Moduli interpretation of Shimura varieties*

Avi Zeff

Our next goal is to interpret Shimura varieties as moduli spaces classifying certain data, just as e.g. modular curves classify (generalized) elliptic curves with some level structure. We'll start by looking at certain special kinds of Shimura varieties which have a good moduli interpretation, and then see how far we can push this.

1. Shimura varieties of PEL type

We first look at Shimura varieties arising from symplectic actions of $\mathbb{Q}$-algebras with involution. In particular, suppose that B is a semisimple $\mathbb{Q}$-algebra with an involution $* : b \mapsto b^*$, and (V, ψ) is a symplectic B-module, i.e. a B-module V with a symplectic form ψ such that $\psi(bu, v) = \psi(u, b^*v)$ for every $b \in B$. This gives rise to a Shimura datum as follows.

Write $\mathrm{GL}_B(V)$ for the group of automorphisms of V as a B-module. We can define an algebraic subgroup G of $\mathrm{GL}_B(V)$ such that $G(\mathbb{Q})$ is the subgroup consisting of elements g such that $\psi(gx, gy) = \mu(g)\psi(x, y)$ for every $x, y \in V$, where $\mu(g)$ is some nonzero rational number depending on g. For the other part of the Shimura datum, we need a homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ satisfying suitable properties, or more precisely a conjugacy class of such homomorphisms; for now assume we have such a homomorphism h, and let X be its $G(\mathbb{R})$-conjugacy class. Then, once we make these conditions more precise and add a few on G, the pair (G, X) is a Shimura datum, called a PEL Shimura datum, and gives rise to a PEL Shimura variety, or a Shimura variety of PEL type.

The letters PEL stand for polarization, endomorphism, and level structure. The reason for this is that PEL Shimura varieties are moduli spaces classifying abelian varieties with polarization, a (class of) endomorphism(s), and some level structure. For motivation, we state the following theorem, even though we haven't yet rigorously defined a PEL Shimura datum.

Theorem 1.1. *Let (G, X) be a PEL Shimura datum, and let $K \subset G(\mathbb{A}_f)$ be a compact open subgroup. Then $\mathrm{Sh}_K(G, X)$ classifies abelian varieties with polarization, an endomorphism, and level structure: in particular on complex points $\mathrm{Sh}_K(G, X)(\mathbb{C})$ classifies isomorphism classes of tuples $(A, i, s, \eta K)$, where*

- *A is a complex abelian variety,*
- *s is a polarization, up to sign, of the Hodge structure $H_1(A, \mathbb{Q})$,*
- *i is a homomorphism $B \to \mathrm{End}^0(A) = \mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$, and*
- *ηK is a K-orbit of $B \otimes \mathbb{A}_f$-linear isomorphisms $\eta : V(\mathbb{A}_f) \to V_f(A)$ sending ψ to an $\mathbb{A}_f^{\times}$-multiple of s,*

such that there exists a B-linear isomorphism $a : H_1(A, \mathbb{Q}) \to V$ sending s to a $\mathbb{Q}^{\times}$-multiple of ψ, such that the cocharacter $\mathbb{S} \ni r \mapsto a \circ h_A(r) \circ a^{-1} \in G$ is in X.

*These notes are based on chapters 6-9 of [1].

Here $V(\mathbb{A}_f) = V \otimes_{\mathbb{Q}} \mathbb{A}_f$, $V_f(A)$ is the Tate module $\mathbb{Q} \otimes \varprojlim_N A[N]$ of A, and h_A is the map given on real points by the map $\mathbb{C}^\times \to \mathrm{GL}(H_1(A, \mathbb{R}))$ associated to the canonical complex structure on $H_1(A, \mathbb{R})$ (since we can write $A(\mathbb{C})$ as $\mathbb{C}^g/\Lambda$, $H_1(A, \mathbb{R}) \simeq H_1(A, \mathbb{Z}) \otimes \mathbb{R} \simeq \Lambda \otimes \mathbb{R} \simeq \mathbb{C}^g$).

This is exactly the sort of moduli structure we hope for. Consider the simple case $B = \mathbb{Q}$, $* = \mathrm{id}$, $V = \mathbb{Q}^2$ with the symplectic form given by the matrix $J = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$. Then $\mathrm{GL}_B(V) = G(\mathbb{Q}) = \mathrm{GL}_2(\mathbb{Q})$ by an explicit computation, and the resulting Shimura variety is just a modular curve (of some level K). The existence of an isomorphism $V(\mathbb{A}_f) \to V_f(A)$ implies that $V_f(A)$ must be rank 2 over $\mathbb{A}_f$, so A must be an elliptic curve, the data of i is trivial since $\mathrm{End}^0(A)$ is already a $\mathbb{Q}$-algebra in a canonical way, and there is a unique polarization of the Hodge structure on $H_1(A, \mathbb{Q})$ for an elliptic curve A so the data of s is also trivial. Thus Theorem 1.1 in this case just gives the modular description of modular curves.

To say more, we need to briefly discuss semisimple $\mathbb{Q}$-algebras with involution. First, we classify semisimple k-algebras with involution for an algebraically closed field.

Proposition 1.2. *Let $(B, *)$ be a semisimple k-algebra with involution for an algebraically closed field k. Then $(B, *)$ is isomorphic to a product of k-algebras with involution of the following types (for some integer n):*

(A) $M_n(k) \times M_n(k)$ *with* $(a, b)^* = (b^\top, a^\top)$*;*

(C) $M_n(k)$ *with* $b^* = b^\top$ *(orthogonal type);*

(BD) $M_{2n}(k)$ *with* $b^* = Jb^\top J^{-1}$*, where* $J = \begin{pmatrix} & -I_n \\ I_n & \end{pmatrix}$ *(symplectic type).*

Proof. Since B is semisimple, we can decompose it as a product $B = B_1 \times \cdots B_r$ of simple k-algebras which are the minimal two-sided ideals of B. Applying $*$, we get $B = B^* = B_1^* \times \cdots B_r^\times$, which by uniqueness just permutes the B_i; so the simple k-algebras with involution are either simple k-algebras or pairs of k-algebras interchanged by $*$.

