Deligne's definition of a Shimura variety

Theo van den Bogaart

9/10/03

These are the notes of my two talks in the local Leiden seminar on Shimura varieties. Closely following Deligne's article [1], we will try to motivate his general definition of a Shimura variety, keeping in mind the examples we have studied in Robert's and Gabor's previous talks. (Notes of these talks are on http://www.math.leidenuniv.nl/~carls/sem_shim.shtml.) Please report errors (of which there are probably a lot) to me.

1 The object

In the previous talks, we looked at sets of the form

$$\mathrm{Sh}_K(G,X)(\mathbf{C}) = G(\mathbf{Q}) \setminus (X \times G(\mathbf{A}_f)/K)\,,$$

where for example in Gabor's talk $G = \mathrm{GL}_2$, $X = \mathbf{H}^{\pm}$ and $K = K_n$. In both talks $\mathrm{Sh}_K(G,X)(\mathbf{C})$ was realized as the set of $\mathbf{C}$-valued points of a (non-connected, algebraic) variety. We also want to do this, but without making specific choices for G, X and K. We begin by generalizing the relation between G and X.

Fix a two dimensional $\mathbf{R}$-vector space V. Recall from Robert's talk that a complex structure on V is an $\mathbf{R}$-algebra homomorphism from $\mathbf{C}$ to $\mathrm{End}(V)$. Robert proved that the set $\mathbf{H}^{\pm}$ (which is by definition $\mathbf{C} - \mathbf{R}$) is in bijective correspondence to the set of complex structures on V. Now restriction to $\mathbf{C}^*$ gives an injection

$$\mathrm{Hom}_{\mathbf{Alg}_{\mathbf{R}}}(\mathbf{C}, \mathrm{End}(V)) \hookrightarrow \mathrm{Hom}_{\mathbf{Gr}}(\mathbf{C}^*, \mathrm{GL}(V,\mathbf{R})).$$

Letting $\mathrm{GL}(V,\mathbf{R})$ act on itself by conjugation, the image of the above injection is a $\mathrm{GL}(V,\mathbf{R})$-orbit (from now on called a $\mathrm{GL}(V,\mathbf{R})$-*conjugacy class*); this is an easy case of the Noether-Skolem theorem. So with $G = \mathrm{GL}_V \simeq \mathrm{GL}_2$ and $X = \mathbf{H}^{\pm}$ (as in Gabor's talk), we conclude then that X corresponds to a $G(\mathbf{R})$-conjugacy class in $\mathrm{Hom}(\mathbf{C}^*, G(\mathbf{R}))$.

Here is a generalisation: Let G be a linear algebraic group over $\mathbf{Q}$, let X be a $G(\mathbf{R})$-conjugacy class in $\mathrm{Hom}(\mathbf{C}^*, G(\mathbf{R}))$ and let K be a compact open subgroup $G(\mathbf{A}_f)$. Then notice that (only as a set) $\mathrm{Sh}_K(G,X)(\mathbf{C})$ is well-defined. However, we have generalized too much! For there is no hope that

$\mathrm{Sh}_K(G, X)(\mathbf{C})$ is the set of $\mathbf{C}$-points of an algebraic variety. The object of this talk, therefore, is to impose some conditions (Deligne's axioms) on G, X, and K such that there does exist such a variety.

The rough idea is this. Recall from Robert's talk that a complex structures on V as mentioned before also corresponds to a decomposition $V \otimes \mathbf{C} = V^{-1,0} \oplus V^{0,-1}$ such that $\overline{V^{-1,0}} = V^{0,-1}$. This is a Hodge decomposition of weight -1 ! We will show that also in the general setting one can relate G and X to Hodge structures. The we impose certain natural conditions on these Hodge structures and see what this means for G and X. Here a 'natural condition' means something as 'Hodge structures coming from algebraic geometry satisfy this condition'.

2 Hodge structures

We start by recalling some theory from last semester's seminar on algebraic groups. Let k be a field, V a k-vector space of finite dimension and T a torus over k. Define the *character group* of T as $X^*(T) = \mathrm{Hom}(T_{k^{\mathrm{sep}}}, \mathbf{G}_{m,k^{\mathrm{sep}}})$. The Galois group $\mathrm{Gal}(k^{\mathrm{sep}}/k)$ acts on the character group via

$$(\sigma\chi)(t) = \sigma^{-1}\chi(\sigma(t)) \qquad (\sigma \in \mathrm{Gal}(k^{\mathrm{sep}}/k),\, \chi \in X^*(T),\, t \in T(k^{\mathrm{sep}})).$$

A representation $\rho\colon T \to \mathrm{GL}_V$ gives a decomposition

$$V \otimes k^{\mathrm{sep}} = \bigoplus_{\chi \in X^*(T)} V_\chi, \qquad \text{with} \quad v \in V_\chi \iff \forall t \in T(k^{\mathrm{sep}})\ \rho(t)v = \chi(t)v.$$

Conversely, given a decomposition as above, we get a representation of $T_{k^{\mathrm{sep}}}$ in $V \otimes k^{\mathrm{sep}}$. To get a representation over k, we need a 'descent datum' identifying $\sigma(V_\chi)$ with $V_{\sigma(\chi)}$ for every element σ in the Galois group of k^{sep}/k. Let us now apply this to the case $k = \mathbf{R}$.

Recall that a *Hodge structure* on an $\mathbf{R}$-vectorspace V is a decomposition

$$V_{\mathbf{C}} := V \otimes \mathbf{C} = \bigoplus_{p,q \in \mathbf{Z}} V^{p,q}, \qquad \text{with } \overline{V^{p,q}} = V^{q,p}.$$

Elements of $V^{p,q}$ are called of *weight* $p+q$. The Hodge structure is *pure of weight* n if all elements are of weight n, i.e., $V^{p,q} \neq 0$ implies $p+q = n$. Its *type* is defined as the set of those pairs of indices (p,q) such that $V^{p,q} \neq 0$.

