SHIMURA VARIETIES

LECTURES BY SOPHIE MOREL

Abstract. These are the notes for the course given by Sophie Morel in 2022 Summer School on the Langlands Program at IHES. The goal of these lectures is to give an introduction to Shimura varieties, to present some examples, and to explain the conjectures on their cohomology (at least in the simplest case). Depending on different points of view, Shimura varieties are a special kind of locally symmetric spaces, a generalization of moduli spaces of abelian schemes with extra structures, or the imperfect characteristic 0 version of moduli spaces of shtuka. They play an important role in the Langlands program because they have many symmetries (the Hecke correspondences) allowing us to link their cohomology to the theory of automorphic representations, and on the other hand they are explicit enough for this cohomology to be computable.

Contents

1. Locally symmetric spaces and Shimura varieties 1
1.1. Locally symmetric spaces 1
1.2. The Siegel modular variety 5
1.3. The Siegel upper half space as a moduli space of abelian varieties over $\mathbb{C}$ 5
1.4. Shimura varieties over $\mathbb{C}$ 12
1.5. Canonical models 14
2. Arithmetic Shimura varieties 17
2.1. PEL type Shimura varieties 17
2.2. Hodge type and abelian type Shimura varieties 22
2.3. General Shimura varieties 24
2.4. Kottwitz's simple Shimura varieties 24
3. The cohomology of Shimura varieties 28
3.1. Matsushima's formula 28
3.2. Étale cohomology of canonical models: the Kottwitz conjecture 29
3.3. Applications of the Kottwitz conjecture 32
3.4. Proving the Kottwitz conjecture 35
References 38

1. Locally symmetric spaces and Shimura varieties

1.1. **Locally symmetric spaces.** Let G be a semisimple algebraic group over $\mathbb{Q}$, for example SL_n, Sp_{2n}, or $\mathrm{SO}(p,q)$. We would like to present some "nice enough" whose cohomology is related to automorphic representations of G. A good reference for locally symmetric spaces is the introductory paper [Ji06].

To simplify the presentation, we will assume here that $G(\mathbb{R})$ is connected. Let K_∞ be a maximal compact subgroup of $G(\mathbb{R})$, and let $X = G(\mathbb{R})/K_\infty$. If Γ is a discrete subgroup of $G(\mathbb{R})$ such that $\Gamma\backslash G(\mathbb{R})$ (or equivalently $\Gamma\backslash X$) is compact and Γ acts properly and freely on X,[1] then there is a classical connection between the cohomology of $\Gamma\backslash X$ and automorphic representations of $G(\mathbb{R})$, called **Matsushima's formula** (see Matsushima's paper [Mat67]). We will state a modern reformulation in

Date: July 12, 2022.

[1] This holds for example if Γ is torsion free, which happens when Γ is small enough.

Section 3, but roughly it relates the Betti numbers of $\Gamma\backslash X$ and the multiplicities of representations of $G(\mathbb{R})$ in $L^2(\Gamma\backslash G(\mathbb{R}))$.

In fact, Matsushima's paper deals with semi-simple real Lie groups. Here, we have an algebraic group defined over $\mathbb{Q}$, so we have a particularly nice way to produce discrete subgroups of $G(\mathbb{R})$. Remember that a subgroup Γ of $G(\mathbb{Q})$ is called an **arithmetic subgroup** if there exists a closed embedding $G \subset \mathrm{GL}_N$ such that, setting $G(\mathbb{Z}) = G(\mathbb{Q}) \cap \mathrm{GL}_N(\mathbb{Z})$, we have that $\Gamma \cap G(\mathbb{Z})$ is of finite index in Γ and in $G(\mathbb{Z})$.[2] If Γ is small enough, then it acts properly and freely on X [Ji06, Proposition 5.5], so the quotient $\Gamma\backslash X$ is a real analytic manifold. Also, the quotient $\Gamma\backslash G(\mathbb{R})$ is compact if and only if G is anisotropic (over $\mathbb{Q}$), which means that G has no nontrivial parabolic subgroup defined over $\mathbb{Q}$ [Ji06, Theorem 5.10]. (A subgroup of G is parabolic if it contains a Borel subgroup B of G.) If $\Gamma\backslash X$ is not compact but $G(\mathbb{R})$ has a discrete series, then there is an extension of Matsushima's formula, due to Borel and Casselman in [BC83], that involves L^2 cohomology of $\Gamma\backslash X$; see Section 3.

We actually would like to see automorphic representations of $G(\mathbb{A})$ (not just $G(\mathbb{R})$) in the cohomology of our spaces, so we will use adelic versions of $\Gamma\backslash X$. Let K be an open compact subgroup of $G(\mathbb{A}_f)$; for example, if we have chosen an embedding $G \subset \mathrm{GL}_N$, then we could take

$$K = G(\mathbb{A}_f) \cap \mathrm{Ker}(\mathrm{GL}_N(\widehat{\mathbb{Z}}) \to \mathrm{GL}_N(\mathbb{Z}/n\mathbb{Z})),$$

for some positive integer n (these are called **principal congruence subgroups**). Let

$$M_K = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K,$$

where the group K acts by right translations on the factor $G(\mathbb{A}_f)$, and the group $G(\mathbb{Q})$ acts by left translations on both factors simultaneously. Choose a system of representatives $(x_i)_{i\in I}$ of the (finite) quotient $G(\mathbb{Q})\backslash G(\mathbb{A}_f)/K$, and set

$$\Gamma_i = G(\mathbb{Q}) \cap x_i K x_i^{-1}$$

for every $i \in I$. Then the Γ_i are arithmetic subgroups of $G(\mathbb{Q})$, and we have

$$M_K = \coprod_{i\in I} \Gamma_i\backslash X,$$

so M_K is a real analytic manifold if K is small enough. But now we have an action of $G(\mathbb{A}_f)$ on the projective system $(M_K)_{K\subset G(\mathbb{A}_f)}$, so we get an action on $\varinjlim_K \mathrm{H}^*(M_K)$, where H^* is any "reasonable" cohomology theory, for example Betti cohomology. If G is anisotropic over $\mathbb{Q}$, then Matsushima's result can be reformulated to give a description of this action in terms of irreducible representations of $G(\mathbb{A})$ appearing in $L^2(G(\mathbb{Q})\backslash G(\mathbb{A}))$, and there is also a version of the Borel–Casselman generalization.

There is another way to think about the action of $G(\mathbb{A}_f)$ on $(M_K)_{K\subset G(\mathbb{A}_f)}$, which does not involve a limit on K. Fix a Haar measure on $G(\mathbb{A}_f)$ such that open compact subgroups of $G(\mathbb{A}_f)$ have rational volume (this is possible because these groups are all commensurable); then every open subset of $G(\mathbb{A}_f)$ has rational volume. The **Hecke algebra** of G is the space $\mathcal{H}_G$ of locally constant functions with compact support from $G(\mathbb{A}_f)$ to $\mathbb{Q}$; if $f, g \in \mathcal{H}_G$, then the convolution product $f * g$ still has rational values by the choice of Haar measure, so convolution defines a multiplication on $\mathcal{H}_G$. For every open compact subgroup K of $G(\mathbb{A}_f)$, the **Hecke algebra at level** K is the subalgebra $\mathcal{H}_{G,K}$ of bi-K-invariant functions in $\mathcal{H}_G$; we have $\mathcal{H}_G = \bigcup_K \mathcal{H}_{G,K}$.

Fix K small enough. Then $\mathrm{H}^*(M_K)$ is basically the set of K-invariant vectors:

$$\mathrm{H}^*(M_K) = \varinjlim_{K'\subset G(\mathbb{A}_f)} \mathrm{H}^*(M_{K'})^K,$$

so it has an action of $\mathcal{H}_{G,K}$.[3] We can describe this action using Hecke correspondences: let $g \in G(\mathbb{A}_f)$, let K' be an open compact subgroup of $G(\mathbb{A}_f)$ such that $K' \subset K \cap gKg^{-1}$, then we have a **Hecke correspondence**

$$\begin{aligned} (T_1, T_g) : M_{K'} &\longrightarrow M_K \times M_K \\ (x, h) &\longmapsto ((x, h), (x, hg)), \end{aligned}$$

[2] We can check that this definition does not depend on the embedding $G \subset \mathrm{GL}_N$, see [Ji06, Proposition 4.2].

[3] In fact, we can recover the action of $G(\mathbb{A}_f)$ on $\varinjlim_{K'\subset G(\mathbb{A}_f)} \mathrm{H}^*(M_{K'})$ from the action of $\mathcal{H}_{G,K}$ on M_K for every K small enough.

and T_1, T_g are both finite covering maps if K is small enough. Up to a scalar,[4] the action $\mathbb{1}_{KgK}$ on $\mathrm{H}^*(M_K)$ is given by pulling back cohomology classes along T_1, then pushing them forward along T_g.

We can also ask whether there is more structure on the spaces $\Gamma\backslash X$ (or M_K). For example, suppose that $G = \mathrm{SL}_2$ and $K_\infty = \mathrm{SO}(2)$. Then, for Γ an arithmetic subgroup of $\mathrm{SL}_2(\mathbb{Z})$, the space $\Gamma\backslash X$ is a modular curve, so it is (the set of complex points of) an algebraic variety defined over a function field F, and we can use the commuting actions of Hecke correspondences and of the absolute Galois group of F on its étale cohomology to construct some instance of the global Langlands correspondence for SL_2 or GL_2.

In order to generalize this picture, we first to know when the spaces $\Gamma\backslash X$ or M_K are the set of $\mathbb{C}$-points of an algebraic variety, and whether this algebraic variety is defined over a number field. As we will see later, another advantage over M_K over $\Gamma\backslash X$ is that, when the answer to the above question is "yes", then the M_K for K varying tend to all be defined over the same field, while this is not the case for the $\Gamma\backslash X$.

Remark 1.1. The first step is to check whether $\Gamma\backslash X$ has the structure of a complex manifold, and there are obvious obstructions to that. For example, if $G = \mathrm{SL}_3$ and $K_\infty = \mathrm{SO}(3)$, then $\Gamma\backslash X$ is 5-dimensional as a real manifold, so it cannot have the structure of a complex manifold. In fact, there is no structure of complex manifold on $\Gamma\backslash X$ for $G = \mathrm{GL}_d$ with $d \geqslant 3$, as we will now see.

Choose a $G(\mathbb{R})$-invariant Riemannian metric on $X = G(\mathbb{R})/K_\infty$ (such a metric is unique up to rescaling on each irreducible factor). Then X is a **symmetric space**, that is, a Riemannian manifold such that:

(a) The group of isometries of X acts transitively on X;
(b) For every $p \in X$, there exists a symmetry s_p of X (i.e. an involutive isometry) such that p is an isolated fixed point of s_p.

Moreover, the symmetric space X is of **noncompact type**, that is, it has negative curvature. For Γ a small enough arithmetic subgroup of $G(\mathbb{Q})$, the Riemannian manifold $\Gamma\backslash X$ is a **locally symmetric space**; in particular, it does not satisfy condition (a) anymore, and it satisfies a variant of condition (b) where we only ask for the symmetry to be defined in a neighborhood of the point. See Ji's notes [Ji06] for a review of locally symmetric spaces.

We say that X is a **Hermitian symmetric domain** if it admits a $G(\mathbb{R})$-invariant Hermitian metric. See Milne's notes [Mil05, §1] for a review of Hermitian symmetric domains.

Example 1.2 (Siegel upper half space)**.** Let d be a positive integer. The **Siegel upper half space** $\mathfrak{h}_d^+$ is the set of symmetric $d \times d$ complex matrices in $Y \in M_d(\mathbb{C})$ such $\mathrm{Im}(Y)$ is positive definite; if $d = 1$, then this is just the usual upper half plane. Then the Siegel upper half space $\mathfrak{h}_d^+$ is a Hermitian symmetric domain. The proofs of the basic properties of $\mathfrak{h}_d^+$ can be found in Siegel's paper [Sie43].

We first need to see $\mathfrak{h}_d^+$ as a symmetric space. Let Sp_{2d} be the symplectic group of the symplectic form with matrix $\begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}$, where $I_d \in \mathrm{GL}_d(\mathbb{Z})$ is the identity matrix. For every commutative ring R, we have

$$\mathrm{Sp}_{2d}(R) = \left\{ g \in \mathrm{GL}_{2d}(R) \,\middle|\, {}^t g \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix} g = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix} \right\}.$$

Note that $\mathrm{Sp}_2 = \mathrm{SL}_2$. We make $\mathrm{Sp}_{2g}(\mathbb{R})$ act on $\mathfrak{h}_d^+$ by the following formula:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot Y = (AY + B)(CY + D)^{-1}, \quad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2d}(\mathbb{R}),$$

where A, B, C, D are $d \times d$ matrices. Then this action is transitive (see [Sie43, Page 9]). Let K_∞ be the stabilizer in $\mathrm{Sp}_{2d}(\mathbb{R})$ of $iI_d \in \mathfrak{h}_d^+$. Then $K_\infty = O(2d) \cap \mathrm{Sp}_{2d}(\mathbb{R})$ (this is easy to check directly), so

[4] Make scalar precise.

it is a maximal compact subgroup of $\mathrm{Sp}_{2d}(\mathbb{R})$,[5] and we have

$$\mathfrak{h}_d^+ \simeq \mathrm{Sp}_{2d}(\mathbb{R})/K_\infty$$

as real analytic manifolds.

Also, the space $\mathfrak{h}_d^+$ is an open subset of the complex vector space of symmetric matrices in $M_d(\mathbb{C})$, so it has an obvious structure of complex manifold. It remains to construct a $\mathrm{Sp}_{2d}(\mathbb{R})$-invariant Hermitian metric on $\mathfrak{h}_d^+$. Let $\mathcal{D}_d$ be the set of symmetric matrices $A \in M_d(\mathbb{C})$ such that $I_d - A^*A$ is positive definite; this is a bounded domain in the complex vector space of symmetric matrices in $M_d(\mathbb{C})$, hence is equipped with a canonical Hermitian metric called the **Bergman metric**, which has negative curvature (see for example, [Mil05, Theorem 1.3]); in particular, this metric is invariant by all holomorphic automorphisms of $\mathcal{D}_d$. Now note that we have an isomorphism

$$h_d^+ \xrightarrow{\sim} \mathcal{D}_d, \quad X \longmapsto (iI_d - X)(iI_d + X)^{-1}$$

(whose inverse sends $A \in \mathcal{D}_d$ to $i(I_d - A)(I_d + A)^{-1}$), see [Sie43, pp. 8-9]. We can give a formula for the resulting Hermitian metric on h_d^+: up to a positive scalar, it is given by

$$ds^2 = \mathrm{Tr}(\mathrm{Im}(Y)^{-2} dY\, \mathrm{Im}(Y)^{-1} d\overline{Y})$$

(see formula (28) on page 17 of [Sie43]).

The isomorphism $h_d^+ \simeq \mathcal{D}_d$ is called a **bounded realization** of h_d^+.

We can give a complete classification of Hermitian symmetric domains (cf. [Mil05, Theorem 1.21]), in terms of real algebraic groups:

Theorem 1.3. *Suppose that $G(\mathbb{R})$ is connected and adjoint. The locally symmetric space X is a Hermitian symmetric domain if and only if there exists a morphism of real Lie groups $u : \mathrm{U}(1) \to G(\mathbb{R})$ such that:*

(a) *The only characters of $\mathrm{U}(1)$ that appear in its representation $\mathrm{Ad} \circ u$ on $\mathrm{Lie}(G(\mathbb{R}))$ are 1, z, and z^{-1};*
(b) *Conjugation by $u(i)$ is a Cartan involution of $G(\mathbb{R})$, which means that $\{g \in G(\mathbb{C}) \mid g = u(i)\overline{g}u(i)^{-1}\}$ is compact;*
(c) *The projection of $u(i)$ to a simple factor of $G(\mathbb{R})$ is never equal to 1.*

Moreover, we can choose u such that K_∞ is the centralizer of u in $G(\mathbb{R})$, which means that X is isomorphic to set of conjugates of u by elements of $G(\mathbb{R})$.

We explain the construction of the morphism u. Suppose that X is a Hermitian symmetric domain, and let $p \in X$. For every $z \in \mathbb{C}$ with $|z| = 1$, multiplication by z on T_pX preserves the Hermitian metric and sectional curvatures, so there exists a unique isometry $u_p(z)$ of D fixing p and such that $T_pu_p(z)$ is multiplication by z. The uniqueness implies that $u_p(z)u_p(z') = u_p(zz')$ if $|z| = |z'| = 1$, so we get a morphism of groups from $\mathrm{U}(1)$ to the group of isometries of X, which is equal to $G(\mathbb{R})^0_{\mathrm{ad}}$.

Example 1.4. (1) If $G = \mathrm{Sp}_{2d}$, let $h : \mathbb{C}^\times \to G(\mathbb{R})$ be defined by

$$h(a + ib) = \begin{pmatrix} aI_d & -bI_d \\ bI_d & aI_d \end{pmatrix}.$$

Then we can take $u : \mathrm{U}(1) \to \mathrm{PSp}_{2d}(\mathbb{R})$ given by $u(z) = h(\sqrt{z})$. Note that u does not lift to a morphism from $\mathrm{U}(1)$ into $G(\mathbb{R})$.

(2) If $G = \mathrm{PGL}_n$ with $n \geqslant 3$, then the centralizer of a character $u : \mathrm{U}(1) \to G(\mathbb{R})$ cannot be a maximal compact subgroup of $G(\mathbb{R})$ (exercise), so the locally symmetric space of maximal compact subgroups of $G(\mathbb{R})$ is not Hermitian.

[5] In fact, we have an isomorphism (with $X, Y \in \mathrm{GL}_d(\mathbb{R})$):

$$U(d) \xrightarrow{\sim} K_\infty, \quad X + iY \longmapsto \begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}.$$

Theorem 1.3 puts some pretty strong restrictions on the root systems of the simple factors of $G(\mathbb{R})$, see [Mil05, Theorem 1.25] and the table following it. In particular, the type A simple factors of $G(\mathbb{R})$ must be of the form $\mathrm{PSU}(p,q)$, and $G(\mathbb{R})$ can have no simple factor of type E_8, F_4 or G_2.

The natural next step would be to wonder for which Hermitian symmetric domains X the quotients $\Gamma\backslash X$ are algebraic varieties, but in fact it turns out that the answer is "for all of them", as was proved by Baily and Borel [BB66].

Theorem 1.5 (Baily–Borel)**.** *Suppose that $X = G(\mathbb{R})/K_\infty$ is a Hermitian symmetric domain. Then, for any torsion free arithmetic subgroup Γ of $G(\mathbb{Q})$, the quotient $\Gamma\backslash X$ has a canonical structure of algebraic variety over $\mathbb{C}$.*

The very rough idea is that the sheaf of automorphic forms on $\Gamma\backslash X$ of sufficiently high weight will define an embedding of $\Gamma\backslash X$ into a projective space.

Remember that we did not just want the locally symmetric spaces $\Gamma\backslash X$ to be algebraic varieties, we also wanted them to be defined over a number field, and we would ideally like the number field in question to only depend on G and K_∞. For this, it will actually be easier to work with reductive groups instead of semi-simple groups. As a motivation for this, and for the definition of Shimura varieties, we now spend some more time on the case of the symplectic group.

1.2. **The Siegel modular variety.** See the end of this subsection (1.3.5) for some background on abelian schemes.

1.3. **The Siegel upper half space as a moduli space of abelian varieties over $\mathbb{C}$.** We use the notation of Example 1.2. It is well-known that $\mathfrak{h}_1^+$ parametrizes elliptic curves over $\mathbb{C}$: an element $\tau \in \mathfrak{h}_1^+$ is sent to the elliptic curve $E_\tau = \mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$, and $E_\tau \simeq E_{\tau'}$ if and only if $\tau, \tau' \in \mathfrak{h}_1^+$ they are conjugated under the action of $\mathrm{SL}_2(\mathbb{Z}) = \mathrm{Sp}_2(\mathbb{Z})$; so we can recover τ from E_τ and the data of a symplectic isomorphism $\mathrm{H}_1(E_\tau, \mathbb{Z}) \simeq \mathbb{Z}^2$ where $\mathbb{Z}^2$ is equipped with the standard symplectic form. We want to give a similar picture for higher-dimensional abelian varieties; in fact, the analogy works best if we consider abelian varieties with a principal polarization (Definition 1.29).

We first introduce some notation about symplectic spaces and recall the definition of the (general) symplectic group as a group scheme over $\mathbb{Z}$. If R is a commutative ring, we denote by ψ_R the perfect symplectic pairing on R^{2d} with matrix $\begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}$. So we have

$$\psi_R((x_1, \dots, x_d, y_1, \dots, y_d), (x'_1, \dots, x'_d, y'_1, \dots, y'_d)) = \sum_{i=1}^d x_i y'_i - \sum_{i=1}^g x'_i y_i.$$

The **general symplectic group** GSp_{2d} is the reductive group scheme over $\mathbb{Z}$ whose points in a commutative ring R are given by

$$\mathrm{GSp}_{2d}(R) = \{g \in \mathrm{GL}_{2d}(R) \mid \exists c(g) \in R^\times,\ \forall v, v' \in R^{2d},\ \psi_R(gv, gv') = c(g)\psi_R(v, v')\}.$$

The scalar $c(g)$ is called the **multiplier** of $g \in \mathrm{GSp}_{2d}(R)$. Sending g to $c(g)$ defines a morphism of group schemes $c : \mathrm{GSp}_{2d} \to \mathrm{GL}_1$, whose kernel Sp_{2d} is called the **symplectic group**.

Example 1.6. We have $\mathrm{GSp}_2 = \mathrm{GL}_2$ in which $c = \det$, and $\mathrm{Sp}_2 = \mathrm{SL}_2$.

1.3.1. *Complex abelian variety.* Let A be complex abelian variety of dimension d; we identify A and its set of complex points. Then A is a connected complex Lie group of dimension d, so we have $A \simeq \mathrm{Lie}(A)/\Lambda$, with $\mathrm{Lie}(A) \simeq \mathbb{C}^d$ the universal cover of A and $\Lambda = \pi_1(A) = \mathrm{H}_1(A, \mathbb{Z}) \simeq \mathbb{Z}^{2d}$ a lattice in the underlying $\mathbb{R}$-vector space. Let $A^\vee$ be the dual abelian variety, i.e., the space of degree 0 line bundles on A (see Definition 1.27). We can identify $\mathrm{Lie}(A^\vee)$ with the space of antilinear forms on $\mathrm{Lie}(A)$ and $\mathrm{H}_1(A^\vee, \mathbb{Z})$ with the subspace $\Lambda^\vee$ of forms whose imaginary part takes integer values on Λ (see [Mum08, §9]). For every positive integer n, we have

$$A[n] = \frac{1}{n}\Lambda/\Lambda, \quad A^\vee[n] = \frac{1}{n}\Lambda^\vee/\Lambda^\vee,$$

and the canonical pairing $A[n] \times A^\vee[n] \to \mu_n(\mathbb{C})$ is given by

$$(v, u) \mapsto e^{-2i\pi n\,\mathrm{Im}(u(v))}$$

(see [Mum08, §24]). We then have a bijection between the set of polarizations on A and the set of positive definite Hermitian forms[6] H on $\mathbb{C}^{2d}$ such that the symplectic form $\mathrm{Im}(H)$ takes integer values on Λ; given such a form H, the corresponding isogeny λ_H from A to $A^\vee$ is given on $\mathbb{C}$-points by:

$$\begin{aligned}\lambda_H : \mathrm{Lie}(A)/\Lambda &\longrightarrow \mathrm{Lie}(A^\vee)/\Lambda^\vee \\ w &\longmapsto (v \mapsto H(v,w)).\end{aligned}$$

It follows that the Weil pairing (see Remark 1.30(2)) corresponding to λ_H is the map

$$\begin{aligned}A[n] \times A[n] &\longrightarrow \mu_n(\mathbb{C}) \\ (v,w) &\longmapsto e^{-2i\pi n\,\mathrm{Im}(H(v,w))}.\end{aligned}$$

Note that we have $v, w \in \frac{1}{n}\Lambda$, so $\mathrm{Im}(H(v,w)) \in \frac{1}{n^2}\mathbb{Z}$.

In particular, the polarization λ_H is principal if and only if Λ is self-dual with respect to the symplectic form $\mathrm{Im}(H)$, that is,

$$\Lambda = \{w \in \mathrm{Lie}(A) \mid \forall v \in \Lambda,\ \mathrm{Im}(H(v,w)) \in \mathbb{Z}\}.$$

In that case, the symplectic $\mathbb{Z}$-module $(\Lambda, \mathrm{Im}(H))$ is isomorphic to $\mathbb{Z}^{2d}$ with the form $\psi_\mathbb{Z}$.

Let $\widetilde{\mathcal{M}}_d$ be the set of isomorphism classes of triples $(A, \lambda, \eta_\mathbb{Z})$, where A is a complex abelian variety of dimension d, λ is a principal polarization on A, and $\eta_\mathbb{Z}$ is an morphism of symplectic spaces from $\mathrm{H}_1(A,\mathbb{Z})$ to $(\mathbb{Z}^{2d}, \psi_\mathbb{Z})$. We have an action of $\mathrm{Sp}_{2d}(\mathbb{Z})$ on $\widetilde{\mathcal{M}}_d$: if $c = (A, \lambda, \eta_\mathbb{Z}) \in \widetilde{\mathcal{M}}_d$ and $x \in \mathrm{Sp}_{2g}(\mathbb{Z})$, set $x \cdot c = (A, \lambda, x \circ \eta_\mathbb{Z})$.

If $(A, \lambda, \eta_\mathbb{Z}) \in \widetilde{\mathcal{M}}_d$, then $\Lambda = \mathrm{H}_1(A,\mathbb{Z})$ is a lattice in the real vector space $\mathrm{Lie}(A)$, we have $A = \mathrm{Lie}(A)/\Lambda$ and we can recover the Hermitian form H_λ corresponding to λ from $\mathrm{Im}(H_\lambda)|_\Lambda$, which is sent to the form $\psi_\mathbb{Z}$ on $\mathbb{Z}^{2d}$ by the isomorphism $\eta_\mathbb{Z} : \Lambda \xrightarrow{\sim} \mathbb{Z}^{2d}$. If we see $\mathbb{R}^{2d}$ as a complex vector space via the isomorphisms (of real vector spaces)

$$\mathrm{Lie}(A) \xrightarrow{\sim} \Lambda \otimes_\mathbb{Z} \mathbb{R} \xrightarrow{\sim} \mathbb{R}^{2d},$$

then the Hermitian form on $\mathbb{R}^{2d}$ corresponding to H_λ is

$$(v,w) \longmapsto \psi_\mathbb{R}(iv, w) + i\psi_\mathbb{R}(v,w).$$

So $\eta_\mathbb{Z}$ determines all the data of the isomorphism class of $(A, \lambda, \eta_\mathbb{Z})$, except for the structure of complex vector space on $\mathbb{R}^{2d}$. This structure of complex vector space is equivalent to the data of an $\mathbb{R}$-linear endomorphism J of $\mathbb{R}^{2d}$ such that $J^2 = -1$ (the endomorphism J corresponds to multiplication by i). We also need the $\mathbb{R}$-bilinear map $\mathbb{R}^{2d} \times \mathbb{R}^{2d} \to \mathbb{C}$ defined by $(v,w) \longmapsto \psi_\mathbb{R}(J(v), w) + i\psi_\mathbb{R}(v,w)$ to be a positive definite Hermitian form on $\mathbb{R}^{2d}$. This is equivalent to the following conditions:

(a) $\psi_\mathbb{R}(J(v), J(w)) = \psi_\mathbb{R}(v,w)$ for all $v, w \in \mathbb{R}^{2d}$;
(b) the $\mathbb{R}$-bilinear form $(v,w) \mapsto \psi_\mathbb{R}(J(v), w)$ on $\mathbb{R}^{2d}$ (which is symmetric by (a)) is positive definite.