If B is a simple k-algebra, since k is algebraically closed it must be of the form $M_n(k)$ for some n, and the theorem of Skolem and Noether says that the involution $*$ must be of the form $b^* = ub^\top u^{-1}$ for some $u \in B = M_n(k)$. In particular $b = (b^*)^* = u(ub^\top u^{-1})^\top u^{-1} = u(u^{-1})^T bu^\top u^{-1} = (u^\top u^{-1})^{-1} bu^\top u^{-1}$ for every b, so $u^\top u^{-1}$ is in the center of $M_n(k)$, which is just k. Call this c, so $u^\top u^{-1} = c$ and so $u^\top = cu$. Since $u = (u^\top)^\top = (cu)^\top = cu^\top = c^2 u$ and u is invertible, we must have $c = \pm 1$, i.e. $u^\top = \pm u$, so either u is symmetric or skew-symmetric. By choosing a suitable basis we can assume u is the identity if it is symmetric or J if skew-symmetric, and thus of either type (C) or (BD) respectively.

On the other hand, if B is the product of two simple k-algebras exchanged by $*$, then $B \simeq M_m(k) \times M_n(k)$ for some m and n, and $*$ gives an isomorphism between the two factors and so $m = n$. We can write $*$ as the composition of some involution $\diamond$ on $M_n(k)$ with the operation of switching between the two factors, i.e. $(a, b)^* = (a^\diamond, b^\diamond)$. Since we can choose the basis of each factor independently, we don't have to worry about the symplectic case, and so we can assume $\diamond$ is just transposition, i.e. this is of type (A). □

For a simple $\mathbb{Q}$-algebra with involution $(B,*)$, we say that it is of type (A), (C), or (BD) if its base change to $\overline{\mathbb{Q}}$ is, so that semisimple $\mathbb{Q}$-algebras with involution again decompose into these types. In particular, the associated algebraic group G is then (at least over a sufficiently large field) a unitary group, i.e. type A, if B is of type (A), a symplectic group if B is of type (C), and an orthogonal group if B is of type (BD). (Note that these last two are exchanged from the obvious guess from the proposition; this is essentially because the form ψ is already symplectic.)

Let k be either $\mathbb{Q}$ or $\mathbb{R}$. We say that a semisimple k-algebra with involution $(B,*)$ is positive if $\mathrm{Tr}_{B/k}(b^*b)$ is positive for every nonzero $b \in B$.

Proposition 1.3. *Let $(B,*)$ be a semisimple $\mathbb{R}$-algebra with positive involution, and let (V,ψ) be a symplectic $(B,*)$-module. Assume that $(B,*)$ has all factors of type* (A) *or* (C), *and let C be the centralizer of B in $\mathrm{End}_{\mathbb{R}}(V)$. Then there exists a homomorphism of $\mathbb{R}$-algebras $h : \mathbb{C} \to C$ such that $h(\bar{z}) = h(z)^*$ and the pairing $(u,v) \mapsto \psi(u, h(i)v)$ is symmetric and positive-definite.*

Proof sketch. The only nontrivial thing is to give a suitable element $J = h(i)$ of C, i.e. a complex structure satisfying the relations $\psi(Ju, Jv) = \psi(u,v)$, $\psi(v, Jv) > 0$ for all nonzero v. It suffices to prove the claim for the simple factors, which are either of type (A) or (C); in either case B is incarnated as either an endomorphism ring of some vector space W with a positive-definite symmetric form or the product of two such endomorphism rings, in either case with the involution given by the adjoint endomorphism. Thus in either case we can explicitly write down a basis and find a suitable complex structure. □

Suppose that we have a semisimple $\mathbb{R}$-algebra with positive involution $(B,*)$ and a symplectic $(B,*)$-module (V,ψ) and corresponding algebraic group G. Let $h : \mathbb{S} \to G_{\mathbb{R}}$ be a homomorphism, which induces a Hodge structure on V, notated (V,h), and a pairing given by $(u,v) \mapsto \psi(u, h(i)v)$; suppose that this Hodge structure is of type $\{(-1,0),(0,-1)\}$, and that the pairing is symmetric and positive-definite. Then if $J = h(i)$ we have a canonical isomorphism of complex vector spaces $(V,J) \simeq V^{-1,0}$, compatible with the actions of B; let $t(b)$ be the trace of $b \in B$ acting on either space, i.e. $t(b) = \mathrm{Tr}_{\mathbb{C}}(b|(V,J))$. Note that t depends on h, through $J = h(i)$.

Let G' be the subgroup of G satisfying the additional requirement that $\mu(g) = \det g = 1$. Let F be the center of B, so that $[B : F] = n^2$ for some n; then V is also an F-vector space, and $\dim_F V$ is divisible by n. Write $m = \frac{1}{n}\dim_F V$; in the case (BD), this is the dimension of the space on which the orthogonal group corresponding to B acts. We will only be interested in the case where m is even, and denote this by (D).

Proposition 1.4. *With the notation above,*

(a) the map $t : B \to \mathbb{C}$ uniquely determines the map $h : \mathbb{S} \to G_{\mathbb{R}}$ up to $G'(\mathbb{R})$-conjugacy;

(b) if $(B,*)$ is of type (A)*, then the isomorphism class of (V,ψ) is uniquely determined by t; if it is of type* (C) *or* (D)*, it is uniquely determined by $\dim_k V$;*

(c) the centralizer of h in $G(\mathbb{R})$ is connected.

