

## Irreducibility of The Igusa Tower

**Haruzo HIDA**

*Department of Mathematics, UCLA, Los Angeles, CA 90095–1555, USA*

*E-mail: hida@math.ucla.edu*

**Abstract** We shall give a simple (basically) purely in characteristic  $p$  proof of the irreducibility of the Igusa tower over Shimura varieties of PEL type. Our result covers Shimura variety of type A and type C classical groups, in particular, the Siegel modular varieties, the Hilbert–Siegel modular varieties, Picard surfaces and Shimura varieties of inner forms of unitary and symplectic groups over totally real fields.

**Keywords** Shimura variety, Igusa tower,  $p$ -adic monodromy

**MR(2000) Subject Classification** 11G15, 11G18, 11G25

Fix a prime  $p$ . In [1], we have shown that the geometric automorphism group of the irreducible component of the mod  $p$  Shimura variety of PEL type (of level away from  $p$ ) associated with a reductive group  $G$  of unitary and symplectic type is almost identical to  $G_1(\mathbb{A}^{(p\infty)})$  modulo global center (cf. Lemma 1.1). Here  $G_1$  is the derived subgroup of  $G$ .

In the present paper, based on this result, we give a (basically) characteristic  $p$  proof of the irreducibility of the Igusa tower over a reduction modulo  $p$  of the Shimura variety by showing that the stabilizer of an irreducible component of the tower is as large as possible under the group theoretic constraints imposed by the PEL data. This is a characteristic  $p$ -version of the proof given in [2] Section 8.4 and [3] where we used mixed-characteristic results to show the maximality of the stabilizer. Here is a general axiomatic approach to prove the irreducibility of an étale covering  $\pi : Ig \rightarrow S$  of an irreducible variety  $S$  over an algebraically closed field  $\mathbb{F}$ . Assume the following two axioms:

(A1) A group  $\mathcal{G} = M_1 \times \mathcal{G}_1$  acts on  $Ig$  and  $S$  compatibly so that  $M_1 \subset \text{Aut}(Ig/S)$ ,  $\mathcal{G}_1 \subset \text{Aut}(S)$  and  $\mathcal{G}_1$  acts trivially on  $\pi_0(Ig)$ .

(A2)  $M_1$  acts on each fiber of  $Ig/S$  **transitively**; so,  $M_1$  acts transitively on  $\pi_0(Ig)$ .

Let  $I/\mathbb{F}$  be an irreducible component of  $Ig/\mathbb{F}$ . Then  $I \rightarrow S$  is a connected Galois étale covering. Condition (A2) provides a canonical group embedding  $\text{Gal}(I/S) \hookrightarrow M_1$ ; moreover, if it is surjective, one concludes by Galois theory that  $Ig = I$ . Let  $H$  be the stabilizer of  $I \in \pi_0(Ig)$  in  $\mathcal{G}$ . It contains  $\mathcal{G}_1$ . Pick a point  $x \in Ig$  (which can be a generic point), and look at the stabilizer  $H_x \subset \mathcal{G}$  of  $x$ . By (A2), there exists  $g_x \in M_1$  such that  $g_x(x) \in I$ ; by connectedness of  $I$ , we have  $g_x H_x g_x^{-1} \subset H$ . Then we show that  $M_1 = \mathcal{G}/\mathcal{G}_1$  is generated by the image modulo  $\mathcal{G}_1$  of  $\{g_x H_x g_x^{-1} | x \in Ig\}$ . This implies that  $M_1 = \text{Gal}(I/S)$ , since  $\langle g_x H_x g_x^{-1} \rangle_x = M_1 \hookrightarrow H/\mathcal{G}_1 \hookrightarrow \text{Gal}(I/S)$ .

When  $S$  is the neutral component  $S^{(p)}$  of the modulo  $p$  Shimura variety  $Sh^{(p)}$  of level away from  $p$  and  $Ig/S$  is the Igusa tower over  $S$ ,  $\mathbb{F} = \overline{\mathbb{F}}_p$  is an algebraic closure of  $\mathbb{F}_p$ ,  $\mathcal{G}_1$  is given by  $G_1(\mathbb{A}^{(\Sigma)})$  for a well chosen finite set  $\Sigma$  of places of  $\mathbb{Q}$  containing  $\{p, \infty\}$ ,  $M_1$  is the Levi-subgroup of the maximal parabolic subgroup of  $G_1(\mathbb{Z}_p)$  associated with the connected-étale exact sequence of the  $p$ -divisible Barsotti–Tate group of the universal abelian scheme over  $S$ . In particular,  $H_x$  for  $x$  moving around  $Ig(\mathbb{F})$  contains a subgroup  $T_x$  isomorphic to a maximal torus in each isomorphism class of maximal tori in  $\mathcal{G}$  anisotropic at  $\infty$ , and from this fact, it is clear that  $\{g_x T_x g_x^{-1} | x \in Ig(\mathbb{F})\}$  and  $\mathcal{G}_1$  topologically generate  $\mathcal{G}$  (though we will give a detailed argument for this fact in Sections §2–3). We say our argument is “basically” of characteristic  $p$ , because we use for simplicity the canonical (characteristic 0) lift of the characteristic  $p$  abelian variety sitting over  $x \in Ig(\mathbb{F})$  to define  $T_x \subset G_1$ , though the use of the lift could be avoided by using the determination of the automorphism group  $\text{Aut}(S^{(p)}/\mathbb{F})$  in [1].

Let us make slightly more precise the reductive group  $G$  and its action on the Igusa tower  $Ig$  limiting ourselves to the extent sufficient to formulate our result (more details follow in the text). The group  $G$  is realized as a subgroup of  $GL(V)$  for  $V = L_{\mathbb{Q}} = L \otimes_{\mathbb{Z}} \mathbb{Q}$  with a well chosen lattice  $L$  (fixed in the PEL data). To be more precise, let  $D$  be a central simple algebra  $D$  over a number field  $F$  with a positive involution. Fix a maximal order  $O_D$  of the algebra  $D$  stable under the involution, and take a projective  $O_D$ -module  $L$  of finite type, and put  $V = L \otimes_{\mathbb{Z}} \mathbb{Q}$ . The reductive group  $G_{/\mathbb{Q}}$  is an appropriate subgroup of  $\text{Aut}_D(V)$ ; in other words,  $G(\mathbb{Q})$  is a classical group inside  $\text{Aut}_D(V)$ .

Fix a prime  $p$ . We introduce the (prime-to- $p$ ) PEL data and the corresponding Shimura variety  $Sh^{(p)}$  of prime-to- $p$  level. The Shimura variety  $Sh^{(p)}$  of PEL type classifies abelian schemes  $A$  with an algebra embedding  $O_D \subset \text{End}(A)$ , a polarization class  $\bar{\lambda}$  and a prime-to- $p$  level structure  $\eta^{(p)} : L \otimes_{\mathbb{Z}} \mathbb{A}^{(p\infty)} \cong \mathcal{T}^{(p)} A \otimes_{\mathbb{Z}} \mathbb{A}^{(p\infty)}$ , where  $\mathcal{T}^{(p)} A = \varprojlim_{p \nmid N} A[N]$  is the prime-to- $p$  Tate module of  $A$  and  $\mathbb{A}^{(p\infty)}$  is the adèle ring away from  $p$  and  $\infty$ . The level structure  $\eta^{(p)}$  is supposed to be  $O_D$ -linear. The Shimura variety  $Sh^{(p)}$  is defined over the  $p$ -integer ring  $O_{E,(p)} = O_E \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$  ( $\mathbb{Z}_{(p)} = \mathbb{Q} \cap \mathbb{Z}_p$ ) of the reflex field  $E$ . Fix embeddings  $i_p : \overline{E} \hookrightarrow \overline{\mathbb{Q}}_p$  and  $i_{\infty} : \overline{E} \hookrightarrow \mathbb{C}$  for an algebraic closure  $\overline{E}/E$ , and let  $\mathcal{W}$  be the strict henselization inside  $\overline{E}$  of the local ring of  $O_{E,(p)}$  corresponding to the embedding  $\overline{E} \hookrightarrow \overline{\mathbb{Q}}_p$ . Let  $\mathbb{F}$  be the residue field  $\mathcal{W}/\mathfrak{m}_{\mathcal{W}}$  (which is an algebraic closure of  $\mathbb{F}_p$ ). Let  $Sh_{/\mathcal{W}}^{(p)} = Sh^{(p)} \times_{O_{E,(p)}} \mathcal{W}$ . Then (under the existence of a smooth toroidal compactification of  $Sh_{/\mathcal{W}}^{(p)}$  over  $\mathcal{W}$ ; for example, see [4]),  $Sh_{/\mathcal{W}}^{(p)}$  is a disjoint union of smooth irreducible subschemes whose generic fiber and special fiber are geometrically irreducible.

The group  $G(\mathbb{A}^{(p\infty)})$  acts on  $Sh^{(p)}$  by sending  $(A, \bar{\lambda}, \eta^{(p)})$  to  $(A, \bar{\lambda}, \eta^{(p)} \circ g)$  for  $g \in G(\mathbb{A}^{(p\infty)})$ . Fix an irreducible component  $S = S_{/\mathbb{F}}^{(p)}$  inside the special fiber  $Sh_{/\mathbb{F}}^{(p)} = Sh_{/\mathcal{W}}^{(p)} \times_{\mathcal{W}} \mathbb{F}$ . Then if  $G$  is of type A or C, the derived group  $G_1(\mathbb{A}^{(p\infty)})$  preserves the component  $S$  of  $Sh_{/\mathbb{F}}^{(p)}$ .

In my book [2], many irreducibility theorems of the Igusa tower  $Ig$  over  $S$  are proven (Theorems 3.3, 4.21, 6.27 and Corollary 8.17). In particular, the proof of the first three theorems Theorems 3.3, 4.21 and 6.27 in [2] consists in showing explicitly that the stabilizer of each irreducible component is as large as possible under the PEL data (using  $p$ -adic mixed characteristic valuation ring of the function field of the characteristic 0 Shimura variety; see [3]

for a more concise account of the argument). In this paper, we shall give a direct proof of the statement of the corollary in line with the proof of the first three theorems (without using mixed characteristic  $p$ -adic valuations).

Proofs purely in characteristic  $p$  of the fact seem important, because as pointed out by Chai, it is plausible that one of such arguments may produce a similar result for the  $p$ -adic monodromy group over each (positive dimensional) leaf inside  $S/\mathbb{F}$  (see [5] 2.2 and 2.8 for a definition of leaves, and see [6] for the  $\ell$ -adic monodromy of leaves with  $\ell \neq p$ ). Indeed, in [7], Chai gives a sketch of three known purely characteristic  $p$  proofs of irreducibility of the Igusa tower over the Siegel modular variety (and their history), and in [5], Chai and Oort give a proof of irreducibility of leaves and the maximality of the  $p$ -adic monodromy of the leaves in the Siegel modular variety. Note here that leaves of non-maximal and non-minimal dimension may not have characteristic 0 lift in general (so the argument using mixed characteristic valuations does not apply to such middle dimensional leaves). See [7] Section 6 for an indication of how to prove maximality of  $p$ -adic monodromy for leaves without characteristic 0 lift. We also note that Boyer [8] has given a proof of such maximality (by cohomological means) in the special cases of the Igusa tower of the first and second kind studied in [9].

Although we concentrate on the groups  $G$  of type A and C and do not touch upon type D groups, our method may be applicable even to the groups of type D, because what we actually need is the transitivity of the action of a subgroup  $\tilde{P}$  in  $G(\mathbb{A}^{(\infty)})$  on  $\pi_0(Ig/\mathbb{F})$  (see Remark 4.1). Indeed, fixing an irreducible component  $I$  of the Igusa tower  $Ig$ , by the transitivity of the action, any closed point  $x \in Ig(\mathbb{F})$  can be brought into  $I$  by an element  $g_x \in \tilde{P}$ ; so, the stabilizer  $H_x$  of the closed point  $x \in Ig$  can be brought into  $\text{Gal}(I/S)$  under the conjugation by  $g_x$ . Since the choice of  $x$  is arbitrary, one expects that  $\text{Gal}(I/S)$  is as large as possible and hence the irreducibility. Though for simplicity we use the strong approximation theorem to pin down the above heuristic into a rigorous proof, the transitivity does not require the strong approximation (the weaker approximation theorem is enough).

The center of the ring  $O_D$  is the integer ring  $O$  of a field  $F$ , and by the definition of PEL structures (see below in the text),  $F$  is either a totally real field  $F_0$  or a CM quadratic extension of a totally real field  $F_0$ . We take the restriction of scalar  $G/\mathbb{Q}$  to  $\mathbb{Q}$  of an inner form of a symplectic or a unitary group (relative to  $F/F_0$ ) over  $F_0$ , and assume that  $G$  has the associated Shimura variety of PEL type (thus  $G$  is of type A or C). In characteristic 0, the full Shimura variety  $Sh/E$  is well defined and  $Sh/G(\mathbb{Z}_p) = Sh^{(p)}$  for a well specified maximal open compact subgroup  $G(\mathbb{Z}_p) \subset G(\mathbb{Q}_p)$ . We will specify  $G(\mathbb{Z}_p)$  later in (H4–5). Then  $Sh_{/O_{E,(p)}}^{(p)}$  classifies a certain type of abelian schemes under the hypothesis (H1–5). Thus we have a universal abelian scheme  $\mathbf{A}$  over  $Sh_{/\mathbb{F}}^{(p)}$ . We assume throughout this paper that

- *The prime  $p$  is unramified in  $F/\mathbb{Q}$  (to have smooth  $Sh_{/\mathcal{W}}^{(p)}$ ),*
- *$G$  is unramified over  $\mathbb{Q}_p$ , and  $G(\mathbb{Q}_p)$  is a subgroup of  $GL_d(F_p)$  for a positive integer  $d$  specified later (this condition follows from the condition (H5) and also to have smooth  $Sh_{/\mathcal{W}}^{(p)}$ ),*
- *$\mathbf{A}$  is ordinary over an open dense subscheme of  $Sh_{/\mathbb{F}}^{(p)}$ .*

The last ordinarity condition can be stated under the unramifiedness conditions as a condition of the signature of the symplectic and unitary group  $G$  (cf. [10] and [2] Lemma 8.10). By the ordinarity condition, the connected-étale exact sequence of the  $p$ -divisible Barsotti–Tate

group  $\mathbf{A}[p^\infty]$  of  $\mathbf{A}$  over the ordinary locus gives rise to a maximal parabolic subgroup  $P$  of  $G/\mathbb{Z}_p$  (and  $\tilde{P} = P(\mathbb{Z}_p)G(\mathbb{A}^{(p^\infty)})$ ). The Levi-factor  $\mathcal{L}$  of  $P$  has the following subgroup  $M_1 = \mathcal{L} \cap G_1$  whose  $\mathbb{Z}_p$ -points provide a group of geometric automorphisms of the Igusa tower  $Ig$  over  $Sh^{(p)}$ . Moreover the main theorem (Theorem 0.1 below) of this paper is equivalent to the equality  $M_1(\mathbb{Z}_p) = \text{Aut}(Ig/Sh^{(p)})$ .

