

MODULAR SHIMURA VARIETIES AND FORGETFUL MAPS

VICTOR ROTGER

ABSTRACT. In this note we consider several maps that occur naturally between modular Shimura varieties, Hilbert-Blumenthal varieties and the moduli spaces of polarized abelian varieties when forgetting certain endomorphism structures. We prove that, up to birational equivalences, these forgetful maps coincide with the natural projection by suitable abelian groups of Atkin-Lehner involutions.

INTRODUCTION

Let F be a totally real number field of degree $[F : \mathbb{Q}] = n$ and let B be a totally indefinite quaternion algebra over F . In this note we will be concerned with certain Shimura varieties X_B attached to an arithmetic datum arising from B and several maps that occur naturally between them.

As complex manifolds, these varieties can be described as quotients of certain bounded symmetric domains by arithmetic groups acting on them, and, by the theory of Baily-Borel ([BaBo]), they become quasi-projective complex algebraic varieties. Shimura discovered a moduli interpretation of these varieties which allowed him to construct canonical models X_B/\mathbb{Q} over the field \mathbb{Q} of rational numbers. Shimura also explored their arithmetic, showing that the coordinates of so-called *Heegner points* on X_B generate certain class fields and that the Galois action on them can be described by explicit reciprocity laws ([Sh1], [Sh2]).

The nature of the Shimura varieties X_B differs notably depending on the existence or absence of zero divisors in B . When B is the split algebra $M_2(F)$, the varieties X_B are classically called *Hilbert-Blumenthal modular varieties*. These are non-complete, and suitable compactifications of them can be constructed, though at the cost of producing new singularities. The literature on them is abundant, especially on the low dimensional cases (in dimension 1, these are called modular curves and have become crucial in many aspects of number theory; in dimension 2, a reference to Hilbert modular surfaces is van der Geer's book [vdGe]).

On the other hand, when B is non-split, that is, it is a division algebra, then the emerging Shimura varieties X_B/\mathbb{Q} are already projective. This fact makes the study of their arithmetic highly difficult since, in the Hilbert-Blumenthal case, much of it is encoded in the added cusps. In remarkable contrast to Hilbert-Blumenthal varieties, Shimura proved that, when $B \not\cong M_2(F)$, the varieties X_B do not have real points and therefore do not have rational points over any number field that

Received by the editors August 19, 2002 and, in revised form, February 12, 2003.

2000 *Mathematics Subject Classification*. Primary 11G18, 14G35.

Key words and phrases. Shimura variety, moduli space, abelian variety, quaternion algebra.

This research was partially supported by a grant FPI from the Ministerio de Ciencia y Tecnología BFM2000-0627.

admits a real embedding. In the last years, there has been increasing interest in Shimura curves arising from rational indefinite quaternion algebras, since they play a crucial role in modularity questions (cf. [HaHaMo] or [Ri], for instance).

From the modular point of view, there are natural maps

$$\pi : X_B \rightarrow \mathcal{A}_{g,(d_1,\dots,d_g)}$$

from X_B into the moduli spaces of polarized abelian varieties that involve *forgetting* some additional structures. Further, for any totally real field L containing F and embedded in B , these maps factor through a morphism $\pi_L : X_B \rightarrow \mathcal{H}_L$ into a Hilbert-Blumenthal modular variety (cf. Section 2).

It is the purpose of this note to describe in detail the nature of these morphisms and their image in the several moduli spaces of abelian varieties. As we will see, π and π_L are quasifinite maps and the mere computation of their degree turns out to be unexpectedly subtle, as it is based on arithmetic questions on B that were recently studied by Chinburg and Friedman in a number of papers in relation with arithmetic hyperbolic 3-orbifolds (cf. [ChFr1], [ChFr2]).

In the first section we briefly recall some basic facts on quaternion algebras that we need throughout the article. Next, in Section 2 we review the construction due to Shimura of the varieties we will be dealing with and the above mentioned maps. In the third section, we introduce several Atkin-Lehner groups of automorphisms acting on them and describe their modular interpretation. In Section 4 we introduce what we call the *stable* and *twisting Atkin-Lehner groups* and we state our main result in Theorem 4.4.

The rest of the note is devoted to presenting the proof of Theorem 4.4. To do so, in Section 5 we introduce what we call the B -part of the Néron-Severi group of an abelian variety whose algebra of endomorphisms $\text{End}(A) \otimes \mathbb{Q}$ contains the quaternion algebra B .

In closing this introduction, we would point out two different applications of the results presented in this note. The first one concerns the geometry of the quaternionic locus $\mathcal{Q}_{\mathcal{O}} = \{[A, \mathcal{L}], \text{End}(A) \supseteq \mathcal{O}\}$ of abelian varieties admitting multiplication by a maximal order \mathcal{O} in the moduli space \mathcal{A}_g of principally polarized abelian varieties of even dimension g . By means of Theorem 4.4 and Eichler's theory on optimal embeddings, the number of irreducible components of $\mathcal{Q}_{\mathcal{O}}$ can be related to certain class numbers, and conditions can be given for its irreducibility. The emerging picture bears a remarkable resemblance to the classical theory of genus of quadratic forms and *idonei numbers* (cf. [Ro4]).

Secondly, Theorem 4.4 can also be used to explore the arithmetic of QM-abelian varieties. Indeed, in [Ro3], we use it to describe the field of moduli of the quaternionic endomorphisms acting on them.

Finally, the results of this paper are also used by Clark in [Cl] in relation to abelian surfaces with potential quaternionic multiplication and their moduli.

1. AN OVERVIEW OF QUATERNION ALGEBRAS

In this section let F denote an arbitrary number field. Let B be a quaternion algebra over F , that is, a simple algebra of $\text{rank}_F(B) = 4$ over its centre F . Elements $i, j \in B$ can be chosen such that $B = F + Fi + Fj + Fij$ with the relations $ij = -ji$ and $i^2 = a, j^2 = b \in F^*$. The classical notation for it is $B = \left(\frac{a,b}{F}\right)$.

A place v of F , archimedean or not, is said to ramify in B if $B \otimes F_v$ is a division algebra over the completion F_v of F at v . When v is archimedean, in this case we also say that B is *definite* at v and otherwise we call it *indefinite*. A quaternion algebra that is indefinite at all archimedean places is called *totally indefinite*.

There are finitely many places of F that ramify in a quaternion algebra B , and they occur in an even number. The discriminant ideal of B is the (square-free) product $\text{disc}(B) = \prod \mathfrak{p}$ of the finite primes that ramify in B .

Let $\beta = x + yi + zj + tij \mapsto \bar{\beta} = x - yi - zj - tij$ denote the conjugation map on B . Then the reduced trace and reduced norm of β are defined as $\text{tr}(\beta) = \beta + \bar{\beta} \in F$ and $\text{n}(\beta) = \beta\bar{\beta} \in F$ respectively. An element $\beta \in B$ of null reduced trace $\text{tr}(\beta) = 0$ is called *pure* and, for any sub-module \mathcal{I} of B , we will denote by $\mathcal{I}_0 \subseteq B_0$ the group of pure quaternions of \mathcal{I} . An element $\beta \in B$ is integral over the ring of integers R_F of F if both $\text{tr}(\beta)$ and $\text{n}(\beta)$ lie in R_F . An order \mathcal{O} of integers in B is an R_F -finitely generated subring of integral elements such that $\mathcal{O} \otimes_{R_F} F = B$. It is called a *maximal order* if it is not properly contained in any other. Unlike number fields, maximal orders are not unique. The *type number* $t(B)$ of B is the number of B^* -conjugation classes of maximal orders in B .

A left ideal of a maximal order \mathcal{O} is an R_F -finitely generated module $\mathcal{I} \subset B$ with $\mathcal{I} \otimes_{R_F} F = B$ and $\mathcal{O} \cdot \mathcal{I} = \mathcal{I}$. The ideal is two-sided if in addition $\mathcal{I} \cdot \mathcal{O} = \mathcal{I}$. We will often consider \mathcal{I} in $\text{Pic}_\ell(\mathcal{O})$, that is, up to principal \mathcal{O} -left ideals $\mathcal{O}\beta$, $\beta \in B^*$. See [Vi], p.25, for details. The conjugate ideal of an ideal \mathcal{I} in B is the ideal $\bar{\mathcal{I}} = \{\bar{\beta}, \beta \in \mathcal{I}\}$. If \mathcal{I} is a left ideal of a maximal order \mathcal{O} , then $\bar{\mathcal{I}}$ is a right \mathcal{O} -ideal. The product ideal $\mathcal{I} \cdot \bar{\mathcal{I}}$ is then a two-sided ideal of \mathcal{O} , and we actually have that $\mathcal{N}(\mathcal{I}) := \text{n}(\mathcal{I})\mathcal{O} = \mathcal{I} \cdot \bar{\mathcal{I}}$.

Finally, if $K \subset F$ is any sub-field of F , the *codifferent* of \mathcal{I} over K is defined to be $\mathcal{I}_{B/K}^\# = \{\beta \in B : \text{tr}_{B/K}(\mathcal{I}\beta) \subseteq R_K\}$. It is a right ideal of \mathcal{O} . In our considerations on Néron-Severi groups, the group $\mathcal{N}(\mathcal{I})_0^\#$ of pure quaternions of the codifferent of the norm ideal $\mathcal{N}(\mathcal{I})$ over \mathbb{Q} will play an important role (cf. Section 5).

