AN EXAMPLE-BASED INTRODUCTION TO SHIMURA VARIETIES

KAI-WEN LAN

Abstract. In this introductory article, we will explain what Shimura varieties are and why they are useful, with an emphasis on examples. We will also describe various compactifications of Shimura varieties, again with an emphasis on examples, and summarize what we know about the integral models of Shimura varieties and their compactifications.

Contents

2010 *Mathematics Subject Classification.* Primary 14G35; Secondary 11G18.

Key words and phrases. locally symmetric varieties, Shimura varieties, examples.

The author was partially supported by the National Science Foundation under agreement No. DMS-1352216, and by an Alfred P. Sloan Research Fellowship. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the author and do not necessarily reflect the views of these organizations.

This article will be published in the proceedings of the ETHZ Summer School on Motives and Complex Multiplication.

1. Introduction

Shimura varieties are generalizations of modular curves, which have played an important role in many recent developments of number theory. Just to mention a few examples with which this author is more familiar, Shimura varieties were crucially used in the proof of the *local Langlands conjecture* for GL_n by Harris and Taylor (see [HT01]); in the proof of the *Iwasawa main conjecture* for GL_2 by Skinner and Urban (see [SU14]); in the construction of *p-adic L-functions* for unitary groups by Eischen, Harris, Li, and Skinner (see [HLS06] and [EHLS20]); in the construction of Galois representations (in the context of the global Langlands correspondence) for *all* cohomological automorphic representations of GL_n by Harris, Taylor, Thorne, and this author (see [HLTT16]); and in the analogous but deeper construction for *torsion* cohomological classes of GL_n by Scholze (see [Sch15]).

One of the reasons that Shimura varieties are so useful is because they carry two important kinds of symmetries—the Hecke symmetry and the Galois symmetry. Roughly speaking, the Hecke symmetry is useful for studying automorphic representations, the Galois symmetry is useful for studying Galois representations, and the compatibility between the two kinds of symmetries is useful for studying the relations between the two kinds of representations, especially in the context of the Langlands program.

However, the theory of Shimura varieties does not have a reputation of being easy to learn. There are important foundational topics such as the notion of Shimura data and the existence of canonical models, and courses or seminars on Shimura varieties often spend a substantial amount of the time on such topics. But learning these topics might shed little light on how Shimura varieties are actually used—in some of the important applications mentioned above, the "Shimura varieties" there were not even defined using any Shimura data or canonical models. In reality, it is still a daunting task to introduce people to this rich and multifaceted subject.

In this introductory article, we will experiment with a somewhat different approach. We still urge the readers to read the excellent texts such as [Del71b], [Del79], and [Mil05]. But we will not try to repeat most of the standard materials, such as variation of Hodge structures, the theory of canonical models, etc, which have already been explained quite well by many authors. We simply assume that people will learn them from some other sources. Instead, we will try to explain the scope of the theory by presenting many examples, and many special facts which are only true for these examples. Our experience is that such materials are rather scattered in the literature, but can be very helpful for digesting the abstract concepts, and for learning how to apply the general results in specific circumstances. We hope that they will benefit the readers in a similar way.

We will keep the technical level as low as possible, and present mainly simple-minded arguments that should be understandable by readers who are willing to see matrices larger than 2×2 ones. This is partly because this article is based on our notes for several introductory lectures targeting audiences of very varied backgrounds. The readers will still see a large number of technical phrases, often without detailed explanations, but we believe that such phrases can be kept as keywords in mind and learned later. We will provide references when necessary, but such references do not have to be consulted immediately—doing that will likely disrupt the flow of reading and obscure the main points. (It is probably unrealistic and unhelpful to try to consult all of the references, especially the historical ones.) We hope that the readers will find these consistent with our goal of explaining the scope of the theory. There will inevitably be some undesirable omissions even with our limited goal, but we hope that our overall coverage will still be helpful.

Here is an outline of the article. In Section 2, we define certain double coset spaces associated with pairs of algebraic groups and manifolds, and introduce the notion of Shimura data. We also summarize the main results concerning the quasi-projectivity of Shimura varieties and the existence of their canonical models. In Section 3, we give many examples of Hermitian symmetric domains, which provide the connected components of the manifolds needed in the definition of Shimura data. Our examples actually exhaust all possibilities of them. In Section 4, we give many examples of rational boundary components and describe the corresponding minimal compactifications. We also summarize the properties of several kinds of useful compactifications. In Section 5, we explain how the so-called *integral models* of Shimura varieties are constructed. This is very sensitive to the types of Shimura varieties we consider, and we also take this opportunity to summarize many special facts and many recent developments for each type of them. For the convenience of the readers, we have also included an index at the end of the article.

2. Double coset spaces

2.1. **Algebraic groups.** Let $\widehat{\mathbb{Z}} := \varprojlim_N (\mathbb{Z}/N\mathbb{Z})$, $\mathbb{A}^\infty := \widehat{\mathbb{Z}} \underset{\mathbb{Z}}{\otimes} \mathbb{Q}$, and $\mathbb{A} := \mathbb{R} \times \mathbb{A}^\infty$. Suppose G is a *reductive algebraic group* over $\mathbb{Q}$. Then there is a natural way to define the topological groups $\mathrm{G}(\mathbb{R})$, $\mathrm{G}(\mathbb{A}^\infty)$, and $\mathrm{G}(\mathbb{A}) = \mathrm{G}(\mathbb{R}) \times \mathrm{G}(\mathbb{A}^\infty)$, with the topology given either as a restricted product (following Weil) (see [PR94, Sec. 5.1]), or as an affine scheme (of finite presentation) over the topological ring $\mathbb{A}$ (following Grothendieck), and the two approaches coincide in this case (see [Con12] for a detailed discussion on this matter). We will not try to define and summarize all the needed properties of reductive linear algebraic groups (from textbooks such as [Bor91] and [Spr98]). Instead, we will just provide many examples of such groups.

Example 2.1.1 (general and special linear groups). For each ring R and each integer $n \geq 0$, we define $\mathrm{GL}_n(R)$ to be the group of invertible $n \times n$ matrices with entries in R, and define $\mathrm{SL}_n(R)$ to be the subgroup of $\mathrm{GL}_n(R)$ formed by matrices of determinant one. When $n = 1$, we have $\mathrm{GL}_1(R) = \mathbf{G}_\mathrm{m}(R) := R^\times$. The assignments of $\mathrm{GL}_n(R)$ and $\mathrm{SL}_n(R)$ are *functorial* in R, and in fact they are the R-points of the *affine group schemes* GL_n and SL_n over $\mathbb{Z}$, respectively. The pullbacks of these group schemes to $\mathbb{Q}$ are affine algebraic varieties serving as prototypical examples of reductive algebraic groups. Moreover, the pullback of SL_n to $\mathbb{Q}$ is a prototypical example of a *semisimple* algebraic group.

For simplicity of exposition, we will often introduce a linear algebraic group G by describing it as the pullback to $\mathbb{Q}$ of some explicitly given *closed subgroup scheme* of GL_m, for some m, over $\mathbb{Z}$. By abuse of notation, we will often denote this group scheme over $\mathbb{Z}$ by the same symbol G, which is then equipped with a fixed choice of a faithful matrix representation $\mathrm{G} \hookrightarrow \mathrm{GL}_m$. The readers should not be too worried about these technical terminologies concerning group schemes over $\mathbb{Z}$. Concretely, they just mean we will explicitly represent the elements of $\mathrm{G}(R)$ as invertible $m \times m$ matrices, for each ring R, whose entries satisfy certain defining conditions given by algebraic equations that are compatible with all base ring extensions $R \to R'$.

Example 2.1.2 (symplectic groups). Let $n \geq 0$ be any integer, and let 0_n and 1_n denote the zero and identity matrices of size n. Consider the *skew-symmetric* matrix

$$J_n := \begin{pmatrix} 0_n & 1_n \\ -1_n & 0_n \end{pmatrix} \tag{2.1.3}$$

of size $2n$. Then the assignments

$$\mathrm{Sp}_{2n}(R) := \{g \in \mathrm{GL}_{2n}(R) : {}^t g J_n g = J_n\} \tag{2.1.4}$$

and

$$\mathrm{GSp}_{2n}(R) := \{(g, r) \in \mathrm{GL}_{2n}(R) \times R^\times = (\mathrm{GL}_{2n} \times \mathbf{G}_\mathrm{m})(R) : {}^t g J_n g = r J_n\} \tag{2.1.5}$$

are functorial in R, and define closed subgroup schemes Sp_{2n} and GSp_{2n} of GL_{2n} and GL_{2n+1}, respectively. Here we view $\mathrm{GL}_{2n} \times \mathbf{G}_\mathrm{m}$ as a closed subgroup scheme of GL_{2n+1} by block-diagonally embedding the matrix entries. To better understand the meaning of these definitions, consider the alternating pairing

$$\langle \cdot \, , \cdot \rangle : R^{2n} \times R^{2n} \to R \tag{2.1.6}$$

defined by setting

$$\langle x, y \rangle = {}^t x J_n y, \tag{2.1.7}$$

for all $x, y \in R^{2n}$ (written as vertical vectors). For each $g \in \mathrm{GL}_{2n}(R)$, we have

$$\langle gx, gy \rangle = {}^t x \, {}^t g J_n g y. \tag{2.1.8}$$

Therefore, g satisfies the condition ${}^t g J_n g = J_n$ exactly when

$$\langle gx, gy \rangle = \langle x, y \rangle, \tag{2.1.9}$$

for all $x, y \in R^{2n}$. That is, g preserves the above pairing $\langle \cdot \, , \cdot \rangle$. Note that we have not used exactly what J_n is, except for its skew-symmetry. Similarly, $(g, r) \in \mathrm{GL}_{2n}(R) \times R^\times$ satisfies the condition ${}^t g J_n g = r J_n$ exactly when

$$\langle gx, gy \rangle = r \langle x, y \rangle, \tag{2.1.10}$$

for all $x, y \in R^{2n}$. That is, g preserves the above pairing $\langle \cdot \, , \cdot \rangle$ up to the scalar factor r. The assignment of r to (g, r) defines a homomorphism

$$\nu : \mathrm{GSp}_{2n}(R) \to \mathbf{G}_\mathrm{m}(R), \tag{2.1.11}$$

called the *similitude character* of $\mathrm{GSp}_{2n}(R)$, whose kernel is exactly $\mathrm{Sp}_{2n}(R)$, for each ring R. We can define similar groups and homomorphisms by replacing J_n with any other matrices (with entries over $\mathbb{Z}$, if we still want to define a group scheme over $\mathbb{Z}$). In fact, many people would prefer to define Sp_{2n} and GSp_{2n} using anti-diagonal matrices (just not a "block anti-diagonal matrix" like the above $J_n = \begin{pmatrix} 0_n & 1_n \\ -1_n & 0_n \end{pmatrix}$).

Example 2.1.12 (orthogonal groups and special orthogonal groups). Consider any integers $n, a, b \geq 0$, and consider the *symmetric* matrices 1_n and

$$1_{a,b} := \begin{pmatrix} 1_a & \\ & -1_b \end{pmatrix}, \tag{2.1.13}$$

with empty slots being filled up by zero matrices of suitable sizes. Then the assignments

$$\mathrm{O}_n(R) := \{g \in \mathrm{GL}_n(R) : {}^t g g = 1_n\}, \tag{2.1.14}$$

$$\mathrm{SO}_n(R) := \mathrm{O}_n(R) \cap \mathrm{SL}_n(R), \tag{2.1.15}$$

$$\mathrm{O}_{a,b}(R) := \{g \in \mathrm{GL}_n(R) : {}^t g 1_{a,b} g = 1_{a,b}\}, \tag{2.1.16}$$

$$\mathrm{SO}_{a,b}(R) := \mathrm{O}_{a,b}(R) \cap \mathrm{SL}_{a+b}(R) \tag{2.1.17}$$

are functorial in R, and define closed subgroup schemes O_n, SO_n, $\mathrm{O}_{a,b}$, and $\mathrm{SO}_{a,b}$ of GL_n, GL_n, GL_{a+b}, and GL_{a+b}, respectively. As in the case of $\mathrm{Sp}_{2n}(R)$ above, $\mathrm{O}_n(R)$ and $\mathrm{O}_{a+b}(R)$ are the group of matrices preserving the *symmetric* bilinear pairings defined by 1_n and $1_{a,b}$, which have signatures $(n, 0)$ and (a, b), respectively.

2.2. **Manifolds.** Let us fix a choice of G, and consider any manifold $\mathcal{D}$ with a smooth transitive action of $\mathrm{G}(\mathbb{R})$. For technical reasons, we shall consider only algebraic groups G over $\mathbb{Q}$ each of whose connected components in the Zariski topology contains at least one rational point. This requirement is satisfied by every connected algebraic group, because the identity element is always rational, and also by all the examples in Section 2.1. Then $\mathrm{G}(\mathbb{Q})$ is dense in $\mathrm{G}(\mathbb{R})$ in the real analytic topology, by applying a special case of the weak approximation theorem as in [PR94, Sec. 7.3, Thm. 7.7] to the identity component of G (namely, the connected component containing the identity element) in the Zariski topology, and by translations between this identity component and other connected components.

Example 2.2.1. For each integer $n \geq 0$, consider the manifold

$$\begin{aligned} \mathcal{S}_n &:= \{X \in \mathrm{Sym}_n(\mathbb{R}) : X > 0, \det(X) = 1\} \\ &= \{X \in \mathrm{M}_n(\mathbb{R}) : {}^t X = X > 0, \det(X) = 1\}, \end{aligned} \tag{2.2.2}$$

where M_n (resp. Sym_n) denotes the space of $n \times n$ matrices (resp. symmetric matrices), and where $X > 0$ means being positive definite. Consider the action of $g \in \mathrm{SL}_n(\mathbb{R})$ on $\mathcal{S}_n$ defined by

$$X \mapsto {}^t g X g, \tag{2.2.3}$$

for each $X \in \mathcal{S}_n$. This action is transitive because each positive definite matrix X is of the form

$$X = Y^2 = {}^t Y Y \tag{2.2.4}$$

for some positive definite matrix Y, and if $\det(X) = 1$ then $\det(Y) = 1$ as well. The stabilizer of $X = 1_n \in \mathcal{S}_n$ is, by definition, $\mathrm{SO}_n(\mathbb{R})$, and so we have

$$\mathcal{S}_n = \mathrm{SL}_n(\mathbb{R}) \cdot 1_n \cong \mathrm{SL}_n(\mathbb{R})/\mathrm{SO}_n(\mathbb{R}). \tag{2.2.5}$$

Note that

$$\dim_{\mathbb{R}}(\mathcal{S}_n) = \tfrac{1}{2} n(n+1) - 1. \tag{2.2.6}$$

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{S}_n$ is a noncompact Riemannian symmetric space of type A I.

Example 2.2.7. When $n = 2$, we also have the familiar example of

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}) \tag{2.2.8}$$

acting on the Poincaré upper half-plane

$$\mathcal{H} := \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\} \tag{2.2.9}$$

by the Möbius transformation

$$z \mapsto gz = \frac{az + b}{cz + d}, \tag{2.2.10}$$

for each $z \in \mathcal{H}$. Note that this is actually induced by the natural $\mathrm{SL}_2(\mathbb{C})$ action on $\mathbb{P}^1(\mathbb{C})$: If we identify $z \in \mathbb{C}$ with $\begin{pmatrix} z \\ 1 \end{pmatrix} \in \mathbb{P}^1(\mathbb{C})$, then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} az + b \\ cz + d \end{pmatrix} \sim \begin{pmatrix} gz \\ 1 \end{pmatrix}. \tag{2.2.11}$$

For readers who have studied modular forms, the factor $(cz + d)$ involved in the above identification between projective coordinates is exactly the same automorphy factor $(cz + d)$ in the definition of holomorphic modular forms. (This is not just a coincidence.) It is well known that this action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathcal{H}$ is transitive (which will be generalized in Section 3.1.2 below), and the stabilizer of $i \in \mathcal{H}$ is

$$\left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}) \right\} = \mathrm{SO}_2(\mathbb{R}) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\}, \tag{2.2.12}$$

so that we have

$$\mathcal{H} = \mathrm{SL}_2(\mathbb{R}) \cdot i \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}). \tag{2.2.13}$$

Example 2.2.14. The isomorphisms $\mathcal{H} \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cong \mathcal{S}_2$ show that $\mathcal{H}$ can be viewed as a special case of $\mathcal{S}_n$ with $n = 2$. But this is the only example of $\mathcal{S}_n$, with $n \geq 1$, such that $\mathcal{S}_n$ has a complex structure. Since $\dim_{\mathbb{R}}(\mathcal{S}_n) = \frac{1}{2}n(n+1) - 1$ is odd when $n = 3$, for example, there is no hope for $\mathcal{S}_3$ to be a complex manifold.

For any $(\mathrm{G}, \mathcal{D})$ as above, and for any open compact subgroup $\mathcal{U}$ of $\mathrm{G}(\mathbb{A}^\infty)$, we can define the double coset space

$$\mathsf{X}_{\mathcal{U}} := \mathrm{G}(\mathbb{Q}) \backslash (\mathcal{D} \times \mathrm{G}(\mathbb{A}^\infty))/\mathcal{U}, \tag{2.2.15}$$

where $\mathrm{G}(\mathbb{Q})$ acts diagonally on $\mathcal{D} \times \mathrm{G}(\mathbb{A}^\infty)$ from the left-hand side, and where $\mathcal{U}$ only acts on $\mathrm{G}(\mathbb{A}^\infty)$ from the right-hand side. The reader might naturally wonder why we need to consider a complicated double quotient as in (2.2.15). One important justification is that the group $\mathrm{G}(\mathbb{A}^\infty)$ has a natural right action on the collection $\{\mathsf{X}_{\mathcal{U}}\}_{\mathcal{U}}$, induced by

$$\mathcal{D} \times \mathrm{G}(\mathbb{A}^\infty) \xrightarrow{\sim} \mathcal{D} \times \mathrm{G}(\mathbb{A}^\infty) : (x, h) \mapsto (x, hg), \tag{2.2.16}$$

for each $g \in \mathrm{G}(\mathbb{A}^\infty)$, which maps $\mathsf{X}_{g\mathcal{U}g^{-1}}$ to $\mathsf{X}_{\mathcal{U}}$ because $h(gug^{-1})g = hgu$, for all $h \in \mathrm{G}(\mathbb{A}^\infty)$ and $u \in \mathcal{U}$. Such an action provides natural *Hecke actions* on, for example, the limit of cohomology groups $\varinjlim_{\mathcal{U}} H^*(\mathsf{X}_{\mathcal{U}}, \mathbb{C})$. Such a symmetry of $\{\mathsf{X}_{\mathcal{U}}\}_{\mathcal{U}}$ is what we meant by *Hecke symmetry* in the introduction (see Section 1). This is crucial for relating the geometry of such double coset spaces to the theory of automorphic representations.

Let $\mathcal{D}^+$ be a connected component of $\mathcal{D}$, which admits a transitive action of $\mathrm{G}(\mathbb{R})^+$, the identity component (namely, the connected component containing the identity element) of $\mathrm{G}(\mathbb{R})$ in the real analytic topology. Let $\mathrm{G}(\mathbb{R})_+$ denote the stabilizer of $\mathcal{D}^+$ in $\mathrm{G}(\mathbb{R})$, which necessarily contains $\mathrm{G}(\mathbb{R})^+$. Let

$$\mathrm{G}(\mathbb{Q})_+ := \mathrm{G}(\mathbb{Q}) \cap \mathrm{G}(\mathbb{R})_+, \tag{2.2.17}$$

which is the subgroup of $\mathrm{G}(\mathbb{Q})$ stabilizing $\mathcal{D}^+$. Then

$$\mathrm{G}(\mathbb{Q})_+ \backslash (\mathcal{D}^+ \times \mathrm{G}(\mathbb{A}^\infty))/\mathcal{U} \to \mathrm{G}(\mathbb{Q}) \backslash (\mathcal{D} \times \mathrm{G}(\mathbb{A}^\infty))/\mathcal{U} \tag{2.2.18}$$

is surjective because $\mathrm{G}(\mathbb{Q})$ is dense in $\mathrm{G}(\mathbb{R})$ (see the beginning of this Section 2.2), and is injective by definition. It is known (see [Bor63, Thm. 5.1]) that

$$\#(\mathrm{G}(\mathbb{Q})_+ \backslash \mathrm{G}(\mathbb{A}^\infty)/\mathcal{U}) < \infty, \tag{2.2.19}$$

which means there exists a subset $\{g_i\}_{i \in I}$ of $\mathrm{G}(\mathbb{A}^\infty)$ indexed by a finite set I such that we have a disjoint union

$$\mathrm{G}(\mathbb{A}^\infty) = \coprod_{i \in I} \mathrm{G}(\mathbb{Q})_+ \, g_i \, \mathcal{U}. \tag{2.2.20}$$

Then

$$\begin{aligned} \mathsf{X}_{\mathcal{U}} &\cong \mathrm{G}(\mathbb{Q})_+ \backslash (\mathcal{D}^+ \times \mathrm{G}(\mathbb{A}^\infty))/\mathcal{U} \\ &= \coprod_{i \in I} \mathrm{G}(\mathbb{Q})_+ \backslash (\mathcal{D}^+ \times \mathrm{G}(\mathbb{Q})_+ \, g_i \, \mathcal{U})/\mathcal{U} \\ &= \coprod_{i \in I} \Gamma_i \backslash \mathcal{D}^+, \end{aligned} \tag{2.2.21}$$

where

$$\Gamma_i := \mathrm{G}(\mathbb{Q})_+ \cap (g_i \mathcal{U} g_i^{-1}), \tag{2.2.22}$$

because $\gamma g_i u = g_i u$ exactly when $\gamma \in g_i \mathcal{U} g_i^{-1}$, for any $\gamma \in \mathrm{G}(\mathbb{Q})_+$, $i \in I$, and $u \in \mathcal{U}$.

Each such Γ_i is an *arithmetic subgroup* of $\mathrm{G}(\mathbb{Q})$; namely, a subgroup *commensurable* with $\mathrm{G}(\mathbb{Z})$ in the sense that $\Gamma_i \cap \mathrm{G}(\mathbb{Z})$ has finite indices in both Γ_i and $\mathrm{G}(\mathbb{Z})$. In fact, since $\mathcal{U}$ is an open compact subgroup of $\mathrm{G}(\mathbb{A}^\infty)$, each such Γ_i is a *congruence subgroup* of $\mathrm{G}(\mathbb{Z})$; namely, a subgroup containing the *principal congruence subgroup* $\ker(\mathrm{G}(\mathbb{Z}) \to \mathrm{G}(\mathbb{Z}/N\mathbb{Z}))$ for some integer $N \geq 1$. By definition, congruences subgroups are arithmetic subgroups. Note that, although the definition of principal congruence subgroups depends on the choice of some faithful matrix representation $\mathrm{G} \hookrightarrow \mathrm{GL}_m$ over $\mathbb{Z}$ (see Section 2.1), the definition of congruence subgroups does not, because congruence subgroups of $\mathrm{G}(\mathbb{Q})$ can be characterized alternatively as the intersections of $\mathrm{G}(\mathbb{Q})$ with open compact subgroups of $\mathrm{G}(\mathbb{A}^\infty)$. More generally, the definition of arithmetic subgroups of $\mathrm{G}(\mathbb{Q})$ does not depend on the choice of the faithful matrix representation either (see [Bor69, 7.13] or [PR94, Sec. 4.1]).

If each Γ_i acts *freely* on $\mathcal{D}^+$, then $\mathsf{X}_{\mathcal{U}}$ is a manifold because $\mathcal{D}^+$ is. This is the case when each Γ_i is a *neat* arithmetic subgroup of $\mathrm{G}(\mathbb{Q})$ (as in [Bor69, (17.1)]), which is in turn the case when $\mathcal{U}$ is a *neat* open compact subgroup of G (as in [Pin89, 0.6]). For reading the remainder of this article, it suffices to know that, when people say such groups are neat (or, much less precisely, *sufficiently small*), they just want to ensure that the stabilizers of geometric actions are all trivial. For simplicity, we shall tacitly assume that the arithmetic subgroups or open compact subgroups we will encounter are all neat, unless otherwise stated.