For each prime factor  $\mathfrak{P}|p$  of  $O$ , the  $O_{D,\mathfrak{P}}$ -rank of the Tate module of each étale quotient  $\mathbf{A}[\mathfrak{P}^\infty]^{et}$  of the  $\mathfrak{P}$ -torsion points (over the ordinary locus) gives a positive integer  $m(\mathfrak{P})$ . By the definition of the moduli problem, for the positive integer  $d$  (as above) independent of  $\mathfrak{P}|p$ , we have  $m(\mathfrak{P}) + m(\overline{\mathfrak{P}}) = d$  if  $F$  is a CM field, where  $\overline{\mathfrak{P}}$  is the conjugate of  $\mathfrak{P}$  under the generator  $c \in \text{Gal}(F/F_0)$ , and  $2m(\mathfrak{P}) = d$  if  $F$  is totally real. The pair of numbers  $(m(\mathfrak{P}), m(\overline{\mathfrak{P}}))$  is the signature of the symplectic or the unitary group  $G$  at the infinite place corresponding to  $\mathfrak{P}$  via  $i_p$  and  $i_\infty$ . We have the following explicit description of  $M_1$ :

$$M_1 = \begin{cases} \left\{ (g_{\mathfrak{P}}) \in \prod_{\mathfrak{P}|p} GL_{m(\mathfrak{P})}(O_{\mathfrak{P}}) \mid \forall \mathfrak{P}|p \det(g_{\mathfrak{P}}) = \det(g_{\overline{\mathfrak{P}}}) \right\}, & \text{if } F \neq F_0, \\ \prod_{\mathfrak{P}|p} GL_{m(\mathfrak{P})}(O_{\mathfrak{P}}), & \text{if } F = F_0. \end{cases}$$

**Theorem 0.1** *Assume the above three conditions and (H1–5) we will specify later. Let  $S_{K/\mathbb{F}}$  be the image of  $S_{/\mathbb{F}}$  in the quotient of  $Sh^{(p)}$  by a compact subgroup  $K$  of  $G(\mathbb{A}^{(p^\infty)})$ . Let  $Ig_K = (Ig/K) \times_{Sh^{(p)}/K} S_K$ , which is the Igusa tower over  $S_K$ . Then for each irreducible component  $I_K$  of  $Ig_K$ , we have a natural isomorphism*

$$\text{Gal}(I_K/S_K) \cong M_1.$$

*In particular, if  $F$  is totally real,  $Ig_K$  is irreducible, and if  $F \neq F_0$  and  $Sh^{(p)}$  is not projective, each irreducible component of  $Ig_K$  contains an unramified cusp over the infinity cusp  $\infty$  of  $S_K$ .*

This is exactly Corollary 8.17 of [2]. By definition,  $\mathcal{L}(\mathbb{Z}_p)$  acts transitively on each fiber of  $Ig_K$  over  $S_K$  (and hence transitively on  $\pi_0(Ig_{K/\mathbb{F}})$ ). Thus  $\text{Gal}(I_K/S_K)$  is the stabilizer of  $I_K \in \pi_0(Ig_{K/\mathbb{F}})$  in  $\mathcal{L}(\mathbb{Z}_p)$ . As shown in the last page of [2] (p. 374), by a geometric argument, we have an inclusion  $\text{Gal}(I_K/S_K) \subset M_1$  (see [3] 3.3 for a simpler argument). We will show  $M_1 \supset \text{Gal}(I_K/S_K)$  by proving basically that the groups  $g_x H_x g_x^{-1}$  ( $x$  running over  $Ig_K(\mathbb{F})$ ) generate  $M_1 \times G_1(\mathbb{A}^{(p^\infty)})$  topologically, where  $H_x \subset \tilde{P}$  is the stabilizer of  $x$  and  $g_x \in G(\mathbb{A}^{(\infty)})$  brings  $x$  into  $I_K(\mathbb{F})$ . The group  $H_x$  may be bigger than  $T_x(\mathbb{Z}_p) \subset M_1 \times G_1(\mathbb{A}^{(p^\infty)})$ , but only the tori  $T_x$ s are actually sufficient for our purpose.

In the following section, we give a short description of the Shimura variety we deal with, state the conditions (H1–5), explain how the number  $m(\mathfrak{P})$  is determined by  $\mathbf{A}[\mathfrak{P}^\infty]^{et}$ , and how the group  $M_1$  is sent isomorphically onto  $\mathcal{L} \cap G_1(\mathbb{Z}_p)$ . In the next two sections §2–3, we prepare necessary lemmas from group theory and algebraic number theory (in order to show definitely the groups  $\{g_x T_x(\mathbb{Z}_p) g_x^{-1} \mid x \in Ig(\mathbb{F})\}$  generate  $M_1$ ). In the final section (§4), we give the proof of the first assertion of the theorem.

As for the last two assertions, when  $F = F_0$ , by the construction of  $Ig$ ,  $M_1$  acts transitively on  $\pi_0(Ig_{K/\mathbb{F}})$ , and hence the irreducibility of  $Ig_K$  follows. If  $F \neq F_0$ , the group  $M = \prod_{\mathfrak{P}|p} GL_{m(\mathfrak{P})}(O_{\mathfrak{P}}) \subset \mathcal{L}$  acts transitively on  $\pi_0(Ig_{K/\mathbb{F}})$ , and hence  $Ig_K$  is not irreducible.

Since the action of  $M$  sends the infinity cusp  $\infty$  to an unramified cusp of each irreducible component (if  $Sh^{(p)}$  is not projective), the last assertion follows from the first.

## Contents

1	Shimura variety	5
2	Group theory	12
3	Algebraic number theory	14
4	Irreducibility	18
	Reference	20

### 1 Shimura Variety

We recall the definition and the assumptions governing the Shimura variety we are looking into. Let  $D$  be a central simple algebra over the field  $F$  with a positive involution  $\rho$  (thus  $\text{Tr}_{D/\mathbb{Q}}(xx^\rho) > 0$  for all  $0 \neq x \in D$ ). Let  $F_0$  be the subfield of  $F$  fixed by  $\rho$ . Thus  $F_0$  is necessarily a totally real field, and either  $F = F_0$  or  $F$  is a CM quadratic extension of  $F_0$  ( $\rho$  restricts to the generator  $c \in \text{Gal}(F/F_0)$ ). We write  $O_0$  (resp.  $O$ ) for the integer ring of  $F_0$  (resp.  $F$ ). We have fixed the algebraic closure  $\mathbb{F}$  of the prime field  $\mathbb{F}_p$  which is the residue field of  $\mathcal{W}$ .

Take a finite subset  $\Sigma$  of rational places including  $\infty$  and  $p$ . Let  $F_+^\times$  be the subset of totally positive elements in  $F_0$ , and  $O_{(\Sigma)}$  denotes the localization of  $O$  at  $\Sigma$  (disregarding the infinite place in  $\Sigma$ ) and  $O_\Sigma$  is the completion of  $O$  at  $\Sigma$  (again disregarding the infinite place); so, we have  $O_{(\Sigma)} = O_\Sigma \cap F$  in  $F_\Sigma = \prod_{v \in \Sigma} F_v$  with the completion  $F_v$  of  $F$  at  $v \in \Sigma$ . We write  $O_{(\Sigma)_+}^\times = F_+^\times \cap O_{(\Sigma)}$ .

Let  $O_D$  be a maximal order of  $D$  stable under  $\rho$ . Let  $L$  be a projective  $O_D$ -module with a non-degenerate  $F_0$ -linear alternating form  $\langle \cdot, \cdot \rangle : L_\mathbb{Q} \times L_\mathbb{Q} \rightarrow F_0$  for  $L_R = L \otimes_{\mathbb{Z}} R$  such that  $\langle bx, y \rangle = \langle x, b^\rho y \rangle$  for all  $b \in D$ . We often write  $V$  for  $L_\mathbb{Q}$ .

Let  $C$  be the opposite algebra of  $C^\circ = \text{End}_D(L_\mathbb{Q})$ . Then  $C$  is a central simple algebra over  $F$  in the same Brauer class of  $D$ . We write  $C_R = C \otimes_{\mathbb{Q}} R$ ,  $D_R = D \otimes_{\mathbb{Q}} R$  and  $F_{0,R} = F_0 \otimes_{\mathbb{Q}} R$ . The algebra  $C$  has involution  $*$  given by  $\langle cx, y \rangle = \langle x, c^*y \rangle$  for  $c \in C$ . The involution  $*$  (resp.  $\rho$ ) induces the involution  $* \otimes 1$  (resp.  $\rho \otimes 1$ ) on  $C_R$  (resp. on  $D_R$ ) which we write as  $*$  (resp.  $\rho$ ) simply. Define an algebraic group  $G_{/\mathbb{Q}}$  by

$$G(R) = \{g \in C_R \mid \nu(g) := gg^* \in (F_{0,R})^\times\} \quad \text{for } \mathbb{Q}\text{-algebras } R. \quad (1.1)$$

We write  $G_1$  for the derived group of  $G$ ; thus,

$$G_1 = \{g \in G \mid N_C(g) = \nu(g) = 1\}$$

for the reduced norm  $N_C$  of  $C$  over  $F$ . We write  $Z^G = G/G_1$  for the cocenter of  $G$ . Then  $g \mapsto (\nu(g), N_C(g))$  identifies  $Z^G$  with a sub-torus of  $\text{Res}_{F_0/\mathbb{Q}}\mathbb{G}_m \times \text{Res}_{F/\mathbb{Q}}\mathbb{G}_m$ . If  $F = F_0$ ,  $G_1$  is equal to the kernel of the similitude map  $g \mapsto \nu(g)$ ; so, in this case, we ignore the right factor  $\text{Res}_{F/\mathbb{Q}}\mathbb{G}_m$  and regard  $Z^G \subset \text{Res}_{F/\mathbb{Q}}\mathbb{G}_m$ . We write  $Z^G(\mathbb{R})^+$  for the identity connected component of  $Z^G(\mathbb{R})$  and put  $Z^G(\mathbb{Z}_{(\Sigma)})^+ = Z^G(\mathbb{R})^+ \cap (O_{0(\Sigma)}^\times \times O_{(\Sigma)}^\times)$ ; so,  $Z^G(\mathbb{Z}_{(\Sigma)})^+ = O_{(\Sigma)_+}^\times$  if  $F = F_0$ . Write  $\mathbb{Z}_\Sigma = \prod_{\ell \in \Sigma - \{\infty\}} \mathbb{Z}_\ell$  and  $\mathbb{Z}_\Sigma^{(p)} = \mathbb{Z}_\Sigma / \mathbb{Z}_p$ .

We suppose to have an  $\mathbb{R}$ -algebra homomorphism  $h : \mathbb{C} \rightarrow C_{\mathbb{R}}$  such that  $h(\bar{z}) = h(z)^*$  and (H1)  $\langle x, y \rangle = \langle x, h(i)y \rangle$  induces a positive definite hermitian form on  $L_{\mathbb{R}}$ .

We define  $X$  to be the conjugacy classes of  $h$  under  $G(\mathbb{R})$ . Then  $X$  is a finite disjoint union of copies of the hermitian symmetric domain isomorphic to  $G(\mathbb{R})^+/C_h$  (see [11] Lemma 4.1), where  $C_h$  is the stabilizer of  $h$  and the superscript “+” indicates the identity connected component of the Lie group  $G(\mathbb{R})$ . Then the pair  $(G, X)$  satisfies the three axioms (see [12] 2.1.1.1–3 or [2] 7.2.1) specifying the data for defining the Shimura variety  $Sh$  (and its field of definition, the reflex field  $E$ ; see below). In [12], two more axioms are stated to simplify the situation: (2.1.1.4-5). These two extra axioms may not hold generally for our choice of  $(G, X)$  (see [13] Remark 2.2 on p. 322).

The complex points of  $Sh$  are given by

$$Sh(\mathbb{C}) = G(\mathbb{Q}) \backslash (G(\mathbb{A}^{(\infty)}) \times X) / \overline{Z(\mathbb{Q})}.$$

The morphism  $h$  gives rise to a complex structure on the real vector space  $L_{\mathbb{R}}$  (by putting  $z \cdot v = h(z)v$  for  $v \in L_{\mathbb{R}}$  and  $z \in \mathbb{C}$ ). Thus we can split  $L_{\mathbb{C}} = V_1 \oplus V_2$  so that  $z \in \mathbb{C}$  acts on  $V_1$  (resp.  $V_2$ ) via multiplication by  $z$  (resp.  $\bar{z}$ ). Since the action of  $D$  commutes with  $h$ , we get a complex representation  $\rho_1 : D \hookrightarrow \text{End}_{\mathbb{C}}(V_1)$ . Then  $\{\text{Tr}(\rho_1(\delta)) \mid \delta \in D\}$  generates a number field  $E$  over  $\mathbb{Q}$ , which is the field of rationality of  $\rho_1$ . The field  $E$  is called the reflex field of  $(G, X)$ , and the quasi-projective variety  $Sh_{/\mathbb{C}}$  canonically descends to a quasi-projective variety  $Sh_{/E}$  (by the work of Shimura and Deligne; e.g., [2] Section 7.1). This variety  $Sh_{/E}$  can be characterized as a moduli variety over  $E$  of abelian varieties up to isogeny with multiplication by  $O_D$ . For each  $x \in X$ , we have  $h_x : \mathbb{C} \rightarrow C_{\mathbb{R}}$  given by  $z \mapsto g \cdot h(z)g^{-1}$  for  $g \in G(\mathbb{R})/C_h$  sending  $h$  to  $x$ . Then  $v \mapsto h_x(z)v$  for  $z \in \mathbb{C}$  gives rise to a complex vector space structure on  $L_{\mathbb{R}}$ , and  $A_x(\mathbb{C}) = L_{\mathbb{R}}/L$  is an abelian variety (see [14] I.3), because by (H1),  $\langle \cdot, \cdot \rangle$  induces a Riemann form on  $L$  and a polarization on  $A_{x/\mathbb{C}}$ . The multiplication by  $b \in O_D$  is given by  $(v \bmod L) \mapsto (b \cdot v \bmod L)$ .

We assume the following conditions:

(H2) The derived subgroup  $G_1$  is simply connected; so,  $G_1(\mathbb{R})$  is of type A (unitary groups) or of type C (symplectic groups);

(H3) The prime  $p$  is unramified in  $F/\mathbb{Q}$ , and  $\Sigma$  contains  $\infty$  and  $p$ ;

(H4) The pairing  $\langle \cdot, \cdot \rangle$  induces  $L_p = L \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong \text{Hom}(L_p, O_{0,p})$ ;

(H5)  $O_{D,p} = O_D \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong M_n(O_p)$  for  $O_{\ell} = O \otimes_{\mathbb{Z}} \mathbb{Z}_{\ell}$ .

Put  $O_C = \{g \in C \mid gL \subset L\}$ . Then  $O_C$  is an order of  $C$ . Under (H5),  $O_{C,p} \cong M_d(O_p)$ ; in particular,  $O_{C,p}$  is a maximal order of  $C_p$ . Let

$$G(\mathbb{Z}_{\Sigma}) = \{g \in G(\mathbb{Q}_{\Sigma}) \mid g \cdot L_{\Sigma} = L_{\Sigma}\}$$

for  $\mathbb{Q}_{\Sigma} = \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}_{\Sigma}$  and  $L_{\Sigma} = L \otimes_{\mathbb{Z}} \mathbb{Z}_{\Sigma}$ . At  $p$ , under (H3–5),  $G(\mathbb{Z}_p)$  is a hyper-special maximal compact subgroup of  $G(\mathbb{Q}_p)$ . Under (H1–5), a  $p$ -integral moduli interpretation due to Kottwitz [11] allows us to have a well defined  $p$ -integral model  $Sh^{(p)}$  over  $O_{E,(p)}$  of  $Sh/G(\mathbb{Z}_p)$  over  $E$  of level away from  $\{p, \infty\}$  (see below for a brief description of the moduli problem, and a more complete description can be found in [2] 7.1.3). We then define  $Sh_{/O_{E,(p)}}^{(\Sigma)} = Sh^{(p)}/G(\mathbb{Z}_{\Sigma}^{(p)})$ , where  $\mathbb{Z}_{\Sigma}^{(p)} = \mathbb{Z}_{\Sigma - \{p\}} = \mathbb{Z}_{\Sigma}/\mathbb{Z}_p$ . In other words,

$$Sh^{(\Sigma)}(\mathbb{C}) = G(\mathbb{Q}) \backslash (G(\mathbb{A}^{(\infty)}) \times X) / \overline{Z(\mathbb{Q})} G(\mathbb{Z}_{\Sigma}) \quad (1.2)$$

has a well defined smooth model over  $O_{E,(p)} := O_E \otimes_{\mathbb{Z}} \mathbb{Z}_{(p)}$  over which  $Sh_{O_{E,(p)}}^{(p)}$  is an étale covering. We write  $Sh^{(p)}$  for  $Sh^{(\Sigma)}$  when  $\Sigma = \{p, \infty\}$ . We also write  $\mathbb{Q}_{\Sigma}^{(p)} = \mathbb{Q}_{\Sigma}/\mathbb{Q}_p$ .