Conversely, if we have a complex structure J on $\mathbb{R}^{2d}$ satisfying (a) and (b), then we get a positive definite Hermitian form H on $\mathbb{R}^{2d}$ whose imaginary part takes integer values on the lattice $\mathbb{Z}^{2d}$, so the complex torus $\mathbb{R}^{2d}/\mathbb{Z}^{2d}$ has a polarization, hence is an abelian variety (for example by the Kodaira embedding theorem), and we gat an element of $\widetilde{\mathcal{M}}_d$.

So we get a bijection from $\widetilde{\mathcal{M}}_d$ to the set X' of endomorphisms J of $\mathbb{R}^{2d}$ such that $J^2 = -1$ and that J satisfies condition (a) and (b).

Now observe that, if W is a $\mathbb{R}$-vector space, then the data of an endomorphism J of W such that $J^2 = -1$ (i.e. of the structure of a $\mathbb{C}$-vector space on W) is equivalent to the data of a $\mathbb{C}$-linear endomorphism $J_\mathbb{C}$ of $W \otimes_\mathbb{R} \mathbb{C}$ such that

$$\mathrm{Ker}(J_\mathbb{C} - i \cdot \mathrm{id}) = \overline{\mathrm{Ker}(J_\mathbb{C} + i \cdot \mathrm{id})},$$

[6] We take Hermitian forms to be semi-linear in the first variable and linear in the second variable.

where $v \mapsto \overline{v}$ is the involution of $W \otimes_{\mathbb{C}} \mathbb{C}$ induced by complex conjugation on $\mathbb{C}$. This is equivalent to giving a $\mathbb{C}$-vector subspace E of $W \otimes_{\mathbb{R}} \mathbb{C}$ such that $W \otimes_{\mathbb{R}} \mathbb{C} = E \oplus \overline{E}$, i.e., a d-dimensional complex subspace E of $W \otimes_{\mathbb{R}} \mathbb{C}$ such that $E \cap \overline{E} = \{0\}$.[7]

We apply this to $W = \mathbb{R}^{2d}$. Let J be a complex structure on $\mathbb{R}^{2d}$, and let E be the corresponding $\mathbb{C}$-vector subspace of $\mathbb{C}^{2d}$. Then condition (a) on J is equivalent to the fact that:

(a') $\psi_{\mathbb{C}}(v, w) = 0$ for all $v, w \in E$,

(i.e., to the fact that E is a Lagrangian subspace[8] of $\mathbb{C}^{2d}$), and condition (b) on J is equivalent to the fact that

(b') $-i\psi_{\mathbb{C}}(v, \overline{v}) \in \mathbb{R}_{>0}$ for all $v \in E \backslash \{0\}$.

Note that these two conditions on a $\mathbb{C}$-vector subspace E of $\mathbb{C}^{2d}$ imply that $E \cap \overline{E} = \{0\}$. So we get a bijection from X' to the set of Lagrangian subspaces E of $V_{\mathbb{C}}$ satisfying (b').

If we represent Lagrangian subspaces of $\mathbb{C}^{2d}$ by their bases, sees as complex matrices of size $d \times 2d$, then the action of $\mathrm{Sp}_{2d}(\mathbb{R})$ is just left multiplication. For example, the subspace E_0 corresponding to $J_d \in X'$ is the one with basis $\begin{pmatrix} iI_d \\ I_d \end{pmatrix}$.

More generally, if $Y \in \mathfrak{h}_d^+$, the subspace of $\mathbb{C}^{2d}$ with basis $\begin{pmatrix} Y \\ I_d \end{pmatrix}$ is a Lagrangian subspace satisfying condition (b'), and every such Lagrangian subspace is of that form. So we get bijections

$$\widetilde{\mathcal{M}}_d \simeq X' \simeq \mathfrak{h}_d^+,$$

and we can check that the second bijection is $\mathrm{Sp}_{2d}(\mathbb{R})$-equivariant. Unraveling the definitions, we see that $Y \in \mathfrak{h}_d^+$ corresponds to the element $(A_Y, \lambda_Y, \eta_{\mathbb{Z},Y})$ of $\widetilde{\mathcal{M}}_d$ such that $A_Y = \mathbb{C}^d/(\mathbb{Z}^d + Y\mathbb{Z}^d)$, λ_Y is the principal polarization given by the Hermitian form with matrix $\mathrm{Im}(Y)^{-1}$ on $\mathbb{C}^d$, and $\eta_{\mathbb{Z},Y} : \mathbb{Z}^d + Y\mathbb{Z}^d \xrightarrow{\sim} \mathbb{Z}^{2d}$ is the isomorphism sending $a \in \mathbb{Z}^d$ to $(a, 0) \in \mathbb{Z}^d \times \mathbb{Z}^d = \mathbb{Z}^{2d}$ and $Ya \in Y\mathbb{Z}^d$ to $(0, a) \in \mathbb{Z}^d \times \mathbb{Z}^d = \mathbb{Z}^{2d}$.

Now we want an interpretation of the quotients $\Gamma \backslash \mathfrak{h}_d^+$, for Γ an arithmetic subgroup of $\mathrm{Sp}_{2d}(\mathbb{Q})$. We will do this for the groups $\Gamma(n) = \mathrm{Ker}(\mathrm{Sp}_{2d}(\mathbb{Z}) \to \mathrm{Sp}_{2d}(\mathbb{Z}/n\mathbb{Z}))$, where n is a positive integer (and $\Gamma(n)$ is called the **principal congruence subgroup** at level n). Note that any arithmetic group contains $\Gamma(n)$ for n divisible enough.

We will need the notion of a level structure; we give the general definition here.

Definition 1.7. Let S be a scheme, (A, λ) be a principally polarized abelian scheme of relative dimension d over S, and n be a positive integer. Than a **level n structure** on (A, λ) is a couple (η, φ), where

$$\eta : A[n] \xrightarrow{\sim} \underline{\mathbb{Z}/n\mathbb{Z}}_S^{2g}, \quad \varphi : \underline{\mathbb{Z}/n\mathbb{Z}}_S \xrightarrow{\sim} \mu_{n,S}$$

are isomorphisms of group schemes such that $\varphi \circ \psi_{\mathbb{Z}/n\mathbb{Z}} \circ \eta$ is the Weil pairing associated to λ on $A[n]$.

Remark 1.8. A level n structure on (A, λ) can only exist if n is invertible on S and $\mu_{n,S}$ is a constant group scheme.

Note that isomorphisms $\varphi : \underline{\mathbb{Z}/n\mathbb{Z}}_S \xrightarrow{\sim} \mu_{n,S}$ correspond to sections $\zeta \in \mu_n(S)$ generating $\mu_{n,S}$ (i.e. to primitive nth roots of 1 over S), by sending φ to $\zeta = \varphi(1)$. So we will also see level structures as couples (η, ζ), with $\zeta \in \mu_n(S)$ primitive.

Let $\zeta_n = e^{-2i\pi/n} \in \mu_n(\mathbb{C})$. If $Y \in \mathfrak{h}_d^+$, then $\frac{1}{n}\eta_{\mathbb{Z},Y}$ defines an isomorphism of groups $\eta_Y : A_Y[n] \xrightarrow{\sim} (\mathbb{Z}/n\mathbb{Z})^{2d}$, and it follows from formula

$$A[n] \times A[n] \to \mu_n(\mathbb{C}), \quad (v, w) \mapsto e^{-2i\pi n \,\mathrm{Im}(H(v,w))}$$

[7]In fancy terms, we are saying that putting a structure of complex vector space on W is the same as putting a pure Hodge structure of type $\{(-1, 0), (0, -1)\}$ on it (or of type $\{(1, 0), (0, 1)\}$, depending on your normalization). When $W = \mathbb{R}^{2d}$ and the complex structure comes from an element $(A, \lambda, \eta_{\mathbb{Z}})$ of $\widetilde{\mathcal{M}}_d$, then this Hodge structure is the one induced by the isomorphism

$$\mathrm{H}_1(A, \mathbb{R}) \xrightarrow{\eta_{\mathbb{Z}} \otimes \mathbb{R}} \mathbb{Z}^{2d} \otimes_{\mathbb{Z}} \mathbb{R} \xrightarrow{\sim} \mathbb{R}^{2d}.$$

[8]By definition, a maximal isotropic subspace.

that (η, ζ_n) is a level n structure on (A_Y, η_Y).

Using the fact that $\mathrm{Sp}_{2d}(\mathbb{Z}) \to \mathrm{Sp}_{2d}(\mathbb{Z}/n\mathbb{Z})$ is surjective for every $n \in \mathbb{N}$, which follows from strong approximation for Sp_{2d},[9] we finally get:

Proposition 1.9. *Let n be a positive integer. The map $Y \mapsto (A_Y, \lambda_Y, \eta_Y)$ induces a bijection from $\Gamma(n)\backslash\mathfrak{h}_d^+$ to the set of isomorphism classes of triples (A, λ, η), where (A, λ) is a principally polarized complex abelian variety of dimension d and $\eta : A[n] \xrightarrow{\sim} (\mathbb{Z}/n\mathbb{Z})^{2d}$ is an isomorphism of groups such that (η, ζ_n) is a level n structure on (A, λ).*

Now there is an obvious way to make $\Gamma(n)\backslash\mathfrak{h}_d^+$ into an algebraic variety.

1.3.2. *The connected Siegel modular variety.* Let $\mathcal{O}_n = \mathbb{Z}[1/n][T]/(T^n - 1)$. If S is a scheme over $\mathcal{O}_n$, we denote by $\varphi_0 : \underline{\mathbb{Z}/n\mathbb{Z}}_S \xrightarrow{\sim} \mu_{n,S}$ the isomorphism sending 1 to the class of T.

Definition 1.10. Let $\mathcal{M}'_{d,n}$ be the functor from the category of $\mathcal{O}_n$-schemes to the category of sets sending S to the set of isomorphisms classes of triples (A, λ, η), where (A, η) is a principally polarized abelian scheme of relative dimension d over S and $\eta : A[n] \xrightarrow{\sim} \underline{\mathbb{Z}/n\mathbb{Z}}_S^{2g}$ is an isomorphism of group schemes such that (η, φ_0) is a level n structure on (A, λ).

An isomorphism from (A, λ, η) to (A', λ', η') is an isomorphism of abelian varieties $u : A \xrightarrow{\sim} A'$ such that $\lambda' \circ u = u^\vee \circ \lambda$ and $\eta' = \eta \circ (u, u)$.

Theorem 1.11 (Mumford, cf. [FC90])**.** *Suppose that $n \geqslant 3$. Then the functor $\mathcal{M}'_{d,n}$ is representable by a smooth quasi-projective $\mathcal{O}_n$-scheme purely of dimension $d(d+1)/2$ and with connected geometric fibers, which we still denote by $\mathcal{M}'_{d,n}$ and call the* ***connected Siegel modular variety*** *of level n.*

Remark 1.12. If $n \in \{1, 2\}$, then triples (A, λ, η) as in Definition 1.10 may have automorphisms, so we should see $\mathcal{M}'_{d,n}$ as a stack. This stack will then be representable by a smooth Deligne–Mumford stack over $\mathcal{O}_n$ that is a finite étale quotient of the scheme $\mathcal{M}'_{d,3n}$.

We can now reformulate Proposition 1.9 in the following way.

Proposition 1.13. *Let $n \geqslant 3$ be an integer. Then the map $Y \mapsto (A_Y, \lambda_Y, \eta_Y)$ induces an isomorphism of complex manifolds from $\Gamma(n)\backslash\mathfrak{h}_d^+$ to $\mathcal{M}'_{d,n}(\mathbb{C})$.*

The fact that this is an isomorphism of complex manifolds is clear on the explicit formula for the bijection $\Gamma(n)\backslash\mathfrak{h}_d^+ \to \mathcal{M}'_{d,n}(\mathbb{C})$.

In particular, we showed that $\Gamma(n)\backslash\mathfrak{h}_d^+$ is the set of complex points of an algebraic variety defined over the number field $\mathbb{Q}(\zeta_n)$. Unfortunately, this number field depends on the level n. The issue is that we need a fixed primitive nth root of 1 in order to define the moduli problem $\mathcal{M}'_{d,n}$, so we need to be over a basis where such a primitive nth root exists. To fix this problem, we will allow the primitive nth root of 1 to vary.

1.3.3. *The Siegel modular variety.*

Definition 1.14. Let n be a positive integer. The **Siegel modular variety** $\mathcal{M}_{d,n}$ is the functor from the category of $\mathbb{Z}/n\mathbb{Z}$-schemes to the category of sets sending a scheme S to the set of isomorphism classes of triples $(A, \lambda, \eta, \varphi)$, where (A, λ) is a principally polarized abelian scheme of relative dimension d over S and (η, φ) is a level n structure on (A, λ).

An isomorphism from (A, λ, η) to (A', λ', η') is an isomorphism of abelian varieties $u : A \xrightarrow{\sim} A'$ such that $\lambda' \circ u = u^\vee \circ \lambda$ and $\eta' = \eta \circ (u, u)$.

The group $\mathrm{GSp}_{2d}(\widehat{\mathbb{Z}})$ acts on $\mathcal{M}_{d,n}$: if $g \in \mathrm{GSp}_{2d}(\widehat{\mathbb{Z}})$ and $(A, \lambda, \eta, \varphi) \in \mathcal{M}_{d,n}(S)$, then

$$g \cdot (A, \lambda, \eta, \varphi) = (A, \lambda, g \circ \eta, c(g)^{-1}\varphi).$$

The kernel of this action is the group

$$K(n) = \mathrm{Ker}(\mathrm{GSp}_{2d}(\widehat{\mathbb{Z}}) \to \mathrm{GSp}_{2d}(\mathbb{Z}/n\mathbb{Z})).$$

[9] See [Pla69] and [Pla70].

If n divides m, then we have a morphism $\mathcal{M}_{d,m} \to \mathcal{M}_{d,n}$ that forgets part of the level m structure; this morphism is (representable) finite étale, and in fact it is a torsor under the finite group $K(n)/K(m)$.

We have the following variant of Theorem 1.11.

Theorem 1.15 (Mumford, cf. [FC90])**.** *Suppose that $n \geqslant 3$. Then the functor $\mathcal{M}_{d,n}$ is representable by a smooth quasi-projective $\mathcal{O}_n$-scheme purely of dimension $d(d+1)/2$, which we will denote by $\mathcal{M}_{d,n}$ and call the* ***Siegel modular variety*** *of level n.*

Remark 1.16. Let $K(n) = \operatorname{Ker}(\mathrm{GSp}_{2d}(\widehat{\mathbb{Z}}) \to \mathrm{GSp}_{2d}(\mathbb{Z}/n\mathbb{Z}))$. Then $\mathcal{M}_{d,n}$ is the Shimura variety for GSp_{2d} with level $K(n)$, or rather its integral model. If K is an open compact subgroup of $\mathrm{GSp}_{2d}(\mathbb{A}_f)$ that is small enough,[10] then we can also define the Shimura variety $\mathcal{M}_{d,K,\mathbb{Q}}$ with level K: choose n such that $K(n) \subset K$. Then $K(n)$ is a normal subgroup of K, so the group $K/K(n) \subset \mathrm{GSp}_{2d}(\mathbb{Z}/n\mathbb{Z})$ acts on $\mathcal{M}_{d,K,\mathbb{Q}}$, and we set $\mathcal{M}_{d,K,\mathbb{Q}} = \mathcal{M}_{d,n,\mathbb{Q}}/(K/K(n))$. It is easy to check that this does not depend on the choice of n.

In fact, for K an open compact subgroup of $\mathrm{GSp}_{2d}(\mathbb{A}_f)$, we have a direct definition of a level K structure on a principally polarized abelian scheme (see [Kot92b, §5]). For K small enough, the scheme $\mathcal{M}_{d,K,\mathbb{Q}}$ is the moduli space of a principally polarized abelian schemes with level K structure. In general, this moduli space is representable by a Deligne–Mumford stack. We can also define this moduli schemes over a localization of $\mathbb{Z}$, but the primes that we invert depend on K; see the discussion in Subsubsection 2.1.4.

Let us explain the relationship between $\mathcal{M}_{d,n}$ and $\mathcal{M}'_{d,n}$. We define a map

$$s : (\mathbb{Z}/n\mathbb{Z})^\times \longrightarrow \mathrm{GSp}_{2d}(\mathbb{Z}/n\mathbb{Z}), \quad s(\alpha) = \begin{pmatrix} 0 & \alpha I_d \\ I_d & 0 \end{pmatrix};$$

note that s is a section of the multiplier $c : \mathrm{GSp}_{2d}(\mathbb{Z}/n\mathbb{Z}) \to (\mathbb{Z}/n\mathbb{Z})^\times$, and that it is not a morphism of groups.

Proposition 1.17. *The morphism*

$$\begin{aligned} \mathcal{M}'_{d,n} \times (\mathbb{Z}/n\mathbb{Z})^\times &\longrightarrow \mathcal{M}_{d,n,\mathcal{O}_n} \\ ((A,\lambda,\eta),\alpha) &\longmapsto (A,\lambda,s(\alpha)\circ\eta,\varphi_0\circ\alpha), \end{aligned}$$

where we see α as an automorphism of $\underline{\mathbb{Z}/n\mathbb{Z}}_S$ for any scheme S, is an isomorphism.

As a corollary, we get a description of the complex points of $\mathcal{M}_{d,n}$. Let $\mathfrak{h}_d = \mathfrak{h}_d^+ \cup (-\mathfrak{h}_d^+)$ be the set of symmetric matrices $Y \in M_d(\mathbb{C})$ such that $\mathrm{Im}(Y)$ is positive definite or negative definite. The action of $\mathrm{Sp}_{2d}(\mathbb{R})$ on $\mathfrak{h}_d$ extends to a transitive action of $\mathrm{GSp}_{2d}(\mathbb{R})$, given by the same formula. The stabilizer of $iI_d \in \mathfrak{h}_d$ in $\mathrm{GSp}_{2d}(\mathbb{R})$ is $\mathbb{R}_{>0}K_\infty$, so $\mathfrak{h}_d \simeq \mathrm{GSp}_{2d}(\mathbb{R})/\mathbb{R}_{>0}K_\infty$ as real analytic manifolds.

Corollary 1.18. *We have an isomorphism of complex manifolds*

$$\mathcal{M}_{d,n}(\mathbb{C}) \simeq \mathrm{GSp}_{2d}(\mathbb{Q})\backslash(\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n))$$

extending the isomorphism of Proposition 1.9, where $K(n) = \operatorname{Ker}(\mathrm{GSp}_{2d}(\widehat{\mathbb{Z}}) \to \mathrm{GSp}_{2d}(\mathbb{Z}/n\mathbb{Z}))$ and $\mathrm{GSp}_{2d}(\mathbb{Q})$ acts diagonally on $\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)$.

This follows from the fact that

$$\mathrm{GSp}_{2d}(\mathbb{Q})\backslash(\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n)) \simeq \mathrm{GSp}_{2d}(\mathbb{Q})^+\backslash(\mathfrak{h}_d^+ \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n)),$$

where $\mathrm{GSp}_{2d}(\mathbb{Q})^+ = \{g \in \mathrm{GSp}_{2d}(\mathbb{Q}) \mid c(g) > 0\}$, and from strong approximation for Sp_{2d},[11] which implies that c induces a bijection

$$\mathrm{GSp}_{2d}(\mathbb{Q})^+\backslash\mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n) \xrightarrow{\sim} \mathbb{Q}_{>0}\backslash\mathbb{A}_f^\times/c(K(n)) \simeq \widehat{\mathbb{Z}}^\times(1+n\widehat{\mathbb{Z}}) \simeq (\mathbb{Z}/n\mathbb{Z})^\times.$$

For every $i \in (\mathbb{Z}/n\mathbb{Z})^\times$, we choose $x_i \in \mathrm{GSp}_{2d}(\mathbb{A}_f)$ lifting i and we set

$$\Gamma(n)_i = \mathrm{GSp}_{2d}(\mathbb{Q})^+ \cap x_iK(n)x_i^{-1} = \mathrm{Sp}_{2d}(\mathbb{Q}) \cap x_iK(n)x_i^{-1}.$$

[10] For example, $K \subset K(n)$ with $n \geqslant 3$.

[11] See [Pla69] and [Pla70].

Then the $\Gamma(n)_i$ are arithmetic subgroups of $\mathrm{Sp}_{2d}(\mathbb{Q})$, and we have

$$\mathrm{GSp}_{2d}(\mathbb{Q})\backslash(\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n)) \simeq \coprod_{i\in(\mathbb{Z}/n\mathbb{Z})^\times} \Gamma(n)_i\backslash\mathfrak{h}_d^+$$

as complex manifolds.

In fact, for $i \in (\mathbb{Z}/n\mathbb{Z})^\times$, we are supposed to take $x_i = \begin{pmatrix} 0 & a_iI_d \\ I_d & 0 \end{pmatrix}$ with $a_i \in \widehat{\mathbb{Z}}^\times$ lifting i. In particular, we have $x_i \in \mathrm{GSp}_{2d}(\widehat{\mathbb{Z}})$; as $K(n)$ is a normal subgroup of $\mathrm{GSp}_{2d}(\widehat{\mathbb{Z}})$, we get $x_iK(n)x_i^{-1} = K(n)$, hence $\Gamma(n)_i = \Gamma(n)$, and finally

$$\mathrm{GSp}_{2d}(\mathbb{Q})\backslash(\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n)) \simeq \coprod_{i\in(\mathbb{Z}/n\mathbb{Z})^\times} \Gamma(n)_i\backslash\mathfrak{h}_d^+.$$

Remark 1.19. If K is a small enough open compact subgroup of $\mathrm{GSp}_{2d}(\mathbb{A}_f)$, then we get an isomorphism of complex manifolds:

$$\mathcal{M}_{d,K}(\mathbb{C}) \simeq \mathrm{GSp}_{2d}(\mathbb{Q})\backslash(\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K).$$

1.3.4. *Heche correspondence.* We can also descend the Hecke correspondences before to morphisms of schemes over $\mathbb{Z}[1/n]$.

We proceed as in [Lau05, §3]. Let $g \in \mathrm{GSp}_{2d}(\mathbb{A}_f)$, and let K, K' be small enough open compact subgroups of $\mathrm{GSp}_{2d}(\mathbb{A}_f)$ such that $K' \subset K \cap gKg^{-1}$. We want to define finite étale morphisms $T_1, T_g : \mathcal{M}_{d,K'} \to \mathcal{M}_{d,K}$, and the Hecke correspondence associated to (g, K, K') is the couple (T_1, T_g).

Choose $n \geqslant 3$ such that $K(n) \subset K'$; then $\mathcal{M}_{d,K'} = \mathcal{M}_{d,n}/(K'/K(n))$ and $\mathcal{M}_{d,K} = \mathcal{M}_{d,n}/(K/K(n))$. The morphism T_1 just forgets part of the level structure: as $K'/K(n) \to K/K(n)$, we have an obvious morphism $T_1 : \mathcal{M}_{d,K'} \to \mathcal{M}_{d,K}$.

To define T_g, we first consider the following special case: if $g \in \mathrm{GL}_{2d}(\widehat{\mathbb{Z}}) \cap \mathrm{GSp}_{2d}(\mathbb{A}_f)$, let $x = (A, \lambda, \eta, \varphi) \in \mathcal{M}_{d,n}(S)$. Let u be the endomorphism of $\underline{\mathbb{Z}/n\mathbb{Z}}^{2d}$ with matrix g. Then T_g sends the class of x in $\mathcal{M}_{d,K'}(S)$ to the class of $(A', \lambda', \eta', \varphi) \in \mathcal{M}_{d,n}(S)$, where A' is the quotient of A by the finite flat subgroup scheme $\eta^{-1}(\mathrm{Ker}\, u)$ of $A[n]$ and λ', η' are the morphisms deduced from λ, $g \circ \eta$.

Note that, if $g = aI_{2d}$ with $a \in \widehat{\mathbb{Z}} \cap \mathbb{A}_f^\times$ and $K' = K$, then the morphisms $T_g : \mathcal{M}_{d,K} \to \mathcal{M}_{d,K}$ is an isomorphism.

Finally, for a general $g \in \mathrm{GSp}_{2d}(\mathbb{A}_f)$, we write $g = a^{-1}g_0$ with $a \in (\widehat{\mathbb{Z}} \cap \mathbb{A}_f^\times)I_{2d}$ and $g_0 \in \mathrm{GL}_{2d}(\widehat{\mathbb{Z}}) \cap \mathrm{GSp}_{2d}(\mathbb{A}_f)$, and we set $T_g = T_{g_0} \circ T_a^{-1}$.

Remark 1.20. If we use instead the general definition of a level K structure from [Kot92b, §5], then it becomes much easier to define the Hecke correspondences; see Section 6 ibid..

Remark 1.21. We have two ways to think of $\mathcal{M}_{d,n}(\mathbb{C})$: as an adelic double quotient or as finite disjoint union of spaces $\Gamma(n)\backslash\mathfrak{h}_d^+$, which are locally symmetric spaces associated to the semi-simple group Sp_{2d}. The first description is more convenient to see the action of adelic Hecke operators, and the second description is a bit more concrete and has simpler combinatorics. Note also that the complex manifold $\Gamma(n)\backslash\mathfrak{h}_d^+$ is isomorphic to the set of $\mathbb{C}$-points of the algebraic variety $\mathcal{M}'_{d,n}$, but this algebraic variety is defined over the field $\mathbb{Q}[T]/(T^n - 1)$, which depends on n. On the other hand, the adelic double quotient $\mathrm{GSp}_{2d}(\mathbb{Q})\backslash(\mathfrak{h}_d \times \mathrm{GSp}_{2d}(\mathbb{A}_f)/K(n)$ is isomorphic to the set of $\mathbb{C}$-points of the algebraic variety $\mathcal{M}_{d,n}$, which is defined over $\mathbb{Q}$. So if we want to consider Shimura varieties as a projective system of algebraic varieties over a number field, then it makes sense to use the adelic double quotients, because they are all defined on the same field.

1.3.5. *Background on abelian schemes.* Let S be a scheme. We denote by Sch/S the category of S-schemes.

Definition 1.22. (1) An **abelian scheme** over S is an S-group scheme $A \to S$ which is smooth and proper with geometrically connected fibers. If S is the spectrum of a field k, an abelian scheme over S is also called an **abelian variety** over k.

(2) A **morphism of abelian schemes** over S is a morphism of S-group schemes between abelian schemes over S.

Proposition 1.23. *Let A be an abelian scheme over S. Then*

(1) *The S-group scheme $A \to S$ is commutative;*
(2) *The morphism $A \to S$ has connected fibers.*

Definition 1.24. Let A be an abelian scheme over S, and let $e : S \to A$ be its zero section. We consider the following two functors from $(\mathsf{Sch}/S)^{\mathrm{op}}$ to the category of sets:

(a) The functor $\mathrm{Pic}_{A/S,e}$ sending an S-scheme $T \to S$ to the set of isomorphism classes of couples $(\mathcal{L}, \varphi)$, where $\mathcal{L}$ is an invertible sheaf on $A \times_S T$, $e_T = e \times_S T : T \to A \times_S T$, and $\varphi : \mathcal{O}_T \xrightarrow{\sim} e_T^* \mathcal{L}$ is an isomorphism. An isomorphism from $(\mathcal{L}, \varphi)$ to $(\mathcal{L}', \varphi')$ is an isomorphism of $\mathcal{O}_T$-modules $\alpha : \mathcal{L} \xrightarrow{\sim} \mathcal{L}'$ such that $(e_T^*\alpha) \circ \varphi = \varphi'$;
(b) The subfunctor $\mathrm{Pic}^0_{A/S,e}$ of $\mathrm{Pic}_{A/S,e}$ sending sending an S-scheme $T \to S$ to the set of isomorphism classes of couples $(\mathcal{L}, \varphi)$ as in (a) such that, for every point t of T, every smooth projective curve C over the residue field $\kappa(t)$ of t, and every morphism of $\kappa(t)$-schemes $f : C \to A \times_S t$, the line bundle $f^*(\mathcal{L}|_{A \times_S t})$ is of degree 0 on C.

