Canonical models*

Avi Zeff

1. Definition of canonical models

Our next goal is to show that Shimura varieties admit models over number fields. In particular, we saw before that for every Shimura datum (G, X) and compact open subgroup $K \subset G(\mathbb{A}_f)$ we get a map

$$\mathrm{Sh}_K(G, X) \to \pi_K$$

where π_K is a "zero-dimensional Shimura variety" and the fibers of the map are the connected components, themselves connected Shimura varieties attached to (G, X^+) where X^+ is a connected component of X. Our goal is to find a model of this map over some number field E depending on (G, X), called the reflex field. This gives a description of the action of $\mathrm{Aut}(\mathbb{C}/E)$ on the set of connected components, which defines a model of each connected component over a finite extension of E.

Why should we expect Shimura varieties to have algebraic models at all? If not, i.e. the smallest model of a Shimura variety is transcendental over $\mathbb{Q}$, then we could write it as $\mathrm{Sh}(G, X) \to \operatorname{Spec} K(t)$ for some extension K of $\mathbb{Q}$, which extends to a family over K of varieties indexed by t, which gives nontrivial deformations of our Shimura variety. On the other hand we expect Shimura varieties to be locally rigid, which would imply they would have to be defined over $\overline{\mathbb{Q}}$: up to isomorphism, there are only countably many arithmetic locally symmetric varieties at all. It is possible to make this argument rigorous, but we will not do so.

The terminology "reflex field" recalls the reflex field of a CM type, and indeed our reflex field will be a direct generalization. Let (E, Φ) be a CM type, and let $T = \mathrm{Res}_{E/\mathbb{Q}}(\mathbb{G}_{\mathrm{m}})$. We can choose an isomorphism $T(\mathbb{R}) = (E \otimes_{\mathbb{Q}} \mathbb{R})^\times \simeq (\mathbb{C}^\Phi)^\times$ and a homomorphism $h_\Phi : \mathbb{S} \to T_{\mathbb{R}}$ defined on real points by

$$z \mapsto (z, \ldots, z) \in (\mathbb{C}^\Phi)^\times,$$

i.e. with the Φ-action of E. On $\mathbb{C}$-points, we get points in both $(\mathbb{C}^\Phi)^\times$ and $(\mathbb{C}^{\bar{\Phi}})^\times$, i.e.

$$(z_1, z_2) \mapsto (z_1, \ldots, z_2, z_2, \ldots, z_2) \in (\mathbb{C}^\Phi)^\times \times (\mathbb{C}^{\bar{\Phi}})^\times.$$

This gives a cocharacter $\mathbb{C}^\times \to T(\mathbb{C})^\times \simeq (\mathbb{C}^\Phi)^\times \times (\mathbb{C}^{\bar{\Phi}})^\times$ by

$$z \mapsto (z, \ldots, z, 1, \ldots, 1),$$

defined over $\overline{\mathbb{Q}}$.

The reflex field of (E, Φ) is the subfield of $\overline{\mathbb{Q}}$ fixed the automorphisms $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ which fix Φ; since the action on $T(\mathbb{C})^\times$ is by Φ, these are the same ones which fix μ. This suggests that we define the reflex field of a Shimura datum (G, X) as the field fixed by the automorphisms fixing $\mu = h(-, 1)$ for some $h \in X$.

*These notes are based on chapters 12-14 of [1].

This has two problems: first of all, such a cocharacter μ is not well-defined since it depends on a choice of $h \in X$; and second, it is a priori defined only over $\mathbb{C}$, not $\overline{\mathbb{Q}}$.

To fix the first, we note that h is well-defined up to $G(\mathbb{R})$-conjugacy, so we instead look at conjugacy classes of cocharacters. To get such a class over $\overline{\mathbb{Q}}$, we look at the set of conjugacy classes varying over fields: for any subfield k of $\mathbb{C}$, let $\mathcal{C}(k)$ be the set of $G(k)$-conjugacy classes of cocharacters of G defined over k, i.e. $\mathcal{C}(k) = G(k)\backslash \operatorname{Hom}(\mathbb{G}_{\mathrm{m}}, G_k)$ with action by conjugation.

This is functorial in k, but how exactly it depends on the field is unclear; we'd like to relate it to something where we can understand the relationship better. Suppose that G splits over k (e.g. when k is algebraically closed), so that it contains a split maximal torus T. Let N be the normalizer of T in G_k and $W = N/T$ be the Weyl group. There is a map $\operatorname{Hom}(\mathbb{G}_{\mathrm{m}}, T_k) \to \operatorname{Hom}(\mathbb{G}_{\mathrm{m}}, G_k)$ by postcomposing with the inclusion; after quotienting by $G(k)$ on the right, the map descends to the quotient by the Weyl group on the left.

Lemma 1.1. *The map $W\backslash \operatorname{Hom}(\mathbb{G}_m, T_k) \to G(k)\backslash \operatorname{Hom}(\mathbb{G}_m, G_k) = \mathcal{C}(k)$ is bijective.*

Proof. Any two maximal tori are conjugate, so any map $\mathbb{G}_{\mathrm{m}} \to G_k$ can up to conjugacy be taken to have image in T_k, i.e. the map is surjective. Suppose that μ and μ' are cocharacters of T over k which are $G(k)$-conjugate, say $\mu = \operatorname{ad}(g) \circ \mu'$ for $g \in G(k)$. Then $\operatorname{ad}(g)(T)$ and T are both maximal split tori in the centralizer C of $\mu(\mathbb{G}_{\mathrm{m}})$ (T centralizes $\mu(\mathbb{G}_{\mathrm{m}})$ by commutativity, and since it also centralizes $\mu'(\mathbb{G}_{\mathrm{m}})$ for the same reason $\operatorname{ad}(g)(T)$ centralizes $\operatorname{ad}(g)(\mu'(\mathbb{G}_{\mathrm{m}})) = \mu(\mathbb{G}_{\mathrm{m}})$), a connected reductive group. Therefore they agree up to $C(k)$-conjugacy, so $\operatorname{ad}(c)(\operatorname{ad}(g)(T)) = \operatorname{ad}(cg)(T) = T$ for some $c \in C(k)$. Thus cg normalizes T, so is in $N(k)$, and $\operatorname{ad}(cg) \circ \mu' = \operatorname{ad}(c)(\operatorname{ad}(g) \circ \mu') = \operatorname{ad}(c) \circ \mu = \mu$ since $c \in C(k)$, so μ and μ' are in the same N-orbit and thus in the same W-orbit. ∎

Now, $W(\overline{\mathbb{Q}})$ and $\operatorname{Hom}(\mathbb{G}_{\mathrm{m}}, T_{\overline{\mathbb{Q}}})$ agree with $W(\mathbb{C})$ and $\operatorname{Hom}(\mathbb{G}_{\mathrm{m}}, T_{\mathbb{C}})$, so the lemma implies that $\mathcal{C}(\mathbb{C}) = \mathcal{C}(\overline{\mathbb{Q}})$. Since $h(-,1) : \mathbb{S}(\mathbb{R}) \to \mathbb{S}(\mathbb{C}) \to G(\mathbb{C})$ gives a unique element of $\mathcal{C}(\mathbb{C})$ depending only on X, it follows that we get a well-defined conjugacy class of cocharacters μ_X defined over $\overline{\mathbb{Q}}$.

Definition 1.2. The reflex field E of a Shimura datum (G, X) is the field of definition of μ_X, i.e. the field fixed by the automorphisms in $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing μ_X in $\mathcal{C}(\overline{\mathbb{Q}})$, or equivalently stabilizing it as a subset of $\operatorname{Hom}(\mathbb{G}_{\mathrm{m}}, G_{\overline{\mathbb{Q}}})$.

From the above in the case where $G = \operatorname{Res}_{E/\mathbb{Q}}(\mathbb{G}_{\mathrm{m}})$ and X is the singleton with element h_Φ coming from a CM-type (E, Φ), the reflex field of (G, X) is just the reflex field of the CM-type. More generally, if T is any torus over $\mathbb{Q}$ and $h : \mathbb{S} \to T_{\mathbb{R}}$ is a homomorphism, then the reflex field E is the fixed field of the subgroup of $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing the cocharacter μ_h.