By definition,  $O_{E,(p)}$  is contained in  $\mathscr{W}$ . The scheme  $Sh_{/R}^{(\Sigma)}$  classifies, for any  $R$ -scheme  $T$ , quadruples  $(A, \bar{\lambda}, i, \eta^{(\Sigma)})_{/T}$  defined as follows:  $A$  is an abelian scheme of dimension  $\frac{1}{2} \text{rank}_{\mathbb{Z}} L$  for which we define the Tate module  $\mathcal{T}(A) = \varprojlim_N A[N]$ ,  $\mathcal{T}^{(p)}(A) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^{(p)}$ ,  $\mathcal{T}_{\Sigma}(A) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \mathbb{Z}_{\Sigma}$  and  $V^{(\Sigma)}(A) = \mathcal{T}(A) \otimes_{\mathbb{Z}} \mathbb{A}^{(\Sigma)}$ . The symbol  $i$  stands for an algebra embedding  $i : O_D \hookrightarrow \text{End}(A)$  taking the identity to the identity map on  $A$ . Though  $\eta^{(\Sigma)}$  is called a level structure away from  $\Sigma$ , it contains some information at places in  $\Sigma$ ; indeed, it is an  $O_D$ -linear map  $\eta^{(p)} : L \otimes_{\mathbb{Z}} \mathbb{A}^{(p)} \cong V^{(p)}(A)$  modulo  $G(\mathbb{Z}_{\Sigma}^{(p)})$  (that is,  $\eta^{(\Sigma)}$  is a sheaf theoretic orbit  $\eta^{(p)}G(\mathbb{Z}_{\Sigma}^{(p)})$ ; see [15] 4.3.1 about sheaf theoretic orbits). The polarization class  $\bar{\lambda}$  is a class of polarizations  $\lambda$  up to scalar multiplication by  $i(O_{(\Sigma)+}^{\times})$  which induces the Riemann form  $\langle \cdot, \cdot \rangle$  on  $L$  up to scalar multiplication by  $O_{(\Sigma)+}^{\times}$ . There is one more condition (cf. [11] Section 5 or [2] 7.1.1 (det)) specifying the module structure of  $\Omega_{A/T}$  over  $O_D \otimes_{\mathbb{Z}} \mathcal{O}_T$  (which we do not recall, but over  $E$ , it imposes that  $\text{Lie}(A_{/T}) \cong \rho_1$  as  $D$ -modules). The group  $N(G(\mathbb{Z}_{\Sigma}^{(p)})) \times G(\mathbb{A}^{(\Sigma)})$  (for the normalizer  $N(G(\mathbb{Z}_{\Sigma}^{(p)}))$  of  $G(\mathbb{Z}_{\Sigma}^{(p)})$  in  $G(\mathbb{Q}_{\Sigma}^{(p)})$ ) acts on  $Sh^{(\Sigma)}$  by

$$\eta^{(\Sigma)} = \eta^{(p)}G(\mathbb{Z}_{\Sigma}^{(p)}) \mapsto \eta^{(p)}G(\mathbb{Z}_{\Sigma}^{(p)}) \circ g = (\eta^{(p)} \circ g)G(\mathbb{Z}_{\Sigma}^{(p)}) = \eta^{(\Sigma)} \circ g.$$

If  $G(\mathbb{Z}_{\ell})$  is not quasi-split (for example, is anisotropic) at  $\ell \in \Sigma$ , the normalizer  $N(G(\mathbb{Z}_{\ell}))$  is often bigger than  $G(\mathbb{Z}_{\ell})$ ; for example, if  $F = F_0 = \mathbb{Q}$  and  $L = O_D$  for an indefinite quaternion algebra  $D$  ramified at  $\ell$ , then  $C \cong D$ ,  $G$  is the multiplicative group of  $D$ , and  $N(G(\mathbb{Z}_{\ell})) = G(\mathbb{Q}_{\ell})$ . Over  $\mathbb{C}$ , this action of  $g$  coincides with the right multiplication by  $g$  on  $Sh^{(\Sigma)}(\mathbb{C})$  in (1.2).

Recall the strict henselization  $\mathscr{W} \subset \overline{E}$  of  $\mathbb{Z}_{(p)}$ ; then,  $\mathscr{W}$  is an unramified valuation ring with residue field  $\mathbb{F} = \overline{\mathbb{F}}_p$ . Under the five conditions (H1–5) for  $\Sigma = \{p, \infty\}$ , the existence of smooth (integral) toroidal compactification of  $Sh^{(p)}$  is now known as an algebraic space after the work of Rapoport [16], Chai–Faltings [17] and Fujiwara [18] (see [4] for a very detailed proof). Although our moduli problem is slightly different from the one Kottwitz considered in [11], as was done in [2] 7.1.3, following [11] closely, the  $p$ -integral moduli  $Sh^{(p)}$  over  $O_{E,(p)}$  is proven to be a quasi-projective scheme; so, the compactified algebraic space is a projective scheme (if we choose the toroidal compactification data well). We can make the smooth toroidal compactification of  $Sh^{(p)}$  equivariant under the action of the compact group  $G(\mathbb{Z}_{\Sigma}^{(p)})$  (cf. [17] Chapter IV), and hence, for a well chosen toroidal compactification data, once a smooth compactification  $\overline{Sh}_{/\mathscr{W}}^{(p)}$  is established for  $Sh_{/\mathscr{W}}^{(p)}$ , for any  $\Sigma \supset \{p, \infty\}$ ,  $Sh_{/\mathscr{W}}^{(\Sigma)}$  has smooth compactification  $\overline{Sh}_{/\mathscr{W}}^{(\Sigma)}$  such that  $\overline{Sh}_{/\mathscr{W}}^{(\Sigma)} = \overline{Sh}_{/\mathscr{W}}^{(p)} / G(\mathbb{Z}_{\Sigma}^{(p)})_{/\mathscr{W}}$ .

We have fixed a geometrically irreducible component  $S_{/E}^{(p)}$  of  $Sh_{/E}^{(p)}$ . The pro-variety  $S_{/E}^{(p)}$  is defined over the quotient field  $\mathscr{K}$  of  $\mathscr{W}$  as a geometrically irreducible variety. Write simply  $S_{/\mathscr{K}}^{(\Sigma)}$  for the image  $S_{G(\mathbb{Z}_{\Sigma}^{(p)})}$  in  $Sh_{/\mathscr{K}}^{(\Sigma)}$ . Then we define  $S_{/\mathscr{W}}^{(\Sigma)}$  by the schematic closure of the image of  $S_{/\mathscr{K}}^{(\Sigma)}$  in  $Sh_{/\mathscr{W}}^{(\Sigma)}$ . By Zariski's connectedness theorem combined with the existence of a normal projective compactification (either minimal or smooth) of  $Sh_{/\mathscr{W}}^{(\Sigma)}$ , the reduction  $S_{/\mathbb{F}}^{(\Sigma)} = S^{(\Sigma)} \times_{\mathscr{W}} \mathbb{F}$  is a geometrically irreducible component of  $Sh_{/\mathbb{F}}^{(\Sigma)} = Sh_{/\mathscr{W}}^{(\Sigma)} \otimes_{\mathscr{W}} \mathbb{F}$ . The following result is a version of the result in [1]:

**Lemma 1.1** *Assume the five conditions (H1–5). Then the stabilizer of  $S_{\mathbb{F}}^{(\Sigma)}$  (and of  $S_{\mathcal{W}}^{(\Sigma)}$ ) in  $G(\mathbb{A}^{(\Sigma)})$  of the Shimura variety of level away from  $\Sigma$  is given by  $\mathcal{G}^{(\Sigma)}/G(\mathbb{Z}_{\Sigma}^{(p)})\overline{Z(\mathbb{Z}_{(p)})}$  for the following locally compact subgroup  $\mathcal{G}^{(\Sigma)}$  of  $G(\mathbb{A}^{(p)})$ :*

$$\mathcal{G}^{(\Sigma)} = \{x \in N(G(\mathbb{Z}_{\Sigma}^{(p)})) \times G(\mathbb{A}^{(\Sigma)}) \mid \mu(x) \in \overline{Z^G(\mathbb{Z}_{(p)})^+}\}, \quad (1.3)$$

where  $\overline{Z^G(\mathbb{Z}_{(p)})^+}$  (resp.  $\overline{Z(\mathbb{Z}_{(p)})}$ ) is the topological closure of  $Z^G(\mathbb{Z}_{(p)})^+$  (resp.  $Z(\mathbb{Z}_{(p)})$ ) in  $Z^G(\mathbb{A}^{(p)})$  (resp. in  $G(\mathbb{A}^{(p)})$ ), and  $N(G(\mathbb{Z}_{\Sigma}^{(p)}))$  is the normalizer of  $G(\mathbb{Z}_{\Sigma}^{(p)})$  in  $G(\mathbb{Q}_{\Sigma}^{(p)})$ . In particular, the action of  $G_1(\mathbb{A}^{(\Sigma)})$  leaves the irreducible component  $S_{\mathbb{F}}^{(\Sigma)}$  stable.

Since  $G(\mathbb{Z}_{\Sigma}^{(p)})$  is compact, we have  $G(\mathbb{Z}_{\Sigma}^{(p)})\overline{Z(\mathbb{Z}_{(\Sigma)})} = \overline{G(\mathbb{Z}_{\Sigma}^{(p)})Z(\mathbb{Z}_{(\Sigma)})}$  inside  $G(\mathbb{A}^{(p)})$ .

In [1], the full scheme automorphism group  $\text{Aut}(S_{\mathbb{F}}^{(\Sigma)})$  is identified with the group in the lemma (up to an explicit finite error term). Determination of the stabilizer in  $G(\mathbb{A}^{(\Sigma)})$  (not in the full automorphism group which may contain an extra automorphism outside  $G(\mathbb{A}^{(\Sigma)})$ ) is far easier, and we shall give a self-contained proof of this lemma deducing it from the characteristic 0 version in [19], though one can give an argument purely in characteristic  $p$  similar to the one in [1].

*Proof* The existence of the smooth toroidal compactification of  $Sh_{\mathcal{W}}^{(\Sigma)}$  tells us, by Zariski's connectedness theorem (e.g., [20] Corollary 11.3 in Chapter III), that the reduction map gives a bijection between  $\pi_0(Sh^{(\Sigma)}(\mathbb{C}))$  and  $\pi_0(Sh_{\mathbb{F}}^{(\Sigma)})$  for the residue field  $\mathbb{F} = \overline{\mathbb{F}}_p$  of  $\mathcal{W}$ . As already mentioned, the isogeny action  $\eta^{(\Sigma)} \mapsto \eta^{(\Sigma)} \circ g$  on  $Sh^{(\Sigma)}$  coincides with the right multiplication by  $g \in G(\mathbb{A}^{(\Sigma)})$  on the homogeneous space (1.2). Then by the strong approximation theorem applied to  $G_1(\mathbb{A}^{(\Sigma)}) \supset G_1(\mathbb{Q})$ , the homomorphism  $\mu : G \rightarrow Z^G$  induces an isomorphism  $\pi_0(Sh^{(\Sigma)}(\mathbb{C})) \cong Z^G(\mathbb{A}^{(\Sigma)})/\overline{Z^G(\mathbb{Z}_{(\Sigma)})^+}$  (see [19] and [2] 7.2.3). Since  $g \in G(\mathbb{A}^{(\Sigma)})$  acts on  $Z^G(\mathbb{A}^{(\Sigma)})/\overline{Z^G(\mathbb{Z}_{(\Sigma)})^+}$  via the right multiplication by  $\mu(g)$ , if  $\Sigma = \{p, \infty\}$ , the stabilizer of each  $x \in \pi_0(Sh^{(p)}(\mathbb{C})) \cong Z^G(\mathbb{A}^{(p)})/\overline{Z^G(\mathbb{Z}_{(p)})^+}$  is given by the group given in the lemma. Since the bijection

$$\pi_0(Sh^{(p)}(\mathbb{C})) \leftrightarrow \pi_0(Sh_{\mathbb{F}}^{(p)})$$

induced by the reduction map is  $G$ -equivariant, the stabilizer of the reduction modulo  $\mathfrak{m}_{\mathcal{W}}$  of the component  $x$  is given by the same group.

Now assume  $\Sigma \supsetneq \{p, \infty\}$ . In this case, there could be a contribution of the normalizer of  $G(\mathbb{Z}_{\Sigma}^{(p)})$  which slightly complicates the setting. The group  $N(G(\mathbb{Z}_{\Sigma}^{(p)})) \times G(\mathbb{A}^{(\Sigma)})$  acts on  $Sh^{(\Sigma)}$  naturally. If  $g \in G(\mathbb{Q}_{\Sigma}^{(p)})$  is outside the normalizer  $N(G(\mathbb{Z}_{\Sigma}^{(p)}))$ , the action of  $g$  brings  $Sh_{\mathbb{F}}^{(\Sigma)}$  isomorphically onto  $Sh^{(p)}/g^{-1}G(\mathbb{Z}_{\Sigma}^{(p)})g$ . Recall  $S^{(p)}$  (the fixed geometrically irreducible component of  $Sh_{\mathbb{F}}^{(p)}$ ). Let  $S^{(\Sigma)}$  be the irreducible component of  $Sh^{(\Sigma)}$  covered by  $S^{(p)}$ . Write  $S_g^{(\Sigma)}$  for the image of  $S^{(p)}$  in  $Sh^{(p)}/g^{-1}G(\mathbb{Z}_{\Sigma}^{(p)})g$  over  $\mathbb{F}$ . Then  $S_g^{(\Sigma)}$  is a geometrically irreducible component of  $Sh^{(p)}/g^{-1}G(\mathbb{Z}_{\Sigma}^{(p)})g$ . Note that  $S^{(p)}/S^{(\Sigma)}$  is a Galois étale covering with Galois group  $G_1(\mathbb{Z}_{\Sigma}^{(p)})$ . For  $K = G_1(\mathbb{Z}_{\Sigma}^{(p)}) \cap g^{-1}G_1(\mathbb{Z}_{\Sigma}^{(p)})g \subsetneq G_1(\mathbb{Z}_{\Sigma}^{(p)})$ ,  $S^{(p)}/K$  is a nontrivial étale covering of  $S^{(\Sigma)}$  (because the action of  $g \in G(\mathbb{Z}_{\Sigma}^{(p)})$  is an étale action). This shows that  $S^{(\Sigma)} \neq S_g^{(\Sigma)}$ . Thus if  $g \in G(\mathbb{A}^{(p)})$  stabilizes  $S^{(\Sigma)}$ ,  $g_p$  has to be in the normalizer of  $G(\mathbb{Z}_{\Sigma}^{(p)})$ . Then again by the bijection:  $\pi_0(Sh^{(\Sigma)}(\mathbb{C})) \leftrightarrow \pi_0(Sh_{\mathbb{F}}^{(\Sigma)})$ , we arrive at the assertion in the same manner as in the case where  $\Sigma = \{p, \infty\}$ .

**Remark 1.2** We have taken full polarization classes under scalar multiplication by  $O_{(p)+}^\times$  in our moduli problem (while Kottwitz's choice in [11] is a partial class of multiplication by  $\mathbb{Z}_{(p)+}^\times$ ). By our choice, the group  $G$  is the full similitude group, while Kottwitz choice is a partial rational similitude group. Our choice is convenient for our purpose because  $G$  has cohomologically trivial center, and the special fiber at  $p$  of the characteristic 0 Shimura variety  $Sh^{(p)} = Sh/G(\mathbb{Z}_p)$  gives rise to the mod  $p$  moduli of abelian varieties of the specific type we study (as shown in [2] Theorem 7.5), while Kottwitz's mod  $p$  moduli is a disjoint union of the reduction modulo  $p$  of finitely many characteristic 0 Shimura varieties associated with finitely many different pairs  $(G_i, X_i)$  with  $G_i$  locally isomorphic to each other at every place ([11] Section 8).

Hereafter, we fix  $\Sigma$  and write  $S$  for  $S^{(\Sigma)}$  if confusion is unlikely. We have a universal quadruple  $\underline{\mathbf{A}}^{(\Sigma)} = (\mathbf{A}, \bar{\lambda}, \mathbf{i}, \eta^{(\Sigma)})$  with the universal abelian scheme  $\mathbf{A}$  over  $Sh^{(\Sigma)}$ . We write  $\mathbf{A}_S = \mathbf{A} \times_{Sh^{(\Sigma)}} S$ . Suppose that the Hasse invariant  $H$  does not vanish identically on  $Sh_{\mathbb{F}}^{(\Sigma)}$  (and hence on  $S_{\mathbb{F}}$ ). Then we define the ordinary locus to be the open subscheme over which  $H$  does not vanish. We write them as  $S_{\mathbb{F}}^{\text{ord}} = S[\frac{1}{H}]$  and  $Sh_{\mathbb{F}}^{(\Sigma), \text{ord}} = Sh^{(\Sigma)}[\frac{1}{H}]$ . Over  $Sh^{(\Sigma), \text{ord}}$ , the connected component  $\mathbf{A}[p^\alpha]^\circ$  of the finite flat group scheme  $\mathbf{A}[p^\alpha] = \text{Ker}(p^\alpha : \mathbf{A} \rightarrow \mathbf{A})$  is isomorphic to  $\mu_{p^\alpha} \otimes_{\mathbb{Z}} L^\circ$  étale locally as  $O_{D,p}$ -modules for a  $O_{D,p}$ -direct summand  $L^\circ$  of  $L_p = L_{\mathbb{Z}_p}$ . The isomorphism class (as hermitian  $O_{D,p}$ -modules) of  $L_p$  is unique as explained in [2] 7.1.5 (2), which determines  $L^\circ$  as a maximally isotropic subspace (see the following paragraph for an explicit description of  $L^\circ$ ). We have  $\text{rank}_{\mathbb{Z}_p} L^\circ = \frac{1}{2} \text{rank}_{\mathbb{Z}} L$ .

