The formalism of Shimura varieties

November 25, 2013

1 Shimura varieties: general definition

Definition 1.0.1. A *Shimura datum* is a pair (G, X), where $G/\mathbb{Q}$ is a reductive group, and X is a $G(\mathbb{R})$-conjugacy class of homomorphisms $h\colon S \to G_{\mathbb{R}}$ satisfying the conditions:

1. (SV1): For $h \in X$ the Hodge structure on $\operatorname{Lie} G_{\mathbb{R}}$ induced by the adjoint representation is of type $\{(1,-1),(0,0),(-1,1)\}$.

2. (SV2): The involution $\operatorname{ad} h(i)$ is Cartan on G^{ad}.

3. (SV3): G has no $\mathbb{Q}$-factor on which the projection of h is trivial.

(If the conditions above pertain to one $h \in X$, they pertain to all of them.)

For a compact open subgroup $K \subset G(\mathbb{A}_f)$, define the *Shimura variety* $\mathrm{Sh}_K(G, X)$ by

$$\mathrm{Sh}_K(G,X) = G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K).$$

Generally this will be a disjoint union of locally symmetric spaces $\Gamma\backslash X^+$, where X^+ is a connected component of X and $\Gamma \subset G(\mathbb{Q})^+$ is an arithmetic subgroup. By the Baily-Borel theorem, $\mathrm{Sh}_K(G,X)$ is a quasi-projective variety.

Let's also define the Shimura variety at infinite level:

$$\mathrm{Sh}(G,X) = G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)) = \varprojlim_K \mathrm{Sh}_K(G,X).$$

This is also a scheme, but not generally of finite type. Note that $\mathrm{Sh}(G,X)$ has a right action of $G(\mathbb{A}_f)$.

1.1 An orienting example: the tower of modular curves

It's good to keep in mind the case of modular curves. Let $G = \mathrm{GL}_2/\mathbb{Q}$, and let X be the conjugacy class of $h\colon a+bi \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. Then we have a $G(\mathbb{R})$-equivariant bijection $X \cong \mathbb{C}\backslash\mathbb{R}$ sending h to i. We think of X as the set of complex structures on the rational vector space $V = \mathbb{Q}^2$. It so happens that these complex structures are all polarized (up to sign) by a common alternating form ψ on V, but since such a form is unique up to scaling, we can ignore the polarizations in this discussion.

The right way to think about $\mathrm{Sh}(G,X)$ is in terms of *isogeny classes* of elliptic curves eqipped with full (adelic) level structure. For an elliptic curve $E/\mathbb{C}$, the (adelic) Tate module is $TE = \varprojlim_N E[N] \approx \hat{Z} \times \hat{\mathbb{Z}}$, and the rational (adelic) Tate module is

$$V_f(E) = \varprojlim_N E[N] \otimes \mathbb{A}_f = H_1(E,\mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{A}_f$$

Proposition 1.1.1. $\mathrm{Sh}(G,X)$ *classifies isogeny classes of pairs* (E,η)*, where* $E/\mathbb{C}$ *is an elliptic curve and*

$$\eta\colon \mathbb{A}_f \times \mathbb{A}_f \to V_f(E)$$

is an $\mathbb{A}_f$*-linear isomorphism. (Two pairs* (E,η) *and* (E',η') *are isogenous if there is an* $f \in \mathrm{Hom}(E,E') \otimes \mathbb{Q}$ *carrying* η *onto* η'*.)*

Proof. Let (E,η) be given. Let $W = H_1(E,\mathbb{Q})$. Choose any isomorphism $afromV \to W$. E gives W a complex structure, which gets transferred via a to a complex structure on V, and thus we a point $h \in X$.

The rational Tate module V_fE is canonically isomorphic to $W \otimes_{\mathbb{Q}} \mathbb{A}_f$, and the level structure η becomes an isomorphism

$$\eta\colon \mathbb{A}_f \times \mathbb{A}_f \to L_{\mathbb{Q}} \otimes_Q \mathbb{A}_f.$$

On the other hand we have the isomorphism

$$a \otimes 1\colon \mathbb{A}_f \times \mathbb{A}_f \to L_{\mathbb{Q}} \otimes_Q \mathbb{A}_f.$$

Let $g = (a \otimes 1)^{-1}\eta$.

