 0.1. *Let q be a primitive n-th root of unity for some $n \geq 2$. Then the discriminant of A_q over its center $Z(A_q)$ is*

$$d(A_q/Z(A_q)) = c(nm)^{n^2}((1-q)^n x^n y^n - 1)^{n(n-1)},$$

where c is some element in $k^\times$ and $m = \prod_{i=2}^{n-1}(1+q+\cdots+q^{i-1})$. By convention, $m = 1$ when $n = 2$.

Theorem 0.1 answers [CPWZ 2016, Conjecture 5.3] affirmatively.

For $n \geq 2$, let W_n be the k-algebra generated by $x_1, \ldots, x_n$ and subject to the relations $x_i x_j + x_j x_i = 1$ for all $i \neq j$ [CPWZ 2015a, Introduction]. This algebra is called a (-1)-*quantum Weyl algebra* [CPWZ 2015b, Introduction]. Let

$$M := \begin{pmatrix} 2x_1^2 & 1 & \cdots & 1 \\ 1 & 2x_2^2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2x_n^2 \end{pmatrix}.$$

Let Z denote the central subalgebra $k\left[x_1^2, \ldots, x_n^2\right] \subseteq W_n$. Our second result is:

Theorem 0.2. *Suppose 2 is invertible in k. Then the discriminant of W_n over the subalgebra Z is*

$$d(W_n/Z) = c(\det M)^{2^{n-1}},$$

where c is an element in $k^\times$.

Theorem 0.2 answers [CPWZ 2015a, Question 4.12(2)] affirmatively.

These results suggest that the discriminant has elegant expressions in some situations. Because of its usefulness, more discriminant formulas should be established; see Lemma 6.4.

This paper contains other related results which we now describe. Let T be a commutative algebra over k and let $\boldsymbol{q} := \{q_{ij} \in T^\times \mid 1 \leq i < j \leq n\}$ and

$\mathcal{A} := \{a_{ij} \in T \mid 1 \le i < j \le n\}$ be sets of elements in T. The skew polynomial ring $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is a T-algebra generated by $x_1, \dots, x_n$ and subject to the relations

$$x_j x_i = q_{ij} x_i x_j \quad \text{for all } 1 \le i < j \le n. \tag{E0.2.1}$$

A generalized quantum Weyl algebra associated to $(\boldsymbol{q}, \mathcal{A})$ is a T-central filtered algebra of the form

$$V_n(\boldsymbol{q}, \mathcal{A}) = \frac{T\langle x_1, \dots, x_n\rangle}{(x_j x_i - q_{ij} x_i x_j - a_{ij} \mid i < j)} \tag{E0.2.2}$$

such that the associated graded ring $\operatorname{gr} V_n(\boldsymbol{q}, \mathcal{A})$ is naturally isomorphic to the skew polynomial ring $T_{\boldsymbol{q}}[x_1, \dots, x_n]$. Another way of constructing $V_n(\boldsymbol{q}, \mathcal{A})$ is to use an iterated Ore extension starting with T. To calculate the discriminant of $V_n(\boldsymbol{q}, \mathcal{A})$ over its center, one needs to determine the center of $V_n(\boldsymbol{q}, \mathcal{A})$. The discriminant is defined whenever $V_n(\boldsymbol{q}, \mathcal{A})$ is a finite module over a central subring Z [CPWZ 2016], and it is most useful when $V_n(\boldsymbol{q}, \mathcal{A})$ is a free module over Z [CPWZ 2015a]. Since $\operatorname{gr} V_n(\boldsymbol{q}, \mathcal{A})$ is isomorphic to $T_{\boldsymbol{q}}[x_1, \dots, x_n]$, it is a finite module over its center if and only if each q_{ij} is a root of unity. Using this, we can show that the algebra $V_n(\boldsymbol{q}, \mathcal{A})$ is a finite module over its center if and only if the parameters q_{ij} are all nontrivial roots of unity. Also, when the center of $V_n(\boldsymbol{q}, \mathcal{A})$ is a polynomial ring, $V_n(\boldsymbol{q}, \mathcal{A})$ is a finitely generated free module over its center. The following useful result concerns the centers of $V_n(\boldsymbol{q}, \mathcal{A})$ and $T_{\boldsymbol{q}}[x_1, \dots, x_n]$.

To state it, we need some notation. When q_{ij} is a root of unity, there are two integers k_{ij} and d_{ij} such that

$$q_{ij} = \exp\bigl(2\pi\sqrt{-1}k_{ij}/d_{ij}\bigr),$$

where $d_{ij} := o(q_{ij}) < \infty$, $|k_{ij}| < d_{ij}$ and $(k_{ij}, d_{ij}) = 1$. Further, we can choose k_{ij} so that $k_{ij} = -k_{ji}$, since $q_{ji} = q_{ij}^{-1}$. Let $L_i = \operatorname{lcm}\{d_{ij} \mid j = 1, \dots, n\}$. Let $\overline{Y}$ be the $n \times n$ matrix $(k_{ij} L_i / d_{ij})_{n\times n}$. For each prime p, define $\overline{Y}_p = \overline{Y} \otimes \mathbb{F}_p$. Let m be any natural number. Let $I_{p,m}$ be the set containing i such that $L_i \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}$. Finally, let $\overline{Y}_{p,m}$ be the submatrix of $\overline{Y}_p$ taken from the rows and columns with indices $i \in I_{p,m}$.

Theorem 0.3. *Suppose q_{ij} is a root of unity and not 1 for all $i < j$.*

(1) *The center of $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is a polynomial ring if and only if it is of the form $T\bigl[x_1^{L_1}, \dots, x_n^{L_n}\bigr]$, if and only if $\det(\overline{Y}_{p,m}) \ne 0$ in $\mathbb{F}_p$ for all primes p and all integers $m > 0$ such that $I_{p,m} \ne \varnothing$.*

(2) *If the center of $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is the subalgebra $T\bigl[x_1^{L_1}, \dots, x_n^{L_n}\bigr]$, then the center of $V_n(\boldsymbol{q}, \mathcal{A})$ is the same subalgebra and $V_n(\boldsymbol{q}, \mathcal{A})$ is finitely generated and free over it.*

The above criterion can be simplified when $n = 3$ or 4 [Corollaries 5.4 and 5.5]. The point of Theorem 0.3 is that it provides an explicit linear algebra criterion for when the center of $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is isomorphic to a polynomial ring.

Question 0.4. Suppose that $A := V_n(\boldsymbol{q}, \mathcal{A})$ is finitely generated and free over its center Z. What is the discriminant $d(A/Z)$?

Theorems 0.1 and 0.2 answer this question for two special cases.

A secondary goal of this paper is to provide some quick applications. These discriminant formulas have potential applications in algebraic geometry, number theory and the study of Clifford algebras. In Section 8 (the final section), we give some immediate applications of discriminants to the cancellation problem and the automorphism problem for several classes of noncommutative algebras.

Let us briefly review some definitions. An algebra A is called *cancellative* if $A[t] \cong B[t]$ for some algebra B implies $A \cong B$. Let $\operatorname{Aut}(A)$ be the group of all algebra automorphisms of A. Let A be connected graded. An algebra automorphism g of A is called *unipotent* if

$$g(v) = v + \text{(higher degree terms)}$$

for all homogeneous elements $v \in A$. Let $\operatorname{Aut}_{\text{uni}}(A)$ denote the subgroup of $\operatorname{Aut}(A)$ consisting of all unipotent automorphisms [CPWZ 2016, after Theorem 3.1]. When $\operatorname{Aut}_{\text{uni}}(A)$ is trivial, $\operatorname{Aut}(A)$ is usually small and easy to handle. We will give a criterion on when $\operatorname{Aut}_{\text{uni}}(A)$ is trivial.

Let A be a domain and let F be a subset of A. Let $\operatorname{Sw}(F)$ be the set of $g \in A$ such that $f = agb$ for some $a, b \in A$ and $0 \neq f \in F$. Let $D_1(F)$ be the k-subalgebra of A generated by $\operatorname{Sw}(F)$. For $n > 2$, we define $D_n(F) = D_1(D_{n-1}(F))$ inductively, and define $D(F) = \bigcup_{n \geq 1} D_n(F)$. This algebra is called the *F-divisor subalgebra of A*. When $F = \{d(A/Z)\}$, $D(F)$ is called the *discriminant-divisor subalgebra of A* and is denoted by $\mathbb{D}(A)$. The main result in Section 8 is the following.

Theorem 0.5. *Suppose k is a field of characteristic zero. Let A be a connected graded domain of finite Gelfand–Kirillov dimension. Assume that A is finitely generated and free over its center. If $\mathbb{D}(A) = A$, then A is cancellative and* $\operatorname{Aut}_{\text{uni}}(A) = \{1\}$.

The above theorem can be applied to some Artin–Schelter regular algebras of global dimension 4 in Examples 6.3 and 8.4. Further applications are certainly expected.

This paper is organized as follows. Background material about discriminants is provided in Section 1. We prove Theorem 0.1 in Section 2 and Theorem 0.2 in Section 3. Sections 4–6 concern the question of when $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ and $V_n(\boldsymbol{q}, \mathcal{A})$ are finitely generated and free over their centers and contain the proof of Theorem 0.3. In Section 7, we review and introduce some invariants related to discriminants, locally nilpotent derivations, and automorphisms, which will be used in Section 8. In Section 8, some applications are provided and Theorem 0.5 is proven.

1. Preliminaries

In this section we recall some definitions and basic properties of the discriminant. A basic reference is [CPWZ 2015a, Section 1].

Throughout, let k be a base commutative domain and let everything be over k. Let A be an algebra and let Z be a central subalgebra of A such that A is finitely generated and free over Z. A modified version of the discriminant was introduced in [CPWZ 2016] when A is not free over Z; however, in this paper, we only consider the case when A is finitely generated and free over Z. Let r be the rank of A over Z.

We embed A in the endomorphism ring $\mathrm{End}(A_Z)$ by sending $a \in A$ to the left multiplication $l_a : A \to A$. Since A is free over Z of rank r, $\mathrm{End}(A_Z) \cong M_{r\times r}(Z)$. Define the trace function

$$\mathrm{tr} : A \longrightarrow \mathrm{End}(A_Z) \cong M_{r\times r}(Z) \xrightarrow{\mathrm{tr}_m} Z, \tag{E1.0.1}$$

where tr_m is the usual matrix trace. The trace function tr is independent of the choice of basis of A over Z.

Definition 1.1. [CPWZ 2015a, Definition 1.3(3)] Retain the above notation. Suppose that A is a free module over a central subalgebra Z with a Z-basis $\{z_1, \dots, z_r\}$. The discriminant of A over Z is

$$d(A/Z) = \det(\mathrm{tr}(z_i z_j))_{r\times r} \in Z.$$

By [CPWZ 2015a, Proposition 1.4(2)], $d(A/Z)$ is unique up to a scalar in $Z^\times$. For $x, y \in Z$, we use the notation $x =_{Z^\times} y$ to indicate that $x = cy$ for some $c \in Z^\times$. So $d(A/Z) =_{Z^\times} \det(\mathrm{tr}(z_i z_j))_{r\times r}$ as in [CPWZ 2015a, Definition 1.3(3)]. The following lemma is easy.

Lemma 1.2. *Retain the notation of Definition 1.1. Let (A', Z') be another pair of algebras such that Z' is a central subalgebra of A' and A' is a free Z'-module of rank r. Let $g : A \to A'$ be an algebra homomorphism such that*:

(a) $g(Z) \subseteq Z'$.

(b) $\{g(z_1), \dots, g(z_r)\}$ *is a Z'-basis of A'.*

Then $g(d(A/Z)) =_{(Z')^\times} d(A'/Z')$.

Proof. For any $a \in A$, we define $a' = g(a)$. Write $az_i = \sum_{j=1}^r a_{ij} z_j$ for all i. By applying g to the last equation, we have $a' z'_i = \sum_{j=1}^r a'_{ij} z'_j$. By definition (E1.0.1), $\mathrm{tr}(a) = \sum_i a_{ii}$ and

$$\mathrm{tr}(g(a)) = \mathrm{tr}(a') = \sum_i a'_{ii} = g\left(\sum_i a_{ii}\right) = g(\mathrm{tr}(a))$$

for all $a \in A$. By Definition 1.1 and the above equation,

$$g(d(A/Z)) = g(\det(\operatorname{tr}(z_i z_j))_{r\times r}) = \det(\operatorname{tr}(z'_i z'_j))_{r\times r} =_{(Z')^\times} d(A'/Z'). \qquad \square$$

Let Z be a central subalgebra of A and consider an Ore set $C \subset Z$. Then the localization ZC^{-1} is central in AC^{-1}.

Lemma 1.3. *Let Z be a central subalgebra of A. Suppose A is free over Z of rank r. Then AC^{-1} is free over ZC^{-1} of rank r. As a consequence,*

$$d\big(AC^{-1}/ZC^{-1}\big) =_{(ZC^{-1})^\times} d(A/Z).$$

Proof. Let $\{z_1, \dots, z_r\}$ be a Z-basis of A. Then it is also a ZC^{-1}-basis of AC^{-1}. The consequence follows from Lemma 1.2. $\square$

We will need the following result from [CPWZ 2016]. We use T in place of k to denote a commutative domain.

Proposition 1.4. *Let T be a commutative domain and let $A = T_{\boldsymbol{q}}[x_1, \dots, x_n]$. Suppose $Z := T\big[x_1^{\alpha_1}, \dots, x_n^{\alpha_n}\big]$ is a central subalgebra of A, where the α_i are positive integers.*

(1) [CPWZ 2016, Proposition 2.8] *Let $r = \prod_{i=1}^n \alpha_i$. Then*

$$d(A/Z) =_{T^\times} r^r \Big(\prod_{i=1}^n x_i^{\alpha_i - 1}\Big)^r.$$

(2) *If $n = 2$, $Z = T\big[x_1^m, x_2^m\big]$, and q_{12} is a primitive m-th root of unity, then*

$$d(A/Z) =_{T^\times} m^{2m^2}\big(x_1^m x_2^m\big)^{m(m-1)}.$$

(3) *If $q_{ij} = -1$ for all $i < j$ and $\alpha_i = 2$ for all i, then*

$$d(A/Z) =_{T^\times} 2^{n2^n}\Big(\prod_{i=1}^n x_i^2\Big)^{2^{n-1}}.$$

Proof. Parts (2) and (3) are special cases of part (1). $\square$

The next lemma is a special case of [CPWZ 2016, Proposition 4.10]. Suppose Z is a central subalgebra of A and A is free over Z of rank $r < \infty$. We fix a Z-basis of A, say $b := \{b_1 = 1, b_2, \dots, b_r\}$. Suppose A is an $\mathbb{N}$-filtered algebra such that the associated graded ring $\operatorname{gr} A$ is a domain. For any element $f \in A$, let $\operatorname{gr} f$ denote the associated element in $\operatorname{gr} A$. Let $\operatorname{gr} b$ denote the set $\{\operatorname{gr} b_1, \dots, \operatorname{gr} b_r\}$, which is a subset of $\operatorname{gr} A$.

Lemma 1.5. *Retain the above notation. Suppose that* $\operatorname{gr} A$ *is finitely generated and free over* $\operatorname{gr} Z$ *with basis* $\operatorname{gr} b$. *Then*

$$\operatorname{gr}(d(A/Z)) =_{(\operatorname{gr} Z)^\times} d(\operatorname{gr} A/\operatorname{gr} Z).$$

2. Discriminant of A_q over its center

Let T be a commutative domain and let $q \in T^{\times}$ be a primitive n-th root of unity for some $n \geq 2$. Let A_q be the q-quantum Weyl algebra over T generated by x and y and subject to the relation $yx = qxy + a$ for some $a \in T$. This agrees with the definition of A_q given in the Introduction when $T = k$ and $a = 1$. It is easy to check that the center of A_q, denoted by $Z(A_q)$, is $T[x^n, y^n]$, and that A_q is free over $Z(A_q)$ of rank n^2. A $Z(A_q)$-basis of A_q is $\mathcal{B} := \{x^i y^j \mid 0 \leq i,\ j \leq n-1\}$. The aim of this section is to compute the discriminant $d(A_q/Z(A_q))$.

Let A' be the T-subalgebra of A_q generated by $x' := (1-q)x$ and y. Since $yx' = qx'y + (1-q)a$ and $(1-q)$ may not be invertible, there is no obvious algebra homomorphism from A_q to A'. Let Z' be the subalgebra $T[(x')^n, y^n]$ which is the center of A'.

Lemma 2.1. *Retain the above notation. Then*

$$d(A'/Z') = (1-q)^{n^2(n-1)} d(A_q/Z(A_q)).$$

Proof. Let $\mathrm{tr}' : A' \to Z'$ be the trace function defined in (E1.0.1). We use this trace function to compute the discriminant $d(A'/Z')$.

Let $\mathcal{B}' := \{(x')^i y^j\}_{0 \leq i, j \leq n-1}$. Then $\mathcal{B}'$ is a Z'-basis of A'. Note that A' and A_q have the same ring of fractions and $Z(A_q)$ and Z' have the same fraction field. Since the trace function is independent of the choice of basis, we have $\mathrm{tr}'(a) = \mathrm{tr}(a)$ for all $a \in A'$.

Picking any two elements $b_s = x^{i_s} y^{j_s}$ and $b_t = x^{i_t} y^{j_t}$ in $\mathcal{B}$, we have corresponding elements $b'_s = (x')^{i_s} y^{j_s}$ and $b'_t = (x')^{i_t} y^{j_t}$ in $\mathcal{B}$. Hence

$$\mathrm{tr}'(b'_s b'_t) = \mathrm{tr}\big((1-q)^{i_s+i_t} b_s b_t\big) = (1-q)^{i_s+i_t}\, \mathrm{tr}(b_s b_t).$$

By definition, $d(A'/Z') = \det\big[\mathrm{tr}'(b'_s b'_t)_{b'_s, b'_t \in \mathcal{B}'}\big]$. Hence we have

$$\begin{aligned} d(A'/Z') &= \det\big[(\mathrm{tr}'(b'_s b'_t))_{b'_s, b'_t \in \mathcal{B}'}\big] = \det\big[\big((1-q)^{i_s+i_t}\, \mathrm{tr}(b_s b_t)\big)_{b_s, b_t \in \mathcal{B}}\big] \\ &= (1-q)^N \det\big[(\mathrm{tr}(b_s b_t))_{b_s, b_t \in \mathcal{B}}\big] = (1-q)^N d(A_q/Z(A_q)), \end{aligned}$$

where

$$N = \sum_{\text{all } i_s, i_t} (i_s + i_t) = 2 \sum_{\text{all } i_s} i_s = 2n(0 + 1 + 2 + \cdots + (n-1)) = n^2(n-1).$$

The assertion follows. □

Following the above lemma, we first compute $d(A'/Z')$. We can rewrite A' as $T\langle x', y\rangle/(yx' - qx'y - (1-q)a)$ so that the positions of x' and y are more symmetrical.

Let $C = \{(y^n)^i \mid i \geq 1\}$. Consider the localizations $Z'' := Z'C^{-1}$ and $A'' := A'C^{-1}$. Let

$$x'' := x' - ay^{-1} = (1-q)x - (ay^{-n})y^{n-1} \in A''.$$

Lemma 2.2. *Retain the above notation. The following hold*:

(1) $yx'' - qx''y = 0$.

(2) $A'' := A'C^{-1}$ *is generated by* T, $(y^n)^{-1}$, x'' *and* y.

(3) $(x'')^n$ *is central and* $d(A''/Z'') =_{(Z'')^\times} n^{2n^2}((x'')^n y^n)^{n(n-1)}$.

(4) $d(A''/Z'') =_{(Z'')^\times} n^{2n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}$.

Proof. (1) We have $yx'' - qx''y = y((1-q)x - ay^{-1}) - q((1-q)x - ay^{-1})y = 0$.

(2) This is clear.

(3) Since $q^n = 1$, $(x'')^n$ commutes with y by part (1). By part (2), $(x'')^n$ commutes with every element in A''.

Consider an algebra homomorphism $g : T_q[x_1, x_2] \to A''$ determined by $g(x_1) = x''$ and $g(x_2) = y$. Then the center of $B := T_q[x_1, x_2]$ is $R := T\left[x_1^n, x_2^n\right]$ and $\left\{x_1^i x_2^j \mid 0 \leq i,\ j \leq n-1\right\}$ is an R-basis of B. It is clear that A'' is free of rank n^2 and that $A'' = \sum_{0 \leq i, j \leq n-1} (x')^i y^j Z''$. Hence $\{(x'')^i y^j \mid 0 \leq i,\ j \leq n-1\}$ is a Z''-basis of A''. Then the hypotheses of Lemma 1.2 hold. Applying Lemma 1.2 to g, we have $g(d(B/R)) =_{(Z')^\times} d(A''/Z'')$. By Proposition 1.4(2), $d(B/R) = n^{2n^2}\left(x_1^n x_2^n\right)^{n(n-1)}$. Therefore, $d(A''/Z'') =_{(Z')^\times} n^{2n^2}((x'')^n y^n)^{n(n-1)}$.

(4) In the following, we will let $\psi = y^{-1}$, $z = x''$ and $p = q^{-1}$. The commutation relation between x' and ψ is

$$\psi x' = (1-q)\psi x = (1-q)(px\psi - pa\psi^2) = px'\psi - (p-1)a\psi^2. \qquad \text{(E2.2.1)}$$

Recall that $z = x'' = x' - a\psi$. Write $z^n = \sum_{i=0}^n c_i (x')^i \psi^{n-i}$. Since z^n is central (see part (3)), we have $c_i = 0$ unless $i = 0, n$. It is clear that $c_n = 1$. Next we determine c_0. Since A'' is a free module over Z'' with basis $\{(x')^i \psi^j \mid 0 \leq i,\ j \leq n-1\}$, we can work modulo the right Z''-submodule W generated by $(x')^i \psi^j$, where $0 < i < n$ and $0 \leq j < n$. Let $\equiv$ denote equivalence mod W.

By induction, for $i = 1, \ldots, n-1$, we have

$$\psi^i x' = p^i x' \psi^i - (p^i - 1)(a\psi^{i+1}). \qquad \text{(E2.2.2)}$$

Then $\psi^i x' \equiv -(p^i - 1)(a\psi^{i+1})$. For each $1 \leq j \leq n-1$, write

$$z^j = \sum_{i=0}^{j} c_i^j (x')^i \psi^{j-i}.$$

Then $x'z^j \in W$ for all $j < n-1$ and $x'z^{n-1} \equiv (x')^n$. For each j, we have $\psi^{j-1}z^{n-j} = \sum_{i=0}^{n-j} d_i^j (x')^i \psi^{n-1-i}$ for some $d_i^j \in Z'$, so

$$x'\psi^{j-1}z^{n-j} \in W \tag{E2.2.3}$$

for all $j \geq 2$. By the above computation and (E2.2.1)–(E2.2.3), we have

$$\begin{aligned}
z^n - (x')^n &= (x' - a\psi)z^{n-1} - (x')^n \\
&= x'z^{n-1} - (x')^n - a\psi z^{n-1} \\
&\equiv -a\psi(x' - a\psi)z^{n-2} \\
&\equiv -a(px'\psi - (p-1)a\psi^2 - a\psi^2)z^{n-2} \\
&\equiv -a(-pa)\psi^2 z^{n-2} - apx'\psi z^{n-2} \\
&\equiv -a(-pa)\psi^2 z^{n-2} \\
&\equiv -a(-pa)(\psi^2 x - a\psi^3)z^{n-3} \\
&\equiv -a(-pa)(-p^2 a)\psi^3 z^{n-3} \\
&\vdots \\
&\equiv -a(-pa)(-p^2 a)\cdots(-p^{n-1}a)\psi^n \\
&= (-a)^n p^{(n-1)n/2}\psi^n = -a^n\psi^n.
\end{aligned}$$

Therefore,

$$z^n \equiv -a^n\psi^n + (x')^n.$$

Hence $c_0 = -a^n$ and $z^n = (x')^n - a^n\psi^n$. Combining all of the above, we have

$$(x'')^n y^n = ((x')^n - a^n\psi^n)y^n = (x')^n y^n - a^n = (1-q)^n x^n y^n - a^n.$$

Part (4) follows from part (3) and the above formula. □

Lemma 2.3. *The discriminant of A' over its center Z' is*

$$d(A'/Z') =_{T^\times} n^{2n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}.$$

Proof. Let g be the embedding of A' into $A'' = A'C^{-1}$, viewed as an inclusion. By Lemma 1.2, g sends $d(A'/Z')$ to $d(A''/Z'')$. Combining this fact with Lemma 2.2(4), we have

$$\begin{aligned}
d(A'/Z') &=_{(Z'')^\times} g(d(A'/Z(A'))) =_{(Z'')^\times} d(A''/Z'') \\
&=_{(Z'')^\times} n^{2n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}.
\end{aligned}$$

Let Φ be the element $d(A'/Z')\{n^{2n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}\}^{-1}$, which can be viewed as an element in the quotient ring of A'. By the above equation, Φ is in $(Z'')^\times$. Since $Z'' = T[(x')^n, y^{\pm n}]$, Φ is of the form αy^{sn} for some $\alpha \in T^\times$

and some s. By symmetry, Φ is also of the form $\beta(x')^{tn}$ for some $\beta \in T^\times$ and some t. Hence $s=t=0$, $\alpha=\beta \in T^\times$ and $\Phi=\alpha \in T^\times$. Therefore, $d(A'/Z') = \alpha n^{2n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}$ and the assertion follows. $\square$

Now let

$$m := \prod_{i=2}^{n-1}(1+q+\cdots+q^{i-1}). \tag{E2.3.1}$$

We can show that $n=(1-q)^{n-1}m$ by first factoring the polynomial $x^n-1 \in T[x]$ and dividing by $(x-1)$:

$$x^n - 1 = \prod_{i=0}^{n-1}(x-q^i) \quad \Longrightarrow \quad \sum_{i=0}^{n-1} x^i = \frac{x^n-1}{x-1} = \prod_{i=1}^{n-1}(x-q^i).$$

We then substitute 1 for x as follows:

$$n = \prod_{i=1}^{n-1}(1-q^i) = (1-q)^{n-1}\prod_{i=2}^{n-1}(1+q+\cdots+q^{i-1}) = (1-q)^{n-1}m. \tag{E2.3.2}$$

Now we are ready to prove the main result of this section, which also recovers Theorem 0.1.

Theorem 2.4. *Retain the above notation. The discriminant of A_q over its center $Z(A_q)$ is*

$$d(A_q/Z(A_q)) =_{T^\times} (nm)^{n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}.$$

Proof. Using Lemmas 2.1 and 2.3 and equation (E2.3.2), we have

$$(1-q)^{n^2(n-1)} d(A_q/Z(A_q)) =_{T^\times} (nm(1-q)^{n-1})^{n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}.$$

Since A_q is a domain, we obtain

$$d(A_q/Z(A_q)) =_{T^\times} (nm)^{n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}. \qquad \square$$

Remark 2.5. (1) By [CPWZ 2016, Lemma 2.7(7)], the integer n in Theorem 2.4 is nonzero in T. However, n and m may not be invertible in general.

(2) Theorem 0.1 is clearly a consequence of Theorem 2.4.

A slight generalization of Theorem 2.4 is the following.

Theorem 2.6. *Let T be a commutative domain and $q \in T^\times$ be a primitive n-th root of unity. Let B be the T-algebra of the form*

$$\frac{T\langle x, y\rangle}{(yx - qxy = a,\ x^n = b,\ y^n = c)},$$

where $a, b, c \in T$. Suppose that B is a free module over T with basis $\{x^i y^j \mid 0 \le i,\ j \le n-1\}$. Then $d(B/T) =_{T^\times} (nm)^{n^2}((1-q)^n x^n y^n - a^n)^{n(n-1)}$, where m is given in (E2.3.1).

Proof. First note that it is well-known and easy to check that T is the center of B.

Recall that A_q is the algebra of the form $T\langle x, y\rangle/(yx - qxy = a)$. There is a natural algebra homomorphism g from A_q to B sending x to x and y to y and $t \in T$ to $t \in T$. Then the hypotheses in Lemma 1.2 hold. By Lemma 1.2, $g(d(A_q/Z(A_q))) = d(B/T)$. Now the assertion follows from Theorem 2.4. □

3. Discriminant of Clifford algebras

In this section we assume that $2^{-1} \in k$. We fix an integer $n \ge 2$.

Let T be a commutative domain and let $\mathcal{A} := \{a_{ij} \mid 1 \le i < j \le n\}$ be a set of scalars in T. We write $a_{ji} = a_{ij}$ if $i < j$. Let $V_n(\mathcal{A})$ be the T-algebra generated by $x_1, \dots, x_n$ and subject to the relations

$$x_i x_j + x_j x_i = a_{ij} \quad \text{for all } i \ne j.$$

This algebra was studied in [CPWZ 2015a; 2015b]. Some basic properties of $V_n(\mathcal{A})$ are given in [CPWZ 2015a, Section 4]. Let M_1 be the matrix

$$M_1 := \begin{pmatrix} 2x_1^2 & a_{12} & \cdots & a_{1n} \\ a_{21} & 2x_2^2 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 2x_n^2 \end{pmatrix}. \tag{E3.0.1}$$

This is a symmetric matrix with entries in $Z := T\left[x_1^2, \dots, x_n^2\right]$. We will define a sequence of matrices M_i later. Note that Z is a central subalgebra of $V_n(\mathcal{A})$. If we write $M_1 = (m_{ij,1})_{n\times n}$, then $m_{ij,1} = x_j x_i + x_i x_j$ for all i, j.

The algebra $V_n(\mathcal{A})$ is a Clifford algebra over Z. We will recall the definition of the Clifford algebra associated to a quadratic form in the second half of this section. In the next few lemmas, we are basically diagonalizing the quadratic form, which is elementary and well-known in the classical case; see [Lam 2005, Chapter I, Corollary 2.4] for some related material. Since we need an explicit construction to complete the proof of our main result, details will be provided below.

We will introduce a sequence of new variables starting with

$$x_{i,1} = x_i \quad \text{for all } i = 1, \dots, n,$$

and

$$a_{ij,1} = a_{ij} \quad \text{for all } i \ne j, \quad \text{and} \quad a_{ii,1} = 2x_i^2 \quad \text{for all } i.$$

So we have $x_{j,1}x_{i,1}+x_{i,1}x_{j,1}=a_{ij,1}$ for all $i,\ j$. Let

$$x_{1,2}:=x_{1,1} \quad \text{and} \quad x_{i,2}:=x_{i,1}-\tfrac{1}{2}a_{1i,1}x_{1,1}^{-2}x_{1,1} \quad \text{for all } i\geq 2. \tag{E3.0.2}$$

Lemma 3.1. *Retain the above notation.*

(1) $x_{i,2}x_{1,2}+x_{1,2}x_{i,2}=0$ *for all* $i\geq 2$.

(2) $x_{i,2}^2=x_{i,1}^2-\frac{1}{4}a_{1i,1}^2x_{1,1}^{-2}$ *for all* $i\geq 2$.

(3) $x_{i,2}x_{j,2}+x_{j,2}x_{i,2}=a_{ij,1}-\frac{1}{2}a_{1i,1}a_{1j,1}x_{1,1}^{-2}$ *for all* $2\leq i<j\leq n$.

(4) *Let* M_2 *be the matrix* $(x_{i,2}x_{j,2}+x_{j,2}x_{i,2})_{1\leq i,j\leq n}$. *Then* $\det M_2=\det M_1$.

(5) *Let*
$$C_1=\{x_{1,1}^{2i}\}_{i\geq 1}.$$
Then the localization $V_n(\mathcal{A})[C_1^{-1}]$ *is free over* $Z[C_1^{-1}]$ *with basis* $\{x_{1,2}^{d_1}\cdots x_{n,2}^{d_n} \mid d_s=0,1\}$.

Proof. (1)–(3) These follow by direct computation.

(4) Let N be the matrix

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\frac{1}{2}a_{12,1}x_{1,1}^{-2} & 1 & 0 & \cdots & 0 \\ -\frac{1}{2}a_{13,1}x_{1,1}^{-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{2}a_{1n,1}x_{1,1}^{-2} & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

By linear algebra and part (3), one can check that $NM_1N^T=M_2$. Since $\det N=1$, we have $\det M_2=\det M_1$.

(5) First of all, $V_n(\mathcal{A})$ is free over Z with basis $\{x_{1,1}^{d_1}\cdots x_{n,1}^{d_n} \mid d_s=0,1\}$. In the localization $V_n(\mathcal{A})[C_1^{-1}]$, this basis can be transformed to a basis $\{x_{1,2}^{d_1}\cdots x_{n,2}^{d_n} \mid d_s=0,1\}$ by using (E3.0.2). □

After we have $x_{i,2}$, define $a_{ij,2}$ to be $x_{i,2}x_{j,2}+x_{j,2}x_{i,2}$ for all $i,\ j$. Now we define $x_{i,s}$ and $a_{ij,s}$ inductively.

Definition 3.2. Let $s\geq 3$ and suppose that $x_{i,s-1}$ and $a_{ij,s-1}$ are defined inductively. Define

$$\begin{aligned} x_{i,s} &:= x_{i,s-1} && \text{for all } i<s, \\ x_{i,s} &:= x_{i,s-1}-\tfrac{1}{2}a_{s-1i,s-1}x_{s-1,s-1}^{-1} && \text{for all } i\geq s. \end{aligned} \tag{E3.2.1}$$

Define $a_{ij,s}:=x_{i,s}x_{j,s}+x_{j,s}x_{i,s}$ for all $i,\ j$.

Similar to Lemma 3.1, we have the following lemma. Its proof is also similar to the proof of Lemma 3.1, so it is omitted.

Lemma 3.3. *Retain the above notation. Let* $2 \le s \le n$.

(1) $x_{i,s}x_{j,s} + x_{j,s}x_{i,s} = 0$ *for all* $i < j$ *and* $i < s$.

(2) $x_{i,s} = x_{i,s-1}$ *if* $i < s$ *and* $x_{i,s}^2 = x_{i,s-1}^2 - \frac{1}{4}a_{s-1i,s-1}^2 x_{s-1,s-1}^{-2}$ *for all* $i \ge s$.

(3) $x_{i,s}x_{j,s} + x_{j,s}x_{i,s} = a_{ij,s-1} - \frac{1}{2}a_{s-1i,s-1}a_{s-1j,s-1}x_{s-1,s-1}^{-2}$ *for all* $s \le i < j \le n$.

(4) *Let* M_s *be the matrix* $(x_{i,s}x_{j,s} + x_{j,s}x_{i,s})_{1 \le i,j \le n}$. *Then* $\det M_s = \det M_1$.

(5) *Let* C_{s-1} *be the Ore set*

$$\{x_{1,1}^{2i_1}x_{2,2}^{2i_1}\cdots x_{s-1,s-1}^{2i_{s-1}}\}_{i_1,\ldots,i_{s-1}\ge 1}.$$

Then the localization $V_n(\mathcal{A})[C_{s-1}^{-1}]$ *is free over* $Z[C_{s-1}^{-1}]$ *with basis* $\{x_{1,s}^{d_1}\cdots x_{n,s}^{d_n} \mid d_s = 0, 1\}$.

We need two more lemmas before we prove the main result.

Lemma 3.4. *Let* T *be a commutative domain. Let* A *be a* T*-algebra containing* T *as a subalgebra, generated by* $x_1, \ldots, x_n$ *and satisfying the relations* $x_jx_i + x_ix_j = 0$ *for all* $i < j$ *and* $x_i^2 = a_i \in T$. *Suppose that* A *is a free module over* T *with basis* $\{x_1^{d_1}\cdots x_n^{d_n} \mid d_s = 0, 1\}$. *Then*

$$d(A/T) =_{T^\times} \left(\prod_{i=1}^n 2x_i^2\right)^{2^{n-1}} =_{T^\times} \left(\prod_{i=1}^n x_i^2\right)^{2^{n-1}}.$$

Proof. Let $B = T_{-1}[x_1, \ldots, x_n]$ and $Z = T[x_1^2, \ldots, x_n^2]$. Then B is a free module over Z with basis $\{x_1^{d_1}\cdots x_n^{d_n} \mid d_s = 0, 1\}$. Let g be the algebra map from B to A sending T to T, x_i to x_i. Then the hypotheses in Lemma 1.2 holds. By Lemma 1.2, $g(d(B/Z)) =_{T^\times} d(A/T)$. Note that $d(B/Z)$ was computed in Proposition 1.4(3) to be $\left(\prod_{i=1}^n 2x_i^2\right)^{2^{n-1}}$, as we assume that 2 is invertible. Now the assertion follows. □

Let A be an Ore domain and let $Q(A)$ denote the skew field of fractions of A. Let Z be the commutative subalgebra $T[x_1^2, \ldots, x_n^2] \subset V_n(\mathcal{A})$. For each $1 \le 1 \le n$, let Z_i be the subring of $Q(Z)$ of the form

$$Q(T[x_1^2, \ldots, \widehat{x_i^2}, \ldots, x_n^2])[x_i^2].$$

Lemma 3.5. *Retain the above notation.*

(1) $\bigcap_{i=1}^n Z_i = Q(T)[x_1^2, \ldots, x_n^2]$.

(2) $Z[C_{n-1}^{-1}] \subseteq Z_n$, *where* $Z[C_{n-1}^{-1}]$ *is defined in Lemma 3.3(5).*

Proof. (1) This is an easy commutative algebra fact.

(2) By Lemma 3.3(2) and induction, each $x_{i,s}^2$, for all $1 \le i < n$ and all $1 \le s \le n$, is in $Q(T[x_1^2, \ldots, x_{n-1}^2])$. So $Z[C_{n-1}^{-1}] \subseteq Z_n$. □

Theorem 3.6. *Suppose 2 is invertible. Let $Z = T\left[x_1^2, \dots, x_n^2\right]$. Then*

$$d(V_n(\mathcal{A})/Z) =_{T^\times} (\det M_1)^{2^{n-1}},$$

where M_1 is given in (E3.0.1).

Proof. Consider the variables $\{x_{i,n}\}_{i=1}^n$ defined in Lemma 3.3. By Lemma 3.3(5), $V_n(\mathcal{A})\left[C_{n-1}^{-1}\right]$ is free over $Z\left[C_{n-1}^{-1}\right]$ with basis $\left\{x_{1,s}^{d_1} \cdots x_{n,s}^{d_n} \mid d_s = 0, 1\right\}$. By Lemma 3.4, the discriminant

$$d\left(V_n(\mathcal{A})\left[C_{n-1}^{-1}\right]/Z\left[C_{n-1}^{-1}\right]\right)$$

is of the form $\left(\prod_{i=1}^n x_i^2\right)^{2^{n-1}}$ up to a unit in $Z\left[C_{n-1}^{-1}\right]$. By Lemma 3.3(4), we have

$$d\left(V_n(\mathcal{A})\left[C_{n-1}^{-1}\right]/Z\left[C_{n-1}^{-1}\right]\right) = \left(\prod_{i=1}^n x_i^2\right)^{2^{n-1}} = (\det M_n)^{2^{n-1}} = (\det M_1)^{2^{n-1}}.$$

By Lemma 1.3,

$$d(V_n(\mathcal{A})/Z) =_{(Z[C_{n-1}^{-1}])^\times} d\left(V_n(\mathcal{A})\left[C_{n-1}^{-1}\right]/Z\left[C_{n-1}^{-1}\right]\right) =_{(Z[C_{n-1}^{-1}])^\times} (\det M_1)^{2^{n-1}}.$$

Let Φ be the element $d(V_n(\mathcal{A})/Z)^{-1}(\det M_1)^{2^{n-1}}$. Then $\Phi \in \left(Z\left[C_{n-1}^{-1}\right]\right)^\times$. This means that both Φ and Φ^{-1} are in $Z\left[C_{n-1}^{-1}\right] \subseteq Z_n$. By symmetry, Φ is Z_i for all i. Thus Φ is in $\bigcap_{i=1}^n Z_i = Q(T)\left[x_1^2, \dots, x_n^2\right]$. Similarly, Φ^{-1} is in $Q(T)\left[x_1^2, \dots, x_n^2\right]$. Therefore, $\Phi, \Phi^{-1} \in Q(T)$.

Write $d(V_n(\mathcal{A})/Z) = c(\det M_1)^{2^{n-1}}$, where $c = \Phi^{-1} \in Q(T)$. It remains to show $c \in Z^\times$. Note that $V_n(\mathcal{A})$ is a filtered algebra such that $\operatorname{gr} V_n(\mathcal{A}) \cong T_{-1}[x_1, \dots, x_n]$. By Lemma 1.5,

$$\operatorname{gr} d(V_n(\mathcal{A})/Z) =_{Z^\times} d(\operatorname{gr} V_n(\mathcal{A})/\operatorname{gr} Z).$$

The left-hand side of the above is $c\left(\prod_{i=1} x_i^2\right)^{2^{n-1}}$ and the right-hand side of the above is $\left(\prod_{i=1} x_i^2\right)^{2^{n-1}}$ by Proposition 1.4(3) (assuming 2 is invertible). Thus $c \in Z^\times$, as required. □

Theorem 0.2 is a special case of Theorem 3.6 by taking $a_{ij} = 1$ for all $i < j$.

The algebras $V_n(\mathcal{A})$ and W_n are special Clifford algebras. Now we consider a Clifford algebra in a more general setting. Let T be a commutative domain and let V be a free T-module of rank n. Given a quadratic form $q : V \to T$, we can associate to this data the Clifford algebra

$$C(V, q) = \frac{T\langle V\rangle}{(x^2 - q(x) \mid x \in V)}.$$

Note that this q is different from the parameter q in the definition of the q-quantum Weyl algebra A_q and the parameter set $\boldsymbol{q}$ in the $V_n(\boldsymbol{q}, \mathcal{A})$ and $T_{\boldsymbol{q}}[x_1, \dots, x_n]$. Consider the bilinear form associated to q,

$$b(x, y) = \tfrac{1}{2}(q(x+y) - q(x) - q(y)) \tag{E3.6.1}$$

for all $x, y \in V$. If we choose a T-basis $x_1, \dots, x_n$ for V and let

$$\mathfrak{B} := (b_{ij}) = (b(x_i, x_j))_{n\times n} \in T^{n\times n} \tag{E3.6.2}$$

be the symmetric matrix which represents b with respect to this basis, then the relations of $C(V, q)$ are

$$x_i x_j + x_j x_i = 2b_{ij} \quad \text{for all } i, j. \tag{E3.6.3}$$

Define $\det(q)$ to be $\det(\mathfrak{B})$.

The following main result is a consequence of Theorem 3.6 and Lemma 1.2.

Theorem 3.7. *Let $A := C(V, q)$ be a Clifford algebra over a commutative domain T defined by a quadratic form $q : V \to T$. Pick a T-basis of V, say $\{x_i\}_{i=1}^n$. Then*

$$d(A/T) =_{T^\times} (\det(x_i x_j + x_j x_i)_{n\times n})^{2^{n-1}} =_{T^\times} \det(q)^{2^{n-1}}. \tag{E3.7.1}$$

Proof. Let $b : V^{\otimes 2} \to T$ be the symmetric bilinear form associated to the quadratic form q. Let $a_{ij} = 2b(x_i, x_j)$ for all $i < j$ and $\mathcal{A} = \{a_{ij}\}_{1\le i<j\le n}$. Then there is a canonical algebra surjection $\pi : V_n(\mathcal{A}) \to C(V, q)$ sending $x_i \to x_i$ for all $i = 1, \dots, n$ and $t \to t$ for all $t \in T$, and the kernel of π is the ideal generated by $\{x_i^2 - b_{ii}\}_{i=1}^n$. Clearly, $\pi\left(T\left[x_1^2, \dots, x_n^2\right]\right) = T$ and the matrix $(x_i x_j + x_j x_i)_{n\times n}$ equals M_1. It is easy to check that $\left\{x_1^{d_1} \cdots x_n^{d_n} \mid d_i = 0, 1\right\}$ is a basis of $V_n(\mathcal{A})$ over $T\left[x_1^2, \dots, x_n^2\right]$ and a basis of $C(V, q)$ over T. The first equation of (E3.7.1) follows from Theorem 3.6 and Lemma 1.2 and the second equation follows from the fact that $2\mathfrak{B} = (x_i x_j + x_j x_i)_{n\times n}$ and 2 is invertible. $\square$

In the rest of this section we briefly discuss “generic Clifford algebras”, which will appear again in Section 8. (This generic Clifford algebra should be called a “universal Clifford algebra”, but the term “universal Clifford algebra” has already been used).

Fix an integer n. Let I be the set $\{(i, j) \mid 1 \le i \le j \le n\}$ that can be thought of as the quotient set $\{(i, j) \mid 1 \le i,\ j \le n\}/((i, j) \sim (j, i))$. Let w denote the integer $\frac{1}{2}n(n+1)$. There is a bijection between I and the set of the first w integers $\{1, 2, \dots, w\}$. Let T_g be the commutative domain $k[t_{(i,j)} \mid (i, j) \in I]$, which is isomorphic to $k[t_1, \dots, t_w]$. Define a T_g-algebra A_g generated by $x_1, \dots, x_n$ and subject to the relations

$$x_i x_j + x_j x_i = 2t_{(i,j)} \quad \text{for all } 1 \le i \le j \le n. \tag{E3.7.2}$$

Let $V_g = \bigoplus_{i=1}^n T_g x_i$. Define a bilinear form $b_g : V_g \otimes V_g \to T_g$ by $b_g(x_i, x_j) = t_{(i,j)}$ and the associated quadratic form by $q_g(x) = b_g(x, x)$ for all $x \in V_g$. The "generic Clifford algebra" A_g is defined to be the Clifford algebra associated to (V_g, q_g). For any Clifford algebra $C(V, q)$ over a commutative ring T, by comparing (E3.6.3) with (E3.7.2), one sees that there is an algebra map $A_g \to C(V, q)$ sending $x_i \to x_i$ and $t_{(i,j)} \to b_{ij}$. Define $\deg x_i = 1$ for all i and $\deg t_{(i,j)} = 2$ for all $(i, j) \in I$. Then A_g is a connected graded algebra over k.

We also define some factor algebras of A_g. Let J be a subset of $\{(i, j) \mid 1 \le i < j \le n\}$ and let w_J denote the integer $w - |J|$. Let $T_{g,J}$ be the commutative polynomial ring $k[t_{i,j} \mid (i, j) \in I \setminus J]$, which is isomorphic to $k[t_1, \dots, t_{w_J}]$. Define a $T_{g,J}$-algebra $A_{g,J}$ generated by $x_1, \dots, x_n$ and subject to the relations

$$x_i x_j + x_j x_i = \begin{cases} 2t_{(i,j)}, & (i, j) \in I \setminus J, \\ 0, & (i, j) \in J. \end{cases} \tag{E3.7.3}$$

Let $V_{g,J} = \bigoplus_{i=1}^n T_{g,J} x_i$. Define a bilinear form $b_{g,J} : V_{g,J} \otimes V_{g,J} \to T_{g,J}$ by

$$b_{g,J}(x_i, x_j) = \begin{cases} t_{(i,j)}, & (i, j) \in I \setminus J, \\ 0, & (i, j) \in J, \end{cases}$$

and the associated quadratic form by $q_{g,J}(x) = b_g(x, x)$ for all $x \in V_{g,J}$. Then $A_{g,J}$ is the Clifford algebra associated to $(V_{g,J}, q_{g,J})$. If $J \subseteq J' \subseteq \{(i, j) \mid 1 \le i < j \le n\}$, there is an algebra map $A_{g,J} \to A_{g,J'}$ sending $x_i \to x_i$ and

$$t_{(i,j)} \to \begin{cases} t_{(i,j)}, & (i, j) \notin J', \\ 0, & (i, j) \in J' \setminus J. \end{cases}$$

In particular, $A_{g,J}$ is a connected graded factor ring of A_g.

In part (4) of the next lemma, we will use a few undefined concepts that are related to the homological properties of an algebra. We refer to [Levasseur 1992; Lu et al. 2007; Rogalski and Zhang 2012] for definitions.

Lemma 3.8. *Retain the above notation. Assume that k is a field of characteristic not 2. Let J' be subset of $\{(i, j) \mid 1 \le i < j \le n\}$ and let $J = J' \setminus \{(i_0, j_0)\}$ for some $(i_0, j_0) \in J'$.*

(1) *The Hilbert series of A_g is*

$$H_{A_g}(t) = \frac{(1+t)^n}{(1-t^2)^w}, \quad \textit{where } w = \tfrac{1}{2}n(n+1).$$

(2) *The Hilbert series of $A_{g,J}$ is*

$$H_{A_{g,J}}(t) = \frac{(1+t)^n}{(1-t^2)^{w_J}}, \quad \textit{where } w_J = w - |J|.$$

(3) *$t_{(i_0,j_0)}$ is a central regular element in $A_{g,J'}$, and $A_{g,J} = A_{g,J'}/(t_{(i_0,j_0)})$.*

(4) *A_g and $A_{g,J}$ are connected graded Artin–Schelter regular, Auslander regular, Cohen–Macaulay noetherian domains.*

Proof. (1) Note that A_g is a free module over T_g with basis $\{x_1^{d_1}\cdots x_n^{d_n} \mid d_s = 0, 1\}$. Recall that $\deg x_i = 1$ and $\deg t_{(i,j)} = 2$. We have

$$H_{A_g}(t) = (1+t)^n H_{T_g}(t) = \frac{(1+t)^n}{(1-t^2)^w}.$$

(2) The proof is similar. Use the fact that $H_{T_{g,J}}(t) = 1/(1-t^2)^{w_J}$.

(3) It is clear that $t_{(i_0,j_0)}$ is central in $A_{g,J'}$ and that $A_{g,J} = A_{g,J'}/(t_{(i_0,j_0)})$. So the ideal $(t_{(i_0,j_0)})$ is the left ideal $t_{(i_0,j_0)}A_{g,J'}$ and the right ideal $A_{g,J'}t_{(i_0,j_0)}$. By parts (1) and (2), the Hilbert series of $(t_{(i_0,j_0)})$ is $t^2 H_{A_{g,J'}}(t)$. So $t_{(i_0,j_0)}$ is regular.

(4) We only provide a proof for A_g. The proof for $A_{g,J}$ is similar.

From part (3), $J_M := \{t_{(i,j)} \mid 1 \le i < j \le n\}$ is a sequence of regular central elements in A_g of positive degree. It is easy to see that $A_{g,J_M} (= A_g/(J_M))$ is isomorphic to the skew polynomial ring $k_{-1}[x_1, \ldots, x_n]$, which is an Artin–Schelter regular, Auslander regular, Cohen–Macaulay noetherian domain. Applying [Lu et al. 2007, Lemma 7.6] repeatedly, A_g has finite global dimension. Applying [Levasseur 1992, Proposition 3.5, Theorem 5.10] repeatedly, A_g is a noetherian Auslander Gorenstein and Cohen–Macaulay domain. By [Levasseur 1992, Theorem 6.3], A_g is Artin–Schelter Gorenstein. Since A_g has finite global dimension, it is Auslander regular and Artin–Schelter regular. □

Remark 3.9. Retain the above notation. (1) Some homological properties of the algebra A_g are given in Lemma 3.8. It would be interesting to work out combinatorial and geometric invariants (and properties) of A_g. For example, what are the point-module and line-module schemes of A_g? Definitions of these schemes can be found in [Vancliff and Van Rompay 2000; Vancliff et al. 1998].

(2) Another way of presenting A_g is the following. Let S be a k-vector space of dimension n. Define A_g to be $k\langle S\rangle/([x^2, y] = 0 \mid \text{for all } x, y, \in S)$. By using this new expression, one can easily see that the group of graded algebra automorphisms of A_g, denoted by $\mathrm{Aut}_{\mathrm{gr}}(A_g)$, is isomorphic to $\mathrm{GL}_n(k)$.

(3) Suppose $n \ge 2$. The full automorphism group $\mathrm{Aut}(A_g)$ has not been determined. It is known that $\mathrm{Aut}(A_g)$ is not affine. For example, if $f(t)$ is a polynomial in t, then

$$x_i \to \begin{cases} x_i, & i > 1, \\ x_1 + f([x_1, x_2]^2)x_2, & i = 1, \end{cases}$$

extends to an algebra automorphism of A_g.

(4) It seems interesting to study the "cubic algebra" $k\langle S\rangle/([x^3, y] = 0 \mid \text{for all } x, y \in S)$ and higher-degree analogues.

(5) The quotient division ring of A_g, denoted by D_g, is called the "generic Clifford division algebra of rank n". It would be interesting to study algebraic properties or invariants of D_g.

4. Center of skew polynomial rings

To use the discriminant most effectively, one needs to first understand the center of an algebra. In this section we give a criterion for when $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is free over its center and when the center of $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is a polynomial ring.

Recall that T is a commutative domain and $\boldsymbol{q} := \{q_{ij} \in T^\times \mid 1 \le i < j \le n\}$ is a set of invertible scalars. Let $P := T_{\boldsymbol{q}}[x_1, \dots, x_n]$ be the skew polynomial ring over T and subject to the relations (E0.2.1). We assume that $d_{ij} := o(q_{ij}) < \infty$ and write

$$q_{ij} = \exp\bigl(2\pi\sqrt{-1}k_{ij}/d_{ij}\bigr), \tag{E4.0.1}$$

where $|k_{ij}| < d_{ij}$ and $(k_{ij}, d_{ij}) = 1$. Note that, by our convention, $q_{ij} = q_{ji}^{-1}$ for all i, j. Hence, we choose $k_{ij} = -k_{ji}$ and $d_{ij} = d_{ji}$. We also adopt the convention that if $q_{ij} = 1$ then $k_{ij} = 0$ and $d_{ij} = 1$. In particular, $k_{ii} = 0$ and $d_{ii} = 1$. We can extend P to $P\bigl[x_1^{-1}, \dots, x_n^{-1}\bigr]$, with an inverse for each x_i, with the expected relations

$$x_i x_i^{-1} = x_i^{-1} x_i = 1, \quad x_j x_i^{-1} = q_{ij}^{-1} x_i^{-1} x_j, \quad \text{and} \quad x_j^{-1} x_i^{-1} = q_{ij} x_i^{-1} x_j^{-1}.$$

We need to do some analysis to understand the center of P. Let η_i denote conjugation by x_i, sending $f \mapsto x_i^{-1} f x_i$, and let $\xi = x_1^{s_1} \cdots x_n^{s_n}$. Then

$$\eta_i(\xi) = \exp\bigl(2\pi\sqrt{-1}\boldsymbol{e}_i^T Y \boldsymbol{s}\bigr)\xi,$$

where $Y \in \mathfrak{so}_n(\mathbb{Q})$ has (i, j)-th entry k_{ij}/d_{ij}, $\boldsymbol{s}$ is the column vector whose i-th entry is s_i appearing in the powers of ξ, and $\boldsymbol{e}_i$ is the i-th standard basis vector in $\mathbb{Q}^n$.

Lemma 4.1. *Retain the above notation. Then ξ is in the center $Z(P)$ of P if and only if $Y\boldsymbol{s} \in \mathbb{Z}^n$.*

Proof. Since P is generated by $\{x_i\}$, we have $\xi \in Z(P)$ if and only if $\eta_i(\xi) = \xi$ for all i, if and only if $\exp\bigl(2\pi\sqrt{-1}\boldsymbol{e}_i^T Y\boldsymbol{s}\bigr) = 1$, if and only if $\boldsymbol{e}_i^T Y\boldsymbol{s} \in \mathbb{Z}$ for all i, and finally, if and only if $Y\boldsymbol{s} \in \mathbb{Z}^n$. □

By choosing the standard basis for $\mathbb{Q}^n$, we can consider Y as a linear transformation $\mathbb{Q}^n \to \mathbb{Q}^n$ by sending $\boldsymbol{s} \mapsto Y\boldsymbol{s}$. Here we view $\mathbb{Q}^n$ as column vectors and Y as a left multiplication. We can restrict this map to $\mathbb{Z}^n \subset \mathbb{Q}^n$ (embedded via the standard basis) and compose with the quotient $\mathbb{Q}^n \to \mathbb{Q}^n/\mathbb{Z}^n$ to obtain a $\mathbb{Z}$-module homomorphism $Y' : \mathbb{Z}^n \to \mathbb{Q}^n/\mathbb{Z}^n$.

Lemma 4.2. *Retain the above notation. Then $\xi \in Z(P)$ if and only if $\boldsymbol{s} \in \ker(Y')$.*

Proof. By Lemma 4.1, $\xi \in Z(P)$ if and only if $Y\boldsymbol{s} \in \mathbb{Z}^n$, which is equivalent to $Y'(\boldsymbol{s}) = 0$ by the definition of Y'. □

Let D be the matrix $(d_{ij})_{n\times n}$ and let L_i be the lcm of the entries in the i-th row of D, namely, $L_i = \operatorname{lcm}\{d_{ij} \mid j = 1, \dots, n\}$. Since D is a symmetric matrix, L_i is also the lcm of the entries in i-th column. Observe that $Z(P)$ contains the central subring $P' := k\big[x_1^{L_1}, \dots, x_n^{L_n}\big]$. In other words, $\ker(Y')$ contains the $\mathbb{Z}$-lattice Λ spanned by $L_i \boldsymbol{e}_i$ for $i = 1, \dots, n$. Therefore, Y' factors through

$$\mathbb{Z}^n \to M := \mathbb{Z}^n/\Lambda = \bigoplus_{i=1}^{n} \mathbb{Z}/L_i\mathbb{Z}.$$

For each $\boldsymbol{s} \in \mathbb{Z}^n$, the i-th entry of $Y'(\boldsymbol{s})$ is $\sum_j k_{ij}s_j/d_{ij} \in \mathbb{Q}/\mathbb{Z}$, which is L_i-torsion, or equivalently, in $L_i^{-1}\mathbb{Z}/\mathbb{Z}$. Therefore, Y' induces a map

$$M \to M' := \bigoplus_{i=1}^{n} L_i^{-1}\mathbb{Z}/\mathbb{Z}.$$

Since M' is naturally isomorphic to M, we can define an endomorphism

$$\overline{Y} : M \to M$$

by setting

$$\overline{Y}\boldsymbol{s} = \left(\sum_{j=1}^{n} L_i(k_{ij}s_j/d_{ij})\right)_{i=1}^{n}.$$

In particular, $\overline{Y}\boldsymbol{e}_j = \sum_{i=1}^{n}(k_{ij}L_i/d_{ij})\boldsymbol{e}_i$. Sometimes we think of $\overline{Y}$ as a matrix:

$$\overline{Y} = (k_{ij}L_i/d_{ij})_{n\times n} = \operatorname{diag}(L_1, \dots, L_n)Y.$$

The following lemma is a reinterpretation of [CPWZ 2016, Lemma 2.3].

Lemma 4.3. *Retain the above notation. The following are equivalent.*

(1) *The center $Z(P)$ of P is a polynomial ring.*

(2) $Z(P) = P'$.

(3) $\ker(\overline{Y}) = 0$.

(4) $\overline{Y}$ *is an isomorphism.*

Proof. (1) $\Leftrightarrow$ (2): One implication is clear. For the other implication, we assume that the center $Z(P)$ is a polynomial ring. By [CPWZ 2016, Lemma 2.3], $Z(P)$ is of the form $T\big[x_1^{a_i}, \dots, x_n^{a_i}\big]$. It is easy to check that $L_i \mid a_i$ for all i. Since $Z(P) \supseteq P'$, $a_i = L_i$ for all i. The assertion follows.

