Moduli spaces for abelian varieties over $\mathbf{C}$.

Kevin Buzzard

November 5, 2019

Last modified August 2005.

1 Introduction.

These are some notes I wrote in order to teach myself the classical analytic theory of moduli spaces for abelian varieties. They may well contain mistakes, and they might have a "lop-sided" feel because they emphasize only the parts of the theory that I wasn't so sure about. They contain lots of elementary calculations and a few examples of things. Perhaps the notes as a whole are of some use to other people nonetheless. Thanks to Toby Gee for comments on initial versions of this stuff.

References I've used for the general theory include: Swinnerton-Dyer's book on abelian varieties; Mumford's book on abelian varieties; and articles by Rosen and Milne in Cornell-Silverman. I also used various other more specialised references when, for example, trying to fathom out the Hilbert case and so on.

My goal was a concrete reference for all the well-known explicit bijections between the complex points of certain Shimura varieties, and isomorphism classes of certain abelian varieties equipped with extra structures (polarizations, endomorphisms, level structures (which as it happens will always be full here, due to laziness on my part))—up to and including the case of a quaternion algebra with exactly one split infinite place over a totally real field. I didn't quite get this far, but I got far enough to realise that I wasn't scared of this case.

The point is that knowing Riemann's theorem makes the whole business of moduli spaces very elementary over $\mathbf{C}$—it just turns the game into linear algebra. Unfortunately I couldn't face reading all those Shimura Annals papers[1], and didn't know any other references for all these standard facts, and so I thought I'd do the linear algebra myself, to help me appreciate some of the issues going on here.

2 Linear algebra and Riemann's theorem.

Let's start by recalling Riemann's theorem and some basic definitions.

[1] although I did end up reading some of some of them

A *piece of linear algebra data* is a pair (V, L), where V is a finite-dimensional complex vector space, of dimension g say, and L is a $\mathbf{Z}$-lattice in V—that is, a subgroup of V, isomorphic to $\mathbf{Z}^{2g}$, and such that the induced map $L \otimes \mathbf{R} \to V$ is an isomorphism. We say that g is the dimension of (V, L).

A *non-degenerate Riemann form* on a piece of linear algebra data (V, L) is a positive definite Hermitian sesquilinear (that is, linear in the first variable, and antilinear in the second) form H on V with a certain "rationality" property, which we now explain: For $v, w \in V$, we can write $H(v, w) = S(v, w) + iE(v, w)$ with S and E real-valued functions. One checks easily that E is an alternating bilinear form on V considered as a real vector space. The property we require is that E is $\mathbf{Q}$-valued on $L \times L$.

We say that a piece of linear algebra data (V, L) is *polarizable* if it admits a non-degenerate Riemann form.

Theorem 2.1 (Riemann)**.** *The category of abelian varieties over* $\mathbf{C}$ *is equivalent to the category of polarizable pieces of linear algebra data.*

Proof. Construct enough theta functions. See any of the references. $\square$

Maybe I should say what a morphism of linear algebra data is, so that this theorem makes sense: a map $(V, L) \to (V', L')$ is just a $\mathbf{C}$-linear map $V \to V'$ sending L into L', and hence inducing a map $V/L \to V'/L'$. Maybe also I should give the equivalence: to (V, L) one associates the abelian variety whose associated complex manifold is V/L. There is a lot more work to be done here, but I won't do it.

3 Non-degenerate Riemann forms, and polarizations.

I only understood non-degenerate Riemann forms after working through several examples. I will give several examples in a minute, but first I want to make some elementary observations about non-degenerate Riemann forms.

Let (V, L) be a piece of linear algebra data, and let H be a non-degenerate Riemann form on V. Write $H = S + iE$ as before. Note that one can reconstruct S and H from E, as follows. Note first that $S(iv, w) + iE(iv, w) = H(iv, w) = iH(v, w) = iS(v, w) - E(v, w)$ and hence $S(v, w) = E(iv, w)$. We deduce that $H(v, w) = E(iv, w) + iE(v, w)$.

Note next that the non-degeneracy of H implies the non-degeneracy of E on V considered as a real vector space: if $w \in V$ and $E(v, w) = 0$ for all $v \in V$ then $H(v, w) = E(iv, w) + iE(v, w) = 0$ for all $v \in V$ and hence $w = 0$. The same argument shows that E is a non-degenerate $\mathbf{Q}$-valued alternating form on $L \otimes \mathbf{Q}$.

Let (V, L) be a piece of linear algebra data, and let H be a non-degenerate Riemann form on V. Multiplying H by a positive rational number gives another non-degenerate Riemann form. Say that two non-degenerate Riemann forms on

V are *equivalent*, or $\mathbf{Q}_{>0}$-*equivalent*, if they differ by a positive rational factor in this way. A *polarization* of (V, L) is an equivalence class of such Hs.

Consider a polarizable (V, L), and fix a non-degenerate Riemann form $H = S + iE$ on V. Then E induces a map from $L \otimes \mathbf{Q}$ to its dual, which is an isomorphism by the non-degeneracy of E. We say that H is a *principal* Riemann form if E induces a bijection from L to its $\mathbf{Z}$-dual. We say that a polarization is *principal* if it contains a principal Riemann form.

Here are some examples. Recall that g is the dimension of V as a complex vector space. I remark that computing these examples gave me some sort of feeling for the kind of condition that the existence of a non-degenerate Riemann form imposes on a lattice.

If $g = 1$ then for any (V, L) there will be a non-degenerate Riemann form H: WLOG $V = \mathbf{C}$ and $L = \mathbf{Z} \oplus \mathbf{Z}\tau$; WLOG $\mathrm{Im}(\tau) > 0$. Then $H(z, w) = z\overline{w}/\mathrm{Im}(\tau)$ works. On the other hand, if $H(z, w) = zt\overline{w}$ is any non-degenerate Riemann form, then t is a positive real with the property $t\tau$ has rational imaginary part, and hence t is a positive rational multiple of $1/\mathrm{Im}(\tau)$. Hence, from the definitions, an elliptic curve has exactly one polarization, and it's principal. Note that even if E has CM and lots of automorphisms, there is only one polarization.

On the other hand, a dimension counting argument shows that if $g > 1$ then most (V, L) will admit no H. Let's make this explicit for $g = 2$. Set $V = \mathbf{C}^2$.

Non-example 1) (write down general L and check no H). Let L be the lattice spanned by $e_1 = (1, 0)$, $e_2 = (0, 1)$, $e_3 = (i, 0)$ and $e_4 = (i\sqrt{2}, i\sqrt{3})$. Claim: no H exists. Proof: Assume it did. Then H would be represented by a hermitian matrix $\left(\begin{smallmatrix} a & b+ic \\ b-ic & d \end{smallmatrix}\right)$ with a, b, c, d real. Now $E(e_1, e_3)$ and $E(e_2, e_3)$ in $\mathbf{Q}$ implies that a and b are rational. But $E(e_1, e_4) \in \mathbf{Q}$ implies that $\sqrt{2}a + \sqrt{3}b \in \mathbf{Q}$, and this must imply $a = b = 0$, and hence this hermitian matrix suddenly has determinant $-c^2 \leq 0$ and can't be positive definite.

Example 2) I got this by writing down $H = \left(\begin{smallmatrix} \sqrt{2} & 1+i \\ 1-i & 2 \end{smallmatrix}\right)$, working out a "general" L for which this H worked, and then forgetting H.

Let L be the lattice spanned by $e_1 = (1, 0)$, $e_2 = (0, 1)$, $e_3 = (i, \sqrt{2})$ and $e_4 = (x + iy, u + iv)$ with $x = -\sqrt{3} - 2\sqrt{2}$, $y = \sqrt{3}$, $u = \sqrt{2} + \sqrt{6}$ and $v = 1 + \sqrt{2}$. What can H be? We set $H = \left(\begin{smallmatrix} a & b+ic \\ b-ic & d \end{smallmatrix}\right)$ with a, b, c, d real, and write $H = S + iE$. From $H(e_i, e_j)$ with $i, j \leq 3$ we deduce that b, c are rational, and that $a - c\sqrt{2}$ is too. Now rationality of $E(e_1, e_4)$ tells us that $\sqrt{3}a + (1 + \sqrt{2})b - (\sqrt{2} + \sqrt{6})c$ is rational. This looks messy, but recall that b and c are rational, so this implies that $\sqrt{3}(a - c\sqrt{2})$ is in the field $\mathbf{Q}(\sqrt{2})$, and we already know that it's in the line $\sqrt{3}\mathbf{Q}$ and hence it must be 0. So $a = c\sqrt{2}$. Now tidying up we can deduce more: we see that $(\sqrt{2})(b - c)$ is rational, but we know $b - c$ is too, and hence $b = c$.

We now go on to look at $H(e_2, e_4)$ and $H(e_3, e_4)$. Recall that we know that c has to be rational. Rationality of $E(e_2, e_4)$ tells us that $2\sqrt{2}c - (1 + \sqrt{2})d$ is rational, and rationality of $E(e_3, e_4)$ tells us that $(1 + 2\sqrt{2})c - (2 + \sqrt{2})d$ and hence $2\sqrt{2}c - (2 + \sqrt{2})d$ is rational. Taking the difference, we deduce that d is rational. We deduce that $\sqrt{2}(2c - d)$ is rational and so $d = 2c$; things are now looking hairy, but in fact fortunately $d = 2c$ happens to satisfy all the remaining rationality constraints!

We deduce that H can only be a positive rational multiple of $\begin{pmatrix} \sqrt{2} & 1+i \\ 1-i & 2 \end{pmatrix}$ but conversely this does work.

My feeling is that this is what should happen for the "generic" polarizable (V, L): there will be precisely one polarization. My procedure to find such (V, L) gives quite messy numbers though.

Example 3) (a "nice" lattice). This is misleading. Let L be spanned by $e_1 = (1, 0)$, $e_2 = (0, 1)$, $e_3 = (\sqrt{2}i, 0)$ and $e_4 = (0, \sqrt{3}i)$. Then C/L is a product of two elliptic curves. If $H = \begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix}$ then our six equations $E(e_i, e_j) \in \mathbf{Q}$ give that c, $\sqrt{2}a$, $\sqrt{3}b$, $\sqrt{2}b$, $\sqrt{3}d$ and $\sqrt{6}c$ are rational, so $b = 0 = c$ but a and d can vary over positive elements of $\sqrt{2}\mathbf{Q}$ and $\sqrt{3}\mathbf{Q}$ respectively, giving a "1-dimensional space" of polarizations. The point is that both elliptic curves have a polarization, and one can choose their ratio arbitrarily, as it were.

