Special points and linearity properties of Shimura varieties

Ben Moonen

Special points and linearity properties of Shimura varieties

Speciale punten en lineariteitseigenschappen van Shimura variëteiten

(Met een samenvatting in het Nederlands)

Proefschrift ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de Rector Magnificus, Prof. Dr. J.A. van Ginkel, ingevolge het besluit van het College van Decanen in het openbaar te verdedigen op dinsdag 5 september 1995 des middags te 12.45 uur

door

Bernardus Julius Johannes Moonen

geboren op 24 augustus 1968 te Geleen.

Promotor: Prof. Dr. F. Oort

Copromotor: Dr. A.J. de Jong,
Harvard University (Cambridge, USA)

Contents

Introduction **v**

Some conventions and definitions **xiii**

I Hodge theory and Shimura varieties **1**

1 Mumford-Tate groups . . . 1
2 Variation of Hodge structure . . . 3
3 Shimura varieties . . . 7
4 Hermitian symmetric domains . . . 17

II A characterization of subvarieties of Hodge type **21**

1 Properties of subvarieties of Hodge type . . . 21
2 Decomposition of the adjoint group . . . 23
3 Totally geodesic subvarieties . . . 29
4 Non-rigid families of abelian varieties . . . 34

III Formal linearity, special points and Shimura varieties **41**

1 Local moduli of abelian varieties . . . 41
2 Isogenies . . . 50
3 Formal linearity . . . 52
4 A theorem of Noot . . . 61
5 Formal linearity and Shimura varieties of Hodge type . . . 64

IV Some applications **71**

1 Oort's conjecture . . . 71
2 The canonical lifting of a moduli point . . . 79

Bibliography **87**

Samenvatting 93

Curriculum vitae 97

Nawoord 99

Introduction

If for some class of algebraic objects there exists a moduli space, then we often find interesting subvarieties of this space by considering the locus of points for which the corresponding objects have some additional structure, or some extra symmetry. For example, when we study the moduli space of abelian varieties, we could consider abelian varieties on which a given ring acts by endomorphisms. In characteristic zero we can generalize this by considering abelian varieties with a given collection of Hodge classes. This leads to so-called Shimura subvarieties of the moduli space.

These Shimura subvarieties are very rich in structure. In this thesis, we will add to this some new results. We prove that Shimura subvarieties are characterized by certain linearity properties. Over the field of complex numbers, this is the property of being "totally geodesic", which we study in some detail. Our main results, however, concern an analogue of this in mixed characteristic, called "formal linearity". It was shown by Noot that Shimura subvarieties (in mixed characteristic) are formally linear. We show that they are in fact characterized by this property. Our proof reveals that formal linearity and total geodesicness are more directly related than one might expect at first glance. Applying our main results, we prove a conjecture of Oort under some additional hypotheses.

To make this more concrete, let us write $\mathsf{A}_{g,1} \otimes \mathbb{C}$ for the moduli space of principally polarized complex abelian varieties of dimension g. Question: how can we describe its Shimura subvarieties?

Experts on Shimura varieties will immediately start to formulate an answer. After all, $\mathsf{A}_{g,1} \otimes \mathbb{C}$ is "the" Shimura variety associated with the group $\mathrm{CSp}_{2g,\mathbb{Q}}$ and the Siegel double space $\mathfrak{H}_g^{\pm}$, and to answer the question we should therefore study closed immersions $i\colon (G, X) \hookrightarrow (\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ of Shimura data. The classification problem is discussed in Satake's book [55] and Section 1.8 of Deligne's paper [18].

But this was not really what we had in mind. It is nice to have a description in terms of Shimura data, but for some problems this is not very useful. To illustrate this, here is another question. Write $\mathcal{J}_g^{\mathrm{o}} \subseteq \mathsf{A}_{g,1} \otimes \mathbb{C}$ for the open Jacobi locus, i.e., the image of the Torelli morphism $j\colon \mathcal{M}_g \otimes \mathbb{C} \to \mathsf{A}_{g,1} \otimes \mathbb{C}$. Are there Shimura varieties $S \hookrightarrow \mathsf{A}_{g,1} \otimes \mathbb{C}$ (say for $g \geq 4$) with $\dim(S) > 0$ which are contained in $(\mathcal{J}_g^{\mathrm{o}})^{\mathrm{Zar}}$ or which have a positive-dimensional intersection with $\mathcal{J}_g^{\mathrm{o}}$? (This question is formulated in [47]. It is suggested by a conjecture of Coleman that for $g \geq 4$ the answer should be negative; however, for $g = 4$ and $g = 6$ it is shown in [12] that the answer is "yes".)

In the above problems it should be clear that we write "Shimura variety" where we really mean an irreducible component at some fixed finite level. To avoid confusion, let us introduce some terminology. Given a Shimura variety

$$Sh_K(G, X) = G(\mathbb{Q}) \setminus X \times G(\mathbb{A}_f) \;/\; K$$

in the sense of Deligne, we say that an irreducible subvariety $S \hookrightarrow Sh_K(G, X)$ is a subvariety of Hodge type if, up to a Hecke correspondence, it is an irreducible component of a Shimura subvariety. (See I, § 3 for a more precise definition.)

One of the main themes of this thesis is the question: "how can such subvarieties of Hodge type be characterized?" The following properties play a central role: if $S \hookrightarrow Sh_K(G, X)$ is of Hodge type, then (i) the special points on S are dense for the Zariski topology (even for the analytic topology), and (ii) S is a totally geodesic subvariety, i.e., it is an algebraic subvariety which is covered by a totally geodesic submanifold of X. If $Sh_K(G, X)$ is the moduli space of abelian varieties in mixed characteristics, we have an additional property (iii) S is formally linear at (most of) its ordinary points in characteristic p; this will be explained later on.

As for the first property, the following conjecture of Oort has been an important motivation for our research. Note that there is an obvious generalization of this conjecture, replacing $\mathsf{A}_{g,1,n} \otimes \mathbb{C}$ with an arbitrary Shimura variety.

Conjecture. (F. Oort) *Let $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ be an irreducible algebraic subvariety such that the CM-points on Z are dense for the Zariski topology. Then Z is a subvariety of Hodge type, in the sense of Definition I.3.8.*

Before we further discuss this conjecture, let us consider property (ii). In general, if $Z \hookrightarrow Sh_K(G, X)$ (over $\mathbb{C}$) is a totally geodesic subvariety, then it is not necessarily of Hodge type. For example, take Z to be a single point, then Z is totally geodesic, but it is of Hodge type only if it is a special point. However, it was suggested to us by D. Kazhdan that it should be possible to characterize the subvarieties of Hodge type in terms of total geodesicness.

Apparently, an additional condition is needed for such characterization. In fact, a special case of this was discussed in the late 1960s, when the question arose which of "Kuga's families of abelian varieties" are of Hodge type. The matter was settled by Mumford in his paper [42], where he showed that a Kuga fibre variety is of Hodge type if and only if it contains at least one fibre which is of CM-type. The same condition works in our more general setting: we show that a totally geodesic subvariety $Z \hookrightarrow Sh_K(G,X)$ is of Hodge type if and only if it contains at least one special point.

One would also like to have a description of general totally geodesic subvarieties. So, suppose $Z \hookrightarrow Sh_K(G,X)$ is an irreducible, totally geodesic subvariety. We compare it with the smallest subvariety of Hodge type containing it. This subvariety, call it S, is associated to some sub-Shimura datum $(M,Y_M) \hookrightarrow (G,X)$. If $u_S\colon Y_M^+ \twoheadrightarrow S$ is the natural covering map (the superscript $^+$ indicating a connected component) then, by assumption, Z is covered by a totally geodesic submanifold $Y' \subseteq Y_M^+$. Essentially, the problem is therefore to decide which totally geodesic submanifolds Y' give rise to algebraic subvarieties of S.

Phrased like this, the answer is not difficult to formulate: for Y' we can take all totally geodesic submanifolds $Y_1^+ \times \{y_2\} \subseteq Y_M^+$ arising from a decomposition $M^{\mathrm{ad}} = M_1 \times M_2$ of algebraic groups over $\mathbb{Q}$. More precisely, if such a decomposition exists, then $Y_M = Y_1 \times Y_2$, where $M_i(\mathbb{R})$ acts transitively on Y_i. Given a point $y_2 \in Y_2$ and a class $\eta K \in G(\mathbb{A}_f)/K$, we define $S_{\eta K}(Y_1 \times \{y_2\})$ as the image of $(Y_1 \times \{y_2\}) \times \eta K$ in $Sh_K(G,X)$. It is a totally geodesic algebraic subvariety and in Chapter II, §2 we show that, conversely, all totally geodesic subvarieties are obtained in this way.

Geometrically, the picture is very simple. If we pass to a suitable level in the Shimura variety $Sh(G,X)$, the subvariety of Hodge type S is a product variety: $S = S_1 \times S_2$ and the totally geodesic subvariety Z takes the form $S_1 \times \{a\}$ for some point $a \in S_2$. In particular, if $\dim(S_2) > 0$ (e.g., if Z is not of Hodge type), then Z can be deformed inside S, by varying the point $a \in S_2$. If $Sh(G,X)$ is a Shimura variety of Hodge type this leads to so-called non-rigid families of abelian varieties: we have a family of abelian varieties over S_1 which, fixing the base space S_1, admits non-trivial deformations.

This non-rigidity phenomenon was studied by several mathematicians. Faltings was the first to give an example showing that there exist non-rigid families of abelian varieties. The relative dimension in his example is 8; if we exclude isotrivial factors this is indeed the lowest dimension for which non-rigidity occurs. More recently, Masa-Hiko Saito classified the algebras which occur as the endomorphism algebra of the local system $\mathrm{R}^1 f_* \mathbb{Q}_{\mathcal{X}}$ associated to a non-rigid family $f\colon \mathcal{X} \to Z$.

We present a description of non-rigid families of abelian varieties in terms of subvarieties of the moduli space. Basically, we show that if $\mathcal{X} \to Z$ is such a non-rigid family (say, principally polarized and equipped with a level n structure), then there exists a sub-Shimura datum $(M, Y_M) \hookrightarrow (\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ and a decomposition $M^{\mathrm{ad}} = M_1 \times M_2$ as above, such that Z maps into a subvariety $S_{\eta K}(Y_1 \times \{y_2\}) \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ and such that all deformations of $\mathcal{X}$ over Z are obtained by varying the point y_2. For a more precise formulation, we refer to Chapter II, § 4.

In Chapter III the décor is slightly different. The "large" Shimura variety, the subvarieties of which we study, is the moduli space of abelian varieties. For most of the chapter, the base field $\mathbb{C}$ can rest in the dressing-room, until its reappearance at a crucial point in the fifth section. Enter the stage: base schemes of the form $\mathrm{Spec}(\widehat{\mathcal{O}}_{\mathfrak{p}})$, where $\widehat{\mathcal{O}}_{\mathfrak{p}}$ is the completion at a finite prime $\mathfrak{p}$ of the ring of integers of a number field F.

Our goal is to study a "linearity property" of subvarieties of Hodge type in mixed characteristic, which turns out to be a nice analogue of totally geodesicness. The relevance of this property was first shown by Rutger Noot in his PhD thesis [43]; his results have greatly stimulated our research. Let us explain the main ideas in their simplest form, referring to Chapter III for more precise statements.

Write $\mathcal{A}_g$ for the moduli space over $\mathbb{Z}_p$ of g-dimensional abelian varieties (together with a principal polarization and possibly a level structure, which we ignore for the moment). Consider an ordinary abelian variety X over a finite field k. Let $x \in \mathcal{A}_g \otimes \mathbb{F}_p$ be its moduli point and write $W = W(k)$ for the ring of infinite Witt vectors of k. Define $\mathfrak{A}_x$ as the formal completion of $\mathcal{A}_g$ at the point x. It is a formal scheme over $\mathrm{Spf}(W)$, called the formal deformation space of X. It was shown by Serre and Tate that $\mathfrak{A}_x$ has a natural structure of a formal torus over $\mathrm{Spf}(W)$.

Next consider an algebraic subvariety $S \hookrightarrow \mathcal{A}_g \otimes \mathbb{Q}$ of Hodge type. We obtain a model $\mathcal{S} \hookrightarrow \mathcal{A}_g$ by taking the Zariski closure inside $\mathcal{A}_g$. Assume that the moduli point x is a point of $\mathcal{S}$. By taking the formal completion at x we get a closed formal subscheme $\mathfrak{S}_x \hookrightarrow \mathfrak{A}_x$ over $\mathrm{Spf}(W)$. Noot's results essentially say that $\mathfrak{S}_x$ is a formal subtorus of $\mathfrak{A}_x$, in which case we say that $\mathcal{S}$ is formally linear at x. In general the situation may be slightly more complicated, but to fix ideas this description is sufficiently accurate.

A large part of Chapter III is devoted to a more detailed study of this notion of "formal linearity" and, keeping Oort's conjecture at the back of our minds, to explain its relation to certain collections of CM-points. This relation stems from the fact that the CM-liftings of X (say over the ring W) are precisely those corresponding to the torsion points of $\mathfrak{A}_x(W)$. These liftings, which are mutually all isogenous, are called quasi-canonical. The lifting X^{can} which corresponds to the identity element of $\mathfrak{A}_x$ is called the canonical lifting. Up to isomorphism it is uniquely determined by its property that all endomorphisms of X lift to X^{can}.

Let $Z \hookrightarrow \mathcal{A}_g \otimes \mathbb{Q}$ be an irreducible algebraic subvariety. Write $\mathcal{Z}$ for the Zariski closure inside $\mathcal{A}_g$, and suppose x (still assumed to be a closed ordinary moduli point) is a point of $\mathcal{Z} \otimes \mathbb{F}_p$. Let us moreover assume that $\mathcal{Z}$ is formally linear at x, i.e., the formal completion $\mathfrak{Z}_x \hookrightarrow \mathfrak{A}_x$ of $\mathcal{Z}$ at the point x is a formal subtorus. It then easily follows from the preceding remarks that the CM-points on Z are dense for the Zariski topology. According to Oort's conjecture Z should therefore be a subvariety of Hodge type. In Chapter III, § 5 we show that this is indeed the case (under slightly weaker assumptions):

Theorem. *Let $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ be an irreducible algebraic subvariety of the moduli space $\mathsf{A}_{g,1,n}$, defined over a number field F. Suppose there is a prime $\mathfrak{p}$ of $\mathcal{O}_F$ such that the model $\mathcal{Z}$ of Z (as in Section III.3.5) has formally quasi-linear components at some closed ordinary point $x \in (\mathcal{Z} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$. Then Z is of Hodge type, i.e., every irreducible component of $Z \otimes_F \mathbb{C}$ is a subvariety of Hodge type.*

Together with Noot's result (to which it is a converse) this theorem provides a characterization of subvarieties of Hodge type in terms of formal linearity. In the proof many ingredients of the first two chapters reappear and we see that formal linearity is not just an analogue of totally geodesicness, but it is indeed strongly related to it.

As explained above, one of the points to notice is that formal linearity is directly related to certain sets of CM-points. That is, if $\mathcal{Z}$ as above is formally linear at x, then the torsion points of $\mathfrak{Z}_x$ correspond to a Zariski dense collection of CM-points on Z. Using our theorem, Oort's conjecture would be proved if, conversely, we could deduce from the existence of a Zariski dense set of CM-points on Z that $\mathcal{Z}$ is formally linear at some ordinary point x. Unfortunately, we do not see how to prove this in full generality. The difficulty lies in the fact that, given a set of CM-points which is dense in Z, there is no obvious "common property", using which we can conclude something about the formal completions of Z at its ordinary points in characteristic p.

The "common property" that we want to assume is the following. Given an abelian variety X of CM-type over a number field F, the set of finite places v where X has ordinary reduction X_v is of density 1 (if F is sufficiently large). At all but finitely many such places v we have that X is the canonical lifting of X_v, in which case we say that X is canonical at v. To our assumption that there exists a Zariski dense collection of CM-points $\{X_t\}_{t\in T}$ on Z we add the hypotheses that there is a residue characteristic p such that each X_t is canonical at a prime above p. This extra hypotheses is quite reasonable: it is not very difficult to show that it is satisfied if $\mathcal{Z}$ is formally linear at some closed ordinary point.

Under these assumptions on Z we prove in Chapter III, § 3 that Z is formally linear at some of its ordinary points in characteristic p, hence, by our characterization, Z is a subvariety of Hodge type. Technically speaking this is one of the hardest parts of our work.

We conclude our work with some applications. In Chapter IV, § 1 we apply the main results of Chapter III to prove Oort's conjecture in a particular situation. Namely, we start with an abelian variety X over a number field, and we assume that the set $\mathcal{P}^{\text{o}}$ of places where X has ordinary reduction has density 1. It is a conjecture of Serre that this holds if we take the base field large enough. For each $v \in \mathcal{P}^{\text{o}}$ we consider the moduli point x_v^{can} of the canonical lifting of X_v and we define $Z \subseteq \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ as the Zariski closure of this collection of points $\{x_v^{\text{can}} \mid v \in \mathcal{P}^{\text{o}}\}$. Then we show that, up to a finite number of "exceptional" CM-points, Z is the smallest subvariety of Hodge type containing the moduli point of X. The trick is that, using the Galois representation on the ℓ-torsion of X (for a suitable prime number ℓ), we are able to find primes p such that sufficiently many of our abelian varieties of CM-type X_v are canonical at a prime above p. Once we have this, a straightforward application of the results of Chapter III proves our statement.

In the last section we study the Zariski closure of the moduli point of X^{can}, where X is an ordinary abelian variety in characteristic p (not necessarily defined over a finite field). First we show that this Zariski closure, call it Z, is a subvariety of

Hodge type. Knowing this, one wonders how $\dim(Z)$ compares to the dimension of the Zariski closure $\{x\}^{\mathrm{Zar}} \subseteq \mathsf{A}_{g,1,n} \otimes \mathbb{F}_p$ of the moduli point of X. Clearly, if x is a closed point, then both dimensions are zero. In general, $\dim(Z) \geq \dim(\{x\}^{\mathrm{Zar}})$. Our last result, joint work with A.J. de Jong and F. Oort, shows that there exist ordinary moduli points x with $\dim(\{x\}^{\mathrm{Zar}}) = 1$ and $\dim(Z) = g(g+1)/2$.

Some conventions and definitions

0.1 General. We write $\hat{\mathbb{Z}}$ for the profinite completion of $\mathbb{Z}$ and $\mathbb{A}_f = \hat{\mathbb{Z}} \otimes \mathbb{Q}$ for the ring of finite adèles of $\mathbb{Q}$. Given a commutative ring R, we write $W(R)$ for its ring of infinite Witt vectors.

A superscript ^{0} indicates a connected component for the Zariski topology; in case of an algebraic group it refers to the connected component of the identity element. A superscript $^+$ denotes a connected component for other topologies, usually of analytic nature.

We use a superscript $^{\mathrm{o}}$ to denote the open subschemes of $\mathsf{A}_{g,1,n} \otimes \mathbb{Z}_p$ and $\mathcal{I}sog(p^{eg})$ obtained by deleting the non-ordinary locus in characteristic p, see III.1 and III.2. The same symbol $^{\mathrm{o}}$ is used to indicate the dual of a category.

0.2 Schemes, varieties and manifolds. We do not use distinct notations for an algebraic variety over $\mathbb{C}$ and the associated analytic space; in most cases (hopefully all) it should be clear from the context what is meant. We use a superscript $^{\mathrm{Zar}}$ to denote a closure for the Zariski topology and a superscript $^{\mathrm{reg}}$ for the non-singular locus.

Manifolds are assumed to be paracompact. The word subvariety is used only for closed subvarieties.

0.3 Algebraic groups. Reductive groups are assumed to be connected. In line with the above, we do not carefully distinguish between an algebraic group over $\mathbb{C}$ (or $\mathbb{R}$) and the associated complex (resp. real) Lie group.

The superscripts $^{\mathrm{ab}}$ (maximal abelian quotient), $^{\mathrm{ad}}$ (adjoint group), $^{\mathrm{alg}}$ (algebraic envelope), $^{\mathrm{der}}$ (derived subgroup) and the symbols Lie (Lie algebra) and Res (restriction of scalars à la Weil) are used as is customary in the literature.

Let G be a reductive group over $\mathbb{Q}$ and consider a compact open subgroup $K \subset G(\mathbb{A}_f)$. For the definition of when K is called neat we refer to [52, 0.6]. If $K \subset G(\mathbb{A}_f)$ is neat then $G(\mathbb{Q}) \cap K$ is a neat arithmetic subgroup in the sense of [4, §17].

Suppose K is a neat subgroup of $G(\mathbb{A}_f)$. Every subgroup $K' \subseteq K$ is again neat. If $H \subseteq G$ is an algebraic subgroup of G then $H(\mathbb{A}_f) \cap K$ is a neat subgroup of $H(\mathbb{A}_f)$ and if $f\colon G \to G'$ is a homomorphism of algebraic groups over $\mathbb{Q}$ then $f(K)$ is a neat subgroup of $G'(\mathbb{A}_f)$. For every $n \geq 3$ the group $K_n = \{g \in \mathrm{GL}_m(\hat{\mathbb{Z}}) \mid g \equiv 1 \pmod n\}$ is a neat subgroup of $\mathrm{GL}_m(\mathbb{A}_f)$. It follows that for every compact open subgroup $K \subset G(\mathbb{A}_f)$ there is a neat subgroup $K' \subseteq K$ of $G(\mathbb{A}_f)$ which has finite index in K.

0.4 Formal schemes. The theory of formal schemes is set up in [27, I, §10 and III, §§3–5]. Unfortunately, not everything we need is treated there. Lacking a good reference, let us briefly discuss some definitions. Convention: *all formal schemes we use are noetherian and adic.*

Consider a formal scheme $(\mathfrak{X}, \mathcal{O}_{\mathfrak{X}})$, for which we usually simply write $\mathfrak{X}$. We write $\mathfrak{X}_{\mathrm{red}}$ for the associated reduced scheme ([27, I, §10.5]), which has the same underlying topological space as $\mathfrak{X}$. As a particular case we consider $\mathfrak{X} = \mathrm{Spf}(A)$, where A is a noetherian ring which is complete for the I-adic topology for some ideal $I \subseteq A$.

We call $\mathfrak{X}$ connected if $\mathfrak{X}_{\mathrm{red}}$ is a connected scheme, i.e., if the underlying topological space is connected. If $\mathfrak{X} = \mathrm{Spf}(A)$ then $\mathfrak{X}$ is connected if and only if I is a primary ideal.

We call $\mathfrak{X}$ formally reduced, if for all points $x \in \mathfrak{X}$ the local ring $\mathcal{O}_x$ is reduced. Define a sheaf of ideals $\mathcal{N}il \subseteq \mathcal{O}_{\mathfrak{X}}$ by $\mathcal{N}il(\mathfrak{U}) = \mathrm{nil}(\Gamma(\mathfrak{U}, \mathcal{O}_{\mathfrak{X}}))$. If this is a coherent $\mathcal{O}_{\mathfrak{X}}$-module (e.g., if $\mathfrak{X}$ is an excellent formal scheme) then it defines a closed formal subscheme $\mathfrak{X}_{\mathrm{fred}} \subseteq \mathfrak{X}$ which is formally reduced and has the same associated reduced scheme as $\mathfrak{X}$. In this case, we call $\mathfrak{X}_{\mathrm{fred}}$ the formal reduction of $\mathfrak{X}$; for every closed point $x \in \mathfrak{X}$, the local ring of $\mathfrak{X}_{\mathrm{fred}}$ at x is isomorphic to $(\mathcal{O}_x)_{\mathrm{red}} = \mathcal{O}_x/\mathrm{nil}(\mathcal{O}_x)$.

Let $\mathfrak{Y}_1$ and $\mathfrak{Y}_2$ be closed formal subschemes of $\mathfrak{X}$, defined by coherent ideal sheaves $\mathcal{I}_1$ and $\mathcal{I}_2$ respectively. We define the closed formal subscheme $\mathfrak{Y}_1 \cup \mathfrak{Y}_2 \subseteq \mathfrak{X}$ by the sheaf of ideals $\mathcal{I}_1 \cap \mathcal{I}_2$, which again is coherent.

Suppose $\mathfrak{X}$ is formally reduced. We say $\mathfrak{X}$ is formally irreducible if for all closed formal subschemes $\mathfrak{Y}_1$, $\mathfrak{Y}_2$ with $\mathfrak{X} = \mathfrak{Y}_1 \cup \mathfrak{Y}_2$ we have $\mathfrak{Y}_1 = \mathfrak{X}$ or $\mathfrak{Y}_2 = \mathfrak{X}$. An excellent formal scheme $\mathfrak{X}$ has a well-defined decomposition into (formal) irreducible components.

Let $f\colon \mathfrak{X} \to \mathfrak{Y}$ be a finite morphism of formal (noetherian and adic) schemes (as defined in [27, III, §4.8]). The $\mathcal{O}_{\mathfrak{Y}}$-algebra $f_*\mathcal{O}_{\mathfrak{X}}$ is coherent (loc. cit., Proposition

4.8.6), hence $\mathcal{K} = \mathcal{K}er(\mathcal{O}_{\mathfrak{Y}} \to f_*\mathcal{O}_{\mathfrak{X}})$ is also coherent (op. cit., 0_{I}, Corollaire 5.3.4). We define the image of f (in the sense of formal schemes) as the closed formal subscheme $f(\mathfrak{X}) \subseteq \mathfrak{Y}$ defined by the ideal $\mathcal{K}$. It is the smallest closed formal subscheme of $\mathfrak{Y}$ through which f factors.

0.5 Conventions related to Hodge theory. For more details we refer to I, §1.

We define the field $\mathbb{C}$ as $\mathbb{C} = \mathbb{R}[i]/(i^2+1)$, i.e., we choose an element $i \in \mathbb{C}$ with $i^2 = -1$. An $\mathbb{R}$-Hodge structure with underlying vector space $V_{\mathbb{R}}$ corresponds to a homomorphism of algebraic groups $h\colon \mathbb{S} \to \mathrm{GL}(V_{\mathbb{R}})$. This correspondence involves the choice of a sign. We choose to let $\mathbb{S}$ act on $V^{p,q} \subseteq V_{\mathbb{C}}$ via the character $z^{-p}\bar{z}^{-q}$. The reader can convince herself that these are really just conventions, which play no crucial role in the rest of this work.

Formally speaking, the Mumford-Tate group of a $\mathbb{Q}$-Hodge structure on a vector space V is defined as an algebraic subgroup of $\mathrm{GL}(V) \times \mathbb{G}_{\mathrm{m}}$. In practice we work with its image $\mathrm{MT} \subseteq \mathrm{GL}(V)$ under the first projection. The difference between the two groups is "at most" a central factor $\mathbb{G}_{\mathrm{m}}$, which in most of our statements can be safely ignored.

0.6 Shimura varieties. For more details, see I, §3.

For the definition of a Shimura variety we refer to Deligne's papers [14] and [18]. We follow most conventions and notations of [18], most of which are recalled in this section.

A pair (G, X) satisfying conditions ibid., (2.1.1.1-3) is called a Shimura datum. We write $Sh_K(G, X)$ for the canonical model of the associated Shimura variety over the reflex field $E(G, X)$. However, we often use the same notation for the weakly canonical model over fields $F \supseteq E(G, X)$ (including the case $F = \mathbb{C}$), whenever it is (or should be) clear from the context what is meant. A Shimura datum (G, X) (as well as the associated Shimura variety) is called "of Hodge type" if there exists a closed immersion $i\colon G \hookrightarrow \mathrm{CSp}_{2g}$ such that X is mapped into the Siegel double space $\mathfrak{H}_g^{\pm}$. This should not be confused with the notion of "subvariety of Hodge type", defined in I.3.8.

We have made no use of the terminology "Kuga variety". It should be clear though that for algebraic subvarieties $S \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ there is considerable overlap between our concepts and the theory of Kuga varieties (or group-theoretic abelian schemes).

Chapter I

Hodge theory and Shimura varieties

§1 Mumford-Tate groups

1.1 Consider a pure rational Hodge structure of weight n with underlying $\mathbb{Q}$-vector space V. Following Deligne (cf. [16, Section 2]) we can describe it by giving a homomorphism of algebraic groups

$$h\colon \mathbb{S} \to \mathrm{GL}(V)_{\mathbb{R}}\,,$$

where $\mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_{\mathrm{m}}$. We follow the sign conventions of [19, Section 3]. To be precise: let z and $\bar{z}$ be the two standard generators of the character group $X^*(\mathbb{S})$, then $\mathbb{S}$ acts on $V^{p,q}$ via the character $z^{-p}\bar{z}^{-q}$. With this convention, the Tate structure $\mathbb{Q}(1)$, as defined in [16, (2.1.13)], corresponds to the Norm homomorphism

$$\mathrm{Nm}\colon \mathbb{S} \to \mathbb{G}_{m,\mathbb{R}}\,,$$

which is just the character $z\bar{z}$.

1.2 The Mumford-Tate group of (V,h), denoted $\mathrm{MT}(h)$, is defined as the smallest algebraic subgroup of $\mathrm{GL}(V) \times \mathbb{G}_{\mathrm{m}}$ defined over $\mathbb{Q}$ such that $\mathrm{MT}(h)_{\mathbb{R}}$ contains the image of $h \times \mathrm{Nm}$. This can be formulated both in more fancy and in more down-to-earth terminology. For a fancy description, consider the category $\mathsf{Hdg}_{\mathbb{Q}}$ of rational Hodge structures, and write $\langle V, \mathbb{Q}(1)\rangle^{\otimes}$ for the Tannakian subcategory generated by (V,h) and $\mathbb{Q}(1)$. It is a neutral Tannakian category over $\mathbb{Q}$ for which the forgetful functor $\omega\colon \langle V, \mathbb{Q}(1)\rangle^{\otimes} \to \mathsf{Vec}_{\mathbb{Q}}$ is a fibre functor. Then $\mathrm{MT}(h) = \mathrm{Aut}^{\otimes}(\omega)$, the automorphism group of this fibre functor. We refer to [22] for an explanation of these terms and notations.

In more down-to-earth terms, we can consider the action of $\mathrm{GL}(V) \times \mathbb{G}_\mathrm{m}$ on the various spaces of the form

$$V(m_1, m_2, m_3) = V^{\otimes m_1} \otimes (V^*)^{\otimes m_2} \otimes \mathbb{Q}(m_3)\,,$$

where $\mathbb{G}_\mathrm{m}$ acts trivially on V and V^*, and acts via the character $z \mapsto z^{m_3}$ on $\mathbb{Q}(m_3) = \mathbb{Q}(1)^{\otimes m_3}$, and where $\mathrm{GL}(V)$ acts trivially on $\mathbb{Q}(m_3)$. Then $\mathrm{MT}(h)$ is the algebraic subgroup of $\mathrm{GL}(V) \times \mathbb{G}_\mathrm{m}$ characterized by the fact that, for all $m_1, m_2 \in \mathbb{Z}_{\geq 0}$, $m_3 \in \mathbb{Z}$, the Hodge classes in $V(m_1, m_2, m_3)$—defined as the rational classes purely of type $(0, 0)$—are precisely the invariants of $\mathrm{MT}(h)$.

Instead of the Mumford-Tate group as defined here, we could also consider the smallest algebraic subgroup $\mathrm{MT}'(h) \subseteq \mathrm{GL}(V)$ defined over $\mathbb{Q}$ such that $\mathrm{MT}'(h)_\mathbb{R}$ contains the image of h. It is the image of $\mathrm{MT}(h)$ under the first projection map $\mathrm{pr}_1\colon \mathrm{GL}(V) \times \mathbb{G}_\mathrm{m} \to \mathrm{GL}(V)$. The homomorphism $\mathrm{pr}_1\colon \mathrm{MT}(h) \to \mathrm{MT}'(h)$ is an isogeny if and only if the weight n is non-zero.

From now on, let us suppose the Hodge structure (V, h) is polarizable. We fix an element $i \in \mathbb{C}$ with $i^2 = -1$, and use this to identify the $\mathbb{Z}$-modules $\mathbb{Z}$ and $\mathbb{Z}(n)$ as in [16, (2.1.14)]. Via this identification a polarization $\psi\colon V^{\otimes 2} \to \mathbb{Q}(-n)$ gives a bilinear form on V, which we again call ψ. This form is symmetric if n is even, skew-symmetric if n is odd, and $\mathrm{MT}'(h)$ is a reductive subgroup of the group $\mathbb{G}\mathrm{U}(V, \psi)$ of elements $g \in \mathrm{GL}(V)$ which preserve ψ up to a scalar. For n even this is the group $\mathbb{G}\mathrm{O}(V, \psi)$ of orthogonal similitudes, for n odd it is the group $\mathrm{CSp}(V, \psi)$ of symplectic similitudes. We use the notation $\mathbb{G}\mathrm{U}(V, \psi)$ to treat both cases simultaneously.

There is a natural character c of $\mathbb{G}\mathrm{U}(V, \psi)$, called the multiplier, defined by the relation $\psi(gx, gy) = c(g) \cdot \psi(x, y)$. If we define the character ν as the restriction to $\mathrm{MT}(h)$ of the second projection $\mathrm{pr}_2\colon \mathrm{GL}(V) \times \mathbb{G}_\mathrm{m} \to \mathbb{G}_\mathrm{m}$, then $\nu^n = c^{-1} \circ \mathrm{pr}_1$ as characters of $\mathrm{MT}(h)$.

In what follows we will mostly be interested in Mumford-Tate groups of polarizable Hodge structures of weight n equal to 1 or -1. In this case the map $\mathrm{Id} \times c^{-n}\colon \mathrm{MT}'(h) \to \mathrm{CSp}(V, \psi) \times \mathbb{G}_\mathrm{m}$ gives an inverse of the projection $\mathrm{pr}_1\colon \mathrm{MT}(h) \to \mathrm{MT}'(h)$, which therefore is an isomorphism. This enables us to identify $\mathrm{MT}'(h)$ and $\mathrm{MT}(h)$, and to view $\mathrm{MT}(h)$ as a subgroup of $\mathrm{CSp}(V, \psi)$. If $n \neq \pm 1$ then $\mathrm{MT}(h) \to \mathrm{MT}'(h)$ is not an isomorphism. Its kernel is a diagonalizable group of dimension at most one, which is contained in the center of $\mathrm{MT}(h)$. In most of our statements, though, we can safely ignore this difference. Therefore, we do not carefully distinguish between $\mathrm{MT}(h)$ and $\mathrm{MT}'(h)$, and we often write MT where, strictly speaking, we mean MT'.

The group $\mathrm{Hg}(h) = \mathrm{Ker}(\nu)^0$ is called the Hodge group, or the special Mumford-Tate group. It is the smallest algebraic subgroup of $\mathrm{GL}(V)$ defined over $\mathbb{Q}$ such that $\mathrm{Hg}(h)_{\mathbb{R}}$ contains the image under h of the circle group $\mathrm{U}_1 = \mathrm{Ker}(\mathrm{Nm}) \subset \mathbb{S}$. The Mumford-Tate group $\mathrm{MT}(h)$ is isogenous to $\mathrm{Hg}(h) \times \mathbb{G}_{\mathrm{m}}$. Since we assumed (V, h) to be polarizable, the inner automorphism $\mathrm{Inn}(h(i))$ of $\mathrm{Hg}_{\mathbb{R}}$ is a Cartan involution (cf. [18, 1.1.15]). In particular, the center $Z(\mathrm{Hg}_{\mathbb{R}})$ (on which $\mathrm{Inn}(h(i))$ is trivial) is compact, i.e., $Z(\mathrm{Hg})(\mathbb{R})$ is compact.

§2 Variation of Hodge structure

2.1 A polarized variation of $\mathbb{Z}$-Hodge structure ($\mathbb{Z}$-VHS) of weight n over a complex manifold S is a triplet $\mathcal{V} = (\mathcal{V}_{\mathbb{Z}}, \mathcal{F}^{\bullet}, Q)$ consisting of a locally constant sheaf $\mathcal{V}_{\mathbb{Z}}$ of free $\mathbb{Z}$-modules of finite rank, a filtration $\mathcal{F}^{\bullet}$ of $\mathcal{V}_{\mathcal{O}} = \mathcal{O}_S \otimes_{\mathbb{Z}} \mathcal{V}_{\mathbb{Z}}$ by holomorphic subbundles, and a polarization form $Q\colon \mathcal{V}_{\mathbb{Z}} \times_S \mathcal{V}_{\mathbb{Z}} \to \mathbb{Z}_S$, such that

- (i) $\mathcal{F}^{\bullet}$ satisfies $\nabla \mathcal{F}^p \subseteq \Omega^1_S \otimes_{\mathcal{O}_S} \mathcal{F}^{p-1}$ with respect to the Gauß-Manin connection ∇ (Griffiths transversality), and
- (ii) $\mathcal{V}$ induces a polarized $\mathbb{Z}$-Hodge structure $(\mathcal{V}_{\mathbb{Z},s}, \mathcal{F}^{\bullet}_s, Q_s)$ of weight n on every fibre.

If we replace $\mathbb{Z}$ by $\mathbb{Q}$ in this definition then we get the notion of a polarized variation of Hodge structure (VHS) of weight n over S.

The polarizable variations of Hodge structure over S (i.e., direct sums of pure ones) are the objects of a category VHS_S, which is Tannakian if S is connected. In particular, given a polarizable VHS $\mathcal{V}$, we can form $\mathcal{V}(m_1, m_2, m_3)$ as in the case of Hodge structures, where $\mathbb{Q}(m)_S$ is the constant structure $\mathbb{Q}(m)$ over S, and $\mathcal{V}^* = \mathcal{H}om_S(\mathcal{V}, \mathbb{Q}(0)_S)$. In the sequel we often use the short-hand notation $\mathcal{V}(\mathbf{m})$, where $\mathbf{m}$ stands for a triplet $\mathbf{m} = (m_1, m_2, m_3)$ with $m_1, m_2 \in \mathbb{Z}_{\geq 0}$, $m_3 \in \mathbb{Z}$. Also, let us explicitly remark that a sub-VHS $\mathcal{W} \subseteq \mathcal{V}$ of a polarizable VHS $\mathcal{V}$ is again polarizable: the restriction to $\mathcal{W}$ of a polarization form Q on $\mathcal{V}$ is a polarization of $\mathcal{W}$.

2.2 Given a polarized VHS $\mathcal{V}$ of weight n over a connected complex manifold S, let MT_s denote the Mumford-Tate group of the Hodge structure on the fibre at $s \in S$.

We write $\mathcal{V}_s$ for $\mathcal{V}_{\mathbb{Q},s}$.

Let $\pi\colon \widetilde{S} \to S$ be a universal covering, and consider a VHS $\mathcal{W}$ of the form $\mathcal{W} = \mathcal{V}(\mathbf{m})$. For a global section w of $\pi^*\mathcal{W}_{\mathbb{Q}}$, let $H(w)$ be the locus of points $s \in \widetilde{S}$ where w_s is a Hodge class. Since w is holomorphic as a section of $\pi^*\mathcal{W}_{\mathcal{O}}$, the locus $H(w)$ is a countable union of irreducible complex analytic subspaces of $\widetilde{S}$. We define

$$\widetilde{\Sigma} = \bigcup H(w)\,,$$

where the union is taken over all $\mathcal{W} = \mathcal{V}(\mathbf{m})$ and all global sections w of $\pi^*\mathcal{W}_{\mathbb{Q}}$ such that $H(w)$ is not all of $\widetilde{S}$. Then $\widetilde{\Sigma}$ has the following properties:

(i) $\widetilde{\Sigma}$ is a countable union of proper, irreducible analytic subspaces of $\widetilde{S}$,

(ii) $\widetilde{\Sigma}$ is stable under the action of the group $\mathrm{Cov}(\widetilde{S}/S) \cong \pi_1(S,s)$ of covering transformations,

(iii) for $s \in \widetilde{S} \setminus \widetilde{\Sigma}$, $\mathcal{W} = \mathcal{V}(\mathbf{m})$, and w a global section of $\pi^*\mathcal{W}_{\mathbb{Q}}$, w_s is a Hodge class if and only if w is a Hodge class at every fibre.

Choose a basepoint $b \in S$ and a point $\tilde{b} \in \widetilde{S}$ with $\pi(\tilde{b}) = b$, then we get a trivialization $\pi^*\mathcal{V}_{\mathbb{Q}} \xrightarrow{\sim} \widetilde{S} \times \mathcal{V}_b$ such that the polarization form π^*Q corresponds fibrewise to the form Q_b on $\mathcal{V}_b$. Let $s \in S$, then the choice of a point $\tilde{s} \in \widetilde{S}$ with $\pi(\tilde{s}) = s$ determines an injective homomorphism $\mathrm{MT}_s \hookrightarrow \mathbb{GU}(\mathcal{V}_b, Q_b) \times \mathbb{G}_{\mathrm{m}}$. Let Σ be the image of $\widetilde{\Sigma}$ in S, then it follows from the above properties of $\widetilde{\Sigma}$ that Σ is a countable union of proper analytic subspaces of S, and that for $s \in S \setminus \Sigma$, the image of $\mathrm{MT}_s \hookrightarrow \mathbb{GU}(\mathcal{V}_b, Q_b) \times \mathbb{G}_{\mathrm{m}}$ does not depend on s, nor on the choice of $\tilde{s}$. Moreover, if we write M for this image, then the image of MT_s in $\mathbb{GU}(\mathcal{V}_b, Q_b) \times \mathbb{G}_{\mathrm{m}}$ is contained in M for every $s \in S$ (and every choice of $\tilde{s}$). We call M the generic Mumford-Tate group of the VHS $\mathcal{V}$. Sometimes it is more convenient not to relate this to a specific base point b, and we simply say that MT_s for $s \in S \setminus \Sigma$ is the generic Mumford-Tate group.

For $\mathcal{W} = \mathcal{V}(\mathbf{m})$ as before, we conclude from the above considerations that there is a well-defined sub-VHS $\mathcal{W}' \subseteq \mathcal{W}$ such that for $s \in S \setminus \Sigma$ the stalk $\mathcal{W}'_s \subseteq \mathcal{W}_s$ is the subspace of Hodge classes. As remarked before, $\mathcal{W}'$ is again a polarizable VHS. For a suitable direct sum $\oplus_i \mathcal{V}(\mathbf{m}_i)$, the generic Mumford-Tate group MT_s ($s \in S \setminus \Sigma$) is the subgroup of $\mathbb{GU}(\mathcal{V}_s, Q_s) \times \mathbb{G}_{\mathrm{m}}$ of elements acting trivially on $\oplus_i \mathcal{V}(\mathbf{m}_i)' \subseteq \oplus_i \mathcal{V}(\mathbf{m}_i)$.

In the language of Tannakian categories the generic Mumford-Tate group corresponds to the Tannakian subcategory $\langle \mathcal{V}, \mathbb{Q}(1)_S \rangle^{\otimes}$ of VHS_S generated by $\mathcal{V}$ and

$\mathbb{Q}(1)_S$. The functor ω_b which associates to a VHS $\mathcal{V}$ the $\mathbb{Q}$-vector space $\mathcal{V}_{\mathbb{Q},b}$ is a tensor functor on $\langle \mathcal{V}, \mathbb{Q}(1)_S \rangle^{\otimes}$, and $M \cong \mathrm{Aut}^{\otimes}(\omega_b)$.

If S is a nonsingular complex algebraic variety, then Σ is in fact a countable union of algebraic subvarieties of S, as was shown in [8].

2.3 We keep the above notations. From now on we assume that S is a connected, nonsingular complex algebraic variety. The local system $\mathcal{V}_{\mathbb{Q}}$ underlying $\mathcal{V}$ corresponds to a representation ρ: $\pi_1(S, s) \to \mathrm{GL}(\mathcal{V}_s)$, called the monodromy representation. The algebraic monodromy group is defined as the smallest algebraic subgroup of $\mathrm{GL}(\mathcal{V}_s)$ defined over $\mathbb{Q}$ which contains the image of ρ. We write H_s for its connected component of the identity, called the connected algebraic monodromy group.

From now on we suppose a $\mathbb{Z}$-structure $\mathcal{V}_{\mathbb{Z}}$ on $\mathcal{V}$ is given, making it a polarized $\mathbb{Z}$-VHS. As shown in [15, Proposition 7.5], the connected algebraic monodromy group H_s is contained in MT_s for $s \in S \setminus \Sigma$. In fact, with all the notations introduced so far the argument is easy to give: in the direct sum $\oplus_i \mathcal{V}(\mathbf{m}_i)$ as above, the subspace $\oplus_i \mathcal{V}(\mathbf{m}_i)'_s$ is purely of type $(0, 0)$, so that the polarization form on it is a positive-definite quadratic form, invariant under $\pi_1(S, s)$. Since $\pi_1(S, s)$ acts through the discrete group $\mathrm{GL}(\oplus_i \mathcal{V}_{\mathbb{Z}}(\mathbf{m}_i)'_s)$, we conclude that $H_s \times \{\mathrm{Id}\}$ acts trivially on $\oplus_i \mathcal{V}(\mathbf{m}_i)'_s$, hence $H_s \times \{\mathrm{Id}\} \subseteq \mathrm{MT}_s$.

Since we assumed S to be a nonsingular complex algebraic variety, we have even stronger results. The following theorem is taken from André's paper [2], where it is in fact stated in the more general context of variation of mixed Hodge structure.