Suppose that (G, X) is a simple PEL datum of type (A) or (C), corresponding to a simple algebra B with involution over $\mathbb{Q}$. Then the reflex field E is generated by the traces of the elements of B; the automorphisms fixing E are exactly those fixing the abelian varieties, as well as the associated PEL structure, parametrized by the Shimura varieties for (G, X). Thus the reflex field is the natural field of definition of the corresponding moduli problem.

Consider for example the case of a quaternion algebra B over a totally real number field F. In this case the cocharacter μ is (up to conjugacy, i.e. base change) the one sending

$z \in \mathbb{G}_m(\mathbb{R})$ to the tuple with entry 1 for the factors for which the quaternion algebra becomes isomorphic to $\mathbb{H}$ after tensoring with $\mathbb{R}$ and $\begin{pmatrix} z & \\ & 1 \end{pmatrix}$ for those where it becomes isomorphic to $M_2(\mathbb{R})$. This cocharacter is defined over $\overline{\mathbb{Q}}$, and the reflex field is the fixed field of the subgroup of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ stabilizing the split places of B over F. For example, in the case where there is exactly one split place, so that the Shimura variety is a curve, then the reflex field E is given by the corresponding embedding of F into $\mathbb{R}$.

We can now observe some properties of the reflex field. One important one is that it is always a number field: we can always find a finite extension of $\mathbb{Q}$ splitting G, and any such field k must contain E since in that case $W(k) = W(\overline{\mathbb{Q}}) = W(\mathbb{C})$ and so μ descends to k and so is defined over some subfield.

The functoriality of $\mathcal{C}$ in G implies that μ_X is functorial under inclusions of Shimura data, and in particular an inclusion $(G, X) \hookrightarrow (G', X')$ induces a reverse inclusion $E(G, X) \supset E(G', X')$. For example, Shimura varieties of Hodge type all embed into Siegel modular varieties, so one might hope that to get the largest set of possibilities for reflex fields of Shimura varieties of Hodge type we should have the reflex field of Siegel modular varieties equal to $\mathbb{Q}$. This is in fact the case, since $\mathrm{GSp}(\psi)$ is already split over $\mathbb{Q}$ and so the corresponding reflex fields must be subfields of $\mathbb{Q}$, i.e. $\mathbb{Q}$.

Now, we have defined the reflex field E over which we hope a given Shimura variety $\mathrm{Sh}(G, X)$ to be defined, and our strategy is to find the action by $\mathrm{Aut}(\mathbb{C}/E)$ to ensure that there must be a model defined over E using our theory of the corresponding action on abelian varieties. Now, at least for many kinds of Shimura varieties, their complex points parametrize abelian varieties; but in fact we only know how to compute the action on abelian varieties with complex multiplication. Thus we need to look at points on the Shimura variety corresponding to abelian varieties with complex multiplication: these are special points.

For $x \in X$, we say that x is special if the corresponding cocharacter $h_x : \mathbb{S} \to G_{\mathbb{R}}$ has image in a torus T (defined over $\mathbb{Q}$), i.e. $h_x(\mathbb{C}^\times) \subseteq T(\mathbb{R})$. In this case we say that (T, x) is a special pair. If we assume that the Shimura datum (G, X) additionally satisfies axioms (4) and (6), i.e. the weight homomorphism is rational and $Z(G)^\circ$ splits over a CM field, then we call x a CM point, and (T, x) a CM pair.

Consider for example the Shimura datum given by GL_2 acting on the upper and lower half-planes $\mathcal{H}^\pm = \mathbb{C} - \mathbb{R}$, corresponding to the connected Shimura variety $(\mathrm{SL}_2, \mathcal{H})$. A point $z = x + iy \in \mathcal{H}^\pm$ corresponds to the cocharacter $h_z : \mathbb{S} \to (\mathrm{GL}_2)_{\mathbb{R}}$ given by the conjugation of $h_0 : a + bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ by some g such that $gi = z$, e.g. $\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}$; the stabilizer in $\mathrm{GL}_2(\mathbb{R})$ of z is exactly $h_z(\mathbb{C}^\times)$. Thus z is special if its stabilizer is contained in a torus. All maximal tori of GL_2 are of the form $\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_m$ where E is some degree 2 étale $\mathbb{Q}$-algebra; such a torus fixes z if and only if it contains z after embedding into $\mathbb{C}$, in which case $E = \mathbb{Q}[z]$. Thus the torus cannot be split and so E is some quadratic extension of $\mathbb{Q}$, necessarily totally imaginary, and the elliptic curve $\mathbb{C}/(\mathbb{Z} + z\mathbb{Z})$ has complex multiplication by E.

Conversely, if z generates a quadratic imaginary extension $E/\mathbb{Q}$, we can embed $\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_m$ into G, where its real points stabilize z. Therefore the special points of $\mathcal{H}^\pm$ are exactly those corresponding to elliptic curves with complex multiplication.

This works more generally: whenever the Shimura variety classifies abelian varieties with some additional structure, the special points correspond to abelian varieties of CM type. In this case the theory of complex multiplication describes an action of an open subgroup of $\mathrm{Aut}(\mathbb{C})$ on these abelian varieties with structure, and thus on the corresponding special points of the Shimura variety. Our next goal is to define a general action on the special points, which does not depend on the moduli interpretation (since not all Shimura varieties have one) but agrees with the action defined by complex multiplication when there is a moduli interpretation.

Let T be a torus over $\mathbb{Q}$, with $\mu : \mathbb{S} \to T_{\mathbb{R}}$ a cocharacter defined over some finite extension E of $\mathbb{Q}$, i.e. descending to $\mu : \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{\mathrm{m}} \to T$. For $Q \in T(E)$, we can take the product over embeddings $\prod_{\rho: E \to \overline{\mathbb{Q}}} \rho(Q)$ to get something $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-invariant and therefore in $T(\mathbb{Q})$. Let $r = r(T, \mu) : \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{\mathrm{m}} \to T$ given by composing this operation with μ, i.e.

$$r(P) = \prod_{\rho: E \to \overline{\mathbb{Q}}} \rho(\mu(P))$$

for $P \in E^{\times}$.

For any special pair $(T, x) \subset (G, X)$, we can define $E(x)$ to be the field of definition of μ_x; this is the reflex field of $(T, \{x\})$ as a Shimura variety, and so is a number field. We define r_x to be the composition

$$\mathbb{A}_{E(x)}^{\times} \xrightarrow{r(\mathbb{A}_{\mathbb{Q}})} T(\mathbb{A}_{\mathbb{Q}}) \to T(\mathbb{A}_f)$$

where the last map is just the projection and the first is $r(T, \mu_x)$. For $a = (a_f, a_\infty) \in \mathbb{A}_{E(x)}^{\times} = (E(x) \otimes_{\mathbb{Q}} \mathbb{R})^{\times} \times \mathbb{A}_{E(x),f}^{\times}$, we have

$$r_x(a) = \prod_{\rho: E \to \overline{\mathbb{Q}}} \rho(\mu_x(a_f)).$$

We can now define canonical models. For a Shimura datum (G, X) and compact open subgroup $K \subset G(\mathbb{A}_f)$, write $[x, a]_K$ for the element of $\mathrm{Sh}_K(G, X) = G(\mathbb{Q}) \backslash X \times G(\mathbb{A}_f)/K$ represented by $(x, a) \in X \times G(\mathbb{A}_f)$.

Definition 1.3. A model $M_K(G, X)$ of $\mathrm{Sh}_K(G, X)$ over the reflex field $E = E(G, X)$ is canonical if for every special pair $(T, x) \subset (G, X)$ and $a \in G(\mathbb{A}_f)$, the point $[x, a]_K$ has coordinates in $E(x)^{\mathrm{ab}}$, and for every $\sigma \in \mathrm{Gal}(E(x)^{\mathrm{ab}}/E(x))$ and $s \in \mathbb{A}_{E(x)}^{\times}$ with s corresponding to σ under the map of class field theory we have

$$\sigma[x, a]_K = [x, r_x(s)a]_K.$$

In other words, $M_K(G, X)$ is canonical if every automorphism $\sigma \in \mathrm{Aut}(\mathbb{C}/E(x))$ acts on $[x, a]_K$ according to the same rule as for complex multiplication.