Note that this example demonstrates that the product of two polarized abelian varieties doesn't inherit a canonical polarization! The point is that if one chooses non-degenerate Riemann forms on both factors, one gets a form on the product, but both forms are only defined up to a rational factor, and one can't control both of them at once.

Example 4) (quite a lot of symmetry). Let $\tau = x + iy$ be a generic element of the upper half plane, e.g. $e + i\pi$, and let L be spanned by $(1, 0)$, $(0, 1)$, $(\tau, 0)$ and $(0, \tau)$. So V/L is the square of a non-CM elliptic curve. Then we see that $c, ay, cx - by, cx + by, dy$ and $c(x^2 + y^2)$ must all be rational. Hence $c = 0$ and a, b, d are all in $\mathbf{Q}/y$. Multiplying up, we see that a polarization here is just $(1/y)$ times a positive definite symmetric element of $\mathrm{M}_2(\mathbf{Q})$, defined up to a positive rational.

Example 5) (too much symmetry). Let L be spanned by $(1, 0)$, $(0, 1)$, $(i, 0)$ and $(0, i)$. Then we see that a, b, c, d have to be rational and the only constraint is that a, d and $ad - b^2 - c^2$ are positive. A is the square of an elliptic curve here. Some calculations I did made me realise that the correct way to think of this, analogous to example 4 above, is that H can be any positive definite Hermitian two by two matrix with coefficients in $\mathbf{Q}(i)$.

4 The dual abelian variety.

Recall that the dual of an abelian variety (i.e., of a polarizable (V, L)) can be thought of as (V^*, L^*) where V^* is the $\mathbf{R}$-linear maps $f : V \to \mathbf{C}$ such that $f(zv) = \overline{z}f(v)$ for $v \in V$ and $z \in \mathbf{C}$, and L^* is the elements f of V^* such that $\mathrm{Im}(f(l))$ is integral for all $l \in L$. Note that if $f \in V^*$ then f is determined by $\mathrm{Im}(f)$, and conversely any real linear map $e : V \to \mathbf{R}$ is the imaginary part of a unique element f of V^*, defined by $f(v) = e(-iv) + ie(v)$. Note also that the induced pairing $L \times L^* \to \mathbf{Z}$ is a perfect pairing.

Endow V^* with the obvious complex structure: $(\lambda f)(v) = \lambda(f(v))$. Note also that (V^*, L^*) is indeed polarizable: choose a non-degenerate Riemann form H on V; then the map $t \mapsto H(t, .)$ is an isomorphism of vector spaces $\phi : V \to V^*$, and one can define H^* on V^* by $H^*(f,g) = H(\phi^{-1}(f), \phi^{-1}(g))$. Note that $\phi(L)$ is commensurable with L^* (that is, their intersection has finite index in both) and that this is enough to ensure that H^* is a non-degenerate Riemann form on (V^*, L^*).

An elementary remark: The dual of the dual of A is just A again. This is tricky in characteristic p but is formal over $\mathbf{C}$: the only catch is that the isomorphism $V \to V^{**}$ which is $\mathbf{C}$-linear is the one that sends $v \in V$ to the map $V^* \to \mathbf{C}$ sending f to $\overline{f(v)}$.

If $\alpha : V \to W$ is a $\mathbf{C}$-linear map of vector spaces, then we get a dual map $\alpha^* : W^* \to V^*$ where $*$ denotes those antilinear spaces as above. The definition is the obvious one: $(\alpha^*(f))(v) = f(\alpha(v))$ and check this works. Hence one checks that for a map of abelian varieties $A \to B$ one gets a dual map $B^* \to A^*$.

If H on (V, L) happens to be $\mathbf{Z}$-valued on $L \times L$ then ϕ as above induces an isogeny $V/L \to V^*/L^*$. In this way, a polarization on $A = V/L$ can be thought of as a certain kind of isogeny from A to its dual, apart from the initially annoying fact that the isogeny is only defined "up to a positive rational factor". We can make this a bit more conceptual by introducing one facet of the category of abelian varieties up to isogeny—that of tensoring all our endomorphism rings with $\mathbf{Q}$.

Formally, let's define the abelian group $\mathrm{Hom}^0(A, B) = \mathrm{Hom}(A, B) \otimes \mathbf{Q}$, and the ring $\mathrm{End}^0(A) = \mathrm{Hom}^0(A, A)$. In this optic, a polarization is a half-line in $\mathrm{Hom}^0(A, A^*)$, that is, something of the form $\mathbf{Q}_{>0}v$, although not every half line is a polarization (consider a CM elliptic curve to see lots of examples of half-lines that aren't polarizations). Note that the half-line determines the polarization, because if $\phi : V \to V^*$ is an element of the half-line, then $H(v, w) = (\phi(v))(w)$.

A morphism of polarized abelian varieties: it really is the following slightly surprising thing. Think of a polarization as being a map from A to its dual, defined up to a positive rational. Then a morphism of polarized abelian varieties $m : A \to B$ will have to be one for which a diagram commutes, and that diagram is quite strange: first go from A to B via m, and then from B to its dual via the polarization, and then back from B^* to A^* via the dual of m. The resulting map must be the polarization of A, where we recall that polarizations are only defined up to a positive rational factor, so we are really asking for an equality of two half-lines in the $\mathbf{Q}$-endomorphism ring. This all implies that m has finite kernel, which is perhaps a bit strange at first sight: it means that things like projection maps and zero maps aren't morphisms of polarizations of abelian varieties.

Elementary exercise: if $\phi : A \to A^*$ is the morphism induced by a polarization, then $\phi^* = \phi$ (once one has identified A^{**} with A canonically, and used the remark above about how this is done!).

One key thing that we will use later, when we move onto endomorphisms, is the Rosati involution. If H is a non-degenerate Riemann form on (V, L) then we get an element ϕ of $\mathrm{Hom}^0(A, A^*)$ defined up to a positive rational. This element has an inverse $\phi^{-1} \in \mathrm{Hom}^0(A^*, A)$. Now if $\alpha \in \mathrm{End}^0(A)$ then $\alpha^\iota := \phi^{-1}\alpha^*\phi \in \mathrm{End}^0(A)$ too. The involution, or rather antiautomorphism, $\alpha \mapsto \alpha^\iota$, is called the Rosati involution on $\mathrm{End}^0(A)$. Note that if one changes ϕ by a positive rational, one does not change the involution, and hence the involution depends only on the polarization induced by H. On the other hand, different polarizations can induce different involutions. For example, in Example 4 of a polarizable (V, L) above, we had a non-CM elliptic curve $E = \mathbf{C}/\langle 1, \tau\rangle$ with $\tau = x + iy$ in the upper half plane; then a polarization on $E \times E$ was a positive definite symmetric matrix in $(1/y)\,\mathrm{M}_2(\mathbf{Q})$. Let's work out everything explicitly here.

Let $V = \mathbf{C}^2$, and let L be the lattice as in Example 4 above. An element f of V^* is determined by $f(1, 0) = s$ and $f(0, 1) = t$, which can be arbitrary complex numbers; and let's think of (s, t) as giving an isomorphism $V^* = \mathbf{C}^2$. Explicitly, if $(s, t) \in \mathbf{C}^2$, then the corresponding $f \in V^*$ is given by $f(z, w) = \overline{z}s + \overline{w}t$. The lattice $L^* \subset V^*$ is the functions such that $f(1, 0)$ and $f(\tau, 0)$ and $f(0, 1)$ and $f(0, \tau)$ all have integral imaginary parts, so s and $(x - iy)s$ and t and $(x - iy)t$ all have integral imaginary parts. One checks easily that this is equivalent to $s, t \in (\mathbf{Z} \oplus \mathbf{Z}\tau)/y$. So, surprise surprise, V/L is canonically self-dual, although we won't use this.

Let's take $H = M$ with My a two by two symmetric positive definite element of $\mathrm{M}_2(\mathbf{Q})$. Then H induces a polarization on A. The induced map $V \to V^*$ sends (u, v) to the f defined by the pair (s, t) as above, where $\binom{s}{t} = M\binom{u}{v}$. So ϕ is represented by M. Note that this map ϕ sends $L \otimes \mathbf{Q}$ isomorphically onto $L^* \otimes \mathbf{Q}$. Now one has to work out the dual of ϕ, but it's a fact that ϕ is self-dual, as all polarizations are. Finally, one has to work out explicitly the dual of a map $V \to V$ given by a matrix P; it turns out to be the conjugate of the transpose of P.

In summary then: $\mathrm{End}^0(A) = \mathrm{M}_2(\mathbf{Q})$ and the Rosati involution induced by $H = M$ is the map $P \mapsto M^{-1}P^tM$. Note that this only depends on $\mathbf{Q}M$ but on the other hand it does depend on this; distinct choices of M do give distinct Rosati involutions. On the other hand, it's not too difficult to give examples of distinct polarizations which induce the same Rosati—for example if A is the product of two non-isogenous non-CM curves, or an abelian surface whose endomorphism ring is the integers in a real quadratic field, then all the (infinitely many) polarizations induce the same (trivial) Rosati.

Here's a pleasant exercise, which probably implies what we've just done above: let (V, L) be polarizable, let H be a non-degenerate Riemann form, and let $^\iota$ be the induced Rosati. Then $H(\alpha v, w) = H(v, \alpha^\iota w)$. So this Rosati business is just the adjoint with respect to H.

5 The elliptic curve case.

We barely indicate the proof here; this is the one case that one can find in lots of places. The only reason I'm saying anything at all is that I will give two constructions of the bijection between elliptic curves and $\mathrm{GL}_2(\mathbf{Z})\backslash(\mathbf{C}\backslash\mathbf{R})$; one obtained by fixing the complex structure and varying the lattice, the other by fixing the lattice and varying the complex structure.

The first construction goes as follows. If A is an elliptic curve, $A = \mathbf{C}/L$, with WLOG $L = \mathbf{Z} \oplus \mathbf{Z}\tau$, with $\tau \in \mathbf{C}\backslash\mathbf{R}$. Two such curves are isomorphic if the corresponding lattices are homothetic, and the usual calculation shows that this is equivalent to the two τ's being in the same $\mathrm{GL}_2(\mathbf{Z})$-orbit.