Proof sketch. We can similarly go through type by type and check that if we view B as a suitable endomorphism ring, the trace reduces to the trace on the corresponding space, which is simple and so has a unique form of the appropriate type. □

Proposition 1.5. *Assume that $(B, *)$ is simple of type (A) or (C), and let (V, ψ) be a symplectic $(B, *)$-module. Then there exists a homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ such that (V, h) has type $\{(-1, 0), (0, -1)\}$ and $2\pi i\psi$ is a polarization of (V, h), and h is unique up to $G(\mathbb{R})$-conjugation.*

Proof. Tensoring with $\mathbb{R}$ decomposes V as a direct sum of real vector spaces over real places of some number field over $\mathbb{Q}$ contained in B; thus it suffices to find a suitable component on each real factor. But this is done in Proposition 1.3. We can then observe that the conditions on h from Proposition 1.3 give us the desired ones for (V, h) and ψ here; finally for each type we can construct the trace on $V_{\mathbb{R}}$ independently of h via this decomposition, which by Proposition 1.4 determines h up to unique isomorphism. □

More generally, suppose we have $(B, *)$ and (V, ψ) such that the resulting algebraic group G is connected, and that there exists a homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ such that (V, h) has type $\{(-1, 0), (0, -1)\}$ and $(u, v) \mapsto \psi(u, h(i)v)$ is symmetric and positive-definite. Then we write X for the $G(\mathbb{R})$-conjugacy class, and call (G, X) a PEL Shimura datum. Thus Proposition 1.5 tells us that if $(B, *)$ is simple of type (A) or (C), then the data $(B, *, V, \psi)$ uniquely gives rise to a Shimura datum.

This makes Theorem 1.1 precise. However we will defer its proof to section 2 for a more general discussion of moduli structure.

One thing we have not shown yet is that the data (G, X) we have called PEL Shimura data are actually Shimura data at all. To see this, consider the following example, generalizing the example of modular curves.

Let $B = \mathbb{Q}$, with the trivial involution, and $V = \mathbb{Q}^{2n}$, with ψ given by the matrix J, or more generally any symplectic form. Then $G = \mathrm{GSp}_{2n}$, or $\mathrm{GSp}(\psi)$, of type C, so by Proposition 1.5 there is a unique conjugacy class $X(\psi)$ of homomorphisms $\mathbb{S} \to \mathrm{GSp}(\psi)_{/\mathbb{R}}$, corresponding to the set of complex structures J on $V(\mathbb{R})$ such that $\psi(Ju, Jv) = \psi(u, v)$ for $u, v, \in V$. We can check that $(G, X(\psi))$ satisfies the conditions to be a Shimura datum. First, we can decompose $V(\mathbb{C})$ as $V^+ \oplus V^-$, where $V^+ = V^{-1,0}$ and $V^- = V^{0,-1}$, so that $h(z)$ acts on V^+ and V^- by multiplication by z and $\bar{z}$ respectively (since they are complex conjugates). Then $\mathrm{End}(V(\mathbb{C})) = \mathrm{Hom}(V^+, V^+) \oplus \mathrm{Hom}(V^+, V^-) \oplus \mathrm{Hom}(V^-, V^+) \oplus \mathrm{Hom}(V^-, V^-)$, and $h(z)$ acts on the factors by $1, \bar{z}/z, z/\bar{z}, 1$ respectively, so the action is of the type permitted for Shimura data. The image of $h(i) = J$ is a Cartan involution since J is a complex structure giving a polarization, and $\mathrm{GSp}(\psi)$ is simple and $\mathrm{GSp}(\psi)^{\mathrm{ad}} = \mathrm{Sp}_{2n}$ is noncompact, so all three axioms hold.

In fact, more is true: the "extra" axioms (4), (5), (6) all hold as well! The weight homomorphism on r is just multiplication by r, which is defined over $\mathbb{Q}$; the center of G is $\mathbb{G}_{\mathrm{m}}$, and $\mathbb{Q}^\times$ is discrete in $\mathbb{A}_f^\times$; and the center $\mathbb{G}_{\mathrm{m}}$ is already split over $\mathbb{Q}$. Thus this is as nice a Shimura datum as we could ask for, and is called the Siegel Shimura datum; by the above, it is also a PEL Shimura datum.

In particular, for any $(B, *, V, \psi)$, the action of B defines the algebraic group G as a subgroup of $\mathrm{GSp}(V, \psi)$, and composing with this inclusion sends any suitable homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ to an element of the Siegel Shimura datum $X(\psi)$. This embedding of Shimura data immediately implies that (G, X) satisfies conditions (1) through (4) since $(\mathrm{GSp}(\psi), X(\psi))$ does. In particular any PEL Shimura datum is in fact a Shimura datum, and so has an associated Shimura variety.

We can also look at the Shimura variety for the Siegel Shimura datum, and try to verify Theorem 1.1 explicitly in this case.

The data that Theorem 1.1 claims corresponds to a point of $\mathrm{Sh}_K(\mathrm{GSp}(\psi), X(\psi)) = \mathrm{GSp}(\psi)(\mathbb{Q})\backslash X(\psi) \times \mathrm{GSp}(\psi)(\mathbb{A}_f)/K$ is a tuple $(A, i, s, \eta K)$, where A is a complex abelian variety with a polarization s up to sign and ηK is a K-orbit of isomorphisms $V(\mathbb{A}_f) \to V_f(A)$ sending ψ to an $\mathbb{A}_f^\times$-multiple of s, such that there is an isomorphism $a : H_1(A, \mathbb{Q}) \to V$ sending s to a $\mathbb{Q}^\times$-multiple of ψ and that the cocharacter $r \mapsto a \circ h_A(r) \circ a^{-1}$ is in $X(\psi)$. (As in the case of modular curves, the data of i is trivial since $B = \mathbb{Q}$.