Let $\mathbf{S}$ be the Weil-restriction of $\mathbf{G}_{m,\mathbf{C}}$ to $\mathbf{R}$ (thus $\mathbf{S}(R) = \mathbf{G}_m(R \otimes_{\mathbf{R}} \mathbf{C})$ for every $\mathbf{R}$-algebra R).

Proposition 1. *Let V be a finite dimensional $\mathbf{R}$-vector space. There is a bijection between the set of Hodge structures on V and the set $\mathrm{Hom}(\mathbf{S}, \mathrm{GL}_V)$ (morphisms of algebraic groups), which we can characterize as follows : if $h\colon \mathbf{S} \to \mathrm{GL}_V$ corresponds to a decomposition $V_{\mathbf{C}} = \bigoplus V^{p,q}$, then for all $z \in \mathbf{S}(\mathbf{R}) = \mathbf{C}^* \subset \mathbf{S}(\mathbf{C})$ and for all $v \in V^{p,q}$ we have $h(z)v = z^{-p}\bar{z}^{-q}v$.*

Proof. Notice that $\mathbf{S}$ is a torus of rank 2 over $\mathbf{R}$. A basis of $X^*(\mathbf{S})$ is formed of maps f, g such that composition with $\mathbf{S}(\mathbf{R}) \subset \mathbf{S}(\mathbf{C})$ gives the identity resp. complex conjugation. Thus mapping $f^{-p}g^{-q}$ to (p,q) defines an isomorphism $X^*(\mathbf{S}) \simeq \mathbf{Z}^2$. Since complex conjugation interchanges f and g, its action on $\mathbf{Z}^2$ is given by permutation. The proposition now follows from the general theory above. $\square$

Let us also consider what the weight corresponds to. Given $h\colon \mathbf{S} \to \mathrm{GL}_V$, denote by w_h^{-1} the restriction of h to $\mathbf{G}_{m,\mathbf{R}}$ (which sits inside $\mathbf{S}$ via $\mathbf{R}^* \subset \mathbf{C}^*$). As the notation suggests, we let w_h be the composition of w_h^{-1} with the 'inverse element' map $\mathbf{G}_m \to \mathbf{G}_m$. Again by the general theory above, since $\mathbf{G}_m$ is a split torus of dimension 1, the w_h's for different h give rise to a decomposition (over $\mathbf{R}$) of V indexed by $\mathbf{Z}$. It is called the *weight decomposition* and w_h is called the *weight homomorphism*; names which are motivated by the next proposition, whose proof is immediate from the constructions.

Proposition. *Under the correspondence of proposition 1, an element $v \in V$ (or $v \in V_{\mathbf{C}}$) has weight n if and only if the corresponding weight homomorphism w_h satisfies $w_h(x)v = x^n v$ for all $x \in \mathbf{R}^*$.* $\square$

Now let G be a real linear algebraic group. Fix a faithful representation V of G (we will always assume V of finite dimension). This representation being a morphism of algebraic groups with trivial kernel, it realizes G as a closed subgroup of GL_V. Notice that the representation of G on V determines a representation of G on $\bigoplus_i V^{m_i} \otimes (V^\vee)^{n_i}$ for every finite collection of non-negative integers m_i, n_i.

Theorem (Chevalley). *There exists a finite collection of non-negative integers m_i, n_i and a line L in $\bigoplus_i V^{m_i} \otimes (V^\vee)^{n_i}$, such that $G = \mathrm{Stab}(L)$, the stabilizer of L.*

Proof (by Bas Edixhoven [2, p. 37]). Put $F = \mathrm{GL}_V$; we consider G as a closed algebraic subgroup of F. Let I_G be its ideal in the coordinate ring $k[F]$ of F. By right translation, $k[F]$ is an F-module (this is the regular representation, which is infinite dimensional if V is of positive dimension). We observe that G is the stabilizer of I_G.

Choose generators $f_1, \ldots, f_r$ of I_G. Let W be a finite dimensional F-module containing the f_i's. Put $J = W \cap I_G$. It is an $\mathbf{R}$-subspace of W; denote its dimension by n. Now if $g \in \mathrm{Stab}(J) \subset F$, then $f_i g \in I_G$ since I_G is an ideal which contains f_i. Since the f_i's generate I_G, $g \in \mathrm{Stab}(I_G) = G$. We conclude $G = Stab(J)$, the converse inclusion being obvious. Put $L = \bigwedge^n J$, which is a line. It is contained in $\bigwedge^n W$, which is a finite dimensional representation of F since the action of F preserves the ideal generated by elements of the form $a \otimes a$ in $k[F]$. Since every finite dimensional representation of F embeds in a finite sum of copies of $k[F]$ (just look at the associated comodules), it remains to show that $k[F]$ can be decomposed into subrepresentations of the form $\bigoplus_i V^{m_i'} \otimes (V^\vee)^{n_i'}$ (for a finite collection m_i', n_i').

Put $d = \dim V$ and denote by $k[\mathrm{End}(V)]$ the coordinate ring of $\mathrm{End}(V)$; a special element of it is the determinant det. We have isomorphisms

$$k[F] \simeq k[\mathrm{End}(V)][\det^{-1}] \simeq \mathrm{Sym}_k(V^d)[\det^{-1}]$$

of F-modules. But the symmetric algebra on the right-hand side has a natural decomposition given by its grading. Now notice that

$$\mathrm{Sym}_k^i(V^d) \subset (V^d)^{\otimes i} = (V^{\otimes i})^{d^i} \quad \text{and}$$
$$k.\det^{-1} = (\bigwedge^d V)^\vee \subset (V^\vee)^{\otimes d}.$$

So $k[F]$ decomposes into tensor products of V and its dual. □

For a vector space W equipped with a Hodge structure and for a line L in W, let us call L an $\mathbf{R}$-*Hodge class* if $L_{\mathbf{C}} \subset W^{p,p}$ for some p.