Remark 1.25. (1) The functor $\mathrm{Pic}_{A/S,e}$ can be made into a functor into the category of abelian groups: if T is an S-scheme and $(\mathcal{L}, \varphi)$, $(\mathcal{L}', \varphi')$ represent elements of $\mathrm{Pic}_{A/S,e}(T)$, their product is represented by $(\mathcal{L} \otimes_{\mathcal{O}_T} \mathcal{L}', \varphi \otimes \varphi')$, where $\varphi \otimes \varphi'$ is the isomorphism

$$\mathcal{O}_T = \mathcal{O}_T \otimes_{\mathcal{O}_T} \mathcal{O}_T \xrightarrow[\varphi\otimes\varphi']{\sim} (e_T^*\mathcal{L}) \otimes_{\mathcal{O}_T} (e_T^*\mathcal{L}') = e_T^*(\mathcal{L} \otimes_{\mathcal{O}_T} \mathcal{L}').$$

Moreover, for every S-scheme T, the set $\mathrm{Pic}^0_{A/S,e}(T)$ is a subgroup of $\mathrm{Pic}_{A/S,e}(T)$.

(2) If $X \to S$ is a scheme over S, then the relative Picard functor on Sch/S is the fppf sheafification of the functor $T \mapsto \mathrm{Pic}(X \times_S T)$ (where, for Y a scheme, we denote by $\mathrm{Pic}(Y)$ the set of isomorphism classes of line bundles on Y, that is an abelian group for the tensor product); see [Sta, Situation 0D25]. We can also define a subfunctor $\mathrm{Pic}^0_{X/S}$ of $\mathrm{Pic}_{X/S}$ as in Definition 1.24. By [Sta, Lemma 0D28], if A is an abelian scheme over S, then there is an isomorphism of functors in abelian groups $\mathrm{Pic}_{A/S} \xrightarrow{\sim} \mathrm{Pic}_{A/S,e}$, inducing an isomorphism $\mathrm{Pic}^0_{A/S} \xrightarrow{\sim} \mathrm{Pic}^0_{A/S,e}$.

(3) We can upgrade $A \mapsto \mathrm{Pic}_{A/S,e}$ and $A \mapsto \mathrm{Pic}^0_{A/S,e}$ to contravariant functors in A: if $f : A \to B$ is a morphism of abelian schemes over S, then it induces a natural transformation $f^* : \mathrm{Pic}_{B/S,e} \to \mathrm{Pic}_{A/S,e}$ sending $(\mathcal{L}, \varphi)$ to $(f^*(\mathcal{L}), f^*(\varphi))$, and f^* sends $\mathrm{Pic}^0_{B/S,e}$ to $\mathrm{Pic}^0_{A/S,e}$.

Theorem 1.26. *Let A be an abelian scheme over S. Then $\mathrm{Pic}^0_{A/S,e}$ is representable by an abelian scheme over S.*

Proof. We know that $\mathrm{Pic}^0_{A/S,e}$ is representable by an algebraic space by a result of M. Artin (see [Art69] or [Sta, Lemma 0D2C]). We can check on the moduli problem that this algebraic space is proper and smooth, and its fibers over points of S are abelian varieties by the classical theory of the dual abelian variety (see sections II.8 and III.13 of Mumford's book [Mum08]). It remains to prove that the algebraic space representing $\mathrm{Pic}^0_{A/S,e}$ is a scheme; this is due to Raynaud, and a proof is given in [FC90, Theorem 1.9]. □

Definition 1.27. Let A be an abelian scheme over S. The abelian scheme over S representing $\mathrm{Pic}^0_{A/S,e}$ is called the **dual abelian scheme** of A and denoted by $A^\vee$. In particular, we get a couple $(\mathcal{P}_A, \varphi_A)$ representing the element of $\mathrm{Pic}^0_{A/S,e}(A^\vee)$ corresponding to $\mathrm{id}_{A^\vee}$, with $\mathcal{P}_A$ a line bundle on $A \times_S A^\vee$, called the **Poincaré line bundle**.

If $f : A \to B$ is a morphism of abelian schemes over S, we denote by $f^\vee : B^\vee \to A^\vee$ the morphism corresponding to the natural transformation $f^* : \mathrm{Pic}^0_{B/S,e} \to \mathrm{Pic}^0_{A/S,e}$ of Remark 1.25.

Remark 1.28. Let $e : S \to A^\vee$ be the unit section. Then the pullback of $\mathcal{P}_A$ by $A \times_S e : A = A \times_S A \to A \times_S A^\vee$ is the line bundle on A corresponding to the element e of $A^\vee(S) = \mathrm{Pic}^0_{A/S,e}(S)$; in other words, it is isomorphic to the trivial line bundle $\mathcal{O}_A$. So $\mathcal{P}_A$ defines an element of $\mathrm{Pic}^0_{A/S,e}(A)$, that is, a morphism of S-schemes $A \to A^{\vee\vee}$, called the **biduality morphism**. The **biduality theorem** says that the biduality morphism is an isomorphism. For S the spectrum of a field, this is proved in [Mum08, Section III.13], and the general case reduces to this by looking at the fibers of points of S.

Let A be an abelian scheme over S and let $\mathcal{L}$ be a line bundle on A. We denote by $\mu, p_1, p_2, \varepsilon : A \times_S A \to A$ the addition morphism, the first projection, the second projection and the zero morphism respectively. Then the line bundle $(\mu^*\mathcal{L}) \otimes (p_1^*\mathcal{L}^{\otimes -1}) \otimes (p_2^*\mathcal{L}^{\otimes -1}) \otimes (\varepsilon^*\mathcal{L})$ on $A \times_S A$ is trivial when restricted to $S \times_S A$ via the zero section of A, hence it defines an element of $\mathrm{Pic}^0_{A/S,e}(A)$, corresponding to a morphism of S-schemes $\lambda(\mathcal{L}) : A \to A^\vee$. Moreover, the theorem of the cube (see for example [Mum08, Section III.10]).

Definition 1.29. Let A be an abelian scheme over S. A **polarization** on A is a morphism of abelian schemes $\lambda : A \to A^\vee$ such that, for every algebraically closed field k and every morphism $\operatorname{Spec} k \to S$, the morphism $\lambda \times_S \operatorname{Spec} k : A \times_S \operatorname{Spec} k \to A^\vee \times_S \operatorname{Spec} k = (A \times_S \operatorname{Spec} k)^\vee$ is of the form $\lambda(\mathcal{L})$, for $\mathcal{L}$ an ample line bundle on $A \times_S \operatorname{Spec} k$. We say that a polarization is **principal** if it is an isomorphism.

A **principally polarized abelian scheme** over S is a pair (A, λ), where A is an abelian scheme over S and λ is a principal polarization on A.

Remark 1.30.
(1) A polarization on A is always an isogeny, i.e. finite and faithfully flat.
(2) Let λ be a polarization on A, and let n be a positive integer. Then, composing $\lambda : A[n] \to A^\vee[n]$ with the canonical pairing $A[n] \times A^\vee[n] \to \mu_{n,S}$, we get a pairing $A[n] \times A[n] \to \mu_{n,S}$, called the **Weil pairing** associated to λ. If λ is principal, this is a perfect pairing.

1.4. **Shimura varieties over $\mathbb{C}$.** Remember the upshot of Subsection 1.2: if we want algebraic varieties that are all defined over the same number field, and Hecke correspondences that are also defined on this number field, it is better to work with adelic double quotient for a reductive group such as GSp_{2d} rather than with locally symmetric spaces for a semi-simple group such as Sp_{2d}. This (and Theorem 1.3) motivates the definition of Shimura data, due to Deligne in [Del71].

1.4.1. *The Serre torus.* Let $\mathbb{S}$ be $\mathbb{C}^\times$ seen as an algebraic group over $\mathbb{R}$; this is called the **Serre torus**. In other words, the group $\mathbb{S}$ is the Weil restriction of scalars from $\mathbb{C}$ to $\mathbb{R}$ of GL_1, so that $\mathbb{S}(R) = (R \otimes_\mathbb{R} \mathbb{C})^\times$ for every $\mathbb{R}$-algebra R. We denote by w the injective morphism $\mathrm{GL}_{1,\mathbb{R}} \to \mathbb{S}$ corresponding to the inclusion $\mathbb{R}^\times \subset \mathbb{C}^\times$.

We have $\mathbb{S}(\mathbb{C}) = (\mathbb{C} \otimes_\mathbb{R} \mathbb{C})^\times \xrightarrow{\sim} \mathbb{C}^\times \times \mathbb{C}^\times$, where the isomorphism sends $a \otimes 1 + b \otimes i$ to $(a+ib, a-ib)$. So the abelian group $\mathrm{Hom}(\mathbb{S}_\mathbb{C}, \mathrm{GL}_{1,\mathbb{C}})$ of characters of $\mathbb{S}$ is free of rank 2 and generated by the characters z and $\overline{z}$ corresponding to the two projections of $\mathbb{C}^\times \times \mathbb{C}^\times$ on $\mathbb{C}^\times$. We denote by $r : \mathrm{GL}_{1,\mathbb{C}} \to \mathbb{S}_\mathbb{C}$ the injective morphism corresponding to the injection of the first factor in $\mathbb{C}^\times \times \mathbb{C}^\times$.

If V is a real vector space and $\rho : \mathbb{S} \to \mathrm{GL}(V)$ is a morphism of algebraic groups (i.e. a representation of $\mathbb{S}$ on V), then we have $V_\mathbb{C} := V \otimes_\mathbb{R} \mathbb{C} = \bigoplus_{p,q \in \mathbb{Z}} V^{p,q}$, where $V^{p,q}$ is the subspace of $V_\mathbb{C}$ on which $\mathbb{S}_\mathbb{C}$ acts by the character $z^{-p}\overline{z}^{-q}$; moreover, as ρ is defined over $\mathbb{R}$, we have $\overline{V^{p,q}} = V^{q,p}$ for all $p, q \in \mathbb{Z}$. Let $m \in Z$. We say that ρ is **of weight** m if $\rho \circ r : \mathrm{GL}_1 \to \mathrm{GL}(V)$ is equal to $x \mapsto x^m \mathrm{id}_V$.

Remark 1.31. If $\rho : \mathbb{S} \to \mathrm{GL}(V)$ is of weight m, we have $V^{p,q} = 0$ unless $p+q = m$, so the decomposition $V_\mathbb{C} = \bigoplus_{p,q} V^{p,q}$ is a pure Hodge structure of weight m on V. In fact, representations of weight m of $\mathbb{S}$ on V are in bijection with pure Hodge structures of weight m on V.

1.4.2. *Shimura data.*

Definition 1.32. A **Shimura datum** is a couple (G, h), where G is a connected reductive algebraic group over $\mathbb{Q}$ and $h : \mathbb{S} \to G_\mathbb{R}$ is a morphism of real algebraic groups such that:
(a) The image of $h \circ w : \mathrm{GL}_{1,\mathbb{R}} \to G_\mathbb{R}$ is central;
(b) If $\mathfrak{g} = \mathrm{Lie}(G_\mathbb{R})$ and $\mathfrak{g}_\mathbb{C} = \bigoplus_{p,q \in \mathbb{Z}} \mathfrak{g}^{p,q}$ is the decomposition induced by the representation $\mathrm{Ad} \circ h : \mathbb{S} \to \mathrm{GL}(\mathfrak{g})$, then we have $\mathfrak{g}^{p,q} = 0$ unless $(p,q) \in \{(-1,1), (0,0), (1,-1)\}$;
(c) Conjugation by $h(i)$ induces a Cartan involution of $G_{\mathrm{der}}(\mathbb{R})$ (see Theorem 1.3);
(d) G_{der} has no normal subgroup (defined over $\mathbb{Q}$) whose group of R-points is compact.[12]

Let (G, h) be a Shimura datum. We denote by K_∞ the centralizer of h in $G(\mathbb{R})$ and by X the set of $G(\mathbb{R})$-conjugates of h. Then K_∞ contains the center of $G(\mathbb{R})$, and $K_\infty \cap G_{\mathrm{der}}(\mathbb{R})^0$ is equal to the centralizer of $h_0(i)$ in $G_{\mathrm{der}}(\mathbb{R})^0$, hence is a maximal compact subgroup of $G_{\mathrm{der}}(\mathbb{R})^0$ by condition (c).

[12] Note that $G_{\mathrm{der}}(\mathbb{R})$ could still have compact normal algebraic subgroups, as long as they are not defined over $\mathbb{Q}$.

We have $X \simeq G(\mathbb{R})/K_\infty$, and Theorem 1.3 implies that there is a $G(\mathbb{R})$-invariant complex structure on X such that the connected components of X are Hermitian symmetric domains.

Example 1.33. Take $G = \mathrm{GSp}_{2d}$. Up to conjugation, there exists a unique morphism $h : \mathbb{S} \to \mathrm{GSp}_{2d}$ satisfying conditions (a)–(c) of Definition 1.32 and such that $h \circ w = xI_{2d}$ for every $x \in \mathbb{R}^\times$. An element of that class is given by

$$h(a+ib) = \begin{pmatrix} aI_d & -bI_d \\ bI_d & aI_d \end{pmatrix}.$$

For this h, we have $K_\infty = \mathrm{GSp}_{2d}(\mathbb{R}) \cap \mathrm{GO}(2d)$, and we can check that the map $\mathrm{GU}(d) \to K_\infty$ sending $X + iY \in \mathrm{GU}(d)$ (with $X, Y \in M_d(\mathbb{R})$) to $\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}$ is an isomorphism of Lie groups. So $K_\infty = \mathbb{R}_{>0}K'_\infty$, where $K_\infty = \mathrm{Sp}_{2d}(\mathbb{R}) \cap \mathrm{O}(d)$ is the maximal compact subgroup of $\mathrm{Sp}_{2d}(\mathbb{R})$ that was called K_∞ in Subsections 1.1 and 1.2. This implies that $X \simeq \mathfrak{h}_d$.

The couple (GSp_{2d}, h) is called a **Siegel Shimura datum**.

Let K be an open compact subgroup of $G(\mathbb{A}_f)$. We set

$$M_K(G,h)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K,$$

where the group K acts by right translations on the factor $G(\mathbb{A}_f)$, and the group $G(\mathbb{Q})$ acts by left translations on both factors simultaneously. This is the **Shimura variety at level** K associated to the Shimura datum K.

As in Subsection 1.1, if $(x_i)_{i\in I}$ is a system of representatives of the (finite) quotient $G(\mathbb{Q})\backslash G(\mathbb{A}_f)/K$, and if $\Gamma_i = G(\mathbb{Q}) \cap x_i K x_i^{-1}$ for every $i \in I$, then the Γ_i are arithmetic subgroups of $G(\mathbb{Q})$, and we have

$$M_K(G,h)(\mathbb{C}) = \coprod_{i\in I} \Gamma_i\backslash X.$$

Hence it follows from Theorem 1.5 that $M_K(G,h)(\mathbb{C})$ is (the set of complex points of) a quasi-projective algebraic variety over $\mathbb{C}$, smooth if K is small enough.

Again as in Subsection 1.1, we have Hecke correspondences between the $M_K(G,h)(\mathbb{C})$, which are finite maps, hence morphisms of algebraic varieties. This defines an action of $G(\mathbb{A}_f)$ on the projective system $(M_K(G,h)(\mathbb{C}))_{K\subset G(\mathbb{A}_f)}$, or on its limit

$$M(G,h)(\mathbb{C}) = \varprojlim M_K(G,h)(\mathbb{C}) = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f),$$

and we have $M_K(G,h)(\mathbb{C}) = M(G,h)(\mathbb{C})/K$ for every open compact subgroup K of $G(\mathbb{A}_f)$.

1.4.3. *Morphisms of Shimura varieties.* Let (G_1, h_1) and (G_2, h_2) be Shimura data, and let $u : G_1 \to G_2$ be a morphism of algebraic groups such that $u \circ h_1$ and h_2 are conjugated under $G_2(\mathbb{R})$; we say that u is a **morphism of Shimura data**. Then u induces a morphism of complex manifolds $X_1 \to X_2$, so, for all $K_1 \subset G_1(\mathbb{A}_f)$, $K_2 \subset G_2(\mathbb{A}_f)$ open compact subgroups such that $u(K_1) \subset K_2$, we get a morphism of quasi-projective varieties $u(K_1, K_2) : M_{K_1}(G_1, h_1)(\mathbb{C}) \to M_{K_2}(G_2, h_2)(\mathbb{C})$. We can also think of this as a morphism of $\mathbb{C}$-schemes $u : M(G_1, h_1)(\mathbb{C}) \to M(G_2, h_2)(\mathbb{C})$.

Proposition 1.34 (See [Del71, Proposition 1.15])**.** *If G_1 is an algebraic subgroup of G_2 and u is the inclusion, then, for every open compact subgroup K_1 of $G_1(\mathbb{A}_f)$, there exists an open compact subgroup $K_2 \supset K_1$ of $G_2(\mathbb{A}_f)$ such that $u(K_1, K_2) : M_{K_1}(G_1, h_1)(\mathbb{C}) \to M_{K_2}(G_2, h_2)(\mathbb{C})$ is a closed immersion.*

1.4.4. *Connected components.* For (G,h) equal to the Shimura datum of Example 1.33 and $K = \mathrm{Ker}(\mathrm{GSp}_{2d}(\widehat{\mathbb{Z}}) \to \mathrm{GSp}(\mathbb{Z}/n\mathbb{Z}))$, we have seen that the multiplier $c : \mathrm{GSp}_{2d} \to \mathrm{GL}_1$ induces a bijection

$$\pi_0(M_K(G,h)(\mathbb{C})) \xrightarrow{\sim} \mathbb{Q}_{>0}\backslash \mathbb{A}_f^\times / c(K) = \mathbb{Q}^\times \backslash \mathbb{A}^\times / c(K_\infty K).$$

In fact, it follows from real approximation and the Hasse principle that this works for many Shimura varieties:

Theorem 1.35 (Deligne, see [Del71, 2.7])**.** *Let $\nu : G \to T := G/G_{\mathrm{der}}$ be the quotient morphism, and suppose that G_{der} is simply connected. Then, for every open compact subgroup K of $G(\mathbb{A}_f)$, the map ν induces an isomorphism of groups*

$$\pi_0(M_K(G,h)(\mathbb{C})) \xrightarrow{\sim} T(\mathbb{Q})\backslash T(\mathbb{A})/\nu(K_\infty K).$$

In other words, ν induces an isomorphism

$$\pi_0(M(G,h)(\mathbb{C})) \xrightarrow{\sim} \pi_0(T(\mathbb{A})/T(\mathbb{Q}))/\pi_0(K_\infty).$$

1.5. **Canonical models.** In the situation of Example 1.33, we have seen that the algebraic varieties $M_K(G,h)(\mathbb{C})$ and all the Hecke correspondences are defined over $\mathbb{Q}$. We would like to generalize this kind of result to other Shimura varieties.

1.5.1. *Model of a Shimura variety.* First we need to say what we mean by a model.

Definition 1.36. Let (G,h) be a Shimura datum, and F be a subfield of $\mathbb{C}$. A **model** of the projective system $(M_K(G,h)(\mathbb{C}))_K$ over F is the data:

- for every open compact subgroup K of $G(\mathbb{A}_f)$, of a quasi-projective variety M_K over F and isomorphism $\iota_K : M_K \otimes_F \mathbb{C} \xrightarrow{\sim} M_K(G,h)(\mathbb{C})$;
- for every $g \in G(\mathbb{A}_f)$ and all open compact subgroups K, K' of $G(\mathbb{A}_f)$ such that $gK'g^{-1} \subset K$, of a morphism of F-varieties $T_{g,K,K'} : M_{K'} \to M_K$,

such that:

(i) For all g, K, K' as above, the morphism $\iota_K \circ T_{g,K,K',\mathbb{C}} \circ \iota_{K'}^{-1} : M_{K'}(G,h)(\mathbb{C}) \to M_K(G,h)(\mathbb{C})$ sends the class of (x,h) in $M_{K'}(G,h)(\mathbb{C})$ to the class of (x,hg) in $M_K(G,h)(\mathbb{C})$;
(ii) If K is an open compact subgroup of $G(\mathbb{A}_f)$ and $g \in K$, then $T_{g,K,K} = \mathrm{id}_{M_K}$;
(iii) If K, K', K'' are open compact subgroups of $G(\mathbb{A}_f)$ and $g, h \in G(\mathbb{A}_f)$ are such that $gK'g^{-1} \subset k$ and $hK''h^{-1} \subset K'$, then $T_{g,K,K'} \circ T_{h,K',K''} = T_{gh,K,K''}$;
(iv) If K, K' are open compact subgroups of $G(\mathbb{A}_f)$ such that K' is a normal subgroup of K, then the morphism $T_{g,K',K'}$ for $g \in K$ define an action of K/K' on $M_{K'}$ (this follows from (ii) and (iii)), and $T_{1,K,K'} : M_{K'} \to M_K$ induces an isomorphism $M_{K'}/(K/K') \to M_K$.

If we have a model $(M_K)_K$ of $(M_K(G,h)(\mathbb{C}))_K$ over F, we write $M = \varprojlim_K M_K$ (where the transition morphisms are given by the $T_{1,K',K}$). This is an F-scheme with an action of $G(\mathbb{A}_f)$, and we have a $G(\mathbb{A}_f)$-equivariant isomorphism $M \otimes_F \mathbb{C} \xrightarrow{\sim} M(G,h)(\mathbb{C})$.

In particular we get an action of $\mathrm{Gal}(\overline{F}/F)$ on $\pi_0(M(G,h)(\mathbb{C})) \xrightarrow{\sim} \pi_0(M \otimes_F \overline{F})$, which must commute with the action of $G(\mathbb{A}_f)$. Under the hypothesis of Theorem 1.35, we have

$$\pi_0(M \otimes_F \overline{F}) \xrightarrow{\sim} \pi_0(T(\mathbb{A})/T(\mathbb{Q}))/\pi_0(K_\infty)$$

with $T = G/G^{\mathrm{der}}$, and $G(\mathbb{A}_f)$ acts transitively on this set of connected components [Del71, Proposition 2.2]. So every element of $\mathrm{Gal}(\overline{F}/F)$ acts by translation by an element of $\pi_0(T(\mathbb{A})/T(\mathbb{Q}))/\pi_0(K_\infty)$, and the action of $\mathrm{Gal}(\overline{F}/F)$ comes from a morphism of groups $\mathrm{Gal}(\overline{F}/F) \to \pi_0(T(\mathbb{A})/T(\mathbb{Q}))/\pi_0(K_\infty)$, that necessarily factors through the maximal abelian quotient $\mathrm{Gal}(\overline{F}/F)^{\mathrm{ab}}$.

Suppose that F is a number field. Then global class field theory[13] gives an isomorphism

$$\mathrm{Gal}(\overline{F}/F)^{\mathrm{ab}} \xrightarrow{\sim} \pi_0(F^\times \backslash \mathbb{A}_F^\times)$$

where $\mathbb{A}_F$ is the ring of adeles of F, so the action of $\mathrm{Gal}(\overline{F}/F)$ on $\pi_0(M \otimes_F \overline{F})$ comes from a morphism of groups

$$\lambda_M : \pi_0(F^\times \backslash \mathbb{A}_F^\times) \to \pi_0(T(\mathbb{Q}) \backslash T(\mathbb{A}))/\pi_0(K_\infty),$$

called the **reciprocity law of the model**.

1.5.2. *The case of tori.* We consider the case where $G = T$ is a torus. Let $h : \mathbb{S} \to T_{\mathbb{R}}$ be any morphism of real algebraic groups. Then h trivially satisfies the conditions of Definition 1.32, so we get a Shimura datum (T,h), and $X = T(\mathbb{R})/\mathrm{Cent}_{T(\mathbb{R})}(h)$ is a singleton. For every open compact subgroup K of $T(\mathbb{A}_f)$,

$$M_K(T,h)(\mathbb{C}) = T(\mathbb{Q}) \backslash T(\mathbb{A}_f)/K$$

is a finite set, and we have

$$M(T,h)(\mathbb{C}) = T(\mathbb{Q}) \backslash T(\mathbb{A}_f),$$

which is a profinite set. Giving a model of the Shimura variety of (T,h) over a subfield F of $\mathbb{C}$ is the same as giving an action of $\mathrm{Gal}(\overline{F}/F)$ over $T(\mathbb{Q}) \backslash T(\mathbb{A}_f)$ (commuting with the action of $T(\mathbb{A}_f)$

[13] Normalized so that local uniformizers correspond to geometric Frobenius elements.

by translation), i.e. a morphism of groups $\mathrm{Gal}(\overline{F}/F) \to T(\mathbb{Q})\backslash T(\mathbb{A}_f)$. If F is a number field, this is equivalent to giving a morphism of groups

$$\lambda : \pi_0(F^\times\backslash\mathbb{A}_F^\times) \to \pi_0(T(\mathbb{Q})\backslash T(\mathbb{A})).$$

It is natural to construct such a morphism from a morphism of algebraic groups $F^\times \to T$, where $F^\times$ is seen as an algebraic group over $\mathbb{Q}$ (so that, for example, we have $F^\times(\mathbb{A}) = (\mathbb{A} \otimes_\mathbb{Q} F)^\times = \mathbb{A}_F^\times$). We already have a morphism $h : \mathbb{S} \to T_\mathbb{R}$, which gives a morphism of complex algebraic groups $h_\mathbb{C} : \mathbb{S}_\mathbb{C} \to T_\mathbb{C}$. Remember that $\mathbb{S}_\mathbb{C} \simeq \mathrm{GL}_{1,\mathbb{C}} \times \mathrm{GL}_{1,\mathbb{C}}$, and that we denoted by $r : \mathrm{GL}_{1,\mathbb{C}} \to \mathbb{S}_\mathbb{C}$ the embedding of the first factor (see Subsubsection 1.4.1). We get a morphism $h_\mathbb{C} \circ r : \mathrm{GL}_{1,\mathbb{C}} \to T_\mathbb{C}$. As T is an algebraic group over $\mathbb{Q}$, this morphism is defined over a finite extension of $\mathbb{Q}$ in $\mathbb{C}$, and we call this extension F. We get a morphism of F-algebraic groups $\mathrm{GL}_{1,F} \to T_F$, hence a morphism of $\mathbb{Q}$-algebraic groups $F^\times \to \mathrm{Res}_{F/\mathbb{Q}} T_F$, where $\mathrm{Res}_{F/\mathbb{Q}} T_F$ is the algebraic group that sends a $\mathbb{Q}$-algebra R to $T(R \otimes_\mathbb{Q} F)$. Composing this with the norm $N_{F/\mathbb{Q}} : \mathrm{Res}_{F/\mathbb{Q}} T_F \to T$, we finally get a morphism $r(h) : F^\times \to T$, called the **reciprocity morphism for** (T, h). We take λ_M to be induced by $r(h)$.

So if (G, h) is a Shimura datum with G a torus, we get a canonically defined model of the associated Shimura variety over the field of definition of $h_\mathbb{C} \circ r$.

1.5.3. *The reflex field.* Let (G, h) be a Shimura datum satisfying the hypothesis of Theorem 1.35, and let $\nu : G \to T := G/G_{\mathrm{der}}$ be the quotient morphism. We have an isomorphism, induced by ν:

$$\pi_0(M(G,h))(\mathbb{C}) \xrightarrow{\sim} \pi_0(T(\mathbb{A})/T(\mathbb{Q}))/\pi_0(K_\infty).$$

Suppose that we have a model $(M_K)_K$ of the Shimura variety of (G, h) over a number field $F \subset \mathbb{C}$. We would expect the Shimura variety of $(T, \nu \circ h)$ to also have a model over F, and the isomorphism above to be $\mathrm{Gal}(\overline{F}/F)$-equivariant, where the action on the right hand side is given by the morphism $r(\nu \circ h) : F^\times \to T$ constructed in Subsubsection 1.5.2.