Assume now that F is totally real and let F_+^* denote the group of totally positive elements. We will also let $R_{F_+}^*$ stand for the group of totally positive units in R_F^* . A positive (anti-)involution ϱ on B is a map $\varrho : B \rightarrow B$ such that $(\beta_1 + \beta_2)^\varrho = \beta_1^\varrho + \beta_2^\varrho$ and $(\beta_1 \cdot \beta_2)^\varrho = \beta_2^\varrho \cdot \beta_1^\varrho$ for any $\beta_1, \beta_2 \in B$, and such that $\text{tr}(\beta \cdot \beta^\varrho) \in F_+^*$ for any $\beta \in B^*$.

By the Skolem-Noether Theorem, if ϱ is a positive (anti-)involution, there exists $\mu \in B^*$ such that $\beta^\varrho = \mu^{-1}\bar{\beta}\mu$. Further, it is easily shown ([Ro2]) that the positive-ness of ϱ implies that $\text{tr}(\mu) = 0$ and $\text{n}(\mu) \in F_+^*$. The element μ is determined up to multiplication by elements of F^* , and we will sometimes use the notation $\varrho = \varrho_\mu$.

2. SHIMURA VARIETIES AND MAPS BETWEEN THEM

Let F be a totally real number field of degree n and B a totally indefinite quaternion algebra over it. Call $\mathcal{D} = \text{disc}(B) = \mathfrak{p}_1 \cdot \dots \cdot \mathfrak{p}_{2r}$ its discriminant.

Fix a datum $(\mathcal{O}, \mathcal{I}, \varrho)$ consisting of a maximal order \mathcal{O} in B , a left \mathcal{O} -ideal \mathcal{I} (or rather its class in $\text{Pic}_\ell(\mathcal{O})$) and a positive involution ϱ_μ with $\mu \in \mathcal{O}$, $\mu^2 + \delta = 0$, $\delta \in F_+^*$. Attached to the datum $(\mathcal{O}, \mathcal{I}, \varrho)$ there is the following moduli problem over \mathbb{Q} : classifying isomorphism classes of triplets (A, ι, \mathcal{L}) where

- A is an abelian variety of dimension $g = 2n$.

- $\iota : \mathcal{O} \hookrightarrow \text{End}(A)$ is a ring homomorphism such that $H_1(A, \mathbb{Z})$, regarded as a left \mathcal{O} -module, is isomorphic to the left ideal \mathcal{I} .
- \mathcal{L} is a primitive polarization on A such that the Rosati involution $\circ : \text{End}^0(A) \rightarrow \text{End}^0(A)$ with respect to \mathcal{L} on $\text{End}^0(A) = \text{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$ coincides with ϱ when restricted to $\iota(\mathcal{O})$: $\circ|_{\iota(\mathcal{O})} = \varrho \cdot \iota$.

This triplet will be referred to as a polarized abelian variety with multiplication by \mathcal{O} . Two triplets $(A_1, \iota_1, \mathcal{L}_1)$, $(A_2, \iota_2, \mathcal{L}_2)$ are isomorphic if there exists an isomorphism $\alpha \in \text{Hom}(A_1, A_2)$ such that $\alpha\iota_1(\beta) = \iota_2(\beta)\alpha$ for any $\beta \in \mathcal{O}$ and $\alpha^*(\mathcal{L}_2) = \mathcal{L}_1 \in \text{NS}(A_1)$. We recall that a polarization $\mathcal{L} \in \text{NS}(A)$ is called primitive if $\mathcal{L} \notin d \cdot \text{NS}(A)$ for any $d \in \mathbb{Z}, d \geq 2$. Its type is then $(1, d_2, \dots, d_g)$ for $d_i \mid d_{i+1}$, $i = 2, \dots, g - 1$. Remark also that, since a priori there is not a canonical structure of R_F -algebra on $\text{End}(A)$, the immersion $\iota : \mathcal{O} \hookrightarrow \text{End}(A)$ is just a homomorphism of rings (or \mathbb{Z} -algebras).

The corresponding moduli functor is coarsely represented by an irreducible and reduced quasi-projective scheme $X_B/\mathbb{Q} = X_{(\mathcal{O}, \mathcal{I}, \varrho)}/\mathbb{Q}$ over \mathbb{Q} and of dimension $n = [F : \mathbb{Q}]$. Moreover, if B is division (that is, $r > 0$), the Shimura variety X_B is complete (cf. [Sh1], [Sh2]).

Complex analytically, the manifold $X_B(\mathbb{C})$ can be described as the quotient of a symmetric space by the action of a discontinuous group as follows. Since B is totally indefinite, we may fix an embedding $B \hookrightarrow B \otimes_{\mathbb{Q}} \mathbb{R} \xrightarrow{\cong} \text{M}_2(\mathbb{R}) \oplus \dots \oplus \text{M}_2(\mathbb{R})$ and regard the group $\Gamma_B = \mathcal{O}^\times = \{\gamma \in \mathcal{O}^*, n(\gamma) = 1\}$ as a discrete subgroup of $\text{SL}_2(\mathbb{R})^n$. An element $\gamma = (\gamma_1, \dots, \gamma_n) \in \Gamma_B$ acts on the cartesian product \mathfrak{H}^n of n copies of Poincaré’s upper half plane $\mathfrak{H} = \{x + yi, x, y \in \mathbb{R}, y > 0\}$ by Moebius transformations: $\gamma \cdot (\tau_1, \dots, \tau_n)^t = \left(\frac{a_1\tau_1+b_1}{c_1\tau_1+d_1}, \dots, \frac{a_n\tau_n+b_n}{c_n\tau_n+d_n}\right)^t$ where $\gamma_i = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \in \text{SL}_2(\mathbb{R})$. Then

$$\Gamma_B \backslash \mathfrak{H}^n \simeq X_B(\mathbb{C}).$$

2.1. The maps into $\mathcal{A}_g, \mathcal{H}_F$ and \mathcal{H}_L . Now let (A, ι, \mathcal{L}) be a (primitively) polarized abelian variety with multiplication by \mathcal{O} with respect to the datum $(\mathcal{O}, \mathcal{I}, \varrho)$. Note that the type $(1, d_2, \dots, d_g)$ of \mathcal{L} was not specified when posing the above moduli problem. However, since $X_{(\mathcal{O}, \mathcal{I}, \varrho)}$ is connected, the type of \mathcal{L} depends only on the datum $(\mathcal{O}, \mathcal{I}, \varrho)$ and not on the particular triplet (A, ι, \mathcal{L}) . Thus $(1, d_2, \dots, d_g)$ will often be referred to as the type of $(\mathcal{O}, \mathcal{I}, \varrho)$. This observation allows us to define a natural morphism

$$\pi : \begin{array}{ccc} X_{(\mathcal{O}, \mathcal{I}, \varrho)} & \longrightarrow & \mathcal{A}_{g, (1, d_2, \dots, d_g)} \\ (A, \iota, \mathcal{L}) & \longmapsto & (A, \mathcal{L}) \end{array}$$

from the Shimura variety to the moduli space of polarized abelian varieties of type $(1, d_2, \dots, d_g)$ that consists of *forgetting* the quaternionic endomorphism structure. This morphism is representable, proper and defined over \mathbb{Q} . Moreover, as we now explain, the morphism $\pi : X \rightarrow \mathcal{A}_{g, (1, d_2, \dots, d_g)}$ factors in a natural way through certain Hilbert modular varieties.

Definition 2.1. An Eichler pair (S, φ) for \mathcal{O} is a pair consisting of

- an order S over R_F in a quadratic extension L of F and
- an R_F -embedding $\varphi : S \hookrightarrow \mathcal{O}$ such that $\varphi(S) = \varphi(L) \cap \mathcal{O}$.

An Eichler pair is totally real if L is.

Note that not all orders S in quadratic extensions L of F can be embedded in \mathcal{O} . Namely, by Eichler’s theory on optimal embeddings ([Vi], 5. C), there exists

an embedding φ of the ring of integers R_L of L into \mathcal{O} iff any prime ideal \mathfrak{p} of F that ramifies in B either remains inert or ramifies in L . Here, the fact that B is division and splits at least at one archimedean place makes the condition for the embeddability of R_L in \mathcal{O} particularly neat. Otherwise, it depends heavily on the conjugation class of \mathcal{O} .

Given a totally real Eichler pair (S, φ) , we can then consider the Hilbert modular variety \mathcal{H}_S that classifies isomorphism classes of triplets (A, i, \mathcal{L}) where

- A is an abelian variety of dimension $[L : \mathbb{Q}] = 2n$,
- $i : S \hookrightarrow \text{End}(A)$ is a ring homomorphism, and
- \mathcal{L} is a polarization of type $(1, d_2, \dots, d_g)$ on A .