2.4 Theorem. *(i) For all $s \in S \setminus \Sigma$ the group H_s is a normal subgroup of the derived group M^{der} of the generic Mumford-Tate group $M = \mathrm{MT}_s$.*

(ii) Suppose there is a point $s_0 \in S$ such that MT_{s_0} is abelian (hence a torus). Then $H_s = M^{\mathrm{der}}$ for every $s \in S \setminus \Sigma$.

Proof. In proving the theorem, we may pass to a finite covering of S, as this does not change the generic Mumford-Tate group and the connected algebraic monodromy group. Therefore, we may assume that the algebraic monodromy group of the local system $\mathcal{V}_{\mathbb{Q}}$ is connected.

Take $s \in S \setminus \Sigma$. By Chevalley's theorem (see [19, Proposition 3.1(b)]) we can find a finite-dimensional representation V of MT_s and a line $l \subseteq V$ such that H_s is the stabilizer of l. By ibid., Proposition 3.1(a) we can take V to be a direct sum of representations of the form $\mathcal{V}(\mathbf{m})_s$. Let

$$W \subseteq V = \bigoplus_{j=1}^{r} \mathcal{V}(\mathbf{m}_j)_s$$

be the MT_s-submodule generated by l, and consider the natural morphism

$$\varphi\colon \mathrm{MT}_s \to \mathrm{GL}(\mathrm{End}(W)) .$$

We claim that $H_s = \mathrm{Ker}(\varphi)$.

First of all, if $g \in \mathrm{Ker}(\varphi)$ then g commutes with every endomorphism of W, so g acts as a scalar on W. In particular, g stabilizes l, so $g \in H_s$. To prove the converse it suffices to show that H_s acts trivially on W.

Since $\pi_1(S, s)$ acts on the free $\mathbb{Z}$-modules $\mathcal{V}_{\mathbb{Z}}(\mathbf{m}_j)_s$, the action on the line l must factor through $\{-1, 1\}$, hence the action of H_s on l is trivial.The space of H_s-invariants in V is the fibre at s of the largest constant $\mathbb{Q}$-sub-local system of $\oplus_{j=1}^{r} \mathcal{V}_{\mathbb{Q}}(\mathbf{m}_j)$, which we call $\mathcal{Y}_{\mathbb{Q}}$. By the theorem of the fixed part (we apply Schmid's version [57, Theorem 7.22], and we use that S is algebraic) $\mathcal{Y}_{\mathbb{Q}}$ underlies a constant sub-VHS of $\oplus_{j=1}^{r} \mathcal{V}_{\mathbb{Q}}(\mathbf{m}_j)$. In particular, the fibre $\mathcal{Y}_s \subseteq \oplus_{j=1}^{r} \mathcal{V}_{\mathbb{Q}}(\mathbf{m}_j)_s = V$ is a MT_s-submodule, so $W \subseteq \mathcal{Y}_s$ (since $l \subseteq \mathcal{Y}_s$). This proves that $H_s = \mathrm{Ker}(\varphi)$.

To conclude the proof of (i) we have to show that $H_s \subseteq (\mathrm{MT}_s)^{\mathrm{der}}$. From $H_s \subseteq \mathrm{U}(\mathcal{V}_s, Q_s)$ and $H \lhd \mathrm{MT}_s$ we conclude that H_s is a normal subgroup of the Hodge group Hg_s. Since Hg_s is reductive this implies that H_s^{ab} is isogenous to a subtorus of the center $Z(\mathrm{Hg})$. In particular, $H_s^{\mathrm{ab}}(\mathbb{R})$ is compact. We can find a sequence of $\mathbf{m}_k$ such that the subspace X of H_s^{der}-invariants in $\oplus_k \mathcal{V}(\mathbf{m}_k)_s$ is a faithful representation of $\mathrm{Hg}_s^{\mathrm{ab}}$. The fundamental group $\pi_1(S, s)$ acts through $\mathrm{GL}(X \cap (\oplus_k \mathcal{V}_{\mathbb{Z}}(\mathbf{m}_k)_s))$, which is discrete. It also acts through the compact torus $H_s^{\mathrm{ab}}(\mathbb{R})$. From the connectedness of H_s^{ab} it then follows that H_s^{ab} is trivial. This proves that H_s is a semi-simple group and $H_s \lhd (\mathrm{MT}_s)^{\mathrm{der}}$.

To prove (ii) we have to show that every H_s-invariant element $v_s \in \oplus_j \mathcal{V}(\mathbf{m}_j)_s$ is also invariant under $(\mathrm{MT}_s)^{\mathrm{der}}$. Now, v_s being H_s-invariant means that it is the stalk at s of a global section v of $\mathcal{Y}_{\mathbb{Q}} \subseteq \oplus_j \mathcal{V}(\mathbf{m}_j)$. For every $p \in S$ there is a natural homomorphism of algebraic groups $r\colon \mathrm{MT}_p \to \mathrm{GL}(\mathcal{Y}_p) \times \mathbb{G}_{\mathrm{m}}$ with $\mathrm{Im}(r) = \mathrm{MT}(\mathcal{Y}_p)$. Since $\mathrm{MT}(\mathcal{Y}_p)$ does not depend on p ($\mathcal{Y}$ being a constant VHS) and $\mathrm{MT}(\mathcal{Y}_{s_0})$ is abelian, we conclude that $(\mathrm{MT}_s)^{\mathrm{der}}$ acts trivially on $\mathcal{Y}_s$, hence on v_s. □

§3 Shimura varieties

3.1 For the definition of a Shimura variety and a wealth of background information, we refer to Deligne's papers [14] and [18], in particular [18, Section 2]. We follow most notations and conventions of ibid., Section 0. In particular, we write a superscript ^{0} to indicate algebraic connected components, and a superscript $^+$ for topological connected components. For a reductive group G over $\mathbb{Q}$, we write $G(\mathbb{Q})_+$ for the intersection of $G(\mathbb{Q})$ and the inverse image of $G^{\mathrm{ad}}(\mathbb{R})^+$ under the adjoint map.

A pair (G, X) satisfying the axioms (2.1.1.1-3) of ibid. is called a Shimura datum. We write $Sh_K(G, X)$ for the canonical model over the reflex field $E(G, X) \subset \mathbb{C}$ (the definition of which can be found in [18, 2.2.1]) of the Shimura variety associated to a Shimura datum (G, X) and a compact open subgroup $K \subset G(\mathbb{A}_f)$. If F is a field extension of $E(G, X)$ then we write $Sh_K(G, X)_F$ for $Sh_K(G, X) \otimes_{E(G,X)} F$. Thus, we write $Sh_K(G, X)_{\mathbb{C}}$ for the scheme called ${}_K M_{\mathbb{C}}(G, X)$ in [18]. However, in an attempt to keep the notations simple we mostly omit the subscript "F", whenever we think it is clear from the context what is meant. In particular, we often write $Sh_K(G, X)$ for the Shimura variety over $\mathbb{C}$.

For a given choice of a connected component X^+ we denote the corresponding connected Shimura variety by $Sh^0_K(G, X)$.

A morphism $f\colon (G_1, X_1) \to (G_2, X_2)$ of Shimura data is defined as a homomorphism $f\colon G_1 \to G_2$ of algebraic groups over $\mathbb{Q}$ which induces a map from X_1 to X_2. We call f a closed immersion if it identifies G_1 with a closed subgroup of G_2. In this case $Sh(G_1, X_1)$ is called a Shimura subvariety of $Sh(G_2, X_2)$.

If $f\colon (G_1, X_1) \to (G_2, X_2)$ is a morphism of Shimura data, and $K_1 \subseteq G_1(\mathbb{A}_f)$, $K_2 \subseteq G_2(\mathbb{A}_f)$ are compact open subgroups with $f(K_1) \subseteq K_2$, then we write

$$f_{(K_1,K_2)}\colon Sh_{K_1}(G_1, X_1) \to Sh_{K_2}(G_2, X_2)$$

for the morphism induced by f. In the particular case that $(G_1, X_1) = (G_2, X_2)$ and f is the identity, we write $Sh_{(K_1,K_2)}$ instead of $f_{(K_1,K_2)}$.

3.2 An important example of a Shimura variety is the Siegel modular variety. To describe it, let us first fix some notations. We write CSp_{2g} for the Chevalley group scheme over $\mathbb{Z}$ of symplectic similitudes of the space $\mathbb{Z}^{2g}$ with its standard symplectic form ψ. The subgroup scheme of symplectic automorphisms is denoted by Sp_{2g}. There is a unique $\mathrm{CSp}_{2g}(\mathbb{R})$-conjugacy class of homomorphisms $h\colon \mathbb{S} \to \mathrm{CSp}_{2g,\mathbb{R}}$

defining a Hodge structure of type $(-1,0)+(0,-1)$ on $\mathbb{Q}^{2g}$ for which $\pm 2\pi i\cdot\psi$ is a polarization. As a Hermitian symmetric domain it can naturally be identified with the Siegel double space $\mathfrak{H}_g^{\pm}$; we choose this such that the connected component $\mathfrak{H}_g$ corresponds to the homomorphisms h for which $+2\pi i\cdot\psi$ is a polarization.

As explained in [14, Section 4], the Shimura variety $Sh(\mathrm{CSp}_{2g},\mathfrak{H}_g^{\pm})$ has an interpretation as a moduli space for abelian varieties with certain extra structures. Let us recall how this works.

A Jacobi level n structure on a polarized abelian variety (X,λ) over a base scheme S is an isomorphism of sheaves

$$\theta\colon X[n]\xrightarrow{\sim}(\mathbb{Z}/n)_S^{2g},$$

such that there exists a sheaf isomorphism $\nu\colon(\mathbb{Z}/n)_S\xrightarrow{\sim}\mu_{n,S}$ making the diagram

$$\begin{array}{ccc} X[n]\times X[n] & \xrightarrow{e_\lambda} & \mu_{n,S} \\ \downarrow{\scriptstyle\theta\times\theta} & & \uparrow{\scriptstyle\nu} \\ (\mathbb{Z}/n)_S^{2g}\times(\mathbb{Z}/n)_S^{2g} & \xrightarrow[\psi]{} & (\mathbb{Z}/n)_S \end{array}$$

commutative. Here e_λ is the Weil pairing and ψ denotes the standard alternating bilinear form on $(\mathbb{Z}/n)^{2g}$. We write $\mathsf{A}_{g,1,n}$ for the (coarse) moduli scheme over $\mathbb{Z}[1/n]$ of principally polarized abelian varieties with a Jacobi level n structure, often omitting the "$1,n$" in the notation. For $n\geq 3$ this is a fine moduli scheme.

For $n\in\mathbb{Z}_{\geq 1}$ we define $K_n=\{g\in\mathrm{CSp}_{2g}(\hat{\mathbb{Z}})\mid g\equiv 1\pmod{n}\}$. We get an isomorphism

$$f\colon Sh_{K_n}(\mathrm{CSp}_{2g},\mathfrak{H}_g^{\pm})\xrightarrow{\sim}\mathsf{A}_{g,1,n}(\mathbb{C})$$

in the following way. First we rewrite

$$\begin{aligned} Sh_{K_n}(\mathrm{CSp}_{2g},\mathfrak{H}_g^{\pm})(\mathbb{C}) &= \mathrm{CSp}_{2g}(\mathbb{Q})\setminus\mathfrak{H}_g^{\pm}\times\mathrm{CSp}_{2g}(\mathbb{A}_f)\;/\;K_n \\ &\cong \mathrm{Sp}_{2g}(\mathbb{Z})\setminus\mathfrak{H}_g\times\mathrm{CSp}_{2g}(\hat{\mathbb{Z}})\;/\;K_n. \end{aligned}$$

The map f is then obtained by using the dictionary between principally polarized $\mathbb{Z}$-Hodge structures of type $(-1,0)+(0,-1)$ and principally polarized complex abelian varieties. For details see [14, Section 4], where it is also shown that the projective system of the $\mathsf{A}_{g,1,n}\otimes\mathbb{Q}$ is a canonical model of $Sh(\mathrm{CSp}_{2g},\mathfrak{H}_g^{\pm})$. In the sequel we identify $Sh_{K_n}(\mathrm{CSp}_{2g},\mathfrak{H}_g^{\pm})$ and $\mathsf{A}_{g,1,n}\otimes\mathbb{Q}$.

The variety $\mathsf{A}_{g,1,n} \otimes \mathbb{C}$ has $\varphi(n)$ irreducible components, corresponding to the various possible choices of the isomorphism $\nu\colon (\mathbb{Z}/n) \xrightarrow{\sim} \mu_n$. Level structures with respect to a fixed choice of ν (or, in other words: a fixed choice for a primitive nth root of unity) are called symplectic level n structures in [25]. By considering abelian varieties with such symplectic level structures we obtain a moduli subscheme $\mathsf{A}_{g,1,(n)}$ of $\mathsf{A}_{g,1,n} \otimes \mathbb{Z}[\zeta_n, 1/n]$. The geometric fibres of $\mathsf{A}_{g,1,(n)} \to \mathrm{Spec}(\mathbb{Z}[\zeta_n, 1/n])$ are irreducible.

3.3 When working with abelian varieties, it will mostly suffice for our purposes to consider only principal polarizations and (Jacobi) level n structures. It should be remarked though that one can define the notion of a "level K structure" for arbitrary compact open subgroups $K \subset \mathrm{CSp}_{2g}(\mathbb{A}_f)$, generalizing the Jacobi level n structures (corresponding to $K = K_n$).

Also let us remark that our interpretation of $Sh(\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ depends on the choice of a "$\mathbb{Z}$-structure" on CSp_{2g}. With a different choice of this "$\mathbb{Z}$-structure" one would get an interpretation of $Sh_K(\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ as a moduli spaces of abelian varieties with a level K structure and some other type of polarization. Since we do not need much of this in the sequel we will not go into details.

3.4 For the study of Shimura varieties it is an important fact that $G(\mathbb{A}_f)$ acts on $Sh(G,X)$ (on the right) by automorphisms. This action is defined as follows. Take $g \in G(\mathbb{A}_f)$ and let $K \subset G(\mathbb{A}_f)$ be a compact open subgroup. Then we define a morphism $Sh_K(G,X) \to Sh_{g^{-1}Kg}(G,X)$ by sending $[x, \theta K]$ to $[x, \theta g(g^{-1}Kg)]$. By taking the limit over all K we obtain an automorphism of $Sh(G,X)$. This action of $G(\mathbb{A}_f)$ also exists on the canonical model of $Sh(G,X)$. (As mentioned before, the notation $Sh_K(G,X)$ is used both for the canonical model and for the Shimura variety over other fields, such as the field $\mathbb{C}$). The existence of an action of $G(\mathbb{A}_f)$ is in fact a crucial part of the definition of a model.

On a fixed finite level we no longer have the full action of $G(\mathbb{A}_f)$; in fact, on $Sh_K(G,X)$ we only have an action of the finite group $N_G(K)/K$ (where $N_G(K)$ is the normalizer of K in $G(\mathbb{A}_f)$). Nevertheless, the group $G(\mathbb{A}_f)$ does "act" on $Sh_K(G,X)$ in the sense of correspondences, as we will explain. Again let $g \in G(\mathbb{A}_f)$. Let $K_1, K_2 \subset G(\mathbb{A}_f)$ be compact open subgroups and, for the moment, write $K' = K_1 \cap gK_2g^{-1}$. The Hecke correspondence $\mathcal{T}_g$ from $Sh_{K_1}(G,X)$ to $Sh_{K_2}(G,X)$ is defined by the diagram

$$\begin{array}{ccccc} & Sh_{K'}(G,X) & \ni & [x,\theta K'] & \\ {}^{Sh_{(K',K_1)}}\swarrow & & \searrow & & \searrow \\ Sh_{K_1}(G,X) & & Sh_{K_2}(G,X) & \ni & [x,\theta g K_2] \end{array}$$

In general, we will not indicate K_1 and K_2 in the notation; this should not cause any confusion. Even though the $\mathcal{T}_g$ are correspondences, we will apply the usual terminology for morphisms to them. In particular, for a subvariety Z of $Sh_{K_1}(G,X)$ we write $\mathcal{T}_g(Z) \subseteq Sh_{K_2}(G,X)$ for the image of Z in the sense of correspondences.

3.5 Consider a Shimura datum (G,X) whose weight is defined over $\mathbb{Q}$. For every representation $\xi\colon G \to \mathrm{GL}(V)$ of G we naturally obtain a polarizable VHS over X with underlying bundle $X \times V$. In general we cannot expect that this descends to a VHS on $Sh_K(G,X)_{\mathbb{C}}$ for K sufficiently small. To repair this one needs to impose conditions, either on the Shimura datum or on the representation ξ.

Let us first analyze the situation. Choose a connected component X^+ of X and a compact open subgroup $K \subset G(\mathbb{A}_f)$. Then

$$Sh_K(G,X) = \coprod_g \Gamma_g \backslash X^+ ,$$

where the (finite) sum runs over a set of representatives $g \in G(\mathbb{A}_f)$ for the double coset space $G(\mathbb{Q})_+\backslash G(\mathbb{A}_f)/K$, and where $\Gamma_g = G(\mathbb{Q})_+ \cap gKg^{-1}$. The condition to get a VHS over $Sh_K(G,X)$ is that for every point $x \in X^+$ the stabilizer $\Gamma_{g,x}$ acts trivially on V. Notice that $\Gamma_g \cap Z(\mathbb{Q}) = Z(\mathbb{Q}) \cap gKg^{-1} \subseteq \Gamma_{g,x}$, where $Z = Z(G)$ is the center of G.

Let $K_\infty(x) \subset G^{\mathrm{ad}}(\mathbb{R})$ be the stabilizer inside G^{ad} of the point $x \in X$, then $K_\infty(x)$ is compact (it is in fact a maximal compact subgroup of $G^{\mathrm{ad}}(\mathbb{R})$), so that $\mathrm{ad}(gKg^{-1}) \times K_\infty(x)$ is a compact subgroup of $G^{\mathrm{ad}}(\mathbb{A})$. Since $G^{\mathrm{ad}}(\mathbb{Q})$ is discrete in $G^{\mathrm{ad}}(\mathbb{A})$ (G being affine as a variety) we conclude that for K sufficiently small we have $\Gamma_{g,x} \subseteq Z(G)(\mathbb{Q})$, hence $\Gamma_{g,x} = gKg^{-1} \cap Z(G)(\mathbb{Q})$. In fact, we can do this for all x simultaneously, since it suffices to take K neat.

One of the extra conditions we can impose is the requirement that the center Z is the almost direct product of a $\mathbb{Q}$-split torus and a torus of compact type defined over $\mathbb{Q}$. If T is the maximal $\mathbb{Q}$-split torus of Z^0, then this is equivalent to the condition that $\mathrm{Inn}(h_x(i))$ is a Cartan involution of G/T for one (equivalently: all) $x \in X$. If this is the case then $Z(\mathbb{Q})$ is discrete in $Z(\mathbb{A}_f)$. For K sufficiently small we therefore get $gKg^{-1} \cap Z(\mathbb{Q}) = \{1\}$, and ξ induces a polarizable VHS on $Sh_K(G, X)$, which we denote $\mathcal{V}(\xi)$.

Another possibility is to consider representations $\xi\colon G \to \mathrm{GL}(V)$ which are induced from a representation of G^{ad}. Again we get a polarizable VHS $\mathcal{V}(\xi)$ on $Sh_K(G, X)$ for K sufficiently small. This also works if the weight of the Shimura datum is not defined over $\mathbb{Q}$.

In the sequel we will only consider representations ξ with $\mathrm{Ker}(\xi) \subseteq Z(G)$, and for which the weight homomorphism $\mathbb{G}_{\mathrm{m}} \to G/\mathrm{Ker}(\xi)$ is defined over $\mathbb{Q}$. If ξ induces a polarizable VHS over $Sh_K(G, X)$ we denote this by $\mathcal{V}(\xi)$.

3.6 Deligne's definition of a Shimura variety as a projective system (indexed by the compact open subgroups $K \subseteq G(\mathbb{A}_f)$) of generally non-irreducible varieties $Sh_K(G, X)$ is very suited for global aspects, such as the existence of a canonical model over the reflex field $E(G, X)$. However, one would sometimes like to work with a single irreducible component of $Sh_K(G, X)$ for a fixed sufficiently small K, rather than working with the whole projective system. In particular, one would like to have a definition of which irreducible subvarieties are "Shimura subvarieties". The notion of a closed immersion of Shimura data leads to the following definition.

3.7 Definition. Let (G, X) be a Shimura datum whose weight is defined over $\mathbb{Q}$, and let K be a compact open subgroup of $G(\mathbb{A}_f)$. Suppose S is an irreducible algebraic subvariety of $Sh_K(G, X)_F$, where $F \supseteq E(G, X)$. Then S is called a subvariety of Shimura type if there exists a closed immersion $i\colon (G', X') \hookrightarrow (G, X)$ of Shimura data with $E(G', X') \subseteq F$, and a compact open subgroup $K' \subset G'(\mathbb{A}_f)$ with $i(K') \subseteq K$, such that S is an irreducible component of the image of the (finite) morphism $i_{(K',K)}\colon Sh_{K'}(G', X')_F \to Sh_K(G, X)_F$.

Another way of thinking about this is that a "Shimura subvariety" should in some sense be "defined" by the existence of certain Hodge classes in the various VHS attached to representations ξ of G.

To make this more precise, consider a representation $\xi\colon G \to \mathrm{GL}(V)$, and let H' be an algebraic subgroup of $\mathrm{GL}(V)$. We can look at the locus of points $x \in X$ such that $\xi \circ h_x$ factors through H', or equivalently, the locus $\{x \in X \mid \mathrm{MT}(\xi \circ h_x) \subseteq H'\}$. Clearly, this only depends on $H = \xi^{-1}(H')^0 \subseteq G$, which suggests that we do not really need the representation ξ. So, for an algebraic subgroup $H \subseteq G$ we define $Y_H = \{x \in X \mid h_x \text{ factors through } H_{\mathbb{R}}\}$. Then Y_H is a union of $H(\mathbb{R})$-conjugacy classes in $\mathrm{Hom}(\mathbb{S}, H_{\mathbb{R}}) \subseteq \mathrm{Hom}(\mathbb{S}, G_{\mathbb{R}})$. We claim that it is a finite union. To see this we argue as follows.

We use the fact that for an algebraic group $H_{\mathbb{R}}$ over $\mathbb{R}$ there are finitely many $H(\mathbb{R})$-conjugacy classes of maximal tori $S \subseteq H_{\mathbb{R}}$ defined over $\mathbb{R}$. Let $S_1, \ldots, S_k$ (where $S_i \subseteq H_{\mathbb{R}}$ is a maximal torus) be representatives for these conjugacy classes, and choose maximal tori $T_i \subset G_{\mathbb{R}}$ with $S_i \subseteq T_i$. Write $W_{\mathbb{R}}(S_i)$ for the real Weyl group corresponding to S_i, i.e., $W_{\mathbb{R}}(S_i) = N_H(S_i)(\mathbb{R})/Z_H(S_i)(\mathbb{R})$. Similarly, we have real Weyl groups $W_{\mathbb{R}}(T_i)$.

Let $Y_\alpha \subseteq Y_H$ be an $H(\mathbb{R})$-conjugacy class. We can choose an index i such that every $h_x\colon \mathbb{S} \to H_{\mathbb{R}}$ with $x \in Y_\alpha$ is $H(\mathbb{R})$-conjugate to a homomorphism that factors through S_i. In this way, Y_α gives rise to a well-determined element

$$\mathrm{cl}_H(Y_\alpha) \in W_{\mathbb{R}}(S_i)\backslash \mathrm{Hom}(X^*(S_i), X^*(\mathbb{S}))\,,$$

by which it is determined. In a similar way it gives rise to a class $\mathrm{cl}_G(Y_\alpha)$ in $W_{\mathbb{R}}(T_i)\backslash \mathrm{Hom}(X^*(T_i), X^*(\mathbb{S}))$. However, since all Y_α are contained in a single $G(\mathbb{R})$-conjugacy class X, the element $\mathrm{cl}_G(Y_\alpha)$ is independent of α. Therefore, $\mathrm{cl}_H(Y_\alpha)$ is in the image under $X^*(T_i) \twoheadrightarrow X^*(S_i)$ of the finite set of representatives for this class $\mathrm{cl}_G(Y_\alpha)$. It follows that there are only finitely many possibilities for $\mathrm{cl}_H(Y_\alpha)$, hence Y_H is a finite union of $H(\mathbb{R})$-conjugacy classes.

3.8 Definition. An irreducible algebraic subvariety $S \subseteq Sh_K(G,X)_{\mathbb{C}}$ is called a subvariety of Hodge type if there exist an algebraic subgroup $H \subseteq G$ (defined over $\mathbb{Q}$), an element $\eta \in G(\mathbb{A}_f)$ and a connected component Y_H^+ of Y_H such that S is the image of $Y_H^+ \times \eta K$ in $Sh_K(G,X)$. For a subfield F of $\mathbb{C}$ which contains $E(G,X)$ we say that an irreducible algebraic subvariety $S \subseteq Sh_K(G,X)_F$ is a subvariety of Hodge type if every irreducible component of $S \otimes_F \mathbb{C}$ is of Hodge type in the sense just defined.

3.9 Remark. The reader should not confuse subvarieties of Hodge type and Shimura varieties of Hodge type (see 0.6).

For $(G,X) = (\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$, a subvariety of Hodge type is the same as that which is called a Kuga variety of Hodge type in some literature.

For the rest of this section we consider all Shimura varieties over $\mathbb{C}$; in Chapter III we will return to Shimura varieties over number fields.

There are a couple of obvious remarks to be made. If $K_1 \subseteq K_2$ are two compact open subgroups of $G(\mathbb{A}_f)$ then $S \subseteq Sh_{K_2}(G,X)$ is a subvariety of Shimura type if and only if it is the image under the natural map $Sh_{K_1}(G,X) \to Sh_{K_2}(G,X)$ of a subvariety of Shimura type of $Sh_{K_1}(G,X)$. The same holds for subvarieties of Hodge type. Every irreducible component of $Sh_K(G,X)$ is of Shimura type. An irreducible component of an intersection of subvarieties of Hodge type is again of Hodge type.

One could guess that the two concepts are equivalent. Although this is not quite true in general, we will show that a subvariety of Shimura type is of Hodge type, and that, conversely, a subvariety of Hodge type is the "image" of a subvariety of Shimura type under a Hecke correspondence.

3.10 Lemma. *Let H be a subgroup of G such that Y_H is non-empty. Then $Y_H = Y_{Z(G)\cdot H}$.*

Proof. Since h_x factors through $Z(G)\cdot H$ if and only if it factors through the connected component $(Z(G)\cdot H)^0 \subseteq Z(G)^0 \cdot H$, it suffices to show that $Y_H = Y_{Z(G)^0\cdot H}$. Write $Z^0 = Z(G)^0$, $T = (Z^0 \cap H)\cdot(Z^0 \cap G^{\mathrm{der}}) \subseteq Z^0$, and, for $y \in Y_{Z(G)^0\cdot H}$, consider the homomorphisms

$$f_1\colon\ \mathbb{S} \xrightarrow{h_y} G_{\mathbb{R}} \longrightarrow G_{\mathbb{R}}^{\mathrm{ab}} \cong Z^0/(Z^0\cap G^{\mathrm{der}}) \longrightarrow Z^0/T\,,$$

$$f_2\colon\ \mathbb{S} \xrightarrow{h_y} Z^0\cdot H \longrightarrow (Z^0\cdot H)/H \cong Z^0/(Z^0\cap H), \quad \text{and}$$

$$\pi\colon\ Z^0/(Z^0\cap H) \longrightarrow Z^0/T\,.$$

One easily checks that $\pi\circ f_2 = f_1$. The image of f_1 does not depend on y, since the image of $\mathbb{S}$ in $G_{\mathbb{R}}^{\mathrm{ab}}$ is already independent of y. The condition that Y_H is non-empty means that there exists a $y \in Y_{H\cdot Z(G)^0}$ for which the image of f_2 is the identity element in $Z^0/(Z^0\cap H)$. Therefore the image of f_2 is contained in $\mathrm{Ker}(\pi)$ for every $y \in Y_{H\cdot Z(G)^0}$. Since π is an isogeny ($Z^0\cap G^{\mathrm{der}}$ being a finite group) and $\mathbb{S}$ is connected, we conclude that f_2 is trivial for every y, which proves the lemma. □

3.11 Proposition. *(i) Given a subvariety $S \subseteq Sh_K(G,X)$ of Hodge type then there exists a compact open subgroup K' of $G(\mathbb{A}_f)$ contained in K, a representation $\xi\colon G \to \mathrm{GL}(V)$ which is induced from a faithful representation of G^{ad}, and an algebraic subgroup $M \subseteq \mathrm{GL}(V)$, such that*

1. *ξ induces a polarizable VHS $\mathcal{V}(\xi)$ over $Sh_{K'}(G,X)$,*
2. *S is the image under the natural map $Sh_{K'}(G,X) \to Sh_K(G,X)$ of an irreducible subvariety $S' \subseteq Sh_{K'}(G,X)$ such that S' is a maximal irreducible subvariety with generic Mumford-Tate group M.*

(ii) For every algebraic subgroup $H \subseteq G$ and $\eta \in G(\mathbb{A}_f)$, the image of $Y_H \times \eta K$ in $Sh_K(G,X)$ is an algebraic subvariety.

Note that it makes sense to state that $\mathcal{V}(\xi)|_{S'}$ has generic Mumford-Tate group M, since the irreducible component of $Sh_{K'}(G,X)$ containing S' is a quotient of X, and over X the bundle underlying $\mathcal{V}(\xi)$ is just $X \times V$.

Proof. Using the lemma we can describe S as the image of some $Y_H^+ \times \eta K$, where H contains $Z(G)$. Choose a faithful representation $\xi^{\mathrm{ad}}\colon G^{\mathrm{ad}} \to \mathrm{GL}(V)$, and write ξ for the induced representation of G. For $K' \subseteq K$ sufficiently small we get a polarizable VHS $\mathcal{V}(\xi)$ over $Sh_{K'}(G,X)$, see 3.5. Let S' be an irreducible component of $Sh_{(K',K)}^{-1}(S)$, and define $M \subseteq \xi(H)$ as the generic Mumford-Tate group of $\mathcal{V}(\xi)|_{S'}$. Clearly, $\xi^{-1}(M)^0 \subseteq H$ and $Y_H^+ \subseteq Y_{\xi^{-1}(M)} \subseteq Y_H$, hence Y_H^+ is a connected component of $Y_{\xi^{-1}(M)}$. Statement (i) readily follows.

With a similar argument, the second statement follows from the results of [8]. □

3.12 Proposition. *(i) A subvariety of Shimura type is of Hodge type.*

(ii) If S is a subvariety of a connected Shimura variety $Sh_K^0(G,X)$ then S is of Hodge type if and only if it is of Shimura type.

(iii) A subvariety S is of Hodge type if and only if there is a subvariety S' of Shimura type such that S is an irreducible component of $\mathcal{T}_\eta(S')$ for some $\eta \in G(\mathbb{A}_f)$.

Proof. If $i\colon (G',X') \hookrightarrow (G,X)$ is a closed immersion of Shimura data then, as we have shown before, $Y_{G'}$ is a finite union of $G'(\mathbb{R})$-conjugacy classes in $\mathrm{Hom}(\mathbb{S},G')$, hence X' is a union of components of $Y_{G'}$, which implies (i).

For (ii), choose a connected component $X^+ \subseteq X$, and consider a subvariety of Hodge type $S \subseteq Sh_K^0(G,X)$. By definition, S is the image of some $Y_H^+ \times \eta K$. The class of η in $G(\mathbb{Q})_+\backslash G(\mathbb{A}_f)/K$ is the same as the class of the identity element $e \in G(\mathbb{A}_f)$. Possibly after replacing H by a conjugate subgroup of G, we may therefore assume that $Y_H^+ \subseteq X^+$ and $\eta = e$.

We choose K', ξ, M and S' as in Proposition 3.11 (i). According to this proposition and its proof, Y_H^+ is a connected component of $Y_{G'}$, where $G' = \xi^{-1}(M)^0$. From the description of M as the generic Mumford-Tate group of a polarizable VHS it easily follows that M satisfies axioms (2.1.1.2-3) of [18] with respect to the h_y for $y \in Y^+$.

Let X' be the $G'(\mathbb{R})$ conjugacy class containing $Y^+ = Y_H^+ \subseteq Y_{G'}$. We are done if we show that (G', X') is a Shimura datum. Axiom (2.1.1.1) of loc. cit. follows from the inclusion $G' \subseteq G$ and the fact that (G, X) is a Shimura datum. Axioms (2.1.1.2) and (2.1.1.3) follow from the corresponding properties of M, since we have a surjective homomorphism $G' \to M$ with kernel $Z(G) \cap G'$, which is a torus.

For the last statement, consider a subvariety $S \subseteq Sh_K(G, X)$ of Hodge type, say the image of some $Y_H^+ \times \eta K$. If $X^+ \subseteq X$ is the connected component containing Y_H^+, then the image of $Y_H^+ \times e \cdot \eta K \eta^{-1}$ in $Sh^0_{\eta K \eta^{-1}}(G, X)$ is an algebraic subvariety S' (Proposition 3.11 (ii)) which by (ii) is of Shimura type. Clearly, S is a component of the image of S' under the Hecke correspondence $\mathcal{T}_\eta$ from $Sh_{\eta K \eta^{-1}}(G, X)$ to $Sh_K(G, X)$. The statement in the opposite direction is clear. □

3.13 Remark. Even though we have formulated and proved the proposition over the field $\mathbb{C}$, the conclusions are valid over all fields $F \subseteq \mathbb{C}$ which are "large enough" in a suitable sense. For example, in (ii) one has the statement that if F is a field containing the field of definition of $Sh^0_K(G, X)$ (which depends on K) and if S is a subvariety of Hodge type of $Sh^0_K(G, X)_F$, then there is a finite field extension $F \subseteq F'$ such that all irreducible components of $S_{F'}$ are of Shimura type.

3.14 In general, a subvariety of Hodge type is not of Shimura type. The point is that if we try to construct a Shimura datum (G', X') as in the proof of (ii), then S may lie in a component of $Sh_K(G, X)$ which is not in the image of $Sh(G', X')$. We can give an example where this is actually the case.

Let E be an imaginary quadratic subfield of $\mathbb{C}$. The choice of a $\mathbb{Q}$-basis for E determines a CM-point $x \in \mathfrak{H}_1^\pm$, and for $\eta \in \mathrm{GL}_2(\mathbb{A}_f)$ the image $S = S_{\eta K}(\{x\})$ of $\{x\} \times \eta K$ in $Sh_K(\mathrm{GL}_2, \mathfrak{H}_1^\pm)$ is a subvariety of Hodge type. Let $T_E = \mathrm{Res}_{E/\mathbb{Q}}\mathbb{G}_\mathrm{m}$, which, by our choice of a basis, can be viewed as a subtorus of GL_2. It has the property that

$\{x\}$ is a connected component of Y_{T_E}, and it is in fact the unique connected subgroup of GL_2 with this property. If $S_{\eta K}(\{x\})$ is of Shimura type we must therefore have that the class of η in

$$\pi_0(Sh_K(\mathrm{GL}_2, \mathfrak{H}_1^{\pm})) \cong \pi_0(\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}))/(K \times K_\infty)$$

is in the image of

$$\pi_0(Sh(T_E, \{x\})) \cong \pi_0(E^*\backslash\mathbb{A}_E^*)\,,$$

where K_∞ is the stabilizer in $\mathrm{GL}_2(\mathbb{R})$ of a point in $\mathfrak{H}_1^{\pm}$. The determinant homomorphism det: $\mathrm{GL}_2 \to \mathbb{G}_\mathrm{m}$ induces an isomorphism $\pi_0(\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A})) \xrightarrow{\sim} \pi_0(\mathbb{Q}^*\backslash\mathbb{A}^*)$ ([14, Théorème 2.4]), and class field theory gives us isomorphisms

$$\mathrm{rec}_\mathbb{Q}\colon\ \pi_0(\mathbb{Q}^*\backslash\mathbb{A}^*) \xrightarrow{\sim} \mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q})\,,$$
$$\mathrm{rec}_E\colon\ \pi_0(E^*\backslash\mathbb{A}_E^*) \xrightarrow{\sim} \mathrm{Gal}(E^{\mathrm{ab}}/E)\,,$$

such that the map $\pi_0(E^*\backslash\mathbb{A}_E^*) \to \pi_0(\mathbb{Q}^*\backslash\mathbb{A}^*)$ corresponds to the natural homomorphism on the abelian Galois groups. The image of $(K \times K_\infty)$ in $\mathbb{Q}^*\backslash\mathbb{A}^*$ is an open subgroup, corresponding to a class field $F \subseteq \mathbb{Q}^{\mathrm{ab}}$, which, for K sufficiently small, contains E. In this case we see that the natural map $\mathrm{Gal}(E^{\mathrm{ab}}/E) \to \mathrm{Gal}(F/\mathbb{Q})$ is not surjective. This means that we can find η and K such that $S_{\eta K}(\{x\})$ is not of Shimura type.

3.15 Let Z be an irreducible algebraic subvariety of a Shimura variety $Sh_K(G, X)$. It is clear from the remarks after Definition 3.8 that there exists a smallest subvariety of Hodge type, say S, containing Z. By definition, it is an irreducible component of the image of $Y_M \times \eta K$ in $Sh_K(G, X)$, where $M \subseteq G$ is an algebraic subgroup (over $\mathbb{Q}$) and $\eta \in G(\mathbb{A}_f)$. If $Y_M^+ \subset Y_M$ is a connected component such that S is the image of $Y_M^+ \times \eta K$ then we write $S = S_{\eta K}(Y_M^+)$.

This description does not uniquely determine the group M. However, as we have seen, we can take for M the "generic Mumford-Tate group on Z". More precisely, let $K' \subset G(\mathbb{A}_f)$ be a compact open subgroup contained in K and let $\xi\colon G \to \mathrm{GL}(V)$ be a representation such that we obtain a polarizable VHS $\mathcal{V}(\xi)$ over $Sh_{K'}(G, X)$. Let $Z' \hookrightarrow S'$ be irreducible components of $Sh_{(K',K)}^{-1}(Z)$ and $Sh_{(K',K)}^{-1}(S)$ respectively. The generic Mumford-Tate group MT of $\mathcal{V}(\xi)|_{Z'}$ is equal to that of $\mathcal{V}(\xi)|_{S'}$ and we may choose M in the above such that MT is conjugated to $\xi(M)$ (for all representations ξ which induce a VHS for K' sufficiently small). Up to conjugation by elements of $G(\mathbb{Q})$ this uniquely determines $Z(G) \cdot M \subseteq G$.

§4 Hermitian symmetric domains

The purpose of this section is to give a brief account of some facts on Hermitian symmetric domains that can be found in the literature. The basic references are [18], [29], [32] and [55]. We often identify a linear algebraic group H over $\mathbb{R}$ ($\mathbb{C}$) with the real (complex) Lie group $H(\mathbb{R})$ ($H(\mathbb{C})$).

4.1 A Hermitian symmetric space is a connected complex manifold X with a Hermitian structure on its tangent bundle such that for every $x \in X$ there exists a holomorphic involutive isometry s_x of X having x as an isolated fixed point.

Let X be a Hermitian symmetric space, and let $A(X)$ denote the group of holomorphic isometries of X. It is a real Lie group acting transitively on X. The same is true for its connected component of the identity $A(X)^+$. If $K_x \subseteq A(X)^+$ is the stabilizer of a point x we therefore get an $A(X)^+$-equivariant diffeomorphism $A(X)^+/K_x \xrightarrow{\sim} X$. The universal covering space $\widetilde{X}$ is again a Hermitian symmetric space. It has a canonical decomposition (de Rham decomposition)

$$\widetilde{X} = \widetilde{X}_f \times \widetilde{X}_1 \times \cdots \widetilde{X}_r\,,$$

where $\widetilde{X}_f$ is isomorphic to $\mathbb{C}^n$ for some $n \geq 0$ (the flat part), and where $\widetilde{X}_i$ (for $i \in \{1, \ldots, r\}$) is an irreducible non-flat Hermitian symmetric space.

If $\widetilde{X}$ is compact then X is said to be of the compact type; if the flat part is reduced to a point and all factors $\widetilde{X}_i$ are non-compact then X is said to be of the non-compact type. In both cases X is simply connected and the Lie group $A(X)^+$ can be described as $A(X)^+ = G(\mathbb{R})^+$ for a semi-simple algebraic group G over $\mathbb{R}$, which is simple if and only if X is irreducible.

4.2 Suppose X is of non-compact type. Choose a point $x \in X$, let $K = K_x$, and write $\mathfrak{g} = \mathrm{Lie}(G) = \mathrm{Lie}(A(X))$, $\mathfrak{k} = \mathrm{Lie}(K)$. We have $X \cong G/K$. The element $\theta = \mathrm{Ad}(s_x)$ is a Cartan involution of $\mathfrak{g}$ with $\mathfrak{k} = \mathfrak{g}(\theta; 1)$. Put $\mathfrak{p} = \mathfrak{g}(\theta; -1)$, then the decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ is called a Cartan decomposition, and there is a natural identification of $\mathfrak{p}$ with the tangent space $T_x(X)$. The complex structure J_0 on $\mathfrak{p}$ can be described as $J_0 = \mathrm{ad}_{\mathfrak{p}}(H_0)$, where H_0 is an element of the center of $\mathfrak{k}$ with $\mathrm{ad}_{\mathfrak{p}}(H_0)^2 = -\mathrm{Id}$. The triplet $(\mathfrak{g}, \theta, H_0)$ is called a Lie algebra of Hermitian type.

Let $\mathfrak{p}_+ = \mathfrak{p}_{\mathbb{C}}(J_0, i)$, $\mathfrak{p}_- = \mathfrak{p}_{\mathbb{C}}(J_0, -i)$, then $\mathfrak{g}_{\mathbb{C}} = \mathfrak{p}_+ + \mathfrak{k}_{\mathbb{C}} + \mathfrak{p}_-$. There is an embedding, called the Harish-Chandra embedding, of X into $\mathfrak{p}_+$, which identifies X with a bounded domain $\mathcal{D} \subset \mathfrak{p}_+$. Conversely, for every bounded domain $\mathcal{D} \subset \mathbb{C}^n$ which is symmetric (in the sense that for every point $x \in \mathcal{D}$ there is a holomorphic map $s_x \colon \mathcal{D} \to \mathcal{D}$ with $s_x^2 = \mathrm{Id}$ having x as an isolated fixed point) there is a holomorphic diffeomorphism of $\mathcal{D}$ to a Hermitian symmetric space of the non-compact type. For this reason the Hermitian symmetric spaces of the non-compact type are often called bounded symmetric domains, or Hermitian symmetric domains, as we shall henceforth do.

The involution $\theta = \mathcal{A}d(s_x)$ is a Cartan involution on G. This means that the form $G^{(\theta)}$ of G with complex conjugation $g \mapsto \theta(\bar{g})$ is compact. The Lie algebra of $G^{(\theta)}$ is given by $\mathrm{Lie}(G^{(\theta)}) = \mathfrak{k} + i \cdot \mathfrak{p} \subset \mathfrak{g}_{\mathbb{C}}$. If K^0 denotes the connected component of the identity of K (which is also connected in the analytic topology) then $\check{X} = G^{(\theta)}/K^0$ is a Hermitian symmetric space of the compact type, called the compact dual of X. If P_- is the connected algebraic subgroup of $G_{\mathbb{C}}$ with Lie algebra $\mathfrak{p}_-$ then there is an isomorphism $\check{X} = G^{(\theta)}/K^0 \cong G_{\mathbb{C}}/K^0_{\mathbb{C}}P_-$. Note however that $G^{(\theta)}$ acts by isometries on $\check{X}$, whereas $G_{\mathbb{C}}$ in general does not.

We have $K^0 = G^0_{\mathbb{R}} \cap K^0_{\mathbb{C}}P_-$, so we obtain an embedding of X into $\check{X}$ by

$$X \cong G^0/K^0 \hookrightarrow G_{\mathbb{C}}/K^0_{\mathbb{C}}P_- \cong \check{X}\,.$$

This identifies X with an open submanifold of $\check{X}$ which is contained in the big open cell obtained from the Bruhat decomposition with respect to a suitable Borel subgroup of $G^0_{\mathbb{C}}$.

4.3 Let $G_{\mathbb{R}}$ be a linear algebraic group over $\mathbb{R}$, and let X^+ be a connected component of the space $\mathrm{Hom}(\mathbb{S}, G)$ of homomorphisms (of algebraic groups) from $\mathbb{S}$ to $G_{\mathbb{R}}$. For every representation $G \to \mathrm{GL}(V)$ we get a collection of $\mathbb{R}$-Hodge structures on V parametrized by X^+. We impose the conditions on the pair $(G_{\mathbb{R}}, X^+)$ that there exist a faithful representation V of $G_{\mathbb{R}}$ such that (i) the weight filtration on V does not depend on $h \in X^+$, (ii) there exists a structure of complex manifold on X^+ such that the Hodge structures (V, h) form a variation of Hodge structure $\mathcal{V}$ over X^+, and (iii) every homogeneous component of $\mathcal{V}$ is polarizable.

