J. DIFFERENTIAL GEOMETRY
66 (2004) 233-287

A CHARACTERIZATION OF CERTAIN SHIMURA CURVES IN THE MODULI STACK OF ABELIAN VARIETIES

Eckart Viehweg & Kang Zuo

Abstract

Let $f : X \to Y$ be a semistable family of complex abelian varieties over a curve Y of genus $g(Y)$, and smooth over the complement of s points. If ℓ denotes the rank of the nonflat $(1,0)$ part $F^{1,0}$ of the corresponding variation of Hodge structures, the Arakelov inequality says that

$$2 \cdot \deg(F^{1,0}) \leq \ell \cdot (2g(Y) - 2 + s).$$

The family reaches this bound if and only if the Higgs field of the variation of Hodge structures is an isomorphism. The latter is reflected in the existence of special Hodge cycles in the general fibre, and the base of such a family is a Shimura curve.

In particular, for $s \neq 0$, such a family must be isogenous to the ℓ-fold product of a modular family of elliptic curves, and a constant abelian variety.

For $s = 0$, if the flat part of the variation of Hodge structures is defined over the rational numbers, one finds the family to be isogenous to the product of several copies of a family $h : Z \to Y$, and a constant abelian variety. In this case, $h : Z \to Y$ is obtained from the corestriction of a quaternion algebra A, defined over a totally real numberfield F, and ramified over all infinite places but one.

In case the flat part is not defined over the rational numbers, one still can classify the corresponding variations of Hodge structures.

This work has been supported by the "DFG-Schwerpunktprogramm Globale Methoden in der Komplexen Geometrie". The second named author was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK 4239/01P).

Received 06/15/2003.

0. Introduction

Throughout this article, Y will denote a nonsingular complex projective curve, U an open dense subset, and $X_0 \to U$ a smooth family of abelian varieties. We choose a projective nonsingular compactification X of X_0 such that the family extends to a morphism $f : X \to Y$, which we call again a family of abelian varieties although some of the fibres are singular. We write $S = Y \setminus U$, and $\Delta = f^{-1}(S)$. Consider the weight 1 variation of Hodge structures given by $f : X_0 \to U$, i.e., $R^1 f_* \mathbb{Z}_{X_0}$. We will always assume that the monodromy of $R^1 f_* \mathbb{Z}_{X_0}$ around all points in S is unipotent. The Deligne extension of $(R^1 f_* \mathbb{Z}_{X_0}) \otimes \mathcal{O}_U$ to Y carries a Hodge filtration. Taking the graded sheaf one obtains the Higgs bundle

$$(E, \theta) = (E^{1,0} \oplus E^{0,1}, \theta_{1,0})$$

with $E^{1,0} = f_* \Omega^1_{X/Y}(\log \Delta)$ and $E^{0,1} = R^1 f_* \mathcal{O}_X$. The Higgs field $\theta_{1,0}$ is given by the edge morphisms

$$f_* \Omega^1_{X/Y}(\log \Delta) \longrightarrow R^1 f_* \mathcal{O}_X \otimes \Omega^1_Y(\log S)$$

of the tautological sequence

$$0 \to f^* \Omega^1_Y(\log S) \to \Omega^1_X(\log \Delta) \to \Omega^1_{X/Y}(\log \Delta) \to 0.$$

By [**14**] E can be decomposed as a direct sum $F \oplus N$ of Higgs bundles with $E^{1,0} \cap F$ ample and with N flat, hence for $F^{i,j} = E^{i,j} \cap F$ and $N^{i,j} = E^{i,j} \cap N$ the Higgs bundle E decomposes in

$$(F = F^{1,0} \oplus F^{0,1}, \theta_{1,0}|_{F^{1,0}}) \quad \text{and} \quad (N^{1,0} \oplus N^{0,1}, 0). \tag{0.0.1}$$

For $g_0 = \operatorname{rank}(F^{1,0})$ the Arakelov inequalities ([**7**], generalized in [**20**], [**12**]) say that

$$2 \cdot \deg(F^{1,0}) \leq g_0 \cdot \deg(\Omega^1_Y(\log S)). \tag{0.0.2}$$

In this note we will try to understand $f : X \to Y$, for which (0.0.2) is an equality, or as we will say, of families reaching the Arakelov bound. By Proposition 1.2, this property is equivalent to the maximality of the Higgs field for F, saying that $\theta|_{F_{1,0}} : F^{1,0} \to F^{0,1} \otimes \Omega^1_Y(\log S)$ is an isomorphism.

As it will turn out, the base of a family of abelian varieties reaching the Arakelov bound is a Shimura curve, and the maximality of the Higgs field is reflected in the existence of special Hodge cycles on the general fibre. Before formulating a general result, let us consider two examples.

For families of elliptic curves, the maximality of the Higgs field just says that the family is modular (see Section 2).

Proposition 0.1. *Let $f : E \to Y$ be a semistable family of elliptic curves, smooth over $U \subset Y$. If $E \to Y$ is nonisotrivial and reaching the Arakelov bound, $E \to Y$ is modular, i.e., U is the quotient of the upper half plane $\mathcal{H}$ by a subgroup of $\mathrm{Sl}_2(\mathbb{Z})$ of finite index, and the morphism $U \to \mathbb{C} = \mathcal{H}/\mathrm{Sl}_2(\mathbb{Z})$ is given by the j-invariant of the fibres.*

For $S \neq \emptyset$ the only families of abelian varieties reaching the Arakelov bound are built up from modular families of elliptic curves.

Theorem 0.2. *Let $f : X \to Y$ be a family of abelian varieties smooth over U, and such that the local monodromies around $s \in S$ are unipotent. If $S \neq \emptyset$, and if $f : X \to Y$ reaches the Arakelov bound, then there exists an étale covering $\pi : Y' \to Y$ such that $f' : X' = X \times_Y Y' \to Y'$ is isogenous over Y' to a product*

$$B \times E \times_{Y'} \cdots \times_{Y'} E,$$

where B is abelian variety defined over $\mathbb{C}$ of dimension $g - g_0$, and where $h : E \to Y'$ is a family of semistable elliptic curves reaching the Arakelov bound.

Results parallel to 0.2 have been obtained in [**30**] for families of $K3$-surfaces, and the methods and results of [**30**] have been a motivation to study the case of abelian varieties.

As we will see in Section 4, Theorem 0.2 follows from the existence of too many endomorphisms of the general fibre of $f : X \to Y$, which in turn implies the existence of too many cycles on the general fibre of $X \times_Y X$. We give an elementary proof of Theorem 0.2 in Section 4, although it is nothing but a first example for the relation between the maximality of Higgs fields, and the moduli of abelian varieties with a given special Mumford–Tate group Hg, constructed in [**17**] and [**18**] (see Section 2).

Proposition 0.3. *Let $f : X \to Y$ be a family of g-dimensional abelian varieties reaching the Arakelov bound. Assume that $g = g_0$, or more generally that the largest unitary local subsystem $\mathbb{U}_1$ of $R^1 f_* \mathbb{C}_{X_0}$ is defined over $\mathbb{Q}$. Then there exists a finite cover $Y' \to Y$, étale over U, and a $\mathbb{Q}$-algebraic subgroup $\mathrm{Hg} \subset \mathrm{Sp}(2g, \mathbb{R})$, such that pullback family $f' : X' = X \times_Y Y' \to Y'$ is a semistable compactification of the universal family of polarized abelian varieties with special Mumford–Tate group contained in* Hg*, and with a suitable level structure.*

As a preparation for the proof of Proposition 0.3 we will show in Section 1, using Simpson's correspondence between Higgs bundles and local systems, that the maximality of the Higgs field enforces a presentation of the local systems $R^1f_*\mathbb{C}_{X_0}$ and $\mathbb{E}\mathrm{nd}(R^1f_*\mathbb{C}_{X_0})$ using direct sums and tensor products of one weight one complex variation of Hodge structures $\mathbb{L}$ of rank two and several unitary local systems.

Proposition 0.3 relates families reaching the Arakelov bound to totally geodesic subvarieties of the moduli space of abelian varieties, as considered by Moonen in [**16**], or to the totally geodesic holomorphic embeddings, studied by Abdulali in [**1**] (see Remark 2.5 b). As in [**21**] one could use the classification of Shimura varieties due to Satake [**22**] to obtain a complete list of those families, and to characterize them in terms of properties of their variation of Hodge structures.

We choose a different approach, less relying on the theory of Shimura varieties, and more adapted to handle the remaining families of abelian varieties (see Remark 2.5 c), as well as some other families of varieties of Kodaira dimension zero (see [**33**]). We first show that the decompositions of $R^1f_*\mathbb{C}_{X_0}$ and $\mathbb{E}\mathrm{nd}(R^1f_*\mathbb{C}_{X_0})$ mentioned above are defined over $\overline{\mathbb{Q}} \cap \mathbb{R}$. In case $S \neq \emptyset$ it is then easy to see, that the unitary parts of the decompositions trivialize, after replacing Y by a finite étale cover Y' (see 4.4).

For $S = \emptyset$, let us assume first that the assumptions made in 0.3 hold true. By [**9**] (see Proposition 6.3) they imply that the family is rigid, i.e., that the morphism from Y to the moduli stack of polarized abelian varieties has no nontrivial deformation, except those obtained by deforming a constant abelian subvariety.

Mumford gave in [**18**] countably many moduli functors of abelian fourfolds, where Hg is obtained via the corestriction of an quaternion algebra, defined over a totally real cubic number field F. Generalizing his construction one considers quaternion division algebras A defined over any totally real number field F, which are ramified at all infinite places except one. Choose an embedding

$$D = \mathrm{Cor}_{F/\mathbb{Q}}A \subset M(2^m, \mathbb{Q}),$$

with m minimal. As we will see in Section 5 writing $d = [F : \mathbb{Q}]$ one finds $m = d$ or $m = d+1$. By 5.9 and 5.10 we get the following types of moduli functors of abelian varieties with special Mumford–Tate group

$$\mathrm{Hg} = \{x \in D^*;\ x\overline{x} = 1\}$$

and with a suitable level structure, which are represented by a smooth family $Z_A \to Y_A$ over a compact Shimura curve Y_A. Since we did not fix the level structure, Y_A is not uniquely determined by A. So it rather stands for a whole class of possible base curves, two of which have a common finite étale covering.

Example 0.4. Let Z_η denote the generic fibre of $Z_A \to Y_A$. Then one of the following holds true:

i. $1 < m = d$ odd.
   In this case $\dim(Z_\eta) = 2^{d-1}$ and $\mathrm{End}(Z_\eta) \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q}$.
ii. $m = d + 1$. Then $\dim(Z_\eta) = 2^d$ and:
   a. For d odd, $\mathrm{End}(Z_\eta) \otimes_{\mathbb{Z}} \mathbb{Q}$ a totally indefinite quaternion algebra over $\mathbb{Q}$.
   b. For d even, $\mathrm{End}(Z_\eta) \otimes_{\mathbb{Z}} \mathbb{Q}$ a totally definite quaternion algebra over $\mathbb{Q}$.

Let us call the family $Z_A \to Y_A$ a family of Mumford type.

For $d = 1$ or 2 the examples in 0.4 include the only two Shimura curves of PEL-type, parameterizing:

- Moduli schemes of false elliptic curves, i.e., polarized abelian surfaces B with $\mathrm{End}(B) \otimes_{\mathbb{Z}} \mathbb{Q}$ a totally indefinite quaternion algebra over $\mathbb{Q}$ (see also [**25**]).
- Moduli schemes of abelian fourfolds B with $\mathrm{End}(B) \otimes_{\mathbb{Z}} \mathbb{Q}$ a totally definite quaternion algebra over $\mathbb{Q}$.

We will see in Section 6 that for $g = g_0$, up to powers and isogenies, the families of Mumford type are the only smooth families of abelian varieties over curves reaching the Arakelov bound.

Theorem 0.5. *Let $f : X \to Y$ be a smooth family of abelian varieties. If the largest unitary local subsystem $\mathbb{U}_1$ of $R^1 f_* \mathbb{C}_X$ is defined over $\mathbb{Q}$ and if $f : X \to Y$ reaches the Arakelov bound, then there exist:*

a. *a quaternion division algebra A, defined over a totally real number field F, and ramified at all infinite places except one,*
b. *an étale covering $\pi : Y' \to Y$,*
c. *a family of Mumford type $h : Z = Z_A \to Y' = Y_A$, as in Example* 0.4*, and an abelian variety B such that $f' : X' = X \times_Y Y' \to Y'$ is isogenous to*

$$B \times Z \times_{Y'} \cdots \times_{Y'} Z \longrightarrow Y'.$$

Things get more complicated if one drops the condition on the unitary local subsystem $\mathbb{U}_1$ of $R^1 f_* \mathbb{C}_X$. For one quaternion algebra A, there

exist several nonisogenous families. Hence it will no longer be true, that up to a constant factor $f : X \to Y$ is isogenous to the product of one particular family.

Example 0.6 (See 5.11)**.** Let A be a quaternion algebra defined over a totally real number field F, ramified at all infinite places but one, and let L be a subfield of F. Let $\beta_1, \dots, \beta_\delta : L \to \overline{\mathbb{Q}}$ denote the different embeddings of L. For $\mu = [F : L] + 1$ (or may be $\mu = [F : \mathbb{Q}]$ in case that $L = \mathbb{Q}$ and μ odd) there exists an embedding

$$\mathrm{Cor}_{F/L} A \subset M(2^\mu, L).$$

As is well-known (see Section 5) for some Shimura curve Y' such an embedding gives rise to a representation of $\pi_1(Y', *)$ in $M(2^\mu, L)$, hence to a local L system $\mathbb{V}_L$. Moreover there exists an irreducible $\mathbb{Q}$-local system $\mathbb{X}_\mathbb{Q} = \mathbb{X}_{A,L;\mathbb{Q}}$ for which $\mathbb{X}_\mathbb{Q} \otimes \overline{\mathbb{Q}}$ is a direct sum of the local systems $\mathbb{V}_L \otimes_{L,\beta_\nu} \overline{\mathbb{Q}}$.

There exist nonisotrivial families $h : Z \to Y'$ with a geometrically simple generic fibre, such that $R^1 h_* \mathbb{Q}$ is a direct sum of ι copies of $\mathbb{X}_\mathbb{Q}$. Such examples, for $g = 4$ and 8 have been considered in [**9**]. Here F is a quadratic extension of $\mathbb{Q}$, $L = F$ and $\iota = 1$ or 2. For $g = 8$, i.e., $\iota = 2$, this gives the lowest-dimensional example of a nonrigid family of abelian varieties without a trivial subfamily [**9**]. A complete classification of such families is given in [**21**].

Theorem 0.7. *Let $f : X \to Y$ be a smooth family of abelian varieties. If $f : X \to Y$ reaches the Arakelov bound, then there exist an étale covering $\pi : Y' \to Y$, a quaternion algebra A, defined over a totally real number field F and ramified at all of the infinite places except one, an abelian variety B, and ℓ families $h_i : Z_i \to Y'$ of abelian varieties with geometrically simple generic fibre, such that:*

i. *$f' : X' = X \times_Y Y' \to Y'$ is isogenous to*

$$B \times Z_1 \times_{Y'} \cdots \times_{Y'} Z_\ell \longrightarrow Y'.$$

ii. *For each $i \in \{1, \dots, \ell\}$ there exists a subfield L_i of F such that the local system $R^1 h_{i*} \mathbb{Q}_{Z_i}$ is a direct sum of copies of the irreducible $\mathbb{Q}$-local system $\mathbb{X}_{A,L_i;\mathbb{Q}}$ defined in Example* 0.6.

iii. *For each $i \in \{1, \dots, \ell\}$ the following conditions are equivalent:*
   a. *$L_i = \mathbb{Q}$.*
   b. *$h_i : Z_i \to Y'$ is a family of Mumford type, as defined in Example* 0.4.

c. $\mathrm{End}(\mathbb{X}_{A,L_i;\mathbb{Q}})^{0,0} = \mathrm{End}(\mathbb{X}_{A,L_i;\mathbb{Q}})$.

Moreover, if one of those conditions holds true, $R^1h_{i*}\mathbb{Q}_{Z_i}$ *is irreducible, hence* $R^1h_{i*}\mathbb{Q}_{Z_i} = \mathbb{X}_{A,L_i;\mathbb{Q}}$.

Here, contrary to 0.5, we do not claim that a component $h_i : Z_i \to Y'$ is uniquely determined up to isogeny by $\mathbb{X}_{A,L_i;\mathbb{Q}}$ and by the rank of $R^1h_{i*}\mathbb{Q}_{Z_i}$.

We do not know for which g there are families of Jacobians among the families of abelian varieties considered in 0.2, 0.5 or 0.7, i.e., whether one can find a family $\varphi : Z \to Y$ of curves of genus g such that $f : J(Z/Y) \to Y$ reaches the Arakelov bound.

For $Y = \mathbb{P}^1$ the Arakelov inequality (0.0.2) implies that $\#S \geq 4$. Our hope, that a family of abelian varieties with $\#S = 4$ can not be a family of Jacobians, broke down when we found an example of a family of genus 2 curves over the modular curve $X(3)$ in [**13**] (see also [**37**]), whose Jacobian is isogenous to the product of a fixed elliptic curve B with the modular curve $E(3) \to X(3)$ (see Section 7).

As mentioned already, this article owes a lot to the recent work of the second named author with Xiao-Tao Sun and Sheng-Li Tan. We thank Ernst Kani for explaining his beautiful construction in [**13**], and for sharing his view about higher genus analogs of families of curves with splitting Jacobians. It is also a pleasure to thank Ben Moonen, Hélène Esnault and Frans Oort for their interest and help, Ngaiming Mok, for explaining us differential geometric properties of base spaces of families and for pointing out Mumford's construction in [**18**], Bruno Kahn and Claus Scheiderer for their help to understand quaternion algebras and their corestriction, and Martin Möller and the referee for pointing out several misprints and ambiguities.

This note grew out of discussions started when the first named author visited the Institute of Mathematical Science and the Department of Mathematics at the Chinese University of Hong Kong. His contributions to the final version (in particular to the proof of Theorems 0.5 and 0.7) were written during his visit to the I.H.E.S., Bures sur Yvette. He would like to thank the members of those three Institutes for their hospitality.

1. Splitting of $\mathbb{C}$-local systems

We will frequently use C. Simpson's correspondence between polystable Higgs bundles of degree zero and representations of the fundamental group $\pi_1(U, *)$.

Theorem 1.1 (C. Simpson [**27**])**.** *There exists a natural equivalence between the category of direct sums of stable filtered regular Higgs bundles of degree zero, and of direct sums of stable filtered local systems of degree zero.*

We will not recall the definition of a "filtered regular" Higgs bundle ([**27**], p. 717), just remark that for a Higgs bundle corresponding to a local system $\mathbb{V}$ with unipotent monodromy around the points in S the filtration is trivial, and automatically $\deg(\mathbb{V}) = 0$. By ([**27**], p. 720) the latter also holds true for local systems $\mathbb{V}$ which are polarisable $\mathbb{C}$-variations of Hodge structures.

For example, 1.1 implies that the splitting of Higgs bundles (0.0.1) corresponds to a decomposition over $\mathbb{C}$

$$(R^1 f_* \mathbb{Z}_{X_0}) \otimes \mathbb{C} = \mathbb{V} \oplus \mathbb{U}_1$$

where $\mathbb{V}$ corresponds to the Higgs bundle $(F = F^{1,0} \oplus F^{0,1}, \tau = \theta|_{F_{0,1}})$ and $\mathbb{U}_1$ to $(N = N^{1,0} \oplus N^{0,1}, \theta_N = \theta|_N = 0)$. Let $\Theta(N, h)$ denote the curvature of the Hodge metric h on $E^{1,0} \oplus E^{0,1}$ restricted to N, then by [**11**], chapter II we have

$$\Theta(N, h|_N) = -\theta_N \wedge \overline{\theta}_N - \overline{\theta}_N \wedge \theta_N = 0.$$

This means that $h|_N$ is a flat metric. Hence, $\mathbb{U}_1$ is a unitary local system.

In general, if $\mathbb{U}$ is a local system, whose Higgs bundle is a direct sum of stable Higgs bundles of degree zero and with a trivial Higgs field, then $\mathbb{U}$ is unitary.

As a typical application of Simpson's correspondence one obtains the polystability of the components of certain Higgs bundles. We just formulate it in the weight one case.

Recall that $F^{1,0}$ is polystable, if there exists a decomposition

$$F^{1,0} \simeq \bigoplus_i \mathcal{A}_i$$

with $\mathcal{A}_i$ stable, and

$$\frac{\deg \mathcal{A}_i}{\mathrm{rank}\mathcal{A}_i} = \frac{\deg F^{1,0}}{\mathrm{rank}F^{1,0}}.$$

Proposition 1.2. *Let $\mathbb{V}$ be a direct sum of stable filtered local systems of degree zero with Higgs bundle $(F = F^{1,0} \oplus F^{0,1}, \tau)$. Assume that $\tau|_{F^{0,1}} = 0$, that*

$$\tau_{1,0} = \tau|_{F^{1,0}} : F^{1,0} \longrightarrow F^{0,1} \otimes \Omega^1_Y(\log S) \subset F \otimes \Omega^1_Y(\log S),$$

and that for $g_0 = \operatorname{rank}(F^{1,0})$

$$2 \cdot \deg(F^{1,0}) = g_0 \cdot \deg(\Omega^1_Y(\log S)). \tag{1.2.1}$$

Then $\tau_{1,0}$ *is an isomorphism, and the sheaf* $F^{1,0}$ *is polystable.*

Proof of 1.2. Let $\mathcal{A} \subset F^{1,0}$ be a subsheaf, and let $\mathcal{B} \otimes \Omega^1_Y(\log S)$ be its image under $\tau_{1,0}$. Then $\mathcal{A} \oplus \mathcal{B}$ is a Higgs subbundle of $F^{1,0} \oplus F^{0,1}$, and applying 1.1 one finds $\deg(\mathcal{A}) + \deg(\mathcal{B}) \leq 0$. Hence

$$\begin{aligned} \deg(\mathcal{A}) &= \deg(\mathcal{B}) + \operatorname{rank}(\mathcal{B}) \cdot \deg(\Omega^1_Y(\log S)) \\ &\leq \deg(\mathcal{B}) + \operatorname{rank}(\mathcal{A}) \cdot \deg(\Omega^1_Y(\log S)) \\ &\leq -\deg(\mathcal{A}) + \operatorname{rank}(\mathcal{A}) \cdot \deg(\Omega^1_Y(\log S)). \end{aligned}$$

The equality (1.2.1) implies that

$$\frac{\deg(\mathcal{A})}{\operatorname{rank}(\mathcal{A})} \leq \frac{1}{2} \deg(\Omega^1_Y(\log S)) = \frac{\deg(F^{1,0})}{g_0},$$

and $F^{1,0}$ is semistable. If

$$\frac{\deg(\mathcal{A})}{\operatorname{rank}(\mathcal{A})} = \frac{\deg(F^{1,0})}{g_0},$$

$\operatorname{rank}(\mathcal{A}) = \operatorname{rank}(\mathcal{B})$ and $\deg(\mathcal{B}) = -\deg(\mathcal{A})$. The Higgs bundle $(F^{1,0} \oplus F^{0,1}, \tau)$ splits by 1.1 as a direct sum of stable Higgs bundles of degree zero. Hence $(\mathcal{A} \oplus \mathcal{B}, \tau|_{\mathcal{A}\oplus\mathcal{B}})$ is a direct factor of $(F^{1,0} \oplus F^{0,1}, \tau)$. In particular, $\mathcal{A}$ is a direct factor of $F^{1,0}$. For $\mathcal{A} = F^{1,0}$ one also obtains that $\tau_{1,0}$ is injective and by (1.2.1) it must be an isomorphism. q.e.d.