Identify  $L_p$  with  $O_p^d = (O_p^n)^{d/n}$  (with column vectors in  $O_p^n$ ) so that the action of  $O_{D,p} = M_n(O_p)$  is just the left matrix multiplication on each entry in  $O_p^n$ . Pick a closed point  $x \in S^{\text{ord}}(\mathbb{F})$  carrying an abelian variety  $(A_x, \bar{\lambda}, \mathbf{i}, \eta^{(\Sigma)})$ . Then we have a canonical lift  $\underline{A}_{x/W} = (A_x, \bar{\lambda}, \mathbf{i}, \eta^{(\Sigma)})_{/W}$  with  $\text{End}(A_{x/\mathbb{F}}) = \text{End}(A_{x/W})$  for  $W = W(\mathbb{F}) = \varprojlim_{\alpha} \mathcal{W}/p^\alpha \mathcal{W}$  by the Serre–Tate deformation theory. The quadruple  $\underline{A}_{x/W}$  descends to  $\underline{A}_{x/\mathcal{W}}$  by the theory of complex multiplication ([21] and [22] combined). By the fixed embedding  $\bar{\mathbb{E}} \hookrightarrow \mathbb{C}$ , we consider  $A_{x/\mathbb{C}} = A_x \times_{\mathcal{W}} \mathbb{C}$ . Note that  $\mathcal{T}A_{x/\bar{\mathbb{E}}} \cong H_1(A_x(\mathbb{C}), \mathbb{A}^{(\infty)}) = H_1(A_x(\mathbb{C}), \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{A}^{(\infty)}$ . Then we have an isomorphism  $f : H_1(A_x(\mathbb{C}), \mathbb{Q}) \cong L_{\mathbb{Q}}$  as  $D$ -modules so that after tensoring  $\mathbb{A}^{(\Sigma)}$ ,  $\eta^{(\Sigma)}$  is induced by  $f$  and that the self-duality of  $L_p$  in (H4) is given by the Riemann form on  $L_{\mathbb{Q}}$  induced by  $f$  and  $\bar{\lambda}$  (up to  $p$ -unit multiple; see [2] 7.1.5). Identifying  $\mathcal{T}_p A_{x/\mathbb{C}}$  with  $L_p = H_1(A_x(\mathbb{C}), \mathbb{Z}_p)$ , the exact sequence  $A_{x/\mathbb{F}}[p^\infty]^\circ \hookrightarrow A_x[p^\infty] \rightarrow A_x[p^\infty]^{et}$  (under the reduction map) induces an exact sequence of  $O_{D,p}$ -modules  $L^\circ \hookrightarrow L_p \rightarrow L^{et}$ . Since  $A_x$  has complex multiplication, we can split  $L_p$  as  $L^\circ \oplus L^{et}$  so that the lift of  $p$ -power Frobenius has unit eigenvalue on  $L^{et}$  and nonunit eigenvalues on  $L^\circ$ . Then we have an isomorphism  $L^\circ \cong \prod_{\mathfrak{p}|p} M_{n,m(\mathfrak{p})}(O_{\mathfrak{p}})$  of  $O_{D,p}$ -modules (where  $O_{D,p} \cong M_n(O_p)$  acts by left matrix multiplication), and by the duality induced by the polarization, we have  $L^{et} \cong \prod_{\mathfrak{p}|p} M_{n,m(\mathfrak{p})}(O_{\mathfrak{p}})$ . The number  $(m(\mathfrak{p}), m(\bar{\mathfrak{p}}))$  gives the signature of the unitary or symplectic group  $G$  at the infinite place  $i_\infty \circ \sigma$  if  $i_p \circ \sigma$  gives the  $p$ -adic place  $\mathfrak{p} = \mathfrak{P} \cap F_0$  (by the condition (det) in [2] p. 308). The module structure of  $L^\circ$  and  $L^{et}$  under  $O_{D,p}$  and  $M = \prod_{\mathfrak{p}|p} GL_m(\mathfrak{p})(O_{\mathfrak{p}}) \subset GL_d(O_p) \cong O_{C,p}^\times$  can be made explicit as follows: Identifying  $L_{\mathfrak{p}}^{et}$  with  $n \times m(\mathfrak{p})$  matrices (suitably),  $O_{D,\mathfrak{p}} \cong M_n(O_{\mathfrak{p}})$  acts by left multiplication on  $L_{\mathfrak{p}}^{et} = M_{n,m(\mathfrak{p})}(O_{\mathfrak{p}})$ , and  $GL_m(\mathfrak{p})(O_{\mathfrak{p}})$  acts by right multiplication through the involution  $O_{C,\mathfrak{p}} \xrightarrow{\sim} O_{C^\circ,\mathfrak{p}}$ .

(P) The parabolic subgroup  $P$  in the introduction is the stabilizer in  $G(\mathbb{Z}_p)$  of the exact

sequence  $L^\circ \hookrightarrow L_p \rightarrow L^{et}$  (and is independent of the choice of  $x$ ).

We define the  $\alpha$ -th layer ( $\alpha \in \mathbb{N}$ ) of the Igusa tower over  $Sh_{/\mathbb{F}}^{(\Sigma),ord}$  by

$$Ig_\alpha = \text{Isom}_{O_D}(\mu_{p^\alpha} \otimes_{\mathbb{Z}_p} L^\circ, \mathbf{A}[p^\alpha]^\circ).$$

Note that  $O_D/p^\alpha O_D \cong M_n(O/p^\alpha O)$ , and arrange this isomorphism so that  $\rho$  corresponds to  $x \mapsto {}^t x^c$ . Fixing such an isomorphism and writing  $\varepsilon = \text{diag}[1, 0, \dots, 0]$  for the diagonal matrix in  $M_n(O_p)$  whose nonzero entry is only the top entry which is equal to 1, we have

$$Ig_\alpha = \text{Isom}_{O_D}(\mu_{p^\alpha} \otimes_{\mathbb{Z}_p} L^\circ, \mathbf{A}[p^\alpha]^\circ) \cong \text{Isom}(\mu_{p^\alpha} \otimes_{\mathbb{Z}_p} \varepsilon L_p^\circ, \varepsilon(\mathbf{A}[p^\alpha]^\circ)).$$

By the Cartier duality combined with the polarization  $\lambda$ , each isomorphism

$\eta_p^\circ \in \text{Isom}_{O_D}(\mu_{p^\alpha} \otimes_{\mathbb{Z}_p} L^\circ, \mathbf{A}[p^\alpha]^\circ)$  induces a unique isomorphism  $\eta_p : L^{et} \otimes_{\mathbb{Z}/p^\alpha \mathbb{Z}} \hookrightarrow \mathbf{A}/\mathbf{A}[p^\alpha]^\circ$ .

Thus

$$Ig_\alpha = \text{Isom}_{O_D}(L^{et}/p^\alpha L_p^{et}, \mathbf{A}[p^\alpha]^{et}) \cong \text{Isom}(\varepsilon(L_p^{et}/p^\alpha L_p^{et}), \varepsilon(\mathbf{A}[p^\alpha]^{et})).$$

By this expression,  $Ig_n$  is an étale finite covering of  $Sh_{/\mathbb{F}}^{(\Sigma),ord}$  and has a natural action of  $M = \prod_{\mathfrak{p}|p} GL_m(\mathfrak{p})(O_{\mathfrak{p}})$ , because  $\varepsilon L^{et} \cong \prod_{\mathfrak{p}} O_{\mathfrak{p}}^{m(\mathfrak{p})}$ . Here by our choice of isomorphism  $O_{D,p} \cong M_n(O_p)$  sending  $\rho$  to  $(x \mapsto {}^t x^c)$ ,  $\varepsilon$  is invariant under the Rosati involution induced by  $\bar{\lambda}$ .

Since  $M$  acts on  $Ig := Ig_\infty = \text{Isom}(\varepsilon L_p^{et} \otimes (\mathbb{Q}_p/\mathbb{Z}_p), \varepsilon \mathbf{A}[p^\infty]^{et})$  through matrix multiplication on  $\varepsilon L_p^{et} = \prod_{\mathfrak{p}|p} O_{\mathfrak{p}}^{m(\mathfrak{p})}$ ,  $M$  acts transitively on  $\pi_0(Ig_{\infty/\mathbb{F}})$  for the Igusa tower  $Ig_\infty$  over  $S$ . Fix an irreducible component  $I_{\infty/\mathbb{F}}$  of  $Ig_{\infty/\mathbb{F}}$ . Our claim is that the stabilizer in  $M$  of  $I_\infty$  contains  $M_1$  and hence  $\text{Gal}(I_\infty/S) = M_1$ .

Let  $X^+$  be the identity connected component of  $X$ . Then we may assume that  $S_K(\mathbb{C})$  is covered by  $1 \times X^+ \cong X^+$  inside  $G(\mathbb{A}^{(\infty)}) \times X$ ; so,  $S_K(\mathbb{C}) \cong \Gamma_K \backslash X^+$  for a congruence subgroup  $\Gamma_K = G(\mathbb{Q})^+ \cap \mathcal{G}^{(\Sigma)}$ . Fix a connected component  $I_{\infty/\mathbb{F}}$  of  $Ig_{\infty/\mathbb{F}}$ .

Recall a part of the axiom (A1):

(St)  $G_1(\mathbb{A}^{(\Sigma)})$  stabilizes  $I_\infty$ .

**Lemma 1.3** *Assume (St). Let  $x \in Sh^{(\Sigma)}(\mathbb{F})$  be a closed point carrying an ordinary abelian variety  $(A_{x/\mathbb{F}}, \bar{\lambda}, \eta^{(\Sigma)})$ . If  $B \subset \text{End}(A_{x/\mathbb{F}_p}) \otimes_{\mathbb{Z}} \mathbb{Q}$  is a maximal commutative semi-simple  $F$ -subalgebra stable under the Rosati involution  $j$  of  $\bar{\lambda}$ , the stabilizer  $\mathcal{D}$  in  $M_1$  of  $I_\infty$  contains a group conjugate (under an element of  $M$ ) to*

$$T_x(\mathbb{Z}_p) := \{g \in O_{B,p}^\times | gg^j = N_{B/F}(g) = 1\}$$

for  $O_{B,p} = B_p \cap \text{End}(A_{x/\mathbb{F}_p}) \otimes_{\mathbb{Z}} \mathbb{Z}_p$ .

*Proof* Take a level  $p$ -structure  $\eta_p$  of  $(A_{x/\mathbb{F}}, \bar{\lambda}, \eta^{(\Sigma)})$ , and regard  $A_x$  as sitting over  $x \in Ig_\infty(\mathbb{F})$ . The level structure  $\eta^{(\Sigma)} : L_{\mathbb{A}(p^\infty)} \cong V^{(p)}(A) \bmod G(\mathbb{Z}_\Sigma^{(p)})$  is a  $O_D$ -linear isomorphism, and  $\eta_p$  is the collection of the ordinary level  $\mathfrak{P}$ -structure  $\eta_{\mathfrak{p}} : O_{\mathfrak{p}}^{m(\mathfrak{p})} \cong \varepsilon \cdot \mathcal{S}A[\mathfrak{P}^\infty]^{et}$ . By the Cartier duality combined with the polarization, the level  $\mathfrak{P}$ -structure  $\eta_{\mathfrak{p}}$  gives  $\eta_{\mathfrak{p}}^\circ : \mu_{p^\infty} \otimes_{\mathbb{Z}} O_{\mathfrak{p}}^{m(\mathfrak{p})} \cong \varepsilon \cdot A[\mathfrak{P}^\infty]^\circ$  for the connected component  $A[\mathfrak{P}^\infty]^\circ$  of  $A[\mathfrak{P}^\infty]$ .

Slightly more generally than the setting of the lemma, let  $B \subset \text{End}(A_{x/\mathbb{F}_p}) \otimes_{\mathbb{Z}} \mathbb{Q}$  be a commutative semi-simple  $F$ -subalgebra stable under the Rosati involution  $j$  of  $\bar{\lambda}$ . Put  $O_B = B \cap \text{End}_{O_D}(A/\mathbb{F})$ , which is an order of  $B$ . Then we have an embedding  $\rho^{(\Sigma)}$  of  $B$

into  $\text{End}_D(L_{\mathbb{A}(\Sigma)}) \xrightarrow{\text{involution}} C_{\mathbb{A}(\Sigma)}^{(\Sigma)}$  given by  $\alpha \circ \eta^{(\Sigma)} = \eta^{(\Sigma)} \circ \rho^{(\Sigma)}(\alpha)$ . Though slightly more complicated, we define the  $p$ -component of the embedding as follows. Consider  $\prod_{\mathfrak{P}|p} (M_{m(\mathfrak{P})}(O_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(O_{\mathfrak{P}})) \subset \prod_{\mathfrak{P}} M_d(O_{\mathfrak{P}}) = O_{C,p}$  (embedded diagonally in blocks), where  $\mathfrak{P}$  runs over the prime factors of  $p$  in  $F$ . Thus if  $F = F_0$  or  $\overline{\mathfrak{P}} = \mathfrak{P}$ ,  $d = 2m(\mathfrak{P})$ , we have two identical copies of  $M_{m(\mathfrak{P})}(O_{\mathfrak{P}}) = M_{m(\overline{\mathfrak{P}})}(O_{\overline{\mathfrak{P}}})$  in the product (which is diagonally embedded into  $C_{\mathfrak{P}} = M_d(O_{\mathfrak{P}})$ ). If  $\mathfrak{P} \neq \overline{\mathfrak{P}}$ ,  $d = m(\mathfrak{P}) + m(\overline{\mathfrak{P}})$ , and again  $M_{m(\mathfrak{P})}(O_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(O_{\mathfrak{P}})$  is diagonally embedded in  $C_{\mathfrak{P}} = M_d(O_{\mathfrak{P}})$ . We define the embedding  $\rho_{\mathfrak{P}} : O_{B,\mathfrak{P}} = O_B \otimes_{O_0} O_{\mathfrak{P}} \hookrightarrow (M_{m(\mathfrak{P})}(O_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(O_{\overline{\mathfrak{P}}}))$  by  $\rho_{\mathfrak{P}} \times \rho_{\mathfrak{P}}^{\circ}$  given by  $\alpha \eta_{\mathfrak{P}} = \eta_{\mathfrak{P}} \circ \rho_{\mathfrak{P}}(\alpha)$  and  $\alpha \eta_{\mathfrak{P}}^{\circ} = \eta_{\mathfrak{P}}^{\circ} \circ \rho_{\mathfrak{P}}^{\circ}(\alpha)$ . We may assume that  $\rho_{\overline{\mathfrak{P}}}^{\circ} = {}^t \rho_{\mathfrak{P}}^{\circ}$ . Then we define  $\rho_p = \prod_{\mathfrak{P}} \rho_{\mathfrak{P}}$ . We put

$$\mathcal{H}_p = \rho_p(O_{B,p}^{\times}) \cap M_1 \text{ and } \mathcal{H}^{(\Sigma)} = \rho^{(\Sigma)}((O_B \otimes_{\mathbb{Z}} (\mathbb{Z}_{\Sigma}^{(p)} \times \mathbb{A}^{(\Sigma)}))^{\times}) \cap G_1(\mathbb{Z}_{\Sigma}^{(p)} \times \mathbb{A}^{(\Sigma)}).$$

The product  $\mathcal{H}_p \times \mathcal{H}^{(\Sigma)}$  is isomorphic to the group of  $(\mathbb{Z}_{\Sigma} \times \mathbb{A}^{(\Sigma)})$ -points of a reductive subgroup  $\mathcal{H}_x$  of  $G_1$  with  $\mathcal{H}_x(\mathbb{Z}_{(\Sigma)}) = (\mathcal{H}_p \times \mathcal{H}^{(\Sigma)}) \cap (O_B \otimes_{\mathbb{Z}} \mathbb{Q})$ .

Since  $M$  acts transitively on  $\pi_0(Ig_{\infty/\mathbb{F}})$ , we find  $g \in M$  such that  $x \in g \cdot I_{\infty}(\mathbb{F})$ . Each element in  $\mathcal{H}_x(\mathbb{Z}_{(\Sigma)})$  is of the form  $\rho(\alpha)$  for a prime-to- $p$  isogeny  $\alpha : A_x \rightarrow A_x$ , where  $\rho = \rho_p \times \rho^{(\Sigma)}$ . Thus  $(A_x, \overline{\lambda}, \eta \circ \rho(\alpha)) = (A_x, \overline{\lambda}, \alpha \circ \eta)$  for  $\eta = \eta_p \times \eta^{(\Sigma)}$  with  $\eta_p = \prod_{\mathfrak{P}} \eta_{\mathfrak{P}}$ , and  $\mathcal{H}_x(\mathbb{Z}_{(\Sigma)})$  fixes  $x$ . Thus  $\mathcal{H}_x(\mathbb{Z}_{(\Sigma)})$  (diagonally embedded in  $G_1(\mathbb{Z}_{\Sigma} \times \mathbb{A}^{(\Sigma)})$ ) leaves  $g \cdot I_{\infty}$  stable. Write  $\mathcal{D}$  for the stabilizer of  $I_{\infty}$  in  $M_1$ . Note that  $h \in G_1(\mathbb{A}^{(\Sigma)})$  acts on  $Ig_{\infty}$  by  $(\eta_p, \eta^{(\Sigma)}) \mapsto (\eta_p, \eta^{(\Sigma)} \circ h)$ . By (St),  $\mathcal{H}^{(\Sigma)} \subset G_1(\mathbb{A}^{(\Sigma)})$  stabilizes each element of  $\pi_0(Ig_{\infty/\mathbb{F}})$ . Then the closure of the group  $\mathcal{H}_x(\mathbb{Z}_{(\Sigma)})\mathcal{H}^{(\Sigma)}$  in  $G_1(\mathbb{A}^{(\Sigma)} \times \mathbb{Z}_{\Sigma})$  stabilizes  $g \cdot I_{\infty} \in \pi_0(Ig_{\infty/\mathbb{F}})$ . Here  $\mathcal{H}_x(\mathbb{Z}_{(\Sigma)})$  is embedded diagonally into  $G_1(\mathbb{A}^{(\Sigma)} \times \mathbb{Z}_{\Sigma})$ . The closure is equal to  $\mathcal{H}_p \times \mathcal{H}^{(\Sigma)}$ , and in particular,  $\mathcal{H}_p = \mathcal{H}_x(\mathbb{Z}_p) \subset g\mathcal{D}g^{-1}$ .