The second construction looks more complicated, but in fact seems to be the one to generalise when the polarization is not totally canonical: one "fixes L" and the alternating form E, and demands conditions on the complex structure on $L\otimes\mathbf{R}$ that make H positive definite. In general, it's easier to do this than it is to demand conditions on a lattice that guarantee the existence of a Hermitian H. The motivation behind this construction, and a lot of the details, are in the next section, which makes it much clearer why such a construction is of use.

So here's the second construction: consider an elliptic curve as $\mathbf{R}^2/\mathbf{Z}^2$ with a complex structure, that is, $I : \mathbf{R}^2 \to \mathbf{R}^2$ such that $I^2 = -1$. The set of such I is just the matrices in $\mathrm{GL}_2(\mathbf{R})$ whose square is -1; if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is such a matrix then $\gamma^2 + 1 = 0$ must be the minimal polynomial of γ, so $\det(\gamma) = 1$ and $\mathrm{tr}(\gamma) = 0$. Conversely any matrix with trace zero and determinant 1 gives us an I. One now checks that γ can be written $\delta\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\delta^{-1}$; the element $\delta i \in (\mathbf{C}\backslash\mathbf{R})$ can be checked to be well-defined (see Lemma 6.7), and two elements of $\mathbf{C}\backslash\mathbf{R}$ give the same elliptic curve iff the corresponding γ's are conjugate by an element of $\mathrm{GL}_2(\mathbf{Z})$, which is iff the elements of $\mathbf{C}\backslash\mathbf{R}$ are in the same $\mathrm{GL}_2(\mathbf{Z})$-orbit. It is this argument which we are about to generalise.

6 Symplectic groups and Siegel space.

Let $g \geq 1$ be an integer. We habitually write $2g$ by $2g$ matrices using the notation $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with A, B, C, D g by g matrices. Let J denote the $2g$ by $2g$ matrix $\left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$.

Definition 6.1. *If R is a ring, then $\mathrm{GSp}_{2g}(R)$ is the $2g$ by $2g$ invertible matrices $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ over R such that $MJM^t = \lambda(M)J$, with $\lambda(M) \in R^\times$.*

I claim that these matrices form a group, and λ is a group homomorphism to $\mathrm{GL}_1(R)$. The fact that $\mathrm{GSp}_{2g}(R)$ is closed under products and inverses, and that λ preserves multiplication, is an easy formal verification. Almost as easy is the fact that if M is in $\mathrm{GSp}_{2g}(R)$ then so is M^t, and that $\lambda(M) = \lambda(M^t)$. This follows because $M^tJMJM^t = \lambda(M)M^tJ^2 = -\lambda(M)M^t = \lambda(M)J^2M^t$ and M^t is invertible, so $M^tJM = \lambda(M)J$.

Definition 6.2. *The subgroup $\mathrm{Sp}_{2g}(R)$ of $\mathrm{GSp}_{2g}(R)$ is the kernel of λ, that is, the invertible matrices M such that $MJM^t = J$.*

The arguments above show that if $M \in \mathrm{Sp}_{2g}(R)$ then so is M^t. One checks easily that $J \in \mathrm{Sp}_{2g}(R)$. Note however that, contrary to what I first expected, if $M \in \mathrm{GSp}_{2g}(R)$ then $\lambda(M)$ is not necessarily the determinant of M. For example, if $M = \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right)$ then $\lambda(M) = -1$ but $\det(M) = (-1)^g$. In fact, if $\gamma \in \mathrm{GSp}_{2g}(R)$ then we have $\det(\gamma) = \lambda(\gamma)^g$ always. There is a proof of this in Liebeck's book on finite simple groups that works in all characteristics, but I'll sketch a characteristic 0 proof: set $V = R^{2g}$; think of J as being a map $\Lambda^2_R(V) \to R$, and γ as a map $V \to V$. Then γ induces a map from $\mathrm{Hom}(\Lambda^2_R(V), R)$ to itself, and it sends J to $\lambda(\gamma)J$. Now $\Lambda^g(J) \in \Lambda^{2g}(V)$ is non-zero and γ multiplies it by $\lambda(\gamma)^g$, so γ must induce multiplication by $\lambda(\gamma)^g$ on $\Lambda^{2g}(V)$ and hence $\det(\gamma) = \lambda(\gamma)^g$.

The proof in characteristic p proves first that Sp is generated by certain "elementary" matrices and hence everything in it has determinant 1; then it bootstraps itself up. The proof that elementary matrices generate Sp_{2g} is done by induction on g, proving that every symplectic matrix can be multiplied by certain elementary matrices so that the product fixes a certain hyperplane, and then using the inductive hypothesis. I've not read the details.

If $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is in $\mathrm{GSp}_{2g}(R)$ (resp. $\mathrm{Sp}_{2g}(R)$ then elementary messing around with the definition shows that AB^t and CD^t are symmetric, and that $AD^t - BC^t = \lambda(M)I$ (resp. I); conversely, these equations are necessary and sufficient for M to be in $\mathrm{GSp}_{2g}(R)$ (resp. $\mathrm{Sp}_{2g}(R)$). One deduces easily from these equations that the inverse of $M \in \mathrm{GSp}_{2g}(R)$ is $\lambda(M)^{-1}\begin{pmatrix} D^t & -B^t \\ -C^t & A^t \end{pmatrix}$.

Recall that $M \in \mathrm{GSp}_{2g}(R)$ iff $M^t \in \mathrm{GSp}_{2g}(R)$, and hence $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is also in $\mathrm{GSp}_{2g}(R)$ (resp. $\mathrm{Sp}_{2g}(R)$) iff A^tC and B^tD are symmetric, and $A^tD - C^tB = \lambda(M)I$ (resp. I). This did my head in a bit when I first saw it, because the order of B and C has swapped, but not the order of A and D—but on the other hand, this is exactly what happens when one takes the transpose of M so it's not that surprising.

Definition 6.3. Siegel space $\mathbf{H}_g$ *is the set of g by g complex matrices $\Omega = X + iY$ such that $\Omega^t = \Omega$ and Y is either positive definite or negative definite (we call such a matrix a* definite *matrix).* $\mathbf{H}_g^+$ *is the subset of $X + iY$ such that Y is positive definite. Letting $X = 0$ and $Y = I$ gives the scalar matrix $i \in \mathbf{H}_g^+$, which we shall refer to as i or i_g; this should not be confused with the identity matrix I.*

One useful fact:

Lemma 6.4. *If Y is a symmetric real g by g matrix, then the real symmetric bilinear form on $\mathbf{R}^g$ represented by Y is positive definite if and only if the Hermitian sesquilinear form on $\mathbf{C}^g$ represented by Y is positive definite.*

Proof. If Y is positive definite in the real sense, then $P^tYP = I_g$ for some real invertible P, and hence $P^tY\overline{P} = I_g$ and so Y is positive definite in the complex sense. Conversely, if Y is positive definite in the complex sense, then $v^tY\overline{v} > 0$ for all non-zero $v \in \mathbf{C}^g$ then certainly $v^tYv > 0$ for all non-zero $v \in \mathbf{R}^g$ and so Y is positive definite in the real sense. □

Definition 6.5. *We define an action of* $\mathrm{GSp}_{2g}(\mathbf{R})$ *on* $\mathbf{H}_g$ *thus: $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ sends Ω to $(A\Omega + B)(C\Omega + D)^{-1}$. The same definition defines an action of subgroup* $\mathrm{Sp}_{2g}(\mathbf{R})$ *on* $\mathbf{H}_g^+$.

We have to check a lot before we see that this last definition makes sense. For $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{GSp}_{2g}(\mathbf{R})$ and $\Omega = X + iY \in \mathbf{H}_g$, define $M = \overline{(C\Omega + D)}^t(A\Omega + B)$.

An easy check, remembering that Ω is symmetric and A, B, C, D are real, shows that $\overline{M}^t = (B^t + \overline{\Omega}A^t)(C\Omega + D)$, and hence (multiply out) that $M - \overline{M}^t = \lambda(\gamma)(\Omega - \overline{\Omega}^t) = 2i\lambda(\gamma)Y$. Now if $v \in \mathbf{C}^g$ and $(C\Omega + D)v = 0$, we deduce that $\overline{v}^t M$ and $\overline{M}^t v$ are zero, and hence $\overline{v}^t(M - \overline{M}^t)v = 0$. Hence $v^t Y \overline{v} = 0$. But $\pm Y$ represents a positive definite Hermitian form by the previous lemma, so we deduce $v = 0$. This shows that $C\Omega + D$ has zero kernel, and hence $C\Omega + D$ is invertible.

So the rule gives a well-defined matrix $(A\Omega + B)(C\Omega + D)^{-1}$. We must next check that this matrix is still in Siegel space, which boils down to checking symmetry and definiteness. Symmetry boils down to checking that $(B^t + \Omega A^t)(C\Omega + D) = (\Omega C^t + D^t)(A\Omega + B)$. Multiplying out and using lots of consequences of the fact that $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ is in $\mathrm{GSp}(\mathbf{R})$ now gives it to us. It remains to prove that the imaginary part of $\gamma\Omega$ is definite. The trick here is to go back to M, which, now $\gamma\Omega$ is well-defined, we may now write as $\overline{(C\Omega + D)}^t \gamma\Omega(C\Omega + D)$. We have seen that $M - \overline{M}^t = 2i\lambda(\gamma)Y$, but now writing $\gamma\Omega = S + iT$ we see that it also equals $\overline{(C\Omega + D)}^t(2iT)(C\Omega + D)$. Hence $\overline{C\Omega + D}^t T(C\Omega + D) = \lambda(\gamma)Y$ and this means that T and $\lambda(\gamma)Y$ represent equivalent Hermitian forms. We know Y is definite as a Hermitian form, hence T is too, so T is definite as a real form.

Finally, we check that it's an action y multiplying everything out. One has to check that $((AE + BG)\Omega + (AF + BH))((CE + DG)\Omega + (CF + DH))^{-1} = (A(E\Omega + F)(G\Omega + H)^{-1} + B)(C(E\Omega + F)(G\Omega + H)^{-1} + D)^{-1}$. Using the fact that $XY^{-1} = XZ(YZ)^{-1}$ on the right hand side with $Z = (G\Omega + H)$, it now all drops out. Note that one didn't use the fact that any matrices were symplectic for this last formal verification.

That does everything for GSp, and for Sp it's all the same with $\lambda = 1$ all the way through.