The scheme \mathcal{H}_S is $2n$ -dimensional, non-complete and defined over \mathbb{Q} . Actually, this is a particular case of the Shimura varieties mentioned above: \mathcal{H}_S is the union of several irreducible components, all them isomorphic to Shimura varieties $X_{(\text{M}_2(S), \mathcal{I}, \varrho)}$ for several $\text{M}_2(S)$ -left ideals \mathcal{I} and involutions ϱ . Imitating the construction of π we then obtain, for any totally real Eichler pair (S, φ) , a morphism

$$\pi_{(S, \varphi)} : \begin{array}{ccc} X_{(\mathcal{O}, \mathcal{I}, \varrho)} & \longrightarrow & \mathcal{H}_S \\ (A, \iota, \mathcal{L}) & \mapsto & (A, \iota \cdot \varphi, \mathcal{L}) \end{array}$$

where $\iota \cdot \varphi : S \hookrightarrow \mathcal{O} \hookrightarrow \text{End}(A)$. Finally, and in a similar way, we also have a morphism

$$\pi_F : X_{(\mathcal{O}, \mathcal{I}, \varrho)} \longrightarrow \mathcal{H}_F$$

from X_B into a Hilbert variety \mathcal{H}_F . However, this time \mathcal{H}_F does not correspond to any of the Shimura varieties X_B introduced above for any quaternion algebra B (not even $\text{M}_2(F)$). On this occasion, by \mathcal{H}_F we mean the (reduced) scheme over \mathbb{Q} that coarsely represents the functor attached to the moduli problem of classifying $(1, d_2, \dots, d_g)$ -polarized abelian varieties A of dimension $2n$ together with a homomorphism $R_F \hookrightarrow \text{End}(A)$. The variety \mathcal{H}_F has dimension $3n$, and $\mathcal{H}_F(\mathbb{C})$ is the quotient of n copies of the 2-Siegel half space \mathfrak{H}_2 by a suitable discontinuous group (cf. [Sh1], [LaBi], Chapter 9). Notice that, when $F = \mathbb{Q}$, $\mathcal{H}_F = \mathcal{A}_{2, (1, d)}$ is Igusa's three-fold of level $d \geq 1$, the moduli space of $(1, d)$ -polarized abelian surfaces.

Then, the map $\pi_F : X_{(\mathcal{O}, \mathcal{I}, \varrho)} \longrightarrow \mathcal{H}_F$ is constructed as above: by restricting the endomorphism structure of a triplet (A, ι, \mathcal{L}) to $\iota|_{R_F} : R_F \hookrightarrow \text{End}(A)$. The whole picture can be summarized with the following commutative diagram of morphisms between Shimura varieties:

$$\pi : X_B \begin{array}{ccc} \nearrow^{\pi_{(S_1, \varphi_1)}} & \mathcal{H}_{S_1} & \searrow \\ & \vdots & \\ \xrightarrow{\pi_{(S_2, \varphi_2)}} & \mathcal{H}_{S_2} & \longrightarrow \mathcal{H}_F \longrightarrow \mathcal{A}_{g, (1, d_2, \dots, d_g)} \\ & \searrow_{\pi_{(S_n, \varphi_n)}} & \nearrow \\ & \mathcal{H}_{S_n} & \end{array}$$

Note that, while there is a canonical forgetful map from X_B to \mathcal{H}_F , we obtain distinct maps from X_B to \mathcal{H}_S as $\varphi : S \hookrightarrow \mathcal{O}$ varies among all possible Eichler embeddings.

3. THE ATKIN-LEHNER GROUP OF A SHIMURA VARIETY

As before, let F be a totally real number field of degree n . We denote by $\text{Pic}(F)$ the class group of fractional ideals up to principal ideals of F . Similarly, $\text{Pic}_+(F)$ will stand for the narrow class group of fractional ideals up to totally positive principal ideals aR_F , $a \in F_+^*$, of F .

Let B be a totally indefinite quaternion algebra over F of discriminant $\mathcal{D} = \mathfrak{p}_1 \cdot \dots \cdot \mathfrak{p}_{2r}$ and let \mathcal{O} be a maximal order in B . Define the groups $\mathcal{O}^* \supset \mathcal{O}_+^* \supseteq \mathcal{O}^1$ of units in \mathcal{O} , units in \mathcal{O} of totally positive reduced norm and units in \mathcal{O} of reduced norm 1 respectively. By the Hasse-Schilling-Maass Theorem (cf. [HaSc], [Vi], p. 90) we have the exact sequence

$$1 \rightarrow \mathcal{O}^1 \rightarrow \mathcal{O}_+^* \xrightarrow{n} R_{F_+}^* \rightarrow 1.$$

Definition 3.1. Let $N_{B^*}(\mathcal{O}) = \{\gamma \in B^*, \gamma^{-1}\mathcal{O}\gamma \subseteq \mathcal{O}\}$ be the normalizer of \mathcal{O} in B^* . The *Atkin-Lehner group* W of \mathcal{O} is

$$W = N_{B^*}(\mathcal{O})/F^*\mathcal{O}^*.$$

Let B_+^* denote the group of invertible quaternions of B of totally positive reduced norm. We define the *positive Atkin-Lehner groups*

$$W_+ = N_{B_+^*}(\mathcal{O})/F^*\mathcal{O}_+^*$$

and

$$W^1 = N_{B^*}(\mathcal{O})/F^*\mathcal{O}^1.$$

Let us note that, by the Skolem-Noether Theorem, $N_{B^*}(\mathcal{O})/F^*$ coincides with the group of automorphisms of \mathcal{O} . The Atkin-Lehner group W is identified with the group of principal two-sided ideals of \mathcal{O} by the assignation $\omega \in W \mapsto \mathcal{O} \cdot \omega$ and it is a finite abelian 2-group. More precisely, by Eichler’s results ([Ei1], [Ei2], [Vi], Theorem 5.7), the reduced norm $n : B^* \rightarrow F^*$ induces an isomorphism $W \simeq \mathbb{Z}/2\mathbb{Z} \times \dots \times \mathbb{Z}/2\mathbb{Z}$. Moreover, any involution $[\omega]$ on W can be represented by an element $\omega \in \mathcal{O}$ whose reduced norm $n(\omega)$ is supported at the prime ideals $\mathfrak{p} \mid \mathcal{D}$. The group W_+ may and will be regarded as the subgroup $W_+ = \{[\omega] \in W : n(\omega) \in F_+^*\}$ of W , and both coincide whenever $\text{Pic}_+(F) \simeq \text{Pic}(F)$. From the above, we obtain the exact sequence

$$1 \rightarrow R_{F_+}^*/R_F^{*2} \rightarrow W^1 \rightarrow W_+ \rightarrow 1.$$

Here, a totally positive unit $u \in R_{F_+}^*$ is mapped to any $\alpha_u \in \mathcal{O}_+^*$ with $n(\alpha_u) = u$, whose existence is guaranteed by the Hasse-Schilling-Maass Theorem in its integral version (cf. [HaSc], [Vi], p. 90). In this way, we obtain that W^1 is isomorphic to the *direct* product

$$W^1 \simeq R_{F_+}^*/R_F^{*2} \times W_+ \simeq (\mathbb{Z}/2\mathbb{Z})^s$$

for some positive integer $s \leq (n - 1) + 2r$. The bound for s follows from Dirichlet’s unit theorem and the inclusion $W_+ \subseteq W$. Its precise value can actually be determined in terms of the behaviour of signatures of elements of R_F^* .

3.1. Modular interpretation of the Atkin-Lehner group. In his Ph.D. thesis, Jordan described the modular interpretation of the action of the positive Atkin-Lehner group W_+ on Shimura curves ([Jo]). This can be extended to the higher dimensional cases, and we do so now. The action of $B_+^* \subset \text{GL}_2^+(\mathbb{R})^n$ on \mathfrak{H}^n by Moebius transformations descends to a free action of W^1 on the set $\mathcal{O}^1 \backslash \mathfrak{H}^n$ of

complex points of the Shimura variety $X_B = X_{(\mathcal{O}, \mathcal{I}, \varrho)}$, for any choice of a left \mathcal{O} -ideal \mathcal{I} and a positive involution ϱ .

The action of an involution $\omega \in W^1$ can be modularly interpreted as follows. Let $P = [(A, \iota, \mathcal{L})]$ denote the isomorphism class of a polarized abelian variety with multiplication by \mathcal{O} viewed as a closed point of $X_{(\mathcal{O}, \mathcal{I}, \varrho)}$. Then ω acts on the triplet by *keeping* the same isomorphism class of the underlying abelian variety A but conjugating the endomorphism structure $\iota : \mathcal{O} \hookrightarrow \text{End}(A)$ and switching the polarization \mathcal{L} . Namely, $\omega(P) = [(A, \iota_\omega, \mathcal{L}_\omega)]$ where

$$\begin{aligned} \iota_\omega : \mathcal{O} &\hookrightarrow \text{End}(A) \\ \beta &\mapsto \omega^{-1} \iota(\beta) \omega \end{aligned}$$

and $\mathcal{L}_\omega := \frac{\omega^*(\mathcal{L})}{\mathfrak{n}(\omega)}$ is $\frac{1}{\mathfrak{n}(\omega)}$ -times the pull-back of the primitive polarization \mathcal{L} by the isogeny $\omega \in \mathcal{O} \xrightarrow{\iota} \text{End}(A)$. In other words, if we regard the first Chern class $E = c_1(\mathcal{L})$ of \mathcal{L} as an alternate bilinear form on $V = \text{Lie}(A(\mathbb{C}))$, then $c_1(\mathcal{L}_\omega) : V \times V \rightarrow \mathbb{R}$, $(u, v) \mapsto E((\omega/\mathfrak{n}(\omega))(u), \omega(v))$. See Section 5 and [Ro2] to check that this does correspond to the first Chern class of a primitive polarization A compatible with ι_ω . From this interpretation and by standard moduli considerations, it follows that $W^1 \subseteq \text{Aut}(X_B)$ acts on X_B as a subgroup of algebraic involuting automorphisms over \mathbb{Q} .