Remember that $r : \mathrm{GL}_{1,\mathbb{C}} \to \mathbb{S}_\mathbb{C} \simeq \mathrm{GL}_{1,\mathbb{C}} \times \mathrm{GL}_{1,\mathbb{C}}$ is the embedding of the first factor (see 1.4.1 again). By the previous paragraph, we would expect F to contain the field of definition of $\nu \circ h \circ r$. In fact it would make sense to take F to be the field of definition of $h_\mathbb{C} \circ r : \mathrm{GL}_{1,\mathbb{C}} \to G_\mathbb{C}$, except that h is only significant up to conjugation. This motivates the following definition.

Definition 1.37. Let (G, h) be a Shimura datum. The **reflex field** $F(G, h)$ of (G, h) is the field of definition of the conjugacy class of $h_\mathbb{C} \circ r : \mathrm{GL}_{1,\mathbb{C}} \to G_\mathbb{C}$.

Let $F = F(G, h)$. Then F is a finite extension of $\mathbb{Q}$ in $\mathbb{C}$, and, for every morphism ρ of G into a commutative algebraic group, the morphism $\rho_\mathbb{C} \circ h_\mathbb{C} \circ r$ is defined over F. Note that $h_\mathbb{C} \circ r$ itself is not necessarily defined over F.

Example 1.38. Let $E = \mathbb{Q}[\sqrt{-d}]$ be an imaginary quadratic extension of $\mathbb{Q}$, let $p \geqslant q \geqslant 1$ be integers, and set $n = p + q$. Let

$$J = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \in \mathrm{GL}_n(\mathbb{Z}).$$

For every commutative ring R, we denote by $x \mapsto \overline{x}$ the involution of $R \otimes_\mathbb{Z} \mathcal{O}_E$ induced by the nontrivial element of $\mathrm{Gal}(E/\mathbb{Q})$, and, for every $Y \in M_n(R \otimes_\mathbb{Q} E)$, we write $Y^* = {}^t\overline{Y}$.

The general unitary group $\mathrm{GU}(p, q)$ is the $\mathbb{Z}$-group scheme defined by

$$\mathrm{GU}(p,q)(R) = \{g \in \mathrm{GL}_n(R \otimes_\mathbb{Z} \mathcal{O}_E) \mid \exists c(g) \in R^\times,\ g^* J g = c(g) J\}$$

for every commutative ring R. Then $\mathrm{GU}(p, q)_\mathbb{Q}$ is a connected reductive algebraic group, and we have a morphism of group schemes $c : \mathrm{GU}(p, q) \to \mathrm{GL}_1$, whose kernel is the unitary group $\mathrm{U}(p, q)$.

Let $h : \mathbb{S} \to \mathrm{GU}(p, q)_\mathbb{R}$ be the morphism defined by

$$h(z) = \begin{pmatrix} zI_p & 0 \\ 0 & \overline{z}I_q \end{pmatrix} \in \mathrm{GU}(p,q)(\mathbb{R}).$$

Then (G, h) is a Shimura datum, and K_∞ is the set of matrices $\begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix}$ such that $g_1 \in \mathrm{GL}_p(\mathbb{C})$, $g_2 \in \mathrm{GL}_q(\mathbb{C})$ and there exists $c \in \mathbb{R}$ with $g_1^* g_1 = cI_p$ and $g_2^* g_2 = cI_q$. We use the isomorphism $\mathbb{C} \otimes_\mathbb{Q} E \xrightarrow{\sim} \mathbb{C} \times \mathbb{C}$ sending $x \otimes 1 + y \otimes \sqrt{-d}$ to $(x + \sqrt{-d}y, x - \sqrt{-d}y)$ to identify $\mathrm{GU}(p, q)(\mathbb{C})$ to a subgroup of $\mathrm{GL}_n(\mathbb{C})\times\mathrm{GL}_n(\mathbb{C})$; note that the involution $g\mapsto\overline{g}$ of $\mathrm{GU}(p,q)(\mathbb{C})$ corresponds to switching of the two factors. With this convention, we have for every $z\in\mathbb{C}^\times$ that

$$h_{\mathbb{C}}\circ r(z)=\left(\begin{pmatrix} zI_p & 0\\ 0 & I_q\end{pmatrix},\begin{pmatrix} I_p & 0\\ 0 & zI_q\end{pmatrix}\right).$$

It is easy to check that $h_{\mathbb{C}}\circ r$ is defined over E but not over $\mathbb{Q}$. On the other hand, the reflex field of (G,h) is E if $p>q$ and $\mathbb{Q}$ if $p=q$.

Example 1.39. Let (G,h) be the Shimura datum of Example 1.33 (so that $G=\mathrm{GSp}_{2d}$). For every $z\in\mathbb{C}^\times$, we have

$$h_{\mathbb{C}}\circ r(z)=\begin{pmatrix}\frac{1}{2}(z+1)I_d & -\frac{1}{2i}(z-1)I_d\\ \frac{1}{2i}(z-1)I_d & \frac{1}{2}(z+1)I_d\end{pmatrix}=P\begin{pmatrix} zI_d & 0\\ 0 & I_d\end{pmatrix}P^{-1},$$

where

$$P=\begin{pmatrix}\frac{1}{\sqrt{2}}I_d & \frac{i}{\sqrt{2}}I_d\\ \frac{i}{\sqrt{2}}I_d & \frac{1}{\sqrt{2}}I_d\end{pmatrix}.$$

So the reflex field of (G,h) is $\mathbb{Q}$.

1.5.4. *Canonical models.* We are now ready to define canonical models.

Definition 1.40. Let (G,h) be a Shimura datum and let $F=F(G,h)$. A **canonical model** of $M(G,h)(\mathbb{C})$ is a model $(M_K)_K$ over F such that, for every torus $u:H\subset G$ and every $h':\mathbb{S}\to H_{\mathbb{R}}$ such that $u\circ h'$ and h are $G(\mathbb{R})$-conjugated (i.e. such that u induces a morphism of Shimura data from (H,h') to (G,h)), the morphism

$$u:M(H,h')(\mathbb{C})\to M(G,h)(\mathbb{C})$$

is defined over the compositum $F\cdot F(H,h')\subset\mathbb{C}$, where we use as model of $M(H,h')(\mathbb{C})$ over $F(H,h')$ the one defined in Subsubsection 1.5.2.

Example 1.41.
(1) If G is a torus, then the model of Subsubsection 1.5.2 is a canonical model of $M(G,h)(\mathbb{C})$.
(2) If (G,h) is the Shimura datum of Example 1.33 (so that $G=\mathrm{GSp}_{2d}$), then the schemes $(\mathcal{M}_{d,K,\mathbb{Q}})_{K\subset\mathrm{GSp}_{2d}(\mathbb{A}_f)}$ of Subsubsection 1.3.3 form a canonical model of $M(G,h)(\mathbb{C})$. This is not obvious but follows from the main theorem of complex multiplication; see [Del71, §4].

At the time of Deligne's paper [Del71], it was not known whether all Shimura varieties has canonical models (spoiler: this is now known to be true, see Theorem 2.27), but it was possible to prove their uniqueness. If $u:(H,h')\to(G,h)$ is a morphism of Shimura varieties as in Definition 1.40 (so that H is a subtorus of G), the image in $M(G,h)(\mathbb{C})$ of the points of $M(H,h')(\mathbb{C})$ are called **special points**. The fact that canonical models are uniquely characterized relies on the following two points:

(i) Special points are dense in $M(G,h)(\mathbb{C})$;
(ii) For every finite extension $F'\subset\mathbb{C}$ of $F(G,h)$, there exists $u:(H,h')\to(G,h)$ as above such that $F(H,h')$ and F' are linearly disjoint over $F(G,h)$ [Del71, Théorème 5.1].

From this, we can deduce:

Theorem 1.42 (See [Del71], Corollaire 5.4)**.** *Let $u:(G_1,h_1)\to(G_2,h_2)$ be a morphism of Shimura data. Then the corresponding morphism $M(G_1,h_1)(\mathbb{C})\to M(G_2,h_2)(\mathbb{C})$ of Shimura is defined over any common extension F of $F(G_1,h_1)$ and $F(G_2,h_2)$ in $\mathbb{C}$.*

Corollary 1.43. *Let (G,h) be a Shimura datum. Then a canonical model of $M(G,h)(\mathbb{C})$ is unique up to unique isomorphism if it exists.*

Corollary 1.44. *Let (G,h) be a Shimura datum satisfying the hypothesis of Theorem 1.35, let $\nu:G\to T:=G/G_{\mathrm{der}}$ be the quotient morphism, and let $F=F(G,h)$. Then the action of $\mathrm{Gal}(\overline{F}/F)$ on $\pi_0(M(G,h)(\mathbb{C}))\xrightarrow{\sim}\pi_0(T(\mathbb{Q})\backslash T(\mathbb{A}))/\pi_0(K_\infty)$ is given by the inverse of the reciprocity morphism for $(T,\nu\circ h)$.*

Apply Theorem 1.42 to $\nu:(G,h)\to(T,\nu\circ h)$. Using the same techniques as for Theorem 1.42, we also get the following very useful result.

Proposition 1.45 (See [Del71], Corollarie 5.7)**.** *Let $u : (G_1, h_1) \to (G_2, h_2)$ be a morphism of Shimura data such that the underlying morphism of algebraic groups is a closed immersion and that $F(G_1, h_1) \subset F(G_2, h_2)$. If $M(G_2, h_2)(\mathbb{C})$ has a canonical model, then so does $M(G_1, h_1)(\mathbb{C})$.*

2. Arithmetic Shimura varieties

In Section 1, we have defined Shimura varieties over $\mathbb{C}$ and introduced the notion of canonical model of a Shimura variety. We also discussed in some detail the example of the Siegel modular varieties, who have canonical models coming from their modular interpretation. In this lecture, we first want to present different types of Shimura varieties, which *each type being contained in the next one*:

(1) **The Siegel modular variety:** It is a moduli space of principally polarized abelian schemes with some level structure;
(2) **PEL type Shimura varieties:** They are have an interpretation as moduli spaces of polarized abelian schemes with multiplication by the ring of integers $\mathcal{O}$ of some number field and some level structure (here "P" means "polarization", "E" means "endomorphisms" in reference to the action of $\mathcal{O}$ and "L" means "level structure");
(3) **Hodge type Shimura varieties:** They come from Shimura data (G, h) that have an injective morphism into a Siegel Shimura datum (Example 1.33);
(4) **Abelian type Shimura varieties:** Their Shimura datum is "isogenous" to a Hodge type Shimura datum (in a way to be made precise later);
(5) **General Shimura varieties:** All Shimura varieties.

The further we go down in the list, the less is known about the geometry of the Shimura variety (and the higher the price for what we know), because many of the techniques we have rely on the interpretation of the Shimura varieties as moduli problems of abelian schemes, and this is only really available for PEL type Shimura varieties (there is something for Hodge type Shimura varieties, but it is harder to use).

For example, PEL type Shimura variety naturally come with an integral model defined over a localization of the ring of integers of their reflex field, but it took a lot of effort to construct integral models for Hodge type and abelian type Shimura varieties, and to formulate their properties; as far as we know, nothing is known for general Shimura varieties.

We are not claiming that the classification above is the only measure of the complexity of a Shimura variety, whatever that means. For example, as we will discuss in Section 3, if one wants to study the cohomology of Shimura varieties and their zeta functions, then the simplest case is not the Siegel modular varieties, but rather compact PEL type Shimura varieties whose group has no endoscopy and a simply connected derived subgroup; we will introduce some examples of these, known as **Kottwitz's simple Shimura varieties**.

2.1. PEL type Shimura varieties.

These Shimura varieties were introduced by Kottwitz in [Kot92b], but we will follow the presentation of Lan for the moduli problems (cf. [Lan13, 1.4.1]), which is closer to our definition of the Siegel moduli problem. The equivalence between the two definitions is proved in [Lan13, 1.4.3].

2.1.1. *PEL data.*

Definition 2.1 (see [Lan20, 5.1] or [Lan13, Definition 1.2.13])**.** An **(integral) PEL datum** is a quintuple $(\mathcal{O}, *, \Lambda, \langle\cdot,\cdot\rangle, h)$, where

(1) $\mathcal{O}$ is an order in a finite-dimensional semisimple $\mathbb{Q}$-algebra B (that is, $\mathcal{O}$ is a subring of B that is a free $\mathbb{Z}$-module and spans the $\mathbb{Q}$-vector space $\mathbb{Q}$);
(2) $*$ is a **positive involution** of $\mathcal{O}$, i.e. an anti-automorphism of rings of order 2 such that, for every $x \in \mathcal{O}\backslash\{0\}$, we have $\mathrm{Tr}_{(B\otimes_{\mathbb{Q}}\mathbb{R})/\mathbb{R}}(xx^*) > 0$;
(3) Λ is an $\mathcal{O}$-module that is finitely generated and free as a $\mathbb{Z}$-module;
(4) $\langle\cdot,\cdot\rangle : \Lambda \times \Lambda \to \mathbb{Z}$ is an alternating bilinear map such that, for all $x, y \in \Lambda$ and $b \in \mathcal{O}$, we have
$$\langle bx, y\rangle = \langle x, b^* y\rangle;$$
(5) $h : \mathbb{C} \to \mathrm{End}_{B\otimes_{\mathbb{Q}}\mathbb{R}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R})$ is an $\mathbb{R}$-algebra morphism such that:

(a) For $z \in \mathbb{C}$ and $x, y \in \Lambda \otimes_{\mathbb{Z}} \mathbb{R}$, we have
$$\langle h(z)(x), y\rangle = \langle x, h(\overline{z})(y)\rangle;$$
(b) The $\mathbb{R}$-bilinear pairing $\langle \cdot, h(i)(\cdot)\rangle$, which is symmetric by (a), is also positive definite.

Let $(\mathcal{O}, *, \Lambda, \langle\cdot,\cdot\rangle, h)$ be a PEL datum. We define a group scheme G over $\mathbb{Z}$ by
$$G(R) = \{g \in \mathrm{End}_{\mathcal{O}\otimes_{\mathbb{Z}} R}(\Lambda \otimes_{\mathbb{Z}} R) \mid \exists c(g) \in R^\times,\ \langle g(\cdot), g(\cdot)\rangle = c(g)\langle \cdot,\cdot\rangle\}$$
for every commutative ring R. We also get a morphism of group schemes $c : G \to \mathrm{GL}_1$. The morphism of $\mathbb{R}$-algebras $h : \mathbb{C} \to \mathrm{End}_{B\otimes_{\mathbb{Q}}\mathbb{R}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R})$ induces a morphism of $\mathbb{R}$-algebraic groups $h : \mathbb{S} \to G_{\mathbb{R}}$.

Proposition 2.2. *The couple $(G^0_{\mathbb{Q}}, h)$ satisfies conditions (a)–(c) in Definition 1.32 of a Shimura datum.*

Let us explain why this proposition is true. Condition (a) of Definition 1.32 follows from the fact that $h : \mathbb{C}^\times \to \mathrm{End}_{B\otimes_{\mathbb{Q}}\mathbb{R}}(\Lambda\otimes_{\mathbb{Z}}\mathbb{R})$ is a morphism of $\mathbb{R}$-algebra, and Condition (c) follows from Condition (5)(b) of Definition 2.1. We have to show how to prove Condition (b), on the decomposition of $\mathrm{Lie}(G_{\mathbb{C}}$ into eigenspaces for the action of $\mathrm{Ad} \circ h_{\mathbb{C}}$. Let $V = \Lambda \otimes_{\mathbb{Z}} \mathbb{R}$, a finite-dimensional $\mathbb{R}$-vector space. Then $h : \mathbb{C} \to \mathrm{End}_{\mathbb{R}}(V)$ defines a structure of $\mathbb{C}$-vector space on V, so $A := V/\Lambda$ is a complex torus. We would like this torus to be an abelian variety, so we need a polarization on it, that is, a positive definite Hermitian form H on V such that $\mathrm{Im}(H)$ takes integer values on Λ. But we already have an alternating form on Λ, so we are already know from Subsubsection 1.3.1 how to proceed: define H by
$$H(v, w) = \langle h(i)(v), w\rangle + i\langle v, w\rangle.$$
The fact that $\mathrm{Im}(H)$ takes integral values on Λ is clear, the pairing H is Hermitian by condition (5)(a) and positive definite by condition (5)(b). So the torus A is an abelian variety, with dual abelian variety $A^\vee = V/\Lambda^\vee$, where $\Lambda^\vee = \{v \in V \mid \forall w \in \Lambda, \langle v, w\rangle \in \mathbb{Z}\}$ is the dual lattice of Λ. The polarization $\lambda : A \to A^\vee$ defined by H is then just the map induced by $\Lambda \subset \Lambda^\vee$. We can see h as a morphism of algebraic groups $\mathbb{S} \to \mathrm{GL}(V)$, and the decomposition $V \otimes_{\mathbb{R}} \mathbb{C} = \bigoplus_{p,q\in\mathbb{Z}} V^{p,q}$ induced by this morphism (see Subsubsection 1.4.1) is the same as the Hodge structure coming from the isomorphism $V \simeq \mathrm{H}_1(A, \mathbb{R})$. As A is an abelian variety, we have $V^{p,q} = 0$ unless $(p, q) \in \{(0, -1), (-1, 0)\}$. As $\mathfrak{g} := \mathrm{Lie}(G_{\mathbb{R}}) \subset \mathrm{End}_{\mathbb{R}}(V)$, for the decomposition $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} = \bigoplus_{p,q\in\mathbb{Z}} \mathfrak{g}^{p,q}$ induced by $\mathrm{Ad} \circ h$, we have $\mathfrak{g}^{p,q} = 0$ unless $(p, q) \in \{(0, 0), (1, -1), (-1, 1)\}$, as desired.

Note also that the action of $\mathcal{O}$ on Λ defines a morphism of rings $\iota : \mathcal{O} \to \mathrm{End}_{\mathbb{C}}(A)$ satisfying the **Rosati condition**, which says that, for every $b \in \mathcal{O}$, we have
$$\lambda \circ i(b^*) = i(b)^\vee \circ \lambda.$$

Remark 2.3 (c.f. [Kot92b, §5, §7]). Let $(\mathcal{O}, *, \Lambda, \langle\cdot,\cdot\rangle, h)$ be a PEL datum, and let G be the associated group scheme. If the $\mathbb{Q}$-algebra $B = \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q}$ is simple, then all the simple factors of $G(\mathbb{C})^{\mathrm{der}}$ are of the same type, which is A, C or D. If we are in type A or C, then $G_{\mathbb{Q}}$ is connected and reductive. In type D, the group $G_{\mathbb{Q}}$ is reductive and has $2^{[F_0:\mathbb{Q}]}$ connected components, where F_0 is the field of the fixed points of $*$ in the center of $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q}$, so it is never connected.

We now give an example of each type.

Example 2.4. Take $\mathcal{O} = \mathbb{Z}$, $*$ the trivial involution, $\Lambda = \mathbb{Z}^{2d}$, $\langle\cdot,\cdot\rangle$ the perfect symplectic pairing with matrix $\begin{pmatrix} 0 & I_d \\ I_d & 0 \end{pmatrix}$ in the canonical basis of $\mathbb{Z}^{2d}$, and $h : \mathbb{C} \to M_{2d}(\mathbb{R})$ defined by
$$h(a + ib) = \begin{pmatrix} aI_d & -bI_d \\ bI_d & aI_d \end{pmatrix}.$$
Then we get a PEL datum, and the couple (G, h) is the Siegel Shimura datum of Example 1.33.

Example 2.5. Let $B = E \subset \mathbb{C}$ be an imaginary quadratic extension of $\mathbb{Q}$, $\mathcal{O}$ be an order in E (for example the ring of integers), $*$ be the restriction to $\mathcal{O}$ of complex conjugation,[14] $\Lambda = \mathcal{O}^{p+q}$

[14] If $(1, a)$ is a $\mathbb{Z}$-basis of $\mathcal{O}_E$, then every order $\mathcal{O}$ of E is contained in $\mathcal{O}_E$ and of the form $\mathbb{Z} \oplus fa\mathbb{Z}$, where $f = [\mathcal{O}_E : \mathcal{O}]$ is the **conductor** of $\mathcal{O}$. In particular, the order $\mathcal{O}$ is stable by the nontrivial element of $\mathrm{Gal}(E/\mathbb{Q})$.

with $p \geqslant q \geqslant 0$. Choose $\varepsilon \in \mathcal{O}$ such that $-i\varepsilon \in \mathbb{R}_{>0}$, let H be the Hermitian pairing on Λ with matrix $\begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$ and $\langle \cdot, \cdot \rangle$ be the alternating pairing $\mathrm{Tr}_{\mathcal{O}/\mathbb{Z}}(\varepsilon H)$ on Λ. Finally, define $h : \mathbb{C} \to \mathrm{End}_{\mathbb{C}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R}) = M_{p+q}(\mathbb{C})$ by $h(z) = \begin{pmatrix} zI_p & 0 \\ 0 & \overline{z} I_q \end{pmatrix}$. Then we get a PEL datum, and the couple (G, h) is the Shimura datum of Example 1.38.

If $p = q$, then (G, h) does not satisfy condition (d) of Definition 1.32.

Example 2.6. Let B be a quaternion algebra over $\mathbb{Q}$ such that $B \otimes_{\mathbb{Q}} \mathbb{R} \simeq \mathbb{H}$, and let $\mathcal{O}$ be an order in B that is stable by the involution of $\mathbb{H}$ defined by $(x + iy + jz + kt)^* = x - iy - jz - kt$. Let $\Lambda = \mathcal{O}^{2n}$, let $\langle \cdot, \cdot \rangle$ be $\mathrm{Tr}_{\mathcal{O}/\mathbb{Z}} \circ H$, where H is the skew-Hermitian pairing on Λ with matrix $\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$. Define $h : \mathbb{C} \to \mathrm{End}_{B \otimes_{\mathbb{Q}} \mathbb{R}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R}) = M_n(\mathbb{H})$ by $h(a + ib) = \begin{pmatrix} aI_n & -bI_n \\ bI_n & aI_n \end{pmatrix}$.

The group $G_{\mathbb{R},\mathrm{der}}$ is often denoted by SO^*_{2n}; it is a quasi-split outer form of the split orthogonal group SO_{2n}, hence is of type D_n. But note that the algebraic group $G_{\mathbb{Q}}$ is not connected.

Definition 2.7. We say that a Shimura datum is of **PEL type** if it is of the form $(G^0_{\mathbb{Q}}, h)$, where (G, h) comes from a PEL datum. The corresponding Shimura varieties are called **PEL type Shimura varieties**.

2.1.2. *PEL moduli problems.* Just as in the case of the Siegel modular variety, PEL type Shimura varieties are the solution of a moduli problem,[15] know as a PEL moduli problem. We now discuss these.

Let $(\mathcal{O}, *, \Lambda, \langle \cdot, \cdot \rangle, h)$ be a PEL datum, and let (G, h) be defined as before. If $V = \Lambda \otimes_{\mathbb{Z}} \mathbb{R}$, then we saw the morphism $h : \mathbb{C}^\times \to \mathrm{End}_{\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{R}}(V)$ defines a decomposition $V \otimes_{\mathbb{R}} \mathbb{C} = V^{-1,0} \oplus V^{0,-1}$. The **reflex field** F of the PEL datum is the field of definition of the isomorphism class of $V^{-1,0}$ as an $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{C}$-module, that is, the subfield of $\mathbb{C}$ generated by the $\mathrm{Tr}(b, V^{1,0})$, for $b \in \mathcal{O}$. It is also equal to the reflex field of (G, h).

Definition 2.8. We say that a prime number p is **good** for the PEL datum if:

- p is unramified in $\mathcal{O}$ (i.e. it does not divide the discriminant of $\mathcal{O}/\mathbb{Z}$);
- p does not divide $[\Lambda^\vee : \Lambda]$, where $\Lambda^\vee$ is as before the lattice $\{v \in V \mid \forall w \in \Lambda, \langle v, w \rangle \in \mathbb{Z}\}$;
- $p \neq 2$ if the PEL datum has a factor of type D (i.e. if $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{R}$ has a simple factor isomorphic to an algebra $M_n(\mathbb{H})$ with its canonical positive involution).

If p is not good we say that it is **bad**. Note that, if p is good, then F and $G_{\mathbb{Q}}$ are unramified at p; in fact, the group $G(\mathbb{Z}_p)$ is then a hyperspecial maximal compact subgroup of $G(\mathbb{Q}_p)$.

Let T be a set of good primes (finite or infinite), and let $\mathcal{O}_{F,T}$ be the localization $\mathcal{O}_F[1/p, p \notin T]$.

Definition 2.9 (PEL moduli problem)**.** Let n be a positive integer that is prime to all the elements of T. Then the PEL moduli problem at level n defined by the fixed PEL datum is the functor $\mathcal{M}_n$ from the category of $\mathcal{O}_{F,T}$-schemes to the category of sets sending an $\mathcal{O}_{F,T}$-scheme S to the set of isomorphism classes of quadruples $(A, \lambda, \iota, (\eta, \varphi))$, where

- A is an abelian scheme over S;
- $\lambda : A \to A^\vee$ is a polarization whose degree is prime to nN;
- $\iota : \mathcal{O} \to \mathrm{End}_S(A)$ is a morphism of rings satisfying the Rosati condition: for every $b \in \mathcal{O}$, we have $\lambda \circ \iota(b^*) = \iota(b)^\vee \circ \lambda$;
- (η, φ) is a level n structure on A, i.e., $\eta : A[n] \xrightarrow{\sim} (\Lambda/n\Lambda)_S$ is an $\mathcal{O}$-equivariant isomorphism of group schemes and $\varphi : \underline{\mathbb{Z}/n\mathbb{Z}}_S \xrightarrow{\sim} \mu_{n,S}$ is an isomorphism of group schemes such that $\varphi \circ \langle \cdot, \cdot \rangle, \circ\eta$ is the Weil pairing defined by λ on $A[n]_S$, and moreover (η, φ) are liftable to level m structures for every prime-to-T multiple m of n, in the sense of [Lan13, Definition 1.3.6.2].

We furthermore require that this quadruple satisfy the following **determinant condition** (see [Lan13, Definition 1.3.4.1]): let $\alpha_1, \dots, \alpha_t$ be a basis of the $\mathbb{Z}$-module $\mathcal{O}$, and let $X_1, \dots, X_t$ be indeterminates. Then $\det(\iota(\alpha_1)X_1 + \cdots + \iota(\alpha_t)X_t, \mathrm{Lie}(A))$ is a polynomial in $\mathcal{O}_S[X_1, \dots, X_t]$, and the

[15] Well, almost. See below for a more precise statement.

condition says that this polynomial is equal to the image by the map $\mathcal{O}_{F,T} \to \mathcal{O}_S$ of the polynomial $\det(\alpha_1 X_1 + \cdots + \alpha_t X_t, V^{1,0}) \in \mathcal{O}_{F,T}[X_1, \ldots, X_t]$.[16]

Remark 2.10. (1) We did not specify the relative dimension of the abelian scheme A in the moduli problem, because it is already determined by the determinant condition.

(2) For $n \in \{1, 2\}$, the objects of the moduli problem $\mathcal{M}_n$ can have nontrivial automorphisms, so it would make more sense to see $\mathcal{M}_n$ as a functor with values in groupoids, i.e. as a stack. See Remark 1.12.

(3) As in Remark 1.16, it is also possible to define the moduli problem $\mathcal{M}_K$ for more general levels K, i.e. for open compact subgroups K of $G(\mathbb{A}_f)$, though there is a condition on K corresponding to the condition that n be prime to T. Let $\mathbb{A}_f^T$ be the ring of prime-to-T adeles of $\mathbb{Q}$, i.e. the restricted product of the $\mathbb{Q}_p$ for $p \notin T$, and $\mathbb{A}_T$ be the restricted product of the $\mathbb{Q}_p$ for $p \in T$. We have $\mathbb{A}_f = \mathbb{A}_T \times \mathbb{A}_f^T$, and the condition on K is that $K = K_T K^T$, where $K^T \subset G(\mathbb{A}_f^T)$ and $K_T = \prod_{p \in T} K_p$ with $K_p = G(\mathbb{Z}_p)$ for every $p \in T$.