The local system $R^1 f_* \mathbb{Q}_{X_0}$ on $U = Y \setminus S$ is a $\mathbb{Q}$-variation of Hodge structures with unipotent local monodromies around $s \in S$, obviously having a $\mathbb{Z}$-form. By Deligne's semisimplicity theorem [**4**] it decomposes as a direct sum of irreducible polarisable $\mathbb{Q}$-variation of Hodge structures $\mathbb{V}_{i\mathbb{Q}}$.

More generally, if $\mathbb{V}$ is a polarized $\mathbb{C}$-variation of Hodge structures, and

$$\mathbb{V} = \bigoplus_i \mathbb{V}_i,$$

a decomposition with $\mathbb{V}_i$ an irreducible $\mathbb{C}$-local system, then by [**7**] each $\mathbb{V}_i$ again is a polarisable $\mathbb{C}$-variation of Hodge structures.

In both cases, taking the grading of the Hodge filtration, one obtains a decomposition of the Higgs bundle

$$(E, \theta) = (F^{1,0} \oplus F^{0,1}, \tau) \oplus (N^{1,0} \oplus N^{0,1}, 0)$$

as a direct sum of sub-Higgs bundles, as stated in 1.1. Obviously, each of the $\mathbb{V}_{i\mathbb{Q}}$ again reaches the Arakelov bound.

Our next constructions will not require the local system to be defined over $\mathbb{Q}$. So by abuse of notations, we will make the following assumptions:

Assumption 1.3. For a number field $L \subset \mathbb{C}$ consider a polarized L-variation of Hodge structures $\mathbb{X}_L$ of weight one over $U = Y \setminus S$ with unipotent local monodromies around $s \in S$. Assume that the local system $\mathbb{X} = \mathbb{X}_L \otimes_L \mathbb{C}$ has a decomposition $\mathbb{X} = \mathbb{V} \oplus \mathbb{U}_1$, with $\mathbb{U}_1$ unitary, corresponding to the decomposition

$$(E, \theta) = (F, \tau_{1,0}) \oplus (N, 0) = (F^{1,0} \oplus F^{0,1}, \tau_{1,0}) \oplus (N^{1,0} \oplus N^{0,1}, 0)$$

of Higgs fields. Assume that $\mathbb{V}$ (or (F, τ)) has a maximal Higgs field, i.e., that

$$\tau_{1,0} = \tau|_{F^{1,0}} : F^{1,0} \to F^{0,1} \otimes \Omega^1_Y(\log S)$$

is an isomorphism. Obviously, for $g_0 = \operatorname{rank}(F^{1,0})$ this is equivalent to the equality (1.2.1). Hence we will also say, that $\mathbb{X}$ (or (E, θ)) reaches the Arakelov bound.

Proposition 1.4. *If* $\deg \Omega^1_Y(\log S)$ *is even there exists a tensor product decomposition of variations of Hodge structures*

$$\mathbb{V} \simeq \mathbb{L} \otimes_{\mathbb{C}} \mathbb{T},$$

with:

a. $\mathbb{L}$ *is a rank-2 local system. For some invertible sheaf* $\mathcal{L}$*, with* $\mathcal{L}^2 = \Omega^1_Y(\log S)$ *the Higgs bundle corresponding to* $\mathbb{L}$ *is*

$$(\mathcal{L} \oplus \mathcal{L}^{-1}, \tau'),$$

with $\tau'|_{\mathcal{L}^{-1}} = 0$ *and* $\tau'|_{\mathcal{L}}$ *given by an isomorphism*

$$\tau'_{1,0} : \mathcal{L} \longrightarrow \mathcal{L}^{-1} \otimes \Omega^1_Y(\log S).$$

$\mathcal{L}$ *has bidegree* $1, 0$*, and* $\mathcal{L}^{-1}$ *has bidegree* $0, 1$*.*

b. *If* g_0 *is odd,* $\mathcal{L}^{g_0} = \det(F^{1,0})$ *and* $\mathcal{L}$ *is uniquely determined.*

c. *For* g_0 *even, there exists some invertible sheaf* $\mathcal{N}$ *of order two in* $\operatorname{Pic}^0(Y)$ *with* $\mathcal{L}^{g_0} = \det(F^{1,0}) \otimes \mathcal{N}$*.*

d. $\mathbb{T}$ *is a unitary local system and a variation of Hodge structures of pure bidegree* $0, 0$*. If* $(\mathcal{T}, 0)$ *denotes the corresponding Higgs field, then* $\mathcal{T} = F^{1,0} \otimes \mathcal{L}^{-1} = F^{0,1} \otimes \mathcal{L}$*.*

In Section 6 we will need a slightly stronger statement.

Addendum 1.5. *If in* 1.4, *there exists a presentation* $\mathbb{V} = \mathbb{T}_1 \otimes_{\mathbb{C}} \mathbb{V}_1$ *with* $\mathbb{T}_1$ *unitary and a variation of Hodge structures of pure bidegree* $0, 0$, *then there exists a unitary local system* $\mathbb{T}_2$ *with* $\mathbb{T} = \mathbb{T}_1 \otimes_{\mathbb{C}} \mathbb{T}_2$.

In fact, write $(\mathcal{T}_1, 0)$ and $(F_1^{1,0} \oplus F_1^{0,1}, \tau_1)$ for Higgs fields corresponding to $\mathbb{T}_1$ and $\mathbb{V}_1$, respectively. Then $\deg(\mathcal{T}_1) = 0$ and

$$\begin{aligned} 2 \cdot \deg(F_1^{1,0}) \cdot \operatorname{rank}(\mathcal{T}) &= 2 \cdot \deg(F^{1,0}) \\ &= g_0 \cdot \deg(\Omega^1_Y(\log S)) \\ &= \operatorname{rank}(F_1^{1,0}) \cdot \operatorname{rank}(\mathcal{T}) \cdot \deg(\Omega^1_Y(\log S)). \end{aligned}$$

So $(F_1^{1,0} \oplus F_1^{0,1}, \tau_1)$ again satisfies the assumptions made in 1.4.

Proof of 1.4. Taking the determinant of

$$\tau_{1,0} : F^{1,0} \xrightarrow{\simeq} F^{0,1} \otimes \Omega^1_Y(\log S),$$

one obtains an isomorphism

$$\det \tau_{1,0} : \det F^{1,0} \xrightarrow{\simeq} \det F^{0,1} \otimes \Omega^1_Y(\log S)^{g_0},$$

By assumption there exists an invertible sheaf $\mathcal{L}$ with $\mathcal{L}^2 = \Omega^1_Y(\log S)$. Since $F^{1,0} \simeq F^{0,1\vee}$,

$$(\det F^{1,0})^2 \simeq \Omega^1_Y(\log S)^{g_0} = \mathcal{L}^{2 \cdot g_0},$$

and $\det F^{1,0} \otimes \mathcal{L}^{-g_0} = \mathcal{N}$ is of order two in $\operatorname{Pic}^0(Y)$.

If g_0 is even, $\mathcal{L}$ is uniquely determined up to the tensor product with two torsion points in $\operatorname{Pic}^0(Y)$.

If g_0 is odd, one replaces $\mathcal{L}$ by $\mathcal{L} \otimes \mathcal{N}$ and obtains $\det F^{1,0} = \mathcal{L}^{g_0}$.

By 1.2 the sheaf

$$\mathcal{T} = F^{1,0} \otimes \mathcal{L}^{-1}$$

is polystable of degree zero. 1.1 implies that the Higgs bundle $(\mathcal{T}, 0)$ corresponds to a local system $\mathbb{T}$, necessarily unitary.

Choose $\mathbb{L}$ to be the local system corresponding to the Higgs bundle

$$(\mathcal{L} \oplus \mathcal{L}^{-1}, \tau'), \quad \text{with} \quad \tau'_{1,0} : \mathcal{L} \xrightarrow{\simeq} \mathcal{L}^{-1} \otimes \Omega^1_Y(\log S).$$

The isomorphism

$$\tau_{1,0} : \mathcal{T} \otimes \mathcal{L} = F^{1,0} \xrightarrow{\simeq} F^{0,1} \otimes \Omega^1_Y(\log S) \xrightarrow{\simeq} F^{0,1} \otimes \mathcal{L}^2$$

induces an isomorphism

$$\phi : \mathcal{T} \otimes \mathcal{L}^{-1} \xrightarrow{\simeq} F^{0,1},$$

such that $\tau_{1,0} = \phi \circ (\mathrm{id}_{\mathcal{T}} \otimes \tau'_{1,0})$. Hence the Higgs bundles $(F^{1,0} \oplus F^{0,1}, \tau)$ and $(\mathcal{T} \otimes (\mathcal{L} \oplus \mathcal{L}^{-1}), \mathrm{id}_{\mathcal{T}} \otimes \tau')$ are isomorphic, and $\mathbb{V} \simeq \mathbb{T} \otimes_{\mathbb{C}} \mathbb{L}$. q.e.d.

Remark 1.6.

i. If $\deg \Omega^1_Y(\log S)$ is odd, hence $S \neq \emptyset$, and if the genus of Y is not zero, one may replace Y by an étale covering, in order to be able to apply 1.4. Doing so one may also assume that the invertible sheaf $\mathcal{N}$ in 1.4 c), is trivial.

ii. For $Y = \mathbb{P}^1$ and for $\mathbb{X}$ reaching the Arakelov bound, $\#S$ is always even. This, together with the decomposition 1.4, for $\mathbb{U} = \mathbb{C}^{g_0}$, can easily obtained in the following way: by 1.2, $F^{1,0}$ must be the direct sum of invertible sheaves $\mathcal{L}_i$, all of the same degree, say ν. Since $\tau_{1,0}$ is an isomorphism, the image $\tau_{1,0}(\mathcal{L}_i)$ is $\mathcal{O}_{\mathbb{P}^1}(2-s+\nu) \otimes \Omega_{\mathbb{P}^1}(\log S)$. Since $F^{0,1}$ is dual to $F^{1,0}$ one obtains $-\nu = 2-s+\nu$, and writing $\mathcal{L}_i^{-1} = \tau_{1,0}(\mathcal{L}_i)$,

$$(F^{1,0} \oplus F^{0,1}, \tau) \simeq \left(\bigoplus_i \mathcal{O}_{\mathbb{P}^1}(\nu) \oplus \mathcal{O}_{\mathbb{P}^1}(-\nu), \bigoplus_i \tau \right).$$

Consider now the local system of endomorphism $\mathbb{E}\mathrm{nd}(\mathbb{V})$ of $\mathbb{V}$, which is a polarized weight zero L-variation of Hodge structures. The Higgs bundle

$$(F^{1,0} \oplus F^{0,1}, \tau)$$

for $\mathbb{V}$ induces the Higgs bundle

$$(F^{1,-1} \oplus F^{0,0} \oplus F^{-1,1}, \tau_{1,-1} \oplus \tau_{0,0})$$

corresponding to $\mathbb{E}\mathrm{nd}(\mathbb{V}) = \mathbb{V} \otimes_{\mathbb{C}} \mathbb{V}^\vee$, by choosing

$$F^{1,-1} = F^{1,0} \otimes F^{0,1^\vee}, \quad F^{0,0} = F^{1,0} \otimes F^{1,0^\vee} \oplus F^{0,1} \otimes F^{0,1^\vee}$$
$$\text{and} \quad F^{-1,1} = F^{0,1} \otimes F^{1,0^\vee}.$$

The Higgs field is given by

$$\tau_{1,-1} = (-\mathrm{id}) \otimes {\tau_{1,0}}^\vee \oplus \tau_{1,0} \otimes \mathrm{id} \quad \text{and} \quad \tau_{0,0} = \tau_{1,0} \otimes \mathrm{id} \oplus (-\mathrm{id}) \otimes {\tau_{1,0}}^\vee.$$

Lemma 1.7. *Assume as in* 1.3 *that* $\mathbb{X}$ *reaches the Arakelov bound or equivalently that the Higgs field of* $\mathbb{V}$ *is maximal. Let*

$$F^{0,0}_u := \mathrm{Ker}(\tau_{0,0}) \quad \textit{and} \quad F^{0,0}_m = \mathrm{Im}(\tau_{1,-1}).$$

Then there is a splitting of the Higgs bundle

$$(F^{1,-1} \oplus F^{0,0} \oplus F^{-1,1}, \tau_{1,-1} \oplus \tau_{0,0})$$
$$= (F^{1,-1} \oplus F^{0,0}_m \oplus F^{-1,1}, \tau_{1,-1} \oplus \tau_{0,0}) \oplus (F^{0,0}_u, 0),$$

which corresponds to a splitting of the local system over $\mathbb{C}$

$$\mathbb{E}\mathrm{nd}(\mathbb{V}) = \mathbb{W} \oplus \mathbb{U}.$$

$\mathbb{U}$ is unitary of rank g_0^2 and a variation of Hodge structures concentrated in bidegree $0,0$, whereas $\mathbb{W}$ is a $\mathbb{C}$-variation of Hodge structures of weight zero and rank $3g_0^2$.

$$\tau_{1,-1} : F^{1,-1} \longrightarrow F_m^{0,0} \otimes \Omega^1_Y(\log S) \qquad \text{and}$$
$$\tau_{0,0} : F_m^{0,0} \longrightarrow F^{-1,1} \otimes \Omega^1_Y(\log S)$$

are both isomorphisms.

Proof. By definition, $(F_u^{0,0}, 0)$ is a sub-Higgs bundle of

$$(F^{1,-1} \oplus F^{0,0} \oplus F^{-1,1}, \tau_{1,-1} \oplus \tau_{0,0}).$$

We have an exact sequence

$$0 \longrightarrow F_u^{0,0} \longrightarrow F^{0,0} \xrightarrow{\tau_{0,0}} F^{-1,1} \otimes \Omega^1_Y(\log S).$$

Since $\tau_{1,0} \otimes \mathrm{id}$ is surjective, $\tau_{0,0}$ is surjective, and

$$\deg(F_u^{0,0}) = \deg(F^{0,0}) - \deg(F^{-1,1}) - \mathrm{rank}(F^{-1,1}) \cdot \deg(\Omega^1_Y(\log S)).$$

By the Arakelov equality,

$$\begin{aligned}\deg(F^{-1,1}) &= g_0 \cdot \deg(F^{0,1}) + g_0 \cdot \deg(F^{1,0^\vee}) \\ &= g_0^2 \cdot \deg(\Omega^1_Y(\log S)) = \mathrm{rank}(F^{-1,1}) \cdot \deg(\Omega^1_Y(\log S))\end{aligned}$$

and one finds $\deg(F_u^{0,0}) = \deg(F^{0,0}) = 0$.

By 1.1 $(F_u^{0,0}, 0)$, as a Higgs subbundle of degree zero with trivial Higgs field, corresponds to a unitary local subsystem $\mathbb{U}$ of $\mathbb{E}\mathrm{nd}(\mathbb{V})$. The exact sequence

$$0 \to F_u^{0,0} \to F^{0,0} \to F^{-1,1} \otimes \Omega^1_Y(\log S) \to 0$$

splits, and one obtains a direct sum decomposition of Higgs bundles

$$(F^{1,-1} \oplus F^{0,0} \oplus F^{-1,1}, \tau) = (F^{1,-1} \oplus F_m^{0,0} \oplus F^{-1,1}, \tau) \oplus (F_u^{0,0}, 0),$$

which induces the splitting on $\mathbb{E}\mathrm{nd}(\mathbb{V})$ with the desired properties.

q.e.d.

In 1.7 the local subsystem $\mathbb{W}$ of $\mathbb{E}\mathrm{nd}(\mathbb{V})$ has a maximal Higgs field in the following sense:

Definition 1.8. Let $\mathbb{W}$ be a $\mathbb{C}$-variation of Hodge structures of weight k, and let

$$(F, \tau) = \left(\bigoplus_{p+q=k} F^{p,q}, \bigoplus \tau_{p,q} \right)$$

be the corresponding Higgs bundle. Recall that the width is defined as

$$\text{width}(\mathbb{W}) = \text{Max}\{|p-q|;\ F^{p,q} \neq 0\}.$$

i. $\mathbb{W}$ (or (F,τ)) has a generically maximal Higgs field if $\text{width}(\mathbb{W}) > 0$ and if:
   a. $F^{p,k-p} \neq 0$ for all p with $|2p-k| \leq \text{width}(\mathbb{W})$.
   b. $\tau_{p,k-p} : F^{p,k-p} \to F^{p-1,k-p+1} \otimes \Omega^1_Y(\log S)$ is generically an isomorphism for all p with $|2p-k| \leq \text{width}(\mathbb{W})$ and $|2p-2-k| \leq \text{width}(\mathbb{W})$.

ii. $\mathbb{W}$ (or (F,τ)) has a maximal Higgs field if the $\tau_{p,k-p}$ in i) b. are all isomorphisms.

In particular, a variation of Hodge structures with a maximal Higgs field can not be unitary.

Properties 1.9.

a. *If $\mathbb{W}$ is a $\mathbb{C}$-variation of Hodge structures with a (generically) maximal Higgs field, and if $\mathbb{W}' \subset \mathbb{W}$ is a direct factor, then* $\text{width}(\mathbb{W}') = \text{width}(\mathbb{W})$ *and $\mathbb{W}'$ has again a (generically) maximal Higgs field.*

b. *Let $\mathbb{L}$ and $\mathbb{T}$ be two variations of Hodge structures with $\mathbb{L} \otimes \mathbb{T}$ of weight 1 and width 1, and with a (generically) maximal Higgs field. Then, choosing the bidegrees for $\mathbb{L}$ and $\mathbb{T}$ in an appropriate way, either $\mathbb{L}$ is a variation of Hodge structures concentrated in degree $0,0$, and $\mathbb{T}$ is a variation of Hodge structures of weight one and width one with a (generically) maximal Higgs field, or vice versa.*

Proof. For a) consider the Higgs field $(\bigoplus F'^{p,q}, \tau'_{p,q})$ of $\mathbb{W}'$, which is a direct factor of the one for $\mathbb{W}$. Since the $\tau_{p,q}$ are (generically) isomorphisms, a) is obvious.

In b) denote the components of the Higgs fields of $\mathbb{L}$ and $\mathbb{T}$ by $\mathcal{L}^{p_1,q_1}$ and $\mathcal{T}^{p_2,q_2}$, respectively. Shifting the bigrading one may assume that $p_1 = 0$ and $p_2 = 0$ are the smallest numbers with $\mathcal{L}^{p_1,q_1} \neq 0$ and $\mathcal{T}^{p_2,q_2} \neq 0$ and moreover that the corresponding $q_i \geq 0$. Since $q_1 + q_2 = 1$, one of q_i must be zero, let us say the first one.

Then $\mathcal{T}^{p_2,q_2}$ can only be nonzero, for $(p_2,q_2) = (0,1)$ or $= (1,0)$ and $\mathbb{L}$ is concentrated in degree $0,0$.

Obviously this forces the Higgs field of $\mathbb{L}$ to be zero. Then the Higgs field of $\mathbb{L} \otimes \mathbb{T}$ is the tensor product of the Higgs field

$$\mathcal{T}^{1,0} \to \mathcal{T}^{0,1} \otimes \Omega^1_Y(\log S)$$

with the identity on $\mathcal{L}^{0,0}$, hence the first one has to be (generically) an isomorphism. q.e.d.

Remark 1.10. The splitting in 1.7 can also be described by the tensor product decomposition $\mathbb{V} = \mathbb{T} \otimes_{\mathbb{C}} \mathbb{L}$ in 1.4 with $\mathbb{T}$ unitary and $\mathbb{L}$ a rank two variation of Hodge structures of weight one and with a maximal Higgs field. For any local system $\mathbb{M}$ one has a natural decomposition $\mathbb{E}\mathrm{nd}(\mathbb{M}) = \mathbb{E}\mathrm{nd}_0(\mathbb{M}) \oplus \mathbb{C}$, where $\mathbb{C}$ acts on $\mathbb{M}$ by multiplication. Applying 1.7 to $\mathbb{L}$ instead of $\mathbb{V}$, gives exactly the decomposition $\mathbb{E}\mathrm{nd}(\mathbb{L}) = \mathbb{E}\mathrm{nd}_0(\mathbb{L}) \oplus \mathbb{C}$. One obtains

$$\begin{aligned}\mathbb{E}\mathrm{nd}(\mathbb{V}) &= \mathbb{T} \otimes_{\mathbb{C}} \mathbb{T}^{\vee} \otimes_{\mathbb{C}} \mathbb{L} \otimes_{\mathbb{C}} \mathbb{L}^{\vee} \\ &= (\mathbb{E}\mathrm{nd}_0(\mathbb{T}) \oplus \mathbb{C}) \otimes_{\mathbb{C}} (\mathbb{E}\mathrm{nd}_0(\mathbb{L}) \oplus \mathbb{C}) \\ &= \mathbb{E}\mathrm{nd}_0(\mathbb{T}) \oplus \mathbb{C} \oplus \mathbb{E}\mathrm{nd}(\mathbb{T}) \otimes_{\mathbb{C}} \mathbb{E}\mathrm{nd}_0(\mathbb{L}).\end{aligned}$$

Here $\mathbb{E}\mathrm{nd}_0(\mathbb{T}) \oplus \mathbb{C}$ is unitary and $\mathbb{W} = \mathbb{E}\mathrm{nd}(\mathbb{T}) \otimes_{\mathbb{C}} \mathbb{E}\mathrm{nd}_0(\mathbb{L})$ has again a maximal Higgs field.

Remark 1.11. If one replaces $\mathbb{E}\mathrm{nd}(\mathbb{V})$ by the isomorphic local system $\mathbb{V} \otimes_{\mathbb{C}} \mathbb{V}$, one obtains the same decomposition. However, it is more natural to shift the weights by two, and to consider this as a variation of Hodge structures of weight 2.