Remark 3.2. If $\omega \in \mathcal{O}_+^*$, $\mathfrak{n}(\omega) = u \in R_{F_+}^*$, then ω can be viewed as an automorphism $\omega \in \text{Aut}(A) \simeq \mathcal{O}^*$ of A . This automorphism induces an isomorphism of triplets $\omega(A, \iota, \mathcal{L}) = (A, \omega^{-1} \iota \omega, \omega^*(\mathcal{L})/u) \xrightarrow{\omega} (A, \iota, \mathcal{L}_u)$ where \mathcal{L}_u is a principal polarization on A such that, although $\mathcal{L} \not\cong \mathcal{L}_u$, both induce the same Rosati involution on $\text{End}(A) \otimes \mathbb{Q}$.

In the literature (cf. [Sh1], [Mi]), \mathcal{L} and \mathcal{L}_u are called weakly isomorphic. It is easy to see that $\mathcal{O}_+^*/\mathcal{O}^1$ acts freely and transitively on the set of isomorphism classes of a given weak polarization class on A .

Remark 3.3. There is a natural moduli theory for polarized abelian varieties with QM up to weak isomorphism which is also considered in [Sh1]. Both theories coincide in dimension 2 (because it corresponds to $F = \mathbb{Q}$), but for higher dimensions the latter is coarser and less suitable for our purposes.

Remark 3.4. We wonder in what circumstances W^1 is the full group of automorphisms of X_B . The impression is that, *generically*, it does hold that $\text{Aut}(X_B) = W^1$, but of course the term *generic* should be made precise in any case. The split case of an order \mathcal{O} in $B = M_2(\mathbb{Q})$ was classically studied by Ogg, who found that the modular curve $X_0(37)$ is an interesting exception to the prediction that, whenever the genus of a modular curve is at least 2, $W = W^1 = \text{Aut}(X_{\mathcal{O}})$. For Shimura curves $X_{\mathcal{O}}$ with \mathcal{O} in a rational division quaternion algebra B/\mathbb{Q} , this question was investigated in [Ro1]. For higher dimensional Shimura varieties, it seems to hold that whenever \mathcal{O}^* does not contain torsion units (besides ± 1), $W^1 = \text{Aut}(X_{\mathcal{O}, \mathcal{I}, \varrho})$ (cf. [Ro1], Theorem 2).

4. MAIN THEOREM AND COROLLARIES

As remarked in the introduction, the type of the primitive polarizations of the abelian varieties parametrized by the Shimura variety $X_{(\mathcal{O}, \mathcal{I}, \varrho)}$ is determined by the datum $(\mathcal{O}, \mathcal{I}, \varrho)$. The following was proved in [Ro2] and makes this observation explicit.

Proposition 4.1. *The polarizations of the abelian varieties with quaternionic multiplication (A, ι, \mathcal{L}) parametrized by $X_{(\mathcal{O}, \mathcal{I}, \varrho)}$ are principal if and only if:*

- $\text{Disc}(B) = (D)$ is a principal ideal of F generated by a totally positive element $D \in F_+^*$.
- $\mathfrak{n}_{B/F}(\mathcal{I})$ and the codifferent $\mathfrak{v}_{F/\mathbb{Q}}^{-1} = \{x \in F, \text{tr}_{F/\mathbb{Q}}(xR_F) \subseteq \mathbb{Z}\}$ of F over \mathbb{Q} lie in the same ideal class in $\text{Pic}(F)$.
- The positive anti-involution on B is $\varrho = \varrho_\mu : B \rightarrow B, \beta \mapsto \mu^{-1}\bar{\beta}\mu$ for some $\mu \in \mathcal{O}$ such that $\mu^2 + D = 0$.

For the rest of the note, we will focus on moduli spaces of principally polarized abelian varieties, and therefore we place ourselves under the assumptions on $(\mathcal{O}, \mathcal{I}, \varrho_\mu)$ of Proposition 4.1. We thus assume in particular that $\text{disc}(B) = (D)$ for some $D \in F_+^*$ and that $\mu \in \mathcal{O}$ satisfies $\mu^2 + D = 0$. In this case we say that $(\mathcal{O}, \mathcal{I}, \varrho_\mu)$ is of *principal type* and we will simply denote by \mathcal{A}_g the moduli space $\mathcal{A}_{g,(1,\dots,1)}$.

Our main Theorem 4.4 below describes how the modular maps introduced in Section 2.1 factor through the quotient of X_B by certain subgroups of W^1 of Atkin-Lehner involutions that we introduce first.

Let us say that an element $\chi \in \mathcal{O} \cap N_{B^*}(\mathcal{O})$ is a twist of (\mathcal{O}, μ) if $\chi^2 + \mathfrak{n}(\chi) = 0, \mu\chi = -\chi\mu$ and therefore

$$B = F + F\mu + F\chi + F\mu\chi = \left(\frac{-D, -\mathfrak{n}(\chi)}{F}\right).$$

Further, we will say that a pair (\mathcal{O}, μ) is *twisting* if it admits some twist χ and that a quaternion algebra B is *twisting* if it contains a twisting pair (\mathcal{O}, μ) . Note that a totally indefinite quaternion algebra is twisting if and only if $B = \left(\frac{-D, m}{F}\right)$ for some $m \in F_+^*, m \mid D$.

Now let $\omega \in W^1$ be an Atkin-Lehner involution of \mathcal{O} . We say that it is a twisting involution with respect to (\mathcal{O}, μ) if the class of ω in $W = N_{B^*}(\mathcal{O})/F^*\mathcal{O}^*$ is represented by a twist $\chi \in \mathcal{O} \cap N_{B^*}(\mathcal{O})$ of (\mathcal{O}, μ) . For any subring $S \subset \mathcal{O}$ we say that a twist χ of (\mathcal{O}, μ) is in S if $\chi \in S$. An involution $\omega \in W^1$ is a twisting involution in S if it can be represented by a twist $\chi \in S$.

Definition 4.2. (1) The *twisting group* V_0 attached to (\mathcal{O}, μ) is the subgroup of W^1 generated by the twisting involutions of (\mathcal{O}, μ) .

(2) The *stable group* attached to (\mathcal{O}, μ) is the subgroup

$$W_0 = V_0 \cdot U_0$$

of W^1 generated by

$$U_0 = N_{F(\mu)^*}(\mathcal{O})/F^*\Omega(S),$$

where $\Omega(S) = \{\xi \in S, \xi^f = 1, f \geq 1\}$ denotes the finite group of roots of unity in the CM-order $S = F(\mu) \cap \mathcal{O}$ and the twisting group V_0 .

For any subring $S \subseteq \mathcal{O}$, we shall let $V_0(S)$ denote the subgroup of W^1 generated by the twisting involutions of (\mathcal{O}, μ) in S .

Let

$$\tilde{X}_B = \tilde{X}_{(\mathcal{O}, \mathcal{I}, \varrho)} = \pi(X_{(\mathcal{O}, \mathcal{I}, \varrho)}) \hookrightarrow \mathcal{A}_g$$

denote the image in \mathcal{A}_g by π of the Shimura variety $X_{(\mathcal{O}, \mathcal{I}, \varrho)}$. Similarly, define $\tilde{X}_{B/F} \subset \mathcal{H}_F$ and $\tilde{X}_{B/(S, \varphi)} \subset \mathcal{H}_S$ respectively to be the algebraic subvarieties

$\pi_F(X_{(\mathcal{O}, \mathcal{I}, \varrho)})$ and $\pi_{(S, \varphi)}(X_{(\mathcal{O}, \mathcal{I}, \varrho)})$ of the Hilbert modular varieties \mathcal{H}_F and \mathcal{H}_S . Note that the classical Humbert varieties arising from Hilbert modular varieties can be considered as a degenerate case of the above defined \tilde{X}_B , since they would correspond to the split algebra $B = M_2(F)$.

Definition 4.3. A closed point $[(A, \iota, \mathcal{L})]$ on X_B and its image on $\tilde{X}_{B/(S, \varphi)}$, $\tilde{X}_{B/F}$ or \tilde{X}_B will be called a *Heegner point* if $\text{End}(A) \otimes \mathbb{Q} \simeq M_2(L)$ for a CM-field L over F .

The set of Heegner points on these varieties is discrete and dense, $\text{End}(A) \otimes \mathbb{Q} = B$ being the generic case.

We say that a morphism between two schemes of not necessarily the same dimension is *quasifinite* if it has finite fibres (cf. [Ha], p. 91).

Theorem 4.4. *Let $(\mathcal{O}, \mathcal{I}, \varrho)$ be a quaternionic datum of principal type and let $X_B = X_{(\mathcal{O}, \mathcal{I}, \varrho)}$ be the corresponding Shimura variety. For any totally real Eichler pair (S, φ) , let*

$$\begin{array}{ccccc} & & \mathcal{H}_S & & \\ & \nearrow^{\pi_{(S, \varphi)}} & & \searrow & \\ \pi : X_B & & \xrightarrow{\pi_F} & \mathcal{H}_F & \rightarrow \mathcal{A}_g \end{array}$$

be the diagram of forgetful morphisms introduced before. Then

- The map $\pi_F : X_B \rightarrow \mathcal{H}_F$ is a quasifinite map that factors over \mathbb{Q} into the natural projection $X_B \rightarrow X_B/W_0$ from X_B onto its quotient by the stable group $W_0 \subseteq \text{Aut}(X_B)$ and a birational morphism $b_F : X_B/W_0 \dashrightarrow \tilde{X}_{B/F}$ onto the image of X_B in \mathcal{H}_F .