If  $B$  is maximally commutative in  $\text{End}_{O_D}(A/\mathbb{F}) \otimes_{\mathbb{Z}} \mathbb{Q}$ ,  $\mathcal{H}_x$  is a maximal torus  $T_x$  in  $G_1$ . The  $\mathbb{Z}_p$ -points of  $T_x = \mathcal{H}_x$  is given as in the lemma.

We can reverse the argument to prove the above lemma: Give ourselves a semi-simple commutative  $F$ -algebra  $B$  with nontrivial positive involution  $j$  embeddable into  $C$  such that the embedding factors through  $\prod_{\mathfrak{P}|p} (M_{m(\mathfrak{P})}(F_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(F_{\mathfrak{P}})) \subset C_p$ . We suppose that the projection of  $B_p$  to each  $M_{m(\mathfrak{P})}(F_{\mathfrak{P}})$  is maximal among commutative semi-simple  $F_{\mathfrak{P}}$ -subalgebras in  $M_{m(\mathfrak{P})}(F_{\mathfrak{P}})$ ; thus,  $\dim_F B = d$ . Since  $C = \text{End}_D(L_{\mathbb{Q}})^{\circ}$ , we have  $B \cap D = F$ . Suppose that  $B$  is stable under the involution  $*$  and that the involution  $j$  is induced by  $*$ . By the maximality of  $B_p$ ,  $B$  is its own commutant inside  $C$ . Since  $\dim_F B = d = \dim_F L_{\mathbb{Q}}$ ,  $L_{\mathbb{R}}$  is free of rank 1 over  $B_{\mathbb{R}} = B \otimes_{\mathbb{Q}} \mathbb{R}$ . Since  $j$  is positive acting nontrivially on all simple components of  $B$ , we have  $B_{\mathbb{R}} \cong F_{\mathbb{R}}^d \cong \mathbb{C}^{\lfloor F:\mathbb{Q} \rfloor d/2}$  as  $\mathbb{C}$ -algebras. The  $\mathbb{C}$ -algebra structure  $\mathbb{C} \hookrightarrow B_{\mathbb{R}}$  induces  $h_B : \mathbb{C} \hookrightarrow C_{\mathbb{R}}$ . Since  $j$  is positive, identifying  $L_{\mathbb{R}}$  with  $B_{\mathbb{R}}$ , the inner product  $(x, y) = \langle x, h_B(\sqrt{-1}y) \rangle$  is definite. Replacing  $h_B$  by its complex conjugate  $h_B \circ c$  if necessary, we may assume that  $(\cdot, \cdot)$  is positive definite; so,  $h$  satisfies (H1) (and hence in  $X$  by [11] Section 2 or [2] Lemma 7.3). Thus we have proven

**Lemma 1.4** *Let the notation and the assumption be as above. Then there exists  $x \in X$  such that  $h_B = h_x$ , replacing  $h_B$  by its complex conjugate  $h_B \circ c$  if necessary.*

Write  $j$  for the involution of  $\prod_{\mathfrak{P}|p} (M_{m(\mathfrak{P})}(F_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(F_{\mathfrak{P}}))$  sending  $(x, y) \in M_{m(\mathfrak{P})}(F_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(F_{\mathfrak{P}})$  to  $({}^t y^c, {}^t x^c) \in M_{m(\overline{\mathfrak{P}})}(F_{\overline{\mathfrak{P}}}) \oplus M_{m(\mathfrak{P})}(F_{\overline{\mathfrak{P}}})$ . Going in the reverse direction, if we start with a maximal semi-simple commutative subalgebra  $B_p \subset \prod_{\mathfrak{P}|p} (M_{m(\mathfrak{P})}(F_{\mathfrak{P}}) \oplus M_{m(\overline{\mathfrak{P}})}(F_{\mathfrak{P}}))$

stable under  $j$ , we can choose (infinitely many nonisomorphic) global  $F$ -algebra  $B$  inside  $C$  stable under  $*$  so that

- $B$  is a product of CM fields,
- $*$  induces  $j$  on  $B$ , and
- $j$  is the complex conjugation on each CM simple component of  $B$ .

For a proof of this fact, see Lemma 3.5. Then by the above argument  $L_{\mathbb{R}} \cong B_{\mathbb{R}}$  has complex structure (given by  $h$  as above), and  $L_{\mathbb{R}}/L$  gives a complex torus  $A$  with  $O_B \subset \text{End}(A)$ , where  $O_B = \{\alpha \in B \mid \alpha L \subset L\}$ . By the Riemann form  $\langle \cdot, \cdot \rangle$ ,  $A(\mathbb{C})$  is the complex manifold of the complex points of a CM abelian variety  $A$  over  $\mathbb{C}$ . We can then descend it to  $\mathscr{W}$  by the theory of complex multiplication (see [21] and [22]). By our choice,  $A/W = A/\mathscr{W} \otimes_{\mathscr{W}} W$  is ordinary; so, we can fix a level structure  $\eta_p$  and  $\eta^{(\Sigma)}$ . Taking its reduction modulo  $p$ , we get a point  $x = x(B_p)$  in  $Ig_{\infty}(\mathbb{F})$  such that  $T_{x(B_p)} \cong M_1 \cap B_p^{\times}$  and  $A_x \cong A \times_W \mathbb{F}$ . The point  $x(B_p)$  lies on one of the components  $g \cdot I_{\infty}$  of  $Ig_{\infty}$ . Replacing  $x(B_p)$  by  $g^{-1}(x(B_p))$ , we may assume that  $x(B_p) \in I_{\infty}(\mathbb{F})$ .

In the following two sections, we will prove that  $M_1$  is generated by  $T_{x(B_p)}$  where  $B_p$  runs through sufficiently many  $F_p$ -algebras. Then in the final section, we prove (St) for a suitable choice of  $\Sigma$  (see Lemma 4.2), and then we reduce the proof of Theorem 0.1 to this slightly weaker statement.

## 2 Group Theory

We prepare necessary items (for the proof of the theorem) from group theory. Fix a local field  $K/\mathbb{Q}_p$  (a finite extension) with  $p$ -adic integer ring  $O_K$  and residue field  $\kappa$ . Write  $\mathfrak{m}$  for the maximal ideal of  $O_K$ ; so,  $O_K/\mathfrak{m} = \kappa$ . Let  $\mathscr{B}$  be a complete representative set for the isomorphism classes of all commutative semi-simple algebras of dimension  $d$  over  $K$ . Fix a basis over  $O_K$  of the integral closure  $O_{B_p}$  of  $O_K$  in  $B_p \in \mathscr{B}$  and embed  $O_{B_p}$  into  $M_d(O_K)$  by the regular representation  $\rho_{B_p}$ . If confusion is unlikely, we write  $B$  for  $B_p$ .

We start with

**Lemma 2.1** *If  $B$  and  $B'$  are linearly disjoint field extensions of  $K$  of degrees  $d \leq d'$ , respectively, we have  $M_{d \times d'}(K) = \rho_B(B)x\rho_{B'}(B')$  for any  $0 \neq x \in M_{d \times d'}(K)$ , where  $M_{d \times d'}(K)$  is the ring of  $d \times d'$  matrices with coefficients in  $K$ .*

*Proof* By the assumption,  $B \otimes_K B'$  has dimension  $dd'$ . Then we can make  $M_{d \times d'}(K)$  as  $B \otimes_K B'$ -module by  $(b \otimes b')(x) = \rho_B(b)x\rho_{B'}(b')$ . Since  $B$  and  $B'$  are commutative, this gives a module structure of  $M_{d \times d'}(K)$  over the semi-simple algebra  $B \otimes_K B'$ . Since  $B$  and  $B'$  are linearly disjoint,  $B \otimes_K B'$  is a field extension of  $K$  of degree  $dd'$ . Thus  $M_{d \times d'}(K)$  is a vector space of dimension 1 over  $B \otimes_K B'$ ; in particular, any nonzero  $d \times d'$  matrix  $x$  is a basis; so, we have  $M_{d \times d'}(K) = (B \otimes_K B')x = \rho_B(B)x\rho_{B'}(B')$ .

We write  $M_d$  for  $M_{d \times d}$ .

**Lemma 2.2** *Let  $B/K$  be an unramified semi-simple extension of degree  $d$  and  $B'/K$  be a fully ramified field extension of degree  $d = d'$ . Then  $M_d(O_K) = \rho_B(O_B)\mathbf{1}_d\rho_{B'}(O_{B'})$  for the  $d \times d$  identity matrix  $\mathbf{1}_d$ , where  $\rho_B$  (resp.  $\rho_{B'}$ ) is the regular representation of  $O_B$  (resp.  $O_{B'}$ ) over  $O_K$ .*

*Proof* Split  $B = \bigoplus_i B_i$  for simple components  $B_i$ . Since we may replace  $\rho_B$  and  $\rho_{B'}$  by their conjugates  $g\rho_B g^{-1}$  and  $g\rho_{B'} g^{-1}$  for  $g \in GL_d(O_K)$ , we may assume that  $\rho_B = \bigoplus_i \rho_i$  for  $\rho_i : B_i \rightarrow M_{d_i}(K)$  for  $d_i = [B_i : K]$ . Then by Lemma 2.1, we have  $\rho_i(B_i)x_i\rho_{B'}(B') = M_{d_i \times d}(K)$ , where  $x_i$  is any non-zero  $d_i \times d$  matrix. Thus we need to show  $\rho_i(O_{B_i})x_i\rho_{B'}(O_{B'}) = M_{d_i \times d}(O_K)$  for  $x_i \in M_{d_i \times d}(O_K)$  with invertible reduction modulo the maximal ideal  $\mathfrak{m}$  of  $O_K$ . Thus we may replace  $B$  by  $B_i$  and write  $d$  for  $d_i \leq d'$ . By Nakayama's lemma, we only need to show  $\rho_B(O_B/\mathfrak{m}O_B)x\rho_{B'}(O_{B'}/\mathfrak{m}O_{B'}) = M_{d \times d'}(\kappa)$  for a nonzero element  $x \in M_{d \times d'}(\kappa)$  of rank  $d$ . Since

$$O_{B'}/\mathfrak{m}O_{B'} \cong \kappa[X]/(X^{d'})$$

for the residue field  $\kappa = O_K/\mathfrak{m}$ , over the field extension  $O_B/\mathfrak{m}O_B$  of  $\kappa$ ,  $\rho_{B'}(X^i)$  for  $i = 0, \dots, d-1$  are linearly independent. Thus by counting dimension, we find the desired identity.

**Corollary 2.3** *The Lie algebra  $\mathfrak{gl}_d(O_K) := M_d(O_K)$  is generated by  $\{\rho_B(O_B)\}_{B \in \mathcal{B}}$ .*

*Proof* We take a fully ramified field  $B'/K$  of degree  $d$  and an unramified field extension  $B/K$  of degree  $d$ . Then any element  $\mathfrak{gl}_d(O_K)$  can be written as a linear combination of  $XY$  for  $X \in \rho_B(O_B)$  and  $Y \in \rho_{B'}(O_{B'})$  (by the above lemma); so,  $\mathfrak{sl}_d(O_K) = \{x \in \mathfrak{gl}_d(O_K) \mid \text{Tr}(x) = 0\}$  is generated by

$$\{[X, Y] \mid X \in \rho_B(O_B), Y \in \rho_{B'}(O_{B'})\}.$$

Since  $\mathfrak{gl}_d(O_K) = \mathfrak{z} \oplus \mathfrak{sl}_d(O_K)$  for the center  $\mathfrak{z} \subset \rho_B(O_B) \subset M_d(O_K)$ , we get the desired result.

**Lemma 2.4** *The group  $GL_d(O_K)$  (resp.  $SL_d(O_K)$ ) is generated topologically by  $\{\rho_B(O_B^\times)\}_{B \in \mathcal{B}}$  (resp.  $\{\rho_B(\text{Ker}(N_{B/K}))\}_{B \in \mathcal{B}}$ ), where  $N_{B/K} : O_B^\times \rightarrow O_K^\times$  is the norm map.*

*Proof* Let  $H \subset GL_d(O_K)$  be the closed subgroup topologically generated by  $\{\rho_B(O_B^\times)\}_{B \in \mathcal{B}}$ . The Lie algebra  $\mathfrak{gl}_d(O_K)$  is generated by  $\{\rho_B(O_B)\}_{B \in \mathcal{B}}$  by Corollary 2.3, because we have an unramified extension  $B/K$  of degree  $d$  and a fully ramified extension  $B'/K$  of degree  $d$  in  $\mathcal{B}$ . Thus applying the  $p$ -adic exponential map, we find that an open subgroup  $S \subset GL_n(O_K)$  is contained in  $H$ . If  $p$  is unramified in  $K/\mathbb{Q}_p$ ,  $\exp$  converges on  $p \cdot \mathfrak{gl}_d(O_K)$  and hence, we may assume that  $S = 1_d + p \cdot \mathfrak{gl}_d(O_K)$ ; so,  $GL_n(O_K)/S = GL_d(\kappa)$  for the residue field  $\kappa$  of  $O_K$ .

Now we show that  $GL_n(\kappa)$  is generated by  $\rho_B((O_B/\mathfrak{m}O_B)^\times)$  for  $B$  running through  $\mathcal{B}$ . It is enough to show that the image  $\overline{H}$  of  $H$  in  $GL_d(\kappa)$  intersects nontrivially with every conjugacy class of  $GL_d(\kappa)$ . If  $\overline{x} \in GL_n(\kappa)$  is a unipotent element, we may assume that  $\overline{x}$  is in the Jordan canonical form. Then it is easy to find semi-simple ramified extension  $B \in \mathcal{B}$  such that the image of  $O_B^\times$  in  $GL_n(\kappa)$  contains a conjugate of  $\overline{x}$ . If  $\overline{x} \in GL_n(\kappa)$  is semisimple, we can of course find an unramified  $B \in \mathcal{B}$  such that the image of  $O_B^\times$  in  $GL_n(\kappa)$  contains a conjugate of  $\overline{x}$ . Thus  $\overline{H} = GL_d(\kappa)$ . This shows the result for  $GL_d(O_K)$  if  $K/\mathbb{Q}_p$  is unramified.

If  $K/\mathbb{Q}_p$  ramifies, by the above proof in the unramified case, we have  $H \supset GL_d(\mathbb{Z}_p)$ . Since  $H$  is  $O_K$ -Lie group, for any unipotent subgroup  $U \subset GL_d(K)$  generated by a root, the Lie algebra of  $U \cap H$  is an  $O_K$ -module. Since  $H \supset GL_d(\mathbb{Z}_p)$ , any simple roots are in  $U \cap H$ ; so, any integral upper triangular unipotent elements are in  $H$ . Since  $H$  contains an open subgroup  $S$ , by [1] Proposition 3.2,  $H$  contains  $SL_d(O_K)$ . Taking an unramified extension  $B/K$ , we find that  $N_{B/F}(O_B^\times) = O_K^\times$ ; so,  $\det(H) = O_K^\times$ , and we have  $H = GL_d(O_K)$ .

The above proof for  $GL_d$  can be performed also for  $SL_d$  without much modification, and the work is left to the reader.