Definition 1.4. Let (G, X) be a Shimura datum. A model of $\mathrm{Sh}(G, X)$ over a field $k \subset \mathbb{C}$ is an inverse system $M(G, X) = (M_K(G, X))_K$ of varieties over k with a right action of $G(\mathbb{A}_f)$ whose base change to $\mathbb{C}$ recovers $(\mathrm{Sh}_K(G, X))_K = \mathrm{Sh}(G, X)$, compatibly with the $G(\mathbb{A}_f)$-action. A model $M(G, X)$ of $\mathrm{Sh}(G, X)$ over $E(G, X)$ is canonical if each $M_K(G, X)$ is canonical.

Consider for example Shimura varieties corresponding to tori. If T is a torus over $\mathbb{Q}$ and $h : \mathbb{S} \to T_{\mathbb{R}}$ is a homomorphism, then (T, h) is a Shimura datum, and $E = E(T, h)$ is the field of definition of $\mu_h = h(-, 1)$. As we saw in our section on first properties of Shimura varieties, $\mathrm{Sh}_K(T, h) = T(\mathbb{Q})\backslash\{h\} \times T(\mathbb{A}_f)/K$ is a finite set for every compact open $K \subset T(\mathbb{A}_f)$, and we have a continuous action of $\mathrm{Gal}(E^{\mathrm{ab}}/E)$ on $\mathrm{Sh}_K(T, h)$ by $\sigma[h, a]_K = [x, r_x(s)a]_K$ for any s corresponding to σ via class field theory. This action gives a model of $\mathrm{Sh}_K(T, h)$ over E, which is canonical by definition.

In fact, we can define canonical models for the Shimura pairs for CM types. We looked above at the Shimura datum (T, h_Φ) associated to a CM type (E, Φ), where $T = \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{\mathrm{m}}$ and $h_\Phi : \mathbb{S} \to (\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{\mathrm{m}})_{\mathbb{R}}$ is given on real points by $z \mapsto \Phi(z) \in (E \otimes_{\mathbb{Q}} \mathbb{R})^\times = (\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{\mathrm{m}})(\mathbb{R})$. The reflex field $E(T, h_\Phi)$ is the reflex field E^*, and $r(T, \mu_\Phi)$ is a homomorphism $\mathrm{Res}_{E^*/\mathbb{Q}} \mathbb{G}_{\mathrm{m}} \to \mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_{\mathrm{m}}$, namely the reflex norm N_{Φ^*}.

The Shimura variety for this datum has a moduli structure: for any compact open $K \subset T(\mathbb{A}_f)$, $\mathrm{Sh}_K(T, h_\Phi)$ classifies isomorphism classes of triples $(A, i, \eta K)$, where (A, i) is an abelian variety over $\mathbb{C}$ of CM-type (E, Φ) and η is a level structure, i.e. an $E \otimes \mathbb{A}_f$-linear isomorphism $V(\mathbb{A}_f) \to V_f(A)$, with isomorphisms E-linear isogenies compatible with the level structures.

We can exhibit this classification by writing down the bijection: let V be a one-dimensional vector space over E, viewed as a vector space over $\mathbb{Q}$, so that the action of E embeds T in $\mathrm{GL}_{\mathbb{Q}}(V)$. Given an abelian variety (A, i) of CM-type (E, Φ), it is E-isogenous to $A_\Phi = \mathbb{C}^\Phi/\Phi(\mathcal{O}_E)$. On cohomology, this isogeny induces an isomorphism $a : H_1(A, \mathbb{Q}) \to H_1(A_\Phi, \mathbb{Q})$; the latter is a $\mathbb{Q}$-vector space with an action of E by Φ, and so can be identified with V such that a carries h_A to h_Φ. Tensoring with $\mathbb{A}_f$, we get a composition of $E \otimes \mathbb{A}_f$-linear isomorphisms

$$V(\mathbb{A}_f) \xrightarrow{\eta} V_f(A) \xrightarrow{a} V(\mathbb{A}_f)$$

of one-dimensional $E\otimes\mathbb{A}_f$-modules, so it is an element of $\mathrm{GL}_1(E\otimes\mathbb{A}_f) = (E\otimes\mathbb{A}_f)^\times = T(\mathbb{A}_f)$. This gives a map $(A, i, \eta) \mapsto [g]$ from the set of such tuples to $T(\mathbb{Q})\backslash\{h_\Phi\} \times T(\mathbb{A}_f)$; one can check that taking η to be defined only up to K induces the right-hand side being defined up to K, and that it descends to the level of isomorphism classes and after this is a bijection.

We've seen that abelian varieties of given CM-type (E, Φ) form equivalent categories over $\overline{\mathbb{Q}}$ and over $\mathbb{C}$, so we can just as well take our triples $(A, i, \eta K)$ with (A, i) defined over $\overline{\mathbb{Q}}$. We have an action of $\mathrm{Gal}(\overline{\mathbb{Q}}/E^*)$ on the set $\mathcal{M}_K$ of such triples $\sigma(A, i, \eta K) = (\sigma A, \sigma i, \sigma\eta K)$, where $\sigma\eta$ refers to the composition

$$V(\mathbb{A}_f) \xrightarrow{\eta} V_f(A) \xrightarrow{\sigma} V_f(\sigma A).$$

Since σ fixes E^*, it does not change the CM type: $(\sigma A, \sigma i)$ is again of CM-type (E, Φ).

On the other hand, we also have an action on $\mathrm{Sh}_K(T, h_\Phi)$ by the same rule as above,

$$\sigma[g] = [r_{h_\Phi}(s)g]_K$$

for s corresponding to σ under class field theory. This defines a model of $\mathrm{Sh}_K(T, h_\Phi)$ over E^*. To see that it is really a canonical model, we need to know that this action is compatible with the moduli structure:

Proposition 1.5. *The map $\mathcal{M}_K \to \mathrm{Sh}_K(T, h_\Phi)$ sending $(A, i, \eta) \mapsto [a \circ \eta]_K$ as defined above commutes with the actions of $\mathrm{Gal}(\overline{\mathbb{Q}}/E^*)$.*

Proof. Fix $(A, i, \eta) \in \mathcal{M}_K$ and $a : H_1(A, \mathbb{Q}) \to V$ as above, and let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/E^*)$. By the main theorem of complex multiplication, there exists an E-linear isogeny $\alpha : A \to \sigma A$ such that $\alpha(N_{\Phi^*}(s)x) = \sigma x$ for s corresponding to σ and $x \in V_f(A)$, and so

$$\sigma(A, i, \eta) = (\sigma A, \sigma i, \sigma \eta) \mapsto [a \circ V_f(\alpha)^{-1} \circ \sigma \circ \eta]_K.$$

But $V_f(\alpha)^{-1} = N_{\Phi^*}(s) = r_{h_\Phi}(s)$ and so

$$\sigma(A, i, \eta) \mapsto [a \circ V_f(\alpha)^{-1} \circ \sigma \circ \eta] = [r_{h_\Phi}(s) \cdot (a \circ \eta)]_K,$$

i.e. the actions of $\mathrm{Gal}(\overline{\mathbb{Q}}/E^*)$ on $\mathcal{M}_K$ and $\mathrm{Sh}_K(T, h_\Phi)$ commute with the map $\mathcal{M}_K \to \mathrm{Sh}_K(T, h_\Phi)$. □

2. Uniqueness of canonical models

With a definition of canonical models in hand, our next goal is to show that (as the name implies) they are unique whenever they exist (up to unique isomorphism). Our strategy will be to use special points to show that certain morphisms are defined over the reflex field E, which will yield the desired isomorphisms of models; then we just need to show that we always have special points.

The first thing to do is to observe that a variety V over a field k of characteristic zero is uniquely determined (up to unique isomorphism) by the data of its base change to V_K for any algebraically closed $K \supset k$ together with an action of $\mathrm{Aut}(K/k)$ on $V(K)$. This is a corollary of the fact that the functor sending V to the data of V_K plus the action of $\mathrm{Aut}(K/k)$ on $V(K)$ is fully faithful, which is essentially an application of Zorn's lemma.

Fix a Shimura datum (G, X), let $g \in G(\mathbb{A}_f)$ and $K, K' \subset G(\mathbb{A}_f)$ be compact open subgroups such that $K' \supset g^{-1}Kg$. Then the action of g defines a map

$$\mathcal{T}(g) : \mathrm{Sh}_K(G, X)(\mathbb{C}) \to \mathrm{Sh}_{K'}(G, X)(\mathbb{C})$$

sending $[x, a]_K \mapsto [x, ag]_{K'}$, which is a map of complex algebraic varieties.