The maximal domain of definition of b_F^{-1} is $\tilde{X}_{B/F} \setminus \mathcal{T}_F$, where \mathcal{T}_F is a finite set of Heegner points.

- The map $\pi_{(S, \varphi)} : X_B \rightarrow \mathcal{H}_S$ is a quasifinite map that factors over \mathbb{Q} into the projection $X_B \rightarrow X/V_0(\varphi(S))$ of X_B onto its quotient by the finite 2-group $V_0(\varphi(S)) \subseteq W_0$ and a birational morphism $b_{(S, \varphi)} : X_B/V_0(\varphi(S)) \dashrightarrow \tilde{X}_{B/(S, \varphi)}$ into the image of X_B in \mathcal{H}_S by $\pi_{(S, \varphi)}$.

As before, $b_{(S, \varphi)}^{-1}$ is defined on the whole $\tilde{X}_{B/(S, \varphi)}$ but at a finite set $\mathcal{T}_{(S, \varphi)}$ of Heegner points.

We call the sets \mathcal{T}_F and $\mathcal{T}_{(S, \varphi)}$ the *singular Heegner loci* of $\tilde{X}_{B/F}$ and $\tilde{X}_{B/(S, \varphi)}$ respectively, as they are indeed sets of singular points (of quotient type) on these varieties. We present the proof of the theorem in Sections 5 and 6; we now derive some corollaries of it.

Being subgroups of W^1 , both the twisting subgroups $V_0(\varphi(S))$ and the stable subgroup W_0 are isomorphic to the direct product of a certain number of copies of C_2 . We refer the reader to [Ro3] for a detailed description of them. As a consequence of the results proved in [Ro3] and our main Theorem 4.4, we obtain the following.

Corollary 4.5. *Let (\mathcal{O}, μ) be a non-twisting polarized order. Then, the maps $\pi_{(S, \varphi)} : X_B \rightarrow \mathcal{H}_S$ are birational equivalences for any totally real quadratic order embedded in \mathcal{O} and*

$$\text{deg}(\pi_F : X_B \rightarrow \mathcal{H}_F) = 2^{\omega_{\text{odd}}},$$

where $\omega_{\text{odd}} = \#\{\xi \in R_\mu, \xi^f = 1, f \text{ odd}\}$.

Let (\mathcal{O}, μ) be a twisting polarized order and let $\chi \in \mathcal{O}$ be any twist of (\mathcal{O}, μ) . Then, $\pi_{R_F[\chi]} : X_B \rightarrow \mathcal{H}_{R_F[\chi]}$ is a 2 : 1 map, whereas for any totally real Eichler pair (S, φ) such that $\varphi(S)$ does not contain any twist, $\pi_{(S, \varphi)}$ is birational.

Moreover,

$$\deg(\pi_F : X_B \rightarrow \mathcal{H}_F) = 2^{2\omega_{\text{odd}}}.$$

In particular, note that the forgetful map $\pi_{\mathbb{Q}} : X_B \rightarrow \mathcal{A}_2$ from a Shimura curve into Igusa’s moduli space of principally polarized abelian surfaces is either of degree 2 or of degree 4 and that a necessary condition for the latter is that $B \simeq (\frac{-\text{disc}(B), m}{\mathbb{Q}})$ for some $m > 0, m \mid \text{disc}(B)$. It turns out, for instance, that for any choice of a datum $(\mathcal{O}, \mathcal{I}, \varrho)$ of principal type of discriminant $D = 6$ or 10 , the corresponding forgetful map of the Shimura curve $X_{(\mathcal{O}, \mathcal{I}, \varrho)}$ into \mathcal{A}_2 has degree 4.

5. THE NÉRON-SEVERI GROUP OF AN ABELIAN VARIETY WITH QUATERNIONIC MULTIPLICATION

In this section we introduce some ingredients that may be useful to make the forgetful maps and the singular Heegner loci \mathcal{T}_F and $\mathcal{T}_{(S, \varphi)}$ of Theorem 4.4 explicit in particular cases, although we will not pursue this purpose in this note. The content of this section will also be used in the proof of Theorem 4.4 in Section 6.

For an abelian variety A , the Néron-Severi group $\text{NS}(A)$ is the group $\text{Pic}(A)/\text{Pic}^0(A)$ of invertible sheaves on A up to algebraic equivalence. By Néron’s Basis Theorem, $\text{NS}(A)$ is a \mathbb{Z} -module of finite rank, and we denote $\text{NS}^0(A) = \text{NS}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$.

As always, let \mathcal{O} denote a maximal order in a totally indefinite quaternion algebra B over a totally real number field $F, [F : \mathbb{Q}] = n$. Assume that A is an abelian variety over \mathbb{C} of dimension $2n$ and that $\text{End}(A)$ contains an order isomorphic to \mathcal{O} . Then A is either simple or isogenous to the square A_0^2 of a simple abelian variety A_0 of dimension n . In the former case we actually have that $\text{End}(A) \simeq \mathcal{O}$, while in the latter $\text{End}(A)$ is an order in a CM-field L over F and $\text{End}(A)$ can be identified with an order in $M_2(L)$.

More precisely, and in a unified way, fix a ring homomorphism $\iota : \mathcal{O} \hookrightarrow \text{End}(A)$ and let $R = \text{End}_{\mathcal{O}}(A) = \{\gamma \in \text{End}(A) : \gamma\beta = \beta\gamma \text{ for all } \beta \in \iota(\mathcal{O})\}$ be the commutator of \mathcal{O} in $\text{End}(A)$. Then either $R = R_F$ or R is an order in L , but in any case $\text{End}(A) = \mathcal{O} \otimes_{R_F} R$.

We now wish to select a suitable piece of the Néron-Severi group of A and describe it also in a unified way, regardless of whether A is simple or isogenous to the square of a CM abelian variety. Recall that a line bundle $\mathcal{L} \in \text{NS}^0(A)$ induces an involution $\circ_{\mathcal{L}}$ on $\text{End}^0(A)$, called the Rosati involution: $\circ_{\mathcal{L}} : \text{End}^0(A) \rightarrow \text{End}^0(A), \beta \mapsto \beta^{\circ} = \varphi_{\mathcal{L}}^{-1} \hat{\beta} \varphi_{\mathcal{L}}$. Here, if \hat{A} is the dual abelian variety of A , we let $\hat{\beta}$ denote the dual endomorphism $\hat{\beta} : \hat{A} \rightarrow \hat{A}$ of β on \hat{A} and we let $\varphi_{\mathcal{L}} : A \rightarrow \hat{A}, x \mapsto \tau_x^*(\mathcal{L}) \otimes \mathcal{L}^{-1}$ be the morphism that \mathcal{L} induces from A to \hat{A} . By τ_x we mean the translation-by- x map on A .

Definition 5.1. Let (A, ι) be an abelian variety of dimension $2n$ endowed with a multiplication by \mathcal{O} . The B -part of the (rational) Néron-Severi group of A is

$$\text{NS}_B^0(A) = \{\mathcal{L} \in \text{NS}^0(A), \circ_{\mathcal{L}}(B) = B\}.$$

The B -part of the Néron-Severi group of A is the group $\text{NS}_B(A) = \text{NS}_B^0(A) \cap \text{NS}(A)$ of line bundles $\mathcal{L} \in \text{NS}(A)$ on A whose Rosati involution $\circ_{\mathcal{L}}$ leaves B invariant.

If A is simple (and thus $\iota : \text{End}(A) \simeq \mathcal{O}$ is an isomorphism), we do not obtain anything new: $NS_B(A) = NS(A)$ and $NS_B^0(A) = NS^0(A)$. However, when A is not simple (and thus $\text{End}^0(A) \simeq M_2(L)$ for some CM-field L), these are proper subgroups of the usual Néron-Severi groups. The following extends a result proved by the author in [Ro2] and allows us to see line bundles on A as pure quaternions in B in a canonical way. See Section 1 for notations.

Proposition 5.2. *Let (A, ι) be a complex abelian variety of dimension $2n$ together with a multiplication by \mathcal{O} . Let \mathcal{I} be the isomorphism class of the left \mathcal{O} -module $H_1(A, \mathbb{Z})$, represented by a left \mathcal{O} -ideal $\mathcal{I} \subset B$. Then the first Chern class induces a natural isomorphism of groups*

$$NS_B(A) \xrightarrow{\cong} \mathcal{N}(\mathcal{I})_0^\sharp \hookrightarrow B_0.$$

Explicitly, the isomorphism is constructed as follows: write $A(\mathbb{C}) = V/\Lambda$ for a vector space V and a lattice $\Lambda = \mathcal{I}v_0$, $v_0 \in V$, and let $\mathcal{L} \in NS_B(A)$. Let also $E = c_1(\mathcal{L}) : V \times V \rightarrow \mathbb{R}$ be its first Chern class regarded as an \mathbb{R} -alternate Riemann form on V . Since $E(\Lambda \times \Lambda) \subseteq \mathbb{Z}$, we may consider the linear map $B \rightarrow \mathbb{Q}$ that sends an element $\beta \in B$ to $E(\iota(\beta)v_0, v_0)$. By the non-degeneracy of the reduced trace $\text{tr}_{B/\mathbb{Q}} = \text{tr}_{F/\mathbb{Q}} \cdot \text{tr}_{B/F}$, we can find an element $\mu \in B$ such that $E(\iota(\beta)v_0, v_0) = \text{tr}_{B/\mathbb{Q}}(\mu\beta)$ for any $\beta \in B$. It follows from the alternateness of E that μ is pure: $\mu \in B_0$ (cf. [Ro2] for details).