The existence of the isomorphism a implies that V and $H_1(A, \mathbb{Q})$ are isomorphic as $\mathbb{Q}$-vector spaces, i.e. $\dim A = n$. Thus the only data about A we need to keep track of, since we only remember isomorphism classes, is the rational Hodge structure on $H_1(A, \mathbb{Q})$ with its polarization and symplectic form s, which amounts to a vector space $W = H_1(A, \mathbb{Q})$ with a Hodge structure h of type $\{(-1, 0), (0, -1)\}$ and an isomorphism $a : W \to V$ satisfying this property on cocharacters.[1]

Given this data, we can form a point of $X(\psi)$ from a: by definition, each a gives rise to an element ah of $X(\psi)$ by $r \mapsto a \circ h(r) \circ a^{-1}$, where $h = h_A$ can be thought of as the cocharacter corresponding to the Hodge structure on W. From η, we can find an element of $\mathrm{GSp}(\psi)(\mathbb{A}_f)$ by $a_{\mathbb{A}_f} \circ \eta : V(\mathbb{A}_f) \to W(\mathbb{A}_f) \to V(\mathbb{A}_f)$, all of which are compatible with the symplectic structures. However, a is not part of our datum, so we want to quotient by the choice of a; and η is only defined up to the right action of K, so we also need to quotient on the right by K. Two possible choices of a differ by the automorphism group of (V, ψ), i.e. $\mathrm{GSp}(\psi)(\mathbb{Q})$, and so we have a well-defined map from the set of such tuples to $\mathrm{GSp}(\psi)(\mathbb{Q})\backslash X(\psi) \times \mathrm{GSp}(\psi)(\mathbb{A}_f)/K$.

First, we check that this map descends to the set of isomorphism class of tuples. Suppose that $f : (W, h, \eta K) \to (W', h', \eta' K)$ is an isomorphism of triples as above, and let $a : W \to V$ be a suitable isomorphism in the first case and $a' : W' \to V$ similarly in the second case. These have images $[ah, a_{\mathbb{A}_f} \circ \eta]$ and $[a'h', a'_{\mathbb{A}_f} \circ \eta']$. Since f is an isomorphism, it induces an isomorphism $(W, h) \to (W', h')$ of Hodge structures such that $f \circ \eta : V(\mathbb{A}_f) \to W(\mathbb{A}_f) \to W'(\mathbb{A}_f)$ and $\eta' : V(\mathbb{A}_f) \to W'(\mathbb{A}_f)$ agree up to the action of K, i.e. $\eta' K = f \circ \eta K$. In particular $a'_{\mathbb{A}_f} \circ \eta' K = a'_{\mathbb{A}_f} \circ f \circ \eta K = a \circ \eta K$ since f takes a to a', and so since f gives an isomorphism of Hodge structures up to rational multiples $[ah, a_{\mathbb{A}_f} \circ \eta = [a'h', a'_{\mathbb{A}_f} \circ \eta']$.

Next, we check that the map is injective on isomorphism classes. Suppose that $(W, h, \eta K)$ and $(W', h', \eta' K)$ map to the same class, i.e. (choosing a and a' as above) we have $(ah, a_{\mathbb{A}_f} \circ \eta) = (qa'h', q \circ a'_{\mathbb{A}_f} \circ \eta' \circ k)$ for some $q \in \mathrm{GSp}(\psi)(\mathbb{Q})$ and $k \in K$. Since a' is only defined up to the action of $\mathrm{GSp}(\psi)(\mathbb{Q})$, we can replace a' by $q^{-1}a'$, so the right-hand side is $(a'h', a'_{\mathbb{A}_f} \circ \eta' \circ k)$, and then $a' \circ a^{-1}$ gives an isomorphism between the two triples.

Finally, for any $[x, g]$, we can define a triple by $W = V$, $h = x$, and $\eta = g$ which maps to

[1] This relies on an omitted theorem, that the category of complex abelian varieties up to isogeny is equivalent to the category of polarizable rational Hodge structures of type $\{(-1, 0), (0, -1)\}$. This is a consequence of GAGA plus a result on polarizability of complex tori.

$[x, g]$, so the map is also surjective. Thus we obtain Theorem 1.1 in this case.

In the case where ψ is given by J, we can describe $X(\psi)$ very explicitly as the Siegel upper half plane $\mathcal{H}_n^{\pm}$, the space of symmetric complex $n \times n$ matrices with imaginary part either positive-definite or negative-definite. The Shimura varieties $\mathrm{Sh}_K(\mathrm{GSp}(\psi), X(\psi))$ are then the higher-dimensional analogues of modular curves.

2. Shimura varieties of Hodge type

In section 1, we looked at Shimura data arising from symplectic actions of semisimple $\mathbb{Q}$-algebras with involution; the key property from which we deduced that these were in fact Shimura data was that there was an embedding of Shimura data into some Siegel Shimura datum $(\mathrm{GSp}(\psi), X(\psi))$. In fact we can work more generally using just this property:

Definition 2.1. A Shimura datum (G, X) is of Hodge type if there is a symplectic space (V, ψ) over $\mathbb{Q}$ and an injection $\rho : G \hookrightarrow \mathrm{GSp}(\psi)$ carrying X to $X(\psi)$. The corresponding Shimura varieties $\mathrm{Sh}_K(G, X)$ are also said to be of Hodge type.

Thus in particular Shimura varieties of PEL type are also of Hodge type.

If $\mu : \mathrm{GSp}(\psi) \to \mathbb{G}_\mathrm{m}$ is the character sending g to the multiplier $\mu(g)$ such that $\psi(gx, gy) = \mu(g)\psi(x, y)$, then $\mu \circ \rho : G \to \mathbb{G}_\mathrm{m}$ is a character of G which we will again denote by μ. Write $\mathbb{Q}(r)$ for the vector space $\mathbb{Q}$ with the action of G by $g \cdot q = \mu(g)^r q$ for $q \in \mathbb{Q}(r)$ and $g \in G$. For any $h \in X$, one can show that $(\mathbb{Q}(r), \mu \circ h)$ is a rational Hodge structure of type $(-r, -r)$, so this is the standard notation for Hodge structures.