To (E,η) we assign the class of (h,g) in $\mathrm{Sh}(G,X)$. Let's see that this is well-defined. If a different isomorphism $a'\colon V \to W$ was chosen, then

$a' = \gamma a$ for some $\gamma \in G(\mathbb{Q})$. The complex structure h and the element g would both get translated by the same γ, so that the resulting class in $\mathrm{Sh}(G,X) = G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)$ would not change. Further, if we have two pairs (E,η) and (E',η') and an isogeny $f \in \mathrm{Hom}(E,E') \otimes \mathbb{Q}$ between them, then f induces isomorphisms $H_1(E,\mathbb{Q}) \to H_1(E',\mathbb{Q})$ and $V_f E \to V_f E'$. If we choose $a' \colon V \to H_1(E',\mathbb{Q})$ to be the composite of a with $H_1(E,\mathbb{Q}) \to H_1(E',\mathbb{Q})$, then the resulting pair (h,g) does not change at all.

The opposite direction is an exercise (or see below where it is written up in the context of Siegel modular varieties). □

Note that $\mathrm{Sh}_K(G,X)$ is disconnected. One way to see this is that we have a projection

$$\begin{array}{ccc} \mathrm{Sh}_K(G,X) = G(\mathbb{Q})^+\backslash(X^+ \times G(\mathbb{A}_f))/K & \longrightarrow & G(\mathbb{Q})^+\backslash G(\mathbb{A}_f)/K \\ & & \Big\downarrow {\scriptstyle \sim}\ {\scriptstyle \det} \\ & & \mathbb{Q}^\times_{>0}\backslash\mathbb{A}_f^\times/\det K \end{array}$$

of $\mathrm{Sh}_K(G,X)$ onto a finite set (note that $\mathbb{Q}^\times_{>0}\backslash\mathbb{A}_f^\times = \hat{\mathbb{Z}}^\times$ is profinite, and $\det K$ is open).

Proposition 1.1.2. *Suppose that $K \subset \mathrm{GL}_2(\hat{\mathbb{Z}})$ and $\det K = \hat{\mathbb{Z}}^\times$. Then there is an isomorphism*

$$\mathrm{Sh}_K(G,X) \to \Gamma\backslash X^+,$$

where X^+ is the upper half-plane.

Proof. We only give the map in one direction. Let (h,g) represent a class in $\mathrm{Sh}_K(G,X)$. By the assumption on K we have $G(\mathbb{A}_f) = G(\mathbb{Q})^+K$, so that we can write $g = sk$ with $s \in G(\mathbb{Q})^+$, $k \in K$. Map (h,g) to $s^{-1}h \in X^+$. Let's see that this is well-defined: if $g = s'k'$ with $s' \in G(\mathbb{Q})^+$, $k' \in K$, then $(s')^{-1} = \gamma s^{-1}$, where $\gamma = s(s')^{-1} = k'k^{-1} \in G(\mathbb{Q})^+ \cap K = \Gamma$. This shows that $(s')^{-1}h = \gamma s^{-1}h$ represents the same class in $\Gamma\backslash X^+$. □

Exercise: work out what happens when $\det K$ is smaller than $\hat{\mathbb{Z}}^\times$ (for instance, when K is the subgroup of $\mathrm{GL}_2(\hat{\mathbb{Z}})$ consisting of matrices congruent to 1 modulo N).

2 The Siegel modular variety

This is the most direct generalization of the tower of modular curves. But to generalize correctly, we really do need to take care of polarizations, which forces us to work with GSp_{2n}.

Let (V, ψ) be a symplectic space over $\mathbb{Q}$: V is a finite-dimensional vector space and $\psi\colon V \times V \to \mathbb{Q}$ is a nondegenerate symplectic form. Let $G = \mathrm{GSp}(V)$ be the group of symplectic similitudes. For a $\mathbb{Q}$-algebra k:

$$G(k) = \left\{ g \in GL(V \otimes k) \;\middle|\; \psi(gu, gv) = \nu(g)\psi(u, v), \text{ some } \nu(g) \in k^\times \right\}$$

Thus G comes packaged with a character $\nu\colon G \to \mathbb{G}_m$, the *similitude character*.