Since $\mathcal{L} \in NS_B(A)$, the Rosati involution $\circ_{\mathcal{L}}$ on $\text{End}^0(A)$ descends to a well-defined involution $\circ : B \rightarrow B$ such that, for any $u, v \in V$, $E(u, \iota(\beta)v) = E(\iota(\beta^\circ)u, v)$. By means of the Skolem-Noether Theorem it can be seen that $\beta^\circ = \mu^{-1}\bar{\beta}\mu$ (cf. [Ro2] again for details). This already determines the Riemann form $E = c_1(\mathcal{L})$:

$$\begin{aligned} E : V \times V &\longrightarrow \mathbb{R} \\ (u, v) &\longmapsto \text{tr}_{B \otimes_{\mathbb{Q}} \mathbb{R} / \mathbb{R}}(\mu\gamma\bar{\beta}) \end{aligned}$$

where β and $\gamma \in B \otimes_{\mathbb{Q}} \mathbb{R} \simeq M_2(\mathbb{R})^n$ are uniquely determined elements such that $\beta v_0 = u$ and $\gamma v_0 = v$. Observe that, since $\Lambda = \mathcal{I}v_0$ and $E(\Lambda \times \Lambda) \subseteq \mathbb{Z}$, $\mu \in \mathcal{N}(\mathcal{I})_0^\sharp$.

In what follows, the above Riemann form will be denoted as E_μ . By an abuse of notation, we will also sometimes use the notation $c_1(\mathcal{L}) = \mu$. It was shown in [Ro2], Theorem 2.2, that the assignation $\mathcal{L} \mapsto c_1(\mathcal{L})$ is an isomorphism of groups between $NS_B(A)$ and $\mathcal{N}(\mathcal{I})_0^\sharp$ in the case that A is simple. As the reader may check, the proof is exactly the same as in the CM split case.

Since the rank of $NS_B(A)$ over \mathbb{Z} is $3[F : \mathbb{Q}] = 3n$, any \mathbb{Z} -module of the same rank is actually isomorphic to it. The interesting part of the above description of the B -part of the Néron-Severi group $NS_B(A)$ is that the isomorphism of Proposition 5.2 behaves nicely with respect to the degree of the line bundles $\mathcal{L} \in NS_B(A)$ and also with respect to the pull-back $\alpha^*(\mathcal{L})$ by endomorphisms $\alpha \in \text{End}(A)$ of A . Indeed, let us define the degree $\text{deg}(\mathcal{L})$ of a non-degenerate line bundle $\mathcal{L} \in NS(A)$ to be the degree of the finite map $\varphi_{\mathcal{L}} : A \rightarrow \hat{A}$ from A to the dual abelian variety \hat{A} induced by \mathcal{L} .

Proposition 5.3. *If the first Chern class of \mathcal{L} is represented by the Riemann form E_μ with $\mu \in B_0$ and $\mu^2 + \delta = 0$, $\delta \in F^*$, then*

$$\text{deg}(\varphi_{\mathcal{L}}) = N_{F/\mathbb{Q}}(\vartheta_{F/\mathbb{Q}}^2 \cdot n_{B/F}(\mathcal{I})^2 \cdot \mathcal{D} \cdot \delta)^2.$$

The proof of Proposition 5.3 was given in [Ro2], Proposition 3.1, for non-CM abelian varieties, but the same arguments are valid for this case as well. In order

to describe now how the first Chern class behaves under pull-backs of line bundles by endomorphisms, we denote by $\lambda \mapsto \bar{\lambda} \in L$ the complex conjugation on L . There should be no confusion with the quaternionic conjugation $\beta \mapsto \bar{\beta}$ on B . Both conjugation maps combine to endow $B \otimes_F L \simeq M_2(L)$ with a conjugation map $\lambda\beta \mapsto \bar{\lambda}\bar{\beta}$ (and extended linearly). In matrices, $\overline{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} = \begin{pmatrix} \bar{d} & -\bar{b} \\ -\bar{c} & \bar{a} \end{pmatrix} \in M_2(L)$.

Having fixed a homomorphism of algebras $\iota : B \hookrightarrow \text{End}^0(A)$, identify $\text{End}^0(A)$ with either B or $M_2(L)$ with B as an embedded sub-algebra.

Proposition 5.4. *Let (A, ι) be an abelian variety with quaternionic multiplication by a maximal order in B and let $\mathcal{L} \in \text{NS}_B^0(A)$ be a line bundle on A compatible with B . Denote by $c_1(\mathcal{L}) = \mu \in B_0$ its first Chern class regarded as a pure quaternion in B .*

Then the pull-back polarization $\alpha^(\mathcal{L})$ of \mathcal{L} by a rational endomorphism $\alpha \in \text{End}^0(A)$ is compatible with B , that is, $\alpha^*(\mathcal{L}) \in \text{NS}_B^0(A)$, if and only if $\bar{\alpha}\mu\alpha \in B_0$. In this case, $c_1(\alpha^*(\mathcal{L})) = \bar{\alpha}\mu\alpha$.*

Obviously, when A is simple the above condition is empty and all polarizations are compatible with B . Moreover, in the simple case the relation $c_1(\alpha^*(\mathcal{L})) = \bar{\alpha}c_1(\mathcal{L})\alpha$ was already proved in [Ro2], Theorem 2.2. However, the proof of this fact in the CM case is not trivial and deserves a more careful inspection.

We hence assume that $\text{End}^0(A) = M_2(L)$ for a CM-field L over F and we first claim that $\alpha^*(\mathcal{L})$ induces the Rosati involution $\beta \mapsto (\bar{\alpha}\mu\alpha)^{-1}\bar{\beta}(\bar{\alpha}\mu\alpha)$ on $M_2(L)$. Denote by $E_{\mathcal{L}}$ and $E_{\alpha^*(\mathcal{L})}$ the Riemann forms on $V = \text{Lie}(A)$ induced by \mathcal{L} and $\alpha^*(\mathcal{L})$ respectively. The Rosati involution $\circ_{\alpha^*(\mathcal{L})}$ restricts to complex conjugation on L and is characterized by the fact that $E_{\alpha^*(\mathcal{L})}(u, \beta v) = E_{\alpha^*(\mathcal{L})}(\beta^{\circ_{\alpha^*(\mathcal{L})}}u, v)$ for any $u, v \in V$ and $\beta \in \text{End}^0(A)$.

The above given anti-involution is indeed complex conjugation on L , and it holds that

$$E_{\alpha^*(\mathcal{L})}(u, \beta v) = E_{\mathcal{L}}(\alpha u, \alpha \beta v) = E_{\mathcal{L}}(\mu^{-1}\bar{\beta}\bar{\alpha}\mu\alpha u, v) = E_{\alpha^*(\mathcal{L})}((\bar{\alpha}\mu\alpha)^{-1}\bar{\beta}(\bar{\alpha}\mu\alpha)u, v)$$

for any $u, v \in V$ and $\beta \in M_2(L)$.

Now that we know this, $\alpha^*(\mathcal{L}) \in \text{NS}_B^0(A)$ if and only if $\bar{\alpha}\mu\alpha \in M_2(L)$ normalizes B . But the Skolem-Noether Theorem implies that an element in $M_2(L)$ normalizes B if and only if it belongs to the multiplicative group $L^* \cdot B^*$.

In order to see that actually $\bar{\alpha}\mu\alpha \in B$ (and thus also in B_0), write $L = F(\eta)$, $\eta^2 = -\Delta \in F_+^*$, and express $\alpha = \alpha_0 + \alpha_1\eta \in B \otimes_F L \simeq M_2(L)$ with $\alpha_i \in B$. Then the reader may check that $\bar{\alpha}\mu\alpha = (\bar{\alpha}_0\mu\alpha_0 + \Delta\bar{\alpha}_1\mu\alpha_1) + \text{tr}_{B/F}(\bar{\alpha}_0\mu\alpha_1)\eta$. Since $\bar{\alpha}\mu\alpha$ must belong to $L^* \cdot B^*$, then either $\bar{\alpha}\mu\alpha = (\bar{\alpha}_0\mu\alpha_0 + \Delta\bar{\alpha}_1\mu\alpha_1)$ or $\bar{\alpha}\mu\alpha = \text{tr}_{B/F}(\bar{\alpha}_0\mu\alpha_1)\eta$.

We claim that the latter cannot occur and therefore $\bar{\alpha}\mu\alpha \in B_0$, as we wish. For if $\bar{\alpha}_0\mu\alpha_0 = -\Delta\bar{\alpha}_1\mu\alpha_1$, this would imply that

- $n_{B/F}(\alpha_0) = \Delta n_{B/F}(\alpha_1)$ and
- $\mu\alpha_0\alpha_1^{-1} = -\Delta\bar{\alpha}_0^{-1}\alpha_1\mu = -\alpha_0\alpha_1^{-1}\mu$.

It would then follow that $\text{tr}_{B/F}(\alpha_0\alpha_1^{-1}) = 0$ and hence $B = (\frac{-\Delta, -\Delta}{F})$, which is a contradiction because B is totally indefinite while $(\frac{-\sigma(\Delta), -\sigma(\Delta)}{\mathbb{R}})$ is the skew field of Hamilton quaternions, for any real embedding $\sigma : F \hookrightarrow \mathbb{R}$.