Pairs $(G_{\mathbb{R}}, X^+)$ satisfying these conditions were studied by Deligne in [18, Section 1]. He showed that for such a pair $(G_{\mathbb{R}}, X^+)$, the space X^+ (with its structure of complex manifold which is uniquely determined by the second condition) is a Hermitian symmetric domain and that, conversely, every Hermitian symmetric domain arises in this way. Furthermore, Deligne showed that (i), (ii) and (iii) can be formulated purely in terms of the group $G_{\mathbb{R}}$ and the action of $\mathbb{S}$ on $G_{\mathbb{R}}^{\mathrm{ad}}$ (via $h \in X^+$); this is the stepping stone to the "abstract" definition of a Shimura datum as in [18, Section 2].

In this context, the compact dual of a Hermitian symmetric domain has a natural interpretation as a flag manifold. For a given representation of $G_{\mathbb{R}}$ the embedding of X^+ into its compact dual $\check{X}^+$ corresponds to the application

$$\begin{pmatrix} \text{the Hodge structure on} \\ V \text{ defined by } h \in X^+ \end{pmatrix} \mapsto \begin{pmatrix} \text{the corresponding Hodge} \\ \text{filtration on } V \end{pmatrix}.$$

4.4 Let X and X' be two Hermitian symmetric domains with base points o and o', and write $(\mathfrak{g}, \theta, H_0)$ and $(\mathfrak{g}', \theta', H_0')$ for the associated Lie algebras of Hermitian type. A holomorphic map $f\colon X \to X'$ with $f(o) = o'$ is called strongly equivariant if there exists a homomorphism of Lie algebras $\rho\colon \mathfrak{g} \to \mathfrak{g}'$ such that $\rho \circ \theta = \theta' \circ \rho$ and such that $f(\exp(\alpha) \cdot x) = \exp(\rho(\alpha)) \cdot f(x)$ for all $\alpha \in \mathfrak{g}$ and $x \in X$. One can show that ρ (if it exists) is uniquely determined by f; its restriction $\rho_{|\mathfrak{p}}\colon \mathfrak{p} \to \mathfrak{p}'$ is $\mathbb{C}$-linear and therefore satisfies

$$\rho([H_0, \alpha]) = [H_0', \rho(\alpha)] \qquad \text{for all } \alpha \in \mathfrak{g}\,.$$

A homomorphism $\rho\colon \mathfrak{g} \to \mathfrak{g}'$ satisfying the latter condition and the condition $\rho \circ \theta = \theta' \circ \rho$ is called a (H_1)-homomorphism of Lie algebras of Hermitian type.

We thus see that a strongly equivariant holomorphic map f gives rise to an (H_1)-homomorphism ρ. Conversely, every such ρ corresponds to a strongly equivariant holomorphic map f. For $\alpha \in \mathfrak{g}$ and $x \in X$, the geodesic $\exp(t\alpha) \cdot x$ in X is mapped under f to the geodesic $\exp(t\rho(\alpha)) \cdot f(x)$ in X'. One can show that X decomposes as a product $X = X^{(1)} \times X^{(2)}$ such that f factors as $f = j \circ \mathrm{pr}_1$, where $\mathrm{pr}_1\colon X \to X^{(1)}$ is the first projection map, and $j\colon X^{(1)} \hookrightarrow X'$ is a strongly equivariant holomorphic map identifying $X^{(1)}$ with a totally geodesic submanifold of X'. (For the definition of a totally geodesic submanifold, see [29, Ch. 1, §14].)

For a complete totally geodesic complex submanifold $Y \subseteq X$, the subgroup $A(X)_Y \subseteq A(X)$ defined by $A(X)_Y = \{\varphi \in A(X) \mid \varphi(Y) = Y\}$ acts transitively on Y (see [32, IX, Theorem 4.2]). It follows that Y is a symmetric subspace of X, and by considering the sectional curvature (see [29, III, Theorem 3.1]) we see that it has non-positive curvature. However, since X has a realization as a bounded domain, the complex submanifold Y cannot have a non-trivial flat part, and we conclude that Y is a Hermitian symmetric subdomain of X.

Chapter II

A characterization of subvarieties of Hodge type

§1 Properties of subvarieties of Hodge type

1.1 Shimura varieties are interesting for various reasons. Although their definition seems to be a matter of Hodge theory, it turns out they have highly non-trivial arithmetic properties as well. In some cases Shimura varieties have an interpretation (depending on the choice of a representation) as moduli spaces of abelian varieties with certain additional structures; more generally they are believed to have a meaning as moduli spaces for motives. In Deligne's words: "Pour interpréter des structures de Hodge de type plus compliqué, on aimerait remplacer les variétés abéliennes par des "motifs" convenables, mais il ne s'agit encore que d'un rêve." ([18, p. 248]).

One and a half decades later, these words still apply; see for example the discussion in Sections 1.6–8 of Deligne's paper [20]. The theory of motives has not yet been developed far enough to formulate a modular interpretation of Shimura varieties in general. It therefore seems very difficult to set up a good theory in mixed and positive characteristics. To our knowledge there is no completely satisfactory definition of an "integral canonical model" of a Shimura variety. Note for example that [37, Proposition 2.13] is based on [25, Theorem V.6.7], which, however, is false in general. This last fact is discussed in some detail in the forthcoming paper [13]. Nevertheless, there are many interesting results on Shimura varieties in relation to motives; we refer to [38] for more on this subject and to [65] for important new results on integral models.

The main problem that we want to study here is the question what varieties can arise as Shimura subvarieties in a larger Shimura variety. For example, we could ask how to characterize the Shimura subvarieties of the moduli space A_g of abelian varieties. It is because we are interested in Shimura varieties as subvarieties that we introduced in Chapter I the notions of subvariety of Shimura type and of Hodge type.

In this chapter we give a characterization of subvarieties of Hodge type as totally geodesic subvarieties containing at least one special point. We also try to understand totally geodesic subvarieties in general. On the basis of our results we can give some complements to the work of Faltings, Saito and others on non-rigid families of abelian varieties.

But first, let us inspect our material and describe some of its properties.

1.2 Fix a Shimura datum (G, X) and a compact open subgroup $K \subset G(\mathbb{A}_f)$. As explained in [18, Section 1] the manifolds X that occur as part of a Shimura datum are finite unions of Hermitian symmetric domains. For K sufficiently small, the Shimura variety $Sh_K(G, X)$ is non-singular and inherits a natural structure of locally symmetric space.

If $S \hookrightarrow Sh_K(G, X)$ is a subvariety of Hodge type then there is a closed immersion of Shimura data $(G', X') \hookrightarrow (G, X)$ such that S is covered by a connected component of X'. The space X' is a Hermitian symmetric subdomain of X. In particular, it is a complete totally geodesic submanifold.

We say that an irreducible algebraic subvariety S of $Sh_K(G, X)$ is totally geodesic if it is covered by a totally geodesic submanifold $X' \subseteq X$, which then is a Hermitian symmetric subdomain of X (cf. the discussion in I.4.4). We can also express this condition in terms of the metric on $Sh_K(G, X)$, provided that we take possible singularities into account. For example, if $Sh_K(G, X)$ is non-singular, then $S \hookrightarrow Sh_K(G, X)$ is a totally geodesic subvariety if and only if every geodesic in $Sh_K(G, X)$ which is tangent to S at a regular point $P \in S^{\mathrm{reg}}$ is a curve in S.

The above remarks show that a subvariety of Hodge type is totally geodesic. Clearly, it is not true that, conversely, every totally geodesic subvariety is of Hodge type. For a trivial example, a single point is a totally geodesic subvariety but it is of Hodge type only if it is special. One of the main goals of this chapter is to clarify the relation between subvarieties of Hodge type and totally geodesic subvarieties.

1.3 Another interesting property of subvarieties of Hodge type is that the special points on them lie dense. Recall that if (G, X) is a Shimura datum, then a point $x \in X$ is called a special point if the corresponding homomorphism $h_x \colon \mathbb{S} \to G_{\mathbb{R}}$

factors through a subtorus of G, defined over $\mathbb{Q}$. A point in $Sh_K(G,X)$ is called special if it is of the form $[g,x]$ with $x \in X$ special. In case there is a closed immersion of Shimura data $i\colon (G,X) \hookrightarrow (\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ (see Section I.3.2) we call such points CM-points, since they are precisely the points that correspond to abelian varieties of CM-type.

In [14, §5], it is shown that the special points are dense in $Sh_K(G,X)$ (even for the analytic topology). The first step is to show that there is at least one special point; this is done (in a much stronger form) in ibid., Théorème 5.1, also see [42, §3]. Once we know that there are special points, we can use the fact that the set of special points in X is stable under the action of $G(\mathbb{Q})$. Notice that $G(\mathbb{Q})$ is analytically dense in $G(\mathbb{R})$; this follows from the unirationality of G and the fact that the (analytic) closure of $G(\mathbb{Q})$ in $G(\mathbb{R})$ is a subgroup (see also op. cit., 0.4). Similarly, the special points of $Sh(G,X)$ are stable under the action of $G(\mathbb{A}_f)$ and every $G(\mathbb{A}_f)$-orbit in $Sh(G,X)$ is Zariski dense (ibid., Proposition 5.2). These arguments also show that the special points lie dense on every subvariety $S \hookrightarrow Sh_K(G,X)$ of Hodge type.

One could ask whether the subvarieties of Hodge type are characterized by this property. For subvarieties of A_g it is in fact a conjecture of Oort that this is the case. We will discuss CM-points and Oort's conjecture in greater detail in the next two chapters. It turns out that they are strongly related to another "linearity property" of subvarieties of Hodge type, which can be considered as an analogue of "totally geodesicness" in mixed characteristic.

§2 Decomposition of the adjoint group

2.1 Consider a closed immersion $i\colon (M,Y) \hookrightarrow (G,X)$ of Shimura data. Let $M^{\mathrm{ad}} = M_1 \times M_2$ be a decomposition of the adjoint group of M. (We do not assume M_1 and M_2 to be non-trivial or $\mathbb{Q}$-simple.) There is a corresponding decomposition $Y = Y_1 \times Y_2$, where Y_i ($i \in \{1,2\}$) is a union of Hermitian symmetric domains, and $M_i(\mathbb{R})$ acts transitively on Y_i. One easily checks that (M_i, Y_i) is a Shimura datum, so we have a decomposition of Shimura data $(M^{\mathrm{ad}}, Y) = (M_1, Y_1) \times (M_2, Y_2)$.

Choose compact open subgroups $C_i \subset M_i(\mathbb{A}_f)$, and $C \subset M(\mathbb{A}_f)$, with $\mathrm{ad}(C) \subseteq C_1 \times C_2$. For C_1 and C_2 sufficiently small the associated morphism

$$\mathrm{ad}_{(C, C_1\times C_2)}\colon Sh_C(M,Y) \longrightarrow Sh_{C_1\times C_2}(M^{\mathrm{ad}}, Y) = Sh_{C_1}(M_1,Y_1) \times Sh_{C_2}(M_2,Y_2)$$

is finite étale on irreducible components. Given a connected component $Y_1^+ \subseteq Y_1$, a point $y_2 \in Y_2$ and a class $\theta C \in M(\mathbb{A}_f)/C$, let

$$S_{\theta C}(Y_1^+ \times \{y_2\}) \subseteq Sh_C(M, Y)$$

denote the image of $(Y_1^+ \times \{y_2\}) \times \theta C$ in $Sh_C(M, Y)$. If $\theta_1 C_1 \times \theta_2 C_2$ is the image of θC in $M_1(\mathbb{A}_f)/C_1 \times M_2(\mathbb{A}_f)/C_2$, then $S_{\theta C}(Y_1^+ \times \{y_2\})$ is an irreducible component of the inverse image of $Sh_{C_1}(M_1, Y_1) \times [y_2, \theta_2 C_2]$ under $\mathrm{ad}_{(C, C_1 \times C_2)}$. In other words, the Shimura variety $Sh_{C_1 \times C_2}(M^{\mathrm{ad}}, Y)$ is a disjoint union

$$Sh_{C_1 \times C_2}(M^{\mathrm{ad}}, Y) = Sh_{C_1}(M_1, Y_1) \times Sh_{C_2}(M_2, Y_2) = \coprod_{g_i, g_j} (\Gamma_{g_i} \backslash Y_1^+) \times (\Gamma_{g_j} \backslash Y_2^+)$$

of product varieties, and $S_{\theta C}(Y_1^+ \times \{y_2\})$ is an irreducible subvariety of $Sh_C(M, Y)$ covering some $(\Gamma_{g_i} \backslash Y_1^+) \times [y_2]$.

More generally, if $K \subset G(\mathbb{A}_f)$ is a compact open subgroup and $\eta K \in G(\mathbb{A}_f)/K$, then we define

$$S_{\eta K}(Y_1^+ \times \{y_2\})$$

as the image of $(Y_1^+ \times \{y_2\}) \times \eta K$ in $Sh_K(G, X)$. Notice that $S_{\eta K}(Y_1^+ \times \{y_2\})$ is an algebraic subvariety. This follows from the remark that $S_{\eta K}(Y_1^+ \times \{y_2\})$ is an irreducible component of $\mathcal{T}_\eta(S_{eK}(Y_1^+ \times \{y_2\}))$, and for C small enough $S_{eK}(Y_1^+ \times \{y_2\})$ is the image of $S_{eC}(Y_1^+ \times \{y_2\}) \subseteq Sh_C(M, Y)$ under the finite morphism $i_{(C,K)}$.

The subvarieties of the form $S_{\eta K}(Y_1^+ \times \{y_2\}) \subseteq Sh_K(G, X)$ are totally geodesic, since $Y_1^+ \times \{y_2\}$ is a complete totally geodesic submanifold of $Y = Y_1 \times Y_2$, and Y is totally geodesic in X. Note that $S_{\eta K}(Y_1^+ \times \{y_2\})$ contains special points if and only if y_2 is a special point of Y_2, in which case the special points are dense.

2.2 As before, let (G, X) be a Shimura datum and let $K \subset G(\mathbb{A}_f)$ be a compact open subgroup. We consider an irreducible algebraic subvariety Z of $Sh_K(G, X)$. As discussed in section I.3.15 there is a smallest subvariety of Hodge type containing Z, which we denote by $S_{\eta K}(Y_M^+)$. It corresponds to a closed immersion $i\colon (M, Y_M) \hookrightarrow (G, X)$ of Shimura data, a connected component $Y_M^+ \subseteq Y_M$ and a class $\eta K \in G(\mathbb{A}_f)/K$. If there is no risk of confusion we simply write $Y = Y_M$ and $S = S_{\eta K}(Y_M^+)$. We write $j\colon Z \hookrightarrow S$ for the inclusion map.

Possibly after replacing K by a subgroup K' of finite index and Z by an irreducible component of its preimage in $S_{\eta K'}(Y_M^+)$ we may assume that the following conditions hold.

1. there exists a representation $\xi\colon G \to \mathrm{GL}(V)$ with $\mathrm{Ker}(\xi) \subseteq Z(G)$ which induces a polarizable VHS $\mathcal{V}(\xi)$ on $Sh_K(G,X)$ such that S is a maximal irreducible subvariety with generic Mumford-Tate group $\xi(M)$ (cf. Proposition I.3.11).

2. K is neat; in particular, $Sh_K(G,X)$ is a union of quotients $\Gamma_i\backslash X^+$ such that the natural maps $X^+ \twoheadrightarrow \Gamma_i\backslash X^+$ are topological coverings and the algebraic monodromy group associated to the VHS $\mathcal{V}(\xi)$ over $\Gamma_i\backslash X^+$ is connected.

3. the natural map $u_S\colon Y^+ \twoheadrightarrow S = S_{\eta K}(Y^+)$ is a topological covering.

For the last condition we need a lemma.

2.3 Lemma. *(i) For K sufficiently small the natural map $u_S\colon Y^+ = Y_M^+ \twoheadrightarrow S = S_{\eta K}(Y_M^+)$ is a topological covering.*

(ii) Let a decomposition $(M^{\mathrm{ad}}, Y) = (M_1, Y_1) \times (M_2, Y_2)$ be given. For K sufficiently small the map $Y_1^+ \twoheadrightarrow S_{\eta K}(Y_1^+ \times \{y_2\})$ is a covering map for every $y_2 \in Y_2$ and $\eta \in G(\mathbb{A}_f)$.

Proof. Take a compact open subgroup $C \subseteq M(\mathbb{A}_f)$ such that $Sh_C(M,Y)$ is non-singular. By [14, Proposition 1.15] there exists a compact open subgroup $K \subset G(\mathbb{A}_f)$ such that $i(C) \subseteq K$ and such that $i_{(C,K)}\colon Sh_C(M,Y) \to Sh_K(G,X)$ is a closed immersion. For this choice of C and K the map $Y^+ \twoheadrightarrow S_{eK}(Y^+)$ therefore is a topological covering.

Let $K' = K \cap \eta^{-1}K\eta$. Clearly, both K' and $\eta K'\eta^{-1}$ are contained in K; in particular, $Y^+ \twoheadrightarrow S_{e(\eta K'\eta^{-1})}(Y^+)$ is a covering. Let X^+ be the component of X containing Y^+. As one easily verifies, the map $S_{e(\eta K'\eta^{-1})}(X^+) \to S_{\eta K'}(X^+)$ obtained by sending the class $[x, eK]$ to the class $[x, \eta K]$ is an isomorphism, compatible with the uniformization maps from X^+. Restricting this to Y^+ we conclude that $Y^+ \twoheadrightarrow S_{\eta K'}(Y^+)$ is a covering, which proves the first part of the lemma.

The second statement easily follows from the first one and [4, Corollaire 8.10]. □

Let $\widetilde{Z}$ be a connected component of $u_S^{-1}(Z)$. The map $u_Z = u_S|_{\widetilde{Z}}\colon \widetilde{Z} \to Z$ is again a topological covering. If $\mathrm{Cov}(u_S)$ and $\mathrm{Cov}(u_Z)$ denote the groups of covering transformations, then $\mathrm{Cov}(u_Z) = \{\gamma \in \mathrm{Cov}(u_S) \mid \gamma\widetilde{Z} = \widetilde{Z}\}$. In general the analytic space $\widetilde{Z}$ is not irreducible[1], and in some arguments to follow this causes problems. To

[1]We do not know whether it is possible in general to choose K small enough such that $\widetilde{Z}$ is irreducible.

circumvent these, we consider the normalization $n\colon Z^n \to Z$ of Z. Let $u_{Z^n}\colon \widetilde{Z^n} \twoheadrightarrow Z^n$ be a universal covering of Z^n. (Caution: our notations may be somewhat misleading, since $\widetilde{Z}$ need not be a *universal* covering of Z, whereas we do write $\widetilde{Z^n}$ for a universal covering of Z^n.) The analytic space $\widetilde{Z^n}$ is connected and normal; in particular it is irreducible.

Let $\mathcal{C} \subseteq \widetilde{Z}$ be an irreducible component. Choose a Hodge generic and regular base point $z \in Z$, i.e., a regular point outside the locus Σ as in Section I.2.2 (applied to the VHS $\mathcal{V}(\xi)$ over S). This is possible, since $Z \subseteq \Sigma$ would contradict the fact that S is the smallest subvariety of Hodge type containing Z. Also choose base points $\tilde{z} \in \mathcal{C}$, $\zeta \in Z^n$ and $\tilde{\zeta} \in \widetilde{Z^n}$ with $u_Z(\tilde{z}) = z$, $n(\zeta) = z$ and $u_{Z^n}(\tilde{\zeta}) = \zeta$. There is a well-determined morphism $\tilde{n}\colon \widetilde{Z^n} \to \widetilde{Z}$ with $\tilde{n}(\tilde{\zeta}) = \tilde{z}$ and $u_Z \circ \tilde{n} = n \circ u_{Z^n}$. We have $\tilde{n}(\widetilde{Z^n}) = \mathcal{C} \subseteq \widetilde{Z}$.

In a diagram, the situation looks as follows:

$$\begin{array}{ccccccccc}
\widetilde{Z^n} & \xrightarrow{\tilde{n}} & \mathcal{C} & \hookrightarrow & \widetilde{Z} & \hookrightarrow & Y^+ & \hookrightarrow & X \\
\downarrow{\scriptstyle u_{Z^n}} & & & \searrow & \downarrow{\scriptstyle u_Z} & & \downarrow{\scriptstyle u_S} & & \downarrow \\
Z^n & \xrightarrow{n} & & & Z & \hookrightarrow & S = S_{\eta K}(Y^+) & \hookrightarrow & Sh_K(G,X)
\end{array}$$

The choice of the point $\tilde{z}$ above z gives an identification of the fibre $\mathcal{V}(\xi)_z$ with V, and we identify the Mumford-Tate group at z with $\xi(M) \subseteq \mathrm{GL}(V)$. This is also the generic Mumford-Tate group of the VHS $n^*\mathcal{V}(\xi)$ over Z^n, via the identification $(n^*\mathcal{V}(\xi))_\zeta \cong \mathcal{V}(\xi)_z = V$. Since $\mathrm{Ker}(\xi) \subseteq Z(G)$ there is a natural surjective homomorphism $f\colon \xi(M) \to M^{\mathrm{ad}}$, and composing the monodromy representation $\rho_S\colon \pi_1(S,z) \to \xi(M)(\mathbb{Q}) \subseteq \mathrm{GL}(V)$ with f we obtain a homomorphism $f \circ \rho_S\colon \pi_1(S,z) \to M^{\mathrm{ad}}(\mathbb{Q}) \subset M^{\mathrm{ad}}(\mathbb{R})$.

2.4 Lemma. *We have* $\mathrm{Im}(f \circ \rho_S) \subseteq M^{\mathrm{ad}}(\mathbb{R})^+$, *and there is a commutative diagram*

$$\begin{array}{ccccccccc}
\pi_1(Z^n,\zeta) & & \xrightarrow{n_*} & & \pi_1(Z,z) & \xrightarrow{j_*} & \pi_1(S,z) & \xrightarrow{f\circ\rho_S} & M^{\mathrm{ad}}(\mathbb{R})^+ \\
\downarrow{\scriptstyle\cong} & & & & \downarrow & & \downarrow{\scriptstyle\cong} & & \downarrow \\
\mathrm{Cov}(u_{Z^n}) & \longrightarrow & \mathrm{Cov}(u_Z)_{\mathcal{C}} & \hookrightarrow & \mathrm{Cov}(u_Z) & \hookrightarrow & \mathrm{Cov}(u_S) & \hookrightarrow & \mathrm{Aut}(Y^+)
\end{array}$$

where $\mathrm{Cov}(u_Z)_{\mathcal{C}} = \{\gamma \in \mathrm{Cov}(u_Z) \mid \gamma\mathcal{C} = \mathcal{C}\}$.

Proof. It is clear that $\pi_1(Z^n,\zeta) \cong \mathrm{Cov}(u_{Z^n})$ maps into $\mathrm{Cov}(u_Z)_{\mathcal{C}}$ and that the two diagrams on the left are commutative. We therefore only have to consider the right-hand square. Let $\Gamma\backslash X^+$ be the irreducible component of $Sh_K(G,X)$ containing Z, where Γ is of the form $\Gamma = G(\mathbb{Q})_+ \cap gKg^{-1}$. The local system $\mathcal{V}(\xi)$ over $\Gamma\backslash X^+$ is the quotient of the trivial bundle $X^+ \times V$ over X^+ under the action of Γ given by $\gamma(x,v) = (\gamma \cdot x, \xi(\gamma)\cdot v)$.

Take $\alpha \in \pi_1(S,z)$ and let $\gamma \in \Gamma$ be an element mapping to $i_*(\alpha) \in \pi_1(\Gamma\backslash X^+, z)$. Then $\gamma Y^+ = Y^+$, and the image of α in $\mathrm{Aut}(Y^+)$ is given by the action of γ. The given description of $\mathcal{V}(\xi)$ shows that $\rho_S(\alpha) = \xi(\gamma)$. In particular, $\gamma \in Z(G)\cdot M$. It now readily follows that $f \circ \rho_S(\alpha) \in M^{\mathrm{ad}}(\mathbb{R})^+$ (since the action of γ stabilizes Y^+), and that the right-hand square is commutative. □

2.5 As in the proof of the lemma, let $\Gamma\backslash X^+$ be the irreducible component of $Sh_K(G,X)$ containing Z. The choice of a Γ-stable lattice in V induces a $\mathbb{Z}$-structure on $\mathcal{V}(\xi)$, which enables us to apply Theorem I.2.4. The connected algebraic monodromy group H_ζ associated to the VHS $n^*(\mathcal{V}(\xi)_{|Z})$ is therefore a normal subgroup of $\xi(M)^{\mathrm{der}} = \xi(M^{\mathrm{der}})$. Since M is reductive, we can find a normal algebraic subgroup $H_2 \triangleleft M$ (defined over $\mathbb{Q}$) such that M is the almost direct product of $\xi^{-1}(H_\zeta)$ and H_2. In this way we obtain a decomposition

$$(M^{\mathrm{ad}}, Y) = (H_\zeta^{\mathrm{ad}}, Y_1) \times (H_2^{\mathrm{ad}}, Y_2)\,.$$

2.6 Proposition. *The image of $\mathcal{C}$ under the projection map $\mathrm{pr}_2\colon Y \to Y_2$ is a single point, say $y_2 \in Y_2$. We have $Z \subseteq S_{\eta K}(Y_1^+ \times \{y_2\})$ for some connected component $Y_1^+ \subseteq Y_1$ and a class $\eta K \in G(\mathbb{A}_f)/K$.*

Proof. It follows from the lemma that $\mathrm{Cov}(u_{Z^n})$ acts trivially on Y_2, hence the composition $\widetilde{Z^n} \twoheadrightarrow \mathcal{C} \subseteq Y \twoheadrightarrow Y_2$ factors through Z^n. Because Y_2 has a realization as a bounded domain in some $\mathbb{C}^N$, the map $Z^n \to Y_2$ is given by an N-tuple of bounded holomorphic functions. Since Z^n is a connected quasi-projective variety these must be constant functions hence the image of $\mathcal{C}$ is a single point.

The last assertion is an immediate consequence of the first. □

If Z contains a regular special point then $H_\zeta = \xi(M)^{\mathrm{der}}$, by the second statement of Theorem I.2.4. This means that $Y \cong Y_1$, Y_2 is a point and that $S_{\eta K}(Y_1^+ \times \{y_2\}) = S$ is a subvariety of Hodge type. In this case, the proposition does not give us any information. However, the very fact that $H_\zeta = \xi(M)^{\mathrm{der}}$ can be used to establish a second decomposition of (M^{ad}, Y).

Consider the group $\{m \in M(\mathbb{Q})_+ \mid m\mathcal{C} = \mathcal{C}\}$, and write $\mathcal{N}$ for its closure inside $M(\mathbb{R})$ for the analytic topology. By Cartan's theorem $\mathcal{N}$ and its connected component of the identity $\mathcal{N}^+$ are Lie subgroups of $M(\mathbb{R})$. Clearly they are contained in $M(\mathbb{R})_\mathcal{C} = \{m \in M(\mathbb{R}) \mid m\mathcal{C} = \mathcal{C}\}$.

2.7 Proposition. *Assume that Z contains a non-singular special point. Then there exists a normal, reductive algebraic subgroup $N_Z \triangleleft M$, defined over $\mathbb{Q}$, such that $\mathcal{N}^+ = N_Z(\mathbb{R})^+$.*

Proof. The center $Z(M)(\mathbb{R})$ of $M(\mathbb{R})$ acts trivially on Y, and $Z(M)(\mathbb{Q})$ is analytically dense in $Z(M)(\mathbb{R})$, so $Z(M)(\mathbb{R}) \subseteq \mathcal{N}$. Furthermore, Lemma 2.4 shows that

$$\mathrm{ad}^{-1}(\mathrm{Im}(f \circ \rho_S \circ j_* \circ n_*\colon \pi_1(Z^n, \zeta) \to M^{\mathrm{ad}}(\mathbb{Q})^+)) \subseteq \mathcal{N}\,,$$

and since $H_\zeta = \xi(M)^{\mathrm{der}}$ it follows that $\mathcal{N}$ is Zariski dense in $M_\mathbb{R}$.

Let $\varphi\colon M_\mathbb{R} \to \mathrm{GL}(W)$ be a finite-dimensional irreducible representation of $M_\mathbb{R}$ with $\mathrm{Ker}(\varphi) \subseteq Z(M)$ (which exists, since M is reductive). Let $W' \subseteq W$ be the largest fully reducible $\mathcal{N}^+$-submodule of W. Then W' is a $\mathcal{N}$-submodule, since $\mathcal{N}$ normalizes $\mathcal{N}^+$. But $\mathcal{N}$ is Zariski dense in $M_\mathbb{R}$, so $W' = W$. Therefore, $\mathcal{N}^+/(\mathcal{N}^+ \cap \mathrm{Ker}(\varphi))$ has a faithful, fully reducible representation. Since $\mathcal{N}^+ \cap \mathrm{Ker}(\varphi)$ is contained in the center of $\mathcal{N}^+$, this implies that $\mathcal{N}^+$ is analytically reductive, i.e., $\mathrm{Lie}(\mathcal{N}^+)$ is reductive.

Write $\mathfrak{n} = \mathrm{Lie}(\mathcal{N}^+)$, which can be decomposed as $\mathfrak{n} = \mathfrak{c} \oplus \mathfrak{n}^{\mathrm{der}}$, where $\mathfrak{c}$ is the center and $\mathfrak{n}^{\mathrm{der}}$ is the derived algebra. From [9, Chapitre II, Théorème 15] (alternatively, [5, Chapter II, Corollary 7.9]) we know that $\mathfrak{n}^{\mathrm{der}}$ is algebraic, so $\mathfrak{n}^{\mathrm{alg}} = \mathfrak{c}^{\mathrm{alg}} \oplus \mathfrak{n}^{\mathrm{der}}$.

Let $\mathcal{N}^{+,\mathrm{alg}} \subseteq M_\mathbb{R}$ be the algebraic envelop of $\mathcal{N}^+$, and let $\mathfrak{N}$ be the normalizer of $\mathcal{N}^{+,\mathrm{alg}}$ inside $M_\mathbb{R}$. Clearly, $\mathcal{N} \subseteq \mathfrak{N}(\mathbb{R})$. On the other hand, $\mathfrak{N}$ is an algebraic subgroup of $M_\mathbb{R}$ and $\mathcal{N}$ is Zariski dense, so $\mathfrak{N} = M_\mathbb{R}$ and $\mathcal{N}^{+,\mathrm{alg}}$ is a normal subgroup. This implies that $\mathfrak{n}^{\mathrm{alg}} = \mathrm{Lie}(\mathcal{N}^{+,\mathrm{alg}})$ is an ideal of $\mathrm{Lie}(M_\mathbb{R})$, hence $\mathfrak{c} \subseteq \mathfrak{c}^{\mathrm{alg}} \subseteq \mathrm{Lie}(Z(M)(\mathbb{R}))$. But, as remarked above, $Z(M)(\mathbb{R}) \subseteq \mathcal{N}$, so $\mathfrak{c} = \mathfrak{c}^{\mathrm{alg}} = \mathrm{Lie}(Z(M)(\mathbb{R}))$. We conclude that $\mathcal{N}^{+,\mathrm{alg}}$ is a normal, reductive algebraic subgroup of $M_\mathbb{R}$ and $\mathcal{N}^+$ is the connected component of the identity of the Lie group $\mathcal{N}^{+,\mathrm{alg}}(\mathbb{R})$.

Since $\mathcal{N}^+$ is open in $\mathcal{N}$ and $\{m \in M(\mathbb{Q})_+ \mid m\mathcal{C} = \mathcal{C}\}$ is analytically dense in $\mathcal{N}$, the group $\{m \in M(\mathbb{Q})_+ \cap \mathcal{N}^+ \mid m\mathcal{C} = \mathcal{C}\}$ is dense in $\mathcal{N}^+$ for the analytic topology and hence is Zariski dense in $\mathcal{N}^{+,\mathrm{alg}}$. Since it consists of $\mathbb{Q}$-valued points of M we conclude that $\mathcal{N}^{+,\mathrm{alg}}$ is defined over $\mathbb{Q}$, which proves the proposition. □

In the next section we need the following, similar, statement.

2.8 Variant. *Consider the inclusion* $Z \subseteq S_{\eta K}(Y_1^+ \times \{y_2\})$ *as in Proposition 2.6, corresponding to the decomposition* $(M^{\mathrm{ad}}, Y) = (H_\zeta^{\mathrm{ad}}, Y_1) \times (H_2^{\mathrm{ad}}, Y_2)$ *and let* $\mathcal{C} \subseteq \widetilde{Z}$ *be an irreducible (analytic) component, as introduced after Lemma 2.3. There exists a normal algebraic subgroup* $H_{\zeta,\mathcal{C}} \triangleleft H_{\zeta,\mathbb{R}}$ *such that* $H_{\zeta,\mathcal{C}}(\mathbb{R})^+ = \{h \in H_\zeta(\mathbb{R}) \mid h\mathcal{C} = \mathcal{C}\}^+$.

Proof. The arguments are analogous to those in the previous proof, except that we leave out the last few lines. □

Assume that Z contains a regular special point. Choose a normal algebraic subgroup $N_2 \triangleleft M$ such that M is the almost direct product of N_Z and N_2. From this we obtain a decomposition $(M^{\mathrm{ad}}, Y) = (N_Z^{\mathrm{ad}}, Y_1') \times (N_2^{\mathrm{ad}}, Y_2')$. (We write Y_1' and Y_2' to avoid confusion with the decomposition $Y = Y_1 \times Y_2$ introduced before.)

From the remark that $N_Z^{\mathrm{ad}}(\mathbb{R})^+$ stabilizes $\mathcal{C}$ it easily follows that there exists a component $Y_1'^+$ and a class $\eta' K \in G(\mathbb{A}_f)/K$ such that $S_{\eta' K}(Y_1'^+ \times \{P\}) \subseteq Z$ for every point P in the image of the projection map $\mathcal{C} \to Y_2'$. Notice that this gives interesting information only if $Z(M)$ is a proper subgroup of N_Z.

§3 Totally geodesic subvarieties

Building on the results of the previous sections, we can now establish one of the main results of this chapter. This result was suggested to us by D. Kazhdan.

3.1 Theorem. *Let* (G, X) *be a Shimura datum, and let* K *be a compact open subgroup of* $G(\mathbb{A}_f)$. *An algebraic subvariety* $Z \hookrightarrow Sh_K(G, X)$ *is totally geodesic if and only if there exists a closed immersion of Shimura data* $i\colon (M, Y) \hookrightarrow (G, X)$, *a decomposition* $(M^{\mathrm{ad}}, Y) = (M_1, Y_1) \times (M_2, Y_2)$, *a component* Y_1^+, *a point* $y_2 \in Y_2$ *and a class* $\eta K \in G(\mathbb{A}_f)/K$ *such that* $Z = S_{\eta K}(Y_1^+ \times \{y_2\})$ *(as defined in Section 2.1).*

If Z *contains a special point, then* Z *is totally geodesic if and only if it is of Hodge type.*

Proof. It suffices to prove the theorem for K sufficiently small, so we may assume that the notations and results of Section 2 apply. Take a totally geodesic subvariety Z, and consider the inclusion $Z \subseteq S_{\eta K}(Y_1^+ \times \{y_2\})$ as in Proposition 2.6. By (ii) of Lemma 2.3 we may assume that $Y_1^+ \twoheadrightarrow S_{\eta K}(Y_1^+ \times \{y_2\})$ is a covering map. By Variant 2.8 there is a normal algebraic subgroup $H_{\zeta,\mathcal{C}} \triangleleft H_{\zeta,\mathbb{R}}$ such that

$$H_{\zeta,\mathcal{C}}(\mathbb{R})^+ = \{h \in H_\zeta(\mathbb{R}) \mid h\mathcal{C} = \mathcal{C}\}^+ . \tag{1}$$

If H' is a complement for $H_{\zeta,\mathcal{C}}$ in $H_{\zeta,\mathbb{R}}$ then we obtain a decomposition $Y_1^+ = W_1 \times W_2$, such that $H_{\zeta,\mathcal{C}}(\mathbb{R})^+$ acts transitively on W_1 and $H'(\mathbb{R})^+$ acts transitively on W_2.

By assumption, $\mathcal{C}$ is a complete, totally geodesic submanifold of Y_1^+. The group $H_{\zeta,\mathcal{C}}(\mathbb{R})^+$ therefore acts transitively on $\mathcal{C}$ (see Section I.4.4). It follows that

$$\mathcal{C} = W_1 \times \{w_2\} \tag{2}$$

for some $w_2 \in W_2$.

Let $H'_{w_2} \subseteq H'$ be the stabilizer subgroup of the point w_2, which is an algebraic subgroup of H'. Combining (1) and (2) we see that $H'_{w_2}(\mathbb{R})^+ = \{1\}$. On the other hand, H' is a semi-simple group over $\mathbb{R}$ and $H'_{w_2}(\mathbb{R})^+$ is a maximal compact subgroup of $H'(\mathbb{R})^+$. We conclude that $H' = \{1\}$, hence W_2 is reduced to the single point w_2. It follows that $\mathcal{C} = Y_1^+ \times \{y_2\}$, which proves the first statement of the theorem.

Next, suppose Z contains a special point. Since $S_{\eta K}(Y_1^+ \times \{y_2\})$ is non-singular for K sufficiently small we may assume this special point to be regular. As remarked after Proposition 2.6, this implies that $Z = S_{\eta K}(Y_1^+ \times \{y_2\})$ is of Hodge type. □

3.2 Corollary. *Let $Z \hookrightarrow Sh_K(G,X)$ be a totally geodesic subvariety, then there exists a subgroup $K' \subseteq K$ of finite index, algebraic varieties S_1, S_2, a closed immersion $g\colon S_1 \times S_2 \hookrightarrow Sh_{K'}(G,X)$ and points $a, b \in S_2$ such that*

1. *$S_1 \times \{s_2\}$ and $\{s_1\} \times S_2$ are totally geodesic subvarieties of $Sh_{K'}(G,X)$ for every $s_1 \in S_1$, $s_2 \in S_2$,*
2. *Z is the image of $S_1 \times \{a\}$ under $Sh_{(K',K)}$,*
3. *$S_1 \times \{b\}$, hence also $Sh_{(K',K)}(S_1 \times \{b\})$ is a subvariety of Hodge type.*

Proof. The map $\varphi\colon Y_1^+ \times Y_2^+ \to Sh_K(G,X)$ obtained by sending (y_1, y_2) to the class $[y_1, y_2, \eta K]$ factors through a finite morphism of algebraic varieties $\varphi'\colon \Gamma\backslash(Y_1^+ \times Y_2^+) \to Sh_K(G,X)$, where Γ is an arithmetic subgroup of $M^{\mathrm{ad}}(\mathbb{Q})$. There are arithmetic subgroups $\Gamma_1 \subset H_z^{\mathrm{ad}}(\mathbb{Q})$ and $\Gamma_2 \subseteq H_2^{\mathrm{ad}}(\mathbb{Q})$ such that $\Gamma_1 \times \Gamma_2$ is of finite index in Γ ([4, Corollaire 8.10]). Taking $S_i = \Gamma_i \backslash Y_i^+$ we arrive at the corollary. □

3.3 To conclude this section, let us discuss an example. The example concerns a subvariety $S \hookrightarrow \mathsf{A}_{4d,1,n}$ $(d \geq 2)$ of Shimura type, such that for a generic point $\eta \in S$, the abelian variety Y_η is simple, whereas the generic Mumford-Tate group G

on S has a non-simple adjoint group. This then leads to non-trivial totally geodesic subvarieties Z which are not of Hodge type, and for which the connected algebraic monodromy group H_z is a proper subgroup of G^{der}. We also obtain a negative answer to two problems formulated in [1, Chapter X].

A sketch of a special case of the example can be found in [2], where a reference is given to Borovoi's paper [6]. However, we have not had the opportunity to read (a translation of) Borovoi's paper, so it is not clear to us to whom the example is due.

Let F be a totally real field of degree $d \geq 2$ over $\mathbb{Q}$, and write $\infty_1, \ldots, \infty_d$ for its places at infinity. Take two quaternion algebras D_1, D_2 which both have at least one invariant 0 at infinity and which moreover have "complementary" invariants at infinity, i.e., $\mathrm{inv}_{\infty_i}(D_1) = 0$ if and only if $\mathrm{inv}_{\infty_i}(D_2) = 1/2$. Then $D_1 \otimes_F D_2 \cong \mathrm{M}_2(D)$ for some other quaternion algebra D over F (using that $\mathrm{inv}_v(D_1 \otimes D_2) = \mathrm{inv}_v(D_1) + \mathrm{inv}_v(D_2)$ in $\mathbb{Q}/\mathbb{Z}$, and the fact that D_1 and D_2 have different invariants at infinity). Let $G_1 = \mathrm{Res}_{F/\mathbb{Q}} D_1^*, G_2 = \mathrm{Res}_{F/\mathbb{Q}} D_2^*$, let $V = D \oplus D$ as a $\mathbb{Q}$-vector space, and define the homomorphism

$$f\colon G_1 \times G_2 \to \mathrm{GL}(V)$$

as the composition of $G_1 \times G_2 \hookrightarrow \mathrm{Res}_{F/\mathbb{Q}}(\mathrm{GL}_2(D))$ and the natural map

$$\mathrm{Res}_{F/\mathbb{Q}}(\mathrm{GL}_2(D)) \to \mathrm{GL}(\mathrm{Res}_{F/\mathbb{Q}}(D \oplus D)) = \mathrm{GL}(V)\,.$$

Let X_1 be the $G_1(\mathbb{R})$-conjugacy class in $\mathrm{Hom}(\mathbb{S}, G_{1,\mathbb{R}})$ of the homomorphism h_1 given on $\mathbb{R}$-valued points by

$$\mathbb{C}^* \ni\ a + bi \mapsto \Big(\prod_{i \in I_1} \mathrm{Id}\ \times \prod_{i \in I_2} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \Big)\ \in \prod_{i \in I_1} \mathbb{H}^* \times \prod_{i \in I_2} \mathrm{GL}_2(\mathbb{R})\ \cong\ G_1(\mathbb{R})\,,$$

where $I_1 = \{i \mid \mathrm{inv}_{\infty_i}(D_1) = 1/2\}$, $I_2 = \{i \mid \mathrm{inv}_{\infty_i}(D_1) = 0\}$. Notice that X_1 is well-defined since all automorphisms of $\mathrm{GL}_2(\mathbb{R})$ are inner. Likewise we get a $G_2(\mathbb{R})$-conjugacy class X_2 in $\mathrm{Hom}(\mathbb{S}, G_{2,\mathbb{R}})$. One easily checks that (G_i, X_i) is a Shimura datum, i.e., a pair satisfying conditions (2.1.1.1-3) of [18, Section 2.1]. In this way we get Shimura varieties $Sh(G_1, X_1)$ and $Sh(G_2, X_2)$. Notice that these are not of Hodge type, since their weight is not defined over $\mathbb{Q}$.

Let G be the image of $G_1 \times G_2$ under f and consider the $G(\mathbb{R})$-conjugacy class X in $\mathrm{Hom}(\mathbb{S}, G_\mathbb{R})$ which is the image of $X_1 \times X_2$ under the natural map $\mathrm{Hom}(\mathbb{S}, G_{1,\mathbb{R}}) \times \mathrm{Hom}(\mathbb{S}, G_{2,\mathbb{R}}) \to \mathrm{Hom}(\mathbb{S}, G_\mathbb{R})$. Notice that $X_1 \times X_2 \cong X$. Again one easily checks that the pair (G, X) thus obtained is a Shimura datum. For a compact open subgroup

K of $G(\mathbb{A}_f)$ let $K_i = f_i^{-1}(K)$, where f_i is the restriction of f to G_i. Then K_i is a compact open subgroup of $G_i(\mathbb{A}_f)$ and we get a morphism

$$f_{(K_1 \times K_2, K)}\colon Sh_{K_1 \times K_2}(G_1 \times G_2, X_1 \times X_2) \longrightarrow Sh_K(G, X)\,.$$

We choose connected components X_1^+, X_2^+ and denote the corresponding connected Shimura varieties by $Sh^0(-)$. For K_1, K_2 sufficiently small the map $f^0_{(K_1 \times K_2, K)}$ on connected Shimura varieties is finite étale.

3.4 Lemma. *The weight homomorphism* $w\colon \mathbb{G}_{\mathrm{m}} \to G_{\mathbb{R}}$ *is defined over* $\mathbb{Q}$.

Proof. Choose subfields $L_i \subset D_i$ of degree $2d$ over $\mathbb{Q}$, and let $T_i \subset G_i$ be the associated maximal torus (so $T_i = \mathrm{Res}_{L_i/\mathbb{Q}} \mathbb{G}_{\mathrm{m}, L_i}$). Let $\mathrm{Hom}(L_1, \mathbb{C}) = \{\sigma_{11}, \sigma_{12}, \ldots, \sigma_{d1}, \sigma_{d2}\}$ and $\mathrm{Hom}(L_2, \mathbb{C}) = \{\tau_{11}, \tau_{12}, \ldots, \tau_{d1}, \tau_{d2}\}$, where σ_{ij} and τ_{ij} extend ∞_i. The character group $X^*(T_1)$ is canonically isomorphic to the free abelian group on $\mathrm{Hom}(L_1, \mathbb{C})$, so it has a standard basis corresponding to the elements σ_{ij}. Let $e_{11}, e_{12}, \ldots, e_{d1}, e_{d2}$ be the dual basis for the cocharacter group $X_*(T_1)$. Likewise we get a standard basis $f_{11}, f_{12}, \ldots, f_{d1}, f_{d2}$ for $X_*(T_2)$.