Theorem 2.11 (Kottwitz and Lan, see Corollaries 1.4.1.12 and 7.2.3.10 of [Lan13])**.** *If $n \geqslant 3$, then the functor $\mathcal{M}_n$ is representable by a smooth quasi-projective scheme over $\mathcal{O}_{F,T}$.*

In fact, by Theorem 1.4.1.11 and Corollary 7.2.3.10 of [Lan13], the functor $\mathcal{M}_K$ is also representable by a smooth quasi-projective scheme for K small enough.

Remark 2.12. We also have an action of the Hecke operators defined by elements $g \in G(\mathbb{A}_f^T)$ (i.e. the Hecke operators that are trivial at primes of T) on the tower $(\mathcal{M}_K)$; as for the Siegel moduli problem, the element g acts on the level structure. See [Lan13, Remark 1.4.3.11] (and the comparison result of Proposition 1.4.3.4 ibid.).

2.1.3. *PEL moduli problems and canonical models.* Consider the PEL datum of Example 2.4. The associated couple (G, h) is the Siegel Shimura datum of Example 1.33, and we have seen in Example 1.41(2) that the associated moduli problem, with $T = \emptyset$, defines a canonical model of the Shimura variety of (G, h).

We would like something like this to be true for general PEL data. One obstacle is that the moduli problem only depends on the group $G(\mathbb{A})$; this is not obvious on Definition 2.9, but it becomes so if we use the moduli problem of [Lan13, Definition 1.4.2.1] or of [Kot92b, §5] (the equivalence of the two moduli problems is proved in [Lan13, 1.4.3]). So, if we have another reductive algebraic group G' over $\mathbb{Q}$ such that $G'_{\mathbb{Q}_v} \simeq G_{\mathbb{Q}_v}$ for every place v of $\mathbb{Q}$, we get a Shimura datum (G', h') by taking $h' = h : \mathbb{S} \to G_{\mathbb{R}} \simeq G'_{\mathbb{R}}$, and we should also see the canonical models for the Shimura variety of (G', h') in the PEL moduli problem. In fact, we have the following result.

Proposition 2.13. *Suppose that the semisimple $\mathbb{Q}$-algebra $\mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q}$ has no simple factor of type D (see Definition 2.8). Then:*

(i) *[Lan13, Remark 1.4.4.4]. Let $T \subset T'$ be two sets of good prime numbers, let $K^{T'}$ be an open compact subgroup of $G(\mathbb{A}_f^{T'})$, and let $K = K^{T'} \prod_{p \in T'} G(\mathbb{Z}_p)$. Let $\mathcal{M}_K$ (resp. $\mathcal{M}'_K$) be the moduli problem over $\operatorname{Spec}(\mathcal{O}_{F,T})$ (resp. $\operatorname{Spec}(\mathcal{O}_{F,T'})$) from Definition 2.9, where we take the set of good primes to be T (resp. T'). Then the forgetful functor $\mathcal{M}_K \to \mathcal{M}'_K \times_{\operatorname{Spec}(\mathcal{O}_{F,T'})} \operatorname{Spec}(\mathcal{O}_{F,T})$ is an isomorphism.*

(ii) *[Kot92b, Sections 7-8]. Let $\operatorname{Ker}^1(\mathbb{Q}, G)$ be the kernel of the diagonal map $\mathrm{H}^1(\mathbb{Q}, G) \to \prod_v \mathrm{H}^1(\mathbb{Q}_v, G)$, where we take the product over all places v of $\mathbb{Q}$; this is a finite set. Suppose that every $(G_{\mathbb{Q}}, h)$ is a Shimura datum, and use $T = \emptyset$ to define the moduli problems of Definition 2.9, so that they are moduli problems over $\operatorname{Spec}(F)$.[17] Then the projective system $(\mathcal{M}_K)_{K \subset G(\mathbb{A}_f)}$ is a disjoint union indexed by $i \in \operatorname{Ker}^1(\mathbb{Q}, G)$ of projective systems $(\mathcal{M}_K^{(i)})_{K \subset G(\mathbb{A}_f)}$, and each $(\mathcal{M}_K^{(i)})_{K \subset G_i(\mathbb{A}_f)}$ is the canonical model of the Shimura variety of $(G_{\mathbb{Q}}, h)$.*

[16] It is not totally obvious that the second polynomial, which is a priori in $F[X_1, \ldots, X_t]$, has its coefficients in $\mathcal{O}_{F,T}$. But it is also not too hard to check. See for example [Kot92b, pp. 389–390].

[17] By (i), these moduli problems are the generic fibers of the moduli problems defined by nonempty sets T.

Remark 2.14.
(1) For every $i \in \mathrm{Ker}^1(\mathbb{Q}, G)$, let G_i be the corresponding inner form of $G_\mathbb{Q}$; we have $G_{i,v} \simeq G_v$ for every place v of $\mathbb{Q}$, and we denote by h_i the morphism $\mathbb{S} \xrightarrow{h} G_\mathbb{R} \simeq G_{i,\mathbb{R}}$. If $(G_\mathbb{Q}, h)$ is a Shimura datum, then all (G_i, h_i) are, and the projective system of projective systems $(\mathcal{M}_K^{(i)})_{K \subset G(\mathbb{A}_f)}$ in (ii) of the proposition is really a canonical model of the Shimura variety of (G_i, h_i) (note that $G_i(\mathbb{A}_f) \simeq G(\mathbb{A}_f)$). But, as noted by Kottwitz at the end of [Kot92b, §8], under our hypothesis that $\mathcal{O} \otimes_\mathbb{Z} \mathbb{Q}$ has no factor of type D, all the groups G_i are isomorphic to $G_\mathbb{Q}$.
(2) There is some information about $\mathrm{Ker}^1(\mathbb{Q}, G)$ in [Kot92b, §7], and ways to calculate it. For example, if $\mathcal{O} \otimes_\mathbb{Z} \mathbb{Q}$ is simple of type C, we have $\mathrm{Ker}^1(\mathbb{Q}, G) = \{1\}$, which explains why the moduli problem of Definition 1.14 gives a canonical model of the Siegel Shimura variety and not of a finite disjoint union of copies of it. If $\mathcal{O} \otimes_\mathbb{Z} \mathbb{Q}$ is simple of type A, then $\mathrm{Ker}^1(\mathbb{Q}, G)$ is automatically trivial "half" of the cases, and it is always isomorphic to $\mathrm{Ker}^1(\mathbb{Q}, Z(G))$; for example, it is trivial for the Shimura datum of Example 1.38.
(3) Kottwitz doesn't say much about the case where $\mathcal{O} \otimes_\mathbb{Z} \mathbb{Q}$ has simple factors of type D. The situation is complicated for many reasons: the group $G_\mathbb{Q}$ is not connected, $\mathrm{Ker}(\mathbb{Q}, G)$ is not trivial, and point (i) of Proposition 2.13 is not true in general, so we must also be careful about the choice of T. In any case, it is still true that the Shimura varieties $M_K(G^0, h)(\mathbb{C})$ are open and closed subschemes of $\mathcal{M}_{K,\mathbb{C}}$ (see [Lan12, 2.5]), and we might even canonical models out of this, but we won't pursue that here because there are other ways to get canonical models for these Shimura data (see Subsubsection 2.2.1).
(4) If we want (G, h) to be a PEL Shimura datum, then this puts pretty strict conditions on the center of G. For example, if F is a nontrivial totally real extension of $\mathbb{Q}$, then the group $\mathrm{Res}_{F/\mathbb{Q}} \mathrm{GL}_2$ (defined by $\mathrm{Res}_{F/\mathbb{Q}} \mathrm{GL}_2(R) = \mathrm{GL}_2(R \otimes_\mathbb{Q} F)$) is part of a Shimura datum, but this Shimura datum cannot be PEL; more generally, we have the issue with the group $\mathrm{Res}_{F/\mathbb{Q}} \mathrm{GSp}_{2d}$. This is somewhat annoying, as sometimes we really do want to consider the Shimura varieties for these precise groups (see for example Nekovar and Scholl's [NS16]). Fortunately, these Shimura data are of Hodge type (see Subsubsection 2.2.1), so their Shimura varieties are still understood reasonably well.

2.1.4. *Canonical integral models.* PEL moduli problems don't just give canonical models of Shimura varieties, they also give models over various localizations of the ring of integers of the reflex field, at least when $\mathrm{Ker}^1(\mathbb{Q}, G)$ is trivial. Here are some things that we can learn from this example:

(1) The ring of integers over which we can expect to have a "good" integral model depends on the level K. More precisely, to have a good integral model defined over $\mathcal{O}_{F,(p)}$, we need K to be of the form $K^p K_p$, where $K^p \subset G(\mathbb{A}_f^p)$ and K_p is a hyperspecial maximal compact subgroup of $G(\mathbb{Q}_p)$. Likewise, the Hecke correspondences that will extend to finite étale morphisms between integral models over $\mathcal{O}_{F,(p)}$ are the ones that are trivial at p, i.e. defined by elements of $G(\mathbb{A}_f^p)$.
(2) We need a notion of what a "good" integral model is. If the Shimura varieties for (G, h) are compact, then we can just ask for the integral model to be projective smooth over the localization of $\mathcal{O}_F$ that are using, and to have a dense generic fiber. But this does not suffice in the noncompact case.

To fix problem (2), Milne suggested only looking at models with a certain extension property. Let (G, h) be a Shimura datum, let $F = F(G, h)$, let p be a prime number at which G is unramified, and let $K_p \subset G(\mathbb{Q}_p)$ be a hyperspecial maximal compact subgroup.[18] For every level K, write $M_K = M_K(G, h)$. We want to say what a canonical integral model $\mathcal{M}_{K_p}$ over $\mathcal{O}_{F,(p)}$ of the projective system $(M_{K^p K_p})_{K^p \subset G(\mathbb{A}_f^p)}$, or of its limit M_{K_p}. The idea, first suggested by Milne in [Mil92] (see also Moonen's paper [Moo98]), is to require that, for every S in a class of "admissible test schemes" over $\mathcal{O}_{F,(p)}$, any morphism $S \otimes_{\mathcal{O}_{F,(p)}} F \to M_{K_p}$ should extend to a morphism $S \to \mathcal{M}_{K_p}$. The problem is to decide what class of admissible test schemes one should use. We will follow Kisin's presentation in [Kis10].

[18]The condition on p means that G extends to a reductive group scheme $\mathcal{G}$ over $\mathbb{Z}_p$, and then we can take $K_p = \mathcal{G}(\mathbb{Z}_p)$.

Definition 2.15 (See [Kis10, 2.3.7])**.** A **canonical integral model** of $(M_{K_pK^p})_{K^p\subset G(\mathbb{A}_f^p)}$ or of M_{K_p} over $\mathcal{O}_{F,(p)}$ is a projective system

$$(\mathcal{M}_{K_pK^p})_{K^p\subset G(\mathbb{A}_f^p)}$$

of smooth $\mathcal{O}_{F,(p)}$-schemes with finite étale transition maps, with finite étale morphisms T_{g,K^p,K'^p} : $\mathcal{M}_{K_pK'^p}\to\mathcal{M}_{K_pK^p}$ for all $g\in G(\mathbb{A}_f^p)$ and K^p,K'^p open compact subgroups of $G(\mathbb{A}_f^p)$ such that $K'^p\subset K^p\cap gK^pg^{-1}$, and with an isomorphism of projective systems $\iota:(\mathcal{M}_{K_pK^p})\otimes_{\mathcal{O}_{F,(p)}}F\xrightarrow{\sim}(M_{K_pK^p})$, such that:

(a) The morphisms T_{g,K^p,K'^p} satisfy the analogue of conditions (ii), (iii) and (iv) of Definition 1.36, and they correspond to the morphisms $T_{g,K_pK^p,K_pK'^p}$ between canonical models by the isomorphism ι (in other words, ι is $G(\mathbb{A}_f^p)$-equivariant);

(b) The scheme

$$\mathcal{M}_{K_p}:=\varprojlim_{K^p}\mathcal{M}_{K_pK^p}$$

satisfies the following extension property: if S is a regular formally smooth $\mathcal{O}_{F,(p)}$-scheme, then any morphism $S\otimes_{\mathcal{O}_{F,(p)}}F\to M_{K_p}$ extends to a morphism $S\to\mathcal{M}_{K_p}$.

In particular, by applying the extension property with $S=\mathcal{M}_{K_p}$, we see that integral canonical models are unique up to unique isomorphism. Now the problem is existence. For PEL type Shimura varieties of type A or C (satisfying the condition that $\mathrm{Ker}^1(\mathbb{Q},G)$ is trivial), the PEL moduli problem will give some a canonical integral model. But what about other Shimura varieties?

2.2. **Hodge type and abelian type Shimura varieties.**

2.2.1. *Hodge type Shimura varieties.* For every $d\geqslant 1$, we denote by (GSp_{2d},h_d) the Siegel Shimura datum of Example 1.33.

The following condition was introduced in Deligne's paper [Del79, Section 2.3] and named in Milne's paper [Mil90] (at the end of Section 3).

Definition 2.16. A Shimura datum (G,h) is of **Hodge type** if there exists an integer $d\geqslant 1$ and a morphism of Shimura data $u:(G,h)\to(\mathrm{GSp}_{2d},h_d)$ such that the underlying morphism of algebraic groups $G\to\mathrm{GSp}_{2d}$ is injective. We also say that the corresponding Shimura variety is of Hodge type.

Example 2.17. (1) Every Shimura datum of PEL type is of Hodge type, pretty much by definition (or by [Del79, Proposition 2.3.2]): if $(\mathcal{O},*,\Lambda,\langle\cdot,\cdot\rangle,h)$ is a PEL datum and G is the corresponding group scheme, then G embeds into the group scheme H defined by

$$H(R)=\{g\in\mathrm{End}_R(\Lambda\otimes_{\mathbb{Z}}R)\mid\exists c(g)\in R^\times,\ \langle g(\cdot),g(\cdot)\rangle=c(g)\langle\cdot,\cdot\rangle\}$$

for every commutative ring R. As the alternating pairing $\langle\cdot,\cdot\rangle$ is nondegenerate by condition (5)(b) of Definition 2.1, and as all nondegenerate alternating pairing on $\Lambda\otimes_{\mathbb{Z}}\mathbb{Q}$ are equivalent, we have $H_{\mathbb{Q}}\simeq\mathrm{GSp}_{2d,\mathbb{Q}}$ for $2d=\dim_{\mathbb{Q}}(\Lambda\otimes_{\mathbb{Z}}\mathbb{Q})$. Let h' be he composition of $h:\mathbb{S}\to G_{\mathbb{R}}$ and of the embedding $G_{\mathbb{R}}\to H_{\mathbb{R}}=\mathrm{GSp}_{2d,\mathbb{R}}$. We have seen in the discussion after Proposition 2.2 that $h:\mathbb{C}\to\mathrm{End}_{\mathbb{R}}(\Lambda\otimes_{\mathbb{Z}}\mathbb{R})$ induces a Hodge structure of type $\{(-1,0),(0,-1)\}$ on $\Lambda\otimes_{\mathbb{Z}}\mathbb{C}$,[19] so h' satisfies condition (b) of Definition 1.32. It also satisfies condition (c) of Definition 1.32 because the $\mathbb{R}$-bilinear pairing $\langle\cdot,h(i)(\cdot)\rangle$ on $\Lambda\otimes_{\mathbb{Z}}\mathbb{R}$ is symmetric definite positive. Finally, for every $a\in\mathbb{R}^\times$, the element $h(a)$ of $\mathrm{End}_{\mathbb{R}}(\Lambda\otimes_{\mathbb{Z}}\mathbb{R})$ is $a\cdot\mathrm{id}$ (because h is a morphism of $\mathbb{R}$-algebras), so $h'(a)=aI_{2d}\in\mathrm{GSp}_{2d}(\mathbb{R})$. This implies that h' and h_d are conjugated by $\mathrm{GSp}_{2d}(\mathbb{R})$.

(2) The list of groups G that have Shimura data of Hodge type is given (at least in theory) in [Del79, §2.3]. For example, the group G can be of type B, while that is not possible for PEL type Shimura data.

The following result is due to Deligne; it follows from [Del71, Corollaire 5.7]), which was already cited as Proposition 1.45.

[19]Which means that, in the decomposition $\Lambda\otimes_{\mathbb{Z}}\mathbb{C}=\bigoplus_{p,q\in\mathbb{Z}}V^{p,q}$ induced by h, we have $V^{p,q}=0$ unless $(p,q)\in\{(-1,0),(0,-1)\}$.

Proposition 2.18. *Every Shimura variety of Hodge type admits a canonical model.*

Remark 2.19. There is a general philosophy that Shimura varieties should be moduli spaces of motives (the conditions that we put on a Shimura datum (G,h) are basically there to force the $G(\mathbb{R})$-conjugacy class of h to be a parameter space of Hodge structures); see [Mil90, §3] for more precise hopes.

For (G,h), we are a bit closed to that hope: we can prove that $M_K(G,h)(\mathbb{C})$ is a moduli space of abelian varieties with Hodge cycles of a certain type and level structure (see [Mil90, Theorem 3.11]). As Hodge cycles on complex abelian varieties are absolute by a theorem of Deligne (see [DMOS82, Chapter I, Theorem 2.11]), we can also see the action of $\mathrm{Aut}(\mathbb{C}/F)$ (where $F = F(G,h)$) on this modular interpretation. So far, this has not allowed people to give a moduli interpretation of the integral models of the Shimura variety of (G,h), but it does help with the construction of integral models, that we now discuss.

Let (G,h) be a Shimura datum of Hodge type, and let $u : G \to \mathrm{GSp}_{2d}$ be an injective morphism inducing a morphism of Shimura data $(G,h) \to (\mathrm{GSp}_{2d}, h_d)$. Let $F = F(G,h)$ be the reflex field of (G,h). We fix a prime number p such that $G_{\mathbb{Q}_p}$ extends to a reductive group scheme $\mathcal{G}$ over $\mathbb{Z}_p$. To simplify the presentation, we will assume that the embedding $G_{\mathbb{Q}_p} \to \mathrm{GSp}_{2d,\mathbb{Q}_p}$ extends to an embedding $\mathcal{G} \to \mathrm{GSp}_{2d,\mathbb{Z}_p}$, though that is not necessary. We set $K_p = \mathcal{G}(\mathbb{Z}_p)$ and $K'_p = \mathrm{GSp}_{2d}(\mathbb{Z}_p)$.

For each sufficiently small open compact subgroup K^p of $G(\mathbb{A}_f^p)$, we fix an open compact subgroup K'^p of $\mathrm{GSp}_{2d}(\mathbb{A}_f^p)$ such that, setting $K = K_pK^p$ and $K' = K'_pK'^p$, the morphism u defines a closed immersion $M_K(G,h) \to M_{K'}(\mathrm{GSp}_{2d}, h_d)_F$ (this is possible by [Del71, Proposition 1.15]).

Let $\mathcal{M}_{K'}$ be the integral model of $M_{K'}(G,h)$ over $\mathcal{O}_{F,(p)}$ given by base change from its canonical integral model. We denote by $\mathcal{M}_K$ the normalization of the closure of the image of $M_K(G,h)$ in $\mathcal{M}_{K'} \supset M_{K'}(G,h)_F$. Kisin proved the following result.

Theorem 2.20 (See [Kis10], Theorem 2.3.8)**.** *Suppose that $p > 2$. The limit $\varprojlim_{K^p} \mathcal{M}_{K_pK^p}$ is a canonical integral model of the Shimura variety of (G,h).*

In particular, the schemes $\mathcal{M}_K$ do not depend on the choice of K'^p or on the embedding $G \to \mathrm{GSp}_{2d}$.

Remark 2.21. In fact, [Kis10, Theorem 2.3.8] is more general, and gives a construction of a canonical integral model without the assumption on the embedding $G_{\mathbb{Q}_p} \to \mathrm{GSp}_{2d,\mathbb{Q}_p}$. It even allows the case $p = 2$ under some conditions.

2.2.2. *Abelian type Shimura varieties.*

Definition 2.22. Let (G,h) be a Shimura datum. We say that (G,h) is of **abelian type** if there exists a Shimura datum of Hodge type (G_1,h_1) and a central isogeny $G_1^{\mathrm{der}} \to G_{\mathrm{der}}$ that induces an isomorphism of Shimura data $(G_{1,\mathrm{ad}}, h_{1,\mathrm{ad}}) \xrightarrow{\sim} (G_{\mathrm{ad}}, h_{\mathrm{ad}})$, where $h_{1,\mathrm{ad}}$ (resp. h_{ad}) is the composition of h_1 (resp. h) and of the quotient morphism $G_1 \to G_{1,\mathrm{ad}}$ (resp. $G \to G_{\mathrm{ad}}$). In that case, we also say that the corresponding Shimura varieties are of the abelian type.

Another way to formulate the definition is to say that a Shimura variety $M_K(G,h)(\mathbb{C})$ is of abelian type if all its connected components are finite quotients of connected components of Shimura varieties of Hodge type (see [Del79, §9]). Deligne has classified all connected Shimura varieties of abelian type in [Del79, 2.3]. The very rough upshot is that all Shimura data (G,h) with G of type A, B and C are of abelian type; if G is of type D, it is complicated, and if G is of type E_6 or E_7, then the Shimura datum is never of abelian type. See page 61 of [Lan20] for more details.

Theorem 2.23 (Deligne, see Corollaire 2.7.21 of [Del79])**.** *Let (G,h) be a Shimura datum of abelian type. Then the Shimura variety of (G,h) admits a canonical model.*

In fact, Deligne reduces the construction of a canonical model of $M_K(G,h)(\mathbb{C})$ to that of canonical models of its connected components (over finite extensions of $F(G,h)$). See for example [Del79, Corollaire 2.7.18].

Remark 2.24. Shimura varieties of abelian type are *not* moduli spaces of abelian varieties in general. However, Milne proved in [Mil94] that they are moduli spaces of motives if $h \circ w : \mathrm{GL}_{1,\mathbb{R}} \to G_{\mathbb{R}}$ is defined over $\mathbb{Q}$, and this leads to a more direct proof of existence of their canonical models even without that condition.

Theorem 2.25 (Milne, see Theorem 3.31 of [Mil94]; see also Brylinski's paper [Bry83])**.** *Let (G, h) be a Shimura datum of abelian type such that $h \circ w : \mathrm{GL}_{1,\mathbb{R}} \to G_{\mathbb{R}}$ is defined over $\mathbb{Q}$. Then each $M_K(G, h)$ is a moduli space of abelian motives (over the reflex field of (G, h)).*

By reducing to the case of Shimura varieties of Hodge type, Kisin is able to prove the existence of canonical integral models of Shimura varieties of abelian type.

Theorem 2.26 (See [Kis10], Corollary 3.4.14)**.** *Let (G, h) be a Shimura datum of abelian type, let $p > 2$ be a prime number such that $G_{\mathbb{Q}_p}$ extends to a reductive group scheme $\mathcal{G}$ over $\mathbb{Z}_p$, and let $K_p = G(\mathbb{Z}_p)$. Then $M_{K_p} := \varprojlim_{K^p \subset G(\mathbb{A}_f^p)} M_{K_p K^p}(H, h)$ admits a canonical integral model over $\mathcal{O}_{F,(p)}$, where $F = F(G, h)$.*

2.3. **General Shimura varieties.** Let (G, h) be a Shimura datum that is not of abelian type. Then we know very little, but we do know that canonical models exist. This results was first claimed in the paper [Mil83] of Milne (based on earlier results of Kazhdan and Borovoi), but there was a gap in the proof, which was fixed by Moonen in [Moo98, §2].

Theorem 2.27 (Milne–Moonen, cf. [Moo98, Section 2])**.** *Let (G, h) be a Shimura datum. Then the Shimura variety of (G, h) admits a canonical model.*

In fact, [Moo98, §2] contains a good summary of the different construction methods of canonical models. The proof in the general case does not proceed by reduction to the case of Siegel modular varieties (unlike the previous proofs in the abelian type case), but uses results of Borovoi, Deligne, Milne and Shih on a conjecture of Langlands, that says that a conjuate of a Shimura variety over $\mathbb{C}$ by an automorphism of $\mathbb{C}$ is still a Shimura variety (Langlands's conjecture is much more precise than this, see for example [Moo98, Theorem 2.4]).

2.4. **Kottwitz's simple Shimura varieties.** After considering more and more complicated Shimura varieties in the previous subsections, we will now introduce a very simple PEL case, that has been studied by Kottwitz in [Kot92a]. These Shimura varieties are simple for several reasons:

- they are compact;
- they are PEL of type A, hence moduli spaces of abelian schemes with extra structures;
- their reductive group has no endoscopy (see below for a more precise statement).

2.4.1. *Inner forms of unitary groups.* We fix a totally real extension F_0 of $\mathbb{Q}$ and a totally imaginary quadratic extension F of F_0. Such an extension F of $\mathbb{Q}$ is called a **CM extension**. We denote by $z \mapsto \overline{z}$ the nontrivial element of $\mathrm{Gal}(F/F_0)$.

Let n be a positive integer. The quasi-split unitary group $\mathrm{U}^*(n)$ over F_0 is defined to be the unitary group of the Hermitian F_0-space F^n, with the form

$$((x_1, \ldots, x_n), (y_1, \ldots, y_n)) \mapsto \sum_{i=1}^{n} x_i \overline{y_{n+1-i}}.$$

In other words, for every commutative F_0-algebra R, we have

$$\mathrm{U}^*(n) = \{g \in \mathrm{GL}_n(F \otimes_{F_0} R) \mid {}^t\overline{g} J g = J\},$$

where J is the $n \times n$ anti-diagonal matrix with nonzero coefficients equal to 1.

We want to describe all inner forms of $\mathrm{U}^*(n)$. Remember that, if G and H are algebraic groups over a field k, we say that they are **inner forms** of each other if there exists an isomorphism $\varphi : G_{\overline{k}} \xrightarrow{\sim} H_{\overline{k}}$ such that, for every $\sigma \in \mathrm{Gal}(\overline{k}/k)$, the automorphism $\varphi^{-1} \circ {}^\sigma\varphi$ of $G(\overline{k})$ is inner (where ${}^\sigma\varphi$ is the isomorphism $\sigma \circ \varphi : G_{\overline{k}} \xrightarrow{\sim} H_{\overline{k}}$). Inner forms of G are in bijection with elements of $\mathrm{H}^1(k, G_{\mathrm{ad}}) := \mathrm{H}^1(\mathrm{Gal}(\overline{k}/k), G_{\mathrm{ad}})$, where $G_{\mathrm{ad}} = G/Z(G)$ with $Z(G)$ the center of G.

Here, observing that $\mathrm{U}^*(n)_{\mathrm{ad}}$ is the group of automorphisms of the couple formed by the central simple algebra $M_n(F)$ over F with the positive involution $g \mapsto J g^* J$,[20] we see that inner forms of

[20]See Definition 2.1 for the definition of a positive involution.

$\mathrm{U}^*(n)$ are exactly the groups $\mathrm{U}(B,*)$, where B is a central simple algebra over F and $*$ is a positive involution on B extending the involution $z \mapsto \overline{z}$ on F, and $\mathrm{U}(B,*)$ is the F_0-group defined by

$$G(R) = \{g \in B \otimes_{F_0} R \mid gg^* = 1\},$$

for R a commutative F_0-algebra.