Lemma 6.6. *The action of* $\mathrm{Sp}_{2g}(\mathbf{R})$ *on* $\mathbf{H}_g^+$ *is transitive, and the action of* $\mathrm{GSp}_{2g}(\mathbf{R})$ *on* $\mathbf{H}_g$ *is too.*

Proof. For the first part, all I have to do is to make sure that for any $\Omega = X + iY \in \mathbf{H}_g$ there is $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ in $\mathrm{Sp}_{2g}(\mathbf{R})$ such that $A\Omega + B = i(C\Omega + D)$, because i is in $\mathbf{H}_g$. So I have to solve $AX + B = -CY$ and $AY = CX + D$. This is easy though: Y is symmetric and positive definite, so there exists a change of basis such that the corresponding bilinear form is represented by the identity matrix. Hence there exists A such that $AYA^t = I_g$. Now one easily checks that $M = \left(\begin{smallmatrix} A & -AX \\ 0 & AY \end{smallmatrix}\right)$ is in $\mathrm{Sp}_{2g}(\mathbf{R})$ and works. For the GSp case, one uses the same trick but perhaps letting AYA^t be $-I_g$. □

Define a map h_0 from $\mathbf{C}^*$ to $\mathrm{GSp}_{2g}(\mathbf{R})$ by sending $x + iy$ to $\left(\begin{smallmatrix} x & y \\ -y & x \end{smallmatrix}\right)$. Note that $h_0(i) = J$. Deligne talks about the "conjugacy class of h_0" but I'll just talk about the conjugacy class of $h_0(i)$, which boils down to the same things in the cases I look at.

Lemma 6.7. *The conjugacy class of* J *is in natural bijection with* $\mathbf{H}_g$*.*

Proof. If $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathrm{GSp}_{2g}(\mathbf{R})$, let's send MJM^{-1} to $Mi_g \in \mathbf{H}_g$. We need to check that this is well-defined. This boils down to checking that if an element of $\mathrm{GSp}_{2g}(\mathbf{R})$ commutes with J then it fixes i_g. One checks that $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ commutes with J iff $A = D$ and $B = -C$, which is true iff $Mi = i$. Hence the map is well-defined, and running the argument backwards shows that if $M \in \mathrm{GSp}_{2g}(R)$ fixes i_g then it commutes with J, which now easily gives injectivity. Finally, surjectivity is clear because the action is transitive by Lemma 6.6. □

Next we note

Lemma 6.8. *If V is a finite-dimensional complex vector space and E is an alternating real bilinear form on V, then there is at most one Hermitian sesquilinear form with imaginary part E; such a form exists iff $E(iv, iw) = E(v, w)$, and if it exists then it's non-degenerate iff E is non-degenerate.*

Proof. If H exists then $H(v, w) = H(iv, iw)$ so the condition on E is necessary for existence of H. Also, $\mathrm{Im}(H(v, w)) = E(v, w)$ and $\mathrm{Re}(H(v, w)) = \mathrm{Im}(iH(v, w)) = E(iv, w)$ so the definition of H is forced upon us: $H(v, w) = E(iv, w) + iE(v, w)$. Say $E(iv, iw) = E(v, w)$. We must check that the H defined above is Hermitian and sesquilinear. Well, H is $\mathbf{R}$-bilinear, $H(iv, w) = E(-v, w) + iE(iv, w) = iH(v, w)$ and $H(v, iw) = E(iv, iw) + iE(v, iw) = E(v, w) + iE(-iv, w) = -iH(v, w)$. Finally $H(w, v) = E(iw, v) + iE(w, v) = -E(v, iw) - iE(v, w) = -E(-iv, w) - iE(v, w) = E(iv, w) - IE(v, w) = \overline{H(v, w)}$ so H is indeed Hermitian and sesquilinear.

Finally, if $H(v, w) = 0$ for all v then $E(v, w) = 0$ for all v, so non-degeneracy of E implies $w = 0$ implies non-degeneracy of H. Conversely, if $E(v, w) = 0$ for all v then $H(v, w) = E(iv, w) + iE(v, w) = 0$ for all v, so non-degeneracy of H implies non-degeneracy of E. □

Before we state and prove the main theorem in this section, we need a lemma which must be implicit in Deligne's "Travaux de Shimura", last paragraph of §1.6, but whether he is just leaving the proof to the reader or is thinking about things in such a conceptual way that for him the lemma needs no proof, I don't know.

Lemma 6.9. *The matrices $M \in \mathrm{Sp}_{2g}(\mathbf{R})$ with the following properties:*
(i) $M^2 = -I$
(ii) M^tJ (which is symmetric by (i)) is definite
are precisely the conjugates of J in $\mathrm{GSp}_{2g}(\mathbf{R})$.

Proof. Firstly we'll justify that if M is as in the lemma then M^tJ is symmetric: the point is that $M^tJ = M^tJMM^{-1} = JM^{-1} = -JM = J^tM$.

Next we'll show that $M = \gamma J\gamma^{-1}$ works: certainly this is in $\mathrm{Sp}_{2g}(\mathbf{R})$ and its square is -1; for definiteness we note that $M^tJ = -JM = -J\gamma J\gamma^{-1}$ and $\gamma \in \mathrm{GSp}_{2g}(\mathbf{R})$ so $\gamma^tJ = \lambda(\gamma)J\gamma^{-1}$. Hence $M^tJ = -\lambda(\gamma)J\gamma\gamma^tJ = \lambda(\gamma)(J\gamma)(J\gamma)^t$ is indeed definite.

Finally we check that a matrix M with the above property is a conjugate of J in $\mathrm{GSp}_{2g}(\mathbf{R})$. Write $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$. We have $M \in \mathrm{Sp}_{2g}(\mathbf{R})$ and hence

$-M = M^{-1} = \left(\begin{smallmatrix} D^t & -B^t \\ -C^t & A^t \end{smallmatrix}\right)$. So $A = -D^t$ and B, C are symmetric. Next note that $M^tJ = \left(\begin{smallmatrix} -C^t & -D^t \\ A^t & B^t \end{smallmatrix}\right)$ is definite, and hence $-C$ is symmetric and definite. Conjugating M by $\left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right)$ if necessary, we may assume that $-C$ is positive definite. Now choose $H \in \mathrm{GL}_g(\mathbf{R})$ with $HH^t = -C$, and define $E = (H^t)^{-1}$ and $F = -AE$. Set $\gamma = \left(\begin{smallmatrix} E & F \\ 0 & H \end{smallmatrix}\right)$. Then $EF^t = C^{-1}A^t$ is symmetric, as A^tC is, and $EH^t = I$, so $\gamma \in \mathrm{Sp}_{2g}(\mathbf{R})$. Hence $\gamma^{-1} = \left(\begin{smallmatrix} H^t & -F^t \\ 0 & E^t \end{smallmatrix}\right)$ and so $\gamma J \gamma^{-1} = \left(\begin{smallmatrix} -FH^t & EE^t+FF^t \\ -HH^t & HF^t \end{smallmatrix}\right) = \left(\begin{smallmatrix} A & B' \\ C & D \end{smallmatrix}\right)$ with $B' = -C^{-1} - AC^{-1}A^t$. We are done if we can prove $B' = B$ but this follows immediately from the fact that $AD^t - B'C^t = I$. □

Finally, we are ready to prove

Theorem 6.10. *There is a bijection between the set of isomorphism classes of principally polarized abelian varieties of dimension g, and* $\mathrm{GSp}_{2g}(\mathbf{Z})\backslash\mathbf{H}_g$.

Proof. By Riemann's theorem, the left hand side is isomorphism classes of (V, L) which admit a principal Riemann form H. Let's vary the complex structure on V instead; we can choose a basis of L such that E is represented by J and now we have to parameterise complex structures h on $\mathbf{Z}^{2g} \otimes \mathbf{R}$ such that

(i) $E(h(i)z, h(i)w) = E(z, w)$,

(ii) The induced H (see above lemma) is definite,

modulo isomorphism. Now h is determined by the matrix $M = h(i)$, which must be in $\mathrm{Sp}_{2g}(\mathbf{R})$ by (i). The induced H will be definite iff M^tJ is definite. By Lemma 6.9, such M are precisely the $\mathrm{GSp}_{2g}(\mathbf{R})$-conjugates of $h_0(i)$. Hence the complex structure h gives us an abelian variety iff h is conjugate to h_0 in $\mathrm{GSp}_{2g}(\mathbf{R})$ (that is, iff $h(i)$ and $h_0(i)$ are conjugate). By Lemma 6.7, the set of such h is naturally in bijection with $\mathbf{H}_g$. Now "modulo isomorphism" means "modulo a change of basis of L" which means "modulo conjugation by an element of $\mathrm{GSp}_{2g}(\mathbf{Z})$". By the construction of the bijection between the conjugacy classes and $\mathbf{H}_g$, conjugation on the conjugacy class side corresponds to left action on the $\mathbf{H}_g$ side, and now we are done. □

We remark that we could have just parameterised positive definite H, in which case all the GSp's change to Sp's and all the $\mathbf{H}$'s change to $\mathbf{H}^+$'s and the proof remains the same.

7 Endomorphisms.

Before we can go any further, we must understand the "philosophy" of endomorphisms of polarized abelian varieties. In a nutshell, it appears to be the following: if A is a polarized abelian variety, then there is an induced involution, the Rosati involution, on $\mathrm{End}^0(A)$ (I defined this above, in the section on the dual abelian variety). The philosophy seems to be that if R is a semi-simple $\mathbf{Q}$-algebra equipped with an antiautomorphism, then an action of R on a polarized abelian variety A should be an injection $R \to \mathrm{End}^0(A)$ such that

Rosati preserves the image of R and induces the given antiautomorphism on it. In fact, one does something slightly more delicate than this: one also fixes a g-dimensional complex representation of R and demands that this is the induced representation of R on the tangent space V of A. Actually, one even does a little more: it seems that somehow one wants to take some kind of maximal order in R and only look at abelian varieties on which this order actually acts, as opposed to acting via isogeny.

However, from personal experience, I know that all this will become clearer with some examples, many of which are forthcoming. Before that, let's recall what possibilities there are for the endomorphism ring $\mathrm{End}^0(A)$, A an abelian variety (up to isogeny). This is quite messy, and I'm finding it difficult to summarise. On the other hand, it can be skipped on first reading, apart from the definition at the end.

Note first that if A is an abelian variety, then up to isogeny it's a product of powers of simple abelian varieties $A_1^{e_1} \times A_2^{e_2} \times \ldots$, with the A_i non-isogenous. In this case, $\mathrm{End}^0(A) = \mathrm{M}_{e_1}(\mathrm{End}^0(A_1)) \times \mathrm{M}_{e_2}(\mathrm{End}^0(A_2)) \times \ldots$ where one has to be careful talking about matrices over non-commutative rings, but it's all OK. Hence it suffices to compute the possibilities for $\mathrm{End}^0(A)$ with A simple (i.e., not containing any non-trivial sub-abelian-varieties).