Suppose J is a complex structure on $V_\mathbb{R}$ such that $\psi(Ju, Jv) = \phi(u, v)$ for all $u, v \in V_\mathbb{R}$. As we know, a complex structure is the same thing as a Hodge structure of type $\{(-1, 0), (0, -1)\}$. Let $h\colon S \to \mathrm{GL}(V_\mathbb{R})$ be the associated morphism of real algebraic groups. For $a + bi \in \mathbb{C}^\times = S(\mathbb{R})$, we have

$$h(a + bi)(v) = av + bJv.$$

Lemma 2.0.3. *h factors through $G_\mathbb{R} = \mathrm{GSp}(V)$.*

Proof. For $z = a + bi \in C^\times$ we have

$$\begin{aligned} \psi(h(z)u, h(z)v) &= \psi((a + bJ)u, (a + bJ)v) \\ &= \psi(au, av) + \psi(bJu, bJv) + \psi(bJu, av) + \psi(au, bJv) \\ &= |z|^2 \psi(u, v) \end{aligned}$$

where in the last step we used $\psi(au, bJv) = \psi(aJu, -bv) = -\psi(bJu, av)$. Thus $h(z)$ is a simplectic similitude with factor $|z|^2$. ☐

In the example where $\psi(u, v) = u^t \Psi v$ for the matrix $\Psi = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}$, an example of a complex structure J satisfying $\psi(Ju, Jv) = \psi(u, v)$ is $J = \Psi$ itself. Then $h(z) = aI_{2g} + bJ \in G(\mathbb{R})$.

When does X satisfy the axioms necessary to define a hermitian symmetric domain? We need to know that $\operatorname{Lie} G$ is of type $\{(1, -1), (0, 0), (1, -1)\}$ (condition (SV1)). We have

$$\operatorname{Lie} G_\mathbb{R} = \left\{ f \in \operatorname{End} V_\mathbb{R} \;\middle|\; \psi(f(u), v) = -\psi(u, f(v)) \right\}.$$

Under the Hodge structure induced by $h\colon S \to G_{\mathbb{R}}$, this gets an action of $\operatorname{ad} h(z)$, namely $f \mapsto h(z)^{-1} \circ f \circ h(z)$. We have

$$\operatorname{Hom}(V_{\mathbb{C}}, V_{\mathbb{C}}) = \bigoplus_{\{\pm 1, \pm 1\}} \operatorname{Hom}(V^{\pm}, V^{\pm}))$$

where $V^{+} = V^{-1,0}$ and $V^{-} = V^{0,-1}$. On $V^{\pm}$, $h(z)$ acts as z (resp., $\overline{z}$), so $h(z)$ acts on these factors as 1, $z/\overline{z}$, and $\overline{z}/z$ as required.

Condition $(SV2)$ is that $\operatorname{ad} h(i) = \operatorname{ad} J$ be Cartan on the adjoint group of $G_{\mathbb{R}}$. But this is equivalent to the condition that every representation (or one faithful representation) of $G_{\mathbb{R}}$ be $h(i)$-polarizable. Let's examine the condition that the tautological representation $V_{\mathbb{R}}$ is J-polarized by ψ. This just means that $\psi(u, Jv)$ is positive definite.

The corresponding real form of $G_{\mathbb{R}}$ is

$$G^{\sigma}(\mathbb{R}) = \left\{ g \in \operatorname{GSp}(V_{\mathbb{C}}) \;\middle|\; \overline{g} = JgJ^{-1} \right\}$$

For g to live in here we need $g^t \Psi g = \nu(g)\Psi$.

Let $H\colon V_{\mathbb{C}} \times V_{\mathbb{C}} \to \mathbb{C}$ be the hermitian form

$$H(u, v) = \psi(u, Jv) + i\psi(u, v)$$

Then $G^{\sigma} \cong U(H) \cap \operatorname{Sp}(V_{\mathbb{C}})$. For this to be compact, we need H to be definite, which is to say that (for some choice of sign) $\pm\psi(u, Ju) > 0$ for all nonzero $u \in V_{\mathbb{R}}$. This is the same as saying that the complex structure h is polarized by $\pm\psi$. If one $h \in X$ has this property, then they all do.

For the remainder of the section, let X be the set of complex structures on $V_{\mathbb{R}}$ which are polarized by $\pm\psi$. We have $X = X^{+} \cup X^{-}$, where X^{+} is the Siegel upper half-plane.