Theorem 2.1. *If $\mathrm{Sh}_K(G, X)$ and $\mathrm{Sh}_{K'}(G, X)$ have canonical models over $E = E(G, X)$, then $\mathcal{T}(g)$ is defined over E.*

Proof. From the full faithfulness of the above functor, it suffices to show that every $\sigma \in \mathrm{Aut}(\mathbb{C}/E)$ fixes $\mathcal{T}(g)$. Suppose that $x_0 \in X$ is a special point. Then the cocharacter μ_x is defined over $E(x_0)$, and so $E(x_0) \supset E$, so our first goal is to show that $\sigma(\mathcal{T}(g)) = \mathcal{T}(g)$ for $\sigma \in \mathrm{Aut}(E(x_0)/E)$.

Choose some $s \in \mathbb{A}^\times_{E(x_0)}$ corresponding to σ (or to its restriction to $E(x_0)^{\mathrm{ab}}$). We have a commutative diagram

$$\begin{array}{ccc} \mathrm{Sh}_K(G, X) & \xrightarrow{\mathcal{T}(g)} & \mathrm{Sh}_{K'}(G, X) \\ \big\downarrow \sigma & & \big\downarrow \sigma \\ \mathrm{Sh}_K(G, X) & \xrightarrow{\mathcal{T}(g)} & \mathrm{Sh}_{K'}(G, X) \end{array}$$

sending

$$\begin{array}{ccc} [x_0, a]_K & \longrightarrow & [x_0, ag]_{K'} \\ \downarrow & & \downarrow \\ [x_0, r_{x_0}(s)a]_K & \longrightarrow & [x_0, r_{x_0}(s)ag]_{K'} \end{array}$$

by the description of the action of σ on a canonical model. Thus σ and $\mathcal{T}(g)$ commute on all points of the form $[x_0, a]$ for $a \in G(\mathbb{A}_f)$ and x_0 special. By Lemma 2.2 below, such points form a dense subset of $\mathrm{Sh}_K(G, X)$ (in the Zariski topology), and so since σ and $\mathcal{T}(g)$ are continuous this is true at every point. We've restricted to σ fixing $E(x_0)$ for some special point x_0, but Lemma 2.4 below shows that in fact such σ generate all of $\mathrm{Aut}(\mathbb{C}/E)$. □

Lemma 2.2. *The set of points of the form $[x, a]_K$ for $a \in G(\mathbb{A}_f)$ and x special is dense in $\mathrm{Sh}_K(G, X)$.*

Proof. In fact, $\{[x, a]_K | a \in G(\mathbb{A}_f)\}$ is dense for any x; the result then follows from the existence of a special point, which we split off as Lemma 2.3 below. To see this, note that by real approximation $G(\mathbb{Q})x$ is dense in X for the complex topology, and so $G(\mathbb{Q})x \times G(\mathbb{A}_f)$ is dense in $X \times G(\mathbb{A}_f)$, so so is its image in $\mathrm{Sh}_K(G, X)(\mathbb{C})$. Since the complex topology is stronger than the Zariski topology, kal vachomer the image of $G(\mathbb{Q})x \times G(\mathbb{A}_f)$ is also dense in the Zariski topology. Since $[gx, b]_K = [x, g^{-1}b]_K$, this image consists of elements of the form $[x, a]_K$, and clearly contains all of them. □

Lemma 2.3. *For every Shimura datum (G, X), there exists a special point in X.*

Proof. Choose some $x \in X$. Since $h_x(\mathbb{C}^\times)$ is abelian, it is contained in some maximal torus T (defined over $\mathbb{R}$) in $G_\mathbb{R}$. As a maximal torus, T is the centralizer on $G_\mathbb{R}$ of some regular $\lambda \in \mathrm{Lie}(G_\mathbb{R})$. For any $\lambda_0 \in \mathrm{Lie}(G)$ sufficiently close to λ, it will also be regular and so has centralizer another maximal torus T_0, now defined over $\mathbb{Q}$. Since λ_0 and λ are chosen very close together, they are in particular in the same connected component and so their stabilizers are conjugate over $\mathbb{R}$: there is some $g \in G(\mathbb{R})$ such that $(T_0)_\mathbb{R} = gTg^{-1}$. Then gh_xg^{-1} has image in $(T_0)_\mathbb{R}$; this is the cocharacter corresponding to the action of g on x by conjugation, i.e. $h_{gx} = gh_xg^{-1}$, and so h_{gx} has image in a torus defined over $\mathbb{Q}$ and so gx is a special point. □

A similar but more involved argument (using the Hilbert irreducibility theorem) proves the following result.

Lemma 2.4. *Let (G, X) be a Shimura datum. For every finite extension L of $E(G, X)$ in $\mathbb{C}$, there exists a special point $x \in X$ such that $E(x)$ is linearly disjoint from L.*

In particular, by taking a chain of (finite subfields of) $E(x)$ we can generate all extensions.

For example, for GL_2, as in our example above, this is just the statement that for any finite extension of $\mathbb{Q}$ in $\mathbb{C}$, there exists a quadratic imaginary extension $E/\mathbb{Q}$ linearly disjoint from L. In this case this is easy to see: take any prime p unramified in L, and let $E = \mathbb{Q}[\sqrt{-p}]$. In general the proof is more delicate.

We can now conclude:

Theorem 2.5. *A canonical model of* $\mathrm{Sh}_K(G,X)$*, if it exists, is unique up to unique isomorphism. If* $\mathrm{Sh}_K(G,X)$ *has a canonical model for every compact open subgroup* $K \subset G(\mathbb{A}_f)$*, then so does* $\mathrm{Sh}(G,X)$*, and it is unique up to unique isomorphism.*

Proof. The second statement follows immediately from the first, together with the result from Theorem 2.1 that the natural morphisms between the $\mathrm{Sh}_K(G,X)$ are defined over $E(G,X)$. To see the first, suppose that $M_K(G,X)$ and $M'_K(G,X)$ are canonical models of $\mathrm{Sh}_K(G,X)$, with isomorphisms $\varphi : M_K(G,X)_{\mathbb{C}} \xrightarrow{\sim} \mathrm{Sh}_K(G,X) \xleftarrow{\sim} M'_K(G,X) : \varphi'$ compatible with the Galois action. We can take the situation of Theorem 2.1 with $K = K'$ but the two canonical models on each side, with an isomorphism $\mathcal{T}(1)$ between them (namely $\varphi'^{-1} \circ \varphi$); then the theorem shows that it must be defined over $E(G,X)$. Thus $M_K(G,X)$ and $M'_K(G,X)$ are uniquely isomorphic over $E(G,X)$. □

A canonical model for $\mathrm{Sh}_K(G,X)$ also defines an action of $\mathrm{Aut}(C/E)$ on the set of connected components $\pi_0(\mathrm{Sh}_K(G,X))$. We saw that when G^{der} is simply connected this is a "zero-dimensional Shimura variety"

$$\pi_0(\mathrm{Sh}_K(G,X)) \simeq T(\mathbb{Q})\backslash Y \times T(\mathbb{A}_f)/\nu(K),$$

where $\nu : G \to T$ is the quotient by G^{der} and Y is the quotient of $T(\mathbb{R})$ by the image $T(\mathbb{R})^{\dagger}$ of $Z(\mathbb{R})$ in $T(\mathbb{R})$. For any $x \in X$, let $h = \nu \circ h_x$. The corresponding cocharacter μ_h is defined over E, and so defines a homomorphism

$$r = r(T, \mu_h) : \mathbb{A}_E^{\times} \to T(\mathbb{A}).$$

If s corresponds to σ (or its restriction to E^{ab}), write $r(s) = r(s)_f \times r(s)_\infty \in T(\mathbb{A}) \simeq T(\mathbb{R}) \times T(\mathbb{A}_f)$. Then $\sigma \in \mathrm{Aut}(\mathbb{C}/E)$ acts on $\pi_0(\mathrm{Sh}_K(G,X)) \simeq T(\mathbb{Q})\backslash Y \times T(\mathbb{A}_f)/\nu(K)$ by

$$\sigma[y,a]_K = [r(s)_\infty y, r(s)_f \cdot a]_K$$

for all $y \in Y$ and $a \in T(\mathbb{A}_f)$. This is compatible with the usual notion: π_0 of a canonical model of $\mathrm{Sh}_K(G,X)$ is the canonical model of $\mathrm{Sh}_{\nu(K)}(T,Y)$. This formula can be deduced from the usual formula for canonical models from Definition 1.3 for σ sending a special point x_0 to y; a slight improvement to Lemma 2.4 shows that such σ generate $\mathrm{Aut}(\mathbb{C}/E)$.