This proves the first part of the proposition. Assume now that $\alpha^*(\mathcal{L}) \in \text{NS}_B^0(A)$ and thus $\bar{\alpha}\mu\alpha = \bar{\alpha}_0\mu\alpha_0 + \Delta\bar{\alpha}_1\mu\alpha_1 \in B_0$. In order to show that indeed $c_1(\alpha^*(\mathcal{L})) = \bar{\alpha}\mu\alpha$, recall that for a polarization $\mathcal{L} \in \text{NS}_B^0(A)$, the Riemann form associated to it

is $E_\mu : V \times V \rightarrow \mathbb{R}, (u, v) \mapsto \text{tr}_{B/\mathbb{Q}}(\mu\gamma\bar{\beta})$ where $\mu = c_1(\mathcal{L})$ and $\beta, \gamma \in B \otimes_{\mathbb{Q}} \mathbb{R}$ are the only elements such that $\beta v_0 = u$ and $\gamma v_0 = v$ respectively.

We must show that $E_{\bar{\alpha}\mu\alpha}(u, v) = E_\mu(\alpha u, \alpha v)$ for any $u, v \in V$. This follows from direct computation provided one takes into account that, if we let $\zeta \in B \otimes_{\mathbb{Q}} \mathbb{R}$ denote the only element such that $\zeta v_0 = \eta v_0$, then for any $\beta \in B \otimes_{\mathbb{Q}} \mathbb{R}$ we have that $\alpha\beta v_0 = (\alpha_0\beta + \alpha_1\beta\zeta)v_0$.

6. PROOF OF THEOREM 4.4

Let $(\mathcal{O}, \mathcal{I}, \varrho)$ be a quaternionic datum attached to a totally indefinite quaternion algebra B over a totally real number field F . We assume that it is of principal type (cf. Proposition 4.1 and below). This means in particular that $\text{disc}(B) = (D)$ can be generated by an element $D \in F_+^*$ such that $\varrho = \varrho_\mu$ for some $\mu \in \mathcal{O}, \mu^2 + D = 0$.

Let $X_B = X_{(\mathcal{O}, \mathcal{I}, \varrho)}$ be the Shimura variety attached to it. By Section 3, the elements in the totally positive Atkin-Lehner group W^1 act as a group of rational automorphisms on X_B . Let us first show that $\pi_F : X_B \rightarrow \mathcal{H}_F$ is a quasifinite map that factors into the natural projection

$$X_B \rightarrow X_B/W_0$$

from X_B onto its quotient by the stable group $W_0 = U_0 \cdot V_0 \subseteq \text{Aut}_{\mathbb{Q}}(X_B)$ attached to (\mathcal{O}, ϱ) .

To this end, we claim that there is an action of the stable group W_0 on the geometric fibres of the morphism $\pi_F : X_B \rightarrow \mathcal{H}_F$ that is free and transitive on a certain open Zariski-dense subset \mathcal{U}_F of X_B .

In order to prove the claim, let (A, j_F, \mathcal{L}) be a principally polarized abelian variety together with a homomorphism of \mathbb{Z} -algebras $j_F : R_F \hookrightarrow \text{End}(A)$. Without making any further mention of it, we will regard $\text{End}(A)$ as an R_F -algebra through the given immersion j_F .

The isomorphism class $[(A, j_F, \mathcal{L})]$ may be interpreted as a closed point in \mathcal{H}_F . If it is non-empty, the elements in X_B of the fibre of π_F at this point can then be interpreted as those principally polarized abelian varieties with quaternionic multiplication whose isomorphism class is represented by a triplet (A, ι, \mathcal{L}) where $\iota : \mathcal{O} \hookrightarrow \text{End}(A)$ is a homomorphism of R_F -algebras such that the Rosati involution that \mathcal{L} induces on \mathcal{O} , via ι , coincides with ϱ_μ .

Choose a triplet (A, ι, \mathcal{L}) as above. Proposition 5.2 identifies $NS(A)$ with a lattice in B_0 by means of a map that we have called c_1 . We warn that this identification depends on ι , and it allows us to regard polarizations on the abelian variety A as pure quaternions of B . In particular, since $\circ|_{\iota(\mathcal{O})} = \varrho_\mu \cdot \iota$, we have that $c_1(\mathcal{L}) = \mu$ up to multiplication by elements in F^* .

Recall now that $W_0 = U_0 \cdot V_0$ for certain subgroups $U_0, V_0 \subseteq W^1$ that were defined in Section 3. Let $\omega \in U_0$. Let $L = F(\mu) \simeq F(\sqrt{-D})$ be the CM-field generated by $\mu \in B$ over F and let $S \supseteq R_F[\mu]$ be the order in L at which μ is optimal. Then ω is represented by an element (that we still denote) $\omega \in S$, and we wish to show that the closed points $[(A, \iota, \mathcal{L})]$ and $\omega[(A, \iota, \mathcal{L})]$ in X_B lie on the same fibre of π_F .

We recall (cf. Section 3.1) that the isomorphism class of $\omega[(A, \iota, \mathcal{L})]$ is represented by a triplet $(A, \iota_\omega, \mathcal{L}_\omega)$ where $\iota_\omega = \omega^{-1}\iota\omega : \mathcal{O} \hookrightarrow \text{End}(A)$ and $\mathcal{L}_\omega = \frac{\omega^*(\mathcal{L})}{n(\omega)}$ is a principal polarization on A . By Proposition 5.4, it is such that $c_1(\mathcal{L}_\omega) = \frac{\bar{\omega}}{n(\omega)}c_1(\mathcal{L})\omega = \omega^{-1}c_1(\mathcal{L})\omega = c_1(\mathcal{L})$. Here, the last equality holds because $c_1(\mathcal{L}) = \mu \in B^*/Fj$ and

μ and ω belong to the same quadratic order S embedded in \mathcal{O} and hence commute. Therefore $c_1(\mathcal{L}) = c_1(\mathcal{L}_\omega)$, and we conclude that (A, \mathcal{L}) and (A, \mathcal{L}_ω) are isomorphic polarized varieties. Moreover, ι and ι_ω coincide when restricted to the centre R_F of \mathcal{O} , and we obtain that $(A, \iota|_{R_F}, \mathcal{L}) \simeq (A, \iota_\omega|_{R_F}, \mathcal{L}_\omega)$, which is what we needed to show.

Now let $\omega \in V_0$ be an element represented by $\omega \in \mathcal{O}_+$ with $n(\omega) = m \in F_+^*$ and such that $B = F + F\mu + F\chi + F\mu\chi$ for some $\chi \in \mathcal{O}$, $\chi^2 = m$, $\mu\chi = -\chi\mu$. Let $[(A, \iota, \mathcal{L})] \in X_B$ be a closed point over $[(A, j_F, \mathcal{L})] \in \mathcal{H}_F$. We have that $\omega[(A, \iota, \mathcal{L})] = [(A, \iota_\omega, \mathcal{L}_\omega)]$ and we must show that $(A, \iota|_{R_F}, \mathcal{L}) \simeq (A, \iota_\omega|_{R_F}, \mathcal{L}_\omega)$. Again, since $\iota|_{R_F} = \iota_\omega|_{R_F}$, we only need to see that \mathcal{L} and \mathcal{L}_ω are isomorphic polarizations on A . To do so, we first note that $\alpha = \omega^{-1} \cdot \chi$ is a unit in \mathcal{O} of reduced norm $n(\alpha) = -1$. This is due to the fact that the principal ideal mR_F is supported at the prime ideals $\mathfrak{p} \mid D$, as can be checked locally. This unit, or rather $\iota(\alpha)$, is an automorphism of A , and we now show that $\alpha^*\mathcal{L}_\omega = \mathcal{L} \in \text{NS}(A)$. Indeed, by Proposition 5.4 and since $c_1(\mathcal{L})$ and χ anti-commute, $c_1(\alpha^*(\mathcal{L}_\omega)) = \bar{\alpha}\omega^{-1}c_1(\mathcal{L})\omega\alpha = \frac{1}{m}\bar{\chi}c_1(\mathcal{L})\chi = -\chi^{-1}c_1(\mathcal{L})\chi = \chi^{-1}\chi c_1(\mathcal{L}) = c_1(\mathcal{L})$. Since $W_0 = U_0 \cdot V_0$, this proves part of our claim.

In order to show that the action of W_0 on the fibres of π_F is free and transitive in an open and dense subset \mathcal{U}_F of $\tilde{X}_{B/F} \subset \mathcal{H}_F$, assume that $[(A, j_F, \mathcal{L})]$ is a point in $\tilde{X}_{B/F}$ that is represented by an abelian variety A such that $\iota : \mathcal{O} \xrightarrow{\sim} \text{End}(A)$ is an isomorphism. In other words, we consider the fibre of π_F at a non-Heegner point. In this case, if $[(A, \iota, \mathcal{L})]$ lies over $[(A, j_F, \mathcal{L})]$, then, by the Skolem-Noether Theorem, any element on the fibre of π_F at this point must be represented by a triplet $(A, \omega^{-1}\iota\omega, \mathcal{L}_\omega)$ for some $\omega \in \text{Norm}_{B^*}(\mathcal{O})$. Moreover, for \mathcal{L}_ω to be a polarization on A , ω must have totally positive reduced norm (cf. [Ro2], Theorem 5.1). Since we clearly have that $(A, \iota, \mathcal{L}) \simeq (A, \iota_\omega, \mathcal{L}_\omega)$ for any $\omega \in F^*\mathcal{O}^1$, we deduce that there exists a subgroup of W^1 that acts freely and transitively on the fibre of π_F at any non-Heegner point.