2.1 The moduli interpretation of the Siegel modular variety

2.2 Quick review of abelian varieties

Let AV be the category of abelian varieties over $\mathbb{C}$. By now we know that this category is equivalent to the category of polarizable integral Hodge structures of type $\{(-1, 0), (0, -1)\}$:

Proposition 2.2.1. *$A \mapsto H_1(A, \mathbb{Z})$ is an equivalence from* AV *to the category of polarizable integral Hodge structures of type* $\{(-1, 0), (0, -1)\}$.

Recall that an integral Hodge structure is a free $\mathbb{Z}$-module $V_{\mathbb{Z}}$ together with a Hodge structure on the real vector space $V_{\mathbb{Z}} \otimes \mathbb{R}$. Hodge structures of type $\{(-1, 0), (0, -1)\}$ are the same as complex structures J on $L \otimes \mathbb{R}$.

For an integral Hodge structure $V_{\mathbb{Z}}$ of weight n to be *polarizable* means there exists a nondegenerate alternating morphism of integral Hodge structures

$$\psi \colon V_{\mathbb{Z}} \times V_{\mathbb{Z}} \to \mathbb{Z}(n)$$

(a *polarization*) for which the induced Hermitian form

$$H(v, w) = \psi_{\mathbb{C}}(v, iw) + i\psi_C(v, w)$$

is definite on V^{pq} of sign i^{p-q-k}. (Further review: $\mathbb{Z}(n)$ is the integral Hodge structure on $\mathbb{Z}$ of type $\{(-n, n)\}$; in terms of the h we have $h(z) = |z|^{-n}$ on $\mathbb{R}(n) = \mathbb{Z}(n) \otimes \mathbb{R}$. The condition that ψ be a morphism of Hodge structures is equivalent to the condition that $\psi(V^{pq}, V^{rs}) = 0$ unless $p = s$ and $q = r$.

In the case of Hodge structures of type $\{(-1, 0), (0, -1)\}$, the weight is -1. An alternating map

$$\psi \colon V_{\mathbb{Z}} \times V_{\mathbb{Z}} \to \mathbb{Z}(-1)$$

is a morphism of Hodge structures exactly when $\psi(V^{-1,0}, V^{-1,0}) = 0$ and similarly for $V^{0,-1}$.

The corresponding abelian variety is then $V_C^{-1,0}/V_{\mathbb{Z}}$, so that $V_{\mathbb{Z}} = H_1(A, \mathbb{Z})$ and $\operatorname{Lie} A = V_C^{-1,0} = V_C / \operatorname{Fil}^0 V_{\mathbb{C}}$.

Let AV^0 be the *isogeny* category: objects are still abelian varieties over $\mathbb{C}$, but morphisms are defined by

$$\operatorname{Hom}_{\mathrm{AV}^0}(A, B) = \operatorname{Hom}_{\mathrm{AV}}(A, B) \otimes \mathbb{Q}$$

An isogeny is a morphism $A \to B$ which divides $n \colon A \to A$ for some nonzero $n \in \mathbb{Z}$; thus isogenies in AV are the same as isomorphisms in AV^0.

Lemma 2.2.2. *The functor $A \mapsto H_1(A, \mathbb{Q})$ is an isomorphism from* AV^0 *to the category of polarizable rational Hodge structures of type* $\{(-1, 0), (0, -1)\}$.

Proof. Certainly the functor is well-defined. Let's construct the inverse functor. Let $V_{\mathbb{Q}}$ be a polarizable rational Hodge structure of type $\{(-1,0),(0,-1)\}$. Suppose that $\psi\colon V_{\mathbb{Q}} \times V_{\mathbb{Q}} \to \mathbb{Q}(-1)$ is the polarization. Let $V_{\mathbb{Z}} \subset V_{\mathbb{Q}}$ be a lattice small enough so that ψ is integral on $V_{\mathbb{Z}}$. Then $V_{\mathbb{Z}}$ is a polarizable integral Hodge structure and thus defines an abelian variety A. Suppose $V_{\mathbb{Z}}$ and $V'_{\mathbb{Z}}$ are two different such lattices, corresponding to A and A', then there exists a lattice $V''_{\mathbb{Z}}$ contained in their intersection, and then the resulting A'' receives isogenies from both A and A'. Thus A and A' are isomorphic in AV^0. □

Note that in the lemma, polarizations of the Hodge structure correspond to polarizations of A.