Now the question is what each $\Gamma_i\backslash\mathcal{D}^+$ is, after knowing that it is a manifold because Γ_i is neat. What additional structures can we expect from such quotients?

Example 2.2.23. When $\mathrm{G} = \mathrm{SL}_2$ and $\mathcal{D} = \mathcal{H}$ as in Example 2.2.7, the groups Γ_i are congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$, each of which contains the principal congruence subgroup

$$\Gamma(N) := \ker(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})) \tag{2.2.24}$$

for some integer $N \geq 1$. The quotients $\Gamma_i\backslash\mathcal{H}$ are familiar objects called *modular curves*, each of which admits a good compactification into a *compact Riemann surface* (or, in other words, *complete algebraic curves* over $\mathbb{C}$) by adding a finite number of points, which are called *cusps*. (We will revisit this example in more detail in Section 4.1.1 below.)

Remark 2.2.25. We may replace the manifold $\mathcal{D}$ above with other geometric objects over $\mathcal{D}$ that still have a smooth (but not necessarily transitive) action of $\mathrm{G}(\mathbb{R})$. For example, when $\mathrm{G} = \mathrm{SL}_2$ and $\mathcal{D} = \mathcal{H}$, we may consider the space $\mathcal{H} \times \mathrm{Sym}^k(\mathbb{C}^2)$, and obtain natural vector bundles or local systems over $\mathsf{X}_{\mathcal{U}}$.

2.3. **Shimura data.** Shimura varieties are not just any double coset spaces $\mathsf{X}_{\mathcal{U}}$ as above. For arithmetic applications, it is desirable that each $\mathsf{X}_{\mathcal{U}}$ is an algebraic variety over $\mathbb{C}$ with a model over some canonically determined number field.

This led to the notion of a *Shimura datum* $(\mathrm{G}, \mathcal{D})$ (see [Del79, 2.1.1] and [Mil05, Def. 5.5]). Consider the *Deligne torus* $\mathbb{S}$, which is the *restriction of scalars* $\mathrm{Res}_{\mathbb{C}/\mathbb{R}}\,\mathbf{G}_{\mathrm{m},\mathbb{C}}$ (as an algebraic group over $\mathbb{R}$), so that $\mathbb{S}(\mathbb{R}) = \mathbb{C}^\times$. Then we require G to be a *connected reductive algebraic group*, and require $\mathcal{D}$ to be a $\mathrm{G}(\mathbb{R})$-*conjugacy class* of homomorphisms

$$h : \mathbb{S} \to \mathrm{G}_{\mathbb{R}}, \tag{2.3.1}$$

where the subscript "$\mathbb{R}$" means "base change to $\mathbb{R}$" as usual, satisfying the following conditions:

(1) The representation defined by h and the *adjoint representation* of $\mathrm{G}(\mathbb{R})$ on $\mathfrak{g} = \mathrm{Lie}\,\mathrm{G}(\mathbb{C})$ induces a decomposition

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}^+ \oplus \mathfrak{p}^- \tag{2.3.2}$$

such that

$$z \in \mathbb{S}(\mathbb{R}) = \mathbb{C}^\times \tag{2.3.3}$$

acts by 1, $z/\bar{z}$, and $\bar{z}/z$, respectively, on the three summands.

(2) $h(i)$ induces a *Cartan involution* on $\mathrm{G}^{\mathrm{ad}}(\mathbb{R})$. (Here G^{ad} denotes the *adjoint quotient* of G; namely, the quotient of G by its center.)

(3) $\mathrm{G}^{\mathrm{ad}}_{\mathbb{R}}$ has no nontrivial $\mathbb{Q}$-simple factor H such that $\mathrm{H}(\mathbb{R})$ is *compact* (or, equivalently, such that the composition of h with the projection to H is trivial, because of the previous condition (2)).

In this case, any finite dimensional $\mathbb{R}$-representation of $\mathrm{G}(\mathbb{R})$ defines a *variation of* (*real*) *Hodge structures*, and $\mathcal{D}$ is a finite union of *Hermitian symmetric domains*. Without going into details, let us just emphasize that $\mathcal{D}$ and its quotients $\mathsf{X}_{\mathcal{U}}$, for neat open compact subgroups $\mathcal{U}$ of $\mathrm{G}(\mathbb{A}^\infty)$, are then *complex manifolds*, not just real ones.

For studying only the connected components $\mathcal{D}^+$ (of some $\mathcal{D}$ which appeared in the last paragraph) and their quotients, it is easier to work with what is called a *connected Shimura datum* $(\mathrm{G}, \mathcal{D}^+)$ (see [Mil05, Def. 4.22]), which requires G to be a *connected semisimple algebraic group*, and requires $\mathcal{D}^+$ to be a $\mathrm{G}^{\mathrm{ad}}(\mathbb{R})^+$*-conjugacy class* of homomorphisms

$$h : \mathbb{S} \to \mathrm{G}^{\mathrm{ad}}_{\mathbb{R}}, \tag{2.3.4}$$

where $\mathrm{G}^{\mathrm{ad}}(\mathbb{R})^+$ denotes (as before) the identity component of $\mathrm{G}^{\mathrm{ad}}(\mathbb{R})$ in the real analytic topology, satisfying the same three conditions (1), (2), and (3) as above.

A homomorphism as in (2.3.4) determines and is determined by a homomorphism

$$\overline{h} : \mathrm{U}_1 \to \mathrm{G}^{\mathrm{ad}}_{\mathbb{R}}, \tag{2.3.5}$$

where

$$\mathrm{U}_1 := \{z \in \mathbb{C} : |z| = 1\} \tag{2.3.6}$$

is viewed as an algebraic group over $\mathbb{R}$ here, via the canonical homomorphism

$$\mathbb{S} \to \mathrm{U}_1 \tag{2.3.7}$$

induced by $z \mapsto z/\overline{z}$. In many cases we will study, the restriction of the homomorphism (2.3.4) to U_1 (now viewed as a subgroup of $\mathbb{S}$) lifts to a homomorphism

$$h : \mathrm{U}_1 \to \mathrm{G}_{\mathbb{R}}, \tag{2.3.8}$$

whose composition with $\mathrm{G}_{\mathbb{R}} \to \mathrm{G}^{\mathrm{ad}}_{\mathbb{R}}$ then provides a square-root of the $\overline{h}$ in (2.3.5). This is nonstandard, but sometimes more convenient, when we prefer to describe $\mathrm{G}_{\mathbb{R}}$ instead of $\mathrm{G}^{\mathrm{ad}}_{\mathbb{R}}$ in many practical examples (to be given in Section 3 below).

2.4. **Shimura varieties and their canonical models.** Now we are ready to state the following beautiful results, whose assertions can be understood without knowing anything about their proofs:

Theorem 2.4.1 (Satake, Baily–Borel, and Borel; see [BB66] and [Bor72])**.** *Suppose* $(\mathrm{G}, \mathcal{D})$ *is a Shimura datum as in Section* 2.3*. Then the whole collection of complex manifolds* $\{\mathsf{X}_{\mathcal{U}}\}_{\mathcal{U}}$*, with* $\mathcal{U}$ *varying among neat open compact subgroups of* $\mathrm{G}(\mathbb{A}^\infty)$*, is the complex analytification* (*see* [Ser56, §2]) *of a canonical collection of smooth* **quasi-projective varieties** *over* $\mathbb{C}$*. Moreover, the analytic covering maps*

$$\mathsf{X}_{\mathcal{U}} \to \mathsf{X}_{\mathcal{U}'}, \tag{2.4.2}$$

when $\mathcal{U} \subset \mathcal{U}'$*, are also given by the complex analytifications of canonical finite étale algebraic morphisms between the corresponding varieties.*

Alternatively, suppose $(\mathrm{G}, \mathcal{D}^+)$ *is a connected Shimura datum as in Section* 2.3*. Then analogous assertions hold for the collection* $\{\Gamma\backslash\mathcal{D}^+\}_\Gamma$*, with* Γ *varying among neat arithmetic subgroups of* $\mathrm{G}^{\mathrm{ad}}(\mathbb{Q})$ *that are contained in* $\mathrm{G}^{\mathrm{ad}}(\mathbb{R})^+$*.* (*If we consider also* $(\mathrm{G}, \mathcal{D})$*, where* $\mathcal{D}$ *is the corresponding* $\mathrm{G}(\mathbb{R})$*-conjugacy class of any* h *in* $\mathcal{D}^+$*, then the connected components of* $\mathsf{X}_{\mathcal{U}}$*, where* $\mathcal{U}$ *is a neat open compact subgroup of* $\mathrm{G}(\mathbb{A}^\infty)$*, are all of the form* $\Gamma\backslash\mathcal{D}^+$*, as explained in Section* 2.2*.*)

Theorem 2.4.3 (Shimura, Deligne, Milne, Borovoĭ, and others; see [Shi70], [Del71b], [Del79], [Mil83], [Bor84], and [Bor87])**.** *Suppose* $(\mathrm{G}, \mathcal{D})$ *is a Shimura datum as in Section* 2.3*. Then there exists a number field* F_0*, given as a subfield of* $\mathbb{C}$ *depending only on* $(\mathrm{G}, \mathcal{D})$*, called the* **reflex field** *of* $(\mathrm{G}, \mathcal{D})$*, such that the whole collection of complex manifolds* $\{\mathsf{X}_{\mathcal{U}}\}_{\mathcal{U}}$*, with* $\mathcal{U}$ *varying among neat open compact subgroups of* $\mathrm{G}(\mathbb{A}^\infty)$*, is the complex analytification of the pullback to* $\mathbb{C}$ *of a canonical collection of smooth quasi-projective varieties over* F_0*, which satisfies* *certain additional properties qualifying them as the* **canonical models** *of* $\{\mathsf{X}_{\mathcal{U}}\}_{\mathcal{U}}$. *Moreover, the analytic covering maps*

$$\mathsf{X}_{\mathcal{U}} \to \mathsf{X}_{\mathcal{U}'}, \tag{2.4.4}$$

when $\mathcal{U} \subset \mathcal{U}'$, *are also given by the complex analytifications of canonical finite étale algebraic morphisms defined over* F_0 *between the corresponding canonical models.*

Remark 2.4.5. Theorem 2.4.1 is proved by constructing the so-called *Satake–Baily–Borel* or *minimal* compactifications of $\mathsf{X}_{\mathcal{U}}$ or $\Gamma\backslash\mathcal{D}^+$, which are projective varieties over $\mathbb{C}$. (See Sections 4.2.2 and 4.2.7 below.) The assertion that the quasi-projective varieties $\mathsf{X}_{\mathcal{U}}$ and $\Gamma\backslash\mathcal{D}^+$ are defined over $\overline{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$, can be shown without using the theory of canonical models, as in [Fal84]. But these general results cannot explain why each $\mathsf{X}_{\mathcal{U}}$ is defined over some particular number field.

Remark 2.4.6. In Theorem 2.4.3, since each $\mathsf{X}_{\mathcal{U}}$ is algebraic and defined over F_0, its arithmetic invariants such as the étale cohomology group $H^*_{\text{ét}}(\mathsf{X}_{\mathcal{U}}, \mathbb{Q}_\ell)$, for any prime number ℓ, carries a canonical action of $\mathrm{Gal}(\overline{\mathbb{Q}}/F_0)$, where the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ in $\mathbb{C}$ contains F_0 because F_0 is given as a subfield of $\mathbb{C}$. (Recall that any base change between separably closed fields induces a canonical isomorphism between the étale cohomology groups—see, for example, [Del77, Arcata, V, 3.3].) This is an instance of what we meant by *Galois symmetry* in the introduction (see Section 1). Since the maps $\mathsf{X}_{\mathcal{U}} \to \mathsf{X}_{\mathcal{U}'}$ are also algebraic and defined over F_0, we have a canonical action of $\mathrm{Gal}(\overline{\mathbb{Q}}/F_0)$ on $\varinjlim_{\mathcal{U}} H^*_{\text{ét}}(\mathsf{X}_{\mathcal{U}}, \mathbb{Q}_\ell)$, which is compatible with the Hecke action of $\mathrm{G}(\mathbb{A}^\infty)$. This is an instance of what we meant by the *compatibility* between Hecke and Galois symmetries in the introduction (see Section 1).

Remark 2.4.7. In Theorem 2.4.3, if we assume instead that $(\mathrm{G}, \mathcal{D}^+)$ is a connected Shimura datum, and consider the collection $\{\Gamma\backslash\mathcal{D}^+\}_\Gamma$ as in Theorem 2.4.1, then it is only true that each $\Gamma\backslash\mathcal{D}^+$ is defined over a number field depending on both $(\mathrm{G}, \mathcal{D}^+)$ and Γ, but in general we cannot expect the whole collection $\{\Gamma\backslash\mathcal{D}^+\}_\Gamma$, or just the subcollection formed by those Γ whose preimage in $\mathrm{G}(\mathbb{Q})$ are congruence subgroups, to be defined over a single number field depending only on $(\mathrm{G}, \mathcal{D})$. This is a generalization of the following classical phenomenon: When $(\mathrm{G}, \mathcal{D}) = (\mathrm{SL}_2, \mathcal{H})$, and when interpreting the quotient $\Gamma(N)\backslash\mathcal{H}$ as in Example 2.2.23 as a parameter space for complex elliptic curves $E_z = \mathbb{C}/(\mathbb{Z}z + \mathbb{Z})$ with level N-structures

$$\begin{aligned} \alpha_N : (\mathbb{Z}/N\mathbb{Z})^2 \xrightarrow{\sim} E_z[N] : \\ (a, b) \mapsto (\tfrac{a}{N}z + \tfrac{b}{N}) \bmod (\mathbb{Z}z + \mathbb{Z}), \end{aligned} \tag{2.4.8}$$

where $z \in \mathcal{H}$ and $E_z[N]$ denotes the N-torsion subgroup of E_z, by identifying E_z with $E_{\gamma z}$ for all $\gamma \in \Gamma(N)$, the value of the Weil pairing of $\alpha_N((1,0),(0,1))$ is a constant primitive N-th root ζ_N of unity on $\Gamma(N)\backslash\mathcal{H}$. (See the introduction of [DR73].) We can only expect $\Gamma(N)\backslash\mathcal{H}$ to have a model over $\mathbb{Q}(\zeta_N)$, rather than $\mathbb{Q}$.

Remark 2.4.9. In the context of Theorem 2.4.3, people may refer to either the varieties $\mathsf{X}_{\mathcal{U}}$ over $\mathbb{C}$, or their canonical models over F_0, or the projective limits of such varieties over either $\mathbb{C}$ or F_0, as the *Shimura varieties* associated with the Shimura datum $(\mathrm{G}, \mathcal{D})$. For the sake of clarity, people sometimes say that $\mathsf{X}_{\mathcal{U}}$ and its canonical model are Shimura varieties at level $\mathcal{U}$. In the context of Remark 2.4.7, people often refer to quotients of the form $\Gamma\backslash\mathcal{D}^+$, where Γ are arithmetic subgroups of

$$\mathrm{G}^{\mathrm{ad}}(\mathbb{Q})^{+} := \mathrm{G}^{\mathrm{ad}}(\mathbb{Q}) \cap \mathrm{G}^{\mathrm{ad}}(\mathbb{R})^{+} \tag{2.4.10}$$

whose preimages under the canonical homomorphism $\mathrm{G}(\mathbb{Q}) \to \mathrm{G}^{\mathrm{ad}}(\mathbb{Q})$ are congruence subgroups of $\mathrm{G}(\mathbb{Q})$, as the *connected Shimura varieties* associated with the connected Shimura datum $(\mathrm{G}, \mathcal{D}^{+})$. We allow all such subgroups Γ of $\mathrm{G}^{\mathrm{ad}}(\mathbb{Q})^{+}$, rather than only the images of the congruence subgroups of $\mathrm{G}(\mathbb{Q})$, so that the connected Shimura varieties thus defined are useful for studying all connected components of Shimura varieties.

Remark 2.4.11. By a morphism of Shimura data

$$(\mathrm{G}_1, \mathcal{D}_1) \to (\mathrm{G}_2, \mathcal{D}_2), \tag{2.4.12}$$

we mean a group homomorphism $\mathrm{G}_1 \to \mathrm{G}_2$ mapping $\mathcal{D}_1$ into $\mathcal{D}_2$. If $\mathcal{D} = \mathrm{G}_1(\mathbb{R}) \cdot h_0$ is the conjugacy class of some homomorphism $h_0 : \mathbb{S} \to \mathrm{G}_{1,\mathbb{R}}$, then this means the composition of h_0 with $\mathrm{G}_{1,\mathbb{R}} \to \mathrm{G}_{2,\mathbb{R}}$ lies in the conjugacy class $\mathcal{D}_2$. If $\mathcal{U}_1$ and $\mathcal{U}_2$ are open compact subgroups of $\mathrm{G}_1(\mathbb{A}^{\infty})$ and $\mathrm{G}_2(\mathbb{A}^{\infty})$, respectively, such that $\mathcal{U}_1$ is mapped into $\mathcal{U}_2$, then we obtain the corresponding morphism $\mathsf{X}_{\mathcal{U}_1} \to \mathsf{X}_{\mathcal{U}_2}$ between Shimura varieties. This morphism is defined over the subfield of $\mathbb{C}$ generated by the reflex fields of $(\mathrm{G}_1, \mathcal{D}_1)$ and $(\mathrm{G}_2, \mathcal{D}_2)$. There are also analogues of these assertions for connected Shimura data and the corresponding connected Shimura varieties.

Remark 2.4.13. An important special case of a morphism of Shimura data as in Remark 2.4.11 is when the homomorphism $\mathrm{G}_1 \to \mathrm{G}_2$ is *injective*, and when $\mathcal{U}_1$ is exactly the pullback of $\mathcal{U}_2$ under the induced homomorphism $\mathrm{G}_1(\mathbb{A}^{\infty}) \to \mathrm{G}_2(\mathbb{A}^{\infty})$. In this case, up to replacing $\mathcal{U}_2$ with an open compact subgroup (satisfying the same conditions), $\mathsf{X}_{\mathcal{U}_1}$ is a closed subvariety of $\mathsf{X}_{\mathcal{U}_2}$. (It is always possible to find such a $\mathcal{U}_2$ for any given $\mathcal{U}_1$—see [Del71b, Prop. 1.15]. The neatness of $\mathcal{U}_2$ can be arranged when $\mathcal{U}_1$ is neat, by taking $\mathcal{U}_2$ to be generated by the image of $\mathcal{U}_1$ and some suitable neat open compact subgroup of $\mathrm{G}_2(\mathbb{A}^{\infty})$.) Subvarieties of this kind are called *special subvarieties*. The most important examples are given by *special points* or (with some additional assumptions on the Shimura data) *CM points*, which are *zero-dimensional* special subvarieties defined by subgroups G_1 of G_2 that are *tori* (see [Mil05, Def. 12.5]). These are generalizations of the points of modular curves parameterizing CM elliptic curves with level structures, which are represented by points of the Poincaré upper half-plane with coordinates in imaginary quadratic field extensions of $\mathbb{Q}$. Zero-dimensional Shimura varieties and special points are important because their canonical models can be defined more directly, and hence they are useful for characterizing the canonical models of Shimura varieties of positive dimensions.

We will not try to explain the proofs of Theorems 2.4.1 and 2.4.3 in this introductory article. In fact, we will not even try to explain in any depth the background notions such as variations of Hodge structures or Hermitian symmetric domains, or key notions such as zero-dimensional Shimura varieties, the Shimura reciprocity law, and the precise meaning of canonical models, because most existing texts already present them quite well, and there is little we can add to them. Instead, we will start by presenting many examples of $(\mathrm{G}, \mathcal{D})$ and $(\mathrm{G}, \mathcal{D}^{+})$, so that the readers can have some idea to where the theory applies. We hope this will benefit not only readers who would like to better understand the existing texts, but also readers who have concrete arithmetic or geometric problems in mind.

3. Hermitian symmetric domains

In this section, we will explore many examples of Hermitian symmetric domains, which can be interpreted as the generalizations of the Poincaré upper half-plane $\mathcal{H}$ (see Example 2.2.7) in many different ways. Most of these will give the $(\mathrm{G}(\mathbb{R}), \mathcal{D})$ or $(\mathrm{G}(\mathbb{R}), \mathcal{D}^+)$ for some Shimura data $(\mathrm{G}, \mathcal{D})$ or connected Shimura data $(\mathrm{G}, \mathcal{D}^+)$, respectively, as in Section 2.3, although we will not be using the structure of $\mathrm{G}(\mathbb{Q})$ in this section. (Later we will indeed use the structure of $\mathrm{G}(\mathbb{Q})$ in Section 4.)

However, we will *not* define Hermitian symmetric domains. Our excuse for this omission is that there is a complete classification of all *irreducible* ones (namely, those that are not isomorphic to nontrivial products of smaller ones), and we will see all of them in our examples. (For a more formal treatment, there are general introductions such as [Mil05, Sec. 1], with references to more advanced texts such as [Hel01].)

In our personal experience, especially in number-theoretic applications requiring actual calculations, the abstract theory of Hermitian symmetric domains cannot easily replace the knowledge of these explicit examples. (Already in the classical theory for SL_2 or GL_2, it is hard to imagine studying modular forms without ever introducing the Poincaré upper half-plane.)

We shall begin with the case of $\mathrm{Sp}_{2n}(\mathbb{R})$ in Section 3.1, followed by the case of $\mathrm{U}_{a,b}$ in Section 3.2, the case of SO^*_{2n} in Section 3.3, the case of $\mathrm{SO}_{a,2}(\mathbb{R})$ in Section 3.4, the case of E_7 in Section 3.5, the case of E_6 in Section 3.6, and a summary of all these cases in Section 3.7.

3.1. The case of $\mathrm{Sp}_{2n}(\mathbb{R})$.

3.1.1. *Siegel upper half-spaces.* Let $n \geq 0$ be any integer. Consider the *Siegel upper half-space*

(3.1.1.1)
$$\begin{aligned}\mathcal{H}_n &:= \{Z \in \mathrm{Sym}_n(\mathbb{C}) : \mathrm{Im}(Z) > 0\}\\ &= \{Z \in \mathrm{M}_n(\mathbb{C}) : {}^t Z - Z = 0,\ \tfrac{1}{2i}({}^t\overline{Z} - Z) < 0\}\\ &= \left\{Z \in \mathrm{M}_n(\mathbb{C}) : {}^t\begin{pmatrix} Z \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} Z \\ 1 \end{pmatrix} = 0,\ \tfrac{1}{2i}\,{}^t\begin{pmatrix} \overline{Z} \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} Z \\ 1 \end{pmatrix} < 0\right\},\end{aligned}$$

where M_n (resp. Sym_n) denotes the space of $n \times n$ matrices (resp. symmetric matrices) as before, and where the notation > 0 (resp. < 0) means *positive definiteness* (resp. *negative definiteness*) of matrices, which admits a natural (left) action of $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2n}(\mathbb{R})$:

$$Z \mapsto gZ := (AZ + B)(CZ + D)^{-1} \tag{3.1.1.2}$$

(note that $(CZ + D)^{-1}$ is multiplied to the right of $(AZ + B)$). Again, such an action should be understood in generalized projective coordinates:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} Z \\ 1 \end{pmatrix} = \begin{pmatrix} AZ + B \\ CZ + D \end{pmatrix} \sim \begin{pmatrix} gZ \\ 1 \end{pmatrix} \tag{3.1.1.3}$$

(by multiplying $(CZ + D)^{-1}$ to the right).

Remark 3.1.1.4. By *generalized projective coordinates* given by an $m \times n$ matrix, we mean an $m \times n$ matrix of rank n that is identified up to the right action of invertible

$n \times n$ matrices. Giving such a matrix is equivalent to defining a point of the *Grassmannian* of n-dimensional subspaces of an m-dimensional vector space. Later we will encounter some similar generalized projective coordinates when defining other Hermitian symmetric domains, and the readers can interpret them as defining points on some other Grassmannians. This is closely related to the interpretation of Hermitian symmetric domains as parameter spaces of variations of Hodge structures, where the Grassmannians parameterize some *Hodge filtrations*.