(3) $\Rightarrow$ (2): Let $\xi := x_1^{s_1} \cdots x_n^{s_n} \in Z(P)$ and $\boldsymbol{s} = (s_i)_{i=1}^{n}$. By Lemma 4.2, $\boldsymbol{s} \in \ker(Y')$. Since $\overline{Y}$ is induced by Y', we have $\overline{Y}(\boldsymbol{s}) = 0$. By part (3), $\boldsymbol{s} = 0$ in $M = \mathbb{Z}^n/\Lambda$. So $\boldsymbol{s} \in \Lambda$, which is equivalent to $\xi \in P'$. Therefore, $Z(P) = P'$, as desired.

(2) $\Rightarrow$ (3): Let $\xi := x_1^{s_1} \cdots x_n^{s_n} \in P$, where $\boldsymbol{s} := (s_i)_{i=1}^{n} \in \ker(\overline{Y})$, viewed as a vector in M. By the definition of M, we might assume that each s_i is nonnegative and less

than L_i. Since $\overline{Y}$ is induced by Y', we have $\boldsymbol{s} \in \ker(Y')$. By Lemma 4.2, $\xi \in Z(P)$. By part (2) and our choice of $0 \le s_i < L_i$, we have $\xi = 1$ or $\boldsymbol{s} = 0$, as desired.

(3) $\Leftrightarrow$ (4): This is clear since M is finite. □

The advantage of working with $\overline{Y}$ is that $\ker(\overline{Y}) = 0$ is equivalent to $\overline{Y}$ being an isomorphism. Next we need to understand when $\overline{Y}$ is an isomorphism. For the rest of this section we use $\otimes$ for $\otimes_{\mathbb{Z}}$ and $\mathbb{F}_p$ for $\mathbb{Z}/p\mathbb{Z}$.

Lemma 4.4. *The morphism $\overline{Y}$ is an isomorphism if and only if $\overline{Y} \otimes \mathbb{F}_p$ is an isomorphism for all primes p.*

Proof. As a $\mathbb{Z}$-module, M is finite, and it suffices to show that $\overline{Y}$ is surjective if and only if $\overline{Y} \otimes \mathbb{F}_p$ is surjective for each prime p. This is clear since $- \otimes \mathbb{F}_p$ is right exact, so surjectivity of a map can be checked on closed fibers. □

Fix any prime p. Let $M_p = M \otimes \mathbb{F}_p$ and $\overline{Y}_p = \overline{Y} \otimes \mathbb{F}_p$. For any $\boldsymbol{e}_i$, if $L_i \notin p\mathbb{Z}$, then the image of $\boldsymbol{e}_i$ is zero in M_p. We can therefore use $\{\boldsymbol{e}_i \mid L_i \in p\mathbb{Z}\}$ as a basis of M_p. Consequently, M_p is a vector space over $\mathbb{F}_p$ of dimension at most n, and we can write $\overline{Y}_p$ as a matrix over $\mathbb{F}_p$. Next we will decompose the vector space M_p and the matrix $\overline{Y}_p$.

For each positive integer m, let $M_{p,m}$ denote the subspace of M_p generated by $\{\boldsymbol{e}_i \mid L_i \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}\}$. Let $\overline{Y}_{p,m}$ be the endomorphism

$$M_{p,m} \longrightarrow M_p \xrightarrow{\overline{Y}_p} M_p \longrightarrow M_{p,m},$$

where the first map is the inclusion and the last map is the natural projection using the given basis $\{\boldsymbol{e}_i \mid L_i \in p\mathbb{Z}\}$. Then $\overline{Y}_{p,m}$ can be expressed as the submatrix of $\overline{Y}$ taken from the rows and columns with indices i such that $\boldsymbol{e}_i \in M_{p,m}$. For all but finitely many values of m, we have $M_{p,m} = 0$, and in this case, $\overline{Y}_{p,m}$ is a 0×0 matrix. We adopt the convention that the determinant of a 0×0 matrix is 1. In general, $\det(\overline{Y}_{p,m})$ is in $\mathbb{F}_p$.

Lemma 4.5. *The following are equivalent.*

(1) *The map $\overline{Y}_p$ is an isomorphism.*

(2) *For all positive integers m, $\overline{Y}_{p,m}$ is an isomorphism.*

(3) $\det(\overline{Y}_{p,m}) \neq 0$ *for all positive integers m.*

Proof. It is clear that (2) and (3) are equivalent, so we need only show that (1) and (2) are equivalent.

Let $m > 0$, and let i, j be such that $L_i \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}$ and $L_j \notin p^m\mathbb{Z}$. Since $L_j = \operatorname{lcm}\{d_{kj} \mid k = 1, \ldots, n\}$, we have $d_{ij} \notin p^m\mathbb{Z}$ and $k_{ij}L_i/d_{ij} \in p\mathbb{Z}$. Therefore, the $\boldsymbol{e}_i$-component of $\overline{Y}_p\boldsymbol{e}_j$ is zero. We can extend this to show that, for any $m > m' > 0$, the $M_{p,m'}$-component of $\overline{Y}_p(M_{p,m})$ is zero, or equivalently,

$$\overline{Y}_p(M_{p,m}) \subseteq \bigoplus_{n\geq m} M_{p,n} =: N_m.$$

This implies that, for any $m > 0$, $\overline{Y}_p$ acts as an endomorphism on N_m. Since each M_p is finite dimensional, $\overline{Y}_p$ is an isomorphism if and only if it acts as an isomorphism on each subquotient $N_m/N_{m+1} \cong M_{p,m}$. This action is already given by $\overline{Y}_{p,m}$, so the assertion follows. $\square$

Combining all the lemmas in this section we have:

Theorem 4.6. *The center of the skew polynomial ring $T_{\boldsymbol{q}}[x_1, \ldots, x_n]$ is a polynomial ring if and only if $\det(\overline{Y}_{p,m}) \neq 0$ for all primes p and all integers $m > 0$.*

Theorem 4.6 is a slight generalization of Theorem 0.3(a) without the hypothesis that $q_{ij} \neq 1$ for all $i \neq j$. The definition of the matrices $\overline{Y}_{p,m}$ is not straightforward, so we give an example below. Hopefully, the example will show that this matrix is not hard to understand.

Example 4.7. We start with the following skew-symmetric matrix with entries in $\mathbb{Q}$:

$$Y := \begin{pmatrix} 0 & \frac{4}{27} & \frac{2}{9} & 0 & \frac{2}{3} & \frac{3}{5} \\ -\frac{4}{27} & 0 & \frac{1}{3} & \frac{7}{9} & \frac{1}{3} & \frac{1}{5} \\ -\frac{2}{9} & -\frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{7}{9} & -\frac{1}{6} & 0 & \frac{2}{3} & 0 \\ -\frac{2}{3} & -\frac{1}{3} & -\frac{1}{2} & -\frac{2}{3} & 0 & \frac{5}{8} \\ -\frac{3}{5} & -\frac{1}{5} & -\frac{1}{2} & 0 & -\frac{5}{8} & 0 \end{pmatrix}.$$

One can easily construct q_{ij} by (E4.0.1) and the skew polynomial ring $T_{\boldsymbol{q}}[x_1, \ldots, x_6]$ by (E0.2.1), but the point of this example is to work out the matrices $\overline{Y}_{p,m}$ for all primes p and all $m > 0$. By considering the denominators of the entries of Y, one sees that

$$(L_1, L_2, L_3, L_4, L_5, L_6) = (3^3 \cdot 5, 3^3 \cdot 5, 2 \cdot 3^2, 2 \cdot 3^2, 2^3 \cdot 3, 2^3 \cdot 5).$$

This implies that $\overline{Y}_{p,m}$ is a trivial matrix (or a 0×0 matrix) except for $p = 2, 3, 5$. Next we consider

$$\overline{Y} = \operatorname{diag}(L_1, \ldots, L_6)Y = \begin{pmatrix} 0 & 20 & 30 & 0 & 90 & 81 \\ -20 & 0 & 45 & 105 & 45 & 27 \\ -4 & -6 & 0 & 3 & 9 & 9 \\ 0 & -14 & -3 & 0 & 12 & 0 \\ -16 & -8 & -12 & -16 & 0 & 15 \\ -24 & -8 & -20 & 0 & -25 & 0 \end{pmatrix}.$$

Recall that $M_{p,m}$ has a basis $\{\boldsymbol{e}_i \mid L_i \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}\}$ and $\overline{Y}_{p,m}$ is the square submatrix of $\overline{Y}$ with indices $\{i \mid L_i \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}\}$ and with entries evaluated in $\mathbb{F}_p$.

For $p = 2$, $\overline{Y}_{2,m}$ are the following:

- $\overline{Y}_{2,1}$ is the principle $(3, 4)$-submatrix of Y, and is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.
- $\overline{Y}_{2,3}$ uses indices 5, 6, and is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.
- For all $m = 2$ or $m > 3$, $\overline{Y}_{2,m}$ is trivial.

Therefore, $\overline{Y}_2$ is an isomorphism by Lemma 4.5.

For $p = 3$, $\overline{Y}_{3,m}$ are the following:

- $\overline{Y}_{3,1}$ uses only index 5, and is the 1×1 zero matrix.
- $\overline{Y}_{3,2}$ uses indices 3, 4, and is the 2×2 zero matrix.
- $\overline{Y}_{3,3}$ uses indices 1, 2, and is $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.
- For all $m > 3$, $\overline{Y}_{3,m}$ is trivial.

Since $\det(\overline{Y}_{3,1}) = \det(\overline{Y}_{3,2}) = 0$, $\overline{Y}_3$ is not an isomorphism by Lemma 4.5. Consequently, the center of $T_{\boldsymbol{q}}[x_1, \ldots, x_6]$ is not a polynomial ring by Theorem 4.6.

For $p = 5$, $\overline{Y}_{5,m}$ are the following:

- $\overline{Y}_{5,1}$ uses indices 1, 2, 6, and
$$\overline{Y}_{5,1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \\ -1 & -2 & 0 \end{pmatrix}.$$
- For all $m > 1$, $\overline{Y}_{5,m}$ is trivial.

It is easy to check that $\det(\overline{Y}_{5,1}) = 0$. Therefore, $\overline{Y}_5$ is not an isomorphism.

For $p > 5$, $\overline{Y}_{p,m}$ is trivial for all $m > 0$.

5. Low dimensional cases

We start with some easy consequences of Theorem 4.6 and then discuss the case when n is 3 or 4.

Corollary 5.1. *Suppose there are a prime p and an $m > 0$ such that $M_{p,m}$ is odd dimensional. Then $\overline{Y}_p$ is not an isomorphism. As a consequence, the center of $T_{\boldsymbol{q}}[x_1, \ldots, x_n]$ is not a polynomial ring.*

Proof. If $\overline{Y}_{p,m}$ is a skew-symmetric matrix of odd size, its determinant is zero (this is true even when $p = 2$). The rest follows from Lemma 4.5 and Theorem 4.6. □

Corollary 5.2. *Suppose there is a prime p such that M_p is odd dimensional. Then $\overline{Y}_p$ is not an isomorphism. As a consequence, the center of $T_{\boldsymbol{q}}[x_1, \ldots, x_n]$ is not a polynomial ring.*

Proof. Since $M_p = \bigoplus_{m=1}^{\infty} M_{p,m}$, if it is odd dimensional, at least one $M_{p,m}$ must be odd dimensional. The assertion follows from Corollary 5.1. □

Corollary 5.3. *Suppose, for each prime p, that $p \mid d_{ij}$ for at most one pair (i, j), $1 \le i < j \le n$. Then $\overline{Y}_p$ is an isomorphism for each p. As a consequence, the center of $T_{\boldsymbol{q}}[x_1, \dots, x_n]$ is a polynomial ring.*

Proof. If $d_{ij} \notin p\mathbb{Z}$ for all i, j, then $L_i \notin p\mathbb{Z}$ for all i, $M_p = 0$ and $\overline{Y}_p$ is trivially an isomorphism.

If $d_{ij} \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}$ for some i, j and some positive integer m, and each of every other term $d_{k\ell}$ is not in $p\mathbb{Z}$, then $L_i, L_j \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}$, and each of every other L_k is not in $p\mathbb{Z}$. This shows that $\overline{Y}_{p,m}$ is a nonzero 2×2 skew-symmetric matrix (i.e., $\det(\overline{Y}_{p,m}) \neq 0$) and $M_{p,m'} = 0$ for each $m' \neq m$. The rest follows from Lemma 4.5 and Theorem 4.6. □

Next we give simple criteria for $\overline{Y}$ to be an isomorphism in the cases $n = 3, 4$.

Corollary 5.4. *The center of $T_{\boldsymbol{q}}[x_1, x_2, x_3]$ is a polynomial ring if and only if $(d_{ij}, d_{ik}) = 1$ for all different i, j, k.*

Proof. There are only three d terms — d_{12}, d_{13}, and d_{23}. If each (d_{ij}, d_{ik}) equals 1, then no prime is a factor of more than one term in $\{d_{ij}\}$. By Corollary 5.3, the center of $T_{\boldsymbol{q}}[x_1, x_2, x_3]$ is a polynomial ring.

Conversely, suppose that p is a prime such that $d_{ij}, d_{ik} \in p\mathbb{Z}$ for some i, j, k. Then $L_1, L_2, L_3 \in p\mathbb{Z}$. This implies that M_p has dimension 3. Hence, by Corollary 5.2, $\overline{Y}_p$ is not an isomorphism. So $\overline{Y}$ is not an isomorphism. Therefore, the center of $T_{\boldsymbol{q}}[x_1, x_2, x_3]$ is not a polynomial ring by Lemma 4.3. □

Corollary 5.5. *The center of $T_{\boldsymbol{q}}[x_1, x_2, x_3, x_4]$ is a polynomial ring if and only if, for each prime p, one of the following holds:*

(a) *$L_i \notin p\mathbb{Z}$ for all i.*

(b) *For some positive integer m, $\overline{Y}_{p,m}$ is 4×4 with nonzero determinant.*

(c) *There are distinct indices $i, j, k, \ell \in \{1, 2, 3, 4\}$ and a nonnegative integer m such that $d_{ij} \in p^{m+1}\mathbb{Z}$, $d_{k\ell} \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}$, and every other d term is not in $p^{m+1}\mathbb{Z}$.*

Proof. Let $P = T_{\boldsymbol{q}}[x_1, x_2, x_3, x_4]$. By Lemmas 4.3 and 4.4, $Z(P)$ is a polynomial ring if and only if $\overline{Y}_p$ is an isomorphism for all p. It remains to show that, for each p, $\overline{Y}_p$ is an isomorphism if and only if one of (a), (b), or (c) holds. Now we fix p and prove the assertion in three cases according to the shape of M_p.

First we prove the "if" part.

(a) If $L_i \notin p\mathbb{Z}$ for all i, then $M_p = 0$ and $\overline{Y}_p$ is trivially an isomorphism. This handles the case when $M_p = 0$.

(b) If for some $m > 0$, $\overline{Y}_{p,m}$ is 4×4 with nonzero determinant, then every other $\overline{Y}_{p,r}$ (for all $r \neq m$) is a 0×0 matrix and, consequently, $\overline{Y}_p$ is an isomorphism. This is the case when $M_p = M_{p,m}$ is 4-dimensional for one m.

(c) Assume the hypotheses in part (c). Let $m' > m$ be the integer such that $d_{ij} \in p^{m'}\mathbb{Z} - p^{m'+1}\mathbb{Z}$. If $m = 0$, then d_{ij} is the only d term divisible by p. Hence $\overline{Y}_{p,m'}$ is a skew-symmetric 2×2 nonzero matrix and $\overline{Y}_{p,r}$ is trivial for all $r \neq m'$. Therefore, $\overline{Y}_p$ is an isomorphism. If $m > 0$, then $\overline{Y}_{p,m}$ and $\overline{Y}_{p,m'}$ are both skew-symmetric and 2×2, and (because $k_{k\ell}L_k/d_{k\ell} \notin p\mathbb{Z}$) nonzero. Furthermore, every other $\overline{Y}_{p,r}$ is 0×0 for all $r \neq m, m'$. Therefore, $\overline{Y}_p$ is an isomorphism.

For the rest we prove the "only if" part.

Suppose that $\overline{Y}_p$ is an isomorphism. By Corollary 5.2, M_p is even dimensional, that is, $\dim M_p = 0, 2$ or 4.

The $\dim M_p = 0$ case coincides with the case when $L_i \notin p\mathbb{Z}$ for all i, so we obtain case (a).

For the $\dim M_p = 2$ case, at least one d_{ij} lies in $p\mathbb{Z}$ and L_i, L_j lie in $p\mathbb{Z}$, and no other d term is a multiple of p, so $\overline{Y}_p$ is necessarily an isomorphism. We can set $m = 0$, so that $d_{ij} \in p^{m+1}\mathbb{Z}$, and all other d_{ab} are not in $p^{m+1}\mathbb{Z}$. So we obtain (c).

All that remains is the $\dim M_p = 4$ case. Each $M_{p,m}$ is even dimensional by Corollary 5.1. If $\dim M_{p,m} = 4$ for some m, then $\overline{Y}_{p,m}$ is 4×4 and $\overline{Y}_p$ is an isomorphism if and only if $\det(\overline{Y}_{p,m}) \neq 0$. So we obtain case (b).

Finally, suppose there exist $m' > m > 0$ such that $\dim M_{p,m} = \dim M_{p,m'} = 2$. Let i, j, k, ℓ be distinct such that $L_i, L_j \in p^{m'}\mathbb{Z} - p^{m'+1}\mathbb{Z}$ and $L_k, L_\ell \in p^m\mathbb{Z} - p^{m+1}\mathbb{Z}$. We must have that $d_{ij} \in p^{m'}\mathbb{Z} \subseteq p^{m+1}\mathbb{Z}$ and every other d term is not in $p^{m+1}\mathbb{Z}$. If $d_{k\ell} \notin p^m\mathbb{Z}$, then $k_{k\ell}L_k/d_{k\ell}, k_{\ell k}L_\ell/d_{\ell k} \in p\mathbb{Z}$ and $\overline{Y}_{p,m}$ is the 2×2 zero matrix, yielding a contradiction. Therefore, $d_{k\ell}$ must be in $p^m\mathbb{Z}$. So we obtain case (c) again. □

6. Center of generalized Weyl algebras

Let T be a commutative k-domain. In this section we assume that $\boldsymbol{q} := \{q_{ij}\}$ is a set of roots of unity in T and let $\mathcal{A} := \{a_{ij} \mid 1 \le i < j \le j\}$ be a subset of T. Define the generalized Weyl algebra associated to $(\boldsymbol{q}, \mathcal{A})$ to be the central T-algebra

$$V(\boldsymbol{q}, \mathcal{A}) := \frac{T\langle x_1, \dots, x_n\rangle}{(x_j x_i - q_{ij} x_i x_j - a_{ij} \mid i \neq j)}.$$

Consider a filtration on $V(\boldsymbol{q}, \mathcal{A})$ with $\deg x_i = 1$ and $\det t = 0$ for all $t \in T$. Suppose

$$\operatorname{gr} V(\boldsymbol{q}, \mathcal{A}) \text{ is naturally isomorphic to } T_{\boldsymbol{q}}[x_1, \dots, x_n]. \tag{E6.0.1}$$

Consider the hypothesis that,

$$\text{for any pair } (i, j),\ a_{ij} = 0 \text{ whenever } q_{ij} = 1. \tag{E6.0.2}$$

Proposition 6.1. *Suppose* (E6.0.1) *and* (E6.0.2) *and let* $A = V(\boldsymbol{q}, \mathcal{A})$. *If the center* $Z(\operatorname{gr} A)$ *is a polynomial ring, then so is* $Z(A)$, *and* $Z(A) \cong Z(\operatorname{gr} A)$.

Proof. If $Z(\operatorname{gr} A)$ is a polynomial ring, then $Z(\operatorname{gr} A) = T\big[x_1^{L_1}, \dots, x_n^{L_n}\big]$, where $L_i = \operatorname{lcm}\{d_{ij} \mid j = 1, \dots, n\}$ (Lemma 4.3). Recall that d_{ij} is the order of q_{ij}.

First we claim that $x_i^{L_i}$ is in the center of A. For each j, we have the equation $x_j x_i = q_{ij} x_i x_j + a_{ij}$. If $q_{ij} = 1$, then x_j commutes with x_i by hypothesis (E6.0.2), so x_j commutes with $x_i^{L_i}$. If $q_{ij} \neq 1$, then the order of q_{ij} is d_{ij}. The equation $x_j x_i = q_{ij} x_i x_j + a_{ij}$ implies that x_j commutes with $x_i^{d_{ij}}$, as each $x_j x_i^k$ is equal to $q_{ij}^k x_i^k x_j + \big(1 + q_{ij} + \dots + q_{ij}^{k-1}\big) a_{ij}$. Since d_{ij} divides L_i, x_j commutes with $x_i^{L_i}$ for all $j \neq i$. This shows that $x_i^{L_i}$ is central.

Since $\operatorname{gr} A$ is the skew polynomial ring $T_{\boldsymbol{q}}[x_1, \dots, x_n]$, it is easy to check that $\operatorname{gr} Z(A) \subset Z(\operatorname{gr} A)$. Since $Z(\operatorname{gr} A)$ is generated by $\big\{x_i^{L_i}\big\}_{i=1}^n$, induction on the degree of element $f \in Z(A)$ shows that f is generated by $x_i^{L_i}$. Therefore, the assertion follows. □

Proposition 6.2. *Retain the above notation and suppose* (E6.0.1). *If* $a_{ij} \neq 0$ *for some* $i \neq j$, *then* $q_{ik} q_{jk} = 1$ *for all* $k \neq i$ *or* j.

Proof. We resolve $x_k x_j x_i$ in two different ways:

$$\begin{aligned}
(x_k x_j) x_i &= (q_{jk} x_j x_k + a_{jk}) x_i \\
&= q_{jk} x_j (x_k x_i) + a_{jk} x_i \\
&= q_{jk} x_j (q_{ik} x_i x_k + a_{ik}) + a_{jk} x_i \\
&= q_{jk} q_{ik} (x_j x_i) x_k + q_{jk} a_{ik} x_j + a_{jk} x_i \\
&= q_{jk} q_{ik} (q_{ij} x_i x_j + a_{ij}) x_k + q_{jk} a_{ik} x_j + a_{jk} x_i \\
&= q_{jk} q_{ik} q_{ij} x_i x_j x_k + q_{jk} q_{ik} a_{ij} x_k + q_{jk} a_{ik} x_j + a_{jk} x_i,
\end{aligned}$$

and similarly

$$\begin{aligned}
x_k (x_j x_i) &= x_k (q_{ij} x_i x_j + a_{ij}) \\
&= q_{ij} (x_k x_i) x_j + a_{ij} x_k \\
&= q_{ij} (q_{ik} x_i x_k + a_{ik}) x_j + a_{ij} x_k \\
&= q_{ij} q_{ik} x_i (x_k x_j) + q_{ij} a_{ik} x_j + a_{ij} x_k \\
&= q_{ij} q_{ik} q_{jk} x_i x_j x_k + q_{ij} q_{ik} a_{jk} x_i + q_{ij} a_{ik} x_j + a_{ij} x_k.
\end{aligned}$$

Comparing the coefficients of x_k gives the result. □

When an algebra A is finitely generated and free over its center (as in the situation of Proposition 6.1), one should be able to compute the discriminant of A over its center. We give an example here.

Example 6.3. Let A be generated by x_1, x_2, x_3, x_4 and subject to the relations

$$\begin{aligned} x_3x_1 - x_1x_2 &= 0, & x_4x_2 + x_2x_4 &= 0,\\ x_3x_2 - x_2x_3 &= 0, & x_3x_4 + x_4x_3 &= 0,\\ x_4x_1 + x_1x_4 &= 0, & x_1x_2 + x_2x_1 &= x_3^2 + x_4^2. \end{aligned} \tag{E6.3.1}$$

This is the example in [Vancliff and Van Rompay 2000, Lemma 1.1] (with $\lambda = 0$). It is an iterated Ore extension, and therefore Artin–Schelter regular of global dimension 4.

It is not hard to check that the center of A is generated by x_i^2. This algebra is a factor ring of the algebra B over $T := k[t]$ generated by x_1, x_2, x_3, x_4 and subject to the relations

$$\begin{aligned} x_3x_1 - x_1x_2 &= 0, & x_4x_2 + x_2x_4 &= 0,\\ x_3x_2 - x_2x_3 &= 0, & x_3x_4 + x_4x_3 &= 0,\\ x_4x_1 + x_1x_4 &= 0, & x_1x_2 + x_2x_1 &= t. \end{aligned} \tag{E6.3.2}$$

Note that $\operatorname{gr} B$ is a skew polynomial ring over T with the above relations by setting $t = 0$. The Y-matrix is

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2}\\ -\frac{1}{2} & 0 & 0 & \frac{1}{2}\\ 0 & 0 & 0 & \frac{1}{2}\\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}.$$

By Corollary 5.5(b), B has center $T\big[x_1^2, x_2^2, x_3^2, x_4^2\big]$. The discriminant of B over its center is $2^{48}\big(4x_1^2x_2^2 - t^2\big)^8 x_3^{16}x_4^{16}$, by the next lemma. By Lemma 1.2, the discriminant of A over its center is $2^{48}\big(4x_1^2x_2^2 - \big(x_3^2+x_4^2\big)^2\big)^8 x_3^{16}x_4^{16}$. We will see in the next sections that $\mathbb{D}(A) = A$. As a consequence of Theorem 0.5, A is cancellative and the automorphism group of A is affine.

Lemma 6.4. *Suppose the $k[t]$-algebra B is generated by x_1, x_2, x_3, x_4 and subject to the six relations given (E6.3.2). Then the discriminant of B over its center is $2^{48}\big(4x_1^2x_2^2 - t^2\big)^8 x_3^{16}x_4^{16}$.*

Sketch of the proof. It is routine to check that the center of B is

$$Z(B) = k[t]\big[x_1^2, x_2^2, x_3^2, x_4^2\big].$$

The algebra B is a free module over $Z(B)$ of rank 16 with a $Z(B)$-basis $\big\{x_1^a x_2^b x_3^c x_4^d \mid a, b, c, d = 0, 1\big\}$. Let $\{z_1, \dots z_{16}\}$ be the above $Z(B)$-basis. Then we can compute the matrix $(\mathrm{tr}(z_i z_j))_{16\times 16}$:

$$\begin{pmatrix}
16 & 0 & 0 & 0 & 0 & 8t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 16a & 8t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 8t & 16b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 16c & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8ct & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 16d & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8dt & 0 & 0 & 0\\
8t & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 16ac & 0 & 8ct & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -16ad & 0 & -8dt & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 8ct & 0 & 16bc & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -8dt & 0 & -16bd & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -16cd & 0 & 0 & 0 & 0 & -8cdt\\
0 & 0 & 0 & 8ct & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 8dt & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 16acd & 8cdt & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8cdt & 16bcd & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8cdt & 0 & 0 & 0 & 0 & \delta
\end{pmatrix}$$

Here $\alpha=-16ab+8t^2$, $\beta=-16abc+8ct^2$, $\gamma=-16abd+8dt^2$, $\delta=16abcd-8cdt^2$, and $a=x_1^2$, $b=x_2^2$, $c=x_3^2$, $d=x_4^2$. We skip the details in computing the above traces. By using Maple, its determinant is $2^{48}\big(4x_1^2x_2^2-t^2\big)^8 x_3^{16}x_4^{16}$. $\square$

7. Three subalgebras

In this section we discuss three (possibly different) subalgebras of A, all of which are helpful for the applications in the next section.

Makar-Limanov invariants. The first subalgebra is the Makar-Limanov invariant of A [Makar-Limanov 1996]. This invariant has been very useful in commutative algebra. For any k-algebra A, let $\mathrm{Der}(A)$ denote the set of all k-derivations of A and let $\mathrm{LND}(A)$ denote the set of locally nilpotent k-derivations of A.

Definition 7.1. Let A be an algebra over k.

(1) The *Makar-Limanov invariant* of A is

$$\mathrm{ML}(A)=\bigcap_{\delta\in\mathrm{LND}(A)}\ker(\delta). \tag{E7.1.1}$$

(2) We say that A is LND-*rigid* if $\mathrm{ML}(A)=A$, or $\mathrm{LND}(A)=\{0\}$.

(3) We say that A is *strongly* LND-*rigid* if $\mathrm{ML}(A[t_1,\dots,t_d])=A$ for all $d\geq 0$.

The following lemma is clear. Part (2) follows from the fact that $\partial\in\mathrm{LND}(A)$ if and only if $g^{-1}\partial g\in\mathrm{LND}(A)$.

Lemma 7.2. *Let A be an algebra.*

(1) $\mathrm{ML}(A)$ *is a subalgebra of* A.

(2) *For any* $g \in \mathrm{Aut}(A)$, *we have* $g(\mathrm{ML}(A)) = \mathrm{ML}(A)$.

Divisor subalgebras. Throughout this subsection let A be a domain containing $\mathbb{Z}$. Let F be a subset of A. Let $\mathrm{Sw}(F)$ be the set of $g \in A$ such that $f = agb$ for some $a, b \in A$ and $0 \neq f \in F$. Here Sw stands for "subword", which can be viewed as a divisor.

Definition 7.3. Let F be a subset of A.

(1) Let $D_0(F) = F$. Inductively define $D_n(F)$ as the k-subalgebra of A generated by $\mathrm{Sw}(D_{n-1}(F))$. The subalgebra $D(F) = \bigcup_{n \geq 0} D_n(F)$ is called the *F-divisor subalgebra of A*. If F is the singleton $\{f\}$, we simply write $D(\{f\})$ as $D(f)$.

(2) If $f = d(A/Z)$ (if it exists), we call $D(f)$ *the discriminant-divisor subalgebra of A*, or *DDS of A*, and write it as $\mathbb{D}(A)$.

The following lemma is well-known [Makar-Limanov 2008, p. 4].

Lemma 7.4. *Let x, y be nonzero elements in A and let* $\partial \in \mathrm{LND}(A)$. *If* $\partial(xy) = 0$, *then* $\partial(x) = \partial(y) = 0$.

Proof. Let m and n be the largest integers such that $\partial^m(x) \neq 0$ and $\partial^n(y) \neq 0$. Then the product rule and the choice of m, n imply that

$$\partial^{m+n}(xy) = \sum_{i=0}^{m+n} \binom{n+m}{i} \partial^i(x) \partial^{m+n-i}(y) = \binom{n+m}{m} \partial^m(x) \partial^n(y) \neq 0.$$

So $m + n = 0$. The assertion follows. □

Lemma 7.5. *Let F be a subset of* $\mathrm{ML}(A)$. *Then* $D(F) \subseteq \mathrm{ML}(A)$.

Proof. Let ∂ be any element in $\mathrm{LND}(A)$. By hypothesis, $\partial(f) = 0$ for all $f \in F$. By Lemma 7.4, $\partial(x) = 0$ for all $x \in \mathrm{Sw}(F)$. So $\partial = 0$ when restricted to $D_1(F)$. By induction, $\partial = 0$ when restricted to $D(F)$. The assertion follows by taking arbitrary $\partial \in \mathrm{LND}(A)$. □

Lemma 7.6. *Suppose $d(A/Z)$ is defined. Then the DDS $\mathbb{D}(A)$ is preserved by all* $g \in \mathrm{Aut}(A)$.

Proof. By [CPWZ 2015a, Lemma 1.8(6)] or [CPWZ 2016, Lemma 1.4(4)], $d(A/Z)$ is g-invariant up to a unit. So, if $g \in \mathrm{Aut}(A)$, then g maps $\mathrm{Sw}(d(A/Z))$ to $\mathrm{Sw}(d(A/Z))$ and $D_1(d(A/Z))$ to $D_1(d(A/Z))$. By induction, one sees that g maps $D_n(d(A/Z))$ to $D_n(d(A/Z))$. So the assertion follows. □

We need to find some elements $f \in A$ so that $\partial(f) = 0$ for all $\partial \in \mathrm{LND}(A)$. The next lemma was proven in [CPWZ 2016, Proposition 1.5].

Lemma 7.7. *Let Z be the center of A and let $d \geq 0$. Suppose $A^\times = k^\times$. Assume that A is finitely generated and free over Z. Then we have $\partial(d(A/Z)) = 0$ for all $\partial \in \mathrm{LND}(A[t_1, \ldots, t_d])$.*

Proof. Let f denote the element $d(A[t_1, \ldots, t_d]/Z[t_1, \ldots, t_d])$ in $Z[t_1, \ldots, t_d]$. By [CPWZ 2016, Proposition 1.5], $\partial(f) = 0$. By [CPWZ 2015a, Lemma 5.4],

$$f =_{k^\times} d(A/Z).$$

The assertion follows. □

Here is the first relationship between the two subalgebras.

Proposition 7.8. *Retain the hypothesis of Lemma 7.7. Let $d \geq 0$. Then*

$$\mathbb{D}(A) \subseteq \mathrm{ML}(A[t_1, \ldots, t_d]) \subseteq A.$$

Proof. It is clear that $\mathrm{ML}(A[t_1, \ldots, t_d]) \subseteq A$ by [Bell and Zhang 2016]. Let f equal $d(A/Z)$, which is in $A \subseteq A[t_1, \ldots, t_d]$. By Lemma 7.7, $f \in \mathrm{ML}(A[t_1, \ldots, t_d])$. Let $D'(f)$ be the discriminant-divisor subalgebra of f in $A[t_1, \ldots, t_d]$. By Lemma 7.5, $D'(f) \subseteq \mathrm{ML}(A[t_1, \ldots, t_d])$. It is clear from the definition that $D(f) \subseteq D'(f)$. Therefore, the assertion follows. □

In particular, by taking $d = 0$, we have $\mathbb{D}(A) \subseteq \mathrm{ML}(A)$.

Aut-*bounded subalgebra.* In this subsection we assume that A is filtered such that the associated graded ring $\mathrm{gr}\, A$ is a connected graded domain. Later we further assume that A is connected graded. Since $\mathrm{gr}\, A$ is a connected graded domain, we can define $\deg f$ to be the degree of $\mathrm{gr}\, f$, and the degree satisfies the equation

$$\deg(xy) = \deg x + \deg y$$

for all $x, y \in A$.

Definition 7.9. Retain the above hypotheses. Let G be a subgroup of $\mathrm{Aut}(A)$ and let V be a subset of A.

(1) Let x be an element in A. The *G-bound* of x is

$$\deg_G(x) := \sup\{\deg(g(x)) \mid g \in G\}.$$

(2) Let g be in $\mathrm{Aut}(A)$. The *V-bound* of g is

$$\deg_g(V) := \sup\{\deg(g(x)) \mid x \in V\}.$$

(3) The *G-bounded subalgebra* of A, denoted by $\beta_G(A)$, is the set of elements x in A with finite G-bound. It is clear that $\beta_G(A)$ is a subalgebra of A (Lemma 7.10(1)). In particular, the Aut-*bounded subalgebra* of A, denoted by $\beta(A)$, is the set of elements x in A with finite $\mathrm{Aut}(A)$-bound.

The following lemma is easy, so we omit the proof.

Lemma 7.10. *Retain the above notation. Let G be a subgroup of* $\mathrm{Aut}(A)$.

(1) *The set $\beta_G(A)$ is a subalgebra of A.*

(2) *$g(\beta_G(A)) = \beta_G(A)$ for all $g \in G$.*

Here is the relation between the two subalgebras $\mathbb{D}(A)$ and $\beta(A)$. Let V be a subset of A. We say V is of *bounded degree* if there is an N such that $\deg(v) < N$ for all $v \in V$.

Proposition 7.11. *Let A be a filtered algebra such that* $\mathrm{gr}\, A$ *is a connected graded domain. Suppose that $G \subseteq \mathrm{Aut}(A)$ and $F \subseteq A$.*

(1) *If $G(F)$ has bounded degree, then $D(F) \subseteq \beta_G(A)$.*

(2) *If $f \in A$ is such that $g(f) =_{Z(A)^\times} f$ for all $g \in G$, then $D(f) \subseteq \beta_G(A)$.*

(3) *Assume that A is finitely generated and free over its center Z. Let $f = d(A/Z)$. Then $\mathbb{D}(A) = D(f) \subseteq \beta(A)$.*

Proof. (1) We have $D_0(F) = F \subseteq \beta_G(A)$ by assumption and use induction on n. Suppose that $D_{n-1}(F) \subseteq \beta_G(A)$. Assume that $D_n(F)$ is not contained in $\beta_G(A)$. Then there exists an $x \in D_n(A)$ such that $G(x)$ does not have bounded degree. Since $D_n(A)$ is generated by $\mathrm{Sw}(D_{n-1}(A))$ as an algebra, there is an $f \in \mathrm{Sw}(D_{n-1}(A))$ such that $G(f)$ does not have bounded degree. By definition of $\mathrm{Sw}(D_{n-1}(A))$, there exists a nonzero $f' \in D_{n-1}(A)$ and $a, b \in A$ such that $f' = afb$. Since $\mathrm{gr}\, A$ is a domain, we have $\deg(g(f')) = \deg(g(a)) + \deg(g(f)) + \deg(g(b))$ for all $g \in G$. Hence $G(f')$ does not have bounded degree, which is a contradiction. Hence $D_n(F) \subseteq \beta_G(A)$ for all $n \geq 1$. Therefore, $D(F) \subseteq \beta_G(A)$.

(2) Since $Z(A)^\times \subseteq A_0$, we see that $G(f)$ has bounded degree, hence part (2) follows from part (1).

(3) The third assertion is a special case of part (2) by Lemma 1.2. □

Under the hypotheses of Propositions 7.8 and 7.11 (and assuming that A is finitely generated and free over its center Z), we have:

$$\begin{tikzcd}[column sep=small] & \mathbb{D}(A) & \\ \mathrm{ML}(A) \arrow[ur, phantom, "\subseteq" sloped] & & \beta(A) \arrow[ul, phantom, "\supseteq" sloped] \\ & A \arrow[ul, phantom, "\supseteq" sloped] \arrow[ur, phantom, "\subseteq" sloped] & \end{tikzcd}$$

For the rest of this section, we assume that A is a connected graded domain and that k contains the field $\mathbb{Q}$. An automorphism g of A is called *unipotent* if

$$g(v) = v + (\text{higher degree terms}) \tag{E7.11.1}$$

for all homogeneous elements $v \in A$. Let $\mathrm{Aut}_{\mathrm{uni}}(A)$ denote the subgroup of $\mathrm{Aut}(A)$ consisting of unipotent automorphisms [CPWZ 2016, after Theorem 3.1]. If $g \in \mathrm{Aut}_{\mathrm{uni}}(A)$, we can define

$$\log g := -\sum_{i=1}^{\infty} \frac{1}{i}(1-g)^i. \tag{E7.11.2}$$

Let C be the completion of A with respect to the graded maximal ideal $\mathfrak{m} := A_{\geq 1}$. Then C is a local ring containing A as a subalgebra. We can define $\deg_l : C \to \mathbb{Z}$ by setting $\deg_l(v)$ to be the lowest degree of the nonzero homogeneous components of $v \in C$. We define *a unipotent automorphism of* C in a similar way to (E7.11.1) by using $\deg_l$. It is clear that if $g \in \mathrm{Aut}_{\mathrm{uni}}(A)$, then it induces a unipotent automorphism of C, which is still denoted by g.

Lemma 7.12. *Let A be a connected graded domain. Let $g \in \mathrm{Aut}_{\mathrm{uni}}(A)$ and let G be any subgroup of $\mathrm{Aut}(A)$ containing g. Let B denote $\beta_G(A)$. Then $(\log g)|_B$ is a locally nilpotent derivation of B. Further, $g|_B$ is the identity if and only if $(\log g)|_B$ is zero.*

Proof. Let C be the completion of A with respect to the graded maximal ideal $\mathfrak{m} := A_{\geq 1}$. Let g also denote the algebra automorphism of C induced by g. Then g is also a unipotent automorphism of C.

Since g is unipotent, $\deg_l(1-g)(v) > \deg_l v$ for any $0 \neq v \in C$. By induction, one has $\deg(1-g)^n(v) \geq n + \deg v$ for all $n \geq 1$. Thus $(\log g)(v)$ converges and therefore is well-defined. It follows from a standard argument that $\log g$ is a derivation of C (this is also a consequence of [Freudenburg 2006, Proposition 2.17(b)]).

Let v be an element in $B := \beta_G(A)$. Note that $g^n(v) \in B$ for all n by Lemma 7.10. Since $v \in B$, there is an N_0 such that $\deg g^n(v) < N_0$ for all n. If $(1-g)^n(v) \neq 0$, then

$$\deg(1-g)^n(v) = \deg\left(\sum_{i=0}^{n} \binom{n}{i} g^i(v)\right) < N_0 \quad \text{for all } n. \tag{E7.12.1}$$

When $n \geq N_0$, the inequalities from the previous paragraph imply that

$$\deg_l(1-g)^n(v) \geq n + \deg v \geq N_0, \tag{E7.12.2}$$

which contradicts (E7.12.1) unless $(1-g)^n(v) = 0$. Therefore,

$$(1-g)^n(v) = 0 \quad \text{for all } n > N_0. \tag{E7.12.3}$$

By (E7.12.3), the infinite sum of $\log g$ in (E7.11.2) terminates when applied to $v \in B$, and $(\log g)(v) \in A$. By Lemma 7.10, $(\log g)(v) \in B$. Since $\log g$ is a derivation of C, it is a derivation when restricted to B.

Next we need to show that it is a locally nilpotent derivation when restricted to B. It suffices to verify that, for any $v \in B$, $(\log g)^N(v) = 0$ for $N \gg 0$, which follows from (E7.11.2) and (E7.12.3).

The final assertion follows from the fact that g is the exponential function of $\log g$ and $\log g$ is locally nilpotent. □

Now we are ready to prove the second part of Theorem 0.5 without the finite GK-dimension hypothesis.

Theorem 7.13. *Let k be a field of characteristic zero and let A be a connected graded domain over k. Assume that A is finitely generated and free over its center Z in part* (2).

(1) *If* $\mathrm{ML}(A) = \beta(A) = A$, *then* $\mathrm{Aut}_{\mathrm{uni}}(A) = \{1\}$.

(2) *If* $\mathbb{D}(A) = A$, *then* $\mathrm{Aut}_{\mathrm{uni}}(A) = \{1\}$.

Proof. (1) By hypothesis, $B := \beta(A)$ equals A. Let $g \in \mathrm{Aut}_{\mathrm{uni}}(A)$. Then $(\log g)|_B$ is a locally nilpotent derivation of B by Lemma 7.12. Hence $\log g \in \mathrm{LND}(A)$. Since $\mathrm{ML}(A) = A$, we have $\mathrm{LND}(A) = \{0\}$. So $\log g = 0$. By Lemma 7.12, g is the identity.

(2) Combining the hypothesis $\mathbb{D}(A) = A$ with Propositions 7.8 and 7.11, we have $\mathrm{ML}(A) = \beta(A) = A$. The assertion follows from part (1). □

8. Applications

In this section we assume that k is a field of characteristic zero.

Zariski cancellation problem. The Zariski cancellation problem for noncommutative algebras was studied in [Bell and Zhang 2016]. We recall some definitions and results.

Definition 8.1. [Bell and Zhang 2016, Definition 1.1] Let A be an algebra.

(1) We call A *cancellative* if $A[t] \cong B[t]$ for some algebra B implies that $A \cong B$.

(2) We call A *strongly cancellative* if, for any $d \geq 1$, $A[t_1, \dots, t_d] \cong B[t_1, \dots, t_d]$ for some algebra B implies that $A \cong B$.

The original Zariski cancellation problem, or ZCP, asks if the polynomial ring $k[t_1, \dots, t_n]$, where k is a field, is cancellative. A recent result of Gupta [2014a; 2014b] settled the question negatively in positive characteristic for $n \geq 3$. The ZCP in characteristic zero remains open for $n \geq 3$. Some history and partial results can be found in [Bell and Zhang 2016], where the authors used discriminants and locally nilpotent derivations to study the ZCP for noncommutative rings.

One of their main results is the following.

Theorem 8.2 [Bell and Zhang 2016, Theorems 0.4 and 3.3]. *Let A be a finitely generated domain of finite Gelfand–Kirillov dimension. If A is strongly* LND*-rigid* (*respectively*, LND*-rigid*), *then A is strongly cancellative* (*respectively*, *cancellative*).

Now we have an immediate consequence, which is the first part of Theorem 0.5. Combining it with Theorem 7.13, we have finished the proof of Theorem 0.5.

Theorem 8.3. *Let A be a finitely generated domain of finite GK-dimension. Let Z be the center of A and suppose $A^\times = k^\times$. Assume that A is finitely generated and free over Z. If $A = \mathbb{D}(A)$, then A is strongly cancellative.*

Proof. Combining the hypothesis $A = \mathbb{D}(A)$ with Proposition 7.8, we have

$$A = \mathbb{D}(A) \subseteq \mathrm{ML}(A[t_1, \dots, t_d]) \subseteq A.$$

So $\mathrm{ML}(A[t_1, \dots, t_d]) = A$, or A is strongly LND-rigid. The assertion follows from Theorem 8.2. □

Next we give two examples.

Example 8.4. Let A be generated by x_1, x_2, x_3, x_4 and subject to the relations

$$\begin{aligned} x_1x_2 + x_2x_1 &= 0, & x_2x_3 + x_3x_2 &= 0, \\ x_1x_3 + x_3x_1 &= 0, & x_3x_4 + x_4x_3 &= 0, \\ x_1x_4 + x_4x_1 &= x_3^2, & x_2x_4 + x_4x_2 &= 0. \end{aligned}$$

This is an iterated Ore extension, so it is Artin–Schelter regular of global dimension 4. This is a special case of the algebra in [Vancliff et al. 1998, Definition 3.1]. Set $x_i^2 = y_i$ for $i = 1, \dots, 4$. Then $Z(A) = k[y_1, y_2, y_3, y_4]$. The M_1-matrix of (E3.0.1) is

$$(a_{ij})_{4\times 4} = \begin{pmatrix} 2y_1 & 0 & 0 & y_3 \\ 0 & 2y_2 & 0 & 0 \\ 0 & 0 & 2y_3 & 0 \\ y_3 & 0 & 0 & 2y_4 \end{pmatrix}.$$

The determinant $\det(a_{ij})$ is $f_0 := 4y_2y_3\big(4y_1y_4 - y_3^2\big)$. By Theorem 3.7, the discriminant $f := d(A/Z)$ is $f_0^{2^3}$. It is clear that $y_2, y_3 \in \mathrm{Sw}(f)$ and $y_1, y_4 \in \mathrm{Sw}(D_1(f))$. Thus $x_i \in \mathrm{Sw}(D_2(f))$ for all i. Consequently, $A = \mathbb{D}(A)$. By Theorem 8.3, A is strongly cancellative.

The next example is somewhat generic.

Example 8.5. Let T be a commutative domain and let $A = C(V, q)$ be the Clifford algebra associated to a quadratic form $q : V \to T$ where V is a free T-module of rank n. Suppose that n is even. Then the center of A is T [Lam 2005, Chapter 5, Theorem 2.5(a)]. We assume that A is a domain with $A^\times = k^\times$. Let $t_1, \dots, t_w$ be a set of generators of T, and suppose that $q(V) \subseteq (t_1 \cdots t_w)T$ or $\det(q) \in (t_1 \cdots t_w)T$.

Then by Theorem 3.7 we have $f := d(A/T) \in (t_1 \cdots t_w)^{2^{n-1}}$. So $t_s \in \mathrm{Sw}(f)$ for all s. This shows that $T \subseteq \mathbb{D}(A)$ and then $A = \mathbb{D}(A)$ (as $x_i^2 \in T$). By Theorem 8.3, A is strongly cancellative.

Remark 8.6. Let A be the algebra in Example 6.3. Using the formula for $d(A/Z)$ given in Lemma 6.4, it is easy to see that $A = \mathbb{D}(A)$. So A is cancellative by Theorem 8.3.

Automorphism problem. By [CPWZ 2015a; 2016], the discriminant controls the automorphism group of some noncommutative algebras. In this section we compute some automorphism groups by using the discriminants computed in previous sections. We first recall some definitions and results.

We modify the definitions in [CPWZ 2015a; 2016] slightly. Let A be an $\mathbb{N}$-filtered algebra such that $\operatorname{gr} A$ is a connected graded domain. Let $X := \{x_1, \ldots, x_n\}$ be a set of elements in A such that it generates A and $\operatorname{gr} X$ generates $\operatorname{gr} A$. We do not require $\deg x_i = 1$ for all i.

Definition 8.7. Let f be an element in A and let $X' = \{x_1, \ldots, x_m\}$ be a subset of X. We say f is *dominating* over X' if, for any subset $\{y_1, \ldots, y_n\} \subseteq A$ that is linearly independent in the quotient k-space A/k, there is a lift of f, say $F(X_1, \ldots, X_n)$, in the free algebra $k\langle X_1, \ldots, X_n\rangle$, such that $\deg F(y_1, \ldots, y_n) > \deg f$ whenever $\deg y_i > \deg x_i$ for some $x_i \in X'$.

The following lemma is easy.

Lemma 8.8. *Retain the above notation. Suppose $f := d(A/Z)$ is dominating over X'. Then for every automorphism $g \in \mathrm{Aut}(A)$, we have $\deg g(x_i) \le \deg x_i$ for all $x_i \in X'$.*

Proof. Let $y_i = g(x_i)$. Then $\{y_1, \ldots, y_n\}$ is linearly independent in A/k (as $\{x_1, \ldots, x_n\}$ is linearly independent in A/k). If $\deg y_i > \deg x_i$ for some i, by the dominating property, there is a lift of f in the free algebra, say $F(X_1, \ldots, X_n)$, such that $\deg F(y_1, \ldots, y_n) > \deg f$. Since g is an algebra automorphism,

$$F(y_1, \ldots, y_n) = F(g(x_1), \ldots, g(x_n)) = g(F(x_1, \ldots, x_n)) = g(f).$$

By [CPWZ 2015a, Lemma 1.8(6)], $g(f) = f$ (up to a unit in Z). Hence

$$\deg F(y_1, \ldots, y_n) = \deg g(f) = \deg f,$$

yielding a contradiction. Therefore, $\deg g(x_i) = \deg y_i \le \deg x_i$ for all i. □

We will study the automorphism group of a class of Clifford algebras; see Example 8.5.

Example 8.9. Let A be the Clifford algebra over a commutative k-domain T as in Example 8.5 and assume that n is even. Let $\{z_1, \ldots, z_n\}$ denote a set of generators for A. We will use $\{x_1, \dots, x_n\}$ for the generators of the generic Clifford algebra A_g defined in Section 3. Then there is an algebra homomorphism from A_g to A sending x_i to z_i for all i. Since n is even, T is the center of A. Assume that A is a filtered algebra such that $\operatorname{gr} A$ is a connected graded domain, so we can define the degree of any nonzero element in A. Further assume that $\deg t_i = 2$ (not 1) for all $i = 1, \dots, w$ and $\deg z_i > 2$ for all $i = 1, 2, \dots, n$. In particular, there is no element of degree 1. Some explicit examples are given later in this example.

Recall that we assumed $q(V) \subseteq (t_1 \cdots t_w)T$. Let $2b_{ij} = z_j z_i + z_i z_j$. Then we can write $b_{ij} = (t_1 \cdots t_w)^N b'_{ij}$ for some $N > 0$. By Theorem 3.7, the discriminant is $f := d(A/T) = \left[\left(\prod_{s=1}^{w} t_s\right)^N d'\right]^{2^{n-1}}$, where $d' = \det(2b'_{ij})_{n\times n}$. We need another hypothesis, which is that

$$\deg d' < N. \tag{E8.9.1}$$

Let $X' = \{t_i\}_{i=1}^{w}$ and $X = \{z_i\}_{i=1}^{n} \cup X'$. Then f is a noncommutative polynomial over X'. We first claim that f is dominating over X'. Let $\{y_i\}_{i=1}^{w}$ be a set of elements in $A \setminus k$. If $\deg y_i > 2$ for some i, then $\deg\left[\left(\prod_{s=1}^{w} y_s\right)^N d'(y_1, \dots, y_w)\right]^{2^{n-1}}$ is strictly larger than the degree of f, as we assume that $\deg d' < N$. This shows the claim.

Now let g be any algebra automorphism of A and let y_i be $g(t_i)$ for all i. Then, by Lemma 8.8, $\deg y_i = 2$. It follows from the relations $z_i z_i = b_{ii}$ that $\deg z_i > 3$. Hence $(\operatorname{gr} A)_2$ is generated by the t_i. This implies that y_i is in the span of X' and k. In some sense, every automorphism of A is affine (with respect to X'). It is a big step in understanding the automorphism group of A.

Below we study the automorphism group of a family of subalgebras of the generic Clifford algebra A_g of rank n that is defined in Section 3. As before, we assume n is even. We have two different sets of variables t, one for A_g and the other for general A. It would be convenient to unify these in the following discussion. So we identify $\{t_{(i,j)} \mid 1 \le i \le j \le n\}$ with $\{t_i\}_{i=1}^{w}$ via a bijection ϕ. Here $w = \frac{1}{2}n(n+1)$ as in the definition of A_g (Section 3).

Let r be any positive integer and let $B_{g,r}$ be the graded subalgebra of A_g generated by $\{t_{(i,j)}\}$ for all $1 \le i \le j \le n$ (or $\{t_i\}_{i=1}^{w}$) and $z_i := x_i\left(\prod_{k=1}^{w} t_k\right)^r$ for all $i = 1, 2, \dots, n$. Since $B_{g,r}$ is a graded subalgebra of A_g, it is a connected graded domain. This is also a Clifford algebra over $T_g := k[t_{(i,j)}]$ generated by $z_1, \dots, z_n$ and subject to the relations

$$z_j z_i + z_i z_j = 2\left(\prod_{k=1}^{w} t_k\right)^{2r} t_{(i,j)} =: 2b_{ij}$$

from which the bilinear form b and associated quadratic form q can easily be recovered. In particular, $q(V) \subseteq \left(\prod_{k=1}^{w} t_k\right)^{2r} T_g$, where $V = \bigoplus_{i=1}^{n} T_g z_i$. By the definition of A_g, we have $\deg t_i = 2$. Then $\deg z_i = 1 + 4rw > 3$. Now we assume that $N := 2r$ is bigger than $2n$, which is the degree of $d' := \det(t_{(i,j)})$. So we have

$$n < r, \quad \text{or equivalently} \quad \deg d' < N,$$

as required by (E8.9.1). See also Remark 8.10.

Let g be an algebra automorphism of $B_{g,d}$. By the above discussion, $g(t_i)$, for each i, is a linear combination of $\{t_j\}_{j=1}^{w}$ and 1. Using the relations $z_i^2 = b_{ii}$, we see that $\deg g(z_i) = \deg(z_i)$ for all i. Thus g must be a filtered automorphism of $B_{g,d}$.

Since g preserves the discriminant f and f is homogeneous in t_i, we have $\deg g(t_i) = 2$. Further, by using the expression of f and the fact that T_g is a UFD, $g(t_i)$ can not be a linear combination of the t_j of more than one term. Thus $g(t_i) = c_i t_j$ for some j and some $c_i \in k^\times$. This implies that there is a permutation $\sigma \in S_w$ and a collection of units $\{c_i\}_{i=1}^{w}$ such that $g(t_i) = c_i t_{\sigma(i)}$ for all i. Since g is filtered (by the last paragraph), $g(z_i) = \sum_{h=1}^{n} d_{ih} z_h + e_i$, where $d_{ih}, e_i \in k$. Applying g to the relation

$$z_i^2 = b_{ii} = \left(\prod_{i=1}^{w} t_i\right)^N t_{\phi(i,i)}, \quad \text{where } N := 2r,$$

we obtain that

$$\left(\sum_h d_{ih} z_h\right)^2 + 2e_i\left(\sum_h d_{ih} z_h\right) + e_i^2 = \left(\prod_{i=1}^{w} c_i t_i\right)^N g(t_{\phi(i,i)}).$$

Since $\left(\sum_h d_{ih} z_h\right)^2 \in T$, we have $e_i\left(\sum_h d_{ih} z_h\right) = 0$. Consequently, $e_i = 0$ and $g(z_i) = \sum_{h=1}^{n} d_{ih} z_h$. Applying g to the relations

$$z_i z_j + z_j z_i = 2b_{ij} = 2\left(\prod_{i=1}^{w} t_i\right)^N t_{\phi(i,j)}$$

and expanding the left-hand side, we obtain

$$\sum_{h,l} d_{ih} d_{jl}(z_h z_l + z_l z_h) = 2\left(\prod_{i=1}^{w} c_i t_i\right)^N g(t_{\phi(i,j)}).$$

Hence $d_{ih} d_{jl}$ is nonzero for only one pair (h, l). Thus there is a set of units $\{d_i\}_{i=1}^{n}$ and a permutation $\psi \in S_n$ such that $g(z_i) = d_i z_{\psi(i)}$ for all $i = 1, \ldots, n$. Then the above equation implies that

$$d_i d_j \left(\prod_{i=1}^{w} t_i\right)^N t_{\phi(\psi(i),\psi(j))} = \left(\prod_{i=1}^{w} c_i\right)^N \left(\prod_{i=1}^{w} t_i\right)^N c_{\phi(i,j)} t_{\sigma(\phi(i,j))}$$

for all i, j. Therefore,

$$\phi(\psi(i), \psi(j)) = \sigma(\phi(i, j)) \tag{E8.9.2}$$

and

$$d_i d_j = \left(\prod_{i=1}^{w} c_i\right)^{N} c_{\phi(i,j)} \tag{E8.9.3}$$

for all i, j.

By (E8.9.2), σ is completely determined by $\psi \in S_n$. Let $\bar{d}_i = d_i\left(\prod_{i=1}^{w} c_i\right)^{-r}$. Then (E8.9.3) says that $\bar{d}_i\bar{d}_j = c_{\phi(i,j)}$. So $\prod_{i=1}^{w} c_i = \prod_{1\le i\le j\le n} \bar{d}_i\bar{d}_j$. This means the $c_{\phi(i,j)}$ and d_i are completely determined by the $\bar{d}_i$. In conclusion,

$$\mathrm{Aut}(B_{g,r}) \cong \{\psi \in S_n\} \ltimes \left\{\bar{d}_i \in k^{\times} \mid i = 1, \ldots, n\right\} \cong S_n \ltimes (k^{\times})^n.$$

In particular, every algebra automorphism of $B_{g,r}$ is a graded algebra automorphism.

Remark 8.10. As a consequence of the computation in Example 8.9, $\mathrm{Aut}(B_{g,r})$ is independent of the parameter r when $r > n$. In fact, this assertion holds for all $r > 0$, but its proof requires a different and longer analysis, so it is omitted. On the other hand, $\mathrm{Aut}(B_{g,0}) = \mathrm{Aut}(A_g)$ is very different; see Remark 3.9(3).

We will work out one more automorphism group below.