Corollary. *The bijection of proposition 1 identifies* $\mathrm{Hom}(\mathbf{S}, G)$ *with the set of Hodge structures on V such that L is an* $\mathbf{R}$*-Hodge class.*

Proof. A Hodge class on V corresponds to $h\colon \mathbf{S} \to \mathrm{GL}_V$. Now L is a $\mathbf{R}$-Hodge class if and only if $h(z)l = |z|^{2p}l$ for all $z \in \mathbf{C}^*$, $l \in L$. Hence $h(z)$ stabilizes L for every z. The converse is also true, since the norm generates $\mathrm{Hom}(\mathbf{S}, \mathbf{G}_m)$. But then all this is equivalent with the condition that h factors over G. □

3 HS from algebraic geometry

Let G be a linear algebraic group defined over $\mathbf{R}$. Fix a $G(\mathbf{R})$-conjugacy class X' in $\mathrm{Hom}(\mathbf{S}, G)$. We give $\mathrm{Hom}(\mathbf{S}, G)$ (and X') the subspace topology via the injection

$$\mathrm{Hom}(\mathbf{S}, G) \hookrightarrow \mathrm{Hom}(\mathrm{Lie}(\mathbf{S}), \mathrm{Lie}(G)),$$

where the right-hand side has a natural topology as subset of some real vector space of finite dimension.

Let X be a connected component of X'. Clearly, $G(\mathbf{R})^\circ$, the identity component of the real Lie group $G(\mathbf{R})$, acts on X via conjugation. By using rigidity of the torus $\mathbf{S}$, we see that any two element of X' are conjugated. Hence X is a $G(\mathbf{R})^\circ$-conjugacy class. (Note that Deligne's notation for $G(\mathbf{R})^\circ$ is $G(\mathbf{R})^+$.)

We will now introduce three properties (or axioms) which we will demand of G and X. The motivation for these come from algebraic geometry, as is explained in the following basic example.

Main example. Let S be a smooth complex variety. Consider a family $f\colon Y \to S$ of smooth, projective complex varieties. Since each fibre is Kähler, its cohomology has a Hodge structure — see for example the book by Griffiths and Harris [3]. To be precise: for every point s of S, look at the n-th cohomology $H^n(Y_s, \mathbf{R})$ (either singular, De Rham, or sheaf cohomology; they are all the same here). Then the finite dimensional vector space $H^n(Y_s, \mathbf{R})$ carries a Hodge structure of weight n. Furthermore, for different points s, the cohomology spaces are all isomorphic (using that S is connected) : the collection $H^n(Y_s, \mathbf{R})$ parametrised by S forms a *local system*, i.e., a locally trivial sheaf of $\mathbf{R}$-vector spaces on S. (In fact, this local system is $R^n f_* \mathbf{R}_Y$.)

Now let H_s be a sum of such $H^n(Y_s, \mathbf{R})$ for a collection of n's. Varying s, we get a collection of Hodge structures parametrised by S. This is our guiding example from which we derive properties a collection of Hodge structures has to satisfy. The first one (α) is immediate from the above : it tells us that the weight filtration does not depend on $s \in S$.

(α) *For all* h, $h' \in X$ *we have* $w_h = w_{h'}$.

Typically, Deligne converts a statement like (α) to a statement about Lie algebras. We will also do this now. Denote by $\mathfrak{g}$ the Lie algebra of G. For $g \in G$, denote by $\mathrm{Ad}(g)\colon \mathfrak{g} \to \mathfrak{g}$ the derivative at the unit element of $h \mapsto ghg^{-1}$; we obtain the *adjoint representation* Ad of G in $\mathfrak{g}$. Now the nice thing is that a Hodge structure on G induces one on $\mathfrak{g}$, by composing $\mathbf{S} \to G$ with the adjoint representation.

Proposition 2. *Condition* (α) *is equivalent to the condition that for a* $h \in X$ *the induced Hodge structure on* $\mathfrak{g}$ *is of weight* 0.

PROOF. Pick $h \in X$. Since X is a conjugacy class, $w_{h'}$ is independent of $h' \in X$ if and only if $h(\mathbf{R}^*)$ is in the center of $G(\mathbf{R})^\circ$. Equivalent to this are: conjugation by $G(\mathbf{R})^\circ$ in $h(\mathbf{R}^*)$ is trivial; the image of $h(\mathbf{R}^*)$ in $\mathrm{GL}(\mathrm{Lie}(G(\mathbf{R})^\circ)) = \mathrm{GL}(\mathfrak{g})$ is trivial; $\mathfrak{g}$ has weight 0 in the Hodge structure induced by h. $\square$

For the rest of this section, we fix a faithful representation $\rho\colon G \to \mathrm{GL}_V$ *(as always, of finite dimension) and furthermore assume that* (α) *holds.* Note that since the weight decomposition is defined over $\mathbf{R}$, we can decompose V into subrepresentations V_n which are pure of weight n.

We will now discuss the second axiom (β). In order to do this, we need to recall two things.

Firstly, let S be a complex manifold. Remember that giving a local system H on S is equivalent to giving a locally free sheaf of $\mathcal{O}_S$-modules $\mathcal{H}$ together with an integrable connection ∇ on $\mathcal{H}$. The correspondence is $\mathcal{H} = H \otimes_{\mathbf{R}} \mathcal{O}_S$ with the horizontal sections of ∇ being exactly the sections of H.