If A is an abelian variety over $\mathbb{C}$, we have the rational Tate module $V_f A = TA \otimes \mathbb{A}_f$, where $TA = \varprojlim A[N]$. This is a free $\mathbb{A}_f$-module of rank 2; note that it is isomorphic to $H_1(A, \mathbb{A}_f) = H_1(A, \mathbb{Q}) \otimes \mathbb{A}_f$. A polarization λ on A induces a polarization on $H_1(A, \mathbb{Z})$ and thus on $V_f A$, which we will still call λ.

Suppose as before that (V, ψ) is a fixed symplectic space over $\mathbb{Q}$, and let $G = \mathrm{GSp}(V)$. Recall that X is the set of complex structures on V which are polarized by $\pm\psi$. For $K \subset G(\mathbb{A}_f)$ we get

$$\mathrm{Sh}_K(G, X) = G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K).$$

Proposition 2.2.3. $\mathrm{Sh}_K(G,X)$ *classifies isomorphism classes of triples of the form* $(A, \lambda, \eta K)$*, where*

- A *is an object of* AV^0*,*
- $\pm\lambda$ *is a polarization of* A*,*
- ηK *is a* K*-orbit of* $\mathbb{A}_f$*-linear isomorphims*

  $$\eta\colon V \otimes \mathbb{A}_f \to V_f(A)$$

  which carry ψ *onto an* $\mathbb{A}_f^\times$*-multiple of* λ*.*

Here, an isomorphism between two triples $(A, \lambda, \eta) \to (A', \lambda', \eta')$ *is an isomophism* $A \to A'$ *in* AV^0 *(thus an isogeny between the abelian varieties) which sends* λ *to a* $\mathbb{Q}^\times$*-multiple of* λ' *and* ηK *to* $\eta' K$*.*

Proof. Suppose $(A, \lambda, \eta K)$ is a given triple. By the lemma, A corresponds to a polarizable rational Hodge structure $W = H_1(A, \mathbb{Q})$, and λ corresponds to a $\pm$-polarization of W, which we'll still call λ. Meanwhile, we have the isomorphism η, which identifies $V \otimes \mathbb{A}_f$ with $V_f(A) = W \otimes \mathbb{A}_f$. This shows that V and W are rational symplectic modules of the same dimension. There is only one isomorphism class of these: let

$$a \colon V \to W$$

be an isomorphism such that $\lambda(av, aw) = \psi(v, w)$ for all $v, w \in V$. Use a to pull back the Hodge structure on W to a Hodge structure h on V. Since a preserves the symplectic structures, h is polarized by $\pm\psi$ and therefore represents an element of X.

Let $g \in \mathrm{GL}(V \otimes \mathbb{A}_f)$ be $g(v) = (a \otimes 1)^{-1}\eta(v)$, where $a \otimes 1 \colon V \otimes \mathbb{A}_f \to W \otimes \mathbb{A}_f$ is the adelic version of a. The hypothesis that η carries ψ onto a multiple of λ implies that $g \in G(\mathbb{A}_f)$.

The triple $(A, \lambda, \eta K)$ then gets sent to the class of (h, g) in $\mathrm{Sh}_K(G, X)$. Let us check that this is well-defined. If we chose a different isomorphism $a' \colon V \to W$, then $a' = a\gamma$ for some $\gamma \in G(\mathbb{Q})$. Then we would have ended up with the pair $(\gamma^{-1}h, \gamma^{-1}g)$, which is the same as (h, g) in $\mathrm{Sh}_K(G, X)$.

In the opposite direction, if (h, g) is a class in $\mathrm{Sh}_K(G, X)$, then h corresponds to a Hodge structure on V for which $\pm\psi$ is a polarization. This gives a pair (A, λ), where A is an abelian variety and λ is a $\pm$-polarization. Then $V = H_1(A, \mathbb{Q})$ and $V \otimes \mathbb{A}_f \cong V_f(A)$. Let η be the composition

$$V \otimes \mathbb{A}_f \xrightarrow{g} V \otimes \mathbb{A}_f \xrightarrow{\sim} V_f(A).$$

Since g is a symplectic similitude, it carries ψ onto a multiple of λ. $\square$

3 Shimura varieties of Hodge type

If (G, X) is a general Shimura datum, one might wonder whether $\mathrm{Sh}_X(G, X)$ also classifies abelian varieties with some extra structure.