3. Existence of canonical models

Next, we want to outline the proof that any Shimura variety has a canonical model, which by Theorem 2.5 is unique. The strategy is as follows: we noted in the last section that the functor sending a variety V over some $k \subset \mathbb{C}$ to its base change $V_{\mathbb{C}}$ together with an action of $\mathrm{Aut}(\mathbb{C}/k)$ on $V(\mathbb{C})$ is fully faithful. Then given $\mathrm{Sh}_K(G,X)$, a complex variety, we want to be able to show that it is in the essential image of this functor; thus our first goal is to be able to describe this essential image, i.e. the theory of descent. We'll then apply this to Siegel modular varieties and then to Shimura varieties of PEL, Hodge, and abelian type, and sketch how the argument works in general.

3.1. Descent: conditions for the existence of a model

Given a variety V over $\mathbb{C}$ and an action of some $G = \mathrm{Aut}(\mathbb{C}/k)$ on $V(\mathbb{C})$, in order for this action to give descent to k it should come from an action on algebraic varieties. In particular for each $\sigma \in G$ we get another complex variety σV, given by applying σ to the coefficients of the polynomials defining V, and similarly $P \in V(\mathbb{C})$ corresponds to a point σP in $(\sigma V)(\mathbb{C})$, given by applying σ to the coordinates of P. Since we want this to arise from an algebraic variety, the morphism $P \mapsto \sigma P$ should be a regular map, and indeed a regular isomorphism since the action of σ is invertible (since G is a group). In particular we should have a regular isomorphism $\sigma V \mapsto V$ sending σP to P. To have any hope of having such things descend to k, they should respect the action of G; to get a map $\sigma V \to V$ respecting G, we compose with the action of σ on the right, to get a map $(\sigma V)(\mathbb{C}) \to V(\mathbb{C})$ sending $\sigma P \mapsto \sigma \cdot P$ which we hope will descend to an isomorphism of varieties over k. At the least, it should come from an isomorphism of varieties, i.e.

$$\sigma P \mapsto \sigma \cdot P$$

should be a regular isomorphism for every σ; this is called the regularity condition for the action of G, and is one of the conditions we will need in order for descent to work.

This is equivalent to requiring that $\sigma P \mapsto \sigma \cdot P$ be induced by a regular map $f_\sigma : \sigma V \to V$; the families of such f_σ are the descent data. In the case where V arises from a variety over k, then σ fixes V, i.e. $\sigma V = V$, and the action of σ is just given by the action on coordinates, i.e. $\sigma \cdot P = \sigma P$, so each f_σ is the identity.

Generally we say that a Galois action is continuous if it factors through a finite extension; in this case, we say that the action by $G = \mathrm{Aut}(\mathbb{C}/k)$ is continuous if there is a model V_0 of V over some intermediate extension $k \subset L \subset \mathbb{C}$ finitely generated over k such that the action of $\mathrm{Aut}(\mathbb{C}/L)$ agrees with the restriction of G to L. Certainly if V has a model over k this is true, with $L = k$; what makes it useful to assume is the following relationship between continuity and regularity.

Proposition 3.1. *A regular action of G on $V(\mathbb{C})$ is continuous if there exist points $P_1, \ldots, P_n$ in $V(\mathbb{C})$ such that the only automorphism of V fixing every P_i is the identity, and there exists a finitely generated L/k in $\mathbb{C}$ such that every $\sigma \in \mathrm{Aut}(\mathbb{C}/L)$ fixes each P_i.*

Proof. Let V_0 be any model of V over some subfield L of $\mathbb{C}$ finitely generated over k, and write $\varphi : (V_0)_\mathbb{C} \to V$ for the corresponding isomorphism. By the second assumption, for L sufficiently large we can assume that the P_i are L-points of V, and thus descend to L-points of V_0. Now, we have a map $\sigma V \to V$ sending

$$\sigma P \mapsto \sigma \cdot P,$$

which is a regular isomorphism since the action is regular; for $P_0 \in V_0(\mathbb{C})$ mapping to P under φ, precomposing with $P \mapsto \varphi^{-1}(P) = P_0 \mapsto (\sigma\varphi)(P_0)$ gives an automorphism of V, where $\sigma\varphi$ can be applied to P_0 for $\sigma \in \mathrm{Aut}(\mathbb{C}/L)$ since V_0 is defined over L and so σ fixes V_0. Since such σ fix each P_i, this whole automorphism fixes each P_i, and so by the first assumption must be the identity, i.e. $f_\sigma((\sigma\varphi)(\varphi^{-1}(P))) = P$ where $f_\sigma : \sigma P \mapsto \sigma \cdot P$ is the descent datum mentioned above. In particular, applying this with $P = \varphi(\sigma P_0)$ gives

$$\varphi(\sigma P_0) = f_\sigma((\sigma\varphi)(\sigma P_0)) = f_\sigma(\sigma(\varphi(P_0))) = \sigma \cdot \varphi(P_0)$$

and so the action of $\sigma \in \mathrm{Aut}(\mathbb{C}/L)$ induced by the model V_0 over L agrees with the action of G restricted to $\mathrm{Aut}(\mathbb{C}/L)$, i.e. the action is continuous. □

A theorem of Weil states that these two conditions, in addition to being necessary as we have seen, are also sufficient: if V is a quasiprojective variety over $\mathbb{C}$ with a regular and continuous action of $\mathrm{Aut}(\mathbb{C}/k)$, then it arises from a model over k. Combining this result with Proposition 3.1, we get the following condition.

Corollary 3.2. *A complex variety V with an action of $\mathrm{Aut}(\mathbb{C}/k)$ arises from a variety over k if V is quasiprojective, the action is regular, and there exist points $P_1, \ldots, P_n$ satisfying the conditions of Proposition 3.2.*

We next want to try to apply our criteria to see that our simplest interesting case, the Siegel modular variety, has a canonical model. First we need to say something about families of Hodge structures and abelian varieties.

3.2. Variations of integral Hodge structures

Let S be a complex manifold, and F be a local system of $\mathbb{Z}$-modules on S, i.e. a sheaf which is locally isomorphic to the constant sheaf $\mathbb{Z}^n$ for some n. Suppose that for every $s \in S$ we have a Hodge structure h_s on $F_s \otimes \mathbb{R}$. We say that F, together with these Hodge structures, is a variation of integral Hodge structures on S if for every open subset $U \subset S$ on which F is trivial $(F \otimes \mathbb{R}, (h_s))$ is a variation of Hodge structures in the usual sense, or equivalently the pullback of $(F \otimes \mathbb{R}, (h_s))$ to the universal cover of S is a variation of Hodge structures. A polarization of a variation of Hodge structures $(F, (h_s))$ is a pairing $\psi : F \times F \to \mathbb{Z}$ such that ψ_s is a polarization of (F_s, h_s) for every s.

Let V be a smooth complex algebraic variety. A family of abelian varieties over V is a regular map $f : A \to V$ of smooth varieties plus a multiplication map $A \times_V A \to A$ over V inducing the structure of an abelian variety of constant dimension on each fiber of f. In this case A is also called an abelian scheme over V. Given any abelian scheme $f : A \to V$ and local system F on A, the analogy of the first homology is the dual of the "first relative cohomology," i.e. $(R^1 f_* F)^\vee$. Whereas the first homology of an abelian variety carries the structure of a polarizable integral Hodge structure, this relative homology is now a variation of such structures.

Theorem 3.3. *Let V be a smooth variety over $\mathbb{C}$. Then $(A, f) \mapsto (R^1 f_* F)^\vee$ is an equivalence between the categories of families of abelian varieties over V and the category of polarizable integral variations of Hodge structures of type $\{(-1,0),(0,-1)\}$.*

The special case where $V = \operatorname{Spec} \mathbb{C}$ is a point recovers the equivalence between complex abelian varieties and the associated Hodge structures.

3.3. The Siegel modular variety

Let's now try to apply this theory. Let (V, ψ) be a symplectic space over $\mathbb{Q}$, with $(G, X) = (\mathrm{GSp}(\psi), X(\psi))$ the associated Shimura datum. Since in this subsection we are focused on this Shimura datum, we will drop it from the notation and write Sh_K for $\mathrm{Sh}_K(G, X)$.