Next recall that a Hodge decomposition on a vector space V can also be described by giving the weight decomposition and the *Hodge filtration* $F^i V_{\mathbf{C}} = \bigoplus_{p \geq i} V^{p,q}$. Conversely we get $V^{p,q}$ as the weight $p+q$ part of $F^p \cap \overline{F^q}$. Note that the Hodge filtration is a decreasing filtration.

(main example continued.) Let $\mathcal{H}$ correspond to the local system obtained from H_s parametrised by $s \in S$ (so by the projection formula it is a sum of $R^n f_* \mathcal{O}_Y$ for different n). Then $\mathcal{H}$ carries a Hodge filtration. One can ask how this filtration varies up to first order movements in the base space. The answer is given by the work of Griffiths (see for example Deligne's Bourbaki talk [4]), who showed that the above gives a so called variation of Hodge structure, which we will define below.

Recall that the *flag variety* $\mathcal{F}$ of $V_{\mathbf{C}}$ of *size* $(f_i)_i$ as a set is the collection of decreasing filtrations $F^\bullet$ of $V_{\mathbf{C}}$ such that $F^i V_{\mathbf{C}}$ is a subspace of dimension f_i. It naturally forms a closed subset inside a product of Grassmannians; in this way $\mathcal{F}$ obtains the structure of a complex projective variety. Given a pure Hodge structure V, it naturally determines a point in the flag variety with

$f_i := \dim F^i V_{\mathbf{C}}$. An arbitrary Hodge structure can be decomposed by its weight, and in this way we obtain a point in a product of flag varieties.

Definition. An *almost variation of Hodge structure* on a complex manifold S consists of a local system H, a decomposition $H = \bigoplus_n H_n$ of local systems, and a filtration $F^\bullet$ on $H \otimes_{\mathbf{R}} \mathcal{O}_S$ such that:

i) fibrewise the filtration is a Hodge filtration (of weight n on H_n);

ii) the filtration is holomorphic in the following sense: for every $s \in S$, consider an open neighbourhood U over which H is trivialized by some vector space V; then the Hodge filtration gives a map from U to a product of flag varieties of $V_{\mathbf{C}}$ as explained above; we demand that this map is a morphism of complex manifolds.

It is a *variation of Hodge structure* if moreover *Griffiths transversality* is satisfied:

iii) $\nabla F^i \mathcal{H} \subset \Omega^1_{S/\mathbf{C}} \otimes_{\mathcal{O}_S} F^{i-1}\mathcal{H}$.

Proposition 3. *(Still assuming (α).) There is a unique complex structure on X such that it has an almost variation of Hodge structure with Hodge filtration in the fibre of $h \in X$ being the one given by h on V (as described in proposition 1). Furthermore, Griffiths transversality is satisfied if and only if the natural Hodge structure of the adjoint representation is of type $\{(-1,1),(0,0),(1,-1)\}$.*

PROOF. Treating each weight seperately, we may assume that the Hodge structure on V is pure of weight n (this condition is independent of the choice of $h \in X$ by (α)). Given $h \in X$, it determines a filtration $F_h^\bullet$ of $V_{\mathbf{C}}$. Since X is connected, $f_i := \dim F_h^i V_{\mathbf{C}}$ is independent of h. Now we have a map

$$\phi\colon X \to \mathcal{F}$$

sending h to its associated Hodge filtration. As V is a faithful representation, ϕ is injective. We will show below that ϕ is in fact an immersion onto a complex submanifold. This induces a complex structure on X such that the filtration varies holomorphically. This structure is clearly unique. This gives the first part of the proposition.

First we will describe the tangent space $T_X(h)$ at $h \in X$. Since $G(\mathbf{R})^\circ$ acts differentiably and transitively on X, we obtain a diffeomorphism

$$X \simeq G(\mathbf{R})^\circ / \operatorname{Stab}(h).$$

Denoting the Lie algebra of $G(\mathbf{R})$ by $\mathfrak{g}$ this gives

$$T_h(X) \simeq \mathfrak{g} / \operatorname{Lie}(\operatorname{Stab}(h)).$$

In terms of the Hodge structure induced by h on $\mathfrak{g}$, we can describe $\operatorname{Lie}(\operatorname{Stab}(h))$ as $\mathfrak{g}_{\mathbf{R}}^{0,0}$, meaning the points of $\mathfrak{g}_{\mathbf{C}}$ of type $(0,0)$ defined over $\mathbf{R}$. To convince us of this fact, it is maybe better to make an intuitive picture for ourselves; but anyway, the next paragraph contains a rigorous proof, which is just doing a messy calculations.

For $z \in \mathbf{S}(\mathbf{R})$, define $A_z, B_z: G(\mathbf{R}) \to G(\mathbf{R})$ by $A_z(g) = gh(z)g^{-1}$ (not a homomorphism) and $B_z(g) = h(z)gh(z)^{-1}$. Pick $\gamma \in \operatorname{Lie}(\operatorname{Stab}(h))$. This means that $T_{B_z}(1)\gamma = 0$ for all $z \in \mathbf{S}(\mathbf{R})$; so (using e.g. the embedding of $G(\mathbf{R})$ in $\operatorname{GL}(V)$) this means

$$\begin{aligned} B_z(1+\epsilon\gamma) - B_z(1) &= (1+\epsilon\gamma)h(z)(1-\epsilon\gamma) - h(z) + \mathcal{O}(\epsilon^2) \\ &= \epsilon(\gamma h(z) - h(z)\gamma) + \mathcal{O}(\epsilon^2) \qquad (\epsilon \to 0) \end{aligned}$$

has to vanish up to first order, i.e., γ commutes with $h(z)$ for all z. On the other hand, $\gamma \in \mathfrak{g}_{\mathbf{R}}^{(0,0)}$ means $T_{A_z}(e)\gamma = \gamma$ for all $z \in \mathbf{S}(\mathbf{R})$; that is

$$A_z(1+\epsilon\gamma) - A_z(1) = \epsilon h(z)\gamma h(z)^{-1} = \epsilon\gamma,$$

which indeed again means that γ and $h(z)$ commute for every h.