Let us briefly explain why the action is indeed defined. Since

$$(3.1.1.5)\qquad {}^t\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

we have

$$(3.1.1.6)\qquad \begin{aligned} {}^t\begin{pmatrix} Z \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} Z \\ 1 \end{pmatrix} &= {}^t\begin{pmatrix} Z \\ 1 \end{pmatrix}{}^tg\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}g\begin{pmatrix} Z \\ 1 \end{pmatrix} \\ &= {}^t(CZ+D)\,{}^t\begin{pmatrix} gZ \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} gZ \\ 1 \end{pmatrix}(CZ+D) = 0 \end{aligned}$$

if and only if

$$(3.1.1.7)\qquad {}^t\begin{pmatrix} gZ \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} gZ \\ 1 \end{pmatrix} = 0;$$

and

$$(3.1.1.8)\qquad \begin{aligned} \tfrac{1}{2i}\,{}^t\begin{pmatrix} \overline{Z} \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} Z \\ 1 \end{pmatrix} &= \tfrac{1}{2i}\,{}^t\begin{pmatrix} \overline{Z} \\ 1 \end{pmatrix}{}^tg\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}g\begin{pmatrix} Z \\ 1 \end{pmatrix} \\ &= \tfrac{1}{2i}\,{}^t(C\overline{Z}+D)\,{}^t\begin{pmatrix} g\overline{Z} \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} gZ \\ 1 \end{pmatrix}(CZ+D) < 0 \end{aligned}$$

if and only if

$$(3.1.1.9)\qquad \tfrac{1}{2i}\,{}^t\begin{pmatrix} g\overline{Z} \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} gZ \\ 1 \end{pmatrix} < 0.$$

Thus, $Z \in \mathcal{H}_n$ if and only if $gZ \in \mathcal{H}_n$, and the action is indeed defined.

3.1.2. *Transitivity of action.* If $Z = X + iY \in \mathcal{H}_n$, then

$$(3.1.2.1)\qquad \begin{pmatrix} 1 & -X \\ 0 & 1 \end{pmatrix} Z = iY,$$

with $Y > 0$, and so there exists some $A \in \mathrm{GL}_n(\mathbb{R})$ such that

$$(3.1.2.2)\qquad {}^tAYA = 1_n,$$

which shows that

$$(3.1.2.3)\qquad \begin{pmatrix} {}^tA & 0 \\ 0 & A \end{pmatrix}(iY) = i1_n.$$

These show that

$$(3.1.2.4)\qquad \mathcal{H}_n = \mathrm{Sp}_{2n}(\mathbb{R}) \cdot (i1_n).$$

The stabilizer of $i1_n$ is

$$(3.1.2.5)\qquad K := \left\{\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathrm{Sp}_{2n}(\mathbb{R})\right\} \cong \mathrm{U}_n := \left\{g = A + iB :\, {}^t\overline{g}g = 1_n\right\}$$

(where the conditions on A and B are the same for $\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathrm{Sp}_{2n}(\mathbb{R})$ and for $g = A + iB \in \mathrm{U}_n$), which is a maximal compact subgroup of $\mathrm{Sp}_{2n}(\mathbb{R})$. (Certainly, the precise matrix expressions of these elements depend on the choice of the point $i1_n$.) Thus, we have shown that

$$\mathcal{H}_n \cong \mathrm{Sp}_{2n}(\mathbb{R})/\mathrm{U}_n. \tag{3.1.2.6}$$

When $n = 1$, this specializes to the isomorphism $\mathcal{H} = \mathcal{H}_1 \cong \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$ in Example 2.2.7, because $\mathrm{SL}_2(\mathbb{R}) \cong \mathrm{Sp}_2(\mathbb{R})$ and $\mathrm{SO}_2(\mathbb{R}) \cong \mathrm{U}_1$. Note that

$$\dim_{\mathbb{C}} \mathcal{H}_n = \dim_{\mathbb{C}}\big(\mathrm{Sym}_n(\mathbb{C})\big) = \tfrac{1}{2}n(n+1), \tag{3.1.2.7}$$

for all $n \geq 0$.

3.1.3. *Bounded realization.* There is also a *bounded realization* of $\mathcal{H}_n$. Note that

$$\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2i \end{pmatrix} \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.1.3.1}$$

Set

$$U := (Z - i1_n)(Z + i1_n)^{-1}, \tag{3.1.3.2}$$

so that

$$\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} Z \\ 1 \end{pmatrix} = \begin{pmatrix} Z - i1_n \\ Z + i1_n \end{pmatrix} \sim \begin{pmatrix} U \\ 1 \end{pmatrix} \tag{3.1.3.3}$$

in generalized projective coordinates (cf. Remark 3.1.1.4). Note that this assignment maps $Z = i1_0$ to $U = 0$. Since

$$\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \begin{pmatrix} 0 & 2i \\ -2i & 0 \end{pmatrix} = 2i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{3.1.3.4}$$

we have

$${}^tZ = Z \iff {}^tU = U, \tag{3.1.3.5}$$

and

$$\begin{aligned} &\tfrac{1}{2i}({}^t\overline{Z} - Z) = \tfrac{1}{2i}\,{}^t\begin{pmatrix} \overline{Z} \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} Z \\ 1 \end{pmatrix} \\ &= \tfrac{1}{8i}\,{}^t\begin{pmatrix} \overline{Z} \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} Z \\ 1 \end{pmatrix} < 0 \\ &\iff {}^t\overline{U}U - 1_n = {}^t\begin{pmatrix} \overline{U} \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} U \\ 1 \end{pmatrix} < 0, \end{aligned} \tag{3.1.3.6}$$

and so

$$\mathcal{H}_n \cong \mathcal{D}_n := \{U \in \mathrm{Sym}_n(\mathbb{C}) : {}^t\overline{U}U - 1_n < 0\}. \tag{3.1.3.7}$$

This generalizes the Cayley transforms that move the Poincaré upper half-plane $\mathcal{H}$ to the open unit disc $\{u \in \mathbb{C} : |u| < 1\}$. (But the formula for the action of $\mathrm{Sp}_{2n}(\mathbb{R})$ has to be conjugated, and is more complicated.)

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{H}_n \cong \mathcal{D}_n$ is a Hermitian symmetric domain of type C I. In Cartan's classification, it is a bounded symmetric domain of type III_n.

3.1.4. *Interpretation as conjugacy classes.* Now consider the homomorphism

$$h_0 : \mathrm{U}_1 \to \mathrm{Sp}_{2n}(\mathbb{R}) : x + yi \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}. \tag{3.1.4.1}$$

For simplicity, in the matrix at the right-hand side, we denoted by x and y the scalar multiples of the $n \times n$ identity matrix 1_n by x and y, respectively. (We will adopt similar abuses of notation later, without further explanation.) Then we have

$$\begin{aligned} & \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ \iff & \begin{pmatrix} -B & A \\ -D & C \end{pmatrix} = \begin{pmatrix} C & D \\ -A & -B \end{pmatrix} \\ \iff & A = D,\ B = -C, \end{aligned} \tag{3.1.4.2}$$

which shows that

$$\mathrm{Cent}_{h_0}(\mathrm{Sp}_{2n}(\mathbb{R})) = \mathrm{U}_n \tag{3.1.4.3}$$

as subgroups of $\mathrm{Sp}_{2n}(\mathbb{R})$, where Cent denotes the centralizer, and so that

$$\mathcal{H}_n = \mathrm{Sp}_{2n}(\mathbb{R}) \cdot h_0, \tag{3.1.4.4}$$

with the action given by conjugation. That is, $\mathcal{H}_n$ is the $\mathrm{Sp}_{2n}(\mathbb{R})$-conjugacy class of h_0. In this case, $(\mathrm{Sp}_{2n}, \mathcal{H}_n)$ is a connected Shimura datum as in Section 2.3 (where we use Sp_{2n} to denote the algebraic group over $\mathbb{Q}$, not the group scheme over $\mathbb{Z}$), taking into account the variant (2.3.8) of (2.3.5) as explained there. (For simplicity, we shall not repeat such an explanation in what follows.)

The above is compatible with the similarly defined action of $\mathrm{GSp}_{2n}(\mathbb{R})$ on

$$\mathcal{H}_n^{\pm} := \{Z \in \mathrm{M}_n(\mathbb{C}) : {}^t Z = Z, \text{ and either } \mathrm{Im}(Z) > 0 \text{ or } \mathrm{Im}(Z) < 0\}. \tag{3.1.4.5}$$

Then

$$\mathcal{H}_n^{\pm} = \mathrm{GSp}_{2n}(\mathbb{R}) \cdot h_0, \tag{3.1.4.6}$$

where

$$h_0 : \mathbb{C}^\times \to \mathrm{GSp}_{2n}(\mathbb{R}) : x + yi \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \tag{3.1.4.7}$$

also encodes the action of $\mathbb{R}^\times \subset \mathbb{C}^\times$, and is better for the purpose of studying variations of Hodge structures (and also for various other reasons). In this case, $(\mathrm{GSp}_{2n}, \mathcal{H}_n^{\pm})$ is a Shimura datum as in Section 2.3 (where we use GSp_{2n} to denote the algebraic group over $\mathbb{Q}$, not the group scheme over $\mathbb{Z}$).

3.1.5. *Moduli of polarized abelian varieties.* For each $Z \in \mathcal{H}_n$, we have a lattice

$$L_Z := \mathbb{Z}^n Z + \mathbb{Z}^n \tag{3.1.5.1}$$

in $\mathbb{C}^n$ such that

$$A_Z := \mathbb{C}^n / L_Z \tag{3.1.5.2}$$

is not only a complex torus, but also a *polarized abelian variety*. (See [Mum70, Sec. 1–3] and [Mum83, Ch. II, Sec. 1 and 4].) When $n = 1$, this is just an elliptic curve.

The point is that the abelian variety $A_Z = \mathbb{C}^n / L_Z$ (as a complex torus) can be identified with the fixed real torus

$$\mathbb{R}^{2n} / \mathbb{Z}^{2n} \tag{3.1.5.3}$$

with its *complex structure* induced by the one on $\mathbb{R}^{2n}$ defined by

$$h(i) \in \mathrm{Sp}_{2n}(\mathbb{R}) \subset \mathrm{GL}_{2n}(\mathbb{R}), \tag{3.1.5.4}$$

where the homomorphism

$$h = g \cdot h_0 \in \mathcal{H}_n = \mathrm{Sp}_{2n}(\mathbb{R}) \cdot h_0 \tag{3.1.5.5}$$

corresponds to the point

$$Z = g \cdot (i1_n) \in \mathcal{H}_n = \mathrm{Sp}_{2n}(\mathbb{R}) \cdot (i1_n). \tag{3.1.5.6}$$

Accordingly, varying the lattice L_Z in $\mathbb{C}^n$ with $Z = g \cdot (i1_n)$ corresponds to varying the complex structure $J = h(i)$ on $\mathbb{R}^{2n}$ with $h = g \cdot h_0$. (Such a variation of complex structures can be interpreted as a *variation of Hodge structures*.)

The pairing on $\mathbb{Z}^{2n}$ defined by the perfect pairing $\langle \cdot\,,\,\cdot \rangle$ defines (up to a sign convention) a *principal polarization* λ_Z on A_Z (and the existence of such a polarization was the reason that the complex torus A_Z is an abelian variety). If $(A_Z, \lambda_Z) \cong (A_{Z'}, \lambda_{Z'})$ for Z and Z' corresponding to h and h', respectively, then we have an induced isomorphism $\mathbb{R}^{2n} \xrightarrow{\sim} \mathbb{R}^{2n}$ preserving $\mathbb{Z}^{2n}$ and $\langle \cdot\,,\,\cdot \rangle$, which defines an element $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2n}(\mathbb{Z})$ conjugating h to h' and mapping Z to Z'. Conversely, if we have

$$Z' = gZ = (AZ + B)(CZ + D)^{-1} \tag{3.1.5.7}$$

for some $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2n}(\mathbb{Z})$, then we have a *homothety of lattices*

$$\begin{aligned} L_Z &= \mathbb{Z}^{2n} Z + \mathbb{Z}^{2n} = \mathbb{Z}^{2n}(AZ + B) + \mathbb{Z}^{2n}(CZ + D) \\ &= (\mathbb{Z}^{2n} Z' + \mathbb{Z}^{2n})(CZ + D) = L_{Z'}(CZ + D) \sim L_{Z'}, \end{aligned} \tag{3.1.5.8}$$

which implies that $(A_Z, \lambda_Z) \cong (A_{Z'}, \lambda_{Z'})$. If g lies in a congruence subgroup Γ of $\mathrm{Sp}_{2n}(\mathbb{Z})$, then the above isomorphism $(A_Z, \lambda_Z) \cong (A_{Z'}, \lambda_{Z'})$ also respect certain (symplectic) level structures. Therefore, $\Gamma \backslash \mathcal{H}_n$ (or the manifolds $\mathsf{X}_{\mathcal{U}}$ defined using $\mathrm{G} = \mathrm{GSp}_{2n}$ and $\mathcal{D} = \mathcal{H}_n^{\pm}$) are *Siegel modular varieties* parameterizing polarized abelian varieties with certain level structures (depending on Γ or $\mathcal{U}$). The spaces $\Gamma \backslash \mathcal{H}_n$ are, a priori, just complex manifolds. The name “Siegel modular varieties” will be justified in Sections 4.1.3 and 4.2.2 below, where we shall see that they are indeed the complex analytifications of some canonical algebraic varieties.

The fact that spaces like $\Gamma \backslash \mathcal{H}_n$ parameterize polarized abelian varieties with level structures is perhaps just a coincidence, but such a coincidence is extremely important. One can redefine them as the $\mathbb{C}$-points of certain moduli problems of polarized abelian schemes with level structures; and then their constructions can be extended over the integers in number fields, because the moduli schemes of polarized abelian schemes, called *Siegel moduli schemes*, have been constructed over the integers. (See [MFK94, Ch. 7] for the construction based on geometric invariant theory. Alternatively, geometric invariant theory can be avoided by first constructing algebraic spaces or Deligne–Mumford stacks by verifying Artin’s criterion—see [Art69] and [Lan13, Sec. B.3] for such a criterion, and see [FC90, Ch. I, Sec. 4] and [Lan13, Ch. 2] for its verification for the moduli of polarized abelian schemes; and then showing that they are quasi-projective schemes by constructing their minimal compactifications as in [FC90, Ch. V] or [Lan13, Ch. 7]—see Section 5.1.4 below.)

Such *integral models* of $\Gamma\backslash\mathcal{H}_n$ are very useful because we can understand their reductions in positive characteristics by studying polarized abelian varieties over finite fields, for which we have many powerful tools. We will encounter more general cases of $\Gamma\backslash\mathcal{D}$ later, but we have good methods for constructing integral models for them (namely, methods for which we can say something useful about the outputs) only in cases that can be related to Siegel modular varieties at all.

3.2. **The case of** $\mathrm{U}_{a,b}$**.** We learned some materials here from [Shi98] and [Shi78].

3.2.1. *Bounded realization.* Let $a \geq b \geq 0$ be any integers. Recall that we have introduced the symmetric matrix $1_{a,b} = \begin{pmatrix} 1_a & \\ & -1_b \end{pmatrix}$ in (2.1.13), which we now view as a *Hermitian* matrix, which defines a Hermitian pairing of signature (a, b). Consider the group

$$\mathrm{U}_{a,b} := \{g \in \mathrm{GL}_{a+b}(\mathbb{C}) : {}^t\overline{g}1_{a,b}g = 1_{a,b}\}, \tag{3.2.1.1}$$

which acts on

$$\mathcal{D}_{a,b} := \left\{U \in \mathrm{M}_{a,b}(\mathbb{C}) : {}^t\begin{pmatrix} \overline{U} \\ 1 \end{pmatrix} \begin{pmatrix} 1_a & \\ & -1_b \end{pmatrix} \begin{pmatrix} U \\ 1 \end{pmatrix} = {}^t\overline{U}U - 1_b < 0\right\}, \tag{3.2.1.2}$$

where $\mathrm{M}_{a,b}$ denotes the space of $a \times b$ matrices, by

$$U \mapsto gU = (AU + B)(CU + D)^{-1} \tag{3.2.1.3}$$

for each $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{U}_{a,b}$. This action is based on the identity

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} U \\ 1 \end{pmatrix} = \begin{pmatrix} AU + B \\ CU + D \end{pmatrix} \sim \begin{pmatrix} (AU + B)(CU + D)^{-1} \\ 1 \end{pmatrix} \tag{3.2.1.4}$$

in generalized projective coordinates (cf. Remark 3.1.1.4). Note that $\mathrm{U}_{a,b}$ is a real Lie group, but not a complex one, despite the use of complex numbers in its coordinates; and that

$$\dim_{\mathbb{C}} \mathcal{D}_{a,b} = \dim_{\mathbb{C}}\big(\mathrm{M}_{a,b}(\mathbb{C})\big) = ab. \tag{3.2.1.5}$$

For any $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{U}_{a,b}$, as soon as $B = 0$, we must also have $C = 0$, by the defining condition in (3.2.1.1). Hence, the stabilizer of $0 \in \mathcal{D}_{a,b}$ is

$$K := \left\{\begin{pmatrix} A & \\ & D \end{pmatrix} \in \mathrm{U}_{a,b}\right\} \cong \mathrm{U}_a \times \mathrm{U}_b : \begin{pmatrix} A & \\ & D \end{pmatrix} \mapsto (A, D). \tag{3.2.1.6}$$

Remark 3.2.1.7. For the purpose of introducing Hermitian symmetric domains, it might be more natural to consider the semisimple Lie group $\mathrm{SU}_{a,b}$—namely, the subgroup of $\mathrm{U}_{a,b}$ of elements of determinant one—or its adjoint quotient, but we still consider the reductive Lie group $\mathrm{U}_{a,b}$, because it is easier to write down homomorphisms of the form $\mathrm{U}_1 \to \mathrm{U}_{a,b}$, as in (2.3.8).

However, to show the transitivity of the action of $\mathrm{U}_{a,b}$ on $\mathcal{D}_{a,b}$, and for many other computations for which it is easier to work with block upper-triangular matrices, it will be easier to switch to some *unbounded realization*.

3.2.2. *Unbounded realization.* Consider

$$\mathrm{U}'_{a,b} := \{g \in \mathrm{GL}_{a+b}(\mathbb{C}) : {}^t\overline{g} J_{a,b} g = J_{a,b}\}, \tag{3.2.2.1}$$

where the matrix

$$J_{a,b} := \begin{pmatrix} & & 1_b \\ & S & \\ -1_b & & \end{pmatrix} \tag{3.2.2.2}$$

is *skew-Hermitian* with $-iS > 0$, for some choice of a skew-Hermitian matrix S (depending on the context), so that there exists some $T \in \mathrm{GL}_n(\mathbb{C})$ such that $-iS = {}^t\overline{T}T$. We warn the readers that the notation of $J_{a,b}$ and $\mathrm{U}'_{a,b}$ here is rather nonstandard. Then $\mathrm{U}'_{a,b}$ acts on

$$\mathcal{H}_{a,b} := \left\{ \begin{array}{c} \begin{pmatrix} Z \\ W \end{pmatrix} \in \mathrm{M}_{a,b}(\mathbb{C}) \cong \mathrm{M}_b(\mathbb{C}) \times \mathrm{M}_{a-b,b}(\mathbb{C}) : \\ -i\,{}^t\begin{pmatrix} \overline{Z} \\ \overline{W} \\ 1 \end{pmatrix} J_{a,b} \begin{pmatrix} Z \\ W \\ 1 \end{pmatrix} = -i({}^t\overline{Z} - Z + {}^t\overline{W}SW) < 0 \end{array} \right\} \tag{3.2.2.3}$$

by

$$\begin{pmatrix} Z \\ W \end{pmatrix} \mapsto g\begin{pmatrix} Z \\ W \end{pmatrix} = \begin{pmatrix} (AZ+EW+B)(CZ+HW+D)^{-1} \\ (FZ+MW+G)(CZ+HW+D)^{-1} \end{pmatrix}, \tag{3.2.2.4}$$

for each $g = \begin{pmatrix} A & E & B \\ F & M & G \\ C & H & D \end{pmatrix} \in \mathrm{U}'_{a,b}$, based on the identity

$$g\begin{pmatrix} Z \\ W \\ 1 \end{pmatrix} = \begin{pmatrix} AZ+EW+B \\ FZ+MW+G \\ CZ+HW+D \end{pmatrix} \sim \begin{pmatrix} (AZ+EW+B)(CZ+HW+D)^{-1} \\ (FZ+MW+G)(CZ+HW+D)^{-1} \\ 1 \end{pmatrix} \tag{3.2.2.5}$$

in generalized projective coordinates (cf. Remark 3.1.1.4).

If we consider

$$\mathrm{Herm}_b(\mathbb{C}) := \{X \in \mathrm{M}_b(\mathbb{C}) : {}^t\overline{X} = X\}, \tag{3.2.2.6}$$

where Herm_b denotes the space of $b \times b$ Hermitian matrices (which are over $\mathbb{C}$ here), then we have a canonical isomorphism

$$\mathrm{M}_b(\mathbb{C}) \xrightarrow{\sim} \mathrm{Herm}_b(\mathbb{C}) \underset{\mathbb{R}}{\otimes} \mathbb{C} : X \mapsto \mathrm{Re}(X) + i\,\mathrm{Im}(X), \tag{3.2.2.7}$$

where

$$\mathrm{Re}(X) := \tfrac{1}{2}(X + {}^t\overline{X}) \tag{3.2.2.8}$$

and

$$\mathrm{Im}(X) := \tfrac{1}{2i}(X - {}^t\overline{X}) \tag{3.2.2.9}$$

are the *Hermitian* real and imaginary parts (abusively denoted by the usual symbols), and we can rewrite the condition

$$-i({}^t\overline{Z} - Z + {}^t\overline{W}SW) < 0 \tag{3.2.2.10}$$

as

$$\mathrm{Im}(Z) > \tfrac{1}{2}\,{}^t\overline{W}(-iS)W. \tag{3.2.2.11}$$

3.2.3. *The special case where $a = b$: Hermitian upper half-spaces.* In this case, we can omit the second coordinate W in the above unbounded realization, and we have

$$\mathcal{H}_{b,b} = \{Z \in \mathrm{M}_b(\mathbb{C}) \cong \mathrm{Herm}_b(\mathbb{C}) \underset{\mathbb{R}}{\otimes} \mathbb{C} : \mathrm{Im}(Z) > 0\}, \tag{3.2.3.1}$$

which can be viewed as a generalization of

$$\mathcal{H}_n = \{Z \in \mathrm{Sym}_n(\mathbb{C}) \cong \mathrm{Sym}_n(\mathbb{R}) \underset{\mathbb{R}}{\otimes} \mathbb{C} : \mathrm{Im}(Z) > 0\}. \tag{3.2.3.2}$$

That is, we can view the symmetric matrices in $\mathrm{Sym}_n(\mathbb{R})$ as being Hermitian with respect to the trivial involution of $\mathbb{R}$, which then generalizes to the usual complex Hermitian matrices in $\mathrm{Herm}_n(\mathbb{C})$ as being Hermitian with respect to the complex conjugation of $\mathbb{C}$ over $\mathbb{R}$. These two examples will be further generalized with $\mathbb{R}$ and $\mathbb{C}$ replaced with the other two normed (but possibly nonassociative) division algebras $\mathbb{H}$ and $\mathbb{O}$ over $\mathbb{R}$, the rings of Hamilton and Cayley numbers, respectively, in Sections 3.3.2 and 3.5.4 below.