A statement similar to 1.7 holds true for $\wedge^2(\mathbb{V})$. Here the Higgs bundle is given by

$$F'^{2,0} = F^{1,0} \wedge F^{1,0}, \quad F^{1,1} = F'^{1,0} \otimes F^{0,1} \quad \text{and} \quad F'^{0,2} = F^{0,1} \wedge F^{0,1}.$$

2. Shimura curves and the special Mumford–Tate group

Lemma 2.1. *Let $\mathbb{L}$ be a real variation of Hodge structures of weight 1 dimension 2, with a nontrivial Higgs field. Let*

$$\gamma_{\mathbb{L}} : \pi_1(U, *) \to \mathrm{Sl}(2, \mathbb{R})$$

be the corresponding representation and let $\Gamma_{\mathbb{L}}$ denote the image of $\gamma_{\mathbb{L}}$. Assume that the local monodromies around the points $s \in S$ are unipotent. Then the Higgs field of $\mathbb{L}$ is maximal if and only if

$$U = Y \setminus S \simeq \mathcal{H}/\Gamma_{\mathbb{L}}.$$

Proof. Writing $\mathcal{L}$ for the $(1,0)$ part, we have an nontrivial map

$$\tau_{1,0} : \mathcal{L} \longrightarrow \mathcal{L}^{-1} \otimes \Omega^1_Y(\log S). \tag{2.1.1}$$

Since $\mathcal{L}$ is ample, $\Omega^1_Y(\log S)$ is ample, hence the universal covering $\widetilde{U}$ of $U = Y \setminus S$ is the upper half plane $\mathcal{H}$. Let

$$\widetilde{U} = \mathcal{H} \xrightarrow{\widetilde{\varphi}} \mathcal{H}$$

be the period map. The tangent sheaf of the period domain $\mathcal{H}$ is given by the sheaf of homomorphisms from the $(1,0)$ part to the $(0,1)$ part of the variation of Hodge structures. Therefore $\tau_{1,0}$ is an isomorphism if and only if $\widetilde{\varphi}$ is a local diffeomorphism. Note that by Schmid [**23**] the Hodge metric on the Higgs bundle corresponding to $\mathbb{L}$ has logarithmic growth at S and bounded curvature. By the remarks following [**28**], Propositions 9.8 and 9.1, $\tau_{1,0}$ is an isomorphism if and only if $\widetilde{\varphi} : \widetilde{U} \to \mathcal{H}$ is a covering map, hence an isomorphism.

Obviously the latter holds true in case $Y \setminus S \simeq \mathcal{H}/\Gamma_{\mathbb{L}}$.

Assume that $\widetilde{\varphi}$ is an isomorphism. Since $\widetilde{\varphi}$ is an equivariant with respect to the $\pi_1(U,*)$-action on $\widetilde{U}$ and the $\mathrm{P}\rho_{\mathbb{L}}(\pi_1(U,*))$-action on $\mathcal{H}$, the homomorphism

$$\rho_{\mathbb{L}_{\mathbb{Z}}} : \pi_1(U,*) \longrightarrow \mathrm{P}\rho_{\mathbb{L}_{\mathbb{Z}}}(\pi_1(U,*)) \subset \mathrm{PSl}_2(\mathbb{R})$$

must be injective, hence an isomorphism. So $\widetilde{\varphi}$ descend to an isomorphism

$$\varphi : Y \setminus S \simeq \mathcal{H}/\Gamma_{\mathbb{L}}. \qquad \text{q.e.d.}$$

Proof of Proposition 0.1. $h : E \to Y$ be the semistable family of elliptic curves, reaching the Arakelov bound, smooth over U. Hence $\mathbb{L}_{\mathbb{Z}} = R^1 h_* \mathbb{Z}_{E_0}$ is a $\mathbb{Z}$-variation of Hodge structures of weight one and of rank two. Writing $\mathcal{L}$ for the $(1,0)$ part, we have an isomorphism

$$\tau_{1,0} : \mathcal{L} \longrightarrow \mathcal{L}^{-1} \otimes \Omega^1_Y(\log S). \tag{2.1.2}$$

Since $\mathcal{L}$ is ample, $\Omega^1_Y(\log S)$ is ample, hence the universal covering of U is the upper half plane $\mathcal{H}$. One obtains a commutative diagram

$$\begin{array}{ccc} \mathcal{H} & \xrightarrow{\widetilde{\varphi}} & \mathcal{H} \\ {\scriptstyle \psi'}\downarrow & & \downarrow{\scriptstyle \psi} \\ U & \xrightarrow{j} & \mathbb{C} \end{array}$$

where j is given by the j-invariant of the fibres of $E_0 \to U$, where ψ is the quotient map $\mathcal{H} \to \mathcal{H}/\mathrm{Sl}_2(\mathbb{Z})$, and where $\widetilde{\varphi}$ is the period map. 2.1 implies that

$$\varphi : U \longrightarrow \mathcal{H}/\rho_{\mathbb{L}_{\mathbb{Z}}}(\pi_1(U,*))$$

is an isomorphism, hence $\rho_{\mathbb{L}_{\mathbb{Z}}}(\pi_1(U,*)) \subset \mathrm{Sl}_2(\mathbb{Z})$ is of finite index, and $E \to Y$ is a semistable model of a modular curve. q.e.d.

Let us recall the description of wedge products of tensor products (see [**10**], p. 80). We will write $\lambda = \{\lambda_1, \ldots, \lambda_\nu\}$ for the partition of g_0 as $g_0 = \lambda_1 + \cdots + \lambda_\nu$. The partition λ defines a Young diagram and the Schur functor $\mathbb{S}_\lambda$. Assuming as in 1.4 that $\mathbb{L}$ is a local system of rank 2, and $\mathbb{T}$ a local system of rank g_0, both with trivial determinant, one has

$$\wedge^k(\mathbb{L} \otimes \mathbb{T}) = \bigoplus \mathbb{S}_\lambda(\mathbb{L}) \otimes \mathbb{S}_{\lambda'}(\mathbb{T})$$

where the sum is taken over all partitions λ of k with at most 2 rows and at most g_0 columns, and where λ' is the partition conjugate to λ. Similarly,

$$S^k(\mathbb{L} \otimes \mathbb{T}) = \bigoplus \mathbb{S}_\lambda(\mathbb{L}) \otimes \mathbb{S}_\lambda(\mathbb{T})$$

where the sum is taken over all partitions λ of k with at most 2 rows.

The only possible λ are of the form $\{k-a, a\}$, for $a \leq \frac{k}{2}$. By [**10**], 6.9 on p. 79,

$$\mathbb{S}_{\{k-a,a\}}(\mathbb{L}) = \begin{cases} \mathbb{S}_{\{k-2a\}}(\mathbb{L}) = S^{k-2a}(\mathbb{L}) \otimes \det(\mathbb{L})^a & \text{if} \quad 2a < k \\ \mathbb{S}_{\{a,a\}}(\mathbb{L}) = \det(\mathbb{L})^a & \text{if} \quad 2a = k. \end{cases}$$

For $k = g_0$ one obtains:

Lemma 2.2. *Assume that* $\det(\mathbb{L}) = \mathbb{C}$ *and* $\det(\mathbb{T}) = \mathbb{C}$.

a. *If g_0 is odd, then for some partitions λ_c,*

$$\bigwedge^{g_0}(\mathbb{L} \otimes \mathbb{T}) = \bigoplus_{c=0}^{\frac{g_0-1}{2}} S^{2c+1}(\mathbb{L}) \otimes \mathbb{S}_{\lambda_{2c}}(\mathbb{T}).$$

In particular, for $c = \frac{g_0-1}{2}$ one obtains

$$S^{g_0}(\mathbb{L}) \otimes \bigwedge^{g_0}(\mathbb{T}) = S^{g_0}(\mathbb{L})$$

as a direct factor.

b. *If g_0 is even, then for some partitions λ_c,*

$$\bigwedge^{g_0}(\mathbb{L} \otimes \mathbb{T}) = S^{g_0}(\mathbb{L}) \oplus \mathbb{S}_{\{2,\ldots,2\}}(\mathbb{T}) \oplus \bigoplus_{c=1}^{\frac{g_0}{2}-1} S^{2c}(\mathbb{L}) \otimes \mathbb{S}_{\lambda_{2c}}(\mathbb{T}).$$

Lemma 2.3. *Assume that* $\mathbb{L}$ *and* $\mathbb{T}$ *are variations of Hodge structures, with* $\mathbb{L}$ *of weight one, width one and with a maximal Higgs field, and with* $\mathbb{T}$ *pure of bidegree* $0, 0$.

a. *If k is odd,*

$$H^0\left(Y, \bigwedge^k(\mathbb{L}\otimes\mathbb{T})\right) = 0.$$

b. *If k is even, say $k = 2c$, then for some λ_c*

$$H^0\left(Y, \bigwedge^k(\mathbb{L}\otimes\mathbb{T})\right) = H^0(Y, \det(\mathbb{L})^c \otimes \mathbb{S}_{\lambda_c}(\mathbb{T})).$$

c. *For $k = 2$ one has in* b) $\lambda_1 = \{2\}$, *hence* $\mathbb{S}_{\lambda_1}(\mathbb{T}) = \bigwedge^2(\mathbb{T})$.

d. $H^0(Y, S^2(\mathbb{L}\otimes\mathbb{T})) = H^0(Y, \det(\mathbb{L}) \otimes \bigwedge^2(\mathbb{T}))$.

Proof. $S^\ell(\mathbb{L})$ has a maximal Higgs field for $\ell > 0$, whereas for all partitions λ' the variation of Hodge structures $\mathbb{S}_{\lambda'}(\mathbb{T})$ is again pure of bidegree $0, 0$. By 1.9 a), $S^\ell(\mathbb{L}) \otimes \mathbb{S}_{\lambda'}(\mathbb{T})$ has no global sections. Hence $\bigwedge^k(\mathbb{L}\otimes\mathbb{T})$ can only have global sections for k even. In this case, the global sections lie in

$$\det(\mathbb{L})^c \otimes \mathbb{S}_{\lambda_c}(\mathbb{T}),$$

for some partition λ_c, and one obtains a) and b). For $k = 2$ one finds $\lambda_1 = \{2\}$. For d) one just has the two partitions $\{1, 1\}$ and $\{2\}$. Again, the direct factor $S^2(\mathbb{L}) \otimes S^2(\mathbb{T})$, corresponding to the first one, has no global section. q.e.d.

Let us shortly recall Mumford's definition of the Hodge group, or as one writes today, the special Mumford–Tate group (see [**17**], [**18**], and also [**5**] and [**24**]). Let B be an abelian variety and $H^1(B, \mathbb{Q})$ and Q the polarization on V. The special Mumford–Tate group $\mathrm{Hg}(B)$ is defined in [**17**] as the smallest $\mathbb{Q}$-algebraic subgroup of $\mathrm{Sp}(H^1(B, \mathbb{R}), Q)$, which contains the complex structure. Equivalently $\mathrm{Hg}(B)$ is the largest $\mathbb{Q}$-algebraic subgroup of $\mathrm{Sp}(H^1(B, \mathbb{Q}), Q)$, which leaves all Hodge cycles of $B \times \cdots \times B$ invariant, hence all elements

$$\eta \in H^{2p}(B \times \cdots \times B, \mathbb{Q})^{p,p} = \left[\bigwedge^{2p}(H^1(B, \mathbb{Q}) \oplus \cdots \oplus H^1(B, \mathbb{Q}))\right]^{p,p}.$$

For a smooth family of abelian varieties $f : X_0 \to U$ with $B = f^{-1}(y)$ for some $y \in U$, and for the corresponding $\mathbb{Q}$-variation of polarized Hodge structures $R^1 f_* \mathbb{Q}_{X_0}$ consider Hodge cycles η on B which remain Hodge cycles under parallel transform. One defines the special Mumford–Tate group $\mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$ as the largest $\mathbb{Q}$-subgroup of $\mathrm{Sp}(H^1(B, \mathbb{Q}), Q)$ which leaves all those Hodge cycles invariant ([**5**], §7, or [**24**], 2.2).

Lemma 2.4.

a. *For all $y \in U$ the special Mumford–Tate group* $\mathrm{Hg}(f^{-1}(y))$ *is a subgroup of* $\mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$. *For all y in the complement U' of the union of countably many proper closed subsets it coincides with* $\mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$.

b. *Let G^{Mon} denote the smallest reductive $\mathbb{Q}$-subgroup of*

$$\mathrm{Sp}(H^1(B, \mathbb{R}), Q),$$

containing the image Γ of the monodromy representation

$$\gamma : \pi_0(U) \longrightarrow \mathrm{Sp}(H^1(B, \mathbb{R}), Q).$$

Then the connected component G_0^{Mon} of one in G^{Mon} is a subgroup of $\mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$.

c. *If $f : X \to Y$ reaches the Arakelov bound, and if $R^1 f_* \mathbb{C}_X$ has no unitary part, then* $G_0^{\mathrm{Mon}} = \mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$.

Proof. The first statement of a) has been verified in [**24**], 2.3., and the second in [**16**] 1.2. As explained in [**5**], §7, or [**24**], 2.4, the Mumford–Tate group contains a subgroup of Γ of finite index, hence b) holds true. It is easy to see, that the same holds true for the special Mumford–Tate group (called Hodge group in [**17**]) by using the same argument.

Since the special Mumford–Tate group of an abelian variety is reductive, a) implies that $\mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$ is reductive. So $G_0^{\mathrm{Mon}} \subset \mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0})$ is an inclusion of reductive groups. The proof of 3.1, (c), in [**6**] carries over to show that both groups are equal, if they leave the same tensors

$$\eta \in \left[\bigwedge^{2p} (H^1(B, \mathbb{Q}) \oplus \cdots \oplus H^1(B, \mathbb{Q})) \right]$$

invariant.

Let $\eta \in H^k(B, \mathbb{Q})$ be invariant under Γ, and let $\widetilde{\eta}$ be the corresponding global section of

$$\bigwedge^k (R^1 f_* \mathbb{Q}_{X_0}) = \bigwedge^k (\mathbb{L} \otimes \mathbb{T}).$$

By 2.3 a) and b), one can only have global sections for $k = 2c$, and those lie in

$$\det(\mathbb{L})^c \otimes \mathbb{S}_{\lambda_c}(\mathbb{T}).$$

In particular they are of pure bidegree c, c.

The same argument holds true, if one replaces B and $f : X \to Y$ by any product, which implies c). q.e.d.

For the Hodge group $\mathrm{Hg}(R^1 f_* \mathbb{Q}_X) = \mathrm{Hg} \subset \mathrm{Sp}(2g, \mathbb{Q})$, as in Lemma 2.4 Mumford considers the moduli functor $\mathcal{M}(\mathrm{Hg})$ of isomorphy classes of polarized abelian varieties with special Mumford–Tate group equal to a subgroup of Hg. He shows that $\mathcal{M}(\mathrm{Hg})$ admits a quasi-projective coarse moduli space $M(\mathrm{Hg})$, which lies in the coarse moduli space of polarized abelian varieties A_g. By Mumford ([**17**], Section 3, [**18**], Sections 1–2)

$$M(\mathrm{Hg}) = \Gamma \backslash \mathrm{Hg}(\mathbb{R})/K$$

where K is a maximal compact subgroup of $\mathrm{Hg}(\mathbb{R})$, and Γ an arithmetic subgroup of $\mathrm{Hg}(\mathbb{Q})$. The embedding $M(\mathrm{Hg}) \hookrightarrow A_g$ is a totally geodesic embedding, and $M(\mathrm{Hg})$ is a Shimura variety of Hodge Type Hg.

Let $f : X_0 \to U$ be a family of abelian varieties with the special Mumford–Tate group $\mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_0}) = \mathrm{Hg}$. By Lemma 2.4 a), f induces a morphism

$$U \to M(\mathrm{Hg}).$$

Proof of 0.3. By Proposition 1.4 the image of the monodromy representation of f lies in $\mathrm{Sl}_2(\mathbb{R}) \times \mathbf{G}$, for some compact group $\mathbf{G}$, and its Zariski closure is $\mathrm{Sl}_2(\mathbb{R}) \times \mathbf{G}$. Hence, $G_0^{\mathrm{Mon}}(\mathbb{R})$ is again the product of $\mathrm{Sl}_2(\mathbb{R})$ with a compact group. Lemma 2.4 c), implies that

$$Hg(\mathbb{R}) = \mathrm{Sl}_2(\mathbb{R}) \times \mathbf{G}'$$

for a compact group $\mathbf{G}'$, hence $\mathrm{Hg}(\mathbb{R})/K \simeq \mathrm{Sl}_2(\mathbb{R})/\mathrm{SO}_2$ is the upper half plane $\mathcal{H}$.

In particular, $\dim M(\mathrm{Hg}) = 1$. Since we assumed the family to be nonisotrivial and semistable, the morphism $U \to M(\mathrm{Hg})$ is surjective.

Consider the composition $\phi : U \to M(\mathrm{Hg}) \to A_g$. Replacing U by an étale covering, we may assume that $X_0 \to U$ is the pullback of a universal family of abelian varieties, defined over an étale covering A'_g of A_g. The pull back of the tangent bundle on A_g via ϕ is just

$$\phi^* T_{A_g} = S^2 E^{0,1} \subset E^{0,1 \otimes 2}.$$

The differential $d\phi : T_U \to \phi^* T_{A_g} \subset E^{0,1 \otimes 2}$ is induced by the Kodaira-Spencer map $E^{1,0} \otimes T_U \to E^{0,1}$. By Proposition 1.4

$$E^{1,0} \oplus E^{0,1} = (\mathcal{L} \oplus \mathcal{L}^{-1}) \otimes \mathbb{T},$$

and the map $d\phi : T_U \to E^{0,1 \otimes 2}$ lies in the component

$$d\phi : T_U \simeq \mathcal{L}^{\otimes -2} \subset \mathcal{L}^{\otimes -2} \otimes \mathrm{End}(\mathbb{T}).$$

This implies that the differential of the map $U \to M(\mathrm{Hg})$ is no where vanishing, hence $U \to M(\mathrm{Hg})$ is étale. q.e.d.

Remarks 2.5.

a) As is well-known (see [**17**], [**18**]) the moduli space of abelian varieties with a given special Mumford–Tate group is necessarily a Satake holomorphic embedding. Hence the assumptions made in Proposition 0.3 imply in particular that the period map from U to the corresponding moduli space of abelian varieties with a fixed level structure is a Satake holomorphic embedding.

b) Presumably Proposition 0.3 can also be obtained using [**1**]. Using Proposition 1.4 the maximality of the Higgs field should imply that the period map from $U = \mathcal{H}/\Gamma$ to the Siegel upper half plane is a rigid, totally geodesic, and equivariant holomorphic map. Then [**1**], Theorem 3.4, implies that $f : X \to Y$ is a family of Mumford type, and as mentioned in the introduction one can finish the proof of Theorem 0.5 going through the classification of Shimura varieties.

c) Without the assumption of rigidity, hidden behind the one saying that the maximal unitary local subsystem is defined over $\mathbb{Q}$, we do not see a way to show directly, that the families are Kuga fibre spaces. One needs a precise description of the $\mathbb{Z}$-structure on the decompositions of the variation of Hodge structures. On the other hand, the latter will allow one to prove Theorems 0.5 and 0.7 directly.

d) Theorems 0.5 and 0.7 imply that all families $f : X \to Y$ with maximal Higgs fields are Kuga fibre spaces, and that the period map is again a Satake holomorphic embedding.

3. Splitting over $\overline{\mathbb{Q}}$

Up to now, we considered local systems of $\mathbb{C}$-vector spaces induced by the family of abelian varieties. We say that a $\mathbb{C}$-local system $\mathbb{M}$ is defined over a subring R of $\mathbb{C}$, if there exists a local system $\mathbb{M}_R$ of torsion-free R-modules with $\mathbb{M} = \mathbb{M}_R \otimes_R \mathbb{C}$. In different terms, the representation

$$\gamma_{\mathbb{M}} : \pi_0(U, *) \longrightarrow \mathrm{Gl}(\mu, \mathbb{C})$$

is conjugate to one factoring like

$$\gamma_{\mathbb{M}} : \pi_0(U, *) \longrightarrow \mathrm{Gl}(\mu, R) \longrightarrow \mathrm{Gl}(\mu, \mathbb{C}).$$

If $\mathbb{M}$ is defined over R, and if $\sigma : R \to R'$ is a ring isomorphism, we will write $\mathbb{M}_R^\sigma$ for the local system defined by

$$\gamma_{\mathbb{M}} : \pi_0(U, *) \longrightarrow \mathrm{Gl}(\mu, R) \xrightarrow{\sigma} \mathrm{Gl}(\mu, R'),$$

and $\mathbb{M}^\sigma = \mathbb{M}_R^\sigma \otimes_{R'} \mathbb{C}$. In this section we want to show, that the splittings $\mathbb{X} = \mathbb{V} \oplus \mathbb{U}_1$ and $\mathbb{E}\mathrm{nd}(\mathbb{V}) = \mathbb{W} \oplus \mathbb{U}$ considered in the last section are defined over $\overline{\mathbb{Q}}$, i.e., that there exists a number field K containing the field of definition for $\mathbb{X}$ and local K-subsystems

$$\mathbb{V}_K \subset \mathbb{X}_K,\ \mathbb{U}_{1K} \subset \mathbb{X}_K,\ \mathbb{W}_K \subset \mathbb{E}\mathrm{nd}(\mathbb{X}_K) \ \text{ and } \ \mathbb{U}_K \subset \mathbb{E}\mathrm{nd}(\mathbb{X}_K)$$

with

$$\mathbb{X}_K = \mathbb{X}_L \otimes_L K = \mathbb{V}_K \oplus \mathbb{U}_{1K}, \quad \mathbb{V}_K = \mathbb{W}_K \oplus \mathbb{U}_K, \quad \text{and with}$$
$$\mathbb{V} = \mathbb{V}_K \otimes_K \mathbb{C}, \quad \mathbb{U}_1 = \mathbb{U}_{1K} \otimes_K \mathbb{C} \quad \mathbb{W} = \mathbb{W}_K \otimes_K \mathbb{C}, \quad \mathbb{U} = \mathbb{U}_K \otimes_K \mathbb{C}.$$

We start with a simple observation. Suppose that $\mathbb{M}$ is a local system defined over a number field L. The local system $\mathbb{M}_L$ is given by a representation $\rho : \pi_1(U, *) \to \mathrm{Gl}(M_L)$ for the fibre M_L of $\mathbb{M}_L$ over the base point $*$.

Fixing a positive integer r, let $\mathcal{G}(r, \mathbb{M})$ denote the set of all rank-r local subsystems of $\mathbb{M}$ and let $\mathrm{Grass}(r, M_L)$ be the Grassmann variety of r-dimensional subspaces. Then $\mathcal{G}(r, \mathbb{M})$ is the subvariety of

$$\mathrm{Grass}(r, M_L) \times_{\mathrm{Spec}(L)} \mathrm{Spec}(\mathbb{C})$$

consisting of the $\pi_1(U, *)$-invariant points. In particular, it is a projective variety defined over L. A K-valued point of $\mathcal{G}(r, \mathbb{M})$ corresponds to a local subsystem of $\mathbb{M}_K = \mathbb{M}_L \otimes_L K$. One obtains the following well-known property:

Lemma 3.1. *If $[\mathbb{W}] \in \mathcal{G}(r, \mathbb{M})$ is an isolated point, then $\mathbb{W}$ is defined over $\overline{\mathbb{Q}}$.*

In the proof of 3.7 we will also need:

Lemma 3.2. *Let $\mathbb{M}$ be a variation of Hodge structures defined over L, and let $\mathbb{W} \subset \mathbb{M}$ be an irreducible local subsystem of rank r defined over $\mathbb{C}$,. Then $\mathbb{W}$ can be deformed to a local subsystem $\mathbb{W}_t \subset \mathbb{M}$, which is isomorphic to $\mathbb{W}$ and which is defined over a finite extension of L.*

Proof. By [**4**] $\mathbb{M}$ is completely reducible over $\mathbb{C}$. Hence we have a decomposition $\mathbb{M} = \mathbb{W} \oplus \mathbb{W}'$.