Let  $K$  now be a semi-simple commutative algebra over  $\mathbb{Q}_p$ . We take a complete set  $\mathcal{B}$  of representatives of isomorphism classes of  $K$ -algebras  $B$  free of rank  $d$  over  $K$ . We fix a regular representation  $\rho_B : O_B \rightarrow M_d(O_K)$  for each  $B \in \mathcal{B}$ . Since  $GL_d(O_K) = \prod_i GL_d(O_{K_i})$  and  $SL_d(O_K) = \prod_i SL_d(O_{K_i})$  for simple components  $K_i$  of  $K$ , we have the following obvious version of the above lemma for the semi-simple base ring  $K$ :

**Corollary 2.5** *The group  $GL_d(O_K)$  (resp.  $SL_d(O_K)$ ) is generated topologically by  $\{\rho_B(O_B^\times)\}_{B \in \mathcal{B}}$  (resp.  $\{\rho_B(\text{Ker}(N_{K/B}))\}_{B \in \mathcal{B}}$ ), where  $N_{B/K} : O_B^\times \rightarrow O_K^\times$  is the norm map.*

### 3 Algebraic Number Theory

We prepare necessary items (for the proof of the theorem) from algebraic number theory. In this section,  $K$  is a local field. We start with the following well-known fact:

**Lemma 3.1** *Let  $K/k/\mathbb{Q}_p$  be finite field extensions. Then the  $p$ -adic integer ring  $O_K$  is generated by a single element over the  $p$ -adic integer ring  $O_k$ .*

*Proof* First suppose that  $K/k$  is fully ramified. Let  $\varpi$  be the uniformizer of  $K$  and  $\mathbb{F}_q$  be the common residue field of  $O_K$  and  $O_k$ . Let  $R = \mu_{q-1} \sqcup \{0\} \subset O_k$  be the Teichmüller lift of  $\mathbb{F}_q$ . Then by  $\varpi$ -adic expansion, any  $x \in O_K$  can be written as  $\sum_{n=0}^{\infty} a_n \varpi^n$  with  $a_n \in R$ . Thus  $O_K = O_k[\varpi] = O_k[\varpi + a]$  for any  $a \in O_k$ . In particular, we may assume that the generator is a unit in  $O_K$ . If  $K/k$  is unramified, writing  $Q = |k|$  (so,  $\kappa = \mathbb{F}_Q$ ) and taking an element  $\theta_0 \in O_K$  whose image in  $\kappa$  generate  $\kappa$  over the residue field  $\mathbb{F}_q$  of  $O_k$ , we have  $O_k[\theta_0] = O_K$ . In general, we choose  $\theta_0$  as above in the maximal unramified extension  $K^{ur}/k$  inside  $K$  so that  $O_{K^{ur}} = O_k[\theta_0]$ . Taking the fixed subfield  $K^r$  in  $K$  of any lift of the  $Q$ -power Frobenius map in  $\text{Gal}(K^{gal}/k)$  for the Galois closure  $K^{gal}$  of  $K/k$ , we find a subfield  $K^r/k$  of  $K$  fully ramified. We choose a uniformizer  $\varpi$  of  $K^r$ . Put  $\theta = \varpi + \theta_0$ . Since  $O_K = O_{K^r} \otimes_{O_k} O_{K^{ur}}$ , the different of  $\theta$  and the different of  $\varpi$  over  $K^{ur}$  generate the same ideal, and  $\theta$  generates  $O_K$  over  $O_k$ .

**Corollary 3.2** *For any semi-simple commutative extension  $B/K$ ,  $B$  is generated over  $K$  by a single element, and  $O_B$  is generated at most by  $m$  elements if  $m$  is the number of simple components of  $B$ . Moreover, we can choose an element  $\xi \in O_B$  so that  $O_K[\xi]$  projects surjectively to the  $p$ -adic integer ring of each simple component of  $B$ .*

*Proof* Let  $B_i$  be a simple component of  $B$ . Choose a generator  $\theta_i$  of  $O_{B_i}$  over  $O_K$  by the above lemma, then  $\theta = (\theta_i) \in B$  generates an order of  $B$  whose projection to any simple component is maximal. In particular,  $\theta$  generate  $B$  over  $K$ . The  $m$  elements  $\theta_i$  generates  $O_B$  over  $O_k$ .

We take the central simple algebra  $D/F$  and projective  $O_D$ -module  $L$  as in Section 1, and recall that  $C$  is the opposite algebra of  $\text{End}_D(L_{\mathbb{Q}})$ . Recall  $d^2 = \dim_F C$ . Since  $D_{\mathfrak{p}} \cong M_n(F_{\mathfrak{p}})$ , for each prime factor  $\mathfrak{p}|p$ ,  $V_{\mathfrak{p}} = L \otimes_{O_D} F_{\mathfrak{p}} \cong M_{n \times m_{\mathfrak{p}}}(F_{\mathfrak{p}})$  on which  $D_{\mathfrak{p}} = M_n(F_{\mathfrak{p}})$  acts by right multiplication. Since  $L_{\mathbb{Q}}$  is  $D$ -projective,  $m = m_{\mathfrak{p}}$  is independent of  $\mathfrak{p}|p$ . Then we have  $\text{End}_D(M_{n \times m}(F_{\mathfrak{p}})) \cong M_m(F_{\mathfrak{p}}) = C_{\mathfrak{p}}^{\circ}$ , and  $m = d$ .

We take a semi-simple commutative  $F_p$ -algebra  $B_p$  with an involution  $j$  inducing  $c$  on  $F$  (we make a convention that  $c$  is the identity map if  $F = F_0$ ). We assume that

- (I0)  $B_p$  is free of rank  $d$  over  $F_p$ , where  $d^2 = \dim_F C$ ;
- (I1)  $j$  does not fix any simple component of  $B_p$ .

If  $\mathfrak{p}$  is a prime factor of  $p$  in  $F_0$  splitting into  $\mathfrak{p} = \mathfrak{P}\overline{\mathfrak{P}}$ , then  $B_{\mathfrak{p}} = B_{\mathfrak{P}} \oplus B_{\overline{\mathfrak{P}}}$  and without

assuming (I1),  $j$  interchanges  $B_{\mathfrak{P}}$  and  $B_{\overline{\mathfrak{P}}}$ , since  $j = c$  on  $F$ . If  $F = F_0$  or  $\mathfrak{P}|\mathfrak{p}$  for a prime  $\mathfrak{P}$  with  $\mathfrak{P}^c = \mathfrak{P}$ , we have  $d = 2m(\mathfrak{P})$ . In such a case, (I0–1) implies

(I2) The  $\mathfrak{p}$ -component  $B_{\mathfrak{p}}$  is isomorphic to  $B_{\mathfrak{P}} \oplus B_{\mathfrak{P}}$  for two copies of semi-simple commutative algebra  $B_{\mathfrak{P}}$  interchanged by  $j$  (and  $\dim_{F_{\mathfrak{p}}} B_{\mathfrak{P}} = m(\mathfrak{P})$ ).

We write  $\mathcal{B}$  for the isomorphism classes of commutative semi-simple algebras  $B_p$  over  $F_p$  with involution  $j$  satisfying the conditions (I0–1) (and hence also (I2)).

**Lemma 3.3** *If  $B_p$  satisfies (I0–1), then  $B_p$  can be embedded into  $C_p = M_d(F_p)$  so that the image is stable under  $*$  and  $j$  coincides with  $*$  on  $B_p$ .*

*Proof* Let  $J \in GL_d(F_p)$  be defined as follows:  $J_{\mathfrak{p}} = 1_d$  if  $\mathfrak{p} = \mathfrak{P}\overline{\mathfrak{P}}$  with  $\mathfrak{P} \neq \overline{\mathfrak{P}}$  in  $F$ , and  $J_{\mathfrak{p}} = \begin{pmatrix} 0 & 1_{m(\mathfrak{P})} \\ 1_{m(\mathfrak{P})} & 0 \end{pmatrix}$  if  $\mathfrak{P}|\mathfrak{p}$  is a prime factor of  $\mathfrak{p}$  in  $F$  with  $\mathfrak{P} = \overline{\mathfrak{P}}$ . Since  $C_p = M_d(F_p)$  has two involutions  $*$  and  $x \mapsto x^\delta := J^t x^c J^{-1}$ , the automorphism  $x \mapsto (x^*)^\delta$  is an inner conjugation by the theorem of Skolem–Noether. Thus modifying the identification, we may assume that  $x^* = x^\delta$ . If  $\mathfrak{p} = \mathfrak{P}\overline{\mathfrak{P}}$  ( $\mathfrak{P} \neq \overline{\mathfrak{P}}$ ), we embed  $B_{\mathfrak{P}}$  into  $M_d(F_{\mathfrak{P}})$  by a regular representation  $\rho_{\mathfrak{P}}$  over  $F_{\mathfrak{P}}$  and  $B_{\overline{\mathfrak{P}}}$  into  $M_d(F_{\overline{\mathfrak{P}}})$  by  $x \mapsto {}^t \rho_{\mathfrak{P}}(x^j)^c$ . Thus  $B_{\mathfrak{p}} = B_{\mathfrak{P}} \oplus B_{\overline{\mathfrak{P}}}$  is embedded into  $C_{\mathfrak{p}}$  as desired. If  $\mathfrak{P} = \overline{\mathfrak{P}}$  or  $F = F_0$ ,  $B_{\mathfrak{p}} = B_{\mathfrak{P}} \oplus B_{\mathfrak{P}}$  as in (I2). We take a regular representation  $\rho_{\mathfrak{P}} : B_{\mathfrak{P}} \rightarrow M_{m(\mathfrak{P})}(F_{\mathfrak{P}})$  of the left component  $B_{\mathfrak{P}}$  and embed  $(x, y) \in B_{\mathfrak{P}} \oplus B_{\mathfrak{P}}$  into  $C_{\mathfrak{p}} = M_{2m(\mathfrak{P})}(F_{\mathfrak{P}})$  by  $(x, y) \mapsto \begin{pmatrix} \rho(x) & 0 \\ 0 & {}^t \rho(y^j)^c \end{pmatrix} \in M_d(F_{\mathfrak{P}}) = C_{\mathfrak{p}}$ . This embedding satisfies the desired property.

For each  $B_p \in \mathcal{B}$ , we embed  $B_p$  into  $C_p$  as in Lemma 3.3. Then  $B_p$  is a maximal commutative semisimple subalgebra of  $C_p$ , and we have  $V_p = V \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong D_p \otimes_{F_p} B_p$ . Indeed, picking  $v \in V_p$  which is not killed by any simple component of  $D_p$  and  $B_p$ , we have an injection  $D_p \otimes_{F_p} B_p \rightarrow V_p$  given by  $d \otimes b \mapsto dbv$ . Since  $\dim_{F_p} V_p = nd = \dim_{F_p} D \otimes_F B_p$ , we have the isomorphism  $V_p \cong D_p \otimes_{F_p} B_p$ .

**Lemma 3.4** *Choosing an identification  $V_p \cong D_p \otimes_{F_p} B_p$  as above, we can find  $S_D \in D_p^\times$  and  $S_B \in B_p^\times$  with  ${}^t S_D^c = \pm S_D$  and  ${}^t S_B^c = \mp S_B$  such that the involution  $\rho \otimes *$  on  $D_p \otimes_{F_p} B_p$  is the adjoint involution with respect to the pairing given by*

$$(d \otimes b, d' \otimes b') \mapsto \mathrm{Tr}_{F/F_0}(\mathrm{Tr}_{D_p/F_p}({}^t d^c S_D d') \mathrm{Tr}_{B_p/F_p}(b^c S_B b'))$$

for  $b, b' \in B_p$  and  $d, d' \in D_p$ .

In other words, to define the algebraic group  $G/\mathbb{Q}_p$ , we may assume that the original pairing  $\langle \cdot, \cdot \rangle$  on  $V_p$  is given by the pairing in the lemma after identifying  $V_p$  with  $D_p \otimes_{F_p} B_p$ .

*Proof* Since  $B_p/F_p$  is semi-simple, identifying  $D_p \otimes_{F_p} B_p$  with  $M_n(B_p)$ , we can write  $\langle x, y \rangle = \mathrm{Tr}_{B_p/F_0, p} \mathrm{Tr}_{D_p \otimes_{F_p} B_p/B_p}({}^t x^c a y)$  for  $a \in D_p \otimes_{F_p} B_p$ , where writing  $x \in D_p \otimes_{F_p} B_p$  as  $(x_{ij}) \in M_n(B_p)$ ,  $x \mapsto {}^t x^c = (x_{ji}^c)$  is the involution  $\rho \otimes *$  of  $D_p \otimes_{F_p} B_p$ . Since  $\langle \cdot, \cdot \rangle$  is alternating, we have  ${}^t a^c = -a$ . Then  $x^{\rho \otimes *} = {}^t a^{-c} x^c a^c$ . Since  $D_p \cong M_n(F_p)$ ,  $x \mapsto {}^t (x^\rho)^c$  for the generator  $c$  of  $\mathrm{Gal}(F/F_0)$  is a central automorphism of  $D_p$ . Thus by the theorem of Skolem–Noether, it is an inner automorphism, and  $x^\rho = s^t x^c s^{-1}$  for a unique  $s \in D_p^\times / F_{0, p}^\times$  with  ${}^t s^c = \pm s$ . Applying the same argument to  $D_p \otimes_{F_p} B_p$  and  $x \mapsto x^*$ , we find

$$x^{\rho \otimes *} = S^t x^c S^{-1}$$

for  $S \in (D_p \otimes_{F_p} B_p)^\times / B_0^\times$  with  ${}^t S^c = \pm S$ , where  $B_0 \subset B_p$  is the subalgebra fixed by  $c$ . Thus  $S \equiv {}^t a^{-c} = -a^{-1} \pmod{B_0^\times}$  and  $S \equiv s \otimes 1 \pmod{B_0^\times}$ . Since  $S \equiv s \alpha \otimes \alpha^{-1} \pmod{B_0^\times}$  for  $\alpha \in F_p^\times$ ,

we may write  $S = S_D \otimes S_B \pmod{B_0^\times}$  with  $S_D \in D_p^\times$  and  $S_B \in B_p^\times$  such that  ${}^t S_D^c = \pm S_D$  and  ${}^t S_B^c = \mp S_B$ . Then, to define  $G(\mathbb{Q}_p)$ , we may assume  $\langle x, y \rangle = \text{Tr}_{D_p \otimes B_p / F_p}({}^t x^c S_D \otimes S_B y)$  on  $D_p \otimes_{F_p} B_p = V_p$ , because our group  $G$  only depends on the involution  $*$ . In particular, we may assume that  $\langle d \otimes b, d' \otimes b' \rangle = \text{Tr}_{F/F_0}(\text{Tr}_{D_p/F_p}({}^t d^c S_D d') \text{Tr}_{B_p/F_p}(b^c S_B b'))$  for  $b, b' \in B$  and  $d, d' \in D$ .

**Lemma 3.5** *Let  $\mathcal{B}$  be a complete set of representatives for isomorphism classes of semi-simple commutative  $F_p$ -algebras  $B_p$  with involution  $j$  satisfying (I0–1). If  $O_{D,p} \cong M_n(O_p)$ , for any  $B_p \in \mathcal{B}$ , there exist infinitely many semisimple commutative  $F$ -algebras  $B$  inside  $C$  such that*

- (1)  $O_{B,p} \hookrightarrow O_{C,p}$  for the integral closure  $O_B$  of  $O$  in  $B$ ,
- (2)  $B \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong B_p$ ,
- (3)  $B$  is stable under the involution  $*$  which induces the involution  $j$  on  $B_p$ ,
- (4)  $B$  is a product of CM fields with  $j$  inducing the complex conjugation on each simple component of  $B$ .

For each infinite place  $v$  of  $F$ , the  $\mathbb{C}$ -algebra structure  $\mathbb{C} \rightarrow B_{\mathbb{R}}$  induces  $h = h_B : \mathbb{C} \hookrightarrow C_{\mathbb{R}}$  satisfying (H1), which gives rise to a CM point  $x \in X$ .

*Proof* The last assertion follows from the first by Lemma 1.4.

By Lemma 3.3, any  $B_p \in \mathcal{B}$  as above can be embedded into  $C_p$  stable under  $*$  which induces  $j$ . For each infinite place  $v$  of  $F_0$ ,  $(D_v = D \otimes_{F_0} F_{0,v}, \rho)$  is isomorphic to  $(M_n(\mathbb{R}), x \mapsto {}^t x)$  (in the type C case) and is isomorphic to  $(M_n(\mathbb{C}), x \mapsto {}^t \bar{x})$  (in the type A case; see [11] Lemma 2.11). In these two cases, we have

$$(C_v, *) \cong \begin{cases} (M_d(\mathbb{C}), x \mapsto I {}^t \bar{x} I^{-1}) & \text{in the type A case,} \\ (M_d(\mathbb{R}), x \mapsto J {}^t x J^{-1}) & \text{in the type C case,} \end{cases}$$

where  $I = \begin{pmatrix} 1_m & 0 \\ 0 & -1_{m'} \end{pmatrix}$  and  $J = \begin{pmatrix} 0 & -1_\delta \\ 1_\delta & 0 \end{pmatrix}$  for  $\delta = d/2$ . See [11] Section 4 or [23] Sections 1 and 2 for these facts. If  $F_0 \neq F$  and  $v$  corresponds to a  $p$ -adic place  $\mathfrak{P}$  by  $i_p$  and  $i_\infty$ , we have  $(m, m') = (m(\mathfrak{P}), m(\overline{\mathfrak{P}}))$ , and if  $F = F_0$  with  $v$  corresponding to  $\mathfrak{P}$ , we have  $\delta = d/2 = m(\mathfrak{P})$  (see [2] Lemma 8.10). Thus we may choose a maximal semi-simple commutative  $F_v$ -algebra  $B_v \cong \mathbb{C}^s$  inside  $C_v$  so that  $B_v$  is stable under  $*$  and  $*$  is complex conjugation on each factor  $\mathbb{C}$  of  $B_v$ . Indeed, in the type A case, we may choose  $(B_v, j)$  to be the algebra of diagonal matrices in  $(M_d(\mathbb{C}), x \mapsto I {}^t \bar{x} I^{-1})$  and  $F_v = \mathbb{C}$  is embedded into  $B_v$  by  $z \mapsto \begin{pmatrix} z & 0 \\ 0 & \bar{z} \end{pmatrix}$ . In the type C case, we choose  $(B_v, j) \subset (M_d(\mathbb{R}), x \mapsto J {}^t x J^{-1})$  so that  $a + b\sqrt{-1} \in \mathbb{C}^\delta$  with  $a, b \in \mathbb{R}^\delta$  is sent to  $\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in M_d(\mathbb{R})$ . Then  $\{x \in B_v^\times | xx^* \in F_{0,v}^\times\} / F_v^\times$  is compact, the  $\mathbb{C}$ -algebra structure of  $B_\infty = \prod_v B_v$  induces  $h : \mathbb{C} \hookrightarrow C_{\mathbb{R}}$  in  $X$ , and the involution  $*$  induces a positive involution on  $B_v$ .