Example 8.11. We continue to study Example 8.4 and prove that every algebra automorphism of A in Example 8.4 is graded. Some unimportant details are omitted due to the length.

Claim 1: $\mathfrak{m} := A_{\ge 1}$ is the only ideal of codimension 1 satisfying $\dim \mathfrak{m}/\mathfrak{m}^2 = 4$. Suppose $I = (x_1 - a_1, x_2 - a_2, x_3 - a_3, x_4 - a_4)$ is an ideal of A of codimension 1 such that $\dim_k I/I^2 = 4$. Then the map $\pi : x_i \to a_i$ for all i extends to an algebra homomorphism $A \to k$. Applying π to the relations of A in (E8.4.1), we obtain

$$a_1a_2 = 0, \quad a_1a_3 = 0, \quad 2a_1a_4 = a_3^2, \quad a_2a_3 = 0, \quad a_3a_4 = 0, \quad a_2a_4 = 0.$$

Therefore, (a_i) is either $(a_1, 0, 0, 0)$, or $(0, a_2, 0, 0)$, or $(0, 0, 0, a_4)$. By symmetry, we consider the first case and the details of the other cases are omitted. Let $z_i = x_i - a_i$ for all i. Then the first relation of (E8.4.1) becomes

$$z_1z_2 + z_2z_1 = (x_1 - a_1)x_2 + x_2(x_1 - a_1) = -2a_1x_2 = -2a_1z_2.$$

So $2a_1z_2 \in I^2$. Since $\dim I/I^2 = 4$, we have $a_1 = 0$. Thus we have proved Claim 1.

One of the consequences of Claim 1 is that any algebra automorphism of A preserves $\mathfrak{m}$. So we have a short exact sequence

$$1 \to \mathrm{Aut}_{\mathrm{uni}}(A) \to \mathrm{Aut}(A) \to \mathrm{Aut}_{\mathrm{gr}}(A) \to 1,$$

where $\mathrm{Aut}_{\mathrm{gr}}(A)$ is the group of graded algebra automorphisms of A and $\mathrm{Aut}_{\mathrm{uni}}(A)$ is the group of unipotent algebra automorphisms of A.

Claim 2: If f is a nonzero normal element in degree 1, then $B := A/(f)$ is an Artin–Schelter regular domain of global dimension 3. By [Rogalski and Zhang 2012, Lemma 1.1], B has global dimension 3. Since A satisfies the χ-condition [Artin and Zhang 1994], so does B. As a consequence, B is AS regular of global dimension 3 [Artin and Schelter 1987]. It is well-known that every Artin–Schelter regular algebra of global dimension 3 is a domain (following by the Artin–Schelter–Tate–Van den Bergh classification [Artin and Schelter 1987; Artin et al. 1991; 1990]).

Claim 3: If $f \in A_1$ is a normal element, then $f \in kx_2$ or $f \in kx_3$. First of all, both x_2 and x_3 are normal elements by the relations (E8.4.1). Note that $x_i g = \eta_{-1}(g)x_i$ for $i = 2, 3$, where η_{-1} is the algebra automorphism of A sending x_i to $-x_i$ for all i.

Suppose that f is nonzero normal and $f \notin kx_3 \cup kx_4$. Then the image $\bar{f}$ of f is normal in $A/(x_3)$. Since $A/(x_3)$ is a skew polynomial ring, by [Kirkman et al. 2010, Lemma 3.5(d)], $\bar{f}$ is a scalar multiple of x_i for some $i = 1, 2$, or 4. This implies that f is either $ax_1 + bx_3$, or $ax_2 + bx_3$, or $ax_4 + bx_3$ for some $a, b \in k$. If $b = 0$, then $f = x_1$ or x_4. The relation $x_1x_4 + x_4x_1 = x_3^2$ implies that $A/(f)$ is not a domain (as $x_3^2 = 0$ in $A/(f)$). This contradicts Claim 2. So the only possible case is $f = x_2$ (again yielding a contradiction). Now assume that $b \neq 0$ (and $a \neq 0$ because $f \notin kx_3 \cup kx_4$). We consider the first case and the details of the other cases are similar and omitted. Since $f = ax_1 + bx_3$, the relation $x_1x_3 + x_3x_1 = 0$ implies that $x_1^2 = 0$ in $A/(f)$, which contradicts Claim 2. In all these cases, we obtain a contradiction, and therefore $f \in kx_2$ or $f \in kx_3$.

Since $A/(x_2)$ is not isomorphic to $A/(x_3)$, there is no algebra automorphism sending x_2 to x_3. As a consequence, any graded automorphism ψ of A maps $x_2 \to c_2x_2$ and $x_3 \to c_3x_3$. Let g be any graded algebra automorphism of A. Let $\bar{g}$ be the induced algebra automorphism of $A/(x_3)$. By [Kirkman et al. 2010, Lemma 3.5(e)], $\bar{g}$ sends $x_1 \to c_1x_1$ and $x_4 \to c_4x_4$, or $x_1 \to c_1x_4$ and $x_4 \to c_4x_1$. Then, by using the original relations in (E8.4.1), one can check that g is of the form

$$x_1 \to c_1x_1, \quad x_2 \to c_2x_2, \quad x_3 \to c_3x_3, \quad x_4 \to c_4x_4,$$

where $c_1c_2 = c_3^2 = c_4^2$, or

$$x_1 \to c_1x_4, \quad x_2 \to c_2x_2, \quad x_3 \to c_3x_3, \quad x_4 \to c_4x_1,$$

where $c_1c_2 = c_3^2 = c_4^2$. So

$$\mathrm{Aut}_{\mathrm{gr}}(A) \cong \{(c_1, c_2, c_3, c_4) \in (k^\times)^4 \mid c_1c_2 = c_3^2 = c_4^2\},$$

which is completely determined.

Claim 4: $\mathrm{Aut}_{\mathrm{uni}}(A)$ is trivial. Recall that the discriminant of A over its center is

$$d := \big(x_2^2x_3^2(4x_1^2x_4^2 - x_3^4)\big)^8.$$

By Example 8.4, the DDS subalgebra $\mathbb{D}(A)$ is the whole algebra A. The assertion follows from Theorem 0.5.

Combining all these claims, one sees that $\mathrm{Aut}(A) = \mathrm{Aut}_{\mathrm{gr}}(A)$, which is described in Claim 3.

Remark 8.12. Ideas as in Remark 8.10 also apply to Example 6.3 and a similar conclusion holds. The interested reader can fill out the details.

Acknowledgements

The authors would like to thank the referees for their careful reading and valuable comments. A. A. Young was supported by the US National Science Foundation (NSF Postdoctoral Research Fellowship, No. DMS-1203744) and J. J. Zhang was supported by the US National Science Foundation (Nos. DMS-0855743 and DMS-1402863).

References

[Artin and Schelter 1987] M. Artin and W. F. Schelter, "Graded algebras of global dimension 3", *Adv. Math.* **66**:2 (1987), 171–216. MR 917738 Zbl 0633.16001

[Artin and Zhang 1994] M. Artin and J. J. Zhang, "Noncommutative projective schemes", *Adv. Math.* **109**:2 (1994), 228–287. MR 1304753 Zbl 0833.14002

[Artin et al. 1990] M. Artin, J. Tate, and M. Van den Bergh, "Some algebras associated to automorphisms of elliptic curves", pp. 33–85 in *The Grothendieck Festschrift*, vol. 1, edited by P. Cartier et al., Progress in Mathematics **86**, Birkhäuser, Boston, MA, 1990. MR 1086882 Zbl 0744.14024

[Artin et al. 1991] M. Artin, J. Tate, and M. Van den Bergh, "Modules over regular algebras of dimension 3", *Invent. Math.* **106**:2 (1991), 335–388. MR 1128218 Zbl 0763.14001

[Bell and Zhang 2016] J. Bell and J. J. Zhang, "Zariski cancellation problem for noncommutative algebras", preprint, 2016. arXiv 1601.04625

[CPWZ 2015a] S. Ceken, J. H. Palmieri, Y.-H. Wang, and J. J. Zhang, "The discriminant controls automorphism groups of noncommutative algebras", *Adv. Math.* **269** (2015), 551–584. MR 3281142 Zbl 06374153

[CPWZ 2015b] S. Ceken, J. H. Palmieri, Y.-H. Wang, and J. J. Zhang, "Invariant theory for quantum Weyl algebras under finite group action", preprint, 2015. arXiv 1501.07881

[CPWZ 2016] S. Ceken, J. H. Palmieri, Y.-H. Wang, and J. J. Zhang, "The discriminant criterion and automorphism groups of quantized algebras", *Adv. Math.* **286** (2016), 754–801. MR 3415697 Zbl 06506332

[Freudenburg 2006] G. Freudenburg, *Algebraic theory of locally nilpotent derivations*, Encyclopaedia of Mathematical Sciences **136**, Springer, Berlin, 2006. MR 2259515 Zbl 1121.13002

[Gupta 2014a] N. Gupta, "On the cancellation problem for the affine space $\mathbb{A}^3$ in characteristic p", *Invent. Math.* **195**:1 (2014), 279–288. MR 3148104 Zbl 1309.14050

[Gupta 2014b] N. Gupta, "On Zariski's cancellation problem in positive characteristic", *Adv. Math.* **264** (2014), 296–307. MR 3250286 Zbl 1325.14078

[Kirkman et al. 2010] E. Kirkman, J. Kuzmanovich, and J. J. Zhang, "Shephard–Todd–Chevalley theorem for skew polynomial rings", *Algebr. Represent. Theory* **13**:2 (2010), 127–158. MR 2601538 Zbl 1215.16032

[Lam 2005] T. Y. Lam, *Introduction to quadratic forms over fields*, Graduate Studies in Mathematics **67**, American Mathematical Society, Providence, RI, 2005. MR 2104929 Zbl 1068.11023

[Levasseur 1992] T. Levasseur, "Some properties of noncommutative regular graded rings", *Glasgow Math. J.* **34**:3 (1992), 277–300. MR 1181768 Zbl 0824.16032

[Lu et al. 2007] D.-M. Lu, J. H. Palmieri, Q.-S. Wu, and J. J. Zhang, "Regular algebras of dimension 4 and their A_∞-Ext-algebras", *Duke Math. J.* **137**:3 (2007), 537–584. MR 2309153 Zbl 1193.16014

[Makar-Limanov 1996] L. Makar-Limanov, "On the hypersurface $x + x^2y + z^2 + t^3 = 0$ in $\mathbb{C}^4$ or a $\mathbb{C}^3$-like threefold which is not $\mathbb{C}^3$", *Israel J. Math.* **96**:2 (1996), 419–429. MR 1433698 Zbl 0896.14021

[Makar-Limanov 2008] L. Makar-Limanov, "Locally nilpotent derivations, a new ring invariant and applications", preprint, 2008, available at http://www.math.wayne.edu/~lml/lmlnotes.pdf.

[Reiner 1975] I. Reiner, *Maximal orders*, London Mathematical Society Monographs **5**, Academic Press, London, 1975. Reprinted by Oxford University Press, 2003. MR 0393100 Zbl 0305.16001

[Rogalski and Zhang 2012] D. Rogalski and J. J. Zhang, "Regular algebras of dimension 4 with 3 generators", pp. 221–241 in *New trends in noncommutative algebra*, edited by P. Ara et al., Contemporary Mathematics **562**, American Mathematical Society, Providence, RI, 2012. MR 2905562 Zbl 1254.16020

[Vancliff and Van Rompay 2000] M. Vancliff and K. Van Rompay, "Four-dimensional regular algebras with point scheme, a nonsingular quadric in $\mathbf{P}^3$", *Comm. Algebra* **28**:5 (2000), 2211–2242. MR 1757458 Zbl 0961.16019

[Vancliff et al. 1998] M. Vancliff, K. Van Rompay, and L. Willaert, "Some quantum $\mathbf{P}^3$s with finitely many points", *Comm. Algebra* **26**:4 (1998), 1193–1208. MR 1612220 Zbl 0915.16035

Communicated by Efim Zelmanov
Received 2015-04-07 Revised 2016-02-07 Accepted 2016-03-10

kenhchan@math.washington.edu
Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350, United States

ayoung@digipen.edu
Department of Mathematics, DigiPen Institute of Technology, 9931 Willows Road NE, Redmond, WA 98052, United States

zhang@math.washington.edu
Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350, United States

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.597

Regularized theta lifts and (1,1)-currents on GSpin Shimura varieties

Luis E. Garcia

We introduce a regularized theta lift for reductive dual pairs $(\mathrm{Sp}_4, O(V))$, for V a quadratic vector space over a totally real number field. The lift takes values in the space of (1,1)-currents on the Shimura variety attached to $\mathrm{GSpin}(V)$; we show its values are cohomologous to currents given by integration on special divisors against automorphic Green functions. A later paper will show how to evaluate the new lift on differential forms obtained as usual (nonregularized) theta lifts.

1. Introduction

1A. ***Background and main results.*** The theory of the theta correspondence provides one of the most powerful tools to construct automorphic forms on classical groups. In recent years, the work of many authors has led to a geometric version of this theory describing the behavior of various spaces of so-called special cycles. Namely, the arithmetic quotients of symmetric spaces attached to classical groups $\mathrm{SO}(p,q)$ and $U(p,q)$ are equipped with a large collection of cycles coming from the subgroups that fix a given rational subspace; these are generally known as special cycles. After the work of Kudla and Millson [1986; 1987; 1990] constructing theta functions that represent their Poincaré dual forms, it has become clear that their cohomological properties are very closely connected with the theta correspondence; see, e.g., [Kudla 1997] for a description of their cup products and intersection numbers for the group $\mathrm{SO}(n,2)$.

In cases where these arithmetic quotients are naturally quasiprojective algebraic varieties (e.g., for the group $\mathrm{SO}(n,2)$ just mentioned), some of these special cycles

MSC2010: primary 11F27; secondary 11F67, 11G18, 14G35.
Keywords: Shimura varieties, theta series, Weil representation.

define complex subvarieties, and it is interesting to ask about more refined invariants, such as Green currents for them, or their image in the appropriate Chow groups. The work of Borcherds [1998; 1999] and its generalization by Bruinier [2002; 2012] successfully addressed these questions for the case of special divisors on arithmetic quotients of $\mathrm{SO}(n, 2)$. Their construction relies again on the theta correspondence and is based on considering theta lifts with respect to the reductive dual pair $(\mathrm{SL}_2, O(V))$. The automorphic forms on $\mathrm{SL}_2(\mathbb{A})$ used in their work as an input are not of moderate growth; thus, the integrals defining the theta lifts are not convergent and need to be regularized. With the proper regularization procedure, one can construct Green functions for special divisors, and also meromorphic automorphic forms, as theta lifts.

One might wonder if regularized theta lifts for reductive dual pairs of the form $(\mathrm{Sp}_{2n}, O(V))$ for $n \geq 2$ can be defined and whether one can construct interesting currents on arithmetic quotients of the symmetric space associated with $\mathrm{SO}(V_{\mathbb{R}})$ in this way. Consider such a quotient X_Γ associated with a lattice $\Gamma \subset \mathrm{SO}(n, 2)$, and let (Y, f) be a pair consisting of a subvariety $Y \subset X_\Gamma$ and a meromorphic function $f \in \mathbb{C}(Y)^\times$. In view of the explicit description of motivic cohomology and regulator maps in terms of higher Chow groups (see, e.g., [Goncharov 2005]), it is interesting to consider the current $\log|f| \cdot \delta_Y$, whose value on a differential form $\alpha \in \mathscr{A}_c^*(X_\Gamma)$ is given by

$$(\log|f| \cdot \delta_Y, \alpha) = \int_Y \log|f| \cdot \alpha. \tag{1-1}$$

The first goal of this paper is to show that, for many pairs (Y, f) such that Y is a special subvariety and f has divisor supported in special cycles, the current $\log|f| \cdot \delta_Y$ can be obtained as a regularized theta lift for $(\mathrm{Sp}_4, O(V))$. This follows from Theorem 1.1 below. For motivation, note that conjectures by Beilinson [1984] relate the values of dd^c-closed $\mathbb{Q}$-linear combinations of such currents with the values at certain integral points of L-functions attached to X_Γ. Our construction allows us to compute the values of some more general currents by using the theta correspondence; a followup paper will relate them to special values of standard L-functions of automorphic representations of Sp_4. Let us now describe more precisely the main objects involved in the statement of the theorem.

Let F be a totally real number field and V be a quadratic vector space over F. We assume that the signature of V is $((n, 2), (n+2, 0), \ldots, (n+2, 0))$ with n positive and even. Let $H = \mathrm{Res}_{F/\mathbb{Q}} \mathrm{GSpin}(V)$. Attached to H there is a Shimura variety X of dimension n whose complex points at a finite level determined by a neat open compact subgroup $K \subset H(\mathbb{A}_f)$ are given by

$$X_K = H(\mathbb{Q}) \backslash (\mathbb{D} \times H(\mathbb{A}_f))/K. \tag{1-2}$$

Here $\mathbb{D}$ denotes the hermitian symmetric space attached to the Lie group $\mathrm{SO}(V_{\mathbb{R}})$. For fixed K, the complex manifold X_K is a finite union of arithmetic quotients of the form $X_\Gamma := \Gamma\backslash\mathbb{D}^+$, where $\mathbb{D}^+$ denotes one of the connected components of $\mathbb{D}$. Consider two vectors $v, w \in V$ spanning a totally positive definite plane in V and write Γ_v (resp. $\Gamma_{v,w}$) for the stabilizer of v (resp. of both v and w) in Γ. One can define complex submanifolds $\mathbb{D}_v^+ \subset \mathbb{D}^+$ and $\mathbb{D}_{v,w}^+ \subset \mathbb{D}_v^+$, each of complex codimension one, and holomorphic maps

$$\begin{tikzcd}
\mathbb{D}^+_{v,w} \arrow[r] \arrow[d] & \mathbb{D}^+_v \arrow[r] \arrow[d] & \mathbb{D}^+ \arrow[d] \\
X(v,w)_\Gamma = \Gamma_{v,w}\backslash\mathbb{D}^+_{v,w} \arrow[r, "\iota"] & X(v)_\Gamma = \Gamma_v\backslash\mathbb{D}^+_v \arrow[r, "f"] & X_\Gamma
\end{tikzcd} \tag{1-3}$$

where the maps in the bottom row are proper and generically one-to-one. In Section 3B we recall the construction of a function

$$G(v,w)_\Gamma \in \mathscr{C}^\infty\big(X(v)_\Gamma - \iota(X(v,w)_\Gamma)\big)$$

that is a Green function for the divisor $[\iota(X(v,w)_\Gamma)] \in \mathrm{Div}(X(v)_\Gamma)$; this function is locally integrable and hence defines a current $[G(v,w)_\Gamma] \in \mathscr{D}^0(X(v)_\Gamma)$. Define the current

$$[\Phi(v,w)_\Gamma] = 2\pi i \cdot f_*([G(v,w)_\Gamma]) \in \mathscr{D}^{1,1}(X_\Gamma), \tag{1-4}$$

where $f_* : \mathscr{D}^0(X(v)_\Gamma) \to \mathscr{D}^{1,1}(X_\Gamma)$ denotes the pushforward map. Note that the $\mathbb{Q}$-linear span of the currents $[\Phi(v,w)_\Gamma]$ for varying w and fixed v includes all the currents of the form $2\pi i \cdot \log|f| \cdot \delta_{X(v)_\Gamma}$, where $f \in \mathbb{C}(X(v)_\Gamma)^\times \otimes_{\mathbb{Z}} \mathbb{Q}$ is one of the meromorphic functions constructed by Bruinier [2012, Theorem 6.8]. Given a totally positive definite symmetric matrix $T \in \mathrm{Sym}_2(F)$ and a Schwartz function $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ fixed by K, in Section 3G we define a current $[\Phi(T,\varphi)_K] \in \mathscr{D}^{1,1}(X_K)$ as a finite sum of currents $[\Phi(v,w)_\Gamma]$ weighted by the values of φ. As an example, consider the case treated in Section 4B, where $X_K = X_0^B \times X_0^B$ is a self-product of a full level Shimura curve X_0^B attached to an indefinite quaternion algebra B over $\mathbb{Q}$. Here the currents $[\Phi(T,\varphi)_K]$ admit a description in terms of Hecke correspondences and CM points on X_K. Namely, if p is a prime not dividing the discriminant of B such that $p \equiv 1 \pmod 4$, and writing $L = \mathbb{Q}[\sqrt{-p}]$, then for a certain choice of $\varphi = \varphi_0$ we have

$$\left[\Phi\left(\begin{pmatrix} 1 & \\ & p \end{pmatrix}, \varphi_0\right)_K\right] = 2\pi i \cdot (X_0^B \xrightarrow{\Delta} X_0^B \times X_0^B)_*([G_{t_{L/\mathbb{Q}}[\mathrm{CM}(\mathscr{O}_L)]}]). \tag{1-5}$$

where Δ denotes the diagonal embedding and $G_{t_{L/\mathbb{Q}}[\mathrm{CM}(\mathscr{O}_L)]}$ denotes a Green function for the divisor $t_{L/\mathbb{Q}}[\mathrm{CM}(\mathscr{O}_L)]$ of points in X_0^B with CM by $\mathscr{O}_L$ (see (4-28)).

Our first main result will show that the currents $[\Phi(T,\varphi)_K]$ are cohomologous to some currents obtained by a process of regularized theta lifting. Let us now introduce these theta lifts. In Section 3H we define, for $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ fixed by K and $g \in \mathrm{Sp}_4(\mathbb{A}_F)$, a theta function $\theta(g;\varphi)_K$ valued in the space of smooth $(1,1)$-forms on X_K. In the same section, we introduce a function

$$\mathscr{M}_T(s) : N(F)\backslash N(\mathbb{A}) \times A(\mathbb{R})^0 \longrightarrow \mathbb{C}. \tag{1-6}$$

Here T denotes a totally positive definite symmetric 2-by-2 matrix, s is a complex number, $N \subset \mathrm{Sp}_{4,F}$ denotes the unipotent radical of the Siegel parabolic of $\mathrm{Sp}_{4,F}$ and $A(\mathbb{R})^0$ denotes the connected component of the identity of the real points of the subgroup $A \subset \mathrm{Sp}_{4,F}$ of diagonal matrices in $\mathrm{Sp}_{4,F}$. This function grows exponentially along $A(\mathbb{R})^0$. We define the regularized theta lift

$$\big(\mathscr{M}_T(s), \theta(\cdot,\varphi)_K\big)^{\mathrm{reg}} = \int_{A(\mathbb{R})^0}\int_{N(F)\backslash N(\mathbb{A})} \mathscr{M}_T(na,s)\theta(na,\varphi)_K \, dn\, da, \tag{1-7}$$

with appropriate measures dn and da.

Theorem 1.1. (1) *The regularized integral* $(\mathscr{M}_T(s), \theta(\cdot;\varphi)_K)^{\mathrm{reg}}$ *converges for* $\mathrm{Re}(s) \gg 0$ *on an open dense set of* X_K *whose complement has measure zero and defines a locally integrable* $(1,1)$*-form* $\Phi(T,\varphi,s)_K$ *on* X_K.

(2) *Let* $\tilde{\mathscr{D}}^{1,1}(X_K) = \mathscr{D}^{1,1}(X_K)/(\mathrm{Im}(\partial)+\mathrm{Im}(\bar\partial))$. *The current* $[\Phi(T,\varphi,s)_K] \in \tilde{\mathscr{D}}^{1,1}(X_K)$ *defined by* $\Phi(T,\varphi,s)_K$ *admits meromorphic continuation to* $s \in \mathbb{C}$; *moreover*, *its constant term at* $s = s_0 = (n-1)/2$ *satisfies*

$$\mathrm{CT}_{s=s_0}[\Phi(T,\varphi,s)_K] = [\Phi(T,\varphi)_K]$$

as elements of $\tilde{\mathscr{D}}^{1,1}(X_K)$.

In fact, Proposition 3.19 shows that the currents in the theorem are compatible under the maps $\mathscr{D}^{1,1}(X_{K'}) \to \mathscr{D}^{1,1}(X_K)$ induced from inclusions $K' \subset K$ of open compact subgroups, so that we obtain currents

$$[\Phi(T,\varphi)] = ([\Phi(T,\varphi)_K])_K \in \mathscr{D}^{1,1}(X) := \varprojlim_K \mathscr{D}^{1,1}(X_K) \tag{1-8}$$

and similarly $[\Phi(T,\varphi,s)] \in \mathscr{D}^{1,1}(X)$ that agree on closed differential forms.

A particularly interesting subspace of $\mathscr{D}^{1,1}(X_K)$ is the image of the regulator map

$$r_{\mathscr{D}} : \mathrm{CH}^2(X_K,1) \to \mathscr{D}^{1,1}(X_K) \tag{1-9}$$

whose definition we recall in Section 3I; in particular, we would like to characterize the currents $[\Phi_K]$ in the $\mathbb{Q}$-linear span of the currents $[\Phi(T,\varphi)_K]$ that belong to the image of $r_{\mathscr{D}}$. We will prove in Proposition 3.23 that, when $\dim X_K \geq 4$, we have for such a current Φ_K,

$$[\Phi_K] \in r_{\mathscr{D}} \Longleftrightarrow dd^c[\Phi_K] = 0. \tag{1-10}$$

Let us assume from now on that V is anisotropic over F; this implies that X_K is compact. Once the currents $[\Phi(T, \varphi)]$ have been constructed, we would like to evaluate them on differential forms $\alpha \in \mathscr{A}_c^{n-1,n-1}(X_K)$. Since the form $\Phi(T, \varphi, s)_K$ is obtained as a (regularized) integral, it is natural to try to do so by interchanging the integrals. However, the regularized integral is not absolutely convergent, and the exchange is not justified. To get around this problem, we introduce some locally integrable $(1,1)$-forms $\tilde{\Phi}(T, \varphi, s)_K$ related to the $\Phi(T, \varphi, s)_K$ in Theorem 1.1. They are also obtained as regularized theta lifts and the associated currents $[\tilde{\Phi}(T, \varphi, s)_K]$ are compatible under the maps induced by inclusions $K' \subset K$, thus defining a current $[\tilde{\Phi}(T, \varphi, s)] \in \mathscr{D}^{1,1}(X)$. As before, these currents enjoy a property of meromorphic continuation to $s \in \mathbb{C}$, and their constant terms satisfy

$$\mathrm{CT}_{s=s_0}[\tilde{\Phi}(T_1, \varphi_1, s)] - [\tilde{\Phi}(T_2, \varphi_2, s)] \equiv [\Phi(T_1, \varphi_1)] - [\Phi(T_2, \varphi_2)] \tag{1-11}$$

modulo $\mathrm{Im}(\partial) + \mathrm{Im}(\bar{\partial})$ for pairs (T_1, φ_1), (T_2, φ_2) related by a certain involution ι (see (3-82)). Here, at a finite level K, the current on the right hand side is a finite sum of currents of the form $[\Phi(v, w)_\Gamma] - [\Phi(w, v)_\Gamma]$, with $[\Phi(v, w)_\Gamma]$ given by (1-4); see Remark 3.24 for some motivation on these currents. Moreover, using ideas of Bruinier and Funke [2004], we show that the values $[\tilde{\Phi}(T, \varphi, s)_K](\alpha)$ for large $\mathrm{Re}(s)$ can be computed by reversing the order of integration; the precise statement is the following.

Proposition 3.27. *Let $K \subset H(\mathbb{A}_f)$ be an open compact subgroup that fixes φ and let $\alpha \in \mathscr{A}_c^{n-1,n-1}(X_K)$. Then, for* $\mathrm{Re}(s) \gg 0$, *we have*

$$([\tilde{\Phi}(T, \varphi, s)_K], \alpha) = \int_{A(\mathbb{R})^0} \int_{N(F)\backslash N(\mathbb{A})} \widetilde{\mathscr{M}}_T(na, s) \int_{X_K} \theta(na; \varphi \otimes \tilde{\varphi}_\infty)_K \wedge \alpha \, dn \, da.$$

This result also gives information on the values of the currents $[\Phi(T, \varphi)]$; see Corollary 3.28.

1B. *Outline of the paper.* We now describe the contents of each section in more detail. Section 2 is a review of definitions and basic facts about Shimura varieties X attached to GSpin groups. In it we recall the definition of the relevant Shimura datum, describe the connected components of X_K at a finite level K and introduce the tautological line bundle $\mathscr{L}$ and its canonical metric. Then we recall the definition of special cycles in X_K and their weighted versions introduced by Kudla.

In Section 3 we construct currents in $\mathscr{D}^{1,1}(X_K)$. Sections 3A and 3B first review previous work by Oda, Tsuzuki and Bruinier on secondary spherical functions on the symmetric space $\mathbb{D}$ attached to $\mathrm{SO}(n,2)$, and on automorphic Green functions for special divisors on arithmetic quotients $\Gamma\backslash\mathbb{D}^+$ (here $\mathbb{D}^+$ denotes one of the connected components of $\mathbb{D}$). In Section 3C we introduce some differential forms with singularities on $\mathbb{D}$. These forms depend on a complex parameter s and are used in Section 3D to define (1,1)-forms on $\Gamma\backslash\mathbb{D}^+$ with singularities on special divisors. We prove that these (1,1)-forms are locally integrable and therefore define currents in $\mathscr{D}^{1,1}(\Gamma\backslash\mathbb{D}^+)$. Section 3E then shows that these currents admit meromorphic continuation to $s \in \mathbb{C}$ and that their regularized value at a certain value s_0 is cohomologous to the pushforward of the automorphic Green function in Section 3B defined on a certain special divisor. An adelic formulation of the above constructions is provided in Section 3F. After this, in Section 3G, we introduce weighted currents; their behavior under pullbacks induced by inclusions of open compact subgroups $K' \subset K$ and under the Hecke algebra of the GSpin group is described. Section 3H explains how these weighted currents can be constructed as regularized theta lifts for the dual pair $(\mathrm{Sp}_4, O(V))$. In Section 3I we give a necessary and sufficient condition for the currents above to belong to the image of the regulator map from the higher Chow group $\mathrm{CH}^2(X_K, 1)$. Section 3J introduces some related currents on X_K and uses their presentation as regularized theta lifts to prove that they can be evaluated on differential forms by interchanging the order of integration.

The example of a product of Shimura curves described above is considered in Section 4. This section starts with some definitions and basic facts on Shimura curves in Section 4A. In Section 4B, we describe several of the currents introduced in Section 3 in terms of Hecke correspondences and CM divisors.

1C. *Notation.* The following conventions will be used throughout the paper.

- We write $\hat{\mathbb{Z}} = \varprojlim_n (\mathbb{Z}/n\mathbb{Z})$ and $\hat{M} = M \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}$ for any abelian group M. We write $\mathbb{A}_f = \mathbb{Q} \otimes_{\mathbb{Z}} \hat{\mathbb{Z}}$ for the finite adeles of $\mathbb{Q}$ and $\mathbb{A} = \mathbb{A}_f \times \mathbb{R}$ for the full ring of adeles.
- For a number field F, we write $\mathbb{A}_F = F \otimes_{\mathbb{Q}} \mathbb{A}$, $\mathbb{A}_{F,f} = F \otimes_{\mathbb{Q}} \mathbb{A}_f$ and $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$. We will suppress F from the notation if no ambiguity can arise.
- For a finite set of places S of F, we will denote by $\mathbb{A}_S$ and $\mathbb{A}^S$ the subset of adeles in $\mathbb{A}_F$ supported on S and away from S, respectively.
- We denote by $\psi_{\mathbb{Q}} = \bigotimes_v \psi_{\mathbb{Q}_v} : \mathbb{Q}\backslash\mathbb{A}_{\mathbb{Q}} \to \mathbb{C}^\times$ the standard additive character of $\mathbb{A}_{\mathbb{Q}}$, defined by

$$\begin{aligned} \psi_{\mathbb{Q}_p}(x) &= e^{-2\pi i x} \qquad \text{for } x \in \mathbb{Z}[p^{-1}], \\ \psi_{\mathbb{R}}(x) &= e^{2\pi i x} \qquad \text{for } x \in \mathbb{R}. \end{aligned}$$

If F_v is a finite extension of $\mathbb{Q}_v$, we set $\psi_v = \psi_{\mathbb{Q}_v}(\mathrm{tr}(x))$, where $\mathrm{tr} : F_v \to \mathbb{Q}_v$ is the trace map. For a number field F, we write $\psi = \bigotimes_v \psi_v : F\backslash\mathbb{A}_F \to \mathbb{C}^\times$ for the resulting additive character of $\mathbb{A}_F$.

- For a locally compact, totally disconnected topological space X, the symbol $\mathscr{S}(X)$ denotes the Schwartz space of locally constant, compactly supported functions on X. For X a finite dimensional vector space over $\mathbb{R}$, the symbol $\mathscr{S}(X)$ denotes the Schwartz space of all $\mathscr{C}^\infty$ functions on X all whose derivatives are rapidly decreasing.
- For a ring R, we denote by $\mathrm{Mat}_n(R)$ the set of all n-by-n matrices with entries in R. The symbols 1_n and 0_n denote the identity and zero matrices in $\mathrm{Mat}_n(R)$.
- The transpose of a matrix $x \in \mathrm{Mat}_n(R)$, is denoted ${}^t x$, and the set of all symmetric matrices in $\mathrm{Mat}_n(R)$ is $\mathrm{Sym}_n(R) = \{x \in \mathrm{Mat}_n(R) \mid x = {}^t x\}$.
- $X \amalg Y$ denotes the disjoint union of X and Y.
- If an object $\phi(s)$ depends on a complex parameter s and is meromorphic in s, we denote by $\mathrm{CT}_{s=s_0}\,\phi(s)$ the constant term of its Laurent expansion at $s = s_0$.

2. Shimura varieties and special cycles

2A. ***Shimura varieties.*** We recall the facts about orthogonal Shimura varieties that we will need. We follow [Kudla 1997] closely, to which the reader is referred for further details. Let F be a totally real number field of degree d with embeddings $\sigma_i : F \to \mathbb{R}$, $i = 1, \ldots, d$. Let (V, Q) a quadratic vector space over F of dimension $n+2$ (with $n \geq 1$); we assume that $V_1 = V \otimes_{F,\sigma_1} \mathbb{R}$ has signature $(n, 2)$ and that $V_{\sigma_i} = V \otimes_{F,\sigma_i} \mathbb{R}$ is positive definite for $i = 2, \ldots, d$.

Let $H = \mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{GSpin}(V)$. The group H fits into a short exact sequence

$$1 \longrightarrow \mathrm{Res}_{F/\mathbb{Q}}\,\mathbb{G}_m \longrightarrow H \longrightarrow \mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{SO}(V) \longrightarrow 1. \tag{2-1}$$

Denote by $\mathbb{D}$ the set of oriented negative definite planes in V_1. We will fix once and for all a point $z_0 \in \mathbb{D}$ and will denote by $\mathbb{D}^+$ the connected component of $\mathbb{D}$ containing z_0. The group $\mathrm{SO}(V_1) \cong \mathrm{SO}(n, 2)$ acts transitively on $\mathbb{D}$, and the stabilizer K_{z_0} of z_0 is isomorphic to $\mathrm{SO}(n) \times \mathrm{SO}(2)$. We have

$$\mathbb{D} \cong \mathrm{SO}(n, 2)/(\mathrm{SO}(n) \times \mathrm{SO}(2)). \tag{2-2}$$

To the pair $(H, \mathbb{D})$ one can attach a Shimura variety $\mathrm{Sh}(H, \mathbb{D})$ that has a canonical model over $\sigma_1(F)$. Namely, in [Kudla 1997, p. 44] a homomorphism

$$h_0 : \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\,\mathbb{G}_m = \mathbb{C}^\times \longrightarrow H(\mathbb{R}) = \prod_{i=1,\ldots,d} \mathrm{GSpin}(V_{\sigma_i}) \tag{2-3}$$

is defined such that $\mathbb{D}$ becomes identified with the space of conjugates of h_0 by $H(\mathbb{R})$; the resulting action of $H(\mathbb{R})$ on $\mathbb{D}$ factors through the projection $H(\mathbb{R}) \to \mathrm{SO}(V_1)$. For any compact open subgroup $K \subset H(\mathbb{A}_f)$, we have

$$X_K = \mathrm{Sh}(H, \mathbb{D})_K(\mathbb{C}) = H(\mathbb{Q})\backslash(\mathbb{D} \times H(\mathbb{A}_f))/K. \tag{2-4}$$

Thus X_K is the complex analytification of a quasiprojective variety $\mathrm{Sh}(H, \mathbb{D})_K$ of dimension n defined over $\sigma_1(F)$. If V is anisotropic over F, then $\mathrm{Sh}(G, \mathbb{D})_K$ is actually projective.

We recall the description of the connected components of X_K. Let $H^{\mathrm{der}} \cong \mathrm{Res}_{F/\mathbb{Q}} \mathrm{Spin}(V)$ be the derived subgroup of H. There is an exact sequence

$$1 \longrightarrow H^{\mathrm{der}} \longrightarrow H \xrightarrow{\nu} T \longrightarrow 1, \tag{2-5}$$

where $T = \mathrm{Res}_{F/\mathbb{Q}} \mathbb{G}_m$ and ν is given by the spinor norm. Let $T(\mathbb{R})^+ = (\mathbb{R}_{>0})^d \subset T(\mathbb{R})$ and $H_+(\mathbb{R}) = \nu^{-1}(T(\mathbb{R})^+)$ be the set of elements of $H(\mathbb{R})$ of totally positive spinor norm; this is the subgroup of $H(\mathbb{R})$ stabilizing $\mathbb{D}^+$. Define

$$H_+(\mathbb{Q}) = H(\mathbb{Q}) \cap H_+(\mathbb{R}). \tag{2-6}$$

By the strong approximation theorem, we can find $h_1 = 1, \dots, h_r \in H(\mathbb{A}_f)$ such that

$$H(\mathbb{A}_f) = \coprod_{j=1}^{r} H_+(\mathbb{Q}) h_j K. \tag{2-7}$$

For $j = 1, \dots, r$, let $\Gamma_{h_j} = H_+(\mathbb{Q}) \cap h_j K h_j^{-1}$. Then

$$X_K \cong \coprod_{j=1}^{r} \Gamma_{h_j} \backslash \mathbb{D}^+. \tag{2-8}$$

We will also need to consider Shimura varieties attached to (V, Q) as above with $n = 0$. In this case, the symmetric domain associated with $\mathrm{SO}(V_1)$ consists of just one point, while $\mathbb{D} = \mathbb{D}^+ \amalg \mathbb{D}^-$ consists of two points (corresponding to two different orientations of the same negative definite plane z_0). Since it turns out to be more convenient for our purposes, we define X_K as in (2-4) and $\mathrm{Sh}(H, \mathbb{D})_K$ to be the union of two copies of the usual Shimura variety attached to H, so that with these notations we have $X_K = \mathrm{Sh}(H, \mathbb{D})_K(\mathbb{C})$.

For $n \geq 1$, we can introduce a different model for $\mathbb{D}$ that makes the presence of an $\mathrm{SO}(V_1)$-invariant complex structure obvious. Let $\mathscr{Q}$ be the quadric in $\mathbb{P}(V_1(\mathbb{C}))$ given by

$$\mathscr{Q} = \{v \in \mathbb{P}(V_1(\mathbb{C})) \mid (v, v) = 0\}. \tag{2-9}$$

Note that if $\{v_1, v_2\}$ is an orthogonal basis of $z \in \mathbb{D}$ with $(v_1, v_1) = (v_2, v_2) = -1$, then $v := v_1 - i v_2 \in V_1 \otimes \mathbb{C}$ satisfies $(v, v) = 0$ and $(v, \bar{v}) < 0$. Moreover, the line

$[v] := \mathbb{C} \cdot v$ is independent of the orthogonal basis we have chosen. Thus we obtain a well defined map $\mathbb{D} \to \mathscr{Q}$ and one checks that it gives an isomorphism

$$\mathbb{D} \longrightarrow \mathscr{Q}_- = \{w \in \mathbb{P}(V_{\sigma_1}(\mathbb{C})) \mid (w, w) = 0,\ (w, \bar{w}) < 0\} \tag{2-10}$$

onto the open subset $\mathscr{Q}_-$ of the quadric $\mathscr{Q}$.

Consider the tautological line bundle $\mathscr{L}$ over $\mathscr{Q}_-$ defined by

$$\mathscr{L} \setminus \{0\} := \{w \in V_1(\mathbb{C}) \mid (w, w) = 0,\ (w, \bar{w}) < 0\}. \tag{2-11}$$

The action of $H(\mathbb{R})$ on $\mathbb{D}$ lifts naturally to $\mathscr{L}$ and gives it the structure of a $H(\mathbb{R})$-equivariant bundle. Any element $v \in V_1$ defines a section s_v of $\mathscr{L}^\vee$ by the rule $s_v(w) = (v, w)$. We will only consider s_v for v of positive norm. The section s_v defines an analytic divisor

$$\operatorname{div}(s_v) = \{w \in \mathbb{P}(V_1(\mathbb{C})) \mid (v, w) = 0\}. \tag{2-12}$$

Under the isomorphism $\mathbb{D} \cong \mathscr{Q}_-$ described above, $\operatorname{div} s_v$ corresponds to $\mathbb{D}_v \subset \mathbb{D}$, where $\mathbb{D}_v$ denotes the set of negative definite planes in V_1 that are orthogonal to v.

The line bundle $\mathscr{L}$ carries a natural hermitian metric $\|\cdot\|$ defined by $\|w\|^2 = |(w, \bar{w})|$; this metric is $H(\mathbb{R})$-equivariant. We say that a function $f \in \mathscr{C}^\infty(\mathbb{D} - \mathbb{D}_v)$ has a logarithmic singularity along $\mathbb{D}_v$ if $f(z) - \log \|s_v(z)\|^2$ extends to $\mathscr{C}^\infty(\mathbb{D})$.

2B. ***Special cycles.*** Let $U \subset V$ be a totally positive definite subspace and let W be its orthogonal complement in V. Denote by H_U the pointwise stabilizer of U in H. Then $H_U \cong \operatorname{Res}_{F/\mathbb{Q}} \operatorname{GSpin}(W)$; its associated symmetric domain can be identified with $\mathbb{D}_U \cap \mathbb{D}^+$, where $\mathbb{D}_U$ denotes the subset of $\mathbb{D}$ consisting of planes z that are orthogonal to U. For a compact open $K \subset H(\mathbb{A}_f)$ and $h \in H(\mathbb{A}_f)$, let $K_{U,h} = H_U(\mathbb{A}_f) \cap hKh^{-1}$, an open compact subset of $H_U(\mathbb{A}_f)$. Define

$$X(U, h)_K = H_U(\mathbb{Q}) \backslash (\mathbb{D}_U \times H_U(\mathbb{A}_f))/K_{U,h}. \tag{2-13}$$

If $h = 1$, we write $X(U)_K := X(U, 1)_K$. Thus $X(U, h)_K$ is the set of complex points of a variety $\operatorname{Sh}(H_U, \mathbb{D}_U)_{K_{U,h}}$ defined over $\sigma_1(F)$. There is a morphism

$$i_U : \operatorname{Sh}(H_U, \mathbb{D}_U) \longrightarrow \operatorname{Sh}(H, \mathbb{D}) \tag{2-14}$$

defined over $\sigma_1(F)$; on complex points it induces a map

$$i_{U,h,K} : X(U, h)_K \longrightarrow X_K \tag{2-15}$$

that is proper and birational onto its image. Denote by $Z(U, h)_K$ the associated effective cycle on X_K. For a set of vectors $x = (x_1, \ldots, x_r) \in V^r$ spanning a totally positive definite vector space U of dimension r, we will write $Z(x, h)_K$ for $Z(U, h)_K$.

For a description of the connected components of these special cycles, see [Kudla 1997, Sections 3 and 4]; the main result is that these cycles have a finite number of components of the form $Z(U,h)_\Gamma$ that we now define. For $h \in H(\mathbb{A}_f)$, let $\Gamma_h = H_+(\mathbb{Q}) \cap hKh^{-1}$. Define $\Gamma_{U,h} = \Gamma_h \cap H_U(\mathbb{R})$ and consider the map

$$X(U,h)_\Gamma := \Gamma_{U,h}\backslash\mathbb{D}_U^+ \longrightarrow \Gamma_h\backslash\mathbb{D}^+ = X_{\Gamma_h}. \tag{2-16}$$

(For $h=1$, we will just write $X(U)_\Gamma$ for $X(U,1)_\Gamma$). The image defines a connected cycle in X_{Γ_h} that we denote by $Z(U,h)_\Gamma$.

In [Kudla 1997], certain weighted sums of these cycles are defined. Namely, let $r = \dim_F U$ and denote by $\mathrm{Sym}_r(F)_{>0}$ the space of totally positive definite r-by-r matrices with coefficients in F. For $T \in \mathrm{Sym}_r(F)_{>0}$ and $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^r)^K$ with values in a ring R, define

$$Z(T,\varphi)_K = \sum_{h \in H_U(\mathbb{A}_f)\backslash H(\mathbb{A}_f)/K} \varphi(h^{-1}x)Z(x,h)_K, \tag{2-17}$$

where $x = (x_1,\ldots,x_r) \in V^r$ is any vector with $\frac{1}{2}(x_i,x_j) = T$ (if no such x exists, we set $Z(T,\varphi)=0$). Note that the sum is finite and hence defines a cycle in $Z^r(X_K)\otimes_{\mathbb{Z}} R$.

3. Currents and regularized theta lifts

In this section we introduce some differential forms and currents on arithmetic quotients of $\mathbb{D}^+$. Some of these forms will be defined as Poincaré series by summation of Γ-translates of a differential form on $\mathbb{D}^+$. Here and throughout this paper, $\Gamma \subset H_+(\mathbb{R})$ denotes a group of the form $\Gamma = H_+(\mathbb{Q}) \cap K$, where $K \subset H(\mathbb{A}_f)$ is some neat open compact subgroup. If $U \subset V$ is a totally positive definite subspace, we will write $\Gamma_U = \Gamma \cap H_U(\mathbb{R})$, where H_U denotes the pointwise stabilizer of U in H. If U is spanned by vectors $v_1,\ldots,v_r$, we will sometimes write $\Gamma_{v_1,\ldots,v_r}$ for Γ_U.

Several currents defined in this Section will be described explicitly in Section 4B, where we consider the particular case when X_K is a product of Shimura curves. The description given there is in terms of Hecke correspondences and CM points, and the reader is advised to study the examples given there to understand the definitions and properties to follow.

3A. ***Secondary spherical functions on*** $\mathbb{D}$. Recall that $\mathbb{D}$ denotes the set of oriented, negative definite 2-planes in $V_1 = V \otimes_{F,\sigma_1} \mathbb{R}$. For every vector $v \in V_1$ of positive norm we have defined an analytic divisor $\mathbb{D}_v \subset \mathbb{D}$ consisting of those $z \in \mathbb{D}$ that are orthogonal to v. Denote by $H_v(\mathbb{R})$ the stabilizer of v in $H(\mathbb{R})$. Then we have $\mathbb{D}_v \cong H_v(\mathbb{R})/(K \cap H_v(\mathbb{R}))$, so that $\mathbb{D}_v$ can be identified with the hermitian symmetric space associated with $H_v(\mathbb{R})$. We write $\mathbb{D}_v^+ := \mathbb{D}_v \cap \mathbb{D}^+$.

We recall some of the main results of Oda and Tsuzuki [2003] concerning the existence and main properties of secondary spherical functions on $\mathbb{D}$. To state these results, we need to introduce certain subgroups of $G = \mathrm{SO}(V_1)$. Let $\{v_1, \ldots, v_{n+2}\}$ be a basis of V_1 whose quadratic form is $I_{n,2}$ and such that $v = v_1$. Let $z_0 = \langle v_{n+1}, v_{n+2}\rangle$ and denote by K_{z_0} the stabilizer of z_0 in $\mathrm{SO}(V_1)^+$. Let $W \subset V_1$ be the plane generated by v_1 and v_{n+1} and let $A = \mathrm{SO}(W)^0$ be the identity component of its orthogonal group. Then $A = \{a_t \mid t \in \mathbb{R}\}$ where $a_t v_1 = \cosh(t)v_1 + \sinh(t)v_{n+1}$. Let

$$A^+ = \{a_t \mid t \geq 0\} \tag{3-1}$$

and G_v be the stabilizer of v in G. Then there is a double coset decomposition

$$G = G_v A^+ K_{z_0}. \tag{3-2}$$

Proposition 3.1 [Oda and Tsuzuki 2003, Proposition 2.4.2]. *Let $\Delta_{\mathbb{D}}$ be the invariant Laplacian on $\mathbb{D}$ and let $\rho_0 = n/2$. Let s be a complex number with* $\mathrm{Re}(s) > \rho_0$. *There exists a unique function $\phi^{(2)}(v, z, s) \in \mathscr{C}^\infty(\mathbb{D} - \mathbb{D}_v)$ with the following properties*:

(1) $\Delta_{\mathbb{D}}\phi^{(2)}(v, z, s) = (s^2 - \rho_0^2)\phi^{(2)}(v, z, s)$.

(2) $\phi^{(2)}(v, gz, s) = \phi^{(2)}(v, z, s)$ *for every* $g \in G_v$.

(3) *Consider the function $\phi^{(2)}(v, g, s) = \phi^{(2)}(v, gz_0, s)$ for $g \in G$. It belongs to $\mathscr{C}^\infty(G - G_v K_{z_0})$ and satisfies $\phi^{(2)}(v, g'gk, s) = \phi^{(2)}(v, g, s)$ for every $g' \in G_v$, $k \in K_{z_0}$. Writing $G = G_v A^+ K_{z_0}$ as above, we have*

$$\phi^{(2)}(v, a_t, s) = \log(t) + O(1) \qquad \text{as } t \to 0,$$
$$\phi^{(2)}(v, a_t, s) = O(e^{-(\mathrm{Re}(s)+\rho_0)t}) \qquad \text{as } t \to +\infty.$$

It follows that $\phi^{(2)}(hv, hz, s) = \phi^{(2)}(v, z, s)$ for all $h \in H(\mathbb{R})$ and $z \in \mathbb{D}$. For a totally positive vector $v \in V(F)$, we will simply write $\phi^{(2)}(v, z, s)$ for $\phi^{(2)}(v_1, z, s)$, where v_1 denotes the image of v in V_1. We will sometimes write $\phi^{(2)}_{\mathbb{D}}(v, z, s)$ for $\phi^{(2)}(v, z, s)$ if we need to be precise about the domain of definition.

The function $\phi^{(2)}(v, z, s)$ admits an explicit description in terms of the Gaussian hypergeometric function. Namely, for $|z| < 1$, let $F(a, b, c, z)$ be the function given by

$$F(a, b, c, z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!},$$

where we write $(a)_0 = 1$ and $(a)_n = \Gamma(a + n)/\Gamma(a)$ for $n \geq 1$. For a vector $v \in V_1$ and a plane $z \in \mathbb{D}$, denote by $v_{z^\perp}$ the projection of v to the orthogonal complement

$z^\perp$ of z in V_1. Then [Oda and Tsuzuki 2003, (2.5.3)]:

$$\phi^{(2)}(v,z,s) = \\ -\frac{\Gamma\left(\frac{s+\rho_0}{2}\right)\Gamma\left(\frac{s-\rho_0}{2}+1\right)}{2\Gamma(s+1)}\left(\frac{Q(v)}{Q(v_{z^\perp})}\right)^{\frac{s+\rho_0}{2}} F\left(\frac{s+\rho_0}{2}, \frac{s-\rho_0}{2}+1, s+1, \frac{Q(v)}{Q(v_{z^\perp})}\right). \quad (3\text{-}3)$$

3B. ***Green currents for special divisors.*** The functions $\phi^{(2)}(v,z,s)$ can be used to construct Green functions for the special divisors introduced above. Namely, let $\Gamma \subset H(\mathbb{R})$ be of the form $\Gamma = H_+(\mathbb{Q}) \cap K$ and $v \in V(F)$ be a vector of totally positive norm. Recall that we write $\Gamma_v = \Gamma \cap H_v(\mathbb{R})$. For $\mathrm{Re}(s) > \rho_0$, define

$$G(v,z,s)_\Gamma = 2 \sum_{\gamma \in \Gamma_v \backslash \Gamma} \phi^{(2)}(v, \gamma z, s). \quad (3\text{-}4)$$

The sum converges absolutely a.e. and defines an integrable function $G(v,s)_\Gamma$ on X_Γ [Oda and Tsuzuki 2003, Proposition 3.1.1]. Denote by $[G(v,s)_\Gamma]$ the associated current on X_Γ, defined by

$$[G(v,s)_\Gamma](\alpha) = \int_{X_\Gamma} G(v,z,s)_\Gamma \cdot \alpha(z), \quad (3\text{-}5)$$

for $\alpha \in \mathscr{A}_c^{2n}(X_\Gamma)$. This current admits meromorphic continuation to $s \in \mathbb{C}$ with only simple poles [Oda and Tsuzuki 2003, Theorem 6.3.1]. In fact, as shown by Bruinier [2012, Theorem 5.12], one can refine this result to show that the function $G(v,z,s)_\Gamma$ itself has meromorphic continuation to the whole complex plane and that the resulting function is real analytic on $X_\Gamma - Z(v)_\Gamma$. Define

$$G(v)_\Gamma = \mathrm{CT}_{s=\rho_0} G(v,s)_\Gamma \quad (3\text{-}6)$$

to be the constant term of $G(v,s)_\Gamma$ at $s = \rho_0$.

Theorem 3.2 [Bruinier 2012, Theorem 5.14, Corollary 5.16]. *The function $G(v)_\Gamma$ is real analytic on $X_\Gamma - Z(v)_\Gamma$ and has a logarithmic singularity on $Z(v)_\Gamma$. The form $dd^c G(v)_\Gamma = -(2\pi i)^{-1}\partial\bar\partial G(v)_\Gamma$ extends to a $\mathscr{C}^\infty$ form on X_Γ and one has the equation of currents*:

$$dd^c[G(v)_\Gamma] = \delta_{Z(v)_\Gamma} + [dd^c G(v)_\Gamma]. \quad (3\text{-}7)$$

Consider now a pair of vectors v, w spanning a totally positive definite plane U in V. Denote by $p_{v^\perp}(w)$ the projection of w to the orthogonal complement of v. Recall that we write $X(v)_\Gamma = \Gamma_v \backslash \mathbb{D}_v^+$ and $\Gamma_{v,w} = \Gamma \cap H_U(\mathbb{R})$. The map

$$\Gamma_{v,w} \backslash \mathbb{D}_U^+ \to X(v)_\Gamma$$

then defines an effective divisor $Z(v, w)_\Gamma$ in $X(v)_\Gamma$. We define

$$G(v, w, z, s)_\Gamma = 2 \sum_{\gamma \in \Gamma_{v,w} \backslash \Gamma_v} \phi^{(2)}_{\mathbb{D}_v}(p_{v^\perp}(w), \gamma z, s). \tag{3-8}$$

The results described above imply that the sum converges when $\mathrm{Re}(s) \gg 0$ to an integrable function on $X(v)_\Gamma$, and that we have a meromorphic continuation property, so that we can define

$$G(v, w, z)_\Gamma = \mathrm{CT}_{s=(n-1)/2}\, G(v, w, z, s)_\Gamma. \tag{3-9}$$

The function $G(v, w)_\Gamma$ is then real analytic on $X(v)_\Gamma - Z(v, w)_\Gamma$ and has a logarithmic singularity on $Z(v, w)_\Gamma$.

3C. *The functions $\phi(v, w, z, s)$ on $\mathbb{D}$.* For a pair of vectors $v, w \in V_1$, denote by $p_w(v)$ and $p_{w^\perp}(v)$ the projection of v to the line spanned by w and to the orthogonal complement of w, respectively.

Definition 3.3. Let v, w be a pair of vectors in V_1 spanning a positive definite plane and let $s_0 = (n-1)/2$. For $\mathrm{Re}(s) > s_0$, define

$$\begin{aligned}\phi(v, w, z, s) = -\frac{1}{2}\frac{\Gamma\left(\frac{s+s_0}{2}\right)\Gamma\left(\frac{s-s_0}{2}+1\right)}{\Gamma(s+1)}&\left(\frac{Q(v)-Q(p_w(v))}{Q(v_{z^\perp})-Q(p_w(v))}\right)^{\frac{s+s_0}{2}}\\ &\times F\left(\frac{s+s_0}{2}, \frac{s-s_0}{2}+1, s+1, \frac{Q(v)-Q(p_w(v))}{Q(v_{z^\perp})-Q(p_w(v))}\right).\end{aligned} \tag{3-10}$$

The following basic properties of $\phi(v, w, z, s)$ are easily checked.

Lemma 3.4. (1) *For every $h \in H_v(\mathbb{R})$, $\phi(v, w, z, s) = \phi(v, w, hz, s)$.*

(2) *For every $h \in H(\mathbb{R})$, $\phi(hv, hw, hz, s) = \phi(v, w, z, s)$.*

(3) *The restriction of $\phi(v, w, z, s)$ to $\mathbb{D}_w$ equals $\phi^{(2)}_{\mathbb{D}_w}(p_{w^\perp}(v), z, s)$.*

(4) *Consider the function $\phi(v, w, g, s) = \phi(v, w, gz_0, s)$, for $g \in G$. It belongs to $\mathscr{C}^\infty(G - G_v K_{z_0})$ and satisfies $\phi(v, w, g'gk, s) = \phi(v, w, g, s)$ for every $g' \in G_v$, $k \in K_{z_0}$. Writing $G = G_v A^+ K_{z_0}$ as above, we have*

$$\phi(v, w, a_t, s) = \log(t) + O(1) \qquad \text{as } t \to 0, \tag{3-11}$$

$$\phi(v, w, a_t, s) = O(e^{-(\mathrm{Re}(s)+s_0)t}) \quad \text{as } t \to +\infty. \tag{3-12}$$

Note that (1) and (2) imply $\phi(v, w, z, s) = \phi(v, h_v w, z, s)$ for every $h_v \in H_v(\mathbb{R})$, so that for fixed v, z, s, the function $\phi(v, w, z, s)$ only depends on the $H_v(\mathbb{R})$-orbit of w. Moreover, property (3-12) also holds for all partial derivatives of $\phi(v, w, z, s)$. Note also that property (3-11) implies that $\phi(v, w, z, s)$ is locally integrable. Concerning the behavior of the partial derivatives of $\phi(v, w, z, s)$ as z approaches $\mathbb{D}_v$, we have the following lemma.