The space $\mathcal{G}(r, \mathbb{M})$ of rank r local subsystems of $\mathbb{M}$ is defined over L and the subset

$$\{\mathbb{W}_t \in \mathcal{G}(r, \mathbb{M});\ \text{the composit } \mathbb{W}_t \subset \mathbb{W} \oplus \mathbb{W}' \xrightarrow{\mathrm{pr}_1} \mathbb{W} \text{ is nonzero}\}$$

forms a Zariski open subset. So there exists some $\mathbb{W}_t$ in this subset, which is defined over some finite extension of L. Since $p : \mathbb{W}_t \to \mathbb{W}$ is nonzero, $\mathrm{rank}(\mathbb{W}_t) = \mathrm{rank}(\mathbb{W})$, and since $\mathbb{W}$ is irreducible, p is an isomorphism. q.e.d.

Lemma 3.3. *Let $\mathbb{M}$ be the underlying local system of a variation of Hodge structures of weight k defined over a number field L. Assume that there is a decomposition*

$$\mathbb{M} = \mathbb{U} \oplus \bigoplus_{i=1}^{\ell} \mathbb{M}_i \tag{3.3.1}$$

in subvariations of Hodge structures, and let

$$(E, \theta) = (N, 0) \oplus \bigoplus_{i=1}^{\ell} (F_i, \tau_i = \theta|_{F_i}) \tag{3.3.2}$$

be the induced decomposition of the Higgs field. Assume that $\mathrm{width}(\mathbb{M}_i) = i$*, and that the $\mathbb{M}_i$ have all generically maximal Higgs fields. Then the decomposition* (3.3.1) *is defined over $\overline{\mathbb{Q}}$. If L is real, it is defined over $\overline{\mathbb{Q}} \cap \mathbb{R}$. If $\mathbb{M}$ is polarized, then the decomposition* (3.3.1) *can be chosen to be orthogonal with respect to the polarization.*

Proof. Consider a family $\{\mathbb{W}_t\}_{t \in \Delta}$ of local subsystems of $\mathbb{M}$ defined over a disk Δ with $\mathbb{W}_0 = \mathbb{M}_\ell$. For $t \in \Delta$ let $(F_{\mathbb{W}_t}, \tau_t)$ denote the Higgs bundle corresponding to $\mathbb{W}_t$. Hence $(F_{\mathbb{W}_t}, \tau_t)$ is obtained by restricting the F-filtration of $\mathbb{M} \otimes \mathcal{O}_U$ to $\mathbb{W}_t \otimes \mathcal{O}_U$ and by taking the corresponding graded sheaf. So the Higgs map

$$\tau_{p,k-p} : F_t^{p,k-p} \longrightarrow F_t^{p-1,k-p+1} \otimes \Omega^1_Y(\log S)$$

will again be generically isomorphic for t sufficiently closed to 0 and

$$|2p - k| \leq \ell \quad \text{and} \quad |2p - 2 - k| \leq \ell.$$

If the projection

$$\rho : \mathbb{W}_t \longrightarrow \mathbb{M} = \mathbb{U} \oplus \bigoplus_{i=1}^{\ell} \mathbb{M}_i \longrightarrow \mathbb{U} \oplus \bigoplus_{i=1}^{\ell-1} \mathbb{M}_i$$

is nonzero, the complete reducibility of local systems coming from variations of Hodge structures (see [**4**]) implies that $\mathbb{W}_t$ contains an irreducible nontrivial direct factor, say $\mathbb{W}'_t$ which is isomorphic to a direct factor of $\mathbb{U}$ or of one of the local systems $\mathbb{M}_i$, for $i < \ell$.

Restricting again the F filtration and passing to the corresponding graded sheaf, we obtain a Higgs bundle $(F_{\mathbb{W}'_t}, \tau'_t)$ with trivial Higgs field, or whose width is strictly smaller than ℓ. On the other hand, $(F_{\mathbb{W}'_t}, \tau'_t)$ is a sub-Higgs bundle of the Higgs bundle $(F_{\mathbb{W}_t}, \tau_t)$ of width ℓ, a contradiction. So ρ is zero and $\mathbb{W}_t = \mathbb{M}_\ell$.

Thus $\mathbb{M}_\ell$ is rigid as a local subsystem of $\mathbb{M}$, and by Lemma 3.1 $\mathbb{M}_\ell$ is defined over $\overline{\mathbb{Q}}$.

Assume now that L is real, hence $\mathbb{M} = \mathbb{M}_{\mathbb{R}} \otimes \mathbb{C}$. The complex conjugation defines an involution ι on $\mathbb{M}$. Let $\overline{\mathbb{M}}_\ell$ denote the image of $\mathbb{M}_\ell$ under ι. Then $\overline{\mathbb{M}}_\ell$ has again generically isomorphic Higgs maps $\tau_{p,k-p}$, for

$$|2p - k| \leq \ell \quad \text{and} \quad |2p - 2 - k| \leq \ell.$$

If $\overline{\mathbb{M}}_\ell \neq \mathbb{M}_\ell$, repeating the argument used above, one obtains a map

$$\overline{\mathbb{M}}_\ell \longrightarrow \mathbb{U} \oplus \bigoplus_{i=1}^{\ell-1} \mathbb{M}_i,$$

from a Higgs bundle of width ℓ and with a maximal Higgs field to one with trivial Higgs field or of lower width. Again such a morphism must be zero, hence $\mathbb{M}_\ell = \overline{\mathbb{M}}_\ell$ in this case.

So we can find a number field K, real in case L is real, and a local system $\mathbb{M}_{\ell,K} \subset \mathbb{M}_K$ with $\mathbb{M}_\ell = \mathbb{M}_{\ell,K} \otimes_K \mathbb{C}$. The polarization on $\mathbb{M}$ restricts to a nondegenerated intersection form on $\mathbb{M}_K$. Choosing for $\mathbb{M}^\perp_{\ell,K}$ the orthogonal complement of $\mathbb{M}_{\ell,K}$ in $\mathbb{M}_K$ we obtain a splitting

$$\mathbb{M}_K = \mathbb{M}_{\ell,K} \oplus \mathbb{M}^\perp_{\ell,K}$$

inducing over $\mathbb{C}$ the splitting of the factor $\mathbb{M}_\ell$ in (3.3.1). By induction on ℓ we obtain 3.3. q.e.d.

For a reductive algebraic group G and for a finitely generated group Γ let $\mathcal{M}(\Gamma, G)$ denote the moduli space of reductive representations of Γ in G.

Theorem 3.4 (Simpson, [**29**], Corollary 9.18)**.** *Suppose Γ is a finitely generated group. Suppose $\phi : G \to H$ is a homomorphism of reductive*

algebraic groups with finite kernel. Then the resulting morphism of moduli spaces

$$\phi : \mathcal{M}(\Gamma, G) \longrightarrow \mathcal{M}(\Gamma, H)$$

is finite.

Corollary 3.5. *Let Γ be $\pi_1(Y, *)$ of a projective manifold, and $\gamma : \Gamma \to G$ be a reductive representation. If $\phi\gamma \in \mathcal{M}(\Gamma, H)$ comes from a $\mathbb{C}$-variation of Hodge structures, then γ comes from a $\mathbb{C}$-variation of Hodge structures as well.*

Proof. By Simpson a reductive local system is coming from a variation of Hodge structures if and only if the isomorphism class of the corresponding Higgs bundle is a fix point of the $\mathbb{C}^*$-action. Since the $\mathbb{C}^*$-action contains the identity and since it is compatible with ϕ, the finiteness of the preimage $\phi^{-1}\phi(\gamma)$ implies that the isomorphism class of the Higgs bundle corresponding to γ is fixed by the $\mathbb{C}^*$-action, as well. q.e.d.

Definition 3.6. Let $\mathbb{M}$ be a local system of rank r, and defined over $\overline{\mathbb{Q}}$. Let $\gamma_{\mathbb{M}} : \pi_1(U, *) \to \mathrm{Sl}(2, \overline{\mathbb{Q}})$ be the corresponding representation of the fundamental group. For $\eta \in \pi_1(U, *)$ we write $\mathrm{tr}(\gamma_{\mathbb{M}}(\eta)) \in \overline{\mathbb{Q}}$ for the trace of η and

$$\mathrm{tr}(\mathbb{M}) = \{\mathrm{tr}(\gamma_{\mathbb{M}}(\eta));\ \eta \in \pi_1(U, *)\}.$$

Corollary 3.7. *Under the assumptions made in* 1.3:

i. *The splitting $\mathbb{X} = \mathbb{V} \oplus \mathbb{U}_1$ is defined over $\overline{\mathbb{Q}}$, and over $\overline{\mathbb{Q}} \cap \mathbb{R}$ in case L is real. If $\mathbb{X}$ is polarized, it can be chosen to be orthogonal.*
ii. *The splitting $\mathbb{E}\mathrm{nd}(\mathbb{V}) = \mathbb{W} \oplus \mathbb{U}$ constructed in Lemma* 1.7 *is defined over $\overline{\mathbb{Q}}$, and over $\overline{\mathbb{Q}} \cap \mathbb{R}$ in case L is real. If $\mathbb{X}$ is polarized, it can be chosen to be orthogonal.*
iii. *Replacing Y by an étale covering Y', one can choose the decomposition $\mathbb{V} \simeq \mathbb{L} \otimes \mathbb{T}$ in* 1.4 *such that:*
   a. *$\mathbb{L}$ and $\mathbb{T}$ are defined over a number field K, real if L is real.*
   b. *One has an isomorphism $\mathbb{V}_{\overline{\mathbb{Q}}} \simeq \mathbb{L}_{\overline{\mathbb{Q}}} \otimes_{\overline{\mathbb{Q}}} \mathbb{T}_{\overline{\mathbb{Q}}}$.*
   c. *$\mathrm{tr}(\mathbb{L})$ is a subset of the ring of integers $\mathcal{O}_K$ of K.*

Proof. i) and ii) are direct consequences of 3.3. For iii) let us first remark that for L real, passing to an étale covering $\mathbb{L}$ and $\mathbb{T}$ can both be assumed to be defined over $\mathbb{R}$. In fact, the local system $\overline{\mathbb{L}}$ has a maximal Higgs field, hence its Higgs field is of the form $(\mathcal{L}' \oplus \mathcal{L}'^{-1}, \tau')$ where $\mathcal{L}'$ is a theta characteristic. Hence it differs from $\mathcal{L}$ at most by the tensor

product with a two torsion point in $\mathrm{Pic}^0(Y)$. Replacing Y by an étale covering, we may assume $\mathbb{L} = \overline{\mathbb{L}}$. From 1.4 d), we obtain $\overline{\mathbb{T}} = \mathbb{T}$.

Consider the isomorphism of local systems $\phi : \mathbb{L} \otimes \mathbb{T} \xrightarrow{\simeq} \mathbb{V}$ and the induced isomorphism

$$\phi^2 : \mathbb{E}\mathrm{nd}_0(\mathbb{L}\otimes\mathbb{T}) = \mathbb{E}\mathrm{nd}_0(\mathbb{L})\oplus\mathbb{E}\mathrm{nd}_0(\mathbb{T})\otimes\mathbb{E}\mathrm{nd}_0(\mathbb{L})\oplus\mathbb{E}\mathrm{nd}_0(\mathbb{T}) \longrightarrow \mathbb{E}\mathrm{nd}(\mathbb{V}).$$

Since $\phi^2\mathbb{E}\mathrm{nd}_0(\mathbb{T})$ is the unitary part of this decomposition, by 3.3 it is defined over $\overline{\mathbb{Q}} \cap \mathbb{R}$, as well as $\phi^2(\mathbb{E}\mathrm{nd}_0(\mathbb{L}) \oplus \mathbb{E}\mathrm{nd}_0(\mathbb{T}) \otimes \mathbb{E}\mathrm{nd}_0(\mathbb{L}))$. The $1, -1$ part of the Higgs field corresponding to $\phi^2\mathbb{E}\mathrm{nd}_0(\mathbb{L})$ has rank one, and its Higgs field is maximal. Hence $\phi^2\mathbb{E}\mathrm{nd}_0(\mathbb{L})$ is irreducible, and by 3.2 it is isomorphic to a local system, defined over $\overline{\mathbb{Q}}$. Hence $\mathbb{T} \otimes \mathbb{T} \simeq \mathbb{E}\mathrm{nd}(\mathbb{T})$ and $\mathbb{L} \otimes \mathbb{L} \simeq \mathbb{E}\mathrm{nd}(\mathbb{L})$ are both isomorphic to local systems defined over some real number field K'. An $\mathcal{O}_{K'}$-structure can be defined by

$$\phi^2(\mathbb{E}\mathrm{nd}(\mathbb{L}))_{\mathcal{O}_{K'}} = \phi^2(\mathbb{E}\mathrm{nd}(\mathbb{L}))_K \cap \mathbb{V}_{\mathcal{O}_{K'}}.$$

Consider for $\nu = 2$ or $\nu = g_0$ the moduli space $\mathcal{M}(U, \mathrm{Sl}(\nu^2))$ of reductive representations of $\pi(U, *)$ into $Sl(\nu^2)$. It is a quasi-projective variety defined over $\mathbb{Q}$. The fact that $\mathbb{L}\otimes\mathbb{L}$ (or $\mathbb{T}\otimes\mathbb{T}$) is defined over $\overline{\mathbb{Q}}$ implies that its isomorphy class in $\mathcal{M}(U, \mathrm{Sl}(\nu^2))$ is a $\overline{\mathbb{Q}}$ valued point.

Consider the morphism induced by the second tensor product

$$\rho : \mathcal{M}(U, \mathrm{Sl}(\nu)) \longrightarrow \mathcal{M}(U, \mathrm{Sl}(\nu^2))$$

which is clearly defined over $\mathbb{Q}$. By Theorem 3.4, ρ is finite, hence the fibre $\rho^{-1}([\mathbb{L} \otimes \mathbb{L}])$ (or $\rho^{-1}([\mathbb{T} \otimes \mathbb{T}])$) consists of finitely many $\overline{\mathbb{Q}}$-valued points, hence $\mathbb{L}$ and $\mathbb{T}$ can be defined over a number field K. If L is real, as already remarked above, we may choose K to be real.

Obviously, for $\rho \in \pi_1(Y, *)$ one has

$$\mathrm{tr}(\gamma_{\mathbb{L}}(\rho))^2 = \mathrm{tr}(\gamma_{\mathbb{L}\otimes\mathbb{L}}(\rho)).$$

In fact, one may assume that $\gamma_{\mathbb{L}}(\rho)$ is a diagonal matrix with entries a and b on the diagonal. Then $\mathrm{tr}(\gamma_{\mathbb{L}\otimes\mathbb{L}}(\rho))$ has a^2, b^2, ab and ba as diagonal elements. Since $\mathrm{tr}(\gamma_{\mathbb{L}\otimes\mathbb{L}}(\rho)) \in \mathcal{O}_{K'}$ we find $\mathrm{tr}(\gamma_{\mathbb{L}}(\rho)) \in \mathcal{O}_K$.

q.e.d.

4. Splitting over $\mathbb{Q}$ for $S \neq \emptyset$ and isogenies

In this section, we will consider the case $L = \mathbb{Q}$ and $\mathbb{X}_{\mathbb{Q}} = R^1 f_* \mathbb{Q}_{X_0}$, where $f : X \to Y$ is a family of abelian varieties, $S = Y \setminus U \neq \emptyset$, and where the restriction $X_0 \to U$ of f is a smooth family.

Lemma 4.1. *Assume that $S \neq \emptyset$ and let $\mathbb{M}_{\mathbb{Q}}$ be a $\mathbb{Q}$-variation of Hodge structures of weight k and with unipotent monodromy around all points $s \in S$. Assume that over some number field K there exists a splitting*

$$\mathbb{M}_K = \mathbb{M}_{\mathbb{Q}} \otimes_{\mathbb{Q}} K = \mathbb{W}_K \oplus \mathbb{U}_K$$

where $\mathbb{U} = \mathbb{U}_K \otimes_K \mathbb{C}$ is unitary and where the Higgs field of $\mathbb{W} = \mathbb{W}_K \otimes_K \mathbb{C}$ is maximal. Then $\mathbb{W}$, $\mathbb{U}$ and the decomposition $\mathbb{M} = \mathbb{W} \oplus \mathbb{U}$ are defined over $\mathbb{Q}$. Moreover, $\mathbb{U}$ extends to a local system over Y.

Proof. Let $\mathbb{T}$ be a local subsystem of $\mathbb{W}$. Writing

$$\left(\bigoplus_{p+q=k} F_{\mathbb{T}}^{p,q}, \bigoplus_{p+q=k} \tau_{p,q} \right),$$

for the Higgs bundle corresponding to $\mathbb{T}$, the maximality of the Higgs field for $\mathbb{W}$ implies that the Higgs field for $\mathbb{T}$ is maximal, as well. In particular, for all $s \in S$ and for $p > 0$ the residue maps

$$\mathrm{res}_s(\tau_{p,q}) : F_{\mathbb{T},s}^{p,q} \longrightarrow F_{\mathbb{T},s}^{p-1,q+1}$$

are isomorphisms. By [**27**] the residues of the Higgs field at s are defined by the nilpotent part of the local monodromy matrix around s. Hence if γ is a small loop around s in Y, and if $\rho_{\mathbb{T}}(\gamma)$ denotes the image of γ under a representation of the fundamental group, defining $\mathbb{T}$, the nilpotent part $N(\rho_{\mathbb{T}}(\gamma)) = \log \rho_{\mathbb{T}}(\gamma)$ of $\rho_{\mathbb{T}}(\gamma)$ has to be nontrivial.

We may assume that K is a Galois extension of $\mathbb{Q}$. Recall that for $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ we denote the local systems obtained by composing the representation with σ by an upper index σ. Consider the composite

$$p : \mathbb{U}_K^{\sigma} \longrightarrow \mathbb{M}_K = \mathbb{W}_K \oplus \mathbb{U}_K \longrightarrow \mathbb{W}_K,$$

and the induced map $\mathbb{U}^{\sigma} = \mathbb{U}_K^{\sigma} \otimes_K \mathbb{C} \to \mathbb{W}$.

Let γ be a small loop around $s \in S$, and let $\rho_{\mathbb{U}}(\gamma)$ and $\rho_{\mathbb{U}^{\sigma}}(\gamma)$ be the images of γ under the representations defining $\mathbb{U}$ and $\mathbb{U}^{\sigma}$ respectively. Since $\mathbb{U}$ is unitary and unipotent, the nilpotent part of the monodromy matrix $N(\rho_{\mathbb{U}}(\gamma)) = 0$. This being invariant under conjugation, $N(\rho_{\mathbb{U}^{\sigma}}(\gamma))$ is zero, as well as $N(\rho_{p(\mathbb{U}^{\sigma})}(\gamma))$.

Therefore $p(\mathbb{U}^{\sigma}) = 0$, hence $\mathbb{U}^{\sigma} = \mathbb{U}$, and $\mathbb{U}$ is defined over $\mathbb{Q}$. Taking again the orthogonal complement, one obtains the $\mathbb{Q}$-splitting asked for in 4.1.

Since $N(\rho_{\mathbb{U}}(\gamma)) = 0$, the residues of $\mathbb{U}$ are zero in all points $s \in S$, hence $\mathbb{U}$ extends to a local system on Y. q.e.d.

Corollary 4.2. *Suppose that $S \neq \emptyset$. Then the splittings in Corollary* 3.7 i) *and* ii), *can be defined over* $\mathbb{Q}$.

Lemma 4.3. *Let $\mathbb{M}$ be a local system, defined over $\mathbb{Z}$, and let $\mathbb{M}_{\mathbb{Q}} = \mathbb{W}_{\mathbb{Q}} \oplus \mathbb{U}_{\mathbb{Q}}$ be a decomposition, defined over $\mathbb{Q}$. Then there exist local systems $\mathbb{U}_{\mathbb{Z}}$ and $\mathbb{W}_{\mathbb{Z}}$, defined over $\mathbb{Z}$ with*

$$\mathbb{U}_{\mathbb{Q}} = \mathbb{U}_{\mathbb{Z}} \otimes \mathbb{Q}, \quad \mathbb{W}_{\mathbb{Q}} = \mathbb{W}_{\mathbb{Z}} \otimes \mathbb{Q}, \qquad \text{and} \quad \mathbb{M}_{\mathbb{Z}} \supset \mathbb{W}_{\mathbb{Z}} \oplus \mathbb{U}_{\mathbb{Z}}. \tag{4.3.1}$$

Moreover, if $\mathbb{U}_{\mathbb{Q}}$ is unitary with trivial local monodromies around S, then there exists an étale covering $\pi : Y' \to Y$ such that $\pi^*\mathbb{U}_{\mathbb{Q}}$ is trivial.

Proof. Defining a $\mathbb{Z}$-structure on $\mathbb{W}_{\mathbb{Q}}$ and $\mathbb{U}_{\mathbb{Q}}$ by

$$\mathbb{W}_{\mathbb{Z}} = \mathbb{W}_{\mathbb{Q}} \cap \mathbb{M}_{\mathbb{Z}} \quad \text{and} \quad \mathbb{U}_{\mathbb{Z}} = \mathbb{U}_{\mathbb{Q}} \cap \mathbb{M}_{\mathbb{Z}}$$

(4.3.1) obviously holds true.

Since the integer elements of the unitary group form a finite group, the representation defining $\mathbb{U}$ factors through a finite quotient of the fundamental group $\pi_1(U, *) \mapsto G$. The condition on the local monodromies implies that this quotient factors through $\pi_1(Y, *)$, and we may choose Y' to be the corresponding étale covering. q.e.d.

By 4.2 we obtain decompositions

$$R^1 f_* \mathbb{Q}_{X_0} = \mathbb{V}_{\mathbb{Q}} \oplus \mathbb{U}_{1\mathbb{Q}} \quad \text{and} \quad \mathbb{E}\text{nd}(\mathbb{V}_{\mathbb{Q}}) = \mathbb{W}_{\mathbb{Q}} \oplus \mathbb{U}_{\mathbb{Q}}.$$

By 4.1 the local monodromies of the unitary parts $\mathbb{U}_1$ and $\mathbb{U}$ are trivial. Moreover, $\mathbb{U}$ is a subvariation of Hodge structures of weight $0, 0$. Summing up, we obtain:

Corollary 4.4. *Let $f : X \to Y$ be a family of abelian varieties with unipotent local monodromies around $s \in S$, and reaching the Arakelov bound. If $S \neq \emptyset$ there exists a finite étale cover $\pi : Y' \to Y$ with:*

i. *For a $\mathbb{Z}$-variation of Hodge structures $\mathbb{V}'_{\mathbb{Z}}$ of weight* 1 *with maximal Higgs field, we have*

$$\pi^*(R^1 f_*(\mathbb{Z}_{X_0})) \supset \mathbb{V}'_{\mathbb{Z}} \oplus \mathbb{Z}^{2(g-g_0)}, \quad \text{and}$$
$$\pi^*(R^1 f_*(\mathbb{Z}_{X_0})) \otimes \mathbb{Q} = (\mathbb{V}'_{\mathbb{Z}} \oplus \mathbb{Z}^{2(g-g_0)}) \otimes \mathbb{Q}.$$

ii. $\mathbb{E}\text{nd}(\mathbb{V}'_{\mathbb{Z}}) \supset \mathbb{W}'_{\mathbb{Z}} \oplus \mathbb{Z}^{g_0^2}$, $\mathbb{E}\text{nd}(\mathbb{V}'_{\mathbb{Z}}) \otimes \mathbb{Q} = (\mathbb{W}'_{\mathbb{Z}} \oplus \mathbb{Z}^{g_0^2}) \otimes \mathbb{Q}$, *where $\mathbb{W}'_{\mathbb{Z}}$ is an $\mathbb{Z}$-variation of Hodge structures of weight* 0 *with maximal Higgs field, and where $\mathbb{Z}^{g_0^2}$ is a local $\mathbb{Z}$-subsystem of type* $(0, 0)$.