3.2.4. *Generalized Cayley transformations and transitivity of actions.* Since

(3.2.4.1)
$$-i \begin{pmatrix} \frac{1}{\sqrt{2}} & & -\frac{i}{\sqrt{2}} \\ & {}^t\overline{T}^{-1} & \\ \frac{1}{\sqrt{2}} & & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} & & 1 \\ & S & \\ -1 & & \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & & \frac{1}{\sqrt{2}} \\ & T^{-1} & \\ \frac{i}{\sqrt{2}} & & -\frac{i}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix},$$

we have an isomorphism

$$\mathrm{U}'_{a,b} \xrightarrow{\sim} \mathrm{U}_{a,b} : g \mapsto \begin{pmatrix} \frac{1}{\sqrt{2}} & & -\frac{i}{\sqrt{2}} \\ & T & \\ \frac{1}{\sqrt{2}} & & \frac{i}{\sqrt{2}} \end{pmatrix} g \begin{pmatrix} \frac{1}{\sqrt{2}} & & \frac{1}{\sqrt{2}} \\ & T^{-1} & \\ \frac{i}{\sqrt{2}} & & -\frac{i}{\sqrt{2}} \end{pmatrix}, \tag{3.2.4.2}$$

and accordingly an isomorphism

$$\mathcal{H}_{a,b} \xrightarrow{\sim} \mathcal{D}_{a,b} : \begin{pmatrix} Z \\ W \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} \frac{1}{\sqrt{2}} & & -\frac{i}{\sqrt{2}} \\ & T & \\ \frac{1}{\sqrt{2}} & & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} Z \\ W \\ 1 \end{pmatrix} \sim \begin{pmatrix} U \\ 1 \end{pmatrix} \tag{3.2.4.3}$$

equivariant with the above isomorphism (3.2.4.2). This is a generalization of the classical Cayley transformation and maps $\begin{pmatrix} i1_b \\ 0 \end{pmatrix} \in \mathcal{H}_{a,b}$ to $0 \in \mathcal{D}_{a,b}$.

Now, given any $\begin{pmatrix} Z \\ W \end{pmatrix} \in \mathcal{H}_{a,b}$, we have

$$\begin{pmatrix} 1 & * & * \\ & 1 & -W \\ & & 1 \end{pmatrix} \begin{pmatrix} Z \\ W \\ 1 \end{pmatrix} = \begin{pmatrix} Z' \\ 0 \\ 1 \end{pmatrix}, \tag{3.2.4.4}$$

$$\begin{pmatrix} 1 & & -\mathrm{Re}(Z') \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} Z' \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} i\,\mathrm{Im}(Z') \\ 0 \\ 1 \end{pmatrix}, \tag{3.2.4.5}$$

$$\begin{pmatrix} {}^t\overline{A}^{-1} & & \\ & 1 & \\ & & A \end{pmatrix} \begin{pmatrix} i\,\mathrm{Im}(Z') \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} i1_b \\ 0 \\ 1 \end{pmatrix}, \tag{3.2.4.6}$$

for some $Z' \in \mathrm{M}_b(\mathbb{C})$ and some $A \in \mathrm{GL}_b(\mathbb{C})$ such that $\mathrm{Im}(Z') = {}^t\overline{A}A$, where the first matrix in each equation is some element of $\mathrm{U}'_{a,b}$. These show that

$$\mathcal{H}_{a,b} = \mathrm{U}'_{a,b} \cdot \begin{pmatrix} i1_b \\ 0 \end{pmatrix} \cong \mathcal{D}_{a,b} = \mathrm{U}_{a,b} \cdot 0 \cong \mathrm{U}_{a,b}/(\mathrm{U}_a \times \mathrm{U}_b), \tag{3.2.4.7}$$

and hence the actions of $\mathrm{U}_{a,b}$ and $\mathrm{U}'_{a,b}$ on $\mathcal{D}_{a,b}$ and $\mathcal{H}_{a,b}$, respectively, are transitive.

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{H}_{a,b} \cong \mathcal{D}_{a,b}$ is a Hermitian symmetric domain of type A III. In Cartan's classification, it is a bounded symmetric domain of type I_{ba}.

3.2.5. *Interpretation as conjugacy classes.* Consider the homomorphism

$$h_0 : \mathrm{U}_1 \to \mathrm{U}_{a,b} : x + yi \mapsto \begin{pmatrix} x - yi & \\ & x + yi \end{pmatrix}, \tag{3.2.5.1}$$

with $x - yi$ and $x + yi$ denoting multiples of 1_a and 1_b, respectively. Then

$$\mathrm{Cent}_{h_0}(\mathrm{U}_{a,b}) = \mathrm{U}_a \times \mathrm{U}_b \tag{3.2.5.2}$$

(as subgroups of $\mathrm{U}_{a,b}$) and so that

$$\mathcal{D}_{a,b} = \mathrm{U}_{a,b} \cdot h_0. \tag{3.2.5.3}$$

Via the inverse of the isomorphism $\mathrm{U}_{a,b} \xrightarrow{\sim} \mathrm{U}'_{a,b}$ above defined by conjugation, we obtain

$$h'_0 : \mathrm{U}_1 \to \mathrm{U}'_{a,b} : x + yi \mapsto \begin{pmatrix} x & & -y \\ & x - yi & \\ y & & x \end{pmatrix}, \tag{3.2.5.4}$$

so that

$$\mathcal{H}_{a,b} = \mathrm{U}'_{a,b} \cdot h'_0. \tag{3.2.5.5}$$

Both of these are compatible with the extensions of the actions of $\mathrm{U}_{a,b}$ and $\mathrm{U}'_{a,b}$ on $\mathcal{D}_{a,b}$ and $\mathcal{H}_{a,b}$, respectively, to some actions of

$$\mathrm{GU}_{a,b} := \{(g, r) \in \mathrm{GL}_{a+b}(\mathbb{C}) \times \mathbb{R}^\times : {}^t\overline{g}1_{a,b}g = r1_{a,b}\} \tag{3.2.5.6}$$

and

$$\mathrm{GU}'_{a,b} := \{(g, r) \in \mathrm{GL}_{a+b}(\mathbb{C}) \times \mathbb{R}^\times : {}^t\overline{g}J_{a,b}g = rJ_{a,b}\} \tag{3.2.5.7}$$

on some

$$\mathcal{D}^\pm_{a,b} = \mathrm{GU}_{a,b} \cdot h_0 \tag{3.2.5.8}$$

and

$$\mathcal{H}^\pm_{a,b} = \mathrm{GU}_{a,b} \cdot h'_0 \tag{3.2.5.9}$$

(where we omit the explicit descriptions of $\mathcal{D}^\pm_{a,b}$ and $\mathcal{H}^\pm_{a,b}$, for simplicity), where h_0 and h'_0 are now homomorphisms

$$h_0 : \mathbb{C}^\times \to \mathrm{GU}_{a,b} \tag{3.2.5.10}$$

and

$$h'_0 : \mathbb{C}^\times \to \mathrm{GU}'_{a,b} \tag{3.2.5.11}$$

defined by the same expressions as in (3.2.5.1) and (3.2.5.4), respectively. (As in Section 3.1.4, the upshot here is that the pairs $(\mathrm{GU}_{a,b}, \mathcal{D}^\pm_{a,b})$ and $(\mathrm{GU}'_{a,b}, \mathcal{H}^\pm_{a,b})$ can be of the form $(\mathrm{G}(\mathbb{R}), \mathcal{D})$ for some Shimura data $(\mathrm{G}, \mathcal{D})$. We shall omit such remarks in later examples.)

Since the real part of any skew-Hermitian pairing is alternating, there is an injective homomorphism of the form

$$\mathrm{U}_{a,b} \cong \mathrm{U}'_{a,b} \hookrightarrow \mathrm{Sp}_{2n}(\mathbb{R}), \tag{3.2.5.12}$$

with $n = a + b$, whose pre-composition with $h_0 : \mathrm{U}_1 \to \mathrm{U}_{a,b}$ lies in the conjugacy class of the analogous homomorphism for $\mathrm{Sp}_{2n}(\mathbb{R})$ in (3.1.4.1).

3.3. The case of SO^*_{2n}.

3.3.1. *Unbounded realization, and interpretation as a conjugacy class.* Consider the subgroup SO^*_{2n} of

$$\mathrm{SO}_{2n}(\mathbb{C}) = \{g \in \mathrm{SL}_{2n}(\mathbb{C}) : {}^t g g = 1_{2n}\} \tag{3.3.1.1}$$

preserving the skew-Hermitian form defined by $\begin{pmatrix} & 1_n \\ -1_n & \end{pmatrix}$. That is,

$$\mathrm{SO}^*_{2n} = \mathrm{SO}_{2n}(\mathbb{C}) \cap \mathrm{U}'_{n,n} \tag{3.3.1.2}$$

as subgroups of $\mathrm{GL}_{2n}(\mathbb{C})$, where $\mathrm{U}'_{n,n}$ is defined as in (3.2.2.1), with Lie algebra

$$\mathfrak{so}^*_{2n} = \left\{ \begin{pmatrix} A & B \\ -\overline{B} & \overline{A} \end{pmatrix} : A, B \in \mathrm{M}_n(\mathbb{C}),\ {}^t A = -A,\ {}^t \overline{B} = B \right\} \tag{3.3.1.3}$$

as Lie subalgebras of $\mathrm{M}_{2n}(\mathbb{C})$ (see, for example, [Kna01, Ch. I, Sec. 1] or [Kna02, Ch. I, Sec. 17]). By explanations as before, SO^*_{2n} acts transitively on

$$\mathcal{H}_{\mathrm{SO}^*_{2n}} := \left\{ \begin{array}{l} Z \in \mathrm{M}_n(\mathbb{C}) : {}^t\begin{pmatrix} Z \\ 1 \end{pmatrix} \begin{pmatrix} Z \\ 1 \end{pmatrix} = {}^t Z Z + 1 = 0, \\ -i\, {}^t\begin{pmatrix} \overline{Z} \\ 1 \end{pmatrix} \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} \begin{pmatrix} Z \\ 1 \end{pmatrix} = -({}^t\overline{Z} - Z) < 0 \end{array} \right\}, \tag{3.3.1.4}$$

whose stabilizer at $i1_n \in \mathcal{H}_{\mathrm{SO}^*_{2n}}$ is

$$K := \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathrm{SO}^*_{2n} : A, B \in \mathrm{M}_n(\mathbb{R}) \right\} \cong \mathrm{U}_n, \tag{3.3.1.5}$$

so that

$$\mathcal{H}_{\mathrm{SO}^*_{2n}} = \mathrm{SO}^*_{2n} \cdot h_0 \cong \mathrm{SO}^*_{2n}/\mathrm{U}_n, \tag{3.3.1.6}$$

where

$$h_0 : \mathrm{U}_1 \to \mathrm{SO}^*_{2n} : x + yi \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}. \tag{3.3.1.7}$$

As before, there is also another version of SO^*_{2n} which acts on some bounded realization $\mathcal{D}_{\mathrm{SO}^*_{2n}}$, but we shall omit them for simplicity. Since $\mathrm{SO}^*_{2n} \subset \mathrm{U}'_{n,n}$, by the same explanation as in Section 3.2.5, there are injective homomorphisms

$$\mathrm{SO}^*_{2n} \hookrightarrow \mathrm{Sp}_{4n}(\mathbb{R}), \tag{3.3.1.8}$$

whose pre-composition with $h_0 : \mathrm{U}_1 \to \mathrm{SO}^*_{2n}$ lies in the conjugacy class of the analogous homomorphism for $\mathrm{Sp}_{4n}(\mathbb{R})$ as in (3.1.4.1).

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{H}_{\mathrm{SO}^*_{2n}} \cong \mathcal{D}_{\mathrm{SO}^*_{2n}}$ is a Hermitian symmetric domain of type D III. In Cartan's classification, it is a bounded symmetric domain of type II_n.

3.3.2. *The special case where $n = 2k$: quaternion upper half-spaces.* Since we have the isomorphism

$$\mathbb{H} \cong \mathbb{C}^2 : x + yi + j(z + wi) \mapsto \begin{pmatrix} x + yi \\ z + wi \end{pmatrix} \tag{3.3.2.1}$$

of right $\mathbb{C}$-modules, where $\mathbb{H}$ denotes the Hamiltonian quaternion algebra over $\mathbb{R}$, we have a homomorphism

$$\mathbb{H} \to \mathrm{M}_2(\mathbb{C}) : x + yi + j(z + wi) \mapsto \begin{pmatrix} x + yi & -z + wi \\ z + wi & x - yi \end{pmatrix} \tag{3.3.2.2}$$

induced by the left action of $\mathbb{H}$ on itself. Using this, we can show that

$$\mathrm{SO}^*_{4k} \cong \left\{ g \in \mathrm{GL}_{2k}(\mathbb{H}) : {}^t\overline{g} \begin{pmatrix} & 1_k \\ -1_k & \end{pmatrix} g = \begin{pmatrix} & 1_k \\ -1_k & \end{pmatrix} \right\}, \tag{3.3.2.3}$$

namely, the subgroup $\mathrm{Sp}_{2k}(\mathbb{H})$ of $\mathrm{GL}_{2k}(\mathbb{H})$ respecting the skew-Hermitian pairing on $\mathbb{H}^k$ defined by $\begin{pmatrix} & 1_k \\ -1_k & \end{pmatrix}$ (see, for example, [GW09, Sec. 1.1.4], combined with a Gram–Schmidt process as in [Lan13, Sec. 1.2.3]); and that

$$\mathcal{H}_{\mathrm{SO}^*_{4k}} \cong \{ Z \in \mathrm{Herm}_k(\mathbb{H}) \underset{\mathbb{R}}{\otimes} \mathbb{C} : \mathrm{Im}(Z) > 0 \}, \tag{3.3.2.4}$$

with the action of $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{GL}_{2k}(\mathbb{H})$ given by

$$gZ = (AZ + B)(CZ + D)^{-1}, \tag{3.3.2.5}$$

which is similar to the action of $\mathrm{Sp}_{2k}(\mathbb{R})$ on $\mathcal{H}_k$ in Section 3.1.1. This is a *quaternion upper half-space*, further generalizing the examples in Sections 3.1.1 and 3.2.3.

3.4. **The case of $\mathrm{SO}_{a,2}(\mathbb{R})$.** We learned some materials here from [Sat80, Appendix, Sec. 6].

3.4.1. *Projective coordinates.* Let $a, b \geq 0$ be any integers. Consider

$$\mathrm{SO}_{a,b}(\mathbb{R}) := \{ g \in \mathrm{SL}_{a+b}(\mathbb{R}) : {}^t g 1_{a,b} g = 1_{a,b} \}, \tag{3.4.1.1}$$

where $1_{a,b} = \begin{pmatrix} 1_a & \\ & -1_b \end{pmatrix}$ is defined as before. Let us temporarily write $V := \mathbb{R}^{a+b}$ and $V_{\mathbb{C}} := V \underset{\mathbb{R}}{\otimes} \mathbb{C} \cong \mathbb{C}^{a+b}$. Then, by definition, $\mathrm{SO}_{a,b}(\mathbb{R})$ acts on

$$\tilde{\mathcal{H}}_{\mathrm{SO}_{a,b}} := \left\{ v \in V_{\mathbb{C}} : {}^t v \begin{pmatrix} 1_a & \\ & -1_b \end{pmatrix} v = 0,\ {}^t\overline{v} \begin{pmatrix} 1_a & \\ & -1_b \end{pmatrix} v < 0 \right\} / \mathbb{C}^\times. \tag{3.4.1.2}$$

Note that $v \neq 0$ in the above definition, and hence the quotient by $\mathbb{C}^\times$ means we are actually working with a subset of $\mathbb{P}^{a+b-1}(\mathbb{C}) \cong (V_{\mathbb{C}} - \{0\})/\mathbb{C}^\times$. (This is consistent with our use of generalized projective coordinates in previous examples.) However, $\tilde{\mathcal{H}}_{\mathrm{SO}_{a,b}}$ is in general not connected, and even its connected components do not have the structure of Hermitian symmetric domains for arbitrary $a, b \geq 0$.

3.4.2. *The case where $a \geq 1$ and $b = 2$.* In this case, we do get Hermitian symmetric domains. (This follows from the general theory, but since we have not defined Hermitian symmetric domains, we are merely stating this as a fact.)

Suppose we have an orthogonal direct sum

$$V = \mathbb{R}^{a+2} = \mathbb{R}^a \oplus \mathbb{R}^2 \tag{3.4.2.1}$$

such that the pairings on $\mathbb{R}^a$ and $\mathbb{R}^2$ are defined by $\begin{pmatrix} 1_{a-1} & \\ & -1 \end{pmatrix}$ and $\begin{pmatrix} & \frac{1}{2} \\ \frac{1}{2} & \end{pmatrix}$, respectively, which have signatures $(a-1, 1)$ and $(1, 1)$ summing up to $(a, 2)$ in total. Note that we have

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} & \frac{1}{2} \\ \frac{1}{2} & \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}. \tag{3.4.2.2}$$

Then we also have

$$V_{\mathbb{C}} \cong \mathbb{C}^a \oplus \mathbb{C}^2, \tag{3.4.2.3}$$

and we denote the coordinates of $\mathbb{C}^a$ and $\mathbb{C}^2$ by $(z_1, \ldots, z_a)$ and (w_1, w_2), respectively. With these coordinates, we have

$$\tilde{\mathcal{H}}_{\mathrm{SO}_{a,2}} \cong \left\{ \begin{matrix} z_1^2 + \cdots + z_{a-1}^2 - z_a^2 + w_1 w_2 = 0, \\ |z_1|^2 + \cdots + |z_{a-1}|^2 - |z_a|^2 + \mathrm{Re}(w_1 \overline{w}_2) < 0 \end{matrix} \right\}. \tag{3.4.2.4}$$

If $w_2 = 0$, then the two conditions in (3.4.2.4) imply that

$$|z_1|^2 + \cdots + |z_{a-1}|^2 < |z_a|^2 = |z_1 + \cdots + z_{a-1}|^2, \tag{3.4.2.5}$$

which is not possible. Hence, we must have $w_2 \neq 0$, and so, up to scaling of the projective coordinates, we may assume that $w_2 = 1$. Then

$$w_1 = z_a^2 - z_1^2 - \cdots - z_{a-1}^2 \tag{3.4.2.6}$$

by the first condition, and the second condition becomes

$$|z_1|^2 + \cdots + |z_{a-1}|^2 - |z_a|^2 + \mathrm{Re}(w_1) < 0. \tag{3.4.2.7}$$

If we write $z_j = x_j + i y_j$, then

$$|z_j|^2 - \mathrm{Re}(z_j^2) = 2 y_j^2, \tag{3.4.2.8}$$

for each $1 \leq j \leq a$. Thus, we have

(3.4.2.9)

$$\tilde{\mathcal{H}}_{\mathrm{SO}_{a,2}} \cong \mathcal{H}_{\mathrm{SO}_{a,2}} := \left\{ \begin{matrix} z = (z_j) \in \mathbb{C}^a : y = \mathrm{Im}(z) > 0 \text{ in the sense that} \\ {}^t y \begin{pmatrix} 1_{a-1} & \\ & -1 \end{pmatrix} y < 0, \text{ i.e., } y_a^2 > y_1^2 + \cdots y_{a-1}^2 \end{matrix} \right\}.$$

The condition

$$y_a^2 > y_1^2 + \cdots y_{a-1}^2 \tag{3.4.2.10}$$

defines a familiar *light cone* with two connected components, one being defined by

$$y_a > \sqrt{y_1^2 + \cdots y_{a-1}^2}, \tag{3.4.2.11}$$

the other being defined by

$$y_a < -\sqrt{y_1^2 + \cdots y_{a-1}^2}. \tag{3.4.2.12}$$

These two connected components correspond to two connected components $\mathcal{H}^+_{\mathrm{SO}_{a,2}}$ and $\mathcal{H}^-_{\mathrm{SO}_{a,2}}$ of $\mathcal{H}_{\mathrm{SO}_{a,2}}$.

Note that

$$\dim_{\mathbb{C}} \tilde{\mathcal{H}}_{\mathrm{SO}_{a,2}} = \dim_{\mathbb{C}} \mathcal{H}_{\mathrm{SO}_{a,2}} = \dim_{\mathbb{C}} \mathcal{H}^+_{\mathrm{SO}_{a,2}} = \dim_{\mathbb{C}} \mathbb{C}^a = a. \tag{3.4.2.13}$$

Also, $\mathcal{H}_{\mathrm{SO}_{1,2}} \cong \mathcal{H}_1^{\pm}$ and $\mathcal{H}^+_{\mathrm{SO}_{1,2}} \cong \mathcal{H}_1 = \mathcal{H}$, when $a = 1$. In particular, we have obtained yet another generalization of the Poincaré upper half-plane $\mathcal{H}$.

3.4.3. *Connected components and spin groups.* Under the assumption that $a \geq 1$, the real Lie group $\mathrm{SO}_{a,2}(\mathbb{R})$ has two connected components (see, for example, [Hel01, Ch. X, Sec. 2, Lem. 2.4] or [Kna02, Prop. 1.145]), and the stabilizer of $\mathcal{H}^+_{\mathrm{SO}_{a,2}}$ in $\mathrm{G}(\mathbb{R}) = \mathrm{SO}_{a,2}(\mathbb{R})$ is the identity component $\mathrm{G}(\mathbb{R})^+ = \mathrm{SO}_{a,2}(\mathbb{R})^+$; namely, the connected component containing the identity element (in the real analytic topology). The group $\mathrm{SO}_{a,2}(\mathbb{R})$ can be viewed as the group of $\mathbb{R}$-points of the algebraic group $\mathrm{SO}_{a,2}$ defined over $\mathbb{R}$, which admits a simply-connected two-fold covering group $\mathrm{Spin}_{a,2}$. This is only a covering of algebraic groups, which means we can only expect a surjection on $\mathbb{C}$-points. In fact, the induced morphism

$$\mathrm{Spin}_{a,2}(\mathbb{R}) \to \mathrm{SO}_{a,2}(\mathbb{R}) \tag{3.4.3.1}$$

on $\mathbb{R}$-points is not surjective, because $\mathrm{Spin}_{a,2}(\mathbb{R})$ is connected and factors through $\mathrm{SO}_{a,2}(\mathbb{R})^+$. Here we are keeping the language as elementary as possible, without introducing the spin groups in detail—see, for example, [GW09, Ch. 6], for a review of their construction based on Clifford algebras. But the connectedness of $\mathrm{Spin}_{a,2}(\mathbb{R})$, nevertheless, is a general feature of the real points of any isotropic simply-connected connected simple algebraic group over $\mathbb{Q}$—see [PR94, Sec. 7.3, Thm. 7.6].

To understand this issue better, consider the $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$-equivariant short exact sequence

$$1 \to \{\pm 1\} \to \mathrm{Spin}_{a,2}(\mathbb{C}) \to \mathrm{SO}_{a,2}(\mathbb{C}) \to 1, \tag{3.4.3.2}$$

which induces by taking Galois cohomology the long exact sequence

$$1 \to \{\pm 1\} \to \mathrm{Spin}_{a,2}(\mathbb{R}) \to \mathrm{SO}_{a,2}(\mathbb{R}) \to H^1(\mathrm{Gal}(\mathbb{C}/\mathbb{R}), \{\pm 1\}) \cong \{\pm 1\}, \tag{3.4.3.3}$$

where the homomorphism

$$\mathrm{SO}_{a,2}(\mathbb{R}) \to \{\pm 1\} \tag{3.4.3.4}$$

(induced by the connecting homomorphism) is the *spinor norm*. The kernel of the spinor norm is $\mathrm{SO}_{a,2}(\mathbb{R})^+$, which is at the same time the image of $\mathrm{Spin}_{a,2}(\mathbb{R})$.

We shall omit the verification that $\mathrm{SO}_{a,2}(\mathbb{R})^+$ does act transitively on $\mathcal{H}^+_{\mathrm{SO}_{a,2}}$, with stabilizers isomorphic to

$$\mathrm{SO}_a(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R}). \tag{3.4.3.5}$$

Accordingly, $\mathrm{SO}_{a,2}(\mathbb{R})$ acts transitively on $\mathcal{H}_{\mathrm{SO}_{a,2}}$, with stabilizers also isomorphic to $\mathrm{SO}_a(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R})$. We shall also omit the description of some bounded realizations $\mathcal{D}^+_{\mathrm{SO}_{a,2}}$ and $\mathcal{D}_{\mathrm{SO}_{a,2}}$ of $\mathcal{H}^+_{\mathrm{SO}_{a,2}}$ and $\mathcal{H}_{\mathrm{SO}_{a,2}}$, respectively.

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{H}^+_{\mathrm{SO}_{a,2}} \cong \mathcal{D}^+_{\mathrm{SO}_{a,2}}$ (with $a \geq 1$) is a Hermitian symmetric domain of type BD I (with $(p, q) = (a, 2)$ in the notation there). In Cartan's classification, it is a bounded symmetric domain of type IV_a.