Let A be simple. The first obvious fact is that any non-zero map $A \to A$ must then be an isogeny, and hence invertible in $D := \mathrm{End}^0(A)$. So D is a skew field. Next note that if $A = V/L$ then D acts faithfully on $L \otimes \mathbf{Q}$ and hence D is a finite-dimensional $\mathbf{Q}$-algebra. In fact $L \otimes \mathbf{Q}$ is a vector space for D and so the $\mathbf{Q}$-dimension of D divides $2g$. This fails in positive characteristic, by the way; one only has the l-adic representations for l not p, and this isn't quite enough.

Let K denote the centre of D. Then K contains $\mathbf{Q}$ and now D is a K-algebra. K is a commutative skew-field, so it's a number field, and D is now a division algebra over K, so it has dimension d^2 over K. Swinnerton-Dyer talks about such things in his lecture notes for a 1995 course I went to, and no doubt it's all in his book too.

Recall the notion of the "reduced trace". There is a canonical trace map $D \to K$: one can define it by going up to the algebraic closure of K, where D becomes a matrix algebra, and then taking the usual trace. Now compose this with the trace map from K down to $\mathbf{Q}$, and one gets something called the *reduced trace of D over* $\mathbf{Q}$ (by Mumford, at least, on the amazing p180 of his book).

Now let's choose a polarization on A, still assumed simple. The corresponding Rosati endomorphism has the property that for any non-zero $x \in \mathrm{End}^0(A)$, the reduced trace of $xx^\iota \in \mathrm{End}^0(A)$ is positive. This is standard and I'll skip the proof—see Mumford, for example.

We are fortunate in that Albert classified all finite-dimensional division algebras D over $\mathbf{Q}$ equipped with antiautomorphisms satisfying this positivity condition. Let's stick with the notation: K is the centre of D, and ι denotes the antiautomorphism. Note that ι also acts on K, and is a positive involution. From this one can deduce formally that either K is totally real and ι is trivial on K, or K is a CM field and ι is complex conjugation. Here's an example of the kind of arguments one uses: these "trace forms" are all so canonical that any canonical traces must be rational multiples of one another. From this one sees that if F is the subfield of K fixed by ι then for any element of F, its square has positive trace. This is enough to deduce that F is totally real! And so on and so on.

The structure theorem is the following. If K is totally real, this is called "the first kind", and we have the following three possibilities:

Type I: $D = K$ and ι is trivial.

Type II: D/K is a quaternion division algebra which is split at all infinite places, and the Rosati involution sends x to ax^ta^{-1} where a^t denotes the standard involution on D, and a is an element of D whose square is a totally negative element of K. Note that the Rosati involution can't equal the standard involution in this case, as the standard involution isn't positive.

Type III: D/K is a quaternion algebra ramified at all infinite places, and the Rosati involution is the standard involution.

If however K is CM, then D is of "the second kind", and we have

Type IV: D is a division algebra over K, with certain constraints on the local invariants of D in the Brauer groups of the completions of K at the finite places, and Rosati is a bit messy—see Mumford p202.

One can also read off things about the dimension of A: for example, in type I we must have $[K : \mathbf{Q}]$ divides g and in types II and III we must have $2[K : \mathbf{Q}]$ divides g. This stuff was all completely analysed by Albert and Shimura in characteristic 0; they worked out exactly which of these D could occur as the endomorphism ring of a g-dimensional abelian variety and so on. Ugh. See Shimura's Annals 78 paper.

I find all this stuff an absolute headache, and Shimura insists on working in such generalities. I will do the opposite, and creep up on the problem I'm interested in ("modèles étranges") by doing lots of examples, which get more and more intricate.

I will recall one useful fact though: if R is a ring acting on a g-dimensional abelian variety A up to isogeny, that is, equipped with a map $R \to \mathrm{End}^0(A)$, then there are induced actions of R on V and on $L \otimes \mathbf{Q}$, and hence, after choosing bases, maps $\rho_{\mathbf{C}} : R \to \mathrm{M}_g(\mathbf{C})$ and $\rho_{\mathbf{Q}} : R \to \mathrm{M}_{2g}(\mathbf{Q})$. Chasing around the definitions shows that $\rho_{\mathbf{Q}}$ is isomorphic over $\mathbf{C}$ to $\rho_{\mathbf{C}} \oplus \overline{\rho_{\mathbf{C}}}$.

Definition 7.1. *We call $\rho_{\mathbf{C}}$ the* analytic representation *and $\rho_{\mathbf{Q}}$ the* rational representation *associated to the action of R on A.*

8 The Hilbert case.

A useful reference is Katz' "p-adic L-functions for CM fields." Rapoport's thesis is another one. These papers probably both look quite scary if you're a grad student though. Maybe Deligne-Pappas also has some good stuff, although as it happens I didn't look at it, probably because I wasn't doing anything anywhere as near as deep as what they are doing, here.

Let F be a totally real field of degree g, and let $\mathcal{O}_F$ be the ring of integers in F. We are going to consider endomorphisms by $\mathcal{O}_F$. We could use a smaller order—I say epsilon about this later.

Consider first a piece of linear algebra data (V, L) of dimension g, equipped with an action of $\mathcal{O}_F$, that is, a map $\mathcal{O}_F \to \mathrm{End}(V)$ that preserves L. If the linear algebra data is polarizable then such a thing is called a Hilbert-Blumenthal Abelian Variety or HBAV (note: this is only the right definition in characteristic 0—we really need an extra condition in general, which is automatic in characteristic 0). We will see in a bit that these pieces of linear algebra data are in fact always polarizable, canonically, up to some equivalence which is a bit weaker than the equivalence we have been considering up to now.

Back to (V, L) with an action of $\mathcal{O}_F$. Extend the action of $\mathcal{O}_F$ to one of F on V and $L \otimes \mathbf{Q}$. The induced analytic representation is diagonalisable, because each matrix $\rho_{\mathbf{C}}(f)$, $f \in F$ satisfies a polynomial equation with distinct linear factors, and commuting diagonalisable matrices are simultaneously diagonalisable. Hence $\rho_{\mathbf{C}}$ is just the sum of 1-dimensional representations of F, that is, maps $F \to \mathbf{R}$. The fact that $\rho_{\mathbf{Q}}$ is rational now easily implies that $\rho_{\mathbf{C}}$ contains each map $F \to \mathbf{R}$ equally often, and hence the trace of $\rho_{\mathbf{C}}$ is just the trace of F down to $\mathbf{Q}$, and $\rho_{\mathbf{C}}$ is just isomorphic to the natural action of F on $F \cong \mathbf{Q}^g$, tensored up to $\mathbf{C}$. In other words, V is free of rank 1 over $F \otimes \mathbf{C} = \mathcal{O}_F \otimes \mathbf{C}$.

[We remark that if A/S is a g-dimensional abelian scheme over a scheme S, with an action of $\mathcal{O}_F$, then one cannot deduce that the tangent space $(\Omega^1_{A/S})^\vee$ is locally free of rank 1 over $\mathcal{O}_F \otimes \mathcal{O}_S$, even if S is the spectrum of an algebraically closed field (of characteristic p): there's a counterexample due to Rapoport in his thesis, at the end of chapter 1. Hence we could add this condition to our definition of HBAV when we leave the world of characteristic 0. It turns out that in fact even this definition is probably not the best one in characteristic dividing the discriminant of F; see Deligne-Pappas for what to do in this case.]

Let A be an HBAV, that is, a polarizable (V, L) as above. Note that because $\mathcal{O}_F$ is commutative, A^* inherits an action of $\mathcal{O}_F$ too. Consider a non-degenerate Riemann form H on A, inducing a map $\phi : A \to A^*$. This map may or may not commute with the $\mathcal{O}_F$-action. If it doesn't, we're not interested. Let's see what happens if it does—in other words, let's assume that the induced ϕ satisfies $f^*\phi = \phi f : A \to A^*$ for all $f \in \mathcal{O}_F$.

Easy check: a non-degenerate Riemann form H commutes with the $\mathcal{O}_F$-action iff $H(fv, w) = H(v, fw)$ for all $v, w \in V$ and $f \in F$. Hence H commutes with the $\mathcal{O}_F$-action iff the image of F in $\mathrm{End}^0(A)$ is fixed by the Rosati involution corresponding to F.

We know that V is free of rank 1 over $\mathcal{O}_F \otimes \mathbf{C} = F \otimes \mathbf{C} = \mathbf{C}^g$, with the index set being the field homomorphisms $F \to \mathbf{R}$. If a matrix commutes with the $\mathcal{O}_F$-action, this implies that it's diagonal, and conversely every diagonal matrix commutes with the $\mathcal{O}_F$-action. Viewed in this way, one sees that if H is a non-degenerate Riemann form which commutes with the $\mathcal{O}_F$-action, then for a fixed totally positive $\lambda \in F$, the sesquilinear form defined by sending v, w to $H(v, \lambda w)$ is also a non-degenerate Riemann form which commutes with the $\mathcal{O}_F$-action. Say two non-degenerate Riemann forms related in this way are

$F_{>0}$-equivalent.

Definition 8.1. *An* F-polarization *of an HBAV is an $F_{>0}$-equivalence class of non-degenerate Riemann forms which commute with the $\mathcal{O}_F$-action.*

Lemma 8.2. *Any g-dimensional (V, L) with an action of $\mathcal{O}_F$ does possess a non-degenerate Riemann form H which commutes with the $\mathcal{O}_F$-action, and is hence an HBAV. Furthermore, any non-degenerate Riemann form which commutes with the $\mathcal{O}_F$-action is $F_{>0}$-equivalent to H. Hence HBAVs are canonically F-polarized.*[2] 2

Note that this is strongly analogous to the elliptic curve case. Note also that this equivalence class might not contain a principal polarization, for these do not always exist, as we shall see later. Note that for a g-dimensional abelian variety over an arbitrary field with endomorphisms by $\mathcal{O}_F$, there will always exist an F-polarization that commutes with the $\mathcal{O}_F$-action, although it's much messier to prove—see section 1 of Rapoport's thesis—one reduces to the case where A is a power of a simple variety and then uses Albert's classification! Not deep, really, but very messy.