Next we deal with the tangent space of $\mathcal{F}$ at $\phi(h)$. For this we notice that $\operatorname{GL}(V_{\mathbf{C}})$ acts naturally on $\mathcal{F}$, by acting on the left of the parts of the filtration. Therefore, the tangent space is

$$\operatorname{End}(V_{\mathbf{C}}) / \operatorname{Lie}(\operatorname{Stab}(\phi(h))),$$

where we view $\operatorname{End}(V_{\mathbf{C}})$ as the Lie algebra of $\operatorname{GL}(V_{\mathbf{C}})$. The composition

$$\mathbf{S} \xrightarrow{h} G \xrightarrow{\rho} \operatorname{GL}_V \xrightarrow{\operatorname{Ad}} \operatorname{GL}_{\operatorname{Lie}(\operatorname{GL}_V)}$$

gives $\operatorname{End}(V_{\mathbf{C}})$ a Hodge decomposition and similarly to what we have seen before, $\operatorname{Lie}(\operatorname{Stab}(\phi(h)))$ is $F^0 \operatorname{End}(V_{\mathbf{C}})$. For this fact, the next paragraph contains a proof.

Let $\gamma \in \operatorname{Lie}(\operatorname{Stab}(\phi(h)))$. Then for ϵ sufficiently small, $(1+\epsilon\gamma)F_h^\bullet V_{\mathbf{C}} = F_h^\bullet V_{\mathbf{C}}$. By linearity, this means that if $v \in V^{p,n-p}$, then $\gamma v \in \bigoplus_{r \geq p} V^{r,n-r}$. Fix an $r \geq p$. If γ^r is γ composed with the projection to $V^{r,n-r}$ and if γ^r is in $F^0 \operatorname{End}(V_{\mathbf{C}})$ for all $r \geq p$, then γ sits in $F^0 \operatorname{End}(V_{\mathbf{C}})$. Hence, we assume $\gamma v \in V^{r,n-r}$ without loss of generality. For $z \in \mathbf{C}^*$,

$$\begin{aligned} (\operatorname{Ad}(h(z))\gamma)\,(h(z)v) &= h(z)\gamma v \\ &= z^{-r}\overline{z}^{r-n}\gamma v \\ &= \left(z^{p-r}\overline{z}^{r-p}\gamma\right)(h(z)v). \end{aligned}$$

Since $p - r \leq 0$ we conclude, if we vary p and v, that $\gamma \in \bigoplus_{p \geq 0} (\operatorname{End}(V_{\mathbf{C}}))^{t,-t} = F^0 \operatorname{End} V_{\mathbf{C}}$. So

$$\operatorname{Lie}(\operatorname{Stab}(\phi(h))) \subset F^0 \operatorname{End}(V_{\mathbf{C}}).$$

The converse inclusion is the above proof read backwards, except that in advance we choose $\gamma \in \operatorname{End}(V_C)^{t,-t}$ for a fixed t and then we end up with γv not necessarily being purely of type $(r, n-r)$ — we could have given the proof purely as a chain of equivalences, but the restriction to one summand makes the formulas shorter and easier to comprehend.

Applying the funtor Lie to ρ, we get a homomorphism of $G(\mathbf{R})$-modules $p: \mathfrak{g} \to \operatorname{End}(V)$. Since ρ is faithful, p is injective. Furthermore, since p commutes with the action of $G(\mathbf{R})$, it is a homorphism of Hodge structures defined by h (in the obvious way). Therefore p induces an injection

$$p: \mathfrak{g}/\mathfrak{g}_{\mathbf{R}}^{0,0} \to \operatorname{End}(V)/\operatorname{End}(V)_{\mathbf{R}}^{0,0}.$$

We have canonical isomorphisms

$$\begin{aligned}\psi_{\mathfrak{g}} &: \mathfrak{g}/\mathfrak{g}_{\mathbf{R}}^{0,0} \xrightarrow{\sim} \mathfrak{g}_{\mathbf{C}}/F^0\mathfrak{g}_{\mathbf{C}} \qquad \text{and} \\ \psi_{\mathrm{End}(V)} &: \mathrm{End}(V)/\mathrm{End}(V)_{\mathbf{R}}^{0,0} \xrightarrow{\sim} \mathrm{End}(V_{\mathbf{C}})/F^0\,\mathrm{End}(V_{\mathbf{C}}).\end{aligned}$$

This is an application of the following general reasoning : Let W be an $\mathbf{R}$-vector space with a Hodge structure and let $W \to W_{\mathbf{C}}$ be the canonical inclusion. Then $F^0W_{\mathbf{C}} \cap W = W_{\mathbf{R}}^{0,0}$, since W must be invariant under complex conjugation. Hence, the map $\psi_W\colon W/W_{\mathbf{R}}^{0,0} \to W_{\mathbf{C}}/F^0W_{\mathbf{C}}$ is well-defined and injective. To see that it is surjective, take $w \in W_{\mathbf{C}}$. Then either w or $\overline{w}$ lives in $F^0W_{\mathbf{C}}$; but $w + \overline{w} \in W$.

The tangent map $T_\phi(h)$ is $\psi_{\mathrm{End}(V)} \circ p = p \circ \psi_{\mathfrak{g}}$, which is injective (for p is) and $T_X(h)$ is identified with the complex subspace $\mathfrak{g}_{\mathbf{C}}/F^0\mathfrak{g}_{\mathbf{C}}$ of $T_{\mathcal{F}}(\phi(h))$.