3.4.4. *Interpretation as conjugacy classes.* Consider the homomorphism

$$h_0 : \mathrm{U}_1 \to \mathrm{SO}_{a,2}(\mathbb{R}) : e^{i\theta} \mapsto \begin{pmatrix} 1_a & & \\ & \cos 2\theta & -\sin 2\theta \\ & \sin 2\theta & \cos 2\theta \end{pmatrix}. \tag{3.4.4.1}$$

Note that the image of h_0 is connected because U_1 is, and hence lies in $\mathrm{SO}_{a,2}(\mathbb{R})^+$. Moreover,

(3.4.4.2)

$$K := \mathrm{Cent}_{h_0}(\mathrm{SO}_{a,2}(\mathbb{R})^+) = \left\{ \begin{pmatrix} A & \\ & D \end{pmatrix} \in \mathrm{SO}_{a,2}(\mathbb{R})^+ \right\} \cong \mathrm{SO}_a(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R}) : \begin{pmatrix} A & \\ & D \end{pmatrix} \mapsto (A, D).$$

Then K is a maximal compact subgroup of $\mathrm{SO}_{a,2}(\mathbb{R})^+$, and we have

$$\mathcal{H}^+_{\mathrm{SO}_{a,2}} = \mathrm{SO}_{a,2}(\mathbb{R})^+ \cdot h_0 \cong \mathrm{SO}_{a,2}(\mathbb{R})^+/(\mathrm{SO}_a(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R})). \tag{3.4.4.3}$$

Note that there is a coefficient 2 of θ in the above definition of h_0. This h_0 lifts to a homomorphism

$$\widetilde{h}_0 : \mathrm{U}_1 \to \mathrm{Spin}_{a,2}(\mathbb{R}), \tag{3.4.4.4}$$

where $\mathrm{Spin}_{a,2}(\mathbb{R})$ is the above double cover of $\mathrm{SO}_{a,2}(\mathbb{R})^+$, and there is an injective homomorphism

$$\mathrm{Spin}_{a,2}(\mathbb{R}) \hookrightarrow \mathrm{Sp}_{2n}(\mathbb{R}), \tag{3.4.4.5}$$

with $n = 2^a$ (cf. Examples 5.2.1.3 and 5.2.2.11 below), whose pre-composition with $\widetilde{h}_0$ lies in the conjugacy class of the analogous homomorphism for $\mathrm{Sp}_{2n}(\mathbb{R})$ in (3.1.4.1).

3.5. **The case of** E_7**.** Our main references are [Bai70], [Kim93], [GL97], and [Eie09, Ch. 12].

3.5.1. *Cayley numbers.* Consider the algebra $\mathbb{O}$ of Cayley numbers, which is an octonion (i.e., degree eight) division algebra over $\mathbb{R}$ which is normed, distributive, noncommutative, and nonassociative. We have

$$\mathbb{O} = \mathbb{R} \oplus \mathbb{R}e_1 \oplus \cdots \oplus \mathbb{R}e_7 \tag{3.5.1.1}$$

as an $\mathbb{R}$-vector space, which satisfies the multiplication rules given by

$$e_i^2 = -1 \tag{3.5.1.2}$$

and

$$(e_i e_{i+1})e_{i+3} = e_i(e_{i+1}e_{i+3}) = 1 \tag{3.5.1.3}$$

(with the indices periodically identified modulo 7) for all $1 \leq i \leq 7$. We have

$$\overline{x} := x_0 - x_1 e_1 - \cdots - x_7 e_7 \tag{3.5.1.4}$$

and

$$N(x) := x\overline{x} = x_0^2 + x_1^2 + \cdot + x_7^2 \geq 0 \tag{3.5.1.5}$$

for each

$$x = x_0 + x_1 e_1 + \cdots + x_7 e_7 \in \mathbb{O}, \tag{3.5.1.6}$$

where the square-root of $N(x)$ defines the norm of x.

3.5.2. *An exceptional Jordan algebra over* $\mathbb{R}$. Consider the *Albert algebra*

$$\mathfrak{A} := \mathrm{Herm}_3(\mathbb{O}), \tag{3.5.2.1}$$

the 3×3 Hermitian matrices over $\mathbb{O}$, which is a formally real Jordan algebra of exceptional type, with the Jordan algebra multiplication given by

$$X \cdot Y = \tfrac{1}{2}(XY + YX), \tag{3.5.2.2}$$

where XY and YX are the usual matrix multiplications. (See, for example, [FK95, Ch. II–V and VIII] for a general introduction to Jordan algebras and the classification of formally real ones. See also [AMRT75, Ch. II, Rem. at the end of Sec. 1] or [AMRT10, Ch. II, Rem. 1.11], and the sections following there, for why they matter.)

For

$$X = \begin{pmatrix} a & x & y \\ \overline{x} & b & z \\ \overline{y} & \overline{z} & c \end{pmatrix} \in \mathfrak{A} \tag{3.5.2.3}$$

(with $a, b, c \in \mathbb{R}$ and $x, y, z \in \mathbb{O}$), we can define the trace

$$\mathrm{tr}(X) = a + b + c \tag{3.5.2.4}$$

and the determinant

$$\det(X) = abc - aN(z) - bN(y) - cN(x) + \mathrm{tr}((xz)\overline{y}). \tag{3.5.2.5}$$

Then we also have an inner product

$$(X, Y) := \mathrm{tr}(X \cdot Y) \tag{3.5.2.6}$$

and a trilinear form $(\,\cdot\,,\,\cdot\,,\,\cdot\,)$ satisfying the identity

$$(X, X, X) = \det(X). \tag{3.5.2.7}$$

This trilinear form defines a cross product $X \times Y$ such that

$$(X \times Y, Z) = 3(X, Y, Z), \tag{3.5.2.8}$$

for $X, Y, Z \in \mathfrak{A}$, which specializes to

$$X \times X = \begin{pmatrix} bc - N(z) & y\overline{z} - cx & xz - by \\ z\overline{y} - c\overline{x} & ac - N(y) & \overline{x}y - az \\ \overline{z}\overline{x} - b\overline{y} & \overline{y}x - a\overline{z} & ab - N(x) \end{pmatrix} \tag{3.5.2.9}$$

when $X = Y$.

Since $\mathbb{O}$ is nonassociative, it is rather a miracle that the above definitions can be made for 3×3 Hermitian matrices over $\mathbb{O}$—nothing similar works for larger matrices.

3.5.3. *A real Lie group with Lie algebra* $\mathfrak{e}_{7(-25)}$. Consider the Lie group

$$M := \{(g, r) \in \mathrm{GL}(\mathfrak{A}) \times \mathbb{R}^\times : \det(gX) = r \det(X)\}, \tag{3.5.3.1}$$

so that we have a homomorphism

$$\nu : M \to \mathbb{R}^\times : (g, r) \mapsto r. \tag{3.5.3.2}$$

Then we have

$$\ker(\nu) = \{g \in \mathrm{GL}(\mathfrak{A}) : \det(gX) = \det(X)\}. \tag{3.5.3.3}$$

It is known that $\ker(\nu)$ is a connected semisimple real Lie group of real rank 2 with Lie algebra isomorphic to $\mathfrak{e}_{6(-26)}$. The number -26 in the notation of $\mathfrak{e}_{6(-26)}$

means it is the real form of the complex simple Lie algebra of type E_6 such that the Killing form has signature (a, b) with $a-b = -26$; or, equivalent, that every Cartan involution has trace $26 = -(-26)$. For complex simple Lie algebras of exceptional type, such numbers already classify their real forms up to isomorphism. (The same explanation applies to other real Lie algebras of exceptional type below.) Also, the center of M is

$$Z(M) \cong \mathbb{R}^\times \tag{3.5.3.4}$$

with $x \in \mathbb{R}^\times$ acting as multiplication by the scalar x on $\mathfrak{A}$, so that $\nu(x) = x^3$. In this case, the Lie algebra of M is the direct sum of those of $Z(M)$ and $\ker(\nu)$, and $\ker(\nu)$ is the derived group of M. The $\ker(\nu)$-orbit of 1 in $\mathfrak{A}$ is isomorphic to the quotient of $\ker(\nu)$ by a maximal compact subgroup with Lie algebra isomorphic to $\mathfrak{f}_4$. In the classification in [Hel01, Ch. X, Sec. 6], this orbit is a noncompact Riemannian symmetric space of type E IV, which is not Hermitian.

Moreover, $g \in M$ induces $g^* \in M$ by

$$(gX, g^*Y) = (X, Y). \tag{3.5.3.5}$$

Then M acts on

$$\mathfrak{W} := \mathfrak{A} \oplus \mathbb{R} \oplus \mathfrak{A} \oplus \mathbb{R} \tag{3.5.3.6}$$

by

$$g(X, \varepsilon, X', \varepsilon') := (gX, \nu(g)^{-1}\varepsilon, g^*X', \nu(g)\varepsilon'). \tag{3.5.3.7}$$

There is also an additive action

$$\rho : \mathfrak{A} \hookrightarrow \mathrm{GL}(\mathfrak{W}) : B \mapsto \rho_B \tag{3.5.3.8}$$

with ρ_B defined by

$$\rho_B \begin{pmatrix} X \\ \varepsilon \\ X' \\ \varepsilon' \end{pmatrix} := \begin{pmatrix} X + 2B \times X' + \varepsilon B \times B \\ \varepsilon \\ X' + \varepsilon B \\ \varepsilon' + (B, X) + (B \times B, X') + \varepsilon \det(B) \end{pmatrix}, \tag{3.5.3.9}$$

with elements of $\mathfrak{W}$ represented by column vectors (cf. (3.5.3.6)).

Consider the Lie group

$$U := \mathrm{image}(\rho). \tag{3.5.3.10}$$

It is known that M normalizes U. Let $\iota \in \mathrm{GL}(\mathfrak{W})$ be defined by

$$\iota(X, \varepsilon, X', \varepsilon') := (-X', -\varepsilon', X, \varepsilon). \tag{3.5.3.11}$$

Consider the Lie groups

$$P := M \ltimes U \tag{3.5.3.12}$$

and

$$G := \langle \iota, P \rangle \tag{3.5.3.13}$$

(i.e., G is generated by ι and P) as subgroups of $\mathrm{GL}(\mathfrak{W})$. It is known that G is a connected semisimple Lie group of real rank 3 with Lie algebra isomorphic to $\mathfrak{e}_{7(-25)}$, and that P is the maximal parabolic subgroup of G stabilizing the subspace $\{(0, 0, 0, \varepsilon') : \varepsilon' \in \mathbb{R}\}$ of $\mathfrak{W}$.

These definitions might seem rather complicated and unmotivated, but it is worth drawing the analogy of the element ι and the subgroups U, M, and P of

G with the corresponding element $\begin{pmatrix} & 1_n \\ -1_n & \end{pmatrix}$ and the subgroups $\left\{\begin{pmatrix} 1_n & * \\ & 1_n \end{pmatrix}\right\}$, $\left\{\begin{pmatrix} * & \\ & * \end{pmatrix}\right\}$, and $\left\{\begin{pmatrix} * & * \\ & * \end{pmatrix}\right\}$ of $\mathrm{Sp}_{2n}(\mathbb{R})$.

A more systematic definition is that G is the Lie subgroup of $\mathrm{GL}(\mathfrak{W})$ consisting of elements preserving both the quartic form

$$Q(X, \varepsilon, X', \varepsilon') := (X \times X, X' \times X') \\ - \varepsilon \det(X) - \varepsilon' \det(X') - \tfrac{1}{4}((X, X') - \varepsilon\varepsilon')^2 \tag{3.5.3.14}$$

and the skew-symmetric bilinear form

$$\langle (X_1, \varepsilon_1, X'_1, \varepsilon'_1), (X_2, \varepsilon_2, X'_2, \varepsilon'_2) \rangle := (X_1, X'_2) - (X_2, X'_1) + \varepsilon_1\varepsilon'_2 - \varepsilon_2\varepsilon'_1 \tag{3.5.3.15}$$

on $\mathfrak{W}$. But, as we shall see, it is useful to know the explicit facts (3.5.3.13) and (3.5.3.12).

3.5.4. *An octonion upper half-space.* Consider

$$\mathcal{H}_{\mathrm{E}_7} := \{Z \in \mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C} : \mathrm{Im}(Z) > 0\}, \tag{3.5.4.1}$$

where $\mathrm{Im}(Z) > 0$ means the left multiplication of $\mathrm{Im}(Z)$ on $\mathfrak{A}$ induces a positive definite linear transformation, or equivalently that $\mathrm{Im}(Z)$ lies in the interior of the set $\{X^2 : X \in \mathfrak{A}\}$ with its natural real analytic topology (see [FK95, Thm. III.2.1]). (For an analogy, consider the case of symmetric matrices over $\mathbb{R}$, where positive definite matrices are exactly those matrices in the interior of the set of squares of symmetric matrices with its natural real analytic topology.) This is the most exotic generalization of the various classical upper half-spaces in Sections 3.1.1, 3.2.3, and 3.3.2 defined by Hermitian matrices over $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$, respectively. Together with the "semi-classical" examples in Section 3.4.2, these are all the irreducible examples (namely, ones that are not isomorphic to nontrivial products of smaller ones) of the so-called *tube domains*. They are all of the form

$$\{X + iY : X, Y \in \mathfrak{J},\, Y > 0\} \subset \mathfrak{J} \underset{\mathbb{R}}{\otimes} \mathbb{C}, \tag{3.5.4.2}$$

where $\mathfrak{J}$ is a formally real Jordan algebra (see [FK95, Ch. VIII–X]; see also [AMRT75, Ch. III, Sec. 1] or [AMRT10, Ch. III, Sec. 1]), where the symbol i denotes the action of $1 \otimes i$ on $\mathfrak{J} \underset{\mathbb{R}}{\otimes} \mathbb{C}$. In the definition of $\mathcal{H}_{\mathrm{E}_7}$, we have $\mathfrak{J} = \mathfrak{A} = \mathrm{Herm}_3(\mathbb{O})$.

Note that

$$\dim_{\mathbb{C}} \mathcal{H}_{\mathrm{E}_7} = \dim_{\mathbb{C}}(\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}) = \dim_{\mathbb{R}}(\mathfrak{A}) = 27. \tag{3.5.4.3}$$

The group U acts on $\mathcal{H}_{\mathrm{E}_7}$ by the translation

$$Z \mapsto Z + B, \tag{3.5.4.4}$$

for each $\rho_B \in U$ associated with $B \in \mathfrak{A}$. The group M acts on $\mathcal{H}_{\mathrm{E}_7}$ by its natural action on $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$ (up to a sign convention), which preserves the subset $\mathcal{H}_{\mathrm{E}_7}$. The element $\iota \in G$ acts on $\mathcal{H}_{\mathrm{E}_7}$ by

$$\iota(Z) := -Z^{-1}, \tag{3.5.4.5}$$

where Z^{-1} is defined by inverting the Jordan algebra multiplication in $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$. One can check that these actions are compatible with each other, and define an action of G on $\mathcal{H}_{\mathrm{E}_7}$.

Given any $Z = X + iY \in \mathcal{H}_{\mathrm{E}_7}$, where $X, Y \in \mathfrak{A}$ and where the symbol i denotes the action of $1 \otimes i$ on $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$, we have

$$\rho_{-X}(Z) = iY \tag{3.5.4.6}$$

for some $Y > 0$. By Freudenthal's theorem (which intriguingly asserts that matrices in $\mathrm{Herm}_3(\mathbb{O})$ can also be *diagonalized*; see [FK95, Thm. V.2.5]), there exists some $g \in M$ such that

$$gY = 1_3, \tag{3.5.4.7}$$

and so that

$$g(iY) = i1_3. \tag{3.5.4.8}$$

This shows that

$$\mathcal{H}_{\mathrm{E}_7} = G \cdot i1_3 \tag{3.5.4.9}$$

and so that the action of G on $\mathcal{H}_{\mathrm{E}_7}$ is transitive.

Let K denote the stabilizer of $i1_3$ in G. Then

$$K = \mathrm{Cent}_\iota(G) \tag{3.5.4.10}$$

as subgroups of G, which is a maximal compact subgroup of G and is a connected compact Lie group with Lie algebra $\mathfrak{e}_6 \oplus \mathbb{R}$. Consider the homomorphism

$$h_0 : \mathrm{U}_1 \to G \tag{3.5.4.11}$$

defined in block matrix form by setting

$$h_0(\cos\theta + i\sin\theta) = \begin{pmatrix} \cos\theta & & -\sin\theta & \\ & \cos 3\theta & & \sin 3\theta \\ \sin\theta & & \cos\theta & \\ & -\sin 3\theta & & \cos 3\theta \end{pmatrix} \in \mathrm{GL}(\mathfrak{W}) \tag{3.5.4.12}$$

(by left multiplication on column vectors representing elements in $\mathfrak{W}$; cf. (3.5.3.6)). In particular, $h_0(i) = \iota$, so that we also have

$$\mathcal{H}_{\mathrm{E}_7} = G \cdot h_0 \cong G/K. \tag{3.5.4.13}$$

There is no homomorphism of the form $G \to \mathrm{Sp}_{2n}(\mathbb{R})$, for any $n \geq 1$, whose pre-composition with $h_0 : \mathrm{U}_1 \to G$ lies in the conjugacy class of the analogous homomorphism for $\mathrm{Sp}_{2n}(\mathbb{R})$ in (3.1.4.1) (see [Del79, 1.3.10(i)] or [Mil13, Sec. 10, p. 531]; cf. the end of Section 5.2.2 below).

3.5.5. *Bounded realization.* There is also a bounded realization $\mathcal{D}_{\mathrm{E}_7}$ of $\mathcal{H}_{\mathrm{E}_7}$, also with coordinates in $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$, with a change of coordinates

$$\mathcal{H}_{\mathrm{E}_7} \xrightarrow{\sim} \mathcal{D}_{\mathrm{E}_7} : Z \mapsto (Z - i1_3)(Z + i1_3)^{-1}, \tag{3.5.5.1}$$

where $(Z + i1_3)^{-1}$ is (as before) defined by inverting the Jordan algebra multiplication in $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$, which is yet another generalization of the classical Cayley transform (cf. Sections 3.1.3 and 3.2.4).

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{H}_{\mathrm{E}_7} \cong \mathcal{D}_{\mathrm{E}_7}$ is a Hermitian symmetric domain of type E VII. In Cartan's classification, it is a bounded symmetric domain of type VI.

3.6. **The case of** E_6.

3.6.1. *Bounded and unbounded realizations.* There is (up to isomorphism) just one irreducible Hermitian symmetric domain that we have not discussed yet. (Again, being irreducible means not isomorphic to a nontrivial products of smaller ones.) Let us start with a bounded realization, given by

$$\mathcal{D}_{\mathrm{E}_6} := \mathcal{D}_{\mathrm{E}_7} \cap \left\{ \begin{pmatrix} 0 & x & y \\ \overline{x} & 0 & 0 \\ \overline{y} & 0 & 0 \end{pmatrix} : x, y \in \mathbb{O} \underset{\mathbb{R}}{\otimes} \mathbb{C} \right\} \tag{3.6.1.1}$$

in $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$ (see Section 3.5.5), where $\overline{x}$ and $\overline{y}$ are the conjugations induced by the ones in $\mathbb{O}$, which can be embedded as an open submanifold of $\mathbb{O}^2 \underset{\mathbb{R}}{\otimes} \mathbb{C}$. Let us also let $\mathcal{H}_{\mathrm{E}_6}$ denote the preimage of $\mathcal{D}_{\mathrm{E}_6}$ under the isomorphism (3.5.5.1), which is then an unbounded realization.

Essentially by construction, we have

$$\dim_{\mathbb{C}} \mathcal{H}_{\mathrm{E}_6} = \dim_{\mathbb{C}} \mathcal{D}_{\mathrm{E}_6} = 2 \dim_{\mathbb{C}}(\mathbb{O} \underset{\mathbb{R}}{\otimes} \mathbb{C}) = 2 \dim_{\mathbb{R}} \mathbb{O} = 16. \tag{3.6.1.2}$$

By [Hir66, Sec. 3], $\mathcal{D}_{\mathrm{E}_6}$ and hence $\mathcal{H}_{\mathrm{E}_6}$ admit transitive actions of a connected semisimple Lie group G of rank 2 with Lie algebra isomorphic to $\mathfrak{e}_{6(-14)}$, with stabilizers given by maximal compact subgroups of G which are connected Lie groups with Lie algebras isomorphic to $\mathfrak{so}_{10} \oplus \mathbb{R}$.

In the classification in [Hel01, Ch. X, Sec. 6], $\mathcal{H}_{\mathrm{E}_6} \cong \mathcal{D}_{\mathrm{E}_6}$ is a Hermitian symmetric domain of type E III. In Cartan's classification, it is a bounded symmetric domain of type V.

3.6.2. *A real Lie group of Lie algebra* $\mathfrak{e}_{6(-14)}$. It is not easy to describe $\mathcal{D}_{\mathrm{E}_6}$ and $\mathcal{H}_{\mathrm{E}_6}$, together with the actions of certain G as above, in explicit coordinates. Nevertheless, let us define a real Lie group G with Lie algebra $\mathfrak{e}_{6(-14)}$, which can be given the structure of the $\mathbb{R}$-points of an algebraic group over $\mathbb{Q}$, and which acts on $\mathcal{D}_{\mathrm{E}_6}$ and $\mathcal{H}_{\mathrm{E}_6}$ by the abstract theory of Hermitian symmetric domains.

Consider the pairing

$$\langle X, Y \rangle := (X^c, Y), \tag{3.6.2.1}$$

for $X, Y \in \mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$, where c denotes the complex conjugation with respect to the tensor factor $\mathbb{C}$; consider the involution σ on $\mathfrak{A} = \mathrm{Herm}_3(\mathbb{O})$ (which extends linearly to $\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}$) defined by

$$\begin{pmatrix} a & x & y \\ \overline{x} & b & z \\ \overline{y} & \overline{z} & c \end{pmatrix}^{\sigma} = \begin{pmatrix} a & -x & -y \\ -\overline{x} & b & z \\ -\overline{y} & \overline{z} & c \end{pmatrix}; \tag{3.6.2.2}$$

and consider

$$G := \{ g \in \mathrm{GL}(\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}) : \det(gX) = \det(X),\ \langle gX, gY \rangle_{\sigma} = \langle X, Y \rangle_{\sigma} \}, \tag{3.6.2.3}$$

where

$$\langle X, Y \rangle_{\sigma} := \langle X^{\sigma}, Y \rangle. \tag{3.6.2.4}$$

We note that

$$\{g \in \mathrm{GL}(\mathfrak{A} \underset{\mathbb{R}}{\otimes} \mathbb{C}) : \det(gX) = \det(X)\} \tag{3.6.2.5}$$

is a complex Lie group with Lie algebra $\mathfrak{e}_6(\mathbb{C})$, and so the upshot is that the condition

$$\langle gX, gY \rangle_\sigma = \langle X, Y \rangle_\sigma \tag{3.6.2.6}$$

defines a real form of rank 2 of this complex Lie group with Lie algebra isomorphic to $\mathfrak{e}_{6(-14)}$, which is not isomorphic to either its compact form (see (3.5.4.10) and the remark following it) or the noncompact real form

$$\{g \in \mathrm{GL}(\mathfrak{A}) : \det(gX) = \det(X)\} \tag{3.6.2.7}$$

(see (3.5.3.3) and the remark following it) whose Lie algebra is isomorphic to $\mathfrak{e}_{6(-26)}$. (Thus far, we have seen three kinds of real Lie groups of type E_6 in our examples, with one compact and the other two of real rank 2.)