First, we want to compute the reflex field. We've already stated that this should be $\mathbb{Q}$ since $G = \mathrm{GSp}(\psi)$ is already split over $\mathbb{Q}$, but we can see this more directly: any symplectic bases for $V(\mathbb{C})$ determines a pair of complementary Lagrangian subspaces $L, L' \subset V(\mathbb{C})$ such that $L \oplus L' = V(\mathbb{C})$. Given such a pair, we can write down an action of $\mathbb{G}_\mathrm{m}$ on $\mathrm{GL}(V)$ given on complex points by acting on L by multiplication by z and on L' by the identity. As we vary the pair of Lagrangians, we get all possible conjugacy classes of cocharacters. Since V (defined over $\mathbb{Q}$) is already symplectic, we can find a symplectic basis over $\mathbb{Q}$ extending to one over $\mathbb{C}$, so the same is true of our Lagrangians and hence of the cocharacters: we can find a representative in $c(X)$ defined over $\mathbb{Q}$, and so $E(G, X) = \mathbb{Q}$.

Next, we want to study the special points. Recall that for each compact open subgroup $K \subset G(\mathbb{A}_f)$, $\mathrm{Sh}_K(\mathbb{C})$ classifies isomorphism classes of triples $(A, s, \eta K)$ where A is an abelian variety over $\mathbb{C}$, s is an alternating form on $H_1(A, \mathbb{Q})$ which is a polarization up to a sign, and $\eta : V(\mathbb{A}_f) \to V_f(A)$ is an isomorphism sending ψ to a multiple of s, defined only up to the K-orbit. We want to know which triples $(A, s, \eta K)$ correspond to special points, i.e. map to $[x, a] \in \mathrm{Sh}_K(\mathbb{C})$ with x special.

The answer is given by a slight extension of our notion of CM abelian varieties. Define a CM algebra to be a finite product of CM fields. Then we say that an abelian variety A over $\mathbb{C}$ is CM if there exists some CM algebra E with a homomorphism $E \to \mathrm{End}^0(A)$ such that $H_1(A, \mathbb{Q})$ is a free E-module of rank 1. This is equivalent to being a product of abelian varieties of some CM-type.

For any abelian variety A over $\mathbb{C}$, there is a homomorphism $h_A : \mathbb{C}^\times \to \mathrm{GL}(H_1(A, \mathbb{R}))$ giving the natural complex structure on $H_1(A, \mathbb{R})$ (since if $A(\mathbb{C}) \simeq \mathbb{C}^g/\Lambda$ then $H_1(A, \mathbb{R}) = \Lambda \otimes \mathbb{R} \simeq \mathbb{C}^g$).

Proposition 3.4. *An abelian variety A over $\mathbb{C}$ is CM if and only if there exists a torus $T \subset \mathrm{GL}(H_1(A, \mathbb{Q}))$ such that $h_A(\mathbb{C}^\times) \subset T(\mathbb{R})$.*

By the definition of special points and the fact that if (A, s, η_K) corresponds to $[x, a]$ then there is an isomorphism $H_1(A, \mathbb{Q}) \to V$ carrying h_A to h_x, an immediate corollary is that $(A, s, \eta K)$ maps to $[x, a]$ with x special if and only if A is CM.

Proof. Up to isogeny, every abelian variety is a product of simple abelian varieties, and the statements only depend on the isogeny class of A, so we may assume A is simple.

If A is a CM simple abelian variety, the corresponding CM algebra E must be a field, and we've seen that a CM field corresponding to an abelian variety will have degree $2 \dim A$ over $\mathbb{Q}$. On the other hand $\mathrm{End}^0(A)$ is a division algebra over $\mathbb{Q}$ since A is simple and acts simply on $H_1(A, \mathbb{Q})$, which is of dimension $2 \dim A$, so $E = \mathrm{End}^0(A)$ is a CM field of degree $2 \dim A$ over $\mathbb{Q}$. On the other hand if $\mathrm{End}^0(A)$ is such a field then by definition A is of CM-type.

Suppose that A is CM, so that by the above $\mathrm{End}^0(A)$ is a CM field of degree $2 \dim A$ over $\mathbb{Q}$, so $H_1(\mathbb{Q})$ is a one-dimensional E-vector space. Tensoring with $\mathbb{R}$, we get an action of $E \otimes \mathbb{R}$ on $H_1(A, \mathbb{R})$ preserving the Hodge structure, and so $h_A(\mathbb{C}^\times)$ commutes with the action of $E \otimes \mathbb{R}$, so that

$$h_A(\mathbb{C}^\times) \subset (E \otimes \mathbb{R})^\times = (\mathrm{Res}_{E/\mathbb{Q}} \mathbb{G}_\mathrm{m})(\mathbb{R}).$$

This proves one direction of the proposition.

Conversely, suppose that $h_A(\mathbb{C}^\times) \subset T(\mathbb{R})$ for some torus $T \subset \mathrm{GL}(H_1(A,\mathbb{Q}))$. Taking complex points, this is equivalent to $\mu_A(\mathbb{C}^\times) \subset T(\mathbb{C})$, where $\mu_A = h_A(\mathbb{C})(-,1)$ as usual. For any abelian variety, $\mathrm{End}^0(A)$ is the subalgebra of $\mathrm{End}(H_1(A,\mathbb{Q}))$ preserving the Hodge structure, or equivalently (tensoring with $\mathbb{C}$) commuting with $\mu_A(\mathbb{G}_\mathrm{m})$ in $\mathrm{GL}(H_1(A,\mathbb{C}))$. Since we assume $\mu_A(\mathbb{C}^\times) \subset T(\mathbb{C})$ for some torus $T \subset \mathrm{GL}(H_1(A,\mathbb{Q}))$, $\mathrm{End}^0(A) \otimes \mathbb{C}$ contains the subalgebra of $\mathrm{End}(H_1(A,\mathbb{C}))$ commuting with the action of $T_\mathbb{C}$. We can decompose $H_1(A,\mathbb{C})$ over the characters of T; the endomorphisms fixing every component give an étale $\mathbb{C}$-algebra of degree $2\dim A$ in $\mathrm{End}(H_1(A,\mathbb{C}))$, and so $\mathrm{End}^0(A)$ contains an étale $\mathbb{Q}$-algebra of degree $2\dim A$, which by the argument above must be all of $\mathrm{End}^0(A)$. By assuming A is simple we can reduce to the case where $E = \mathrm{End}^0(A)$ is a field.

It remains to see that this is a CM-field. Choose a polarization of A giving a Riemann form ψ on $H_1(A,\mathbb{Q})$. The Rosati involution $e \mapsto e^*$ is determined by

$$\psi(x, ey) = \psi(e^*x, y)$$

for $e \in E$, $x, y \in H_1(A,\mathbb{Q})$, so since $\psi(x,y) = \psi(h(i)x, h(i)y)$ we get

$$h(i)^* = h(i)^{-1} = -h(i).$$

Therefore the Rosati involution is nontrivial on E, and fixes an index 2 subfield F; we can find some $\alpha \in F^\times$ such that $E = F[\sqrt{\alpha}]$ and $\sqrt{\alpha}^* = -\sqrt{\alpha}$, uniquely determined up to multiplication by a square in F, i.e. E is totally imaginary over F. We need only to show that F is totally real. By identifying $H_1(A,\mathbb{Q})$ with E by choosing some basis vector over E, we can write

$$\psi(x,y) = \mathrm{Tr}_{E/\mathbb{Q}}\, \alpha x y^*$$

for $x, y \in E$, and by the positivity of ψ we have

$$\psi_\mathbb{R}(x,x) = \mathrm{Tr}_{E\otimes\mathbb{R}/\mathbb{R}}(\alpha/h(i)\cdot x^2) > 0$$

for $0 \neq x \in F \otimes \mathbb{R}$, i.e. F is totally real, and the image of α in $\mathbb{R}$ for any embedding $F \hookrightarrow \mathbb{R}$ must be negative since otherwise $E \otimes_F \mathbb{R} = \mathbb{R} \times \mathbb{R}$ along this embedding with $(r_1, r_2)^* = (r_2, r_1)$, which makes the positivity condition impossible, so $*$ is complex conjugation for E/F, i.e. E is a CM field. By the equivalence above, this completes the proof. □

Now that we understand special points, we can give a criterion for when a model is canonical. First, we need to understand how $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ acts on the set of tuples $(A, s, \eta K) \in \mathcal{M}_K$. We know how σ acts on A and ηK, but s is something of a mystery. However, we can understand it more explicitly: s is a Hodge tensor in $H^2(A,\mathbb{Q})$, and so is represented (up to coefficients in $\mathbb{Q}$) by some codimension 1 cycle, i.e. a divisor D in A, i.e. $s = r[D]$ for $r \in \mathbb{Q}^\times$ and a divisor D. Then we can define $\sigma s = r[\sigma D]$; the necessary conditions on s are preserved by this action. Then we get an action of $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ on $\mathcal{M}_K$ by $\sigma(A, s, \eta K) = (\sigma A, \sigma s, \sigma\eta K)$.