It turns out that for any embedded G, we can describe it by the stabilizer of certain tensors.

Lemma 2.2. *Let (G, X) be a Shimura datum of Hodge type, embedding into $(\mathrm{GSp}(\psi), X(\psi))$. There exists some finite set of multilinear maps $t_i : V \times \cdots \times V \to \mathbb{Q}(r_i)$ such that G is the subgroup of $\mathrm{GSp}(\psi)$ fixing the t_i.*

Proof. First, observe that a multilinear map $t : V \times \cdots \times V \to \mathbb{Q}$ is the same thing as an element of $(V^{\otimes m})^\vee$ for some integer m. A theorem of Chevalley states that any faithful self-dual representation $G \to \mathrm{GL}(V)$ gives G as the stabilizer in $\mathrm{GL}(V)$ of some one-dimensional subspace L of a representation W of $\mathrm{GL}(V)$. For any nonzero $t \in L \otimes L^\vee \subset W \otimes W^\vee$, the group of $g \in \mathrm{GL}(V)$ fixing t is exactly the stabilizer G of L. We can write the tensor representation $W \otimes W^\vee$ as a subrepresentation of a sum of representations of the form $V^{\otimes m_i} \otimes (V^{\otimes n_i})^\vee$; we can then decompose t into terms t_i from each such factor. But the symplectic form ψ gives an isomorphism of G-representations $\psi : V \times V \to \mathbb{Q}(1)$, since $\psi(gx, gy) = \mu(g)\psi(x, y)$, and thus an isomorphism $V \simeq V^\vee \otimes \mathbb{Q}(1)$, so $V^{\otimes m_i} \otimes (V^{\otimes n_i}) \simeq (V^\vee)^{\otimes(m_i+n_i)} \otimes \mathbb{Q}(m_i) = \mathrm{Hom}(V^{\otimes(m_i+n_i)}, \mathbb{Q}(m_i))$, and so in particular we can write each t_i in the form $\mathrm{Hom}(V^{\otimes m_i}, \mathbb{Q}(r_i))$ for some m_i, r_i, and the group fixing all the t_i is G. □

This is enough to let us state—and prove!—a moduli interpretation for Shimura varieties of Hodge type.

Theorem 2.3. *Let (G, X) be a Shimura datum of Hodge type, with a fixed embedding into the Siegel modular datum $(\mathrm{GSp}(\psi), X(\psi))$ corresponding to some symplectic space (V, ψ) and presentation as the fixed group of some multilinear maps t_i. Then for any compact open subgroup*

$K \subset G(\mathbb{A}_f)$, $\mathrm{Sh}_K(G,X)(\mathbb{C})$ classifies isomorphism classes of tuples $(W, h, s_0, s_1, \dots, s_n, \eta K)$, where

- *h is a rational Hodge structure on W of type $\{(-1,0),(0,-1)\}$,*
- *s_0 is a polarization, up to sign, of (W,h),*
- *$s_1, \dots, s_n$ are multilinear maps $W \times \cdots \times W \to \mathbb{Q}(r_i)$, and*
- *ηK is a K-orbit of isomorphisms $V(\mathbb{A}_f) \to W(\mathbb{A}_f)$ sending ψ to an $\mathbb{A}_f^\times$-multiple of s_0 and each t_i to s_i,*

such that there exists an isomorphism $a : W \to V$ sending s_0 to a $\mathbb{Q}^\times$-multiple of ψ, each s_i to t_i, and h to an element of X.

Here an isomorphism of such tuples is an isomorphism of rational Hodge structures $(W,h) \to (W',h')$ sending s_0 to s_0' up to a $\mathbb{Q}^\times$-multiple, each s_i to s_i', and ηK to $\eta' K$.

Proof. The map from the set of tuples to $\mathrm{Sh}_K(G,X)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K$ is given by choosing an isomorphism $a : W \to V$ satisfying the conditions of the theorem and sending the tuple $(W, h, s_0, s_1, \dots, s_n, \eta K)$ to the pair $(ah, a_{\mathbb{A}_f} \circ \eta)$, where $ah : \mathbb{S} \to G_{\mathbb{R}}$ is defined by $(ah)(z) = a \circ h(z) \circ a^{-1}$ and is assumed to be in X. Since $a_{\mathbb{A}_f} \circ \eta$ by definition fixes the t_i and preserves the symplectic structure, it is in $G(\mathbb{A}_f)$; again η is only defined up to the right action of K and our choice of a is up to the action of $G(\mathbb{Q})$, so this gives a map from the set of isomorphisms into the double quotient $G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K$. Verifying that it descends to isomorphism classes and is a bijection is as in the discussion for Siegel Shimura varieties. □

This solves the problem of providing a moduli interpretation for Shimura varieties of Hodge type, but it is not very satisfying: for example, it doesn't obviously specialize to Theorem 1.1 in the case of PEL Shimura varieties. To improve this interpretation, we introduce the notion of Hodge tensors.

Let $t : V \times \cdots \times V \to \mathbb{Q}(r)$ be a multilinear map, i.e. $t(gv_1, \dots, gv_m) = \mu(g)^r t(v_1, \dots, v_m)$. For $h \in X$, this gives a morphism of Hodge structures $(V,h)^{\otimes m} \to \mathbb{Q}(r)$. Since $\mathbb{Q}(r)$ has weight $-2r$ and (V,h) has weight -1, it follows that t is zero unless $m = 2r$.