3.6.3. *Interpretation as conjugacy classes.* Consider the homomorphism

$$h_0 : \mathrm{U}_1 \to G \tag{3.6.3.1}$$

defined by

$$h_0(\alpha) \begin{pmatrix} a & x & y \\ \overline{x} & b & z \\ \overline{y} & \overline{z} & c \end{pmatrix} = \begin{pmatrix} \alpha^4 a & \alpha x & \alpha y \\ \alpha \overline{x} & \alpha^{-2} b & \alpha^{-2} z \\ \alpha \overline{y} & \alpha^{-2} \overline{z} & \alpha^{-2} c \end{pmatrix}, \tag{3.6.3.2}$$

with α acting as $1 \otimes \alpha$ on $\mathbb{O} \underset{\mathbb{Z}}{\otimes} \mathbb{C}$. Then it can be checked directly that

$$K := \mathrm{Cent}_{h_0}(\mathrm{G}), \tag{3.6.3.3}$$

is a maximal compact subgroup of G, which is a connected Lie group with Lie algebra isomorphic to $\mathfrak{so}_{10} \oplus \mathbb{R}$, so that we have

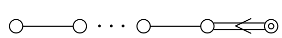
$$\mathcal{H}_{\mathrm{E}_6} = G \cdot h_0 \cong G/K. \tag{3.6.3.4}$$

There is no homomorphism of the form $G \to \mathrm{Sp}_{2n}(\mathbb{R})$, for some $n \geq 1$, whose pre-composition with $h_0 : \mathrm{U}_1 \to G$ lies in the conjugacy class of the analogous homomorphism for $\mathrm{Sp}_{2n}(\mathbb{R})$ in (3.1.4.1) (see [Del79, 1.3.10(i)] or [Mil13, Sec. 10, p. 530]; cf. the end of Section 5.2.2 below).

3.7. **A brief summary.** Let us summarize the groups and symmetric domains we have considered as follows. Let us also include the information of the corresponding *special node* on each Dynkin diagram, marked with a unique double circle on the diagram. Such special nodes have not been explained, but should be useful for the readers when they consult more advanced texts.

(1) $\mathrm{G}(\mathbb{R}) = \mathrm{Sp}_{2n}(\mathbb{R})$, $K \cong \mathrm{U}_n$, $\mathcal{D} = \mathcal{H}_n$:

(3.7.1)

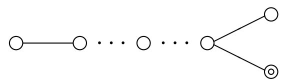


(2) $\mathrm{G}(\mathbb{R}) = \mathrm{U}_{a,b}$, $K \cong \mathrm{U}_a \times \mathrm{U}_b$, $\mathcal{D} = \mathcal{H}_{a,b}$:

(3.7.2)

Here the special node is the b-th from the left, or alternatively, from the right, because it is only up to automorphisms of the Dynkin diagram.

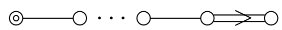
(3) $\mathrm{G}(\mathbb{R}) = \mathrm{SO}^*_{2n}$, $K \cong \mathrm{U}_n$, $\mathcal{D} = \mathcal{H}_{\mathrm{SO}^*_{2n}}$, with $n \geq 2$:

(3.7.3)


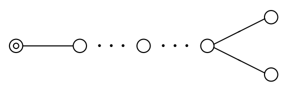
Again, the special node is only up to automorphisms of the Dynkin diagram. Therefore, it can be, alternatively, the other right-most node.

(4) $\mathrm{G}(\mathbb{R})^+ = \mathrm{SO}_{a,2}(\mathbb{R})^+$, $K \cong \mathrm{SO}_a(\mathbb{R}) \times \mathrm{SO}_2(\mathbb{R})$, $\mathcal{D}^+ = \mathcal{H}^+_{\mathrm{SO}_{a,2}}$, with $a \geq 1$ but $a \neq 2$:

(a) $a = 2n - 1$ is odd, with $n \geq 1$:

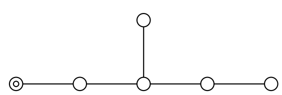
(3.7.4)


(b) $a = 2n - 2$ is even, with $n \geq 3$:

(3.7.5)


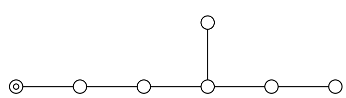
(In the excluded case $a = 2$, we have $\mathrm{Spin}_{2,2}(\mathbb{R}) \cong \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$, and $\mathcal{H}^+_{\mathrm{SO}_{2,2}} \cong \mathcal{H} \times \mathcal{H}$ is not irreducible.)

(5) $\mathrm{Lie}\, G \cong \mathfrak{e}_{6(-14)}$, $\mathrm{Lie}\, K \cong \mathfrak{so}_{10} \oplus \mathbb{R}$, $\mathcal{D} = \mathcal{H}_{\mathrm{E}_6}$:

(3.7.6)


Yet again, the special node is only up to automorphisms of the Dynkin diagram. Therefore, it can be, alternatively, the right-most node.

(6) $\mathrm{Lie}\, G \cong \mathfrak{e}_{7(-25)}$, $\mathrm{Lie}\, K \cong \mathfrak{e}_6 \oplus \mathbb{R}$, $\mathcal{D} = \mathcal{H}_{\mathrm{E}_7}$:

(3.7.7)


These exhaust all the possibilities of irreducible Hermitian symmetric domains.

In the cases (1), (2), and (3), there exist injective homomorphisms of the form $\mathrm{G}(\mathbb{R}) \to \mathrm{Sp}_{2n'}(\mathbb{R})$, for some $n' \geq 1$, whose pre-composition with $h_0 : \mathrm{U}_1 \to \mathrm{G}(\mathbb{R})$ lies in the conjugacy class of the analogous homomorphism for $\mathrm{Sp}_{2n'}(\mathbb{R})$ as in (3.1.4.1). In the case (4), the analogous assertion is true after replacing $\mathrm{SO}_{a,2}(\mathbb{R})^+$ with its double cover $\mathrm{Spin}_{a,2}(\mathbb{R})$. In the cases (5) and (6), the analogous assertion is simply false. (In Section 5.2.2 below, we will encounter this issue again in the classification of abelian-type Shimura data, although the question there is more subtle because we need to also consider the structure of algebraic groups over $\mathbb{Q}$.)

4. Rational boundary components and compactifications

In this section, we will explain how the double coset spaces

$$\mathsf{X}_{\mathcal{U}} = \mathrm{G}(\mathbb{Q}) \backslash (\mathcal{D} \times \mathrm{G}(\mathbb{A}^\infty)) / \mathcal{U} \cong \coprod_{i \in I} \Gamma_i \backslash \mathcal{D}^+$$

as in (2.2.15) and (2.2.21) are compactified and shown to be algebraic varieties in cases where $\mathcal{D}$ are Hermitian symmetric domains or their finite unions.

4.1. **Examples of rational boundary components.** Let us start by explaining the notion of *rational boundary components* in many examples, based on the Hermitian symmetric domains we have seen in Section 3. We will begin with the cases with $\mathcal{D} = \mathcal{H}$ in Section 4.1.1, followed by the cases with $\mathcal{D} = \mathcal{H} \times \mathcal{H}$ in Section 4.1.2; the cases with $\mathcal{D} = \mathcal{H}_n$ in Section 4.1.3; the cases with $\mathcal{D} = \mathcal{H}_{a,b}$, with $a \geq b$, in Section 4.1.4; the cases with $\mathcal{D} = \mathcal{H}_{\mathrm{SO}^*_{2n}}$ in Section 4.1.5; the cases with $\mathcal{D}^+ = \mathcal{H}^+_{\mathrm{SO}_{a,2}}$, with $a \geq 2$, in Section 4.1.6; and the cases with $\mathcal{D} = \mathcal{H}_{\mathrm{E}_6}$ or $\mathcal{D} = \mathcal{H}_{\mathrm{E}_7}$ in Section 4.1.7.

4.1.1. *Cases with $\mathcal{D} = \mathcal{H}$.* Suppose $\mathrm{G} = \mathrm{SL}_2$, so that we have $\mathcal{D} = \mathcal{H}$ as in Examples 2.2.7 and 2.2.23, and we have to consider quotients of the form $\Gamma\backslash\mathcal{D}$, where Γ is an arithmetic subgroup of $\mathrm{SL}_2(\mathbb{Q})$. Note that the action of $\mathrm{SL}_2(\mathbb{Q})$ preserves

$$\mathcal{H}^* := \mathcal{H} \cup \mathbb{P}^1(\mathbb{Q}) \tag{4.1.1.1}$$

as a *subset* of $\mathbb{P}^1(\mathbb{C})$.

In the natural topology of $\mathbb{P}^1(\mathbb{C})$, the boundary of $\mathcal{H}$ is $\mathbb{P}^1(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R}) \cdot \infty$, where ∞ means just $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in projective coordinates. The reason to consider instead the rational subset $\mathbb{P}^1(\mathbb{Q}) = \mathrm{SL}_2(\mathbb{Q}) \cdot \infty$ is the *reduction theory*: The fundamental domain of $\mathcal{H}$ under the action of Γ can be compactified by adding only points in $\mathbb{P}^1(\mathbb{Q}) = \mathrm{SL}_2(\mathbb{Q}) \cdot \infty$. The irrational points are completely irrelevant. Therefore, by topologizing $\mathcal{H}^*$ at the points of $\mathbb{P}^1(\mathbb{Q})$ using open discs bounded by the so-called *horocycles*, namely, circles tangent to $\mathbb{P}^1(\mathbb{R})$ at the points of $\mathbb{P}^1(\mathbb{Q})$ (see [Shi71, Sec. 1.3]), we obtain a compactification

$$\Gamma\backslash\mathcal{H} \hookrightarrow \Gamma\backslash\mathcal{H}^* \tag{4.1.1.2}$$

such that $\Gamma\backslash\mathcal{H}^*$ has the structure of a compact Riemann surface (or, in other words, a complete algebraic curve over $\mathbb{C}$), with finitely many *boundary points*

$$\Gamma\backslash\mathcal{H}^* - \Gamma\backslash\mathcal{H} \cong \Gamma\backslash\mathbb{P}^1(\mathbb{Q}) \tag{4.1.1.3}$$

called *cusps*. (Here the finiteness of $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ can be deduced from the elementary fact that $\mathbb{P}^1(\mathbb{Q}) = \mathrm{SL}_2(\mathbb{Q}) \cdot \infty = \mathrm{SL}_2(\mathbb{Z}) \cdot \infty$.) The points of $\mathbb{P}^1(\mathbb{Q}) = \mathrm{SL}_2(\mathbb{Q}) \cdot \infty$ are the so-called *rational boundary components* of $\mathcal{H}$, which are prototypical examples of their analogues for more general Hermitian symmetric domains.

Note that the notions of *cusps* and *rational boundary components* depend on the structure of $\mathrm{G} = \mathrm{SL}_2$ as an algebraic group over $\mathbb{Q}$ in an essential way. There exist other algebraic groups G' over $\mathbb{Q}$ such that $\mathrm{G}'(\mathbb{R}) \cong \mathrm{G}(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$ and such that the fundamental domains for the actions of the corresponding arithmetic subgroups Γ' of $\mathrm{G}'(\mathbb{Q})$ on $\mathcal{D}$ are already compact, in which case there is no need to add any cusps (cf. the example of *Shimura curves* in Section 5.2.2 below).

4.1.2. *Cases with $\mathcal{D} = \mathcal{H} \times \mathcal{H}$.* There are two important cases of G with $\mathrm{G}(\mathbb{R})$ acting on $\mathcal{D} = \mathcal{H} \times \mathcal{H}$.

The first G is simply $\mathrm{SL}_2 \times \mathrm{SL}_2$. Then we should take

(4.1.2.1)

$$\begin{aligned} \mathcal{D}^* &:= \mathcal{H}^* \times \mathcal{H}^* \\ &= (\mathcal{H} \times \mathcal{H}) \cup \big(\mathrm{G}(\mathbb{Q}) \cdot (\mathcal{H} \times \infty)\big) \cup \big(\mathrm{G}(\mathbb{Q}) \cdot (\infty \times \mathcal{H})\big) \cup \big(\mathrm{G}(\mathbb{Q}) \cdot (\infty \times \infty)\big) \end{aligned}$$

as subsets of $\mathbb{P}^1(\mathbb{C}) \times \mathbb{P}^1(\mathbb{C})$. There are three $\mathrm{G}(\mathbb{Q})$-orbits of rational boundary components (which are those orbits other than $\mathcal{H} \times \mathcal{H}$).

The other G is the *restriction of scalars* $\mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{SL}_2$, for some real quadratic extension field F of $\mathbb{Q}$. It is the group scheme G over $\mathbb{Q}$ whose R-points are

$$\mathrm{G}(R) = \mathrm{SL}_2(R \underset{\mathbb{Q}}{\otimes} F), \tag{4.1.2.2}$$

for each $\mathbb{Q}$-algebra R. In particular, we have $\mathrm{G}(\mathbb{Q}) = \mathrm{SL}_2(F)$. This is the case of *Hilbert modular surfaces*, which are generalized to higher-dimensional *Hilbert modular varieties* by replacing F with totally real extension fields of $\mathbb{Q}$ of higher degrees. Then we should take

$$\mathcal{D}^* := (\mathcal{H}\times\mathcal{H}) \cup \big(\mathrm{G}(\mathbb{Q})\cdot(\infty\times\infty)\big), \tag{4.1.2.3}$$

with only one $\mathrm{G}(\mathbb{Q})$-orbit of rational boundary components (which is the unique orbit other than $\mathcal{H}\times\mathcal{H}$).

In both cases, we can topologize $\mathcal{D}^*$ such that $\Gamma\backslash\mathcal{D}^*$ is a compactification of $\Gamma\backslash\mathcal{D}$ for each arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$. When G is simply $\mathrm{SL}_2\times\mathrm{SL}_2$, this is just doubling the case where $\mathrm{G} = \mathrm{SL}_2$. When G is $\mathrm{Res}_{F/\mathbb{Q}}\,\mathrm{SL}_2$, this is essentially proved in [Shi63a], in a different language. This generalizes the case where $\mathrm{G} = \mathrm{SL}_2$ but will be superseded by the introduction of *Satake topology* in general.

Here are some general principles, which are valid for not only these two examples, but also for the forthcoming ones:

(1) The rational boundary components are in one-one correspondence with (rational) parabolic subgroups P of the algebraic group G over $\mathbb{Q}$ whose pullback to every $\mathbb{Q}$-simple (almost) factor G' of G is either the whole G' or a maximal proper parabolic subgroup P'. When the group G is $\mathbb{Q}$-simple, we will just say that such a P is a maximal parabolic subgroup.

(2) For *classical groups* G defined by nondegenerate pairings over finite-dimensional semisimple algebras over $\mathbb{Q}$, the (rational) parabolic subgroups of G are always stabilizers of increasing sequences of totally isotropic subspaces.

4.1.3. *Cases with $\mathcal{D} = \mathcal{H}_n$.* Suppose $\mathrm{G} = \mathrm{Sp}_{2n}$, so that $\mathrm{G}(\mathbb{R})$ acts on $\mathcal{H}_n$ as in Section 3.1.1.

When $n = 2$, there are three $\mathrm{G}(\mathbb{Q})$-orbits of proper parabolic subgroups of $\mathrm{G}(\mathbb{Q})$, with the following representatives, respectively: (We warn the readers that, although we will write down only the groups of $\mathbb{Q}$-points, we will still be working with algebraic groups over $\mathbb{Q}$. This is a somewhat misleading abuse of language, which we adopt only for simplicity of exposition.)

(1) A *Borel subgroup*

$$\mathrm{B}(\mathbb{Q}) = \left\{ \begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & \\ & & * & * \end{pmatrix} \right\} \tag{4.1.3.1}$$

(which is a minimal parabolic subgroup) stabilizing the increasing sequence

$$\left\{ \begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \subset \left\{ \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix} \right\} \tag{4.1.3.2}$$

of totally isotropic subspaces of $\mathbb{Q}^4$. Note that we have

$$\left\{\begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\}^{\perp} = \left\{\begin{pmatrix} * \\ * \\ 0 \\ * \end{pmatrix}\right\} \tag{4.1.3.3}$$

and

$$\left\{\begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}\right\}^{\perp} = \left\{\begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}\right\} \tag{4.1.3.4}$$

with respect to the alternating pairing

$$\langle \cdot\,, \cdot \rangle : \mathbb{Q}^4 \times \mathbb{Q}^4 \to \mathbb{Q} \tag{4.1.3.5}$$

defined by $J_2 = \begin{pmatrix} & 1_2 \\ -1_2 & \end{pmatrix}$, and so we can complete the above sequence (4.1.3.2) into a flag

$$\left\{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\} \subset \left\{\begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\} \subset \left\{\begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}\right\} \subset \left\{\begin{pmatrix} * \\ * \\ 0 \\ * \end{pmatrix}\right\} \subset \left\{\begin{pmatrix} * \\ * \\ * \\ * \end{pmatrix}\right\} \tag{4.1.3.6}$$

in $\mathbb{Q}^4$. This also explains why the entries of B are not exactly upper-triangular. Sometimes it is convenient to work with upper-triangular matrices—then one should replace J_2 with a matrix that is indeed anti-diagonal, not just block anti-diagonal, as remarked in Example 2.1.2.

(2) The *Klingen parabolic subgroup*

$$\mathrm{P}^{(1)}(\mathbb{Q}) = \left\{\begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & \\ & * & * & * \end{pmatrix}\right\} \tag{4.1.3.7}$$

(which is a maximal parabolic subgroup) stabilizing the totally isotropic subspace

$$\left\{\begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\} \tag{4.1.3.8}$$

of $\mathbb{Q}^4$.

(3) The *Siegel parabolic subgroup*

$$\mathrm{P}^{(2)}(\mathbb{Q}) = \left\{\begin{pmatrix} * & * & * & * \\ * & * & * & * \\ & & * & * \\ & & * & * \end{pmatrix}\right\} \tag{4.1.3.9}$$

(which is a maximal parabolic subgroup) stabilizing the totally isotropic subspace

$$\left\{\begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}\right\} \tag{4.1.3.10}$$

of $\mathbb{Q}^4$.

The Borel subgroup B is not maximal, and hence will not be needed in the consideration of rational boundary components. Nevertheless, we have

$$\mathrm{B}(\mathbb{Q}) = \mathrm{P}^{(1)}(\mathbb{Q}) \cap \mathrm{P}^{(2)}(\mathbb{Q}) \tag{4.1.3.11}$$

by definition, and $\mathrm{P}^{(1)}(\mathbb{Q})$ and $\mathrm{P}^{(2)}(\mathbb{Q})$ will both be needed.

Now that we have seen the maximal parabolic subgroups when $n = 2$, let us consider a more general n. The rational boundary components of $\mathcal{H}_n$ are given by elements of the $\mathrm{G}(\mathbb{Q})$-orbits of

$$\begin{pmatrix} \infty_r & \\ & \mathcal{H}_{n-r} \end{pmatrix} \cong \mathcal{H}_{n-r}, \tag{4.1.3.12}$$

which means

$$\left\{\begin{pmatrix} 1_r & & \\ & & Z \\ 0_r & & \\ & & 1_{n-r} \end{pmatrix} : Z \in \mathcal{H}_{n-r}\right\} \tag{4.1.3.13}$$

in generalized projective coordinates (cf. Remark 3.1.1.4), for $r = 1, \ldots, n$. (When $r = 0$, we just get $\mathcal{H}_n$ itself.)

There is a topology, called the *Satake topology* (see [Sat60], [BB66, Thm. 4.9], and [BJ06, Sec. III.3] for systematic treatments not just for this case), on the set

$$\mathcal{H}_n^* := \mathcal{H}_n \cup \left(\underset{1\leq r\leq n}{\cup} \mathrm{G}(\mathbb{Q}) \cdot \mathcal{H}_{n-r}\right), \tag{4.1.3.14}$$

where $\mathcal{H}_{n-r}$ stands for $\begin{pmatrix} \infty_r & \\ & \mathcal{H}_{n-r} \end{pmatrix}$ as in (4.1.3.12), such that $\mathcal{H}_{n-r}$ lies in the closure of $\mathcal{H}_{n-s}$ exactly when $r \geq s$, and such that $\Gamma\backslash\mathcal{H}_n^*$ is a compactification of $\Gamma\backslash\mathcal{H}_n$ as a topological space, for every arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$ (by the reduction theory, as in [BHC62] and [Bor62]). According to [BB66], $\Gamma\backslash\mathcal{H}_n^*$ is a normal projective variety over $\mathbb{C}$, which is called the *Satake–Baily–Borel* or *minimal* compactification of $\Gamma\backslash\mathcal{H}_n$. (The references given above are not the historical ones mainly for $\mathcal{H}_n^*$ and $\Gamma\backslash\mathcal{H}_n^*$, such as [Sat56], [Bai58], [Car58], and [SC58]—see [Sat01, Sec. 2] for a nice review. Rather, they are more recent ones that also work for our later examples.)

Each $\begin{pmatrix} \infty_r & \\ & \mathcal{H}_{n-r} \end{pmatrix}$, or rather $\left\{\begin{pmatrix} 1_r & & \\ & & Z \\ 0_r & & \\ & & 1_{n-r} \end{pmatrix} : Z \in \mathcal{H}_{n-r}\right\}$ in generalized projective coordinates (see (4.1.3.13)), is stabilized by a *maximal parabolic subgroup* of

G(ℚ) of the form

$$\mathrm{P}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & \\ & * & * & * \end{pmatrix} \right\} \tag{4.1.3.15}$$

because

$$\begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & \\ & * & * & * \end{pmatrix} \begin{pmatrix} 1 & & & \\ & & & * \\ 0 & & & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} * & * \\ & * \\ 0 & \\ & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & & & \\ & & & * \\ 0 & & & \\ & & & 1 \end{pmatrix}, \tag{4.1.3.16}$$

where the sizes of the block matrices are given by r, $n-r$, r, and $n-r$, both horizontally and vertically. This $\mathrm{P}^{(r)}(\mathbb{Q})$ is the stabilizer of the totally isotropic ℚ-subspace

$$\left\{ \begin{pmatrix} x \\ 0 \\ 0 \\ 0 \end{pmatrix} : x \in \mathbb{Q}^r \right\} \tag{4.1.3.17}$$

of

$$\mathbb{Q}^{2n} \cong \mathbb{Q}^r \oplus \mathbb{Q}^{n-r} \oplus \mathbb{Q}^r \oplus \mathbb{Q}^{n-r}. \tag{4.1.3.18}$$

We have

$$\mathrm{P}^{(r)}(\mathbb{Q}) = \mathrm{L}^{(r)}(\mathbb{Q}) \ltimes \mathrm{U}^{(r)}(\mathbb{Q}), \tag{4.1.3.19}$$

where

(4.1.3.20)

$$\begin{aligned} \mathrm{L}^{(r)}(\mathbb{Q}) &= \left\{ \begin{pmatrix} X & & & \\ & A & & B \\ & & {}^tX^{-1} & \\ & C & & D \end{pmatrix} : X \in \mathrm{GL}_r(\mathbb{Q}), \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2n-2r}(\mathbb{Q}) \right\} \\ &\cong \mathrm{GL}_r(\mathbb{Q}) \times \mathrm{Sp}_{2n-2r}(\mathbb{Q}) \end{aligned}$$

is a *Levi subgroup*, and where

$$\mathrm{U}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & * & * & * \\ & 1 & * & \\ & & 1 & \\ & & * & 1 \end{pmatrix} \right\} \tag{4.1.3.21}$$

is the *unipotent radical*. Note that $\mathrm{L}^{(r)}(\mathbb{Q})$ acts on the rational boundary component $\mathcal{H}_{2n-2r}$ via the second factor $\mathrm{Sp}_{2n-2r}(\mathbb{Q})$. This is the so-called *Hermitian part* of the Levi subgroup $\mathrm{L}^{(r)}(\mathbb{Q})$. The unipotent radical $\mathrm{U}^{(r)}(\mathbb{Q})$ is an extension

$$1 \to \mathrm{W}^{(r)}(\mathbb{Q}) \to \mathrm{U}^{(r)}(\mathbb{Q}) \to \mathrm{V}^{(r)}(\mathbb{Q}) \to 1, \tag{4.1.3.22}$$

where

$$\mathrm{W}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & & Y & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} : Y \in \mathrm{Sym}_r(\mathbb{Q}) \right\} \cong \mathrm{Sym}_r(\mathbb{Q}) \tag{4.1.3.23}$$

is the *center* of $\mathrm{U}^{(r)}(\mathbb{Q})$, and where

$$\mathrm{V}^{(r)}(\mathbb{Q}) \cong \mathbb{Q}^{2n-2r} \tag{4.1.3.24}$$

(by viewing $\mathrm{Sym}_r(\mathbb{Q})$ and $\mathbb{Q}^{2n-2r}$ as commutative algebraic groups over $\mathbb{Q}$).