Proof. Just generalise the elliptic curve case. We have seen that V is free of rank 1 over $\mathcal{O}_F \otimes \mathbf{C}$; so WLOG $V = \mathcal{O}_F \otimes \mathbf{C}$, and L is $\mathcal{O}_F$-invariant and so shrinking it a finite amount if necessary, we may assume it's free of rank 2 and then WLOG it's $\mathcal{O}_F \oplus \mathcal{O}_F\tau$. Shrinking L a little more if necessary, we may assume that $\mathrm{Im}(\tau) \in \mathcal{O}_F \otimes \mathbf{R} = \mathbf{R}^g$ has g positive entries. Now writing $\mathcal{O}_F \otimes \mathbf{C} = \mathbf{C}^g$ and $x_i \in \mathbf{C}$ for the ith component of $x \in \mathcal{O}_F \otimes \mathbf{C}$, we may define $H(v, w) = \sum_i v_i \overline{w_i} / \mathrm{Im}(\tau_i)$; note the analogy with the elliptic curve case. This is clearly Hermitian; the fact that $\mathrm{Im}(H)$ is $\mathbf{Q}$-valued on L is a pleasant check and boils down to the fact that the trace of an element of F is in $\mathbf{Q}$. Note that the matrix associated to H is just diagonal, and so H commutes with the $\mathcal{O}_F$-action.

Finally, any non-degenerate Riemann form which commutes with the $\mathcal{O}_F$-action must be diagonal with positive real entries $\alpha_i / \mathrm{Im}(\tau_i)$, and for it to be a Riemann form we must have $\sum_i \alpha_i \lambda_i \in \mathbf{Q}$ for any $\lambda \in F$; but the space of $\mathbf{Q}$-linear maps $F \to \mathbf{Q}$ is g-dimensional as a $\mathbf{Q}$-vector space and hence all such maps must be of the form $f \mapsto \mathrm{tr}_{F/\mathbf{Q}}(\alpha f)$ for some $\alpha \in F$; we deduce that $\alpha_i = (\alpha)_i$ for some α. Positivity implies that α is totally positive, and hence our form is $F_{>0}$-equivalent to H. □

Elliptic curves are canonically principally polarized. HBAVs are canonically F-polarized—but on the other hand it's easy to find examples where the $F_{>0}$-equivalence class of this polarization contains no principal Riemann form. This is because any isomorphism $V/L \to V^*/L^*$ induced by a polarization of an HBAV will commute with the $\mathcal{O}_F$-action and hence give an isomorphism $L \to L^*$ of $\mathcal{O}_F$-modules. On the other hand, by the construction of L^*, one sees that

[2]Note however that I think that people like Deligne-Pappas still actually choose a polarization in this equivalence class when working with Shimura varieties.

it equals $\mathrm{Hom}(L, \mathbf{Z})$, and this is an isomorphism of $\mathcal{O}_F$-modules by definition of the action of $\mathcal{O}_F$ on both sides. Hence we definitely always have $L^* \cong \mathrm{Hom}(L, \mathcal{O}_F) \otimes_{\mathcal{O}_F} \mathrm{Hom}(\mathcal{O}_F, \mathbf{Z})$. Now $L \cong \mathcal{O}_F \oplus I$ for some fractional ideal I, by the general theory of projective modules over Dedekind domains. Hence $L^* \cong (\mathcal{O}_F \oplus I^{-1}) \otimes_{\mathcal{O}_F} \mathfrak{d}^{-1} \cong \mathcal{O}_F \oplus I^{-1}\mathfrak{d}^{-2}$ which won't be isomorphic to L in general—in fact, it definitely won't be isomorphic to L if $I^2\mathfrak{d}^2$ isn't principal.

Here is a discrete invariant attached to an HBAV. Let (V, L) be an HBAV. Consider the group of morphisms $\alpha : V/L \to V^*/L^*$ of abelian varieties, which are symmetric ($\alpha = \alpha^*$) and commute with the $\mathcal{O}_F$-action. Non-trivial such things exist, because the polarizations we constructed above give examples. In fact, to give a map $V \to V^*$ is to give a sesquilinear form on V; for the map to commute with the $\mathcal{O}_F$-action means that it is diagonal when we write $V \cong F \otimes \mathbf{C} = \mathbf{C}^g$; symmetry translates into the form being Hermitian, and hence having real entries on the diagonal; taking $L \otimes \mathbf{Q}$ into $L^* \otimes \mathbf{Q}$ translates into the form being an F-multiple of the canonical polarization H we constructed on an HBAV earlier; finally, taking L into L^* is an integrality condition ensuring that the resulting object is, once one fixes H, a fractional ideal of F. Changing H by a totally positive element might change the fractional ideal but it doesn't change its class in the narrow class group of ideals modulo ideals generated by a totally positive element. Let's call this element of the narrow class group the "type" of the HBAV. This is non-standard terminology, in the sense that I just made it up.

This invariant, taking values in the narrow class group of F, comes out very naturally in the adelic viewpoint of all this, because the Shimura variety that one constructs adelically is some quotient of $\mathrm{GL}_2(\mathbf{A}_F)$ which admits a map onto some quotient of $\mathrm{GL}_1(\mathbf{A}_F)$ which turns out to be precisely the narrow class group of F by class field theory.

I now have several choices. I can write down a bijection between a big space of polarised HBAV's and some kind of adelic space, which is naturally disconnected and has the same number of components as the narrow class group of F (see p27 of van der Geer's book, for example), or I can restrict to HBAV's of some fixed type and then just get more "classical" moduli spaces, related to quotients of products of $\mathbf{H}_1$s or $\mathbf{H}_1^+$s. I have opted to do the latter first, on the basis that it's more in the spirit of everything else I'm doing here, and I say something about the former later.

Let's firstly try and work out this discrete invariant in a concrete situation. Let (V, L) be a piece of linear algebra data with an action of $\mathcal{O}_F$. As we have seen, V is isomorphic to $\mathcal{O}_F \otimes \mathbf{C}$ as an $\mathcal{O}_F$-module, and we can choose the isomorphism so that L becomes isomorphic to $\mathcal{O}_F \oplus \mathfrak{a}\tau$ with $\mathfrak{a}$ a fractional ideal of F and τ an element of $F \otimes \mathbf{C}$ whose imaginary part is invertible in $F \otimes \mathbf{R}$. By changing $\mathfrak{a}$ to $\mathfrak{a}\lambda^{-1}$ and τ to $\lambda\tau$ for some appropriate $\lambda \in \mathcal{O}_F$, we may assume that $\mathrm{Im}(\tau)$ is totally positive. We identify V^* with $F \otimes \mathbf{C}$ too, and the canonical map $V \times V^* \to \mathbf{C}$ is given by $v, w \mapsto \mathrm{tr}(\overline{v}w)$. Now L^* is the $\mathbf{Z}$-dual of L, and one computes that $L^* = (\mathfrak{d}^{-1}\mathfrak{a}^{-1}\mathcal{O}_F \oplus \mathfrak{d}^{-1}\mathcal{O}_F\tau)/\mathrm{Im}(\tau)$. The canonical polarization is the $F_{>0}$-class of H which is represented by $1/\mathrm{Im}(\tau)$, and hence the type of (V, L) is the class in the narrow class group represented by the ideal

$\mathfrak{d}^{-1}\mathfrak{a}^{-1}$.

Conclusion: if $\mathrm{Im}(\tau)$ is totally positive, and $L = \mathcal{O}_F \oplus \mathfrak{a}\tau$, then the type of $(F \otimes \mathbf{C})/L$ is $\mathfrak{d}^{-1}\mathfrak{a}^{-1}$.

We can now prove anything we like in the Hilbert case. For example

Theorem 8.3. *(version 1) There is a canonical bijection between isomorphism classes of (F-polarized) HBAVs of type $\mathfrak{d}^{-1}$ and the set $\mathrm{GL}_2^+(\mathcal{O}_F)\backslash(\mathbf{H}_1^+)^g$, where $\mathrm{GL}_2^+(\mathcal{O}_F)$ is the elements of $\mathrm{GL}_2(\mathcal{O}_F)$ with totally positive determinant.*

Recall that an HBAV has a unique F-polarization and so we will ignore the polarization in what follows.

Proof. An HBAV of type $\mathfrak{d}^{-1}$ can be realised as $V/(\mathcal{O}_F \oplus \mathcal{O}_F\tau)$ with $V = F \otimes \mathbf{C}$ and $\mathrm{Im}(\tau)$ totally positive. This gives us $\tau \in (\mathbf{H}_1^+)^g$. Any $\mathcal{O}_F$-module automorphism of $\mathcal{O}_F^2$ is represented by an element of $\mathrm{GL}_2(\mathcal{O}_F)$; the obvious generalisation of the usual calculation shows that if the abelian varieties $V/(\mathcal{O}_F \oplus \mathcal{O}_F\tau_1)$ and $V/(\mathcal{O}_F \oplus \mathcal{O}_F\tau_2)$ are isomorphic, with both $\mathrm{Im}(\tau_i)$ totally positive, then $\tau_2 = \gamma\tau_1$ with $\gamma \in \mathrm{GL}_2^+(\mathcal{O}_F)$. And that's it. □

Note that $\mathrm{GL}_2^+(F)$ looks a lot more unnatural than $\mathrm{SL}_2(\mathbf{Z})$, but it really is what comes out in the calculations. This is perhaps why one should find the next version more appealing.

Theorem 8.4. *(version 2) There is a canonical bijection between isomorphism classes of (F-polarized) HBAVs which have type some fractional ideal equivalent to $\mathfrak{d}^{-1}$ in the full ideal class group of F, and the set $\mathrm{GL}_2(\mathcal{O}_F)\backslash(\mathbf{H}_1)^g$.*

Proof. We can write an abelian variety of type I as $(F \otimes \mathbf{C})/(\mathcal{O}_F \oplus \mathcal{O}_F\tau)$ with $\tau \in F \otimes \mathbf{C}$ and $\mathrm{Im}(\tau) \in (\mathbf{H}_1)^g$. Two such abelian varieties are isomorphic iff their τs are in the same $\mathrm{GL}_2(\mathcal{O}_F)$-orbit, and that's it. □

Note that the above set might well be disconnected, but even so we don't have all abelian varieties! So this version is still perhaps unnatural—if we allow disconnected moduli spaces then we may as well parameterise all HBAVs.