Therefore, this subgroup must contain W_0 . Let us now see that it cannot be larger than W_0 . Assume that $\omega \in W^1$ is such that $\mathcal{L} \simeq \mathcal{L}_\omega$. Then there exists $\alpha \in \mathcal{O}^* \xrightarrow{\iota} \text{Aut}(A)$ such that $\bar{\alpha}\omega^{-1}\mu\omega\alpha = \mu$. Taking reduced norms, this already implies that $n(\alpha)^2 = 1$.

If $n(\alpha) = 1$, then $\bar{\alpha} = \alpha^{-1}$ and the above yields $\mu\omega\alpha = \omega\alpha\mu$. This means that $\omega\alpha \in F(\mu) \cap \mathcal{O} = S$ and therefore $\omega \in U_0$. If $n(\alpha) = -1$, then $\mu\omega\alpha = -\omega\alpha\mu$. Write $\chi = \omega\alpha$. From the discussion above, we have that $\text{tr}(\mu\chi) = \mu\chi + \bar{\chi}\bar{\mu} = \mu\chi - \bar{\chi}\mu = -\text{tr}(\chi)\mu \in F$. Thus, since $\text{tr}(\chi) \in F$ as well, we deduce that actually $\text{tr}(\chi) = 0$. Since $n(\chi) = n(\omega)n(\alpha) = -n(\omega)$, we obtain that $\chi^2 = n(\omega)$. This says that $\omega \in V_0$.

Now that we know this, we obtain as a consequence that $\pi_F : X_B \rightarrow \mathcal{H}_F$ is a quasifinite map that factors through the natural projection of X_B onto the quotient X_B/W_0 and a morphism $b_F : X_B/W_0 \rightarrow \mathcal{H}_F$ that is one-to-one outside the Heegner locus of X_B/W_0 . It is a well-known fact (see [Sh2]) that Heegner points on X_B (and thus also on the quotient X_B/W_0) are isolated. Since b_F is an algebraic morphism, it must be a birational equivalence between X_B/W_0 and its image in \mathcal{H}_F whose inverse must be defined everywhere but at a finite set \mathcal{T}_F of Heegner points.

Moreover, since $W_0 \subseteq \text{Aut}_{\mathbb{Q}}(\mathcal{S}_B)$, the projection $\mathcal{S}_B \rightarrow \mathcal{S}_B/W_0$ is defined over \mathbb{Q} . Since π_F is also a morphism over \mathbb{Q} which is the composition of the above projection and the birational equivalence b_F , it follows that b_F is also defined over \mathbb{Q} . This finishes the proof of the first part of Theorem 4.4.

Now let (S, φ) be an Eichler pair for \mathcal{O} and identify S with its image $\varphi(S)$ in \mathcal{O} . As we saw, it induces a natural morphism $\pi_{S, \varphi} : X_B \rightarrow \mathcal{H}_S$ from X_B into the Hilbert modular variety \mathcal{H}_S in such a way that we have $\pi_F : X_B \xrightarrow{\pi_{(S, \varphi)}} \mathcal{H}_S \rightarrow \mathcal{H}_F$. As the situation is very similar to the one studied above, we will limit ourselves to showing that the subgroup $V_0(S)$ of the stable group W_0 acts freely and transitively on the fibre of $\pi_{(S, \varphi)}$ at any non-Heegner point of $\pi_{(S, \varphi)}(X_B) = \tilde{X}_{B/(L, \varphi)}$ in \mathcal{H}_L . This already implies that $\pi_{(S, \varphi)}$ is quasifinite.

Let $[(A, \iota, \mathcal{L})]$ be a closed point on X_B represented by a principally polarized abelian variety with quaternionic multiplication. From the above discussion, it is clear that any other point at the same geometric fibre of $\pi_{(S, \varphi)}$ as $[(A, \iota, \mathcal{L})]$ is represented by $(A, \iota_\omega, \mathcal{L}_\omega)$ for some $\omega \in W_0$.

If $(A, \iota \cdot \varphi, \mathcal{L}) \simeq (A, \iota_\omega \cdot \varphi, \mathcal{L}_\omega)$ is an isomorphism given by $\alpha \in \mathcal{O}^*$, then, as we already saw, $n(\alpha) = \pm 1$. Write $\chi = \omega\alpha$.

If $n(\alpha) = 1$, $\chi \in S \subset F(\mu)$. At the same time, we must have that, for any $\beta \in S \subset \mathcal{O}$, $\chi\beta = \beta\chi$. This means that χ commutes element-wise with L and $F(\mu)$. Since the two quadratic extensions are distinct because the first is totally real while the second purely imaginary, $\chi \in R_F^*$ and thus ω is the identity element of W_0 .

If $n(\alpha) = -1$, then $\alpha^*(\mathcal{L}_\omega) = \mathcal{L}$ implies that $B = F + F\mu + F\chi + F\mu\chi$ while $\alpha^{-1}\iota_\omega|_S\alpha = \iota|_S$ says that $\chi \in S$. Thus $\omega \in V_0(S)$. The converse also holds, and the theorem follows as before. \square

ACKNOWLEDGEMENTS

I am indebted to P. Bayer for her assistance throughout the elaboration of this work. I also express my gratitude to B. Mazur and P. Clark for the interest they have shown in this article. Finally, I thank J. Kramer and U. Kuehn for their warm hospitality at the Humboldt-Universität zu Berlin during the fall of 2001.

REFERENCES

- [BaBo] W. L. Baily, A. Borel, Compactification of arithmetic quotients of bounded symmetric domains *Ann. Math.* **84** (1966), 443-507. MR **35**:6870
- [ChFr1] T. Chinburg, E. Friedman, The smallest arithmetic hyperbolic 3-orbifold, *Invent. Math.* **86** (1986), 507-527. MR **88a**:22022
- [ChFr2] T. Chinburg, E. Friedman, Hilbert symbols, class groups and quaternion algebras, *J. Théor. Nombres Bordeaux* **12** (2000), 367-377. MR **2002a**:11131
- [Cl] P.L. Clark, *Local and global points on moduli spaces of abelian surfaces with potential quaternionic multiplication*, Harvard Ph.D. Thesis, in progress.
- [Ei1] M. Eichler, Bestimmung der Idealklassenzahl in gewissen normalen einfachen Algebren, *J. reine angew. Math.* **176** (1937), 192-202.
- [Ei2] M. Eichler, Über die Idealklassenzahl hypercomplexer Systeme, *Math. Z.* **43** (1938), 481-494.
- [vdGe] G. van der Geer, *Hilbert modular surfaces*, *Ergebn. Math. Grenz.* **16**, Springer, 1987. MR **89c**:11073
- [Ha] R. Hartshorne, *Algebraic Geometry*, *Grad. Texts Math.* **52**, Springer, 1977. MR **57**:3116
- [HaHaMo] Y. Hasegawa, K. Hashimoto, F. Momose, Modular conjecture for \mathbb{Q} -curves and QM-surfaces, *International J. Math.* **10** (1999), 1011-1036. MR **2001f**:11095
- [HaSc] H. Hasse, O. Schilling, Die Normen aus einer normalen Divisionsalgebra, *J. reine angew. Math.* **174** (1936), 248-252.
- [Jo] B.W. Jordan, *On the Diophantine arithmetic of Shimura curves*, Harvard Ph.D. Thesis, 1981.
- [LaBi] H. Lange, Ch. Birkenhake, *Complex Abelian Varieties*, *Grundl. math. Wiss.* **302**, Springer, 1992. MR **94j**:14001

- [Mi] J. S. Milne, Points on Shimura varieties mod p , *Proc. Symp. Pure Math.* **33** (1979), 165-184. MR **82e**:10048
- [Ri] K. A. Ribet, On modular representations of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms, *Invent. Math.* **100** (1990), 431-476. MR **91g**:11066
- [Ro1] V. Rotger, On the group of automorphisms of Shimura curves and applications, *Compositio Math.* **132** (2002), 229-241. MR **2003c**:11065
- [Ro2] V. Rotger, Quaternions, polarizations and class numbers, *J. reine angew. Math.* **561** (2003), 177-197.
- [Ro3] V. Rotger, The field of moduli of quaternionic multiplication on abelian varieties, preprint 2002, submitted to publication.
- [Ro4] V. Rotger, Shimura curves embedded in Igusa's threefold, preprint 2002, to appear in *Modular curves and abelian varieties*, Progress in Mathematics, Birkhäuser.
- [Sh1] G. Shimura, On analytic families of polarized abelian varieties and automorphic functions, *Ann. Math.* **78** (1963), 149-192. MR **27**:5934
- [Sh2] G. Shimura, Construction of class fields and zeta functions of algebraic curves, *Ann. Math.* **85** (1967), 58-159. MR **34**:4268
- [Vi] M.F. Vignéras, *Arithmétique des algèbres de quaternions*, Lect. Notes Math. **800**, Springer, 1980. MR **82i**:12016

UNIVERSITAT POLITÈCNICA DE CATALUNYA, DEPARTAMENT DE MATEMÀTICA APLICADA IV (EU-PVG), AV. VÍCTOR BALAGUER S/N, 08800 VILANOVA I LA GELTRÚ, SPAIN
E-mail address: `vrotger@mat.upc.es`