SHIMURA VARIETIES

Lecture 1: Moduli of Elliptic Curves

One useful feature of Shimura varieties is that they translate problems in moduli theory to problems in linear algebra. The linear algebra is rather complicated, because it involves real/complex vector spaces as well as p-adic vector spaces for all p, together with some compatibility among them, but it apparently makes the issues in moduli theory much less complicated. For this reason, it rarely works; not so many varieties can be described completely in terms of linear algebra. The ones that can are abelian varieties, because, over $\mathbb{C}$ at least, they are quotients of complex vector spaces by lattices, with some additional properties to guarantee algebraicity. Surprisingly (or not), abelian varieties over finite fields can also be recovered from linear algebra, together with some rather subtle Galois cohomology.

So before we can talk about Shimura varieties, we need to talk about abelian varieties. I'll begin with elliptic curves, for which most of the theory is straightforward and intuitive.

Over $\mathbb{C}$, an elliptic curve E is given as a quotient $\mathbb{C}/\Lambda$, where $\Lambda \xrightarrow{\sim} \mathbb{Z}^2$ is a lattice in $\mathbb{C}$. More precisely, any smooth projective curve of genus 1 over any field is a group. Over $\mathbb{C}$ it is a topological group, so its universal cover is a simply-connected complex manifold of dimension 1 which is also a group, and therefore must be $\mathbb{C}$; then $\Lambda = \pi_1(E) = H_1(E, \mathbb{Z})$. There are more insightful ways to see this, but I'll remain briefly at this level.

Now if Λ and Λ' are two lattices, the elliptic curves $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ are isomorphic as complex algebraic varieties if and only if they are isomorphic as groups. In other words, any holomorphic automorphism ϕ of $\mathbb{C}$ that takes Λ to Λ' such that $\phi(0) = 0$ is a homomorphism of groups.

Exercise. *Prove this fact using only complex analysis. Prove it again using only algebraic geometry.*

Every holomorphic group automorphism of $\mathbb{C}$ is given by multiplication by an element $\alpha \in \mathbb{C}^\times$. Indeed, any continuous group automorphism of a real vector space is necessarily a linear map, and one checks that $\phi(a + bi) = a\alpha_1 + b\alpha_i$ is holomorphic if and only if $\alpha_i = i\alpha_1$. So $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ if and only if $\Lambda' = \alpha\Lambda$ for some $\alpha \in \mathbb{C}^\times$. Now we orient $\mathbb{C}$, as real vector space, so that $(1, i)$ is a positive orientation. The set Ω of pairs (ω, ω') of positively oriented $\mathbb{R}$-bases of $\mathbb{C}$ can be identified, in the basis $\{1, i\}$, with

$$GL(2, \mathbb{R})^+ = \{g \in GL(2, \mathbb{R}) \mid \det(g) > 0\},$$

but we're not yet ready for this. However, the group $SL(2, \mathbb{Z})$ acts on Ω via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (\omega, \omega') = (a\omega + b\omega', c\omega + d\omega').$$

This action fixes the lattice $\mathbb{Z}\omega + \mathbb{Z}\omega' \subset \mathbb{C}$, and the quotient $SL(2,\mathbb{Z})\backslash\Omega$ is thus identified with the set of all lattices in $\mathbb{C}$. Thus the set of complex elliptic curves is in one-to-one correspondence with the double coset space

$$SL(2,\mathbb{Z})\backslash\Omega/\mathbb{C}^\times$$

where $\alpha \in \mathbb{C}^\times$ takes (ω,ω') to $(\alpha\omega,\alpha\omega')$, or equivalently takes the lattice Λ to $\alpha\Lambda$.