The calculus of the components is as follows. Note that $(\mathbf{H}_1)^g$ has 2^g components; on the other hand, $\mathcal{O}_F^\times$ contains elements that are not totally positive, and hence certain components are identified. In fact, the number of components of the quotient will be $2^g/C$, where C is the size of the image of the natural map $\mathcal{O}_F^\times \to (\pm 1)^g$ where $\pm 1 = \mathbf{R}^\times/\mathbf{R}_{>0}$. Of course the size of the image will also be the size of the kernel of the map from the narrow class group to the class group, by the nature of the space it's parameterising, and hence we recover what is no doubt a standard theorem relating the sizes of the class group and the narrow class group.

What is also true, and I'll say a word about it in the next section, is "version 3" of these results: there is a canonical bijection between isomorphism classes of *all* (F-polarized) HBAVs and some adelic space.

It's unsurprising that adeles are involved: the kind of spaces one considers look like $\mathrm{GL}_2(F)\backslash \mathrm{GL}_2(\mathbf{A}_F)/K_f K_\infty$, where K_∞ is the product of lots of copies of $\mathbf{R}^\times \mathrm{SO}_2(\mathbf{R})$, and this space admits a determinant map down to the group $F^\times \backslash \mathbf{A}_F^\times / \widehat{\mathcal{O}}_F^\times \prod \mathbf{R}_{>0}$ which is just the narrow class group by Class Field Theory.

There is no doubt a "Hilbert-Siegel" theory, where one parameterises abelian varieties of dimension gh equipped with an action of $\mathcal{O}_F$; the answer will then be $\mathrm{GSp}_{2h}(\mathcal{O}_F)\backslash(\mathbf{H}_h)^g$ or something; I won't follow this up at the minute because it's not really where I want to go.

In summary, what I have learnt here is that the good thing to parameterise is abelian varieties of dimension g, equipped with an action of $\mathcal{O}_F$ (which induces the obvious action of F on the tangent space; this comes for free), and such that the Rosati involution preserves $\mathcal{O}_F$ pointwise (this was the statement that the polarization commuted with $\mathcal{O}_F$).

Before I talk about the adelic side of things, I'll say that it now seems to me why the adelic viewpoint is a good one: after a while one gets really bogged down by class groups and so on. For example, let's let $\mathcal{O}$ be an order in F, but not a maximal one. If one wants to parameterise g-dimensional abelian varieties with an action of $\mathcal{O}$, they will be of the form V/L, with L an $\mathcal{O}$-module of rank 2. However, L might presumably not be a sum of two invertible rank 1 $\mathcal{O}$-modules. So the discrete parameter, the "type" of V/L, becomes a lot harder to describe. If one restricts to a "principal" type one can still do something though: for example if we just restrict to L isomorphic to $\mathcal{O} \oplus \mathcal{O}$ then no doubt the same arguments show that the moduli space of such things is $\mathrm{GL}_2(\mathcal{O})\backslash(\mathbf{H}_1)^g$. This makes me think that we should be thinking about such objects as HBAVs with some kind of weird level structure, the likes of which I've never seen because the phenomenon doesn't happen in the elliptic curve case: I guess $\mathrm{GL}_2(\mathcal{O})$ is a congruence subgroup of $\mathrm{GL}_2(\mathcal{O}_F)$. Things are inevitably going to get messier whatever I do, so I'll say a bit about the adeles. Just before I do, I will emphasize this last point, because we'll not see it again: some "level structures" seem to be able to actually control the size of the order that our abelian varieties have endomorphisms by. As we shall see in the section after next, adelic level structures can also sometimes say something about the polarizations of the objects in question. So all of the letters P,E and L in the phrase "PEL type" are related to adelic level structures.

9 The adelic setting: the Hilbert case.

I will just briefly indicate the value of the adelic approach in the next two sections; this one will have as an end result a theorem in the Hilbert case which naturally follows on from the previous section; the next section will re-interpret results from the Siegel case in the adelic setting. I am ashamed to say that after that I'll go back to the classical approach!

Here's the kind of lemma one uses. Let F be a totally real field, and let W be an n-dimensional vector space over F. An $\mathcal{O}_F$*-lattice* in W is an $\mathcal{O}_F$-submodule L of W such that L is finitely-generated as an $\mathcal{O}_F$-module, and contains an F-basis of W. This might not be standard terminology. Another way of saying it is: an $\mathcal{O}_F$-lattice in W is an $\mathcal{O}_F$-submodule which is a $\mathbf{Z}$-lattice in W considered as a $\mathbf{Q}$-vector space.

Let $\mathbf{A}_{F,f}$ denote the finite adeles of F. We define an $\widehat{\mathcal{O}_F}$-lattice in $W \otimes_F \mathbf{A}_{F,f}$ to be a sub-$\widehat{\mathcal{O}_F}$-module of $W \otimes_F \mathbf{A}_{F,f}$ which is isomorphic to $\widehat{\mathcal{O}_F}^n$.

Lemma 9.1. *There's a canonical bijection between $\mathcal{O}_F$-lattices of W and $\widehat{\mathcal{O}_F}$-lattices in $W \otimes_F \mathbf{A}_{F,f}$.*

Proof. Choose a basis[3] for W. Given a lattice L in F^n, its completion at every prime $\mathfrak{p}$ of F is a lattice $L_\mathfrak{p}$ in $F_\mathfrak{p}^n$. This lattice will be the standard lattice $\mathcal{O}_{F,\mathfrak{p}}^n$ for all but finitely many $\mathfrak{p}$, by the standard kind of argument: there's an element of $\mathrm{GL}_n(F)$ sending L into $\mathcal{O}_F^n$ and another element of $\mathrm{GL}_n(F)$ sending $\mathcal{O}_F^n$ into L; now just avoid any primes dividing the indexes of these inclusions, and the determinants of these maps and so on. Hence the elements of $\prod_\mathfrak{p} L_\mathfrak{p}$ are indeed in $\mathbf{A}_{F,f}^n$ and we have an $\widehat{\mathcal{O}_F}$-lattice. Conversely, because $\widehat{\mathcal{O}_F}$ contains the projection map $\mathbf{A}_{F,f}^n \to F_\mathfrak{p}^n$, an $\widehat{\mathcal{O}_F}$-lattice is of the form $\prod L_\mathfrak{p}$ with $L_\mathfrak{p}$ a lattice in $F_\mathfrak{p}$. Choosing a basis of the $\widehat{\mathcal{O}_F}$-lattice gives bases for all the $L_\mathfrak{p}$ and hence all but finitely many of them are $\mathcal{O}_{F,\mathfrak{p}}^n$; intersecting with F^n gives an $\mathcal{O}_F$-submodule L of F^n which certainly contains an F-basis of F^n; the only issue is why L is finitely-generated, and this is because any F-basis of F^n in L will generate an $\mathcal{O}_F$-module which will equal L at all but finitely many primes, and will only be a finite distance away at all primes, which is enough to show that the global index is finite. □

Corollary 9.2. *There is a bijection between $\mathcal{O}_F$-lattices in W and equivalence classes of isomorphisms $\mathbf{A}_{F,f}^n \cong W \otimes_F \mathbf{A}_{F,f}$, where two isomorphisms α and β are equivalent if $\alpha k = \beta$ for some $k \in \mathrm{GL}_n(\widehat{\mathcal{O}_F})$.*

Proof. Now obvious. □

The point is somehow that to give an abelian variety is to give a certain $\mathbf{Z}$-lattice in a complex vector space; to give an abelian variety up to isogeny is to give a $\mathbf{Q}$-lattice; to get back to the $\mathbf{Z}$-lattice one could either give the $\mathbf{Z}$-lattice, or the adelic isomorphism modulo equivalence above, which looks much more complicated but is in fact sufficiently robust that it generalises well.

This kind of argument removes some of these slightly delicate ideal class, or non-free-but-projective, issues. For example, let's go back to that piece of the Hilbert case that I'd not done. To give an HBAV up to isogeny is to give an F-vector space W of dimension 2 equipped with a complex structure on $W \otimes \mathbf{R}$, and to give an F-polarization on the result. The F-polarization will be unique by an earlier lemma, so all we have to do is to understand F-lattices of rank 2 in $F \otimes \mathbf{C}$. WLOG such a lattice is $F \oplus F\tau$ where $\tau \in F \otimes \mathbf{R}$ has invertible imaginary part. Now two such things are isomorphic as polarized abelian varieties up to isogeny iff they are in the same $\mathrm{GL}_2(F)$-orbit (one doesn't have to worry about the polarization because it is forced upon us). One can think of the $\mathcal{O}_F$-lattices

[3]yeah yeah, I know: but the construction won't depend on it.

in F^2 as equivalence classes of isomorphisms $\mathbf{A}_{F,f}^2 \to \mathbf{A}_{F,f}^2$, that is, elements of $\mathrm{GL}_2(\mathbf{A}_{F,f})/\mathrm{GL}_2(\widehat{\mathcal{O}_F})$, and we deduce

Theorem 9.3. *(version 3) There is a canonical bijection between all F-polarized HBAVs for F, and the set*

$$\mathrm{GL}_2(F)\backslash\left(\mathrm{GL}_2(\mathbf{A}_{F,f})/\mathrm{GL}_2(\widehat{\mathcal{O}_F}) \times (\mathbf{H}_1)^g\right).$$

Somehow the point is that we don't parameterise abelian varieties plus extra structure, we parameterise abelian varieties up to isogeny plus extra structure, plus some adelic level structure which nails the abelian variety down within the isogeny class. This sometimes makes our life easier.

10 The adelic approach: the Siegel case.

I added this section to illustrate the following point: an adelic level structure may not only nail down the abelian variety within its isogeny class, it may also say something about the polarization of the variety. I will attempt to explain this in the concrete example of the Siegel case, by reformulating things via the "up-to-isogeny" category.

To give a g-dimensional principally polarized abelian variety, is to give a piece of linear algebra data (V, L) and a Hermitian pairing H on V such that the induced E is integer-valued and induces an isomorphism between L and its $\mathbf{Z}$-dual. On the other hand, it seems hard to give a reasonable definition of "a principally polarized abelian variety up to isogeny"—the problem is that something isogenous to a principally polarized abelian variety could well not be principally polarized, or even principally polarizable! However, it certainly does make sense to talk about a g-dimensional polarized abelian variety up to isogeny; such a thing corresponds to a complex vector space V of dimension g, equipped with a $\mathbf{Q}$-vector subspace $L_{\mathbf{Q}}$ of dimension $2g$ such that the induced map $L_{\mathbf{Q}} \otimes \mathbf{R} \to V$ is an $\mathbf{R}$-linear isomorphism, and equipped with a positive definite H such that the induced E is $\mathbf{Q}$-valued on $L_{\mathbf{Q}}$.