Lemma 3.5. *Each of the partial derivatives* $\partial\phi(v, w, z, s)$, $\bar\partial\phi(v, w, z, s)$ *and* $\partial\bar\partial\phi(v, w, z, s)$ *is locally integrable.*

Proof. Let $U \subset \mathbb{D}^+$ be open with coordinates $\{z_1, \dots, z_n\}$ such that the analytic divisor $\mathbb{D}_v^+ \cap U$ is given by the equation $z_1 = 0$ on U. Choosing a trivialization of $\mathscr{L}$ on U we can write $-Q(v_z) = \|s_v(z)\|^2 = h(z)|z_1|^2$, where $h(z)$ is real analytic on U. It follows from the expansion of the hypergeometric function $F(a, b, a+b, w)$ around $w = 1$ (see [Lebedev 1965, (9.7.5)]) that, for fixed v, w, s and $z \in U$,

$$\phi(v, w, z, s) = \log|z_1| + |z_1|^2 \log|z_1| f(z) + g(z), \tag{3-13}$$

where f and g are real analytic functions on U. Thus at worst the singularities of $\|\partial\phi(v, w, z, s)\|$, $\|\bar\partial\phi(v, w, z, s)\|$ and $\|\partial\bar\partial\phi(v, w, z, s)\|$ are of the form $|z_1|^{-1}$ or $\log|z_1|$, and the statement follows. □

The function $\phi(v, w, z, s)$ can also be obtained as a Laplace transform of a certain Whittaker function that depends on s. Namely, consider Kummer's hypergeometric function:

$$M(a, b, z) = \sum_{n=0}^{+\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}. \tag{3-14}$$

The function

$$M_{\nu,\mu}(z) = e^{-z/2} z^{1/2+\mu} M\big(\tfrac{1}{2} + \mu - \nu, 1 + 2\mu, z\big) \tag{3-15}$$

is then a solution of the Whittaker differential equation

$$\frac{d^2w}{dz^2} + \left(-\frac{1}{4} + \frac{\nu}{z} - \frac{\mu^2 - 1/4}{z^2}\right) w = 0. \tag{3-16}$$

It is characterized among solutions of this equation by its asymptotic behavior, given by:

$$M_{\nu,\mu}(z) = z^{\mu+1/2}(1 + O(z)) \qquad \text{when } z \to 0, \tag{3-17}$$

$$M_{\nu,\mu}(z) = \frac{\Gamma(1+2\mu)}{\Gamma(\mu - \nu + 1/2)} e^{z/2} z^{-\nu}(1 + O(z^{-1})) \qquad \text{when } z \to \infty. \tag{3-18}$$

For a positive definite symmetric matrix $T = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$, define

$$s_0 = (n-1)/2, \quad k = 1 - s_0, \tag{3-19}$$

$$C(T, s) = -\frac{1}{2} \frac{\Gamma\big(\frac{s-s_0}{2} + 1\big)}{\Gamma(s+1)} \left(\frac{4\pi \det T}{c}\right)^{-k/2}, \tag{3-20}$$

and

$$M_T(y, s) = C(T, s)|y|^{-k/2} M_{-k/2, s/2}\left(\left|\frac{4\pi \det T}{c} y\right|\right) e^{\frac{2\pi b^2}{c} y}, \quad \mathrm{Re}(s) > s_0. \tag{3-21}$$

Now consider $v, w \in V_1$ spanning a positive definite plane and denote by

$$T(v, w) = \frac{1}{2}\begin{pmatrix} (v, v) & (v, w) \\ (v, w) & (w, w) \end{pmatrix} \tag{3-22}$$

the associated moment matrix. Then (see [Erdélyi et al. 1954, p. 215, §4.22 (11)])

$$\phi(v, w, z, s) = \int_0^\infty M_{T(v,w)}(y, s)e^{-2\pi y(Q(v_{z^\perp})-Q(v_z))}\,\frac{dy}{y}. \tag{3-23}$$

3D. ***Currents in $\mathscr{D}^{1,1}(X_\Gamma)$.*** We now define some (1,1)-forms and currents on X_Γ by summation over translates by elements of Γ of some differential forms with singularities on $\mathbb{D}$. For vectors $v, w \in V(F)$ spanning a totally positive definite space, consider the (1,1)-form $\omega(v, w, z, s)$ defined for $z \in \mathbb{D}^+ - (\mathbb{D}^+_v \cup \mathbb{D}^+_w)$ by

$$\begin{aligned}\omega(v, w, z, s) &= \bar\partial(\phi(w, v, z, s)\partial\phi(v, w, z, s)) \\ &= \bar\partial\phi(w, v, z, s) \wedge \partial\phi(v, w, z, s) \\ &\quad + \phi(w, v, z, s)\bar\partial\partial\phi(v, w, z, s).\end{aligned} \tag{3-24}$$

We would like to define a (1,1)-form on X_Γ by averaging the form $\omega(v, w, z, s)$ over Γ. Before making such a definition, we need to check that the resulting sums converge in a suitable sense. This is the content of the next result. Note that we have

$$\gamma^*(\omega(v, w, s))(z) = \omega(\gamma^{-1}v, \gamma^{-1}w, z, s)$$

for all $\gamma \in \Gamma$, due to the invariance property in Lemma 3.4(2).

Proposition 3.6. *Let $v, w \in V(F)$ be vectors spanning a totally positive definite plane. Let $U = \mathbb{D}^+ - (\Gamma \cdot \mathbb{D}^+_v \cup \Gamma \cdot \mathbb{D}^+_w)$. For* $\mathrm{Re}(s) \gg 0$, *the sum*

$$\sum_{\gamma \in \Gamma_{v,w}\backslash\Gamma} \omega(\gamma^{-1}v, \gamma^{-1}w, z, s)$$

and all its partial derivatives converge normally for every $z \in U$.

Proof. Since the function $\phi(v, w, \gamma z, s)$ is defined and smooth for every $z \in \mathbb{D} - \mathbb{D}_{\gamma^{-1}v}$, all the terms in the sum are defined whenever $z \in U$. Fix $z_0 \in U$ and let $U_0 \subset U$ be a compact neighborhood of z_0; then there exists $\epsilon > 0$ such that $|Q((\gamma v)_z)| > \epsilon$ and $|Q((\gamma w)_z)| > \epsilon$ for all $\gamma \in \Gamma$ and all $z \in U_0$. It follows from Lemma 3.4 that on U_0 we have

$$\|\omega(\gamma^{-1}v, \gamma^{-1}w, z, s)\| < C_\epsilon |Q((\gamma^{-1}v)_{z^\perp})|^{-(s+s_0)/2}\,|Q((\gamma^{-1}w)_{z^\perp})|^{-(s+s_0)/2}$$

for some constant $C_\epsilon > 0$, and a similar bound holds for the sums of all the partial derivatives of the summands. Thus, for $z \in U_0$, the sums in the statement are dominated by a constant multiple of

$$\sum_{\gamma\in\Gamma_{v,w}\backslash\Gamma} |Q((\gamma^{-1}v)_{z^\perp})|^{-(s+s_0)/2}\,|Q((\gamma^{-1}w)_{z^\perp})|^{-(s+s_0)/2}.$$

Pick a lattice $L\subset V(F)$ such that $\Gamma\cdot(v,w)\subset L^2$; then the above sum is dominated by

$$\Bigg(\sum_{\substack{\lambda\in L\\ Q(\lambda)=Q(v)}} |Q(\lambda_{z^\perp})|^{-(s+s_0)/2}\Bigg)\Bigg(\sum_{\substack{\lambda\in L\\ Q(\lambda)=Q(w)}} |Q(\lambda_{z^\perp})|^{-(s+s_0)/2}\Bigg),$$

which converges normally on U, since the assignment $v\mapsto Q(v_{z^\perp})-Q(v_z)$ defines a positive definite quadratic form on V_1 that depends continuously on z. $\square$

Define

$$\Phi(v,w,z,s)_\Gamma = 2\sum_{\gamma\in\Gamma_{v,w}\backslash\Gamma} \omega(\gamma^{-1}v,\gamma^{-1}w,z,s), \tag{3-25}$$

and note that

$$\Phi(v,w,z,s)_\Gamma = \Phi(\gamma v,\gamma w,z,s)_\Gamma \quad \text{for all } \gamma\in\Gamma. \tag{3-26}$$

Proposition 3.6 shows that $\Phi(v,w,\cdot,s)_\Gamma$ converges and defines a smooth (1,1)-form on $X_\Gamma-(Z(v)_\Gamma\cup Z(w)_\Gamma)$.

Denote the cotangent bundle of a manifold X by T^*X. A section s of a metrized vector bundle $(E,\|\cdot\|)$ over a manifold X endowed with a measure $d\mu(z)$ is said to be L^1 (or integrable) if $\|s\|\in L^1(X,d\mu(z))$. Our next goal is to show that $\Phi(v,w,z,s)_\Gamma$ is integrable on X_Γ; this is the content of Proposition 3.9. The next two lemmas will be used in the proof.

Lemma 3.7. *Let M be a complete, simply connected Riemannian manifold of everywhere nonpositive sectional curvature. Let $X,Y\subset M$ be complete, simply connected, totally geodesic submanifolds that intersect transversely and at a single point $z_0\in M$. For $z\in M$, denote by $d(z,z_0)$ the geodesic distance between z and z_0 and by $d_X(z)$ and $d_Y(z)$ the geodesic distance from z to X and from z to Y, respectively. Then there exists a constant $k>0$ such that $d(z_0,z)\geq t$ implies max$\{d_X(z),d_Y(z)\}\geq kt$ for every $t\geq 0$.*

Proof. Let $d>0$ and suppose that $\max\{d_X(z),d_Y(z)\}<d$. Choose points $z_X\in X$ and $z_Y\in Y$ such that $d(z_X,z)<d$ and $d(z_Y,z)<d$. Let $\gamma(z_X,z_Y)$ be the geodesic segment connecting z_X and z_Y; such a geodesic exists, is unique and minimizes the distance (see [Chavel 2006, Exercise IV.12(a)]), hence its length $l(\gamma(z_X,z_Y))$ satisfies $l(\gamma(z_X,z_Y))<2d$. Let $\gamma(z_0,z_X)$ and $\gamma(z_0,z_Y)$ be the geodesic segments in X and Y connecting z_0 and z_X and z_0 and z_Y, respectively; as before, these geodesics exist and are unique and minimizing.

Consider now the triangle T in M with sides $\{\gamma(z_0, z_X), \gamma(z_0, z_Y), \gamma(z_X, z_Y)\}$. This is a geodesic triangle since X and Y are totally geodesic. Note that the angle at z_0 is bounded below since X and Y are assumed to intersect transversely. By the Cartan–Hadamard theorem (see [Bridson and Haefliger 1999, Theorem II.4.1]), the space M is a $CAT(0)$ space, in other words the (unique up to congruence) triangle in the euclidean plane with same sides as T has larger angles than T (see [Bridson and Haefliger 1999, Proposition II.1.7(4)]). It follows that $d(z_0, z_X) \leq cd(z_X, z_Y)$ for some positive constant c. Hence $d(z_0, z) \leq d(z_0, z_X) + d(z_X, z) < (2c+1)d$ as required. □

Lemma 3.8. *Let M, X, Y be as in Lemma 3.7. Assume that the codimension of X and Y in M is greater than one and that the sectional curvature of M is bounded below. Let $f_{1,s}, f_{2,s} : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be continuous functions defined for* $\mathrm{Re}(s) > s_0 > 0$ *such that*

$$t f_{i,s}(t) = O(1), \qquad as\ t \to 0,$$
$$f_{i,s}(t) = e^{-\mathrm{Re}(s)t}, \quad as\ t \to \infty,$$

for $i = 1, 2$. Let $d\mu(z)$ be the Riemannian volume element of M. Then, with notation as in Lemma 3.7, we have

$$\int_M f_{1,s}(d_X(z)) f_{2,s}(d_Y(z))\, d\mu(z) < \infty,$$

for $\mathrm{Re}(s) \gg 0$.

Proof. Let $U_X = \{z \in M \mid d_X(z) \leq 1\}$ and $U_Y = \{z \in M \mid d_Y(z) \leq 1\}$ be tubular neighborhoods around X and Y of radius 1. Let $U = M - (U_X \cup U_Y)$. It suffices to show that $f_s(z) = f_{1,s}(d_X(z)) f_{2,s}(d_Y(z))$ is integrable when restricted to U, U_X and U_Y.

Consider first the integral over U. By hypothesis, the functions $f_{1,s}(d_X(z))$ and $f_{2,s}(d_Y(z))$ are bounded on U. By Lemma 3.7, there exists a constant $k > 0$ such that

$$f_s(z) = O(e^{-\mathrm{Re}(s) k d(z, z_0)})$$

for $z \in U$. Let $S(z_0, t)$ be the geodesic sphere with center z_0 and radius t and denote by $A(t)$ its area. Since M has curvature that is bounded below, there exists $\rho > 0$ such that $A(t) = O(e^{\rho t})$ (see [Chavel 2006, Theorem III.4.4]). It follows that

$$\int_U f_s(z)\, d\mu(z) < \infty$$

whenever $\mathrm{Re}(s) > \rho/k$.

Now consider the integral over U_X (the same argument works for U_Y). Since $f_s(z)$ is locally integrable, it suffices to integrate over $U_X - (U_X \cap U_Y)$. The inclusion $i : X \subset U_X$ admits a left inverse $\pi : U_X \to X$ whose fibers are diffeomorphic to the closed unit disk in $\mathbb{C}$ (this is because the exponential map from the total space of the normal bundle of X to M is a diffeomorphism). We can compute the integral over U_X by first integrating over the fibers of π and then integrating over X. By hypothesis, the integral of $f_s(z)$ over $\pi^{-1}(z)$ is $O(e^{-\mathrm{Re}(s)d(z,z_0)})$ for every $z \in X - (X \cap U_Y)$. Now the resulting integral over X converges for $\mathrm{Re}(s) \gg 0$ since the area of a sphere of radius t in X is $O(e^{\rho t})$ as above. □

We can now prove that $\Phi(v, w, z, s)_\Gamma$ is integrable on X_Γ. Recall that $\mathbb{D}$ carries an $H(\mathbb{R})$-invariant Riemannian metric; it induces an invariant metric on $\bigwedge^2 T^*\mathbb{D}$ that we denote by $\|\cdot\|$.

Proposition 3.9. *Let $v, w \in V(F)$ be vectors spanning a totally positive plane. For* $\mathrm{Re}(s) \gg 0$, *the sum $\Phi(v, w, z, s)_\Gamma$ converges outside a set of measure zero in X_Γ and defines an L^1 section of $(\bigwedge^2 T^*X_\Gamma, \|\cdot\|)$.*

Proof. The sum converges for $z \notin Z(v)_\Gamma \cup Z(w)_\Gamma$ by Proposition 3.6, and this set has measure zero. Thus it remains to prove integrability. We need to show that

$$\int_{X_\Gamma} \|\Phi(v, w, z, s)\| \, d\mu(z)$$

is convergent, where $d\mu(z)$ denotes an invariant volume form on $\mathbb{D}^+$. By Fubini's theorem, it suffices to show that

$$\int_{\Gamma_{v,w}\backslash\mathbb{D}^+} \|\omega(w, v, z, s)\| \, d\mu(z) < \infty.$$

Let $H'(\mathbb{R}) = (H_v)_+(\mathbb{R}) \cap (H_w)_+(\mathbb{R})$ and let $Z_{H'}(\mathbb{R})$ be the center of $H'(\mathbb{R})$. Since the integrand is left invariant under $H'(\mathbb{R})$ by Lemma 3.4 and $Z_{H'}(\mathbb{R})\Gamma_{v,w}\backslash H'(\mathbb{R})$ has finite volume (see [Borel 1969]), this is equivalent to

$$\int_{H'(\mathbb{R})\backslash\mathbb{D}^+} \|\omega(w, v, z, s)\| \, d\mu(z) < \infty. \tag{$*$}$$

We now apply Lemma 3.8. Namely, let $M = H'(\mathbb{R})\backslash\mathbb{D}^+$. Let $X = H'(\mathbb{R})\backslash\mathbb{D}^+_v$ and $Y = H'(\mathbb{R})\backslash\mathbb{D}^+_w$. Note that there is a map $\pi : \mathbb{D}^+ \to \mathbb{D}^+_v$ that is left inverse to the inclusion $\mathbb{D}^+_v \subset \mathbb{D}^+$ and turns $\mathbb{D}^+$ into an $H_v(\mathbb{R})_+$-equivariant real vector bundle of rank 2 over $\mathbb{D}^+_v$ (see [Kudla and Millson 1988, p. 26]). Hence the inclusions

$$\{*\} = H'(\mathbb{R})\backslash\mathbb{D}^+_{v,w} \subset H'(\mathbb{R})\backslash\mathbb{D}^+_v \subset H'(\mathbb{R})\backslash\mathbb{D}^+$$

are diffeomorphic to zero sections of vector bundles, in particular they are simply connected. Moreover $\mathbb{D}^+_v$ and $\mathbb{D}^+_w$ are totally geodesic submanifolds of $\mathbb{D}^+$, and the latter is known to have sectional curvatures that are bounded below and everywhere nonpositive. Hence X, Y and M satisfy the hypotheses in Lemma 3.7 and Lemma 3.8. Moreover, by Lemma 3.5, the integrand also satisfies the hypotheses in Lemma 3.8; applying it gives $(*)$ and hence the assertion. □

Since $\Phi(v, w, z, s)_\Gamma$ is an integrable section of $\bigwedge^2 T^* X_\Gamma$, its coordinates in any chart $U \subset X_\Gamma$ are locally integrable functions. Thus $\Phi(v, w, z, s)_\Gamma$ defines a current on X_Γ.

Definition 3.10. Let $v, w \in V(F)$ be vectors spanning a totally positive definite plane. For $\mathrm{Re}(s) \gg 0$, define a current $[\Phi(v, w, s)_\Gamma] \in \mathscr{D}^{1,1}(X_\Gamma)$ by

$$[\Phi(v, w, s)_\Gamma](\omega) = \int_{X_\Gamma} \Phi(v, w, s)_\Gamma \wedge \omega, \tag{3-27}$$

for $\omega \in \mathscr{A}_c^{n-1,n-1}(X_\Gamma)$.

Recall that we assume $\Gamma = H_+(\mathbb{Q}) \cap K$ for some open compact $K \subset H(\mathbb{A}_f)$. For $h \in H(\mathbb{A}_f)$, we write $\Gamma_h = H_+(\mathbb{Q}) \cap hKh^{-1}$ and we define

$$\Phi(v, w, h, s)_\Gamma = \Phi(v, w, s)_{\Gamma_h}, \tag{3-28}$$

an L^1 section of $\wedge^2 T^*(\Gamma_h \backslash \mathbb{D}^+)$. As above, we denote by $[\Phi(v, w, h, s)_\Gamma]$ the associated current in $\mathscr{D}^{1,1}(\Gamma_h \backslash \mathbb{D}^+)$.

3E. ***Some properties of*** $[\boldsymbol{\Phi(v, w, h, s)_\Gamma}]$**.** We now introduce another family of currents $[\Phi(v, w)_\Gamma]$ on X_Γ. These currents are obtained by restricting a compactly supported form $\omega \in \mathscr{A}_c^{n-1,n-1}(X_\Gamma)$ to a special divisor $X(v)_\Gamma$ and integrating it against a Green function of the form (3-6). In this section we will prove that the current $[\Phi(v, w, s)_\Gamma]$ introduced above, regarded modulo $\mathrm{Im}(\partial) + \mathrm{Im}(\bar\partial)$, admits meromorphic continuation to the complex plane s and that the current $[\Phi(v, w)_\Gamma]$ is cohomologous to the current obtained as the constant term of the meromorphic continuation of $[\Phi(v, w, s)_\Gamma]$ at a certain value $s = s_0$.

For $v \in V(F)$ of totally positive norm, denote by $\delta_{X(v)_\Gamma} \in \mathscr{D}^{1,1}(X_\Gamma)$ the current of integration along $X(v)_\Gamma$. That is, for $\omega \in \mathscr{A}_c^{n-1,n-1}(X_\Gamma)$, we have

$$\delta_{X(v)_\Gamma}(\omega) = \int_{X(v)_\Gamma} \omega. \tag{3-29}$$

Consider now $v, w \in V(F)$ spanning a totally positive definite plane. In Section 3B we recalled the construction (see [Bruinier 2012; Oda and Tsuzuki 2003]) of a function $G(v, w)_\Gamma \in \mathscr{C}^\infty(X(v)_\Gamma - Z(v, w)_\Gamma)$. The function has a logarithmic singularity along $Z(v, w)_\Gamma$, hence is locally integrable on $X(v)_\Gamma$ and defines an element of $\mathscr{D}^0(X(v)_\Gamma)$ that we denote by $[G(v, w)_\Gamma]$. Recall that there is a pushforward map

$$f_* : \mathscr{D}^0(X(v)_\Gamma) \to \mathscr{D}^{1,1}(X_\Gamma) \tag{3-30}$$

induced by $f : X(v)_\Gamma \to X_\Gamma$ and defined by $(f_*(\alpha), \omega) = (\alpha, f^*(\omega))$ for α in $\mathscr{D}^0(X(v)_\Gamma)$ and ω in $\mathscr{A}_c^{n-1,n-1}(X_\Gamma)$.

Definition 3.11. Let $v, w \in V(F)$ spanning a totally positive definite plane. Define the current $[\Phi(v,w)_\Gamma] \in \mathscr{D}^{1,1}(X_\Gamma)$ by

$$[\Phi(v,w)_\Gamma] = 2\pi i \cdot f_*([G(v,w)_\Gamma]). \tag{3-31}$$

For $h \in H(\mathbb{A}_f)$ and $K \subset H(\mathbb{A}_f)$ such that $\Gamma = H_+(\mathbb{Q}) \cap K$, define

$$[\Phi(v,w,h)_\Gamma] = [\Phi(v,w)_{\Gamma_h}], \tag{3-32}$$

where $\Gamma_h = H_+(\mathbb{Q}) \cap hKh^{-1}$.

That is, for $\omega \in \mathscr{A}_c^{n-1,n-1}(X_\Gamma)$, we have

$$[\Phi(v,w)_\Gamma](\omega) = 2\pi i \int_{X(v)_\Gamma} G(v,w)_\Gamma \cdot \omega. \tag{3-33}$$

See Section 4B2 for an example.

The next proposition relates the currents $[\Phi(v,w)_\Gamma]$ and $[\Phi(v,w,s)_\Gamma]$ and is key to the computation of values of $[\Phi(v,w)_\Gamma]$ on forms obtained as theta lifts as below. Let

$$\tilde{\mathscr{D}}^{1,1}(X_\Gamma) = \mathscr{D}^{1,1}(X_\Gamma)/(\mathrm{Im}(\partial) + \mathrm{Im}(\bar\partial)). \tag{3-34}$$

We let $[\Phi(v,w,s)_\Gamma]$ and $[\Phi(v,w)_\Gamma]$ also denote the classes of $[\Phi(v,w,s)_\Gamma]$ and $[\Phi(v,w)_\Gamma]$ in $\tilde{\mathscr{D}}^{1,1}(X_\Gamma)$.

Proposition 3.12. *The current $[\Phi(v,w,s)_\Gamma] \in \tilde{\mathscr{D}}^{1,1}(X_\Gamma)$ admits meromorphic continuation to $s \in \mathbb{C}$. Let $\mathrm{CT}_{s=s_0}[\Phi(v,w,s)_\Gamma] \in \tilde{\mathscr{D}}^{1,1}(X_\Gamma)$ denote the constant term of $[\Phi(v,w,s)_\Gamma]$ at $s = s_0$. Then*

$$\mathrm{CT}_{s=s_0}[\Phi(v,w,s)_\Gamma] = [\Phi(v,w)_\Gamma] \tag{3-35}$$

as elements of $\tilde{\mathscr{D}}^{1,1}(X_\Gamma)$.

Proof. Let $\alpha \in \mathscr{A}_c^{n-1,n-1}(X_\Gamma)$. By Proposition 3.9, we have

$$[\Phi(v,w,s)_\Gamma](\alpha) = 2\int_{\Gamma_{v,w}\backslash\mathbb{D}^+} \omega(v,w,z,s) \wedge \alpha(z).$$

For fixed s, write $g_v(z) = \phi(v,w,z,s)$ and $g_w(z) = \phi(w,v,z,s)$. We regard g_v as a smooth function defined on $\Gamma_{v,w}\backslash\mathbb{D}^+ - \Gamma_{v,w}\backslash\mathbb{D}^+_v$. If we choose an open $U \subset \Gamma_{v,w}\backslash\mathbb{D}^+$ such that the analytic divisor $(\Gamma_{v,w}\backslash\mathbb{D}^+_v) \cap U$ is given by the equation $z = 0$, then it follows from (3-13) that

$$\partial g_v(z) = \frac{dz}{z} + o(|z|^{-1}), \quad \bar\partial g_v(z) = \frac{d\bar z}{\bar z} + o(|z|^{-1}).$$

(Here $o(|z|^{-1})$ stands for a differential form α on $U - (\Gamma_{v,w}\backslash\mathbb{D}^+_v) \cap U$ such that the components of $|z|\alpha$ extend to continuous functions on U vanishing on $(\Gamma_{v,w}\backslash\mathbb{D}^+_v) \cap U$.) Similar statements hold for $g_w(z)$ when z approaches $\Gamma_{v,w}\backslash\mathbb{D}^+_w$. Denote by $\delta_v \in \mathscr{D}^{1,1}(\Gamma_{v,w}\backslash\mathbb{D}^+)$ the current given by integration on $\Gamma_{v,w}\backslash\mathbb{D}^+_v$. The following identity of currents on $\Gamma_{v,w}\backslash\mathbb{D}^+$ follows from Stokes's theorem applied to $\Gamma_{v,w}\backslash\mathbb{D}^+ - (\Gamma_{v,w}\backslash\mathbb{D}^+_v \cup \Gamma_{v,w}\backslash\mathbb{D}^+_w)$:

$$\bar{\partial}[g_w \partial g_v] = [\bar{\partial} g_w \partial g_v] + [g_w \bar{\partial}\partial g_v] - 2\pi i g_w \delta_v. \tag{3-36}$$

We find that for any closed compactly supported form $\alpha_c \in \mathscr{A}_c^{n-1,n-1}(\Gamma_{v,w}\backslash\mathbb{D}^+)$,

$$\int_{\Gamma_{v,w}\backslash\mathbb{D}^+} \omega(v,w,z,s) \wedge \alpha_c(z) = 2\pi i \int_{\Gamma_{v,w}\backslash\mathbb{D}^+_v} \phi(w,v,z,s)\alpha_c(z). \tag{3-37}$$

The form $\alpha(z)$ is not compactly supported, but we claim that (3-37) is still true for $\mathrm{Re}(s) \gg 0$ when we replace $\alpha_c(z)$ by $\alpha(z)$. Assuming this for now and using that the restriction of $\phi(w,v,z,s)$ to $\mathbb{D}_v$ equals $\phi^{(2)}_{\mathbb{D}_v}(p_{v^\perp}(w), z, s)$, we conclude that for $\mathrm{Re}(s) \gg 0$

$$\begin{aligned}
[\Phi(v,w,s)_\Gamma](\alpha) &\equiv 2\pi i \cdot 2 \int_{\Gamma_{v,w}\backslash\mathbb{D}^+_v} \phi^{(2)}_{\mathbb{D}_v}(p_{v^\perp}(w), z, s)\alpha(z) \\
&= 2\pi i \int_{\Gamma_v\backslash\mathbb{D}^+_v} 2 \sum_{\gamma\in\Gamma_{v,w}\backslash\Gamma_v} \phi^{(2)}_{\mathbb{D}_v}(p_{v^\perp}(w), \gamma z, s)\alpha(z) \\
&= 2\pi i \int_{\Gamma_v\backslash\mathbb{D}^+_v} G(p_{v^\perp}(w), z, s)_{\Gamma_v} \cdot \alpha(z).
\end{aligned}$$

This last equation defines a current on X_Γ that admits meromorphic continuation to $s \in \mathbb{C}$ and whose constant term at $s = s_0$ is given by $[\Phi(v,w)_\Gamma]$; the claim follows from this.

It only remains to show that (3-37) still holds when we replace $\alpha_c(z)$ by $\alpha(z)$. Let $X = \Gamma_{v,w}\backslash\mathbb{D}^+$ and consider the submanifolds $X_v = \Gamma_{v,w}\backslash\mathbb{D}^+_v$ and $X_w = \Gamma_{v,w}\backslash\mathbb{D}^+_w$ of X. Let $X_{v,w} = X_v \cap X_w = \Gamma_{v,w}\backslash\mathbb{D}^+_{v,w}$. As remarked by Kudla and Millson [1988, p. 26], the exponential map of the normal bundle of $X_{v,w} \subset X$ is a diffeomorphism, and hence X carries a natural vector bundle structure $\pi : X \to X_{v,w}$ of rank 4 over $X_{v,w}$ with totally geodesic fibers. For $t > 0$, let $X_v(t) = \{z \in X \mid d_{X_v}(z) \leq t\}$ be the tubular neighborhood of radius t around X_v; here $d_{X_v}(z)$ denotes the geodesic distance between z and X_v. Define $X_w(t)$ and $X_{v,w}(t)$ similarly and let $X(t) = X_{v,w}(t) - (X_v(1/t) \cup X_w(1/t))$. Then we have $X - (X_v \cup X_w) = \bigcup_{t\geq 1} X(t)$ and

$$\int_X \omega(v,w,z,s) \wedge \alpha = \lim_{t\to\infty} \int_{X(t)} \omega(v,w,z,s) \wedge \alpha.$$

Denote by $S_{v,w}(t) = \partial X_{v,w}(t)$ the boundary of $X_{v,w}(t)$. By Stokes's theorem, (3-37) is equivalent to

$$\int_{S_{v,w}(t)-(X_v(1/t)\cup X_w(1/t))} \phi(w,v,z,s)\partial\phi(v,w,z,s) \wedge \alpha \longrightarrow 0,$$

as $t \to \infty$. Since $\|\alpha\|$ is bounded, it suffices to show that

$$\int_{S_{v,w}(t)-(X_v(1/t)\cup X_w(1/t))} |\phi(w,v,z,s)| \cdot \|\partial\phi(v,w,z,s)\| \, d\mu(z) \longrightarrow 0, \tag{3-38}$$

as $t \to \infty$. Now let $H'(\mathbb{R}) = (H_v)_+(\mathbb{R}) \cap (H_w)_+(\mathbb{R})$ and note that the integrand is invariant under $H'(\mathbb{R})$. Let $M = H'(\mathbb{R})\backslash\mathbb{D}^+$ and consider the submanifolds $X = H'(\mathbb{R})\backslash\mathbb{D}_v^+$ and $Y = H'(\mathbb{R})\backslash\mathbb{D}_w^+$ of M, whose intersection is a single point z_0. Let $S(z_0, t)$ be the sphere of geodesic radius t around z_0 and let $X(1/t)$ and $Y(1/t)$ be tubular neighborhoods of X and Y with radius $1/t$. Since $X_{v,w}$ has finite volume by [Borel 1969, Theorem 15.5], to show that the integrals in (3-38) tend to 0 it suffices to show that

$$\int_{S(z_0,t)-(X(1/t)\cup Y(1/t))} |\phi(w,v,z,s)| \cdot \|\partial\phi(v,w,z,s)\| \, d\mu(z) \longrightarrow 0,$$

as $t \to \infty$. Now Lemma 3.7 and Lemma 3.4 show that the integrand is $O(te^{-k\,\mathrm{Re}(s)t})$ for some positive constant $k > 0$. Since the sectional curvatures of M are bounded below, we have $\mathrm{Area}(S(z_0,t)) = O(e^{\rho t})$ for some positive constant $\rho > 0$ and hence (3-37) holds, with α_c replaced by α, for $\mathrm{Re}(s) > \rho/k$. □

3F. ***Currents on X_K.*** We introduce now currents in $\mathscr{D}^{1,1}(X_K)$. Fix a neat open compact subgroup $K \subset H(\mathbb{A}_f)$ and recall that we write

$$X_K = H(\mathbb{Q})\backslash(\mathbb{D} \times H(\mathbb{A}_f))/K.$$

Thus X_K is a compact complex manifold with finitely many components. These were described in Section 2: choose $h_1 = 1, \ldots, h_r \in H(\mathbb{A}_f)$ such that

$$H_+(\mathbb{Q})\backslash H(\mathbb{A}_f)/K = \coprod_{j=1}^{r} H_+(\mathbb{Q})h_jK.$$

For $h \in H(\mathbb{A}_f)$, we write $\Gamma_h = H_+(\mathbb{Q}) \cap hKh^{-1}$ (and $\Gamma = \Gamma_1$). Then

$$X_K \cong \coprod_{j=1}^{r} \Gamma_{h_j}\backslash\mathbb{D}^+.$$

Let $v, w \in V(F)$ span a totally positive definite plane U and recall that we denote by $H_U \subset H$ the pointwise stabilizer of U. For $h \in H(\mathbb{A}_f)$, let $K_{U,h} = H_U(\mathbb{A}_f) \cap hKh^{-1}$. Choose coset representatives $h_i' \in H_U(\mathbb{A}_f)$ such that

$$(H_U)_+(\mathbb{Q})\backslash H_U(\mathbb{A}_f)/K_{U,h} = \coprod_{i=1}^{s} (H_U)_+(\mathbb{Q})h_i'K_{U,h}, \tag{3-39}$$

and write $h_i'h = \gamma_i h_j k_i$ with $\gamma_i \in H_+(\mathbb{Q})$, $k_i \in K$ and $h_j = h_{j(i)}$ a coset representative as in (2-8). Note that the double coset $(H_U)_+(\mathbb{Q})\gamma_i\Gamma_{h_j}$ is well defined, that is, it is independent of the choice of h_i' and decomposition $h_i'h = \gamma_i h_j k_i$.

Definition 3.13. Assume that $n > 2$. We define $\Phi(v, w, h, s)_K$ to be the section of $\bigwedge^2 T^*(X_K)$ whose restriction to the connected component $\Gamma_{h_j}\backslash\mathbb{D}^+$ is

$$\sum_{i\to j} \Phi(\gamma_i^{-1}v, \gamma_i^{-1}w, h_j, s)_\Gamma, \tag{3-40}$$

where the sum runs over those i such that $j(i) = j$.

Note that this is well defined because of the invariance property (3-26). For $n = 2$ we give a different definition. Namely, assume that $n = 2$ and choose γ_0 in $H(\mathbb{Q})$ such that $\gamma_0^{-1}\mathbb{D}_U^+ = \mathbb{D}_U^-$. With h_i' as in (3-39), write $\gamma_0 h_i' h = \gamma_{i_0} h_{j_0} k_{i_0}$ with $\gamma_{i_0} \in H_+(\mathbb{Q})$, $k_{i_0} \in K$ and $h_{j_0} = h_{j_0(i_0)}$ a coset representative as in (2-8). As above, the double coset $(H_U)_+(\mathbb{Q})\gamma_{i0}\Gamma_{h_{j_0}}$ is well defined.

Definition 3.14. Assume that $n = 2$. We define $\Phi(v, w, h, s)_K$ to be the section of $\bigwedge^2 T^*(X_K)$ whose restriction to the connected component $\Gamma_{h_j}\backslash\mathbb{D}^+$ is

$$\sum_{i\to j} \Phi(\gamma_i^{-1}v, \gamma_i^{-1}w, h_j, s)_\Gamma + \sum_{i_0\to j} \Phi(\gamma_{i_0}^{-1}v, \gamma_{i_0}^{-1}w, h_j, s)_\Gamma, \tag{3-41}$$

where the sums run over those i and i_0 such that $j(i) = j$ and $j_0(i_0) = j$, respectively.

The forms $\Phi(v, w, h, s)_K$ are locally integrable on X_K. We denote by

$$[\Phi(v, w, h, s)_K] \in \mathscr{D}^{1,1}(X_K) \tag{3-42}$$

the corresponding current on X_K.

We also define a current

$$[\Phi(v, w, h)_K] \in \mathscr{D}^{1,1}(X_K) \tag{3-43}$$

whose restriction to the connected component $\Gamma_{h_j}\backslash\mathbb{D}^+$ is

$$\begin{cases} \displaystyle\sum_{i\to j}[\Phi(\gamma_i^{-1}v, \gamma_i^{-1}w, h_j)_\Gamma] & \text{if } n > 2, \\ \displaystyle\sum_{i\to j}[\Phi(\gamma_i^{-1}v, \gamma_i^{-1}w, h_j)_\Gamma] + \sum_{i_0\to j}[\Phi(\gamma_{i_0}^{-1}v, \gamma_{i_0}^{-1}w, h_j)_\Gamma] & \text{if } n = 2, \end{cases} \tag{3-44}$$

with the currents in the sum as in (3-32). See Section 4B3 for an example.

Remark 3.15. The above definitions reflect the structure of the connected components of the special cycles $Z(v, w, h)_K$ in Section 2B. Namely, let $v, w \in V(F)$ be vectors spanning a totally positive definite plane and $h \in H(\mathbb{A}_f)$. Attached to such a pair there are Shimura varieties $X(v, w, h)_K$ and $X(v, h)_K$ (see (2-13)) together with proper maps

$$X(v, w, h)_K \xrightarrow{\iota} X(v, h)_K \xrightarrow{f} X_K. \tag{3-45}$$

Then $\iota_*([X(v, w, h)_K])$ defines a divisor on $X(v, h)_K$, and a finite sum of functions of the form (3-9) defines a Green function $G(v, w, h)_K$ on $X(v, h)_K$ with a logarithmic singularity along $\iota_*([X(v, w, h)_K])$. Writing $[G(v, w, h)_K]$ for the current in $\mathscr{D}^0(X_K)$ associated with $G(v, w, h)_K$, it follows from Kudla's description of the connected components of the cycles $Z(v, w, h)_K$ (see [Kudla 1997, Lemma 4.1]) that

$$[\Phi(v, w, h)_K] = 2\pi i \cdot f_*([G(v, w, h)_K]).$$

Some basic properties of the forms $\Phi(v, w, h, s)_K$ are summarized in the next lemma; these properties are analogous to those of special cycles proved in [Kudla 1997, Lemma 2.2]. Recall that for every $h \in H(\mathbb{A}_f)$ there is a map

$$r(h) : X_{hKh^{-1}} \longrightarrow X_K \tag{3-46}$$

sending $H(\mathbb{Q})(z, h')hKh^{-1}$ to $H(\mathbb{Q})(z, h'h)K$. The map $r(h)$ is an isomorphism of complex manifolds, and we denote by $\Phi \mapsto \Phi \cdot h$ the induced map defined on sections of the bundle of differential forms.

Lemma 3.16. (1) $\Phi(v, w, hk, s)_K = \Phi(v, w, h, s)_K$ *for all* $k \in K$.

(2) $\Phi(v, w, h_U h, s)_K = \Phi(v, w, h, s)_K$ *for all* $h_U \in H_U(\mathbb{A}_f)$.

(3) $\Phi(\gamma v, \gamma w, \gamma h, s)_K = \Phi(v, w, h, s)_K$ *for all* $\gamma \in H(\mathbb{Q})$.

(4) $\Phi(v, w, h_1 h^{-1}, s)_{hKh^{-1}} \cdot h = \Phi(v, w, h_1, s)_K$ *for all* $h_1, h \in H(\mathbb{A}_f)$.

Proof. Part (1) is obvious. Part (2) follows from the fact that for any complete set $\{h_i' \mid i = 1, \dots, s\}$ of coset representatives for

$$S(U, h, K) = (H_U)_+(\mathbb{Q})\backslash H_U(\mathbb{A}_f)/K_{U,h},$$

the set $\{h_i' h_U^{-1} \mid i = 1, \dots, s\}$ is a complete set of representatives for $S(U, h_U h, K)$. To prove part (3), note that given any set $\{h_i' \mid i = 1, \dots, s\}$ as above and any $\gamma \in H(\mathbb{Q})$, the elements $\gamma h_i' \gamma^{-1}$ for $i = 1, \dots, s$ form a complete set of representatives for $S(\gamma(U), \gamma h, K)$, so that writing $\gamma h_i' \gamma^{-1} \cdot (\gamma h) = (\gamma\gamma_i) h_j k_i$ with $j = j(i)$ leads to

$$\begin{aligned}
\Phi(\gamma v, \gamma w, \gamma h, s)_K\big|_{\Gamma_{h_j}\backslash\mathbb{D}^+} &= \sum_{i \to j} \Phi\big((\gamma\gamma_i)^{-1}\gamma v, (\gamma\gamma_i)^{-1}\gamma w, z, s\big)_{\Gamma_{h_j}} \\
&= \sum_{i \to j} \Phi(\gamma_i^{-1} v, \gamma_i^{-1} w, z, s)_{\Gamma_{h_j}} = \Phi(v, w, h, s)_K\big|_{\Gamma_{h_j}\backslash\mathbb{D}^+},
\end{aligned}$$

as was to be shown. Finally, (4) follows from the fact that if $\{h_j \mid j = 1, \dots, r\}$ is a set of coset representatives for $H_+(\mathbb{Q})\backslash H(\mathbb{A}_f)/K$, then $\{h_j h^{-1} \mid j = 1, \dots, r\}$ is a set of coset representatives for $H_+(\mathbb{Q})\backslash H(\mathbb{A}_f)/hKh^{-1}$. $\square$

Assume that $K' \subset K$, with K' an open compact subgroup of $H(\mathbb{A}_f)$ and let $\mathrm{pr} : X_{K'} \to X_K$ be the natural projection map. The following lemma computes $\mathrm{pr}^*(\Phi(v, w, h, s)_K)$.

Lemma 3.17. *Let $K' \subset K$ be as above. Then*

$$\mathrm{pr}^*(\Phi(v, w, h, s)_K) = \sum_{k \in h^{-1}K_{U,h}h\backslash K/K'} \Phi(v, w, hk, s)_{K'}.$$

Proof. Note that the sum on the right hand side is well defined by (1) and (2) of Lemma 3.16. Now consider the restriction of $\Phi(v, w, h, s)_K$ to $\Gamma_{h_j}\backslash\mathbb{D}^+$. By definition, this is the sum

$$\sum_{i \in I} \Phi(\gamma_i^{-1}v, \gamma_i^{-1}w, h, s)_{\Gamma_{h_j}},$$

where $\gamma_i \in H_+(\mathbb{Q})$ satisfies $\gamma_i h_j k_i = h_i' h$ for some $k_i \in K$ and $h_i' \in H_U(\mathbb{A}_f)$, with $\{h_i' \mid i \in I\}$ a complete set of representatives of the double coset

$$(H_U)_+(\mathbb{Q})\backslash H_U(\mathbb{A}_f) \cap H_+(\mathbb{Q})h_j K h^{-1}/K_{U,h}.$$

Assume first that $n > 2$. By [Kudla 1997, Lemma 5.7(i)], this double coset is in bijection with the set of Γ_{h_j}-orbits in

$$S(v, w, h_j K h^{-1}) := H_+(\mathbb{Q}) \cdot (v, w) \cap h_j K h^{-1} \cdot (v, w).$$

The bijection sends $\Gamma_{h_j} \cdot (v_i, w_i)$, where $(v_i, w_i) = \gamma_i \cdot (v, w) = h_j k_i h^{-1} \cdot (v, w)$ with $\gamma_i \in H_+(\mathbb{Q})$ and $k \in K$, to the double coset $(H_U)_+(\mathbb{Q})\gamma_i^{-1} h_j k_i h^{-1} K_{U,h}$. Substituting the definition of $\Phi(v, w, h, s)_{\Gamma_{h_j}}$, we see that the restriction of $\frac{1}{2}\Phi(v, w, h, s)_K$ to $\Gamma_{h_j}\backslash\mathbb{D}^+$ is given by

$$\sum_{(v', w') \in S(v, w, h_j K h^{-1})} \omega(v', w', z, s).$$

This sum can be rewritten as

$$\sum_{k \in h^{-1}K_{U,h}h\backslash K/K'} \; \sum_{(v', w') \in S(v, w, h_j K'(hk)^{-1})} \omega(v', w', z, s)$$

and the claim follows directly from this. The proof for $n = 2$ proceeds similarly by using [Kudla 1997, Lemma 5.7(ii)]. $\square$

Analogous statements to those in Lemma 3.16 and Lemma 3.17 hold for the currents $[\Phi(v, w, h, s)_K]$ and $[\Phi(v, w, h)_K]$.

3G. ***Weighted currents.*** Following Kudla's [1997] definition of weighted cycles we introduce currents in $\mathscr{D}^{1,1}(X) = \varprojlim \mathscr{D}^{1,1}(X_K)$ as finite sums of the currents $[\Phi(v, w, h, s)_K]$ above weighted by the values of a Schwartz function $\mathscr{S}(V(\mathbb{A}_f)^2)$.

Given a totally positive definite symmetric matrix $T \in \mathrm{Sym}_2(F)$, let

$$\Omega_T(\mathbb{A}_f) = \{(v, w) \in V(\mathbb{A}_f)^2 \mid T(v, w) = T\}, \tag{3-47}$$

where $T(v, w)$ is defined in (3-22). Assume that $\Omega_T(\mathbb{A}_f) \neq \varnothing$. Then there exists $(v_0, w_0) \in \Omega_T(\mathbb{A}_f) \cap V(F)^2$ by the Hasse principle for quadratic forms. Moreover, the action of $H(\mathbb{A}_f)$ on $\Omega_T(\mathbb{A}_f)$ is transitive by Witt's theorem. Let K be a compact open subgroup of $H(\mathbb{A}_f)$. The orbits of K on $\Omega_T(\mathbb{A}_f)$ are open, and so if $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ is invariant under K, we have

$$\mathrm{Supp}(\varphi) \cap \Omega_T(\mathbb{A}_f) = \coprod_{i=1}^{k} K\xi_i^{-1} \cdot (v_0, w_0) \tag{3-48}$$

for some elements $\xi_1, \dots, \xi_k \in H(\mathbb{A}_f)$.

Definition 3.18. Let $T \in \mathrm{Sym}_2(F)$ be a totally positive definite matrix and let $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ be fixed by K. With (v_0, w_0) and ξ_i as above and $\mathrm{Re}(s) \gg 0$, define

$$\Phi(T, \varphi, s)_K = \sum_{i=1}^{k} \varphi\big(\xi_i^{-1} \cdot (v_0, w_0)\big) \cdot \Phi(v_0, w_0, \xi_i, s)_K.$$

We denote by $[\Phi(T, \varphi, s)_K]$ the corresponding current in $\mathscr{D}^{1,1}(X_K)$.

Note that $\Phi(T, \varphi, s)_K$ is independent of the choice of $\{\xi_1, \dots, \xi_k\}$ by (1) and (2) of Lemma 3.16. The behavior of $\Phi(T, \varphi, s)_K$ under pullbacks coming from compact subgroups $K' \subset K$ is simpler than that of the forms $\Phi(v, w, h, s)_K$ defined above. The next proposition proves this and an equivariance property for the action of $H(\mathbb{A}_f)$.

Proposition 3.19. (1) *Let $K' \subset K$ be an open compact subgroup of $H(\mathbb{A}_f)$ and consider the natural map* $\mathrm{pr} : X_{K'} \to X_K$. *Then*

$$\mathrm{pr}^*(\Phi(T, \varphi, s)_K) = \Phi(T, \varphi, s)_{K'}.$$

(2) *For any $h \in H(\mathbb{A}_f)$, we have*

$$\Phi(T, \omega(h)\varphi, s)_{hKh^{-1}} = \Phi(T, \varphi, s)_K \cdot h^{-1},$$

where $\omega(h)\varphi$ is the Schwartz function given by $\omega(h)\varphi(v, w) = \varphi(h^{-1}v, h^{-1}w)$.

Proof. To prove (1), let $(v_0, w_0) \in \Omega_T(F)$ and denote by H_U the pointwise stabilizer in H of the plane spanned by v_0 and w_0. Note that the map $h \mapsto h^{-1} \cdot (v_0, w_0)$

induces a bijection $H_U(\mathbb{A}_f)\backslash H(\mathbb{A}_f) \cong \Omega_T(\mathbb{A}_f)$. Now we use Lemma 3.17 and obtain

$$\begin{aligned}
&\mathrm{pr}^*(\Phi(T,\varphi,s)_K)\\
&\qquad = \sum_{h\in H_U(\mathbb{A}_f)\backslash H(\mathbb{A}_f)/K} \omega(h)\varphi(v_0,w_0)\,\mathrm{pr}^*(\Phi(v_0,w_0,h,s)_K)\\
&\qquad = \sum_{h\in H_U(\mathbb{A}_f)\backslash H(\mathbb{A}_f)/K}\ \sum_{k\in h^{-1}K_{U,h}h\backslash K/K'} \omega(h)\varphi(v_0,w_0)\Phi(v,w,hk,s)_{K'}\\
&\qquad = \sum_{h\in H_U(\mathbb{A}_f)\backslash H(\mathbb{A}_f)/K'} \omega(h)\varphi(v_0,w_0)\Phi(v_0,w_0,h,s)_{K'} = \Phi(T,\varphi,s)_{K'}.
\end{aligned}$$

Part (2) follows directly from part (4) of Lemma 3.16. □

We can also define a weighted version of the currents $[\Phi(v,w,h)_K]$ in (3-43). Namely, for $T \in \mathrm{Sym}_2(F)_{\gg 0}$ and $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ fixed by K as above and ξ_i as in (3-48), let

$$[\Phi(T,\varphi)_K] = \sum_{i=1}^{k} \varphi(\xi_i^{-1}\cdot(v_0,w_0))\cdot[\Phi(v_0,w_0,\xi_i)_K] \in \mathscr{D}^{1,1}(X_K). \tag{3-49}$$

See Section 4B4 for an example. It follows from (1) in Proposition 3.19 that the currents $[\Phi(T,\varphi,s)_K]$ and $[\Phi(T,\varphi)_K]$ are compatible under inclusions $K' \subset K$ and hence one can define

$$\begin{aligned}
[\Phi(T,\varphi,s)] &= ([\Phi(T,\varphi,s)_K])_K \in \mathscr{D}^{1,1}(X) = \varprojlim_K \mathscr{D}^{1,1}(X_K),\\
[\Phi(T,\varphi)] &= ([\Phi(T,\varphi)_K])_K \in \mathscr{D}^{1,1}(X).
\end{aligned} \tag{3-50}$$

Moreover, the space $\mathscr{D}^{1,1}(X)$ carries a natural left action of $H(\mathbb{A}_f)$ induced by the maps $r(h)^{-1}: X_K \to X_{hKh^{-1}}$; we denote the action of $h \in H(\mathbb{A}_f)$ on $\Phi \in \mathscr{D}^{1,1}(X)$ by $\Phi \cdot r(h)^{-1}$. Then, for any $h \in H(\mathbb{A}_f)$, we have

$$\begin{aligned}
[\Phi(T,\omega(h)\varphi,s)] &= [\Phi(T,\varphi,s)]\cdot r(h)^{-1},\\
[\Phi(T,\omega(h)\varphi)] &= [\Phi(T,\varphi)]\cdot r(h)^{-1}.
\end{aligned} \tag{3-51}$$

That is, the assignments $T\otimes\varphi \mapsto [\Phi(T,\varphi,s)]$ and $T\otimes\varphi \mapsto [\Phi(T,\varphi)]$ induce $H(\mathbb{A}_f)$-equivariant linear maps

$$\mathbb{C}[\mathrm{Sym}_2(F)_{>0}]\otimes\mathscr{S}(V(\mathbb{A}_f)^2) \longrightarrow \mathscr{D}^{1,1}(X). \tag{3-52}$$

3H. ***A regularized theta lift.*** From now on and to avoid dealing with metaplectic groups, we will assume that V has even dimension over F. Our next goal is to show that, for $\mathrm{Re}(s) \gg 0$, the form $\Phi(T,\varphi,s)$ can be obtained as a regularized theta lift. More precisely, below we introduce a function $\mathscr{M}_T(g,s)$ defined on a certain subgroup of $\mathrm{Sp}_4(\mathbb{A}_F)$ and a theta function $\theta(g;\varphi)$ that takes values in $\mathscr{A}^{1,1}(X)$. We then define a regularized theta lift $(\mathscr{M}_T(s),\theta(\,\cdot\,;\varphi))^{\mathrm{reg}}$. The main result of this section (Proposition 3.21) shows that the regularized theta lift converges on an open dense subset of X and moreover agrees with $\Phi(T,\varphi,s)$ there. The next two subsections define the functions just mentioned.

3H1. *Schwartz forms.* For $z\in\mathbb{D}$, note that the map $v\mapsto Q(v_{z^\perp})-Q(v_z)$ defines a positive definite quadratic form on V_1. We write

$$\varphi^0(v,z)=e^{-2\pi(Q(v_{z^\perp})-Q(v_z))} \tag{3-53}$$

for the Gaussian associated with z. Note that $\varphi^0(v,z)$ lies in $\mathscr{S}(V_1)\otimes\mathscr{C}^\infty(\mathbb{D})$ and that it is fixed by $H(\mathbb{R})$, i.e., $\varphi^0(hx,hz)=\varphi^0(x,z)$ for every $h\in H(\mathbb{R})$. Now define

$$\varphi^{1,1}(v,w,z)\in[\mathscr{S}(V_1^2)\otimes\mathscr{A}^{1,1}(\mathbb{D})]^{H(\mathbb{R})} \tag{3-54}$$

by

$$\begin{aligned}\varphi^{1,1}(v,w,z)&=\bar\partial(\varphi^0(w,z)\partial\varphi^0(v,z))\\&=\bar\partial\varphi^0(w,z)\wedge\partial\varphi^0(v,z)+\varphi^0(w,z)\bar\partial\partial\varphi^0(v,z).\end{aligned} \tag{3-55}$$

For a quadratic vector space (W,Q) with positive definite quadratic form, let $\varphi^0_+(v,w)\in\mathscr{S}(W^2)$ be the standard Gaussian defined by

$$\varphi^0_+(v,w)=e^{-2\pi(Q(v)+Q(w))}. \tag{3-56}$$

For $v\in V(\mathbb{R})$, denote by v_i, $i=1,\dots,d$ the image of v under the natural map $V(\mathbb{R})\to V\otimes_{F,\sigma_i}\mathbb{R}$. Define

$$\varphi^{1,1}_\infty\in[\mathscr{S}(V(\mathbb{R})^2)\otimes\mathscr{A}^{1,1}(\mathbb{D})]^{H(\mathbb{R})} \tag{3-57}$$

by

$$\varphi^{1,1}_\infty(v,w,z)=\varphi^{1,1}(v_1,w_1,z)\otimes\varphi^0_+(v_2,w_2)\otimes\cdots\otimes\varphi^0_+(v_d,w_d). \tag{3-58}$$

Denote by $\omega=\omega_\psi$ the Weil representation of $\mathrm{Sp}_4(\mathbb{A}_F)$ on $\mathscr{S}(V(\mathbb{A})^2)$ with respect to our fixed character ψ (see, e.g., [Kudla and Rallis 1988] for explicit formulas). For $g=(g_f,g_\infty)\in\mathrm{Sp}_4(\mathbb{A}_F)$, $h\in H(\mathbb{A}_f)$ and $\varphi\in\mathscr{S}(V(\mathbb{A}_f)^2)$ fixed by an open compact subgroup K of $H(\mathbb{A}_f)$, the theta function

$$\theta(g;\varphi)_K=\sum_{(v,w)\in V(F)^2}\omega(g_f)\varphi(v,w)\cdot\omega(g_\infty)\varphi^{1,1}_\infty(v,w) \tag{3-59}$$

defines a (1,1)-form on X_K.

3H2. *Regularized lifts.* Let $\kappa = \frac{n+2}{2}$. For $a \in \mathbb{R}_{>0}$, define

$$W_a(y) = \frac{(4\pi a)^{\kappa-1}}{\Gamma(\kappa-1)} \cdot y^{\kappa/2} e^{-2\pi a y}, \quad y > 0. \tag{3-60}$$

Note that

$$\int_0^\infty W_a(y) y^{\kappa/2} e^{-2\pi a y} \frac{dy}{y^2} = 1. \tag{3-61}$$

Consider the following subgroups of $\mathrm{Sp}_{4,F}$:

$$N(k) = \left\{ n = n(X) = \begin{pmatrix} 1_2 & X \\ & 1_2 \end{pmatrix} \,\middle|\, X = {}^tX \in \mathrm{Sym}_2(k) \right\}, \tag{3-62}$$

$$A(k) = \left\{ a = m(t,v) = \begin{pmatrix} y & & & \\ & t & & \\ & & y^{-1} & \\ & & & t^{-1} \end{pmatrix} \,\middle|\, y, t \in k^\times \right\}. \tag{3-63}$$

Let dn be the unique Haar measure on $N(\mathbb{A})$ such that $\mathrm{Vol}(N(F)\backslash N(\mathbb{A}), dn) = 1$. Denote by $A(\mathbb{R})^0$ the connected component of the identity in $A(\mathbb{R})$. Let da be the measure on $A(\mathbb{R})^0$ defined by

$$\begin{aligned} &\int_{A(\mathbb{R})^0} f(a)da \\ &= \int_{(\mathbb{R}_{>0})^d} \int_{(\mathbb{R}_{>0})^d} f\big(m(y_1^{1/2}, t_1^{1/2}), \dots, m(y_d^{1/2}, t_d^{1/2})\big) \frac{dy_1}{y_1^2} \frac{dt_1}{t_1^2} \cdots \frac{dy_d}{y_d^2} \frac{dt_d}{t_d^2}, \end{aligned} \tag{3-64}$$

where dy_i, dt_i denote the Lebesgue measure.

For a matrix $T \in \mathrm{Sym}_2(F)$, define a character $\psi_T : N(F)\backslash N(\mathbb{A}) \to \mathbb{C}^\times$ by $\psi_T(n(X)) = \psi(\mathrm{tr}(TX))$. For such a symmetric matrix $T = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ and $i = 1, \dots, d$, we write

$$\sigma_i(T) = \begin{pmatrix} a_i & b_i \\ b_i & c_i \end{pmatrix},$$

where $a_i = \sigma_i(a),\ b_i = \sigma_i(b)$ and $c_i = \sigma_i(c)$. We also write $T^\iota = \left(\begin{smallmatrix} c & b \\ b & a \end{smallmatrix}\right)$.

Definition 3.20. For $T = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right) \in \mathrm{Sym}_2(F)$ totally positive definite, the function

$$\mathscr{M}_T(na, s) : N(F)\backslash N(\mathbb{A}) \times A(\mathbb{R})^0 \longrightarrow \mathbb{C}$$

is defined by

$$\begin{aligned} \mathscr{M}_T(nm(y^{1/2}, t^{1/2}), s) = (2\kappa_{\dim(V)}^{-1}) \overline{\psi_T(n)} M_{\sigma_1(T)}(y_1, s) M_{\sigma_1(T)^\iota}(t_1, s) \\ \times (y_1 t_1)^{1-\frac{\kappa}{2}} \prod_{i=2}^d W_{a_i}(y_i) W_{c_i}(t_i), \end{aligned} \tag{3-65}$$

where $\kappa_4 = 2$ and $\kappa_n = 1$ for $n > 5$.

Given a measurable function $f : \mathrm{Sp}_4(\mathbb{A}_F) \to \mathbb{C}$ that satisfies $f(ng) = f(g)$ for all $n \in N(F)$, define

$$(\mathscr{M}_T(s),\ f)^{\mathrm{reg}} = \int_{A(\mathbb{R})^0} \int_{N(F)\backslash N(\mathbb{A})} \mathscr{M}_T(na, s) f(na)\, dn\, da, \tag{3-66}$$

provided that the integral converges.