The next step might be to observe that the map

$$(\omega,\omega') \mapsto \omega'/\omega$$

identifies $\Omega/\mathbb{C}^\times$ with the upper half plane $\mathfrak{H} \subset \mathbb{C}$:

$$\mathfrak{H} = \{x + iy \in \mathbb{C} \mid y > 0\}.$$

Under the identification of Ω with $GL(2,\mathbb{R})^+$, this map takes the identity matrix to the point $i \in \mathfrak{H}$. Moreover, the map is equivariant for the tautological left action of $GL(2,\mathbb{R})^+$ on $\Omega/\mathbb{C}^\times$, if we let $GL(2,\mathbb{R})^+$ act on $\mathfrak{H}$ by linear fractional transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d}.$$

Thus the set of complex elliptic curves is in one-to-one correspondence with the quotient $SL(2,\mathbb{Z})\backslash\mathfrak{H}$, where $SL(2,\mathbb{Z})$ acts via linear fractional transformations. And we can then go on to develop the classical theory of the j-function, which is a holomorphic map from $\mathfrak{H}$ to $\mathbb{C}$, invariant and bijective modulo $SL(2,\mathbb{Z})$, and in this way make the connection with the algebraic theory of elliptic curves. However, this doesn't generalize to higher dimensions, and instead we will do something that brings us closer to the group theory.

In the above discussion, $\mathbb{C}$ was treated as fixed and the lattice Λ was moving around in $\mathbb{C}$. It is much more clever to fix $\Lambda = \mathbb{Z}^2$ and let $\mathbb{C}$ change. More precisely, an inclusion $\mathbb{Z}^2 \subset \mathbb{C}$ identifies $\mathbb{C}$ with $\mathbb{R}^2$; we fix $\mathbb{R}^2$ and its canonical lattice $\mathbb{Z}^2$, and let the complex structure vary.

Definition. *A* complex structure *on $\mathbb{R}^2$ is a homomorphism $h : \mathbb{C}^\times \to GL(2,\mathbb{R}) = Aut(\mathbb{R}^2)$ such that the eigenvalues of $h(z) \in \mathbb{C}^\times$ on $\mathbb{R}^2$ are z and $\bar{z}$. In other words, it is a homomorphism of groups that extends to a homomorphism of $\mathbb{R}$-algebras $\mathbb{C} \to M(2,\mathbb{R})$.*

Choosing the base point $e_0 = (1,0) \in \mathbb{R}^2$, we see that any complex structure h defines an isomorphism $i_h : \mathbb{R}^2 \to \mathbb{C}$ of complex vector spaces, via $i_h^{-1}(z) = h(z)\cdot e_0$. And thus $\mathbb{C}/i_h(\mathbb{Z}^2)$ is an elliptic curve. We will view this in another way.

Example. Let $V = \mathbb{Q}^2$, and let $h : \mathbb{C}^\times \to Aut(\mathbb{R}^2)$ be a complex structure. Then for any $z \in \mathbb{C}$, $z \notin \mathbb{R}$, $h(z)$ has two eigenvalues on $V_\mathbb{C}$, namely z and $\bar{z}$. We let $w = -1$ and let $V^{-1,0} = V_h^{-1,0}$, resp. $V^{0,-1} = V_h^{0,-1}$, denote the z-eigenspace, resp. the $\bar{z}$-eigenspace, for $h(z)$ on $V_\mathbb{C}$. For example, we can define a complex structure by the homomorphism $h_0 : \mathbb{C}^\times \to GL(2,\mathbb{R})$ such that $h_0(x+iy) = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$; this obviously satisfies the hypothesis, with $V^{-1,0} = \mathbb{C}\cdot v_0$, $V^{0,-1} = \mathbb{C}\cdot v_0'$, where

$$v_0 = \begin{pmatrix} -i \\ 1 \end{pmatrix},\ v_0' = \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

(Check this by hand, and note carefully the signs!) For general h, since $\mathbb{C}^\times$ is a commutative group, it is easy to see that the decomposition $V_{\mathbb{C}} = V^{-1,0} \oplus V^{0,-1}$ does not depend on the choice of z; indeed, each $V^{p,q}$ is an eigenspace for the action of $h(\mathbb{C}^\times)$. Furthermore, because $h(z)$ acts by $\mathbb{R}$-linear transformations, it is easy to see that

$$V^{-1,0} = \bar{V}^{0,-1}.$$

Verification: Let $z = x + iy \in \mathbb{C}^\times$, $y \neq 0$. For example, take $z = i$. Incidentally, this choice is not completely innocent: there are two choices of $\sqrt{-1} \in \mathbb{C}$ and *a priori* the constructions that follow depend on the choice. Anyway, there is a basis v, v' of $V \otimes \mathbb{C}$ such that $h(i)v = iv$, $h(i)v' = -iv'$. Thus $V^{-1,0} = \mathbb{C}\cdot v$, $V^{0,-1} = \mathbb{C}\cdot v'$. On the other hand, $h(i) \in Aut(\mathbb{R}^2) = GL(2, \mathbb{R})$ is a real matrix with eigenvalues i, $-i$, hence there is a real matrix γ such that

$$\gamma^{-1} h(i) \gamma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = h_0(i).$$