There is an exact sequence

$$1 \to \Delta^- \to G_1 \times G_2 \to G \to 1\,,$$

where Δ^- is (the Weil restriction of) the antidiagonal $\{(f, f^{-1}) \in D_1^* \times D_2^* \mid f \in F^*\}$. Then $T = f(T_1 \times T_2)$ is a maximal torus of G and $X_*(T)$ is the quotient of $X_*(T_1) \times X_*(T_2)$ by the subgroup $X_*(\Delta^-)$ generated by the elements of the form $e_{i1} + e_{i2} - f_{i1} - f_{i2}$ ($i \in \{1, \ldots, d\}$).

In these notations the weight cocharacter $w\colon \mathbb{G}_{\mathrm{m}} \to G_{\mathbb{R}}$ is the element

$$w = \sum_{i=1}^{d} \alpha_i \cdot (e_{i1} + e_{i2}) + (1 - \alpha_i) \cdot (f_{i1} + f_{i2}) \pmod{X_*(\Delta^-)}\,,$$

where $\alpha_i = 0$ if $\mathrm{inv}_{\infty_i}(D_1) = 1/2$ and $\alpha_i = 1$ if $\mathrm{inv}_{\infty_i}(D_1) = 0$. Since $e_{i1} + e_{i2} = f_{i1} + f_{i2}$ in $X_*(T)$ we can rewrite this as

$$w = \sum_{i=1}^{d} e_{i1} + e_{i2} = \sum_{i=1}^{d} f_{i1} + f_{i2}$$

and from this it is clear that w is a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-invariant element. □

By considering the representation of $G_{\mathbb{R}}$ on $V_{\mathbb{R}}$ one sees that there does not exist a symplectic form Ψ on V such that G acts through symplectic similitudes. Essentially the problem is that the center of G is "too large". Therefore, we introduce the algebraic subgroup $G' = w(\mathbb{G}_{\mathrm{m}}) \cdot G^{\mathrm{der}} \subseteq G$, which, by the lemma, is defined over $\mathbb{Q}$. All $h_x\colon \mathbb{S} \to G_{\mathbb{R}}$ for $x \in X$ factor through $G'_{\mathbb{R}}$, and we have a closed immersion of Shimura data $(G', X) \hookrightarrow (G, X)$.

The lemma shows that (G', X) satisfies condition (2.1.1.4) of [18, Section 2.1]. It also satisfies loc. cit., condition (2.1.1.5), as one easily verifies. Furthermore, for $x \in X$ the representation h_x on V is of type $(-1,0)+(0,-1)$. From [18, Proposition 2.3.2] it now follows that there exists a symplectic form Ψ on V such that the inclusion $G' \hookrightarrow \mathrm{GL}(V)$ induces a morphism of Shimura data $i\colon (G', X) \hookrightarrow (\mathrm{CSp}(V, \Psi), \mathfrak{H}^{\pm}_{4d})$. Here we identify the Siegel double space $\mathfrak{H}^{\pm}_{4d}$ as the space of $\mathbb{R}$-Hodge structures $h\colon \mathbb{S} \to \mathrm{GL}(V)$ of type $(-1,0)+(0,-1)$ such that $\pm\Psi$ is a polarization. This shows that $Sh(G', X)$ is a Shimura variety of Hodge type.

For a compact open subgroup $K \subseteq G(\mathbb{A}_f)$, write $K' = G'(\mathbb{A}_f) \cap K$. For K sufficiently small we get a "universal" family $\alpha\colon (Y, \lambda, \theta) \to Sh^0_{K'}(G', X)$ of $4d$-dimensional principally polarized abelian varieties with a level K'-structure. The morphism $f^0_{(K_1 \times K_2, K)}$ on connected Shimura varieties factors through $Sh^0_{K'}(G', X)$. We can choose a point $x_2 \in X_2^+$ such that the subvariety

$$Z_{x_2} = Sh^0(f)(Sh^0_{K_1}(G_1, X_1) \times [x_2, eK_2]) = S_{eK'}(X_1^+ \times \{x_2\}) \subset Sh^0_{K'}(G', X)$$

is not contained in the locus Σ (applying the discussion of Section I.2.2 to $S = Sh^0_{K'}(G', X)$ and the natural VHS with local system $\mathrm{R}^1\alpha_*\mathbb{Z}_Y$). The generic fibre Y_η on $Z = Z_{x_2}$ has Mumford-Tate group G', and the representation of $\mathrm{MT}(Y_\eta)$ on $\mathrm{H}^1(Y_\eta(\mathbb{C}), \mathbb{Q})$ is isomorphic to $G' \hookrightarrow \mathrm{GL}(V)$. In particular, V being an irreducible G'-module, Y_η is simple. On the other hand, it is clear that the connected algebraic monodromy group of the restricted family (Y, λ, θ) over Z is contained in $f(G_2)^{\mathrm{der}}$, so it is strictly contained in $(G')^{\mathrm{der}} = G^{\mathrm{der}}$.

As remarked by André ([2], footnote on p. 13) the example contradicts the conjectural statement IX, 3.1.6 in [1]. We claim that it also gives a negative answer to op. cit., Chapter X, Problems 2 and 3.

Loc. cit., Problem 2 is essentially the following. Consider a subvariety $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ satisfying (i) $\dim(Z) = 1$, (ii) the generic fibre in the family of abelian varieties over Z is simple, (iii) there are infinitely many points on Z which lie on a proper subvariety of Hodge type. Does it follow that Z is of Hodge type? We see that the answer is negative in general: in the above example we choose D_1 and D_2 such that $\#I_2 = 1$, which implies that $\dim(Z_{x_2}) = 1$. As we have seen, Z_{x_2} satisfies condition (ii) and it is not of Hodge type. Finally, for all special points $x_1 \in X_1^+$, the point $Sh^0(f)([x_1, eK_1] \times [x_2, eK_2]) \in Z_{x_2}$ lies on a proper subvariety of Hodge type.

A special case of loc. cit., Problem 3, is the following question. Consider a subvariety $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ satisfying conditions (i) and (ii) and also satisfying (iv) there are infinitely many points on Z such that the corresponding abelian varieties are all isogenous. Does it follow that Z is of Hodge type? Again, the answer is negative. The example is the same as above; for (iv) we only have to remark that for a fixed $x_1 \in X_1^+$, the fibres over the points $Sh^0(f)([g_1 \cdot x_1, eK_1] \times [x_2, eK_2]) \in Z_{x_2}$ with $g_1 \in G_1(\mathbb{Q})$ are all isogenous.

§4 Non-rigid families of abelian varieties

4.1 In his paper [3], Arakelov proved that, given a complete and non-singular curve B over $\mathbb{C}$, a finite set of points $S \subseteq B$ and an integer $g \geq 2$, the set

$$\left\{ \begin{array}{l} \text{isomorphism classes of non-constant families of} \\ \text{non-singular irreducible curves of genus } g \text{ over } B \setminus S \end{array} \right\}$$

is finite. (For $S = \emptyset$ this was done by Paršin.) One of the main steps in the proof is to show that if X is such a non-constant family of non-singular curves of genus g over $B \setminus S$, then X does not have non-trivial deformations.

In [24], Faltings then gave an example showing that the analogous statement for abelian varieties is false in general. His example concerns a non-rigid family of abelian eightfolds. After that, several people came up with related results. As for the non-rigid families of abelian varieties, Masahiko Saito obtained in [54] a classification of the endomorphism algebras of the underlying local systems. In particular, he determined for which (relative) dimensions there exist such non-rigid families (without isotrivial factors).

Using the notations and results discussed before, we can add to Saito's work. We describe the non-rigid families of abelian varieties in terms of the corresponding subvarieties of the moduli space and we "explain" the non-rigidity geometrically. Before we do so, let us first discuss the problem in a rather general setting.

4.2 Suppose we are given a polarized $\mathbb{Z}$-VHS $\mathcal{V} = (\mathcal{V}_{\mathbb{Z}}, \mathcal{F}^{\bullet}, \mathcal{Q})$ of pure weight n over a non-singular, irreducible, complex algebraic variety Z. We can ask if there are non-trivial deformations of $\mathcal{V}$, fixing the base space Z. We are in fact most interested in the infinitesimal deformations. To give this a precise meaning one has to set up some theory.

Part of the structure of $\mathcal{V}$ is discrete, and therefore cannot vary continuously. Specifically, let $\mathcal{T}$ be the set of equivalence classes of 4-tuples $(V_{\mathbb{Z}}, Q, \pi, \rho)$, where $V_{\mathbb{Z}}$ is a free $\mathbb{Z}$-module of finite rank, $Q\colon V_{\mathbb{Z}} \times V_{\mathbb{Z}} \to \mathbb{Z}$ is a bilinear form, π is a group and $\rho\colon \pi \to \mathrm{U}(V_{\mathbb{Z}}, Q)$ is a homomorphism. Two such 4-tuples $(V_{\mathbb{Z}}, Q, \pi, \rho)$ and $(V'_{\mathbb{Z}}, Q', \pi', \rho')$ are said to be equivalent if there exist isomorphisms $\alpha\colon (V_{\mathbb{Z}}, Q) \xrightarrow{\sim} (V'_{\mathbb{Z}}, Q')$ and $\beta\colon \pi \xrightarrow{\sim} \pi'$ such that $\alpha_* \circ \rho = \rho' \circ \beta$. For $z \in Z$ we get such a 4-tuple by taking $(V_{\mathbb{Z}}, Q) = (\mathcal{V}_z, Q_z)$, $\pi = \pi_1(Z, z)$ and $\rho\colon \pi_1(Z, z) \to \mathrm{U}(V, Q)$ the monodromy representation. The class of this 4-tuple in $\mathcal{T}$ does not depend on the choice of z, and therefore we get a well-determined element $\tau(\mathcal{V}) \in \mathcal{T}$ associated to $\mathcal{V}$.

In Peters' paper [51] it is shown that the set

$$\mathbb{Z}\text{-}\mathsf{VHS}_{\tau}(Z) = \left\{ \begin{array}{l} \text{isomorphism classes of polarized } \mathbb{Z}\text{-VHS} \\ \mathcal{V} \text{ over } Z \text{ with } \tau(\mathcal{V}) = \tau \in \mathcal{T} \end{array} \right\}$$

has a natural structure of an analytic variety such that the tangent space to the class $[\mathcal{V}]$ is isomorphic to $(E^Q \otimes_{\mathbb{Q}} \mathbb{C})^{-1,1}$. Here $E = \mathrm{H}^0(Z, \mathcal{E}nd(\mathcal{V}_{\mathbb{Q}}))$, the algebra of global (flat) endomorphisms of the local system $\mathcal{V}_{\mathbb{Q}}$, and

$$E^Q = \{e \in E \mid Q(ev, w) + Q(v, ew) = 0 \quad \text{for all sections } v, w \text{ of } \mathcal{V}_{\mathbb{Q}}\}$$

is the subspace of elements of E which are skew-symmetric with respect to Q. We conclude that the polarized $\mathbb{Z}$-VHS $\mathcal{V}$ over Z is rigid (i.e., it has no infinitesimal deformations over Z) if and only if $(E^Q \otimes_{\mathbb{Q}} \mathbb{C})^{-1,1} = 0$. The problem that we are interested in is to describe, or to classify, the polarizable $\mathbb{Z}$-VHS $\mathcal{V}$ over Z such that $\mathcal{V}$ is non-rigid over some finite covering of Z. This is still somewhat vague. Moreover, there are trivial ways of constructing non-rigid variations of Hodge structure (as we shall discuss in a moment), and part of the problem is to distinguish these from the interesting, non-trivial cases.

4.3 Let $z \in Z$ be a Hodge-generic point, and write $M = \mathrm{MT}_z$, $\mathrm{Hg} = \mathrm{Hg}_z$ for the Hodge group, and $(V, Q) = (\mathcal{V}_{\mathbb{Q},z}, Q_z)$. Since we allow finite coverings of Z we may assume the algebraic monodromy group to be connected. Then we have algebraic groups

$$H_z \lhd M^{\mathrm{der}} = \mathrm{Hg}^{\mathrm{der}} \subseteq \mathrm{U}(V, Q)\,,$$

and E^Q is just the space of H_z-invariants in $\mathrm{End}^Q(V)$.

As in [54, §3] we can reduce the problem to the case that V is isotypical as a H_z-module, in which case the VHS $\mathcal{V}_{\mathbb{Q}}$ is called primary. To see this, let $V = V^{(1)} \oplus \cdots \oplus V^{(k)}$ be the canonical decomposition of V into H_z-isotypical summands, where $V^{(i)}$ is isomorphic to a power of an irreducible H_z-module W_i. Since $H_z \lhd M$ and M acts (by conjugation) on H_z through inner automorphisms, the $V^{(i)}$ are M-submodules of V. The form Q restricts to a polarization form Q_i on $V^{(i)}$. The H_z-module $\mathrm{End}(V)$ is a direct sum of the $\mathrm{End}(V^{(i)})$ and of spaces of the form $\mathrm{Hom}(W_i, W_j)$ with $i \neq j$. By Schur's lemma there are no H_z-invariants in $\mathrm{Hom}(W_i, W_j)$, hence

$$E^Q = (\mathrm{End}^Q(V))^{H_z} = (\mathrm{End}^{Q_1}(V^{(1)}))^{H_z} \oplus \cdots \oplus (\mathrm{End}^{Q_k}(V^{(k)}))^{H_z}\,,$$

and we conclude that $\mathcal{V}$ is non-rigid if and only if one of the $\mathcal{V}^{(i)}$ (the sub-VHS corresponding to $V^{(i)}$) is non-rigid.

As mentioned before there are trivial ways to produce non-rigid VHS. If, for example, the VHS $\mathcal{V}_{\mathbb{Q}}$ splits as $\mathcal{V}_{\mathbb{Q}} = \mathcal{V}_1 \oplus \mathcal{V}_2$ such that the monodromy on one of the two summands, say $\mathcal{V}_1$, is trivial, then we can deform $\mathcal{V}$ by varying the constant VHS $\mathcal{V}_1$. To exclude this we simply require that the monodromy representation on the irreducible components of $\mathcal{V}$ is non-trivial. The problem therefore is to describe the non-rigid VHS over Z for which the underlying local system is primary and non-constant.

4.4 From now on we assume the VHS $\mathcal{V}$ to be of type $(-1,0)+(0,-1)$, which means that it corresponds to a family of abelian varieties over Z. We will make free use of the correspondence between polarizable $\mathbb{Z}$-VHS of type $(-1,0)+(0,-1)$ and families of abelian varieties; in particular, we say that a family of abelian varieties is non-rigid if the corresponding VHS is non-rigid.

One of the main advantages of restricting our attention to abelian varieties is that in this case the endomorphism algebra $E = \mathrm{End}(V)^{H_z}$ is of type $(-1,1)+(0,0)+(1,-1)$. Therefore, the family is rigid if and only if E^Q is purely of type $(0,0)$, which is equivalent to saying that H_z and Hg have the same invariants in $\mathrm{End}^Q(V)$. Notice that the polarization form Q induces an isomorphism $V \xrightarrow{\sim} V^*$ as Hg-modules, from which we get an isomorphism $\mathrm{End}^Q(V) \xrightarrow{\sim} \mathrm{Sym}^2(V)$ as representations of Hg, so we can also say that the family is rigid if and only if H_z and Hg have the same invariants in $\mathrm{Sym}^2(V)$.

In Saito's paper [54] we find a classification of the endomorphism algebras E of the local systems underlying a non-rigid family of abelian varieties. As one of the corollaries one obtains restrictions on the relative dimensions for which there exist such non-rigid families. It is known that the smallest relative dimension for which there exist non-rigid families is 8; as mentioned before, an example for relative dimension 8 was given in [24].

Our purpose here is to elucidate the structure of the families themselves, in terms of subvarieties of the moduli space of abelian varieties.

4.5 Consider a family $f\colon X \to Z$ of abelian varieties over a normal variety Z, with a polarization $\lambda\colon X \to X^t$ (X^t denoting the dual abelian scheme). It corresponds to the polarized $\mathbb{Z}$-VHS $\mathcal{V}$ over Z. We assume that the algebraic monodromy group is connected, that the family is non-rigid, and that the local system $\mathcal{V}_\mathbb{Q}$ is primary and non-constant. We keep the notations $z \in Z$, H_z, M, Hg and (V, Q) as above.

Fix an integer $n \geq 3$ which is relatively prime with the degree of λ. Possibly after passing to a finite covering of Z we can choose a Jacobi level n structure on X over Z. The family (X, λ) plus the choice of this level structure corresponds to a morphism $\varphi_f\colon Z \to \mathsf{A}_{g,\delta,n}(\mathbb{C})$, where δ is the type of the polarization λ. The fact that the generic Mumford-Tate group is M means that φ_f maps Z into a subvariety of Hodge type $S_{\eta K}(Y_M^+)$ (with $K = K_n$ as in Section I.3.2).

Recall the decomposition of Shimura data $(M^{\mathrm{ad}}, Y_M) = (H_\zeta^{\mathrm{ad}}, Y_1) \times (H_2^{\mathrm{ad}}, Y_2)$ that was introduced in Section 2.5, where $H_\zeta \lhd M$ is the connected algebraic monodromy group of the family $X \to Z$. It follows from Proposition 2.6 that $\varphi_f(Z) \subseteq S_{\eta K}(Y_1^+ \times \{y_2\})$ for some component $Y_1^+ \subseteq Y_1$ and a point $y_2 \in Y_2$. Let $\mathcal{W}$ be the $\mathbb{Z}$-VHS corresponding to the universal family over $S = S_{\eta K}(Y_1^+ \times \{y_2\})$, then we can identify the fibres $\mathcal{V}_z$ and $\mathcal{W}_{\varphi_f(z)}$, and the monodromy representation of $\mathcal{V}$ factors through that of $\mathcal{W}$. Pulling back by φ_f defines a map $\varphi_f^*\colon \mathbb{Z}\text{-}\mathsf{VHS}_{\tau(\mathcal{W})}(S) \to \mathbb{Z}\text{-}\mathsf{VHS}_{\tau(\mathcal{V})}(Z)$ which induces an isomorphism on the tangent spaces at $[\mathcal{W}]$ and $[\mathcal{V}]$ respectively. Therefore, it remains to explain the non-rigidity of the family over S. In order to keep the notations as clear as possible, we assume from now on that $Z = S = S_{\eta K}(Y_1^+ \times \{y_2\})$ and $\mathcal{V} = \mathcal{W}$.

The basic idea now becomes apparent: Z is totally geodesic, and after Corollary 3.2 there is a product variety $S_1 \times S_2$ covering $S_{\eta K}(Y_1^+ \times Y_2^+)$, such that Z is the image of $S_1 \times \{a\}$ for some $a \in S_2$. If Y_2^+ (hence also S_2) is not reduced to a single point then we can vary the point $a \in S_2$, and this gives global deformations of the VHS over $S_1 \times \{a\}$. However, if Y_2^+ is a single point (which happens if Z is of Hodge type) then this idea does not seem to work. It does, but in general we first have to replace (M, Y_M) by a "larger" Shimura datum (N, Y_N). We do this as follows.

Consider the algebraic group $C = C_{\mathrm{Sp}(V,Q)}(H_z)$, the centralizer of H_z inside $\mathrm{Sp}(V,Q)$. Its connected component C^0 is a reductive subgroup of $\mathrm{Sp}(V,Q)$. Notice that $E^Q = \mathrm{Lie}(C)$. Similarly, $(E^Q)^{0,0}$ is isomorphic to the Lie algebra of $C_{\mathrm{Sp}(V,Q)}(\mathrm{Hg})$, so the fact that E^Q is not purely of type $(0,0)$ is equivalent to saying that $C_{\mathrm{Sp}(V,Q)}(\mathrm{Hg})^0 \subsetneq C^0 = C_{\mathrm{Sp}(V,Q)}(H_z)^0$.

The reductive group C^0 is the almost direct product of its center Z_{C^0} and a number of $\mathbb{Q}$-simple semi-simple factors C_i^0. Let C_c be the product of the factors C_i^0 for which $C_i^0(\mathbb{R})$ is compact, and let C' be the product of Z_C^0 and the factors C_i^0 which are not compact over $\mathbb{R}$. The intersection $H_z \cap C^0$ is contained in the center of H_z and is therefore finite, so $H_z \cdot C^0 \subseteq \mathrm{Sp}(V,Q)$ is the almost direct product of H_z and C^0. Clearly, $M \subseteq \mathbb{G}_{\mathrm{m}} \cdot H_z \cdot C^0 \subseteq \mathrm{CSp}(V,Q)$ (using that H_z is normal in M), and it follows from [18, 1.1.15] that $\mathrm{Inn}(h(i))$ is a Cartan involution of $H_z \cdot C^0$ (where $h\colon \mathbb{S} \to M_{\mathbb{R}} \subseteq \mathrm{CSp}(V,Q)_{\mathbb{R}}$ is the homomorphism giving the Hodge structure on $V = \mathcal{V}_{\mathbb{Q},z}$). The composite map

$$\mathbb{S} \to (\mathbb{G}_{\mathrm{m}} \cdot H_z \cdot C^0)_{\mathbb{R}} \xrightarrow{\mathrm{ad}} H_z^{\mathrm{ad}} \times C_c^{\mathrm{ad}} \times (C')^{\mathrm{ad}} \xrightarrow{\mathrm{pr}} C_c^{\mathrm{ad}}$$

must therefore be trivial, hence h factors through $\mathbb{G}_{\mathrm{m}} \cdot H_z \cdot C'$.

Let $N = \mathbb{G}_{\mathrm{m}} \cdot H_z \cdot C'$, and let Y_N be the $N(\mathbb{R})$-conjugacy class of h. The above arguments show that (N, Y_N) is a Shimura datum. We have a decomposition $(N^{\mathrm{ad}}, Y_N) = (H_z^{\mathrm{ad}}, Y_1) \times (C'^{\,\mathrm{ad}}, Y_{C'})$, and $S = S_{\eta K}(Y_1^+ \times \{y_2\})$, which we can also write as $S_{\eta K}(Y_1^+ \times \{v\})$ for a point $v \in Y_{C'}$. Notice that $Y_{C'}$ is not reduced to a single point, or, equivalently: C' has non-trivial semi-simple factors. This follows from the fact that $\mathrm{Lie}(C)$ is not purely of type $(0,0)$. Alternatively: C' being a torus would contradict the above remark that $C_{\mathrm{Sp}(V,Q)}(\mathrm{Hg})^0 \subsetneq C^0$, since $\mathrm{Hg} \subseteq H_z \cdot C'$.

This brings us to a situation where we can apply Corollary 3.2. We have a commutative diagram

$$\begin{array}{ccc} S_1 \times \{a\} & \hookrightarrow & S_1 \times S_2 \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle g'} \\ S = S_{\eta K}(Y_1^+ \times \{v\}) & \hookrightarrow & S_{\eta K}(Y_1^+ \times Y_{C'}^+) \end{array}$$

where g and g' are finite surjective morphisms, and where $S_1 = Y_1^+/\Gamma_1$, $S_2 = Y_{C'}^+/\Gamma_2$ for some arithmetic subgroups $\Gamma_1 \subset H_z^{\mathrm{ad}}(\mathbb{Q})$ and $\Gamma_2 \subset C'^{\mathrm{ad}}(\mathbb{Q})$. We may take Γ_1 and Γ_2 small enough such that S_1 and S_2 are non-singular.

The morphism $g^*\colon \mathbb{Z}\text{-}\mathsf{VHS}_{\tau(\mathcal{W})}(S) \to \mathbb{Z}\text{-}\mathsf{VHS}_{\tau(g^*\mathcal{W})}(S_1)$ induces an isomorphism on tangent spaces at $\mathcal{W}$ and $g^*\mathcal{W}$ respectively; this follows from the description of the tangent space given above and the fact that both the generic Mumford-Tate group and the connected algebraic monodromy group on S and S_1 are the same.

Varying the point $a \in S_2$ then gives global deformations of the VHS $g^*\mathcal{W}$ over S_1. We remark that this indeed "explains all deformations": we have seen that the tangent space to $\mathbb{Z}\text{-}\mathsf{VHS}_\tau(S)$ at the point $[\mathcal{V}]$ is isomorphic to $\mathrm{Lie}(C)^{-1,1}$, and since $\mathrm{Lie}(C_c)$ is purely of type $(0,0)$ this is equal to $\mathrm{Lie}(C')^{-1,1}$. Our remark then follows from the fact that there are natural isomorphism $T_aS_2 \cong T_vY_{C'} \xrightarrow{\sim} \mathrm{Lie}(C')^{-1,1}$ (where $Y_{C'} \ni v \mapsto a \in S_2$), which is an infinitesimal version of the correspondence

$$\text{varying the point } v \in Y_{C'} \quad \rightsquigarrow \quad \text{deformations of the VHS } g^*\mathcal{V} \text{ over } S_1\,,$$

see [51, Sections 1 and 2] (notice that in our case the horizontal tangent bundle to the period domain $Y_{C'}$ is equal to the full tangent bundle).

In summary, we obtain the following result.

4.6 Theorem. *Let $f\colon X \to Z$ be a principally polarized abelian scheme over a normal irreducible complex algebraic variety Z, and assume that X admits a Jacobi level n structure θ for some $n \geq 3$. Let $\varphi_f\colon Z \to \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ be the corresponding morphism of Z into the moduli space. Then there exists a closed immersion of Shimura data $i\colon (N, Y_N) \hookrightarrow (\mathrm{CSp}_{2g}, \mathfrak{H}_g^\pm)$, a decomposition $(N^{\mathrm{ad}}, Y_N) = (N_1, Y_1) \times (N_2^{\mathrm{ad}}, Y_2)$ and a diagram*

$$\begin{array}{ccccccc} & & S_1 \times \{a\} & \hookrightarrow & S_1 \times S_2 & \hookrightarrow & \mathsf{A}_{g,1,K'} \otimes \mathbb{C} \\ & & {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle g'} & & \downarrow \\ Z & \xrightarrow{\varphi_f} & S = S_{\eta K}(Y_1^+ \times \{v\}) & \hookrightarrow & S_{\eta K}(Y_1^+ \times Y_{C'}^+) & \hookrightarrow & \mathsf{A}_{g,1,n} \otimes \mathbb{C} \end{array}$$

such that there are natural isomorphisms

$$T_vY_2 \cong T_aS_2 \xrightarrow{\sim} T_{[g^*\mathcal{W}]}(\mathbb{Z}\text{-}\mathsf{VHS}(S_1)) \xleftarrow{\sim} T_{[\mathcal{W}]}(\mathbb{Z}\text{-}\mathsf{VHS}(S)) \xrightarrow{\sim} T_{[\mathcal{V}]}(\mathbb{Z}\text{-}\mathsf{VHS}(Z))\,.$$

Here we write $\mathcal{V}$, $\mathcal{W}$ and $\mathcal{U}$ for the VHS corresponding to the first homology of the abelian schemes over Z, S and $S_1 \times S_2$ respectively, and the map $T_aS_2 \xrightarrow{\sim} T_{[g^*\mathcal{W}]}(\mathbb{Z}\text{-}\mathsf{VHS}_{\tau(g^*\mathcal{W})}(S_1))$ is the map on tangent spaces induced by the map $S_2 \to \mathbb{Z}\text{-}\mathsf{VHS}_{\tau(g^*\mathcal{W})}(S_1)$ sending $s_2 \in S_2$ to the class of $\mathcal{U}_{|S_1\times\{s_2\}}$ in $\mathbb{Z}\text{-}\mathsf{VHS}_{\tau(g^*\mathcal{W})}(S_1)$.

4.7 Corollary. *Let X be a simple abelian variety over $\mathbb{C}$ with $\dim(X) \leq 7$. Then $\mathrm{MT}(X)^{\mathrm{ad}}$ is either trivial or it is a $\mathbb{Q}$-simple algebraic group.*

Proof. This now follows immediately from the fact that there are no (non-trivial) non-rigid families of abelian varieties of relative dimension ≤ 7, see [54, Corollary 8.4]. □

Chapter III

Formal linearity, special points and Shimura varieties

§1 Local moduli of abelian varieties

1.1 Fix an integer $n \geq 3$ and a prime number p with $p \nmid n$. We also fix a finite field k of characteristic p. Write $W = W(k)$ for its ring of (infinite) Witt vectors, and write $\mathcal{A}_g$ for $\mathsf{A}_{g,1,n} \otimes W$. Let $(\mathcal{A}_g \otimes k)^{\mathrm{o}}$ be the ordinary locus in characteristic p. This is a locally closed subscheme of $\mathcal{A}_g$, hence we can take the formal completion along it to obtain a formal scheme $\widehat{\mathcal{A}}_g = \mathcal{A}_{g\,/(\mathcal{A}_g \otimes k)^{\mathrm{o}}}$ over $\mathrm{Spf}(W)$.

Let $U \subset \mathcal{A}_g$ be an open subscheme such that the ordinary locus $U \cap (\mathcal{A}_g \otimes k)^{\mathrm{o}}$ is a closed subscheme of U, defined by an ideal sheaf $\mathcal{J}$. For $m \geq 0$, let Y_m be the subscheme of U defined by $\mathcal{J}^m$ and let $(X_m, \lambda_m, \theta_m)$ be the universal object over Y_m. Then X_m is an ordinary abelian scheme over Y_m and the multiplicative part $X_m[p]_\mu$ of its p-torsion is a finite, locally free subgroup scheme of X_m of rank p^g, which moreover is maximal totally isotropic for the Weil pairing e_{λ_m}. It then follows that the abelian scheme $X'_m = X_m/X_m[p]_\mu$ has a principal polarization λ'_m such that $\pi^*\lambda'_m = p \cdot \lambda_m$, where $\pi\colon X_m \to X'_m$ is the canonical map. Also, since $p \nmid n$, the level n structure θ_m naturally induces a level n structure θ'_m on X'_m.

The new triplet $(X'_m, \lambda'_m, \theta'_m)$ corresponds to a morphism $\Phi_m\colon Y_m \to \mathcal{A}_g$, which factors through Y_m. These morphisms Φ_m form a projective system. Taking the inverse limit we obtain an endomorphism $\Phi_U\colon \widehat{U} \to \widehat{U}$ on the formal completion. Finally, we can glue these Φ_U to obtain a morphism $\Phi_{\mathrm{can}}\colon \widehat{\mathcal{A}}_g \to \widehat{\mathcal{A}}_g$ over $\mathrm{Spf}(W)$. (Alternatively, we can take for U the complement of the non-ordinary locus in characteristic p, in which case $\widehat{U} = \widehat{\mathcal{A}}_g$.) It lifts the endomorphism of $(\mathcal{A}_g \otimes k)^{\mathrm{o}}$ which is obtained by pulling back the Frobenius endomorphism of $(\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p)^{\mathrm{o}}$ via $\mathrm{Spec}(k) \to \mathrm{Spec}(\mathbb{F}_p)$.

1.2 Let $(X_0, \lambda_0, \theta_0)$ be a principally polarized abelian variety of dimension g with a Jacobi level n structure over $\mathrm{Spec}(k)$. It corresponds to some closed point x of $\mathcal{A}_g \otimes k$. Let $\mathfrak{A}_x \to \mathrm{Spf}(W)$ be the formal completion of $\mathcal{A}_g$ at x. If x is an ordinary point (i.e., the corresponding abelian variety X_0 is ordinary) then $\mathfrak{A}_x$ has the structure of a formal torus over $\mathfrak{S} = \mathrm{Spf}(W)$. (See Section 3.1 for a brief discussion on formal tori.) Let us recall how this is defined.

1.3 We can define the group structure on $\mathfrak{A}_x$ over $\mathfrak{S}$ by working out the following steps:

1. Identify $\mathfrak{A}_x$, as a functor on affine formal schemes $\mathrm{Spf}(\mathcal{R})$ for $\mathcal{R}$ in a certain category $\widehat{\mathcal{C}}_W$, with the formal deformation functor of the pair (X_0, λ_0) which is studied in [48] and [56].

2. Show that this deformation functor is a group functor, using a theorem by Serre and Tate, and using that x is an ordinary point.

3. Conclude that there exist morphisms s: $\mathfrak{A}_x \times \mathfrak{A}_x \to \mathfrak{A}_x$ (multiplication), ι: $\mathfrak{A}_x \to \mathfrak{A}_x$ (inverse) and ε: $\mathfrak{S} \to \mathfrak{A}_x$ (identity element) in the category of formal schemes over $\mathfrak{S} = \mathrm{Spf}(W)$, giving the desired group structure.

1.4 To explain this in more detail, let us introduce categories $\mathcal{C}_W$ and $\widehat{\mathcal{C}}_W$, following [56]. The objects of $\mathcal{C}_W$ are the artinian local W-algebras R such that the structure homomorphism $W \to R$ is local and induces an isomorphism $k \xrightarrow{\sim} R/\mathfrak{m}_R$. The morphisms in $\mathcal{C}_W$ are the homomorphisms of W-algebras. Then $\widehat{\mathcal{C}}_W$ is defined as the category of complete noetherian local W-algebras $\mathcal{R}$ such that $\mathcal{R}/\mathfrak{m}_{\mathcal{R}}^i$ is in $\mathcal{C}_W$ for all i. Again the morphisms are just the homomorphisms of W-algebras. Notice that $\mathcal{C}_W$ is a full subcategory of $\widehat{\mathcal{C}}_W$. We consider the rings $\mathcal{R}$ in $\widehat{\mathcal{C}}_W$ with their $\mathfrak{m}_{\mathcal{R}}$-adical topology; for R in $\mathcal{C}_W$ this is just the discrete topology.

Next we define a formal deformation functor $\mathcal{D}efo_{X_0}$: $\mathcal{C}_W \to \mathsf{Sets}$, given by

$$\mathcal{D}efo_{X_0}(R) = \left\{ \begin{array}{l} \text{isomorphism classes of pairs } (X, \varphi)\text{, where } X \text{ is an abelian} \\ \text{scheme over } \mathrm{Spec}(R) \text{ and } \varphi \text{ is an isomorphism } \varphi\colon X \otimes k \xrightarrow{\sim} X_0 \end{array} \right\}.$$

Similarly, we have deformation functors $\mathcal{D}efo_{(X_0,\lambda_0)}$ and $\mathcal{D}efo_{(X_0,\lambda_0,\theta_0)}$ of the pair (X_0, λ_0), and the triplet $(X_0, \lambda_0, \theta_0)$, respectively, where in each case an isomorphism

φ over $\mathrm{Spec}(k)$ is part of the data. We extend these deformation functors to the category $\widehat{\mathcal{C}}_W$ by defining $\mathcal{D}efo_*(\mathcal{R})$ as the projective limit of the $\mathcal{D}efo_*(\mathcal{R}/\mathfrak{m}_{\mathcal{R}}^i)$.

1.5 Lemma. *The natural morphism $\mathcal{D}efo_{(X_0,\lambda_0,\theta_0)} \to \mathcal{D}efo_{(X_0,\lambda_0)}$ is an isomorphism.*

Proof. Let R be an object of $\mathcal{C}_W$, and suppose we have a polarized abelian scheme (X, λ) over $\mathrm{Spec}(R)$ and an isomorphism $\varphi\colon (X,\lambda) \otimes k \xrightarrow{\sim} (X_0, \lambda_0)$. Since $p \nmid n$, the n-torsion $X[n]$ is an étale group scheme over $\mathrm{Spec}(R)$, which implies that there is a unique level n structure on X lifting θ_0. From this the statement immediately follows. □

The functor $\mathcal{D}efo_{(X_0,\lambda_0,\theta_0)}$ is represented by $\mathfrak{A}_x$. More precisely: the composed functor $\mathfrak{A}_x \circ \mathrm{Spf}(-)\colon \widehat{\mathcal{C}}_W \to \mathsf{Sets}$ is naturally isomorphic to $\mathcal{D}efo_{(X_0,\lambda_0,\theta_0)}$. This is a formal consequence of the fact that $\mathsf{A}_{g,1,n}$ is a fine moduli scheme. Let us also remark that every triplet (X, λ, θ) over an affine formal W-scheme $\mathrm{Spf}(R)$ is algebraizable: it is the formal completion of a polarized abelian scheme with level structure over $\mathrm{Spec}(R)$.

The next step is based on a theorem of Serre and Tate, which, roughly speaking, tells us that the deformations of X_0 are completely determined by the deformations of the associated p-divisible group. To state this more precisely, we have to introduce some notations. For a given ring R and an ideal $I \subseteq R$, let $R_0 = R/I$. Write $\mathsf{AS}(R)$ for the category of abelian schemes over R and write $\mathsf{BT\text{-}Defo}(R, R_0)$ for the category of triplets (A_0, G, α), where A_0 is an abelian scheme over R_0, G is a p-divisible (or Barsotti-Tate) group over R and α is an isomorphism $\alpha\colon G_0 = G \otimes R_0 \xrightarrow{\sim} A_0[p^\infty]$.

A proof of the following theorem can be found in [31] and also in [35, Chapter 5].

1.6 Theorem. (Serre and Tate) *Let R be a ring in which the prime p is nilpotent, let $I \subseteq R$ be a nilpotent ideal and write $R_0 = R/I$. Then the functor $\mathsf{AS}(R) \to \mathsf{BT\text{-}Defo}(R, R_0)$ obtained by sending A to the triplet*

$$(A_0 = A \otimes R_0, A[p^\infty], \text{the natural isomorphism } \alpha)$$

is an equivalence of categories.

Now we start using the assumption that X_0 is an ordinary abelian variety. In this case the p-divisible group $X_0[p^\infty]$ is a direct sum $X_0[p^\infty] = X_0[p^\infty]_\mu \oplus X_0[p^\infty]_{\text{ét}}$ of a toroidal and an étale part, which are both of height g. For R in $\mathcal{C}_W$ these summands both have a unique lifting to a p-divisible group over R, say $G_{\mu,R}$ and $G_{\text{ét},R}$. The deformations of X_0 over $\mathrm{Spec}(R)$ correspond to the extension classes in $\mathrm{Ext}(G_{\text{ét},R}, G_{\mu,R})$ which are trivial over $R/\mathfrak{m}_R$. Since the set of these classes has a natural group structure, this shows that $\mathcal{D}efo_{X_0}$ is a group functor. The functor $\mathcal{D}efo_{(X_0,\lambda_0,\theta_0)} \cong \mathcal{D}efo_{(X_0,\lambda_0)}$ is a closed sub-group functor.

1.7 The existence of morphisms $s\colon \mathfrak{A}_x \times_{\mathfrak{S}} \mathfrak{A}_x \to \mathfrak{A}_x$, $\iota\colon \mathfrak{A}_x \to \mathfrak{A}_x$ and $\varepsilon\colon \mathfrak{S} \to \mathfrak{A}_x$ in the category of formal schemes over $\mathfrak{S} = \mathrm{Spf}(W)$, giving the group structure on $\mathfrak{A}_x$ over $\mathfrak{S}$, now is a formal consequence of the existence of products in this category.

1.8 We still have to show that $\mathfrak{A}_x$ is a formal torus. Let $\overline{k}$ be an algebraic closure of k, and write $\overline{W} = W(\overline{k})$ for its ring of Witt vectors. Write $T_pX_0 = T_pX_0(\overline{k})$ and $T_pX_0^t = T_pX_0^t(\overline{k})$ for the "physical" Tate modules of X_0 and X_0^t, respectively. Over $\overline{k}$ there are isomorphisms

$$X_0[p^\infty]_{\mu,\overline{k}} \cong \mathrm{Hom}(T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}}) \quad \text{and} \quad X_0[p^\infty]_{\text{ét},\overline{k}} \cong T_pX_0 \otimes_{\mathbb{Z}_p} (\mathbb{Q}_p/\mathbb{Z}_p)\,.$$

This leads to a description of $\mathcal{D}efo_{X_0}(R)$ (with R in $\mathcal{C}_{\overline{W}}$) as the category of extensions

$$\mathrm{Ext}_R(T_pX_0 \otimes_{\mathbb{Z}_p} (\mathbb{Q}_p/\mathbb{Z}_p), \mathrm{Hom}(T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}}))\,.$$

Since $\mathrm{Ext}((\mathbb{Q}_p/\mathbb{Z}_p), \widehat{\mathbb{G}}_{\mathrm{m}}) \cong \widehat{\mathbb{G}}_{\mathrm{m}}$ (considering both sides as functors on $\mathcal{C}_{\overline{W}}$) it follows that $\mathcal{D}efo_{X_0\otimes\overline{k}}$ is a formal torus.

It can be shown that every class in $\mathrm{Ext}(T_pX_0 \otimes_{\mathbb{Z}_p} (\mathbb{Q}_p/\mathbb{Z}_p), \mathrm{Hom}(T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}}))$ (over $\overline{W}$) is obtained from the exact sequence

$$0 \to T_pX_0 \to T_pX_0 \otimes \mathbb{Q}_p \to T_pX_0 \otimes (\mathbb{Q}_p/\mathbb{Z}_p) \to 0$$

by pushing out along a homomorphism $T_pX_0 \to \mathrm{Hom}(T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}})$. This leads to an isomorphism of functors $\mathcal{D}efo_{X_0\otimes\overline{k}} \xrightarrow{\sim} \mathrm{Hom}(T_pX_0 \otimes T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}})$.

We can also describe liftings of homomorphisms between ordinary abelian varieties. So, suppose X_0 and X_0' are ordinary abelian varieties over k, and let $f\colon X_0 \to X_0'$ be a homomorphism. For $\mathcal{R}$ in $\widehat{\mathcal{C}}_W$, let $G_{\text{ét}}$ and $G'_{\text{ét}}$ be the liftings over $\mathcal{R}$ of $X_0[p^\infty]_{\text{ét}}$ and $X_0'[p^\infty]_{\text{ét}}$ respectively, and, similarly, write G_μ and G'_μ for the liftings of the toroidal p-divisible groups. Then f induces homomorphisms $f_{\text{ét}}\colon G_{\text{ét}} \to G'_{\text{ét}}$ and $f_\mu\colon G_\mu \to G'_\mu$. As we have seen, liftings of X_0 and X_0' are given by extensions of p-divisible group, and f lifts to a homomorphism $\tilde{f}\colon X \to X'$ if and only if it extends to a homomorphism of the corresponding extensions. If $[X]$ and $[X']$ are the extension classes of X and X' respectively, then this is the case precisely if $(f_\mu)_*[X] = (f_{\text{ét}})^*[X']$ in $\operatorname{Ext}(G'_\mu, G_{\text{ét}})$. For $\mathcal{R}$ in $\widehat{\mathcal{C}}_{\overline{W}}$ this becomes the condition that the elements $q_X \in \operatorname{Hom}(T_pX_0 \otimes T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}})(\mathcal{R})$ and $q_{X'} \in \operatorname{Hom}(T_pX'_0 \otimes T_p(X'_0)^t, \widehat{\mathbb{G}}_{\mathrm{m}})(\mathcal{R})$ satisfy

$$q_X(\alpha, f^t(\beta)) = q_{X'}(f(\alpha), \beta)$$

for all elements $\alpha \in T_pX_0$, $\beta \in T_pX'_0$.

As an application, consider the isomorphism $T_pX_0 \xrightarrow{\sim} T_pX_0^t$ induced by the polarization λ_0 (which we assumed to be principal). From the preceding remarks we derive that $\mathfrak{A}_x \otimes \overline{W}$ is the formal subtorus $\operatorname{Hom}(\operatorname{Sym}^2(T_pX_0), \widehat{\mathbb{G}}_{\mathrm{m}})$ of $\operatorname{Hom}(T_pX_0^{\otimes 2}, \widehat{\mathbb{G}}_{\mathrm{m}}) \cong \operatorname{Hom}(T_pX_0 \otimes T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}})$. For proofs we refer to [31, esp. Theorem 2.1].

1.9 Let $\sigma\colon W \to W$ be the Frobenius automorphism of W. The Frobenius morphism $\mathrm{Frob}\colon X_0 \to X_0^{(p)}$ induces an isomorphism $X_0/X_0[p]_\mu \xrightarrow{\sim} X_0^{(p)}$. Using this we see that the formal completion of $\mathcal{A}_g$ at the point $\Phi_{\mathrm{can}}(x)$ is isomorphic to $\mathfrak{A}_x^{(\sigma)} = \mathfrak{A}_x \times_{\mathfrak{S},\sigma} \mathfrak{S}$. The morphism Φ_{can} introduced in 1.1 above therefore induces a morphism $\mathfrak{A}_x \to \mathfrak{A}_x^{(\sigma)}$, which we again call Φ_{can} (cf. [21, p. 135]). It is not difficult to see that this is a group homomorphism. If $N = {}^p\log(\#k)$ (i.e., $k \cong \mathbb{F}_{p^N}$) then Φ_{can}^N is an endomorphism of $\mathfrak{A}_x$. It is the endomorphism "raising to the p^Nth power" in the group $\mathfrak{A}_x$.

The next lemma, which we quote from the appendix to [21] by Katz, shows that the group structure is uniquely determined by the fact that it is compatible with Φ_{can}.