Proof of Theorem 0.2. Let Y' be the étale covering constructed in 4.4 ii). So using the notations introduced there,

$$(4.4.1)\quad R^1 f'_*(\mathbb{Z}_{X'_0}) \otimes \mathbb{Q} = \mathbb{V}'_{\mathbb{Q}} \oplus \mathbb{Z}^{2(g-g_0)} \quad \text{and} \quad \mathbb{E}\mathrm{nd}(\mathbb{V}'_{\mathbb{Q}}) = \mathbb{W}'_{\mathbb{Q}} \oplus \mathbb{Z}^{g_0^2}.$$

The left-hand side of (4.4.1) implies that $f' : X' \to Y'$ is isogenous to a product of a family of g_0-dimensional abelian varieties with a constant abelian variety B of dimension $g - g_0$. By abuse of notations we will assume from now on, that B is trivial, hence $g = g_0$ and $R^1 f'_*(\mathbb{Z}_{X'_0}) \otimes \mathbb{Q} = \mathbb{V}'_{\mathbb{Q}}$, and we will show that under this assumption $f' : X' \to Y'$ is isogenous to a g-fold product of a modular family of elliptic curves.

Let us write

$$\mathrm{End}(*) = H^0(Y', \mathbb{E}\mathrm{nd}(*))$$

for the global endomorphisms. $\mathrm{End}(\mathbb{V}'_{\mathbb{Q}}) = \mathbb{Q}^{g^2}$ is a $\mathbb{Q}$ Hodge structure of weight zero, in our case the Hodge filtration is trivial, i.e.,

$$\mathrm{End}(\mathbb{V}'_{\mathbb{Q}})^{0,0} = \mathrm{End}(\mathbb{V}'_{\mathbb{Q}}).$$

If $X_\eta = X' \times_{Y'} \mathrm{Spec}(\overline{\mathbb{C}(Y')})$ denotes the general fibre of f', one obtains from [**4**], 4.4.6,

$$\mathrm{End}(X_\eta) \otimes \mathbb{Q} = \mathrm{End}(\mathbb{V}'_{\mathbb{Q}})^{0,0} = \mathrm{End}(\mathbb{V}'_{\mathbb{Q}}).$$

By the complete reducibility of abelian varieties, there exists simple abelian varieties $B_1, \ldots, B_r$ of dimension g_i, respectively, which are pairwise nonisogenous, and such that X_η is isogenous to the product

$$B_1^{\times \nu_1} \times \cdots \times B_r^{\times \nu_r}.$$

Moreover, since $\mathbb{V}$ has no flat part, none of the B_i can be defined over $\mathbb{C}$. Let us assume that $g_i = 1$ for $i = 1, \ldots, r'$ and $g_i > 1$ for $i = r'+1, \ldots, r$.

By [**19**], p. 201, $D_i = \mathrm{End}(B_i) \otimes \mathbb{Q}$ is a division algebra of finite rank over $\mathbb{Q}$ with center K_i. Let us write

$$d_i^2 = \dim_{K_i}(D_i) \quad \text{and} \quad e_i = [K_i : \mathbb{Q}].$$

Hence $e_i \cdot d_i^2 = \dim_{\mathbb{Q}}(D_i)$.

By [**19**], p. 202, or by [**15**], p. 141, either $d_i \leq 2$ and $e_i \cdot d_i$ divides g_i, or else $e_i \cdot d_i^2$ divides $2 \cdot g_i$. In both cases, the rank $e_i \cdot d_i^2$ is smaller than or equal to $2 \cdot g_i$. If $i \leq r'$, hence if B_i is an elliptic curve, not defined over $\mathbb{C}$, we have $e_i = d_i = 1$.

Writing $M_{\nu_i}(D_i)$ for the $\nu_i \times \nu_i$ matrices over D_i, one finds ([**19**], p. 174)

$$\mathrm{End}(X_\eta) \otimes \mathbb{Q} = M_{\nu_1}(D_1) \oplus \cdots \oplus M_{\nu_r}(D_r)$$

hence

$$\begin{aligned} g^2 &\leq \dim_{\mathbb{Q}}(\mathrm{End}(X_\eta)\otimes\mathbb{Q}) \\ &= \left(\sum_{i=1}^{r}\nu_i\cdot g_i\right)^2 = \sum_{i=1}^{r}(e_i\cdot d_i^2)\cdot\nu_i^2 \\ &\leq \sum_{i=1}^{r'}\nu_i^2 + \sum_{i=r'+1}^{r}\nu_i^2\cdot 2\cdot g_i \leq \sum_{i=1}^{r}\nu_i^2\cdot g_i^2. \end{aligned}$$

Obviously this implies that $r = 1$ and that $g_1 \leq 2$. If $g_1 = 1$, we are done. In fact, the isogeny extends all over $Y' \setminus S'$ and, since we assumed the monodromies to be unipotent, B_1 is the general fibre of a semistable family of elliptic curves. The Higgs field for this family is again maximal, and 0.2 follows from 0.1.

It remains to exclude the case that $g_1 = 2$, and that $e_1 \cdot d_1^2 = 4$. If the center K_1 is not a totally real number field, e_1 must be lager than 1 and one finds:

I. $d_1 = 1$ and $D_1 = K_1$ is a quadratic imaginary extension of a real quadratic extension of $\mathbb{Q}$.

If K_1 is a real number field, looking again to the classification of endomorphisms of simple abelian varieties in [**19**] or [**15**], one finds that e_1 divides g_1, hence the only possible case is:

II. $d_1 = 2$ and $e_1 = 1$, and D_1 is a quaternion algebra over $\mathbb{Q}$.

The abelian surface B_1 over $\mathrm{Spec}(\mathbb{C}(Y'))$ extends to a nonisotrivial family of abelian varieties $B' \to Y'$, smooth outside of S and with unipotent monodromies for all $s \in S$. This family again has a maximal Higgs field, and thereby the local monodromies in $s \in S$ are nontrivial. As we will see below, in both cases, I and II, the moduli scheme of abelian surfaces with the corresponding type of endomorphisms turns out to be a compact subvariety of the moduli scheme of polarized abelian varieties, a contradiction.

I. By [**15**], Example 6.6 in Chapter 9, there are only finitely many g_1-dimensional abelian varieties with a given type of complex multiplication, i.e., with D_1 a quadratic imaginary extension of a real number field of degree g_1 over $\mathbb{Q}$.

II. By [**15**], Exercise (1) in Chapter 9, there is no abelian surface for which D_1 is a totally definite quaternion algebra. If $D_1 = \mathrm{End}(B)\otimes\mathbb{Q}$ is

totally indefinite, B is a false elliptic curve, as considered in Example 0.4, ii, for $d = 1$. Such abelian surfaces have been studied in [**25**], and their moduli scheme is a compact Shimura curve. The latter follows from Shimura's construction of the moduli scheme as a quotient of the upper half plane $\mathcal{H}$ (see [**15**], §8 in Chapter 9, for example) and from [**26**], Chapter 9. q.e.d.

5. Quaternion algebras and Fuchsian groups

Let A denote a quaternion algebra over a totally real algebraic number field F with d distinct embeddings

$$\sigma_1 = id,\ \sigma_2, \ldots, \sigma_d : F \to \mathbb{R},$$

which satisfies the following extra condition: for $1 \leq i \leq d$ there exists $\mathbb{R}$-isomorphism

$$\rho_1 : A^{\sigma_1} \otimes \mathbb{R} \simeq M(2, \mathbb{R}), \quad \rho_i : A^{\sigma_i} \otimes \mathbb{R} \simeq \mathbb{H}, \quad 2 \leq i \leq d,$$

where $\mathbb{H}$ is the quaternion algebra over $\mathbb{R}$. An order $\mathcal{O} \subset A$ over F is a subring of A containing 1 which is a finitely generated $\mathcal{O}_F$-module generating the algebra A over F. The group of units in $\mathcal{O}$ of reduced norm 1 is defined as

$$\mathcal{O}^1 = \{x \in \mathcal{O};\ \mathrm{Nrd(x)} = 1\}.$$

As Shimura has shown $\rho_1(\mathcal{O}^1) \subset Sl_2(\mathbb{R})$ is a discrete subgroup and for a torsion-free subgroup $\Gamma \subset \mathcal{O}^1$ of finite index $\mathcal{H}/\rho_1(\Gamma)$ is a quasi-projective curve, called Shimura curve. Furthermore, if A is a division algebra $\mathcal{H}/\rho_1(\Gamma)$ is projective (see [**26**], Chapter 9).

Remark 5.1. We will say that over some field extension F' of F the quaternion algebra splits, if $A_{F'} = A \otimes_F F' \simeq M(2, F')$. If $F' = F_v$ is the completion of F with respect to a place v of F, one says that F is ramified at v, if $A_v = A_{F_v}$ does not split. As well-known, there exists some $a \in F$ for which $A_{F(\sqrt{a})}$ splits. As explained in [**34**], for example, we can choose such $a \in F$ in the following way:

Fix one nonarchimedian prime p^0 of $\mathbb{Q}$, such that A is unramified over all places of F lying over p^0. Then choose a such that for all places v of F not lying over p^0 the quaternion algebra A ramifies at v if and only if $F_v(\sqrt{a}) \neq F_v$. Moreover one may assume, that the product over all conjugates of a is not a square in $\mathbb{Q}$.

Definition 5.2. If $\widetilde{\Gamma} \in \mathrm{PSl}_2(\mathbb{R})$ is a subgroup of finite index of some $P\rho_1(\mathcal{O}^1)$, then we call $\widetilde{\Gamma}$ a Fuchsian group derived from a quaternion algebra A.

Theorem 5.3 (Takeuchi [**31**])**.** *Let $\widetilde{\Gamma} \subset \mathrm{PSl}_2(\mathbb{R})$ be a discrete subgroup such that $\mathcal{H}/\widetilde{\Gamma}$ is quasi-projective. Then $\widetilde{\Gamma}$ is derived from a quaternion algebra A over a totally real number field F with d distinct embeddings*

$$\sigma_1 = id, \sigma_2, \ldots, \sigma_d : F \longrightarrow \mathbb{R},$$

with

$$\rho_1 : A^{\sigma_1} \otimes \mathbb{R} \simeq M(2, \mathbb{R}), \quad \rho_i : A^{\sigma_i} \otimes \mathbb{R} \simeq \mathbb{H}, \quad 2 \leq i \leq d$$

if and only if $\widetilde{\Gamma}$ satisfies the following conditions:

(I) *Let k be the field generated by the set $\mathrm{tr}(\widetilde{\Gamma})$ over $\mathbb{Q}$. Then k is an algebraic number field of finite degree, and $\mathrm{tr}(\widetilde{\Gamma})$ is contained in the ring of integers of k, $\mathcal{O}_k$.*

(II) *Let σ be any embedding of k into $\mathbb{C}$ such that $\sigma \neq \mathrm{id}_k$. Then $\sigma(\mathrm{tr}(\widetilde{\Gamma}))$ is bounded in $\mathbb{C}$.*

In the proof of Theorem 5.3 one gets, in fact, $k = F$. If A is a division algebra, for example if $d > 1$, then $Y = \mathcal{H}/\widetilde{\Gamma}$ is projective, and it is determined by A, and by the choice of the order $\mathcal{O} \subset A$ up to finite étale coverings.

Assumption 5.4. Let $\mathbb{X}_{\mathbb{Q}}$ be an irreducible $\mathbb{Q}$-variation of Hodge structures of weight one and width one, and with a maximal Higgs field. Assume moreover, that $\mathbb{X}_{\mathbb{Q}}$ is polarized. There are isomorphisms

$$\psi : \mathbb{X} = \mathbb{X}_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{C} \xrightarrow{\simeq} \mathbb{V} \oplus \mathbb{U}_1 \quad \text{and} \quad \phi : \mathbb{V} \xrightarrow{\simeq} \mathbb{L} \otimes \mathbb{T}$$

where $\mathbb{U}_1$ and $\mathbb{T}$ are both unitary, and where $\mathbb{L}$ is a rank two variation of Hodge structures of weight one and width one, with a maximal Higgs field. Moreover $\mathbb{V}$, $\mathbb{U}_1$, $\mathbb{L}$, $\mathbb{T}$ and ψ are defined over some real number field K, and ϕ over some number field K'. We fix an embedding of K' into $\mathbb{C}$ and denote by $k \subset K \subset K' \subset \mathbb{C}$ the field spanned by $\mathrm{tr}(\mathbb{L})$ over $\mathbb{Q}$.

Proposition 5.5. *Keeping the notations and assumptions made in* 5.4*, replacing Y by an étale covering, one may assume that:*

i. *$\Gamma_{\mathbb{L}}$ is derived from a quaternion algebra A over a totally real number field F with d distinct embeddings*

$$\sigma_1 = id, \sigma_2, \ldots, \sigma_d : F \longrightarrow \mathbb{R}.$$

ii. *For* $1 \leq i \leq d$ *there exists* $\mathbb{R}$*-isomorphism*

$$\rho_1 : A^{\sigma_1} \otimes \mathbb{R} \simeq M(2,\mathbb{R}), \quad \textit{and} \quad \rho_i : A^{\sigma_i} \otimes \mathbb{R} \simeq \mathbb{H}, \quad \textit{for} \quad 2 \leq i \leq d.$$

iii. *The representation* $\gamma_{\mathbb{L}} : \pi_1(Y,*) \to \mathrm{Sl}(2,\mathbb{R})$ *defining the local system* $\mathbb{L}$ *factors as*

$$\pi_1(Y,*) \xrightarrow{\simeq} \Gamma \subset \rho_1(\mathcal{O}^1) \longrightarrow \mathrm{Sl}(2,\mathbb{R} \cap \overline{\mathbb{Q}}) \subset \mathrm{Sl}(2,\mathbb{R}),$$

and $Y \simeq \mathcal{H}/\Gamma$.

iv. *For* a *as in* 5.1 $F(\sqrt{a})$ *is a field of definition for* $\mathbb{L}$.

v. *If* τ_i, $1 \leq i \leq d$ *are extension of* σ_i *to* $F(\sqrt{a})$, *and if* $\mathbb{L}_i$ *denotes the local system defined by*

$$\pi_1(Y,*) \longrightarrow \mathrm{Sl}(2,F(\sqrt{a})) \xrightarrow{\tau_i} \mathrm{Sl}(2,\overline{\mathbb{Q}}),$$

then $\mathbb{L}_i$ *is a unitary local system, for* $i > 1$, *and* $\mathbb{L}_1 \simeq \mathbb{L}$.

vi. *Up to isomorphism,* $\mathbb{L}_i$ *does not depend on the extension* τ_i *chosen.*

Proof. i) and ii): By Corollary 3.7 iii), $\Gamma_{\mathbb{L}}$ satisfies Condition (I) in Theorem 5.3. So, we only have to verify Condition (II) for $\Gamma_{\mathbb{L}}$. Let σ be an embedding of k into $\mathbb{C}$ which is not the identity, and let $\widetilde{\sigma} : K' \to \mathbb{C}$ be any extension of σ.

By 3.5 $\psi^{-1}\mathbb{V}^{\widetilde{\sigma}}$ is a subvariation of Hodge structures of $\mathbb{X}$, hence of width zero or one. On the other hand, $\mathbb{V}^{\widetilde{\sigma}}$ is isomorphic to $\mathbb{L}^{\widetilde{\sigma}} \otimes \mathbb{T}^{\widetilde{\sigma}}$. Both factors are variations of Hodge structures, hence at least one of them has a trivial Higgs field.

Assume both have a trivial Higgs field, hence $\mathbb{V}^{\widetilde{\sigma}}$ as well. By 1.9 a), the composite

$$\psi^{-1}\mathbb{V}^{\widetilde{\sigma}} \longrightarrow \mathbb{X} \xrightarrow{\psi} \mathbb{V} \oplus \mathbb{U}_1 \longrightarrow \mathbb{V}$$

has to be zero. Hence $\mathbb{V}^{\widetilde{\sigma}}$ is a sublocal system of the unitary system $\mathbb{U}_1$, hence unitary itself. The $\overline{\mathbb{Q}}$-isomorphism $\phi : \mathbb{V} \xrightarrow{\simeq} \mathbb{L} \otimes \mathbb{T}$ induces an isomorphism

$$\phi^{\otimes} : \bigotimes^{g_0} \mathbb{V}^{\widetilde{\sigma}} \longrightarrow \left(\bigotimes^{g_0} \mathbb{L}^{\widetilde{\sigma}} \right) \otimes \left(\bigotimes^{g_0} \mathbb{T}^{\widetilde{\sigma}} \right).$$

The right-hand side contains $S^{g_0}(\mathbb{L}^{\widetilde{\sigma}})$ as a direct factor, hence $S^{g_0}(\mathbb{L}^{\widetilde{\sigma}})$ is unitary, as well as $\mathbb{L}^{\widetilde{\sigma}}$. So $\mathrm{tr}(\mathbb{L}^{\widetilde{\sigma}}) = \sigma(\mathrm{tr}(\mathbb{L}))$ is bounded in this case.

If the Higgs field of $\mathbb{V}^{\widetilde{\sigma}}$ is nontrivial, it is generically maximal. This implies that the composite

$$\mathbb{V}^{\widetilde{\sigma}} \longrightarrow \mathbb{X} \longrightarrow \mathbb{U}_1$$

is zero. Hence $\mathbb{V}^{\widetilde{\sigma}} \simeq \mathbb{V}$. If the Higgs field of $\mathbb{L}^{\widetilde{\sigma}}$ is an isomorphism, by 1.4 replacing Y by an étale covering, $\mathbb{L}^{\widetilde{\sigma}} \simeq \mathbb{L}$. Hence up to conjugation the representations $\gamma_{\mathbb{L}^{\widetilde{\sigma}}}$ and $\gamma_{\mathbb{L}}$ coincide and for all $\eta \in \pi_1(Y, *)$

$$\operatorname{tr}(\gamma_{\mathbb{L}^{\widetilde{\sigma}}}(\eta)) = \operatorname{tr}(\gamma_{\mathbb{L}}(\eta)).$$

So σ is the identity, a contradiction.

It remains to consider the case that $\mathbb{V}^{\widetilde{\sigma}} \simeq \mathbb{V}$ and that $\mathbb{L}^{\widetilde{\sigma}}$ is concentrated in degree $0, 0$.

For g_0 even, one has a $\overline{\mathbb{Q}}$-isomorphism

$$\wedge^{g_0}\phi : \wedge^{g_0}\mathbb{V} \simeq S^{g_0}(\mathbb{L}) \oplus \mathbb{S}_{\{2,\dots,2\}}(\mathbb{T}) \oplus \bigoplus_{c=1}^{\frac{g_0}{2}-1} S^{2c}(\mathbb{L}) \otimes \mathbb{S}_{\lambda_{2c}}(\mathbb{T}),$$

where $\mathbb{S}_{\{2,\dots,2\}}(\mathbb{T})$ is of width zero, where $S^{g_0}\mathbb{L}$ has a maximal Higgs field of width g_0, and where all other factors have a maximal Higgs field of width between 2 and $g_0 - 2$. Let K denote the field of definition $\Gamma_{\mathbb{L}}$. Then $K \supset k$ is a finite extension of k. Let σ be an embedding of k into $\mathbb{C}$ which is not identity, and let $\widetilde{\sigma} : K \to \mathbb{C}$ be an extension of σ. Via the isomorphisms $\wedge^{g_0}\phi$ and $\wedge^{g_0}\phi^{\widetilde{\sigma}}$ we obtain an embedding

$$S^{g_0}\mathbb{L}^{\widetilde{\sigma}} \longrightarrow S^{g_0}(\mathbb{L}) \oplus \mathbb{S}_{\{2,\dots,2\}}(\mathbb{T}) \oplus \bigoplus_{c=1}^{\frac{g_0}{2}-1} S^{2c}(\mathbb{L}) \otimes \mathbb{S}_{\lambda_{2c}}(\mathbb{T}).$$

The projection of $S^{g_0}\mathbb{L}^{\widetilde{\sigma}}$ into $S^{g_0}\mathbb{L}$ must be zero, for otherwise, we would get an isomorphism $S^{g_0}\mathbb{L}^{\widetilde{\sigma}} \simeq S^{g_0}\mathbb{L}$. By Corollary 3.5 $\mathbb{L}^{\widetilde{\sigma}}$ is a subvariation of Hodge structures, hence it has a maximal Higgs field.

The projection

$$S^{g_0}\mathbb{L}^{\widetilde{\sigma}} \longrightarrow \bigoplus_{c=1}^{\frac{g_0}{2}-1} S^{2c}(\mathbb{L}) \otimes \mathbb{S}_{\lambda_{2c}}(\mathbb{T})$$

must be also zero, for otherwise, by applying again Corollary 3.5 to $S^{g_0}\mathbb{L}^{\widetilde{\sigma}}$ one would find $\mathbb{L}^{\widetilde{\sigma}}$ to have a maximal Higgs field, hence $S^{g_0}\mathbb{L}^{\widetilde{\sigma}}$ to have a maximal Higgs field of width g_0. But, then it can not be embedded in a local system of width $< g_0$.

Thus, the projection

$$S^{g_0}\mathbb{L}^{\widetilde{\sigma}} \longrightarrow S_{\{2,2,\dots,2\}}\mathbb{T}$$

is an embedding. This implies that $\mathbb{L}^{\widetilde{\sigma}}$ is unitary. In particular, again $\operatorname{tr}(\mathbb{L}^{\widetilde{\sigma}}) = \sigma(\operatorname{tr}(\mathbb{L}))$ is bounded in $\mathbb{C}$.

Finally, the assumption $\mathbb{V}^{\tilde{\sigma}} \simeq \mathbb{V}$ and $\mathbb{L}^{\tilde{\sigma}}$ unitary does not allow $g_0 = \operatorname{rank}(\mathbb{V})$ to be odd:

The $\overline{\mathbb{Q}}$-isomorphism $\phi : \mathbb{V} \simeq \mathbb{L} \otimes \mathbb{T}$, induces a $\overline{\mathbb{Q}}$-isomorphism

$$\wedge^{g_0}\phi : \wedge^{g_0}\mathbb{V} \simeq \bigoplus_{c=0}^{\frac{g_0-1}{2}} S^{2c+1}(\mathbb{L}) \otimes \mathbb{S}_{\lambda_{2c}}(\mathbb{T})$$

(see 2.2). The left-hand side contains a local subsystem isomorphic to $S^{g_0}(\mathbb{L}^{\tilde{\sigma}})$, hence with a trivial Higgs field, whereas the right-hand side only contains factors of width > 0, with a maximal Higgs field, a contradiction.