Proposition 3.21. *Let $T \in \mathrm{Sym}_2(F)$ be a positive definite symmetric matrix, φ be a Schwartz form in $\mathscr{S}(V(\mathbb{A}_f)^2)$ fixed by an open compact subgroup $K \subset H(\mathbb{A}_f)$ and $\theta(g; \varphi)_K$ be the theta function defined in (3-59). Then there is a dense open set $U \subseteq X_K$ with complement of measure zero such that for $\mathrm{Re}(s) \gg 0$, the regularized theta lift*

$$(\mathscr{M}_T(s), \theta(\,\cdot\,; \varphi)_K)^{\mathrm{reg}}$$

converges and equals $\Phi(T, \varphi, s)_K$ on U.

Proof. The sum defining $\theta(na; \varphi)_K$ and the inner integral in $(\mathscr{M}_T(s), \theta(\,\cdot\,, h; \varphi)_K)^{\mathrm{reg}}$ unfold to

$$\int_{A(\mathbb{R})^0} \mathscr{M}_T(a, s) \sum_{(v,w) \in \Omega_T(F)} \varphi(v, w)\omega(a)\varphi_\infty^{1,1}(v, w, z)\, da.$$

Let

$$\tilde{U} = \mathbb{D} - \bigcup_{(v,w) \in \Omega_T(F) \cap \mathrm{Supp}(\varphi)} (\mathbb{D}_v \cup \mathbb{D}_w),$$

so that $\tilde{U}$ is an open dense subset of $\mathbb{D}$ whose complement has measure zero. By Fubini's theorem and Lemma 3.22 below, the sum and the integral can be interchanged whenever $z \in \tilde{U}$; thus the above equals

$$\sum_{(v,w) \in \Omega_T(F)} \varphi(v, w) \int_{A(\mathbb{R})^0} \mathscr{M}_T(a, s)\omega(a)\varphi_\infty^{1,1}(v, w, z)\, da.$$

The integral can be computed using equations (3-23) and (3-61). We obtain

$$\int_{A(\mathbb{R})^0} \mathscr{M}_T(a, s)\omega(a)\varphi_\infty^{1,1}(v, w, z)\, da = 2\omega(v, w, z, s).$$

Assume first that $n > 2$. Then $H_+(\mathbb{Q})$ acts transitively on $\Omega_T(F)$, by [Kudla 1997, Lemma 5.5]; fixing $(v_0, w_0) \in \Omega_T(F)$ we see that for $z \in \tilde{U}$, the integral $(\mathscr{M}_T(s), \theta(\,\cdot\,; \varphi)_K)^{\mathrm{reg}}$ equals

$$I(v_0, w_0, \varphi, s) := \sum_{(v,w) \in H_+(\mathbb{Q}) \cdot (v_0, w_0)} \varphi(v, w) \cdot \omega(v, w, z, s).$$

With h_j, $j = 1, \dots, r$ as in (2-8), we have

$$I(v_0, w_0, \varphi, s)|_{\Gamma_{h_j}\backslash\mathbb{D}^+} = \sum_{(v,w)\in H_+(\mathbb{Q})\cdot(v_0,w_0)} \omega(h_j)\varphi(v, w)\cdot\omega(v, w, z, s).$$

Let ξ_i, $i = 1, \dots, k$ be as in (3-48) and define

$$S_{j,i}(v_0, w_0) = H_+(\mathbb{Q})\cdot(v_0, w_0) \cap h_j K \xi_i^{-1}\cdot(v_0, w_0).$$

Note that $\omega(h_j)\varphi(v, w) = \varphi(\xi_i^{-1}\cdot(v_0, w_0))$ for every $(v, w) \in S_{j,i}(v_0, w_0)$ and

$$H_+(\mathbb{Q})\cdot(v_0, w_0) \cap \mathrm{Supp}(\omega(h_j)\varphi) = \coprod_{i=1}^{k} S_{j,i}(v_0, w_0).$$

Hence

$$I(v_0, w_0, \varphi, s)|_{\Gamma_{h_j}\backslash\mathbb{D}^+} = \sum_{i=1}^{k} \varphi(\xi_i^{-1}\cdot(v_0, w_0)) \sum_{(v,w)\in S_{j,i}(v_0,w_0)} \omega(v, w, z, s).$$

Note that the set $S_{j,i}(v_0, w_0)$ is stable under $\Gamma_{h_j} = H_+(\mathbb{Q}) \cap h_j K h_j^{-1}$, so that we can write

$$\begin{aligned}\sum_{(v,w)\in S_{j,i}(v_0,w_0)} \omega(v, w, z, s) &= \sum_{(v,w)\in \Gamma_{h_j}\backslash S_{j,i}(v_0,w_0)} \ \sum_{\gamma\in(\Gamma_{h_j})_{v,w}\backslash\Gamma_{h_j}} \omega(\gamma^{-1}v, \gamma^{-1}w, z, s) \\ &= \sum_{(v,w)\in \Gamma_{h_j}\backslash S_{j,i}(v_0,w_0)} \Phi(v, w, z, s)_{\Gamma_{h_j}}.\end{aligned}$$

By [Kudla 1997, Lemma 5.7(i)], the set of orbits $\Gamma_{h_j}\backslash S_{j,i}(v_0, w_0)$ is in bijection with the double coset (where we write H_U for H_{v_0,w_0})

$$(H_U)_+(\mathbb{Q})\backslash H_U(\mathbb{A}_f) \cap H_+(\mathbb{Q}) h_j K \xi_i^{-1} / K_{U,\xi_i}.$$

Moreover, the bijection is as follows: Suppose $(v, w) \in S_{j,i}(v_0, w_0)$ is of the form $\gamma\cdot(v_0, w_0) = h_j k \xi^{-1}\cdot(v_0, w_0)$. Then $\Gamma_{h_j}\cdot(v, w)$ corresponds to the double coset $(H_U)_+(\mathbb{Q})\gamma^{-1} h_j k \xi^{-1} K_{U,\xi_i}$. Thus, by definition of $\Phi(v, w, h, s)_K$ we have

$$\sum_{(v,w)\in \Gamma_{h_j}\backslash S_{j,i}(v_0,w_0)} \Phi(v, w, z, s)_\Gamma = \Phi(v_0, w_0, \xi_i, s)_K|_{\Gamma_{h_j}\backslash\mathbb{D}^+},$$

and hence

$$I(v_0, w_0, \varphi, s)|_{\Gamma_{h_j}\backslash\mathbb{D}^+} = \sum_{i=1}^{k} \varphi(\xi_i^{-1}\cdot(v_0, w_0))\cdot\Phi(v_0, w_0, \xi_i, s)_K|_{\Gamma_{h_j}\backslash\mathbb{D}^+}$$

for every j, as was to be shown.

Now assume that $n = 2$. By [Kudla 1997, Lemma 5.5], the group $H_+(\mathbb{Q})$ acts with two orbits on $\Omega_T(F)$, and we have $\Omega_T(F) = H_+(\mathbb{Q})\cdot(v_0, w_0) \amalg H_+(\mathbb{Q})\gamma_0\cdot(v_0, w_0)$

for any $\gamma_0 \in H(\mathbb{Q})$ that fixes the plane U_0 spanned by (v_0, w_0) but reverses its orientation given by the ordered basis $\{v_0, w_0\}$. Thus, for $z \in \tilde{U}$, the integral $(\mathscr{M}_T(s), \theta(\cdot\,; \varphi)_K)^{\mathrm{reg}}$ equals

$$I(v_0, w_0, \varphi, s) + I(\gamma_0 \cdot (v_0, w_0), \varphi, s).$$

Define

$$S_{j,i}(v_0, w_0, \gamma_0) = H_+(\mathbb{Q})\gamma_0 \cdot (v_0, w_0) \cap h_j K \xi_i^{-1} \cdot (v_0, w_0).$$

Then $S_{j,i}(v_0, w_0, \gamma_0)$ is stable under Γ_{h_j} and one shows as above that

$$I(\gamma_0 \cdot (v_0, w_0), \varphi, s)|_{\Gamma_{h_j}\backslash\mathbb{D}^+} = \sum_{i=1}^{k} \varphi(\xi_i^{-1} \cdot (v_0, w_0)) \sum_{(v,w)\in\Gamma_{h_j}\backslash S_{j,i}(v_0,w_0,\gamma_0)} \Phi(v, w, z, s)_{\Gamma_{h_j}}.$$

Note that we can choose γ_0 so that $\gamma_0 \cdot v_0 = v_0$ and $\gamma_0 \cdot w_0 = -w_0$. Since $\omega(v, w, z, s) = \omega(v, -w, z, s)$, we conclude from [Kudla 1997, Lemma 5.7(ii)] that

$$\Phi(v_0, w_0, \xi_i, s)_K|_{\Gamma_{h_j}\backslash\mathbb{D}^+} = \sum_{(v,w)\in\Gamma_{h_j}\backslash S_{j,i}(v_0,w_0)} \Phi(v, w, z, s)_{\Gamma_{h_j}} + \sum_{(v,w)\in\Gamma_{h_j}\backslash S_{j,i}(v_0,w_0,\gamma_0)} \Phi(v, w, z, s)_{\Gamma_{h_j}},$$

and the claim follows from this. $\square$

The next lemma completes the proof of Proposition 3.21.

Lemma 3.22. *Let $T \in \mathrm{Sym}_2(F)$ be totally positive definite, φ be a Schwartz form in $\mathscr{S}(V(\mathbb{A}_f)^2)$ and let*

$$\tilde{U} = \mathbb{D} - \bigcup_{(v,w)\in\Omega_T(F)\cap\mathrm{Supp}(\varphi)} (\mathbb{D}_v \cup \mathbb{D}_w).$$

Then, for $\mathrm{Re}(s) \gg 0$, the sum

$$\sum_{(v,w)\in\Omega_T(F)} |\varphi(v, w)| \int_{A(\mathbb{R})^0} |\mathscr{M}_T(a, s)| \cdot \|\omega(a)\varphi_\infty^{1,1}(v, w, z)\|\, da \tag{3-67}$$

converges for every $z \in \tilde{U}$.

Proof. Let $(v, w) \in \Omega_T(F)$. It is enough to show that, for $\mathrm{Re}(s) \gg 0$ and any $\Gamma \subset H_+(\mathbb{R})$, the sum

$$\sum_{\gamma\in\Gamma_{v,w}\backslash\Gamma} \int_{A(\mathbb{R})^0} |\mathscr{M}_T(a, s)| \cdot \|\omega(a)\varphi_\infty^{1,1}(\gamma^{-1}v, \gamma^{-1}w, z)\|\, da$$

converges for $z \in \mathbb{D}^+ - (\Gamma \cdot \mathbb{D}_v^+ \cup \Gamma \cdot \mathbb{D}_w^+)$, since (3-67) is a finite linear combination of sums of this form. Note that if $\omega(v, z)$ is any of the forms $\partial\varphi^0(v, z)$, $\bar{\partial}\varphi^0(v, z)$ or $\partial\bar{\partial}\varphi^0(v, z)$, then we can write

$$\|\omega(v, z)\| = \sum_i \|P_i(v, z)\| \cdot \varphi^0(v, z),$$

where the sum over i is finite and the functions $P_i(v, z)$ are polynomial functions of v for fixed z satisfying $\|P_i(hv, hz)\| = \|P_i(v, z)\|$ for every $h \in H(\mathbb{R})$. In particular, there exists a positive constant C and a natural number k (in fact, $k = 2$ will do) such that $\|P_i(v, z)\| \leq C\, Q(v_{z^\perp})^k$ for every $z \in \mathbb{D}^+$ and every v of fixed positive norm $Q(v) = m > 0$. Now choose $\epsilon > 0$ such that $|Q(\gamma^{-1}v)_z| > \epsilon$ and $|Q(\gamma^{-1}w)_z| > \epsilon$ for all $\gamma \in \Gamma$. Then there exists a constant $C_\epsilon > 0$ such that

$$\int_{A(\mathbb{R})^0} |\mathscr{M}_T(a, s)| \cdot \|\omega(a)\varphi_\infty^{1,1}(\gamma^{-1}v, \gamma^{-1}w, z)\|\, da \\ < C_\epsilon \big| Q(\gamma^{-1}v)_{z^\perp} \cdot Q(\gamma^{-1}w)_{z^\perp} \big|^{-\frac{s+s_0}{2}+k},$$

and the claim follows as in the proof of Proposition 3.6. □

Theorem 1.1 now follows from Proposition 3.12 and Proposition 3.21.

3I. ***Higher Chow groups and regulators.*** We next focus on the relationship between the currents $\Phi(T, \varphi)$ introduced above and the currents in the image of the regulator map

$$r_{\mathscr{D}} : \mathrm{CH}^2(X_K, 1) \to \mathscr{D}^{1,1}(X_K). \tag{3-68}$$

Let us first recall the definitions of the higher Chow group $\mathrm{CH}^2(X_K, 1)$ and of the above map.

Let Y be an irreducible algebraic variety defined over a field k. The group $\mathrm{CH}^2(Y, 1)$ is defined as a quotient

$$\mathrm{CH}^2(Y, 1) = Z^2(Y, 1)/B^2(Y, 1). \tag{3-69}$$

An element $c \in Z^2(Y, 1)$ is a finite linear combination

$$c = \sum_i a_i \cdot (\pi_i : Z_i \to Y, f_i), \tag{3-70}$$

where Z_i is a normal variety over k of dimension $\dim(Y) - 1$, π_i is a generically finite proper map, f_i is a meromorphic function on Z_i, and $a_i \in \mathbb{Q}$; it is also required that

$$\sum_i a_i \cdot (\pi_i)_*(\operatorname{div} f_i) = 0 \tag{3-71}$$

as a cycle of codimension 2 in Y. For a description of $B^2(Y, 1)$, see [Voisin 2002].

Suppose that $k \subseteq \mathbb{C}$. Define a map

$$\begin{aligned} r_{\mathscr{D}} : \mathrm{CH}^2(Y, 1) &\longrightarrow \mathscr{D}^{1,1}(Y_{\mathbb{C}}) \\ \sum_i a_i \cdot (\pi_i : Z_i \to Y, f_i) &\longmapsto 2\pi i \sum a_i \cdot (\pi_i)_*([\log |f_i|]), \end{aligned} \tag{3-72}$$

where $(\pi_i)_*(\log |f_i|) \in \mathscr{D}^{1,1}(Y_{\mathbb{C}})$ is the current defined by

$$\big((\pi_i)_*(\log |f_i|), \alpha\big) = \int_{Z_i} \pi_i^*(\alpha) \cdot \log |f_i| \tag{3-73}$$

for $\alpha \in \mathscr{A}_c^{2\dim(Y)-2}(Y_{\mathbb{C}})$. The map $r_{\mathscr{D}}$ is known as a regulator map; it is linear and its image defines a rational vector subspace of $\mathscr{D}^{1,1}(Y_{\mathbb{C}})$. Note also that for any $c \in \mathrm{CH}^2(Y, 1)$, the current $r_{\mathscr{D}}(c)$ is dd^c-closed: this follows from the identity of currents

$$dd^c(\pi_i)_*(\log |f_i|^2) = \delta_{\operatorname{div} f_i} \tag{3-74}$$

and condition (3-71).

Note that the currents $[\Phi(T, \varphi)]$ in (3-50) are not dd^c-closed. In fact, for the currents $[\Phi(v, w)_\Gamma]$ in (3-31), we have

$$dd^c[\Phi(v, w)_\Gamma] = \delta_{Z(v,w)_\Gamma} + dd^c G(v, w)_\Gamma \cdot \delta_{X(v)_\Gamma}. \tag{3-75}$$

Here $dd^c G(v, w)_\Gamma$ extends to a smooth 2-form defined on $X(v)_\Gamma$, and the current $dd^c G(v, w)_\Gamma \cdot \delta_{X(v)_\Gamma} \in \mathscr{D}^{1,1}(X_\Gamma)$ is defined by

$$\big(dd^c G(v, w)_\Gamma \cdot \delta_{X(v)_\Gamma}, \alpha\big) = \int_{X(v)_\Gamma} dd^c G(v, w)_\Gamma \wedge \alpha,$$

for $\alpha \in \mathscr{A}_c^{n-1,n-1}(X_\Gamma)$.

Since $\Phi(T, \varphi)$ is not dd^c-closed, it is not in the image of the regulator map defined above. It is natural to ask for necessary and sufficient conditions for a finite linear combination $\sum_{T,\varphi} a(T, \varphi)[\Phi(T, \varphi)]$ with $a(T, \varphi) \in \mathbb{Q}$ to belong to the image of the regulator. The next proposition proves a weak result in this direction when $n \geq 4$. It turns out that in this case being dd^c-closed is also sufficient.

Proposition 3.23. *Assume that $n \geq 4$. Let $\Phi_K = \sum_{T,\varphi} a(T, \varphi)[\Phi(T, \varphi)_K] \in \mathscr{D}^{1,1}(X_K)$, where the sum is finite and $a(T, \varphi) \in \mathbb{Q}$. Then $dd^c \Phi_K = 0$ if and only if $\Phi_K = r_{\mathscr{D}}(c)$ for some $c \in \mathrm{CH}^2(X_K, 1)$.*

Proof. Above we showed that $r_{\mathscr{D}}(c)$ is dd^c-closed for any $c \in \mathrm{CH}^2(X_K, 1)$. Now let $\Phi_K = \sum_{T,\varphi} a(T, \varphi)[\Phi(T, \varphi)_K]$ as in the statement and assume that $dd^c \Phi_K = 0$. We compute

$$0 = dd^c \Phi_K = \sum_{T,\varphi} a(T, \varphi) \cdot (\delta_{Z(T,\varphi)_K} + \Psi(T, \varphi)_K), \tag{3-76}$$

where $\Psi(T,\varphi)_K \in \mathscr{D}^{1,1}(X_K)$ is a current whose support is a finite union of special divisors on X_K. More precisely, we have

$$\Psi(T,\varphi)_K = \sum_i dd^c G_i \cdot \delta_{X(v_i,h_i)_K},$$

where the sum is finite and G_i is a finite linear combination of Green functions of the form (3-6) on $X(v_i,h_i)_K$ with logarithmic singularities on special divisors. Since the currents $\Psi(T,\varphi)_K$ and $\delta_{Z(T,\varphi)_K}$ are supported in different codimensions, it follows from (3-76) that

$$\sum_{T,\varphi} a(T,\varphi)\cdot \delta_{Z(T,\varphi)_K} = 0, \tag{3-77}$$

$$\sum_{T,\varphi} a(T,\varphi)\cdot \Psi(T,\varphi)_K = 0. \tag{3-78}$$

(To see this, pick a basis of open neighborhoods $(U_j)_{j\geq 1}$ of $\bigcup_{T,\varphi} Z(T,\varphi)_K$ and compactly supported smooth functions $\phi_j : U_j \to [0,1]$ such that $\phi_j|_{Z(T,\varphi)_K} \equiv 1$. For $\alpha \in \mathscr{A}_c^{n-2,n-2}(X_K)$, evaluate (3-76) on the sequence $(\phi_j\alpha)_{j\geq 1}$ and apply dominated convergence on each $X(v_i,h_i)_K$.) Write

$$\sum_{T,\varphi} a(T,\varphi)[\Phi(T,\varphi)_K] = \sum_i G_i \delta_{X(v_i,h_i)_K},$$

where the sum over i is finite and G_i is a Green function on $X(v_i,h_i)_K$. Now Equation (3-78) implies that the summand corresponding to a connected special divisor $X(v)_\Gamma$ in this sum is of the form $G(v,\Gamma)\delta_{X(v)_\Gamma}$, where $G(v,\Gamma)$ is a Green function on $X(v)_\Gamma$ that satisfies $dd^c G(v,\Gamma) = 0$. Since $n \geq 4$, we have $H^1(X(v)_\Gamma, \mathbb{C}) = 0$ (see [Vogan and Zuckerman 1984, Theorem 8.1]) and it follows that $G(v,\Gamma) = a(v,\Gamma)\log|f_{v,\Gamma}|$ for some meromorphic function $f_{v,\Gamma} \in k(X(v)_\Gamma)^\times$ and some $a(v,\Gamma) \in \mathbb{Q}$. Thus, denoting by $\pi_{v,\Gamma}$ the map $X(v)_\Gamma \to X_K$, we find that $\Phi_K = \sum_{v,\Gamma} a(v,\Gamma)\cdot(\pi_{v,\Gamma})_*(\log|f_{v,\Gamma}|)$, where the sum is finite. Consider now the formal sum $\sum a(v,\Gamma)\cdot(\pi_{v,\Gamma}, f_{v,\Gamma})$. By Equation (3-77), we have $\sum_{v,\Gamma} a(v,\Gamma)\cdot(\pi_{v,\Gamma})_*(\operatorname{div} f_{v,\Gamma}) = 0$ and hence it defines an element $c \in \mathrm{CH}^2(X_K,1)$ satisfying $r_{\mathscr{D}}(c) = \Phi_K$. □

3J. *Evaluating currents on differential forms.* Let $\alpha \in \mathscr{A}_c^{n-1,n-1}(X_K)$ be a compactly supported form. Since Proposition 3.21 shows that the forms $\Phi(T,\varphi,s)_K$ are theta lifts, one can try to evaluate

$$[\Phi(T,\varphi,s)_K](\alpha) = \int_{X_K} \Phi(T,\varphi,s)_K \wedge \alpha$$

by interchanging the integrals. However, this interchange is not justified since the resulting integrals are not absolutely convergent. In this section, we will introduce certain currents $[\tilde{\Phi}(T,\varphi,s)]$ closely related to the $[\Phi(T,\varphi,s)]$. These currents will be meromorphic in $s\in\mathbb{C}$ (modulo $\mathrm{Im}(\partial)+\mathrm{Im}(\bar{\partial})$ as before) and we will show that their constant term at $s=s_0$ is a certain $\mathbb{Q}$-linear combination of the $[\Phi(T,\varphi)]$. Moreover, following ideas in [Bruinier and Funke 2004], we will give an expression of these currents as regularized theta lifts that allows us to evaluate them by interchanging the integrals (see Proposition 3.27).

For a pair of vectors $(v,w)\in V(F)^2$ spanning a totally positive definite plane, consider the $(1,1)$-form

$$\tilde{\omega}(v,w,z,s)=\phi(v,w,z,s)\partial\bar{\partial}\phi(w,v,z,s) \tag{3-79}$$

in $\mathscr{A}^{1,1}(\mathbb{D}-(\mathbb{D}_v\cup\mathbb{D}_w))$. The form $\tilde{\omega}(v,w,z,s)$ is related to the form $\omega(v,w,z,s)$ as follows:

$$\begin{aligned}\omega(v,w,z,s)&+\overline{\omega(w,v,z,s)}\\ &=\bar{\partial}\phi(w,v,z,s)\wedge\partial\phi(v,w,z,s)+\phi(w,v,z,s)\bar{\partial}\partial\phi(v,w,z,s)\\ &\qquad+\partial\phi(v,w,z,s)\wedge\bar{\partial}\phi(w,v,z,s)+\phi(v,w,z,s)\partial\bar{\partial}\phi(w,v,z,s)\\ &\qquad\qquad=\tilde{\omega}(v,w,z,s)-\tilde{\omega}(w,v,z,s).\end{aligned} \tag{3-80}$$

For $\Gamma\subset H_+(\mathbb{R})$, define a $(1,1)$-form on X_Γ by

$$\tilde{\Phi}(v,w,z,s)_\Gamma=\sum_{\gamma\in\Gamma_{v,w}\backslash\Gamma}\tilde{\omega}(\gamma^{-1}v,\gamma^{-1}w,z,s). \tag{3-81}$$

The proofs of Propositions 3.6 and 3.9 apply to this sum and show that it converges normally on $X_\Gamma-(X(v)_\Gamma\cup X(w)_\Gamma)$ and defines a locally integrable $(1,1)$-form on X_Γ. We define forms $\tilde{\Phi}(v,w,h,s)_K$ and $\tilde{\Phi}(T,\varphi,s)_K$ as in Section 3F and Section 3G by replacing $\omega(v,w,z,s)$ with $\tilde{\omega}(v,w,z,s)$ throughout. As before, denote by $[\tilde{\Phi}(T,\varphi,s)_K]$ the current in $\mathscr{D}^{1,1}(X_K)$ corresponding to the form $\tilde{\Phi}(T,\varphi,s)_K$. The proof of Proposition 3.19 shows that the currents $[\tilde{\Phi}(T,\varphi,s)_K]$ for varying K form a compatible system under the maps induced by inclusions $K'\subset K$, so that we obtain a current $[\tilde{\Phi}(T,\varphi,s)]\in\mathscr{D}^{1,1}(X)=\varprojlim_K\mathscr{D}^{1,1}(X_K)$.

Let us now describe the relation of the currents $[\tilde{\Phi}(T,\varphi,s)]$ with the currents $[\Phi(T,\varphi)]$. For $T=\left(\begin{smallmatrix}a&b\\b&c\end{smallmatrix}\right)\in\mathrm{Sym}_2(F)_{>0}$ and $\varphi\in\mathscr{S}(V(\mathbb{A}_f)^2)$, define

$$T^\iota=\begin{pmatrix}c&b\\b&a\end{pmatrix},\qquad\varphi^\iota(v,w)=\varphi(w,v). \tag{3-82}$$

Then it follows from Proposition 3.12 and Equation (3-80) that the image of the current $[\tilde{\Phi}(T,\varphi,s)]-[\tilde{\Phi}(T^\iota,\varphi^\iota,s)]$ in $\tilde{\mathscr{D}}^{1,1}(X)$ admits meromorphic continuation to $s\in\mathbb{C}$ and that its constant term at $s=s_0=(n-1)/2$ is given by

$$\mathrm{CT}_{s=s_0}[\tilde{\Phi}(T,\varphi,s)]-[\tilde{\Phi}(T^\iota,\varphi^\iota,s)]\equiv[\Phi(T,\varphi)]-[\Phi(T^\iota,\varphi^\iota)], \tag{3-83}$$

where $\equiv$ denotes equality of currents modulo $\partial+\bar{\partial}$. See Section 4B5 for an example of a current of this form.

Remark 3.24. From the point of view of regulator maps $r_{\mathscr{D},K} : \mathrm{CH}^2(X_K, 1) \to \mathscr{D}^{1,1}(X_K)$, the currents on the right hand side of Equation (3-83) are quite natural objects. Namely, let $[\Phi] \in \mathscr{D}^{1,1}(X_\Gamma)$ be any current of the form $[\Phi] = r_{\mathscr{D}}(c)$ with $c = \sum n_i(C_i, f_i) \in \mathrm{CH}^2(X_\Gamma, 1)$ (see Section 3I for definitions) such that the C_i are special divisors and the $f_i \in k(C_i)^\times \otimes \mathbb{Q}$ are (pushforwards of) the meromorphic functions constructed by Bruinier [2012, Theorem 6.8]. Then condition (3-71) implies that $[\Phi]$ is a linear combination with $\mathbb{Q}$-coefficients of currents $[\Phi(v, w)_\Gamma] - [\Phi(w, v)_\Gamma]$ for some pairs $(v, w) \in V(F)^2$. The current $[\Phi(T, \varphi)] - [\Phi(T^\iota, \varphi^\iota)]$ is just a finite sum of such currents, weighted by the values of φ.

Our next goal is to obtain an expression of $\tilde{\Phi}(T, \varphi, s)_K$ as a regularized theta lift with good convergence properties. To do so, we will use a relation between $\partial\bar{\partial}\varphi^0(v, z)$ and $\varphi_{KM}(v, z)$ established by Bruinier and Funke.

Denote by

$$\varphi_{KM} \in [\mathscr{S}(V_1) \otimes \mathscr{A}^{1,1}(\mathbb{D})]^{H(\mathbb{R})} \tag{3-84}$$

the $\mathscr{S}(V_1)$-valued, closed (1,1)-form constructed by Kudla and Millson [1986]. We have

$$\varphi_{KM}(v, z) = P(v, z)\varphi^0(v, z), \tag{3-85}$$

where $P(v, z) \in [\mathscr{C}^\infty(V_1) \otimes \mathscr{A}^{1,1}(\mathbb{D})]^{H(\mathbb{R})}$ is, for fixed z, a polynomial in v of degree 2 (see [Kudla 1997, (7.16)] for an explicit description of $P(v, z)$; our φ_{KM} is denoted $\varphi^{(1)}$ there).

Let $\tau = x + iy$ be an element of the upper half plane and let $g_\tau = \left(\begin{smallmatrix} y^{1/2} & xy^{-1/2} \\ 0 & y^{-1/2} \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$. Define

$$\varphi^0(v, \tau, z) = y^{-(n-2)/4}\omega(g_\tau)\varphi^0(v, z) = y \cdot e\big(Q(v_{z^\perp})\tau + Q(v_z)\bar{\tau}\big), \tag{3-86}$$

$$\varphi_{KM}(v, \tau, z) = y^{-(n+2)/4}\omega(g_\tau)\varphi_{KM}(v, z). \tag{3-87}$$

Here ω denotes the Weil representation of $\mathrm{SL}_2(\mathbb{R})$ on $\mathscr{S}(V_1)$ and $e(x) = e^{2\pi i x}$. Our presentation of $[\tilde{\Phi}(T, \varphi, s)_K]$ as a regularized theta lift will use the following result.

Proposition 3.25 [Bruinier and Funke 2004, Theorem 4.4]. *Let $L = -2i\,\mathrm{Im}(\tau)^2 \dfrac{\partial}{\partial\bar{\tau}}$ be the Maass lowering operator. Then*

$$dd^c\varphi^0(v, \tau, z) = -L\varphi_{KM}(v, \tau, z) \tag{3-88}$$

where d and $d^c = (4\pi i)^{-1}(\partial - \bar{\partial})$ are the usual differential operators on $\mathbb{D}$.

Using this result, we can find a different expression for the form $\bar{\partial}\partial\phi(v, w, z, s)$. Let L be the lowering operator in Proposition 3.25. For a symmetric positive definite matrix $T = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ and $\tau = x + iy \in \mathbb{H}$, define

$$\widetilde{M}_T(\tau, s) = 4\pi y^2 \frac{\partial}{\partial \bar{\tau}} (M_T(y, s) e^{-2\pi i a x}). \tag{3-89}$$

One computes

$$\widetilde{M}_T(\tau, s) = \widetilde{C}(T, s) y^{1-k/2} M_{1-k/2, s/2}\left(\left|\frac{4\pi \det T}{c} y\right|\right) e^{\frac{2\pi b^2}{c} y} e^{-2\pi i a x}, \tag{3-90}$$

with $\widetilde{C}(T, s) = \pi i C(T, s) \cdot (s + s_0)$.

Lemma 3.26. *For* $v, w \in \Omega_T(V_1)$ *and* $\mathrm{Re}(s) \gg 0$, *we have*

$$\bar{\partial}\partial\phi(v, w, z, s) = \int_0^\infty \widetilde{M}_T(y, s) \varphi_{KM}(v, y, z) \frac{dy}{y^2}.$$

Proof. Recall the integral expression for $\phi(v, w, z, s)$ given in (3-23). In terms of $\varphi^0(v, \tau, z)$, we have

$$\begin{aligned} \phi(v, w, z, s) &= \int_0^\infty M_T(y, s) \varphi^0(v, y, z) \frac{dy}{y^2} \\ &= \int_0^\infty \int_0^1 M_T(y, s) e^{-2\pi i Q(v) x} \varphi^0(v, \tau, z) \frac{dx\, dy}{y^2}. \end{aligned}$$

Using (3-88), we obtain

$$\begin{aligned} dd^c \phi(v, w, z, s) &= \int_0^\infty \int_0^1 M_T(y, s) e^{-2\pi i Q(v) x} dd^c \varphi^0(v, \tau, z) \frac{dx\, dy}{y^2} \\ &= -\int_0^\infty \int_0^1 M_T(y, s) e^{-2\pi i Q(v) x} L\varphi_{KM}(v, \tau, z) \frac{dx\, dy}{y^2} \\ &= -\int_0^\infty \int_0^1 M_T(y, s) e^{-2\pi i Q(v) x}\, \bar{\partial}(\varphi_{KM}(v, \tau, z)\, d\tau) \\ &= -\lim_{N \to \infty} \int_{\mathscr{F}_N} M_T(y, s) e^{-2\pi i Q(v) x}\, \bar{\partial}(\varphi_{KM}(v, \tau, z)\, d\tau), \end{aligned}$$

where $\mathscr{F}_N = [0, 1] \times [N^{-1}, N] \subset \mathbb{H}$. Applying Stokes's Theorem, we find

$$\begin{aligned} dd^c \phi(v, w, z, s) = \int_0^\infty \int_0^1 \bar{\partial}(M_T(y, s) e^{-2\pi i Q(v) x}) \wedge \varphi_{KM}(v, \tau, z)\, d\tau \\ - \lim_{N \to \infty} \big(M_T(N, s) \varphi_{KM}(v, N, z) - M_T(N^{-1}, s) \varphi_{KM}(v, N^{-1}, z)\big). \end{aligned}$$

Since $dd^c = -(2\pi i)^{-1} \partial\bar{\partial}$, we see that to establish the claim it suffices to show that the second term in the right hand side vanishes. This follows for $z \notin \mathbb{D}_v$ from the asymptotic behavior of $M_T(y, s)$ given by (3-17) and (3-18). □

We can now express $\tilde{\Phi}(T,\varphi,s)_K$ as a regularized theta lift. Namely, for $T = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right) \in \mathrm{Sym}_2(F)$ totally positive definite, define a function

$$\widetilde{\mathscr{M}}_T(na,s) : N(F)\backslash N(\mathbb{A}) \times A(\mathbb{R})^0 \longrightarrow \mathbb{C}$$

by

$$\widetilde{\mathscr{M}}_T(nm(y^{1/2},t^{1/2}),s) = 2\kappa_{\dim(V)}^{-1}\overline{\psi_T(n)}M_{\sigma_1(T)}(y_1,s)\widetilde{M}_{\sigma_1(T)^\iota}(t_1,s) \\ \times y_1^{1-\kappa/2}t_1^{-\kappa/2}\prod_{i=2}^{d} W_{a_i}(y_i)W_{c_i}(t_i). \quad (3\text{-}91)$$

We also need to specify a Schwartz form

$$\tilde{\varphi}_\infty \in [\mathscr{S}(V(\mathbb{R})^2)\otimes\mathscr{A}^{1,1}(\mathbb{D})]^{H(\mathbb{R})}$$

to define the regularized theta lift. Define

$$\tilde{\varphi}^{1,1}(v,w,z) = \varphi^0(v,z)\cdot\varphi_{KM}(w,z) \in \mathscr{S}(V_1^2)\otimes\mathscr{A}^{1,1}(\mathbb{D}) \quad (3\text{-}92)$$

and

$$\tilde{\varphi}_\infty(v,w,z) = \tilde{\varphi}^{1,1}(v_1,w_1,z)\otimes\varphi_+^0(v_2,w_2)\otimes\cdots\otimes\varphi_+^0(v_d,w_d), \quad (3\text{-}93)$$

so that for every $g \in \mathrm{Sp}_4(\mathbb{A}_F)$ and $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ fixed by K, the theta function

$$\theta(g;\varphi\otimes\tilde{\varphi}_\infty)_K = \sum_{(v,w)\in V(F)^2}\omega(g_f)\varphi(v,w)\cdot\omega(g_\infty)\tilde{\varphi}_\infty(v,w) \quad (3\text{-}94)$$

defines a (1,1)-form on X_K. Given a measurable function $f : \mathrm{Sp}_4(\mathbb{A}_F) \to \mathbb{C}$ that satisfies $f(ng) = f(g)$ for all $n \in N(F)$, define

$$(\widetilde{\mathscr{M}}_T(s), f)^{\mathrm{reg}} = \int_{A(\mathbb{R})^0}\int_{N(F)\backslash N(\mathbb{A})}\widetilde{\mathscr{M}}_T(na,s)f(na)\,dn\,da, \quad (3\text{-}95)$$

provided that the integral converges. Then we have the identity

$$\tilde{\Phi}(T,\varphi,s)_K = (\widetilde{\mathscr{M}}_T(s), \theta(\,\cdot\,;\varphi\otimes\tilde{\varphi}_\infty)_K)^{\mathrm{reg}}, \quad (3\text{-}96)$$

valid in an open set $U \subset X_K$ whose complement has measure zero. This is proved in the same way as Proposition 3.21.

Here is the desired result that shows that one can evaluate $[\tilde{\Phi}(T,\varphi,s)]$ by interchanging the order of integration.

Proposition 3.27. *Let $K \subset H(\mathbb{A}_f)$ be an open compact subgroup that fixes φ and let $\alpha \in \mathscr{A}_c^{n-1,n-1}(X_K)$. Then, for* $\mathrm{Re}(s) \gg 0$, *we have*

$$([\tilde{\Phi}(T,\varphi,s)_K],\alpha) = \int_{A(\mathbb{R})^0}\int_{N(F)\backslash N(\mathbb{A})}\widetilde{\mathscr{M}}_T(na,s)\int_{X_K}\theta(na;\varphi\otimes\tilde{\varphi}_\infty)_K\wedge\alpha\,dn\,da.$$

Proof. Performing the integration over $N(F)\backslash N(\mathbb{A})$, we find

$$([\tilde{\Phi}(T,\varphi,s)_K],\alpha)=\int_{X_K}\int_{A(\mathbb{R})^0}\widetilde{\mathscr{M}}_T(a,s)\sum_{(v,w)\in\Omega_T(F)}\varphi(v,w)\cdot\omega(a)\tilde{\varphi}_\infty(v,w,z)\wedge\alpha,$$

and we need to prove that this expression is absolutely convergent. Since K has only finitely many orbits on the support of φ, it suffices to show that

$$\int_{\Gamma_{v,w}\backslash\mathbb{D}^+}\int_{A(\mathbb{R})^0}\widetilde{\mathscr{M}}_T(a,s)\cdot\omega(a)\tilde{\varphi}_\infty(v,w,z)\wedge\eta$$

is absolutely convergent, for any vectors $v,w\in\Omega_T(F)$ and any compactly supported form $\eta\in\mathscr{A}_c^{n-1,n-1}(\Gamma\backslash\mathbb{D}^+)$. This will follow if we can show that

$$\int_{\Gamma_{v,w}\backslash\mathbb{D}^+}\int_{A(\mathbb{R})^0}|\widetilde{\mathscr{M}}_T(a,s)|\cdot\|\omega(a)\tilde{\varphi}_\infty(v,w,z)\|\,da\,d\mu(z)<\infty,$$

that is, we need to show that the inner integral in this expression yields an integrable function on $\Gamma_{v,w}\backslash\mathbb{D}^+$. Denote this inner integral by $f(v,w,z,s)$. Note that

$$\|\tilde{\varphi}_\infty(v,w,z)\|=\sum_i\|P_i(w,z)\|\cdot\varphi^0(v,z)\varphi^0(w,z),$$

where the sum over i is finite and, for fixed z, the $P_i(w,z)$ are polynomials in w (valued in differential forms). These polynomials satisfy $\|P_i(hw,hz)\|=\|P_i(w,z)\|$ for all $h\in H(\mathbb{R})$ and have degree 2; see [Kudla 1997, (7.16)]. Hence

$$\|\tilde{\varphi}_\infty(v,w,z)\|<C\cdot Q(w_{z^\perp})\cdot\varphi^0(v,z)\varphi^0(w,z),$$

for some constant $C>0$. Using this estimate, we find that

$$\begin{aligned}
f(v,w,z,s)&=O\big(Q(v_{z^\perp})^{-\frac{s+s_0}{2}}\cdot Q(w_{z^\perp})^{-\frac{s+s_0}{2}}\big) &&\text{if } |Q(v_z)|,|Q(w_z)|>\epsilon>0,\\
f(v,w,z,s)&=O\big(\log(|Q(v_z)|)\cdot|Q(w_{z^\perp})|^{-\frac{s+s_0}{2}+1}\big) &&\text{as } |Q(v_z)|\to 0,\\
f(v,w,z,s)&=O\big(\log(|Q(w_z)|)\cdot|Q(v_{z^\perp})|^{-\frac{s+s_0}{2}}\big) &&\text{as } |Q(w_z)|\to 0.
\end{aligned}$$

Since $f(v,w,h'z,s)=f(v,w,z,s)$ for $h'\in H'(\mathbb{R})=(H_v)_+(\mathbb{R})\cap(H_w)_+(\mathbb{R})$ and the quotient $\Gamma_{v,w}\backslash\mathbb{D}^+_{v,w}$ has finite volume, the claim follows from these estimates by Lemma 3.8 applied to $H'(\mathbb{R})\backslash\mathbb{D}^+$. □

Corollary 3.28. *Let $K\subset H(\mathbb{A}_f)$ be an open compact subgroup that fixes φ and let $\alpha\in\mathscr{A}_c^{n-1,n-1}(X_K)$ be a closed form. For $g\in\mathrm{Sp}_4(\mathbb{A}_F)$, write*

$$\theta(g;\varphi,\alpha)=\int_{X_K}\theta(g;\varphi\otimes\tilde{\varphi}_\infty)\wedge\alpha.$$

Then

$$([\Phi(T,\varphi)_K]-[\Phi(T^\iota,\varphi^\iota)_K],\alpha)$$
$$=\mathrm{CT}_{s=(n-1)/2}[(\widetilde{\mathscr{M}}_T(s),\theta(\,\cdot\,;\varphi,\alpha))^{\mathrm{reg}}-(\widetilde{\mathscr{M}}_{T^\iota}(s),\theta(\,\cdot\,;\varphi^\iota,\alpha))^{\mathrm{reg}}].$$

Proof. This follows from (3-83) and Proposition 3.27. □

4. An example: products of Shimura curves

The goal of this section is to illustrate the main constructions and results above in one of the simplest cases: when the Shimura variety attached to $\mathrm{GSpin}(V)$ is a product of Shimura curves attached to a quaternion algebra B over $\mathbb{Q}$. In this case, the currents in Section 3 can be described in the more familiar language of Hecke correspondences and CM points. We give this description in Section 4B.

Throughout this section, we fix an indefinite quaternion algebra B over $\mathbb{Q}$; we assume that $B\not\cong M_2(\mathbb{Q})$. We write S for the set of places where B ramifies and $d(B)$ for the discriminant of B. Denote by $n: B\to F$ the reduced norm and let V be B endowed with the quadratic form given by $Q(v)=n(v)$. Then (V,Q) is a nondegenerate quadratic space over $\mathbb{Q}$ with signature $(2,2)$ and $\chi_V=1$.

4A. ***Quaternion algebras and Shimura curves.*** The group $H=\mathrm{GSpin}(V)$ can in this case be described more concretely. Namely, consider $B^\times$ as an algebraic group over $\mathbb{Q}$ defined by

$$B^\times(R)=(B\otimes_{\mathbb{Q}}R)^\times \tag{4-1}$$

for any $\mathbb{Q}$-algebra R and let

$$B^\times\times_{\mathrm{GL}_1}B^\times=\{(g_1,g_2)\in B^\times\times B^\times \mid n(g_1)=n(g_2)\}. \tag{4-2}$$

The group $B^\times\times B^\times$ acts on V by $(g_1,g_2)\cdot x=g_1xg_2^{-1}$. This induces an exact sequence

$$1\longrightarrow \mathbb{G}_m\longrightarrow B^\times\times_{\mathrm{GL}_1}B^\times\longrightarrow \mathrm{SO}(V)\longrightarrow 1 \tag{4-3}$$

showing that

$$\mathrm{SO}(V)\cong\mathbb{G}_m\backslash(B^\times\times_{\mathrm{GL}_1}B^\times),\qquad \mathrm{GSO}(V)\cong\mathbb{G}_m\backslash(B^\times\times B^\times) \tag{4-4}$$

and in fact one has

$$H\cong B^\times\times_{\mathrm{GL}_1}B^\times. \tag{4-5}$$

The theory in Section 2 applies to this case. If we denote by $\mathbb{H}$ the Poincaré upper half plane, we have

$$\mathbb{D}^+\cong\mathbb{H}\times\mathbb{H}. \tag{4-6}$$

Fix once and for all an isomorphism $\iota : B \otimes_{\mathbb{Q}} \mathbb{A}^S \cong M_2(\mathbb{A}^S)$. For $p \in S$, denote by $\mathscr{O}_{B,p}$ the maximal order of $B \otimes_{\mathbb{Q}} \mathbb{Q}_p$. Let

$$\hat{\mathscr{O}}_B = \iota^{-1}\biggl(M_2\biggl(\prod_{p\notin S}\mathbb{Z}_p\biggr)\biggr) \times \prod_{p\in S}\mathscr{O}_{B,p}, \quad K_B = \hat{\mathscr{O}}_B^\times. \tag{4-7}$$

Then $\hat{\mathscr{O}}_B$ is a maximal order of $B \otimes_{\mathbb{Q}} \mathbb{A}_f$ and K_B is a maximal compact subgroup of $B(\mathbb{A}_f)^\times$.

Define the (full level) Shimura curve attached to B to be

$$X_{B,K} = B^\times(\mathbb{Q})\backslash(\mathbb{H}^\pm \times B^\times(\mathbb{A}_f))/K. \tag{4-8}$$

Then $X_{B,K}$ is the set of complex points of a complete curve C_K defined over $\mathbb{Q}$. Let $K = (K_B \times K_B) \cap H(\mathbb{A}_f)$ and define the (full level) Shimura variety

$$X_K = H(\mathbb{Q})\backslash(\mathbb{D} \times H(\mathbb{A}_f))/K. \tag{4-9}$$

Thus $X_{B,K}$ is the set of complex points of the surface $C_K \times C_K$. By (2-8), the surface $X_{B,K}$ is connected.

Given $v \in V$ of positive norm and denoting by $W \subset V$ its orthogonal complement, we have

$$H_v = \mathrm{GSpin}(W) \cong B^\times \tag{4-10}$$

as algebraic groups over $\mathbb{Q}$. The special divisors $Z(v,h)_K$ are hence given by embedded Shimura curves in X_K.

4B. ***Examples of* (1,1)*-currents.*** Let us give some explicit examples of the currents introduced in Section 3 in the case when X_K is a product of Shimura curves, in the more classical language of Hecke correspondences and CM points. Assume that $F = \mathbb{Q}$ for simplicity and denote by $d(B) = p_1 \cdots p_{2r}$ the discriminant of B. Let $\hat{\mathscr{O}}_B$ and $K_B = \hat{\mathscr{O}}_B^\times$ be as in (4-7) and let $K = (K_B \times K_B) \cap H(\mathbb{A}_f)$. Then $\mathscr{O}_B = B \cap \hat{\mathscr{O}}_B$ is a maximal order in B. Denote by $\mathscr{O}_B^1 \subset \mathscr{O}_B^\times$ be the subgroup of units of reduced norm 1. The group $\mathscr{O}_B^1$ acts on $\mathbb{H}$ through the embedding $\iota_\infty : \mathscr{O}_B^1 \to \mathrm{SL}_2(\mathbb{R})$ and we conclude that

$$X_{B,K} \cong \mathscr{O}_B^1\backslash\mathbb{H} =: X_0^B \tag{4-11}$$

is the full level Shimura curve X_0^B and that $X_K = X_0^B \times X_0^B$.

4B1. *Special divisors.* Consider the vector $v_1 = 1 \in B = V$ of norm 1. Then the inclusion $H_{v_1} \subset H$ corresponds to the diagonal embedding $\Delta : B^\times \to B^\times \times_{\mathrm{GL}_1} B^\times$ and hence the map $i_{v_1,1,K} : X(v_1)_K \to X_K$ defined in (2-15) is just the diagonal

$$\Delta : X_0^B \longrightarrow X_0^B \times X_0^B. \tag{4-12}$$

More generally, suppose $v \in \mathscr{O}_B$ has reduced norm d and consider the map $i_{v,1,K} : X(v)_K \to X_K$. If d equals a prime $p \nmid d(B)$ then the intersection $H_v(\mathbb{Q}) \cap K$ is an Eichler order $\mathscr{O}_B(p)$ of level p in B and the map $i_{v,1,K} : X(v)_K \to X_K$ equals the map

$$X_0^B(p) \longrightarrow X_0^B \times X_0^B \tag{4-13}$$

whose image is the Hecke correspondence $T(p)$. Similarly, if d is a divisor of $d(B)$, we obtain the graph of the Atkin–Lehner involution w_d.

4B2. *Currents for connected cycles: $G(v,w)_\Gamma$ and $[\Phi(v,w)_\Gamma]$.* Consider now $v, w \in B$ spanning a positive definite plane. To simplify matters, let us assume that $v = 1$ and that $w \in \mathscr{O}_B$ is such that $R := \mathbb{Z}[w]$ is the full ring of integers of an imaginary quadratic field $L = R \otimes_{\mathbb{Z}} \mathbb{Q}$; such an R is then automatically optimally embedded in $\mathscr{O}_B$ (recall that an embedding $j : R \hookrightarrow \mathscr{O}_B$ is said to be optimal if $j(L) \cap \mathscr{O}_B = j(R)$). The diagram (1-3) in this case becomes

$$\begin{tikzcd}
\{\tau_{p_{v^\perp}(w)}\} \arrow[r] \arrow[d] & \mathbb{H} \arrow[r, "\Delta"] \arrow[d, "\mathrm{pr}"] & \mathbb{H} \times \mathbb{H} \arrow[d, "\mathrm{pr} \times \mathrm{pr}"] \\
\{P_w := \mathrm{pr}(\tau_{p_{v^\perp}(w)})\} \arrow[r] & X_0^B \arrow[r, "\Delta"] & X_0^B \times X_0^B
\end{tikzcd}$$

and $P_w \in X_0^B$ is a point with CM by R (for one of the two CM-types of R). The function $G(v,w)_\Gamma \in \mathscr{C}^\infty(X_0^B - \{P_w\})$ defined by (3-9) is a Green function for the divisor $[P_w] \in \mathrm{Div}(X_0^B)$; we denote this function by $G_{[P_w]}$ and the associated current in $\mathscr{D}^{0,0}(X_0^B)$ by $[G_{[P_w]}]$. The current $[\Phi(v,w)_\Gamma]$ in (3-33) is given by

$$[\Phi(v,w)_\Gamma] = 2\pi i \cdot \Delta_*([G_{[P_w]}]), \tag{4-14}$$

so that for $\alpha \in \mathscr{A}^{1,1}(X_0^B \times X_0^B)$ we have

$$[\Phi(v,w)_\Gamma](\alpha) = 2\pi i \int_{X_0^B} G_{[P_w]} \cdot \Delta^*(\alpha). \tag{4-15}$$

4B3. *The current $[\Phi(v,w,1)_K]$.* Our next goal is to write down an explicit example of the current $[\Phi(v,w,1)_K]$ in (3-43). We have

$$H_{v,w} = \mathrm{GSpin}(\mathbb{Q}\langle v,w\rangle) = L^\times \tag{4-16}$$

as an algebraic group over $\mathbb{Q}$. The embeddings $H_{v,w} \to H_v \to H$ correspond to embeddings of algebraic groups

$$L^\times \longrightarrow B^\times \xrightarrow{\Delta} B^\times \times_{\mathrm{GL}_1} B^\times, \tag{4-17}$$

defined over $\mathbb{Q}$, where the second embedding is just the diagonal. Note that $H_{v,w}(\mathbb{R}) = (K \otimes_{\mathbb{Q}} \mathbb{R})^\times = \mathbb{C}^\times$ with spinor norm the usual norm on $\mathbb{C}$. In particular, every element of this group has positive spinor norm and hence $(H_{v,w})_+(\mathbb{R}) = H_{v,w}(\mathbb{R})$ and $(H_{v,w})_+(\mathbb{Q}) = H_{v,w}(\mathbb{Q}) = L^\times$. Moreover, since $R \to \mathscr{O}$ is optimal, we have

$$(H_{v,w})_+(\mathbb{Q})\backslash H_{v,w}(\mathbb{A}_f)/K_U \cong L^\times\backslash\mathbb{A}_{L,f}^\times/\hat{\mathscr{O}}_L^\times = \mathrm{Pic}(\mathscr{O}_L). \tag{4-18}$$

Let $\{h'_i \mid i = 1, \ldots, s\}$ be a set of representatives for this double coset and write $h'_i = \gamma_i k_i$ with $\gamma_i \in H_+(\mathbb{Q})$ and $k_i \in K$. Note that we can find $\gamma_i \in (H_v)_+(\mathbb{Q})$ and $k_i \in K \cap H_v(\mathbb{A}_f)$. With such choices, we have

$$\sum_i [\Phi(\gamma_i^{-1}v, \gamma_i^{-1}w)_\Gamma] = \sum_i [\Phi(v, \gamma_i^{-1}w)_\Gamma]. \tag{4-19}$$

The sum $\sum_i [\gamma_i^{-1} \cdot P_w]$ defines a divisor on X_0^B of degree $h(\mathscr{O}_L)$. In fact, by Shimura's description of the Galois action, this divisor coincides with the orbit under $\mathrm{Gal}(H/L)$ of $P_w \in X_0^B(H)$, with H the Hilbert class field of L. Hence we can write

$$\sum_i [\gamma_i^{-1} P_w] = t_{H/L}[P_w]. \tag{4-20}$$

(Here $t_{H/L}$ stands for taking the trace from H to L). Since in this case $n = 2$, the current $[\Phi(v, w, 1)_K]$ involves an additional sum. Namely, we need to choose $\gamma_0 \in H(\mathbb{Q})$ such that $\gamma_0 \cdot \mathbb{D}_U^+ = \mathbb{D}_U^-$; we can find such an element satisfying additionally that $\gamma_0 \cdot v = v$ and $\gamma_0 \cdot w = -w$. Now we have to find $k_{i_0} \in K$ and $\gamma_{i_0} \in H_+(\mathbb{Q})$ such that $\gamma_0 h'_i = \gamma_{i_0} k_{i_0}$. With our choice of K this is easy to do explicitly: let $\epsilon \in \mathscr{O}_B^\times$ be a unit of norm -1; such an element always exists by [Vignéras 1980, Corollary 5.9]. Then $(\epsilon, \epsilon) \in H(\mathbb{Q}) \cap K$. If $h'_i = \gamma_i k_i$ as above, then we can choose $\gamma_{i_0} = \gamma_0 \gamma_i \cdot (\epsilon, \epsilon)^{-1}$ and $k_{i_0} = (\epsilon, \epsilon) k_i$. Then we have $\gamma_{i_0}^{-1} \cdot v = v$ and $\gamma_{i_0}^{-1} \cdot w = -(\epsilon, \epsilon) \cdot \gamma_i^{-1} \cdot w$ and hence

$$\sum_i [\Phi(\gamma_{i_0}^{-1}v, \gamma_{i_0}^{-1}w)_\Gamma] = \sum_i [\Phi(v, (\epsilon, \epsilon)\gamma_i^{-1}w)_\Gamma], \tag{4-21}$$

since $[\Phi(v, w)_\Gamma] = [\Phi(v, -w)_\Gamma]$. By Shimura's reciprocity law [Ogg 1983, Equation (5)], if $P_{w'}$ is the point of X_0^B corresponding to $\mathbb{D}_{w'} \subset \mathbb{D} = \mathbb{H}^\pm$, then its complex conjugate $\bar{P}_{w'}$ corresponds to $\mathbb{D}_{(\epsilon,\epsilon)\cdot w'}$. It follows that

$$\sum_i [\gamma_i^{-1} P_w] + \sum_i [\gamma_{i_0}^{-1} P_w] = t_{H/\mathbb{Q}}[P_w] \tag{4-22}$$

and hence

$$[\Phi(v, w, 1)_K] = 2\pi i \cdot \Delta_*([G_{t_{H/\mathbb{Q}}[P_w]}]), \tag{4-23}$$

with $G_{t_{H/\mathbb{Q}}[P_w]}$ a Green function for the divisor $t_{H/\mathbb{Q}}[P_w]$ on X_0^B.

4B4. *The current* $[\Phi(T,\varphi)_K]$. Consider now an order $R = \mathbb{Z}[\alpha]$ in an imaginary quadratic field $L \subset \mathbb{C}$ and let $x^2 - tx + n$ be the minimal polynomial of α. We assume that L admits an embedding into B and (for simplicity) that $(d(L), d(B)) = 1$ and that $R = \mathscr{O}_L$ is the ring of integers of L. Define

$$\begin{aligned} T &= \begin{pmatrix} 1 & t/2 \\ t/2 & n \end{pmatrix}, \\ \varphi_{\mathscr{O}_B^2} &= \text{ characteristic function of } \hat{\mathscr{O}}_B^2, \end{aligned} \tag{4-24}$$

and let us describe the current $[\Phi(T, \varphi_{\mathscr{O}_B^2})]$ in (3-49). To do so, we need to describe the set of K-cosets of $\mathrm{Supp}(\varphi_{\mathscr{O}_B^2}) \cap \Omega_T(\mathbb{A}_f)$. We have

$$\begin{aligned} K\backslash[\mathrm{Supp}(\varphi_{\mathscr{O}_B^2}) \cap \Omega_T(\mathbb{A}_f)] &= (\hat{\mathscr{O}}_B^\times \times_{\hat{\mathbb{Z}}^\times} \hat{\mathscr{O}}_B^\times)\backslash \Omega_T(\hat{\mathscr{O}}_B^2) \\ &= \prod_{v \nmid \infty} (\mathscr{O}_{B,v}^\times \times_{\mathbb{Z}_v^\times} \mathscr{O}_{B,v}^\times)\backslash \Omega_T(\mathscr{O}_{B,v}^2). \end{aligned} \tag{4-25}$$

Note that the assignment $j \mapsto j(\alpha)$ induces a bijection between the (optimal) embeddings $j : R \to \mathscr{O}_B$ and the set of elements $w \in \mathscr{O}_B$ with $t(w) = t$ and $n(w) = n$, and this statement holds true locally too. It follows that the map $(1, w) \mapsto w$ induces a one-to-one correspondence

$$(\mathscr{O}_{B,v}^\times \times_{\mathbb{Z}_v^\times} \mathscr{O}_{B,v}^\times)\backslash \Omega_T(\mathscr{O}_{B,v}^2) \longleftrightarrow \{j : R \to \mathscr{O}_{B,v} \text{ optimal}\}/\mathscr{O}_{B,v}^\times, \tag{4-26}$$

where the equivalence in the right-hand side is with respect to conjugation by $\mathscr{O}_{B,v}^\times$. The set on the right-hand side has cardinality 1 if $B_v \cong M_2(\mathbb{Q}_v)$ and 2 if B_v is division; moreover, in the latter case the local Atkin–Lehner involution permutes the two elements (see [Vignéras 1980, Theorems II.3.1, II.3.2]). Hence the set

$$K\backslash[\mathrm{Supp}(\varphi_{\mathscr{O}_B^2}) \cap \Omega_T(\mathbb{A}_f)] \tag{4-27}$$

is a torsor under the Atkin–Lehner group W_B. Since the set $\mathrm{CM}(\mathscr{O}_L)$ of points in X_0^B with CM by $\mathscr{O}_L$ is a torsor under $\mathrm{Pic}(\mathscr{O}_L) \times W_B$, we conclude that

$$\left[\Phi\left(\begin{pmatrix} 1 & t/2 \\ t/2 & n \end{pmatrix}, \varphi_{\mathscr{O}_B^2}\right)_K\right] = 2\pi i \cdot (X_0^B \xrightarrow{\Delta} X_0^B \times X_0^B)_*([G_{t_{L/\mathbb{Q}}[\mathrm{CM}(\mathscr{O}_L)]}]) \tag{4-28}$$

is the pushforward along the diagonal of a Green current $[G_{t_{L/\mathbb{Q}}[\mathrm{CM}(\mathscr{O}_L)]}]$ for the divisor $t_{L/\mathbb{Q}}[\mathrm{CM}(\mathscr{O}_L)]$.