We write  $B_\infty^-$  for the open subset of  $\{x \in B_\infty | x^* = -x\}$  made up of elements  $x$  generating  $B_\infty$  over  $F_\infty$ . Similarly, at  $p$ , we write  $B_p^-$  for the open subset of the set  $\{x \in O_{B,p} | x^* = -x\}$  made up of elements  $x$  generating  $B_p$  over  $F_p$ . By Corollary 3.2,  $B_\ell^-$  for  $\ell = p, \infty$  is nonempty. We put  $C^- = \{x \in O_{C,(\Sigma)} | x^* = -x\}$  which is an  $O_{0,(\Sigma)}$ -free module. Choose a norm  $|\cdot|_\ell$  on  $C_\ell^-$  for each  $\ell \in \{p, \infty\}$  so that  $C_\ell^-$  is the completion of  $C^-$  under  $|\cdot|_\ell$  and  $|\cdot|_\ell$  induces the standard absolute value on  $F_\ell$ . Pick a generator  $g_\ell \in B_\ell^-$  of  $B_\ell$  over  $F_\ell$  for each  $\ell \in \{p, \infty\}$  (this is possible by Corollary 3.2). Since  $C^- \cong C^- + (C_{\mathbb{A}}^-)^{(\Sigma)} / (C_{\mathbb{A}}^-)^{(\Sigma)} \subset C_\Sigma^-$  is dense under these norms, for any given  $\epsilon > 0$ , we can think of an infinite set

$$U_\epsilon = \{x \in C^- \mid |x_\ell - g_\ell|_\ell < \epsilon \text{ for all } \ell \in \{p, \infty\}\}.$$

Then  $\{F[\xi]\}_{\xi \in U_\epsilon}$  contains infinitely many isomorphism classes of  $F$ -algebras (because we may choose a prime  $\mathfrak{q}$  outside  $\Sigma$  and  $\xi, \xi' \in U_\epsilon$  so that  $F_{\mathfrak{q}}[\xi] \not\cong F_{\mathfrak{q}}[\xi']$  (and moves  $\mathfrak{q}$  around). If  $\epsilon$  is sufficiently small, any  $x \in U_\epsilon$  generates an extension of  $F_\ell$  isomorphic to the given  $B_\ell$  for all  $\ell \in \{p, \infty\}$  (by [24] XI.3 Lemma 1).

Choose  $\epsilon$  small so that  $(F[\xi]_p, *) \cong (B_p, j)$  and  $(F[\xi]_\infty, *) \cong (B_\infty, j)$ . We now show that we can modify  $\xi$  so that  $O_{B,p} \hookrightarrow O_{C,p}$  for  $F[\xi]$ . Since  $O_C$  is maximal, we may replace  $V_p = L \otimes_{\mathbb{Z}} \mathbb{Q}_p$  by  $\varepsilon V_p \cong F_p^d$  and regard  $C_p = \text{End}_{F_p}(\varepsilon V_p)$  for the idempotent  $\varepsilon \in D_p$ . Then  $G(\mathbb{Q}_p)$  is the unitary group of the hermitian vector space  $\varepsilon \cdot V_p$ . Pick  $\xi \in U_\epsilon$ , and consider  $F[\xi] \subset C$ . Write  $O_{F[\xi]}$  for the maximal order of  $F[\xi]$ . Then  $O_{F[\xi]}$  is contained a maximal order  $R$  of  $C_\Sigma$ . Since  $O_{C,p}$  is maximal in  $C_p$ ,  $R = gRg^{-1} = O_{C,p}$  for  $g \in C_p^\times$ . We want to show that we can take  $g$  inside  $G(\mathbb{Q}_p)$ . If this is the case, by conjugating  $F[\xi]$  in  $C$  by  $g$ , we have the desired subalgebra  $gF[\xi]g^{-1}$  with  $gO_{F[\xi]}g^{-1} \subset O_{C,p}$  whose involution  $c$  is induced by  $*$ . Let  $L'$  be the  $O_{F[\xi]}$ -span of  $\varepsilon L_p$ :  $L' = \sum_{\alpha \in O_{F[\xi]}} \alpha(\varepsilon L_p)$ , which is a lattice in  $\varepsilon V_p \cong \varepsilon(D_p \otimes_F F[\xi]) = F_p[\xi]$  stable under  $O_{F_p[\xi]}$ . By Lemma 3.4,  $G/\mathbb{Q}_p$  is defined with respect to the pairing on  $V_p$  of the following form:

$$\langle f, f' \rangle = \text{Tr}_{F[\xi]/F_0}(f^c \delta f')$$

for  $f, f' \in F_p[\xi]$  with  $\delta = -\delta^c \in F[\xi]_p^\times$ . Since the different of  $O_{F_p[\xi]}$  is principal (and  $L_p$  is self dual),  $\varepsilon L_p$  and  $L'$  are maximal lattices in the sense of [25] Section 4.7. Then by Lemma 5.9 in [25], we find  $g \in G(\mathbb{Q}_p)$  such that  $gL' = L$ . Since  $G(\mathbb{Q})$  is dense in  $G(\mathbb{Q}_p)$ , take  $\alpha \in G(\mathbb{Q})$  sufficiently close to  $g$  in  $B_p$  and replacing  $\xi$  by  $\alpha\xi\alpha^{-1}$ , we get  $B = \alpha F[\xi]\alpha^{-1}$  satisfying  $O_B \subset O_{C,p} = \{x \in C \mid xL_p \subset L_p\}$ .

#### 4 Irreducibility

We are now ready to prove Theorem 0.1. Let  $\mathcal{S}_0$  be the set of prime ideals of  $F_0$  over  $p$  and  $\mathcal{S}$  be the set of all prime factors of  $p$  in  $F$ . If  $F \neq F_0$ , we can decompose  $\mathcal{S} = \mathcal{S}^{sp} \sqcup \mathcal{S}^i$  so that  $\mathcal{S}^{sp}$  is made up of all prime factors of  $p$  split over  $F_0$ . We split  $\mathcal{S}^{sp} = \mathcal{S}_{sp} \sqcup \overline{\mathcal{S}}_{sp}$  so that  $\overline{\mathcal{S}}_{sp} = \{\mathfrak{P}^c \mid \mathfrak{P} \in \mathcal{S}_{sp}\}$ . If  $F \neq F_0$ , we fix a splitting:  $\text{Hom}_{\text{field}}(F, \overline{E}) = \Phi \sqcup \overline{\Phi}$  compatible with  $\mathcal{S}^{sp} = \mathcal{S}_{sp} \sqcup \overline{\mathcal{S}}_{sp}$ . In other words,  $i_p \circ \sigma$  induces a place  $\mathfrak{P}_\sigma$  in  $\mathcal{S}_{sp}$  if and only if  $\sigma \in \Phi$  (as long as  $i_p \circ \sigma$  induces a place in  $\mathcal{S}^{sp}$ ). For the place in  $\mathcal{S}^i$ , we just choose splitting  $\Phi$  as above arbitrarily. Then  $\sum_{\sigma \in \Phi} i_\infty \circ \sigma$  is a CM type of  $F$ . The  $p$ -adic group  $G_1(\mathbb{Q}_p)$  is inside  $C_p^\times \cong M_d(O_p)$ , and if  $F \neq F_0$ , the projection  $\pi_\Phi : M_d(O_p) \rightarrow \prod_{\mathfrak{P} \in \mathcal{S}_{sp} \sqcup \mathcal{S}^i} M_d(O_{\mathfrak{P}})$  induces an isomorphism of  $G_1(\mathbb{Q}_p) \subset M_d(O_p)$  onto its image. We often regard  $G_1(\mathbb{Q}_p) \subset \prod_{\mathfrak{P} \in \mathcal{S}_{sp} \sqcup \mathcal{S}^i} GL_d(O_{\mathfrak{P}})$  if  $F \neq F_0$ . Hereafter we often write  $\overline{X}$  for the conjugate  $X^c$  for an object  $X$  defined over  $F$  (for example,  $\overline{\mathfrak{P}} = \mathfrak{P}^c$ ). In particular, we often write  $\mathfrak{P}$  for a prime in  $\mathcal{S}_{sp}$  above  $\mathfrak{p} \in \mathcal{S}_0$  and hence  $\overline{\mathfrak{P}} \in \overline{\mathcal{S}}_{sp}$ . Decompose  $G(\mathbb{Z}_p) = \prod_{\mathfrak{p} \in \mathcal{S}_0} G_{\mathfrak{p}}$  with

$$G_{\mathfrak{p}} = \{x \in O_{C,\mathfrak{p}}^\times \mid gg^* \in O_{0,\mathfrak{p}}^\times\}.$$

Similarly we can decompose  $G_1(\mathbb{Z}_p) = \prod_{\mathfrak{p} \in \mathcal{S}_0} G_{\mathfrak{p}}^1$  for  $G_{\mathfrak{p}}^1 = G_1(\mathbb{Z}_p) \cap G_{\mathfrak{p}}$ .

Recall the universal abelian scheme  $\mathbf{A}$  over the ordinary locus of  $Sh^{(p)}$ . For each  $\mathfrak{p} \mid p$  in  $F_0$ , we consider the exact sequence  $\mathbf{A}[\mathfrak{p}^\infty]^\circ \rightarrow \mathbf{A}[\mathfrak{p}^\infty] \rightarrow \mathbf{A}^{et}$ . The collection of  $g \in G_{\mathfrak{p}}$  whose action on the level  $p$ -structure preserves this exact sequence forms a maximal parabolic subgroup  $P_{\mathfrak{p}}$  of  $G_{\mathfrak{p}}$  (as seen in (P) in Section 1). We put  $P = \prod_{\mathfrak{p} \in \mathcal{S}_0} P_{\mathfrak{p}}$ . The Levi subgroup of  $P \cap G_1(\mathbb{Z}_p)$  is

isomorphic to  $M_1$ . Here for each  $\mathfrak{p} \in \mathcal{S}_0$ , if  $\mathfrak{p} = \mathfrak{P}\overline{\mathfrak{P}}$  in  $F$  with  $\mathfrak{P} \in \mathcal{S}_{sp}$ ,  $(g_{\mathfrak{P}}, g_{\overline{\mathfrak{P}}})$  is embedded into  $P$  in the upper triangular form via

$$(g_{\mathfrak{P}}, g_{\overline{\mathfrak{P}}}) \mapsto \begin{pmatrix} g_{\mathfrak{P}} & * \\ 0 & {}^t g_{\overline{\mathfrak{P}}}^{-1} \end{pmatrix}.$$

If  $\mathfrak{P} \in \mathcal{S}^i$ , the embedding is given by

$$g_{\mathfrak{P}} \mapsto \begin{pmatrix} g_{\mathfrak{P}} & * \\ 0 & {}^t \overline{g_{\mathfrak{P}}}^{-1} \end{pmatrix}$$

identifying the involution  $*$  with  $x \mapsto {}^t \overline{x} := {}^t x^c$ .

As mentioned in the introduction, our strategy to prove Theorem 0.1 is to produce sufficiently many closed points  $x \in I_{\infty}(\mathbb{F})$  whose stabilizers in the Levi subgroup  $M$  of  $P$  span the group  $M_1$  in order to show  $M_1 \subset \text{Gal}(I_{\infty}/S)$ . Note that  $\text{End}_{O_D}(A_x) \otimes \mathbb{Q}$  is a semi-simple  $F$ -algebra for each point  $x \in I$ . We will therefore first list commutative semi-simple algebras  $B_p/F_p$  inside  $C_p$  which is the  $p$ -completion of a global semi-simple algebra  $B$  embeddable into  $\text{End}_{O_D}(A_x) \otimes \mathbb{Q}$  for a good choice of  $x$ ; thereby, we produce maximal commutative subgroups  $T_x$  (indexed by a suitable finite set of points  $x \in I$ ) whose  $\mathbb{Z}_p$ -points taken modulo  $G_1(\mathbb{A}^{(\Sigma)})$  generate  $M_1(\mathbb{Z}_p)$ . Since  $B_p = \prod_{\mathfrak{P} \in \mathcal{S}} B_{\mathfrak{P}}$ , we shall describe the possible algebras  $B_{\mathfrak{P}}$  over  $F_{\mathfrak{P}}$ . To describe the data  $\{B_{\mathfrak{P}}\}$ , we first suppose  $F \neq F_0$ . Pick  $\mathfrak{p} \in \mathcal{S}_0$  splitting into  $\mathfrak{p} = \mathfrak{P}\overline{\mathfrak{P}}$  in  $F$ . In this splitting case, we take  $K_{\mathfrak{P}}$  and  $K_{\overline{\mathfrak{P}}}$  such that  $B_{\mathfrak{P}} = K_{\mathfrak{P}} \oplus K_{\overline{\mathfrak{P}}} = B_{\overline{\mathfrak{P}}}$  for a finite semi-simple extension  $K_{\mathfrak{P}}$  and  $K_{\overline{\mathfrak{P}}}$  of  $F_{\mathfrak{P}} \cong F_{\overline{\mathfrak{P}}}$  with  $\dim_{F_{\mathfrak{P}}} K_{\mathfrak{P}} = m(\mathfrak{P})$  and  $\dim_{F_{\overline{\mathfrak{P}}}} K_{\overline{\mathfrak{P}}} = m(\overline{\mathfrak{P}})$ . Note that

$$\dim_{F_{\mathfrak{P}}} B_{\mathfrak{P}} = \dim_{F_{\overline{\mathfrak{P}}}} B_{\overline{\mathfrak{P}}} = m(\mathfrak{P}) + m(\overline{\mathfrak{P}}) = d.$$

If  $\mathfrak{P} \in \mathcal{S}^i$ , we take  $K_{\mathfrak{P}}/F_{\mathfrak{P}}$  semisimple of dimension  $m(\mathfrak{P}) = d/2$  and put  $B_{\mathfrak{P}} = K_{\mathfrak{P}} \oplus K_{\mathfrak{P}}$  (so,  $\dim_{F_{\mathfrak{P}}} B_{\mathfrak{P}} = 2m(\mathfrak{P}) = d$  for  $\mathfrak{P} \in \mathcal{S}^i$  above  $\mathfrak{p}$ ). Then we put  $B_p = \bigoplus_{\mathfrak{P} \in \mathcal{S}} B_{\mathfrak{P}}$ . If  $\mathfrak{P} \in \mathcal{S}^i$ , by the unramifiedness of  $F_{\mathfrak{P}}/F_{0,\mathfrak{p}}$ , we have the Frobenius automorphism  $\sigma \in \text{Aut}(K_{\mathfrak{P}}/F_{0,\mathfrak{p}})$  corresponding to the generator  $c$  of  $\text{Gal}(F_{\mathfrak{P}}/F_{0,\mathfrak{p}})$ . We embed  $F_{\mathfrak{P}}$  into  $K_{\mathfrak{P}} \oplus K_{\mathfrak{P}}$  by  $x \mapsto (x \oplus x^{\sigma})$ , and write  $j \in \text{Aut}(B_{\mathfrak{P}}/F_{0,\mathfrak{p}})$  given by  $B_{\mathfrak{P}} = K_{\mathfrak{P}} \oplus K_{\mathfrak{P}} \ni (x, y) \mapsto (y, x) \in K_{\mathfrak{P}} \oplus K_{\mathfrak{P}} = B_{\mathfrak{P}}$ . Thus  $j$  induces  $c \in \text{Gal}(F/F_0)$  on  $F_{\mathfrak{P}} \subset B_{\mathfrak{P}}$ . When  $\mathfrak{P} \in \mathcal{S}^{sp}$ , we have two identical factors  $B_{\mathfrak{P}}$  and  $B_{\overline{\mathfrak{P}}}$  in  $B_p$ , and we write  $j$  for the interchange of the two components. Thus we have well defined involution  $j \in \text{Aut}(B_p/F_{0,p})$  inducing the generator  $c$  of  $\text{Gal}(F/F_0)$ .