Definition 3.0.4. (G, X) is of *Hodge type* if there exists a symplectic space (V, ψ) over $\mathbb{Q}$ and a closed embedding $G \to G(\psi) = \mathrm{GSp}(\psi)$ which carries X onto $X(\psi)$, where $X(\psi)$ is the space of Hodge structures on V of type $\{(-1, 0), (0, -1)\}$ for which ψ is a $\pm$-polarization.

3.1 Example: A unitary Shimura variety

An example of this arises when G is a unitary group over $\mathbb{Q}$. Let $F = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field, let V_0 be a finite-dimensional F-vector space, and let $H\colon V_0 \times V_0 \to F$ be a nondegenerate hermitian form (relative to complex conjugation on F). Let $G = \mathrm{GU}(V_0)$ be the group of unitary similitudes of V_0. For a $\mathbb{Q}$-algebra k, we have

$$G(k) = \left\{ g \in \mathrm{GL}(V_0 \otimes_{\mathbb{Q}} k) \;\middle|\; H(gu, gv) = \nu(g)H(u,v) \right\}$$

where $\nu(g)$ is a scalar. Note that $\nu(g)$ lies in k (set $u = v$).

We have $G_{\mathbb{R}} = \mathrm{GU}(H_{\mathbb{C}})$, where $H_{\mathbb{C}}$ is the hermitian form on $V_0 \otimes_F \mathbb{C}$ induced from H. Its isomorphism class only depends on the signature (p,q) of $H_{\mathbb{C}}$.

Let V be the underlying $\mathbb{Q}$-vector space of V_0, and let $\psi\colon V \times V \to \mathbb{Q}$ be

$$\psi(v,w) = \operatorname{im} H(v,w),$$

where im means project onto the $\sqrt{-d}$-coordinate. Then ψ is an alternating form on V. Note that $\dim_{\mathbb{Q}} V = 2\dim_F V_0$. We get an embedding of algebraic groups $G \to G(\psi)$, because if $g \in G$, then $\psi(gv, gw) = \operatorname{im} H(gv, gw) = \operatorname{im} H(v,w) = \psi(v,w)$.

What's the image? Since V_0 was an F-vector space, we get an endormophism $\alpha\colon V \to V$ which corresponds to multiplication by $\sqrt{d}$. Then everything in the image of $G(\psi)$ has to commute with α. The endomorphism α corresponds to a map $V \otimes V^\vee \to \mathbb{Q}$, or (via ψ) a map $t\colon V \otimes V \to \mathbb{Q}(-1)$, which works out as $t(v,w) = \psi(v, \alpha w)$. Such morphisms have an action of $G(\psi)$, by $t^g(v,w) = \nu(g)^{-1} t(gv, gw)$, and for g to commute with α it is necessary and sufficient that $t^g = t$.

If g commutes with α and lies in $G(\psi)$, then since

$$\operatorname{re} H(v,w) = \operatorname{im} \sqrt{-d} H(w,v) = \operatorname{im} H(\alpha w, v) = \psi(\alpha w, v),$$

we find that $g \in G$. Thus $G \subset G(\psi)$ is exactly the subgroup that preserves the tensor t.

Let $\tau\colon F \to \mathbb{C}$ be an embedding, and let $V_{0,\mathbb{C}} = V_0 \otimes_{F,\tau} \mathbb{C}$, so that $V_{0,\mathbb{C}}$ is a hermitian space under H of signature (p,q). Let $e_1, \ldots, e_n$ be an orthogonal basis for $V_{0,\mathbb{C}}$ with $H(e_j, e_j) = \pm 1$ (depending on whether $j \le p$ or not). Let J be the complex structure on $V_{\mathbb{R}}$ coming from the endomorphism

$\sqrt{-1} \in \mathbb{C} = F \otimes_{\mathbb{Q}} \mathbb{R}$. The underlying real vector space of $V_{0,\mathbb{C}}$ is $V_{\mathbb{R}}$, which therefore has basis $\{e_j, Je_j\}$. Then $\psi(e_j, Je_j) = \pm 1$ depending on whether $j \leq p$.

Now let J' be the complex structure on $V_{\mathbb{R}}$ which satisfies $J^2 = -I_{V_{\mathbb{R}}}$, $JJ' = J'J$, and

$$J'e_j = \begin{cases} Je_j & 1 \leq j \leq p \\ -Je_j & p+1 \leq j \leq n. \end{cases}$$

Then J' defines a Hodge structure which is polarized by ψ. We have to check that for nonzero $v \in V_{\mathbb{R}}$ we have $\psi(v, J'v) > 0$. It is enough to check this for the e_j: we have $\psi(e_j, J'e_j) = 1$.