Under the extra assumption of Griffiths transversality, the above shows the image to be in $F^{-1}\,\mathrm{End}\,V_{\mathbf{C}}/F^0\,\mathrm{End}\,V_{\mathbf{C}}$ and vice versa. □

With this proposition and the above example in mind, we can state our next axiom :

(β) *The complex structure on X given by proposition 3 (and implicitly by V) satisfies Griffiths transversality, i.e., X and V give a variation of Hodge structure.*

Proposition 3 gives a translation of the above statement in terms of Lie algebras: the adjoint representation is of type $\{(-1,1),(0,0),(1,-1)\}$. Notice that assuming this statement, (α) is automatically satisfied. Also, (β) is now clearly independent of the choice of V.

Definition. Let V be a vector space equipped with a Hodge structure $h\colon \mathbf{S} \to \mathrm{GL}_V$. A *polarization* on V is a bilinear form

$$\psi\colon V \otimes_{\mathbf{R}} V \to \mathbf{R}$$

such that:

i) it is invariant under $h(U^1)$, with U^1 the unit circle in $\mathbf{C}^*$;

ii) the map $(x,y) \mapsto \psi(x, h(i)y)$ is symmetric and positive definite;

iii) if $n \neq m$, then $\psi(V_n, V_m) = 0$.

(main example continued.) On $H^n(X_s,\mathbf{R})$ there is a canonical polarization, coming from the Kähler metric (see [4] again). This polarization is independent of s.

(γ) *There is a map $V \otimes_{\mathbf{R}} V \to \mathbf{R}$ which gives a polarization for every $h \in X$.*

In fact, it is equivalent to demand the existence of a locally constant bilinear form on the local system $(V_n)_{\mathbf{R}}$ which defines a polarization fibrewise. This is the usual definition of a *polarization on a variation of Hodge structure*.

Again, we will translate (γ) in a statement which does not involve the representation V. Therefore, suppose that (γ) holds, and denote by ψ a polarization of V. Let G_2 be the smallest algebraic subgroup of G through which $h|_{U^1}$ factorizes for all $h \in X$. Note that ψ is invariant under G_2, invariance being a closed condition. Also note that G_2 is connected, as every morphism $\mathbf{S} \to G_2$ must map the connected group $\mathbf{C}^*$ into $G_2(\mathbf{R})^\circ \subset G_2^\circ(\mathbf{R})$, so by minimality $G_2 = G_2^\circ$.

Extend ψ to a map $\psi_{\mathbf{C}}\colon V_{\mathbf{C}} \times V_{\mathbf{C}} \to \mathbf{C}$ which is linear in the first and anti-linear in the second variable. Then $\psi(\,—\,, h(i)\,—\,)$ is a positive definite Hermitian form on $V_{\mathbf{C}}$, which satisfies

$$\psi_{\mathbf{C}}\bigl(gx, h(i)\,(h(i))^{-1}\,\overline{g}\,h(i)y\bigr) = \psi_{\mathbf{C}}(x, h(i)y),$$

for all $x, y \in V$ and $g \in G_2(\mathbf{R})$. This last equality can be described by saying that $\psi_{\mathbf{C}}(\,—\,, h(i)\,—\,)$ is invariant under the group

$$G_2^{(\operatorname{ad} h(i))}(\mathbf{R}) := \bigl\{g \in G_2(\mathbf{C}) \bigm| g = (h(i))^{-1}\,\overline{g}\,h(i)\bigr\}.$$

We will see in the next proposition that this group is compact; and vice versa: if this group is compact, then there is a polarization. But first we will abstractly define this property:

Definition/Construction. Let G be a connected algebraic group over $\mathbf{R}$ and let ι be an involution of G (i.e., $\iota^2 = 1$). Then ι decomposes the Lie algebra $\mathfrak{g}$ of G into the $+1$ and -1 eigenspace $\mathfrak{g}^+$ resp. $\mathfrak{g}^-$. Put $\mathfrak{g}^{(\iota)} = \mathfrak{g}^+ \oplus i\mathfrak{g}^- \subset \mathfrak{g}_{\mathbf{C}}$; it is the Lie algebra of a real algebraic group $G^{(\iota)}$ which satisfies $G^{(\iota)}_{\mathbf{C}} = G_{\mathbf{C}}$. In terms of group functors: $G^{(\iota)}(\mathbf{C}) = G(\mathbf{C})$; complex conjugation is mapped to $G^{(\iota)}(z \mapsto \overline{z})$, which sends $g \in G(\mathbf{C})$ to $\iota(\overline{g})$; and the elements of $G^{(\iota)}(\mathbf{R})$ are the elements invariant under the image of complex conjugation, i.e.,

$$G^{(\iota)}(\mathbf{R}) = \bigl\{g \in G(\mathbf{C}) \bigm| g = \iota(\overline{g})\bigr\}.$$

Now by definition, ι is a *Cartan involution* of G if $G^{(\iota)}(\mathbf{R})$ is compact.

Define $G^{\mathrm{ad}} = G/\mathcal{C}$, with $\mathcal{C}$ the center of G. This is called the *adjoint group* of G. Note that for G connected, the adjoint representation injects G^{ad} into $\mathrm{GL}(\mathfrak{g})$ (this is only true in characteristic 0).

Proposition 4. *Let G_1 be the smallest algebraic subgroup of G through which all $h \in X$ factor. A necessary and sufficient condition for (γ) to hold is: G_1 is reductive and $\operatorname{ad} h(i)$ is a Cartan involution of G_1^{ad}.*

Proof. Assuming (γ), we already constructed a positive definite Hermitian form on $V_{\mathbf{C}}$ which is invariant under the action of $G_2^{(\operatorname{ad} h(i))}$. This means that $G_2^{(\operatorname{ad} h(i))}$ is an algebraic subgroup of the group of unitary tranformations of $V_{\mathbf{C}}$. But this last group is compact, hence so is $G_2^{(\operatorname{ad} h(i))}(\mathbf{R})$. Thus $\operatorname{ad} h(i)$ is a Cartan involution of G_2.