Proposition 3.5. *Suppose that* Sh_K *has a model over* $\mathbb{Q}$ *for which the map* $\mathcal{M}_K \to M_K(\mathbb{C})$ *commutes with the actions of* $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$*. Then* M_K *is canonical.*

Proof. By Lemma 2.2, it suffices to prove that we have the correct action on points $[x, a]$ with x special, since the action of each σ is continuous and such points are dense in $\mathrm{Sh}_K(\mathbb{C})$. For

$[x, a]$ corresponding to an abelian variety A with complex multiplication by a CM field, the correct formula for the action of any σ is given by the main theorem of complex multiplication. For more general CM abelian varieties, the same formula follows in the same way with a bit more difficulty, after decomposing into factors. □

We are now ready to outline the proof of the existence of a canonical model for Siegel modular varieties using Corollary 3.2.

The action of $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ on $\mathcal{M}_K$ induces such an action on $\mathrm{Sh}_K(\mathbb{C})$. Any Shimura variety is quasiprojective, so we only need to check that this action is regular and continuous in order to apply Corollary 3.2 to conclude that it has a model over $\mathbb{Q}$; if we can find such a model, Proposition 3.5 shows that it is canonical.

Let's first check continuity. For any $x \in X$, the proof of Lemma 2.2 showed that the set of points $[x, a]$ for $a \in G(\mathbb{A}_f)$ is dense in $\mathrm{Sh}_K(\mathbb{C})$, and so any automorphism of Sh_K fixing all such points must be the identity. In fact Sh_K has only finitely many automorphisms, so we can pick a finite subset $[x, a_1], \dots, [x, a_n]$ such that any automorphism fixing all of them must be the identity: if f is an automorphism other than the identity, pick some $[x, a_1]$ not fixed by f, which we can assume is sent to another point of the form $[x, a_2]$; this is also not a fixed point of f, since if we apply f sufficiently many times we recover $[x, a_1]$, so we can continue applying f until we recover $[x, a_1]$; we can do the same thing for each f and take the union. We can restrict x to be special by density, and so by the main theorem of complex multiplication for $\sigma \in \mathrm{Aut}(\mathbb{C}/L)$ for some fixed finite extension of $E(x)$ we have $\sigma \cdot [x, a_i] = [x, a_i]$, so the action is continuous.

Regularity is a bit more involved. We want to show that the map $f_\sigma : \sigma\,\mathrm{Sh}_K(\mathbb{C}) \to \mathrm{Sh}_K(\mathbb{C})$ sending $\sigma P \mapsto \sigma \cdot P$ is regular. We can assume that K is small, since if $K' \supset K$ then $\mathrm{Sh}_{K'}$ is a quotient of Sh_K and so the claim for Sh_K implies the claim for $\mathrm{Sh}_{K'}$.

We want to work one connected component at a time. The largest commutative quotient $\nu : G \to T$ of $G = \mathrm{GSp}(\psi)$ is just $\mathbb{G}_\mathrm{m}$, and so the connected components of Sh_K are given by the double quotient

$$\mathbb{Q}_{>0} \backslash \mathbb{A}_f^\times / \nu(K).$$

Let ϵ be an element of this quotient, with corresponding connected component Sh_K^ϵ. This is some quotient of the form $\Gamma_\epsilon \backslash X^+$ for a connected component X^+ of X and $\Gamma_\epsilon = G(\mathbb{Q}) \cap K_\epsilon$ for some conjugate K_ϵ of K.

If $a : H_1(A, \mathbb{Q}) \to V$ is an isomorphism sending s to a multiple of ψ, as must exist for each tuple $(A, s, \eta K) \in \mathcal{M}_K$, then we can write the image of $(A, s, \eta K)$ in $\mathbb{Q}_{>0} \backslash \mathbb{A}_f^\times / \nu(K)$ as $[\nu(a \circ \eta)]$, where we view $a \circ \eta : V(\mathbb{A}_f) \to V(\mathbb{A}_f)$ as an element of $G(\mathbb{A}_f) = \mathrm{Aut}_\psi(V)$. Let $\mathcal{M}_K^\epsilon$ be the fiber of the moduli problem over ϵ, i.e. the tuples $(A, s, \eta K)$ with $\nu(a \circ \eta) \in \epsilon$. When we let $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ act on $\mathbb{Q}_{>0} \backslash \mathbb{A}_f^\times / \nu(K)$ via the cyclotomic character

$$\sigma[\alpha] = [\chi(\sigma)\alpha]$$

for $\alpha \in \mathbb{A}_f^\times$ and $\chi(\sigma) \in \widehat{\mathbb{Z}}^\times$ determined by sending roots of unity to $\zeta^{\chi(\sigma)} = \sigma\zeta$, the map $\mathcal{M}_K \to \mathbb{Q}_{>0} \backslash \mathbb{A}_f^\times / \nu(K)$ is $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$-equivariant, so it suffices to prove the claim on each fiber.

Let $\sigma(\Gamma_\epsilon \backslash X^+)$ be the algebraic variety given by change of base field by σ as usual; let U be the universal cover of $\sigma(\Gamma_\epsilon \backslash X^+)$. For any Γ_ϵ-stable lattice Λ in V, since X^+ is the universal cover of $\Gamma_\epsilon\backslash X^+$ the Γ_ϵ-action on Λ gives a local system F of $\mathbb{Z}$-modules on $\Gamma_\epsilon\backslash X^+$ since π_1-modules are equivalent to local systems; this is a polarized integral variation of Hodge structures. By Theorem 3.3, this is equivalent to a family of abelian varieties $f : A \to \Gamma_\epsilon\backslash X^+$. This is a morphism of complex algebraic varieties and so we can apply σ to get a family of abelian varieties $\sigma f : \sigma A \to \sigma(\Gamma_\epsilon\backslash X^+)$, or equivalently a polarized integral variation of Hodge structures $(R^1(\sigma f)_*\mathbb{Z})^\vee$ on $\Gamma_\epsilon\backslash X^+$, which pulls back to a polarized integral variation of Hodge structures on U. Tensoring with $\mathbb{Q}$ to eliminate the dependence on the lattice Λ, we get a variation of polarized rational Hodge structures on U, i.e. a rational local system $\tilde{F}$ on U together with a complex structure h_u on $\tilde{F}_u$ for every $u \in U$. The local system is just the one coming from $\Lambda \otimes \mathbb{Q}$, i.e. the constant local system with value V; if we keep track of orientations correctly, each h_u gives a Hodge structure on V which is positive for ψ, i.e. a point of X^+. This gives a map $U \to X$, $u \mapsto h_u$, making the diagram

$$\begin{array}{ccc} U & \xrightarrow{u\mapsto h_u} & X^+ \\ \downarrow & & \downarrow \\ \sigma(\Gamma_\epsilon\backslash X^+) & \xrightarrow{f_\sigma} & \Gamma_\epsilon\backslash X^+ \end{array}$$

commute. The map $u \mapsto h_u$ is holomorphic, so so is f_σ; by Borel's theorem it follows that it is regular.

We've shown that Sh_K satisfies the conditions of Corollary 3.2 over $\mathbb{Q}$, so it has a model over $\mathbb{Q}$; Proposition 3.5 shows that this is canonical. Therefore we've proven that Sh_K, and generally $\mathrm{Sh}(\mathrm{GSp}(\psi), X(\psi))$, has a canonical model over $\mathbb{Q}$.