Applying 5.3 we obtain a quaternion algebra A satisfying i), ii) and the first part of iii). By 2.1 one has $Y \simeq \mathcal{H}/\Gamma$.

For iv) we recall that by the choice of a the quaternion algebra A splits over $F(\sqrt{a})$. So v) follows from i) and ii).

To see that $\mathbb{L}_i$ is independent of the extension of σ_i to $\tau_i : F(\sqrt{a}) \to \overline{\mathbb{Q}}$ it is sufficient to show vi) for $i = 1$. Let $\overline{\mathbb{L}}$ denote the local system obtained by composing the representation with the involution on $F(\sqrt{a})$. Then both, $\mathbb{L}$ and $\overline{\mathbb{L}}$ have a maximal Higgs field, hence by 1.4 c), their Higgs fields differ at most by the product with a two torsion element in $\operatorname{Pic}^0(Y)$. Replacing Y by an étale covering, we may assume both to be isomorphic. q.e.d.

Given a quaternion algebra A as in 5.5 i) and ii) allows us to construct certain families of abelian varieties. To this aim we need some well-known properties of quaternion algebras A defined over number fields F. Let us fix a subfield L of F.

Notations 5.6. Let us write $\delta = [L : \mathbb{Q}]$, $\delta' = [F : L]$ and

$$\beta_1 = \mathrm{id}_L, \beta_2, \ldots, \beta_\delta : L \longrightarrow \mathbb{C}$$

for the different embeddings. We renumber the embeddings $\sigma_i : F \to \mathbb{C}$ in such a way, that

$$\sigma_i|_L = \beta_\nu \quad \text{for} \quad (\nu - 1)\delta' < i \leq \nu\delta'.$$

Recall, that the corestriction $\operatorname{Cor}_{F/L}(A)$ is defined (see [**34**], p. 10) as the subalgebra of $\operatorname{Gal}(\overline{\mathbb{Q}}/L)$-invariant elements of

$$\bigotimes_{i=1}^{\delta'} A^{\sigma_i} = \bigotimes_{i=1}^{\delta'} A \otimes_{F,\sigma_i} \overline{\mathbb{Q}}.$$

Lemma 5.7. *Let A be a quaternion division algebra defined over a totally real number field F, of degree d over $\mathbb{Q}$. Assume that A is ramified at all infinite places of F except one. For some subfield L of F let $D_L = \mathrm{Cor}_{F/L}(A)$ be the corestriction of A to L. Finally let $a \in F$ be an element, as defined in* 5.1*, and*

$$b = a \cdot \sigma_2(a) \cdot \cdots \cdot \sigma_{\delta'}(a) \in L.$$

a. *If $L = \mathbb{Q}$, i.e., if $d = \delta'$, then either*
   i. *$D_{\mathbb{Q}} \simeq M(2^d, \mathbb{Q})$, and d is odd, or*
   ii. *$D_{\mathbb{Q}} \not\simeq M(2^d, \mathbb{Q})$. Then*

$$D_{\mathbb{Q}} \simeq M(2^d, \mathbb{Q}(\sqrt{b})).$$

   $\mathbb{Q}(\sqrt{b})$ is a quadratic extension of $\mathbb{Q}$, real if and only if d is odd.

b. *If $L \neq \mathbb{Q}$, then $D_L \not\simeq M(2^{\delta'}, L)$, and:*
   i. *$L(\sqrt{b})$ is an imaginary quadratic extension of L.*
   ii. *$D_L \otimes_L L(\sqrt{b}) \simeq M(2^{\delta'}, L(\sqrt{b}))$.*

In a) ii)*, or in* b)*, choosing an embedding $L(\sqrt{b}) \to M(2, L)$, one obtains an embedding*

$$D_L \longrightarrow M(2^{d+1}, L).$$

Proof. For $\delta = [L : \mathbb{Q}] \geq 1$, choose δ different embeddings $\beta_\nu : L \to \overline{\mathbb{Q}}$, corresponding to infinite places $v_1, \ldots, v_\delta$. We may assume that β_1 extends to the embedding σ_1 of F. Hence A is ramified over $\delta' - 1$ extensions of v_1 to F, and over all δ' extensions of v_ν to F, for $\nu \neq 2$. Writing L_v for the completion of L at v, one has

$$D_{v_\nu} = \mathrm{Cor}_{F/L} A \otimes_L L_{v_\nu} = \begin{cases} M(2, \mathbb{R}) \otimes \bigotimes^{\delta'-1} \mathbb{H} & \text{for } \nu = 1 \\ \bigotimes^{\delta'} \mathbb{H} & \text{for } \nu \neq 1. \end{cases}$$

Recall that the r-fold tensor product of $\mathbb{H}$ is isomorphic to $M(2^r, \mathbb{R})$ if and only if r is even. By our choice of a and b this holds true, if and only if $L_v(\sqrt{b}) = L_v$. In fact, the image of b in L_{v_ν} has the sign $(-1)^{\delta'-1}$, for $\nu = 1$ and $(-1)^{\delta'}$ otherwise.

In particular $D_L \simeq M(2^{\delta'}, L)$ can only hold true for $L = \mathbb{Q}$ and $d = \delta'$ odd.

For $L = \mathbb{Q}$, one also finds $b > 0$ if and only if d is odd.

For all but finitely many nonarchimedian places v of L, in particular for those dominating the prime p^0 in 5.1, and for the completion L_v

with respect to v, one has

$$D_v = \mathrm{Cor}_{F/L} A \otimes_L L_v = M(2^{\delta'}, L_v).$$

If this is not the case, consider the extension $L'_v = L(\sqrt{b}) \otimes_L L_v$ of L_v. One finds

$$D_v \otimes_{L_v} L'_v = M(2^{\delta'}, L'_v).$$

In fact, let $v_1, \ldots, v_\ell$ be the places of F, lying over v, and let $F_1, \ldots, F_\ell$ be the corresponding local fields. Then

$$D_v = \bigotimes_{i=1}^{\ell} \mathrm{Cor}_{F_i/L_v}(A \otimes_F F_i),$$

and it is sufficient to show that $D_i = \mathrm{Cor}_{F_i/L_v}(A \otimes_F F_i)$ splits over L'_v. If L'_v is a subfield of F_i

$$D_i \otimes L'_v = \mathrm{Cor}_{F_i/L_v}(A \otimes_F F_i)^{\otimes 2}$$

splits, since $(A \otimes_F F_i)^{\otimes 2}$ does. The same holds true, if L'_v is not a field.

If L'_v is a field, not contained in F_i, then $F'_i = F_i \otimes_{L_v} L'_v$ is a field extension of F_i of degree two, and

$$D_i \otimes L'_v = \mathrm{Cor}_{F'_i/L_v}(A \otimes_F F'_i)$$

splits again, since $(A \otimes_F F'_i)$ does. By [**35**], Chapter XI, §2, Theorem 2 (p. 206),

$$D_{L(\sqrt{b})} = D \otimes_L L(\sqrt{b}) = M(2^{\delta'}, L(\sqrt{b})). \qquad \text{q.e.d.}$$

Choose again an order $\mathcal{O}$ in A, and let $\mathcal{O}^1$ be the group of units in $\mathcal{O}$ of reduced norm 1. For any discrete torsion-free subgroup $\widetilde{\Gamma} \subset P\rho_1(\mathcal{O}^1)$ with preimage Γ in $\mathcal{O}^1 \subset \mathrm{Sl}_2(\mathbb{R})$ the diagonal embedding

$$\Gamma \longrightarrow \mathcal{O}^1 \longrightarrow \bigotimes_{i=1}^{\delta'} A^{\sigma_i}$$

induces an embedding

$$\Gamma \longrightarrow \mathcal{O}^1 \longrightarrow D_L = \mathrm{Cor}_{F/L} A. \tag{5.7.1}$$

Construction 5.8. For $L = \mathbb{Q}$ the morphism (5.7.1) and 5.7 a), give a morphism

$$\Gamma \subset D = \mathrm{Cor}_{F/\mathbb{Q}} A \subset D \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{b}) = M(2^d, \mathbb{Q}(\sqrt{b})) \subset M(2^{d+\epsilon}, \mathbb{Q})$$

for $\epsilon = 0$ or 1, where $b \in \mathbb{Q}$ is either a square, or as defined in 5.5. One obtains a representation

$$\eta : \Gamma \longrightarrow \mathrm{Gl}(2^d, \mathbb{Q}(\sqrt{b})) \longrightarrow \mathrm{Gl}(2^{d+\epsilon}, \mathbb{Q}).$$

If $\epsilon = 0$, the degree d must be odd. Over $\mathbb{R}$ one has

$$D \otimes_{\mathbb{Q}} \mathbb{R} \simeq M(2, \mathbb{R}) \otimes \mathbb{H} \otimes \cdots \otimes \mathbb{H}. \tag{5.8.1}$$

The $\mathbb{Q}$-algebraic group $G := \{x \in D^*;\ \mathrm{Nrd}(x) = x\overline{x} = 1\}$ is $\mathbb{Q}$-simple and by (5.8.1) it is a $\mathbb{Q}$-form of the $\mathbb{R}$-algebraic group

$$G(\mathbb{R}) \simeq \mathrm{Sl}(2, \mathbb{R}) \times \mathrm{SU}(2) \times \cdots \times \mathrm{SU}(2).$$

Projection to the first factor, gives a representation of Γ in $\mathrm{Sl}(2, \mathbb{R})$, hence a quotient $Y = \mathcal{H}/\Gamma$ with $\Gamma = \pi_1(Y, *)$.

Let us denote by $\mathbb{V}_{\mathbb{Q}}$ or by $\mathbb{X}_{\mathbb{Q}}$ the $\mathbb{Q}$-local system on Y induced by η. If we want to underline, that the local systems are determined by A we also write $\mathbb{V}_{A\mathbb{Q}}$ and $\mathbb{X}_{A\mathbb{Q}}$, respectively.

Lemma 5.9. *Keeping the assumptions and notations from* 5.8 *one finds:*

a.
$$\dim(\mathrm{End}(X_{A,\mathbb{Q}})) = \begin{cases} 1 & \text{for } \epsilon = 0 \\ 4 & \text{for } \epsilon = 1. \end{cases}$$

b. *For* $\epsilon = 1$ *one has*
$$\dim\left(H^0\left(Y, \bigwedge^2(X_{A,\mathbb{Q}})\right)\right) = \begin{cases} 3 & \text{for } d \text{ odd} \\ 1 & \text{for } d \text{ even.} \end{cases}$$

Proof. Consider for $\epsilon' = 2^\epsilon$

$$\mathbb{X} = \mathbb{X}_{A\mathbb{Q}} \otimes \mathbb{C} = \mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_d \otimes \mathbb{C}^{\epsilon'}$$

where for $\widetilde{\sigma} \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ the local system $\mathbb{L}_i^{\widetilde{\sigma}}$ has a maximal Higgs field if and only if $\widetilde{\sigma}|_F = \sigma_i^{-1}$. Otherwise this local system is unitary and of pure bidegree $0, 0$.

The determinant of each $\mathbb{L}_i$ is $\mathbb{C}$, hence $\mathrm{End}(\mathbb{X})$ contains $\mathbb{C}^{\epsilon'} \otimes \mathbb{C}^{\epsilon'}$ as a direct factor. Then

$$\dim_{\mathbb{Q}}(\mathrm{End}(\mathbb{X}_{A\mathbb{Q}})) = \dim_{\mathbb{C}}(\mathrm{End}(\mathbb{X})) \geq 4^\epsilon. \tag{5.9.1}$$

One has

$$\mathrm{End}(\mathbb{X}_{\mathbb{Q}}) = H^0(Y, \mathbb{E}\mathrm{nd}(\mathbb{X}_{\mathbb{Q}})) \simeq H^0\left(Y, \bigwedge^2(\mathbb{X}_{\mathbb{Q}})\right) \oplus H^0(Y, S^2(\mathbb{X}_{\mathbb{Q}})).$$

By 2.3

$$H^0\left(Y, \bigwedge^2(\mathbb{X})\right) = H^0(Y, S^2(\mathbb{L}_2 \otimes \cdots \otimes \mathbb{L}_d \otimes \mathbb{C}^{\epsilon'})) \qquad \text{and}$$

$$H^0(Y, S^2(\mathbb{X})) = H^0\left(Y, \bigwedge^2(\mathbb{L}_2 \otimes \cdots \otimes \mathbb{L}_d \otimes \mathbb{C}^{\epsilon'})\right).$$

Since $\mathrm{End}(\mathbb{X}_{\mathbb{Q}})$ is invariant under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, for $\widetilde{\sigma}$ with $\widetilde{\sigma}|_F = \sigma_2$ it is for $d > 1$ contained in the direct sum of

$$H^0(Y, S^2(\mathbb{L}_2^{\widetilde{\sigma}} \otimes \cdots \otimes \mathbb{L}_d^{\widetilde{\sigma}} \otimes \mathbb{C}^{\epsilon'})) = H^0\left(Y, \bigwedge^2(\mathbb{L}_3^{\widetilde{\sigma}} \otimes \cdots \otimes \mathbb{L}_d^{\widetilde{\sigma}} \otimes \mathbb{C}^{\epsilon'})\right)$$

and

$$H^0\left(Y, \bigwedge^2(\mathbb{L}_2^{\widetilde{\sigma}} \otimes \cdots \otimes \mathbb{L}_d^{\widetilde{\sigma}} \otimes \mathbb{C}^{\epsilon'})\right) = H^0(Y, S^2(\mathbb{L}_3^{\widetilde{\sigma}} \otimes \cdots \otimes \mathbb{L}_d^{\widetilde{\sigma}} \otimes \mathbb{C}^{\epsilon'})).$$

Repeating this game we find

$$H^0\left(Y, \bigwedge^2(\mathbb{X}_{\mathbb{Q}})\right) \subset \begin{cases} H^0(Y, S^2(\mathbb{C}^{\epsilon'})) & \text{for } d \text{ odd} \\ H^0(Y, \bigwedge^2(\mathbb{C}^{\epsilon'})) & \text{for } d \text{ even} \end{cases} \tag{5.9.2}$$

$$H^0(Y, S^2(\mathbb{X}_{\mathbb{Q}})) \subset \begin{cases} H^0(Y, \bigwedge^2(\mathbb{C}^{\epsilon'})) & \text{for } d \text{ odd} \\ H^0(Y, S^2(\mathbb{C}^{\epsilon'})) & \text{for } d \text{ even.} \end{cases} \tag{5.9.3}$$

For $\epsilon' = 1$ we obtain that $\mathrm{End}(\mathbb{X}_{\mathbb{Q}})$ is a most one-dimensional and for $\epsilon' = 2$ we find

$$\dim_{\mathbb{Q}}\left(H^0\left(Y, \bigwedge^2(\mathbb{X}_{\mathbb{Q}})\right)\right) \leq 3 \quad \text{and} \quad \dim_{\mathbb{Q}}(H^0(Y, S^2(\mathbb{X}_{\mathbb{Q}}))) \leq 1$$

or vice versa. Comparing this with (5.9.1) one obtains 5.9 a) and b).

q.e.d.

Lemma 5.10. *Given a quaternion division algebra A, as in* 5.5 i) *and* ii), *there exists a smooth family of abelian varieties $f : X_A \to Y$ with $R^1 f_* \mathbb{Q}_{X_A} = \mathbb{X}_{A\mathbb{Q}}$. Moreover, the special Mumford–Tate group* Hg *of the general fibre of f is the same as the group G in* 5.8.

Proof. (See [**18**].) The group G in 5.8 and the representation

$$G \longrightarrow D^* \to \mathrm{Gl}(2^{d+\epsilon}, \mathbb{Q})$$

are $\mathbb{Q}$ forms of an $\mathbb{R}$-representation

$$\mathrm{Sl}(2,\mathbb{R}) \times \mathrm{SU}(2)^{\times(d-1)} \longrightarrow \mathrm{Sl}(2,\mathbb{R}) \times \mathrm{SO}(2^{d-1}) \longrightarrow \mathrm{Gl}(2^{d+\epsilon},\mathbb{R}).$$

The group in the middle acts on $\mathbb{R}^2 \times \mathbb{R}^{2^{d-1}}$. Over $\mathbb{R}$, this representation leaves a unique nondegenerate symplectic form $\langle\ ,\ \rangle$ on $\mathbb{R}^{2^d}$ invariant, the tensor product of the $\mathrm{Sl}(2,\mathbb{R})$-invariant symplectic form on $\mathbb{R}^2$ with the $\mathrm{SO}(2^{d-1})$-invariant Hermitian form.

Hence for $\epsilon = 0$ and $V = \mathbb{Q}^{2^d}$ there is a unique symplectic form Q on V, invariant under $\Gamma \subset G$.

For $\epsilon = 1$, one chooses $V = \mathbb{Q}(\sqrt{b})^{2^d}$. Again one has a unique $\mathbb{Q}(\sqrt{b})$ valued symplectic form on V. Regarding V as a $\mathbb{Q}$ vector space, the trace $\mathbb{Q}(\sqrt{b}) \to \mathbb{Q}$ gives a $\mathbb{Q}$ valued symplectic form Q, again invariant under $\Gamma \subset G$.

Note that Γ is the group of units of an order $\mathcal{O}$ in A. Hence Γ leaves a $\mathbb{Z}$-module $L \subset V$ of rank $\dim V$ invariant. For some submodule $H \subset L$ of the form $H = mL$, for $m \gg 0$, one has $Q(H \times H) \subset \mathbb{Z}$. Obviously Γ leaves H again invariant. So one obtains a representation

$$\Gamma \longrightarrow \mathrm{Sp}(H,Q) \otimes \mathbb{Q}.$$

Finally let

$$\phi_0 : T = \{z \in \mathbb{C};\ |z| = 1\} \longrightarrow \mathrm{Sl}(2,\mathbb{R}) \times \mathrm{SO}(2^{d-1}) \subset \mathrm{Sp}(H,Q) \otimes \mathbb{R}$$

be the homomorphism defined by

$$e^{i\theta} \mapsto \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \times \mathrm{I}_{2^{d-1}}.$$

$J_0 = \phi_0(i)$ defines a complex structure on $H \otimes \mathbb{R}$, and

$$Q(x, J_0 x) > 0, \quad \text{for all} \quad x \in H.$$

The image of G in $\mathrm{Sp}(H,Q) \otimes \mathbb{R}$ is normalized by $\phi_0(T)$, i.e., for all $g \in G$ one has

$$g\phi_0(T)g^{-1} = \phi_0(T).$$

So $\mathbb{X}_{A\mathbb{Q}}$ defines a smooth family of abelian varieties $f : X_A \to Y = \mathcal{H}/\Gamma$.

By the construction this family reaches the Arakelov bound and $\mathbb{X}_{A\mathbb{Q}}$ has no unitary part. By Lemma 2.4 c), one knows that

$$G_0^{\mathrm{Mon}} = \mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_A}).$$

On the other hand, G_0^{Mon} is contained in the image of G in $\mathrm{Sp}(H,\mathbb{Q}) \otimes \mathbb{Q}$. Since

$$\mathbb{X}_{A\mathbb{C}} = \mathbb{L}_{1\mathbb{C}} \otimes \mathbb{L}_{2\mathbb{C}} \otimes \cdots \otimes \mathbb{L}_{d\mathbb{C}} \otimes \mathbb{C}^{2^\epsilon}$$

and since all factors are Zariski dense in $\mathrm{Sl}(2,\mathbb{C})$ one finds that

$$G_{0\mathbb{C}}^{\mathrm{Mon}} = \mathrm{Sl}(2,\mathbb{C})^{\times d} = G_{\mathbb{C}},$$

hence

$$G_0^{\mathrm{Mon}} = \mathrm{Hg}(R^1 f_* \mathbb{Q}_{X_A}) = G. \qquad \text{q.e.d.}$$

Let us remark, that in the proof of Theorem 0.5 in Section 6 we will see, that the families $f : X_A \to Y$ in 5.10 are unique up to isogenies, and up to replacing Y by étale coverings, and that they belong to one of the examples described in 0.4.

Construction 5.11. If $L \neq \mathbb{Q}$ choose b as in 5.7. The morphism (5.7.1) and 5.7 b), give a map

$$\Gamma \subset D_L = \mathrm{Cor}_{F/L} A \subset D_L \otimes_L L(\sqrt{b}) = M(2^{\delta'}, L(\sqrt{b})) \subset M(2^{\delta'+1}, L),$$

inducing a representation $\Gamma \to \mathrm{Gl}(2^{\delta'+1}, L)$, hence an L-local system $\mathbb{V}_L$ on $Y = \mathcal{H}/\Gamma$.

An embedding $L \subset M(\delta, \mathbb{Q})$ gives rise to

$$\Gamma \subset D_L = \mathrm{Cor}_{F/L} A \subset M(2^{\delta'+1}, L) \subset M(\delta 2^{\delta'+1}, \mathbb{Q}),$$

hence to a $\mathbb{Q}$-local system $\mathbb{X}_{\mathbb{Q}} = \mathbb{X}_{A,L;\mathbb{Q}}$.

In different terms, choose extensions $\widetilde{\beta}_\nu$ of β_ν to $\overline{\mathbb{Q}}$. For $\mathbb{V}_{\overline{\mathbb{Q}}} = \mathbb{V}_L \otimes_L \overline{\mathbb{Q}}$, the $\overline{\mathbb{Q}}$-local system

$$X_{\overline{\mathbb{Q}}} = \mathbb{X}_{A,L;\overline{\mathbb{Q}}} = \mathbb{V}_{\overline{\mathbb{Q}}} \oplus \mathbb{V}_{\overline{\mathbb{Q}}}^{\widetilde{\beta}_2} \oplus \cdots \oplus \mathbb{V}_{\overline{\mathbb{Q}}}^{\widetilde{\beta}_\delta}$$

is invariant under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, hence defined over $\mathbb{Q}$.

Remark 5.12. Consider any family $X \to Y$ of abelian varieties, with a geometrically simple generic fibre. If $\mathbb{X}_{A,L;\mathbb{Q}}$ is an irreducible component of $R^1 f_* \mathbb{Q}_X$, all irreducible components of $R^1 f_* \mathbb{Q}_X$ are isomorphic to $\mathbb{X}_{A,L;\mathbb{Q}}$. As in [**4**], p. 55, for $\Delta = \mathrm{End}(\mathbb{X}_{A,L;\mathbb{Q}})$ one finds

$$R^1 f_* \mathbb{Q}_X \simeq \mathbb{X}_{A,L;\mathbb{Q}} \otimes_\Delta \mathrm{Hom}(\mathbb{X}_{A,L;\mathbb{Q}}, R^1 f_* \mathbb{Q}_X),$$

and for some m

$$\mathrm{End}(R^1 f_* \mathbb{Q}_X) \simeq M(m, \Delta).$$

In [**21**], Section 9, one finds examples showing that all $m > 0$ occur.