On the other hand, if G is an inner form of $\mathrm{U}^*(n)$, then $G_{F_{0,v}}$ is an inner form of $\mathrm{U}^*(n)_{F_{0,v}}$, for every place v of F_0, which is isomorphic to $\mathrm{U}^*(n)_{F_{0,v}}$ itself for all but finitely many v. Inner forms over local fields are easier to classify (because the absolute Galois groups of local fields are simpler); but then we must be able to decide when a family of inner forms of the $\mathrm{U}^*(n)_{F_{0,v}}$ comes from a "global" inner form of $\mathrm{U}^*(n)$ (defined over F_0). Using Galois cohomology calculations or the calculation of the Brauer groups of local and global fields, we get the following results (see for example Clozel's paper [Clo91, §2]):

Proposition 2.28. *Let v be a place of F_0.*

- *Suppose that v is finite and does not split in F. If n is odd, then the only inner form of $\mathrm{U}^*(n)_{F_{0,v}}$ is $\mathrm{U}^*(n)_{F_{0,v}}$ itself (up to isomorphism). If n is even, then there are two isomorphism classes of inner forms of $\mathrm{U}^*(n)_{F_{0,v}}$.*
- *Suppose that v splits in F (in particular, v is finite), and let w be a place of F above v. Then $\mathrm{U}^*(n)_{F_{0,v}} \simeq \mathrm{GL}_{n,F_{0,v}} \simeq \mathrm{GL}_{n,F_w}$, and its inner forms are (up to isomorphism) the groups $\mathrm{GL}_m(D)$, for m dividing n and D a central division algebra over F_w of dimension $(n/m)^2$.*
- *Suppose that v is infinite, hence a real place of F_0. Then the inner forms of $\mathrm{U}^*(n)_{F_{0,v}}$ are (up to isomorphism) the real unitary groups $\mathrm{U}_{p,q}$ of signature (p,q), for $p+q=n$, and we have $\mathrm{U}_{p,q} \simeq \mathrm{U}_{r,s}$ if and only if $(r,s)=(p,q)$ or (q,p).*

Suppose that n is even, let v be a place of F_0 and let G be an inner form of $\mathrm{U}^*(n)_{F_{0,v}}$.

- If v is finite and does not split in F, set $\epsilon(G)=1$ if $G \simeq \mathrm{U}^*(n)_{F_{0,v}}$ and $\epsilon(G)=-1$ otherwise;
- If v is finite and splits in F, set $\epsilon(G)=(-1)^m$ if $G \simeq \mathrm{GL}_m(D)$ with m dividing n and D a central division algebra over F_w of dimension $(n/m)^2$;
- If v is infinite, set $\epsilon(G)=(-1)^{n/2-p}$ if $G \simeq \mathrm{U}_{p,q}$.

Proposition 2.29. *Consider a family $(G_v)_{v \text{ place of } F_0}$, where G_v is an inner form of $\mathrm{U}^*(n)_{F_{0,v}}$ for every place v of F_0. Suppose that $G_v \simeq \mathrm{U}^*(n)_{F_{0,v}}$ for all but finitely many places v.*

(i) *If n is odd, there exists an inner form G of $\mathrm{U}^*(n)$ such that $G_{F_{0,v}} \simeq G_v$ for every v.*
(ii) *If n is even, there exists an inner form G of $\mathrm{U}^*(n)$ such that $G_{F_{0,v}} \simeq G_v$ for every v if and only if $\prod_v \epsilon(G_v)=1$.*

Remark 2.30. Let G be an inner form of $\mathrm{U}^*(n)$. We know that $G \simeq \mathrm{U}(B,*)$, with B a central simple algebra over F and $*$ a positive involution on B extending the involution $z \mapsto \overline{z}$ on F. In the next subsection, it will be of interest to us to know when B is a division algebra. Let v be a place of F_0, and let $G_v = G_{F_{0,v}}$.

- If v does not split in F, and let w be a place of F above v. The group G_v is a unitary group over $F_{0,v}$, so we have $G_v \simeq \mathrm{U}(B_v,*_v)$, with B a central simple algebra over F_w and $*$ a positive involution on B extending the nontrivial element of $\mathrm{Gal}(F_w/F_{0,v})$. Also, we can write $B_v = M_{m_v}(D_v)$, with m_v dividing n and D_v a central division algebra of dimension $(n/m_v)^2$ over F_w.
- If v splits in F, then we have $G_v \simeq M_{m_v}(D_v)$, with m_v dividing n and D_v a central division algebra of dimension $(n/m_v)^2$ over $F_{0,v}$.

We can now deduce from the classification of central simple algebras over F that B is a division algebra if and only if the gcd of the family (m_v) is equal to 1. The simplest way to make sure that this condition is satisfied is to take one of the m_v equal to 1, for example to take G_v of the form $D_v^\times$ for a place of F_0 split in F, where D_v is a central division algebra over $F_{0,v}$.

2.4.2. *Simple Shimura varieties.*

Definition 2.31 (See [Kot92a], §1)**.** A **Kottwitz simple Shimura variety** is a Shimura variety defined by the Shimura datum (G,h) associated to a PEL datum $(\mathcal{O},*,\Lambda,\langle\cdot,\cdot\rangle,h)$ such that:

- $D := \mathcal{O} \otimes_{\mathbb{Z}} \mathbb{Q}$ is a division algebra with center a CM extension F of $\mathbb{Q}$;
- $*$ extends the nontrivial automorphism of F/F_0, where F_0 is the maximal totally real subextension of F;
- $\Lambda = \mathcal{O}$;
- $\langle x, y\rangle = \mathrm{Tr}_{\mathcal{O}/\mathbb{Z}}(x^* y)$ for all $x, y \in \Lambda$.

We have $\mathrm{End}_{\mathcal{O}}(\Lambda) = \mathcal{O}^{\mathrm{op}}$ (acting by right multiplication), so, for every commutative ring R,

$$G(R) = \{g \in (\mathcal{O}^{\mathrm{op}} \otimes_{\mathbb{Z}} R)^\times \mid gg^* \in R^\times\}.$$

We have a morphism $c : G \to \mathrm{GL}_1$ sending $g \in G(R)$ to $gg^* \in \mathrm{GL}_1(R)$, and we denote its kernel by G_0.

Denote by $\mathrm{Res}_{F_0/\mathbb{Q}}$ **the foncteur of Weil restriction of scalars**[21] from F_0 to $\mathbb{Q}$; in particular, if H is an algebraic group over F_0, then $\mathrm{Res}_{F_0/\mathbb{Q}} H$ is the algebraic group over $\mathbb{Q}$ defined by $\mathrm{Res}_{F_0/\mathbb{Q}} H(R) = H(R \otimes_{\mathbb{Q}} F_0)$ for every commutative $\mathbb{Q}$-algebra R. Then

$$G_{0,\mathbb{Q}} = \mathrm{Res}_{F_0/\mathbb{Q}} \mathrm{U},$$

with U the algebraic group over F_0 defined by

$$\mathrm{U}(R) = \{g \in (D^{\mathrm{op}} \otimes_{F_0} R)^\times \mid gg^* = 1\}$$

for every F_0-algebra R.

The dimension of D over F is of the form n^2 with $n \in \mathbb{N}$, and U is an inner form of the quasi-split unitary group $\mathrm{U}^*(n)$ over F_0 defined by the extension F/F_0. By Propositions 2.28 and 2.29 and Remark 2.30, we have a description of such inner forms. In particular, we know that

$$G_{0,\mathbb{R}} \simeq \prod_{\tau \in \Phi} U_{p_\tau, q_\tau},$$

where Φ is the set of real places of F_0 and, for every $\tau \in \Phi$, p_τ and q_τ are nonnegative integers such that $p_\tau + q_\tau = n$. Note that Definition 2.31 says nothing about $h : \mathbb{S} \to G_{\mathbb{R}}$, but we now know what morphism h to choose: if $z \in \mathbb{C}^\times$ is of norm 1, we take

$$h(z) = \left(\begin{pmatrix} zI_{p_\tau} & 0 \\ 0 & \overline{z} I_{q_\tau} \end{pmatrix} \right)_{\tau \in \Phi} \in \prod_{\tau \in \Phi} U_{p_\tau, q_\tau} \simeq G_0(\mathbb{R}).$$

Then there is a unique extension h to a morphism $h : \mathbb{C}^\times \to G(\mathbb{R})$ such that $h(a) = a \cdot \mathrm{id}$ if $a \in \mathbb{R}^\times$.

Note that we can choose the signatures of $G_{0,\mathbb{R}}$ arbitrarily, by manipulating what happens at finite places. On the other hand, the Shimura varieties of (G, h) are always compact, thanks to the following lemma.

Lemma 2.32. *The group G^{der} is of $\mathbb{Q}$-rank 0.*

Proof. We have $G^{\mathrm{der}} = G_0$, so $G^{\mathrm{der}}_F \simeq D^\times$. Let T be a maximal torus of G_0. Then T_F is a maximal torus in $D^\times$ (seen as an algebraic group over F). As D is a division algebra, there exists a degree n extension F' of F such that $T_F = F'^\times$, i.e. $T_F = \mathrm{Res}_{F'/F} \mathrm{GL}_{1,F'}$. This shows that the maximal split subtorus of T_F is $\{1\}$, so the maximal split subtorus of T is also $\{1\}$. □

We now discuss endoscopy.

2.4.3. *Endoscopy.* In its simplest form, endoscopy is the following phenomenon: let G be an algebraic group over a field k. Then the $G(\overline{k})$-conjugacy classes in $G(k)$ can be larger than the $G(k)$-conjugacy classes. Some vocabulary: $G(\overline{k})$-conjugacy classes in $G(k)$ are often called the **stable conjugacy classes**, and $G(\overline{k})$-conjugate elements are called **stably conjugate** (sometimes this terminology is only used for regular semisimple elements).

Example 2.33.
1. Suppose that $G = \mathrm{GL}_n$. It is then a classical exercise that any two elements of $G(k)$ that are $G(\overline{k})$-conjugate are actually $G(k)$-conjugate. We say that GL_n has no endoscopy.
2. We can generalize (1) to inner forms of GL_n, i.e. algebraic groups of the form $B^\times$, where B is a central simple algebra over k.

[21] Say "foncteur" is the French word for "functor" in English.

(3) Take $k = \mathbb{R}$ and $G = \mathrm{SL}_2$. As $\mathrm{GL}_2(\mathbb{C}) = \mathbb{C}^\times \cdot \mathrm{SL}_2(\mathbb{C})$, elements of $G(\mathbb{R})$ are $G(\mathbb{C})$-conjugate if and only if they are $\mathrm{GL}_2(\mathbb{C})$-conjugate. For example, the matrices $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ are in the same $\mathrm{SL}_2(\mathbb{C})$-conjugacy class, but we can check by a direct calculation that they are not in the same $\mathrm{SL}_2(\mathbb{R})$-conjugacy class. So SL_2 has endoscopy, and we can generalize that example to SL_n for $n \geqslant 2$.

(4) If $k = \mathbb{R}$ and $G = \mathrm{U}_n := \{g \in \mathrm{GL}_n(\mathbb{C}) \mid g^* g = I_n\}$, then again it is a classical exercise to check that $\mathrm{U}_n(\mathbb{C})$-conjugacy classes in $\mathrm{U}_n(\mathbb{R})$ coincide with $\mathrm{U}_n(\mathbb{R})$-conjugacy classes.

Remark 2.34. We care about stable conjugacy classes because, in the Langlands philosophy, groups are related via their L-groups (see 3.2.1 and 3.3.2). As inner forms have the same L-groups, this means that we should be able to move information between inner forms; but we can compare conjugacy classes in two inner forms, only stable conjugacy classes. More generally, if G and H are algebraic groups over k and their is a morphism ${}^L H \to {}^L G$ between their L-groups, then we can use this to transport stable conjugacy classes of semisimple elements from H to G.

It quickly becomes tiring to calculate conjugacy classes by hand, so we need more efficient methods to check for endoscopy. Let $\gamma \in G(k)$, and let $G_\gamma \subset G$ be the centralizer of γ.

Let $\delta \in G(k)$ be stably conjugate to γ, and let $g \in G(\overline{k})$ such that $\delta = g\gamma g^{-1}$. For every $\sigma \in \mathrm{Gal}(\overline{k}/k)$, we have

$$g\gamma g^{-1} = \delta = \sigma(\delta) = \sigma(g)\sigma(\gamma)\sigma(g)^{-1} = \sigma(g)\gamma\sigma(g)^{-1},$$

hence $\sigma(g)^{-1}g \in G_\gamma(\overline{k})$. So we get a 1-cocycle $c : \mathrm{Gal}(\overline{k}/k) \to G_\gamma(\overline{k})$, $\sigma \mapsto \sigma(g)^{-1}g$, and we can check that the image $\mathrm{inv}(\gamma, \delta)$ of this 1-cocycle in $\mathrm{H}^1(k, G_\gamma)$ does not depend on the choice of g. Moreover, the image of $\mathrm{inv}(\gamma, \delta)$ in $\mathrm{H}^1(k, G)$ is trivial.

Proposition 2.35. *Suppose that γ is semisimple. Then the map $\delta \mapsto \mathrm{inv}(\gamma, \delta)$ gives a bijection from the set of stable conjugacy classes in the $G(k)$-conjugacy class of γ to* $\mathrm{Ker}(\mathrm{H}^1(k, G_\gamma) \to \mathrm{H}^1(k, G))$.

This is particularly useful when G^{der} is simply connected and γ is regular and semisimple, because then the centralizer of γ is a maximal torus.

Example 2.36. (1) Take $k = \mathbb{R}$ and $G = \mathrm{U}_{p,q}$, the unitary group of a form with signature (p, q). Any maximal torus T of G is isomorphic to $\mathbb{S}^r \times \mathrm{U}_1^s$, with $2r + s = p + q$. (Remember that $\mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathrm{GL}_{1,\mathbb{C}}$, i.e. it is $\mathbb{C}^\times$ seen as an algebraic group over $\mathbb{R}$.) By Shapiro's lemma and Hilbert's theorem 90, we have

$$\mathrm{H}^i(\mathbb{R}, \mathbb{S}) = \mathrm{H}^i(\mathbb{C}, \mathrm{GL}_{1,\mathbb{C}}) = 0$$

and

$$\mathrm{H}^i(\mathbb{R}, \mathrm{GL}_{1,\mathbb{R}}) = 0$$

for every $i \geqslant 1$. Using the exact sequence

$$1 \to \mathrm{U}_1 \to \mathbb{S} \xrightarrow{\mathrm{Nm}} \mathrm{GL}_{1,\mathbb{R}} \to 1$$

where $\mathrm{Nm} : \mathbb{C}^\times \to \mathbb{R}^\times$ is the map $z \mapsto z\overline{z}$, we see that $\mathrm{H}^i(\mathbb{R}, \mathrm{U}_1) = 0$ if $i \geqslant 2$, and $\mathrm{H}^1(\mathbb{R}, \mathrm{U}_1) = \mathbb{R}^\times / \mathrm{Nm}(\mathbb{C}^\times) = \mathbb{Z}/2\mathbb{Z}$. So

$$\mathrm{H}^1(\mathbb{R}, T) = (\mathbb{Z}/2\mathbb{Z})^s.$$

On the other hand, by the main result of Borovoi's paper [Bor14, Theorem 9], we have

$$\mathrm{H}^1(\mathbb{R}, G) = \mathrm{H}^1(\mathbb{R}, T_0)/W_0,$$

where T_0 is a maximal torus of G with minimal $\mathbb{R}$-rank, i.e. $T_0 \simeq \mathrm{U}_1^{p+q}$, and W_0 is the group of $\mathbb{R}$-points of the Weyl group scheme W_{T_0}. As T_0 is anisotropic, the group scheme W_{T_0} is constant, so $W_0 = \mathfrak{S}_{p+q}$. We finally get that

$$\mathrm{H}^1(\mathbb{R}, G) = (\mathbb{Z}/2\mathbb{Z})^{p+q}/\mathfrak{S}_{p+q},$$

where $\mathfrak{S}_{p+q}$ acts by permuting the entries. We still need to check that the map $\mathrm{H}^1(\mathbb{R}, T) \to \mathrm{H}^1(\mathbb{R}, G)$ is the obvious one, i.e. induced by the map $(\mathbb{Z}/2\mathbb{Z})^s \to (\mathbb{Z}/2\mathbb{Z})^{p+q}$ adding $p + q - s$ entries equal to 0 (it does not matter which ones, as we are taking the quotient by $\mathfrak{S}_{p+q}$), but this is not hard. We conclude that $\mathrm{Ker}(\mathrm{H}^1(\mathbb{R},T) \to \mathrm{H}^1(\mathbb{R},G))$ is trivial. Hence $G(\mathbb{R})$-conjugacy classes of regular semisimple elements coincide with $G(\mathbb{C})$-conjugacy classes.

Note that in the previous paragraph, although $\mathrm{Ker}(\mathrm{H}^1(\mathbb{R},T) \to \mathrm{H}^1(\mathbb{R},G))$ is trivial, the map $\mathrm{H}^1(\mathbb{R},T) \to \mathrm{H}^1(\mathbb{R},G)$ is not injective. This happens because $\mathrm{H}^1(\mathbb{R},G)$ is just a pointed set and not a group (unlike $\mathrm{H}^1(\mathbb{R},T)$).

(2) We now take k to be a finite extension of $\mathbb{Q}_p$, E be an unramified quadratic extension of k, and $G = \mathrm{U}^*(n)_{E/k}$ to be the quasi-split unitary group defined by that extension, i.e. the unitary group of the Hermitian k-space E^n, with the form

$$((x_1,\dots,x_n),(y_1,\dots,y_n)) \mapsto \sum_{i=1}^n x_i\overline{y_{n+1-i}},$$

where $x \mapsto \overline{x}$ is the nontrivial element of $\mathrm{Gal}(E/k)$. Any maximal torus T of G is isomorphic to $(\mathrm{Res}_{E/k}\mathrm{GL}_1)^r \times \mathrm{U}(1)^s_{E/k}$, where $\mathrm{U}(1)_{E/k}$ is the subgroup of norm elements in $\mathrm{Res}_{E/k}\mathrm{GL}_1$ and $2r+s=n$. As before, we have $\mathrm{H}^1(k,\mathrm{Res}_{E/k}\mathrm{GL}_1) = 0$ and $\mathrm{H}^1(k,\mathrm{U}(1)_{E/k}) = k^\times/\mathrm{Nm}_{E/k}(E^\times) = \mathbb{Z}/2\mathbb{Z}$.

But now we have $\mathrm{H}^1(k,G) = \mathbb{Z}/2\mathbb{Z}$. Indeed, the derived group $G^{\mathrm{der}} = \mathrm{SU}^*(n)_{E/k}$ of G is simply connected, so $\mathrm{H}^1(k,G^{\mathrm{der}}) = 0$ by a theorem of Kneser (see [Kne65]), and then we use the fact that $G = G^{\mathrm{der}} \rtimes \mathrm{U}(1)_{E/k}$. The map $\mathrm{H}^1(k,T) \to \mathrm{H}^1(k,G)$ is then just the sum map.

So we see that, for some maximal tori T, the set $\mathrm{Ker}(\mathrm{H}^1(k,T) \to \mathrm{H}^1(k,G))$ is not a singleton, hence there are stable conjugacy classes containing more than one conjugacy class. As in (1), the more anisotropic factors the torus T has, the bigger its H^1 is.

The second example shows that, if G is the general unitary group of a Kottwitz simple Shimura variety as in Definition 2.31, then it is not reasonable to expect that G will behave as GL_n and have absolutely no endoscopy over any field. In fact, what Kottwitz actually proved about these groups is the following:

Proposition 2.37 (c.f. [Kot92a, Lemma 2] and [Kot86, Theorem 6.6])**.** *Let $\overline{\mathbb{A}}$ be the restricted product of the $\overline{\mathbb{Q}}_v$, for v a place of $\mathbb{Q}$. Let γ be a semisimple element of $G(\mathbb{Q})$, and δ be an element of $G(\mathbb{A})$ that is $G(\overline{\mathbb{A}})$-conjugate to γ. Then there exists an element of $G(\mathbb{Q})$ that is $G(\mathbb{A})$-conjugate to δ.*[22]

This implies that in the trace formula for G (see 3.4.4), we will be able to group the orbital integrals on the geometric side by stable conjugacy class and obtain an expression that is easier to transfer between groups. See [Kot92a, §4].

The proof of Proposition 2.37 is a more complicated global version of the Galois cohomology calculations of Example 2.36. First Kottwitz reduces to the case of $G^{\mathrm{der}} = \mathrm{Res}_{F_0/\mathbb{Q}}\mathrm{U}$, where U is the unitary group (over F_0) given by $\mathrm{U}(R) = \{g \in (D \otimes_{F_0} R)^\times \mid gg^* = 1\}$. Then Then he reduces to a similar result for U, now involving F_0-rational points and the ring of adeles of F_0. The main point is that, if T is a maximal torus of U, then T_F is a maximal torus of $D^\times$, and maximal tori of $D^\times$ are all of the form $\mathrm{Res}_{K/F}\mathrm{GL}_1$, for K a degree n extension of F. This allows us to control the Galois cohomology of T.

3. The cohomology of Shimura varieties

In this section, we discuss techniques to calculate the cohomology of Shimura varieties, concentrating on the compact case (see at the end for the general case).

We fix a connected reductive group G over $\mathbb{Q}$; let $Z(G)$ be the center of G, and let S_G be the maximal $\mathbb{Q}$-split torus in $Z(G)$.

We will often assume that G^{der} is of $\mathbb{Q}$-rank 0, which means that G has no proper parabolic subgroup.

3.1. **Matsushima's formula.**

[22] Note however that this new element of $G(\mathbb{Q})$ may not be $G(\mathbb{Q})$-conjugate to γ.

3.1.1. *Discrete automorphic representations.* Let $A_G = S_G(\mathbb{R})^0$. If G^{der} is of $\mathbb{Q}$-rank 0, then the quotient $G(\mathbb{Q})\backslash G(\mathbb{A})/A_G$ is compact; this follows from the theorem on page 461 of the paper [MT62] by Mostow and Tamagawa and from the fact that, with the notation of that paper, we have $G(\mathbb{A}) = A_G G(\mathbb{A})^1$ (which is easy to prove). Without this assumption, the quotient $G(\mathbb{Q})\backslash G(\mathbb{A})/A_G$ is still of finite volume. We denote by L^2_G the space of complex L^2 functions on $G(\mathbb{Q})\backslash G(\mathbb{A})/A_G$. The group $G(\mathbb{A})$ acts on L_G by right translation on the argument, and this defines a continuous unitary representation of $G(\mathbb{A})$.

If $G(\mathbb{Q})\backslash G(\mathbb{A})/A_G$ is compact, the representation L^2_G decomposes as a Hilbertian sum of irreducible representations π of $G(\mathbb{A})$, with finite multiplicities $m(\pi)$. This is not true in general, but we can still consider the part $L^2_{G,\mathrm{disc}}$ of L^2_G that decomposes discretely, and we still denote by $m(\pi)$ the multiplicity of an irreducible representation π of $G(\mathbb{A})$ in $L^2_{G,\mathrm{disc}}$.

We denote by Π_G the set of equivalences classes of irreducible representations π of $G(\mathbb{A})$ such that $m(\pi) \neq 0$. Elements of Π_G are called **discrete automorphic representations** of $G(\mathbb{A})$ (or just of G). If G^{der} is of rank 0, then the concepts of automorphic representations, discrete automorphic representations and cuspidal automorphic representations coincide.

If $\pi \in \Pi_G$, then, as $G(\mathbb{A}) = G(\mathbb{A}_f) \times G(\mathbb{R})$, we can write $\pi = \pi_f \otimes \pi_\infty$, where π_f (resp. π_∞) is an irreducible representation of $G(\mathbb{A}_f)$ (resp. $G(\mathbb{R})$).

3.1.2. *The theorem.* To state Matsushima's formula, we first need some definitions. Let $\mathfrak{g}$ be the Lie algebra of $G_\mathbb{R}$, let K'_∞ be a maximal compact subgroup of $G_\mathbb{R}$, and set $K_\infty = A_G \cdot K'_\infty$. If $\pi \in \Pi_G$, then we denote by $\mathrm{H}^*(\mathfrak{g}, K_\infty; \pi_\infty)$ the $(\mathfrak{g}, K_\infty)$**-cohomology** of π_∞, i.e. the cohomology of the complex $C^q(\mathfrak{g}, K_\infty; \pi_\infty) = \mathrm{Hom}_{K_\infty}(\wedge^q(\mathfrak{g}/\mathfrak{k}), (\pi_\infty)^\infty)$, where $\mathfrak{k} = \mathrm{Lie}(K_\infty)$ and $(\pi_\infty)^\infty)$ is the space of smooth vectors in π_∞ (which is stable by K_∞ because π_∞ is K'_∞-finite and A_G acts trivially).[23]

The following theorem for connected components of complex Shimura varieties is [BW00, Corollary VII.3.4], and the adelic reformulation can be found in Arthur's paper [Art89, §2].

Theorem 3.1 (Matsushima's formula)**.** *Let (G, h) be a Shimura datum with G^{der} of $\mathbb{Q}$-rank 0. Then we have a $G(\mathbb{A}_f)$-equivariant isomorphism of graded $\mathbb{C}$-vector spaces*

$$\varinjlim_K \mathrm{H}^*(M_K(G,h)(\mathbb{C}), \mathbb{C}) \simeq \bigoplus_{\pi \in \Pi_G} \pi_f \otimes \mathrm{H}^*(\mathfrak{g}, K_\infty; \pi_\infty)^{m(\pi)},$$

where $G(\mathbb{A}_f)$ acts on the factors π_f on the right hand side.

Here $\mathrm{H}^*(M_K(G,h)(\mathbb{C}), \mathbb{C})$ is Betti cohomology with coefficients in $\mathbb{C}$.

The theorem is equivalent to the following corollary. Let K be an open compact subgroup of $G(\mathbb{A}_f)$, and remember from page 2 of Section 1 that the **Hecke algebra** $\mathcal{H}_{G,K}$ **at level** K is the space of bi-K-invariant functions from $G(\mathbb{A}_f)$ to $\mathbb{Q}$ with compact support, with the convolution product as multiplication.[24] If $\pi \in \Pi_G$, then π_f^K is a finite-dimensional representation of $\mathcal{H}_{G,K} \otimes_\mathbb{Q} \mathbb{C}$.

Corollary 3.2. *Let (G, h) be as in Theorem 3.1, and let K be an open compact subgroup of $G(\mathbb{A}_f)$. Then we have an isomorphism of graded $\mathcal{H}_{G,K}$-modules*

$$\mathrm{H}^*(M_K(G,h)(\mathbb{C}), \mathbb{C}) \simeq \bigoplus_{\pi \in \Pi_G} \pi_f^K \otimes \mathrm{H}^*(\mathfrak{g}, K_\infty; \pi_\infty)^{m(\pi)}.$$

3.2. **Étale cohomology of canonical models: the Kottwitz conjecture.** Let ℓ be a prime number. If (G, h) is a Shimura datum, then the projective system $(M_K(G,h)(\mathbb{C}))_K$ with its $G(\mathbb{A}_f)$ has a model over the reflex field $F = F(G, h)$. So the ℓ-adic étale cohomology $\mathrm{H}^*_{\text{ét}}(M_K(G,h)_{\overline{F}}, \overline{\mathbb{Q}}_\ell)$ has commuting actions of $\mathcal{H}_{G,K}$ and $\mathrm{Gal}(\overline{F}/F)$. For every isomorphism $\overline{\mathbb{Q}}_\ell \simeq \mathbb{C}$, we have comparison isomorphisms

$$\mathrm{H}^*_{\text{ét}}(M_K(G,h)_{\overline{F}}, \overline{\mathbb{Q}}_\ell) \simeq \mathrm{H}^*(M_K(G,h)(\mathbb{C}), \mathbb{C})$$

equivariant for the action of $\mathcal{H}_{G,K}$, because this action comes from the geometric Hecke correspondences. So, when G^{der} is of $\mathbb{Q}$-rank 0, Matsushima's formula tells us that the action of $\mathcal{H}_{G,K} \otimes_\mathbb{Q} \overline{\mathbb{Q}}_\ell$

[23] See Borel and Wallach's book [BW00, §I.5] for more about $(\mathfrak{g}, K)$-cohomology. The main point of us is that this is something that can in theory be calculated.