3.4. Simple PEL Shimura varieties of type (A) or (C)

This case is similar to the case of the Siegel modular variety, though somewhat more complicated: $\mathrm{Sh}_K(G, X)$ classifies tuples $(A, i, s, \eta K)$ satisfying certain conditions; we can verify that σ fixing the reflex field also fix such tuples, find that special points correspond to CM abelian varieties, apply the main theorem of complex multiplication to compute the action on these points and verify that they satisfy the necessary conditions, and use the Shimura-Taniyama theorem to see that a model with the correct action is canonical.

3.5. Shimura varieties of Hodge type

In this case, $\mathrm{Sh}_K(G, X)(\mathbb{C})$ classifies isomorphism classes of tuples $(A, s_0, s_1, \ldots, s_n, \eta K)$, where the s_i are Hodge tensors. We can apply a similar proof to the case of the Siegel modular variety; first, though, we need to define an action of $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ on such data, and in particular on the s_i. In the Siegel case, we only had one Hodge tensor s, which was an element of $H^2(A, \mathbb{Q})$ and so could be identified, up to a rational scalar, with a divisor, on which $\mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ naturally acts. For general Hodge tensors, though, to make a similar identification of Hodge tensors with cycles we would need to know the Hodge conjecture, which is far from known even for abelian varieties. The fundamental difficulty is that in general there is no natural map $H^n(A, \mathbb{Q}) \to H^n(\sigma A, \mathbb{Q})$ which we could take to be the action of σ on Hodge tensors.

However, if we take coefficients in $\mathbb{A}_f$ the situation is better: if $A \simeq \mathbb{C}^g/\Lambda$, then we have identifications

$$H^n(A, \mathbb{A}_f) \simeq \mathrm{Hom}(\wedge^n \Lambda, \mathbb{A}_f) \simeq \mathrm{Hom}(\wedge^n(\Lambda \otimes \mathbb{A}_f), \mathbb{A}_f) \simeq \mathrm{Hom}(\Lambda^n V_f(A), \mathbb{A}_f),$$

and σ acts on $V_f(A)$ and thus on everything. A (difficult) theorem of Deligne states that this descends to the situation over $\mathbb{Q}$ for Hodge tensors: given a Hodge tensor s on an abelian variety A over $\mathbb{C}$, if $s_{\mathbb{A}_f}$ is the image in the $\mathbb{A}_f$-cohomology, for any $\sigma \in \mathrm{Aut}(\mathbb{C}/\mathbb{Q})$ there is a unique Hodge tensor σs on σA such that $(\sigma s)_{\mathbb{A}_f} = \sigma(s_{\mathbb{A}_f})$. At this point a similar argument to that for the Siegel modular variety shows that any Shimura variety of Hodge type has a canonical model.

Alternatively, we can use the following result, which subsumes the case of PEL Shimura varieties as well:

Proposition 3.6. *Let $(G, X) \hookrightarrow (G', X')$ be an inclusion of Shimura data. If $\mathrm{Sh}(G', X')$ has a canonical model, then so does $\mathrm{Sh}(G, X)$.*

Proof. An inclusion of Shimura data induces a closed immersion of Shimura varieties. The existence of a canonical model M' for $\mathrm{Sh}(G', X')$ over the reflex field E implies that the diagram

$$\begin{array}{ccc} \mathrm{Sh}(G', X') & \longrightarrow & M' \\ \downarrow & & \downarrow \\ \mathrm{Spec}\,\mathbb{C} & \longrightarrow & \mathrm{Spec}\,E \end{array}$$

is Cartesian, so the immersion $\mathrm{Sh}(G, X) \hookrightarrow \mathrm{Sh}(G', X')$ has image a closed subscheme M of M' extending the diagram to

$$\begin{array}{ccc} \mathrm{Sh}(G, X) & \longrightarrow & M \\ \big\downarrow & & \big\downarrow \\ \mathrm{Sh}(G', X') & \longrightarrow & M' \\ \downarrow & & \downarrow \\ \mathrm{Spec}\,\mathbb{C} & \longrightarrow & \mathrm{Spec}\,E \end{array} .$$

The action induced by M commutes with that from M', so it is easy to check that (possibly up to a finite quotient) the Galois action makes M canonical (possibly over some finite extension of E, the reflex field of $\mathrm{Sh}(G', X')$). □

Since Shimura data of Hodge type are those that inject into Siegel modular data (and include PEL Shimura data), it follows that Shimura data of Hodge (and therefore PEL) type have canonical models. This does give less information than repeating the above proofs, though, in that it does not give as detailed a description over arbitrary fields.

3.6. Shimura varieties of abelian type

Recall that by definition a Shimura datum (G, X) of abelian type is one such that (G^{der}, X^+) is of abelian type, which itself means that it is isogenous to a product of Shimura data injecting into Siegel modular data, i.e. it is a (connected) Hodge Shimura datum. To use this definition, we'd like to be able to reduce to the connected setting, so we'd like a notion of canonical models for connected Shimura varieties. Deligne defines such a notion, whose precise definition we omit; the key result is that $\mathrm{Sh}(G, X)$ has a canonical model (in our usual sense) if and only if $\mathrm{Sh}^\circ(G^{\mathrm{der}}, X^+)$ has a canonical model (as a connected Shimura variety, in Deligne's sense).

In fact, Deligne also shows that these canonical models of connected Shimura data behave well with respect to isogenies and products:

Proposition 3.7. *Let* $(G_1, X_1) \to (G_2, X_2)$ *be an isogeny of connected Shimura data. If* $\mathrm{Sh}^\circ(G_1, X_1)$ *has a canonical model, then so does* $\mathrm{Sh}^\circ(G_2, X_2)$*. Further, if* $(G, X) = \prod_i (G_i, X_i) = (\prod_i G_i, \prod_i X_i)$ *and each* $\mathrm{Sh}^\circ(G_i, X_i)$ *has a canonical model* $M^\circ(G_i, X_i)$*, then* $\prod_i M^\circ(G_i, X_i)$ *is a canonical model for* $\mathrm{Sh}^\circ(G_i, X_i)$*.*

The claim then follows: by Deligne's results, to show that Shimura varieties of abelian type have canonical models it suffices to show the claim for Hodge Shimura varieties, which we've seen follows from the Siegel modular variety case and Proposition 3.6.

Note that this tells us only about existence, and doesn't give much in the way of a description. When $\mathrm{Sh}_K(G, X)(\mathbb{C})$ has a moduli interpretation, it's possible (as in previous cases) to extract a description of the canonical model.

3.7. General Shimura varieties

In general, we do not have access to moduli interpretations, and the situation is much more difficult. The key idea is to use Deligne's techniques to deduce the result from sufficient results for certain subvarieties, and in particular to focus on Shimura subvarieties of type A_1, i.e. coming from quaternion algebras over totally real fields F.

By the techniques of Deligne we can reduce to connected Shimura data (G, X^+) with G simple and simply connected, and therefore of the form $\mathrm{Res}_{F/\mathbb{Q}} H$ for some geometrically simple H over F. If we replace F by some sufficiently large extension F' such that $H_{F'}$ splits over a CM field over F', we get a group $G' = \mathrm{Res}_{F'/\mathbb{Q}} H_{F'}$ whose Shimura variety has Shimura subvarieties of type A_1, unlike the original (G, X^+). These A_1-Shimura subvarieties are certainly Hodge so the claim can be proven for them, and it follows for (G', X^+) through difficult arguments; then $\mathrm{Sh}^\circ(G, X^+)$ is a Shimura subvariety of $\mathrm{Sh}^\circ(G', X^+)$ and so also has a canonical model.

In fact, what is proven in this way is not Shimura's conjecture on the existence of canonical models, but a slightly stronger statement due to Langlands, namely that there exist isomorphisms $f_\sigma : \sigma\,\mathrm{Sh}(G, X) \to \mathrm{Sh}(G^\sigma, X^\sigma)$ for new Shimura data (G^σ, X^σ) defined in terms of (G, X) and σ, satisfying certain conditions. For σ fixing the reflex field, these f_σ define descent data and thus give rise to a canonical model.

This method is independent of our previous (moduli problem-based) methods for proving the existence of a canonical model and so does not depend on the moduli interpretations of

Shimura varieties, except for these A_1-Shimura varieties, whose moduli interpretation can be understood explicitly (as we've looked at).

References

[1] James S Milne. Introduction to Shimura varieties. *Harmonic analysis, the trace formula, and Shimura varieties*, 4:265–378, 2005.