Conversely, if $G_2^{(\operatorname{ad} h(i))}(\mathbf{R})$ is compact, it has a Haar measure. Via integration we can then define a positive definite Hermitian form on $V_{\mathbf{C}}$. Some easy linear algebra (the same as we did above, but in the opposite direction) makes this into a polarization.

It remains to show that $\operatorname{ad} h(i)$ is a Cartan involution of G_2 if and only if it is one on G_1^{ad} and G^1 is reductive.

There is a canonical surjection $U^1 \times \mathbf{G}_m \to \mathbf{S}$ (with finite kernel; in fact, generated by $(-1,-1)$). Pushing out we realize G_1 as a quotient of $G_2 \times \mathbf{G}_m$. We can conclude two things. Firstly, since $1 \times \mathbf{G}_m$ is in the center of $G_2 \times \mathbf{G}_m$, we get $G_1^{\mathrm{ad}} = G_2^{\mathrm{ad}}$. Secondly, if G_2 is reductive, so is G_1, being the quotient of a reductive group; the converse statement is obvious.

A group which has a Cartan involution is necessarily reductive, for the form defined in the proces of the above proof can also be used to fully decompose a representation of the group. Hence if G_2 has a Cartan involution, it is reductive.

Suppose G_2 is reductive. Then its connected center C a torus. As we are working over $\mathbf{R}$, this means that C is the product of $\mathbf{G}_m$'s and U^1's. Now any morphism from U^1 to $\mathbf{G}_m$ is trivial, so by minimality, there appear no factors $\mathbf{G}_m$ in C. So $C(\mathbf{R})$ is compact, and hence the (non-connected) center of G_2 is compact. As $\operatorname{ad} h(i)$ acts trivially on this center, we conclude that $\operatorname{ad} h(i)$ is a Cartan involution on G_2 iff it is one on G_2^{ad}. One can check that we now have all the implications we need. $\square$

We still consider X as above with the structure of a complex manifold from proposition 3. Note that $G(\mathbf{R})^\circ$ acts on X. For the next theorem, we suppose $G = H_{\mathbf{R}}$ for some algebraic group H defined over $\mathbf{Q}$. We say that Γ is a *congruence subgroup* of H if there exists a compact open subgroup K in $H(\mathbf{A}_f)$ such that $\Gamma = K \cap H(\mathbf{Q})$. If we furthermore suppose Γ to be torsion free, then it is easily seen that Γ acts freely and discontinuously on X, so that $\Gamma\backslash X$ is also a complex manifold. We can say much more :

Theorem 5. *Let Γ be a torsion free congruence subgroup with $\Gamma(\mathbf{R}) \subset G(\mathbf{R})^\circ$. Then $\Gamma\backslash X$ has a unique structure of a quasi-projective algebraic variety over the complex numbers. This is functorial in the following sense: if $\Gamma' \subset \Gamma$, then the canonical map $\Gamma'\backslash X \to \Gamma\backslash X$ is a morphism of algebraic varieties.*

Proof. This is a well-known and difficult result of Baily and Borel [5, 6]. However, they state their result not in terms of 'conjugacy classes in $\mathrm{Hom}(\mathbf{S}, G)$' but in terms of 'Hermitian symmetric domains'. But Deligne [1, 1.1.17] proves that the conjugacy classes of this section satisfying the axioms (α)–(γ) above, are in bijective correspondence to the Hermitian symmetric domains (and G is then the automorphism group). The proof of this is not difficult, but one needs some results from the theory of these Hermitian symmetric domains. We will not do this here, and content ourselves with the references sited above. Please note that functoriality can be much more generalized, but we do not need that here. $\square$

4 Shimura varieties

Let us look at theorem 5 again. We may ask, following Shimura, whether the variety $\Gamma\backslash X$ can be defined over a number field. Deligne observed that the natural way to approach this problem is via Shimura varieties, which will be defined in this section. *Caution: the symbols G and X will have a slightly different meaning in this section.*

Definition. A *Shimura datum* (G, X) consists of a reductive group G defined over $\mathbf{Q}$ and a $G(\mathbf{R})$-conjugacy class X in $\mathrm{Hom}(\mathbf{S}, G_{\mathbf{R}})$ satisfying:

i) for all $h \in X$, the induced Hodge structure on the adjoint representation $\mathrm{Lie}(G_{\mathbf{R}})$ is of type $\{(-1,1),(0,0),(1,-1)\}$;

ii) $\mathrm{ad}\, h(i)$ is a Cartan involution of $G^{\mathrm{ad}}_{\mathbf{R}}$;

iii) G^{ad} has no factor G' defined over $\mathbf{Q}$ such that, for all $h \in X$, the composition of h with the projection to $G'_{\mathbf{R}}$ is trivial.

A few remarks are in order. First note that, unlike in the previous paragraph, G has to be defined over $\mathbf{Q}$ and X is not necessarily connected. The previous paragraph motivates the demand that G must be reductive (in general, reductiveness is defined by looking at geometric fibres, hence G is reductive iff $G_{\mathbf{R}}$ is) and the axioms (i) and (ii). We cannot yet explain the reasons behind axiom (iii).

Let (G, X) be a Shimura datum. If K is a compact and open subgroup of $G(\mathbf{A}_f)$, define the topological space

$$\mathrm{Sh}_K(G,X)(\mathbf{C}) = G(\mathbf{Q}) \setminus X \times \big(G(\mathbf{A}_f)/K\big).$$

Denote by $G(\mathbf{R})_+$ the inverse image of $G^{\mathrm{ad}}(\mathbf{R})^\circ$ under the canonical map $G(\mathbf{R}) \to G^{\mathrm{ad}}(\mathbf{R})$.