4.1.4. *Cases with $\mathcal{D} = \mathcal{H}_{a,b}$, with $a \geq b$.* Suppose E is some imaginary quadratic extension field of $\mathbb{Q}$. Suppose G is a group scheme over $\mathbb{Q}$ defined by

$$\mathrm{G}(R) := \{g \in \mathrm{GL}_{a+b}(E \underset{\mathbb{Q}}{\otimes} R) : {}^t\overline{g} J_{a,b} g = J_{a,b}\}, \tag{4.1.4.1}$$

for each $\mathbb{Q}$-algebra R, where $J_{a,b}$ is some skew-Hermitian matrix as in (3.2.2.2) with S a skew-Hermitian matrix with entries in E. (We do not need to choose any S when $a = b$.) Then $\mathrm{G}(\mathbb{R})$ is the group $\mathrm{U}'_{a,b}$ defined by $J_{a,b}$ as in (3.2.2.1), which acts on $\mathcal{H}_{a,b}$ as in Section 3.2.2.

In this case, the rational boundary components of $\mathcal{H}_{a,b}$ (with respect to the choice of G) are given by elements of the $\mathrm{G}(\mathbb{Q})$-orbits of

$$\begin{pmatrix} \infty_r & \\ & \mathcal{H}_{a-r,b-r} \end{pmatrix} \cong \mathcal{H}_{a-r,b-r}, \tag{4.1.4.2}$$

which means

$$\left\{ \begin{pmatrix} 1_r & \\ & Z \\ & W \\ 0_r & \\ & 1_{b-r} \end{pmatrix} : \begin{pmatrix} Z \\ W \end{pmatrix} \in \mathcal{H}_{a-r,b-r} \right\} \tag{4.1.4.3}$$

in generalized projective coordinates (cf. Remark 3.1.1.4), for $r = 1, \ldots, b$. (When $r = 0$, we just get $\mathcal{H}_{a,b}$ itself.)

With the Satake topology (see the reference given in Section 4.1.3, which we shall not repeat) on the set

$$\mathcal{H}^*_{a,b} := \mathcal{H}_{a,b} \cup \left(\underset{1 \leq r \leq b}{\cup} \mathrm{G}(\mathbb{Q}) \cdot \mathcal{H}_{a-r,b-r} \right), \tag{4.1.4.4}$$

where $\mathcal{H}_{a-r,b-r}$ stands for $\begin{pmatrix} \infty_r & \\ & \mathcal{H}_{a-r,b-r} \end{pmatrix}$ as in (4.1.4.2), $\mathcal{H}_{a-r,b-r}$ lies in the closure of $\mathcal{H}_{a-s,b-s}$ exactly when $r \geq s$, and the quotient $\Gamma \backslash \mathcal{H}^*_{a,b}$ is a compactification of $\Gamma \backslash \mathcal{H}_{a,b}$ as a topological space, for every arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$. Again, according to [BB66], $\Gamma \backslash \mathcal{H}^*_{a,b}$ is a normal projective variety over $\mathbb{C}$, which is the *minimal compactification* of $\Gamma \backslash \mathcal{H}_{a,b}$.

Each $\begin{pmatrix} \infty_r & \\ & \mathcal{H}_{a-r,b-r} \end{pmatrix}$, or rather $\left\{ \begin{pmatrix} 1_r & \\ & Z \\ & W \\ 0_r & \\ & 1_{b-r} \end{pmatrix} : \begin{pmatrix} Z \\ W \end{pmatrix} \in \mathcal{H}_{a-r,b-r} \right\}$ in generalized projective coordinates (see (4.1.4.3)), is stabilized by a maximal parabolic subgroup of $\mathrm{G}(\mathbb{Q})$ of the form

$$\mathrm{P}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} * & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \\ & & & * & \\ & * & * & * & * \end{pmatrix} \right\}, \tag{4.1.4.5}$$

where the sizes of the block matrices are given by r, $b-r$, $a-b$, r, and $b-r$, both horizontally and vertically. This $\mathrm{P}^{(r)}(\mathbb{Q})$ is the stabilizer of the totally isotropic E-subspace

$$\left\{ \begin{pmatrix} x \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} : x \in E^r \right\} \tag{4.1.4.6}$$

of

$$E^{a+b} \cong E^r \oplus E^{b-r} \oplus E^{a-b} \oplus E^r \oplus E^{b-r}. \tag{4.1.4.7}$$

We have

$$\mathrm{P}^{(r)}(\mathbb{Q}) = \mathrm{L}^{(r)}(\mathbb{Q}) \ltimes \mathrm{U}^{(r)}(\mathbb{Q}), \tag{4.1.4.8}$$

where

(4.1.4.9)

$$\begin{aligned} \mathrm{L}^{(r)}(\mathbb{Q}) &= \left\{ \begin{pmatrix} X & & & & \\ & A & E & & B \\ & F & M & & G \\ & & & {}^t\overline{X}^{-1} & \\ & C & H & & D \end{pmatrix} : X \in \mathrm{GL}_r(E), \begin{pmatrix} A & E & B \\ F & M & G \\ C & H & D \end{pmatrix} \in \mathrm{G}^{(r)}(\mathbb{Q}) \right\} \\ &\cong \mathrm{GL}_r(E) \times \mathrm{G}^{(r)}(\mathbb{Q}) \end{aligned}$$

is a Levi subgroup, where $\mathrm{G}^{(r)}$ is the analogue of G defined by the smaller matrix

$$J_{a-r,b-r} = \begin{pmatrix} & & 1_{b-r} \\ & S & \\ -1_{b-r} & & \end{pmatrix} \tag{4.1.4.10}$$

with the same S; and where

$$\mathrm{U}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & * & * & * & * \\ & 1 & & * & \\ & & 1 & * & \\ & & & 1 & \\ & & & * & 1 \end{pmatrix} \right\} \tag{4.1.4.11}$$

is the unipotent radical. Note that $\mathrm{L}^{(r)}(\mathbb{Q})$ acts on the rational boundary component $\mathcal{H}_{a-r,b-r}$ via the second factor $\mathrm{G}^{(r)}(\mathbb{Q})$. This is the Hermitian part of the Levi subgroup $\mathrm{L}^{(r)}(\mathbb{Q})$. The unipotent radical $\mathrm{U}^{(r)}(\mathbb{Q})$ is an extension

$$1 \to \mathrm{W}^{(r)}(\mathbb{Q}) \to \mathrm{U}^{(r)}(\mathbb{Q}) \to \mathrm{V}^{(r)}(\mathbb{Q}) \to 1, \tag{4.1.4.12}$$

where

$$\mathrm{W}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & & & Y & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} : Y \in \mathrm{Herm}_r(E) \right\} \cong \mathrm{Herm}_r(E) \tag{4.1.4.13}$$

is the center of $\mathrm{U}^{(r)}(\mathbb{Q})$, and where

$$\mathrm{V}^{(r)}(\mathbb{Q}) \cong E^{(a-r)+(b-r)} \tag{4.1.4.14}$$

(by viewing $\mathrm{Herm}_r(E)$ and $E^{(a-r)+(b-r)}$ as commutative algebraic groups over $\mathbb{Q}$). We note that the overall picture is very similar to the case of $\mathcal{H}_n$ in Section 4.1.3, although we need to work with coordinates over E and hence with $\mathrm{GL}_r(E)$ as an algebraic group over $\mathbb{Q}$ (which means we are working with the restriction of scalars $\mathrm{Res}_{E/\mathbb{Q}} \mathrm{GL}_r$), and with the matrix S when $a > b$.

Let us mention two important variants which frequently appear in the literature.

The first variant replaces the imaginary quadratic extension field E of $\mathbb{Q}$ in the above definition of G with a CM extension (namely, a totally imaginary quadratic extension of a totally real extension) F of $\mathbb{Q}$, and replaces $J_{a,b}$ with a nondegenerate skew-Hermitian matrix J over $\mathcal{O}_F$ such that $-iJ$ defines a Hermitian pairing of signatures $(a_1, b_1), (a_2, b_2), \ldots, (a_d, b_d)$ at the d real places of the maximal totally real subextension F^+ of $\mathbb{Q}$ in F, where $d := [F^+ : \mathbb{Q}]$, for some integers $a_j \geq b_j \geq 0$ and $n \geq 0$ such that $a_j + b_j = n$, for all $j = 1, \ldots, d$. Then

$$\mathrm{G}(\mathbb{R}) \cong \mathrm{U}'_{a_1,b_1} \times \mathrm{U}'_{a_2,b_2} \times \cdots \times \mathrm{U}'_{a_d,b_d} \tag{4.1.4.15}$$

acts on

$$\mathcal{D} \cong \mathcal{H}_{a_1,b_1} \times \mathcal{H}_{a_2,b_2} \times \cdots \times \mathcal{H}_{a_d,b_d}, \tag{4.1.4.16}$$

and the rational boundary components of $\mathcal{D}$ are given by elements of the $\mathrm{G}(\mathbb{Q})$-orbits of some

$$\mathcal{D}^{(r)} \cong \mathcal{H}_{a_1-r,b_1-r} \times \mathcal{H}_{a_2-r,b_2-r} \times \cdots \times \mathcal{H}_{a_d-r,b_d-r}, \tag{4.1.4.17}$$

for $r = 1, \ldots, \min(b_1, b_2, \ldots, b_d)$. Note that there is no r at all if at least one of $b_1, b_2, \ldots, b_d$ is zero.

In the special case where $(a_1, b_1) = (a, b)$ but $(a_j, b_j) = (n, 0)$ for all $j \neq 1$, for some integers $a \geq b \geq 0$ and $n = a + b \geq 0$, we simply have

$$\mathcal{D} \cong \mathcal{H}_{a,b}. \tag{4.1.4.18}$$

If $d = 1$, then this is just as before. If $d \geq 2$, then $\mathcal{D}^* = \mathcal{D}$ and the quotient $\Gamma\backslash\mathcal{D}$ is automatically compact, for every arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$. Whether $d \geq 2$ or not, when $(a, b) = (2, 1)$ (and when (4.1.4.18) holds), the quotient $\Gamma\backslash\mathcal{D}$ is called a *Picard modular surface*.

The second variant is a further generalization of the above first variant, by replacing F with a *central simple algebra* B over a CM extension over $\mathbb{Q}$ that is equipped with a *positive involution* $\star$ (namely, an anti-automorphism of order 2) such that $\mathrm{tr}_{B/\mathbb{Q}}(xx^\star) > 0$ for every $x \neq 0$ in B, and by defining G using a nondegenerate skew-Hermitian pairing over a finitely generated B-module. Then we still have an isomorphism as in (4.1.4.16), and the rational boundary components of $\mathcal{D}$ are still of the same form as in (4.1.4.17), but the values of r allowed in (4.1.4.17) are more restrictive, depending on the possible $\mathbb{Q}$-dimensions of totally isotropic

B-submodules. If B is a division algebra and the pairing is defined over a simple B-module, then no $r > 0$ is allowed, in which case $\Gamma\backslash\mathcal{D}$ is again automatically compact, for every arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$.

4.1.5. *Cases with* $\mathcal{D} = \mathcal{H}_{\mathrm{SO}^*_{2n}}$. Starting with this case, we will be less explicit about the choices of G. There exists a group scheme G over $\mathbb{Q}$ such that $\mathrm{G}(\mathbb{R}) \cong \mathrm{SO}^*_{2n}$, such that the rational boundary components of $\mathcal{H}_{\mathrm{SO}^*_{2n}}$ are given by elements of the $\mathrm{G}(\mathbb{Q})$-orbits of some $\mathcal{D}^{(r)} \cong \mathcal{H}_{\mathrm{SO}^*_{2n-4r}}$ (which we shall abusively denote as $\mathcal{H}_{\mathrm{SO}^*_{2n-4r}}$), with $r = 1, \ldots, \lfloor \frac{n}{2} \rfloor$, and such that, with the Satake topology on

$$\mathcal{H}^*_{\mathrm{SO}^*_{2n}} := \mathcal{H}_{\mathrm{SO}^*_{2n}} \cup \left(\underset{1 \leq r \leq \lfloor \frac{n}{2} \rfloor}{\cup} \mathrm{G}(\mathbb{Q}) \cdot \mathcal{H}_{\mathrm{SO}^*_{2n-4r}} \right), \tag{4.1.5.1}$$

$\mathcal{H}_{\mathrm{SO}^*_{2n-4r}}$ lies in the closure of $\mathcal{H}_{\mathrm{SO}^*_{2n-4s}}$ exactly when $r \geq s$, and the quotient $\Gamma\backslash\mathcal{H}^*_{\mathrm{SO}^*_{2n}}$ gives the minimal compactification of $\Gamma\backslash\mathcal{H}_{\mathrm{SO}^*_{2n}}$, for every arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$.

When $n = 2k$, an example of G is given by

$$\mathrm{G}(R) := \{ g \in \mathrm{GL}_{2k}(\mathcal{O} \underset{\mathbb{Z}}{\otimes} R) : {}^t\overline{g} J_k g = J_k \}, \tag{4.1.5.2}$$

for each $\mathbb{Q}$-algebra R, where B is a quaternion algebra over $\mathbb{Q}$ such that $B \underset{\mathbb{Q}}{\otimes} \mathbb{R} \cong \mathbb{H}$, and where $J_k = \begin{pmatrix} & 1_k \\ -1_k & \end{pmatrix}$ is as before. Then we can describe the rational boundary components $\mathcal{H}_{\mathrm{SO}^*_{4k-4r}}$ of $\mathcal{H}_{\mathrm{SO}^*_{4k}}$ as

$$\left\{ \begin{pmatrix} \infty_r & \\ & Z \end{pmatrix} : Z \in \mathrm{Herm}_{k-r}(\mathbb{H}) \underset{\mathbb{R}}{\otimes} \mathbb{C}, \ \mathrm{Im}(Z) > 0 \right\}, \tag{4.1.5.3}$$

which means

$$\left\{ \begin{pmatrix} 1_r & \\ & Z \\ 0_r & \\ & 1_{k-r} \end{pmatrix} : Z \in \mathrm{Herm}_{k-r}(\mathbb{H}) \underset{\mathbb{R}}{\otimes} \mathbb{C}, \ \mathrm{Im}(Z) > 0 \right\} \tag{4.1.5.4}$$

in generalized projective coordinates (cf. Remark 3.1.1.4), for $r = 1, \ldots, k$. Each $\mathcal{H}_{\mathrm{SO}^*_{4k-4r}}$ as above is stabilized by a maximal parabolic subgroup of $\mathrm{G}(\mathbb{Q})$ of the form

$$\mathrm{P}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & \\ & * & * & * \end{pmatrix} \right\} \tag{4.1.5.5}$$

because

$$\begin{pmatrix} * & * & * & * \\ & * & * & * \\ & & * & \\ & * & * & * \end{pmatrix} \begin{pmatrix} 1 & \\ & * \\ 0 & \\ & 1 \end{pmatrix} = \begin{pmatrix} * & * \\ & * \\ 0 & \\ & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & \\ & * \\ 0 & \\ & 1 \end{pmatrix}, \tag{4.1.5.6}$$

where the sizes of the block matrices are given by r, $n-r$, r, and $n-r$, both horizontally and vertically. This $\mathrm{P}^{(r)}(\mathbb{Q})$ is the stabilizer of the totally isotropic

B-submodule

$$\left\{ \begin{pmatrix} x \\ 0 \\ 0 \\ 0 \end{pmatrix} : x \in B^r \right\} \tag{4.1.5.7}$$

of

$$B^{2n} \cong B^r \oplus B^{n-r} \oplus B^r \oplus B^{n-r}, \tag{4.1.5.8}$$

with left B-module structures given by having each $b \in B$ act by right multiplication of its conjugate $\overline{b}$, so that they do not intervene with left multiplications of matrix entries. We have

$$\mathrm{P}^{(r)}(\mathbb{Q}) = \mathrm{L}^{(r)}(\mathbb{Q}) \ltimes \mathrm{U}^{(r)}(\mathbb{Q}), \tag{4.1.5.9}$$

where
(4.1.5.10)

$$\begin{aligned} \mathrm{L}^{(r)}(\mathbb{Q}) &= \left\{ \begin{pmatrix} X & & & \\ & A & & B \\ & & {}^t\overline{X}^{-1} & \\ & C & & D \end{pmatrix} : X \in \mathrm{GL}_r(B), \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{G}^{(r)}(\mathbb{Q}) \right\} \\ &\cong \mathrm{GL}_r(B) \times \mathrm{G}^{(r)}(\mathbb{Q}) \end{aligned}$$

is a Levi subgroup, where $\mathrm{G}^{(r)}$ is the analogue of G defined by the smaller matrix

$$J_{k-r} = \begin{pmatrix} & 1_{k-r} \\ -1_{k-r} & \end{pmatrix}; \tag{4.1.5.11}$$

and where

$$\mathrm{U}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & * & * & * \\ & 1 & * & \\ & & 1 & \\ & & * & 1 \end{pmatrix} \right\} \tag{4.1.5.12}$$

is the unipotent radical. Note that $\mathrm{L}^{(r)}(\mathbb{Q})$ acts on the rational boundary component $\mathcal{H}_{\mathrm{SO}^*_{4n-4r}}$ via the second factor $\mathrm{G}^{(r)}(\mathbb{Q})$. This is the Hermitian part of the Levi subgroup $\mathrm{L}^{(r)}(\mathbb{Q})$. The unipotent radical $\mathrm{U}^{(r)}(\mathbb{Q})$ is an extension

$$1 \to \mathrm{W}^{(r)}(\mathbb{Q}) \to \mathrm{U}^{(r)}(\mathbb{Q}) \to \mathrm{V}^{(r)}(\mathbb{Q}) \to 1, \tag{4.1.5.13}$$

where

$$\mathrm{W}^{(r)}(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & & Y & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} : Y \in \mathrm{Herm}_r(B) \right\} \cong \mathrm{Herm}_r(B) \tag{4.1.5.14}$$

is the center of $\mathrm{U}^{(r)}(\mathbb{Q})$, and where

$$\mathrm{V}^{(r)}(\mathbb{Q}) \cong B^{2k-2r} \tag{4.1.5.15}$$

(by viewing $\mathrm{Herm}_r(B)$ and B^{2k-2r} as commutative algebraic groups over $\mathbb{Q}$). We note that, again, the overall picture is very similar to the one in Section 4.1.3, although we need to work with coordinates over B and hence with $\mathrm{GL}_r(B)$ as an algebraic group over $\mathbb{Q}$.

4.1.6. *Cases with $\mathcal{D}^+ = \mathcal{H}^+_{\mathrm{SO}_{a,2}}$, with $a \geq 2$.* Suppose G is a group scheme over $\mathbb{Q}$ defined by some $\mathbb{Q}$-valued symmetric bilinear form Q on $\mathbb{Q}^{a+2}$ which has signature $(a, 2)$ over $\mathbb{R}$, so that $\mathrm{G}(\mathbb{R}) \cong \mathrm{SO}_{a,2}(\mathbb{R})$ and $\mathrm{G}(\mathbb{R})^+ \cong \mathrm{SO}_{a,2}(\mathbb{R})^+$. (The precise choice of Q will not be important for our exposition.)

The maximal parabolic subgroups of $\mathrm{G}(\mathbb{Q}) = \mathrm{SO}(\mathbb{Q}^{a+2}, Q)$, where $\mathrm{SO}(\mathbb{Q}^{a+2}, Q)$ denotes the subgroup of $\mathrm{SL}_{a+2}(\mathbb{Q})$ consisting of elements preserving the pairing Q (and we shall adopt similar notation in what follows), are in bijection with the nonzero subspaces I of $\mathbb{Q}^{a+2}$ that are totally isotropic with respect to the bilinear form Q. Then we must have $1 \leq \dim_{\mathbb{Q}}(I) \leq 2$ because the signature of Q is $(a, 2)$ over $\mathbb{R}$. (But it might happen that no such I exists.) For each such I, we have a parabolic subgroup of $\mathrm{G}(\mathbb{Q})$ of the form

$$\mathrm{P}_I(\mathbb{Q}) \cong \mathrm{L}_I(\mathbb{Q}) \ltimes \mathrm{U}_I(\mathbb{Q}), \tag{4.1.6.1}$$

where

$$\mathrm{L}_I(\mathbb{Q}) \cong \mathrm{SO}(I^\perp/I, Q|_{I^\perp/I}) \times \mathrm{GL}_{\mathbb{Q}}(I) \tag{4.1.6.2}$$

is the Levi quotient (which is also noncanonically a subgroup), and where $\mathrm{U}_I(\mathbb{Q})$ is the unipotent radical which fits into a short exact sequence

$$1 \to \wedge^2 I \to \mathrm{U}_I(\mathbb{Q}) \to \mathrm{Hom}_{\mathbb{Q}}(I^\perp/I, I) \to 1. \tag{4.1.6.3}$$

When $\dim_{\mathbb{Q}}(I) = 1$, we have

$$\mathrm{L}_I(\mathbb{R}) \cong \mathrm{SO}_{a-1,1}(\mathbb{R}) \times \mathrm{GL}_1(\mathbb{R}), \tag{4.1.6.4}$$

and so $\mathrm{L}_I(\mathbb{R}) \cap \mathrm{G}(\mathbb{R})^+$ acts trivially on some zero-dimensional rational boundary component

$$\mathcal{H}_I^+ \cong \mathcal{H}_0. \tag{4.1.6.5}$$

When $\dim_{\mathbb{Q}}(I) = 2$, we have

$$\mathrm{L}_I(\mathbb{R}) \cong \mathrm{SO}_{a-2}(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}), \tag{4.1.6.6}$$

and so $\mathrm{L}_I(\mathbb{R}) \cap \mathrm{G}(\mathbb{R})^+$ acts via the usual Möbius transformation of

$$\mathrm{GL}_2(\mathbb{R})^+ = \{g \in \mathrm{GL}_2(\mathbb{R}) : \det(g) > 0\} \tag{4.1.6.7}$$

on some one-dimensional rational boundary component

$$\mathcal{H}_I^+ \cong \mathcal{H}_1. \tag{4.1.6.8}$$

For example, suppose (for simplicity) that Q is defined by the symmetric matrix

$$\begin{pmatrix} 1_{a-2} & & & & \\ & 0 & 1 & & \\ & 1 & 0 & & \\ & & & 0 & 1 \\ & & & 1 & 0 \end{pmatrix}, \tag{4.1.6.9}$$

where the sizes of the block matrices are given by $a - 2$, 1, 1, 1, and 1. If

$$I = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ * \\ 0 \end{pmatrix} \right\}, \tag{4.1.6.10}$$

which is one-dimensional, then

$$\mathrm{P}_I(\mathbb{Q}) = \left\{ \begin{pmatrix} * & * & * & & * \\ * & * & * & & * \\ * & * & * & & * \\ * & * & * & * & * \\ & & & & * \end{pmatrix} \right\} \cong \mathrm{L}_I(\mathbb{Q}) \ltimes \mathrm{U}_I(\mathbb{Q}), \tag{4.1.6.11}$$

where

$$\mathrm{L}_I(\mathbb{Q}) = \left\{ \begin{pmatrix} * & * & * & & \\ * & * & * & & \\ * & * & * & & \\ & & & X & \\ & & & & X^{-1} \end{pmatrix} : \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \in \mathrm{SO}_{a-1,1}(\mathbb{Q}),\ X \in \mathbb{Q}^\times \right\} \tag{4.1.6.12}$$

and

$$\mathrm{U}_I(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & & & & x \\ & 1 & & & y \\ & & 1 & & z \\ -{}^t x & -z & -y & 1 & -\frac{1}{2}\,{}^t xx - yz \\ & & & & 1 \end{pmatrix} : x \in \mathbb{Q}^{a-2},\ y, z \in \mathbb{Q} \right\}; \tag{4.1.6.13}$$

and $\mathrm{P}_I(\mathbb{Q})$ stabilizes

$$\mathcal{H}_I^+ = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\} \tag{4.1.6.14}$$

because

$$\begin{pmatrix} * & * & * & & * \\ * & * & * & & * \\ * & * & * & & * \\ * & * & * & * & * \\ & & & & * \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ * \\ 0 \end{pmatrix} \sim \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{4.1.6.15}$$

in projective coordinates. Alternatively, if

$$I = \left\{ \begin{pmatrix} 0 \\ * \\ 0 \\ * \\ 0 \end{pmatrix} \right\}, \tag{4.1.6.16}$$

which is two-dimensional, then

$$\mathrm{P}_I(\mathbb{Q}) = \left\{ \begin{pmatrix} * & & * & & * \\ * & * & * & * & * \\ & & * & & * \\ * & * & * & * & * \\ & & * & & * \end{pmatrix} \right\} \cong \mathrm{L}_I(\mathbb{Q}) \ltimes \mathrm{U}_I(\mathbb{Q}), \tag{4.1.6.17}$$

where
(4.1.6.18)

$$\mathrm{L}_I(\mathbb{Q}) = \left\{ \begin{pmatrix} X & & & & \\ & a & & b & \\ & & a' & & b' \\ & c & & d & \\ & & c' & & d' \end{pmatrix} : \begin{matrix} X \in \mathrm{SO}_{a-2}(\mathbb{Q}), \\ {}^t\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Q}) \end{matrix} \right\}$$

and
(4.1.6.19)

$$\mathrm{U}_I(\mathbb{Q}) = \left\{ \begin{pmatrix} 1 & 0 & x & 0 & y \\ -{}^t x & 1 & -\frac{1}{2}\,{}^t xx & 0 & w \\ 0 & 0 & 1 & 0 & 0 \\ -{}^t y & 0 & -w - {}^t xy & 1 & -\frac{1}{2}\,{}^t yy \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} : w \in \mathbb{Q},\, x, y \in \mathbb{Q}^{a-2} \right\};$$

and $\mathrm{P}_I(\mathbb{Q}) \cap \mathrm{G}(\mathbb{R})^+$ stabilizes

$$\mathcal{H}_I^+ = \left\{ \begin{pmatrix} 0 \\ z \\ 0 \\ 1 \\ 0 \end{pmatrix} : z \in \mathbb{C},\, \mathrm{Im}\, z > 0 \right\} \tag{4.1.6.20}$$

because

$$\begin{pmatrix} * & & * & & * \\ * & a & & b & * \\ & & * & & * \\ * & c & * & d & * \\ & & * & & * \end{pmatrix} \begin{pmatrix} 0 \\ z \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ az+b \\ 0 \\ cz+d \\ 0 \end{pmatrix} \sim \begin{pmatrix} 0 \\ (az+b)(cz+d)^{-1} \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{4.1.6.21}$$

in projective coordinates.