Being a Hodge structure on V which is polarized by ψ, J' determines a morphism $h\colon S \to G(\psi)$. But since J' and J commute, h factors through G. Let X be the $G(\mathbb{R})$-conjugacy class of h. Then X classifies all complex structures on $V_{\mathbb{R}}$ which are polarized by ψ and which commute with J. X is a hermitian symmetric domain (exercise!).

Lemma 3.1.1. *$X = U(p,q)/U(p) \times U(q)$ is the parameter space for positive p-dimensional subspaces of $V_{0,\mathbb{C}}$ (with respect to the hermitian form H).*

Proof. An element of X is a complex structure J' on $V_{\mathbb{R}}$ polarized by ψ which commutes with J. Note that $J'J^{-1}$ is an involution and therefore is diagonalizable with eigenvalues ± 1. To such a J' we associate $W = V_{\mathbb{R}}^{J=J'}$. Then for nonzero $v \in W$ we have $H(v,v) = \psi(v, Jv) = \psi(v, J'v) > 0$, so that W is positive. Similarly $W' = V_{\mathbb{R}}^{J=-J'}$ is negative, and $V_{\mathbb{R}} = W \oplus W'$. Since the signature of H is (p,q) we must have $\dim W = p$ and $\dim W' = q$.

Conversely if $V_{0,\mathbb{C}} = W \oplus W'$ is a decomposition into positive and negative subspaces, then let J' be the unique complex structure on $V_{\mathbb{R}}$ which is J on W and $-J$ on W'.

The set of such subspaces W is permuted transitively by $U(p,q)$, and the stabilizer of any particular W is $U(p) \times U(q)$. □

The above lemma shows that there is an embedding of hermitian symmetric spaces $U(p \times q)/U(p) \times U(q) \to \mathrm{Sp}(2n)/U(n)$ which is compatible with the embedding $G \to \mathrm{GSp}(2n)$.

Now we can give an interpretation of the Shimura variety $\mathrm{Sh}_K(G,X)$.

Proposition 3.1.2. $\mathrm{Sh}_K(G,X)$ *classifies equivalence classes of quadruples $(A, \lambda, \iota, \eta K)$, where*

- *A is an object of* AV^0*,*
- $\lambda\colon A \to A^\vee$ *is a polarization,*
- $\iota\colon F \to \operatorname{End} A$ *is an action of F on A up to isogeny,*
- ηK *is a K-orbit of F-linear isomorphisms*

$$\eta\colon V \otimes \mathbb{A}_f \to V_f(A).$$

These are required to satisfy the following properties:

1. *η carries ψ onto a $\mathbb{A}_f^\times$-multiple of λ.*
2. *There exists an F-linear isomorphism $a\colon V \to H_1(A, \mathbb{Q})$ which carries ψ onto an $\mathbb{Q}^\times$-multiple of λ.*

Two quadruples are considered equivalent when there is an F-linear isogeny $A \to A'$ carrying one λ onto a $\mathbb{Q}^\times$-multiple of the other, and one ηK onto the other.

Proof. Let $(A, \lambda, \iota, \eta K)$ be such a quadruple. The pair (A, λ) determines a rational Hodge structure $W = H_1(A, \mathbb{Q})$ which is polarized by an alternating form; as usual we abuse notation by calling this form λ. W gets the structure of an F-vector space via ι, and the Hodge structure is compatible with it. Using the F-isomorphism $a\colon V \to W$, we get a Hodge structure h on V; since a is compatible with the symplectic structures, h is polarized by ψ. Since a was F-linear, h commutes with the action of $F \otimes \mathbb{R} = \mathbb{C}$. Therefore h belongs to X.

Let $g \in \mathrm{GL}(V \otimes_{\mathbb{Q}} \mathbb{A}_f)$ be defined by $g(v) = (a \otimes 1)^{-1}\eta(v)$, where $a \otimes 1\colon V \otimes \mathbb{A}_f \to W \otimes \mathbb{A}_f$ is the adelic version of a. Since a and η are compatible with the symplectic structures and with F, we see that $g \in G(\mathbb{A}_f)$. The quadruple $(A, \lambda, \iota, \eta K)$ gets sent to the class of (h, g) in $\mathrm{Sh}_K(G, X) = G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f))/K$.