Proposition 6. *Fix a connected component X^+ of X. For $g \in G(\mathbf{A}_f)$, define $K_g = G(\mathbf{Q})_+ \cap gKg^{-1}$. Let $I \subset G(\mathbf{A}_f)$ be a complete set of coset representatives of $G(\mathbf{Q})_+\backslash G(\mathbf{A}_f)/K$. Then I is a finite set and $x \mapsto (x, g)$ canonically induces a homeomorphism*

$$\coprod_{g \in I} K_g\backslash X^+ \simeq \mathrm{Sh}_K(G,X)(\mathbf{C}).$$

Proof. The set $\pi_0(X)$ of connected components of X has a natural action of $G(\mathbf{R})$. Again by using the fact that the connected components of X are Hermitian symmetric domains (cf. the proof of theorem 5), one shows that this action makes $\pi_0(X)$ a principal homogeneous space under $G(\mathbf{R})/G(\mathbf{R})_+$. We will not discuss this; see [1, §1.2].

The real approximation theorem (difficult to prove) tells us that the set $G(\mathbf{Q})$ is dense in the topological space $G(\mathbf{R})$; this is in fact true for all connected linear algebraic groups—see [7, 3.5(iii)]. We have a monomorphism $G(\mathbf{Q})/G(\mathbf{Q})_+ \to G(\mathbf{R})/G(\mathbf{R})_+$ which is in fact an isomorphism : Indeed, let $x \in G(\mathbf{R})$. As $G(\mathbf{R})_+$ is open in $G(\mathbf{R})$, so is $xG(\mathbf{R})_+$. The fact that $G(\mathbf{Q})$ is dense implies the existence of an element $y \in xG(\mathbf{R})_+ \cap G(\mathbf{Q})$, from which surjectivity of the above map follows. It readily follows that

$$\mathrm{Sh}_K(G,X)(\mathbf{C}) \simeq G(\mathbf{Q})_+ \setminus X^+ \times \big(G(\mathbf{A}_f)/K\big).$$

Abbreviate the right-hand side to Sh^+.

Now for $g \in G(\mathbf{A}_f)$, a straightforward check show that $x \mapsto (x,g)$ induces an open and closed immersion $\phi_g\colon K_g\backslash X^+ \hookrightarrow \mathrm{Sh}^+$. Thus, we get a family of open and closed immersions parametrized by $g \in G(\mathbf{A}_f)$. It follows from the characterization of $\pi_0(X)$ in the previous paragraph that $\phi_g = \phi_{g'}$ if and only if g and g' give the same class in $G(\mathbf{Q})_+\backslash G(\mathbf{A}_f)/K$ and if this is not the case then ϕ_g and ϕ_g' have disjoint image.

Finally, the proof that $G(\mathbf{Q})_+\backslash G(\mathbf{A}_f)/K$ is finite uses some very untrivial facts about certain algebraic groups over $\mathbf{Q}$ (Deligne formulates them in (0.1)–(0.3) in [8]. We don't discuss the proof, but refer to [ibed., §2]. $\square$

Assume K is torsion free. Then K_g is a torsion free arithmetic subgroup. By theorem 5 and the above, there is a unique scheme $\mathrm{Sh}_K(G,X)_{\mathbf{C}}$ defined over $\mathbf{C}$, being a finite disjoint union of non-singular quasi-projective varieties, whose associated analytic space is $\mathrm{Sh}_K(G,X)(\mathbf{C})$; this is the *Shimura variety* associated to (G,X) and K. (Note: Deligne also uses this term for $\mathrm{Sh}_K(G,X)(\mathbf{C})$, even when K has torsion; in this case there is still an algebraic structure, but it is not necessarily unique.)

An inclusion $L \subset K$ of torsion-free compact open subgroups of $G(\mathbf{A}_f)$ induces a finite morphism $\mathrm{Sh}_K(G,X)_{\mathbf{C}} \to \mathrm{Sh}_L(G,X)_{\mathbf{C}}$. We get a projective system, whose limit exists (as a scheme) since all the maps are finite; we denote it by $\mathrm{Sh}(G,X)_{\mathbf{C}}$. The advantage of working with this projective limit is that it has a natural action of $G(\mathbf{A}_f)$; indeed, a straightforward check shows that an element g induces an isomorphism $\mathrm{Sh}_K(G,X)_{\mathbf{C}} \to \mathrm{Sh}_{g^{-1}Kg}(G,X)_{\mathbf{C}}$.

References

[1] Pierre Deligne: Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques; *Proceedings of Symposia in Pure Mathematics* Vol. **33**, part 2 (1979), pp. 247–290.

[2] Bas Edixhoven: 'Scratch notes' accompanying On the André-Oort conjecture for Hilbert modular surfaces; http://www.maths.univ-rennes1.fr/∼edix/publications/texel2.html.

[3] Phillip Griffiths and Joseph Harris: *Principles of algebraic geometry*; Wiley-Interscience (1978).

[4] Pierre Deligne: Travaux de Griffiths; *Séminaire Bourbaki vol. 1969/70*, LNM **180** (1971), pp. 213–237.

[5] W.L. Baily, Jr. and A. Borel: *Compactification of arithmetic quotients of bounded symmetric domains*; Ann. of Math. (2) **84** (1966), pp. 442–528.

[6] Armand Borel: *Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem*; J. Differential Geometry **6** (1972), pp. 543–560.

[7] J.-J. Sansuc: *Groupe de Brauer et arithmétique des groupes algébriques linéaires sur un corps de nombres*; J. Reine Angew. Math. 327 (1981), pp. 12–80.

[8] Pierre Deligne: *Travaux de Shimura*; Séminaire Bourbaki 1970/1971, exposé 389; LNM 244 (1971), pp. 123–165.