1.10 Lemma. (Katz, [21, A.1]) *Let k be a perfect field of characteristic $p > 0$, let $W = W(k)$ its ring of Witt vectors, and let $\sigma\colon W \xrightarrow{\sim} W$ be the automorphism induced by the Frobenius automorphism of k. Let $\mathcal{M}$ be a formally smooth affine formal scheme of finite type over W, i.e., $\mathcal{M} \cong \operatorname{Spf}(W[\![t_1, \ldots, t_n]\!])$. Suppose we are given a morphism*

$$\Phi\colon \mathcal{M} \longrightarrow \mathcal{M}^{(\sigma)}$$

of formal schemes over W whose reduction modulo p is the Frobenius morphism

$$\mathrm{Frob}\colon \mathcal{M} \otimes_W k \longrightarrow (\mathcal{M} \otimes_W k)^{(p)}\,.$$

(i) Given $(\mathcal{M}, \Phi)$, there exists at most one structure of commutative formal group over W on $\mathcal{M}$ for which the given map $\Phi\colon \mathcal{M} \to \mathcal{M}^{(\sigma)}$ is a group homomorphism.

(ii) If this structure exists, it makes $\mathcal{M}$ into a formal torus and the given Φ is the unique homomorphism lifting Frobenius.

(iii) If $(\mathcal{M}_1, \Phi_1)$ and $(\mathcal{M}_2, \Phi_2)$ both admit group structures as in (i), then a morphism $f\colon \mathcal{M}_1 \to \mathcal{M}_2$ of formal schemes over W is a group homomorphism if and only if the diagram

$$\begin{array}{ccc} \mathcal{M}_1 & \xrightarrow{f} & \mathcal{M}_2 \\ {\scriptstyle \Phi_1}\downarrow & & \downarrow{\scriptstyle \Phi_2} \\ \mathcal{M}_1^{(\sigma)} & \xrightarrow{f^{(\sigma)}} & \mathcal{M}_2^{(\sigma)} \end{array}$$

is commutative.

1.11 Remark. Fix an integer $m \geq 1$. We can replace (iii) by the slightly stronger statement that f is a group homomorphism if and only if the diagram

$$\begin{array}{ccc} \mathcal{M}_1 & \xrightarrow{f} & \mathcal{M}_2 \\ {\scriptstyle \Phi_1^m}\downarrow & & \downarrow{\scriptstyle \Phi_2^m} \\ \mathcal{M}_1^{(\sigma^m)} & \xrightarrow{f^{(\sigma^m)}} & \mathcal{M}_2^{(\sigma^m)} \end{array}$$

is commutative. The proof is similar to the one given by Katz in [21, p. 130]. A stronger statement is proved in [11].

In particular (the case that $\mathcal{M}_2 = \widehat{\mathbb{G}}_{\mathrm{m}}$), a function q on $\mathcal{M}_1$ is a character if and only if $(\Phi^m)^*(q^{(\sigma^m)}) = q^{p^m}$.

1.12 Choose a $\mathbb{Z}_p$-basis $\alpha_1, \ldots, \alpha_g$ for T_pX_0 and let $\alpha_1^t, \ldots, \alpha_g^t$ be the basis of $T_pX_0^t$ given by $\alpha_i^t = \lambda_0(\alpha_i)$. Define characters $q_{ij} = q(\alpha_i, \alpha_j^t)$ on

$$\mathrm{Hom}(T_pX_0 \otimes T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}}) \cong \mathrm{Hom}(T_pX_0^{\otimes 2}, \widehat{\mathbb{G}}_{\mathrm{m}})$$

by sending $\varphi \in \mathrm{Hom}_R(T_pX_0 \otimes T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}})$ to $\varphi(\alpha_i \otimes \alpha_j^t) \in \widehat{\mathbb{G}}_{\mathrm{m}}(R) = 1 + \mathfrak{m}_R$. This gives an isomorphism

$$\mathrm{Hom}(T_pX_0 \otimes T_pX_0^t, \widehat{\mathbb{G}}_{\mathrm{m}}) \cong \mathrm{Spf}(\overline{W}[\![\{q_{ij} - 1\}_{1 \leq i,j \leq g}]\!])\,.$$

(This is a rather tautological statement.) We can describe $\mathfrak{A} = \mathfrak{A}_x \otimes \overline{W}$ as the formal subtorus $\mathrm{Spf}(\overline{W}[\![q_{ij} - 1]\!]/(q_{ij} - q_{ji}))$. To keep the notations easy we write $A = \overline{W}[\![q_{ij} - 1]\!]/(q_{ij} - q_{ji})$.

Let $\mathfrak{X} \to \mathfrak{A}$ be the universal (polarized) formal abelian scheme. The A-module $H = \mathrm{H}^1_{\mathrm{DR}}(\mathfrak{X}/\mathfrak{A})$ with the Gauß-Manin connection ∇ and its Hodge filtration

$$\mathcal{F}^0 = H \ \supset \ \mathcal{F}^1 = \mathrm{H}^0(X, \Omega^1_{\mathfrak{X}/\mathfrak{A}})$$

has the structure of an ordinary Hodge F-crystal of level 1. Such crystals are studied in [21]. Here we confine ourselves to giving a description of H; proofs of our statements can be found in [31].

To the chosen elements α_i and α_j^t we can associate elements $a_1, \ldots, a_g$ (called $\mathrm{Fix}(\alpha_i^\vee)$ in [31]) and $b_1, \ldots, b_g$ ($\omega(\alpha_j^t)$ in loc. cit.) of H such that $\mathcal{F}^1 = A\cdot b_1 \oplus \cdots \oplus A\cdot b_g$ and H is the direct sum of $U = A \cdot a_1 \oplus \cdots A \cdot a_g$ and $\mathcal{F}^1$. For the connection ∇ we have

$$\nabla(a_i) = 0, \quad \nabla(b_j) = \sum_i a_i \otimes \eta_{ij}$$

for certain forms $\eta_{ij} \in \Omega^1_{\mathfrak{A}/W}$ (continuous differential forms). Furthermore,

$$F(\Phi_{\mathrm{can}})\Phi^*_{\mathrm{can}}(a_i^{(\sigma)}) = a_i,\ F(\Phi_{\mathrm{can}})\Phi^*_{\mathrm{can}}(b_i^{(\sigma)}) = pb_i,\ \Phi^*_{\mathrm{can}}(\eta_{ij}^{(\sigma)}) = p\eta_{ij}, \quad \text{and} \quad \mathrm{d}\eta_{ij} = 0\,.$$

In particular, U is a unit sub-F-crystal of H.

Let K be the fraction field of $W(\overline{k})$ and write $\tau_{ij} = \log(q_{ij}) \in K[\![q_{ij} - 1]\!]$. Let $B = K[\![\tau_{ij}]\!]/(\tau_{ij} - \tau_{ji})$, then we obtain a homomorphism $A \to B$ by sending q_{ij} to $\exp(\tau_{ij})$. We have the identities $\Phi^*_{\mathrm{can}}(q_{ij}^{(\sigma)}) = q_{ij}^p$, $\Phi^*_{\mathrm{can}}(\tau_{ij}^{(\sigma)}) = p\tau_{ij}$ and $\eta_{ij} = \mathrm{d}\tau_{ij}$. If $\mathbf{0}\colon A \to W(\overline{k})$ is the Teichmüller lift of the augmentation map $A \twoheadrightarrow \overline{k}$ with respect to Φ_{can} then $q_{ij}(\mathbf{0}) = 1$ and $\tau_{ij}(\mathbf{0}) = 0$.

Write $c_j = b_j - \sum_i \tau_{ij} a_i$. The elements $a_1, \ldots, a_g, c_1, \ldots, c_g$ form a horizontal B-basis for $H \widehat{\otimes}_A B$, and the Hodge flag $\mathcal{F}^1 \widehat{\otimes}_A B$ is spanned by the elements $c_j + \sum_i \tau_{ij} a_i$.

1.13 As above, let $N = {}^p\log(\#k)$. We still assume that X_0 is an ordinary abelian variety. Write $\pi = \pi_{X_0}\colon X_0 \to X_0$ for the Frobenius endomorphism of X_0 (so "$\pi = \mathrm{Frob}^N$"). Let $\mathcal{R}$ be an object of $\widehat{\mathcal{C}}_W$ and let $s\colon \mathrm{Spf}(\mathcal{R}) \to \mathfrak{A}_x$ be an $\mathcal{R}$-valued point of $\mathfrak{A}_x$ over $\mathrm{Spf}(W)$. Let X_s denote the corresponding abelian scheme over $\mathrm{Spec}(\mathcal{R})$.

We say that an abelian scheme X over $\mathrm{Spec}(\mathcal{R})$ is of CM-type if $\mathrm{End}^0(X \otimes \overline{\mathcal{R}})$ contains a commutative semi-simple $\mathbb{Q}$-subalgebra of rank $2g$ over $\mathbb{Q}$. Here $\overline{\mathcal{R}} = \mathcal{R} \widehat{\otimes} \overline{W}$. If $\mathcal{R}$ is a normal domain then X is of CM-type if and only if its generic fibre is of CM-type.

1.14 Lemma. *The following conditions are equivalent.*

(a) *s is the identity element of $\mathfrak{A}_x(\mathcal{R})$.*

(b) $\mathrm{End}_{\mathcal{R}}(X_s) \xrightarrow{\sim} \mathrm{End}_k(X_0)$.

(c) *π^m lifts to an endomorphism of X_s for some $m \geq 1$.*

Proof. The implications (a) $\Rightarrow$ (b) $\Rightarrow$ (c) are obvious. If π^m lifts to an endomorphism of X_s then this is certainly also the case over $\mathcal{R} \otimes \overline{W}$, and it follows from the discussion in 1.8 that the extension class $q_s = q_{X_s} \in \mathrm{Hom}(T_pX_0^{\otimes 2}, \widehat{\mathbb{G}}_{\mathrm{m}})(\mathcal{R} \otimes \overline{W})$ satisfies $q_s(\pi^m(\alpha), \beta) = q_s(\alpha, (\pi^t)^m(\beta))$ for all $\alpha, \beta \in T_pX_0$. Now π induces an automorphism of T_pX_0, whereas the endomorphism of T_pX_0 induced by π^t is divisible by p. It follows that $q_s(\alpha, \beta)$ is a p^{im}th power for every $i \in \mathbb{Z}_{>0}$, and from this we conclude that s must be the identity element of $\mathfrak{A}_x(\mathcal{R})$. □

1.15 Lemma. *The following conditions are equivalent.*

(a) *s is a torsion element of $\mathfrak{A}_x(\mathcal{R})$.*

(b) $\mathrm{End}_{\mathcal{R}}(X_s) \otimes \mathbb{Z}[1/p] \xrightarrow{\sim} \mathrm{End}_k(X_0) \otimes \mathbb{Z}[1/p]$.

(c) *X_s is isogenous to the lifting X_1 (where 1 is the identity element of $\mathfrak{A}_x(\mathcal{R})$).*

(d) *X_s is of CM-type.*

Proof. Assume s is torsion in $\mathfrak{A}_x(\mathcal{R})$. Its order must be a p-power, since $\mathfrak{A}_x(\mathcal{R})$ is ℓ-divisible for all primes $\ell \neq p$. We therefore have $s^{p^m} = 1$ for some $m \in \mathbb{Z}_{\geq 1}$. Let φ be a multiple of p^m in $\mathrm{End}_k(X_0)$. To prove (b) it suffices to show that φ lifts to an endomorphism of the p-divisible group $X_s[p^\infty]$, using the Serre-Tate theorem 1.6. This reduces to a general statement about extensions: given an extension $0 \to A \to E \to B \to 0$ in an abelian category such that the corresponding class in $\mathrm{Ext}(B, A)$ is n-torsion, then every multiple of n in $\mathrm{End}(A \oplus B)$ induces an endomorphism of E. With a similar argument, taking $\varphi = p^m \cdot \mathrm{Id}$, we conclude that (a) implies (c).

The implications (b) $\Rightarrow$ (d) and (c) $\Rightarrow$ (d) are obvious. Next assume that X_s is of CM-type. Then $X_s \otimes \overline{\mathcal{R}}$ also is, and it suffices to show that the corresponding element $s_{\overline{\mathcal{R}}} \in \mathfrak{A}_x(\overline{\mathcal{R}})$ is torsion. The endomorphism algebra $\mathrm{End}^0(X_s \otimes \overline{\mathcal{R}})$ contains a commutative subalgebra K of rank $2g$ over $\mathbb{Q}$. The image of K under $\mathrm{End}^0(X_s \otimes \overline{\mathcal{R}}) \hookrightarrow \mathrm{End}^0(X_0 \otimes \overline{k})$ contains the center of $\mathrm{End}^0(X_0 \otimes \overline{k})$, and in view of [63, Theorem 2] it follows that $m \cdot \pi$ lifts to an endomorphism of $X_s \otimes \overline{\mathcal{R}}$ for some $m \in \mathbb{Z}_{\geq 1}$. If X' is the lifting of $X_0 \otimes \overline{k}$ given by the class $q' \in \mathrm{Hom}(T_pX_0^{\otimes 2}, \widehat{\mathbb{G}}_{\mathrm{m}})(\overline{\mathcal{R}})$ with $q'(\alpha, \beta) = q_s(\alpha, \beta)^m$ then it follows that π lifts to an endomorphism of X', hence by the previous lemma $q' = 1$. □

1.16 Definition. Suppose $\mathcal{R} \in \widehat{\mathcal{C}}_W$ is a flat W-algebra. The lifting of X_0 over $\mathrm{Spec}(\mathcal{R})$ corresponding to the identity element of $\mathfrak{A}_x(\mathcal{R})$ is called the canonical lifting of X_0. We denote it by X_0^{can}. The liftings of X_0 over $\mathrm{Spec}(\mathcal{R})$ corresponding to the torsion elements of $\mathfrak{A}_x(\mathcal{R})$ are called quasi-canonical liftings; by 1.15 these are precisely the CM-liftings of X_0.

1.17 Remark. The canonical and quasi-canonical liftings of an ordinary abelian variety X_0 were first studied by Serre and Tate; see the report [33].

Property (c) in Lemma 1.14 is a little bit weaker than what is needed for the arguments in Section 3. The statement we need is the following: given an ordinary abelian variety X_0 over a finite field k, then a lifting X of X_0 over $W(k)$ is the canonical lifting if and only if Frob: $X_0 \to X_0^{(p)}$ lifts to a morphism $\widetilde{F}$: $X \to X^{(\sigma)}$. A proof of this statement (which does not make sense for liftings over more general rings $\mathcal{R}$) can be found in [35, Appendix, Corollary (1.2)].

The property that Frob lifts can be formulated in terms of Φ_{can}. Let us give the statement in the form we require.

1.18 Lemma. *Consider the formal scheme $\widehat{\mathcal{A}}_g$ over $\mathrm{Spf}(W)$ as in Section 1.1, with $k \cong \mathbb{F}_{p^N}$. Let $k \to k'$ be a finite field extension, and let s: $\mathrm{Spf}(W(k')) \to \widehat{\mathcal{A}}_g$ be a $W(k')$-valued point, giving rise to a triplet (X, λ, θ) over $\mathrm{Spec}(W(k'))$. Let $\mathrm{Frob} = \mathrm{Frob}_{X_{k'}/k'}$: $X_{k'} \to X_{k'}^{(p)}$ be the Frobenius morphism. If a is a multiple of N then Frob^a lifts to a morphism $\widetilde{F}$: $X \to X^{(\sigma^a)}$ over $\mathrm{Spec}(W)$ if and only if the diagram*

$$\begin{array}{ccc} \mathrm{Spf}(W(k')) & \xrightarrow{s} & \widehat{\mathcal{A}}_g \\ \downarrow{\scriptstyle \sigma^a} & & \downarrow{\scriptstyle \Phi^a_{\mathrm{can}}} \\ \mathrm{Spf}(W(k')) & \xrightarrow{s} & \widehat{\mathcal{A}}_g \end{array}$$

is commutative.

Proof. For $m \in \mathbb{Z}_{>0}$ let s_m: $\mathrm{Spf}(W_m(k')) \to \widehat{\mathcal{A}}_g \otimes W_m$ be the reduction of s modulo p^m, and let Φ_m be the endomorphism of $\widehat{\mathcal{A}}_g \otimes W_m$ induced by Φ_{can}. Then Frob^a lifts over $W(k')$ if and only if it lifts (necessarily uniquely) over each $W_m(k')$. Also, $\Phi^a_{\mathrm{can}} \circ s = s \circ \sigma^a$ if and only if $\Phi^a_m \circ s_m = s_m \circ \sigma^a_m$ for all m. (Notice that this holds for $m = 1$, since a is a multiple of N.)

Let $(X_m, \lambda_m, \theta_m)$ be the reduction of (X, λ, θ) modulo p^m and, similar to the construction in 1.1, let $(X'_m, \lambda'_m, \theta'_m)$ be the triplet obtained by dividing out $X_m[p^a]_\mu$.

Clearly, $\Phi_m^a \circ s_m = s_m \circ \sigma_m^a$ if and only if $(X_m^{(\sigma^a)}, \lambda_m^{(\sigma^a)}, \theta_m^{(\sigma^a)}) \cong (X'_m, \lambda'_m, \theta'_m)$. For $m = 1$, there is a unique such isomorphism, induced by Frob^a. We conclude that $\Phi_m^a \circ s_m = s_m \circ \sigma_m^a$ implies that Frob^a lifts over $W_m(k')$.

Conversely, if Frob^a lifts to $\widetilde{F}_m$ over $W_m(k')$ then $(\widetilde{F}_m)^* \lambda_m^{(\sigma^a)} = p^{2ag} \cdot \lambda_m$, since $\deg(\widetilde{F}_m) = \deg(\mathrm{Frob}^a) = p^{ag}$. Also, $(\widetilde{F}_m)^* \theta_m^{(\sigma^a)} = \theta_m$, since we consider level n structures with $p \nmid n$. It follows that $(X_m^{(\sigma^a)}, \lambda_m^{(\sigma^a)}, \theta_m^{(\sigma^a)}) \cong (X'_m, \lambda'_m, \theta'_m)$, and this proves the lemma. □

1.19 So far we only discussed the "unramified" case, studying formal completions of the scheme $\mathsf{A}_{g,1,n}$ over a ring of Witt vectors. By base change we can extend most of the above results to a slightly more general situation, which is what we need for the next sections. Since most of this is obvious, the following remarks are mainly intended to fix notations, which extend the ones used before.

Let F be a number field with ring of integers $\mathcal{O}_F$, and let $\mathfrak{p}$ be a prime of $\mathcal{O}_F$ lying over p. We write $\widehat{\mathcal{O}}_{\mathfrak{p}}$ for the completion of the local ring $\mathcal{O}_{\mathfrak{p}}$. Write $\mathcal{A}_g = \mathsf{A}_{g,1,n} \otimes \widehat{\mathcal{O}}_{\mathfrak{p}}$, and let $\widehat{\mathcal{A}}_g$ be the formal completion of $\mathcal{A}_g$ along the ordinary locus in characteristic p. We obtain a morphism $\Phi_{\mathrm{can}}\colon \widehat{\mathcal{A}}_g \to \widehat{\mathcal{A}}_g$ over $\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathfrak{p}})$ by pulling back the Φ_{can} defined in 1.1 via $\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathfrak{p}}) \to \mathrm{Spf}(W(\kappa(\mathfrak{p})))$.

Let x be a closed point of the ordinary locus $(\mathcal{A}_g \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$. Consider the ring $\Lambda = W(\kappa(x)) \otimes_{W(\kappa(\mathfrak{p}))} \widehat{\mathcal{O}}_{\mathfrak{p}}$, which is a complete local ring with residue field $\kappa(x)$, and write $\mathfrak{S} = \mathrm{Spf}(\Lambda)$. We let $\mathfrak{A}_x \to \mathfrak{S}$ be the formal completion of $\mathcal{A}_g$ at x (which has a natural morphism to $\mathfrak{S}$). It is obtained via base change $\mathfrak{S} \to \mathrm{Spf}(W(\kappa(x)))$ from a formal deformation space as studied above, and therefore has the structure of a formal torus over $\mathfrak{S}$. Via this base change and the results of 1.12 we also get a description of the de Rham cohomology $\mathrm{H}^1_{\mathrm{DR}}(\mathfrak{X}/\mathfrak{A})$ in this more general setting.

§2 Isogenies

2.1 In this section we recall the definition of an algebraic stack called $\mathsf{Isog}(p^{eg})$ (in the cases we study it is actually a scheme). Here we follow [25, Chapter VII]. We use the scheme $\mathsf{Isog}(p^{eg})$ to give an alternative description of the morphism Φ_{can} introduced in Section 1.1.

Recall that a p-isogeny between principally polarized abelian schemes (X, λ) and (X', λ') of relative dimension g over a base scheme S is an isogeny $f\colon X \to X'$ such that $f^*\lambda' = p^e \cdot \lambda$ for some $e \in \mathbb{Z}_{\geq 1}$. If this holds then f has degree p^{eg}. If X and X' are equipped with level n structures θ and θ' ($p \nmid n$) then we further require that $f^*\theta' = \theta$ (meaning that θ "=" θ' via the isomorphism $T^pX \xrightarrow{\sim} T^pX'$ on the "prime-to-p Tate modules" induced by f).

Let A_g be the moduli stack of pincipally polarized abelian schemes, as in [25, Chapter I, 4.3]. The p-isogenies form a stack p-Isog, with two natural morphisms $\mathrm{pr}_1, \mathrm{pr}_2$: p-$\mathsf{Isog} \to \mathsf{A}_g$ obtained by associating to an isogeny $f\colon (X,\lambda) \to (X',\lambda')$ its source (X,λ), and its target (X',λ'), respectively. Bounding the degree of the isogeny gives a substack of p-Isog which is representable by a relative scheme over $\mathsf{A}_g \times \mathsf{A}_g$. We write $\mathsf{Isog}(p^{eg})$ for the stack of p-isogenies of degree p^{eg} (it is empty if the degree is not a power of p^g).

As a variant, we can take level structures into account. Choose an integer $n \geq 3$ with $p \nmid n$ and, as before, write $\mathsf{A}_{g,1,n}$ for the moduli scheme over $\mathrm{Spec}(\mathbb{Z}[1/n])$ of principally polarized g-dimensional abelian varieties with a Jacobi level n structure. It is a fine moduli scheme ($n \geq 3$). By considering isogenies which respect level structures we obtain a scheme $\mathsf{Isog}(p^{eg})$ over $\mathsf{A}_{g,1,n} \times \mathsf{A}_{g,1,n}$; to keep notations easy we do not write a subscript "$g, 1, n$" to $\mathsf{Isog}(p^{eg})$.

2.2 We use the notations F, $\mathfrak{p}$ and $\widehat{\mathcal{O}}_{\mathfrak{p}}$ of Section 1.19, and we write $\mathcal{A}_g = \mathsf{A}_{g,1,n} \otimes \widehat{\mathcal{O}}_{\mathfrak{p}}$, $\mathcal{I}sog(p^{eg}) = \mathsf{Isog}(p^{eg}) \otimes \widehat{\mathcal{O}}_{\mathfrak{p}}$. Write $\mathcal{A}_g^{\mathrm{o}} \subset \mathcal{A}_g$ for the open subscheme obtained by deleting the non-ordinary locus in characteristic p. The isogenies lying over $\mathcal{A}_g^{\mathrm{o}} \times \mathcal{A}_g^{\mathrm{o}}$ form an open subscheme $\mathcal{I}sog(p^{eg})^{\mathrm{o}}$ of $\mathcal{I}sog(p^{eg})$. The restricted projection morphisms $\mathrm{pr}_i\colon \mathcal{I}sog(p^{eg})^{\mathrm{o}} \to \mathcal{A}_g^{\mathrm{o}}$ are finite and flat.

The ordinary locus $\mathcal{I}sog(p^{eg})^{\mathrm{o}} \otimes \kappa(\mathfrak{p})$ in characteristic p is a locally closed subscheme of $\mathcal{I}sog(p^{eg})$. We can take the formal completion along it to obtain a formal scheme $\mathcal{I}sog(p^{eg})^\wedge$ over $\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathfrak{p}})$, with projection maps $\mathrm{pr}_i\colon \mathcal{I}sog(p^{eg})^\wedge \to \widehat{\mathcal{A}}_g$.

2.3 Proposition. *There is an open and closed formal subscheme $\widehat{\mathcal{I}} \subseteq \mathcal{I}sog(p^{eg})^\wedge$ such that the restriction $\widehat{\mathrm{pr}}_1\colon \widehat{\mathcal{I}} \to \widehat{\mathcal{A}}_g$ is an isomorphism, and such that the composition*

$$\widehat{\mathcal{A}}_g \xrightarrow{\widehat{\mathrm{pr}}_1^{-1}} \widehat{\mathcal{I}} \xrightarrow{\widehat{\mathrm{pr}}_2} \widehat{\mathcal{A}}_g$$

is equal to the morphism Φ_{can}^e, where Φ_{can} is defined as in 1.1 and 1.19. The reduced underlying scheme $\widehat{\mathcal{I}}_{\mathrm{red}}$ is a disjoint union of irreducible components.

Proof. This is essentially [25, Proposition VII.4.1]; in the notation of loc. cit., our $\widehat{\mathcal{I}}$

is the formal completion along the subscheme of $\mathsf{Isog}(p^{eg})$ classifying isogenies of type LmL with $m = \mathrm{diag}(p^e \cdot \mathrm{Id}_g, 1 \cdot \mathrm{Id}_g)$. In other words, $\widehat{\mathcal{I}}$ is the formal completion along the pull-back (via $\mathrm{Spec}(\kappa(\mathfrak{p})) \to \mathrm{Spec}(\mathbb{F}_p)$) of the graph of the Frobenius morphism. Let us nevertheless sketch a proof.

Except for the last statement, it suffices to prove the proposition over $\mathbb{Z}_p$, since all ingredients over $\widehat{\mathcal{O}}_\mathfrak{p}$ are obtained via pull-back over $\mathrm{Spf}(\widehat{\mathcal{O}}_\mathfrak{p}) \to \mathrm{Spf}(\mathbb{Z}_p)$. We only do the case $e = 1$; the general argument only differs in that the notations are more complicated.

Write $(X_m, \lambda_m, \theta_m)$ for the universal object over $\mathcal{A}_g^{\mathrm{o}} \otimes (\mathbb{Z}/p^m)$. In Section 1.1, we have defined Φ_{can} on $\widehat{\mathcal{A}}_g$ as the limit of morphisms Φ_m such that $\Phi_m^*(X_m, \lambda_m, \theta_m) = (X'_m, \lambda'_m, \theta'_m)$, obtained by dividing out $X_m[p]_\mu$. (Here we apply the discussion of Section 1.1 to $U = \mathcal{A}_g^{\mathrm{o}}$, in which case $\mathcal{J} = p \cdot \mathcal{O}_U$.)

We obtain a section s_m of $\mathrm{pr}_1\colon \mathcal{I}sog(p^g)^{\mathrm{o}} \otimes (\mathbb{Z}/p^m) \to \mathcal{A}_g^{\mathrm{o}} \otimes (\mathbb{Z}/p^m)$ by associating to $(X_m, \lambda_m, \theta_m)$ the natural isogeny $\pi_m\colon X_m \to X'_m = X_m/X_m[p]_\mu$ (compatible with polarizations and level structures). Clearly, $\mathrm{pr}_2 \circ s_m = \Phi_m$ on Y_m.

Define $I_m \subseteq \mathcal{I}sog(p^g)^{\mathrm{o}} \otimes (\mathbb{Z}/p^m)$ as the (scheme-theoretic) image of s_m. The section s_m maps into the open subscheme of $\mathcal{I}sog(p^g)^{\mathrm{o}} \otimes (\mathbb{Z}/p^m)$ of isogenies with a kernel of multiplicative type. Over this locus, the first projection is finite étale (by rigidity of group schemes of multiplicative type). It follows that I_m is an open and closed subscheme of $\mathcal{I}sog(p^g)^{\mathrm{o}} \otimes (\mathbb{Z}/p^m)$, with $s_m\colon \mathcal{A}_g^{\mathrm{o}} \otimes (\mathbb{Z}/p^m) \xrightarrow{\sim} I_m$. Moreover, $I_m = I_{m+k} \otimes_{(\mathbb{Z}/p^{m+k})} (\mathbb{Z}/p^m)$ for every $k \geq 0$.

Define $\widehat{\mathcal{I}} \subseteq \mathcal{I}sog(p^{eg})^\wedge$ as the formal subscheme with $\widehat{\mathcal{I}} \otimes (\mathbb{Z}/p^m) = I_m$ for every $m \geq 0$. It follows from the preceding remarks that $\widehat{\mathcal{I}}$ is an open and closed formal subscheme of $\mathcal{I}sog(p^{eg})^\wedge$. The section $s\colon \widehat{\mathcal{A}}_g \to \widehat{\mathcal{I}}$ obtained by taking the limit over all s_m is an isomorphism, and $\widehat{\mathrm{pr}}_2 \circ s = \Phi_{\mathrm{can}}$. This proves the proposition, except for the statement that $\widehat{\mathcal{I}}_{\mathrm{red}}$ is a disjoint union of irreducible components. To see this, remark that the topological space underlying $\widehat{\mathcal{I}}_{\mathrm{red}}$ is homeomorphic to that of $(\mathcal{A}_g \otimes \kappa(\mathfrak{p}))^{\mathrm{o}} = \mathcal{A}_g^{\mathrm{o}} \otimes \kappa(\mathfrak{p})$. Since this is the disjoint union of irreducible components, the same holds for $\widehat{\mathcal{I}}_{\mathrm{red}}$. □

§3 Formal linearity

3.1 Let k be a perfect field of characteristic $p > 0$, let $\overline{k}$ an algebraic closure of k, and write $W = W(k)$, $\overline{W} = W(\overline{k})$. Let Λ be a complete local noetherian ring with residue field k. A formal group $\mathcal{M}$ over $\mathrm{Spf}(\Lambda)$ (defined as in [23, Exposé $\mathrm{VII_B}$]) is called a formal torus if $\mathcal{M}\widehat{\otimes}\overline{\Lambda}$ is isomorphic to $\widehat{\mathbb{G}}_{\mathrm{m}}^d$ for some $d \geq 0$, where $\overline{\Lambda} = \Lambda\widehat{\otimes}_W\overline{W}$ (which again is a complete local noetherian ring). Such a formal torus is completely determined by its fibre $\mathcal{M}\widehat{\otimes}_\Lambda k$. In particular, every formal torus is defined over W, and there is an (anti-)equivalence of categories

$$\left\{ \begin{array}{l} \text{formal tori} \\ \text{over } \mathrm{Spf}(\Lambda) \end{array} \right\}^{\mathrm{o}} \xrightarrow{\text{eq.}} \left\{ \begin{array}{l} \text{free } \mathbb{Z}_p\text{-modules of finite rank with a} \\ \text{continuous action of } \mathrm{Gal}_{\mathrm{cont}}(\overline{\Lambda}/\Lambda) \cong \mathrm{Gal}(\overline{k}/k) \end{array} \right\}$$

by associating to $\mathcal{M}$ its character group $X^*(\mathcal{M}) = \mathrm{Hom}(\mathcal{M}\widehat{\otimes}\overline{\Lambda}, \widehat{\mathbb{G}}_{\mathrm{m}})$.

For a Galois submodule $Y \subseteq X^*(\mathcal{M})$ we write $\mathcal{N}(Y)$ for the common kernel of the characters $\chi \in Y$. If Y is primitive (meaning that the quotient group is torsion free) this is a formal subtorus of $\mathcal{M}$ with character group $X^*(\mathcal{M})/Y$. For general Y it has the form $\mathcal{N}(Y) = \mathfrak{T}\cdot\mathcal{N}$, where $\mathcal{N}$ is a formal subtorus of $\mathcal{M}$ and $\mathfrak{T}$ is a torsion subgroup.

3.2 Lemma. *Let k be a finite field, and let $W, \mathcal{M}$ and Φ be as in Lemma 1.10. Suppose $\mathcal{M}$ has the structure of a formal torus such that $\Phi\colon \mathcal{M} \to \mathcal{M}^{(\sigma)}$ is a group homomorphism. Let Λ be a complete discrete valuation ring with residue field k, and let $\mathcal{N} \subseteq \mathcal{M}_\Lambda$ be an irreducible closed formal subscheme of $\mathcal{M}_\Lambda = \mathcal{M} \times_{\mathrm{Spf}(W)} \mathrm{Spf}(\Lambda)$ which is formally smooth over Λ. Take an integer $m \geq 1$ such that the automorphism σ^m of W lifts to an automorphism τ of Λ. Then the following properties are equivalent.*

(a) *$\mathcal{N}$ is a formal subtorus of $\mathcal{M}_\Lambda$.*

(b) *There is a primitive Galois submodule $Y \subseteq X^*(\mathcal{M})$ such that $\mathcal{N} = \mathcal{N}(Y)_\Lambda$.*

(c) *$\Phi^m|_{\mathcal{N}}\colon \mathcal{N} \to \mathcal{M}_\Lambda^{(\tau)}$, obtained by restricting*

$$\Phi^m \otimes \mathrm{Id}\colon \mathcal{M}_\Lambda \to \mathcal{M}^{(\sigma^m)} \times_{\mathrm{Spf}(W)} \mathrm{Spf}(\Lambda) \cong \mathcal{M}_\Lambda^{(\tau)}$$

to $\mathcal{N}$, factors through $\mathcal{N}^{(\tau)} \hookrightarrow \mathcal{M}_\Lambda^{(\tau)} = (\mathcal{M}^{(\sigma^m)})_\Lambda$.

Proof. Assume $\mathcal{N}$ has property (c). For the implication (c) $\Rightarrow$ (a) it suffices to show that $\mathcal{N}_{\overline{\Lambda}}$ is a formal subtorus of $\mathcal{M}_{\overline{\Lambda}}$. We know that

$$\mathcal{M}_{\overline{\Lambda}} \cong (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}})^d = \mathrm{Spf}(\overline{\Lambda}[\![q_1 - 1, \ldots, q_d - 1]\!])$$

for some $d \geq 0$, with Φ^m given by $q_i \mapsto q_i^{p^m}$.

Since $\mathcal{N}$ is formally smooth over Λ we can take our coordinates such that the composition

$$\mathcal{N}_{\overline{\Lambda}} \hookrightarrow \widehat{\mathbb{G}}_{\mathrm{m}}^d \xrightarrow{\mathrm{pr}_e} \widehat{\mathbb{G}}_{\mathrm{m}}^e$$

is an isomorphism, where pr_e is the projection map onto the first e factors. Let $\alpha\colon \widehat{\mathbb{G}}_{\mathrm{m}}^e \to \mathcal{N}_{\overline{\Lambda}}$ be the inverse isomorphism. Property (c) means that we have a commutative diagram

$$\begin{array}{ccccc}
\mathcal{N}_{\overline{\Lambda}} & \xrightarrow{i} & (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}})^d & \xrightarrow{\mathrm{pr}_e} & (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}})^e \\
\downarrow{\scriptstyle \Phi^m|_{\mathcal{N}}} & & \downarrow{\scriptstyle \Phi^m} & & \downarrow{\scriptstyle \Phi^m} \\
\mathcal{N}_{\overline{\Lambda}}^{(\tau)} & \xrightarrow{i^{(\tau)}} & (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}}^{(\tau)})^d & \xrightarrow{\mathrm{pr}_e} & (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}}^{(\tau)})^e
\end{array}$$

Therefore, the morphism $i \circ \alpha\colon (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}})^e \hookrightarrow (\widehat{\mathbb{G}}_{\mathrm{m},\overline{\Lambda}})^d$ satisfies the relation $\Phi^m \circ (i \circ \alpha) = (i \circ \alpha)^{(\tau)} \circ \Phi^m$. From Remark 1.11 (which is also valid over $\overline{\Lambda}$) it then follows that $i \circ \alpha$ is a group homomorphism, and this proves that $\mathcal{N}$ is a formal subtorus of $\mathcal{M}_\Lambda$. Finally, the implications (a) $\Rightarrow$ (b) $\Rightarrow$ (c) are clear. □

3.3 Remark. It turns out that the condition that $\mathcal{N}$ is "compatible" with Φ is very strong. In [11] it is shown that the above lemma is also true if we replace the assumption that $\mathcal{N}$ is formally smooth over Λ by the weaker condition that $\mathcal{N}$ is flat over Λ. We will use this stronger result, which allows us to formulate the main results of this chapter without restrictive smoothness conditions.

3.4 Lemma. *Let $k, W, \mathcal{M}, \Phi, \Lambda, m$ and τ be as in the previous lemma, and let $\mathcal{N}$ be an irreducible closed formal subscheme of $\mathcal{M}_\Lambda$ which is flat over Λ. The following properties are equivalent.*

(a) *$\mathcal{N} = \mathfrak{T}\cdot\mathcal{N}'$ is the translate of a formal subtorus $\mathcal{N}' \subseteq \mathcal{M}_\Lambda$ over an irreducible closed formal subscheme $\mathfrak{T}$, flat over Λ, contained in the p^n-torsion subgroup $\mathcal{M}_\Lambda[p^n]$ for some $n \geq 0$.*

(b) *There is a (generally non-primitive) Galois submodule $Y \subseteq X^*(\mathcal{M})$ such that $\mathcal{N}$ is an irreducible component of $\mathcal{N}(Y)_\Lambda$.*

(c) *There are integers $k, l \geq 1$ such that the morphism $\Phi^{(k+l)m}|_{\mathcal{N}}\colon \mathcal{N} \to \mathcal{M}_\Lambda^{(\tau^{k+l})}$ factors through $(\Phi^{km}|_{\mathcal{N}})^{(\tau^l)}\colon \mathcal{N}^{(\tau^l)} \to \mathcal{M}_\Lambda^{(\tau^{k+l})}$.*

Proof. Again the implications (a) $\Rightarrow$ (b) $\Rightarrow$ (c) are clear. Assume that (c) holds, and let $\mathcal{N}'$ be the image of $\Phi^{lm}|_{\mathcal{N}}$. Then $\mathcal{N}'$ is mapped into $(\mathcal{N}')^{(\tau^k)}$ under $\Phi^{km}\colon \mathcal{M}_\Lambda^{(\tau^l)} \to \mathcal{M}_\Lambda^{(\tau^{k+l})}$. By the previous lemma (here we use the stronger version of the lemma which is proved in [11]) it follows that $\mathcal{N}'$ is a formal subtorus of $\mathcal{M}^{(\tau^l)}$, compatible with Φ^m. From the description of Φ^m over $\overline{\Lambda}$ it readily follows that

$$(\phi^{lm})^{-1}(\mathcal{N}'_{\overline{\Lambda}}) \subseteq \mathcal{M}_{\overline{\Lambda}}[p^{lm}] \cdot \mathcal{N}'_{\overline{\Lambda}},$$

hence also

$$\mathcal{N} \subseteq (\phi^{lm})^{-1}(\mathcal{N}') \subseteq \mathcal{M}_{\Lambda}[p^{lm}] \cdot \mathcal{N}'.$$

Because we assumed $\mathcal{N}$ to be irreducible we conclude that (a) holds. □

3.5 At this point, let us set up the situation that we will study the rest of this chapter. As before, $\mathsf{A}_g = \mathsf{A}_{g,1,n}$ ($n \geq 3$) denotes the fine moduli scheme over $\mathrm{Spec}(\mathbb{Z}[1/n])$ of principally polarized g-dimensional abelian varieties with a Jacobi level n structure. We consider a closed, absolutely irreducible algebraic subvariety Z of $\mathsf{A}_g \otimes F$, where F is a number field.

Next we introduce models in mixed characteristic. So, let $\mathfrak{p}$ be a finite prime of F with residue field κ of characteristic $p > 0$, with $p \nmid n$. Write $\mathcal{A}_g = \mathsf{A}_{g,1,n} \otimes \widehat{\mathcal{O}}_{\mathfrak{p}}$, and define $\mathcal{Z} \hookrightarrow \mathcal{A}_g$ as the Zariski closure of Z inside $\mathcal{A}_g$. We write $\widehat{\mathcal{Z}} \hookrightarrow \widehat{\mathcal{A}}_g$ for the formal completion along the ordinary locus in characteristic p, and for a closed ordinary point $x \in (\mathcal{Z} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$, let $\mathfrak{Z}_x \hookrightarrow \mathfrak{A}_x$ over $\mathfrak{S} = \mathrm{Spf}(\Lambda)$ (with $\Lambda = W(\kappa(x)) \otimes_{W(\kappa(\mathfrak{p}))} \widehat{\mathcal{O}}_{\mathfrak{p}}$) be the formal completion at x. Notice that here $\widehat{\mathcal{A}}_g$ and $\mathfrak{A}_x$ are defined as formal completions of the scheme $\mathsf{A}_{g,1,n} \otimes \widehat{\mathcal{O}}_{\mathfrak{p}}$, cf. 1.19.

If we want to indicate the dependence on $\mathfrak{p}$, then we use notations such as $\mathcal{Z}_{\mathfrak{p}}$ and $\mathfrak{Z}_{\mathfrak{p},x}$.

3.6 Definition. We say that $\mathcal{Z}$ is formally linear at the closed point $x \in (\mathcal{Z} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$ if $\mathfrak{Z}_x$ is a formal subtorus of $\mathfrak{A}_x$. If all irreducible components of $\mathfrak{Z}_x$ have the properties described in Lemma 3.4 (with respect to $\mathfrak{A}_x$ and Φ_{can}) then we say that $\mathcal{Z}$ is formally quasi-linear at x.

If $\mathfrak{Z}_x$ has at least one irreducible component which is a formal subtorus of $\mathfrak{A}_x$ (respectively the translate of a formal subtorus over a torsion point) then we say that $\mathcal{Z}$ has formally linear (respectively formally quasi-linear) components at x.

3.7 Definition. Let X be an abelian variety of CM-type, defined over a number field K. If $\mathfrak{p}$ is a finite prime of K then we say that X is canonical at $\mathfrak{p}$ if there exists an abelian scheme $\mathcal{X}_{\mathfrak{p}}$ over $\mathrm{Spec}(\mathcal{O}_{K,\mathfrak{p}})$ with generic fibre X and ordinary special fibre

$\mathcal{X}_\mathfrak{p} \otimes \kappa(\mathfrak{p})$, such that $\mathcal{X}_\mathfrak{p}$ is the canonical lifting of $\mathcal{X}_\mathfrak{p} \otimes \kappa(\mathfrak{p})$. We say that a CM-point $t \in \mathsf{A}_{g,1,n}(K)$ is canonical at $\mathfrak{p}$ if the corresponding abelian variety has this property.

3.8 Suppose that $\mathcal{Z}$ has formally quasi-linear components at the closed ordinary point x. Let $\mathcal{R}$ be a complete local algebra which is finite and flat over Λ, and let $\hat{t} \in \mathfrak{Z}_x(\mathcal{R})$ be a torsion point. The formal abelian scheme over $\mathrm{Spf}(\mathcal{R})$ corresponding to $\hat{t}$ is algebraizable, so we get an abelian scheme $\mathcal{X} = \mathcal{X}_t$ over $\mathrm{Spec}(\mathcal{R})$, and $\hat{t}$ extends to a section $t \in \mathcal{Z}(\mathcal{R})$. It follows from Lemma 1.15 that $\mathcal{X}$ is of CM-type.

Let $\widehat{\mathcal{T}}$ be the collection of all points $\hat{t} \in \mathfrak{Z}_x(\mathcal{R})$, where $\mathcal{R}$ ranges over all complete local domains, finite and flat over Λ, and $\hat{t}$ is a torsion point of $\mathfrak{Z}_x(\mathcal{R})$. Let $\mathcal{T}$ be the collection of corresponding points $t \in \mathcal{Z}(\mathcal{R})$. We claim that $\mathcal{T}$ is Zariski dense in $\mathcal{Z}$. To see this, write $\mathcal{Z}' \subseteq \mathcal{Z}$ for the Zariski closure of $\mathcal{T}$. By assumption, there is an irreducible component $\mathfrak{C} \subseteq \mathfrak{Z}_x$ which is the translate of a formal subtorus of $\mathfrak{A}_x$ over a torsion point. From the definition of the set $\mathcal{T}$ we see that $\mathfrak{C}$ is contained in the formal completion of $\mathcal{Z}'$ at x. The claim follows by a dimension argument: $\mathcal{Z}'$ and $\mathcal{Z}$ are flat over $\mathrm{Spec}(\widehat{\mathcal{O}}_\mathfrak{p})$ of relative dimensions $d' \leq d$. Then the closed fibres $\mathcal{Z}' \otimes \kappa(\mathfrak{p})$ and $\mathcal{Z} \otimes \kappa(\mathfrak{p})$ are equidimensional of dimension d' and d respectively, and $\mathfrak{C} \subseteq (\mathcal{Z}')_{/\{x\}}$ implies that $d' = d$. Since $Z \otimes Q(\widehat{\mathcal{O}}_\mathfrak{p})$ is irreducible, the generic fibre of $\mathcal{Z}'$ is equal to $Z \otimes Q(\widehat{\mathcal{O}}_\mathfrak{p})$, and by definition of $\mathcal{Z}$ this implies that $\mathcal{Z}' = \mathcal{Z}$.

Let $Q = Q(\mathcal{R})$ denote the quotient field of a complete domain $\mathcal{R}$ as above, then we have a collection T of CM-points $t \in Z(Q)$, corresponding to the characteristic zero fibres X_t of the abelian schemes $\mathcal{X}_t$. From the fact that $\mathcal{T}$ is Zariski dense in $\mathcal{Z}$ it follows that T is dense in Z. Notice that the abelian varieties X_t are all p-isogenous, i.e., given two torsion points $\hat{t} \in \mathfrak{Z}_x(\mathcal{R})$ and $\hat{t}' \in \mathfrak{Z}_x(\mathcal{R}')$ in the collection $\widehat{\mathcal{T}}$, then over a common field extension of $Q(\mathcal{R})$ and $Q(\mathcal{R}')$ the abelian varieties X_t and $X_{t'}$ are isogenous via an isogeny whose degree is a power of p. This is because X_t and $X_{t'}$ are CM-liftings of the same ordinary abelian variety in characteristic p.