Now we suppose  $F = F_0$ . In this case, we take semi-simple commutative  $B_p/F_p$  of dimension  $m(\mathfrak{p}) = d/2$ . Then we put  $B_p = \prod_{\mathfrak{p}} (B_p \oplus B_p)$ . We write  $j$  for the automorphism interchanging the two identical components of  $B_p$ .

We now describe the embedding of a subgroup of  $B_p^{\times}$  into the parabolic subgroup  $P$ . If  $\mathfrak{P} \in \mathcal{S}_{sp}$  (with  $\mathfrak{p} = \mathfrak{P} \cap F_0$ ),  $G_{\mathfrak{p}}^1 = \text{SL}_d(O_{\mathfrak{P}})$ ,  $C_{\mathfrak{p}} = M_d(O_{\mathfrak{p}})$  and the Levi factor of  $P_{\mathfrak{p}} \cap G_1$  is isomorphic to

$$M_{1,\mathfrak{p}} = \{(g_{\mathfrak{P}}, g_{\overline{\mathfrak{P}}}) \in \text{GL}_{m(\mathfrak{P})}(O_{\mathfrak{P}}) \times \text{GL}_{m(\overline{\mathfrak{P}})}(O_{\overline{\mathfrak{P}}}) \mid \det(g_{\mathfrak{P}}) = \det(g_{\overline{\mathfrak{P}}})^c\} \hookrightarrow G_{\mathfrak{p}}^1,$$

where  $(g_{\mathfrak{P}}, g_{\overline{\mathfrak{P}}})$  is sent to  $\begin{pmatrix} g_{\mathfrak{P}} & 0 \\ 0 & {}^t \overline{g_{\mathfrak{P}}}^{-1} \end{pmatrix} \in \text{SL}_d(O_{\mathfrak{P}}) = G_{\mathfrak{p}}^1$  (identifying  $O_{\mathfrak{P}}$  and  $O_{\overline{\mathfrak{P}}}$  by  $c$ ). We can embed  $B_{\mathfrak{P}}$  into  $C_{\mathfrak{P}}$  so that

$$B_{\mathfrak{P}}^1 := \{(b_{\mathfrak{P}}, b_{\overline{\mathfrak{P}}}^{-j}) \in O_{K_{\mathfrak{P}}}^{\times} \times O_{K_{\overline{\mathfrak{P}}}}^{\times} \mid N_{K_{\mathfrak{P}}/F_{\mathfrak{P}}}(b_{\mathfrak{P}}) = N_{K_{\overline{\mathfrak{P}}}/F_{\mathfrak{P}}}(b_{\overline{\mathfrak{P}}})^c\} \subset B_{\mathfrak{P}}^{\times}$$

is in  $M_{1,p}$ . In other words,  $(b_{\mathfrak{P}}, b_{\mathfrak{P}}^{-j})$  is sent to  $\begin{pmatrix} b_{\mathfrak{P}} & 0 \\ 0 & {}^t b_{\mathfrak{P}}^{-1} \end{pmatrix} \in M_{1,p}$ . If  $\mathfrak{P} \in \mathcal{S}^i$  or  $F = F_0$ ,  $(b_{\mathfrak{P}}, b_{\mathfrak{P}}^{-j})$  in

$$B_{\mathfrak{P}}^1 := \{(b_{\mathfrak{P}}, b_{\mathfrak{P}}^{-j}) \in O_{K_{\mathfrak{P}}}^{\times} \times O_{K_{\mathfrak{P}}}^{\times} \mid N_{K_{\mathfrak{P}}/F_{\mathfrak{P}}}(b_{\mathfrak{P}}) = N_{K_{\mathfrak{P}}/F_{\mathfrak{P}}}(b_{\mathfrak{P}})^c\} \subset B_{\mathfrak{P}}^{\times}$$

is sent to  $\begin{pmatrix} b_p & 0 \\ 0 & {}^t b_p^{-1} \end{pmatrix} \in M_{1,p}$ . In this way, we have an embedding of  $B_p^{\times}$  into  $C_p^{\times}$ .

We now globalize  $B_{p/F_p}$  finding an  $F$ -subalgebra  $B \subset B_p$  such that  $B \otimes_{\mathbb{Q}} \mathbb{Q}_p = B_p$  with  $B \subset \text{End}_D(A_x) \otimes \mathbb{Q}$  for some  $x \in I_{\infty}(\mathbb{F})$ . By Lemma 3.5, we find a CM semi-simple commutative algebra  $B \subset C$  such that  $B^{\times} \cap G_1$  gives the maximal commutative subgroup of the stabilizer of a point  $x \in \text{Sh}(\Sigma)$ . At complex place  $\sigma : F \hookrightarrow \mathbb{C}$ ,  $C \otimes_{F,\sigma} \mathbb{C} \cong M_d(\mathbb{C})$  and  $D \otimes_{F,\sigma} \mathbb{C} \cong M_n(\mathbb{C})$ . Thus  $V \otimes_{F,\sigma} \mathbb{C} \cong M_{n \times d}(\mathbb{C})$ , where  $(\alpha, \beta) \in D_{\sigma} \times C_{\sigma}$  acts on  $V_{\sigma}$  by  $x \mapsto \alpha x^t \beta$ . Thus again we can choose an idempotent  $\varepsilon \in M_n(\mathbb{C}) = D_{\sigma}$  so that  $\varepsilon = \text{diag}[1, 0, \dots, 0]$  stable under  $\rho$  (in this case, we may assume that  $x^{\rho} = {}^t \bar{x}$ ). By our choice of  $B_{\sigma}$  (in the proof of Lemma 3.5), we may assume that the action of  $B_{\sigma}$  for  $\sigma \in \Phi$  on  $\varepsilon \cdot V_{\sigma} = \varepsilon(V \otimes_{F,\sigma} \mathbb{C})$  gives rise to a representation whose trace is given by  $m(\mathfrak{P}_{\sigma})\sigma + m(\overline{\mathfrak{P}}_{\sigma})\sigma c$  for each  $\sigma \in \Phi$ . Thus we can give a complex structure on  $V_{\infty} = L_{\mathbb{R}}$  so that on its tangent space, the action of  $B$  is given by  $\sum_{\sigma \in \Phi} m(\mathfrak{P}_{\sigma})\sigma + m(\overline{\mathfrak{P}}_{\sigma})\sigma c$ . Then  $A_x(\mathbb{C}) = V_{\infty}/L$  is a complex torus with a Riemann form sitting over  $x \in \text{Sh}(\mathbb{C})$ . Since  $H_1(A_x, \mathbb{Q}) = V$  canonically, we get a level structure  $\eta : L_{\mathbb{A}(\infty)} \cong V(A_x)$ . Since  $L$  is naturally a  $D \otimes_F B$ -module,  $A_x$  is an abelian variety with complex multiplication by  $D \otimes_F B$ . Thus  $\underline{A}_x = (A_x, \bar{\lambda}, \eta^{(\Sigma)})$  descends to a finite extension  $\widetilde{\mathcal{W}}$  of  $\mathcal{W}$ . By our choice again,  $\underline{A}_{x/\mathbb{F}} = \underline{A}_x \times_{\widetilde{\mathcal{W}}} \mathbb{F}$  is an ordinary abelian variety giving rise to a point  $x \in \text{Ig}(\mathbb{F})$ . Since the action of  $G(\mathbb{A}^{(\Sigma)}) \times M$  on  $\pi_0(\text{Ig}_{\infty/\mathbb{F}})$  is transitive, by moving  $x$  by an element in  $G(\mathbb{A}^{(\Sigma)}) \times M$ , we may assume that  $x$  is in the connected component  $I_{\infty}$  we have chosen (see Lemma 1.3).

**Remark 4.1** Although we used the strong approximation theorem in the proof of Lemma 1.1, what we really need is the transitivity of the action of  $G(\mathbb{A}^{(\Sigma)}) \times M$  on  $\pi_0(\text{Ig}_{\infty/\mathbb{F}})$  (which is true for all Shimura variety of PEL type by the approximation theorem).

**Lemma 4.2** Choose  $\Sigma$  so that for all  $\ell$  outside  $\Sigma$ ,  $G_{1/\mathbb{Q}_{\ell}}$  is isomorphic to a quasi-split special unitary group or a split symplectic group according as  $F \neq F_0$  or  $F = F_0$ . Then  $G_1(\mathbb{A}^{(\Sigma)})$  fixes  $\pi_0(\text{Ig}_{\infty/\mathbb{F}})$  element by element for the Igusa tower  $\text{Ig}_{\infty}$  over  $S^{(\Sigma)}$ .

*Proof* Write  $S = S^{(\Sigma)}$ . By Lemma 1.1,  $G_1(\mathbb{A}^{(\Sigma)})/\{\text{global center}\} \subset \text{Aut}(S/\mathbb{F})$ , and  $G_1(\mathbb{A}^{(\Sigma)})$  is isomorphic to  $Sp_d(F_{0,\mathbb{A}}^{(\Sigma)})$  if  $F = F_0$  and quasi-split  $SU(m, m')(F_{0,\mathbb{A}}^{(\Sigma)})$  (so,  $|m - m'| \leq 1$ ) if  $F \neq F_0$ . The classical groups  $Sp(d)$  and  $SU(m, m')$  (with  $|m - m'| \leq 1$ ) are almost simple and simply connected, and hence they are generated by (divisible) unipotent subgroups (cf. [1] Section 3). Thus  $G_1(\mathbb{A}^{(\Sigma)})$  does not have nontrivial finite quotient groups. Since  $G_1(\mathbb{A}^{(\Sigma)})$  gives permutations of the finite set  $\pi_0(\text{Ig}_{\alpha/\mathbb{F}})$  ( $\alpha < \infty$ ), it has to fix  $I_{\alpha}$  and hence  $I_{\infty} = I_K$  for  $K = G(\mathbb{Z}_{\Sigma}^{(p)})$ .

Thus we have verified (St) for the above choice of  $\Sigma$ . Since  $T_x(\mathbb{Z}_{(\Sigma)})T_x(\mathbb{A}^{(\Sigma)})$  is dense in  $T_x(\mathbb{Z}_{\Sigma} \times \mathbb{A}^{(\Sigma)})$  and  $T_x(\mathbb{Z}_{(\Sigma)})$  fixes the point  $x \in I_{\infty}$ ,  $T_x(\mathbb{Z}_p)$  has to fix  $I_{\infty}$ , because  $T_x(\mathbb{A}^{(\Sigma)}) \subset G_1(\mathbb{A}^{(\Sigma)})$  fixes  $I_{\infty} \in \pi_0(\text{Ig}_{\infty/\mathbb{F}})$  by (St) now verified. By Lemma 1.3,  $\text{Gal}(I_{\infty}/S)$  contains a group isomorphic to  $\prod_{\mathfrak{P}} B_{\mathfrak{P}}^1$ . By Corollary 2.5,  $\text{Gal}(I_{\infty}/S)$  contains the derived group  $D(M_1) \cong \prod_{\mathfrak{P} \in \mathcal{S}} SL_{m(\mathfrak{P})}(O_{\mathfrak{P}})$  of  $M_1$ . For totally split  $B_p \cong F_p^d$ ,  $\prod_{\mathfrak{P}} B_{\mathfrak{P}}^1$  surjects down to  $M_1/D(M_1)$ , and hence  $\text{Gal}(I_{\infty}/S)$  contains  $M_1$ . Since  $\text{Gal}(I_{\infty}/S) \subset M_1$  by construction (see the argument in [2] p. 374 or [3] 3.3), we get the identity  $\text{Gal}(I_{\infty}/S) = M_1$ .

Let  $I_\infty^{(p)}$  be the irreducible component of the Igusa tower over  $S^{(p)}$ . The function fields of  $S^{(p)}/S$  and  $I_\infty/S$  are linearly disjoint over the function field of  $S$ , and the function field of  $I_\infty^{(p)}$  is the composite of function fields of  $S^{(p)}$  and  $I_\infty$  over the function field of  $S$ . From this we conclude  $\text{Gal}(I_\infty/S^{(p)}) = M_1$ . Then for any compact subgroup  $K$  of  $G(\mathbb{A}^{(p\infty)})$ ,  $S_K$  is a quotient of  $S^{(p)}$ ; so, we get  $\text{Gal}(I_K/S_K) = M_1$ . This finishes the proof of the first assertion of Theorem 0.1. As already explained after the theorem, the rest of Theorem 0.1 follows from the first assertion.

## References

- [1] Hida, H.: Automorphism Groups of Shimura Varieties of PEL type. *Documenta Math.*, **11**, 25–56 (2006)
- [2] Hida, H.:  $p$ -Adic Automorphic Forms on Shimura Varieties, Springer Monographs in Mathematics, 2004, Springer
- [3] Hida, H.: Irreducibility of the Igusa tower over unitary Shimura varieties, to appear in Shahidi anniversary volume from Clay Math. Institute (available at [www.math.ucla.edu/~hida](http://www.math.ucla.edu/~hida)).
- [4] Lan, K. W.: Arithmetic compactifications of PEL-type Shimura varieties, Ph. D. thesis at Harvard, 2008
- [5] Chai, C. L., Oort, F.: Monodromy and irreducibility of leaves (available at: <http://www.math.upenn.edu/~chai>)
- [6] Chai, C. L.: Monodromy of Hecke invariant subvarieties. *Pure Appl. Math. Q.*, **1**, 291–303 (2005)
- [7] Chai, C. L.: Methods for  $p$ -adic monodromy, preprint 2006, to appear in Journal of the Inst. of Math. Jussieu (available at <http://www.math.upenn.edu/~chai>)
- [8] Boyer, P.: Sur l'irréductibilité de quelques variétés d'Igusa, preprint, 2007 (ArXiv:math/0702329v1)
- [9] Harris, M., Taylor, R.: The geometry and cohomology of some simple Shimura varieties, *Ann. of Math. Studies* **151**, Princeton University Press, 2001
- [10] Wedhorn, T.: Ordinarity in good reductions of Shimura varieties of PEL type. *Ann. Scient. Éc. Norm. Sup. 4th Series*, **32**, 575–618 (1999)
- [11] Kottwitz, R.: Points on Shimura varieties over finite fields. *J. Amer. Math. Soc.*, **5**, 373–444 (1992)
- [12] Deligne, P.: Variétés de Shimura: Interprétation modulaire, et techniques de construction de modèles canoniques. *Proc. Symp. Pure Math.*, **33.2**, 247–290 (1979)
- [13] Milne, J.: Canonical models of (mixed) Shimura varieties and automorphic vector bundles. *Perspective Math.*, **10**, 283–414 (1990)
- [14] Mumford, D.: Abelian Varieties, TIFR Studies in Mathematics, Oxford University Press, New York, 1994
- [15] Hida, H.: Hilbert Modular Forms and Iwasawa Theory, Oxford Mathematical Monographs, Oxford University Press, 2006
- [16] Rapoport, M.: Compactifications de l'espace de modules de Hilbert–Blumenthal. *Comp. Math.*, **36**, 255–335 (1978)
- [17] Faltings, G., Chai, C. L.: Degeneration of Abelian Varieties, Springer, New York, 1990
- [18] Fujiwara, K.: Arithmetic compactifications of Shimura varieties, Master's Thesis (University of Tokyo), 1989
- [19] Milne, J., Shih, K.-y.: Automorphism groups of Shimura varieties and reciprocity laws. *Amer. J. Math.*, **103**, 911–935 (1981)
- [20] Hartshorne, R.: Algebraic Geometry, Graduate Texts in Mathematics, 52, Springer, New York, 1977
- [21] Shimura, G.: Abelian Varieties with Complex Multiplication and Modular Functions, Princeton University Press, Princeton, NJ, 1998
- [22] Serre, J. P., Tate, J.: Good reduction of abelian varieties. *Ann. of Math.*, **88**, 452–517 (1968) (Serre's Œuvres II, 472–497, No. 79)
- [23] Shimura, G.: On analytic families of polarized abelian varieties and automorphic functions. *Ann. of Math.*, **78**, 149–192 (1963)
- [24] Weil, A.: Basic Number Theory, Springer, New York, 1974
- [25] Shimura, G.: Euler Products and Eisenstein Series, CBMS Regional Conference Series 93, American Mathematical Society, Providence, RI, 1997
- [26] Chai, C. L.: Compactification of Siegel Moduli Schemes, LMS Lecture Note Series, 107, Cambridge University Press, New York, 1985
- [27] Deligne, P.: Travaux de Shimura, Sem. Bourbaki, Exp. 389. *Lecture Notes in Math.*, **244**, 123–165 (1971)
- [28] Hida, H.:  $p$ -Adic automorphic forms on reductive groups, *Astérisque* **298** (2005), 147–254 (available at [www.math.ucla.edu/~hida](http://www.math.ucla.edu/~hida)).
- [29] Igusa, J.: Kroneckerian model of fields of elliptic modular functions. *Amer. J. Math.*, **81**, 561–577 (1959)