Let A be an abelian variety over $\mathbb{C}$, and let $W = H_1(A,\mathbb{Q})$. The cohomology of abelian varieties over $\mathbb{C}$ is given by $H^m(A,\mathbb{Q}) \simeq \mathrm{Hom}(\bigwedge^m W, \mathbb{Q})$. For $t \in H^{2r}(A,\mathbb{Q})$, we say that t is a Hodge tensor for A if the corresponding map $W^{\otimes 2r} \to \bigwedge^{2r} W \to \mathbb{Q}(r)$ is a morphism of Hodge structures.

We can now rephrase Theorem 2.3.

Theorem 2.4. *Let (G,X) be a Shimura datum of Hodge type, with a fixed embedding into the Siegel modular datum $(\mathrm{GSp}(\psi), X(\psi))$ corresponding to some symplectic space (V,ψ) and presentation as the fixed group of some multilinear maps t_i. Then for any compact open subgroup $K \subset G(\mathbb{A}_f)$, $\mathrm{Sh}_K(G,X)(\mathbb{C})$ classifies isomorphism classes of tuples $(A, s_0, s_1, \dots, s_n, \eta K)$, where*

- *A is a complex abelian variety,*
- *s_0 is a polarization, up to sign, of the rational Hodge structure on $H_1(A,\mathbb{Q})$,*

- *$s_1, \ldots, s_n$ are Hodge tensors for A or its powers, and*
- *ηK is a K-orbit of isomorphisms $V(\mathbb{A}_f) \to V_f(A)$ sending ψ to an $\mathbb{A}_f^\times$-multiple of s_0 and each t_i to s_i,*

such that there exists an isomorphism $a : H_1(A, \mathbb{Q}) \to V$ sending s_0 to a $\mathbb{Q}^\times$-multiple of ψ, each s_i to t_i, and h to an element of X.

Here an isomorphism of tuples is as above; note that isomorphisms of abelian varieties are after inverting isogenies.

Proof. Again, the data of A is equivalent to the data of $H_1(A, \mathbb{Q})$ with its rational Hodge structure, so by the discussion above the data of the Hodge tensors s_i is equivalent to the data of the multilinear maps from Theorem 2.3 and so the claim follows immediately from Theorem 2.3. □

Note that now Theorem 1.1 follows immediately: in the special case of PEL Shimura varieties, for any set of generators $b_1, \ldots, b_m$ of the semisimple $\mathbb{Q}$-algebra B we can use the tensors $t_{b_i} : (x, y) \mapsto \psi(x, by)$. The g which fix all of these t_{b_i} are exactly those which commute with b, i.e. G, so the action of B on A (up to isogeny) is the same as the data of the Hodge tensors (or equivalently multilinear maps to $\mathbb{Q}(r_i)$) s_i sent to the t_{b_i} by the isomorphisms η^{-1} and a of the theorem.

3. Shimura varieties of abelian type

The most general Shimura varieties we've considered so far are those of Hodge type, which are moduli spaces for abelian varieties with tensor and level structures. In order to obtain moduli interpretations for larger classes of Shimura varieties, we need to consider a more general category of objects. These will be abelian motives.

When working with Hodge structures, one straightforward operation is the direct sum. When the Hodge structures come from the cohomology of abelian varieties, this corresponds to disjoint union, so we first want to enlarge our category to complex varieties V whose connected components V_i are abelian varieties. In particular $H^*(V, \mathbb{Q})$ acquires a polarizable Hodge structure from its summands $H^*(V_i, \mathbb{Q})$, which in turn derive their Hodge structures from $H_1(V_i, \mathbb{Q})$. Write $H^*(V, \mathbb{Q})(m)$ for $H^*(V, \mathbb{Q}) \otimes \mathbb{Q}(m)$.

An abelian motive over $\mathbb{C}$ is a triple (V, e, m), where V is a variety over $\mathbb{C}$ whose components are abelian varieties, e is an idempotent of the rational Hodge structure on $H^*(V, \mathbb{Q})$ (i.e. $e^2 = e$, so e induces a splitting of the Hodge structure into the images of e and $1 - e$) and m is an integer. For example, if A is an abelian variety, the projection to the ith component $e^i : H^*(A, \mathbb{Q}) \to H^i(A, \mathbb{Q}) \subset H^*(A, \mathbb{Q})$ is an idempotent; we write the motive $(A, e^i, 0)$ as $h^i(A)$. A morphism of abelian motives $(V, e, m) \to (V', e', m')$ is a map $e' \circ f \circ e : H^*(V, \mathbb{Q}) \to H^*(V', \mathbb{Q})$ where $f : H^*(V, \mathbb{Q}) \to H^*(V', \mathbb{Q}')$ is a ring homomorphism of degree $m' - m$.

We can also define a direct sum (for motives with m in common) and tensor product:

$$(V, e, m) \oplus (V', e', m) = (V \sqcup V', e \oplus e', m),$$

$$(V, e, m) \otimes (V', e', m') = (V \times V', e \otimes e', m + m').$$

We can also define a dual motive: if V is pure of dimension d, then set

$$(V, e, m)^\vee = (V, e^\mathsf{T}, d - m)$$

where e^T denotes the transpose of e as a correspondence.

Given an abelian motive (V, e, m) over $\mathbb{C}$, we can define its cohomology $H(V, e, m)$ to be $eH^*(V, \mathbb{Q})(m)$. For example, $H(h^i(A)) = e^i H^*(A, \mathbb{Q}) = H^i(A, \mathbb{Q})$. This defines a functor H from the category of abelian motives to the category of polarizable rational Hodge structures, and it commutes with all three operations above, direct sum, tensor product, and duality. If a rational Hodge structure (W, h) is in the essential image of this functor, i.e. $(W, h) \simeq H(V, e, m)$ for some abelian motive (V, e, m), we say that (W, h) is abelian. For example, if A is an elliptic curve then $\mathbb{Q}(1) \simeq \bigwedge^2 H_1(A, \mathbb{Q}) \simeq H^2(A, \mathbb{Q})^\vee = h^2(A)^\vee$, so $\mathbb{Q}(1)$ is abelian. Observe that H then gives an equivalence of categories between the category of abelian motives and the category of abelian rational Hodge structures.