Choose one of the points $t \in T$, and consider the corresponding abelian variety X_t. As X_t is of CM-type, it is defined over some number field $K_t \supseteq F$, which we take large enough so that all endomorphisms of $X_t \otimes \overline{K}_t$ are defined over K_t. The endomorphism ring $\mathrm{End}(X_t)$ is an order in $\mathrm{End}^0(X_t)$. It has a well-determined index in a maximal order of $\mathrm{End}^0(X_t)$, which we call the conductor of $\mathrm{End}(X_t)$, and which we denote by $\mathfrak{f}(X_t)$.

Now choose a prime number $\ell \neq p$, with the following properties:

1. ℓ does not divide the conductor $\mathfrak{f}(X_t)$, i.e., $\mathrm{End}(X_t)$ is "maximal at ℓ",
2. the prime ℓ splits completely in the endomorphism algebra $\mathrm{End}^0(X_t)$, i.e., $\mathrm{End}^0(X_t) \otimes \mathbb{Q}_\ell$ is a product of algebras $\mathrm{M}_m(\mathbb{Q}_\ell)$.

Possibly after first replacing K_t by a finite extension, X_t has good reduction $X_{t,\mathfrak{l}}$ at all primes $\mathfrak{l}$ above ℓ. The fact that ℓ splits completely in $\mathrm{End}^0(X_t)$ implies that the reduction is ordinary (using [64, Lemme 5]). By Lemma 1.15, X_t is isogenous to the canonical lifting of $X_{t,\mathfrak{l}}$, so $\mathrm{End}^0(X_t) = \mathrm{End}^0(X_{t,\mathfrak{l}})$. The conductors of the endomorphism rings can only differ by an ℓ-power, see [49, Lemma 2.1], and it then follows from the first condition on ℓ that $\mathrm{End}(X_t) \cong \mathrm{End}(X_{t,\mathfrak{l}})$. We conclude that X_t is canonical at all primes of K_t above ℓ.

As remarked above, all abelian varieties X_t with $t \in T$ are p-isogenous. Therefore, the conditions on ℓ do not depend on the chosen t, and our conclusion is valid for all X_t simultaneously. This shows that if $\mathcal{Z}$ is formally quasi-linear at an ordinary point x in characteristic p, then there is a different characteristic ℓ and a Zariski dense collection T of CM-points $t \in Z(K_t)$ such that each X_t is canonical at all primes $\mathfrak{l}$ of K_t above ℓ.

Conversely, we will now show that this last property implies that $\mathcal{Z}$ has formally linear components at some of its ordinary points in prime characteristic. First we need two lemmas.

3.9 Lemma. *Let $\widehat{\mathcal{O}}$ be a complete discrete valuation ring, and write $\mathfrak{S} = \mathrm{Spf}(\widehat{\mathcal{O}})$. Let $f\colon \mathfrak{X} \to \mathfrak{S}$ and $g\colon \mathfrak{Y} \to \mathfrak{S}$ be $\mathfrak{S}$-formal schemes which are formally reduced, noetherian and adic, flat and of finite type over $\mathfrak{S}$. Suppose the associated reduced schemes $\mathfrak{X}_{\mathrm{red}}$ and $\mathfrak{Y}_{\mathrm{red}}$ are equidimensional, with $\dim(\mathfrak{X}_{\mathrm{red}}) = \dim(\mathfrak{Y}_{\mathrm{red}}) = d$. Let $p\colon \mathfrak{X} \to \mathfrak{Y}$ be a finite $\mathfrak{S}$-morphism. Then for every irreducible component $\mathfrak{C} \subseteq \mathfrak{X}$, the image $p(\mathfrak{C}) \subseteq \mathfrak{Y}$ (in the sense of formal schemes) is an irreducible component of $\mathfrak{Y}$.*

Proof. Replacing $\mathfrak{X}$ by the irreducible component $\mathfrak{C}$ and $\mathfrak{Y}$ by an irreducible component containing $p(\mathfrak{C})$, we may assume $\mathfrak{X}$ and $\mathfrak{Y}$ to be irreducible. The image $p(\mathfrak{X}) \subseteq \mathfrak{Y}$ is defined by the coherent sheaf of ideals $\mathcal{J} = \mathcal{K}er(\mathcal{O}_{\mathfrak{Y}} \to p_*\mathcal{O}_{\mathfrak{X}})$. Let $\mathcal{A} = \mathcal{A}nn_{\mathcal{O}_{\mathfrak{Y}}}(\mathcal{J})$ be the annihilator sheaf of the $\mathcal{O}_{\mathfrak{Y}}$-module $\mathcal{J}$, which is a coherent ideal by [27, 0_{I}, 5.3.10]. Write $\mathfrak{A} \subseteq \mathfrak{Y}$ for the closed formal subscheme defined by $\mathcal{A}$. The assumption that $\mathfrak{Y}$ is formally reduced implies that $\mathfrak{Y} = p(\mathfrak{X}) \cup \mathfrak{A}$. In fact, this immediately follows from the remark that in a reduced ring R we have $I \cap \mathrm{Ann}_R(I) = (0)$ for all ideals I.

Suppose $p(\mathfrak{X}) \neq \mathfrak{Y}$, then there exists an affine open formal subscheme $\mathfrak{U} = \mathrm{Spf}(A) \subseteq \mathfrak{Y}$ such that $\mathcal{J}(\mathfrak{U}) = J \subseteq A$ is not the zero ideal. We are done if we show that $\mathcal{A}(\mathfrak{U}) = \mathrm{Ann}_A(J)$ is also not zero, for this would contradict the assumption that $\mathfrak{Y}$ is irreducible.

The preimage $p^{-1}(\mathfrak{U})$ is again affine (p being finite), say $p^{-1}(\mathfrak{U}) = \mathrm{Spf}(B)$, where $A \to B$ is continuous and finite. By assumption, A and B are flat over $\widehat{\mathcal{O}}$. We claim that A and B are equidimensional of dimension $d+1$. To see this, we argue as follows.

Let $\pi \in \widehat{\mathcal{O}}$ be a uniformizing element, then $g^*(\pi)\mathcal{O}_{\mathfrak{Y}}$ is a defining ideal of $\mathfrak{Y}$ (because g is of finite type, hence adic). It follows that $\mathfrak{U}_{\mathrm{red}} = \mathrm{Spec}(A/\mathrm{rad}(\pi))$. Let $\mathfrak{m} \subseteq A$ be a maximal ideal. Using the flatness of A over $\widehat{\mathcal{O}}$, the equidimensionality of $\mathfrak{Y}_{\mathrm{red}}$ and [34, Theorem 15.1] we see that

$$\dim(A_{\mathfrak{m}}) = \dim(\widehat{\mathcal{O}}) + \dim(A_{\mathfrak{m}} \otimes_{\widehat{\mathcal{O}}} \widehat{\mathcal{O}}/(\pi)) = 1 + \dim((A/\mathrm{rad}(\pi))_{\mathfrak{m}}) = 1 + d\,.$$

For the ring B, the argument is similar.

We conclude that the morphism $\mathrm{Spec}(B) \to \mathrm{Spec}(A)$ is a finite morphism between (reduced) noetherian schemes, both of equidimension $d+1$. The image, which is $\mathrm{Spec}(A/I) \subseteq \mathrm{Spec}(A)$, is a union of irreducible components of $\mathrm{Spec}(A)$. Therefore, if $\mathfrak{q}_1, \ldots, \mathfrak{q}_k$ are the minimal prime ideals of A (taken in a suitable order), we have $I = \mathfrak{q}_1 \cap \ldots \cap \mathfrak{q}_r$. The assumption that $I \neq (0) = \mathrm{nil}(A)$ implies that $r < k$. Therefore, $\mathrm{Ann}_A(J) = \mathfrak{q}_{r+1} \cap \ldots \cap \mathfrak{q}_k \neq (0)$, and this is what we wanted to prove. □

3.10 Lemma. *Let $\widehat{\mathcal{O}}$ be a complete discrete valuation ring, and write $S = \mathrm{Spec}(\widehat{\mathcal{O}})$, $\mathfrak{S} = \mathrm{Spf}(\widehat{\mathcal{O}})$. Let $f\colon X \to S$ be an S-scheme, flat and of finite type over S. Write $\hat{f}\colon \mathfrak{X} \to \mathfrak{S}$ for the formal completion of X along its closed fibre. Let $\mathcal{R}$ be a complete local domain which is finite and flat over $\widehat{\mathcal{O}}$, and let $t\colon \mathrm{Spec}(\mathcal{R}) \to X$ be an S-morphism. Write $\hat{t}\colon \mathrm{Spf}(\mathcal{R}) \to \mathfrak{X}$ for the induced $\mathfrak{S}$-morphism. Assume t maps the generic point of $\mathrm{Spec}(\mathcal{R})$ into the regular locus of X. Then there is a unique irreducible component $\mathfrak{C} \subseteq \mathfrak{X}$ such that $\hat{t}$ factors through $\mathfrak{C}$.*

Proof. Let $U = \mathrm{Spec}(A) \subseteq X$ be an affine open subscheme such that t factors through U. Write $\mathfrak{p} \subset A$ for the kernel of $t^*\colon A \to \mathcal{R}$. The assumption that t maps the generic point of $\mathrm{Spec}(\mathcal{R})$ to the regular locus of X implies that the local ring $A_{\mathfrak{p}}$ is regular. Let $\pi \in \widehat{\mathcal{O}}$ be a uniformizing element, and let $\widehat{A}$ denote the completion of A for the (π)-adic topology.

Since $\mathcal{R}$ is a finite module over $\widehat{\mathcal{O}}$, this is also the case for $A/\mathfrak{p} \subseteq \mathcal{R}$, which therefore is complete and separated for the (π)-adic topology ([34, Theorem 8.7]). It follows that $A/\mathfrak{p} \xrightarrow{\sim} \widehat{A}/\mathfrak{p}\widehat{A}$, and we have a flat homomorphism of local rings $A_{\mathfrak{p}} \to \widehat{A}_{\mathfrak{p}\widehat{A}}$. Write $\mathfrak{m} = \mathfrak{p}A_{\mathfrak{p}}$ and $\mathfrak{n} = \mathfrak{p}\widehat{A}_{\mathfrak{p}\widehat{A}}$ for the maximal ideals. Clearly, $\mathfrak{n}$ is generated by the image of $\mathfrak{m}$, and $(A_{\mathfrak{p}})/\mathfrak{m} \cong \kappa(\mathfrak{p}) \cong (\widehat{A}_{\mathfrak{p}\widehat{A}})/\mathfrak{n}$. Applying [34, Theorem 15.1] we see that $\dim(A_{\mathfrak{p}}) = \dim(\widehat{A}_{\mathfrak{p}\widehat{A}})$, hence

$$\dim_{\kappa(\mathfrak{p})} \mathfrak{n}/\mathfrak{n}^2 \leq \dim_{\kappa(\mathfrak{p})} \mathfrak{m}/\mathfrak{m}^2 = \dim(A_{\mathfrak{p}}) = \dim(\widehat{A}_{\mathfrak{p}\widehat{A}})\,.$$

We conclude that $\widehat{A}_{\mathfrak{p}\widehat{A}}$ is a regular local ring.

Let $\mathfrak{q}_1, \ldots, \mathfrak{q}_k$ be the minimal prime ideals of $\widehat{A}$. If $\hat{t}$ factors through more than one irreducible component of $\mathfrak{X}$, then there are at least two minimal primes $\mathfrak{q}_i \neq \mathfrak{q}_j$ which are contained in $\mathfrak{p}\widehat{A}$. Since the prime ideals in $\widehat{A}_{\mathfrak{p}\widehat{A}}$ are in bijective, inclusion-preserving correspondence with the prime ideals of $\widehat{A}$ contained in $\mathfrak{p}\widehat{A}$, it follows that there is more than one minimal prime in $\widehat{A}_{\mathfrak{p}\widehat{A}}$. This contradicts the regularity of this ring. □

3.11 Proposition. *Let Z, $\mathfrak{p}$, $\mathcal{Z}$ etc. be as in 3.5. Suppose there is a collection T of CM-points $t \in Z(K_t)$ (K_t a number field containing F) which is Zariski dense in Z (over F). Also suppose that each X_t is canonical at some prime $\mathfrak{q}$ of K_t above $\mathfrak{p}$. Then there is a non-empty union U of irreducible components of $(\mathcal{Z} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$ such that Z has formally linear components at all closed points $x \in U$.*

Proof. For each t, choose a prime $\mathfrak{q}$ of K_t above $\mathfrak{p}$ such that X_t is canonical at $\mathfrak{q}$. Write $\mathcal{R}_{\mathfrak{q}} = W(\kappa(\mathfrak{q})) \otimes_{W(\kappa(\mathfrak{p}))} \widehat{\mathcal{O}}_{\mathfrak{p}}$, then X_t gives rise to an $\mathcal{R}_{\mathfrak{q}}$-valued point $t_{\mathfrak{q}}$: $\mathrm{Spec}(\mathcal{R}_{\mathfrak{q}}) \to \mathcal{Z}$, corresponding to the abelian scheme $\mathcal{X}_{t,\mathfrak{q}}$ over $\mathcal{R}_{\mathfrak{q}}$.

Let $N = {}^{p}\!\log(\#\kappa(\mathfrak{p}))$. The automorphism σ^N of $W(\kappa(\mathfrak{q}))$ lifts to an automorphism τ of $\mathcal{R}_{\mathfrak{q}}$. Since $\mathcal{X}_{t,\mathfrak{q}}$ is the canonical lifting of $\mathcal{X}_{t,\mathfrak{q}} \otimes \kappa(\mathfrak{q})$, the morphism

$$\mathrm{Frob}^N\colon \mathcal{X}_{t,\mathfrak{q}} \otimes \kappa(\mathfrak{q}) \to (\mathcal{X}_{t,\mathfrak{q}} \otimes \kappa(\mathfrak{q}))^{(p^N)}$$

lifts to a morphism F_t: $\mathcal{X}_{t,\mathfrak{q}} \to \mathcal{X}_{t,\mathfrak{q}}^{(\tau)}$ over $\mathrm{Spec}(\mathcal{R}_{\mathfrak{q}})$. We consider this as a point $F_t \in \mathcal{I}sog^{\mathrm{o}}(\mathcal{R}_{\mathfrak{q}})$, where $\mathcal{I}sog = \mathcal{I}sog(p^{Ng})$. Define $\mathcal{Y} \subseteq \mathcal{I}sog^{\mathrm{o}}$ as the Zariski closure over $\mathrm{Spec}(\widehat{\mathcal{O}}_{\mathfrak{p}})$ of these points.

For the projection maps pr_i: $\mathcal{I}sog^{\mathrm{o}} \to \mathcal{A}_g^{\mathrm{o}}$ we have $\mathrm{pr}_1(F_t) = t_{\mathfrak{q}}$ and $\mathrm{pr}_2(F_t) = t_{\mathfrak{q}} \circ \tau$, and since the points $t_{\mathfrak{q}}$ (hence also the points $t_{\mathfrak{q}} \circ \tau$) are Zariski dense in $\mathcal{Z}^{\mathrm{o}} = \mathcal{Z} \cap \mathcal{A}_g^{\mathrm{o}}$, it follows that pr_i ($i \in \{1, 2\}$) restricts to a finite surjective morphism pr_i: $\mathcal{Y} \to \mathcal{Z}^{\mathrm{o}}$. Possibly after replacing the collection of CM-points T by a subcollection T' (which is still dense in Z), and replacing $\mathcal{Y}$ by the Zariski closure of the points F_t with $t \in T'$, we may assume that, moreover, every irreducible component of $\mathcal{Y}$ maps surjectively to an irreducible component of $\mathcal{Z}^{\mathrm{o}}$.

Write $k = \kappa(\mathfrak{p})$. By construction, $\mathcal{Y}$ and $\mathcal{Z}^{\mathrm{o}}$ are flat over $\mathrm{Spec}(\widehat{\mathcal{O}}_{\mathfrak{p}})$ and, as remarked above, the projections pr_i are finite. The closed fibres $\mathcal{Y}_k$ and $\mathcal{Z}^{\mathrm{o}}_k$ are therefore equidimensional of the same dimension, and every irreducible component of $\mathcal{Y}_k$ maps surjectively to an irreducible component of $\mathcal{Z}^{\mathrm{o}}_k$.

We have seen before (§2) that there is a disjoint union of irreducible components $\mathcal{I}_k \subset \mathcal{I}sog^{\mathrm{o}}_k$ classifying the Nth power of Frobenius. Then $\mathcal{Y}_k = \mathcal{Y}'_k \sqcup \mathcal{Y}''_k$, where $\mathcal{Y}'_k$ and $\mathcal{Y}''_k$ are unions of connected components of $\mathcal{Y}_k$, chosen such that $\mathcal{Y}'_k \subseteq \mathcal{I}_k$ and $\mathcal{Y}''_k \cap \mathcal{I}_k = \emptyset$. Now $\mathrm{pr}_{1|\mathcal{I}}\colon \mathcal{I}_k \to \mathcal{A}^{\mathrm{o}}_g \otimes k$ is an isomorphism and (by our choice of N) the composition $\mathrm{pr}_2 \circ (\mathrm{pr}_{1|\mathcal{I}})^{-1}\colon \mathcal{A}^{\mathrm{o}}_g \otimes k \to \mathcal{A}^{\mathrm{o}}_g \otimes k$ is the identity on the underlying topological space. The image of $\mathcal{Y}'_k$ under both projections to $\mathcal{Z}^{\mathrm{o}}_k$ is therefore the same; call it $\mathcal{Z}^{\mathrm{o}\prime}_k \subseteq \mathcal{Z}^{\mathrm{o}}_k$. It is a union of irreducible components of $\mathcal{Z}^{\mathrm{o}}_k$, which is non-empty because the special fibre of every F_t factors through $\mathcal{Y}'_k$.

Next we look at formal completions. Write $\widehat{\mathcal{Y}}$, $\widehat{\mathcal{Y}}'$ and $\widehat{\mathcal{Z}}$ for the formal completions of $\mathcal{Y}$, $\mathcal{Y}$, $\mathcal{Z}$ along $\mathcal{Y}_k$, $\mathcal{Y}'_k$ and $\mathcal{Z}^{\mathrm{o}}_k$ respectively. Notice that these formal schemes are formally reduced, noetherian and adic, flat and of finite type over $\mathrm{Spf}(\widehat{\mathcal{O}}_{\mathfrak{p}})$, since $\mathcal{Y}$ and $\mathcal{Z}^{\mathrm{o}}$ are excellent schemes (being of finite type over a complete local ring) with the corresponding properties.

Since $\mathcal{Y}'_k$ and $\mathcal{Y}''_k$ are disjoint, $\widehat{\mathcal{Y}}'$ is an open and closed formal subscheme of $\widehat{\mathcal{Y}}$. Let $\widehat{\mathcal{Z}} = \cup_{\alpha\in A}\widehat{\mathcal{Z}}_\alpha$ be the decomposition of $\widehat{\mathcal{Z}}$ into irreducible components. The projections $\widehat{\mathrm{pr}}_i\colon \widehat{\mathcal{Y}}' \to \widehat{\mathcal{Z}}$ are finite (using [27, III, Corollaire 4.8.4]), so by Lemma 3.9 we have

$$\widehat{\mathrm{pr}}_1\colon \widehat{\mathcal{Y}}' \xrightarrow{\sim} \cup_{\alpha\in A_1}\widehat{\mathcal{Z}}_\alpha, \qquad \widehat{\mathrm{pr}}_2(\widehat{\mathcal{Y}}') = \cup_{\alpha\in A_2}\widehat{\mathcal{Z}}_\alpha$$

for some $A_1 \subseteq A$, $A_2 \subseteq A$. Proposition 2.3 shows that the composition

$$\widehat{\mathrm{pr}}_2 \circ (\widehat{\mathrm{pr}}_1)^{-1}\colon \cup_{\alpha\in A_1}\widehat{\mathcal{Z}}_\alpha \to \cup_{\alpha\in A_2}\widehat{\mathcal{Z}}_\alpha$$

is the restriction of Φ^N_{can}.

At this point we apply Lemma 3.10. We take $t \in T$ (the subcollection with which we replaced the original T) corresponding to a point in the regular locus of Z. This is certainly possible, since the collection T is Zariski dense in Z. We conclude that there is a unique component $\widehat{\mathcal{Z}}_{\alpha(t)} \subseteq \widehat{\mathcal{Z}}$ with $\alpha(t) \in A_1$ such that $\hat{t}_{\mathfrak{q}}\colon \mathrm{Spf}(\mathcal{R}_{\mathfrak{q}}) \to \widehat{\mathcal{Z}}$ factors through $\widehat{\mathcal{Z}}_{\alpha(t)}$. Let $\widehat{\mathcal{Y}}_\alpha(t)$ be the unique irreducible component with $\widehat{\mathrm{pr}}_1\colon \widehat{\mathcal{Y}}_{\alpha(t)} \xrightarrow{\sim} \widehat{\mathcal{Z}}_{\alpha(t)}$. Since $t_{\mathfrak{q}} = \mathrm{pr}_1 \circ F_t$, the section $\hat{F}_t\colon \mathrm{Spf}(\mathcal{R}_{\mathfrak{q}}) \to \widehat{\mathcal{Y}}'$ factors through $\widehat{\mathcal{Y}}_{\alpha(t)}$. The image of $\widehat{\mathcal{Y}}_{\alpha(t)}$ under $\widehat{\mathrm{pr}}_2$ is some irreducible component $\widehat{\mathcal{Z}}_{\alpha'(t)}$ through which $\widehat{\mathrm{pr}}_2 \circ \hat{F}_t$ factors. But $\widehat{\mathrm{pr}}_2 \circ \hat{F}_t = \hat{\tau} \circ \hat{t}_{\mathfrak{q}}$: $\mathrm{Spf}(\mathcal{R}_{\mathfrak{q}}) \to \widehat{\mathcal{Z}}$ and, τ being an automorphism of $\mathcal{R}$, we see that $\hat{t}_{\mathfrak{q}}$ factors through $\widehat{\mathcal{Z}}_{\alpha'(t)}$. By assumption we have $\widehat{\mathcal{Z}}_{\alpha(t)} = \widehat{\mathcal{Z}}_{\alpha'(t)}$.

Let x be a closed point on the component $\widehat{\mathcal{Z}}_{\alpha(t)}$. The formal completion of $\widehat{\mathcal{Z}}_{\alpha(t)}$ at x is the union of a number of irreducible components, say $\mathfrak{C}_1, \ldots, \mathfrak{C}_r$ of $\mathfrak{Z}_x$. If m is a suitable multiple of both N and ${}^p\!\log(\#\kappa(x))$ then Φ^m_{can} induces a finite morphism $\Phi^m_{\mathrm{can}}\colon \mathfrak{A}_x \to \mathfrak{A}_x$ of formal schemes, and it follows from the above that this maps $\cup_j \mathfrak{C}_j \subseteq \mathfrak{A}_x$ onto itself. Then Φ^m_{can} acts by permutations on the set $\{\mathfrak{C}_1, \ldots, \mathfrak{C}_r\}$, so after replacing m by a suitable multiple it preserves all $\mathfrak{C}_j$. By Lemma 3.2 these irreducible components $\mathfrak{C}_j \subseteq \mathfrak{Z}_x$ are therefore formal subtori of $\mathfrak{A}_x$. □

3.12 Corollary. *Let Z, $\mathfrak{p}$, $\mathcal{Z} = \mathcal{Z}_{\mathfrak{p}}$ be as in 3.5 and suppose $\mathcal{Z}_{\mathfrak{p}}$ has formally quasi-linear components at some closed ordinary point x. Then there exist infinitely many primes $\mathfrak{l}$ of $\mathcal{O}_F$ such that the model $\mathcal{Z}_{\mathfrak{l}}$ of Z over $\mathrm{Spec}(\widehat{\mathcal{O}}_{\mathfrak{l}})$ is formally linear at all closed points y in a non-empty open subset of $(\mathcal{Z}_{\mathfrak{l}} \otimes \kappa(\mathfrak{l}))^{\mathrm{o}}$.*

Proof. We start as in 3.8. We have seen that for primes ℓ satisfying conditions 1. and 2. on p. 56, there is a Zariski dense collection T of CM-points $t \in Z(K_t)$ such that each X_t is canonical at all primes of K_t above ℓ.

We consider primes $\mathfrak{l}$ of $\mathcal{O}_F$ such that the residue characteristic ℓ satisfies these conditions 1. and 2., and such that no irreducible component of $\mathcal{Z}_{\mathfrak{l}} \otimes \kappa(\mathfrak{l})$ is contained in the singular locus of $\mathcal{Z}_{\mathfrak{l}}$. This last condition excludes only finitely many primes $\mathfrak{l}$. The model $\mathcal{Z}_{\mathfrak{l}}$ being an excellent scheme, it follows that for generic $y \in \mathcal{Z}_{\mathfrak{l}} \otimes \kappa(\mathfrak{l})$, the completed local ring $\mathcal{O}^{\wedge}_y$ of $\mathcal{Z}_{\mathfrak{l}}$ at y is regular. The corollary now results from the previous proposition and the remark that $(\mathcal{Z}_{\mathfrak{l}})_{/\{y\}}$ is irreducible if $\mathcal{O}^{\wedge}_y$ is regular. □

3.13 Remark. A posteriori we will get a much stronger conclusion, see 5.3

§4 A theorem of Noot

4.1 Our main motivation for introducing the concept of formal linearity is its relation to Shimura varieties, or rather (using the terminology introduced in Chapter I) to subvarieties of $\mathsf{A}_{g,1,n}$ of Hodge type. The first main result in this direction was established in the PhD thesis of Noot ([43], see also [44]). Roughly speaking it says that subvarieties of Shimura type are “almost everywhere” formally linear. In this section we give a precise formulation of Noot's theorem and we deduce some linearity properties of subvarieties of Hodge and of Shimura type.

4.2 Let (M, X_M) be a Shimura datum giving rise to a Shimura variety of Hodge type. This means that there exists a closed immersion of Shimura data $i\colon (M, X_M) \hookrightarrow (\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ for some $g \geq 1$. We fix such a closed immersion. Let

$$K_n = \{g \in \mathrm{CSp}_{2g}(\hat{\mathbb{Z}}) \mid g \equiv 1 \pmod{n}\} \subset \mathrm{CSp}_{2g}(\mathbb{A}_f)\,,$$

and let $K'_n = i^{-1}(K_n) \subset M(\mathbb{A}_f)$. One can show that, for a given prime number p, there exists an integer n with $p \nmid n$ such that the morphism $i_{(K'_n, K_n)}\colon Sh_{K'_n}(M, X_M) \to Sh_{K_n}(\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})$ is a closed immersion ([44, Lemma 3.3]. Here $i_{(K'_n, K_n)}$ is defined over the reflex field $E(M, X_M)$ of the datum (M, X_M).

Fix p and choose n accordingly. Let F be a finite field extension of $E(M, X_M)$, and let $\mathfrak{p}$ be a prime of F lying over p. Let S be an irreducible component of $Sh_{K'_n}(M, X_M)_F$ (which, by definition, is therefore a subvariety of Shimura type of $Sh_{K_n}(\mathrm{CSp}_{2g}, \mathfrak{H}_g^{\pm})_F = \mathsf{A}_{g,1,n} \otimes F$), and let $\mathcal{S}$ be the model over $\mathrm{Spec}(\mathcal{O}_{\mathfrak{p}})$ obtained by taking the Zariski closure of S in $\mathcal{A}_g$. We assume n and F to be large enough such that S is non-singular and geometrically irreducible.

Finally, let $\overline{\kappa}$ be an algebraic closure of $\kappa(\mathfrak{p})$; write $W(\overline{\kappa})$ for its ring of Witt vectors. We write W^{abs} for the completion of the integral closure of $W(\overline{\kappa})$ in an algebraic closure of its field of fractions.

4.3 Theorem. (Noot, [43, Proposition 2.2.3] and [44, Theorem 3.7]) *(i) Using the above notations, suppose that $s \in \mathcal{S}(\mathcal{O}_{\mathfrak{p}})$ is an $\mathcal{O}_{\mathfrak{p}}$-valued point of $\mathcal{S}$ such that the corresponding abelian scheme $\mathcal{X}_s \to \mathrm{Spec}(\mathcal{O}_{\mathfrak{p}})$ has good and ordinary reduction $\mathcal{X}_s \otimes \kappa(\mathfrak{p})$. Let $x \in \mathcal{S}(\kappa(\mathfrak{p}))^{\mathrm{o}}$ be the closed point corresponding to $\mathcal{X}_s \otimes \kappa(\mathfrak{p})$, and let $\mathfrak{S}_x \hookrightarrow \mathfrak{A}_x$ be the formal completion of $\mathcal{S}$ at x. Since S is a regular scheme there is a unique irreducible component $\mathfrak{C} \subset \mathfrak{S}_x$ such that $\hat{s}\colon \mathrm{Spf}(\widehat{\mathcal{O}}_{\mathfrak{p}}) \to \mathfrak{S}_x$ is a section of $\mathfrak{C}$. Then $\mathfrak{C}\widehat{\otimes}_{\widehat{\mathcal{O}}_{\mathfrak{p}}} W^{\mathrm{abs}}$ is the translate of a formal subtorus of $\mathfrak{A}_x \widehat{\otimes} W^{\mathrm{abs}}$ over a torsion point.*

(ii) For F sufficiently large and p outside a finite collection of prime numbers, S is formally linear at all its non-singular points $x \in \mathcal{S}(\kappa(\mathfrak{p}))^{\mathrm{o}}$.

For a proof of this theorem we refer to [43] and [44].

4.4 Corollary. *(i) Let $S \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ be a subvariety of Shimura type, and let $\mathcal{S}_\mathfrak{p}$ be its Zariski closure inside $\mathsf{A}_{g,1,n} \otimes \widehat{\mathcal{O}}_\mathfrak{p}$. Let x be a closed point of $(\mathcal{S}_\mathfrak{p} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$. Then $\mathcal{S}_\mathfrak{p}$ is formally quasi-linear at x. For $\mathfrak{p}$ outside a finite set Σ of primes of $\mathcal{O}_F$, the formal completion $\mathfrak{S}_x$ of $\mathcal{S}_\mathfrak{p}$ at x is a union of formal subtori of $\mathfrak{A}_x$. If S is non-singular then we can choose the (finite) set Σ such that $\mathcal{S}_\mathfrak{p}$ for $\mathfrak{p} \notin \Sigma$ is formally linear at all closed ordinary points $x \in (\mathcal{S}_\mathfrak{p} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$.*

(ii) Let $S \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ be a subvariety of Hodge type, and let $\mathcal{S}_\mathfrak{p}$ and x be as in (i). For $\mathfrak{p}$ outside a finite set of primes of $\mathcal{O}_F$, the formal completion $\mathfrak{S}_x$ is a union of formal subtori of $\mathfrak{A}_x$.

Proof. (i) There is a Shimura datum (M, X_M) as in 4.2 such that S is an irreducible component of $Sh_{K'_n}(M, X_M)_F$. Let $F \subseteq F'$ be a finite field extension, $m \in \mathbb{Z}_{\geq 1}$ an integer with $p \nmid m$, and let $S' \hookrightarrow \mathsf{A}_{g,1,mn} \otimes F'$ be an irreducible component of $Sh_{K'_{mn}}(M, X_M)_{F'}$ which maps to $S \otimes F'$ under the morphism $Sh_{(K'_{mn}, K'_n)}$. Choose a prime $\mathfrak{p}'$ of $\mathcal{O}_{F'}$ with $\mathfrak{p}' \cap \mathcal{O}_F = \mathfrak{p}$, and write $\mathcal{S}'$ for the model of S' over $\mathcal{O}_{\mathfrak{p}'}$. The morphism $f = f_{(mn,n)}\colon \mathcal{A}_{g,1,mn} \to \mathcal{A}_{g,1,n}$ is a quotient morphism for the action of a finite group, and $\mathcal{S}'$ is an irreducible component of the pull-back $f^*(\mathcal{S} \otimes \mathcal{O}_{\mathfrak{p}'})$.

We claim that we can choose F' and m so large that Noot's theorem applies for all points $x' \in (\mathcal{S}' \otimes \kappa(\mathfrak{p}'))^{\mathrm{o}}$ in the preimage of x and all irreducible components of $\mathfrak{S}_{x'}$. This is what we do: first we choose m and F' so large that $S' \hookrightarrow \mathsf{A}_{g,1,mn} \otimes F'$ is non-singular, such that all points $x' \in S' \otimes \kappa(\mathfrak{p}')$ in the preimage of x are rational over $\kappa(\mathfrak{p}')$ and such that all irreducible components of $\mathfrak{S}'_{x'}$ over W^{abs} are defined over $\widehat{\mathcal{O}}_{\mathfrak{p}'}$. After further enlarging F' these properties are maintained and every irreducible component $\mathfrak{C} \subseteq \mathfrak{S}'_{x'}$ (which is flat over $\widehat{\mathcal{O}}_{\mathfrak{p}'}$) has a section $\hat{s}\colon \mathrm{Spf}(\widehat{\mathcal{O}}_{\mathfrak{p}'}) \to \mathfrak{C}$ which is obtained by completion from a section $s\colon \mathrm{Spec}(\mathcal{O}_{\mathfrak{p}'}) \to \mathsf{A}_{g,1,mn}$. Applying Noot's theorem, the first two statements of (i) now readily follow.

The last assertion of (i) then follows from the remark that if S is non-singular, then there exists a finite set Σ of primes such that $\mathcal{S}_\mathfrak{p} \otimes \kappa(\mathfrak{p})$ is non-singular for all $\mathfrak{p} \notin \Sigma$; here we use some results of [27, Ch. IV, §17], in particular (17.7.11)(ii), (17.8.2) and (17.15.2).

(ii) If $S \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ is a subvariety of Hodge type then we can find m and $F \subseteq F'$ as above, $\eta \in \mathrm{CSp}_{2g}(\mathbb{A}_f)$, and a subvariety $S' \hookrightarrow \mathsf{A}_{g,1,mn} \otimes F'$ of Shimura type, such that $K_{mn} \subseteq \eta K_n \eta^{-1}$ and such that $S \otimes F'$ is the image of S' under the Hecke correspondence $\mathcal{T}_\eta\colon \mathsf{A}_{g,1,mn} \to \mathsf{A}_{g,1,n}$ (which in this case is a morphism). At all but finitely many primes $\mathfrak{p}'$ of $\mathcal{O}_{F'}$, the morphism $\mathcal{T}_\eta$ has a natural extension to a morphism

$\mathcal{T}_\eta\colon \mathcal{A}_{g,1,mn} \to \mathcal{A}_{g,1,n}$ on models over $\widehat{\mathcal{O}}_{\mathfrak{p}'}$, which induces a group homomorphism on completions at closed ordinary points in characteristic p. Statement (ii) now follows directly from (i). □

§5 Formal linearity and Shimura varieties of Hodge type

5.1 The main result of this section is a converse to Noot's theorem discussed in the previous section. Together with this and the yoga on formal linearity that we carried out in Section 3, it provides a characterization of subvarieties of $\mathsf{A}_{g,1,n}$ of Hodge type in terms of formal linearity.

5.2 Theorem. *Let $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ be an irreducible algebraic subvariety of the moduli space $\mathsf{A}_{g,1,n}$, defined over a number field F. Suppose there is a prime $\mathfrak{p}$ of $\mathcal{O}_F$ such that the model $\mathcal{Z}$ of Z (as in Section 3.5) has formally quasi-linear components at some closed ordinary point $x \in (\mathcal{Z} \otimes \kappa(\mathfrak{p}))^{\mathrm{o}}$. Then Z is of Hodge type, i.e., every irreducible component of $Z \otimes_F \mathbb{C}$ is a subvariety of Hodge type.*

Proof. We divide the proof in a number of steps.

Step 1. First we reduce to the case that Z is absolutely irreducible and $\mathcal{Z}$ is formally linear at x. Fix a field embedding $F \to \mathbb{C}$. It suffices to show that one of the irreducible components of $Z_{\mathbb{C}} = Z \otimes_F \mathbb{C}$ is of Hodge type. This is because the class of subvarieties of Hodge type is stable under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the (non-connected) variety $\mathsf{A}_{g,1,n} \otimes \mathbb{C}$. To see this, let $S \subseteq \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ be a subvariety of Hodge type, and let $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. By Proposition I.3.12 there is a subvariety S' of Shimura type, and a Hecke correspondence $\mathcal{T}_\eta$ such that S is an irreducible component of $\mathcal{T}_\eta(S')$. Then $S^{(\tau)}$ is a component of $\mathcal{T}_\eta(S'^{(\tau)})$, since the Hecke correspondences are defined over the reflex field, which for A_g is the field $\mathbb{Q}$. It therefore suffices to see that $S'^{(\tau)}$ is again of Shimura type. This follows easily from the main results on conjugation of Shimura varieties, which are discussed in [36, Section II.4].[1]

Choose a finite extension F' of F such that every irreducible component $Z' \hookrightarrow \mathsf{A}_{g,1,n} \otimes F'$ of $Z \otimes_F F'$ is absolutely irreducible. There exists a component Z' and a prime $\mathfrak{p}'$ of $\mathcal{O}_{F'}$ above $\mathfrak{p}$, such that the model $\mathcal{Z}'$ of Z' over $\mathrm{Spec}(\widehat{\mathcal{O}}_{\mathfrak{p}'})$ has formally linear components at some point x' in the preimage of x. By the preceding remarks it suffices to prove the theorem for Z'. However, since Z' is absolutely irreducible, we

[1] It is possible to carry out Step 1 without using the results on conjugation of Shimura varieties.

can apply Corollary 3.12, and it follows that we may even assume that $\mathcal{Z}'$ is formally linear at the point x'.

From now on we may therefore assume that Z is absolutely irreducible and that $\mathcal{Z}$ is formally linear at x.

Step 2. Write $\overline{\Lambda} = \overline{W} \otimes_{W(\kappa(\mathfrak{p}))} \widehat{\mathcal{O}}_{\mathfrak{p}}$, where $\overline{W}$ denotes the ring of Witt vectors of an algebraic closure of $\kappa(x)$. For the rest of the proof we fix compatible embeddings $\overline{\Lambda} \hookrightarrow \mathbb{C}$ and $F \hookrightarrow \mathbb{C}$.

Using these embeddings, let $x_{\mathbb{C}}^{\mathrm{can}} \in Z_{\mathbb{C}}$ be the moduli point of the canonical lifting of x. As in Chapter II, we compare $Z_{\mathbb{C}}$ with the smallest subvariety of Hodge type containing it, which we call $S_{\eta K}(Y_M^+)$ (see Section I.3.15). Here $M \subseteq \mathrm{CSp}_{2g,\mathbb{Q}}$ is the generic Mumford-Tate group on $Z_{\mathbb{C}}$ and Y_M^+ is a $M(\mathbb{R})$-conjugacy class in $\mathrm{Hom}(\mathbb{S}, M_{\mathbb{R}})$. If there is no risk of confusion we write $Y = Y_M^+$, $S = S_{\eta K}(Y_M^+)$. Let $u_S\colon Y \to S$ be the uniformization map, and let $\widetilde{Z}$ be a connected component of $u_S^{-1}(Z_{\mathbb{C}})$. We assume that the integer n of the level structure with which we are working, is large enough such that the conditions formulated in Section II.2.2 are satisfied. For this we may have to pass to finite coverings of $Z_{\mathbb{C}}$ and S, but it suffices to prove the theorem after such modifications. We use the notations $\mathrm{Cov}(u_S)$ and $\mathrm{Cov}(u_Z)$ as in Chapter II.

Let $\mathcal{C} \subseteq \widetilde{Z}$ be an irreducible (analytic) component and choose a point $\tilde{x}_{\mathbb{C}}^{\mathrm{can}} \in \mathcal{C}$ with $u(\tilde{x}_{\mathbb{C}}^{\mathrm{can}}) = x_{\mathbb{C}}^{\mathrm{can}}$. We define a subgroup $\pi_{\mathcal{C}} \subseteq \pi_1(Z_{\mathbb{C}}, x_{\mathbb{C}}^{\mathrm{can}})$ by

$$\pi_{\mathcal{C}} = \{\alpha \in \pi_1(Z_{\mathbb{C}}, x_{\mathbb{C}}^{\mathrm{can}}) \mid \alpha\mathcal{C} = \mathcal{C}\}\,.$$

Let $n\colon Z^n = Z_{\mathbb{C}}^n \to Z_{\mathbb{C}}$ be the normalization and let $u_{Z^n}\colon \widetilde{Z^n} \twoheadrightarrow Z^n$ be a universal covering. We can choose points $\tilde{\zeta} \in \widetilde{Z^n}$ and $\zeta \in Z^n$ with $u_{Z^n}(\tilde{\zeta}) = \zeta$ and $n(\zeta) = x_{\mathbb{C}}^{\mathrm{can}}$. There is a well-determined morphism $\tilde{n}\colon \widetilde{Z^n} \to \widetilde{Z}$ with $\tilde{n}(\tilde{\zeta}) = \tilde{x}_{\mathbb{C}}^{\mathrm{can}}$ and $\tilde{n}(\widetilde{Z^n}) = \mathcal{C}$.

With these choices and notations we are in the situation of the first diagram on page 26 and we can apply Lemma II.2.4.

Step 3. Let $\mathfrak{H}_g^\vee$ be the compact dual of $\mathfrak{H}_g$. This can be described as the flag variety of g-dimensional subspaces of $\mathbb{C}^{2g}$ which are totally isotropic for the (standard) symplectic form ψ.

Suppose $a_1, \ldots, a_g, c_1, \ldots, c_g$ is a basis for $\mathbb{C}^{2g}$ such that $\mathcal{F} = \mathrm{Span}\{c_1, \ldots, c_g\}$ is totally isotropic. Then $\mathfrak{H}_g^\vee \cong \mathrm{Sp}_{2g,\mathbb{C}}/P(\mathcal{F})$, where $P(\mathcal{F})$ is the stabilizer of $\mathcal{F}$, which is a parabolic subgroup of $\mathrm{Sp}_{2g,\mathbb{C}}$. Let $P(\mathcal{F})^- \subset \mathrm{Sp}_{2g,\mathbb{C}}$ be the parabolic subgroup opposite to $P(\mathcal{F})$, and define U as the image of $P(\mathcal{F})^-$ in $\mathfrak{H}_g^\vee$. It is a Zariski open subset of $\mathfrak{H}_g^\vee$, whose complement $D_\infty = \mathfrak{H}_g^\vee \setminus U$ is a divisor.

In terms of flags, U is the open part of $\mathfrak{H}_g^\vee$ corresponding to flags of the form

$$\mathcal{F}_T = \mathrm{Span}(\{c_j + \sum_{i=1}^{n} T_{ij} \cdot a_i\}_{j=1,\ldots,n})\,,$$

where T is a $g \times g$ matrix such that $\mathcal{F}_T$ is totally isotropic. (If ψ has the standard form with respect to the basis $a_1, \ldots, a_g, c_1, \ldots, c_g$ then this is equivalent to the condition that T is symmetric. However, we do not want to assume that ψ has a special form on the given basis.) The coefficients t_{ij} of the matrix T are well-determined regular functions on U and can in fact be described as global sections of the line bundle $\mathcal{L} = \mathcal{O}(kD_\infty)$ for some $k \geq 1$ (a direct computation shows that one can take $k = 1$). Notice that there is a natural action of $\mathrm{Sp}_{2g,\mathbb{C}}$ on $\mathcal{L}$, which makes it into a $\mathrm{Sp}_{2g,\mathbb{C}}$-bundle over $\mathfrak{H}_g^\vee$.

Step 4. Similar to the discussion in Section 1.12, write $\mathfrak{A} = \mathfrak{A}_x \widehat{\otimes} \overline{\Lambda} = \mathrm{Spf}(A)$, with $A = \overline{\Lambda}[\![q_{ij} - 1]\!]/(q_{ij} - q_{ji})$. Our assumption that $\mathcal{Z}$ is formally linear at x implies that $\mathfrak{Z} = \mathfrak{Z}_x \widehat{\otimes} \overline{\Lambda}$ can be described as $\mathfrak{Z} = \mathrm{Spf}(A/\mathfrak{a}) \hookrightarrow \mathfrak{A}$, where $\mathfrak{a} \subseteq A$ is an ideal generated by elements of the form $(\prod_{ij} q_{ij}^{m_{ij}}) - 1$, with $m_{ij} \in \mathbb{Z}_p$.