Note that by choosing $\varphi \in \mathscr{S}(V(\mathbb{A}_f)^2)$ to have support in a single K-orbit of (4-27), we recover all the currents of the form (4-23).

4B5. *The current* $[\Phi(T,\varphi)_K] - [\Phi(T^\iota,\varphi^\iota)_K]$. Recall that we have defined an involution ι on the set of pairs (T,φ), given by (3-82). Our next goal is to give an example of the action of ι.

Let p be a prime, $p \equiv 1 \pmod 4$ and not dividing $d(B)$, and define

$$\begin{aligned} T &= \begin{pmatrix} 1 & \\ & p \end{pmatrix}, \\ \varphi_{\mathscr{O}_B^2} &= \text{ characteristic function of } \hat{\mathscr{O}}_B^2. \end{aligned} \tag{4-29}$$

The previous computation of $[\Phi(T, \varphi_{\mathscr{O}_B^2})]$ shows that this current is supported on the diagonal Δ, and more precisely that

$$\left[\Phi\left(\begin{pmatrix} 1 & \\ & p \end{pmatrix}, \varphi_{\mathscr{O}_B^2}\right)_K\right] = 2\pi i \cdot (X_0^B \xrightarrow{\Delta} X_0^B \times X_0^B)_*([G_{t_{L/\mathbb{Q}}[\mathrm{CM}(\mathbb{Z}[\sqrt{-p}])]}]).$$

Note that $\varphi^{\iota}_{\mathscr{O}_B^2} = \varphi_{\mathscr{O}_B^2}$ and that $T^{\iota} = \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)$. In particular, the current $[\Phi(T^{\iota}, \varphi^{\iota}_{\mathscr{O}_B^2})]$ is different from $[\Phi(T, \varphi_{\mathscr{O}_B^2})]$, as the former is supported on the Hecke correspondence $T(p)$. More precisely, the same argument as above, with trivial modifications, shows that

$$\left[\Phi\left(\begin{pmatrix} p & \\ & 1 \end{pmatrix}, \varphi_{\mathscr{O}_B^2}\right)_K\right] = 2\pi i \cdot (X_0^B(p) \to X_0^B \times X_0^B)_*([G_{t_{L/\mathbb{Q}}[\mathrm{CM}(\mathbb{Z}[\sqrt{-p}])]}]),$$

where here $[\mathrm{CM}(\mathbb{Z}[\sqrt{-p}])]$ denotes the divisor consisting of all points in $X_0^B(p)$ with CM by $\mathbb{Z}[\sqrt{-p}]$ (for some CM type of $\mathbb{Z}[\sqrt{-p}]$).

Acknowledgments

Most of the work on this paper was done during my Ph.D. at Columbia University, as part of my Ph.D. thesis. I would like to express my deep gratitude to my advisor Shou-Wu Zhang, for introducing me to this area of mathematics, for his guidance and encouragement and for many very helpful suggestions. I would also like to thank Stephen S. Kudla for answering my questions about the theta correspondence and for several very inspiring remarks and conversations. This paper has also benefited from comments and discussions with Patrick Gallagher, Yifeng Liu, André Neves, Ambrus Pál, Yiannis Sakellaridis and Wei Zhang; I am grateful to all of them. Finally, I want to thank the referee for his useful suggestions and comments.

References

[Beilinson 1984] A. A. Beilinson, "Vysshie regulatory i znacheniya L-funktsiĭ", *Itogi Nauki i Tekhniki Ser. Sovrem. Probl. Mat. Nov. Dost.* **24** (1984), 181–238. Translated as "Higher regulators and values of L-functions", *Soviet Math.* **20**:2 (1985), 2036–2070. MR 86h:11103 Zbl 0588.14013

[Borcherds 1998] R. E. Borcherds, "Automorphic forms with singularities on Grassmannians", *Invent. Math.* **132**:3 (1998), 491–562. MR 99c:11049 Zbl 0919.11036

[Borcherds 1999] R. E. Borcherds, "The Gross–Kohnen–Zagier theorem in higher dimensions", *Duke Math. J.* **97**:2 (1999), 219–233. MR 2000f:11052 Zbl 0967.11022

[Borel 1969] A. Borel, *Introduction aux groupes arithmétiques*, Actualités Scientifiques et Industrielles **1341**, Hermann, Paris, 1969. MR 39 #5577 Zbl 0186.33202

[Bridson and Haefliger 1999] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften **319**, Springer, Berlin, 1999. MR 2000k:53038 Zbl 0988.53001

[Bruinier 2002] J. H. Bruinier, *Borcherds products on O(2, l) and Chern classes of Heegner divisors*, Lecture Notes in Mathematics **1780**, Springer, Berlin, 2002. MR 2003h:11052 Zbl 1004.11021

[Bruinier 2012] J. H. Bruinier, "Regularized theta lifts for orthogonal groups over totally real fields", *J. Reine Angew. Math.* **672** (2012), 177–222. MR 2995436 Zbl 1268.11058

[Bruinier and Funke 2004] J. H. Bruinier and J. Funke, "On two geometric theta lifts", *Duke Math. J.* **125**:1 (2004), 45–90. MR 2005m:11089 Zbl 1088.11030

[Chavel 2006] I. Chavel, *Riemannian geometry: a modern introduction*, 2-nd ed., Cambridge Studies in Advanced Mathematics **98**, Cambridge University Press, 2006. MR 2006m:53002 Zbl 1099.53001

[Erdélyi et al. 1954] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Tables of integral transforms*, vol. I, McGraw-Hill, New York, 1954. MR 15,868a Zbl 0055.36401

[Goncharov 2005] A. B. Goncharov, "Regulators", pp. 295–349 in *Handbook of K-theory*, vol. 1, edited by E. M. Friedlander and D. R. Grayson, Springer, Berlin, 2005. MR 2006j:11092 Zbl 1101.19004

[Kudla 1997] S. S. Kudla, "Algebraic cycles on Shimura varieties of orthogonal type", *Duke Math. J.* **86**:1 (1997), 39–78. MR 98e:11058 Zbl 0879.11026

[Kudla and Millson 1986] S. S. Kudla and J. J. Millson, "The theta correspondence and harmonic forms, I", *Math. Ann.* **274**:3 (1986), 353–378. MR 88b:11023 Zbl 0594.10020

[Kudla and Millson 1987] S. S. Kudla and J. J. Millson, "The theta correspondence and harmonic forms, II", *Math. Ann.* **277**:2 (1987), 267–314. MR 89b:11041 Zbl 0618.10022

[Kudla and Millson 1988] S. S. Kudla and J. J. Millson, "Tubes, cohomology with growth conditions and an application to the theta correspondence", *Canad. J. Math.* **40**:1 (1988), 1–37. MR 90k:11054 Zbl 0652.10021

[Kudla and Millson 1990] S. S. Kudla and J. J. Millson, "Intersection numbers of cycles on locally symmetric spaces and Fourier coefficients of holomorphic modular forms in several complex variables", *Inst. Hautes Études Sci. Publ. Math.* **71** (1990), 121–172. MR 92e:11035 Zbl 0722.11026

[Kudla and Rallis 1988] S. S. Kudla and S. Rallis, "On the Weil–Siegel formula", *J. Reine Angew. Math.* **387** (1988), 1–68. MR 90e:11059 Zbl 0644.10021

[Lebedev 1965] N. N. Lebedev, *Special functions and their applications*, edited by R. A. Silverman, Prentice-Hall, Englewood Cliffs, NJ, 1965. MR 30 #4988 Zbl 0131.07002

[Oda and Tsuzuki 2003] T. Oda and M. Tsuzuki, "Automorphic Green functions associated with the secondary spherical functions", *Publ. Res. Inst. Math. Sci.* **39**:3 (2003), 451–533. MR 2004f:11046 Zbl 1044.11033

[Ogg 1983] A. P. Ogg, "Real points on Shimura curves", pp. 277–307 in *Arithmetic and geometry*, vol. I, edited by M. Artin and J. Tate, Progr. Math. **35**, Birkhäuser, Boston, 1983. MR 85m:11034 Zbl 0531.14014

[Vignéras 1980] M.-F. Vignéras, *Arithmétique des algèbres de quaternions*, Lecture Notes in Mathematics **800**, Springer, Berlin, 1980. MR 82i:12016 Zbl 0422.12008

[Vogan and Zuckerman 1984] D. A. Vogan, Jr. and G. J. Zuckerman, "Unitary representations with nonzero cohomology", *Compositio Math.* **53**:1 (1984), 51–90. MR 86k:22040 Zbl 0692.22008

[Voisin 2002] C. Voisin, "Nori's connectivity theorem and higher Chow groups", *J. Inst. Math. Jussieu* **1**:2 (2002), 307–329. MR 2003m:14014 Zbl 1036.14004

Communicated by Peter Sarnak
Received 2015-05-08 Revised 2015-12-09 Accepted 2016-02-04

lgarcia@math.toronto.edu *Department of Mathematics, South Kensington Campus, Imperial College London, London, SW7 2AZ, United Kingdom*

Current address: *Department of Mathematics, University of Toronto, 40 St. George Street, BA 6290, Toronto, ON M5S 2E4, Canada*

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.645

Multiple period integrals and cohomology

Roelof W. Bruggeman and Youngju Choie

We give a version of the Eichler–Shimura isomorphism with a nonabelian H^1 in group cohomology. Manin has given a map from vectors of cusp forms to a noncommutative cohomology set by means of iterated integrals. We show that Manin's map is injective but far from surjective. By extending Manin's map we are able to construct a bijective map and remarkably this establishes the existence of a nonabelian version of the Eichler–Shimura map.

1. Introduction

In the theory of modular forms the Eichler–Shimura isomorphism has played an important role, with many applications. For instance, it gives integrality of eigenvalues for Hecke operators and algebraicity of the critical values of the L-functions of modular forms, which, for example, enables the construction of p-adic L-functions and gives a connection to Iwasawa theory as well as the computational aspects of modular form theory. The Eichler–Shimura isomorphism relates spaces of cusp forms of integral weight to a parabolic cohomology group, namely,

$$S_k(\mathrm{SL}_2(\mathbb{Z})) \oplus \bar{S}_k(\mathrm{SL}_2(\mathbb{Z})) \cong H^1_{\mathrm{par}}(\mathrm{SL}_2(\mathbb{Z}), \mathbb{C}_{k-2}[X, Y]),$$

where $\mathbb{C}_{k-2}[X, Y]$ is the $\mathrm{SL}_2(\mathbb{Z})$-module of homogeneous polynomials of degree $k-2$ in the indeterminates X, Y, and where $S_k(\mathrm{SL}_2(\mathbb{Z}))$ (resp. $\bar{S}_k(\mathrm{SL}_2(\mathbb{Z}))$) is the space of holomorphic (resp. antiholomorphic) cusp forms of weight k.

The Eichler–Shimura isomorphism was eventually extended in [Knopp 1974] and [Knopp and Mawi 2010] by establishing a canonical isomorphism between 1-cohomology of cofinite discrete subgroups Γ of $\mathrm{SL}_2(\mathbb{R})$ with appropriate holomorphic coefficients and the space of cusp forms with real weight.

Manin [2005] defined a "nonabelian" H^1 in group cohomology with values in a nonabelian group, and a map from a product of spaces of cusp forms to this cohomology set, in analogy to the Eichler–Shimura map. Manin's construction uses iterated integrals in the spirit of the multiple zeta values which have proved so useful

Choie is partially supported by NRF-2015049582 and NRF-2013R1A2A2A01068676.
MSC2010: primary 11F67; secondary 11F75.
Keywords: cusp form, iterated integral, noncommutative cohomology.

in understanding zeta values and mixed Tate motives, for example. Manin's integrals give a way to express multiple L-values of modular forms and have been studied by the second author [Choie 2014] and independently in the thesis of N. Provost [2014] recently. In [Choie 2014] the period polynomials whose coefficients are multiple L-values were treated as elements in a nonabelian H^1 for the first time.

In a recent talk at ICM, Brown [2014] mentioned a connection between the iterated integrals of Manin and certain mixed motives. He explained how to interpret motivic multiple zeta values as periods of the pro-unipotent fundamental groupoid of the projective line minus three points $X = \mathbb{P}^1 \setminus \{0, 1, \infty\}$ via iterated integral of smooth 1-forms on a differentiable manifold discussed by Chen [1977]. Hain [2015] discussed the relation between Manin's iterated integrals and the Hodge theory of modular groups. However, it was not clear yet how to relate Manin's iterated integral and Eichler–Shimura theory. Manin's map from spaces of cusp forms to cohomology differs in two aspects from the Eichler–Shimura map: the summand $\bar{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ is absent and the map is injective but not surjective.

This paper addresses the second difference by extending Manin's map to more complicated combinations of spaces of cusp forms to obtain a variant of the Eichler–Shimura isomorphism with values in a nonabelian cohomology H^1. Our main result (Theorem 6.7) states that there is an extension of Manin's map that is bijective onto a noncommutative cohomology set. It is remarkable that there exists some nonabelian version of the Eichler–Shimura map.

To obtain our main result we modify Manin's construction [2005] in several ways in the spirit of a variant Eichler–Shimura isomorphism established in [Knopp 1974; Knopp and Mawi 2010]: first replace the finite-dimensional spaces of polynomials by spaces of functions on the lower half-plane. Secondly, unlike in the classical Eichler–Shimura isomorphism, antiholomorphic modular forms are not considered. Thirdly, we allow automorphic forms for cofinite discrete subgroups Γ of $\mathrm{SL}_2(\mathbb{R})$ with arbitrary real weights and multiplier system. Finally, we collapse the number of variables in the iterated integrals.

To be more precise consider the iterated integral

$$\begin{aligned} R_\ell(f_1, \dots, f_\ell; y, x; t_1, \dots, t_\ell) &:= \int_{\tau_1=x}^{y} f_1(\tau_1)(\tau_1 - t_1)^{w_1} \int_{\tau_2=x}^{\tau_1} f_2(\tau_2)(\tau_2 - t_2)^{w_2} \\ &\qquad \cdots \int_{\tau_\ell=x}^{\tau_{\ell-1}} f_\ell(\tau_\ell)(\tau_\ell - t_\ell)^{w_\ell}\, d\tau_\ell \cdots d\tau_2\, d\tau_1, \end{aligned} \tag{1-1}$$

where x, y are in the extended complex upper half-plane and each t_j, $1 \le j \le \ell$, is in the lower half-plane. If the f_j are cusp forms of even integral weight $w_j + 2$, the iterated integral defines a polynomial function in the t_j, $1 \le j \le \ell$, whose coefficients are multiple L-values of f_j. The resulting iterated integral is holomorphic in

$(t_1, \ldots, t_\ell)$ in the product of ℓ copies of the lower half-plane if the f_j are cusp forms of real weight. As the order ℓ of the iterated integral increases, the relations between iterated integrals become more and more complicated. However, the relations between iterated integrals of order ℓ look simple modulo all products of iterated integrals of lower order. Manin [2005; 2006] has shown how to give a neat formulation for all relations among iterated integrals of the type indicated in (1-1). His approach works with formal series in noncommuting variables and can be applied to much more general iterated integrals than studied here.

The factors $(\tau_j - t_j)^{w_j}$ in (1-1) occur also in the definition of cocycles attached to cusp forms. Manin attaches to vectors of cusp forms $(f_1, \ldots, f_\ell)$ a cocycle in a noncommutative cohomology set, and thus gives a generalization of the Eichler–Shimura map. The cohomology has values in a noncommutative subgroup $\boldsymbol{N}(\mathcal{A})$ of the unit group of the noncommutative ring $\mathcal{A}$ of formal power series in noncommuting variables $A_1, \ldots, A_\ell$ with coefficients in spaces of holomorphic functions on the lower half-plane. The variables A_j correspond to spaces of cusp forms $S_{w_j+2}(\Gamma, v_j)$ with positive real weights $w_j + 2$ and corresponding multiplier systems v_j. Then Manin's approach leads to a map

$$\prod_{j=1}^{\ell} S_{w_j+2}(\Gamma, v_j) \longrightarrow H^1(\Gamma; \boldsymbol{N}(\mathcal{A})) \tag{1-2}$$

from a product of finitely many spaces of cusp forms to a noncommutative cohomology set. This (nonlinear) map is far from surjective. In Theorem 6.7 we show that Manin's map can be extended, and that all elements of the cohomology set $H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$ can be related to combinations of cusp forms by means of iterated integrals. The simplification $t_1 = \cdots = t_\ell$ in the iterated integrals is essential for our methods to work.

Sections 2 and 3 have a preliminary nature. We review the approach of Knopp [1974] of associating cocycles to any cusp form of real weight and the definition of the iterated integrals that we use. Sections 4 and 5 discuss Manin's approach of using formal series in noncommuting variables to associate noncommutative cocycles to vectors of cusp forms. In Section 6 we extend this approach in such a way that the resulting map from collections of cusp forms to noncommutative cohomology is bijective.

2. Cusp forms and theorem of Knopp and Mawi

Discrete group. Let Γ be a cofinite discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$ with translations. Without loss of generality we assume that $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \Gamma$. For convenience we conjugate Γ into a position for which ∞ is among its cusps and such that the subgroup Γ_∞ of Γ fixing ∞ is generated by $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

Notation. For $w \in \mathbb{R}$ and v a corresponding *unitary* multiplier system, we denote by $S_{w+2}(\Gamma, v)$ the space of holomorphic cusp forms of weight $w+2$ and multiplier system v. This is the finite-dimensional space of holomorphic functions f on the upper half-plane satisfying $f(\gamma z) = v(\gamma)(cz+d)^{w+2} f(z)$ for $\gamma \in \Gamma$, with exponential decay upon approach of the cusps. If the weight $w+2$ is integral, a multiplier system is a character.

Functions with at most polynomial growth. By $V(v, w)$, with $w \in \mathbb{R}$ and v a corresponding multiplier system, we denote the space of holomorphic functions on the lower half-plane $\mathfrak{H}^-$ with at most polynomial growth at the boundary $\mathbb{P}^1_{\mathbb{R}}$ of $\mathfrak{H}^-$, provided with the action of $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ given by

$$f|_{v,-w}\gamma(t) = v(\gamma)(ct+d)^{-w} f(\gamma t). \tag{2-1}$$

The condition that f has polynomial growth on $\mathfrak{H}^-$ can be formulated as

$$|f(t)| \le C_1 |t|^A + C_2 |\mathrm{Im}\, t|^{-A} \quad \text{for all } t \in \mathfrak{H}^-, \text{ for some } A, C_1, C_2 \ge 0. \tag{2-2}$$

The action $|_{v,-w}$ of Γ preserves this condition.

Remarks. (a) In [Bruggeman et al. 2014, §1.4] we denoted the representation $V(v, w)$ of Γ by $\mathcal{D}^{-\infty}_{v,-w}$ (actually we used $r = w+2$ as the main parameter, and wrote $\mathcal{D}^{-\infty}_{v,2-r}$).

(b) The polynomial growth condition in (2-2) can be formulated in terms of an estimate by one function $Q(t) = |\mathrm{Im}\, t| / |t-i|^2$ as $|f(t)| \le C Q(t)^{-A}$ for some $A, C \ge 0$. See the discussion in [Bruggeman et al. 2014, §1.5].

Knopp's cocycles associated to cusp forms. Knopp [1974] associated to cusp forms $f \in S_{w+2}(\Gamma, v)$ a cocycle $\bar{\psi}_f$ given by

$$\bar{\psi}_{f,\gamma}(z) = \overline{\int_{\gamma^{-1}\infty}^{\infty} f(\tau)(\tau - \bar{z})^w \, d\tau}.$$

This cocycle takes values in the holomorphic functions on the upper half-plane $\mathfrak{H}$ that have at most polynomial growth on $\mathfrak{H}$ in the sense of (2-2) (now with t replaced by $z \in \mathfrak{H}$). We avoid the complex conjugation by taking a cocycle with values in the holomorphic functions on the lower half-plane $\mathfrak{H}^-$ with at most polynomial growth at the boundary:

$$\psi_{f,\gamma}(t) = \int_{\tau=\gamma^{-1}\infty}^{\infty} f(\tau)(\tau - t)^w \, d\tau. \tag{2-3}$$

So ψ_f has values in the Γ-module $V(v, w)$.

Theorem 2.1 [Knopp and Mawi 2010]. *For real weight $w+2$ and corresponding unitary multiplier system v, the map $f \mapsto [\psi_f]$ determines a linear bijection*

$$S_{w+2}(\Gamma, v) \longrightarrow H^1(\Gamma; V(v, w)).$$

Knopp [1974] conjectured this result, and proved it for many cases. Finally, the remaining cases were completed in [Knopp and Mawi 2010].

Remarks. (a) A multiplier system v is called *unitary* if $|v(\gamma)| = 1$ for all $\gamma \in \Gamma$.

(b) Since $S_{w+2}(\Gamma, v) = \{0\}$ for $w+2 \le 0$, the theorem implies that the cohomology groups vanish as well for $w+2 \le 0$.

(c) If $w \in \mathbb{Z}_{\ge 0}$, the cocycles take values in polynomial functions on $\mathfrak{H}^-$, which for the trivial multiplier system form a submodule of $V(1, w)$ isomorphic to $\mathbb{C}_w[X, Y]$.

If the multiplier system v has values only in $\{1, -1\}$ then conjugation gives cocycles in the same module. The Eichler–Shimura theory gives the parabolic cohomology group with values in polynomial functions of degree at most w as the direct sum of the images of the two maps $f \mapsto [\psi_f]$ and $f \mapsto [\bar{\psi}_f]$. However, in the large module of polynomially growing functions, the cocycles $\bar{\psi}_f$ become coboundaries. Also the cocycles associated to Eisenstein series become coboundaries over the module of functions with at most polynomial growth.

(d) Knopp [1974] shows that the parabolic cohomology group $H^1_{\mathrm{par}}(\Gamma; V(v, w))$ is equal to the cohomology group $H^1(\Gamma; V(v, w))$.

3. Iterated integrals

By taking $t_1 = \cdots = t_\ell = t$ we consider the following holomorphic function in t running through the lower half-plane:

$$R_\ell(f_1, \dots, f_\ell; y, x; t) := \int_{\tau_1 = x}^{y} f_1(\tau_1)(\tau_1 - t)^{w_1} \int_{\tau_2 = x}^{\tau_1} f_2(\tau_2)(\tau_2 - t)^{w_2} \cdots \int_{\tau_\ell = x}^{\tau_{\ell-1}} f_\ell(\tau_\ell)(\tau_\ell - t)^{w_\ell}\, d\tau_\ell \cdots d\tau_2\, d\tau_1. \quad (3\text{-}1)$$

It is a multilinear form on $\prod_{j=1}^{\ell} S_{w_j+2}(\Gamma, v_j)$ for ℓ pairs $(v_1, w_1), \dots, (v_\ell, w_\ell)$ of real numbers w_j and corresponding unitary multiplier systems v_j. The parameter t is in the lower half-plane $\mathfrak{H}^-$. The value of the iterated integral does not depend on the path of integration, provided we take care to approach cusps along geodesic half-lines (for instance, vertically).

The most interesting case is $y = \gamma^{-1}\infty$, $\gamma \in \Gamma$, and $x = \infty$. For $\ell = 1$ this gives the value $\psi_{f_1,\gamma}$ of the cocycle in (2-3). That is why we call $R_\ell(f_1, \dots, f_\ell; \gamma^{-1}\infty, \infty; t)$ a *multiple period integral*.

Functions with at most polynomial growth. The condition of polynomial growth in (2-2) is preserved by the action of Γ given for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ by

$$\begin{aligned} h|_{\boldsymbol{v},-\boldsymbol{w}}\gamma(t) &= \boldsymbol{v}(\gamma)^{-1}(ct+d)^{\boldsymbol{w}} h(\gamma t), \\ \boldsymbol{v}(\gamma) &= v_1(\gamma)v_2(\gamma)\cdots v_\ell(\gamma), \\ \boldsymbol{w} &= w_1 + w_2 + \cdots + w_\ell. \end{aligned} \tag{3-2}$$

By $V(\boldsymbol{v}, \boldsymbol{w})$ we denote the vector space of holomorphic functions on $\mathfrak{H}^-$ with the action $|_{\boldsymbol{v},-\boldsymbol{w}}$ given in (3-2). Multiplication of functions gives a bilinear map $V(\boldsymbol{v}; \boldsymbol{w}) \times V(\boldsymbol{v}'; \boldsymbol{w}') \to V(\boldsymbol{v}\boldsymbol{v}'; \boldsymbol{w} + \boldsymbol{w}')$. The action behaves according to the rule

$$(h|_{\boldsymbol{v},-\boldsymbol{w}}\gamma)(h'|_{\boldsymbol{v}',-\boldsymbol{w}'}\gamma) = (hh')|_{\boldsymbol{v}\boldsymbol{v}',-\boldsymbol{w}-\boldsymbol{w}'}\gamma. \tag{3-3}$$

Lemma 3.1. *For $\boldsymbol{f} = (f_1, \dots, f_\ell) \in \prod_{j=1}^{\ell} S_{w_j+2}(\Gamma, v_j)$, the multiple period integral $R_\ell(\boldsymbol{f}; y, x; \cdot)$ defines an element of $V(\boldsymbol{v}, \boldsymbol{w})$.*

Proof. Each cusp form has at most polynomial growth on $\mathfrak{H}$, and has exponential decay at cusps when the cusp is approached along a geodesic half-line. This implies that the iterated integral in (3-1) has at most polynomial growth in t and $\tau_{\ell-1}$. Successively this also implies polynomial growth in τ_{j-1} and t of the further integrals. □

Trivial relation. Directly from the definition we have

$$R_\ell(\boldsymbol{f}; x, x; t) = 0. \tag{3-4}$$

Lemma 3.2. *For $\gamma \in \Gamma$,*

$$R_\ell(\boldsymbol{f}; \gamma^{-1}y, \gamma^{-1}x; t) = R_\ell(\boldsymbol{f}; y, x; \cdot)|_{\boldsymbol{v},-\boldsymbol{w}}\gamma(t). \tag{3-5}$$

Proof. In the following computation all τ_j are replaced by $\gamma\tau_j$, with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$:

$$\begin{aligned} &R_\ell(\boldsymbol{f}; x, y; \cdot)|_{\boldsymbol{v},-\boldsymbol{w}}\gamma(t) \\ &= \prod_{j=1}^{\ell} (v_j(\gamma)^{-1}(ct+d)^{w_j}) \int_{\tau_1=x}^{y} f_1(\tau_1)(\tau_1-\gamma t)^{w_1} \int_{\tau_\ell=x}^{\tau_{\ell-1}} f_\ell(\tau_\ell)(\tau_\ell-\gamma t)^{w_\ell} \\ &\qquad\qquad d\tau_\ell \cdots d\tau_1 \\ &= \prod_{j=1}^{\ell} (v_j(\gamma)^{-1}(ct+d)^{w_j}) \int_{\tau_1=\gamma^{-1}x}^{\gamma^{-1}y} f_1(\gamma\tau_1) \frac{(\tau_1-t)^{w_1}}{(c\tau_1+d)^{w_1}(ct+d)^{w_1}} \\ &\qquad \int_{\tau_\ell=\gamma^{-1}x}^{\gamma^{-1}(\gamma\tau_{\ell-1})} f_\ell(\gamma\tau_\ell) \frac{(\tau_\ell-t)^{w_\ell}}{(c\tau_\ell+d)^{w_\ell}(ct+d)^{w_\ell}} \frac{d\tau_\ell}{(c\tau_\ell+d)^2} \cdots \frac{d\tau_1}{(c\tau_1+d)^2} \\ &= R_\ell(\boldsymbol{f}; \gamma^{-1}y, \gamma^{-1}x; t). \end{aligned}$$

□

Cocycles. For $\ell = 1$ we get the cocycle ψ_f in (2-3):

$$\psi_{f,\gamma}(t) = -R_1(f; \gamma^{-1}\infty, \infty; t). \tag{3-6}$$

Decomposition. It is easy to see that the cocycles in (2-3) satisfy the cocycle relation

$$c_{\gamma\delta} = c_\gamma|\delta + c_\delta$$

for $\gamma, \delta \in \Gamma$: use the decomposition relation $\int_b^a + \int_c^b = \int_c^a$ for integrals together with the invariance relation in Lemma 3.2.

There are *decomposition relations* for the iterated integrals in (3-1), which can be obtained by application of the decomposition relation for integrals of one variable to the subintegrals in (3-1). For the orders 2 and 3 these relations take the form

$$\begin{aligned} &R_2(f_1, f_2; z, y; t) + R_2(f_1, f_2; y, x; t) - R_2(f_1, f_2; z, x; t) \\ &\quad = R_1(f_1; z, y; t) R_1(f_2; y, x; t), \end{aligned} \tag{3-7}$$

$$\begin{aligned} &R_3(f_1, f_2, f_3; z, y; t) + R_3(f_1, f_2, f_3; y, x; t) - R_3(f_1, f_2, f_2; z, x; t) \\ &\quad = -R_1(f_1; z, y; t) R_2(f_2, f_3; y, x; t) + R_2(f_1, f_2; z, y; t) R_1(f_3; y, x; t). \end{aligned} \tag{3-8}$$

We have written these relations in such a way that the quantity on the left should be zero if the standard decomposition would hold. On the right is a correction term consisting of products of iterated integrals of lower order.

Example. The decomposition relations can be used to obtain relations between values of multiple L-functions at special points, as studied in [Choie 2014] and in the thesis by Provost [2014] independently.

Let us take $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, and assume that $v_1 = v_2 = 1$, and $w_1, w_2 \in 2\mathbb{Z}_{\geq 0}$. This implies that the multiple integrals yield polynomial functions in the variable t. We apply (3-7) with $z = x = \infty$ and $y = 0$. With (3-4),

$$R_2(f_1, f_2; \infty, 0; t) + R_2(f_1, f_2; 0, \infty; t) = R_1(f_1; \infty, 0; t) R_1(f_2; 0, \infty; t).$$

Using the binomial theorem, we see that $R_1(f; \infty, 0; t)$ is a polynomial in t with coefficients that can be expressed in values of completed L-functions. In a similar way, $R_2(f_1, f_2; \infty, 0; t)$ is a polynomial in t with coefficients that can be expressed in values of a completed multiple L-function of order 2 as defined in [Choie 2014, (2.6)]. With Lemma 3.2,

$$R_2(f_1, f_2; 0, \infty; \cdot) = R_2(f_1, f_2; \infty, 0; \cdot)|_{-w} S,$$

where $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. In this way, the decomposition relation (3-7) implies the equality of two polynomials. Comparing coefficients leads to the relation in [Choie 2014, Theorem 3.1].

This account is a simplification. The decomposition relations are valid for the iterated integrals in (1-1), and lead for $w_j \in 2\mathbb{Z}_{\geq 0}$ to polynomials in two variables. Choie [2014] works in that generality.

4. Formal series

Manin [2005; 2006] has indicated a way to give structure to the decomposition relations of any order. His approach works in a general context of iterated integrals associated to cusp forms. The factors $(\tau_j - t)^w$ of the kernel in (3-1) and $(\tau_j - t_j)$ in (1-1) may be replaced by more general factors, for instance, by factors leading to iterated L-integrals as studied in [Choie 2014]. Here we use Manin's formalism for the iterated integrals in (3-1).

We keep fixed ℓ combinations of a weight $w_j + 2 \in \mathbb{R}$ and a corresponding unitary multiplier system v_j. For a vector $\boldsymbol{f} = (f_1, \dots, f_\ell) \in \prod_{j=1}^{\ell} S_{w_j+2}(\Gamma, v_j)$ of length ℓ, we form iterated integrals of arbitrary order

$$R_n(f_{m_1}, f_{m_2}, \dots, f_{m_n}; y, x; t) \tag{4-1}$$

for any choice $m = (m_1, \dots, m_n) \in \{1, \dots, \ell\}^n$, for any $n \geq 0$. For $n = 0$ we define this quantity to be 1. The same f_j may occur several times as f_{m_i}. So we do not get linearity in f_j. The result is a holomorphic function on $\mathfrak{H}^-$, and has at most polynomial growth by Lemma 3.1.

To formulate the Γ-equivariance, we put for $m = (m_1, m_2, \dots, m_n)$

$$\boldsymbol{v}(m) := v_{m_1} v_{m_2} \cdots v_{m_n}, \qquad \boldsymbol{w}(m) := w_{m_1} + w_{m_2} + \cdots + w_{m_n}. \tag{4-2}$$

We consider the iterated integral in (4-1) as an element of $V(\boldsymbol{v}(m), \boldsymbol{w}(m))$. For the empty sequence $m = ()$ we put $V(\boldsymbol{v}(), \boldsymbol{w}()) = \mathbb{C}$ with the trivial action $|_{1,0}$. Multiplication follows the rule in (3-3). Lemma 3.2 can be applied.

Power series in noncommuting variables. We choose ℓ spaces of cusp forms $S_{w_j+2}(\Gamma, v_j)$ with $w_j + 2 > 0$ and unitary multiplier systems v_j, for $1 \leq j \leq \ell$. We indicate this choice by the symbol $\mathcal{A}$. For this choice $\mathcal{A}$ we take ℓ noncommuting variables $A_1, A_2, \dots, A_\ell$.

Let $\boldsymbol{O}(\mathcal{A})$ be the set of formal power series in the A_j for which the coefficient of the monomial $A_{m_1} A_{m_2} \cdots A_{m_n}$ is in $V(\boldsymbol{v}(m), \boldsymbol{w}(m))$ for each $m \in \{1, \dots, \ell\}^n$. The constant term is in $V(\boldsymbol{v}(), \boldsymbol{w}()) = \mathbb{C}$. The relation (3-3) implies that $\boldsymbol{O}(\mathcal{A})$ is a ring.

Formal series associated to vectors of cusp forms. Following Manin we combine all iterated integrals in (4-1) as coefficients of an element of the ring $\boldsymbol{O}(\mathcal{A})$. Let

$$S_{\mathcal{A}}(\Gamma) = \prod_{j=1}^{\ell} S_{w_j+2}(\Gamma, v_j). \tag{4-3}$$

For $\boldsymbol{f} = (f_1, \dots, f_\ell) \in S_{\mathcal{A}}(\Gamma)$ define the formal series $J(\boldsymbol{f}; y, x; t) \in \boldsymbol{O}(\mathcal{A})$ by

$$J(\boldsymbol{f}; y, x; t) \\ = 1 + \sum_{n \geq 1} \sum_{m_1, \dots, m_n \in \{1, \dots, \ell\}} R_n(f_{m_1}, f_{m_2}, \dots, f_{m_n}; y, x; t) A_{m_1} A_{m_2} \cdots A_{m_n}. \tag{4-4}$$

Remarks. (a) $J(\boldsymbol{f}; z, w; \cdot)$ is an invertible element of $\boldsymbol{O}(\mathcal{A})$ since it has a nonzero constant term.

(b) The coefficients $R_n(f_{m_1}, \ldots, f_{m_n}; y, x; t)$ are continuous functions of $y, x \in \mathfrak{H}^*$, and are holomorphic in $x, y \in \mathfrak{H}$.

(c) The A_j codes for the space $S_{w_j+2}(\Gamma, v_j)$. This approach differs from that in [Manin 2005, §2]. There the formal variables code for linearly independent elements of the space $\prod_j S_{w_j+2}(\Gamma, v_j)$.

Action of Γ. We define an action of Γ on $\boldsymbol{O}(\mathcal{A})$ by the action $|_{\boldsymbol{v}(m),-\boldsymbol{w}(m)}$ on the coefficient of $A_{m_1} \cdots A_{m_n}$. Lemma 3.2 implies the relation

$$J(\boldsymbol{f}; \gamma^{-1}y, \gamma^{-1}x; \cdot) = J(\boldsymbol{f}; y, x; \cdot)|\gamma \quad \text{for each } \gamma \in \Gamma. \tag{4-5}$$

Multiplication properties. These formal series satisfy for $z, y, x \in \mathfrak{H}^*$

$$J(\boldsymbol{f}; x, x; t) = 1, \tag{4-6}$$

$$J(\boldsymbol{f}; x, y; t) = J(\boldsymbol{f}; y, x; t)^{-1}, \tag{4-7}$$

$$J(\boldsymbol{f}; z, x; t) = J(\boldsymbol{f}; z, y; t)J(\boldsymbol{f}; y, x; t). \tag{4-8}$$

We will prove a more general result in Proposition 6.3.

These relations encapsulate infinitely many relations between multiple period integrals. The reader who takes the trouble to compare the coefficients of A_1A_2 in (4-8) obtains the relation (3-7). Similarly, relation (3-8) is given by the coefficient of $A_1A_2A_3$.

Commutative example. In the modular case we may look at $w = N/2-2$ for some $N \in \mathbb{Z}_{\geq 1}$. As the corresponding multiplier system we choose $v_{N/2}$ determined by

$$v_{N/2}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} = e^{\pi i N/12}, \qquad v_{N/2}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} = e^{-\pi i N/4}. \tag{4-9}$$

For $1 \leq N \leq 24$ the space of cusp forms is one-dimensional, in fact:

$$S_{N/2+2}(\Gamma(1), v_{N/2}) = \mathbb{C}\eta^N,$$

where $\eta(\tau) = q^{1/24}\prod_{n\geq 1}(1-q^n)$, $q = e^{2\pi i\tau}$, is the Dedekind eta function.

We take $\ell = 1$, with $w_1 = N/2-2$ and multiplier system $v_{N/2}$. The ring $\boldsymbol{O}(\mathcal{A})$ is a *commutative* ring of formal power series in one variable A. The coefficient of A^m is in the $\Gamma(1)$-module $V(v_{mN/2}, mN/2-2m)$.

If we take $1 \leq N \leq 24$, then with $\boldsymbol{f} = (\eta^N)$ we get in (4-4)

$$J(\boldsymbol{f}; y, x; t) = 1 + \sum_{n\geq 1} R_n((\eta^N)^{\times n}; y, x; t)A^n, \tag{4-10}$$

where $(\eta^N)^{\times n}$ means a sequence of n copies of η^N.

If $N > 24$ we still can work with $\boldsymbol{f} = (f)$, but now f need not be a multiple of η^N.

5. From cusp forms to noncommutative cohomology

Manin uses relation (4-8) to associate a noncommutative cocycle to the vector $\boldsymbol{f} = (f_1, \dots, f_\ell)$ of cusp forms. We first reformulate Manin's description [2005, §1] of noncommutative cohomology for a right action, and then determine the map from vectors of cusp forms to noncommutative cohomology.

Noncommutative cohomology. Let G and N be groups, written multiplicatively, and suppose that for each $g \in G$ there is an automorphism $n \mapsto n|g$ of N such that the map $g \mapsto |g$ is an antihomomorphism from G to the automorphism group $\mathrm{Aut}(N)$, i.e., $n|(gh) = (n|g)|h$ for $n \in N$ and $g, h \in G$.

A map $\rho : G \to N$ is called a 1-cocycle if it satisfies

$$\rho_{gh} = (\rho_g|h)\rho_h \quad \text{for all } g, h \in G. \tag{5-1}$$

The set of such cocycles is called $Z^1(G; N)$. It is not a group. Nevertheless it contains the special element $1 : g \mapsto 1$.

The group N acts on $Z^1(G; N)$ from the left, by $\rho \mapsto {}^n\rho$ defined by

$${}^n\rho_g = (n|g)\rho_g n^{-1}. \tag{5-2}$$

The cohomology set $H^1(G; N)$ is the set of N-orbits in $Z^1(G; N)$ for this action. The orbit of the cocycle $g \mapsto 1$ is called the set of coboundaries $B^1(G; N)$.

Noncommutative cocycles attached to a sequence of cusp forms. As the group N we use the subgroup $\boldsymbol{N}(\mathcal{A})$ of the group of those units in $\boldsymbol{O}(\mathcal{A})^*$ that have constant term equal to 1. The series $J(\boldsymbol{f}; y, x; \cdot\,)$ in (4-4) is an element of $\boldsymbol{N}(\mathcal{A})$.

Following Manin we define for $\boldsymbol{f} = (f_1, \dots, f_\ell) \in S_{\mathcal{A}}(\Gamma)$ and $x \in \mathfrak{H}^*$

$$\Psi(\boldsymbol{f})^x_\gamma(t) = J(\boldsymbol{f}; \gamma^{-1}x, x; t). \tag{5-3}$$

The properties (4-5) and (4-8) imply that this defines a noncommutative cocycle $\Psi(\boldsymbol{f})^x \in Z^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$, and that its cohomology class $\mathrm{Coh}_{\mathcal{A}}(\boldsymbol{f}) \in H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$ does not depend on the choice of the base-point x. We write $\Psi(\boldsymbol{f}) = \Psi(\boldsymbol{f})^\infty$.

Proposition 5.1. *The map*

$$\mathrm{Coh}_{\mathcal{A}} : S_{\mathcal{A}}(\Gamma) \to H^1(\Gamma; \boldsymbol{N}(\mathcal{A})) \tag{5-4}$$

is injective.

Proof. Suppose that the cocycles $\Psi(f_1, \dots, f_\ell)$ and $\Psi(f_1', \dots, f_\ell')$ are in the same cohomology class. Then there is an $n \in \boldsymbol{N}(\mathcal{A})$ such that for all $\gamma \in \Gamma$

$$\Psi(f_1', \dots, f_\ell')_\gamma = (n|\gamma)\Psi(f_1, \dots, f_\ell)_\gamma n^{-1}. \tag{5-5}$$

We denote the coefficient of A_j in n by $n_j \in V(v_j, w_j)$. In relation (5-5) we consider only the constant term and the term with A_j, and work modulo all other terms:

$$1 - \psi_{f_j',\gamma} A_j \equiv (1 + n_j A_j)(1 - \psi_{f_j,\gamma} A_j)(1 - n_j A_j).$$

Taking the factor of A_j gives

$$-\psi_{f_j',\gamma} = n_j|_{v_j,-w_j}\gamma - \psi_{f_j,\gamma} - n_j.$$

In other words, $\psi_{f_j'}$ and ψ_{f_j} differ by a coboundary. We have used the noncommutative relation (5-5) in $\boldsymbol{N}(\mathcal{A})$ to get a commutative relation in $V(v_j, w_j)$.

By the theorem of Knopp and Mawi (Theorem 2.1) we conclude that $f_j' = f_j$ for all j. Hence $\mathrm{Coh}_{\mathcal{A}}$ is injective. $\square$

Remarks. (a) Implicit in the proof is the quotient of $\mathcal{A}$ by the ideal generated by all monomials in the A_j with degree 2. The corresponding quotient of $\boldsymbol{N}(\mathcal{A})$ is isomorphic to the direct sum of the $V(v_j, w_j)$.

(b) The injectivity of the map from cusp forms to cocycles is a point in common for this result, the theorem of Knopp and Mawi, and the classical Eichler–Shimura result. The bijectivity in the theorem of Knopp and Mawi is not shared by the classical result, where conjugates of cocycles also determine cohomology classes. In the next section we will see that the whole group $H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$ can be described with cusp forms, but in a more complicated way than by the map $\mathrm{Coh}_{\mathcal{A}}$.

Commutative example. Toward the end of page 653 we considered the case $\ell = 1$. Then $\boldsymbol{N}(\mathcal{A})$ is a commutative group, and $H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$ is a cohomology *group*.

When $\Gamma = \Gamma(1)$, with the choices and notations indicated on page 653, the cocycle $\Psi(\eta^N)$ vanishes on $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ (hence may be called a parabolic cocycle), and is determined by its value on $S = \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$:

$$\Psi(\eta^N)_S(t) = J(\eta^N; 0, \infty; t) = 1 + \sum_{n\geq 1} R_n((\eta^N)^{\times n}; 0, \infty; t) A^n. \tag{5-6}$$

The coefficient of A^n is an iterated period integral of η^N. The cocycle satisfies the well known relations $(\Psi(\eta^N)_S|S)\Psi(\eta^N)_S = 1$ and $\Psi(\eta^N)_S = \Psi(\eta^N)_S|T'\Psi(\eta^N)_S|T$, with $T' = \left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right) = TST$.

6. Noncommutative cocycles and collections of cusp forms

The proof of Proposition 5.1 is based on the fact that the vector of cusp forms $\boldsymbol{f}$ can be recovered from the terms of degree 1 in the formal series $J(\boldsymbol{f}; \gamma^{-1}\infty, \infty; t)$. In this section we associate to collections of cusp forms noncommutative cocycles of a more general nature.

We keep fixed the choice $\mathcal{A}$ of positive weights $w_1+2, \dots, w_\ell+2$ and corresponding multiplier systems $v_1, \dots, v_\ell$. To each monomial $B = A_{m_1} \cdots A_{m_d}$ in $\boldsymbol{O}(\mathcal{A})$ we associate the shifted weight $\boldsymbol{w}(B) := \boldsymbol{w}(m)$ and the multiplier system $\boldsymbol{v}(B) := \boldsymbol{v}(m)$ as defined in (4-2) for $m = (m_1, \dots, m_d) \in \{1, \dots, \ell\}^d$. So $\boldsymbol{w}(B)$ and $\boldsymbol{v}(B)$ depend only on the factors A_{m_i} occurring in B, not on their order.

Definition 6.1. We call the *degree* $d(B)$ of the monomial $B = A_{m_1} \cdots A_{m_d}$ the number d of factors A_j $(1 \le j \le \ell)$ occurring in it.

Let $\mathcal{B}(\mathcal{A})$ be the set of all monomials B in $A_1, \dots, A_\ell$ with $d(B) \ge 1$ for which $S_{\boldsymbol{w}(B)+2}(\Gamma, \boldsymbol{v}(B)) \ne \{0\}$. We put

$$S(\mathcal{A}; \Gamma) := \prod_{B \in \mathcal{B}(\mathcal{A})} S_{\boldsymbol{w}(B)+2}(\Gamma, \boldsymbol{v}(B)). \tag{6-1}$$

Remarks. (a) The space of cusp forms $S_{\boldsymbol{w}(B)+2}(\Gamma, \boldsymbol{v}(B))$ may be zero. In fact, this is necessarily the case if $\boldsymbol{w}(B) \le -2$. For $\boldsymbol{w}(B) > -2$ it may also happen to be zero, depending on Γ and $\boldsymbol{v}(B)$.

(b) The set $\mathcal{B}(\mathcal{A})$ is often infinite. We recall that elements of infinite direct sums of vector spaces have zero components at all but finitely many $B \in \mathcal{B}(\mathcal{A})$. Here we use the product. Its elements may have nonzero components for all B.

(c) We denote elements of $S(\mathcal{A}; \Gamma)$ by $\boldsymbol{h}$, with component $\boldsymbol{h}(B)$ in the factor corresponding to the monomial B.

(d) There may be more than one monomial B for which $S_{\boldsymbol{w}(B)+2}(\Gamma, \boldsymbol{v}(B))$ is equal to a given space of cusp forms. See (f2) below for an example where this happens for infinitely many monomials.

(e) The space $S_{\mathcal{A}}(\Gamma) = \prod_{j=1}^{\ell} S_{w_j+2}(\Gamma, v_j)$ in (4-3) may be considered as a subspace of $S(\mathcal{A}; \Gamma)$. To do this we define for a given $\boldsymbol{f} = (f_1, \dots, f_\ell) \in S_{\mathcal{A}}(\Gamma)$ the element $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$ by

$$\boldsymbol{h}(A_j) = f_j, \quad \boldsymbol{h}(B) = 0, \quad \text{if } d(B) \ge 2.$$

(f) In the *commutative case* $\ell = 1$ we have $\mathcal{B}(\mathcal{A}) \subset \{A^n : n \in \mathbb{Z}_{\ge 1}\}$. We consider three specializations of the example on page 653.

(f1) Take $N = 24$. So $\mathcal{B}(\mathcal{A}) = \{A^n : n \in \mathbb{Z}_{\ge 1}\}$, $\boldsymbol{w}(B) = 10n$ and $\boldsymbol{v}(B) = v_{12} = v_0 = 1$. Hence

$$S(\mathcal{A}; \Gamma(1)) = \prod_{n \ge 1} S_{10n+2}(\Gamma(1), 1). \tag{6-2}$$

(f2) Take $N = 4$. So $\boldsymbol{w}(A^n) = 0$ for all $n \ge 1$ and the space $S_2(\Gamma(1), v_{2n})$ is equal to $\mathbb{C}\eta^4$ if $n \equiv 1 \bmod 6$ and zero otherwise. This implies that

$$\mathcal{B}(\mathcal{A}) = \{A^n \ge 1 : n \equiv 1 \bmod 6\}.$$

Since $\boldsymbol{v}(B) = v_n = v_2$ for $n \equiv 1 \bmod 6$, we obtain

$$S(\mathcal{A}; \Gamma(1)) = \prod_{\substack{n \geq 1 \\ n \equiv 1 \bmod 6}} S_2(\Gamma(1), v_2). \tag{6-3}$$

(f3) Take $N = 1$. So $\boldsymbol{w}(B) = w_1 = -\frac{3}{2}$, and $nw_1 < -2$ for $n \geq 2$. Hence $\mathcal{B}(\mathcal{A}) = \{A\}$, $\boldsymbol{v}(B) = v_{1/2}$, and

$$S(\mathcal{A}; \Gamma(1)) = S_{1/2}(\Gamma(1), v_{1/2}) = \mathbb{C}\eta, \tag{6-4}$$

with $\eta(\tau) = e^{\pi i \tau/12} \prod_{n \geq 1}(1 - e^{2\pi i n \tau})$.

Lemma 6.2. *For each $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$, the series*

$$J(\boldsymbol{h}; y, x; t) := 1 + \sum_{n \geq 1} \sum_{B_1, \ldots, B_n \in \mathcal{B}(\mathcal{A})} R_n(\boldsymbol{h}(B_1), \boldsymbol{h}(B_2), \ldots, \boldsymbol{h}(B_n); y, x; t) B_1 B_2 \cdots B_n \tag{6-5}$$

converges and defines an element of $\boldsymbol{N}(\mathcal{A})$.

Proof. The degree of $B_1 B_2 \cdots B_n$ is at least n. For convergence in $\boldsymbol{O}(\mathcal{A})$ there should be for each $D \geq 0$ only finitely many terms with degree at most D. This restricts n to $n \leq D$, and the B_j to monomials of degree bounded by D, of which there are only finitely many.

The terms with $n \geq 1$ cannot contribute to the constant term, hence we obtain an element of $\boldsymbol{N}(\mathcal{A})$. □

Remarks. (a) If $h(B_i) = 0$ for some i in the iterated integral in (6-5), then the integral vanishes. In (6-5) we could have restricted the B_i in the sum by the condition $\boldsymbol{h}(B_i) \neq 0$. In particular, $J(0; y; x; t) = 1$.

(b) Definition (6-5) extends definition (4-4). If $\boldsymbol{f} \in S_{\mathcal{A}}(\Gamma)$ is considered as an element $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$, as in remark (e) to Definition 6.1, then

$$J(\boldsymbol{f}; y, x; t) = J(\boldsymbol{h}; y, x; t). \tag{6-6}$$

Proposition 6.3. *For all $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$, $\gamma \in \Gamma$, $z, y, x \in \mathfrak{H}^*$,*

$$J(\boldsymbol{h}; \gamma^{-1} y, \gamma^{-1} x; \cdot) = J(\boldsymbol{h}; y, x; \cdot)|\gamma, \tag{6-7}$$

$$J(\boldsymbol{h}; x, x; t) = 1, \tag{6-8}$$

$$J(\boldsymbol{h}; z, x; t) = J(\boldsymbol{h}; z, y; t) J(\boldsymbol{h}; y, x; t), \tag{6-9}$$

$$J(\boldsymbol{h}; x, y; t) = J(\boldsymbol{h}; y, x; t)^{-1}. \tag{6-10}$$

Proof. The relations (6-7) and (6-8) follow directly from Lemma 3.2 and (3-1). We will prove relation (6-9) in a sequence of lemmas, and finally will derive relation (6-10) from relation (6-9).

Relation (6-9). This relation holds in a general context of iterated integrals; automorphic properties are not needed. Our proof follows [Manin 2006, Proposition 1.2] closely. We first show relation (6-9) for $x, y, z \in \mathfrak{H}$.

Lemma 6.4. *For $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$ put*

$$\Omega(\boldsymbol{h}; z; t) := \sum_{B \in \mathcal{B}(\mathcal{A})} (z-t)^{\boldsymbol{w}(B)} \boldsymbol{h}(B; z)\, dz \cdot B. \tag{6-11}$$

This formal series of $\boldsymbol{O}(\mathcal{A})$-valued differential forms converges, and for $z \in \mathfrak{H}$

$$d_z J(\boldsymbol{h}; z, x; t) = \Omega(\boldsymbol{h}; z; t) J(\boldsymbol{h}; z, x; t). \tag{6-12}$$

Proof. The sum in (6-11) is infinite in most cases. The convergence follows from the fact that the number of monomials with a given degree is finite. The differential of a nonconstant term in (6-5) is given by

$$\begin{aligned} &d_z R_n(\boldsymbol{h}(B_1), \boldsymbol{h}(B_2), \dots, \boldsymbol{h}(B_n); z, x; t) B_1 B_2 \cdots B_n \\ &\qquad = \boldsymbol{h}(B_1; z)(z-t)^{\boldsymbol{w}(B_1)} B_1 R_{n-1}(\boldsymbol{h}(B_2), \dots, \boldsymbol{h}(B_n); z, x; t) B_2 \cdots B_n. \end{aligned}$$

With a renumbering in the summation this gives (6-12). □

Lemma 6.5. $d_z J(\boldsymbol{h}; z, x; t)^{-1} = -J(\boldsymbol{h}; z, x; t)^{-1} \Omega(\boldsymbol{h}; z; t)$.

Proof. The inverse is defined by the relation

$$1 = J(\boldsymbol{h}; z, x; t)^{-1} J(\boldsymbol{h}; z, x; t).$$

Taking the differential of both sides gives, with (6-12),

$$0 = (d_z J(\boldsymbol{h}; z, x; t)^{-1}) J(\boldsymbol{h}; z, x; t) + J(\boldsymbol{h}; z, x; t)^{-1} \Omega(\boldsymbol{h}; z; t) J(\boldsymbol{h}; z, x; t).$$

Right multiplication by $J(\boldsymbol{h}; z, x; t)$ gives the relation in the lemma. □

Lemma 6.6. *For fixed x and y, put $K(z) = J(\boldsymbol{h}; z, y; t)^{-1} J(\boldsymbol{h}; z, x; t)$. Then $K(z) = J(\boldsymbol{h}; y, x; t)$.*

Proof. With Lemmas 6.4 and 6.5 we find

$$\begin{aligned} d_z K(z) &= -J(\boldsymbol{h};z,y;t)^{-1} \Omega(\boldsymbol{h};z;t) J(\boldsymbol{h};z,x;t) + J(\boldsymbol{h};z,y;t)^{-1} \Omega(\boldsymbol{h};z;t) J(\boldsymbol{h};z,x;t) \\ &= 0. \end{aligned}$$

Hence $K(z)$ is constant. By (6-8), its value is

$$K(y) = J(\boldsymbol{h}; y, y; t)^{-1} J(\boldsymbol{h}(y, x; t) = J(\boldsymbol{h}; y, x; t).$$ □

Completion of the proof of relation (6-9). The relation

$$K(z) = J(\boldsymbol{h}; y, x; t) = J(\boldsymbol{h}; z, y; t)^{-1} J(\boldsymbol{h}; z, x; t)$$

implies the desired result for $z, y, x \in \mathfrak{H}$. By continuity it holds for $z, y, x \in \mathfrak{H}^*$.

Relation (6-10). This relation follows from (6-8) and (6-9). This ends the proof of Proposition 6.3. □

Noncommutative cocycle. From (6-7) and (6-9) it follows that for any $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$

$$\Psi(\boldsymbol{h})_\gamma := J(\boldsymbol{h}; \gamma^{-1}\infty, \infty; t) \tag{6-13}$$

defines a cocycle $\gamma \to \Psi(\boldsymbol{h})_\gamma$ in $Z^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$.

If we replace ∞ in (6-13) by another base-point $x \in \mathfrak{H}^*$ we get a cocycle in the same cohomology class. So $\boldsymbol{h} \mapsto \Psi(\boldsymbol{h})$ induces a map from $S(\mathcal{A}; \Gamma)$ to the noncommutative cohomology set $H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$, extending the map $\mathrm{Coh}_{\mathcal{A}}$ in Proposition 5.1.

The main result of this paper is the bijectivity of this map:

Theorem 6.7. *Let $\mathcal{A}$ denote the choice of finitely many positive weights w_1+2, $w_2+2, \dots, w_\ell+2$ and corresponding multiplier systems $v_1, \dots, v_\ell$ of Γ. For each noncommutative cohomology class $c \in H^1(\Gamma; N(\mathcal{A}))$ there is a unique element $\boldsymbol{h} \in S(\mathcal{A}; \Gamma)$ such that $\Psi(\boldsymbol{h}) \in c$.*

Proof. The induction runs over $k \geq 0$. We start with a cocycle $\mathsf{X}^0 \in Z^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$, and replace it in the course of an induction procedure by cocycles $\mathsf{X}^1, \mathsf{X}^2, \dots$ in the same cohomology class. During the induction we form a sequence $\boldsymbol{h}_0, \boldsymbol{h}_1, \dots, \boldsymbol{h}_k, \dots$ of elements of $S(\mathcal{A}; \Gamma)$, and a strictly increasing sequence of integers $c_0, c_1, \dots$. The connection between the induction quantities X^k, $\boldsymbol{h}_k$ and c_k is given by the requirement that at each stage of the induction the following conditions hold:

(H) $\boldsymbol{h}_k(B) = 0$ for all B with $d(B) > c_k$.

(XPs) If $\mathsf{X}^k_\gamma - \Psi(\boldsymbol{h}_k)_\gamma = \sum_B a(\gamma, B)B$, where the sum B runs over all noncommutative polynomials in $A_1, \dots, A_\ell$, then for each $\gamma \in \Gamma$

$$a(\gamma, B) = 0 \quad \text{for all } B \text{ with } d(B) \leq k.$$

Either at a certain stage k in the induction procedure the process stops, and we take $\boldsymbol{h} = \boldsymbol{h}_k$, or the process goes on indefinitely, in which case we construct $\boldsymbol{h}$ as a limit of the $\boldsymbol{h}_k$. In both cases we show that $\Psi(\boldsymbol{h})$ is in the cohomology class of the X^k, and that the element $\boldsymbol{h}$ is uniquely determined.

Start of the induction. For a given cocycle X^0 in the cohomology class c we put $\boldsymbol{h}_0 = 0$ and $c_k = 0$. Then for all $\gamma \in \Gamma$

$$\mathsf{X}^0_\gamma - \Psi(\boldsymbol{h}_0)_\gamma = \mathsf{X}^0_\gamma - 1$$

has no constant term, and conditions (H) and (XPs) are trivially satisfied.

Has the end of the induction process been reached? If $\mathsf{X}^k = \Psi(\boldsymbol{h}_k)$, we have found a description of the class c as required in the theorem. This may happen already at the start of the induction if X^0 is the trivial cocycle $\gamma \mapsto 1$.