The opposite direction is very similar to what appears in the proof of Prop. 2.2.3. □

We'll finish this section by reinterpreting the condition (2) appearing in Prop. 3.1.2.

Proposition 3.1.3. *Let $(A, \lambda, \iota, \eta K)$ be a quadruple of objects as in Prop. 3.1.2 satisfying the following conditions:*

- *The polarization λ and the action ι are required to be compatible, in the sense that for $\alpha \in F$, the diagram*

$$\begin{array}{ccc} A & \xrightarrow{\lambda} & A^\vee \\ {\scriptstyle \iota(\alpha)}\downarrow & & \downarrow{\scriptstyle \iota(\overline{\alpha})} \\ A & \xrightarrow[\lambda]{} & A^\vee \end{array}$$

  commutes.

- *η carries ψ onto an $\mathbb{A}_f^\times$-multiple of λ.*

Then there exists an F-linear isomorphism $a\colon V \to H_1(A, \mathbb{Q})$ carrying ψ onto a $\mathbb{Q}^\times$-multiple of λ if and only if the action of $\iota(\sqrt{-d})$ on the complex vector space $\operatorname{Lie} A$ *has eigenvalues $\sqrt{-d}$ and $-\sqrt{-d}$ with multiplicities p and q, respectively.*

Proof. The key fact here is a Hasse principle for hermitian inner product spaces: two such spaces are isomorphic if and only if they are isomorphic over every completion. We have the hermitian inner product space V_0. If $(A, \lambda, \iota, \eta K)$ is a quadruple, then $W = H_1(A, \mathbb{Q})$ has the structure of an F-vector space with an alternating $\mathbb{Q}$-bilinear form λ. The condition on the polarization λ shows that $\lambda(\alpha v, w) = \lambda(v, \overline{\alpha} w)$ for $\alpha \in F$. If we define $H'\colon W \times W \to F$ by

$$H'(v, w) = \lambda(v, \sqrt{-d} w) + \sqrt{-d} \lambda(v, w),$$

then H' is F-linear in v and F-semilinear in w (check this), so that W becomes a hermitian inner product space for $F/\mathbb{Q}$. The existence of η shows that the hermitian inner product spaces V_0 and W are isomorphic at every finite place. Over the real place, $V_{0,\mathbb{R}}$ and $W_\mathbb{R}$ will be isomorphic if and only if $W_\mathbb{R}$ has signature (p, q).

The real vector space $W_\mathbb{R}$ has one complex structure J coming from the endomorphism $\sqrt{-1} \in \mathbb{C} \cong F \otimes \mathbb{R}$, with respect to which $H(v, w) = \psi(v, Jw) + i\psi(v, w)$ has signature (p, q). But then A gives $W_\mathbb{R}$ another complex structure (the Hodge structure), with respect to which $H'(v, w) = \psi(v, J'w) + i\psi(v, w)$ is positive definite. We have $\operatorname{Lie} A = W_\mathbb{C}^{J'=i}$. J and J' commute.

Suppose that the signature of $W_\mathbb{R}$ is (p', q'). We have already seen from the proof of Lemma 3.1.1 that $\dim W_\mathbb{R}^{J=J'} = p'$ (dimension as a complex vector space using J). We're done as soon as we observe that $W_\mathbb{R}^{J=J'} = W_\mathbb{C}^{J=J'=i} = (\operatorname{Lie} A)^{J=i}$. $\square$

3.2 PEL Shimura varities

The general PEL setup involves replacing the imaginary quadratic field $F/\mathbb{Q}$ with a semisimple $\mathbb{Q}$-algebra B which comes equipped with an involution $x \mapsto x^*$. One also needs a V, which will be a B-module equipped with an alternating form ψ satisfying $\psi(bv, w) = \psi(v, b^*w)$ for $b \in B$. From here one defines an algebraic group $G/\mathbb{Q}$ as the group of B-linear automorphisms of V which preserve ψ up to a scalar. Considering complex structures on V which are polarized by ψ, one gets a hermitian symmetric domain X, and a family of Shimura varieties $\mathrm{Sh}_K(G, X)$. These parametrize abelian varieties with endomorphisms by B. For details, see Milne's notes.