Let K be the quotient field of $\overline{\Lambda}$. We have an isomorphism

$$K[\![\tau_{ij}]\!]/(\tau_{ij} - \tau_{ji}) \xrightarrow{\sim} K[\![q_{ij} - 1]\!]/(q_{ij} - q_{ji})\,,$$

given by $\tau_{ij} \mapsto \log(q_{ij})$. Under this isomorphism the element $\sum_{ij} m_{ij}\tau_{ij}$ maps to $\sum_{ij} m_{ij} \log(q_{ij}) = \log\big(\prod_{ij} q_{ij}^{m_{ij}}\big)$, which, up to a unit in $K[\![q_{ij} - 1]\!]/(q_{ij} - q_{ji})$, is equal to $(\prod_{ij} q_{ij}^{m_{ij}}) - 1$. Using the chosen embedding of $\overline{\Lambda}$ into $\mathbb{C}$, we obtain a ring homomorphism

$$A = \overline{\Lambda}[\![q_{ij} - 1]\!]/(q_{ij} - q_{ji}) \longrightarrow A_{\mathbb{C}} := \mathbb{C}[\![\tau_{ij}]\!]/(\tau_{ij} - \tau_{ji})$$

such that the ideal $\mathfrak{a}_{\mathbb{C}} = \mathfrak{a} \cdot A_{\mathbb{C}}$ is generated by elements of the form $\sum_{ij} m_{ij}\tau_{ij}$, where the coefficients m_{ij} (as above) are now viewed as elements of $\mathbb{C}$, via the chosen embedding $\mathbb{Z}_p \subseteq \overline{\Lambda} \hookrightarrow \mathbb{C}$.

Pulling back the universal formal abelian scheme $\mathfrak{X} \to \mathfrak{A}$ via the continuous homomorphism $A \to A_{\mathbb{C}}$ yields a formal abelian scheme $\mathfrak{X}_{\mathbb{C}} \to \mathfrak{A}_{\mathbb{C}} = \mathrm{Spf}(A_{\mathbb{C}})$ for which we have a description of the de Rham cohomology in terms of the elements τ_{ij}. Namely, there is a horizontal basis $a_1, \ldots, a_g, c_1, \ldots, c_g$ for the $A_{\mathbb{C}}$-module $H_{\mathbb{C}} = \mathrm{H}^1_{\mathrm{DR}}(\mathfrak{X}_{\mathbb{C}}/\mathfrak{A}_{\mathbb{C}})$ such that the Hodge flag $\mathcal{F}^1$ is spanned by the elements $c_j + \sum_i \tau_{ij} a_i$.

The formal abelian scheme $\mathfrak{X} \to \mathfrak{A}$ is algebraizable, i.e., it is the formal completion along the closed fibre of an abelian scheme $\mathcal{X} \to \mathrm{Spec}(A)$. Write $\mathcal{X}_{\mathbb{C}} \to \mathrm{Spec}(A_{\mathbb{C}})$ for the pull-back via $A \to A_{\mathbb{C}}$; its formal completion along the closed fibre is $\mathfrak{X}_{\mathbb{C}} \to \mathfrak{A}_{\mathbb{C}}$. The corresponding morphism $\mathrm{Spec}(A_{\mathbb{C}}) \to \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ sends the point with coordinates $\tau_{ij} = 0$ to the point $x_{\mathbb{C}}^{\mathrm{can}}$, and it follows from the given description of the de Rham cohomology that $\mathfrak{X}_{\mathbb{C}} \to \mathfrak{A}_{\mathbb{C}}$ is the universal deformation of $X_{\mathbb{C}}^{\mathrm{can}}$. We thus have an isomorphism of $\mathrm{Spf}(A_{\mathbb{C}})$ with the formal completion of $\mathsf{A}_{g,1,n} \otimes \mathbb{C}$ at $x_{\mathbb{C}}^{\mathrm{can}}$.

The assumption that $\mathcal{Z}$ is formally linear at x implies that the point x lies in the locus where the structure morphism $\mathcal{Z}_{\overline{\Lambda}} \to \mathrm{Spec}(\overline{\Lambda})$ is smooth. Since this is an open locus, the same is true for the point x^{can}, and by the results of [27, Chap. IV, §17] it follows that $x_{\mathbb{C}}^{\mathrm{can}}$ is a non-singular point of $Z_{\mathbb{C}}$. We claim that the isomorphism of $\mathrm{Spf}(A_{\mathbb{C}})$ with the formal completion of $\mathsf{A}_{g,1,n} \otimes \mathbb{C}$ at $x_{\mathbb{C}}^{\mathrm{can}}$ restricts to an isomorphism

$$\mathfrak{Z}_{\mathbb{C}} := \mathrm{Spf}(A_{\mathbb{C}}/\mathfrak{a}_{\mathbb{C}}) \xrightarrow{\sim} (Z_{\mathbb{C}})_{/\{x_{\mathbb{C}}^{\mathrm{can}}\}} \subseteq (\mathsf{A}_{g,1,n} \otimes \mathbb{C})_{/\{x_{\mathbb{C}}^{\mathrm{can}}\}}\,.$$

First we remark that the composite morphism

$$\mathrm{Spec}(A_{\mathbb{C}}/\mathfrak{a}_{\mathbb{C}}) \hookrightarrow \mathrm{Spec}(A_{\mathbb{C}}) \to \mathsf{A}_{g,1,n} \otimes \mathbb{C}$$

factors through $Z_{\mathbb{C}}$. It follows that the closed formal subscheme $\mathrm{Spf}(A_{\mathbb{C}}/\mathfrak{a}_{\mathbb{C}}) \hookrightarrow (\mathsf{A}_{g,1,n} \otimes \mathbb{C})_{/\{x_{\mathbb{C}}^{\mathrm{can}}\}}$ is contained in $(Z_{\mathbb{C}})_{/\{x_{\mathbb{C}}^{\mathrm{can}}\}}$. The point $x_{\mathbb{C}}^{\mathrm{can}} \in Z_{\mathbb{C}}$ being non-singular, the claim then follows from the fact that the dimensions of $A_{\mathbb{C}}/\mathfrak{a}_{\mathbb{C}}$ and $Z_{\mathbb{C}}$ are equal.

Since $x_{\mathbb{C}}^{\mathrm{can}}$ is a regular point of $Z_{\mathbb{C}}$, the covering maps $u\colon \mathfrak{H}_g^{\pm} \twoheadrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ and $u_{Z|\mathcal{C}}\colon \mathcal{C} \twoheadrightarrow Z_{\mathbb{C}}$ induce isomorphisms

$$\hat{u}\colon (\mathfrak{H}_g^{\pm})_{/\{\tilde{x}_{\mathbb{C}}^{\mathrm{can}}\}} \xrightarrow{\sim} (\mathsf{A}_{g,1,n} \otimes \mathbb{C})_{/\{x_{\mathbb{C}}^{\mathrm{can}}\}} \quad \text{and} \quad \hat{u}_{|\mathcal{C}}\colon \mathcal{C}_{/\{\tilde{x}_{\mathbb{C}}^{\mathrm{can}}\}} \xrightarrow{\sim} (Z_{\mathbb{C}})_{/\{x_{\mathbb{C}}^{\mathrm{can}}\}}\,.$$

The choice of the point $\tilde{x}_{\mathbb{C}}^{\mathrm{can}} \in \mathcal{C} \subseteq \widetilde{Z}$ gives a symplectic basis for $\mathrm{H}^1(X^{\mathrm{can}}, \mathbb{Q})$, which we use to identify $\mathrm{H}^1(X^{\mathrm{can}}, \mathbb{C})$ with $\mathbb{C}^{2g}$. The comparison isomorphism between de Rham and singular cohomology yields an isomorphism $\mathrm{H}^1(X^{\mathrm{can}}, \mathbb{C}) \xrightarrow{\sim} H_{\mathbb{C}} \otimes_{A_{\mathbb{C}},\mathrm{ev}_0} \mathbb{C}$, where $\mathrm{ev}_0\colon A_{\mathbb{C}} \to \mathbb{C}$ is the evaluation map at zero (i.e., the map with $\tau_{ij} \mapsto 0$). In this way we obtain a basis of elements $a_1, \ldots, a_g, c_1, \ldots, c_g \in \mathbb{C}^{2g}$ such that $\mathcal{F} = \mathrm{Span}\{c_1, \ldots, c_g\}$ is a totally isotropic subspace. As in Step 3 we have global sections $t_{ij} \in \Gamma(\mathfrak{H}_g^{\vee}, \mathcal{L})$, which extend the regular functions τ_{ij} on $A_{\mathbb{C}}$. (Notice that we get the condition $T = T^t$ on the matrix $T = (t_{ij})$, since the polarization form is the standard one with respect to the elements $a_1, \ldots, a_g, c_1, \ldots, c_g$.)

The situation now is as follows. The open part U of $\mathfrak{H}_g^\vee$ can be described as $U = \mathrm{Spec}(\mathbb{C}[t_{ij}]/(t_{ij} - t_{ji}))$. The point

$$\tilde{x}_{\mathbb{C}}^{\mathrm{can}} \in (\mathcal{C} \cap U) \subseteq (\widetilde{Z} \cap U) \subseteq U$$

has coordinates $t_{ij} = 0$, and in the formal completion at $\tilde{x}_{\mathbb{C}}^{\mathrm{can}}$, the locus $\mathcal{C}$ is given by linear equations $\sum_{ij} m_{ij} t_{ij} = 0$. We define a subspace $I \subseteq \Gamma(\mathfrak{H}_g^\vee, \mathcal{L})$ by

$$I = \{ s \in \Gamma(\mathfrak{H}_g^\vee, \mathcal{L}) \mid s_{|\mathcal{C}} = 0 \} .$$

Notice that all sections $\sum_{ij} m_{ij} t_{ij} \in \Gamma(\mathfrak{H}_g^\vee, \mathcal{L})$ (where the $m_{ij} \in \mathbb{C}$ are as before) are in I, since they define holomorphic functions on $\mathcal{C}$ with vanishing Taylor expansion at the point $\tilde{x}_{\mathbb{C}}^{\mathrm{can}}$. If $V(I)$ denotes the zero locus of I then it follows that $\mathcal{C}$ is an irreducible component of $V(I) \cap \mathfrak{H}_g$.

We have a homomorphism $\pi_1(Z^n, \zeta) \to M^{\mathrm{der}}(\mathbb{Q}) \subseteq \mathrm{Sp}_{2g}(\mathbb{Q})$ induced from the composition $Z^n \to Z_{\mathbb{C}} \hookrightarrow S$. This gives an action of $\pi_1(Z^n, \zeta)$ on $\Gamma(\mathfrak{H}_g^\vee, \mathcal{L})$. Since $\mathcal{L}$ is a $\mathrm{Sp}_{2g,\mathbb{C}}$-bundle, the fact that $\pi_1(Z^n, \zeta)$ maps into the subgroup $\pi_{\mathcal{C}}$ of $\pi_1(Z_{\mathbb{C}}, x_{\mathbb{C}}^{\mathrm{can}})$ (by Lemma II.2.4) implies that the subspace $I \subseteq \Gamma(\mathfrak{H}_g^\vee, \mathcal{L})$ is stable under the action of $\pi_1(Z^n, \zeta)$. Then I is also stable under the algebraic envelope of the image of $\pi_1(Z^n, \zeta)$ in $\mathrm{GL}(\Gamma(\mathfrak{H}_g^\vee, \mathcal{L}))$. At this point we can apply André's result, see Theorem I.2.4, from which we conclude that M^{der} is equal to the connected algebraic monodromy group H_ζ of the abelian scheme over Z^n (obtained by pulling back the abelian scheme over $Z_{\mathbb{C}}$). The image of $\pi_1(Z^n, \zeta)$ is therefore Zariski dense in $M^{\mathrm{der}}_{\mathbb{R}}$ (cf. the remark in the penultimate paragraph of the proof of Theorem II.3.1) and I is stable under the action of $M^{\mathrm{der}}(\mathbb{R})$. Consequently, $V(I) \cap \mathfrak{H}_g$ is stable under $M^{\mathrm{der}}(\mathbb{R})^+$, and since this last group acts transitively on $Y \subseteq \mathfrak{H}_g$ we conclude that $\mathcal{C} = Y$, hence $Z_{\mathbb{C}} = S$. This proves that $Z_{\mathbb{C}}$ is of Hodge type. □

5.3 Conclusion. In summary, we have proved that subvarieties $S \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ of Hodge type are characterized by their property of being "formally linear" in a suitable sense. More precisely, if S is of Hodge type then we have seen in 4.4 that for almost all primes $\mathfrak{p}$ of $\mathcal{O}_F$, the model $\mathcal{S}$ of S over $\mathrm{Spec}(\widehat{\mathcal{O}}_{\mathfrak{p}})$ is a union of formally linear components at all closed ordinary points. If S is of Shimura type or if S is non-singular then it is formally linear at all its ordinary points in characteristic p (excluding finitely many p).

Conversely, if S is "formally linear" in the weaker sense that $\mathcal{S}$ has formally quasi-linear components at some closed ordinary point, then the previous theorem shows, via the results of Section 3, that S is of Hodge type.

Chapter IV

Some applications

§1 Oort's conjecture

1.1 In our study of subvarieties $Z \hookrightarrow \mathsf{A}_{g,1,n}$ of Hodge type in the previous two chapters, we have seen that an important role is played by the CM-points. It then seems natural to ask whether subvarieties of Hodge type are characterized by the property that the CM-points on them are dense. It is a conjecture of Oort that this is the case. Before we give a precise statement, let us remark that one could formulate the conjecture for general Shimura varieties (in which case we would use the terminology "special points" rather than "CM-points"). Here, however, we restrict our attention to moduli spaces of abelian varieties.

1.2 Conjecture. (Oort) *Let $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{C}$ be an irreducible algebraic subvariety such that the CM-points on Z are dense for the Zariski topology. Then Z is a subvariety of Hodge type, in the sense of Definition I.3.8.*

1.3 Remark. In [1, Chapter X, §4], a number of problems are suggested, the first of which is equivalent to the above conjecture for $\dim(Z) = 1$. As mentioned, the example discussed in II.3.3 provides a counterexample to loc. cit., Problems 2 and 3.

Let Z be a subvariety as in the conjecture. Then Z is defined over a number field, since it is the Zariski closure of a set of points which are rational over $\overline{\mathbb{Q}}$ (even over the union $\mathbb{Q}^{\mathrm{CM}} \subset \overline{\mathbb{Q}}$ of all CM-subfields). It follows from the results of the previous chapter (in particular Corollary III.3.11 and Theorem III.5.2) that the conjecture is equivalent to the following statement.

1.4 Conjecture. (Variant of 1.2) *Let F be a number field, and let $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes F$ be an irreducible algebraic subvariety such that the CM-points on Z are dense for the*

Zariski topology. Then there is a collection T of CM-points $t \in Z(K_t)$ (K_t a number field containing F) and a prime number p such that the collection T is Zariski dense in Z, and such that every X_t is canonical at a prime $\mathfrak{q}_t$ of K_t lying over p.

Given an abelian variety X_t of CM-type over a number field K_t we know that X_t is canonical at infinitely many primes of K_t. However, it is not true that for any infinite collection of CM-points $t \in \mathsf{A}_g(K_t)$ there is always a prime number p such that infinitely many of the abelian varieties X_t are canonical at a prime $\mathfrak{q}_t$ above p. For example, let $\mathcal{E}_1, \mathcal{E}_2, \ldots$ be elliptic curves of CM-type such that the conductor of $\mathrm{End}(\mathcal{E}_i)$ (by which we mean the index of $\mathrm{End}(\mathcal{E}_i)$ in the ring of integers of $\mathrm{End}^0(\mathcal{E}_i)$) is divisible by the first i rational prime numbers. By [49, Lemma 2.2] we conclude that for every given prime number p, there are only finitely many $\mathcal{E}_i$ which are canonical at a prime above p.

Of course this does not provide a counterexample to the conjecture: the Zariski closure of our collection of CM-points $\mathcal{E}_i$ is the whole moduli space $\mathsf{A}_{1,1,n}$ (assuming that the $\mathcal{E}_i$ were equipped with a level n structure), which is certainly of Hodge type.

1.5 In the proof of the next result we will use the Galois representation on the ℓ-torsion and on the Tate-ℓ-module of an abelian variety X defined over a number field. A lot of useful general theory can be found in [61], [53] and in the letters of Serre to Ribet and Tate [58], [59]; part of the material of these letters is given in Chi's paper [10]. Here we only record some facts needed further on.

Let X be an abelian variety over a number field F, and write $\rho_\ell\colon \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to \mathrm{Aut}(T_\ell X)$ and $\overline{\rho}_\ell\colon \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to \mathrm{Aut}(X(\overline{\mathbb{Q}})[\ell])$ for the Galois representation on its Tate-ℓ-module and its ℓ-torsion respectively. We write G_ℓ for the algebraic envelope of the image of ρ_ℓ. Its connected component of the identity is a reductive algebraic group over $\mathbb{Q}_\ell$ containing the group $\mathbb{G}_{\mathrm{m}} \cdot \mathrm{Id}$ of homotheties. Its Lie algebra does not change if we replace F by a finite extension, but G_ℓ itself may be non-connected and may become smaller after such an extension.

Choose an embedding $\sigma\colon F \to \mathbb{C}$, and write $V = \mathrm{H}_1(X_\sigma(\mathbb{C}), \mathbb{Q})$ and $V_\ell = T_\ell X \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$. There is a natural comparison isomorphism $V_\ell \cong V \otimes \mathbb{Q}_\ell$, and, by the results of Deligne and Piatetski-Shapiro (see [19]), G_ℓ is an algebraic subgroup of $\mathrm{MT}(X_\sigma) \otimes \mathbb{Q}_\ell$. The Mumford-Tate conjecture (stated in [41] and in a more refined version in [60]) asserts that the two groups are equal.

It is known that the representation of $\mathrm{MT}(X_\sigma)$ on V is defined by miniscule weights. By this we mean the following: first we write $\mathrm{MT}(X_\sigma)_{\overline{\mathbb{Q}}}$ as the almost direct product of its center Z and a number of simple factors $M_1, \ldots, M_r$. Then every irreducible submodule $W \subseteq V_{\overline{\mathbb{Q}}}$ decomposes as a tensor product $W \cong \chi \otimes W_1(\varpi_1) \otimes \cdots \otimes W_r(\varpi_r)$, where χ is a character of Z, and where $W_i(\varpi_i)$ is an irreducible M_i-module with highest weight ϖ_i (with respect to a chosen Borel subgroup of M_i). The representation V is said to be defined by miniscule weights if all occurring weights ϖ_i are miniscule in the sense of [7, Ch. VIII, §7, n° 3].

An immediate consequence of the Mumford-Tate conjecture would be that the representation of G_ℓ on V_ℓ is also defined by miniscule weights. So far this has not been proved in general. However, Zarhin proved it under the additional assumption that X has ordinary reduction at a set of places of density 1; see [66, Theorem 4.2].

1.6 Lemma. *Let K be an algebraically closed field of characteristic zero, let G be a reductive algebraic group over K and let V be a finite-dimensional representation of G which is defined by miniscule weights. Write $V \cong W_1^{d_1} \oplus \cdots \oplus W_m^{d_m}$, where $W_1, \ldots, W_m$ are mutually non-isomorphic irreducible representations of G. If ϖ is a weight of W_i then it has multiplicity d_i in the representation V. The total number of different weights that occur in the representation V is therefore equal to* $\dim(W_1) + \cdots + \dim(W_m)$.

Proof. Suppose $\mathfrak{g}$ is a simple Lie algebra over K, ϖ is a miniscule weight of $\mathfrak{g}$ (with respect to a chosen Cartan subalgebra) and W is an irreducible $\mathfrak{g}$-module with highest weight ϖ. The lemma follows directly from the following two facts, proven in [7, Ch. VIII, §7, n° 3]: (i) all weights of W have multiplicity 1, (ii) the Weyl group acts transitively on the set of weights of W. □

1.7 Let X be defined over the number field F, and let v be a finite place of F such that X has good reduction at v. If $\ell \nmid v$ then ρ_ℓ is unramified at v. By the choice of a place $\bar{v}$ of $\overline{\mathbb{Q}}$ extending v we get a well-determined action of a Frobenius element $\rho_\ell(\mathrm{Fr}_{\bar{v}}) \in \mathrm{Aut}(V_\ell X)$. Alternatively, X having good reduction at v means that it extends to an abelian scheme $\mathcal{X}_v$ over $\mathrm{Spec}(\mathcal{O}_v)$, whose special fibre X_v is an abelian variety over the finite field $\kappa(v)$. Let π_v be the Frobenius endomorphism of X_v, which acts on the Tate module $T_\ell X_v$. Via the choice of the place $\bar{v}$ we get an isomorphism $T_\ell X \cong T_\ell X_v$, and the action of π_v on $T_\ell X$ obtained in this manner is given by the element $\rho_\ell(\mathrm{Fr}_{\bar{v}})$.

Associated to X_v there is an algebraic torus T_v over $\mathbb{Q}$, called a Frobenius torus. As a module under $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, its character group $X^*(T_v)$ is isomorphic to $\Gamma_v/\mathrm{Tors}(\Gamma_v)$, where $\Gamma_v \subset \overline{\mathbb{Q}}^*$ is the subgroup generated by the eigenvalues of $\rho_\ell(\mathrm{Fr}_{\bar{v}})$. This description determines T_v up to isomorphism. The choice of a place $\bar{v}$ as above induces an injective homomorphism $T_v \otimes \mathbb{Q}_\ell \hookrightarrow G_\ell$. For more on Frobenius tori we refer to [10, Section 3].

The following facts were proved by Serre (see [10]): (i) the rank of G_ℓ does not depend on ℓ, (ii) we can replace F by a finite extension such that all groups G_ℓ and all Frobenius tori T_v (for places v of good reduction) are connected, and (iii) after replacing F by such an extension, there is a Zariski open and dense subset $U \subseteq G_\ell$ such that if $\ell \nmid v$ and $\rho_\ell(\mathrm{Fr}_{\bar{v}}) \in U(\mathbb{Q}_\ell)$, then $T_v \otimes \mathbb{Q}_\ell$ is a maximal torus of G_ℓ (the set of places v for which this holds thus has density 1).

1.8 Theorem. *Let (X, λ, θ) be a principally polarized abelian variety with a Jacobi level n structure, defined over a number field F. Suppose that for some finite field extension $F \subseteq F'$, the set*

$$\mathcal{P}^{\mathrm{o}}(F') = \{\text{finite places } v \text{ of } F' \mid X \otimes F' \text{ has good and ordinary reduction at } v\}$$

has Dirichlet density 1. For each $v \in \mathcal{P}^{\mathrm{o}}(F)$ with residue characteristic p_v not dividing n, let $(X_v, \lambda_v, \theta_v)$ be the reduction at v, and let $x_v^{\mathrm{can}} \in \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ be the moduli point of its canonical lifting. Define $Z \subseteq \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ as the Zariski closure of the set $\{x_v^{\mathrm{can}} \mid v \in \mathcal{P}^{\mathrm{o}}(F), p_v \nmid n\}$. Then Z is a union of subvarieties of Hodge type; more precisely:

$$Z_{\overline{\mathbb{Q}}} = \bigcup_{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})} S^{(\sigma)} \quad \cup \quad \{s_1, \ldots, s_r\},$$

where $S \hookrightarrow \mathsf{A}_{g,1,n} \otimes \overline{\mathbb{Q}}$ is the smallest subvariety of Hodge type containing the moduli point of $(X, \lambda, \theta) \otimes_F \overline{\mathbb{Q}}$ for some embedding $F \hookrightarrow \overline{\mathbb{Q}}$, and where $s_1, \ldots, s_r$ ($r \in \mathbb{Z}_{\geq 0}$) are CM-points.

Proof. Suppose F' is a finite extension of F such that $\mathcal{P}^{\mathrm{o}}(F')$ has Dirichlet density 1. Let F'' be a Galois extension of F' of degree d, and write $\mathcal{Q}$ for the set of primes of F' which split completely in F''/F'. Using the Čebotarev density theorem we see that the set $\mathcal{P}^{\mathrm{o}}(F') \cap \mathcal{Q}$ has Dirichlet density $1/d$, which means that the function

$$\sum_{\mathfrak{p} \in \mathcal{P}^{\mathrm{o}}(F') \cap \mathcal{Q}} N(\mathfrak{p})^{-s} \; + \; 1/d \cdot \log(s-1)$$

extends to a holomorphic function on $\{s \in \mathbb{C} \mid \mathrm{Re}(s) \geq 1\}$. If $\mathcal{R}$ is the set of primes of F'' lying over $\mathcal{P}^{\mathrm{o}}(F') \cap \mathcal{Q}$, then it it follows that the function

$$\sum_{\mathfrak{p}' \in \mathcal{R}} N(\mathfrak{p}')^{-s} \; + \; \log(s-1)$$

also extends to a holomorphic function of s for $\mathrm{Re}(s) \geq 1$. Clearly, $\mathcal{R} \subseteq \mathcal{P}^{\mathrm{o}}(F'')$, and almost all primes of $\mathcal{P}^{\mathrm{o}}(F'') \setminus \mathcal{R}$ have degree at least 2 over $\mathbb{Q}$. It readily follows from this that $\mathcal{P}^{\mathrm{o}}(F'')$ has Dirichlet density 1.

The preceding remarks show that, in proving the theorem, we may replace F by a finite extension. We claim that, after such an extension, there exists an infinite subset $\mathcal{P}' \subseteq \mathcal{P}^{\mathrm{o}}(F)$ and a prime $\mathfrak{p}$ of $\mathcal{O}_F$ such that each of the abelian varieties X_v^{can} with $v \in \mathcal{P}'$ is canonical at some prime $\mathfrak{q}$ above $\mathfrak{p}$. Before we prove this, let us show how the result would follow from it.

So, suppose we have such a set $\mathcal{P}'$, and write $Z' \subseteq Z$ for the Zariski closure of the corresponding set of CM-points $\{x_v^{\mathrm{can}} \mid v \in \mathcal{P}', p_v \nmid n\}$. It follows from Corollary III.4.4 that almost all points x_v^{can} lie on $\cup_{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})} S^{(\sigma)} \hookrightarrow \mathsf{A}_{g,1,n}$, so $Z'_{\overline{\mathbb{Q}}} \subseteq Z_{\overline{\mathbb{Q}}}$ is contained in the union of $\cup_{\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})} S^{(\sigma)}$ and a finite number of CM-points. On the other hand, from III.3.11 and III.5.2 we see that all irreducible components of Z' are of Hodge type. Therefore, we are done if we show that the moduli point x_F of (X, λ, θ) lies on Z'. This we can see as follows.

First we may replace $\mathcal{P}'$ by an infinite subset such that its Zariski closure Z' is irreducible. Over some open part $U = \mathrm{Spec}(\mathbb{Z}[1/N])$ of $\mathrm{Spec}(\mathbb{Z})$, the point x_F extends to a section $x\colon \mathrm{Spec}(\mathcal{O}_F[1/N]) \to \mathsf{A}_{g,1,n}$. We define $\mathcal{Z}'$ as the Zariski closure of Z' inside $\mathsf{A}_{g,1,n} \otimes \mathbb{Z}[1/N]$. Then we have an infinite collection $\mathcal{P}'' = \{v \in \mathrm{Spec}(\mathcal{O}_F[1/N]) \mid v \in \mathcal{P}', p_v \nmid N\}$ such that (by construction) every

$$x_v\colon \mathrm{Spec}(\kappa(v)) \to \mathsf{A}_{g,1,n} \otimes \mathbb{Z}[1/N]$$

with $v \in \mathcal{P}''$ factors through $\mathcal{Z}'$. Because the collection $\mathcal{P}''$ is dense in $\mathrm{Spec}(\mathcal{O}_F[1/N])$ it then follows that x factors through $\mathcal{Z}'$, which means that x_F is a point of Z'. We conclude that the theorem follows if we can construct a set $\mathcal{P}'$ as above.

From now on we use the notations and results discussed in 1.5 up to 1.7 above. We replace F by a finite extension such that $\mathcal{P}^{\mathrm{o}}(F)$ has Dirichlet density 1 and such that the groups G_ℓ and the Frobenius tori T_v are connected (for all ℓ and all places v where X has good reduction). This implies that all endomorphisms of $X \otimes \overline{\mathbb{Q}}$ and

$X_v \otimes \overline{\kappa(v)}$ are defined over F and $\kappa(v)$, respectively. We write $\mathfrak{f} = \mathfrak{f}(X)$ for the conductor of the endomorphism ring $\mathrm{End}(X)$, i.e., the index of $\mathrm{End}(X)$ in a maximal order of $\mathrm{End}^0(X)$, and if X has good reduction at a place v of F then we simply write $\mathfrak{f}_v$ for the conductor of $\mathrm{End}(X_v)$.

Suppose ℓ is a prime number and v is an element of $\mathcal{P}^{\mathrm{o}}(F)$ such that $\ell \nmid \mathfrak{f}_v$ and such that ℓ splits completely in the field $\mathbb{Q}(\pi_v)^{\mathrm{norm}} \subset \overline{\mathbb{Q}}$ generated by the eigenvalues of π_v. We claim that under these assumptions X_v^{can} (which is defined over some number field $K \supseteq F$) is canonical at all primes λ of K above ℓ (where we take K large enough such that X_v^{can} has good reduction at all primes of K). In fact, the assumption that ℓ splits completely in $\mathbb{Q}(\pi_v)^{\mathrm{norm}}$ implies that the reduction Y_λ of X_v^{can} modulo λ is ordinary (using [64, Lemme 5]) and since ℓ does not divide the conductor of $\mathrm{End}(X_v^{\mathrm{can}})$, the endomorphism rings of X_v^{can} and Y_λ are the same (see Lemma III.1.15), so X_v^{can} is the canonical lifting of Y_λ. Therefore, we are done if we show that there are primes ℓ such that the set

$$\mathcal{P}^{\mathrm{o}}(\ell) = \{v \in \mathcal{P}^{\mathrm{o}}(F) \mid \ell \nmid \mathfrak{f}_v \text{ and } \ell \text{ splits completely in the field } \mathbb{Q}(\pi_v)^{\mathrm{norm}}\}$$

is infinite.

Suppose X' is an abelian variety which is F-isogenous to X, say by an isogeny $f\colon X \to X'$ of degree d. For a place v where X and X' have good reductions X_v and X'_v, the associated fields $\mathbb{Q}(\pi(X_v))^{\mathrm{norm}}$ and $\mathbb{Q}(\pi(X'_v))^{\mathrm{norm}}$ are naturally isomorphic and there is an isogeny $f_v\colon X_v \to X'_v$ of degree d, cf. [25, Ch. I, Proposition 2.7]. It follows that for all ℓ, the sets $\mathcal{P}^{\mathrm{o}}(\ell)$ associated to X and X' differ only by finitely many elements. We may therefore assume that $X = Y_1^{m_1} \times \cdots \times Y_r^{m_r}$, where $Y_1, \ldots, Y_r$ are mutually non-isogenous simple abelian varieties over F and $m_1, \ldots, m_r$ are positive integers.

Choose a place v of F and a place $\bar{v}$ of $\overline{\mathbb{Q}}$ extending v, such that X has good reduction at v and such that $T_v \otimes \mathbb{Q}_\ell \subseteq G_\ell$ is a maximal torus for every $\ell \neq p_v$. Choose a prime p with $p \neq p_v$ which splits completely in $\mathrm{End}^0(X)$, i.e., $\mathrm{End}^0(X) \otimes \mathbb{Q}_p$ is a product of matrix algebras over $\mathbb{Q}_p$. Let Y be one of the simple factors Y_i, let E be the center of $\mathrm{End}^0(Y)$, and let $\mathfrak{p}_1, \ldots, \mathfrak{p}_e$ be the primes of $\mathcal{O}_E$ above p. By the choices we have made, $\mathrm{End}^0(Y) \otimes \mathbb{Q}_p \cong \mathrm{M}_d(\mathbb{Q}_p)^e$ with $e = [E : \mathbb{Q}]$. The representation ρ_p of $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$ on $V_pY = T_pY \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ decomposes as

$$V_pY \cong \mathbb{V}_1^d \oplus \cdots \oplus \mathbb{V}_e^d\,,$$

where $\mathbb{V}_1, \ldots, \mathbb{V}_e$ are mutually non-isomorphic and absolutely irreducible G_p-modules such that E acts on $\mathbb{V}_i$ through its completion $E_{\mathfrak{p}_i} \cong \mathbb{Q}_p$. Write $\tilde{\rho}_p$ for the representation of G_p on $\mathbb{V}_1 \oplus \cdots \oplus \mathbb{V}_e$.

Let $P_v(t) = \det(t \cdot \mathrm{Id} - \tilde{\rho}_p(\mathrm{Fr}_{\bar{v}}) \mid \mathbb{V}_1 \oplus \cdots \oplus \mathbb{V}_e)$ be the characteristic polynomial of $\tilde{\rho}_p(\mathrm{Fr}_{\bar{v}})$, which has coefficients in $\mathbb{Z}$. We define $\delta_Y \in \mathbb{Z}$ as the discriminant of the polynomial $P_v(t)$, and we put $\delta_X = \delta_{Y_1} \cdots \delta_{Y_r}$. We observe that $\delta_X \neq 0$. In fact, the eigenvalues of $\tilde{\rho}_p(\mathrm{Fr}_{\bar{v}})$ are the elements $\varpi(\pi_v) \in \overline{\mathbb{Q}}_p$, where ϖ runs through the set of weights (counted with multiplicities) of $\tilde{\rho}_p$ with respect to the maximal torus $T_v \otimes \mathbb{Q}_p$. Since T_v is generated by the element π_v, we have $\varpi(\pi_v) = \varpi'(\pi_v)$ if and only if $\varpi = \varpi'$. By Zarhin's result [66, Theorem 4.2], the representation ρ_p is defined by miniscule weights, and using Lemma 1.6 we conclude that all eigenvalues of $\tilde{\rho}_p(\mathrm{Fr}_{\bar{v}})$ have multiplicity 1, hence $\delta_X \neq 0$.

Next we consider prime numbers ℓ satisfying the following conditions: (i) ℓ splits completely in $\mathrm{End}^0(X_v)$, (ii) $\ell \nmid \delta_X \cdot \mathfrak{f}(X)$. We claim that for every such ℓ the set $\mathcal{P}^{\circ}(\ell)$ is infinite. To see this, consider the representation $\overline{\rho}_\ell$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$ on $X(\overline{\mathbb{Q}})[\ell]$, and write $\gamma = \overline{\rho}_\ell(\mathrm{Fr}_{\bar{v}}) \in \mathrm{Aut}(X(\overline{\mathbb{Q}})[\ell]) \cong \mathrm{GL}_{2g}(\mathbb{F}_\ell)$, where $g = \dim(X)$.

As above, let Y be one of the simple factors of X, and write γ_Y for the restriction of γ to $Y(\overline{\mathbb{Q}})[\ell]$. The assumption that $\ell \nmid \mathfrak{f}(X)$ implies that $\mathrm{End}(Y) \otimes \mathbb{Z}_\ell$ is a maximal order of $\mathrm{End}(Y) \otimes \mathbb{Q}_\ell$, which by (i) is isomorphic to $\mathrm{M}_d(\mathbb{Q}_\ell)^e$. Thus $\mathrm{End}(Y) \otimes \mathbb{Z}_\ell \cong \mathrm{M}_d(\mathbb{Z}_\ell)^e$, and we conclude from this that there exists a decomposition of the Tate module (as a $\mathbb{Z}_\ell$-module with an action of $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$)

$$T_\ell Y = \mathbb{T}_1^d \oplus \cdots \oplus \mathbb{T}_e^d ,$$

where $\mathbb{T}_i$ is a free $\mathbb{Z}_\ell$-module of rank $2 \cdot \dim(Y)/ed$ and $\mathbb{T}_i \otimes \mathbb{Q}_\ell$ is an absolutely irreducible representation of G_ℓ. The above argument shows that all eigenvalues of $\mathrm{Fr}_{\bar{v}}$ on $(\mathbb{T}_1 \oplus \cdots \oplus \mathbb{T}_e) \otimes \overline{\mathbb{Q}}_\ell$ have multiplicity 1. Moreover, assumption (i) on ℓ implies that all these eigenvalues lie in $\mathbb{Z}_\ell$. Now

$$Y(\overline{\mathbb{Q}})[\ell] \cong ((\mathbb{T}_1/\ell\mathbb{T}_1) \oplus \cdots \oplus (\mathbb{T}_e/\ell\mathbb{T}_e))^d ,$$

and we assumed that $\ell \nmid \delta_X$, so the $2\dim(Y)/d$ different eigenvalues of $\mathrm{Fr}_{\bar{v}}$ are also different modulo ℓ. It follows that $\gamma_Y \in \mathrm{Aut}(Y(\overline{\mathbb{Q}})[\ell]) \cong \mathrm{GL}_{2\dim(Y)}(\mathbb{F}_\ell)$ is diagonalizable over $\mathbb{F}_\ell$ with eigenvalues all of multiplicity d.

Now we start working backwards. Suppose w is a place in the set $\mathcal{P}^{\circ}(F)$, and $\bar{w}$ is an extension of w to $\overline{\mathbb{Q}}$, such that $\ell \neq p_w$, and $\overline{\rho}_\ell(\mathrm{Fr}_{\bar{w}}) = \gamma$. Since $\mathrm{Aut}(X(\overline{\mathbb{Q}})[\ell])$

is finite and $\mathcal{P}^{\mathrm{o}}(F)$ has density 1, there are infinitely many such places w, by the Čebotarev density theorem. Therefore, the proof is finished if we show that all such w lie in the set $\mathcal{P}^{\mathrm{o}}(\ell)$, i.e., $\ell \nmid \mathfrak{f}_w$ and ℓ splits completely in the field $\mathbb{Q}(\pi_w)^{\mathrm{norm}}$.

First we reduce to the case that $X = Y$ is absolutely simple. Since $\mathbb{Q}(\pi_w)^{\mathrm{norm}}$ is the compositum of the fields $\mathbb{Q}(\pi_{Y_i,w})^{\mathrm{norm}}$ associated to the simple factors Y_i, the prime ℓ splits completely in $\mathbb{Q}(\pi_w)^{\mathrm{norm}}$ if and only if it splits completely in each of the fields $\mathbb{Q}(\pi_{Y_i,w})^{\mathrm{norm}}$. Also, we claim that $\ell \mid \mathfrak{f}_w$ if and only if ℓ divides one of the factors $\mathfrak{f}_w(Y_i)$. In the "if" direction this is clear, so let us assume that ℓ does not divide any of the factors $\mathfrak{f}_w(Y_i)$. Suppose $A \subseteq Y_{i,w}$ and $A' \subseteq Y_{j,w}$ are simple factors which are isogenous. Then $\mathrm{End}^0(A) = \mathrm{End}^0(A')$ is a CM-field (since A and A' are ordinary) and, by assumption, both orders $\mathrm{End}(A)$ and $\mathrm{End}(A')$ are maximal at ℓ. Using Tate's theorem that $\mathrm{Hom}(A, A') \otimes \mathbb{Z}_\ell \xrightarrow{\sim} \mathrm{Hom}_{\mathrm{Gal}}(T_\ell A, T_\ell A')$ we conclude that A and A' are prime-to-ℓ isogenous. From this remark and our assumption that X is the product of the factors $Y_i^{m_i}$ it then follows that ℓ does not divide $\mathfrak{f}_w = \mathfrak{f}_w(X)$. From now on we may therefore assume that $X = Y$ is absolutely simple.

The characteristic polynomial $P_w(t)$ of the action of $\mathrm{Fr}_{\bar{w}}$ on $\mathbb{T}_1 \oplus \cdots \oplus \mathbb{T}_e$ has coefficients in $\mathbb{Z}$. Modulo ℓ it is a product of linear factors, and all zeroes have multiplicity 1. By Hensel's lemma, $P_w(t) = (t - \alpha_1) \cdots (t - \alpha_u)$ in $\mathbb{Z}_\ell[t]$, with all $\alpha_i \in \mathbb{Z}_\ell$ different and $u = 2 \cdot \dim(Y)/d$. Let $Y_w^{(1)}, \ldots, Y_w^{(s)}$ be the simple factors of the reduction Y_w, and let $\pi_w^{(i)}$ be the Frobenius automorphism of $Y_w^{(i)}$. Then ℓ splits completely in each of the fields $\mathbb{Q}(\pi_w^{(i)})^{\mathrm{norm}} \subset \overline{\mathbb{Q}}$ generated by the eigenvalues of $\pi_w^{(i)}$, hence it splits completely in $\mathbb{Q}(\pi_w)^{\mathrm{norm}}$. Finally, the eigenvalues α_i of $\mathrm{Fr}_{\bar{w}}$ on $\mathbb{T}_1 \oplus \cdots \oplus \mathbb{T}_e$ are all different and $T_\ell Y \cong (\mathbb{T}_1 \oplus \cdots \oplus \mathbb{T}_e)^d$, so we get

$$\mathrm{End}(Y_w) \otimes \mathbb{Z}_\ell \cong \mathrm{End}_{\mathrm{Fr}_{\bar{w}}}(T_\ell Y) \cong \mathrm{M}_d(\mathbb{Z}_\ell^u) \hookrightarrow \mathrm{M}_d(\mathbb{Q}_\ell^u) \cong \mathrm{End}_{\mathrm{Fr}_{\bar{w}}}(V_\ell Y) \cong \mathrm{End}(Y_w) \otimes \mathbb{Q}_\ell,$$

and we see that $\ell \nmid \mathfrak{f}_w$. This finishes the proof. □

1.9 Remark. Regarding the condition that the set $\mathcal{P}^{\mathrm{o}}(F)$ should have density 1 (possibly after first replacing F by a finite extension), we can say the following. The condition is satisfied if $\dim(X) \leq 2$; see [46, Corollary 2.9]. Also, it is satisfied if X is of CM-type. It was conjectured by Serre in [59] that $\mathcal{P}^{\mathrm{o}}(F)$ has density 1 for all abelian varieties over a number field F, where F should be taken large enough such that the groups G_ℓ are connected. To our knowledge, this is a deep problem.

There is another case where the condition is known to be satisfied. Suppose X is an abelian fourfold with $\mathrm{End}(X_{\overline{\mathbb{Q}}}) = \mathbb{Z}$. Then the Mumford-Tate group $\mathrm{MT}(X_\sigma)$

is either isomorphic to CSp_8 (over $\mathbb{Q}$) or it is isogenous (over $\overline{\mathbb{Q}}$) to $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$. By Mumford's example in [42] both possibilities occur (see also the discussion in [43, Section 1.4], where it is shown that both possibilities occur for abelian varieties defined over a number field). For the group G_ℓ we have the same two possibilities (with $\mathbb{Q}$ replaced by $\mathbb{Q}_\ell$). We see that in this case we can not determine the groups MT and G_ℓ only by knowing the endomorphism algebra of X. If G_ℓ is (isogenous to) a form of $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$ then the Mumford-Tate conjecture of course predicts that this is also the case for MT, but as yet this remains unproved. For abelian varieties of dimension ≤ 4 this is in fact the only type of example where the Mumford-Tate conjecture is not known to be true. This fact, probably known to some experts for a long time (cf. Mumford's remark "...what seems to be the *only* family of this type...", [42, p. 349]) is proved in [39] and [40].

It was shown in Noot's PhD thesis [43] that if X is an abelian fourfold over a number field F such that $G_\ell \otimes \overline{\mathbb{Q}}_\ell \sim \mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$, then the set $\mathcal{P}^{\mathrm{o}}(F)$ (for F sufficiently large, as always) has density 1. If we have a principal polarization λ and a level n structure θ on X then, by our theorem, the Zariski closure Z of the set $\{x_v^{\mathrm{can}} \mid v \in \mathcal{P}^{\mathrm{o}}(F), p_v \nmid n\}$ either is the full moduli space $\mathsf{A}_{4,1,n} \otimes \mathbb{Q}$ (which would contradict the Mumford-Tate conjecture), or $Z_{\mathbb{C}}$ is the union of a number of (conjugated) Shimura curves in $\mathsf{A}_{4,1,n} \otimes \mathbb{C}$ and a finite number of CM-points.

More generally, for every $m \geq 1$ there exist abelian varieties X of dimension 4^m, defined over a number field F, such that $G_\ell \otimes \overline{\mathbb{Q}}_\ell \sim \mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^{2m+1}$. This was discussed in [62]. Noot showed in [45] that for such abelian varieties X the density of $\mathcal{P}^{\mathrm{o}}(F)$ is 1 for F large enough.

1.10 Remark. In the statement of the theorem we must allow a finite number of "exceptional" CM-points $s_1, \ldots, s_r$. For example, let X_0 be an ordinary abelian variety over a finite field, and let X be a quasi-canonical lifting which is not canonical. The density condition in the theorem is satisfied (cf. the preceding remark). The moduli point x^{can} of the canonical lifting of X_0 will occur as one of the exceptional points s_i.

§2 The canonical lifting of a moduli point

2.1 So far we only considered canonical liftings of ordinary abelian varieties over a finite field. However, we can associate a moduli point $x^{\mathrm{can}} \in \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ to any point $x \in (\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p)^{\mathrm{o}}$. Namely, write $\kappa(x)$ for the residue field of x, and let $\overline{\kappa(x)}$ be an algebraic closure. Since x is an ordinary moduli point, we have a canonical lifting of $s_x\colon \mathrm{Spec}(\overline{\kappa(x)}) \to \mathsf{A}_{g,1,n}$ to a section $s\colon \mathrm{Spec}(W(\overline{\kappa(x)})) \to \mathsf{A}_{g,1,n}$, and we define x^{can} as the image under s of the generic point of $\mathrm{Spec}(W(\overline{\kappa(x)}))$. The point x^{can} is easily seen to be independent of any choices.

In order to understand the behaviour of the canonical lifting under specialization, we have to generalize our notion of canonical lifting to abelian schemes over a perfect ring. This is done as follows.