Then we see that, with v_0 and v_0' as above, we have $\gamma\mathbb{C}\cdot v_0 = V_h^{-1,0}$, resp. $\gamma\mathbb{C}\cdot v_0' = V_h^{0,-1}$; in other words there are non-zero complex scalars λ, λ' such that

$$v = \lambda\gamma v_0, v' = \lambda'\gamma v_0'.$$

Say $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$v = \lambda \begin{pmatrix} -ai + b \\ -ci + d \end{pmatrix}, \quad v' = \lambda' \begin{pmatrix} ai + b \\ ci + d \end{pmatrix}$$

Since c and d are real, $\pm ci + d \neq 0$, hence we can normalize v (resp. v') by taking $\lambda = (-ci + d)^{-1}$, (resp. $\lambda' = (ci + d)^{-1}$. Let $\tau_h = \gamma(i) = \frac{ai+b}{ci+d}$. Then

$$v' = \begin{pmatrix} \tau_h \\ 1 \end{pmatrix}; \; v = \begin{pmatrix} \bar{\tau_h} \\ 1 \end{pmatrix} = \bar{v'}.$$

From this it follows that $V^{-1,0} = \bar{V}^{0,-1}$, as claimed.

Indeed, it would have sufficed to carry out this verification for h_0, then to derive the corresponding fact for h from the fact that γ is a real matrix. However, we have obtained a side benefit. Let $\mathfrak{H}^\pm = \mathbb{C} - \mathbb{R}$, the union of the upper and lower half planes. The group $GL(2, \mathbb{R})$ acts by fractional linear transformations on $\mathfrak{H}^\pm$ as above. The complex number $\tau_h = \gamma(i)$ then belongs to $\mathfrak{H}^\pm$. Moreover we can define a map

$$\pi : \{\text{complex structures }\} \to \mathfrak{H}^\pm$$

by $\pi(h) = \tau_h$. This map may appear to depend on the choice of the matrix γ such that $h(i) = \gamma h_0(i)\gamma^{-1}$. We write $\tau_h(\gamma)$ to take provisional account of this dependence. Note first of all that h_0 and h both extend to algebra homomorphisms $\mathbb{C} \to M(2, \mathbb{R})$, and since i generates $\mathbb{C}$ as $\mathbb{R}$-algebra it follows that $\gamma h_0 \gamma^{-1} = h$. If γ' is another choice, then $k = \gamma'^{,-1}\gamma$ belongs to the centralizer in $M(2, \mathbb{R})$ of h_0,

i.e. to the centralizer of $h_0(\mathbb{C})$, which is just $h_0(\mathbb{C})$. Thus $k \in h_0(\mathbb{C}^\times)$, and if $k = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ we have

$$\tau_h(\gamma) = \gamma(i) = \gamma' k(i) = \gamma'(k(i)) = \gamma'(\frac{xi+y}{-yi+x}) = \gamma'(i) = \tau_h(\gamma'),$$

so there is no dependence. In other words, letting $K_\infty = h_0(\mathbb{C}^\times) \subset GL(2,\mathbb{R})$, there is a sequence of identifications

$$\{\text{complex structures }\} \xrightarrow{\sim} GL(2,\mathbb{R})/K_\infty \xrightarrow{\sim} \mathfrak{H}^\pm.$$

The significance of this is that the final term has an obvious complex structure, hence so do the first two terms. Moreover, this complex structure is $GL(2,\mathbb{R})$-invariant.