In this setting, a choice of lattice L in $L_{\mathbf{Q}}$ corresponds, by Corollary 9.2, to an isomorphism $\mathbf{A}_f^{2g} \to L_{\mathbf{Q}} \otimes \mathbf{A}_f$ (defined up to $\mathrm{GL}_n(\widehat{\mathbf{Z}})$). To give a choice of lattice is of course to give an abelian variety in the isogeny class. Then H induces a polarization on the corresponding abelian variety. But most choices will give polarizations that aren't principal. How to pull out the principal ones?

We have an alternating $\mathbf{Q}$-valued pairing on $L_{\mathbf{Q}}$, and this induces an alternating $\mathbf{A}_f$-valued pairing on $L_{\mathbf{Q}} \otimes \mathbf{A}_f$. If we are given an isomorphism $\mathbf{A}_f^{2g} \cong L_{\mathbf{Q}} \otimes \mathbf{A}_f$, then the pairing on the right hand side induces a pairing on the left hand side. The polarized abelian variety induced by this isomorphism will be principally polarized if and only if, up to a positive rational factor, this induced pairing on $\mathbf{A}_f^{2g}$ induces an isomorphism between $\widehat{\mathbf{Z}}^{2g}$ and its dual. Motivated by this, let's fix a pairing on $\mathbf{A}_f^{2g}$, namely the one given by J. Then we see that given a polarized abelian variety up to isogeny, if we choose a principally polarized abelian variety in this isogeny class then we get a lattice L in $L_{\mathbf{Q}}$ and that there will be an isomorphism $\mathbf{A}_f^{2g} \cong L_{\mathbf{Q}} \otimes \mathbf{A}_f$ giving rise to L via Corollary 9.2, and under which the pairings on both sides match up, up to a positive rational factor. Conversely, given an isomorphism $\mathbf{A}_f^{2g} \cong L_{\mathbf{Q}} \otimes \mathbf{A}_f$ under which the pairings match up, up to a rational factor, there is an associated principally-polarized abelian variety (although we might have to change the sign of E to get the principal polarization). Two such isomorphisms give the same lattice iff they differ by an element of $\mathrm{GSp}_{2g}(\widehat{\mathbf{Z}})$. These thoughts are an example of what's going on in Scholie 4.11 of Deligne's "varietes de Shimura".

To give a polarized abelian variety up to isogeny is, as the arguments in the Siegel section show, to give an element of $\mathbf{H}_g$, up to an element of $\mathrm{GSp}_{2g}(\mathbf{Q})$ (I guess I am being a bit sloppy here; one should also do the exercise that says that every abelian variety is isogenous to a principally-polarized one). Hence to give a principally-polarized abelian variety is to give an element of

$$\mathrm{GSp}_{2g}(\mathbf{Q})\backslash \left(\mathrm{GSp}_{2g}(\mathbf{A}_f) \times \mathbf{H}_g\right) / \,\mathrm{GSp}_{2g}(\widehat{\mathbf{Z}}),$$

that is, an element of

$$\mathrm{GSp}_{2g}(\mathbf{Q})\backslash \mathrm{GSp}_{2g}(\mathbf{A})/ \mathrm{GSp}_{2g}(\widehat{\mathbf{Z}})K_\infty,$$

where K_∞ is the stabiliser of I_g in $\mathbf{H}_g$.

The point is that this adelic object really somehow knows that it's parameterising the principally-polarized abelian varieties; its points correspond to abelian varieties up to isogeny equipped with a certain symplectic isomorphism, and the existence of this isomorphism forces the polarization to be principal.

11 Indefinite quaternion algebras over Q.

We follow the same (non-adelic) line of attack as in the Hilbert case. Firstly we check that our endomorphisms force a certain structure on the tangent space (and we would have demanded this structure if our data hadn't forced it). Then we find a notion of a polarization such that there is a canonical one on our objects. Then we parameterise. Thanks to Andrei Yafaev for reminding me of the recipe here.

Let B be an indefinite quaternion algebra over $\mathbf{Q}$, with discriminant $d > 0$. If (V, L) is a 2-dimensional piece of linear algebra data, and we have a map $B \to \mathrm{End}^0(V, L)$, this gives us a map $B \to \mathrm{End}(L \otimes \mathbf{Q})$ making $L \otimes \mathbf{Q}$ into a left B-module. If B is non-split then already this implies that $L \otimes \mathbf{Q}$ is free of rank 1 over B. If B is split then $L \otimes \mathbf{Q}$ is a module for $\mathrm{M}_2(\mathbf{Q})$ and, well, I think of it as Morita equivalence but probably there is a more sensible reason, we see that again $L \otimes \mathbf{Q}$ is a free left B-module of rank 1 because it must be two copies of the standard module.

As for the tangent space and the analytic representation: V gives a map $B \to \mathrm{M}_2(\mathbf{C})$ and all such maps are conjugate by Noether-Skolem, so give isomorphic representations: if we fix an isomorphism $B \otimes \mathbf{C} \cong \mathrm{M}_2(\mathbf{C})$ then V

is isomorphic to the standard 2-dimensional module for $M_2(\mathbf{C})$. This may or may not be automatic in any characteristic; probably we should stay away from characteristic dividing $\operatorname{disc}(B)$ though, as we are almost certainly not asking the right question here.

Let's choose a maximal order $\mathcal{O}_B$ in B. If (V, L) has an action of $\mathcal{O}_B$ and happens also to be polarizable, the induced Rosati might fix the image of $\mathcal{O}_B$ but it will not induce the canonical involution on $\mathcal{O}_B$ because the canonical involution is not positive. What we do instead is as follows: fix $t \in \mathcal{O}_B$ with $t^2 = -\operatorname{disc}(B)$ (I think that such a thing automatically exists; certainly $\mathbf{Q}(\sqrt{-d})$ splits B so there are elements of B whose square is $-d$; no doubt it's standard to make sure one is in $\mathcal{O}_B$). Now define an involution ι on B by $b^\iota = t^{-1}b^*t$, where b^* is the canonical involution. This involution *is* positive, and hence there is some chance that it is induced by Rosati. We will see shortly that this is the case. We remark that $t^* = -t$ and so $t^\iota = -t$.

Let us think of V as being equal to $B \otimes \mathbf{R}$. What possibilities are there for the complex structure? If i is the map $V \to V$ denoting multiplication by $i \in \mathbf{C}$ then i has to commute with the action of B and an elementary computation shows that i must then be right multiplication by an element c of $B \otimes \mathbf{R}$; this element must have square -1 of course. It's easy to check that if $c \in M_2(\mathbf{R})$ has square -1 then c is in $SL_2(\mathbf{R})$ and hence $c^* = -c$.

Lemma 11.1. *If (V, L) has an action of $\mathcal{O}_B$ and t, ι are defined as above, then (V, L) is polarizable, and there is a unique polarization on (V, L) such that $\mathcal{O}_B$ is fixed as a set under Rosati, and the induced involution is $b \mapsto b^\iota$.*

Proof. (see Milne's Corvalis article, and Boutot-Carayol.) We identify V with $B \otimes \mathbf{R}$ and define E on V by $E(v, w) = \operatorname{tr}(vtw^\iota) = \operatorname{tr}(vw^*t)$. Note that $E(iv, iw) = E(vc, wc) = E(v, w)$ because $cc^* = -c^2 = 1$. So there is a unique Hermitian $H(v, w) = E(iv, w) + iE(v, w)$ giving rise to E. To check definiteness of H we must check that $\operatorname{tr}(vctv^\iota) > 0$ for any non-zero $v \in B \otimes \mathbf{R}$, which follows from explicitly writing everything out and using the fact that c is conjugate to J. So (V, L) is polarizable. For this choice of H, the induced Rosati is just the adjoint for E, so to check that the induced Rosati is ι on B we must check $E(bv, w) = E(v, b^\iota w)$ which is immediate from the definitions and the fact that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

It remains to show uniqueness. The trick is that if $H = S + iE$ is any polarization then it induces a non-degenerate pairing $B \times B \to \mathbf{Q}$ and hence $E(1, -)$ is a linear map $B \to \mathbf{Q}$ which must be of the form $w \mapsto \operatorname{tr}(btw^\iota)$ by non-degeneracy of trace. Then $E(v, w) = E(1, v^\iota w) = \operatorname{tr}(btw^\iota v) = \operatorname{tr}(vbtw^\iota)$. Now the fact that E is alternating means that $\operatorname{tr}(wbtv^\iota) = -\operatorname{tr}(vbtw^\iota) = -\operatorname{tr}(wt^\iota b^\iota v^\iota) = \operatorname{tr}(wtb^\iota v^\iota) = \operatorname{tr}(wb^*tv^\iota)$ and hence $b = b^*$ which means $b \in \mathbf{Q}$. So E, and hence H, is just a $\mathbf{Q}$-linear multiple of the form we constructed earlier. □

One can check that this polarization is principal but I am a bit tired of this kind of argument; we won't need it. One just checks locally at every finite prime.

Note that for E as above, we have $E(bv, bw) = \det(b)E(v, w)$ for $b \in B$, and hence B is mapping to the general symplectic group determined by E.

Call a (V, L) equipped with a faithful action of $\mathcal{O}_B$ a "false elliptic curve".

Theorem 11.2. *There is a canonical bijection between isomorphism classes of false elliptic curves and the set $\mathcal{O}_B^\times \backslash \mathbf{H}_1$.*

Proof. We need to know that if (V, L) is a piece of linear algebra data with an action of $\mathcal{O}_B$ then $L \cong \mathcal{O}_B$. This is because L is a locally principal $\mathcal{O}_B$-module of rank 1 and B is indefinite, and one can use some kind of strong approximation result—I haven't checked the details but I think I can reconstruct the argument. If we believe this then WLOG $V = B \otimes \mathbf{R}$ and $L = \mathcal{O}_B$, and we must parameterise the varying complex structure on V; such a thing is $c \in \mathrm{SL}_2(\mathbf{R})$ with $c^2 = -1$ (acting as right multiplication). As we know, such things are all conjugates of J and if $c = \gamma J \gamma^{-1}$ then γi is the element in $\mathbf{H}_1$ associated to c. Two such structures are isomorphic iff there's an element $u \in \mathcal{O}_B^\times$ such that right multiplication by u takes one situation to the other; unravelling this, we see that u acts on the set of all c's by conjugation, and hence the action on $\mathbf{H}_1$ is the usual left action and we are finished. □

One could work over a totally real field and use a totally indefinite quaternion algebra; I won't do this—see Milne's Corvalis article for more details.