6. The proof of Theorems 0.5 and 0.7

In order to prove Theorems 0.5 and 0.7 we will show, that the local subsystem $\mathbb{X}_{\mathbb{Q}}$ in 5.4 is for some $L \subset F$ isomorphic to the one constructed in 5.8 or 5.11.

Let us consider the subgroup H of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ of all β with $(\psi^{-1}\mathbb{V})^{\beta} = \psi^{-1}\mathbb{V}$, and let L denote the field of invariants under H. So $\mathbb{V} = \mathbb{V}_L \otimes_L \mathbb{C}$.

Proposition 6.1. *Let us keep the assumptions made in* 5.4 *and use the notations introduced in* 5.5. *Replacing Y by a finite étale covering, the field of invariants L under H is a subfield of F. Using the notations introduced in* 5.6 *for such a subfield, there exists a decomposition*

$$\mathbb{V}_L \simeq \mathbb{L}_{1L} \otimes \cdots \otimes \mathbb{L}_{\delta' L} \otimes \mathbb{T}'_L$$

with:

i. *For $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$ and $i \leq \delta'$ one has $\mathbb{L}^{\beta} \simeq \mathbb{L}_i$, if and only if $\beta|_F = \sigma_i$.*

ii. *For $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$ the Higgs field of*

$$(\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})^{\beta}$$

is maximal.

iii. *For $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with $\beta|_L \neq \mathrm{id}_L$ the local system*

$$(\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})^{\beta}$$

is unitary.

iv. *For $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$ the local system $\mathbb{T}'^{\beta}$ is a unitary.*

Proof. Replacing Y by an étale covering, we are allowed to apply 5.5. In particular we have the rank 2 local systems $\mathbb{L}_1, \ldots, \mathbb{L}_d$, defined there. Consider any decomposition $\mathbb{V} \simeq \mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_r \otimes \mathbb{T}_r$ with:

i′. If for $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ one has $\beta|_F = \sigma_i$, with $i \in \{1, \ldots, r\}$, then $\mathbb{L}^{\beta} \simeq \mathbb{L}_i$.

For $r = 1$, 1.4 gives a decomposition $\mathbb{V} = \mathbb{L} \otimes \mathbb{T}$. Write again $\mathbb{L}_1 = \mathbb{L}$ and $\mathbb{T}_1 = \mathbb{T}$. By 5.5 iv), the local system is defined over $F(\sqrt{a})$ and by 5.5 vi), $\mathbb{L}_1^{\beta} \simeq \mathbb{L}_1$ if the restriction of β to F is $\sigma_1 = \mathrm{id}_F$. Hence i′) holds true for this decomposition.

Consider for $r \geq 1$ a decomposition satisfying i′).

Step 1. If for some $\beta' \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and for $i \in \{1, \ldots, r\}$ one has $\mathbb{L}^{\beta'} \simeq \mathbb{L}_i$, then necessarily $\beta'|_F = \sigma_i$.

In fact, let $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}})$ be an automorphism with $\beta_F = \sigma_i$. Then $\mathbb{L}^{\beta^{-1}\circ\beta'} \simeq \mathbb{L}$, and 5.5 v), implies that $\beta^{-1} \circ \beta'|_F = \mathrm{id}_F$.

Step 2. There exists no $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with $\mathbb{L}_1^\tau \otimes \cdots \otimes \mathbb{L}_r^\tau$ not unitary and with $\tau|_F \neq \sigma_i$ for $i = 1, \ldots, r$.

Assume the contrary. Renumbering the embeddings of $F \to \mathbb{R}$ one may assume that $\tau|_F = \sigma_{r+1}$. Recall that by 3.5 $\mathbb{L}_i^\tau$ is a variation of Hodge structures of rank 2. It either is of width zero, hence unitary, or of width one, hence with maximal Higgs field. By assumption there exists some $i < r+1$ for which $\mathbb{L}_i^\tau$ has a maximal Higgs field. Choose $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with $\beta|_F = \sigma_i$. Then $\mathbb{L}^{\beta\circ\tau} = \mathbb{L}_i^\tau$ has a maximal Higgs field. 5.5 v), implies $\beta \circ \tau|_F = \mathrm{id}_F$, a contradiction.

Step 3. Assume there exists some $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with $\tau|_F \neq \sigma_i$ for $i = 1, \ldots, r$, with $\mathbb{V}^\tau$ not unitary, but with $\mathbb{L}_1^\tau \otimes \cdots \otimes \mathbb{L}_r^\tau$ unitary. Then (renumbering the embeddings $F \to \mathbb{R}$, if necessary) one finds a decomposition with $r+1$ factors, again satisfying i′).

$\mathbb{L}_i^\tau$ is unitary for $i = 1, \ldots, r$. By 1.4 and by 1.5 over some étale covering of Y we find a splitting $\mathbb{T}_r^\tau \simeq \mathbb{L} \otimes \mathbb{T}''$, with

$$\mathbb{V}^\tau \simeq (\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_r)^\beta \otimes \mathbb{L} \otimes \mathbb{T}''.$$

Apply τ^{-1}. Then one has

$$\mathbb{V} \simeq \mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_r \otimes \mathbb{L}^{\tau^{-1}} \otimes \mathbb{T}_{r+1}.$$

Since $\mathbb{L}_1$ has maximal Higgs field, $\mathbb{L}_{r+1} := \mathbb{L}^{\tau^{-1}}$ must be unitary, as well as $\mathbb{T}_{r+1}$. Applying any extension τ_i of σ_i^{-1} for $i \leq r$, one finds $\mathbb{L}_{r+1}^{\tau_i}$ to be unitary, since otherwise there would be two factors with a maximal Higgs field, $\mathbb{L}_i^{\tau_i}$ and $\mathbb{L}_{r+1}^{\tau_i}$.

So $\tau|_F$ must be one of the remaining σ_j, and renumbering we may assume $\tau|_F = \sigma_{r+1}$.

Step 4. Assume we have found a decomposition as in i′), and of maximal possible length. Then for all $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with $\tau|_F \neq \sigma_i$ for $i = 1, \ldots, r$ the local system

$$\mathbb{V}^\tau \simeq \mathbb{L}_1^\tau \otimes \cdots \otimes \mathbb{L}_r^\tau \otimes \mathbb{T}_r^\tau$$

is unitary. For those τ one has $(\psi^{-1}\mathbb{V})^\tau \neq \psi^{-1}\mathbb{V}$. On the other hand, for all β with $\beta|_F = \sigma_i$ with $1 \leq i \leq r$ the local system $\mathbb{V}^\beta$ has a maximal Higgs field, hence $(\psi^{-1}\mathbb{V})^\beta = \psi^{-1}\mathbb{V}$. So

$$H = \{\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q});\ \beta|_F = \sigma_i \text{ with } 1 \leq i \leq r\}$$

and L as the field of invariants under H is contained in F. Using the notations introduced in 5.6 for such subfields, one finds $r = \delta'$ and $\mathbb{L}_1^{\beta} \otimes \cdots \otimes \mathbb{L}_{\delta'}^{\beta}$ has a maximal Higgs field, for all $\beta \in H$. This in turn implies that $\mathbb{T}'^{\beta}$ is unitary for those β. q.e.d.

Theorem 6.2. *Let us keep the assumption made in* 5.4 *and use the notations introduced in* 5.6. *Replacing Y by an étale covering, there exists some $\epsilon' > 0$ and a decomposition*

$$\psi : \mathbb{X} \xrightarrow{\simeq} \bigoplus_{\nu=1}^{\delta} \left(\bigotimes_{i=(\nu-1)\delta'}^{\nu\delta'} \mathbb{L}_i \right)^{\oplus \epsilon'} \tag{6.2.1}$$

such that:

a. *For $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ the local system $\mathbb{L}_i^{\beta^{-1}}$ has a maximal Higgs field if and only if $\beta|_F = \sigma_i$. Moreover $\mathbb{L}^{\beta} = \mathbb{L}_i$ in this case.*
b. *The direct sum in* (6.2.1) *is orthogonal with respect to the polarization.*
c. *If the local subsystems $\psi^{-1}\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}$ of $\mathbb{X}$ are defined over L then $\epsilon' = 1$, $L = \mathbb{Q}$ and $[F : \mathbb{Q}]$ is odd.*
d. *If $\psi^{-1}\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'} \subset \mathbb{X}$ is not defined over L choose b to be the element defined in* 5.5 *and $\iota \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$ with $\iota(\sqrt{b}) = -\sqrt{b}$. Then $\epsilon' = 2$, the direct factor $\psi^{-1}\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'} \otimes \mathbb{C}^2$ in* (6.2.1) *is defined over L and it decomposes over $L(\sqrt{b})$ like*

$$\psi^{-1}\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'} \oplus (\psi^{-1}\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})^{\iota} \subset \mathbb{X}.$$

e. *$\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}$ is irreducible as a $\mathbb{C}$-local system.*

Proof. Using the notations from 6.1 let us define $\mathbb{L}_i = \mathbb{L}^{\widetilde{\sigma}_i}$, where $\widetilde{\sigma}_i$ is any extension of σ_i to $\overline{\mathbb{Q}}$. Obviously, fixing any extension $\widetilde{\beta}_\nu$ of β_ν one has

$$\mathbb{V}^{\widetilde{\beta}_\nu} = \mathbb{L}_{(\nu-1)\delta'+1} \otimes \cdots \otimes \mathbb{L}_{\nu\delta'} \otimes \mathbb{T}'^{\widetilde{\beta}_\nu}.$$

$\mathbb{V}$ has a maximal Higgs field, whereas $\bigoplus_{\nu=2}^{\delta} \mathbb{V}^{\widetilde{\beta}_\nu}$ is unitary. Hence their intersection is zero. Applying $\widetilde{\beta}_\nu$ one obtains the same for the intersection of $\mathbb{V}^{\widetilde{\beta}_\nu}$ and $\bigoplus_{\mu=1,\mu\neq\nu}^{\delta} \mathbb{V}^{\widetilde{\beta}_\mu}$. So

$$\psi^{-1}\left(\bigoplus_{\nu=1}^{\delta} \mathbb{V}^{\widetilde{\beta}_\nu} \right)$$

is a local subsystem of $\mathbb{X}$, defined over $\mathbb{Q}$. By assumption both must be equal. One obtains

$$\psi : \mathbb{X} \xrightarrow{\simeq} \bigoplus_{\nu=1}^{\delta} \left(\bigotimes_{(\nu-1)\delta'+1}^{\nu\delta'} \mathbb{L}_i \right) \otimes {\mathbb{T}'}^{\widetilde{\beta}_\nu}. \tag{6.2.2}$$

Let us show next, that $\mathbb{T}'$ is a trivial local system. The $\overline{\mathbb{Q}}$-isomorphism in (6.2.2) induces an isomorphism

$$\mathbb{E}\mathrm{nd}(\mathbb{X}) \xrightarrow{\simeq} \mathbb{E}\mathrm{nd}\left(\bigoplus_{\nu=1}^{\delta} \mathbb{V}^{\widetilde{\beta}_\nu} \right).$$

Since $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ permutes the direct factors $\mathbb{V}^{\widetilde{\beta}_\nu}$ of $\mathbb{X}$,

$$\bigoplus_{\nu=1}^{\delta} \mathbb{E}\mathrm{nd}(\mathbb{V}^{\widetilde{\beta}_\nu})$$

is a local subsystem, defined over $\mathbb{Q}$. So ϕ^{-1} induces an embedding

$$\phi' : \bigoplus_{\nu=1}^{\delta} \mathbb{E}\mathrm{nd}(\mathbb{L}_{(\nu-1)\delta'+1}) \otimes \cdots \otimes \mathbb{E}\mathrm{nd}(\mathbb{L}_{\nu\delta'}) \otimes \mathbb{E}\mathrm{nd}({\mathbb{T}'}^{\widetilde{\beta}_\nu}) \longrightarrow \mathbb{E}\mathrm{nd}(\mathbb{X}),$$

Writing $\mathbb{E}\mathrm{nd}(\mathbb{L}_i) = \mathbb{C} \oplus \mathbb{E}\mathrm{nd}_0(\mathbb{L}_i)$ we obtain a decomposition of the left-hand side in direct factors, all of the form

$$\mathbb{E}\mathrm{nd}_0(\mathbb{L}_{j_1}) \otimes \cdots \otimes \mathbb{E}\mathrm{nd}_0(\mathbb{L}_{j_\ell}) \otimes \mathbb{E}\mathrm{nd}({\mathbb{T}'}^{\widetilde{\beta}_\nu}),$$

for some $(\nu-1)\delta' + 1 \leq j_1 < \cdots < j_\ell \leq \nu\delta'$.

The only ones without any $\mathbb{E}\mathrm{nd}_0(\mathbb{L}_i)$ are the $\mathbb{E}\mathrm{nd}({\mathbb{T}'}^{\widetilde{\beta}_\nu})$. We claim that

$$\phi'\left(\bigoplus_{\nu=1}^{\delta} \mathbb{E}\mathrm{nd}\left({\mathbb{T}'}^{\widetilde{\beta}_\nu} \right) \right)^{\beta} = \phi'\left(\bigoplus_{\nu=1}^{\delta} \mathbb{E}\mathrm{nd}\left({\mathbb{T}'}^{\widetilde{\beta}_\nu} \right) \right),$$

for all $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Otherwise, we would get a nonzero projection from $\phi'(\bigoplus_{\nu=1}^{\delta} \mathbb{E}\mathrm{nd}({\mathbb{T}'}^{\widetilde{\beta}_\nu}))$ to an irreducible local system $\mathbb{E}$, containing at least one of the $\mathbb{E}\mathrm{nd}_0(\mathbb{L}_i)$. By construction, there exists an $\beta_i \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, such that $\mathbb{L}_i^{\beta_i}$ has a maximal Higgs field. Hence $\mathbb{E}^{\beta_i}$ has a maximal Higgs field.

Applying β_i we obtain a nonzero map

$$\phi'\left(\bigoplus_{\nu=1}^{\delta} \mathbb{E}\mathrm{nd}\left({\mathbb{T}'}^{\widetilde{\beta}_\nu} \right) \right)^{\beta\beta_i} \longrightarrow \mathbb{E}^{\beta_i}.$$

The right-hand side has a maximal Higgs field induced by the one on $\mathbb{E}\mathrm{nd}_0(\mathbb{L}_i^{\beta_i})$, whereas the left-hand side is unitary, a contradiction.

So $\phi'(\bigoplus_{\nu=1}^{\delta}\mathbb{E}\mathrm{nd}(\mathbb{T}'^{\widetilde{\beta}_\nu}))$ is $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-invariant, hence a unitary local system admitting a $\mathbb{Z}$-structure. This implies that $\phi'(\bigoplus_{\nu=1}^{\delta}\mathbb{E}\mathrm{nd}(\mathbb{T}'^{\widetilde{\beta}_\nu}))$ is trivial, after replacing Y by a finite étale cover. So the same holds true for $\mathbb{E}\mathrm{nd}(\mathbb{T}')$, hence for $\mathbb{T}'$ as well. Let us write $\mathbb{T}' = \mathbb{C}^{\oplus\epsilon'}$. Hence for some ϵ' one has the decomposition (6.2.1), and a) holds true by construction.

Recall that the local system $\mathbb{L}$ is defined over $F(\sqrt{a})$ for a as in 5.1. Hence $\mathbb{L}_i$ is defined over $\sigma_i(F)(\sqrt{\sigma_i(a)})$, and $\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}$ is defined over the compositum F' of those fields, for $i = 1, \ldots, \delta'$.

By 5.5 $\mathrm{Cor}_{F/L}A$ can only split if $L = \mathbb{Q}$ and if $[F : \mathbb{Q}]$ is odd. Let us write $L' = \mathbb{Q}$ in this case. Otherwise it splits over the subfield $L' = L(\sqrt{b})$ of F', where b is given in 5.7 b). In both cases one finds

$$(\mathrm{Cor}_{F/L}A) \otimes_L L' \simeq M(2^{\delta'}, L')$$

and correspondingly $\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}$ is defined over L'.

If $L' = \mathbb{Q}$, this is a local subsystem of $\mathbb{X}_{\mathbb{Q}}$. Since it is a $\mathbb{Q}$-variation of Hodge structures, and since we assumed $\mathbb{X}_{\mathbb{Q}}$ to be irreducible, both coincide.

If $L' \neq L$ consider the L'-local subsystem $(\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})_{L'}$ of $\mathbb{V}_{L'}$. For ι as in d),

$$\mathbb{V}' = (\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}) \oplus (\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})^{\iota},$$

is a local subsystem of $\mathbb{V}$, defined over L, and of rank $2^{\delta'+1}$. Then

$$\bigoplus_{nu=1}^{\delta} \psi^{-1}(\mathbb{V}')^{\widetilde{\beta}_\nu}$$

is a local subsystem of rank $\delta \cdot 2^{\delta'+1}$ of $\mathbb{X}$, defined over $\mathbb{Q}$. It is also a subvariation of Hodge structures. Since we assumed $\mathbb{X}_{\mathbb{Q}}$ to be irreducible, both must coincide and ϵ' is equal to two.

It remains to verify e). Assume that $\mathbb{M}$ is a direct factor of $\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}$. By 3.2 we may assume that $\mathbb{M}$ is defined over $\overline{\mathbb{Q}}$.

1.9 a), implies that $\mathbb{M}$ has a maximal Higgs field. By 1.4 $\mathbb{M} = \mathbb{L}' \otimes \mathbb{T}'_1$, and replacing Y by an étale covering we may assume that $\mathbb{L}' = \mathbb{L}_1 = \mathbb{L}$, and that $\mathbb{T}'$ is a direct factor of $\mathbb{L}_2 \otimes \cdots \otimes \mathbb{L}_{\delta'}$. Using the notations introduced in 5.6, let $\widetilde{\sigma}_i \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$ be an extension of σ_i, for $i =$

$1, \dots, \delta'$. For those i by 6.1

$$(\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})^{\widetilde{\sigma}_i}$$

has again a maximal Higgs field. Applying 1.9 a), one obtains the same for

$$\mathbb{M}^{\widetilde{\sigma}_i} = \mathbb{L}_1^{\widetilde{\sigma}_i} \otimes \mathbb{T}_1'^{\widetilde{\sigma}_i}.$$

For $i = 2$, the first factor is unitary, hence the second has again a maximal Higgs field. 1.5 tell us, that replacing Y again by some étale covering,

$$\mathbb{T}_1'^{\widetilde{\sigma}_i} = \mathbb{L} \otimes \mathbb{T}'',$$

hence for $\mathbb{T}_2' = \mathbb{T}''^{\widetilde{\sigma}_i^{-1}}$

$$\mathbb{M} = \mathbb{L}_1 \otimes \mathbb{L}_2 \otimes \mathbb{T}_2'.$$

Repeating this construction one finds

$$\mathbb{M} = \mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'} \otimes \mathbb{T}_{\delta'}',$$

necessarily with $\mathbb{T}_{\delta'}' = \mathbb{C}$. q.e.d.

Proposition 6.3. *Let $f : X \to Y$ be a family of abelian varieties with general fibre X_η, and reaching the Arakelov bound. Then:*

i. *For a generic fibre X_η of f*

$$\mathrm{End}(X_\eta) \otimes \mathbb{Q} \simeq \mathrm{End}_Y(X) \otimes \mathbb{Q} \simeq \mathrm{End}(R^1 f_* \mathbb{Q}_X)^{0,0}.$$

ii. *If $R^1 f_* \mathbb{C}_X$ has no unitary part then:*
   a. $\mathrm{End}(R^1 f_* \mathbb{Q}_X)^{0,0} = \mathrm{End}(R^1 f_* \mathbb{Q}_X)$.
   b. *If X_η is geometrically simple, $R^1 f_* \mathbb{Q}_X$ is irreducible.*
   c. *$f : X \to Y$ is rigid, i.e., the morphism from Y to the moduli scheme of polarized abelian varieties has no nontrivial deformation.*

Proof. i) is a special case of [**4**], 4.4.6.

If $R^1 f_* \mathbb{C}_X$ has no unitary part, for $\mathbb{V} = R^1 f_* \mathbb{C}_X$ 1.7 gives a decomposition $\mathbb{E}\mathrm{nd}(\mathbb{V}) = \mathbb{W} \oplus \mathbb{U}$ where $\mathbb{W}$ has a maximal Higgs field, and where $\mathbb{U}$ is concentrated in bidegree $0, 0$. Since 1.9 a), implies that $\mathbb{W}$ has no global section, one gets a).

For X_η geometrically simple $\mathrm{End}(X_\eta) \otimes \mathbb{Q} = \mathrm{End}(R^1 f_* \mathbb{Q}_X)^{0,0}$ is a skew field, hence a) implies that $R^1 f_* \mathbb{Q}_X$ is irreducible.

ii) c), follows from [**9**] (see also [**21**]). q.e.d.

Proposition 6.4. *Let $f : X \to Y$ be a family of abelian varieties, with a geometrically simple generic fibre X_η and reaching the Arakelov bound. Assume that (replacing Y by an étale covering, if needed) one has the decomposition* (6.2.1) *in* 6.2. *Then $R^1 f_* \mathbb{C}_X$ has no unitary part if and only if*

$$\mathrm{End}(R^1 f_* \mathbb{Q}_X)^{0,0} = \mathrm{End}(R^1 f_* \mathbb{Q}_X). \tag{6.4.1}$$

Proof. By 6.3 ii), a) and b), if $\mathbb{X}_\mathbb{Q} = R^1 f_* \mathbb{Q}_X$ has no unitary part, $\mathbb{X}_\mathbb{Q}$ is irreducible, and (6.4.1) holds true.

If on the other hand, $R^1 f_* \mathbb{C}_X$ has a unitary part, the same holds true for $\mathbb{X}$. Let us write again $\mathbb{U}_1$ for the unitary part of $\mathbb{X}$. So the field L in 5.4 can not be $\mathbb{Q}$. Recall that the Higgs field of $\mathbb{U}_1$ splits in two components, one of bidegree $1, 0$, the other of bidegree $0, 1$, both with a trivial Higgs field. Correspondingly $\mathbb{U}_1$ is the direct sum of two subsystems, say $\mathbb{U}^{1,0}$ and $\mathbb{U}^{0,1}$.