There is more. The function associating the normalized vector $v' = v'_h \in V^{0,-1}$ to h. is compatible with the complex structure. Now $V_h^{0,-1} \subset V_\mathbb{C}$ is a variable line in $V_\mathbb{C}$, hence defines a variable point $p_h \in \mathbb{P}(V_\mathbb{C}) = \mathbb{P}^1(\mathbb{C})$. If $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ is the homogeneous coordinate of a point in $\mathbb{P}^1$, we use the standard inhomogeneous coordinate $\frac{\alpha}{\beta}$. Then the inhomogeneous coordinate of $V_h^{0,-1}$ is just τ_h. We thus have a holomorphic embedding

$$\{\text{complex structures }\} \xrightarrow{\sim} GL(2,\mathbb{R})/K_\infty \hookrightarrow \mathbb{P}(V_\mathbb{C})$$

obtained by associating the subspace $V_h^{0,-1}$ to h.

Later we will ignore the coordinates and use the embedding in $\mathbb{P}(V_\mathbb{C})$ to define the complex structure. Meanwhile, it is time to return to our family of elliptic curves $E_h = i_h(\mathbb{Z}^2)\backslash\mathbb{C}$, where h is a variable complex structure. The pertinent question is: what is the $\mathbb{C}$ in the numerator? Recall the formula for $i_h : \mathbb{R}^2 \xrightarrow{\sim} \mathbb{C}$:

$$i_h(h(z)e_0) = z = z \cdot i_h(e_0).$$

The map i_h extends by linearity to a surjective homomorphism

$$\mathbb{R}^2 \otimes \mathbb{C} = V_\mathbb{C} \to \mathbb{C}.$$

The left hand side is $V^{-1,0} \oplus V^{0,-1}$, and since the formula shows that i_h commutes with the action of $\mathbb{C}^\times$ on both sides, it follows that the map $V_\mathbb{C} \to \mathbb{C}$ is the projection $V_\mathbb{C} \to V_\mathbb{C}/V^{0,-1}$. In other words, the $\mathbb{C}$ in the numerator is identified with $V^{-1,0}$, and we have the formula

$$E_h = \mathbb{Z}^2\backslash V_\mathbb{C}/V_h^{0,-1}.$$

It can also be verified by hand from the above formulas that $i_h(\mathbb{Z}^2) = \mathbb{Z} \oplus \mathbb{Z} \cdot \tau_h$. Indeed,

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\tau_h \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \tau_h \\ 1 \end{pmatrix} = \tau_h e_0 + v'_h,$$

and since $i_h(v'_h) = 0$ it follows that $i_h(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = -\tau_h$.

(Question: why is i_h orientation-reversing?)

The family $\mathcal{E}$ of E_h is parametrized by the set of complex structures h, or by $GL(2,\mathbb{R})/K_\infty$, or by $\mathfrak{H}^\pm$. But as we saw above, elliptic curves over $\mathbb{C}$ are parametrized by $SL(2,\mathbb{Z})\backslash\mathfrak{H} = GL(2,\mathbb{Z})\backslash\mathfrak{H}^\pm$. The family $\mathcal{E}/\mathfrak{H}^\pm$ does not admit a quotient by $GL(2,\mathbb{Z})$. More precisely, there is an action of $GL(2,\mathbb{Z})$ on the family $\{E_h = V_\mathbb{C}/V_h^{0,-1}\}$ covering the action on $\mathfrak{H}^\pm$, and preserving the subgroup $\mathbb{Z}^2$; we simply let $g \in GL(2,\mathbb{Z}) = Aut(\mathbb{Z}^2)$ act naturally on $\mathbb{Z}^2 \subset V_\mathbb{C}$ and by conjugation on h. However, the element $-I_2 \in GL(2,\mathbb{Z})$ acts as -1 on each E_h and the quotient is no longer a family of elliptic curves; and there are other elliptic fixed points in $\mathfrak{H}^\pm$ whose stabilizers define automorphisms of the corresponding elliptic curves. However, the principal congruence subgroup

$$\Gamma_N = \{g \in GL(2,\mathbb{Z}) \mid g \equiv I_2 \pmod N\}$$

has no fixed points in $\mathfrak{H}^\pm$ for any integer $N \geq 3$. We can see this as follows: if $g(z) = z$ for some $z \in \mathfrak{H}^\pm$, then g is necessarily of finite order. If in addition g has integral coefficients, then $\mathbb{Z}[g]$ is isomorphic to a subring of the ring of algebraic integers in the field $\mathbb{Q}[g]$, which is necessarily of degree ≤ 2. Then g is a root of unity in a field of degree ≤ 2, hence a root of unity of order 2, 3, 4, or 6. Hence g cannot be congruent to 1 modulo N for $N \geq 3$.