[24] Here we fixed any Haar measure on $G(\mathbb{A}_f)$ such that open compact subgroups of $G(\mathbb{A}_f)$ have rational volume.

on the cohomology groups $\mathrm{H}^i_{\text{ét}}(M_K(G,h)_{\overline{F}},\overline{\mathbb{Q}}_\ell)$ is semi-simple, that the only representations of $\mathcal{H}_{G,K}$ that appear are the π_f^K for $\pi\in\Pi_G$, and that the π_f^K-isotypic part $\mathrm{H}^i_K(\pi_f)$ of $\mathrm{H}^i_{\text{ét}}(M_K(G,h)_{\overline{F}},\overline{\mathbb{Q}}_\ell)$ is of dimension

$$\sum_{\substack{\pi'\in\Pi_G\\ \pi'_f\simeq\pi_f}} m(\pi')\dim\mathrm{H}^i(\mathfrak{g},K_\infty;\pi'_\infty).$$

We would like to calculate the action of $\mathrm{Gal}(\overline{K}/K)$ on $\mathrm{H}^i_K(\pi_f)$. We would like to calculate the action of $\mathrm{H}^i_K(\pi_f)$, and we want to state that conjecture in the simplest case. We need some preparation.

3.2.1. *The Langlands group of F.* The Langlands group $\mathcal{L}_F$ of F is a conjectural group scheme over $\mathbb{C}$ whose irreducible representations on n-dimensional vector spaces should classify the cuspidal automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$. Remember that we defined discrete automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)=\mathrm{GL}_n(F\otimes_{\mathbb{Q}}\mathbb{A})$ in Subsubsection 3.1.1; roughly speaking, cuspidal automorphic representations are discrete automorphic representations that don't arise from an automorphic representation of a Levi subgroup of GL_n via parabolic induction. Langlands conjectures that there is a bijection $\pi\mapsto\phi_\pi$ from the set of cuspidal automorphic representations π of $\mathrm{GL}_n(\mathbb{A}_F)$ to the set of equivalence classes of representations $\phi_\pi:\mathcal{L}_F\to\mathrm{GL}_n(\mathbb{C})$, and this correspondence is the one determined by local compatibilities. We call ϕ_π the **Langlands parameter** of π.

More precisely, if π is a discrete automorphic representation of $\mathrm{GL}_n(\mathbb{A}_F)$, then we can write π as a restricted tensor product $\bigotimes'_v\pi_v$ over all places of F, where π_v is an irreducible admissible representation of $\mathrm{GL}_n(F_v)$. On the other hand, for each place of v, we have the (non-conjectural) Langlands groups $\mathcal{L}_{F_v}$ of F_v, with a (conjectural) embedding $\mathcal{L}_{F_v}\subset\mathcal{L}_F$, and the (non-conjectural) local Langlands correspondence relates irreducible admissible representations of $\mathrm{GL}_n(F_v)$ and n-dimensional representations of $\mathcal{L}_{F_v}$ ("local Langlands parameters"). The local Langlands correspondence was proved independently by Harris–Taylor, Henniart and Scholze, but for our purposes it is enough to understand the *unramified local Langlands correspondence*, which is just given by the *Satake isomorphism*. If π is a cuspidal automorphic representation and we write $\pi=\bigotimes'_v\pi_v$, then we expect that, for every place v of F, the restriction $\phi_{\pi|\mathcal{L}_{F_v}}$ corresponds to π_v by the local Langlands correspondence, and that this uniquely determines ϕ_π. In fact, it should be enough to know ϕ_π at the finite places v such that π_v is unramified (hence the corresponding representation of $\mathcal{L}_{F_v}$ is given by the Satake isomorphism).

It is also expected that $\mathcal{L}_F$ canonically surjects to $\mathcal{G}_{F,\mathbb{C}}$, where $\mathcal{G}_F$ is the motivic Galois group of F, a group scheme over $\mathbb{Q}$ defined as the Tannakian group of the conjectural category of mixed motives over F.[25] The irreducible representations of $\mathcal{L}_F$ factoring through $\mathcal{G}_{F,\mathbb{C}}$ are supposed to correspond to automorphic representations satisfying a certain condition at the infinite places of F, called **algebraic automorphic representations**. On the other hand, for ever prime number ℓ, the étale ℓ-adic realization functor defines a continuous morphism of groups $\mathrm{Gal}(\overline{F}/F)\to\mathcal{G}_F(\overline{\mathbb{Q}}_\ell)$, that is supposed to be injective (by the conservativity conjecture) and have dense image (by the Tate conjecture).

Remark 3.3. The philosophy behind the Langlands group of a number field, and its relation to the motivic Galois group, are explained much better in Clozel's paper [Clo90].

Now we come back to the case of a connected reductive group G. We need to define the L-group of G; a good reference for this is [Kot84b, §1]. Let T be a maximal torus of G. The **root datum** of G is the quadruple $(X^*,\Phi,X_*,\Phi^\vee)$, where $X^*=X^*(T_{\mathbb{C}})$, $\Phi\subset X^*$ is the set of roots of $T_{\mathbb{C}}$ in $G_{\mathbb{C}}$, $X_*=X_*(T_{\mathbb{C}})$ and $\Phi^\vee\subset X_*$ is the set of coroots. As G and T are defined over $\mathbb{Q}$, this root datum has an action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. The **dual group** $\widehat{G}$ of G is the complex connected reductive group with root datum $(X_*,\Phi^\vee,X^*,\Phi)$. If we fix a pinning of $\widehat{G}$, then this defines an action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $\widehat{G}$, and the **L-group** of G if ${}^LG=\widehat{G}\rtimes W_{\mathbb{Q}}$, where $W_{\mathbb{Q}}\subset\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is the Weil group of $\mathbb{Q}$.

Remark 3.4. If ℓ is a prime number, we could define $\widehat{G}$ to be the connected reductive group over $\overline{\mathbb{Q}}_\ell$ with root datum $(X_*,\Phi^\vee,X^*,\Phi)$, and we would get a group ${}^LG$ over $\overline{\mathbb{Q}}_\ell$. We write $\widehat{G}(\overline{\mathbb{Q}}_\ell)$ and ${}^LG(\overline{\mathbb{Q}}_\ell)$ for the resulting groups, when we want to distinguish them from the complex versions. Which form

[25]To make this precise, we need a fiber functor. We fix an embedding of F into $\mathbb{C}$ and take the fiber product given by the corresponding Betti realization with $\mathbb{Q}$-coefficients.

of the L-group we use depends on the context: for Langlands parameters defined on $\mathcal{L}_F$, we use the complex form, and for Langlands parameters defined on $\mathrm{Gal}(\overline{F}/F)$, we use the ℓ-adic form.

Example 3.5. If G is of type $A,\ B,\ E,\ F$ or G, then $\widehat{G}$ is of the same type as G. If G is of type B_n (resp. C_n), then $\widehat{G}$ is of type C_n (resp. B_n). We can also relate other properties of G and $\widehat{G}$: for example, the derived group G^{der} is simply connected if and only if $Z(\widehat{G})$ is connected, and in that case $Z(\widehat{G})$ is the dual group of G/G^{der}.

Here are some examples of dual groups:

- $\widehat{\mathrm{GL}_n} = \widehat{\mathrm{U}(p,q)} = \mathrm{GL}_n(\mathbb{C})$ if $p+q=n$;
- $\widehat{\mathrm{SL}_n} = \widehat{\mathrm{SU}(p,q)} = \mathrm{PGL}_n(\mathbb{C})$ if $p+q=n$;
- $\widehat{\mathrm{PGL}_n} = \mathrm{SL}_n(\mathbb{C})$;
- $\widehat{\mathrm{Sp}_{2n}} = \mathrm{SO}(2n+1)$, $\widehat{\mathrm{GSp}_{2n}} = \mathrm{GSpin}_{2n+1}(\mathbb{C})$;
- $\widehat{\mathrm{GU}(p,q)} = \mathrm{GL}_1(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C})$.

We see that different groups can have isomorphic dual groups. In fact, if G' is an inner form of G, then the L-groups ${}^LG$ and ${}^LG'$ are isomorphic. However, as $\mathrm{U}(p,q)$ is not an inner form of GL_n, the actions of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on their dual groups are not the same, and we get non-isomorphic L-groups. If $G = \mathrm{GL}_n$ (or more generally if G split over $\mathbb{Q}$), then $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts trivially on $\widehat{G}$. On the other hand, if $G = \mathrm{U}(p,q)$ and if E is the imaginary quadratic extension of $\mathbb{Q}$ that we used to define G, then $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on $\widehat{G} = \mathrm{GL}_{p+q}(\mathbb{C})$ via its quotient $\mathrm{Gal}(E/\mathbb{Q})$, and the nontrivial element of $\mathrm{Gal}(E/\mathbb{Q})$ acts as a non-inner automorphism of $\mathrm{GL}_{p+q}(\mathbb{C})$, i.e. a conjugate of the automorphism $g \mapsto {}^tg^{-1}$ (which conjugate depends on the choice of the pinning).

Coming back to the Langlands correspondence for G, there are several complications:

- A cuspidal automorphic representation π of $G(\mathbb{A})$ should now have a Langlands parameter ϕ_π with values not in $\mathrm{GL}_n(\mathbb{C})$ but in ${}^LG$;
- There is a still a characterization of algebraic automorphic representations (conjecturally corresponding to the parameters that factors through $\mathcal{G}_{F,\mathbb{C}}$), but it is more complicated;
- Distinct cuspidal automorphic representations can have the same Langlands parameter. We say that they are in the same L-packet.

Also, if π is an algebraic cuspidal automorphic representation of G, then we expect $\phi_\pi : \mathcal{G}_{\mathbb{Q},\mathbb{C}} \to {}^LG$ to be defined over a finite extension L of $\mathbb{Q}$ in $\mathbb{C}$. Choosing a finite place λ of L over a prime number ℓ, we get a morphism from $\mathcal{G}_{\mathbb{Q}}(\overline{\mathbb{Q}}_\ell)$ into the ℓ-adic version of ${}^LG$, and this gives a morphism $\sigma_\pi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to {}^LG(\overline{\mathbb{Q}}_\ell)$, also called the Langlands parameter of π, and whose value on the Frobenius elements at big enough prime numbers p is predicted by the Satake parameter of π_p. Now the conjecture only involves well-defined objects, and we can actually try to prove it!

Remark 3.6. If we are very brave and want to classify all discrete automorphic representations of $G(\mathbb{A})$, then there is an extension of the Langlands conjecture due to Arthur. Now a discrete automorphic representation π should have a parameter $\psi_\pi : \mathcal{L}_F \times \mathrm{SL}_2(\mathbb{C}) \to {}^LG$, satisfying a long list of properties (in particular compatibility with a local version of the Arthur conjectures), and there is a somewhat explicit formula to calculate the multiplicity $m(\pi)$. For a quick review of Arthur's conjectures, see Kottwitz's paper [Kot90, §8]. Warning: if π is cuspidal, then we can recover ϕ_π from ψ_π and vice versa, but the two parameters are not equal.

3.2.2. *The Kottwitz conjecture.* Let (G,h) be a Shimura datum, and let $F = F(G,h)$ be its reflex field. We assume that G^{der} is of $\mathbb{Q}$-rank 0, so that the Shimura varieties $M_K(G,h)(\mathbb{C})$ are compact.

The conjugacy class of the cocharacter $\mu := h_{\mathbb{C}} \circ r : \mathrm{GL}_{1,\mathbb{C}} \to G_{\mathbb{C}}$ is defined over F (see Subsubsection 1.5.3), hence it defines a finite-dimensional representation r_μ of ${}^LG_F(\overline{\mathbb{Q}}_\ell) := \widehat{G}(\overline{\mathbb{Q}}_\ell) \rtimes W_F$ in the following way (see Kottwitz's paper [Kot84a, Lemma 2.1.2]): fix a maximal torus $\widehat{T}$ of $\widehat{G}$ and a Borel subgroup $\widehat{B}$ containing $\widehat{T}$ that are part of a splitting fixed by W_F. The cocharacter $h_{\mathbb{C}} \circ r$ corresponds to a unique dominant character μ of $\widehat{T}$, and we denote by V_μ the corresponding highest weight representation of $\widehat{G}(\overline{\mathbb{Q}}_\ell)$. The action of $\widehat{G}(\overline{\mathbb{Q}}_\ell)$ on V_μ extends to a unique action r_μ of ${}^LG_F(\overline{\mathbb{Q}}_\ell)$ such that W_F acts trivially on the highest weight subspace.

Example 3.7. (1) If $G = \mathrm{GSp}_{2d}$ and (G,h) is the Shimura datum of Example 1.33, then $\widehat{G} = \mathrm{GSpin}_{2d+1}(\mathbb{C})$ and $r_\mu : \widehat{G} \to \mathrm{GL}_{2^d}(\mathbb{C})$ is the spin representation.
- If $d = 1$, then $\mathrm{GSp}_{2d} = \mathrm{GL}_2$ and r_μ is the standard representation of $\widehat{\mathrm{GL}_2} = \mathrm{GL}_2(\mathbb{C})$.
- If $d = 2$, then we have an exceptional isomorphism $\mathrm{GSpin}_5(\mathbb{C}) \simeq \mathrm{GSp}_4(\mathbb{C})$, and r_μ is isomorphic to the standard representation of $\mathrm{GSp}_4(\mathbb{C})$.

(2) If $G = \mathrm{GU}(p,q)$ and (G,h) is the Shimura datum of Example 1.38, then $\widehat{G} \simeq \mathrm{GL}_1(\mathbb{C}) \times \mathrm{GL}_{p+q}(\mathbb{C})$ and r_μ is, up to twists by characters, the qth exterior power of the standard representation of $\mathrm{GL}_{p+q}(\mathbb{C})$.

Let K be an open compact subgroup, and let π be a cuspidal automorphic representation of $G(\mathbb{A})$ such that the π_f^K-isotypic part $\mathrm{H}^i_K(\pi_f)$ of $\mathrm{H}^i(M_K(G,h)(\mathbb{C}),\mathbb{C})$ is nonzero for at least one $i \in \mathbb{Z}$. Then π should be algebraic, so its Langlands parameter should give rise to a Galois representation $\sigma_\pi : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to {}^LG(\overline{\mathbb{Q}}_\ell)$ as above.

Finally, let $d = \dim M_K(G,h)$.

Conjecture 3.8 (Kottwitz, cf. [Kot90, §10] and [Kot92a, §1])**.** *There is an explicitly defined integer $a(\pi_f)$ such that we have a $\mathrm{Gal}(\overline{F}/F)$-equivariant isomorphism of virtual representations*

$$\sum_{i=0}^{2d}(-1)^i \mathrm{H}^i_K(\pi_f)(d/2) \simeq (r_\mu \circ \sigma_\pi)^{a(\pi_f)}.$$

Moreover, the integers i such that $\mathrm{H}^i_K(\pi_f^K) \neq 0$ all have the same parity.

By the Weil conjectures (proved by Deligne), the representation $\mathrm{H}^i_K(\pi_f)$ of $\mathrm{Gal}(\overline{F}/F)$ is pure of weight i for every $i \in \mathbb{Z}$, so we can separated the degrees in the formula of Conjecture 3.8 by using Frobenius weights.

Remark 3.9. (1) The conjecture as stated is in a very naive form, and almost always false. In fact, we only expect it to be true when the Shimura varieties for (G,h) are compact and when G has no endoscopy (i.e. admits no nontrivial elliptic endoscopic triple).

(2) While we cannot hope to prove the Kottwitz conjecture without first constructing the Langlands parameter of σ_π, we do know what the image by σ_π of the Frobenius element $\mathrm{Frob}_\wp$ at a place $\wp$ of F over a nice enough prime number p should be, so we can try to prove that $\mathrm{Frob}_\wp$ has the correct characteristic polynomial on the $\mathrm{H}^i_K(\pi_f)$. For the simple Shimura varieties of Subsection 2.4, Kottwitz proved this consequence of his conjecture in [Kot92a].

(3) If we know the local Langlands correspondence for G, we can also try to check that the restriction to all local Galois groups of the $\mathrm{H}^i_K(\pi_f)$ are as predicted by the Kottwitz conjecture. This is a much harder problem and we won't discuss it here.

3.3. **Applications of the Kottwitz conjecture.**

3.3.1. *The zeta function of a Shimura variety.*

Definition 3.10. Let X be a smooth proper variety over a finite field $\mathbb{F}_q$. The **Hasse–Weil zeta function** of X is the following formal power series in q^{-s}:

$$Z(X,s) = \exp\left(\sum_{n\geqslant 1} \frac{\#X(\mathbb{F}_{q^n})}{n} q^{-sn}\right).$$

Using the Grothendieck–Lefschetz fixed point formula (cf. Theorem 3.15 later), we get the following result.

Theorem 3.11 (Grothendieck, see Theorem 3.1 of [Del77])**.** *Let $\mathrm{Frob}_q \in \mathrm{Gal}(\overline{\mathbb{F}}_q/\mathbb{F}_q)$ be the geometric Frobenius (the inverse of the arithmetic Frobenius $a \mapsto a^q$). Then*

$$Z(X,s) = \prod_{i=0}^{2\dim(X)} \det(1 - q^{-s}\mathrm{Frob}_q, \mathrm{H}^i_{\acute{e}t}(X_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell))^{(-1)^{i+1}}.$$

In particular, the formal power series $Z(X,s)$ is actually a rational function on q^{-s}.

Remark 3.12. We can define the zeta function of any algebraic variety over $\mathbb{F}_q$ by the same formula, and Theorem 3.11 still holds providing we use étale cohomology with proper supports.

Now let X be proper smooth algebraic variety over a number field F. For all but finite places $\wp$ of F, the variety X has a proper smooth model $\mathcal{X}$ over $\mathcal{O}_{F,\wp}$ (we say that X has **good reduction** at $\wp$), and we set

$$\zeta_{X,\wp}(s) = \zeta_{\mathcal{X}_{\kappa(\wp)}}(\#\kappa(\wp)^{-s}),$$

where $\kappa(\wp) = \mathcal{O}_F/\wp$ is the residue field of $\wp$. By Theorem 3.11 and the specialization theorem for étale cohomology, this does not depend on the choice of the model.

If v is a finite place of F where X does not have good reduction or an infinite place, we will not give the definition of $\zeta_{X,v}(s)$; we will just say that $\zeta_{X,v}(s)$ is a rational function of $\#\kappa(v)^{-s}$ if v is finite and a product of Γ functions if v is infinite.

Definition 3.13. The **Hasse–Weil zeta function** of X is the infinite product

$$\zeta_X(s) = \prod_{v \text{ place of } F} \zeta_{X,v}(s).$$

Example 3.14. If $X = \operatorname{Spec}\mathbb{Q}$, then ζ_X is the Riemann zeta function.

A priori this product only makes sense for $\Re(s)$ big enough. The **Hasse–Weil conjecture** predicts that $\zeta_X(s)$ has a meromorphic continuation to $\mathbb{C}$ and a functional equation similar to the one of Riemann zeta function.

The conjecture seems to be out of reach in general, but for Shimura varieties we can approach it using the Kottwitz conjecture. The general idea goes as follows:

(1) Essentially by Theorem 3.11, we have an equality

$$\zeta_X(s) = \prod_{i=0}^{2\dim(X)} L(\mathrm{H}^i_{\text{ét}}(X_{\overline{F}}, \mathbb{Q}_\ell), s)^{(-1)^{i+1}},$$

where, for every continuous representation ρ of $\mathrm{Gal}(\overline{F}/F)$, we denote by $L(\rho, s)$ the **L-function of ρ**.

(2) Suppose that (G, h) is a Shimura datum such that G^{der} has $\mathbb{Q}$-rank 0, that $F = F(G, h)$ and $X = M_K(G, h)$. For every i, we have up to semi-simplification

$$\mathrm{H}^i_{\text{ét}}(M_K(G,h)_{\overline{F}}, \overline{\mathbb{Q}}_\ell) \simeq \bigoplus_{\pi \in \Pi_G} \mathrm{H}^i_K(\pi_f)^{\dim(\pi_f^K)}$$

as representations of $\mathrm{Gal}(\overline{F}/F)$.

(3) The Kottwitz conjecture predicts that $\mathrm{H}^i_K(\pi_f)$ is a sum of copies of $r_\mu \circ \sigma_\pi$, where $\sigma_\pi : \mathrm{Gal}(\overline{F}/F) \to {}^L G(\overline{\mathbb{Q}}_\ell)$ is the Langlands parameter of π and r_μ is the algebraic representation of ${}^L G$ defined in Subsubsection 3.2.2. But the local compatibility between π and σ_π implies immediately that

$$L(r_\mu \circ \sigma_\pi, s) = L(\pi, s, r_\mu),$$

where the L-function $L(\pi, s, r_\mu)$ is defined in Borel's survey [Bor79].

(4) In theory we understand the analytic properties of L-functions of automorphic representations better, so we get some information of the zeta function of $M_K(G, h)$.

In practice the automorphic L-functions that appear are usually not standard L-functions and so our understanding of them is still limited. However, this methods can still go through when r_μ is the standard representation of a classical group, such as in the case of modular curves or Picard modular surface (when $G = \mathrm{GU}(2, 1)$), see [Bor79].

3.3.2. *The global Langlands correspondence.* Let (G, h) be a Shimura datum, and let $F = F(G, h)$. In a way, the Kottwitz conjecture says that the cohomology $\varinjlim_K \mathrm{H}^*_{\text{ét}}(M_K(G,h)_{\overline{F}}, \overline{\mathbb{Q}}_\ell)$ realizes the global Langlands correspondence for those automorphic representations of $G(\mathbb{A})$ that contribute to Matsushima's formula. So we could try to use this cohomology to construct the global Langlands correspondence in that case, and then use results like the main theorem of Kottwitz's [Kot92a] to check that this does satisfy the desired compatibility with the local correspondence (at least in the unramified case). There are several problems with this approach:

(1) The representation of $\mathrm{Gal}(\overline{F}/F)$ that appears in the cohomology of $M_K(G,h)$ is not σ_π (this would not even make sense, as σ_π is a morphism into ${}^L G(\overline{\mathbb{Q}}_\ell)$) but $r_\mu \circ \sigma_\pi$;
(2) There are multiplicities (the integer $a(\pi_f)$ in the Kottwitz conjecture);
(3) We want to construct the Langlands correspondence for GL_n, not some strange unitary group;
(4) We want to get a representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, not $\mathrm{Gal}(\overline{F}/F)$;
(5) This will only ever work for "cohomological" automorphic representations, i.e. those π that appear in the cohomology of Shimura varieties. This is a condition on π_∞: roughly, we need it to have nontrivial $(\mathfrak{g}, K_\infty)$-cohomology.

All of these can be somewhat addressed, at some cost. For point (1), we can choose the group such that $\widehat{G}$ is classical and r_μ is the standard representation. This will for example be the case if $G = \mathrm{GU}(n-1,1)$, although in practice we will rather G to be a more complicated unitary group (defined by a CM extension of $\mathbb{Q}$ of degree $2r > 2$) of signature $(n-1,1) \times (n,0)^{r-1}$ at infinity, so that we can get a simple Shimura variety. Of course, this solution puts even greater restrictions on the groups that we can use, so it seems that we are making problems (3) and (4) worse.

For problem (2), we can sometimes calculate the multiplicities, and they tend to be equal to 1 for nice unitary groups.

Problem (3) can be attacked using the **Langlands functoriality principle.** The idea is that, if discrete automorphic representations of $G(\mathbb{A})$ are parametrized by morphisms $\mathcal{L}_\mathbb{Q} \times \mathrm{SL}_2(\mathbb{C}) \to {}^L G$, then, if H is another connected reductive algebraic group over $\mathbb{Q}$ and if we have a morphism ${}^L G \to {}^L H$, then we should be able to "transfer" discrete automorphic representations from G to H. As with the global Langlands correspondence or the Arthur conjectures, this principle can be made very precise at "good" primes (i.e. primes where both groups and all automorphic representations we consider are unramified) using the Satake isomorphism, so we can pin down the conjectural transfer using a local-global compatibility principle. Of course, things are not so simple: the conjectural Arthur parametrization is not bijective in general so we can only expect to transfer L-packets, so we can only expect to transfer.

One favorable case is when $\widehat{G}$ is the set of fixed points of an automorphism of $\widehat{H}$, because then we can use the (twisted) Arthur–Selberg trace formula to construct the transfer map, although this is very technically difficult; see Subsubsection 3.4.4 for a very simple instance of the (untwisted) trace formula. This is for example the case if $G = \mathrm{Sp}_{2d}$ and $H = \mathrm{GL}_{2d+1}$, if G is the group $\mathrm{U}(p,q)$ constructed using a quadratic imaginary extension E of $\mathbb{Q}$ and $H = \mathrm{GL}_{p+q}(E)$, seen as an algebraic group over $\mathbb{Q}$, or if $G = \mathrm{GL}_n$ and $H = \mathrm{GL}_n(E)$ (seen as an algebraic group over $\mathbb{Q}$ again) for $E/\mathbb{Q}$ as before (or more generally a cyclic extension). Using trace formula techniques, we can transfer discrete automorphic representations of $\mathrm{GL}_{p+q}(\mathbb{A})$ to $\mathrm{GL}_{p+q}(\mathbb{A}_E)$, and then back down to $\mathrm{U}(p,q)(\mathbb{A})$. One caveat is that this only works for representations of $\mathrm{GL}_{p+q}(\mathbb{A}_E)$ that are conjugate self-dual (because their parameter should be stable by the involution of $\widehat{\mathrm{GL}_{p+q}(E)}$ whose set of fixed points is equal to $\widehat{\mathrm{U}(p,q)}$). This kind of construction will also work for more general unitary groups defined using other CM fields. Again, this is difficult and there are many technical problems, including annoying congruence conditions on $p+q$. The first results that we are aware of in this direction are due to Clozel (see [Clo91], which rests on the results of Kottwitz from [Kot92a]). More recent and more powerful results can be found in the papers [Shi11] of Shin and [SS90] of Scholze–Shin [SS13]; these require unitary Shimura varieties that are not simple, so they use a more complicated form of the Kottwitz conjecture, as well as the fundamental lemma.

Problem (4) can be addressed by "gluing" Galois representations constructed using different Shimura varieties. More precisely, suppose that you have a discrete automorphic representation π of $\mathrm{GL}_n(\mathbb{A})$, and that you know that it is self-dual and that π_∞ satisfies all the required conditions for π to transfer to cohomological automorphic representations of unitary groups. By varying the CM field E_i (and the corresponding unitary group), we get a family of representations σ_i of $\mathrm{Gal}(\overline{E}_i/E_i)$, compatible by looking at what happens at "nice" prime numbers, and then glue then if we chose the E_i disjoint enough.

Problem (5) requires entirely new techniques; a related problem is that we can only ever transfer self-dual automorphic representations of GL_n to groups that have Shimura varieties, but we don't have time to explain them here. See for example the papers [CH13] of Chenevier–Harris, [HLTT16] of Harris–Lan–Taylor–Thorne, [Sch15] of Scholze, or [Box15] of Boxer.

3.4. **Proving the Kottwitz conjecture.** We will present the original approach to the Kottwitz conjecture, due to Ihara, Langlands and Kottwitz. We will not talk about the more refined approaches through Igusa varieties or through Scholze's methods, that also allow us to understand the cohomology at ramified primes.