To define Shimura varieties of abelian type, we first go back to connected Shimura varieties. For the Siegel Shimura datum $(\mathrm{GSp}(\psi), X(\psi))$ for (V, ψ), the corresponding connected Shimura data are of the form $(\mathrm{Sp}(\psi), X(\psi)^\pm)$, where $\mathrm{Sp}(\psi)$ is the usual symplectic group for ψ and $X(\psi)^\pm$ corresponds to complex structures J on $V(\mathbb{R})$ such that $\psi(Ju, Jv) = \psi(u, v)$ as usual and additionally $\pm\psi(u, Jv) > 0$, so e.g. $\psi(u, Jv) > 0$ in the positive case.

Definition 3.1. (a) A connected Shimura datum (H, X^+) is of primitive abelian type if H is simple and there exists a symplectic space (V, ψ) over $\mathbb{Q}$ and an injection $H \to \mathrm{Sp}(\psi)$ taking X^+ to $X(\psi)^+$.

(b) A connected Shimura datum (H, X^+) is of abelian type if H is isogenous to a product of H_i with (H_i, X_i^+) each of primitive abelian type with the isogeny $\prod_i H_i \to H$ taking $\prod_i X_i^+$ to X^+.

(c) A Shimura datum (G, X) is of abelian type if (G^{der}, X^+) is of abelian type.

(d) Let $K \subset G(\mathbb{A}_f)$ be a compact open subgroup. A Shimura variety $\mathrm{Sh}_K(G, X)$, resp. a connected Shimura variety $\mathrm{Sh}_K^\circ(G, X^+)$, is of abelian type if (G, X), resp. (G, X^+), is.

The following proposition, due to Milne, relates the abelian condition for Shimura data and Hodge structures.

Proposition 3.2. *Let (G, X) be a Shimura datum satisfying the additional axioms (4) (the weight homomorphism w_X is rational) and (6) (the connected component of the center Z° splits over a CM field), and additionally suppose that there is a character $\mu : G \to \mathbb{G}_m$ such that $\mu \circ w_X : \mathbb{G}_m \to \mathbb{G}_m$ is -2. Then (G, X) is of abelian type if and only if for some (equivalently, all) representation (V, ρ) of G and every $h \in X$, the Hodge structure $(V, \rho \circ h)$ is abelian.*

Let (G, X) be a Shimura datum of abelian type satisfying the conditions of the proposition, and let $\rho : G \to \mathrm{GL}(V)$ be a faithful representation of G. Suppose there exists a pairing $\psi : V \times V \to \mathbb{Q}$ such that $\psi(gx, gy) = \mu(g)^m \psi(x, y)$ for some fixed m, and ψ is a polarization of $(V, \rho \circ h)$ for every $h \in X$. Then as in the proof of Lemma 2.2 the faithful representation ρ allows us to express G as the group satisfying this condition on ψ and fixing some multilinear maps $t_i : V \times \cdots \times V \to \mathbb{Q}(r_i)$. Fix such t_i.

Theorem 3.3. *With the above notation,* $\mathrm{Sh}_K(G, X)(\mathbb{C})$ *classifies the isomorphism classes of tuples* $(A, s_0, s_1, \ldots, s_n, \eta K)$*, where*

- *A is an abelian motive,*
- *s_0 is a polarization, up to sign, for the rational Hodge structure $H(A)$,*
- *$s_1, \ldots, s_n$ are tensors for A, and*
- *ηK is a K-orbit of isomorphisms $V(\mathbb{A}_f) \to V_f(A)$ sending ψ to an $\mathbb{A}_f^\times$-multiple of s_0 and each t_i to s_i,*

such that there exists an isomorphism $a : H(A) \to V$ sending s_0 to a $\mathbb{Q}^\times$-multiple of ψ, each s_i to t_i, and h to an element of X.

Here a tensor for A means a tensor for $H(A)$, i.e. an element of $H(A)^{\otimes m} \otimes (H(A)^{\otimes n})^\vee$ for some integers m, n.

If $A = h^1(A)^\vee$ for an abelian variety A, by an abuse of notation, so that $H(A) = H_1(A, \mathbb{Q})$, then the Hodge condition of Theorem 2.4 on the s_i is automatically satisfied due to the conditions on the t_i and so we recover Theorem 2.4, and thus Theorem 1.1.

Proof sketch. The same proof as for Theorem 2.3 shows that $\mathrm{Sh}_K(G, X)(\mathbb{C})$ classifies tuples $(W, h, s_0, s_1, \ldots, s_n, \eta K)$, with (W, h) a rational Hodge structure replacing $H(A)$ and otherwise the same conditions. The isomorphism a shows that every such (W, h) is isomorphic to $(V, \rho \circ h')$ for some $h' \in X$, so by the conditions we assumed on (G, X) and Proposition 3.2 every (W, h) is abelian. Therefore by the equivalence of categories between abelian rational Hodge structures and abelian motives we can write every (W, h) as $H(A)$ for some abelian motive A. ☐

In fact, it is possible to classify the connected Shimura data of abelian type; this is due to Deligne. Suppose that (G, X^+) is a connected Shimura datum where G is simple. If G^{ad} is of type A, B, or C, then (G, X^+) is of abelian type. If G^{ad} is of type E_6 or E_7, then there is no symplectic embedding and so (G, X^+) is not of abelian type. If G^{ad} is of type D, the situation is ambiguous: (G, X^+) may be of abelian type, or one of two things may go wrong.