Induction, choice of c_{k+1}. If the process has not ended, then the difference $\mathsf{Y}^k := \mathsf{X}^k - \Psi(\boldsymbol{h}_k)$ determines a nonzero map $\gamma \mapsto \mathsf{Y}^k_\gamma$ from Γ to $\boldsymbol{O}(\mathcal{A})$. It is not a cocycle.

We define c_{k+1} as the minimum degree such that $\mathsf{Y}^k_\gamma \in \boldsymbol{O}(\mathcal{A})$ has nonzero terms of degree c_{k+1} in $A_1, \dots, A_\ell$ for some $\gamma \in \Gamma$. Since condition (XPs) holds for k we have $c_{k+1} > c_k$.

Cocycle relation. The cocycle relations for the noncommutative cocycles X^k and $\Psi(\boldsymbol{h}_k)$ give

$$\begin{aligned} \mathsf{Y}^k_{\gamma\delta} &= \big((\mathsf{Y}^k_\gamma + \Psi(\boldsymbol{h}_k)_\gamma)|\delta\big)(\mathsf{Y}^k_\delta + \Psi(\boldsymbol{h}_k)_\delta) - (\Psi(\boldsymbol{h}_k)_\gamma|\delta)\Psi(\boldsymbol{h}_k)_\delta \\ &= (\mathsf{Y}^k_\gamma|\delta)\Psi(\boldsymbol{h}_k)_\delta + (\Psi(\boldsymbol{h}_k)_\gamma|\delta)\mathsf{Y}^k_\delta + (\mathsf{Y}^k_\gamma|\delta)\mathsf{Y}^k_\delta. \end{aligned} \tag{6-14}$$

By condition (XPs) and the choice of c_{k+1}, the element $\mathsf{Y}^k_\gamma \in \boldsymbol{O}(\mathcal{A})$ has no terms with degree less than c_{k+1}. We denote by $\bar{\mathsf{Y}}^k_\gamma$ the sum of the terms of Y^k_γ with exact degree c_{k+1}. We consider relation (6-14) modulo terms with degree strictly larger than c_{k+1}:

$$\bar{\mathsf{Y}}^k_{\gamma\delta} \equiv (\bar{\mathsf{Y}}^k_\gamma|\delta)\Psi(\boldsymbol{h}_k)_\delta + (\Psi(\boldsymbol{h}_k)_\gamma|\delta)\bar{\mathsf{Y}}^k_\delta + 0. \tag{6-15}$$

In the two products only the constant term 1 of $\Psi(\boldsymbol{h}_k)_\gamma$ and $\Psi(\boldsymbol{h}_k)_\delta$ is relevant, and we obtain

$$\bar{\mathsf{Y}}^k_{\gamma\delta} = \bar{\mathsf{Y}}^k_\gamma|\delta + \bar{\mathsf{Y}}^k_\delta. \tag{6-16}$$

So the noncommutative cocycle relations for X^k and $\Psi(\boldsymbol{h}_k)$ imply that $\gamma \mapsto \bar{\mathsf{Y}}^k_\gamma$ is a commutative cocycle with values in the additive group of $\boldsymbol{O}(\mathcal{A})$.

The elements $\bar{\mathsf{Y}}^k_\gamma$ have the form

$$\bar{\mathsf{Y}}^k_\gamma = \sum_{n=1}^{K} \varphi^n_\gamma C_n, \qquad C_n = A_{p_{n,1}} A_{p_{n,2}} \cdots A_{p_{n,c_{k+1}}}, \tag{6-17}$$

with $\varphi^n_\gamma \in V(\boldsymbol{v}(C_n), \boldsymbol{w}(C_n))$. The C_n have degree c_{k+1} in $A_1, \dots, A_\ell$. For each n there is some $\gamma \in \Gamma$ for which $\varphi^n_\gamma \neq 0$. Relation (6-16) implies that each component of $\bar{\mathsf{Y}}^k$ is a cocycle: $\varphi^n \in Z^1\big(\Gamma; V(\boldsymbol{v}(C_n), \boldsymbol{w}(C_n))\big)$. By Theorem 2.1 there exist $a_n \in V(\boldsymbol{v}(C_n), \boldsymbol{w}(C_n))$ and unique cusp forms $g_n \in S_{\boldsymbol{w}(C_n)+2}(\Gamma, \boldsymbol{v}(C_n))$ such that for all $\gamma \in \Gamma$

$$\varphi^n_\gamma = -\psi_{g_n,\gamma} + a_n|_{\boldsymbol{v}(C_n),-\boldsymbol{w}(C_n)}(\gamma - 1). \tag{6-18}$$

Induction, choice of X^{k+1}. Take

$$H_k = 1 - \sum_{n=1}^{K} a_n C_n. \tag{6-19}$$

This is an element of $\boldsymbol{N}(\mathcal{A})$. We define the cocycle X^{k+1} in the same class as X^k by

$$\mathsf{X}^{k+1}_\gamma = (H_k|\gamma)\mathsf{X}^k_\gamma H_k^{-1}. \tag{6-20}$$

Induction, choice of $\boldsymbol{h}_{k+1}$. It may happen that $C_n \notin \mathcal{B}(\mathcal{A})$ for some $n \in \{1, \dots, k\}$. Then $S_{\boldsymbol{w}(C_n)+2}(\Gamma, \boldsymbol{v}(C_n)) = 0$, and φ^n is a coboundary and $g_n = 0$.

By condition (H) we have $\boldsymbol{h}_k(C_n) = 0$ for $1 \le n \le K$. We construct $\boldsymbol{h}_{k+1}$ from $\boldsymbol{h}_k$ by taking $\boldsymbol{h}_{k+1}(C_n) = g_n$ for those n for which $C_n \in \mathcal{B}(\mathcal{A})$, and $\boldsymbol{h}_{k+1}(B) = \boldsymbol{h}_k(B)$ otherwise. So $\boldsymbol{h}_k(B) = \boldsymbol{h}_{k+1}(B)$ for all B with $d(B) > c_{k+1}$, and condition (H) stays valid for $k+1$. If $C_n \notin \mathcal{B}(\mathcal{A})$ for all n, then $\boldsymbol{h}_{k+1} = \boldsymbol{h}_k$.

Induction, check of condition (XPs) for $k+1$. Modulo terms of order larger than c_{k+1}, we have

$$\begin{aligned}
\mathsf{X}_\gamma^{k+1} &\equiv \Big(1 - \sum_n a_n|\gamma C_n\Big)(\Psi(\boldsymbol{h}_k)_\gamma + \bar{\mathsf{Y}}_\gamma^k)\Big(1 + \sum_n a_n C_n\Big) && \text{(by (6-20), (XPs))}\\
&\equiv \Psi(\boldsymbol{h}_k)_\gamma + \bar{\mathsf{Y}}_\gamma^k - \sum_n a_n|\gamma C_n + \sum_n a_n C_n\\
&\equiv \Psi(\boldsymbol{h}_k)_\gamma + \sum_n \big(-\psi_{g_n,\gamma} + a_n|(\gamma-1) - a_n|\gamma + a_n\big) C_n && \text{(by (6-17), (6-18))}\\
&= \Psi(\boldsymbol{h}_k)_\gamma + \sum_n R_1(g_n; \gamma^{-1}\infty, \infty) C_n && \text{(by (3-6))}. \quad (6\text{-}21)
\end{aligned}$$

By (6-13) and (6-5), we have

$$\begin{aligned}
&\Psi(\boldsymbol{h}_{k+1})_\gamma\\
&\quad = 1 + \sum_{m \ge 1} \sum_{B_1, \dots, B_m \in \mathcal{B}(\mathcal{A})} R_m(\boldsymbol{h}_{k+1}(B_1), \boldsymbol{h}_{k+1}(B_2), \dots, \boldsymbol{h}_{k+1}(B_m); \gamma^{-1}\infty, \infty; t)\\
&\qquad\qquad \times B_1 B_2 \cdots B_m,
\end{aligned}$$

in which we can leave out the terms in which a B_i occurs with $d(B_i) > c_{k+1}$, by condition (H). If we leave out the terms with a B_i for which $d(B_i) > c_k$, we obtain $\Psi(\boldsymbol{h}_k)_\gamma$. If there is a B_i with $d(B_i) > c_k$ this is one of the C_n in (6-17), with $d(B_i) = d(C_n) = c_{k+1}$. Working modulo terms with degree larger than c_{k+1} we obtain

$$\begin{aligned}
\Psi(\boldsymbol{h}_{k+1})_\gamma - \Psi(\boldsymbol{h}_k)_\gamma &\equiv \sum_{\substack{1 \le n \le K\\ C_n \in \mathcal{B}(\mathcal{A})}} R_1(\boldsymbol{h}_{k+1}(C_n); \gamma^{-1}\infty, \infty; t) C_n\\
&= \sum_{\substack{1 \le n \le K\\ C_n \in \mathcal{B}(\mathcal{A})}} R_1(g_n; \gamma^{-1}\infty, \infty; t) C_n. \qquad (6\text{-}22)
\end{aligned}$$

A comparison of (6-21) and (6-22) gives condition (XPs) for $k+1$.

The induction may halt. It may happen that the induction stops at stage k; namely, if $\mathsf{X}^k = \Psi(\boldsymbol{h}_k)$. Then we have found an element $\boldsymbol{h} = \boldsymbol{h}_k \in S(\mathcal{A}; \Gamma)$ such that $\Psi(\boldsymbol{h})$ is in the class c.

The induction may have infinitely many steps. It may also happen that we have obtained after infinitely many steps an infinite sequence of cocycles X^k in the class c, an infinite sequence of $\boldsymbol{h}_k$, and a strictly increasing sequence of c_k satisfying conditions (H) and (XPs) for all k. For each monomial $B \in \mathcal{B}(\mathcal{A})$ there is at most one k such that $h_n(B) = 0$ for $n \leq k$, and $h_n(B) = h_{k+1}(B)$ for $n \geq k+1$. So the componentwise limit $\boldsymbol{h} := \lim_{k\to\infty} \boldsymbol{h}_k$ exists.

The construction of the sequence $(\mathsf{X}^k)_k$ implies that

$$\mathsf{X}^k_\gamma = ((H_{k-1}H_{k-2}\cdots H_0)|\gamma)\mathsf{X}^0_\gamma(H_{k-1}H_{k-2}\cdots H_0)^{-1},$$

with H_k as in (6-19). The infinite product $H = \cdots H_2 H_1 H_0$ converges in $\boldsymbol{N}(\mathcal{A})$, since each H_k equals 1 plus a term in degree c_{k+1}. Similarly, $J(\boldsymbol{h}; \gamma^{-1}\infty, \infty; t)$ is the limit of the $J(\boldsymbol{h}_k; \gamma^{-1}\infty, \infty; t)$ as $k \to \infty$, since enlarging k we change only terms of degrees larger than c_k. Condition (XPs) is valid for all k, so the conclusion is that, in the limit,

$$(H|\gamma)\mathsf{X}^0 H^{-1} = \Psi(\boldsymbol{h}). \tag{6-23}$$

Uniqueness. Let $\mathsf{X}^0 = \Psi(\boldsymbol{h}')$ for some $\boldsymbol{h}' \in S(\mathcal{A}; \Gamma)$. We claim that, in the induction procedure described above applied to this cocycle X^0, we have at each stage

$$\boldsymbol{h}_k(B) = \boldsymbol{h}'(B) \quad \text{for all } B \in \mathcal{B}(\mathcal{A}) \text{ with } d(B) \leq c_k; \tag{6-24}$$

$$\mathsf{X}^k \equiv \Psi(\boldsymbol{h}') \quad \text{modulo terms of degree larger than } c_k. \tag{6-25}$$

This is true at the start of the induction (use $c_0 = 0$).

At stage k, the nonzero terms with lowest degree in

$$\mathsf{Y}^k_\gamma = J(\boldsymbol{h}'; \gamma^{-1}\infty, \infty; t) - J(\boldsymbol{h}_k; \gamma^{-1}\infty, \infty; t)$$

are due to the $C \in \mathcal{B}(\mathcal{A})$ with degree equal to c_{k+1}. So

$$\bar{\mathsf{Y}}_\gamma = \sum_{\substack{C \in \mathcal{B}(\mathcal{A}) \\ d(C) = c_{k+1}}} R_1(\boldsymbol{h}'(B); \gamma^{-1}\infty, \infty; t)C. \tag{6-26}$$

Let us number the monomials in this sum as $C_1, \ldots, C_K$. Then φ^n_γ in (6-18) is equal to $-\psi_{\boldsymbol{h}'(C_n),\gamma}$, and $a_n = 0$, $H_k = 1$. This implies that $\boldsymbol{h}_{k+1}(C) = \boldsymbol{h}'(C)$ for the monomials $C \in \mathcal{B}(\mathcal{A})$ with degree c_{k+1}, and $\mathsf{X}^{k+1} = \mathsf{X}^k \equiv \Psi(\boldsymbol{h}')$ modulo terms with degree larger than c_{k+1}.

At the end of the induction process we have $\mathcal{B} = \mathcal{B}'$, thus obtaining uniqueness. □

Concluding remarks. (a) Manin [2005; 2006] used formal series similar to those in (4-4) to get a simple description of relations among iterated integrals. In that approach the noncommutative cohomology set $H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$ is a tool. In this paper we further study the cohomology set $H^1(\Gamma; \boldsymbol{N}(\mathcal{A}))$.

(b) One may apply the approach of this paper to weights in $\mathbb{Z}_{\geq 2}$ and trivial multiplier systems. Then the iterated integrals are polynomial functions. These are in a much smaller Γ-module than the functions with polynomial growth that we employ. The consequence is that the theorem analogous to the theorem of Knopp and Mawi (Theorem 2.1) does not hold. Cocycles attached to conjugates of holomorphic cusp forms have to be considered as well (see [Knopp 1974]). However, iterated integrals in which occur both holomorphic and antiholomorphic cusp forms satisfy more complicated decomposition relations. We think that Manin's formalism does not work in that situation.

(c) The same problem occurs if we use the modules in Theorems B and D of [Bruggeman et al. 2014], unless we pick the weights $w_j + 2$ in such a way that the elements $\boldsymbol{w}(C)$ that occur in the sums defining $J(\mathcal{B}; y, x; t)$ are never in $\mathbb{Z}_{\geq 0}$.

(d) We work with iterated integrals of the type in (3-1). Equation (1-1) defines iterated integrals depending on variables $t_1, \dots, t_\ell$ all running independently through the lower half-plane. It would be nice to have results for the corresponding noncommutative cocycles. These cocycles can be defined, and one can show injectivity of the map from cusp forms to cohomology, like in Proposition 5.1. We did not manage to adapt the proof of Theorem 6.7 to cocycles of this type. The problem is to construct formal sequences of the type in (6-5) such that they have the lowest degree terms in a prescribed summand in the decomposition of the tensor products $V(v_1, w_1) \otimes \cdots \otimes V(v_\ell, w_\ell)$ into submodules.

Acknowledgement

The authors would like to thank the referees for numerous helpful comments and suggestions which greatly improved the exposition of this paper, in particular the introduction.

References

[Brown 2014] F. Brown, "Motivic periods and $\mathbb{P}^1 \setminus \{0, 1, \infty\}$", preprint, 2014. arXiv 1407.5165

[Bruggeman et al. 2014] R. Bruggeman, Y. Choie, and N. Diamantis, "Holomorphic automorphic forms and cohomology", preprint, 2014. arXiv 1404.6718

[Chen 1977] K. T. Chen, "Iterated path integrals", *Bull. Amer. Math. Soc.* **83**:5 (1977), 831–879. MR 0454968 Zbl 0389.58001

[Choie 2014] Y. Choie, "Parabolic cohomology and multiple Hecke L-values", preprint, Pohang University of Science and Technology, 2014, Available at http://math.postech.ac.kr/~yjc/para-Choie.pdf.

[Hain 2015] R. Hain, "The Hodge–de Rham theory of modular groups", preprint, 2015. arXiv 1403.6443

[Knopp 1974] M. I. Knopp, "Some new results on the Eichler cohomology of automorphic forms", *Bull. Amer. Math. Soc.* **80** (1974), 607–632. MR 0344454 Zbl 0292.10022

[Knopp and Mawi 2010] M. Knopp and H. Mawi, "Eichler cohomology theorem for automorphic forms of small weights", *Proc. Amer. Math. Soc.* **138**:2 (2010), 395–404. MR 2557156 Zbl 1233.11058

[Manin 2005] Y. I. Manin, "Iterated Shimura integrals", *Mosc. Math. J.* **5**:4 (2005), 869–881, 973. MR 2266463 Zbl 1215.11052

[Manin 2006] Y. I. Manin, "Iterated integrals of modular forms and noncommutative modular symbols", pp. 565–597 in *Algebraic geometry and number theory*, edited by V. Ginzburg, Progr. Math. **253**, Birkhäuser, Boston, 2006. MR 2263200 Zbl 1184.11019

[Provost 2014] N. Provost, *Valeurs multiples de fonctions L de formes modulaires*, thèse de doctorat, Université Paris Diderot, 2014. arXiv 1604.01913

Communicated by Yuri Manin
Received 2015-06-29 Revised 2016-02-26 Accepted 2016-02-28

r.w.bruggeman@uu.nl *Mathematisch Instituut, Universiteit te Utrecht, Postbus 80010, 3508 TA Utrecht, Netherlands*

yjc@postech.ac.kr *Department of Mathematics and PMI, Postech, Pohang 790-784, South Korea*

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.665

The existential theory of equicharacteristic henselian valued fields

Sylvy Anscombe and Arno Fehm

We study the existential (and parts of the universal-existential) theory of equicharacteristic henselian valued fields. We prove, among other things, an existential Ax–Kochen–Ershov principle, which roughly says that the existential theory of an equicharacteristic henselian valued field (of arbitrary characteristic) is determined by the existential theory of the residue field; in particular, it is independent of the value group. As an immediate corollary, we get an unconditional proof of the decidability of the existential theory of $\mathbb{F}_q((t))$.

1. Introduction

We study the first order theory of a henselian valued field (K, v) in the language of valued fields. For residue characteristic zero, this theory is well-understood through the celebrated *Ax–Kochen–Ershov* (*AKE*) *principles*, which state that, in this case, the theory of (K, v) is completely determined by the theory of the residue field Kv and the theory of the value group vK (see, e.g., [Prestel and Delzell 2011, §4.6]). In other words, if a sentence holds in one such valued field, then it holds in any other with elementarily equivalent residue field and value group (the *transfer principle*). As a consequence, one gets that the theory of (K, v) is *decidable* if and only if the theory of the residue field and the theory of the value group are decidable.

Some of this theory can be carried over to certain mixed characteristic henselian valued fields such as the fields of p-adic numbers $\mathbb{Q}_p$, whose theory was axiomatised and proven to be decidable by Ax–Kochen and Ershov in 1965. However, for henselian valued fields of positive characteristic, no such general principles are available. For example, in [Kuhlmann 2001], it is shown that the theory of characteristic $p > 0$ henselian valued fields with value group elementarily equivalent to $\mathbb{Z}$ and residue field $\mathbb{F}_p$ is incomplete. It is not known whether there is a suitable modification of the AKE principles that hold for arbitrary henselian valued fields

During this research, Anscombe was funded by EPSRC grant EP/K020692/1.
MSC2010: primary 03C60; secondary 12L12, 12J10, 11U05, 12L05.
Keywords: model theory, henselian valued fields, decidability, diophantine equations.

of positive characteristic, and the decidability of the field of formal power series $\mathbb{F}_q((t))$ is a long-standing open problem.

For the first problem, the most useful approximations are AKE principles for certain classes of valued fields, most notably F.-V. Kuhlmann's recently published work [2014] on the model theory of *tame fields*. For the second problem, the best known result is by Denef and Schoutens [2003], who proved that resolution of singularities in positive characteristic would imply that the *existential* theory of $\mathbb{F}_q((t))$ is decidable (i.e., Hilbert's tenth problem for $\mathbb{F}_q((t))$ has a positive solution).

In this work, we take a different approach at deepening our understanding of the positive characteristic case: instead of limiting ourselves to certain classes of valued fields, we attempt to prove results for arbitrary equicharacteristic henselian valued fields, but (having results like Denef–Schoutens in mind) instead restrict to existential or slightly more general sentences. The technical heart of this work is a study of transfer principles for certain universal-existential sentences, which builds on the aforementioned [Kuhlmann 2014]; see the results in Section 5. While some of these general results will have applications for example in the theory of definable valuations (see [Anscombe and Koenigsmann 2014; Cluckers et al. 2013; Fehm 2015; Prestel 2015] for some of the recent developments), in this work we then restrict this machinery to existential sentences and deduce the following result (cf. Theorem 6.5):

Theorem 1.1. *For any field F, the theory T of equicharacteristic henselian nontrivially valued fields with residue field which models both the existential and universal theories of F is $\exists$-complete, i.e., for any existential sentence ϕ either $T \models \phi$ or $T \models \neg\phi$.*

Note that the value group plays no role here: the existential theory of an equicharacteristic henselian nontrivially valued field is determined solely by its residue field. From this theorem, we obtain an AKE principle for $\exists$-sentences (cf. Corollary 7.2):

Corollary 1.2. *Let (K, v), (L, w) be equicharacteristic henselian nontrivially valued fields. If the residue fields Kv and Lw have the same existential theory, then so do the valued fields (K, v) and (L, w).*

Moreover, we conclude the following corollary on decidability (cf. Corollary 7.5):

Corollary 1.3. *Let (K, v) be an equicharacteristic henselian valued field. The following are equivalent:*

(1) *The existential theory of Kv in the language of rings is decidable.*

(2) *The existential theory of (K, v) in the language of valued fields is decidable.*

As an immediate consequence, we get the first unconditional proof of the decidability of the existential theory of $\mathbb{F}_q((t))$ (cf. Corollary 7.7). Note, however, that the conditional result in [Denef and Schoutens 2003] is for a language with a constant for t — Section 7 also contains a brief discussion of this difference.

As indicated above, these results are essentially known in residue characteristic zero (cf. Remark 7.3), but are new in positive characteristic. However, each of the above results fails if "equicharacteristic" is dropped or replaced by "mixed characteristic", in contrast to the mixed characteristic AKE principles mentioned above (cf. Remark 7.4 and Remark 7.6).

2. Valued fields

For a valued field (K, v) we denote by $vK = v(K^\times)$ its value group, by $\mathcal{O}_v$ its valuation ring, and by $Kv = \{av \mid a \in \mathcal{O}_v\}$ its residue field. For standard definitions and facts about henselian valued fields we refer the reader to [Engler and Prestel 2005]. As a rule, if L/K is a field extension to which the valuation v can be extended uniquely, we denote also this unique extension by v. This applies in particular if v is henselian, and for the perfect hull $L = K^{\mathrm{perf}}$ of K. We will make use of the following well-known fact:

Lemma 2.1. *Let (K, v) be a valued field and let F/Kv be any field extension. Then there is an extension of valued fields $(L, w)/(K, v)$ such that Lw/Kv is isomorphic to the extension F/Kv.*

Proof. See, e.g., [Kuhlmann 2004, Theorem 2.14]. □

The next lemma is also probably well known, but for lack of reference we sketch a proof, which closely follows [Kuhlmann 2011, Lemma 9.30].

Definition 2.2. Let (K, v) be a valued field. A *partial section* (of the residue homomorphism) is a map $f : E \to K$, for some subfield $E \subseteq Kv$, which is an $\mathcal{L}_{\mathrm{ring}}$-embedding such that $(f(a))v = a$ for all $a \in E$. It is a *section* if $E = Kv$.

Lemma 2.3. *Let (K, v) be an equicharacteristic henselian valued field, let $E \subseteq Kv$ be a subfield of the residue field, and suppose that there is a partial section $f : E \to K$. If F/E is a separably generated subextension of Kv/E then we may extend f to a partial section $F \to K$.*

Proof. Write $L_1 := f(E)$. Let T be a separating transcendence base for F/E and, for each $t \in T$, choose $s_t \in K$ such that $s_t v = t$. Then $S := \{s_t \mid t \in T\}$ is algebraically independent over L_1. Thus we may extend f to a partial section $E(T) \to L_1(S)$ by sending $t \mapsto s_t$.

Let L_2 be the relative separable algebraic closure of $L_1(S)$ in K. By Hensel's lemma, $L_2 v$ is separably algebraically closed in Kv. Thus F is contained in $L_2 v$. Since v is trivial on L_2, the restriction of the residue map to L_2 is an isomorphism $L_2 \to L_2 v$. Thus the restriction to F of the inverse of the residue map is a partial section $F \to K$ which extends f, as required. □

Recall that a valued field (K, v) of residue characteristic p is *tame* if it is henselian, the value group vK is p-divisible, the residue field Kv is perfect, and (K, v) is defectless, i.e., for every finite extension L/K,

$$[L : K] = [Lv : Kv] \cdot [vL : vK].$$

Proposition 2.4. *Let (K, v) be a valued field. There exists an extension (K^t, v^t) of (K, v) such that (K^t, v^t) is tame, K^t is perfect, $v^t K^t = \frac{1}{p^\infty} vK$, and $K^t v^t = Kv^{\mathrm{perf}}$.*

Proof. In the special case $\mathrm{char}(K) = \mathrm{char}(Kv)$, any maximal immediate extension of K^{perf} satisfies the claim. In general, [Kuhlmann et al. 1986, Theorem 2.1, Proposition 4.1, and Proposition 4.5(i)] gives such a K^t that is in addition algebraic over K. □

3. Model theory of valued fields

Let

$$\mathcal{L}_{\mathrm{ring}} = \{+, -, \cdot, 0, 1\}$$

be the language of rings and let

$$\mathcal{L}_{\mathrm{vf}} = \{+^K, -^K, \cdot^K, 0^K, 1^K, +^\Gamma, <^\Gamma, 0^\Gamma, \infty^\Gamma, +^k, -^k, \cdot^k, 0^k, 1^k, v, \mathrm{res}\}$$

be a three sorted language for valued fields (like the Denef–Pas language, but without an angular component) with a sort K for the field itself, a sort $\Gamma \cup \{\infty\}$ for the value group with infinity, and a sort k for the residue field, as well as both the valuation map v and the residue map res, which we interpret as the constant 0^k map outside the valuation ring. For a field C, we let $\mathcal{L}_{\mathrm{ring}}(C)$ and $\mathcal{L}_{\mathrm{vf}}(C)$ be the languages obtained by adding symbols for elements of C. In the case of $\mathcal{L}_{\mathrm{vf}}(C)$, the constant symbols are added to the field sort K.

A valued field (K, v) gives rise in the usual way to an $\mathcal{L}_{\mathrm{vf}}$-structure

$$(K, vK \cup \{\infty\}, Kv, v, \mathrm{res}),$$

where vK is the value group, Kv is the residue field, and res is the residue map. For notational simplicity, we will usually write (K, v) to refer to the $\mathcal{L}_{\mathrm{vf}}$-structure it induces. For further notational simplicity, we write (K, D) instead of $(K, (d_c)_{c\in C})$, where $D = \{d_c \mid c \in C\}$ is the set of interpretations of the constant symbols. Combining these two simplifications, we write (K, v, D) for the $\mathcal{L}_{\mathrm{vf}}(C)$-structure

$$(K, vK \cup \{\infty\}, Kv, v, \mathrm{res}, (d_c)_{c\in C}).$$

We also write Dv for the set of residues of elements from D.

As usual, we say that an $\mathcal{L}_{\mathrm{vf}}(C)$-formula is an $\exists$-*formula* if it is logically equivalent to a formula in prenex normal form with only existential quantifiers (over any of the three sorts). We say that an $\mathcal{L}_{\mathrm{vf}}(C)$-sentence is an $\forall^k\exists$-*sentence* if it is logically equivalent to a sentence of the form $\forall \boldsymbol{x}\, \psi(\boldsymbol{x})$, where ψ is an $\exists$-formula and the universal quantifiers range over the residue field sort.

Let $(K, v, D) \subseteq (L, w, E)$ be an extension of $\mathcal{L}_{\mathrm{vf}}(C)$-structures. Note that $d_c = e_c$ for all $c \in C$. We say that certain $\mathcal{L}_{\mathrm{vf}}(C)$-sentences ϕ *go up* from K to L if $(K, v, D) \models \phi$ implies that $(L, w, E) \models \phi$. For examples, $\exists$-sentences always go up every extension. Furthermore, if $(L, w)/(K, v)$ is an extension of valued fields such that Lw/Kv is trivial, then $\forall^k\exists$-$\mathcal{L}_{\mathrm{vf}}(K)$-sentences go up from (K, v) to (L, w). Although the previous statement is not referenced directly, it underlies many of the arguments in Section 5.

Lemma 3.1. *Let L/K be an extension of fields. If $K \preceq_\exists L$, then $K^{\mathrm{perf}} \preceq_\exists L^{\mathrm{perf}}$.*

Proof. This is clear, since $K^{\mathrm{perf}} = \bigcup_n K^{p^{-n}}$ and $L^{\mathrm{perf}} = \bigcup_n L^{p^{-n}}$, and the Frobenius gives that $K^{p^{-n}} \preceq_\exists L^{p^{-n}}$ for all n. □

F.-V. Kuhlmann [2014] proves the following on the model theory of tame fields:

Proposition 3.2. *The elementary class of tame fields has the relative embedding property. That is, for tame fields (K, v) and (L, w) with common subfield (F, u), if*

(1) *(F, u) is defectless,*

(2) *(L, w) is $|K|^+$-saturated,*

(3) *vK/uF is torsion-free and Kv/Fu is separable, and*

(4) *there are embeddings $\rho : vK \to wL$ (over uF) and $\sigma : Kv \to Lw$ (over Fu),*

then there exists an embedding $\iota : (K, v) \to (L, w)$ over (F, u) which respects ρ and σ.

Proof. See [Kuhlmann 2014, Theorem 7.1]. (Note that this result is stated in the language

$$\mathcal{L}'_{\mathrm{vf}} = \{+, -, \cdot, {}^{-1}, 0, 1, O\},$$

where O is a binary predicate which is interpreted in a valued field (K, v) so that $O(a, b)$ if and only if $va \geq vb$. However, the exact choice of language does not directly affect us.) □

From Proposition 3.2, Kuhlmann deduces the following AKE principle:

Theorem 3.3. *The class of tame fields is an* AKE$^{\preceq}$*-class: if $(L, w)/(K, v)$ is an extension of tame fields with $vK \preceq wL$ and $Kv \preceq Lw$, then $(K, v) \preceq (L, w)$.*

Proof. See [Kuhlmann 2014, Theorem 1.4]. □

4. Power series fields

For a field F and an ordered abelian group Γ we denote by $F((\Gamma))$ the field of generalised power series with coefficients in F and exponents in Γ; see, e.g., [Efrat 2006, §4.2]. We identify $F((\mathbb{Z}))$ with the field of formal power series $F((t))$ and denote the power series valuation on any subfield of any $F((\Gamma))$ by v_t.

Lemma 4.1. *A field $(F((\Gamma)), v_t)$ of generalised power series is maximal. In particular, it is tame if and only if F is perfect and Γ is p-divisible.*

Proof. See [Efrat 2006, Theorem 18.4.1] and note that maximal implies henselian and defectless. □

Proposition 4.2. *Let A be a complete discrete (i.e., with value group $\mathbb{Z}$) equicharacteristic valuation ring. Let $F \subseteq A$ be a set of representatives for the residue classes which forms a field. Let $s \in A$ be a uniformiser (i.e., an element of least positive value). Then A is isomorphic to $F[\![s]\!]$ by an isomorphism which fixes F pointwise.*

Proof. See [Serre 1979], Chapter 2 Proposition 5 and the discussion following the example. □

Corollary 4.3. *Let F be a field and let $E/F((t))$ be a finite extension such that $Ev_t = F$. Then (E, v_t, F) is isomorphic to $(F((s)), v_s, F)$. This applies in particular to finite extensions of $F((t))$ inside $F((\mathbb{Q}))$.*

Proof. We are already provided with a section since $F \subseteq F((t)) \subseteq E$ and $Ev_t = F$. Since $E/F((t))$ is finite, E is also a complete discrete equicharacteristic valued field (cf. [Serre 1979, Chapter 2 Proposition 3]). By Proposition 4.2, there is an F-isomorphism of valued fields $E \to F((s))$. □

Definition 4.4. We denote by $F(t)^h$ the henselization of $F(t)$ with respect to v_t, i.e., the relative algebraic closure of $F(t)$ in $F((t))$, and by $F((t))^{\mathbb{Q}}$ the relative algebraic closure of $F((t))$ in $F((\mathbb{Q}))$.

Lemma 4.5. *For any field F we have $(F(t)^h, v_t) \preceq_{\exists} (F((t)), v_t)$.*

Proof. See [Kuhlmann 2014, Theorem 5.12]. □

The following proposition may be deduced from the more general [Kuhlmann 2014, Lemma 3.7], but we give a proof in this special case for the convenience of the reader.

Proposition 4.6. *If F is perfect, then $F((t))^{\mathbb{Q}}$ is tame.*

Proof. We have that $F((t))^{\mathbb{Q}} v_t = F$ is perfect and $v_t F((t))^{\mathbb{Q}} = \mathbb{Q}$ is p-divisible. Moreover, as an algebraic extension of the henselian field $F((t))$, $F((t))^{\mathbb{Q}}$ is henselian. It remains to show that $F((t))^{\mathbb{Q}}$ is defectless.

Let $E/F((t))^{\mathbb{Q}}$ be a finite extension of degree n. Since $F((\mathbb{Q}))$ is perfect, so is $F((t))^{\mathbb{Q}}$, and hence $F((\mathbb{Q}))/F((t))^{\mathbb{Q}}$ is regular. Therefore, if $E' = F((\mathbb{Q})) \cdot E$

denotes the compositum of $F((\mathbb{Q}))$ and E in an algebraic closure of $F((\mathbb{Q}))$, then $[E' : F((\mathbb{Q}))] = n$. Since $F((\mathbb{Q}))$ is maximal (Lemma 4.1), $E'/F((\mathbb{Q}))$ is defectless. So since $(F((\mathbb{Q})), v_t)$ is henselian and $v_t F((\mathbb{Q})) = \mathbb{Q}$ is divisible, we get that $[E'v_t : F] = n$. Since $E'v_t/F$ is separable, we can assume without loss of generality that $F' := E'v_t \subseteq E'$ (Lemma 2.3).

$$\begin{tikzcd}
F((\mathbb{Q})) \arrow[r, dash, "n"] \arrow[d, dash, "\text{reg.}"'] & E' \arrow[r, equal] \arrow[d, dash, "\text{reg.}"'] & E' \arrow[d, dash] \\
F((t))^{\mathbb{Q}} \arrow[r, dash, "n"] \arrow[d, dash] & E \arrow[r, dash] & EF' \arrow[d, dash] \\
F \arrow[rr, dash, "n"] & & F'
\end{tikzcd}$$

The extension E'/E is also regular, since $E/F((t))^{\mathbb{Q}}$ is algebraic. In particular, E is relatively algebraically closed in E'; so since EF'/E is algebraic we have that $F' \subseteq E$. Thus $Ev_t = F'$, which shows that $E/F((t))^{\mathbb{Q}}$ is defectless. $\square$

In particular, Theorem 3.3 implies that $F((t))^{\mathbb{Q}} \preceq F((\mathbb{Q}))$. We therefore get the following picture:

$$\begin{tikzcd}
F(t) \arrow[r, dash, "\text{alg.}"] & F(t)^h \arrow[r, dash, "\preceq_\exists"] & F((t)) \arrow[r, dash, "\text{alg.}"] & F((t))^{\mathbb{Q}} \arrow[r, dash, "\preceq"] & F((\mathbb{Q}))
\end{tikzcd}$$

5. The transfer of universal-existential sentences

Throughout this section F/C will be a separable extension of fields of characteristic p. We show that the truth of $\forall^k\exists$-sentences transfers between various valued fields. Usually the valued fields considered will have only elementarily equivalent residue fields. However, for convenience, we will sometimes discuss $\exists$-sentences with additional parameters from the residue field.

Lemma 5.1 (going down from $\boldsymbol{F}((\Gamma))$). *Suppose that F is perfect. Let ϕ be an $\exists$-$\mathcal{L}_{\mathrm{vf}}(F)$-sentence, let $F \preceq \boldsymbol{F}$ be an elementary extension, and let Γ be an ordered abelian group. If $(\boldsymbol{F}((\Gamma)), v_t, F) \models \phi$, then $(F(t)^h, v_t, F) \models \phi$.*

Proof. Without loss of generality we may assume that Γ is nontrivial. For notational simplicity, we suppress the parameters F from the notation. Let Δ be the divisible hull of Γ. Then $(\boldsymbol{F}((\Gamma)), v_t) \subseteq (\boldsymbol{F}((\Delta)), v_t)$, and existential sentences "go up", so $(\boldsymbol{F}((\Delta)), v_t) \models \phi$.

Choose an embedding of $\mathbb{Q}$ into Δ; this induces an embedding $(F((\mathbb{Q})), v_t) \subseteq (\boldsymbol{F}((\Delta)), v_t)$, and therefore $(F((t))^{\mathbb{Q}}, v_t) \subseteq (\boldsymbol{F}((\Delta)), v_t)$. Since the theory of divisible ordered abelian groups is model complete (see, e.g., [Prestel and Delzell 2011, Thm. 4.1.1]),

$$v_t F((t))^{\mathbb{Q}} = \mathbb{Q} \preceq \Delta = v_t \boldsymbol{F}((\Delta)).$$

Moreover,

$$F((t))^{\mathbb{Q}}v_t = F \preceq \boldsymbol{F} = \boldsymbol{F}((\Delta))v_t.$$

Thus, since $(F((t))^{\mathbb{Q}}, v_t)$ is tame by Proposition 4.6 and $(\boldsymbol{F}((\Delta)), v_t)$ is tame by Lemma 4.1, Theorem 3.3 implies that

$$(F((t))^{\mathbb{Q}}, v_t) \preceq (\boldsymbol{F}((\Delta)), v_t).$$

Therefore, $(F((t))^{\mathbb{Q}}, v_t) \models \phi$.

Let E be a finite extension of $F((t))$ that contains witnesses to the truth of ϕ in $(F((t))^{\mathbb{Q}}, v_t)$. Thus $(E, v_t) \models \phi$. By Corollary 4.3, there is an $\mathcal{L}_{\mathrm{vf}}(F)$-isomorphism

$$f : (E, v_t) \to (F((t)), v_t).$$

Thus $(F((t)), v_t) \models \phi$. By Lemma 4.5,

$$(F(t)^h, v_t) \preceq_\exists (F((t)), v_t),$$

hence $(F(t)^h, v_t) \models \phi$, as claimed. □

Definition 5.2. Let $\boldsymbol{H}(F/C)$ be the class of tuples (K, v, D, i), where (K, v, D) is an $\mathcal{L}_{\mathrm{vf}}(C)$-structure and $i : F \to Kv$ is a map such that

(1) (K, v) is an equicharacteristic henselian nontrivially valued field,

(2) $c \mapsto d_c$ is an $\mathcal{L}_{\mathrm{ring}}$-embedding $C \to K$,

(3) the valuation is trivial on D, and

(4) $i : (F, C) \to (Kv, Dv)$ is an $\mathcal{L}_{\mathrm{ring}}(C)$-embedding.

Lemma 5.3 (going up from $F(t)^h$). *Let ϕ be an $\exists$-$\mathcal{L}_{\mathrm{vf}}$-sentence with parameters from C and the residue sort of $(F(t)^h, v_t)$, and suppose that $(F(t)^h, v_t, C) \models \phi$. Then, for all $(K, v, D, i) \in \boldsymbol{H}(F/C)$, we have that $(K, v, D) \models \phi$ (where we replace the parameters from the residue sort by their images under the map i).*

Proof. Write $\phi = \exists \boldsymbol{x}\ \psi(\boldsymbol{x}; \boldsymbol{c}, \beta)$ for some quantifier-free formula ψ and parameters $\boldsymbol{c}$ from C and β from $F(t)^h v_t$. Note that the variables in the tuple $\boldsymbol{x}$ may be from any sorts. Let $\boldsymbol{a}$ be such that

$$(F(t)^h, v_t, C) \models \psi(\boldsymbol{a}; \boldsymbol{c}, \beta).$$

Since $F(t)^h$ is the directed union of fields $E_0(t)^h$ for finitely generated subfields E_0 of F, there exists a subfield E of F containing C such that E/C is finitely generated, $\boldsymbol{a} \in E(t)^h$, and $\beta \in E(t)^h v_t$. Thus

$$(E(t)^h, v_t, C) \models \psi(\boldsymbol{a}; \boldsymbol{c}, \beta).$$

Since F/C is separable and E/C is finitely generated, E is separably generated over C. Thus $i(E)/Dv$ is separably generated. Note that the map $Dv \to D$ given by $d_c v \mapsto d_c$ is a partial section. By Lemma 2.3 we may extend it to a partial section $g : i(E) \to K$. Let $h := g \circ i|_E$ be the composition. Then

$$h : (E, v_0, C) \to (K, v, D)$$

is an $\mathcal{L}_{\mathrm{vf}}(C)$-embedding, where v_0 denotes the trivial valuation on E:

$$\begin{tikzcd}
 & Kv & K \arrow[l, dotted, "\mathrm{res}"'] \arrow[dd, dash] \\
F \arrow[r, "i"] \arrow[d, dash] & i(F) \arrow[u, dash] \arrow[d, dash] & \\
E \arrow[r, "i|_E"] \arrow[d, dash] & i(E) \arrow[r, "g"] \arrow[d, dash] & h(E) \arrow[d, dash] \\
C \arrow[r, "\cong"] & Dv \arrow[r, "\cong"] & D
\end{tikzcd}$$

Since (K, v) is nontrivial, there exists $s \in K^\times$ with $v(s) > 0$, which must be transcendental over $h(E)$, since v is trivial on $h(E)$. As the rational function field $E(t)$ admits (up to equivalence) only one valuation which is trivial on E and positive on t, we may extend h to an $\mathcal{L}_{\mathrm{vf}}(C)$-embedding

$$h' : (E(t), v_t, C) \to (K, v, D)$$

by sending $t \mapsto s$. Since (K, v) is henselian, there is a unique extension of h' to an $\mathcal{L}_{\mathrm{vf}}(C)$-embedding

$$h'' : (E(t)^h, v_t, C) \to (K, v, D).$$

So, since existential sentences "go up",

$$(K, v, D) \models \psi(h''(\boldsymbol{a}); h''(\boldsymbol{c}), h''(\beta)),$$

and thus $(K, v, D) \models \phi$, as claimed. $\square$

Definition 5.4. We let $R_{F/C}$ be the $\mathcal{L}_{\mathrm{ring}}(C)$-theory of F and let $R^1_{F/C}$ be the subtheory consisting of existential and universal sentences. Let $\boldsymbol{T}_{F/C}$ (respectively, $\boldsymbol{T}^1_{F/C}$) be the $\mathcal{L}_{\mathrm{vf}}(C)$-theory consisting of the following axioms (expressed informally about a structure (K, v, D)):

(1) (K, v) is an equicharacteristic henselian nontrivially valued field,

(2) $c \mapsto d_c$ is an $\mathcal{L}_{\mathrm{ring}}$-embedding $C \to K$,

(3) the valuation v is trivial on D, and

(4) (Kv, Dv) is a model of $R_{F/C}$ (respectively, $R^1_{F/C}$).

The "1" is intended to suggest that the sentences considered contain only one type of quantifier. Note that for any $(K, v, D) \models \boldsymbol{T}^1_{F/C}$, the map $d_c v \mapsto d_c$ is a partial section of the residue map. Let ϕ be an $\forall^k\exists$-sentence and write $\phi = \forall^k \boldsymbol{x}\ \psi(\boldsymbol{x})$ for some $\exists$-formula $\psi(\boldsymbol{x})$ with free variables $\boldsymbol{x}$ belonging to the residue field sort. Let ${}^{\boldsymbol{x}}Kv$ denote the set of $\boldsymbol{x}$-tuples from Kv. Then we observe that $(K, v, D) \models \phi$ if and only if ${}^{\boldsymbol{x}}Kv \subseteq \psi(K)$. In this next proposition we show that, roughly, if $\boldsymbol{T}_{F/C}$ is consistent with the property "${}^{\boldsymbol{x}}F \subseteq \psi$" then in fact $\boldsymbol{T}_{F^{\mathrm{perf}}/C^{\mathrm{perf}}}$ entails "${}^{\boldsymbol{x}}F \subseteq \psi$".

Proposition 5.5 (main proposition). *Let $\psi(\boldsymbol{x})$ be an $\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-formula with free variables $\boldsymbol{x}$ belonging to the residue field sort. Suppose there exists*

$$(K, v, D) \models \boldsymbol{T}_{F/C} \cup \{\forall^k \boldsymbol{x}\ \psi(\boldsymbol{x})\}.$$

Then, for all $(L, w, E, i) \in \boldsymbol{H}(F^{\mathrm{perf}}/C^{\mathrm{perf}})$, we have ${}^{\boldsymbol{x}}i(F) \subseteq \psi(L)$.

Proof. Since (K, v, D) models $\boldsymbol{T}_{F/C}$, we have $(Kv, Dv) \equiv (F, C)$. Passing, if necessary, to an elementary extension of (K, v, D), there is an elementary embedding

$$f : (F, C) \overset{\preceq}{\to} (Kv, Dv).$$

As noted after the definition of $\boldsymbol{T}_{F/C}$, the map $g_0 : Dv \to D$ given by $d_c v \mapsto d_c$ is a partial section. Since F/C is separable, $f(F)/Dv$ is also separable. Thus any finitely generated subextension of $f(F)/Dv$ is separably generated. By Lemma 2.3 we may pass again, if necessary, to an elementary extension and extend g_0 to a partial section $g : f(F) \to K$. Note that g is also an $\mathcal{L}_{\mathrm{ring}}(C)$-embedding $(f(F), Dv) \to (K, D)$.

Let $h := g \circ f$. Then $h : (F, C) \to (K, D)$ is an $\mathcal{L}_{\mathrm{ring}}(C)$-embedding. Because g is a section, the valuation v is trivial when restricted to the image of h. Thus, if v_0 denotes the trivial valuation on F, the map h is an $\mathcal{L}_{\mathrm{vf}}(C)$-embedding $(F, v_0, C) \to (K, v, D)$. The induced embedding of residue fields $\bar{h} : Fv_0 \to Kv$ is the composition of the elementary embedding f with an isomorphism. Thus $\bar{h} : Fv_0 \to Kv$ is an elementary embedding. From now on we identify (F, v_0, C) with its image under h as a substructure of (K, v, D), noting that the residue field extension is an elementary extension.

$$\begin{tikzcd}
& Kv & K \arrow[l, dotted, "\mathrm{res}"'] \\
F \arrow[r, "f"] \arrow[d, dash] & f(F) \arrow[r, "g"] \arrow[u, dash] \arrow[d, dash, "\mathrm{sep}"'] & h(F) \arrow[u, dash] \arrow[d, dash] \\
C \arrow[r, "\cong"] & Dv \arrow[r, "g_0"] & D
\end{tikzcd}$$

Choose an extension $(K^t, v^t)/(K, v)$ as in Proposition 2.4. Since K^t is perfect, we can embed D^{perf} into K^t over D so that $(K^t, v^t, D^{\mathrm{perf}})$ is an $\mathcal{L}_{\mathrm{vf}}(C^{\mathrm{perf}})$-structure. Furthermore $(F^{\mathrm{perf}}, v_0, C^{\mathrm{perf}})$ is naturally (identified with) a substructure of $(K^t, v^t, D^{\mathrm{perf}})$. Since $Fv_0 \preceq Kv$, Lemma 3.1 gives that

$$F^{\mathrm{perf}}v_0 = Fv_0^{\mathrm{perf}} \preceq_\exists Kv^{\mathrm{perf}} = K^t v^t.$$

Thus there is an elementary extension $F^{\mathrm{perf}}v_0 \preceq \boldsymbol{F}$ and an embedding $\sigma : K^t v^t \to \boldsymbol{F}$ over $F^{\mathrm{perf}}v_0$; see the diagram below.

Now we consider the two valued fields (K^t, v^t) and $(\boldsymbol{F}((v^t K^t)), v_t)$ with common subfield (F^{perf}, v_0). Note that K^t is tame by definition, and $\boldsymbol{F}((v^t K^t))$ is tame by Lemma 4.1. As a trivially valued field, (F^{perf}, v_0) is defectless. The extension of value groups $v^t K^t / v_0 F^{\mathrm{perf}}$ is isomorphic to $v^t K^t$, thus it is torsion-free. The extension $K^t v^t / F^{\mathrm{perf}} v_0$ is separable since $F^{\mathrm{perf}} v_0$ is isomorphic to F^{perf} which is perfect. Let $(\boldsymbol{F}((v^t K^t)), v_t)^*$ be a $|K|^+$-saturated elementary extension of $(\boldsymbol{F}((v^t K^t)), v_t)$. We have satisfied the hypotheses of Proposition 3.2, thus there exists an embedding

$$\iota : (K^t, v^t) \to (\boldsymbol{F}((v^t K^t)), v_t)^*$$

over (F^{perf}, v_0). As existential sentences "go up", we get that $(\boldsymbol{F}((v^t K^t)), v_t)^*$, and therefore also $(\boldsymbol{F}((v^t K^t)), v_t)$, models the existential $\mathcal{L}_{\mathrm{vf}}(F^{\mathrm{perf}})$-theory of (K^t, v^t).

$$\begin{tikzcd}
& & & \boldsymbol{F}((v^t K^t))^* \arrow[d, dash, "\preceq"] \\
& & & \boldsymbol{F}((v^t K^t)) \arrow[d, dash] \\
& K^t \arrow[r, dotted, "\mathrm{res}"'] \arrow[uurr, "\iota"] \arrow[d, dash] & K^t v^t \arrow[r, "\sigma"] \arrow[dd, dash, "\preceq_\exists"'] & \boldsymbol{F} \arrow[ddl, dash, "\preceq"] \\
K \arrow[r, dash] \arrow[d, dash] & K^{\mathrm{perf}} \arrow[d, dash] & & \\
F \arrow[r, dash] \arrow[d, dash] & F^{\mathrm{perf}} \arrow[r, "\cong"] \arrow[d, dash] & F^{\mathrm{perf}} v_0 & \\
C \arrow[r, dash] & C^{\mathrm{perf}} & &
\end{tikzcd}$$

Our assumption was that $\psi(\boldsymbol{x})$ is an $\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-formula with free variables $\boldsymbol{x}$ belonging to the residue field sort, and that $(K, v, D) \models \forall^k \boldsymbol{x}\ \psi(\boldsymbol{x})$, i.e., ${}^{\boldsymbol{x}}Kv \subseteq \psi(K)$. Then ${}^{\boldsymbol{x}}Fv \subseteq {}^{\boldsymbol{x}}Kv \subseteq \psi(K)$ (note that we write Fv rather than F because we have identified F with a subfield of K). Let

$$\Psi_F := \{\psi(\boldsymbol{a}) \mid \boldsymbol{a} \in {}^{\boldsymbol{x}}Fv\}.$$

Then Ψ_F is a set of $\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-sentences (with additional parameters from Fv) which is equivalent to the property that "${}^{\boldsymbol{x}}Fv \subseteq \psi$". We may now restate our assumption as $(K, v) \models \Psi_F$. Since existential sentences "go up", $(K^t, v^t) \models \Psi_F$. By the result of the previous paragraph, we have $(\boldsymbol{F}((v^t K^t)), v_t) \models \Psi_F$. By an application of Lemma 5.1, $(F^{\mathrm{perf}}(t)^h, v_t) \models \Psi_F$. By Lemma 5.3, $(L, w) \models \Psi_F$ (where we replace the parameters from Fv by their images under the map i). This shows that ${}^{\boldsymbol{x}}i(F) \subseteq \psi(L)$, as claimed. $\square$

Corollary 5.6 (near $\forall^k\exists$-C-completeness). *Let $\psi(\boldsymbol{x})$ be an $\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-formula with free variables $\boldsymbol{x}$ belonging to the residue field sort. Suppose there exists*

$$(K, v, D) \models \boldsymbol{T}_{F/C} \cup \{\forall^k \boldsymbol{x}\ \psi(\boldsymbol{x})\}.$$

Then there exists $n \in \mathbb{N}$ such that ${}^{\boldsymbol{x}}Lw \subseteq \psi(L^{p^{-n}})$ for all $(L, w, E) \models \boldsymbol{T}_{F/C}$.

Proof. Let $(L, w, E) \models \boldsymbol{T}_{F/C}$. As F/C is separable and $(Lw, Ew) \equiv (F, C)$ as $\mathcal{L}_{\mathrm{ring}}(C)$-structures, Lw/Ew is also separable. In particular, both (K, v, D) and (L, w, E) are models of $\boldsymbol{T}_{Lw/Ew}$, and thus we may apply the conclusion of Proposition 5.5 to

$$(L^{\mathrm{perf}}, w, E^{\mathrm{perf}}, \mathrm{id}) \in \boldsymbol{H}(Lw^{\mathrm{perf}}/Ew^{\mathrm{perf}}).$$

Thus we have that ${}^{\boldsymbol{x}}Lw \subseteq \psi(L^{\mathrm{perf}})$. To find n, we use a simple compactness argument, as follows.

Write the formula $\psi(\boldsymbol{x})$ as $\exists \boldsymbol{y}\ \rho(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{c})$, for a quantifier-free $\mathcal{L}_{\mathrm{vf}}$-formula ρ. For each $n \in \mathbb{N}$, let $\psi_n(\boldsymbol{x})$ be the formula $\exists \boldsymbol{y}\ \rho(\boldsymbol{x}^{p^n}, \boldsymbol{y}, \boldsymbol{c}^{p^n})$ and consider the $\mathcal{L}_{\mathrm{vf}}(C)$-structure $(L^{p^{-n}}, w, E)$ which extends (L, w, E). Then, for $\boldsymbol{a} \in {}^{\boldsymbol{x}}Lw$, $\boldsymbol{a} \in \psi(L^{p^{-n}})$ if and only if $\boldsymbol{a} \in \psi_n(L)$. Let $p(\boldsymbol{x})$ be the set of formulas $\{\neg\psi_n(\boldsymbol{x}) \mid n \in \mathbb{N}\}$. If $p(\boldsymbol{x})$ is a type, i.e., $p(\boldsymbol{x})$ is consistent with $\boldsymbol{T}_{F/C}$, then we may realise it by a tuple $\boldsymbol{a}$ in a model $(L, w, E) \models \boldsymbol{T}_{F/C}$. Thus $\boldsymbol{a} \notin \psi(L^{p^{-n}})$, for all $n \in \mathbb{N}$. Since L^{perf} is the directed union $\bigcup_{n\in\mathbb{N}} L^{p^{-n}}$ (even as $\mathcal{L}_{\mathrm{vf}}(C)$-structures), we have that $\boldsymbol{a} \notin \psi(L^{\mathrm{perf}})$. This contradicts the result of the previous paragraph.

Consequently, there exists $n \in \mathbb{N}$ such that $\boldsymbol{T}_{F/C}$ entails $\forall^k \boldsymbol{x}\ \psi_n(\boldsymbol{x})$. Equivalently, for all $(L, w, E) \models \boldsymbol{T}_{F/C}$, we have ${}^{\boldsymbol{x}}Lw \subseteq \psi(L^{p^{-n}})$, as required. $\square$

Corollary 5.7 (perfect residue field, $\forall^k\exists$-C-completeness). *Suppose that F is perfect. Then $\boldsymbol{T}_{F/C}$ is $\forall^k\exists$-C-complete, i.e., for any $\forall^k\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-sentence ϕ, either $\boldsymbol{T}_{F/C} \models \phi$ or $\boldsymbol{T}_{F/C} \models \neg\phi$.*

Proof. Suppose that there is $(K, v, D) \models \boldsymbol{T}_{F/C} \cup \{\phi\}$ and let $(L, w, E) \models \boldsymbol{T}_{F/C}$. Then $(K, v, D) \models \boldsymbol{T}_{Lw/Ew}$ and

$$(L, w, E, \mathrm{id}) \in \boldsymbol{H}(Lw/Ew) = \boldsymbol{H}(Lw^{\mathrm{perf}}/Ew^{\mathrm{perf}}).$$

We write $\phi = \forall^k \boldsymbol{x}\ \psi(\boldsymbol{x})$ for some $\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-formula $\psi(\boldsymbol{x})$ with free variables $\boldsymbol{x}$ belonging to the residue field sort. Then $(K, v, D) \models \phi$ means that ${}^{\boldsymbol{x}}Kv \subseteq \psi(K)$. Applying Proposition 5.5, we have that ${}^{\boldsymbol{x}}Lw \subseteq \psi(L)$. Thus $(L, w, E) \models \phi$. This shows that $\boldsymbol{T}_{F/C} \models \phi$, as required. □

Remark 5.8. We do not know whether the assumption that F is perfect is necessary in Corollary 5.7. However, note that Corollary 5.7 cannot be extended from $\forall^k\exists$-sentences to arbitrary $\forall\exists$-sentences (even without parameters and with only one universal quantifier). For example, the sentence

$$\forall x \exists y\ (v(x) = v(y^2))$$

expresses 2-divisibility of the value group, so is satisfied in $F((\mathbb{Q}))$ but not in $F((t))$.

On the other hand, one could generalise Corollary 5.7 by slightly adapting the proof to allow also sentences with more general quantifiers over the residue field, namely $Q^k\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-sentences, i.e., sentences of the form

$$\exists^k \boldsymbol{x}_1 \forall^k \boldsymbol{y}_1 \ldots \exists^k \boldsymbol{x}_n \forall^k \boldsymbol{y}_n\ \psi(\boldsymbol{x}_1, \boldsymbol{y}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_n)$$

with $\psi(\boldsymbol{x}_1, \boldsymbol{y}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_n)$ an $\exists$-$\mathcal{L}_{\mathrm{vf}}(C)$-formula.

6. The existential theory

We now restrict the machinery of the previous section to existential sentences and prove Theorem 1.1 from the introduction. We fix a field F, let C be the prime field of F, and write $\boldsymbol{T}_F = \boldsymbol{T}_{F/C}$, $\boldsymbol{H}(F) = \boldsymbol{H}(F/C)$.

Lemma 6.1. *$\boldsymbol{T}_F$ is $\exists$-complete, i.e., for any $\exists$-$\mathcal{L}_{\mathrm{vf}}$-sentence ϕ, either $\boldsymbol{T}_F \models \phi$ or $\boldsymbol{T}_F \models \neg\phi$.*

Proof. Suppose that $\boldsymbol{T}_F \cup \{\phi\}$ is consistent. Thus there exists $(K, v) \models \boldsymbol{T}_F \cup \{\phi\}$. Simply viewing ϕ as an $\forall^k\exists$-formula $\forall^k x\ \psi(x)$ with $\psi(x) = \phi$, we have that $Kv \subseteq \psi(K)$. By Corollary 5.6 there exists $n \in \mathbb{N}$ such that, for every $(L, w) \models \boldsymbol{T}_F$, $Lw \subseteq \psi(L^{p^{-n}})$. In particular, $\psi(L^{p^{-n}})$ is nonempty. Since no parameters appear in ψ, we may apply the n-th power of the Frobenius map to get that $\psi(L)$ is nonempty, for every $(L, w) \models \boldsymbol{T}_F$. Viewing ϕ as an $\exists$-sentence again, we have that $(L, w) \models \phi$. Thus $\boldsymbol{T}_F \models \phi$, as required. □

For the proof of Theorem 1.1 it remains to show that $\boldsymbol{T}_F^1$ already entails those existential and universal sentences which are entailed by $\boldsymbol{T}_F$.