Let R be a domain of characteristic $p > 0$ with fraction field K. Write K^{perf} for the perfect closure of K, and let R^{perf} be the integral closure of R in K^{perf}, which is a perfect closure of R. The ring $W(R^{\mathrm{perf}})$ of Witt vectors is a domain, complete and separated for the p-adic topology, and $W(R^{\mathrm{perf}})/p \cong R^{\mathrm{perf}}$. Suppose $X_0 \to \mathrm{Spec}(R)$ is an ordinary abelian scheme. By extending scalars we get an ordinary abelian scheme X over $\mathrm{Spec}(R^{\mathrm{perf}})$, and because R^{perf} is a perfect ring of characteristic p, the p-divisible group $X[p^\infty]$ is the direct sum $X[p^\infty] = X[p^\infty]_\mu \oplus X[p^\infty]_{\text{ét}}$ of a toroidal and an étale part. These summands each have a unique lifting to a p-divisible group, say G_μ and $G_{\text{ét}}$ respectively, over $\mathrm{Spec}(W(R^{\mathrm{perf}}))$, using [27, IV.18.3.4] and Cartier duality. Applying the Serre-Tate theorem III.1.6 we get a lifting X^{can} of X over $\mathrm{Spec}(W(R^{\mathrm{perf}}))$ whose p-divisible group is $G_\mu \oplus G_{\text{ét}}$.

This construction is functorial in the obvious sense. For example, if $\mathfrak{m} \subset R$ is a maximal ideal then the quotient homomorphism $R \twoheadrightarrow \kappa = R/\mathfrak{m}$ naturally extends to a homomorphism $R^{\mathrm{perf}} \twoheadrightarrow \kappa^{\mathrm{perf}}$ and we get a canonical map $W(R^{\mathrm{perf}}) \to W(\kappa^{\mathrm{perf}})$. It is clear from the construction that $X^{\mathrm{can}} \otimes_{W(R^{\mathrm{perf}})} W(\kappa^{\mathrm{perf}})$ is the canonical lifting of $X_0 \otimes_R \kappa$. Likewise, $X^{\mathrm{can}} \otimes_{W(R^{\mathrm{perf}})} W(K^{\mathrm{perf}})$ is the canonical lifting of $X_0 \otimes_R K$.

2.2 Lemma. *Let x, y be points of $(\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p)^{\mathrm{o}}$ such that x specializes to y. Then x^{can} specializes to y^{can}.*

Proof. (See also [50, Proof of Lemma 1.3].) Let $\mathcal{O}_y$ be the local ring of $\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p$ at y, and let $\mathfrak{p}_x \subset \mathcal{O}_y$ be the prime ideal corresponding to the point x. Let $R = \mathcal{O}_y/\mathfrak{p}_x$, then we have an ordinary abelian scheme X over R^{perf}, and, as just explained, we can form a canonical lifting X^{can} of X over $\mathrm{Spec}(W(R^{\mathrm{perf}}))$. The lemma readily follows from the functoriality of this construction (as explained above). □

2.3 Lemma. *Let W be a p-adically complete and separated domain such that $p \in W$ is prime. Let I be an index set, and let $\{\mathfrak{p}_\alpha \subset W \mid \alpha \in I\}$ be a collection of prime ideals such that $p \notin \mathfrak{p}_\alpha$. Assume that the intersection of the ideals $\mathfrak{q}_\alpha = (\sqrt{p + \mathfrak{p}_\alpha} \bmod p)$ in W/p is the zero ideal. Then the set $\{\mathfrak{p}_\alpha \mid \alpha \in I\}$ is Zariski dense in* $\mathrm{Spec}(W)$.

Proof. If $f \in \cap_{\alpha \in I} \mathfrak{p}_\alpha$ then $(f \bmod p) \in \cap_{\alpha \in I} \mathfrak{q}_\alpha = (0)$, hence $f = p \cdot f'$ for some $f' \in W$. Since $p \notin \mathfrak{p}_\alpha$ we have $f' \in \cap_{\alpha \in I} \mathfrak{p}_\alpha$, and by induction we then see that $f \in p^n \cdot W$ for every n. As W is p-adically separated, this implies $f = 0$. □

2.4 Proposition. *Let $x \in (\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p)^{\mathrm{o}}$, and define $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ as the Zariski closure of its canonical lifting x^{can}. Then Z is a subvariety of Hodge type.*[1]

Proof. Let $Y \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{F}_p$ be the Zariski closure of x, and consider the set $\mathcal{Y}$ of closed ordinary points of Y. If $Z' \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ is the Zariski closure of the set $\{y^{\mathrm{can}} \mid y \in \mathcal{Y}\}$ then by Lemma 2.2 we have $Z' \subseteq Z$. First we show that Z and Z' are in fact equal.

Let $U = \mathrm{Spec}(B) \subset (\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p)^{\mathrm{o}}$ be an affine open subscheme with $x \in U$. Write $C = U \cap Y = \mathrm{Spec}(B/J)$, then C is irreducible and $x \in C$. The ring $R = B/J$ is a domain of finite type over $\mathbb{F}_p$. As above, let R^{perf} be a perfect closure of R, let $W(R^{\mathrm{perf}})$ be its ring of Witt vectors, and let $s^{\mathrm{can}}\colon \mathrm{Spec}(W(R^{\mathrm{perf}})) \to \mathsf{A}_{g,1,n} \otimes \mathbb{Z}_p$ be the canonical lifting of $s\colon \mathrm{Spec}(R^{\mathrm{perf}}) \to (\mathsf{A}_{g,1,n} \otimes \mathbb{F}_p)^{\mathrm{o}}$. If $\mathfrak{m} \subset R^{\mathrm{perf}}$ is a maximal ideal with quotient field $k = R^{\mathrm{perf}}/\mathfrak{m}$, then the morphism $g\colon \mathrm{Spec}(k) \to \mathrm{Spec}(R^{\mathrm{perf}})$ lifts to $W(g)\colon \mathrm{Spec}(W(k)) \to \mathrm{Spec}(W(R^{\mathrm{perf}}))$, and $s^{\mathrm{can}} \circ W(g)$ is the canonical lifting of $s \circ g$.

Let $\{\mathfrak{m}_\alpha \mid \alpha \in I\}$ be the set of maximal ideals of R^{perf}. For each $\alpha \in I$ the kernel of $W(R^{\mathrm{perf}}) \to W(R^{\mathrm{perf}}/\mathfrak{m}_\alpha)$ is a prime ideal $\mathfrak{p}_\alpha \subset W(R^{\mathrm{perf}})$. Clearly, the collection $\{\mathfrak{p}_\alpha \mid \alpha \in I\}$ satisfies the assumptions of the previous lemma, and therefore it is Zariski dense in $\mathrm{Spec}(W(R^{\mathrm{perf}}))$. By construction, every $\mathfrak{p}_\alpha$ maps into Z' under s^{can}. It follows that x^{can} also maps into Z', hence $Z = Z'$.

We thus have an irreducible algebraic subvariety $Z \hookrightarrow \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ with a dense collection of CM-points (namely the points y^{can}) which are all canonical at some prime in characteristic p. Applying Corollary III.3.12 we conclude that the model $\mathcal{Z}$ of Z over $\mathbb{Z}_p$ is formally linear at some of its ordinary points, and by Theorem III.5.2 we conclude that Z is of Hodge type. □

[1] This result was obtained independently by M. Nori (unpublished).

Our final results are joint work with A.J. de Jong and F. Oort. The results were announced in [47], where also a sketch of the arguments was given. We keep the above notations, i.e., we fix an integer $n \geq 3$ and we consider an ordinary (but not necessarily closed) moduli point $x \in \mathsf{A}_{g,1,n} \otimes \mathbb{F}_p$. The problem that we are interested in is to compare

$$\mathrm{tr.deg.}_{\mathbb{F}_p}\kappa(x) = \dim(\{x\}^{\mathrm{Zar}}) \quad \text{and} \quad \mathrm{tr.deg.}_{\mathbb{Q}}\kappa(x^{\mathrm{can}}) = \dim(\{x^{\mathrm{can}}\}^{\mathrm{Zar}})\,.$$

We have an inequality $\mathrm{tr.deg.}_{\mathbb{F}_p}\kappa(x) \leq \mathrm{tr.deg.}_{\mathbb{Q}}\kappa(x^{\mathrm{can}})$. Our result shows that in general the two numbers are not equal. Before we state the precise result, we introduce some notations and we formulate a lemma.

Let R be a ring such that n is invertible in R. Given a morphism $f\colon S \to \mathsf{A}_{g,1,n} \otimes \mathrm{Spec}(R)$ of schemes over R, we simply write X_S for the corresponding abelian scheme over S, if it is clear which morphism f we take. Let $\bar{s}$ be a geometric point of S, and let ℓ be a prime number which is invertible in R. The polarization on $X_{\bar{s}}$ induces a non-degenerate alternating bilinear form φ_ℓ on $T_\ell X_{\bar{s}}$, and the image of the monodromy representation

$$\rho_S\colon \pi_1(S, \bar{s}) \longrightarrow \mathrm{Aut}(T_\ell X_{\bar{s}})$$

is an ℓ-adic Lie subgroup $\mathcal{G}_\ell = \mathcal{G}_\ell(S)$ of $\mathrm{CSp}(T_\ell X_{\bar{s}}, \varphi_\ell)$. Via the choice of a symplectic basis for $T_\ell X_{\bar{s}}$ we can identify $\mathcal{G}_\ell$ with a subgroup of $\mathrm{CSp}_{2g}(\mathbb{Z}_\ell)$. If S is connected then, up to conjugation, the group $\mathcal{G}_\ell(S)$ is independent both of the chosen basis and the choice of the base point $\bar{s}$.

If x is a point of $\mathsf{A}_{g,1,n}$ then we write $\mathcal{G}_\ell(x)$ for $\mathcal{G}_\ell(\mathrm{Spec}(\kappa(x)))$. Write $S = \{x\}^{\mathrm{Zar}}$ for the Zariski closure of $\{x\}$ inside $\mathsf{A}_{g,1,n}$, then the monodromy representation $\rho_{\mathrm{Spec}(\kappa(x))}$ factors through ρ_S, hence $\mathcal{G}_\ell(x) = \mathcal{G}_\ell(\{x\}^{\mathrm{Zar}})$. From this we see that if x specializes to a point y, then $\mathcal{G}_\ell(y)$ is conjugated to a subgroup of $\mathcal{G}_\ell(x)$.

2.5 Lemma. *Given a positive integer g and two different prime numbers p and ℓ, not dividing n, there exists an irreducible curve $C \subset \mathsf{A}_{g,1,n} \otimes \overline{\mathbb{F}}_p$ such that C meets the ordinary locus $(\mathsf{A}_{g,1,n} \otimes \overline{\mathbb{F}}_p)^{\mathrm{o}}$ and $\mathcal{G}_\ell(C) = \mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$.*

Proof. Choose a primitive nth root of unity in $\overline{\mathbb{F}}_p$. We will construct C as a subvariety of the moduli space $\mathsf{A}_{g,1,(n)} \otimes_{\mathbb{Z}[\zeta_n,1/n]} \overline{\mathbb{F}}_p$ of abelian varieties with a symplectic level n structure, which can be identified with an irreducible component of $\mathsf{A}_{g,1,n} \otimes_{\mathbb{Z}} \overline{\mathbb{F}}_p$.

Let W be the ring of Witt vectors of $\overline{\mathbb{F}}_p$. By the work of Faltings and Chai, see [25, IV, Theorem 6.7 and V, Theorem 5.8], there exists a scheme $\overline{\mathsf{A}}$, smooth and projective over $\mathrm{Spec}(W)$, such that $\mathsf{A} = \mathsf{A}_{g,1,(n)} \otimes W$ is identified with the complement of a divisor $D \subset \overline{\mathsf{A}}$ with normal crossings relative to $\mathrm{Spec}(W)$, and such that the universal abelian scheme over A extends to a semi-abelian scheme over $\overline{\mathsf{A}}$. Choose a projective embedding $\overline{\mathsf{A}} \hookrightarrow \mathbb{P}^N_W$ over $\mathrm{Spec}(W)$. We consider an intersection of $\overline{\mathsf{A}}_{\overline{\mathbb{F}}_p}$ with hyperplanes

$$\overline{C} = \overline{\mathsf{A}}_{\overline{\mathbb{F}}_p} \cap H_1 \cap \ldots \cap H_{g(g+1)/2-1} \subset \mathbb{P}^N_{\overline{\mathbb{F}}_p}$$

such that (i) $\overline{C}$ is irreducible and smooth, (ii) $\overline{C}$ intersects the divisor $D_{\overline{\mathbb{F}}_p}$ transversally at smooth points and (iii) $\overline{C} \cap (\mathsf{A}_{g,1,(n)} \otimes \overline{\mathbb{F}}_p)^{\mathrm{o}}$ is not empty. Note that such intersections exist by Bertini's Theorem (see [28, II, Theorem 8.18]). Write $C = \overline{C} \cap (\mathsf{A}_{g,1,(n)} \otimes \overline{\mathbb{F}}_p)$. (For $g = 1$ we have $C = \mathsf{A}_{1,1,(n)} \otimes \overline{\mathbb{F}}_p$.) We claim that C has the required properties.

To see this, we choose arbitrary hyperplanes $\mathcal{H}_i \subset \mathbb{P}^N_W$ with $\mathcal{H}_i \otimes \overline{\mathbb{F}}_p = H_i$, and let

$$\overline{\mathcal{C}} = \overline{\mathsf{A}} \cap \mathcal{H}_1 \cap \ldots \cap \mathcal{H}_{g(g+1)/2-1}\,.$$

Then $\overline{\mathcal{C}}$ is a projective curve, which by [27, IV.6.8.7 and IV.17.5.1] and property (i) above is smooth over $\mathrm{Spec}(W)$. Moreover, $\overline{\mathcal{C}} \cap D \subset \overline{\mathcal{C}}$ is a divisor with normal crossings relative to $\mathrm{Spec}(W)$ (it is a union of sections).

Write $\mathcal{C} = \overline{\mathcal{C}} \backslash D$, let η be the generic point of $\mathrm{Spec}(W)$, and write $\bar{\eta}$: $\mathrm{Spec}(\overline{\kappa(\eta)}) \to \mathrm{Spec}(W)$ for the geometric point of $\mathrm{Spec}(W)$ that factors through η. By [26, Exposé XIII, 2.10] there is a specialization homomorphism

$$sp\colon \pi_1^{\mathrm{t}}(\mathcal{C}_{\bar{\eta}}) \to \pi_1^{\mathrm{t}}(C)$$

on tame fundamental groups (omitting base points from the notation). The representations ρ_C and $\rho_{\mathcal{C}_{\bar{\eta}}}$ factor through $\pi_1^{\mathrm{t}}(C)$ and $\pi_1^{\mathrm{t}}(\mathcal{C}_{\bar{\eta}})$ respectively, as follows from the fact that the universal abelian scheme over A extends to a semi-abelian scheme over $\overline{\mathsf{A}}$. From the definition of the specialization homomorphism and the fact that the monodromy representation comes from an abelian scheme over $\mathcal{C}$, it follows that $\rho_{\mathcal{C}_{\bar{\eta}}}$ also factors through sp. We conclude that $\mathcal{G}_\ell(\mathcal{C}_{\bar{\eta}}) \subseteq \mathcal{G}_\ell(C)$ (up to conjugation), so it suffices to prove that $\mathcal{G}_\ell(\mathcal{C}_{\bar{\eta}}) = \mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$. For this, in turn, it suffices to show that $\pi_1(\mathcal{C}_{\bar{\eta}})$ maps surjectively to $\pi_1(\mathsf{A}_{g,1,(n)} \otimes \overline{\kappa(\eta)})$, since $\mathcal{G}_\ell(\mathsf{A}_{g,1,(n)} \otimes \overline{\kappa(\eta)}) = \mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$. This last statement follows from [25, IV, 6.8], and our assumption that $\ell \nmid n$.

Consider curves $\overline{\Gamma} \subseteq \overline{\mathsf{A}}_{\mathbb{C}}$ which satisfy

$(*)$ $\overline{\Gamma}$ is a smooth complete intersection which intersects $D_{\mathbb{C}}$ transversally.

We are done if we show that for any such curve $\overline{\Gamma}$, the map $\pi_1(\Gamma) \to \pi_1(\mathsf{A}_{\mathbb{C}})$ on topological fundamental groups is surjective, writing $\mathsf{A}_{\mathbb{C}} = \mathsf{A}_{g,1,(n)} \otimes \mathbb{C}$ and $\Gamma = \overline{\Gamma} \cap \mathsf{A}_{\mathbb{C}}$. For a generic curve $\overline{\Gamma}$ this is true by [17, Lemma 1.4]. Now, any two curves $\overline{\Gamma}_0$, $\overline{\Gamma}_1$ satisfying $(*)$ can be connected by a continuous family $\overline{\Gamma}_t \subseteq \overline{\mathsf{A}}_{\mathbb{C}}$ of such curves ($t \in [0,1]$), and if $\pi_1(\Gamma_t) \to \pi_1(\mathsf{A}_{\mathbb{C}})$ is surjective for $t = 0$ then this holds for every $t \in [0,1]$. This proves the lemma.

Second proof. Using a stronger Bertini theorem, we can give a different proof of the lemma. First we remark that, for some fixed, sufficiently large integer m, it suffices to construct an irreducible curve C which intersects the ordinary locus and for which $\mathcal{G}_\ell(C)$ maps surjectively to $\mathrm{Sp}_{2g}(\mathbb{Z}/\ell^m)$. (In fact, for $\ell \neq 2$ we can take $m = 2$; for $\ell = 2$ we take $m = 3$. We omit the proof of this fact; a similar statement can be found in [60, Chap. IV, 3.4])

Consider the Galois covering

$$g\colon \mathsf{A}' = \mathsf{A}_{g,1,(\ell^m n)} \to \mathsf{A} = \mathsf{A}_{g,1,(n)}\,,$$

which has Galois group $\mathrm{Sp}_{2g}(\mathbb{Z}/\ell^m)$. Write $d = g(g+1)/2$, which of course is the dimension of A. By first choosing an embedding $\mathsf{A} \hookrightarrow \mathbb{P}^N$ and then projecting from a sufficiently general linear subvariety of codimension $d+1$, we can find an affine open subscheme $\mathsf{U} \subset \mathsf{A}$ for which there exists a finite morphism $f\colon \mathsf{U} \to \mathbb{A}^d$. Write U' for the inverse image of U in A'.

Starting from the morphism $f \circ g\colon \mathsf{U}' \to \mathbb{A}^d$ and applying [30, Theorem 6.3] $d-1$ times, we find a line $L \subset \mathbb{A}^d$ such that $(f \circ g)^{-1}(L)$ is an irreducible curve in U'. Let $C \subset \mathsf{A}$ and $C' \subset \mathsf{A}'$ be the Zariski closure of $f^{-1}(L)$ and $(f \circ g)^{-1}(L)$, respectively. The diagram

$$\begin{array}{ccc} C' & \hookrightarrow & \mathsf{A}' \\ {\scriptstyle g_{|C'}} \downarrow & & \downarrow {\scriptstyle g} \\ C & \hookrightarrow & \mathsf{A} \end{array}$$

is Cartesian and C and C' are irreducible curves. It follows that $g_{|C'}\colon C' \to C$ is a Galois covering with group $\mathrm{Sp}_{2g}(\mathbb{Z}/\ell^m)$. By what was said before, this implies that C has the required properties. □

2.6 Theorem. (A.J. de Jong, B.M., F. Oort) *Given a prime number p not dividing n and an integer $g \geq 1$, there exists a field k of characteristic p and a k-valued* *ordinary moduli point* $x \in \mathsf{A}^{\mathrm{o}}_{g,1,n}(k)$ *such that*

$$\mathrm{tr.deg.}_{\mathbb{F}_p}\kappa(x) = 1 \quad \textit{and} \quad \mathrm{tr.deg.}_{\mathbb{Q}}\kappa(x^{\mathrm{can}}) = \frac{g(g+1)}{2}\,.$$

Proof. We take a curve $C \subset \mathsf{A}_{g,1,n} \otimes \overline{\mathbb{F}}_p$ as in the lemma. For $x \in \mathsf{A}_{g,1,n} \otimes \mathbb{F}_p$ we take the generic point of C, which is an ordinary moduli point. Clearly, $\mathrm{tr.deg.}_{\mathbb{F}_p}\kappa(x) = 1$.

By Proposition 2.4, the Zariski closure Z of the point $x^{\mathrm{can}} \in \mathsf{A}_{g,1,n} \otimes \mathbb{Q}$ is a subvariety of Hodge type. We are done if we show that it is equal to $\mathsf{A}_{g,1,n} \otimes \mathbb{Q}$. To see this we use that, by construction, the monodromy representation of Z has a "large" image.

Write $\mathcal{Z}$ for the Zariski closure of Z over $\mathrm{Spec}(\mathbb{Z}_p)$, and let $\bar{\eta}$ be a geometric point of $\mathcal{Z}$ which factors through the generic point η. Then η specializes to x, and as $\mathcal{G}_\ell(x) \supseteq \mathrm{Sp}_{2g}(\mathbb{Z}_\ell)$ (by construction of C) we have $\mathrm{Sp}_{2g}(\mathbb{Z}_\ell) \subseteq \mathcal{G}_\ell(\eta) = \mathcal{G}_\ell(Z)$.

Next we choose a number field F such that there exists an F-rational point $z\colon \mathrm{Spec}(F) \to Z$. If $\bar{z}$ is a geometric point factoring through z then we have a homomorphism $z_*\colon \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to \pi_1(Z, \bar{z})$, which is a section on $\mathrm{Gal}(\overline{\mathbb{Q}}/F)$ of the natural homomorphism $\pi_1(Z, \bar{z}) \to \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Let M be the generic Mumford-Tate group on Z, and write $\overline{Z} = Z \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}$. The homomorphisms $\rho_{\overline{Z}}\colon \pi_1(\overline{Z}, \bar{z}) \hookrightarrow \pi_1(Z, \bar{z}) \to \mathrm{CSp}(T_\ell X_{\bar{z}}, \varphi_\ell)$ and $\rho \circ z_*\colon \mathrm{Gal}(\overline{\mathbb{Q}}/F) \to \mathrm{CSp}(T_\ell X_{\bar{z}}, \varphi_\ell)$ both factor through $M(\mathbb{Q}_\ell)$. We conclude that there is a subgroup of finite index $\pi \subseteq \pi_1(Z, \bar{z})$ such that $\rho(\pi) \subset M(\mathbb{Q}_\ell) \subseteq \mathrm{CSp}_{2g}(\mathbb{Q}_\ell)$. Since $M \otimes \mathbb{Q}_\ell$ is an algebraic subgroup of $\mathrm{CSp}_{2g} \otimes \mathbb{Q}_\ell$ with $\mathbb{G}_{\mathrm{m}} \cdot \mathrm{Id} \subset M$, and since $\mathrm{Sp}_{2g}(\mathbb{Z}_\ell) \subseteq \mathcal{G}_\ell(Z)$ we conclude that $M = \mathrm{CSp}_{2g} \otimes \mathbb{Q}$ and $Z = \mathsf{A}_{g,1,n}$. This finishes the proof. $\square$

Bibliography

[1] Y. ANDRÉ, *G-functions and geometry*, Aspects of Mathematics, Vol. E13, Vieweg, Braunschweig, 1989.

[2] ——, *Mumford-Tate groups of mixed Hodge structures and the theorem of the fixed part*, Compositio Math., 82 (1992), pp. 1–24.

[3] S. ARAKELOV, *Families of algebraic curves with fixed degeneracies*, Math. USSR Izvestia, 5 (1971), pp. 1277–1302.

[4] A. BOREL, *Introduction aux groupes arithmétiques*, Actualités scientifiques et industrielles, volume 1341, Hermann, Paris, 1969.

[5] ——, *Linear algebraic groups (second edition)*, Graduate texts in mathematics, Vol. 126, Springer-verlag, New York, 1991.

[6] M. BOROVOI, *The Hodge group and the algebra of endomorphisms of an abelian variety*, in Questions of group theory and homological algebra, A. Onishchik, ed., Yaroslav, 1981. (In Russian.)

[7] N. BOURBAKI, *Groupes et algèbres de Lie, Chapitres 7 et 8 (Nouveau tirage)*, Éléments de mathématique, Masson, Paris, 1990.

[8] E. CATTANI, P. DELIGNE, and A. KAPLAN, *On the locus of Hodge classes*, manuscript, 1994.

[9] C. CHEVALLEY, *Théorie des groupes de Lie, Tome II: Groupes Algébriques*, Actualités scientifiques et industrielles, volume 1152, Hermann & C^ie^, Paris, 1952.

[10] W. CHI, *ℓ-adic and λ-adic representations associated to abelian varieties defined over number fields*, American J. of Math., 114 (1992), pp. 315–353.

[11] A. J. de Jong, *A result on formal linearity*, Preprint nr. 911, University of Utrecht, April 1995.

[12] A. J. de Jong and R. Noot, *Jacobians with complex multiplication*, in Arithmetic Algebraic Geometry, G. van der Geer, F. Oort and J. Steenbrink, eds., Progress in Math., Vol. 89, Birkhäuser, Boston, 1991, pp. 177–192.

[13] A. J. de Jong and F. Oort, *On extending families of curves*, in preparation.

[14] P. Deligne, *Travaux de Shimura*, in Séminaire Bourbaki, Exposé 389, Février 1971, pp. 123–165, .

[15] ——, *La conjecture de Weil pour les surfaces K3*, Invent. Math., 15 (1972), pp. 206–226.

[16] ——, *Théorie de Hodge II*, Publ. Math. de l'I.H.E.S., 40 (1972), pp. 5–57.

[17] ——, *Le groupe fondamental du complément d'une courbe plane n'ayant que des points doubles ordinaires est abélien (d'après W. Fulton)*, in Séminaire Bourbaki, Exposé 543, Novembre 1979, pp. 1–10.

[18] ——, *Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques*, in Automorphic Forms, Representations, and *L*-functions, Part 2, A. Borel and W. Casselman, eds., Proc. of Symp. in Pure Math., Vol. XXXIII, American Mathematical Society, 1979, pp. 247–290.

[19] ——, *Hodge cycles on abelian varieties (Notes by J.S. Milne)*, in Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Mathematics 900, Springer-Verlag, Berlin, 1982, pp. 9–100.

[20] ——, *A quoi servent les motifs?*, in Motives, Part 1, U. Jannsen, S. Kleiman, and J-P. Serre, eds., Proc. of Symp. in Pure Math., Vol. 55, American Mathematical Society, 1994, pp. 143–161.

[21] P. Deligne and L. Illusie, *Cristaux ordinaires et coordonnées canoniques (with an appendix by N. Katz)*, Exposé V in Surfaces algébriques, J. Giraud, L. Illusie, and M. Raynaud, eds., Lecture Notes in Mathematics 868, Springer-Verlag, Berlin, 1981, pp. 80–137.

[22] P. Deligne and J. S. Milne, *Tannakian categories*, in Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Mathematics 900, Springer-Verlag, Berlin, 1982, pp. 101–228.

[23] M. Demazure, A. Grothendieck, et al., *SGA3; Schémas en groupes I*, Lecture Notes in Mathematics 151, Springer-Verlag, Berlin, 1970.

[24] G. Faltings, *Arakelov's theorem for abelian varieties*, Inventiones Math., 73 (1983), pp. 337–347.

[25] G. Faltings and C-L. Chai, *Degeneration of abelian varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 22, Springer-Verlag, Berlin, 1990.

[26] A. Grothendieck, *SGA1; Revêtements étales et groupe fondamental*, Lecture Notes in Mathematics 224, Springer-Verlag, Berlin, 1971.

[27] A. Grothendieck and J. Dieudonné, *Éléments de géométrie algébrique*, Publ. Math. de l'I.H.E.S., 4, 8, 11, 17, 20, 24 and 32 (1960–67).

[28] R. Hartshorne, *Algebraic geometry*, Graduate texts in mathematics, Vol. 52, Springer-Verlag, New York, 1977.

[29] S. Helgason, *Differential Geometry, Lie groups, and symmetric spaces*, Academic Press, Inc., Orlando, 1978.

[30] J. P. Jouanolou, *Théorèmes de Bertini et Applications*, Progress in Math., Vol. 42, Birkhäuser, Boston, 1983.

[31] N. M. Katz, *Serre-Tate local moduli*, Exposé $\mathrm{V}^{\mathrm{bis}}$ in Surfaces algébriques, J. Giraud, L. Illusie, and M. Raynaud, eds., Lecture Notes in Mathematics 868, Springer-Verlag, Berlin, 1981, pp. 138–202.

[32] S. Kobayashi and K. Nomizu, *Foundations of differential geometry I, II*, Wiley (Interscience), New York, 1963 and 1969.

[33] J. Lubin, J-P. Serre, and J. Tate, *Elliptic curves and formal groups*, in Lecture Notes Woods Hole Summer Institue, American Mathematical Society, 1964.

[34] H. Matsumura, *Commutative ring theory*, Cambridge studies in advanced mathematics, Vol. 8, Cambridge University Press, Cambridge, 1986.

[35] W. Messing, *The crystals associated to Barsotti-Tate groups: with applications to abelian schemes*, Lecture Notes in Mathematics 264, Springer-Verlag, Berlin, 1972.

[36] J. S. Milne, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, in Automorphic forms, Shimura varieties, and L-functions, L. Clozel and J. S. Milne, eds., Persp. in Math., Vol. 10(I), Academic Press, Inc., 1990, pp. 283–414.

[37] ——, *The points on a Shimura variety modulo a prime of good reduction*, in The Zeta Functions of Picard Modular Surfaces, R. P. Langlands and D. Ramakrishnan, eds., Les Publications CRM, Montréal, 1992, pp. 151–253.

[38] ——, *Shimura varieties and motives*, in Motives, Part 2, U. Jannsen, S. Kleiman, and J-P. Serre, eds., Proc. of Symp. in Pure Math., Vol. 55, American Mathematical Society, 1994, pp. 447–523.

[39] B. Moonen and Yu. Zarhin, *Hodge classes and Tate classes on simple abelian fourfolds*, Duke Math. Journal, 77 (1995), pp. 553–581.

[40] ——, *Hodge classes and Tate classes on abelian fourfolds*, in preparation.

[41] D. Mumford, *Families of abelian varieties*, in Algebraic groups and discontinuous subgroups, A. Borel and G. D. Mostow, eds., Proc. of Symp. in Pure Math., Vol. 9, American Mathematical Society, 1966, pp. 347–351.

[42] ——, *A note of Shimura's paper "Discontinuous groups and abelian varieties"*, Math. Ann., 181 (1969), pp. 345–351.

[43] R. Noot, *Hodge classes, Tate classes, and local moduli of abelian varieties*, PhD thesis, University of Utrecht, 1992.

[44] ——, *Models of Shimura varieties in mixed characteristic*, Prépublication 95-07, Institut de recherche Mathématique de Rennes, Février 1995. To appear in J. of Alg. Geometry.

[45] ——, *Abelian varieties — Galois representations and properties of ordinary reduction*, to appear in Compositio Math. (1995).

[46] A. Ogus, *Hodge cycles and crystalline cohomology*, in Hodge Cycles, Motives, and Shimura Varieties, Lecture Notes in Mathematics 900, Springer-Verlag, Berlin, 1982, pp. 357–414.

[47] F. Oort, *Canonical liftings and dense sets of CM-points*, Symposium Cortona, October 1994.

[48] ——, *Finite group schemes, local moduli for abelian varieties, and lifting problems*, Compositio Math., 23 (1971), pp. 265–296.

[49] ——, *Endomorphism algebras of abelian varieties*, in Algebraic geometry and commutative algebra in honor of Masayoshi Nagata, Vol. II, H. Hijikata et al., eds., Tokyo, 1988, Kinokuniya Company, Ltd., pp. 469–502.

[50] F. Oort and T. Sekiguchi, *The canonical lifting of an ordinary Jacobian variety need not be a Jacobian variety*, J. Math. Soc. Japan, 38 (1986), pp. 427–437.

[51] C. Peters, *Rigidity for variations of Hodge structure and Arakelov-type finiteness theorems*, Compositio Math., 75 (1990), pp. 113–126.

[52] R. Pink, *Arithmetical compactification of mixed Shimura varieties*, PhD thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn, 1989.

[53] K. A. Ribet, *Galois action on division points of abelian varieties with many real multiplications*, American J. of Math., 98 (1976), pp. 751–804.

[54] M.-H. Saito, *Classification of non-rigid families of abelian varieties*, Tôhoku Math. Journal, 45 (1993), pp. 159–189.

[55] I. Satake, *Algebraic structures of symmetric domains*, Publications of the Mathematical Society of Japan, Vol. 14, Iwanami Shoten and Princeton University Press, 1980.

[56] M. Schlessinger, *Functors of Artin rings*, Trans. Amer. Math. Soc., 130 (1968), pp. 208–222.

[57] W. Schmid, *Variation of Hodge structure: the singularities of the period mapping*, Inventiones math., 22 (1973), pp. 211–319.

[58] J-P. Serre, Letters to Ken Ribet, 1 January and 29 January, 1981.

[59] ——, Letter to John Tate, 2 January, 1985.

[60] ——, *Représentations l-adiques*, in Kyoto Symposium on algebraic number theory, Japan Soc. for the promotion of science, 1977, pp. 177–193. (Œuvres 112)

[61] ——, *Abelian ℓ-adic representations and elliptic curves*, Advanced Book Classics, Addison-Wesley Publishing Company, Inc., 1989.

[62] S. Tankeev, *On algebraic cycles on surfaces and abelian varieties*, Math. USSR Izvestia, 18 (1982), pp. 349–380.

[63] J. Tate, *Endomorphisms of abelian varieties over finite fields*, Inventiones math., 2 (1966), pp. 134–144.

[64] ——, *Classes d'isogénie des variétés abéliennes sur un corps fini (d'après T. Honda)*, in Séminaire Bourbaki, Exposé 352, Novembre 1968, pp. 95–110.

[65] A. Vasiu, *Integral canonical models for Shimura varieties of Hodge type*, PhD thesis, Princeton University, November 1994.

[66] Yu. Zarhin, *Weights of simple Lie algebras in the cohomology of algebraic varieties*, Math. USSR Izvestia, 24 (1985), pp. 245–281.

Samenvatting

Voor de studie van abelse variëteiten is het van belang dat we beschikken over een moduliruimte. We schrijven $\mathsf{A}_{g,1,n}$ voor de moduliruimte van g-dimensionale abelse variëteiten (met een hoofdpolarisatie en een niveau n structuur). We vinden vaak interessante deelvariëteiten van $\mathsf{A}_{g,1,n}$ door de locus van punten te beschouwen waarbij de corresponderende abelse variëteiten een bepaalde extra structuur of extra symmetrie hebben. We kunnen bijvoorbeeld abelse variëteiten bestuderen die een gegeven ring van endomorfismen toelaten. In karakteristiek 0 is het een directe generalisatie hiervan om abelse variëteiten te beschouwen waarop bepaalde Hodgeklassen bestaan. Dit geeft aanleiding tot zogeheten Shimuravariëteiten binnen de moduliruimte van abelse variëteiten. De irreducibele componenten van deze Shimuravariëteiten noemen we deelvariëteiten van Hodge type.

Deze deelvariëteiten van Hodge type zijn zeer rijk aan structuur. In dit proefschrift voegen we daar nieuwe inzichten aan toe. We bewijzen dat de deelvariëteiten van Hodge type gekarakteriseerd worden door bepaalde lineariteitseigenschappen. Over de complexe getallen is dit het "totaal geodetisch" zijn. De belangrijkste van onze resultaten betreffen een analogon hiervan in gemengde karakteristiek; we noemen dit "formele lineariteit". Door Noot werd bewezen dat deelvariëteiten van Hodge type (in gemengde karakteristiek) formeel lineair zijn. We bewijzen dat deelvariëteiten van Hodge type door deze eigenschap gekarakteriseerd worden. Hierbij zal blijken dat de beide lineariteitseigenschappen directer verband houden dan men op het eerste gezicht zou vermoeden. We passen onze hoofdresultaten toe om, onder bepaalde extra aannamen, een vermoeden van Oort te bewijzen.

Beschouw, om het bovenstaande wat concreter te maken, een deelvariëteit van Hodge type $S \hookrightarrow \mathsf{A}_{g,1,n}$. Dan heeft S de volgende eigenschappen: (i) de speciale punten van S liggen dicht voor de Zariski-topologie, (ii) S is een totaal geodetische deelvariëteit, in de zin dat S overdekt wordt door een totaal geodetische deel-menigvoud van X en (iii) S is "formeel lineair" bij (de meeste) van zijn gewone modulipunten in karakteristiek $p > 0$ —een uitspraak die toegelicht zal worden.

In Hoofdstuk II van dit proefschrift richten we ons voornamelijk op de tweede eigenschap, waarbij we werken binnen een (algemene) Shimura variëteit $Sh_K(G, X)$. Het is niet waar dat elke totaal geodetische deelvariëteit $Z \hookrightarrow Sh_K(G, X)$ van Hodge type is. We bewijzen dat dit echter wel het geval is wanneer Z tenminste 1 speciaal punt bevat. Tevens geven we een beschrijving van totaal geodetische deelvariëteiten in het algemeen en passen we de gebruikte technieken toe om een beschrijving te geven van zogenaamde niet-starre families van abelse variëteiten.

In Hoofdstuk III bestuderen we de moduliruimte $\mathcal{A}_g$ van g-dimensionale abelse variëteiten over $\mathbb{Z}_p$. Als $x \in \mathcal{A}_g \otimes \mathbb{F}_p$ het modulipunt is van een gewone abelse variëteit over een eindig lichaam k, dan heeft de formele completering $\mathfrak{A}_x$ van $\mathcal{A}_g$ in x de structuur van een formele torus over de Wittring $W(k)$. We noemen een algebraïsche deelvariëteit $\mathcal{Z} \hookrightarrow \mathcal{A}_g$ met $x \in \mathcal{Z} \otimes \mathbb{F}_p$ "formeel lineair" in het punt x als de formele completering $\mathfrak{Z}_x \hookrightarrow \mathfrak{A}_x$ een formele subtorus is.

In het proefschrift [43] van R. Noot werd een tot dan toe onvermoede structuur aan het licht gebracht. Noot bewees dat deelvariëteiten $\mathcal{S} \hookrightarrow \mathcal{A}_g$ van Hodge type formeel lineair zijn in de gewone gesloten punten van $\mathcal{S} \otimes \mathbb{F}_p$. (Hierbij staan we onszelf toe een vereenvoudigde voorstelling van zaken te geven.) We bewijzen hiervan een omkering: als een algebraïsche deelvariëteit $\mathcal{Z} \hookrightarrow \mathcal{A}_g$ formeel lineair is bij een gewoon gesloten punt, dan is $\mathcal{Z}$ van Hodge type.

Tevens bestuderen we de relatie van "formele lineariteit" met Zariski-dichte collecties van CM-punten. Onze belangstelling hiervoor heeft te maken met een vermoeden van Oort, dat de deelvariëteiten van Hodge type karakteriseert als de algebraïsche deelvariëteiten waarop de CM-punten dicht liggen. Wanneer $\mathcal{Z}$ ergens formeel lineair is, dan volgt daaruit vrij eenvoudig dat de CM-punten op $\mathcal{Z}$ dicht liggen in de Zariski-topologie. De omkering hiervan bewijzen we onder een extra aanname: stel er is een residukarakteristiek p en een Zariski-dichte collectie T van CM-punten op Z, zo dat elk van de corresponderende abelse variëteiten X_t canoniek is bij een plaats $\mathfrak{p}$ boven p (waarmee we bedoelen dat X_t de canonieke lifting is van zijn reductie $X_{t,\mathfrak{p}}$ bij $\mathfrak{p}$). Dan is er een gewoon punt $x \in \mathcal{Z} \otimes \mathbb{F}_p$ zo dat $\mathcal{Z}$ formeel lineair is bij x. Onze karakterisering toepassend concluderen we dat $\mathcal{Z}$ van Hodge type is.

In Hoofdstuk IV passen we deze resultaten toe om het vermoeden van Oort te bewijzen in een speciaal geval. We beginnen met een abelse variëteit X (met een hoofdpolarisatie en eventueel een niveaustructuur, die we hier steeds buiten beschouwing laten) die gedefinieerd is over een getallenlichaam F, dat we voldoende groot nemen. Schrijf $\mathcal{P}^{\mathrm{o}}$ voor de collectie van plaatsen van F waar X goede en gewone reductie heeft. We nemen aan dat $\mathcal{P}^{\mathrm{o}}$ dichtheid 1 heeft. Voor $v \in \mathcal{P}^{\mathrm{o}}$ duiden we het modulipunt van de canonieke lifting van X_v aan met $x_v^{\mathrm{can}} \in \mathcal{A}_g \otimes \mathbb{Q}$. Dan bewijzen we dat de Zariski-afsluiting van de verzameling $\{x_v^{\mathrm{can}} \mid v \in \mathcal{P}^{\mathrm{o}}\}$ gelijk is aan de kleinste deelvariëteit van Hodge type die het modulipunt van X bevat, eventueel nog verenigd met een eindig aantal "uitzonderlijke" CM-punten.

Tenslotte bestuderen we deelvariëteiten van de vorm $Z = \{x^{\mathrm{can}}\}^{\mathrm{Zar}} \subseteq \mathcal{A}_g \otimes \mathbb{Q}$, waarbij x een gewoon, maar niet noodzakelijk gesloten, modulipunt in karakteristiek p is. We tonen aan dat dergelijke deelvariëteiten Z van Hodge type zijn. Verder laten we zien (gezamenlijk werk met A.J. de Jong en F. Oort) dat er voorbeelden zijn waarbij $\mathrm{tr.deg.}_{\mathbb{F}_p} \kappa(x) = 1$ en $\dim(Z) = g(g+1)/2$.

Curriculum vitae

Ben Moonen werd op 24 augustus 1968 geboren te Geleen. Hij groeide op in Hoensbroek, waar hij ook de lagere en middelbare school doorliep. Een grote triomf, in november 1978, was het behalen van het zwemdiploma A. Met voetbal ging het minder goed.

Na het behalen van het diploma Atheneum B aan het St. Janscollege, begon hij in 1985 aan de (toen nog Rijks-) Universiteit Utrecht zijn studie wiskunde, met natuurkunde als bijvak. In juni 1986 slaagde hij cum laude voor het propaedeutisch examen. Na afstudeerwerk bij de hoogleraren Steenbrink (Nijmegen) en Oort, ontving hij in augustus 1991 zijn doctoraalbul (cum laude).

In september 1991 werd hij werkzaam als onderzoeker in opleiding (o.i.o.) bij de Utrechtse Vakgroep Wiskunde, binnen een door N.W.O. gefinancierd onderzoeksproject. Naast het doen van onderzoek behoorde het leiden van enkele werkcolleges tot de taken. Gedurende de eerste vijf maanden van 1994 was hij als Visiting Fellow verbonden aan Harvard University in Cambridge (V.S.).

Vanaf 1 oktober 1995 zal Ben Moonen verbonden zijn aan het "Graduiertenkolleg Algebraische Geometrie und Zahlentheorie" van de universiteit Münster. Met het oog hierop bestudeert hij thans voorzetsels en naamvallen in de Duitse taal.

Nawoord

Met dit proefschrift sluit ik een periode af waarin ik als o.i.o. werkzaam was bij de Vakgroep Wiskunde van de Universiteit Utrecht, binnen een onderzoeksproject dat gefinancierd werd door de Nederlandse organisatie voor Wetenschappelijk Onderzoek (N.W.O.) te Den Haag en dat werd geleid door Frans Oort en Jozef Steenbrink. Ik heb in deze tijd veel steun gehad van een aantal mensen, die ik daar hartelijk voor wil bedanken.

Promotor en copromotor van dit proefschrift, Frans Oort en Johan de Jong, hebben een grote invloed gehad op mijn wiskundig denken. Zij delen een ontembaar enthousiasme voor de wiskunde en een buitengewone vakbekwaamheid. Ik wil Frans niet alleen bedanken voor alles wat hij mij geleerd heeft, maar zeker ook voor zijn constante optimisme en vertrouwen, dat ik me tot voorbeeld wil nemen. Johan dank ik hartelijk voor de vele ideeën die het werk aan dit proefschrift op gang hielpen en hielden. Met veel plezier denk ik terug aan de vele discussies die we voerden—qua thematiek variërend van Ben & Jerry's ijs tot het onderwerp van dit proefschrift.

Ik bedank de andere wiskundigen van wie ik veel heb kunnen leren; met name noem ik Bert van Geemen, Rutger Noot, Yuri Zarhin, Bas Edixhoven en Jozef Steenbrink. Laatstgenoemde twee bedank ik samen met Eduard Looijenga en Thomas Zink voor hun bereidheid zitting te nemen in de beoordelingscommissie; in het bijzonder ben ik Bas Edixhoven en Thomas Zink zeer erkentelijk voor hun uitvoerige commentaar bij een eerdere versie van dit werk.

I thank Harvard University for its hospitality and the very stimulating atmosphere during my stay as a Visiting Fellow in the first half of 1994. I thank Kathy Barrett, not just for being a dear friend, but also for her help in translating this thesis from the language in which I wrote it into proper English.

Tot slot wil ik een aantal mensen bedanken die erg belangrijk voor me zijn, zonder dat dat veel met wiskunde van doen heeft. Mijn ouders, mijn vrienden, mijn familie en Linda; ieder met wie ik humor, muziek en gezelligheid kan delen en bij wie ik—

ook in de laatste drukke maanden—mijn proefschrift even kon vergeten: hartelijk bedankt!