By 6.2 $\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_{\delta'}$ is an irreducible $\mathbb{C}$-local system. Let us choose one element of $\mathbb{C}^{\epsilon'}$ and the corresponding local subsystem

$$\mathbb{M} = \psi^{-1}(L_1 \otimes \cdots \otimes \mathbb{L}_{\delta'})$$

of $\mathbb{X}$. There exists some $\beta \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with $\mathbb{M}^\beta$ and $\overline{\mathbb{M}}^\beta$ unitary. Replacing $\mathbb{M}$ by $\overline{\mathbb{M}}$, if necessary we may assume that $\mathbb{M}^\beta$ lies in $\mathbb{U}^{1,0}$ and $\overline{\mathbb{M}}^\beta$ in $\mathbb{U}^{0,1}$. Then

$$\mathbb{M}^\beta \otimes \overline{\mathbb{M}}^{\beta^\vee} \subset \mathbb{U}^{1,0} \otimes \mathbb{U}^{0,1^\vee} \subset \mathbb{E}\mathrm{nd}(R^1 f_* \mathbb{C}_X)^{1,-1}.$$

In 5.5 v), we have seen that $\mathbb{L}_i \simeq \overline{\mathbb{L}}_i$ for all i. Hence

$$\overline{\mathbb{M}} \simeq \overline{\mathbb{L}}_1 \otimes \cdots \otimes \overline{\mathbb{L}}_{\delta'} \simeq \mathbb{M},$$

and $\mathbb{M}^\beta$ and $\overline{\mathbb{M}}^\beta$ are isomorphic. One obtains

$$\mathrm{End}(R^1 f_* \mathbb{C}_X)^{1,-1} \neq 0. \qquad \text{q.e.d.}$$

Proof of 0.5. Replacing Y by an étale covering, we may assume that $Rf_* \mathbb{C}_X$ has no unitary part as all. 1.4 provides us with a local system $\mathbb{L}$, independent of all choices, again after replacing Y by some étale covering.

Hence it is sufficient to consider the case that the generic fibre of $f : X \to Y$ is geometrically simple. By 6.3 i), the local system $\mathbb{X}_\mathbb{Q} = R^1 f_* \mathbb{Q}_X$ is irreducible. In 6.2 the nonexistence of a unitary part implies that $\delta = 1$, hence $L = \mathbb{Q}$, and

$$\mathbb{X} = \mathbb{V} = (\mathbb{L}_1 \otimes \cdots \otimes \mathbb{L}_d)^{\oplus \epsilon'}.$$

For $\epsilon' = 1$, the $\mathbb{Q}$-local system $\mathbb{X}_{\mathbb{Q}}$ is given by the representation

$$\eta : \pi_1(Y, *) \longrightarrow D^* = (\mathrm{Cor}_{F/\mathbb{Q}} A)^* = \mathrm{Gl}(2^d, \mathbb{Q}).$$

By 2.1 $\pi_1(Y, *) \to \Gamma = \eta(\pi_1(Y, *))$ is an isomorphism and $Y = \mathcal{H}/\Gamma$. Hence $\mathbb{X}_{\mathbb{Q}}$ is isomorphic to the local system $\mathbb{X}_{A\mathbb{Q}}$ constructed in 5.8. In particular, $d = [F : \mathbb{Q}]$ is odd, and by 6.3 i), 6.4, and 5.9

$$\mathrm{End}(X_\eta) = \mathrm{End}(\mathbb{X}_{\mathbb{Q}}) = \mathbb{Q} \quad \text{and} \quad H^0(Y, \mathbb{X}_{\mathbb{Q}} \otimes \mathbb{X}_{\mathbb{Q}}) = \mathbb{Q}.$$

The second equality implies that the polarization of $\mathbb{X}_{\mathbb{Q}}$ is unique, up to multiplication with constants, hence $\mathbb{X}_{\mathbb{Q}}$ and $\mathbb{X}_{A\mathbb{Q}}$ are isomorphic as polarized variations of Hodge structures. For some $\mathbb{Z}$-structure on $\mathbb{X}_{A\mathbb{Q}}$ we constructed in 5.10 a smooth family of abelian varieties $X_A \to Y$, and this family is isogenous to $f : X \to Y$. Both satisfy the properties, stated in Example 0.4 i).

For $\epsilon' = 2$ and for b as in 5.5, $\mathbb{X}_{\mathbb{Q}}$ is given by

$$\begin{aligned} \pi_1(Y, *) \longrightarrow D^* &= (\mathrm{Cor}_{F/\mathbb{Q}} A)^* \subset (D \otimes_{\mathbb{Q}} \mathbb{Q}(\sqrt{b}))^* \\ &= \mathrm{Gl}(2^d, \mathbb{Q}(\sqrt{b})) \subset \mathrm{Gl}(2^{d+1}, \mathbb{Q}), \end{aligned}$$

hence again $X_{\mathbb{Q}}$ is isomorphic to the local system $\mathbb{X}_{A\mathbb{Q}}$ constructed in 5.8.

By 6.3 i), 6.4 and 5.9 a), one finds that

$$\mathrm{End}(X_\eta) = \mathrm{End}(\mathbb{X}_{\mathbb{Q}})^{0,0} = \mathrm{End}(\mathbb{X}_{\mathbb{Q}}),$$

is of dimension 4.

For b as in 5.5, consider the local system

$$\mathbb{L}_{1\mathbb{Q}(\sqrt{b})} \otimes \cdots \otimes \mathbb{L}_{d\mathbb{Q}(\sqrt{b})}$$

defined by the representation $\pi_1(Y, *) \to \mathrm{Gl}(2^d, \mathbb{Q}(\sqrt{b}))$, together with a embedding into $\mathbb{X}_{\mathbb{Q}(\sqrt{b})}$. Restricting the polarization, one obtains a polarization Q' on $\mathbb{L}_{1\mathbb{Q}(\sqrt{b})} \otimes \cdots \otimes \mathbb{L}_{d\mathbb{Q}(\sqrt{b})}$, unique up to multiplication with constants. Regarding this local system as a $\mathbb{Q}$-local system, the inclusion

$$\mathrm{Gl}(2^d, \mathbb{Q}(\sqrt{b})) \subset \mathrm{Gl}(2^{d+1}, \mathbb{Q})$$

defines an isomorphism

$$\mathbb{L}_{1\mathbb{Q}(\sqrt{b})} \otimes \cdots \otimes \mathbb{L}_{d\mathbb{Q}(\sqrt{b})} \longrightarrow \mathbb{X}_{\mathbb{Q}}$$

and the restriction of the polarization of $\mathbb{X}_{\mathbb{Q}}$ is the composite of Q' with the trace on $\mathbb{Q}(\sqrt{b})$. In particular, the polarization is uniquely determined, and the family $f : X \to Y$ is isogenous to the family $X_A \to Y_A = Y$ constructed in 5.10.

Since, up to a shift in the bidegrees,

$$\mathbb{R}^2 f_* \mathbb{Q}_X = \bigwedge^2 \mathbb{X}_{\mathbb{Q}}$$

is a subvariation of Hodge structures of $\mathbb{E}\mathrm{nd}(\mathbb{X}_{\mathbb{Q}})$ one obtains the first equality in

$$\dim(H^0(Y, R^2 f_* \mathbb{Q})^{1,1}) = \dim(H^0(Y, R^2 f_* \mathbb{Q})) = \begin{cases} 3 & \text{for } d \text{ odd} \\ 1 & \text{for } d \text{ even,} \end{cases}$$

whereas the second one has been verified in 5.9 b). The dimension of $H^0(Y, R^2 f_* \mathbb{Q})^{1,1}$ is the Picard number of a general fibre of $f : X \to Y$. In fact, the Neron–Severi group of a general fibre is invariant under the special Mumford–Tate group of the fibre, hence by 2.4 a), it coincides with $\dim(H^0(Y, R^2 f_* \mathbb{Q})^{1,1})$.

Looking to the list of possible Picard numbers and to the structure of the corresponding endomorphism algebras for simple abelian varieties (for example in [**15**], p. 141), one finds that $\mathrm{End}(X_\eta) \otimes \mathbb{Q}$ is a quaternion algebra over $\mathbb{Q}$, totally indefinite for d odd, and totally definite otherwise. Hence $f : X \to Y$ satisfies the properties stated in Example 0.4 ii).
q.e.d.

Proof of 0.7. Again we may assume that $R^1 f_* \mathbb{C}_X$ has no nontrivial unitary subbundle defined over $\mathbb{Q}$. Let $\mathbb{V} \oplus \mathbb{U}_1$ be the decomposition of $R^1 f_* \mathbb{C}_X$ in a part with a maximal Higgs field and a unitary bundle. By 1.4 one can write $\mathbb{V} = \mathbb{L} \otimes \mathbb{T}$, where after replacing Y by a finite covering, $\mathbb{L}$ only depends on Y. If $h : Z \to Y$ is a subfamily of $f : X \to Y$ with a geometrically simple generic fibre, then repeating this construction with g instead of f, we obtain the same local system $\mathbb{L}$, hence by 5.3 the same quaternion algebra A. Hence we may assume that $f : X \to Y$ has a geometrically simple generic fibre, and we have to show, that $f : X \to Y$ is one of the families in Example 0.6.

By [**4**], §4, $R^1 f_* \mathbb{Q}_X$ is a direct sum of the same irreducible $\mathbb{Q}$-local system $\mathbb{X}_{\mathbb{Q}}$. From 1.4 and 5.3 we obtain $\mathbb{L}$ and a quaternion algebra A, defined over a totally real number field F. By 6.1, $\mathbb{X}$ contains a local system $\mathbb{V}$, defined over a subfield L of F. which satisfies the conditions stated there. By 6.2, for b as in 5.7, $\mathbb{V}$ is given by the representation $\pi_1(Y, *) \to \mathrm{Gl}(2^{\delta'+1}, L)$ induced by

$$\begin{aligned} \pi_1(Y, *) \longrightarrow D_L &= \mathrm{Cor}_{F/L} A \subset D_L \otimes_L L(\sqrt{b}) \\ &= M(2^{\delta'}, L(\sqrt{b})) \subset M(2^{\delta'+1}, L), \end{aligned}$$

hence it is isomorphic to the local system in 5.11. Then the decomposition of $\mathbb{X}$ in direct factors in 6.2 coincides with the one in 5.11, and $f : X \to Y$ is one of the families in Example 0.6.

In iii), the condition b) implies a) and vice versa. On the other hand, $L_i = \mathbb{Q}$ if and only if $R^1 h_{i*}\mathbb{C}_{Z_i}$ has no unitary part, which by 6.4 is equivalent to c). q.e.d.

7. Families of curves and Jacobians

Let us shortly discuss the relation between Theorems 0.2 and 0.5 and the number of singular fibres for semistable families of curves.

Let Y be a curve, let $h : \mathcal{C} \to Y$ be a semistable nonisotrivial family of curves of genus $g > 1$, smooth over V, and let $f : J(\mathcal{C}/Y) \to Y$ be a compactification of the Neron model of the Jacobian of $h^{-1}(V) \to V$. Let us write S for the points in $Y - V$ with $f^{-1}(y)$ singular, and Υ for the other points in $Y \setminus V$, i.e., for the points y with $h^{-1}(y)$ singular but $f^{-1}(y)$ smooth. Let $g(Y)$ be the genus of Y and $U = Y \setminus S$.

The Arakelov inequality for nonisotrivial families of curves says that

$$0 < 2 \cdot \deg(F^{1,0}) \leq g_0 \cdot (2 \cdot g(Y) - 2 + \#S + \#\Upsilon), \tag{7.0.2}$$

whereas the Arakelov inequality for $f : J(\mathcal{C}/Y) \to Y$ gives the stronger bound

$$0 < 2 \cdot \deg(F^{1,0}) \leq g_0 \cdot (2 \cdot g(Y) - 2 + \#S). \tag{7.0.3}$$

Hence for a family of curves, the right-hand side of (7.0.2) can only be an equality, if Υ is empty. On the other hand, if both, S and Υ are empty, the Miyaoka-Yau inequality for the smooth surface $\mathcal{C}$ implies that

$$\deg(h_*\omega_{\mathcal{C}/Y}) \leq \frac{g-1}{6}(2 \cdot g(Y) - 2).$$

Hence if $h : \mathcal{C} \to Y$ is smooth and if $h_*\mathbb{C}_{\mathcal{C}}$ has no unitary part, the inequalities (7.0.2) and (7.0.3) have both to be strict.

Let us consider the case $g(Y) = 0$, i.e., families of curves over $\mathbb{P}^1$. S.-L. Tan [**32**] has shown that $h : \mathcal{C} \to \mathbb{P}^1$ must have at least 5 singular fibres, hence that $\#S + \#\Upsilon \geq 5$, and (7.0.2) is strict in this special case.

Moreover, he and Beauville [**2**] gave examples of families with exactly 5 singular fibres for all $g > 1$. In those examples one has $\Upsilon = \emptyset$.

On the other hand, (7.0.3) implies that $\#S \geq 4$. For $\#S = 4$, the family $f : J(\mathcal{C}/Y) \to Y$ reaches the Arakelov bound, hence by 0.2 it is isogenous to a product of a constant abelian variety with a product of

modular elliptic curves, again with 4 singular fibres. By [**3**] there are just 6 types of such families, among them the universal family $E(3) \to X(3)$ of elliptic curves with a level 3-structures.

Being optimistic one could hope, that those families can not occur as families of Jacobians, hence that there is no family of curves $h : \mathcal{C} \to \mathbb{P}^1$ with $\#S = 4$. However, a counterexample has been constructed in [**13**]. In fact, the same family had been obtained before by G. Xiao in [**37**].

Example 7.1. Let B be a fixed elliptic curve, defined over $\mathbb{C}$. Consider the Hurwitz functor $\mathcal{H}_{B,N}$ defined in [**13**], i.e., the functor from the category of complex schemes to the category of sets with

$$\mathcal{H}_{B,N}(T) = \{f : C \longrightarrow B \times T;\ f \text{ is a normalized covering of degree } N \text{ and } C \text{ a smooth family of curves of genus 2 over } T\}.$$

The main result of [**13**] says that for $N \geq 3$ this functor is represented by an open subscheme $V = H_{B,N}$ of the modular curve $X(N)$ parameterizing elliptic curves with a level N-structure.

The universal curve $\mathcal{C} \to H_{B,N}$ extends to a semistable curve $\mathcal{C} \to X(N)$ whose Jacobian is isogenous to $B \times E(N)$. Hence writing S for the cusps, $J(\mathcal{C}/X(N))$ is smooth outside of S, whereas $\mathcal{C} \to X(N)$ has singular semistable fibres outside of $H_{B,N}$. Theorem 6.2 in [**13**] gives an explicit formula for the number of points in $\Upsilon = X(N) \setminus (H_{B,N} \cup S)$.

Evaluating this formula for $N = 3$ one finds $\#\Upsilon = 3$. For $N = 3$ the modular curve $X(3)$ is isomorphic to $\mathbb{P}^1$ with 4 cusps. So the number of singular fibres is 4 for $J(\mathcal{C}/\mathbb{P}^1) \to \mathbb{P}^1$ and 7 for $\mathcal{C} \to \mathbb{P}^1$.

We do not know whether similar examples exist for $g > 2$. For $g > 7$ the constant part B in Theorem 0.2 can not be of codimension one. In fact, the irregularity $q(\mathcal{C})$ of the total space of a family of curves of genus g over a curve of genus q satisfies by [**36**], p. 461, the inequality

$$q(\mathcal{C}) \leq \frac{5 \cdot g + 1}{6} + g(Y).$$

If $J(\mathcal{C}/Y) \to Y$ reaches the Arakelov bound, hence if it is isogenous to a product

$$B \times E \times_Y \cdots \times_Y E,$$

one finds

$$\dim(B) \leq \frac{5 \cdot g + 1}{6}.$$

As explained in [**8**] it is not known, whether for $g \gg 2$ there are any curves C over $\mathbb{C}$ whose Jacobian is isogenous to the product of elliptic

curves. Here we are even asking for families of curves whose Jacobian is isogenous to the product of the same nonisotrivial family of elliptic curve, up to a constant factor.

For the smooth families of abelian varieties, considered in 0.5 or 0.7 we do not know of any example, where such a family is a family of Jacobians.

References

[1] S. Abdulali, *Conjugates of strongly equivariant maps*, Pacific J. Math. **165** (1994) 207–216, MR 1300831 (95i:32041), Zbl 0843.14022.

[2] A. Beauville, *Le nombre minimum de fibres singulierès d'une courbe stable sur* P^1 (French), Asterisque **86** (1981) 97–108, Zbl 0502.14009.

[3] A. Beauville, *Les familles stables de courbes elliptiques sur* P^1 *admettant quatre fibres singulières*, C.R. Acad. Sci. Paris Sér. I Math. **294** (1982) 657–660, MR 0664643 (83h:14008), Zbl 0504.14016.

[4] P. Deligne, *Théorie de Hodge* II, I.H.É.S. Publ. Math. **40** (1971) 5–57, MR 0498551 (58 #16653a), Zbl 0219.14007.

[5] P. Deligne, *La conjecture de Weil pour les surfaces* $K3$, Invent. math. **15** (1972) 206–226, MR 0296076 (45 #5137), Zbl 0219.14022.

[6] P. Deligne, *Hodge cycles on abelian varieties*, (Notes by J. S. Milne), Springer Lecture Notes in Math. **900** (1982) 9–100, MR 0654325 (84m:14046), Zbl 0537.14006.

[7] P. Deligne, *Un théorème de finitude pour la monodromie*, Discrete Groups in Geometry and Analysis, Birkhäuser, Progress in Math. **67** (1987) 1–19, MR 0900821 (88h:14013), Zbl 0656.14010.

[8] T. Ekedahl & J.-P. Serre, *Exemples de courbes algébriques à jacobienne complétement décomposable*, C.R. Acad. Sci. Paris Sér. I Math. **317** (1993) 509–513, MR 1239039 (94j:14029), Zbl 0789.14026.

[9] G. Faltings, *Arakelov's theorem for abelian varieties*, Invent. Math. **73** (1983) 337–348, MR 0718934 (85m:14061), Zbl 0588.14025.

[10] W. Fulton & J. Harris, *Representation Theory. A first course*, Graduate Texts in Math., **129**, Springer-Verlag, New-York, 1991, MR 1153249 (93a:20069), Zbl 0744.22001.

[11] P. Griffths, *Topics in transcendental algebraic geometry*, Ann of Math. Stud., **106**, Princeton Univ. Press. Princeton, N.J., 1984, MR 0756842 (86b:14004), Zbl 0528.00004.

[12] J. Jost & K. Zuo, *Arakelov type inequalities for Hodge bundles over algebraic varieties, Part* 1: *Hodge bundles over algebraic curves*, J. Alg. Geom. **11** (2002) 535–546, MR 1894937 (2003b:14016).

[13] E. Kani, *Hurwitz spaces of genus* 2 *covers of an elliptic curve*, preprint.

[14] J. Kollár, *Subadditivity of the Kodaira dimension: fibres of general type*, Algebraic Geometry, Sendai, 1985, Advanced Studies in Pure Mathematics **10** (1987) 361–398, MR 0946244 (89i:14029), Zbl 0659.14024.

[15] H. Lange & Ch. Birkenhage, *Complex Abelian Varieties*, Grundlehren Math. Wiss., **302**, Springer-Verlag, Berlin-Heidelberg, 1992, MR 1217487 (94j:14001), Zbl 0779.14012.

[16] B. Moonen, *Linearity properties of Shimura varieties*, I, J. Algebraic Geometry **7** (1998) 539–567, MR 1618140 (99e:14028), Zbl 0956.14016.

[17] D. Mumford, *Families of abelian varieties*, Proc. Sympos. Pure Math. **9** (1966) 347–351, MR 0206003 (34 #5828), Zbl 0199.24601.

[18] D. Mumford, *A note of Shimura's paper "Discontinuous groups and Abelian varietes"*, Math. Ann. **181** (1969) 345–351, MR 0248146 (40 #1400), Zbl 0169.23301.

[19] D. Mumford, *Abelian Varieties*, Oxford Univ. Press, 1970, Oxford, MR 0282985 (44 #219), Zbl 0223.14022.

[20] C. Peters, *Arakelov-type inequalities for Hodge bundles*, Prépub. de l'Inst. Fourier, **511**, math.AG/0007102.

[21] M.-H. Saito, *Classification of nonrigid families of abelian varieties*, Tohoku Math. J. **45** (1993) 159–189, MR 1215923 (94d:14043), Zbl 0806.14004.

[22] I. Satake, *Symplectic representations of algebraic groups*, in 'Algebraic Groups and Discontinuous Subgroups' (Proc. Sympos. Pure Math., Boulder, Colo., 1965), 352–357, Amer. Math. Soc., Providence, R.I., MR 0209409 (35 #307), Zbl 0214.28903.

[23] W. Schmid, *Variation of Hodge structure: The singularities of the period mapping*, Invent. Math. **22** (1973) 211–319, MR 0382272 (52 #3157), Zbl 0278.14003.

[24] C. Schoen, *Varieties dominated by product varieties*, Int. J. Math. **7** (1996) 541–571, MR 1408839 (97f:14050), Zbl 0907.14002.

[25] I.R. Shafarevich, *On some families of abelian surfaces*, Izv. Math. **60** (1996) 1083–1093, MR 1427401 (97m:14044).

[26] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Publ. Math. Soc. of Japan, **11**, Iwanami Shoten and Princeton University Press, 1971, MR 0314766 (47 #3318), Zbl 0221.10029.

[27] C. Simpson, *Harmonic bundles on noncompact curves*, J. Amer. Math. Soc. **3** (1990) 713–770, MR 1040197 (91h:58029), Zbl 0713.58012.

[28] C. Simpson, *Constructing variations of Hodge structure using Yang–Mills theory and applications to uniformization*, Journal of the AMS **1** (1988) 867–918, MR 0944577 (90e:58026), Zbl 0669.58008.

[29] C. Simpson, *Moduli of representations of the fundamental group of a smooth projective variety* II, IHES **80** (1994) 5–79, MR 1320603 (96e:14013), Zbl 0891.14006.

[30] X.-T. Sun, S.L. Tan & K. Zuo, *Families of $K3$ surfaces over curves reaching the Arakelov-Yau type upper bounds and modularity*, Math. Res. Lett. **10** (2003) 323–342, MR 1981907 (2004h:14017).

[31] K. Takeuchi, *A characterization of arithmetic Fuchsian groups*, J. Math. Soc. Japan **27** (1975) 600–612, MR 0398991 (53 #2842), Zbl 0311.20030.

[32] S.-L. Tan, *The minimal number of singular fibers of a semistable curve over* P^1, J. Alg. Geom. **4** (1995) 591–596, MR 1325793 (96e:14038), Zbl 0864.14003.

[33] E. Viehweg & K. Zuo, *Families over curves with a strictly maximal Higgs field*, Asian J. Math. **7** (2003) 575–598.

[34] M.-F Vignéras, *Arithmetique des algebres de quaternions*, Springer Lect. Notes in Math., **800**, 1980, MR 0580949 (82i:12016), Zbl 0422.12008.

[35] A. Weil, *Basic number theory*, 3rd ed. Grundlehren Math. Wiss. **144**, Springer Verlag, Berlin-Heidelberg-New York, 1974, MR 0427267 (55 #302), Zbl 0326.12001.

[36] G. Xiao, *Fibered algebraic surfaces with low slope*, Math. Ann. **276** (1987) 449–466, MR 0875340 (88a:14046), Zbl 0596.14028.

[37] G. Xiao, *Surfaces fibrées en courbes de genre deux,* Lecture Notes in Mathematics, **1137**, Springer-Verlag, Berlin, 1985, MR 0872271 (88a:14042), Zbl 0579.14028.

Universität Duisburg-Essen
Mathematik, Campus Essen
45117 Essen, Germany
E-mail address: viehweg@uni-essen.de

Universität Mainz
Fachbereich 17, Mathematik
55099 Mainz, Germany
E-mail address: kzuo@informatik.uni-mainz.de