(a) It may be that there is some nontrivial finite algebraic subgroup $N \subset G$ in the kernel of every $(G, X^+) \to (\mathrm{Sp}(\psi), X(\psi)^+)$, so that any such morphism fails to be an embedding. In this case for any normal N' containing N the quotient $(G/N', X^+)$ is of abelian type (at least assuming G is the universal cover of G^{ad}), but (G, X^+) is not.

(b) There may be no homomorphism $G \to \mathrm{Sp}(\psi)$ at all, in which case certainly (G, X^+) cannot be abelian.

This last can happen for Lie-theoretic reasons: over $\mathbb{R}$, in general G^{ad} may decompose as a product of simple groups, and if there is a homomorphism $G_{\mathbb{R}} \to \mathrm{Sp}(\psi)_{\mathbb{R}}$ (and thus on all factors) which descends to $\mathbb{Q}$, the action of the Galois group permutes the Dynkin diagrams of the factors and so all of the factors must have the nodes corresponding to these symplectic representations in the same position. In particular the opposite node, corresponding to a cocharacter from $\mathbb{S}$, must all be in the same position, which is the case for some but not all groups of type D.

4. General abelian varieties

One hopes that, generalizing the examples we've discussed, any Shimura variety with rational weight (axiom (4)) will classify isomorphism classes of motives with some additional structure. More precisely, given a rational representation $\rho : G \to \mathrm{GL}(V)$, we get a family of Hodge structures $\rho_{\mathbb{R}} \circ h$ on V for $h \in X$; if the weight of (G, X) is defined over $\mathbb{Q}$, we hope that these Hodge structures occur naturally in the cohomology of some algebraic varieties. However this is not known[2] for any Shimura varieties more general than of abelian type.

Consider for example a quaternion algebra B over a totally real field F, with G the algebraic group with $G(\mathbb{Q}) = B^\times$, so that $B \otimes \mathbb{R}$ decomposes as a product of algebras, one for each embedding $F \hookrightarrow \mathbb{R}$, which are either $M_2(\mathbb{R})$ or the Hamiltonian quaternions $\mathbb{H}$, and so $G(\mathbb{R})$ is a product of $\mathrm{GL}_2(\mathbb{R})$ and $\mathbb{H}^\times$. We can define a homomorphism $h : \mathbb{C}^\times \to G$ by sending $a + bi$ to 1 in the $\mathbb{H}^\times$ terms and $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ in the $\mathrm{GL}_2(\mathbb{R})$ terms. There are several possibilities:

(a) If B is already split over F, then (G, X) is of PEL type, and by Theorem 1.1 its Shimura variety classifies abelian varieties of dimension $[F : \mathbb{Q}]$ with an F-action since $G(\mathbb{R})$ factors as a product of $[F : \mathbb{Q}]$ copies of $\mathrm{GL}_2(\mathbb{R})$, each of which acts on a 2-dimensional space with a corresponding action of F. These are Hilbert or Hilbert-Blumenthal varieties, and generalize modular curves.

(b) If B does not split over F but every factor base changed to $\mathbb{R}$ splits, then (G, X) is again of PEL type, and we again get an action of each $\mathrm{GL}_2(\mathbb{R})$ factor but now the corresponding semisimple $\mathbb{Q}$-algebra is not F but B, and so the dimension changes: $\mathrm{Sh}_K(G, X)$ classifies abelian varieties of dimension $2[F : G]$ with a B-action (by isogenies).

(c) If B does not split over F and has at least one factor not splitting over $\mathbb{R}$, then (G, X) is of abelian type, but the weight is not rational: w_X sends a real number r to the element of $(F \otimes \mathbb{R})^\times \simeq \prod_{F \hookrightarrow \mathbb{R}} \mathbb{R}$ given by r for the split terms and 1 for the nonsplit terms. Let $T = \mathrm{Res}_{F/\mathbb{Q}} \mathbb{G}_{\mathrm{m}}$ be the torus over $\mathbb{Q}$ such that $T(\mathbb{Q}) = F^\times$. Then $w_X : \mathbb{G}_{\mathrm{m}} \to T_{\mathbb{R}}$ is defined over the subfield of the algebraic numbers fixed by the subgroup of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing the embeddings $F \hookrightarrow \mathbb{R}$ corresponding to the nonsplit factors, which is not $\mathbb{Q}$ for F different from $\mathbb{Q}$. Thus although $\mathrm{Sh}_K(G, X)$ will classify certain Hodge structures, these do not arise from the cohomology of algebraic varieties, i.e. they are not motivic, since then they would be rational.

[2] At least at the time of Milne's notes (originally 2004, most recently revised 2017).

(d) If B is split over $\mathbb{R}$ at exactly one place, then the Shimura variety is a curve; these were Shimura's original examples, namely Shimura curves.

In the setup of PEL Shimura varieties where we have a semisimple $\mathbb{Q}$-algebra with involution $(B, *)$ and a symplectic $(B, *)$-module (V, ψ), if G is the corresponding algebraic group we saw in Theorem 1.1 that under certain conditions (e.g. G is of type A (with V having even reduced dimension) or C) that there is a unique $G(\mathbb{R})$-conjugacy class X giving a Shimura datum and the set of abelian varieties with a B-action and suitable structure are classified by the Shimura varieties for (G, X). In general this fails to be true, but under weaker conditions they are still represented by some algebraic variety, called the PEL modular variety attached to $(B, *, V, \psi)$. In general it is a finite disjoint union of Shimura varieties. This suggests that to get a more general moduli interpretation we should broaden the definition of Shimura varieties to allow G to be nonconnected, and the study of boundaries of Shimura varieties suggests that we should similarly allow X to be a finite cover of a conjugacy class of homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$. This is exactly what we did in the case of zero-dimensional Shimura varieties, and so would make those into genuine Shimura varieties (unsurprisingly, the zero-dimensional ones).

References

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