Definition 6.2. We define two subtheories of $\boldsymbol{T}_F^1$. Let $T_F^{\exists}$ be the $\mathcal{L}_{\mathrm{vf}}$-theory consisting of the following axioms (expressed informally about a structure (K, v)):

(1) (K, v) is an equicharacteristic henselian nontrivially valued field and

(2) Kv is a model of the existential $\mathcal{L}_{\mathrm{ring}}$-theory of F.

Let $T_F^\forall$ be the $\mathcal{L}_{\mathrm{vf}}$-theory consisting of the following axioms (again expressed informally):

(1) (K, v) is an equicharacteristic henselian nontrivially valued field and

(2) Kv is a model of the universal $\mathcal{L}_{\mathrm{ring}}$-theory of F.

Note that $\boldsymbol{T}_F^1 \equiv T_F^\exists \cup T_F^\forall$.

Lemma 6.3. *Let ϕ be an existential $\mathcal{L}_{\mathrm{vf}}$-sentence. If $\boldsymbol{T}_F \models \phi$ then $T_F^\exists \models \phi$.*

Proof. Let $(K, v) \models T_F^\exists$. Then Kv is a model of $\mathrm{Th}_\exists(F)$; equivalently the theory of Kv is consistent with the atomic diagram of F. Thus there is an elementary extension $(K, v) \preceq (K^*, v^*)$ with an embedding $\sigma : F \to K^*v^*$, cf. [Marker 2002, Lemma 2.3.3]. Note that $(K^*, v^*, \sigma) \in \boldsymbol{H}(F)$ and that $(F(t)^h, v_t) \models \boldsymbol{T}_F$, hence $(F(t)^h, v_t) \models \phi$. Therefore, Lemma 5.3 implies that $(K^*, v^*) \models \phi$; thus $(K, v) \models \phi$. This shows that $T_F^\exists \models \phi$. □

Lemma 6.4. *Let ϕ be a universal $\mathcal{L}_{\mathrm{vf}}$-sentence. If $\boldsymbol{T}_F \models \phi$ then $T_F^\forall \models \phi$.*

Proof. Let $(K, v) \models T_F^\forall$. Then $Kv \models \mathrm{Th}_\forall(F)$. There exists $F' \equiv F$ with an embedding $\sigma : Kv \to F'$ (see [Marker 2002, Exercise 2.5.10]). Using Lemma 2.1, we may choose an equicharacteristic nontrivially valued field (L, w) which extends (K, v) and is such that Lw is isomorphic to F'. In particular $Lw \equiv F$. Let $(L, w)^h$ be the henselisation of (L, w); then we have $(L, w)^h \models \boldsymbol{T}_F$, so $(L, w)^h \models \phi$. Since ϕ is universal, we conclude that $(K, v) \models \phi$. □

Theorem 6.5 ($\exists$-completeness). *$\boldsymbol{T}_F^1$ is $\exists$-complete, i.e., for any $\exists$-$\mathcal{L}_{\mathrm{vf}}$-sentence ϕ either $\boldsymbol{T}_F^1 \models \phi$ or $\boldsymbol{T}_F^1 \models \neg\phi$.*

Proof. Let ϕ be an existential $\mathcal{L}_{\mathrm{vf}}$-sentence. By Lemma 6.1, either $\boldsymbol{T}_F \models \phi$ or $\boldsymbol{T}_F \models \neg\phi$. In the first case we apply Lemma 6.3 and find that $T_F^\exists \models \phi$; in the second case we apply Lemma 6.4 and find that $T_F^\forall \models \neg\phi$. Since $\boldsymbol{T}_F^1 \equiv T_F^\exists \cup T_F^\forall$, in either case $\boldsymbol{T}_F^1$ "decides" ϕ, and we are done. □

Remark 6.6. Let $\chi(x)$ be an existential $\mathcal{L}_{\mathrm{ring}}$-formula with one free variable. In [Anscombe and Fehm 2016], we apply Theorem 6.5 to the following $\exists$- or $\forall$-$\mathcal{L}_{\mathrm{vf}}$-sentences:

(1) $\forall x\ (\chi(x) \to v(x) \geq 0)$,

(2) $\forall x\ (\chi(x) \to v(x) > 0)$, and

(3) $\exists x\ (v(x) > 0 \wedge x \neq 0 \wedge \chi(x))$.

We also apply Corollary 5.6 to the $\forall^k\exists$-$\mathcal{L}_{\mathrm{vf}}$-sentence

(4) $\forall^k x \exists y\ (\mathrm{res}(y) = x \wedge \chi(y))$.

7. An "existential AKE principle" and existential decidability

Theorem 6.5 shows that the existential (respectively, universal) theory of an equicharacteristic henselian nontrivially valued field depends only on the existential (respectively, universal) theory of its residue field. We formulate this in the following "existential AKE principle".

Theorem 7.1. *Let (K, v) and (L, w) be equicharacteristic henselian nontrivially valued fields. Then*

$$(K, v) \models \mathrm{Th}_\exists(L, w) \quad \textit{if and only if} \quad Kv \models \mathrm{Th}_\exists(Lw).$$

Proof. $(\Rightarrow)$: Note that the maximal ideal is defined by the quantifier-free formula $v(x) > 0$. Therefore any existential statement about the residue field can be translated into an existential statement about the valued field.

$(\Leftarrow)$: If $Kv \models \mathrm{Th}_\exists(Lw)$ then $(K, v) \models T^\exists_{Lw}$. By Lemma 6.1, $\boldsymbol{T}_{Lw}$ entails the existential theory of (L, w), and by Lemma 6.3, $T^\exists_{Lw}$ entails the existential consequences of $\boldsymbol{T}_{Lw}$. Combining these two statements, we have that $T^\exists_{Lw}$ entails the existential theory of (L, w). Thus (K, v) models the existential theory of (L, w). □

Corollary 7.2. *Let (K, v) and (L, w) be equicharacteristic henselian nontrivially valued fields. Then*

$$\mathrm{Th}_\exists(K, v) = \mathrm{Th}_\exists(L, w) \quad \textit{if and only if} \quad \mathrm{Th}_\exists(Kv) = \mathrm{Th}_\exists(Lw).$$

Proof. This follows from Theorem 7.1, since $\mathrm{Th}_\exists(K, v) = \mathrm{Th}_\exists(L, w)$ if and only if both $(K, v) \models \mathrm{Th}_\exists(L, w)$ and $(L, w) \models \mathrm{Th}_\exists(K, v)$, and $\mathrm{Th}_\exists(Kv) = \mathrm{Th}_\exists(Lw)$ if and only if both $Kv \models \mathrm{Th}_\exists(Lw)$ and $Lw \models \mathrm{Th}_\exists(Kv)$. □

Note that Corollary 7.2 is in fact simply a reformulation of Theorem 6.5. Note moreover that, by the usual duality between existential and universal sentences, the same principle holds with "$\exists$" replaced by "$\forall$".

Remark 7.3. The reader has probably noticed that as opposed to the usual AKE principles, the value group does not occur here. However, since the existential theory of a valued field determines the existential theory of its value group, Corollary 7.2 could also be phrased as

$$\mathrm{Th}_\exists(K, v) = \mathrm{Th}_\exists(L, w) \quad \text{if and only if} \quad \mathrm{Th}_\exists(Kv) = \mathrm{Th}_\exists(Lw) \text{ and } \mathrm{Th}_\exists(vK) = \mathrm{Th}_\exists(wL).$$

In fact, all nontrivial ordered abelian groups have the same existential theory (which follows immediately from the completeness of the theory of divisible ordered abelian groups; see also [Gurevich and Kokorin 1963]). In residue characteristic zero, this special form of the existential AKE principle was known before; see, e.g., [Koenigsmann 2014, p. 192].

Remark 7.4. In mixed characteristic the situation is very different. Fix a prime p and let (K, v) and (L, w) be henselian nontrivially valued fields. Just as in Remark 7.3, the existential theory of a valued field determines the existential theory of the residue field and the value group, i.e.,

$$\mathrm{Th}_\exists(K, v) = \mathrm{Th}_\exists(L, w) \implies \mathrm{Th}_\exists(Kv) = \mathrm{Th}_\exists(Lw) \text{ and } \mathrm{Th}_\exists(vK) = \mathrm{Th}_\exists(wL).$$

However, in mixed characteristic the converse fails. For example, consider the valued fields $(K, v) = (\mathbb{Q}_p, v_p)$ and $(L, w) = (\mathbb{Q}_p(\sqrt{p}), v_p)$. Both residue fields Kv and Lw are equal to $\mathbb{F}_p$ and both value groups are isomorphic to $\mathbb{Z}$, but the existential theories of (K, v) and (L, w) are not equal since $\mathbb{Q}_p$ does not contain a square-root of p. In particular, both Theorem 6.5 and Corollary 7.2 fail if we replace "equicharacteristic" by "mixed characteristic".

One feature of mixed characteristic is that the existential theory of (K, v) determines the existential theory of (vK, vp), which is the ordered abelian group vK together with the distinguished nonzero element vp. Therefore, if (K, v) and (L, w) are both of characteristic zero and residue characteristic p, we have the implication

$$\mathrm{Th}_\exists(K, v) = \mathrm{Th}_\exists(L, w) \implies \mathrm{Th}_\exists(Kv) = \mathrm{Th}_\exists(Lw) \text{ and } \mathrm{Th}_\exists(vK, vp) = \mathrm{Th}_\exists(wL, wp). \quad (*)$$

Note that not all ordered abelian groups with a distinguished nonzero element have the same existential theory. For example, $\mathrm{Th}_\exists(\mathbb{Z}, 1) \neq \mathrm{Th}_\exists(\mathbb{Z}, 2)$. Nevertheless, we claim that the implication $(*)$ is not invertible. To prove this claim we need a new counterexample because $v_p p$ is minimal positive in $v_p\mathbb{Q}_p$ but $v_p p = 2v_p\sqrt{p}$ in $v_p\mathbb{Q}_p(\sqrt{p})$, and so

$$\mathrm{Th}_\exists(v_p\mathbb{Q}_p, v_p p) = \mathrm{Th}_\exists(\mathbb{Z}, 1) \neq \mathrm{Th}_\exists(\mathbb{Z}, 2) = \mathrm{Th}_\exists(v_p\mathbb{Q}_p(\sqrt{p}), v_p p).$$

Instead, we cite the example of two valued fields (L_1, v) and (F_1, v) which were constructed in [Anscombe and Kuhlmann 2016, Theorem 1.5]. Both are tame and algebraic extensions of $(\mathbb{Q}, v_p)$, both residue fields $L_1 v$ and $F_1 v$ are equal to $\mathbb{F}_p$, and both value groups vL_1 and vF_1 are equal to the p-divisible hull of $\frac{1}{p-1}(v_p p)\mathbb{Z}$. Nevertheless $(L_1, v) \not\equiv (F_1, v)$. In fact, since L_1 and F_1 are algebraic, we have that $\mathrm{Th}_\exists(L_1, v) \neq \mathrm{Th}_\exists(F_1, v)$. This example shows that the converse to $(*)$ does not hold, even under the additional hypothesis that (K, v) and (L, w) are tame.

Next we deduce Corollary 1.3 from Theorem 6.5.

Corollary 7.5. *Let (K, v) be an equicharacteristic henselian valued field. The following are equivalent.*

(1) $\mathrm{Th}_\exists(Kv)$ *is decidable.*

(2) $\mathrm{Th}_\exists(K, v)$ *is decidable.*

Proof. $2 \Longrightarrow 1$: As before, residue fields are interpreted in valued fields in such a way that existential statements about Kv remain existential statements about (K, v). Therefore, if (K, v) is $\exists$-decidable, then Kv is $\exists$-decidable.

$1 \Longrightarrow 2$: Write $F := Kv$ and suppose that F is $\exists$-decidable. If v is trivial, then $(K, v) = (F, v)$ is also $\exists$-decidable, so suppose that v is nontrivial. We may recursively enumerate the existential and universal theory R^1_F of F, so $\boldsymbol{T}^1_F$ is effectively axiomatisable. By Theorem 6.5, $\boldsymbol{T}^1_F$ is an $\exists$-complete subtheory of $\mathrm{Th}(K, v)$. Thus we may decide the truth of existential (and universal) sentences in (K, v). □

Remark 7.6. If we replace "equicharacteristic" by "mixed characteristic" then the statement of Corollary 7.5 is no longer true. To see this, let P be an undecidable set of primes, let K be the extension of $\mathbb{Q}_p$ generated by a family of l-th roots of p, for $l \in P$, and let v be the unique extension of v_p to K. Then $Kv = \mathbb{F}_p$, so $\mathrm{Th}_\exists(Kv)$ is decidable, but $\mathrm{Th}_\exists(vK, vp)$ is undecidable, hence so is $\mathrm{Th}_\exists(K, v)$. At present, we do not know of an example of a mixed characteristic henselian valued field (K, v) for which $\mathrm{Th}_\exists(Kv)$ and $\mathrm{Th}_\exists(vK, vp)$ are decidable but $\mathrm{Th}_\exists(K, v)$ is undecidable.

Let $\mathcal{L}_{\mathrm{vf}}(t)$ be the language of valued fields with an additional parameter t, and let q be a prime power. In [Denef and Schoutens 2003], it is shown that resolution of singularities in characteristic p would imply that the existential $\mathcal{L}_{\mathrm{vf}}(t)$-theory of $\mathbb{F}_q((t))$ is decidable. Using our methods we can prove the following weaker but unconditional result.

Corollary 7.7. *The existential theory of $\mathbb{F}_q((t))$ in the language of valued fields is decidable.*

First proof. We can apply Corollary 7.5, noting that $\mathrm{Th}_\exists(\mathbb{F}_q)$ is decidable. □

For the sake of interest, we present a more direct proof of this special case. However, note that this "second proof" uses the decidability of $\mathbb{F}_q$, while the "first proof" used only the decidability of the *existential* theory of $\mathbb{F}_q$.

Second proof. As an equicharacteristic tame field (Proposition 4.6) with decidable residue field and value group, $(\mathbb{F}_q((t))^{\mathbb{Q}}, v_t)$ is decidable, by [Kuhlmann 2014, Theorem 7.7(a)]. Since $(\mathbb{F}_q((t))^{\mathbb{Q}}, v_t)$ is the directed union of structures isomorphic to $(\mathbb{F}_q((t)), v_t)$ (Corollary 4.3), in fact $(\mathbb{F}_q((t)), v_t)$ and $(\mathbb{F}_q((t))^{\mathbb{Q}}, v_t)$ have the same $\exists$-$\mathcal{L}_{\mathrm{vf}}$-theory. Thus, to decide the existential $\mathcal{L}_{\mathrm{vf}}$-theory of $(\mathbb{F}_q((t)), v_t)$, it suffices to apply the decision procedure for the $\mathcal{L}_{\mathrm{vf}}$-theory of $(\mathbb{F}_q((t))^{\mathbb{Q}}, v_t)$. □

Remark 7.8. Since Corollary 7.7 shows decidability of the existential theory of $\mathbb{F}_q((t))$ in the language of *valued* fields $\mathcal{L}_{\mathrm{vf}}$, in which the valuation ring is definable by a quantifier-free formula, we also get decidability of the existential theory of the *ring* $\mathbb{F}_q[\![t]\!]$. It might however be interesting to point out that it was proven only recently that already decidability of the existential theory of $\mathbb{F}_q((t))$ in the language of *rings* would imply decidability of the existential theory of the ring $\mathbb{F}_q[\![t]\!]$; see [Anscombe and Koenigsmann 2014, Corollary 3.4].

Remark 7.9. The $\exists$-$\mathcal{L}_{\mathrm{vf}}(t)$-theory of $(\mathbb{F}_q((t)), v_t)$ is equivalent to the $\forall_1^K\exists$-$\mathcal{L}_{\mathrm{vf}}$-theory of $(\mathbb{F}_q((t)), v_t)$. This "equivalence" is meant in the sense that there is a truth-preserving effective translation between $\exists$-$\mathcal{L}_{\mathrm{vf}}(t)$-sentences and $\forall\exists$-$\mathcal{L}_{\mathrm{vf}}$-sentences which have only one universal quantifier ranging over the valued field sort (and arbitrary existential quantifiers). In this argument we make repeated use of the fact that, for all $a \in \mathbb{F}_q((t))$ with $v_t(a) > 0$ and $a \neq 0$, there is an $\mathcal{L}_{\mathrm{vf}}$-embedding $\mathbb{F}_q((t)) \to \mathbb{F}_q((t))$ which sends $t \mapsto a$.

Let $\phi(t)$ be an existential $\mathcal{L}_{\mathrm{vf}}(t)$-sentence. We claim that $\phi(t)$ is equivalent to the $\forall_1^K\exists$-$\mathcal{L}_{\mathrm{vf}}$-sentence

$$\forall u\ ((v(u) > 0 \wedge u \neq 0) \to \phi(u)).$$

This follows from the fact about embeddings stated above.

On the other hand, let $\psi(x)$ be an $\exists$-$\mathcal{L}_{\mathrm{vf}}$-formula in one free variable x in the valued field sort and consider the $\exists$-$\mathcal{L}_{\mathrm{vf}}(t)$-sentence χ which is defined to be

$$\exists y \exists z_0 \ldots \exists z_{q-1} \Big(yt = 1 \wedge \psi(y) \wedge \bigwedge_j z_j^q = z_j \wedge \bigwedge_{i\neq j} z_i \neq z_j \wedge \bigwedge_j \psi(z_j + t) \wedge \bigwedge_j \psi(z_j) \Big).$$

Written more informally, the sentence χ expresses that

$$\psi(t^{-1}) \wedge \bigwedge_{z\in\mathbb{F}_q} (\psi(z+t) \wedge \psi(z)).$$

We claim that $\forall x\ \psi(x)$ and χ are equivalent. First suppose that $\mathbb{F}_q((t)) \models \forall x\ \psi(x)$. By choosing (z_j) to be an enumeration of $\mathbb{F}_q$, we immediately have that $\mathbb{F}_q((t)) \models \chi$.

In the other direction, suppose that $\mathbb{F}_q((t)) \models \chi$ and let $a \in \mathbb{F}_q((t))$. If $v_t(a) < 0$ then consider the embedding which sends $t \mapsto a^{-1}$. Since $\psi(t^{-1})$ holds, applying the embedding shows that $\psi(a)$ also holds. On the other hand suppose that $v_t(a) \geq 0$. If $a \in \mathbb{F}_q$ then χ already entails that $\psi(a)$. Now suppose that $a \notin \mathbb{F}_q$ and let z be the residue of a. Consider the embedding which sends $t \mapsto a - z$ (note that $a - z \neq 0$). Since $\psi(z+t)$ holds, applying the embedding shows that $\psi(a)$ also holds. This completes the proof that $\mathbb{F}_q((t)) \models \forall x\ \psi(x)$.

Acknowledgements

The authors would like to thank Immanuel Halupczok, Ehud Hrushovski, Jochen Koenigsmann, Dugald Macpherson, and Alexander Prestel for helpful discussions and encouragement.

References

[Anscombe and Fehm 2016] S. Anscombe and A. Fehm, "Characterizing diophantine henselian valuation rings and valuation ideals", preprint, 2016. arXiv

[Anscombe and Koenigsmann 2014] W. Anscombe and J. Koenigsmann, "An existential $\varnothing$-definition of $\mathbb{F}_q[[t]]$ in $\mathbb{F}_q((t))$", *J. Symb. Log.* **79**:4 (2014), 1336–1343. MR Zbl

[Anscombe and Kuhlmann 2016] S. Anscombe and F.-V. Kuhlmann, "Notes on extremal and tame valued fields", 2016, http://math.usask.ca/~fvk/EXTRNOT.pdf. To appear in *J. Sym. Logic*.

[Cluckers et al. 2013] R. Cluckers, J. Derakhshan, E. Leenknegt, and A. Macintyre, "Uniformly defining valuation rings in Henselian valued fields with finite or pseudo-finite residue fields", *Ann. Pure Appl. Logic* **164**:12 (2013), 1236–1246. MR Zbl

[Denef and Schoutens 2003] J. Denef and H. Schoutens, "On the decidability of the existential theory of $\mathbb{F}_p[[t]]$", pp. 43–60 in *Valuation theory and its applications* (Saskatoon, 1999), vol. II, edited by F.-V. Kuhlmann et al., Fields Inst. Commun. **33**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Efrat 2006] I. Efrat, *Valuations, orderings, and Milnor K-theory*, Mathematical Surveys and Monographs **124**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[Engler and Prestel 2005] A. J. Engler and A. Prestel, *Valued fields*, Springer, Berlin, 2005. MR Zbl

[Fehm 2015] A. Fehm, "Existential $\varnothing$-definability of Henselian valuation rings", *J. Symb. Log.* **80**:1 (2015), 301–307. MR Zbl

[Gurevich and Kokorin 1963] Y. S. Gurevich and A. I. Kokorin, "Universal equivalence of ordered Abelian groups", *Algebra i Logika Sem.* **2**:1 (1963), 37–39. In Russian. MR

[Koenigsmann 2014] J. Koenigsmann, "Undecidability in number theory", pp. 159–195 in *Model theory in algebra, analysis and arithmetic*, edited by H. D. Macpherson and C. Toffalori, Lecture Notes in Math. **2111**, Springer, Heidelberg, 2014. MR Zbl

[Kuhlmann 2001] F.-V. Kuhlmann, "Elementary properties of power series fields over finite fields", *J. Symbolic Logic* **66**:2 (2001), 771–791. MR Zbl

[Kuhlmann 2004] F.-V. Kuhlmann, "Value groups, residue fields, and bad places of rational function fields", *Trans. Amer. Math. Soc.* **356**:11 (2004), 4559–4600. MR Zbl

[Kuhlmann 2011] F.-V. Kuhlmann, "Valuation theory", book in progress, 2011, http://math.usask.ca/~fvk/Fvkbook.htm.

[Kuhlmann 2014] F.-V. Kuhlmann, "The algebra and model theory of tame valued fields", *J. Reine Angew. Math.* (online publication May 2014).

[Kuhlmann et al. 1986] F.-V. Kuhlmann, M. Pank, and P. Roquette, "Immediate and purely wild extensions of valued fields", *Manuscripta Math.* **55**:1 (1986), 39–67. MR Zbl

[Marker 2002] D. Marker, *Model theory: an introduction*, Graduate Texts in Mathematics **217**, Springer, New York, 2002. MR Zbl

[Prestel 2015] A. Prestel, "Definable Henselian valuation rings", *J. Symb. Log.* **80**:4 (2015), 1260–1267. MR Zbl

[Prestel and Delzell 2011] A. Prestel and C. N. Delzell, *Mathematical logic and model theory: a brief introduction*, Springer, London, 2011. MR Zbl

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics **67**, Springer, New York, 1979. MR Zbl

Communicated by Jonathan Pila
Received 2015-09-18 Revised 2016-02-09 Accepted 2016-03-15

sanscombe@uclan.ac.uk — *Jeremiah Horrocks Institute, University of Central Lancashire, Preston, PR1 2HE, United Kingdom*

arno.fehm@manchester.ac.uk — *School of Mathematics, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 10:3 (2016)
dx.doi.org/10.2140/ant.2016.10.685

On twists of modules over noncommutative Iwasawa algebras

Somnath Jha, Tadashi Ochiai and Gergely Zábrádi

It is well known that, for any finitely generated torsion module M over the Iwasawa algebra $\mathbb{Z}_p[[\Gamma]]$, where Γ is isomorphic to $\mathbb{Z}_p$, there exists a continuous p-adic character ρ of Γ such that, for every open subgroup U of Γ, the group of U-coinvariants $M(\rho)_U$ is finite; here $M(\rho)$ denotes the twist of M by ρ. This twisting lemma was already used to study various arithmetic properties of Selmer groups and Galois cohomologies over a cyclotomic tower by Greenberg and Perrin-Riou. We prove a noncommutative generalization of this twisting lemma, replacing torsion modules over $\mathbb{Z}_p[[\Gamma]]$ by certain torsion modules over $\mathbb{Z}_p[[G]]$ with more general p-adic Lie group G. In a forthcoming article, this noncommutative twisting lemma will be used to prove the functional equation of Selmer groups of general p-adic representations over certain p-adic Lie extensions.

Introduction

Let us fix an odd prime p throughout the paper. We denote by Γ a p-Sylow subgroup of $\mathbb{Z}_p^\times$. For a compact p-adic Lie group G and the ring $\mathcal{O}$ of integers of a finite extension of $\mathbb{Q}_p$, we denote the Iwasawa algebra $\mathcal{O}[\![G]\!]$ of G with coefficient in $\mathcal{O}$ by $\Lambda_{\mathcal{O}}(G)$.

In this article, we study $\Lambda_{\mathcal{O}}(G)$-modules, motivated by [Coates et al. 2005]. More precisely, we study specializations of certain $\Lambda_{\mathcal{O}}(G)$-modules by two-sided ideals of $\Lambda_{\mathcal{O}}(G)$. Recall that the paper [Coates et al. 2005] establishes a reasonable setting of noncommutative Iwasawa theory in the following situation.

(G) G is a compact p-adic Lie group which has a closed normal subgroup H such that G/H is isomorphic to Γ.

S. Jha gratefully acknowledges the support of a JSPS postdoctoral fellowship and a DST INSPIRE faculty award grant. T. Ochiai is partially supported for this work by KAKENHI (Grant-in-Aid for Exploratory Research: Grant Number 24654004, Grant-in-Aid for Scientific Research (B): Grant Number 26287005). G. Zábrádi was supported by a Hungarian OTKA Research grant K-100291 and by the János Bolyai Scholarship of the Hungarian Academy of Sciences.
MSC2010: primary 11R23; secondary 16S50.
Keywords: Selmer group, noncommutative Iwasawa theory.

According to the philosophy of [Coates et al. 2005], for a reasonable ordinary p-adic representation T of a number field K and a pair of compact p-adic Lie groups $H \subset G$ satisfying the condition (G), the Pontryagin dual $\mathcal{S}_A^\vee$ of the Selmer group $\mathcal{S}_A$ of the Galois representation $A = T \otimes \mathbb{Q}_p/\mathbb{Z}_p$ over a Galois extension K_∞/K with $\mathrm{Gal}(K_\infty/K) \cong G$ seems to be a nice object. The $\Lambda_{\mathcal{O}}(G)$-module $\mathcal{S}_A^\vee$ divided by the largest p-primary torsion subgroup $\mathcal{S}_A^\vee(p)$ is conjectured to belong to the category $\mathfrak{n}_H(G)$ which consists of finitely generated $\Lambda_{\mathcal{O}}(G)$-modules M such that M is also finitely generated over $\Lambda_{\mathcal{O}}(H)$. From such arithmetic background, we are led to study finitely generated $\Lambda_{\mathcal{O}}(G)$-modules for a compact Lie group G with $H \subset G$ satisfying the condition (G).

On the other hand, for any open subgroup U of G and for any arithmetic module $\mathcal{S}_A^\vee$ as above, the largest U-coinvariant quotient $(\mathcal{S}_A^\vee)_U$ is expected to be related to the Selmer group of A over a finite extension L of K with $\mathrm{Gal}(L/K) \cong G/U$. As remarked above, we have the following fact (Tw) when $G = \Gamma$ (i.e., when $H = 1$) which was used quite effectively in the work of Greenberg [1989] and Perrin-Riou [2003].

(Tw) For any finitely generated torsion $\Lambda_{\mathcal{O}}(\Gamma)$-module M, there exists a continuous character $\rho : \Gamma \to \mathbb{Z}_p^\times$ such that the largest U-coinvariant quotient $(M \otimes_{\mathbb{Z}_p} \mathbb{Z}_p(\rho))_U$ of $M \otimes_{\mathbb{Z}_p} \mathbb{Z}_p(\rho)$ is finite for every open subgroup U of Γ, where $\mathbb{Z}_p(\rho)$ is a free $\mathbb{Z}_p$-module of rank one on which Γ acts through the character $\Gamma \xrightarrow{\rho} \mathbb{Z}_p^\times$.

We call such a statement (Tw) a twisting lemma. In this commutative situation of $G = \Gamma$, the twisting lemma is proved in a quite elementary way. For example, we consider the characteristic ideal $\mathrm{char}_{\mathcal{O}[\![\Gamma]\!]} M$. If we take a ρ such that the values $\rho(\gamma)^{-1}\zeta_{p^n} - 1$ do not coincide with any roots of the distinguished polynomial associated to $\mathrm{char}_{\mathcal{O}[\![\Gamma]\!]} M$ when natural numbers n and p^n-th roots of unity ζ_{p^n} vary, the twisting lemma is known to hold.

If we have a twisting lemma in a noncommutative setting, it seems quite useful for some arithmetic applications for noncommutative Iwasawa theory. On the other hand, for a noncommutative G, it was not clear what to do to prove the twisting lemma because we cannot talk about "roots of characteristic polynomials" as we did in commutative setting. We finally succeeded in proving the twisting lemma which is stated as our Main Theorem below.

For a $\Lambda_{\mathcal{O}}(G)$-module M and a continuous character $\rho : \Gamma \to \mathbb{Z}_p^\times$, we denote by $M(\rho)$ the $\Lambda_{\mathcal{O}}(G)$-module $M \otimes_{\mathbb{Z}_p} \mathbb{Z}_p(\rho)$ with diagonal G-action.

Main Theorem. *Let G be a compact p-adic Lie group and let H be a closed normal subgroup such that G/H is isomorphic to Γ. Let M be a $\Lambda_{\mathcal{O}}(G)$-module which is finitely generated over $\Lambda_{\mathcal{O}}(H)$.*

Then there exists a continuous character $\rho : \Gamma \to \mathbb{Z}_p^\times$ such that the largest U-coinvariant quotient $M(\rho)_U$ of $M(\rho)$ is finite for every open normal subgroup U of G.

We give some examples of a pair $H \subset G$ satisfying the condition (G) and a $\Lambda_{\mathcal{O}}(G)$-module M which should appear in arithmetic applications.

Examples. (1) Let us choose a prime $p \geq 5$. Let E be a non-CM elliptic curve over $\mathbb{Q}$ with good ordinary reduction at p. Take $K = \mathbb{Q}(E[p])$ and set $K_\infty = \mathbb{Q}\bigl(\bigcup_{n\geq 1} E[p^n]\bigr)$. Then by a well known result of Serre, $\mathrm{Gal}(K_\infty/K)$ is an open subgroup of $\mathrm{GL}_2(\mathbb{Z}_p)$. By Weil pairing, the cyclotomic $\mathbb{Z}_p$ extension K_{cyc} of K is contained in K_∞. We denote $\mathrm{Gal}(K_\infty/K)$, $\mathrm{Gal}(K_\infty/K_{\mathrm{cyc}})$ and $\mathrm{Gal}(K_{\mathrm{cyc}}/K)$ by G, H and Γ respectively. The pair $H \subset G$ satisfies the condition (G).

Let us consider the Pontryagin dual $\mathcal{S}_A^\vee$ of the Selmer group $\mathcal{S}_A$ of the Galois representation $A = T_pE \otimes \mathbb{Q}_p/\mathbb{Z}_p$ over the Galois extension K_∞/K discussed above. We take M to be the module $\mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p)$. It is conjectured that the module $M = \mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p)$ is in the category $\mathfrak{n}_H(G)$ (see [Coates et al. 2005, Conjecture 5.1]) and there are examples where this conjecture is satisfied (see [loc. cit.]).

(2) Let us choose a p-th power free integer $m \geq 2$. Put $K = \mathbb{Q}(\mu_p)$, $K_{\mathrm{cyc}} = \mathbb{Q}(\mu_{p^\infty})$ and $K_\infty = \bigcup_{n=1}^\infty K_{\mathrm{cyc}}(m^{1/p^n})$. Such an extension K_∞/K is called a false-Tate curve extension. We denote $\mathrm{Gal}(K_\infty/K)$, $\mathrm{Gal}(K_\infty/K_{\mathrm{cyc}})$ and $\mathrm{Gal}(K_{\mathrm{cyc}}/K)$ by G, H and Γ respectively. Note that we have $G \cong \mathbb{Z}_p \rtimes \mathbb{Z}_p$, $H \cong \mathbb{Z}_p$ and $\Gamma \cong \mathbb{Z}_p$. Again the pair $H \subset G$ satisfies the condition (G).

Let us consider the Pontryagin dual $\mathcal{S}_A^\vee$ of the Selmer group $\mathcal{S}_A$ of the Galois representation $A = T \otimes \mathbb{Q}_p/\mathbb{Z}_p$ over a Galois extension K_∞/K discussed above. We take M to be the module $\mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p)$. Under certain assumptions on A, it is expected that $\mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p)$ will be in $\mathfrak{n}_H(G)$. We refer to [Hachimori and Venjakob 2003] for some examples of $\mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p)$ which are in $\mathfrak{n}_H(G)$.

(3) Let K be an imaginary quadratic field in which a rational prime $p \neq 2$ splits. Let K_∞ be the unique $\mathbb{Z}_p^{\oplus 2}$-extension of K. Let $G = \mathrm{Gal}(K_\infty/K)$ and $H = \mathrm{Gal}(K_\infty/K_{\mathrm{cyc}})$. Once again the pair $H \subset G$ satisfies the condition (G).

For the Pontryagin dual $\mathcal{S}_A^\vee$ of the Selmer group $\mathcal{S}_A$ of the Galois representation $A = T \otimes \mathbb{Q}_p/\mathbb{Z}_p$ over a Galois extension $K_\infty/\mathbb{Q}$ in this commutative two-variable situation, similar phenomena as above are expected and we take M to be the module $\mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p)$.

In a forthcoming joint work of two of us [Jha and Ochiai $\geq$ 2016], the Main Theorem above will be applied to establish the functional equation of Selmer groups for general p-adic representations over a general noncommutative p-adic Lie extension. This is a partial motivation for our present work for two of us. Note that the third author proved the functional equation of Selmer groups for elliptic curves over false-Tate curve extension (see [Zábrádi 2008]) and for non-CM elliptic curves in GL_2-extension (see [Zábrádi 2010]). But the main method of the papers [Zábrádi 2008; 2010] is not based on the twisting lemma.

Notation. Unless otherwise specified, all modules over $\Lambda_{\mathcal{O}}(G)$ are considered as left modules. Throughout the paper we fix a topological generator γ of Γ.

1. Preliminary Theorem

In this section, we formulate and prove the Preliminary Theorem below, which gives the same conclusion as the Main Theorem under stronger assumptions (i.e., the hypothesis (H) and nonexistence of nontrivial element of order p in G). In the next section, our Main Theorem is deduced from the Preliminary Theorem and the Key Lemma which is given in the next section.

Preliminary Theorem. *Let G be a compact p-adic Lie group without any element of order p and let H be a closed normal subgroup such that G/H is isomorphic to Γ. Let M be a finitely generated torsion $\Lambda_{\mathcal{O}}(G)$-module satisfying the following condition.*

(H) *There is a $\Lambda_{\mathcal{O}}(H)$-linear homomorphism $M \to \mathbb{Z}_p[\![H]\!]^{\oplus d}$ that induces an isomorphism $M \otimes_{\mathbb{Z}_p} \mathbb{Q}_p \xrightarrow{\sim} (\mathbb{Z}_p[\![H]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d}$ after taking $\otimes_{\mathbb{Z}_p}\mathbb{Q}_p$.*

Then there exists a continuous character $\rho : \Gamma \to \mathbb{Z}_p^\times$ such that the largest U-coinvariant quotient $M(\rho)_U$ of $M(\rho)$ is finite for every open normal subgroup U of G.

Before going into the proof of the Preliminary Theorem, we collect some basic results in noncommutative Iwasawa theory which are relevant for the article.

Lemma 1. *Let $H \subset G$ be a pair satisfying the condition (G) and let M be a finitely generated $\Lambda_{\mathcal{O}}(G)$-module which satisfies the condition (H). Then there exists a matrix $A \in M_d(\mathbb{Z}_p[\![H]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)$ such that $M \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ is isomorphic to*

$$(\mathbb{Z}_p[\![G]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d} \big/ (\mathbb{Z}_p[\![G]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d} (\tilde{\gamma} \mathbf{1}_d - A), \tag{1}$$

where γ is a topological generator of Γ and $\tilde{\gamma} \in G$ is a fixed lift of γ and elements in $(\mathbb{Z}_p[\![G]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d}$ are regarded as row vectors.

Proof. Let us take a basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$ of the free $\mathbb{Z}_p[\![H]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$-module $M \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$. Through the isomorphism $M \otimes_{\mathbb{Z}_p} \mathbb{Q}_p \xrightarrow{\sim} (\mathbb{Z}_p[\![H]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d}$ fixed by the condition (H), $\tilde{\gamma}$ acts on M. Thus we define a matrix $A = (a_{ij})_{1 \le i,j \le d} \in M_d(\mathbb{Z}_p[\![H]\!] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)$ by

$$\tilde{\gamma} \cdot \boldsymbol{v}_i = \sum_{1 \le j \le d} a_{ji} \boldsymbol{v}_j.$$

We denote the module presented in (1) by N_A. By construction, we have a $\mathbb{Z}_p[[H]] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$-linear isomorphism $(\mathbb{Z}_p[[H]] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d} \xrightarrow{\sim} N_A$ on which $\tilde{\gamma}$ acts in the same manner as the action of $\tilde{\gamma}$ on $M \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$. □

We denote by $\mathcal{U}$ the set of all open normal subgroups U of G. We remark that the set $\mathcal{U}$ is a countable set since G is profinite and has a countable base at the identity.

Lemma 2. *For any $U \in \mathcal{U}$, $\mathbb{Z}_p[G/U] \otimes_{\mathbb{Z}_p} \overline{\mathbb{Q}}_p$ is isomorphic to a finite number of products of matrix algebras $\prod_{i=1}^{k(U)} M_{r_i}(\overline{\mathbb{Q}}_p)$.*

Proof. First of all, the algebra $\mathbb{Z}_p[G/U] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p \cong \mathbb{Q}_p[G/U]$ is a semisimple algebra over $\mathbb{Q}_p$ since G/U is a finite group and $\mathbb{Q}_p$ is of characteristic 0. We have an isomorphism

$$\mathbb{Q}_p[G/U] \cong \prod_{i=1}^{l_n} M_{s_i}(D_i),$$

where D_i is a finite dimensional division algebra over $\mathbb{Q}_p$. For each i, the center K_i of D_i is a finite extension of $\mathbb{Q}_p$. It is well-known that $\dim_{K_i} D_i$ is a square of some natural number t_i and $D_i \otimes_{K_i} \overline{\mathbb{Q}}_p$ is isomorphic to $M_{t_i}(\overline{\mathbb{Q}}_p)$. Thus $M_{s_i}(D_i) \otimes \overline{\mathbb{Q}}_p$ is isomorphic to $[K_i : \mathbb{Q}_p]$ copies of $M_{s_i+t_i}(\overline{\mathbb{Q}}_p)$. The lemma follows immediately from this. □

Proof of the Preliminary Theorem. First, we remark that for an open normal subgroup U of G, we have

$$M(\rho)_U \text{ is finite if and only if } M(\rho)_U \otimes_{\mathbb{Z}_p} \mathbb{Q}_p = 0. \tag{2}$$

Since the operation of taking the base extension $\otimes_{\mathbb{Z}_p}\mathbb{Q}_p$ commutes with the operation of taking the largest U-coinvariant quotient, by Lemma 1 we have

$$\begin{aligned} &M(\rho)_U \otimes_{\mathbb{Z}_p} \mathbb{Q}_p \\ &\qquad \cong (\mathbb{Z}_p[G/U] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d} \big/ (\mathbb{Z}_p[G/U] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^{\oplus d} (\tilde{\gamma}_U^{\oplus d} - A_U(\rho)), \end{aligned} \tag{3}$$

where we denote the projection of $\tilde{\gamma} \in G$ to G/U by $\tilde{\gamma}_U$. Here, the matrix $A_U(\rho) \in M_d(\mathbb{Z}_p[G/U] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)$ is defined as the image of $\rho(\gamma)^{-1} A \in M_d(\mathbb{Z}_p[[H]] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)$ via the composite map

$$M_d(\mathbb{Z}_p[[H]] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p) \longrightarrow M_d(\mathbb{Z}_p[[G]] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p) \longrightarrow M_d(\mathbb{Z}_p[G/U] \otimes_{\mathbb{Z}_p} \mathbb{Q}_p).$$

Taking the base extension $\otimes_{\mathbb{Q}_p} \overline{\mathbb{Q}}_p$ of the isomorphism (3), we have a $\overline{\mathbb{Q}}_p$-linear isomorphism by Lemma 2:

$$M(\rho)_U \otimes_{\mathbb{Z}_p} \overline{\mathbb{Q}}_p \cong \prod_{i=1}^{k(U)} M_{r_i}(\overline{\mathbb{Q}}_p)^{\oplus d} \Big/ M_{r_i}(\overline{\mathbb{Q}}_p)^{\oplus d} (\gamma_{U,i}^{\oplus d} - A_{U,i}(\rho)) \tag{4}$$

where $\gamma_{U,i} \in \mathrm{Aut}_{\overline{\mathbb{Q}}_p}(M_{r_i}(\overline{\mathbb{Q}}_p))$ and $A_{U,i}(\rho) \in \mathrm{End}_{\overline{\mathbb{Q}}_p}(M_{r_i}(\overline{\mathbb{Q}}_p)^{\oplus d})$ are defined as follows. We consider the base extension to $\overline{\mathbb{Q}}_p$ of

$$\tilde{\gamma}_U \in \mathrm{Aut}_{\mathbb{Z}_p[G/U]\otimes_{\mathbb{Z}_p}\mathbb{Q}_p}(\mathbb{Z}_p[G/U]\otimes_{\mathbb{Z}_p}\mathbb{Q}_p) \subset \mathrm{Aut}_{\mathbb{Q}_p}(\mathbb{Z}_p[G/U]\otimes_{\mathbb{Z}_p}\mathbb{Q}_p).$$

This is an element of $\mathrm{Aut}_{\overline{\mathbb{Q}}_p}\big(\prod_{i=1}^{k(U)} M_{r_i}(\overline{\mathbb{Q}}_p)\big)$. We denote the projection of this element to the i-th component by $\gamma_{U,i}$. The base extension to $\overline{\mathbb{Q}}_p$ of

$$A_U(\rho) \in M_d(\mathbb{Z}_p[G/U]\otimes_{\mathbb{Z}_p}\mathbb{Q}_p) \subset \mathrm{End}_{\mathbb{Q}_p}\big((\mathbb{Z}_p[G/U]\otimes_{\mathbb{Z}_p}\mathbb{Q}_p)^{\oplus d}\big)$$

is an element of $\mathrm{End}_{\overline{\mathbb{Q}}_p}\big(\prod_{i=1}^{k(U)} M_{r_i}(\overline{\mathbb{Q}}_p)^{\oplus d}\big)$, and we denote the projection of this element to the i-th component by $A_{U,i}(\rho)$.

Now, we denote by $A_{U,i}$ the element $A_{U,i}(\mathbf{1})$. We remark that $A_{U,i}(\rho)$ is equal to $\rho(\gamma)^{-1}A_{U,i}$ for any continuous character $\rho : \Gamma \to \mathbb{Z}_p^\times$. We define $\mathrm{EV}_{U,i}$ to be the set of roots of the characteristic polynomial

$$P_{U,i}(T) := \det(\gamma_{U,i}^{\oplus d} - A_{U,i}T).$$

Since $\gamma_{U,i}^{\oplus d}$ is an automorphism, the polynomial $P_{U,i}(T)$ is not zero. Hence $\mathrm{EV}_{U,i}$ is a finite set. We denote the union of $\mathrm{EV}_{U,i}$ for $1 \le i \le k(U)$ by EV_U, which is again a finite set. If $\rho(\gamma)^{-1}$ is not contained in $\mathrm{EV}_U \cap \mathbb{Z}_p^\times$, the module in (4) is zero and hence the module in (3) is zero. Now, we denote by EV_M the union of $\mathrm{EV}_U \cap \mathbb{Z}_p^\times$ over all $U \in \mathcal{U}$. Since $\mathcal{U}$ is a countable set, EV_M is a countable set. Thus $\mathbb{Z}_p^\times \setminus \mathrm{EV}_M$ is nonempty since $\mathbb{Z}_p^\times$ is uncountable. By choosing $\rho(\gamma)^{-1} \in \mathbb{Z}_p^\times \setminus \mathrm{EV}_M$, we complete the proof. □

2. Proof of the Main Theorem

In this section, we prove the Main Theorem, which relies on the following result.

Key Lemma. *Let G be a compact p-adic Lie group without any element of order p and let H be a closed subgroup such that G/H is isomorphic to Γ. Let M be $\Lambda_{\mathcal{O}}(G)$-module which is finitely generated over $\Lambda_{\mathcal{O}}(H)$. Then, there exists an open subgroup $G_0 \subset G$ containing H, a $\Lambda_{\mathcal{O}}(G_0)$-module N which is a free $\Lambda_{\mathcal{O}}(H)$-module of finite rank, and a surjective $\Lambda_{\mathcal{O}}(G_0)$-linear homomorphism $N \twoheadrightarrow M$.*

Proof. We denote by I the Jacobson radical of $\Lambda_{\mathcal{O}}(H)$. Note that I is a two sided ideal of $\Lambda_{\mathcal{O}}(H)$ such that we have $\Lambda_{\mathcal{O}}(H)/I \cong \mathbb{F}_q$, where $\mathbb{F}_q$ is the residue field of $\mathcal{O}$. We also have $\Lambda_{\mathcal{O}}(G)/I \cong \mathbb{F}_q[\![\Gamma]\!]$ by definition.

Let us take a system of generators $m_1, \dots, m_d$ of M as a $\Lambda_{\mathcal{O}}(H)$-module. Note that M is equipped with a topology obtained by a natural $\Lambda_{\mathcal{O}}(H)$-module structure. The set $\{I^n M\}_{n\in\mathbb{N}}$ forms a system of open neighborhoods of M.

Choose a topological generator γ of Γ and take a lift $\tilde{\gamma} \in G$ of γ. By continuity of the action of G on M, the following two conditions hold true simultaneously for a sufficiently large integer n:

(i) We have $(\tilde{\gamma}^{p^n} - 1)m_i \in IM$ for any i with $1 \le i \le d$.

(ii) The conjugate action of $\tilde{\gamma}^{p^n}$ on I/I^2 is trivial.

We will choose and fix a natural number n satisfying the conditions (i) and (ii). Then we define G_0 to be the preimage of Γ^{p^n} by the surjection $G \twoheadrightarrow \Gamma$. By definition, G_0 is an open subgroup of G which contain H.

Let us consider the set $\{a_{ij} \in I\}_{1 \le i,j \le d}$ such that we have $(\tilde{\gamma}^{p^n} - 1)m_j = \sum_{i=1}^d a_{ij} m_i$. We consider F (resp. F') which is a free $\Lambda_{\mathcal{O}}(G_0)$-module of rank d equipped with a system of generators $f_1, \dots, f_d$ (resp. $f'_1, \dots, f'_d$). We consider a $\Lambda_{\mathcal{O}}(G_0)$-linear homomorphism

$$\varphi : F' \longrightarrow F, \quad f'_j \mapsto (\tilde{\gamma}^{p^n} - 1) f_j - \sum_{i=1}^d a_{ij} f_i .$$

We define a $\Lambda_{\mathcal{O}}(G_0)$-module N to be the cokernel of the map φ above.

Claim. *For each i with $1 \le i \le d$, we denote the image of f_i by $\bar{f}_i$. Then the $\Lambda_{\mathcal{O}}(G_0)$-module N is a free $\Lambda_{\mathcal{O}}(H)$-module of finite rank d with a system of generators $\bar{f}_1, \dots, \bar{f}_d$.*

If the claim holds true, a $\Lambda_{\mathcal{O}}(G_0)$-linear homomorphism $N \twoheadrightarrow M$ sending $\bar{f}_i$ to m_i for each i is surjective. Since N is free over $\Lambda_{\mathcal{O}}(H)$, this is what we want. Thus it remains only to prove the claim.

By applying the functor $\Lambda_{\mathcal{O}}(G_0)/I\Lambda_{\mathcal{O}}(G_0) \otimes_{\Lambda_{\mathcal{O}}(G_0)} \cdot$ to the map φ, and noting that

$$\Lambda_{\mathcal{O}}(G_0)/I \cong \mathbb{F}_q[\![\Gamma]\!],$$

we obtain

$$\varphi_I : \oplus_{j=1}^d \mathbb{F}_q[\![\Gamma^{p^n}]\!] f'_j \xrightarrow{\times(\tilde{\gamma}^{p^n}-1)} \oplus_{j=1}^d \mathbb{F}_q[\![\Gamma^{p^n}]\!] f_j .$$

Since N/IN is isomorphic to the cokernel of the above map φ_I, N/IN is a free $\mathbb{F}_q$-module of rank d. By applying the topological Nakayama lemma (see [Balister and Howson 1997, Corollary in §3]) to the compact $\Lambda_{\mathcal{O}}(H)$-module N, N is generated by $\bar{f}_1, \dots, \bar{f}_d$ over $\Lambda_{\mathcal{O}}(H)$. We will prove that N is free of rank d over $\Lambda_{\mathcal{O}}(H)$ with this system of generators. Let r be an arbitrary natural number. Since we have a natural surjection from the r-fold tensor product of I/I^2 to I^r/I^{r+1}, the conjugate action of $\tilde{\gamma}^{p^n}$ on I^r/I^{r+1} is also trivial. Thus, by applying the functor $I^r/I^{r+1}\Lambda_{\mathcal{O}}(G_0) \otimes_{\Lambda_{\mathcal{O}}(G_0)} \cdot$ to the map φ, we obtain a $\Lambda_{\mathcal{O}}(G_0)$-linear map

$$\varphi \otimes I^r/I^{r+1} : I^r F'/I^{r+1} F' \longrightarrow I^r F/I^{r+1} F$$

which is again defined as a multiplication of $(\tilde{\gamma}^{p^n} - 1)$. This proves

$$\dim_{\mathbb{F}_q} N/I^s N = \sum_{r=0}^{s-1} \dim_{\mathbb{F}_q} I^r N/I^{r+1} N = \sum_{r=0}^{s-1} \dim_{\mathbb{F}_q} (I^r/I^{r+1})^{\oplus d}.$$

Thus the cardinality of $N/I^s N$ is equal to the cardinality of $(\Lambda_{\mathcal{O}}(H)/I)^{\oplus d}$ for any natural number s, which implies that N is free of rank d over $\Lambda_{\mathcal{O}}(H)$. This completes the proof of the claim. □

Proof of Main Theorem. We will use the Key Lemma and the Preliminary Theorem to prove the Main Theorem in two steps.

First, we consider the situation where G is a compact p-adic Lie group without any element of order p and H is a closed subgroup such that G/H is isomorphic to Γ. Thus we dropped the assumption (H) of the Preliminary Theorem but we still keep the assumption of nonexistence of a nontrivial element of order p in G.

Let M be $\Lambda_{\mathcal{O}}(G)$-module which is finitely generated over $\Lambda_{\mathcal{O}}(H)$. By the Key Lemma, for a sufficiently large natural number n, we have a surjective $\Lambda_{\mathcal{O}}(G_0)$-linear homomorphism $N \twoheadrightarrow M$ from a free $\Lambda_{\mathcal{O}}(H)$-module N of finite rank. Here G_0 is a unique open subgroup $G_0 \subset G$ of index p^n containing H. Note that the module N satisfies condition (H) of the Preliminary Theorem. We thus find a continuous character $\rho_0 : \Gamma^{p^n} \to \mathbb{Z}_p^\times$ such that $N(\rho_0)_{U_0}$ is finite for any open normal subgroup U_0 of G_0. By the proof of the Preliminary Theorem, we can choose uncountably many such ρ_0. Thus, we see that we can take ρ_0 as above so that the value of ρ_0 is contained in a open subgroup $1 + p^n\mathbb{Z}_p$ of $\mathbb{Z}_p^\times$. Then, we take a continuous character $\rho : \Gamma \to \mathbb{Z}_p^\times$ whose restriction to Γ^{p^n} coincides with ρ_0. The twist $M(\rho)$ with this character is what we want in our Main Theorem. In fact, for any open normal subgroup U of G, we have a surjection $N(\rho_0)_{U_0} \twoheadrightarrow M(\rho)_U$ taking an open normal subgroup U_0 of G_0 contained in U. Since $N(\rho_0)_{U_0}$ is finite by the Preliminary Theorem, $M(\rho)_U$ must be finite. Thus we finished the proof of our Main Theorem under the assumption of nonexistence of a nontrivial element of order p in G.

Now we deduce our Main Theorem assuming that it is true under the assumption of nonexistence of a nontrivial element of order p in G. We consider the situation where G is a compact p-adic Lie group with elements of order p and H is a closed subgroup such that G/H is isomorphic to Γ. Let M be a $\Lambda_{\mathcal{O}}(G)$-module which is finitely generated over $\Lambda_{\mathcal{O}}(H)$. Let G' be a uniform open normal subgroup of G (see [Lazard 1965, Chapter III, §(3.1)]), which is automatically without any elements of order p. Let H' be the intersection of H and G'. Since M is finitely generated over $\Lambda_{\mathcal{O}}(H)$ and since H' is of finite index in H, M is finitely generated over $\Lambda_{\mathcal{O}}(H')$. According to the result in our first step, there exist a continuous character $\rho' : G'/H' \to \mathbb{Z}_p^\times$ such that $M(\rho')_{U'}$ is finite for every open subgroup U'

of G'. Note that G'/H' is naturally regarded as an open subgroup of G/H. Thus by choosing ρ' so that the image of ρ' is small enough in $\mathbb{Z}_p^\times$ compared to the index of G'/H' in G/H, there exists a continuous character $\rho : G/H \to \mathbb{Z}_p^\times$ whose restriction to G'/H' coincides with ρ'. Now for any open normal subgroup U of G, we take an open normal subgroup U' of G' which is contained in U. We have a natural map $M(\rho')_{U'} \to M(\rho)_U$, where $M(\rho')_{U'}$ is finite by the choice of ρ' and by our discussion above. Unlike in the first step, $M(\rho')_{U'} \to M(\rho)_U$ is not necessarily surjective. However, the cokernel of this map is still finite by construction. We thus deduce that $M(\rho)_U$ is finite, which completes the proof of the Main Theorem. □

Remark (p-torsion modules). For a compact p-adic Lie group G without any element of order p, it is well-known that $\Lambda_O(G)$ is left and right noetherian. Let N be a finitely generated p-primary torsion left $\Lambda_{\mathcal{O}}(G)$ module. Then, we have $N = N[p^r]$ for some $r \in \mathbb{N}$. For any open normal subgroup U of G, N_U is a finitely generated $\mathbb{Z}/p^r\mathbb{Z}[G/U]$ module. In other words, N_U is always finite when N is of p-primary torsion.

For a finitely generated torsion $\Lambda_{\mathcal{O}}(G)$ module M, we consider the exact sequence

$$0 \longrightarrow M(p) \longrightarrow M \longrightarrow M/M(p) \longrightarrow 0,$$

where $M(p)$ is the largest p-primary torsion submodule of M. Then, from the preceding discussion, it is clear that in the situation of the Main Theorem, for any continuous $\rho : \Gamma \to \mathbb{Z}_p^\times$ and for any open normal subgroup U of G, $M(\rho)_U$ is finite if and only if $\big((M/M(p))(\rho)\big)_U$ is finite.

In particular, when we want to apply the Main Theorem to arithmetic situations coming from Selmer groups of certain Galois modules A, we remark that $\mathcal{S}_A^\vee(\rho)_U$ is finite if and only if $\big((\mathcal{S}_A^\vee/\mathcal{S}_A^\vee(p))(\rho)\big)_U$ is finite.

Acknowledgements

A part of the paper was done when S. Jha visited Osaka University. He thanks Osaka University for their hospitality. The paper was finalized when T. Ochiai stayed at the Indian Institute of Technology Kanpur (IITK) on September 2015. He thanks IITK for their hospitality. We are grateful to Prof. John Coates for encouragement and invaluable suggestions leading to improvements on an earlier draft of the paper.

References

[Balister and Howson 1997] P. N. Balister and S. Howson, "Note on Nakayama's lemma for compact Λ-modules", *Asian J. Math.* **1**:2 (1997), 224–229. MR 1491983 Zbl 0904.16019

[Coates et al. 2005] J. Coates, T. Fukaya, K. Kato, R. Sujatha, and O. Venjakob, "The GL_2 main conjecture for elliptic curves without complex multiplication", *Inst. Hautes Études Sci. Publ. Math.* **101** (2005), 163–208. MR 2217048 Zbl 1108.11081

[Greenberg 1989] R. Greenberg, "Iwasawa theory for p-adic representations", pp. 97–137 in *Algebraic number theory*, edited by J. Coates et al., Adv. Stud. Pure Math. **17**, Academic Press, Boston, 1989. MR 1097613 Zbl 0739.11045

[Hachimori and Venjakob 2003] Y. Hachimori and O. Venjakob, "Completely faithful Selmer groups over Kummer extensions", *Doc. Math.* **Extra Vol.** (2003), 443–478. MR 2046605 Zbl 1117.14046

[Jha and Ochiai $\geq$ 2016] S. Jha and T. Ochiai, "Functional equation of Selmer groups over p-adic Lie extension", in preparation.

[Lazard 1965] M. Lazard, "Groupes analytiques p-adiques", *Inst. Hautes Études Sci. Publ. Math.* **26** (1965), 5–219. MR 0209286 Zbl 0139.02302

[Perrin-Riou 2003] B. Perrin-Riou, "Groupes de Selmer et accouplements: cas particulier des courbes elliptiques", *Doc. Math.* **Extra Vol.** (2003), 725–760. MR 2046613 Zbl 1142.11337

[Zábrádi 2008] G. Zábrádi, "Characteristic elements, pairings and functional equations over the false Tate curve extension", *Math. Proc. Cambridge Philos. Soc.* **144**:3 (2008), 535–574. MR 2418704 Zbl 1243.11106

[Zábrádi 2010] G. Zábrádi, "Pairings and functional equations over the GL_2-extension", *Proc. Lond. Math. Soc.* (3) **101**:3 (2010), 893–930. MR 2734964 Zbl 1272.11119

Communicated by John Henry Coates
Received 2015-10-21 Revised 2015-12-22 Accepted 2016-02-01

jhasom@iitk.ac.in — *Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur 208016, India*

ochiai@math.sci.osaka-u.ac.jp — *Department of Mathematics, Graduate School of Science, Osaka University, Machikaneyama 1-1, Toyonaka, Osaka 5600043, Japan*

zger@cs.elte.hu — *Department of Algebra and Number Theory, Mathematical Institute, Eötvös Loránd University, Bertalan Lajos utca 11, 1111 Budapest, Hungary*

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.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

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 10 No. 3 2016