The situation is the following: we have a Shimura datum (G, h) such that G^{der} is of $\mathbb{Q}$-rank 0 (so that the Shimura varieties $M_K(G, h)$ are projective), and a discrete automorphic representation π of $G(\mathbb{A})$. We fix a small enough open compact subgroup K of $G(\mathbb{A}_f)$, and we write $F = F(G, h)$. We are trying to understand the representation of $\mathrm{Gal}(\overline{F}/F)$ on the π_f^K-isotypic components $\mathrm{H}^i_K(\pi_f)$ in the cohomology groups $\mathrm{H}^i_{\text{ét}}(M_K(F, h)_{\overline{F}}, \overline{\mathbb{Q}}_\ell)$.

3.4.1. *The specialization theorem.* We restrict our attention to the action of the local Galois group at finite places $\wp$ of F that are "nice enough", i.e. such that $M_K(G, h)$ has a proper smooth integral model $\mathcal{M}_K$ over $\mathcal{O}_{F,\wp}$. Remember from Section 2 that, if (G, h) is of abelian type, then we have control over these places $\wp$ (in terms of G and K), but in general we only know that this holds for all but finitely many $\wp$.

Fix $\wp$ as in the previous paragraph, and let $\kappa(\wp)$ be the residue field of $\wp$. Then we have an exact sequence

$$1 \to I_\wp \to \mathrm{Gal}(\overline{F}_\wp/F_\wp) \to \mathrm{Gal}(\overline{\kappa(\wp)}/\kappa(\wp)) \to 1,$$

where $I_\wp$ is the **inertia group** at $\wp$. Moreover, as $\kappa(\wp)$ is a finite field, its absolute Galois group is topologically generated by the geometric Frobenius $\mathrm{Frob}_\wp$, which is the inverse of the arithmetic Frobenius $a \mapsto a^{\#\kappa(\wp)}$.

The specialization theorem for étale cohomology (which follows from the proper and smooth base change theorems) tells us that the representations $\mathrm{H}^i_{\text{ét}}(M_K(F, h)_{\overline{F}_\wp}, \overline{\mathbb{Q}}_\ell)$ of $\mathrm{Gal}(\overline{F}_\wp/F_\wp)$ are **unramified**, i.e. that $I_\wp$ acts trivially on them, and that we have isomorphisms of representations of $\mathrm{Gal}(\overline{\kappa(\wp)}/\kappa(\wp))$:

$$\mathrm{H}^i_{\text{ét}}(M_K(F, h)_{\overline{F}_\wp}, \overline{\mathbb{Q}}_\ell) \simeq \mathrm{H}^i_{\text{ét}}(\mathcal{M}_{K,\overline{\kappa(\wp)}}, \overline{\mathbb{Q}}_\ell). \tag{3.1}$$

If we have a Hecke correspondence defined by $g \in G(\mathbb{A}_f)$ and $K' \subset K \cap gKg^{-1}$, then these isomorphisms will be compatible with the corresponding Hecke operator, provided that $M_{K'}(G, h)$ also has a proper smooth model over $\mathcal{O}_{F,\wp}$ and that the Hecke correspondence extends to the model over $\mathcal{O}_{F,\wp}$.

So we can now work over the finite field $\kappa(\wp)$.

3.4.2. *The Grothendieck–Lefschetz fixed point formula and Deligne's conjecture.* Suppose we were only trying to understand the representation of $\mathrm{Gal}(\overline{F}_\wp/F_\wp)$ on $\mathrm{H}^*_{\text{ét}}(M_K(G, h)_{\overline{F}_\wp}, \overline{\mathbb{Q}}_\ell)$ (and not on the Hecke isotypic components), and that we only cared about semisimplications. The, by the isomorphism (3.1), it would suffice to calculate the characteristic polynomial of $\mathrm{Frob}_\wp$ acting on $\mathrm{H}^*_{\text{ét}}(\mathcal{M}_{K,\overline{\kappa(\wp)}}, \overline{\mathbb{Q}}_\ell)$. We can do this thanks to the Grothendieck–Lefschetz fixed point formula (already mentioned in Subsubsection 3.3.1 on the zeta function):

Theorem 3.15 (Grothendieck, cf. Théorème 3.2 of [Del77])**.** *Let $\mathbb{F}_q$ be a finite field, $\overline{\mathbb{F}}_q$ be an algebraic closure of $\mathbb{F}_q$, ℓ be a prime number different from the characteristic of $\mathbb{F}_q$ and Frob_q be the geometric Frobenius automorphic of $\overline{\mathbb{F}}_q$. Then, for every separated $\mathbb{F}_q$-scheme of finite type X and every positive integer r, we have*

$$\sum_{i=0}^{2\dim(X)} (-1)^i \,\mathrm{Tr}(\mathrm{Frob}_q^r, \mathrm{H}^i_{\acute{e}t,c}(X_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)) = \#X(\mathbb{F}_{q^r}),$$

where $\mathbb{F}_{q^r}$ is the unique extension of $\mathbb{F}_q$ of degree r in $\overline{\mathbb{F}}_q$.

To understand the action of $\mathrm{Gal}(\overline{F}_\wp/F_\wp)$ on the Hecke isotypic components, we need to calculate the traces of Hecke operators multiplied by powers of $\mathrm{Frob}_\wp$ on $\mathrm{H}^*_{\text{ét}}(M_K(G,h)_{\overline{F}_\wp}, \overline{\mathbb{Q}}_\ell)$. For this, we use a generalization of the Grothendieck–Lefschetz fixed point formula called Deligne's conjecture. We will not state the most general version here, but just the consequence that we need.

We use the notation of Theorem 3.15. Let X, X' be separated $\mathbb{F}_q$-schemes of finite type, and let $a, b : X' \to X$ be finite morphisms. Suppose that the trace morphism $\mathrm{Tr}_b : b_* b^* \to \mathrm{id}$ exists.[26] Let u be the endomorphism of $\mathrm{H}^*_{\text{ét,c}}(X_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)$ that is the composition of the pullback by a map $\mathrm{H}^*_{\text{ét,c}}(X_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell) \to \mathrm{H}^*_{\text{ét,c}}(X'_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)$ and of the map $\mathrm{H}^*_{\text{ét,c}}(X'_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell) \to \mathrm{H}^*_{\text{ét,c}}(X_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)$ induced by Tr_b. Finally, we denote by $F_X : X \to X$ the Frobenius morphism (which is identity on the underlying topological spaces and raises functions to the qth power).

Theorem 3.16. *For any big enough positive integer r, we have*

$$\sum_{i=0}^{2\dim(X)} (-1)^i \operatorname{Tr}(\mathrm{Frob}_q^r \cdot u, \mathrm{H}^*_{\acute{e}t,c}(X_{\overline{\mathbb{F}}_q}, \mathbb{Q}_\ell)) = \#\{x' \in X'(\overline{\mathbb{F}}_q) \mid a(x') = F_X^r \circ b(x')\}.$$

Although it is still often refereed to as "Deligne's conjecture", this statement is a theorem: if X, X' are reductions of Shimura varieties and a, b are Hecke operators, it was proved by Pink in [Pin92]. In the general case, it was proved independently by Fujiwara [Fuj97] and Varshavsky [Var07].

So now we need to understand the set of points the Shimura varieties $M_K(G,h)$ and $M_{K'}(G,h)$ (or rather of their integral models) over $\overline{\kappa(\wp)}$, as well as the action of the Frobenius and of Hecke correspondences on these points.

3.4.3. *The Langlands–Rapoport conjecture.* The Langlands–Rapoport conjecture gives a purely group-theoretical description of the set of points of a Shimura variety over the algebraic closure of the residue field at a good place $\wp$ of the reflex field, as well as a description of the action of the Frobenius at $\wp$ and of Hecke operators on this set. We will only give a rough statement here; see for example Milne's paper [Mil92] for a precise statement of this conjecture.

Let (G,h) be a Shimura datum with reflex field F, let p be a prime number and $\wp$ a place of F above p. Suppose that G and F are unramified at p, and let $K = K^p K_p$ be a level with $K_p \subset G(\mathbb{Q}_p)$ hyperspecial. Then we expect the Shimura variety $M_K(G,h)$ to have a "nice" model $\mathcal{M}_K$ over $\mathcal{O}_{F,\wp}$, and the Langlands–Rapoport conjecture gives a description of $\mathcal{M}_K(\mathbb{F}_q)$, where $\mathbb{F}_q$ is an extension of $\mathcal{O}/\wp$. The description has the rough shape

$$\mathcal{M}_K(\mathbb{F}_q) = \coprod_\varphi I_\varphi(\mathbb{Q}) \backslash (X^p(\varphi) \times X_p(\varphi)),$$

where φ is in a certain set of parameters, I_φ is an algebraic group over $\mathbb{Q}$ (the centralizer of φ), $X^p(\varphi)$ involves the finite adeles outside of p and $X_p(\varphi)$ is a purely p-adic objects. The conjecture also includes a description of the actions of the Frobenius and of the Hecke operators.

It would take too long to explain what the parameters φ are, but we can say that they give rise to triples $(\gamma_0, \gamma, \delta)$, where:

- γ_0 is a semisimple element of $G(\mathbb{Q})$, given up to $G(\overline{\mathbb{Q}})$-conjugacy;
- $\gamma = (\gamma_\ell)_{\ell \neq p}$ is an element of $G(\mathbb{A}_f^p)$, given up to $G(\mathbb{A}_f^p)$-conjugacy, and such that γ_ℓ and γ_0 are conjugated under $G(\overline{\mathbb{Q}}_\ell)$ for every $\ell \neq p$;
- δ is an element of $G(F)$, where L is the unramified extension of degree $r = [\mathbb{F}_q : \mathcal{O}/\wp]$ for $F_\wp$, such that, if we denote by $\sigma \in \mathrm{Gal}(L, \mathbb{Q}_p)$ the lift of the (arithmetic) Frobenius, then $N(\delta) = \delta\sigma(\delta)\cdots\sigma^{r-1}(\delta)$ is $G(\overline{\mathbb{Q}}_p)$-conjugate to γ_0.

There are some more conditions on the triple $(\gamma_0, \gamma, \delta)$, see for example [Kot90, §2]. In any case, we then want to take $I(\varphi) = G_{\gamma_0}$, the centralizer of γ_0 in G,[27] then

$$X^p(\varphi) = \{g \in G(\mathbb{A}_f^p)/K^p \mid g^{-1}\gamma g \in K^p\}$$

[26] This holds for example if b is flat or if X and X' are both normal.

[27] At least if G^{der} is simply connected, which is a hypothesis that Kottwitz makes in [Kot90].

and

$$X_p(\varphi) = \{g \in G(L)/G(\mathcal{O}_L) \mid g^{-1}\delta\sigma(g) \in G(\mathcal{O}_L)\mu_h(\varpi_L)G(\mathcal{O}_L)\},$$

where we extended G to reductive group scheme over $\mathbb{Z}_p$, ϖ_L is a uniformizer of L, and where μ_h is the morphism $h_{\mathbb{C}} \circ r : \mathrm{GL}_{1,\mathbb{C}} \to G_{\mathbb{C}}$; as seen as a conjugacy class of morphisms $\mathrm{GL}_{1,\overline{\mathbb{Q}}_p} \to G_{\overline{\mathbb{Q}}_p}$ that is stable by $\mathrm{Gal}(\overline{\mathbb{Q}}_p/L)$; as G is quasi-split over L for r big enough, we may assume up to taking r big enough that μ_h is defined over L (see [Kot84a, Lemma 1.1.3]).

If the Shimura datum (G, h) is PEL of type A or C, so that $\mathcal{M}_K$ has a modular description as in Definition 2.9, then the triples $(\gamma_0, \gamma, \delta)$ should parametrize the $\mathbb{Q}$-isogeny classes of triples (A, λ, ι) as in the moduli problem. This parametrization rests on Honda–Tate theory, which classifies abelian varieties A over finite fields using their Frobenius, seen as a central element of $\mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$. See [Kot90, Part III] for the case of Siegel modular varieties, and [Kot92b] for the PEL cases of type A and C.

The Langlands–Rapoport conjecture is not known for general Shimura varieties, because we do not even have integral models for general Shimura varieties. All the proofs that we know about ultimately rest on an interpretation of the Shimura variety as a moduli space of abelian varieties with extra structures, so we need to have such an interpretation. Here are some known results about it:

- If (G, h) is of PEL type and G^{der} is simply connected (i.e. G is of type A or C), then Kottwitz reformulated the Langlands–Rapoport conjecture in [Kot90] and proved this reformulation in [Kot92b];
- For Shimura varieties of abelian type, Kisin proved the Langlands–Rapoport conjecture in [Kis17].

The upshot is that, if $g \in G(\mathbb{A}_f)$ has a trivial component at the prime number p under $\wp$ and $f^\infty = \mathbb{1}_{KgK} \in \mathcal{H}_K$, we get a formula for the trace of $\mathrm{Frob}_\wp^r \cdot f^\infty$ on the ℓ-adic cohomology of $M_K(G, h)$ involving terms such as orbital integrals for f^∞ (i.e. integrals of f over $G(\mathbb{A}_f)$-conjugacy orbits of elements of $G(\mathbb{Q})$) and twisted orbital integrals of a function (depending on r) at p; a priori we only get this for r big enough, but then the identity automatically extends to all non-negative r. This is the kind of input that we can plug into the geometric side of the Arthur–Selberg trace formula, see the next subsubsection. The spectral side of the trace formula will then give us an expression that can be massage into what we want, i.e., in cases when the simplest form of the Kottwitz conjecture applies, the trace of $\mathrm{Frob}_\wp^r \cdot f^\infty$ on the virtual representation

$$\sum_{i \geqslant 0} (-1)^i \sum_{\pi \in \Pi_G} a(\pi_f)\pi_f^K \otimes (r_\mu \circ \sigma_\pi)$$

of $\mathcal{H}_K \times \mathrm{Gal}(\overline{F}/F)$. So we win.

Remark 3.17. (1) The sentence "This is the kind of input that we can plug into the geometric side of the Arthur–Selberg trace formula" in the previous paragraph is sweeping a lot of difficulties under the rug. If we are looking at cases where G has no endoscopy (such as the simple Shimura varieties of Kottwitz), then the traces given by the Langlands–Rapoport conjecture are not too far from the geometric side of the trace formula (see [Kot92a, §4]). In general, we must first perform a complicated process known as "stabilization", which uses difficult results such as the fundamental lemma (not known thanks to work of Laumon–Ngo, Ngo, and Waldspurger). See Kottwitz's paper [Kot90, §4] for an explanation of stabilization in the simpler case when G^{der} is simply connected, and the book [KSZ21] of Kisin–Shin–Zhu for the case of Shimura data of abelian type without this simplifying hypothesis.

(2) Note that we ever used Matsushima's formula in our outline of the proof of the Kottwitz conjecture. In fact, though it serves as a guide, Matsushima's formula is not logically necessary to the proof.

3.4.4. *The Arthur–Selberg trace formula.* We consider the situation of Subsubsection 3.1.1, so G is a connected reductive group over $\mathbb{Q}$, S_G is the maximal $\mathbb{Q}$-split torus in the center of G and $A_G = S_G(\mathbb{R})^0$. We write L^2_G for $L^2(G(\mathbb{Q})\backslash G(\mathbb{A})/A_G, \mathbb{C})$, with the action of $G(\mathbb{A})$ given by right translations.

We also assume that G^{der} is of $\mathbb{Q}$-rank 0, so that the quotient $G(\mathbb{Q})\backslash G(\mathbb{A})/A_G$ is compact. Then the representation L^2_G is semi-simple, and we have

$$L^2_G \simeq \widehat{\bigoplus_{\pi\in\Pi_G}} \pi^{m(\pi)}. \tag{3.2}$$

Fix a Haar measure on $G(\mathbb{A})$. If f is a smooth function with compact support on $G(\mathbb{A})$, then we write $R(f)$ for the action of f on L^2_G by right convolution. Thanks to our hypothesis on G, the operator $R(f)$ is of trace class, and the goal of Arthur–Selberg trace formula is to give two expressions of its trace.

Theorem 3.18 (Arthur, see [Art05, §1])**.** *We have*

$$\begin{aligned}\mathrm{Tr}(R(f)) &= \sum_{\pi\in\Pi_G} m(\pi)\,\mathrm{Tr}(\pi(f))\\ &= \sum_{\gamma\in G(\mathbb{Q})/\sim} \mathrm{vol}(G_\gamma(\mathbb{Q})\backslash G_\gamma(\mathbb{A})/A_G)\int_{G_\gamma(\mathbb{A})\backslash G(\mathbb{A})} f(x^{-1}\gamma x)dx,\end{aligned}$$

where, in the first formula, $\pi(f)$ is the operator $\int_G f(x)\pi(x)dx$ acting on the space of π and, in the second formula, the sum is over all elements of $G(\mathbb{Q})$ modulo conjugation and, if $\gamma\in G(\mathbb{Q})$, we denote by G_γ the centralizer of γ in G.

Note that we need to choose Haar measures on the groups $G_\gamma(\mathbb{A})$ to make sense of the second formula for $R(f)$ (we also use the counting measure on $G_\gamma(\mathbb{Q})$), but that the result does not depend on that choice.

- The first formula for $R(f)$ is called the **spectral side**. It follows from the isomorphism of (3.2).
- The second formula for $R(f)$ is called the **geometric side**. We can deduce it by noting that $R(f)$ is an integral operator with kernel

  $$K(x,y) = \sum_{\gamma\in G(\mathbb{Q})} f(x^{-1}\gamma y),$$

  so the trace of $R(f)$ must be equal to $\int_G K(x,x)dx$. See [Art05, §1] for more details.

Remark 3.19. Of course, the fun really starts when the quotient $G(\mathbb{Q})\backslash G(\mathbb{A})/A_G$ is not compact. To learn more about the various version of the trace formula in that case, see the rest of Arthur's introductory notes [Art05].

References

[Art69] M. Artin. Algebraization of formal moduli (I). *Global Analysis*, pages 21–71, 1969. Papers in Honor of K. Kodaira.

[Art89] James Arthur. The L^2-Lefschetz numbers of Hecke operators. *Invent Math.*, 97(2):257–290, 1989.

[Art05] James Arthur. An introduction to the trace formula. In *Harmonic analysis, the trace formula, and Shimura varieties*, volume 4, pages 1–263. Clay Math Proc., 2005.

[BB66] L. Baily and A. Borel. Compactification of arithmetic quotients of bounded symmetric domains. *Ann. of Math*, 84(2):442–528, 1966.

[BC83] Armand Borel and William Casselman. L^2-cohomology of locally symmetric manifolds of finite volume. *Duke Math. J.*, 50(3):625–647, 1983.

[Bor79] A. Borel. *Automorphic L-functions, Automorphic forms, representations and L-functions.* Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1979. Part 2, pp. 27–61.

[Bor14] Mikhail Borovoi. Galois cohomology of reductive algebraic groups over the field of real numbers. 2014.

[Box15] George Andrew Boxer. *Torsion in the Coherent Cohomology of Shimura Varieties and Galois Representations.* PhD thesis, Harvard University, 2015. ProQuest LLC, Ann Arbor, MI.

[Bry83] Jean-Luc Brylinski. 1-motifs et formes automorphes (théorie arithmétique des domaines de Siegel). In *Conference on automorphic theory (Dijon 1981), vol. 15*, pages 43–106, Paris, 1983. Publ. Math. Univ. Paris VII.

[BW00] A. Borel and N. Wallach. Continuous cohomology, discrete subgroups, and representations of reductive groups. *Mathematical Surveys and Monographs*, 67, 2000. second ed.

[CH13] Gaëtan Chenevier and Michael Harris. Construction of automorphic Galois representations, II. *Camb. J. Math*, (1):53–73, 2013.

[Clo90] Laurent Clozel. Motifs formes automorphes: applications du principe de fonctorialité. In *Automorphic forms, Shimura varieties, and L-functions Vol. I (Ann Arbor, MI, 1988)*, pages 77–159. Academic Press, Boston, MA, 1990. Perspect. Math., 10.

[Clo91] Laurent Clozel. Représentations galoisiennes associées aux représentations automorphes autoduales de $\mathrm{GL}(n)$. *Inst. Hautes Études Sci Publ. Math.*, (73):97–145, 1991.

[Del71] Pierre Deligne. Travaux de shimura. *Lecture Notes in Math.*, 244(389):123–165, 1971. Séminaire Bourbaki, 23ème année (1970/71). Springer, Berlin.

[Del77] P. Deligne. Rapport sur la formule des traces, cohomologie étale. *Lecture Notes in Math.*, 569:76–109, 1977. Springer, Berlin.

[Del79] Pierre Deligne. Interprétation modulaire, et techniques de construction de modèles canoniques, automorphic forms, representations and L-functions. In *Variétés de Shimura*, pages 247–289, Oregon State Univ., Corvallis, 1979. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, R.I., 1979.

[DMOS82] Pierre Deligne, James S. Milne, Arthur Ogus, and Kuang-yen Shih. Hodge cycles, motives, and Shimura varieties. *Lecture Notes in Mathematics*, 900, 1982. Springer-Verlag, Berlin-New York.

[FC90] Gerd Faltings and Ching-Li Chai. Degeneration of abelian varieties. *Ergebnisse der Mathematik und ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)], 22, 1990. Springer-Verlag, Berlin. With an appendix by David Mumford.

[Fuj97] Kazuhiro Fujiwara. Rigid geometry, Lefschetz–Verdier trace formula and Deligne's conjecture. *Invent Math.*, 127(3):489–533, 1997.

[HLTT16] Michael Harris, Kai-Wen Lan, Richard Taylor, and Jack Thorne. On the rigid cohomology of certain Shimura varieties. *Res. Math. Sci*, 3(37):308, 2016.

[Ji06] Lizhen Ji. Lectures on locally symmetric spaces and arithmetic groups, Lie groups and automorphic forms. *AMS/IP Stud. Adv. Math*, 37:87–146, 2006. Amer. Math. Soc., Providence, RI.

[Kis10] Mark Kisin. Integral models for Shimura varieties of abelian type. *J. Amer. Math*, 23(4):967–1012, 2010.

[Kis17] Mark Kisin. Mod p points on Shimura varieties of abelian type. *J. Amer. Math*, 30(3):819–914, 2017.

[Kne65] Martin Kneser. Galois-kohomologie halbeinfacher algebraischer Gruppen über p-adischen Körpern (I). *Math. Z.*, 88:40–47, 1965.

[Kot84a] Robert E. Kottwitz. Shimura varieties and twisted orbital integrals. *Math. Ann.*, 269(3):287–300, 1984.

[Kot84b] Robert E. Kottwitz. Stable trace formula: cuspidal tempered terms. *Duke Math. J.*, 51(3):611–650, 1984.

[Kot86] Robert E. Kottwitz. Stable trace formula: elliptic singular terms. *Math. Ann.*, 275(3):365–399, 1986.

[Kot90] Robert E. Kottwitz. Shimura varieties and λ-adic representations. In *Automorphic forms, Shimura varieties, and L-functions*, pages 161–209. Academic Press, Perspect. Math., vol. 10, Boston, MA, 1990. Vol. I (Ann Arbor, MI, 1988).

[Kot92a] Robert E. Kottwitz. On the λ-adic representations associated to some simple Shimura varieties. *Invent Math.*, 108(3):653–665, 1992.

[Kot92b] Robert E. Kottwitz. Points on some Shimura varieties over finite fields. *J. Amer. Math Soc.*, 5(2):373–444, 1992.

[KSZ21] Mark Kisin, Sug Woo Shin, and Yihang Zhu. The stable trace formula for Shimura varieties of abelian type. 2021.

[Lan12] Kai-Wen Lan. Comparison between analytic and algebraic constructions of toroidal compactifications of PEL-type Shimura varieties. *J. Reine Angew. Math*, 664:163–228, 2012.

[Lan13] Kai-Wen Lan. *Arithmetic compactifications of PEL-type Shimura varieties*, volume 36 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2013.

[Lan20] Kai-Wen Lan. *An example-based introduction to Shimura varieties*. 2020. To appear in the proceedings of the ethz summer school on motives and complex multiplication. Available at https://www-users.cse.umn.edu/kwlan/articles/intro-sh-ex.pdf.

[Lau05] Gérard Laumon. Fonctions zêtas des variétés de Siegel de dimension trois. *Formes automorphes,* II. *Le cas du groupe* GSp(4), (302):1–66, 2005.

[Mat67] Yozô Matsushima. A formula for the Betti numbers of compact locally symmetric Riemannian manifolds. *J. Differential Geometry 1*, pages 99–109, 1967.

[Mil83] James S. Milne. The action of an automorphism of $\mathbf{C}$ on a Shimura variety and its special points. *Arithmetic and geometry, Progr. Math*, I(35):239–265, 1983. vol. Birkhäuser Boston, Boston, MA.

[Mil90] James S. Milne. Canonical models of (mixed) Shimura varieties and automorphic vector bundles. In *Automorphic forms, Shimura varieties, and L-functions*, pages 283–414. Academic Press, Perspect. Math., vol. 10, Boston, MA, 1990. Vol. I (Ann Arbor, MI, 1988).

[Mil92] James S. Milne. The points on a Shimura variety modulo a prime of good reduction. In *The zeta functions of Picard modular surfaces*, pages 151–253. Univ. Montréal, Montreal, QC, 1992.

[Mil94] James S. Milne. Shimura varieties and motives. In *Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math., vol. 55*, pages 447–523, Providence, RI, 1994. Amer. Math. Soc.

[Mil05] James S. Milne. Introduction to Shimura varieties. In *Harmonic analysis, the trace formula, and Shimura varieties, vol. 4*, pages 265–378. Clay Math. Proc., Amer. Math. Soc., Providence, RI, 2005.

[Moo98] Ben Moonen. Models of Shimura varieties in mixed characteristics. In London Math. Soc., editor, *Galois representations in arithmetic algebraic geometry (Durham, 1996)*, pages 267–350. Cambridge Univ. Press, Cambridge, 1998. Lecture Note Ser., vol. 254.

[MT62] G. D. Mostow and T. Tamagawa. On the compactness of arithmetically defined homogeneous spaces. *Ann. of Math*, 76(2):446–463, 1962.

[Mum08] David Mumford. *Abelian varieties*, volume 5 of *Tata Institute of Fundamental Research Studies in Mathematics*. Published for the Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition.

[NS16] J. Nekovář and A. J. Scholl. Introduction to plectic cohomology. *Advances in the theory of automorphic forms and their L-functions, Contemp. Math.*, 664:321–337, 2016. Amer. Math. Soc., Providence, RI.

[Pin92] Richard Pink. On the calculation of local terms in the Lefschetz–Verdier trace formula and its application to a conjecture of Deligne. *Ann. of Math.*, 135(3):483–525, 2 1992.

[Pla69] V. P. Platonov. The problem of strong approximation and the Kneser–Tits hypothesis for algebraic groups. *Izv. Akad.*, 33:1211–1219, 1969.

[Pla70] V. P. Platonov. A supplement to the paper "the problem of strong approximation and the Kneser–Tits hypothesis for algebraic groups". *Izv. Akad.*, 34:775–777, 1970.

[Sch15] Peter Scholze. On torsion in the cohomology of locally symmetric varieties. *Ann. of Math.*, 182(3):945–1066, 2 2015.

[Shi11] Sug Woo Shin. Galois representations arising from some compact Shimura varieties. *Ann. of Math.*, 173(3):1645–1741, 2 2011.

[Sie43] Carl Ludwig Siegel. Symplectic geometry. *Amer. J. Math.*, 65:1–86, 1943.

[SS90] Leslie Saper and Mark Stern. L^2-cohomology of arithmetic varieties. *Ann. of Math.*, 132(1):1–69, 2 1990.

[SS13] Peter Scholze and Sug Woo Shin. On the cohomology of compact unitary group Shimura varieties at ramified split places. *J. Amer. Math. Soc.*, 26(1):261–294, 2013.

[Sta] The Stacks Project. 2022. https://stacks.math.columbia.edu.

[Var07] Yakov Varshavsky. Lefschetz–Verdier trace formula and a generalization of a theorem of Fujiwara. *Geom. Funct. Anal.*, 17(1):271–319, 2007.

School of Mathematical Sciences, Peking University, 100871, Beijing, China
Email address: daiwenhan@pku.edu.cn