It follows that the quotient $\Gamma_N\backslash\mathcal{E}$ is a family of elliptic curves over the open modular curve $Y(N) = \Gamma_N\backslash\mathfrak{H}^\pm$. The classical theory of modular forms shows that we can compactify $Y(N)$ to a projective modular curve $X(N)$, and in particular $Y(N)$ carries a natural structure of complex algebraic curve. Since Γ_N fixes the group $N^{-1}\mathbb{Z}^2/\mathbb{Z}^2$, the basis of points of order N in E_h defined by the generators $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ modulo N is fixed for all $hin\mathfrak{H}^\pm$. In this way we can show that $Y(N)$ is a connected component of the moduli space over $\mathbb{C}$ parametrizing pairs (E, α_N) where E is an elliptic curve and $\alpha_N : (\mathbb{Z}/N\mathbb{Z})^2 \xrightarrow{\sim} E[N]$ is a level N structure. This will later be treated correctly, in more generality, in the adelic setting, so I will not pursue the point here, especially since I have not explained what I mean by a moduli space. Instead I will explain how to generalize the constructions in pure linear algebra underlying the above discussion.

Definition. *Let V be a finite-dimensional vector space over $\mathbb{Q}$, $V_\mathbb{C} = V \otimes \mathbb{C}$. Let $w \in \mathbb{Z}$. A* Hodge structure *on V, pure of weight w, is a decomposition*

$$V_\mathbb{C} = \oplus_{p+q=w} V^{p,q}$$

such that $\bar{V}^{p,q} = V^{q,p}$. A Hodge structure *on V is a decomposition $V = \oplus_{w\in\mathbb{Z}} V_w$ of rational vector spaces, together with a pure Hodge structure of weight w on each V_w. Alternatively, it is a decomposition*

$$V_\mathbb{C} = \oplus V^{p,q}$$

such that $\bar{V}^{p,q} = V^{q,p}$ and such that, for each $w \in \mathbb{Z}$, the sum $\oplus_{p+q=w} V^{p,q}$ is the complexification of a rational subspace V_w of V.

Here $\bar{V}^{p,q}$ is the complex conjugate of $V^{p,q}$ for the $\mathbb{R}$-structure on $V_\mathbb{C}$, which is all that remains of the $\mathbb{Q}$-structure for the moment.

We can define a category (Hodge), whose objects are $\mathbb{Q}$-vector spaces V with Hodge structures, and whose morphisms are linear maps $V \to V'$ whose complexifications respect the p,q-decomposition. The direct sum of two objects in this category is again in this category. This is a simple construction in linear algebra.

Proposition. *There is an equivalence of categories between pairs $(V, h : \mathbb{C}^\times \to Aut(V_\mathbb{R}))$, where V is a rational vector space and h is a homomorphism of real algebraic groups, and Hodge structures.*

Proof. Let V be a rational vector space with a Hodge structure. Define $h : \mathbb{C}^\times \to Aut(V_\mathbb{C})$ by letting $h(z)$ act as $z^{-p}\bar{z}^{-q}$ on $V^{p,q}$ (this is the convention). One then has to check that h is a homomorphism of real algebraic groups. Concretely, this means that we need to check that $h(\bar{z}) = h(\bar{z})$, and this is obvious.

Given a pure Hodge structure of weight w, we can define an apparently coarser object, the *Hodge filtration*

$$V = F^a V \supset \cdots \supset F^j V \supset F^{j+1} V \cdots \supset F^b V = 0.$$

This is defined by setting

$$F^p V = \oplus_{p' \geq p} V^{p', w-p'}.$$

The individual $V^{p,q}$ are recovered using the formula

$$V^{p,q} = F^p V \cap \overline{F^q V},$$

which follows easily from the basic property of a Hodge structure.