There is a systematic way to combine these into a union

$$(\mathcal{H}_{\mathrm{SO}_{a,2}}^+)^* = \mathcal{H}_{\mathrm{SO}_{a,2}}^+ \cup \left(\bigcup_{\substack{0 \neq I \text{ totally} \\ \text{isotropic}}} \mathcal{H}_I^+ \right), \tag{4.1.6.22}$$

equipped with Satake topology, such that $\mathcal{H}_I^+$ lies in the closure of $\mathcal{H}_{I'}^+$ (for nonzero I and I') exactly when $I \subset I'$ (which is qualitatively quite different from the previous cases in Sections 4.1.3, 4.1.4, and 4.1.5, where rational boundary components associated with larger totally isotropic subspaces are smaller) and such that the quotient $\Gamma\backslash(\mathcal{H}_{\mathrm{SO}_{a,2}}^+)^*$ gives the minimal compactification of $\Gamma\backslash\mathcal{H}_{\mathrm{SO}_{a,2}}^+$, for every arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$ that is contained in $\mathrm{G}(\mathbb{R})^+$.

4.1.7. *Cases with $\mathcal{D} = \mathcal{H}_{\mathrm{E}_6}$ or $\mathcal{D} = \mathcal{H}_{\mathrm{E}_7}$.* We shall be very brief in these two cases. For $\mathcal{H}_{\mathrm{E}_6}$, there exists a group scheme G over $\mathbb{Q}$ such that $\mathrm{G}(\mathbb{R})$ is as in Section 3.6.2 and such that the rational boundary components of $\mathcal{H}_{\mathrm{E}_6}$ are isomorphic to either $\mathcal{H}_{5,1}$ or $\mathcal{H}_0$. For $\mathcal{H}_{\mathrm{E}_7}$, there exists a group scheme G over $\mathbb{Q}$ such that $\mathrm{G}(\mathbb{R})$ is as in Section 3.5.3 and such that the rational boundary components of $\mathcal{H}_{\mathrm{E}_7}$ are isomorphic to either $\mathcal{H}_{\mathrm{SO}_{10,2}}^+$, $\mathcal{H}_1$, or $\mathcal{H}_0$. In each of these two cases, by [PR06, Sec. 3, Prop. 3], there exists some G such that $\mathrm{G}(\mathbb{R})$ is as above and such that $\mathrm{rank}_{\mathbb{Q}}\, \mathrm{G}_{\mathbb{Q}} = \mathrm{rank}_{\mathbb{R}}\, \mathrm{G}_{\mathbb{R}}$; or, in other words, $\mathrm{G}_{\mathbb{Q}}$ is as split as $\mathrm{G}_{\mathbb{R}}$ allows. In both cases, we have the minimal compactifications, as usual, for arithmetic subgroups Γ of $\mathrm{G}(\mathbb{Q})$.

4.2. **Compactifications and algebraicity.**

4.2.1. *Overview.* In all cases considered so far, we have the following diagram of useful compactifications of $\Gamma\backslash\mathcal{D}$, with canonical morphisms between them denoted by solid arrows:

$$\begin{tikzcd} (\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma} \arrow[rr, dotted, dash, "\text{not directly related}"] \arrow[ddr] & & (\Gamma\backslash\mathcal{D})^{\mathrm{BS}} \arrow[d] \\ & & (\Gamma\backslash\mathcal{D})^{\mathrm{RBS}} \arrow[dl] \\ & (\Gamma\backslash\mathcal{D})^{\min} & \end{tikzcd} \tag{4.2.1.1}$$

Let us explain the objects in this diagram.

4.2.2. $(\Gamma\backslash\mathcal{D})^{\min} = \Gamma\backslash\mathcal{D}^*$. This is the *Satake–Baily–Borel* or *minimal compactification* of $\Gamma\backslash\mathcal{D}$, which is a normal *projective variety* over $\mathbb{C}$ (see [BB66]). Here $\mathcal{D}^*$ is the union of $\mathcal{D}$ with its rational boundary components, equipped with the Satake topology, as explained in Section 4.1. The transition from the analytic quotient $\Gamma\backslash\mathcal{D}^*$ to an algebraic variety is achieved by first endowing $\Gamma\backslash\mathcal{D}^*$ with the structure of a *complex analytic space* (see [Ser56, §1]) in which the closed complement of the open subset $\Gamma\backslash\mathcal{D}$ of $\Gamma\backslash\mathcal{D}^*$ is a *closed complex analytic subspace*; and then by constructing a closed immersion of $\Gamma\backslash\mathcal{D}^*$ into a projective space $\mathbb{P}^N(\mathbb{C})$ for some N, and by applying Chow's theorem (or "GAGA" for closed complex analytic subspaces of $\mathbb{P}^N_{\mathbb{C}}$, showing that they are all complex analytifications of algebraic subvarieties—see [Ser56, §3, 19, Prop. 13]). Then the open subspace $\Gamma\backslash\mathcal{D}$ of $(\Gamma\backslash\mathcal{D})^{\min} = \Gamma\backslash\mathcal{D}^*$ is a *quasi-projective variety* over $\mathbb{C}$, by applying Chow's theorem again to the closed complement. (This is the only known general method to show that $\Gamma\backslash\mathcal{D}$ is an algebraic variety.) Moreover, it is proved in [Bor72] that any holomorphic map from a quasi-projective variety to $\Gamma\backslash\mathcal{D}$ is the complex analytification of a morphism of algebraic varieties. In particular, the structure of $\Gamma\backslash\mathcal{D}$ as an algebraic variety is *unique* up to isomorphism. (These are the key results used in the proof of Theorem 2.4.1.)

The compactification $(\Gamma\backslash\mathcal{D})^{\min}$ is canonical and does not depend on any choices, but is highly singular in general. As we have seen in the examples in Section 4.1, the boundary $(\Gamma\backslash\mathcal{D})^{\min} - (\Gamma\backslash\mathcal{D})$ has a high codimension in $(\Gamma\backslash\mathcal{D})^{\min}$ in general. When the connected components of $\mathcal{D}$ are irreducible in the sense explained in the beginning of Section 3 but are not one-dimensional, and when $\Gamma\backslash\mathcal{D}$ is not already compact, this codimension is always at least two. The intuition is that we are collapsing a lot into a small neighborhood of each boundary point.

4.2.3. $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$. This is the *toroidal compactification* of $\Gamma\backslash\mathcal{D}$ (see [AMRT75] or [AMRT10], and see also [Mum75]), which is an algebraic space depending on the choice of some combinatorial data Σ called a *cone decomposition* (or *fan*). (More precisely, Σ is a compatible collection of cone decompositions, but we prefer not to go into the details here.) This means $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$ is not a canonical compactification of $\Gamma\backslash\mathcal{D}$ in the sense that it does not just depend on $\Gamma\backslash\mathcal{D}$. Nevertheless, it depends only on $\Gamma\backslash\mathcal{D}$ and Σ, and there exist choices of Σ such that $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$ is a *smooth projective variety*, and such that the boundary $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma} - (\Gamma\backslash\mathcal{D})$ is a *simple normal crossings divisor*. Such choices can always be achieved by replacing any given Σ with a *refinement*. In particular, we obtain meaningful smooth projective compactifications by paying the price of working with the noncanonical Σ's. (As a comparison, we have no precise control on the smooth compactifications obtained by resolving the singularities of $(\Gamma\backslash\mathcal{D})^{\mathrm{min}}$ using Hironaka's general theory in [Hir64a, Hir64b].)

As in the case of $(\Gamma\backslash\mathcal{D})^{\mathrm{min}}$, the boundary of $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$ is stratified by locally closed subsets corresponding to Γ-orbits of the proper parabolic subgroups P of the algebraic group G over $\mathbb{Q}$ whose pullback to every $\mathbb{Q}$-simple (almost) factor G' of G is either the whole G' or a maximal parabolic subgroup P'. This stratification is further refined by a stratification in terms of the cones in the cone decomposition Σ. The incidence relations of the strata in this finer stratification is dual to the incidence relations of the cones in the cone decomposition Σ. The canonical map $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma} \to (\Gamma\backslash\mathcal{D})^{\mathrm{min}}$ in (4.2.1.1) respects such stratifications.

One subtlety is that, except in very special cases, no choice of Σ can be compatible with all Hecke correspondences. For arithmetic applications, it is necessary to systematically work with refinements of Σ.

4.2.4. $(\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$. This is the *Borel–Serre* compactification of $\Gamma\backslash\mathcal{D}$ (see [BS73] and [BJ06, Sec. III.5]; see also [Gor05, Sec. 4]), which is a (real analytic) *manifold with corners*. It can be defined more generally when $\mathcal{D}$ is a Riemannian symmetric space that is not Hermitian, but is almost never an algebraic variety. Already when $\mathrm{G} = \mathrm{SL}_2$ and $\mathcal{D} = \mathcal{H}$ (and when Γ is neat), the compactification $(\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$ is achieved by adding *circles* to $\mathcal{D}$ (instead of adding points to $\mathcal{D}$ as in the case of $\mathcal{D}^*$). In general, its boundary is stratified by locally closed subsets corresponding to the Γ-orbits of *all* proper rational parabolic subgroups P of the algebraic group G over $\mathbb{Q}$, and the construction of the subset associated with P uses the whole P, instead of just its Levi quotient as in the case of $(\Gamma\backslash\mathcal{D})^{\mathrm{min}}$.

The most important feature of $(\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$ is that the canonical open embedding $\Gamma\backslash\mathcal{D} \hookrightarrow (\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$ is a *homotopy equivalence*, which makes $(\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$ useful for studying the cohomology of $\Gamma\backslash\mathcal{D}$, and also that of the arithmetic subgroup Γ of $\mathrm{G}(\mathbb{Q})$, which are closely related to the theory of automorphic representations and Eisenstein series (see, for example, [BW00], [Sch90b], [Fra98], [FS98], and [LS04]). But let us remark here, as an addendum to Section 4.2.2 above, that cuspidal automorphic representations are more closely related to the *L^2-cohomology* of $\Gamma\backslash\mathcal{D}$, and this L^2-cohomology is canonical isomorphic to the *intersection cohomology* of $(\Gamma\backslash\mathcal{D})^{\mathrm{min}}$. This was conjectured by Zucker (see [Zuc82, (6.20)]) and proved by Looijenga (see [Loo88]), Saper–Stern (see [SS90]), and Looijenga–Rapoport (see [LR91]).

4.2.5. $(\Gamma\backslash\mathcal{D})^{\mathrm{RBS}}$. This is the *reductive Borel–Serre* compactification of $\Gamma\backslash\mathcal{D}$ (see [Zuc82, Sec. 4], [GHM94, Part I], and [BJ06, Sec. III.6]; see also [Gor05, Sec. 5]), which is similar to $(\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$ in the following aspects: It can also be defined more generally when $\mathcal{D}$ is a Riemannian symmetric space that is not Hermitian, and its boundary is also stratified by locally closed subsets corresponding to the Γ-orbits of *all* proper rational parabolic subgroups P of G. But there is a key difference—the subset associated with P is constructed using only the (reductive) Levi quotient L of P. The reductive Borel–Serre compactification $(\Gamma\backslash\mathcal{D})^{\mathrm{RBS}}$ is useful for studying, for example, the *weighted cohomology* of Γ (see [GHM94, Parts II–IV]; see also [Gor05, Sec. 6.8]).

4.2.6. *Models over number fields or their rings of integers.* Although $(\Gamma\backslash\mathcal{D})^{\mathrm{BS}}$ and $(\Gamma\backslash\mathcal{D})^{\mathrm{RBS}}$ are very useful analytic objects, when it comes to applications requiring algebraic varieties (or at least algebraic spaces or algebraic stacks) defined over number fields, we need to work with $\Gamma\backslash\mathcal{D}$, $(\Gamma\backslash\mathcal{D})^{\min}$, and $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$, the last being perhaps the most complicated. It is known that $(\Gamma\backslash\mathcal{D})^{\min}$ and $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$ also have models over the same number fields over which $\Gamma\backslash\mathcal{D}$ is defined (see [Pin89]), and they extend to good models over the integers in many important special cases (see Sections 5.1.4 and 5.2.4 below).

All assertions so far in this Section 4.2 have their counterparts for Shimura varieties; i.e., the double coset spaces $\mathsf{X}_{\mathcal{U}}$ defined by Shimura data. (As we explained in Section 4.2.2, the results there are used in the proof of Theorem 2.4.1.) Moreover, the minimal and toroidal compactifications also have canonical models over the same reflex field F_0 determined by $(\mathrm{G},\mathcal{D})$ as in Theorem 2.4.3 (see [Pin89] again).

4.2.7. *Algebro-geometric definition of modular forms.* The projective coordinates of the minimal compactification $(\Gamma\backslash\mathcal{D})^{\min} = \Gamma\backslash\mathcal{D}^*$ are given by (holomorphic) *modular forms*, namely, holomorphic functions $f : \mathcal{D} \to \mathbb{C}$ satisfying the following two conditions:

(1) *Automorphy condition*: $f(\gamma Z) = j(\gamma, Z)f(Z)$, for all $\gamma \in \Gamma$ and $Z \in \mathcal{D}$, for some *automorphy factor* $j : \Gamma \times \mathcal{D} \to \mathbb{C}^\times$ defining the *weight* of f, satisfying the cocycle condition $j(\gamma'\gamma, Z) = j(\gamma', Z)j(\gamma, Z)$ for all $\gamma, \gamma' \in \Gamma$ and $Z \in \mathcal{D}$. (This means f represents a section of an *automorphic line bundle* over $\Gamma\backslash\mathcal{D}$ defined by j. The line bundles used in [BB66] are powers of the *canonical bundle*, namely, the top exterior power of the cotangent bundle, of the complex manifold $\Gamma\backslash\mathcal{D}$.)
(2) *Growth condition*: $f(Z)$ stays bounded near $\mathcal{D}^* - \mathcal{D}$. (When Γ is neat, this implies that the section represented by f extends to a section of a line bundle over $(\Gamma\backslash\mathcal{D})^{\min} = \Gamma\backslash\mathcal{D}^*$.) If f satisfies the additional condition that $f(Z) \to 0$ as $Z \to \mathcal{D}^* - \mathcal{D}$, then f is called a *cusp form*.

Ratios of such modular forms f define *modular functions* which are in general highly transcendental as functions of the coordinates of $\mathcal{D}$. Therefore, to give models of $\Gamma\backslash\mathcal{D}$ and $(\Gamma\backslash\mathcal{D})^{\min}$ over number fields or over their rings of integers, we cannot just naively use the coordinates of $\mathcal{D}$.

Not all modular forms are defined as above using only automorphy factors valued in $\mathbb{C}^\times$. It is possible to consider automorphy factors valued in $\mathrm{Cent}_{h_0}(\mathrm{G}(\mathbb{C}))$, so that it makes sense to consider modular forms with weights given by all finite-dimensional representations W of $\mathrm{Cent}_{h_0}(\mathrm{G}(\mathbb{C}))$. But there is a subtlety here: For such a W, in general, the associated *automorphic vector bundle* $\underline{W}$ over $\Gamma\backslash\mathcal{D}$ does not extend to a vector bundle over $(\Gamma\backslash\mathcal{D})^{\min}$. (This is consistent with the fact that, in general, the boundary $(\Gamma\backslash\mathcal{D})^{\min} - (\Gamma\backslash\mathcal{D})$ has a high codimension in $(\Gamma\backslash\mathcal{D})^{\min}$.) Nevertheless, the automorphic vector bundle $\underline{W}$ has *canonical extensions* (see [Mum77] and [Har89]) to vector bundles $\underline{W}^{\mathrm{can}}$ over the toroidal compactifications $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$, whose sections are independent of Σ. (When W is the restriction of a representation of $\mathrm{G}(\mathbb{C})$, the vector bundle $\underline{W}$ is equipped with an integrable connection, and its canonical extension $\underline{W}^{\mathrm{can}}$ is equipped with an integrable connection with log poles along the boundary and coincides with the one given in [Del70].) The global sections or more generally the cohomology classes of such canonical extensions $\underline{W}^{\mathrm{can}}$ over $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}_{\Sigma}$ then provide an algebro-geometric definition of modular forms with coefficients in W. There are also the *subcanonical extensions* $\underline{W}^{\mathrm{sub}} := \underline{W}^{\mathrm{can}}(-\infty)$, where ∞ abusively denotes the boundary divisor $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}} - (\Gamma\backslash\mathcal{D})$ with its reduced structure, whose global sections are (roughly speaking) useful for studying cusp forms. The cohomology of $\underline{W}^{\mathrm{can}}$ and $\underline{W}^{\mathrm{sub}}$ is often called the *coherent cohomology* of $(\Gamma\backslash\mathcal{D})^{\mathrm{tor}}$. It is this approach that will allow us to define modular forms in mixed or positive characteristics and study congruences among them using integral models. Even in characteristic zero, there have been some results which have only been proved with such an algebro-geometric definition. (See [Har90], [Mil90], [Lan12b], and [Lan19] for some surveys on this topic.)

4.2.8. *Mixed Shimura varieties.* Just as Shimura varieties parameterize certain variations of Hodge structures with additional structures, there is also a notion of *mixed Shimura varieties* parameterizing certain variations of *mixed Hodge structures* with additional structures. Then there is also a theory of canonical models for such mixed Shimura varieties, and even for their *toroidal compactifications*—see [Mil90] and [Pin89]. Many mixed Shimura varieties naturally appear along the boundary of the toroidal compactifications of *pure* Shimura varieties (i.e., the usual ones associated with reductive groups as in Sections 2.3 and 2.4).

The theory of mixed Shimura varieties is even heavier in notation than the theory of (pure) Shimura varieties, and is beyond the scope of this introductory article. But let us at least mention the following prototypical example. Consider the Shimura varieties $\mathsf{X}_{\mathcal{U}}$ associated to $(\mathrm{G}, \mathcal{D}) = (\mathrm{GL}_2, \mathcal{H}^{\pm})$, where $\mathcal{H}^{\pm} := \mathcal{H}^{\pm}_1$, which are modular curves. For each integer $m \geq 1$, consider the (non-reductive) semidirect product $\widetilde{\mathrm{G}} = \mathrm{GL}_2 \ltimes (\mathbf{G}_{\mathrm{a}}^2)^m$, with $\mathrm{G} = \mathrm{GL}_2$ acting diagonally on $(\mathbf{G}_{\mathrm{a}}^2)^m$, where $\mathbf{G}_{\mathrm{a}}$ is the additive group scheme over $\mathbb{Z}$, so that $\mathrm{GL}_2(R)$ acts by left multiplication on $(\mathbf{G}_{\mathrm{a}}^2)(R) = R^2$ for every ring R. Consider the corresponding action of $\widetilde{\mathrm{G}}(\mathbb{R})$ on $\widetilde{\mathcal{D}} = \mathcal{H}^{\pm} \times (\mathbb{R}^2)^m$. Then we obtain by considering double coset spaces as in Section 2.2 the *mixed Shimura varieties* $\widetilde{\mathsf{X}}_{\widetilde{\mathcal{U}}}$ over $\mathsf{X}_{\mathcal{U}}$ for neat open compact subgroups $\widetilde{\mathcal{U}}$ with image in $\mathcal{U}$ under the canonical homomorphism $\widetilde{\mathrm{G}}(\mathbb{A}^{\infty}) \to \mathrm{G}(\mathbb{A}^{\infty})$, which are torsors under abelian schemes over $\mathsf{X}_{\mathcal{U}}$ that are isogenous to the m-fold self-fiber product of the universal elliptic curve over $\mathsf{X}_{\mathcal{U}}$. Specifically, there exists some $\widetilde{\mathcal{U}}$ such that $\widetilde{\mathsf{X}}_{\widetilde{\mathcal{U}}}$ is exactly isomorphic to the m-fold self-fiber product of the universal elliptic curve over $\mathsf{X}_{\mathcal{U}}$. The projective smooth toroidal compactifications of such mixed Shimura varieties then provide generalizations of the classical *Kuga–Sato varieties*, whose cohomology with trivial coefficients are useful for studying the cohomology with nontrivial coefficients over the usual modular curves $\mathsf{X}_{\mathcal{U}}$, or even for the construction of related motives—see, for example, [Del71a, Sec. 5] and [Sch90a]. (We can also reduce questions about the cohomology of Shimura varieties with nontrivial coefficients to questions about the cohomology of mixed Shimura varieties with trivial coefficients in many more general cases—see the end of Section 5.1.4 below.)

5. Integral models

5.1. **PEL-type cases.** All currently known constructions of integral models of Shimura varieties (which we can describe in some reasonable detail) rely on the important coincidence that, as explained in Section 3.1.5, when $\mathrm{G} = \mathrm{Sp}_{2n}$ and Γ is a congruence subgroup of $\mathrm{G}(\mathbb{Z})$, the Siegel modular variety $\Gamma\backslash\mathcal{H}_n$ parameterizes polarized abelian varieties with level structures; and on the fact that we can define moduli problems over the integers that naturally extend such complex varieties. More generally, we can define the so-called *PEL moduli problems*, which parameterize abelian schemes with additional *PEL structures*. As we shall see below, the abbreviation *PEL* stands for *polarizations*, *endomorphism structures*, and *level structures*. This is the main topic of this subsection.

5.1.1. *PEL datum.* The definition of PEL moduli problems require some linear algebraic data. Suppose we are given an *integral PEL datum*

$$(\mathcal{O}, \star, L, \langle\,\cdot\,,\,\cdot\,\rangle, h_0), \tag{5.1.1.1}$$

whose entries can be explained as follows:

(1) $\mathcal{O}$ is an order in a finite-dimensional semisimple algebra over $\mathbb{Q}$.
(2) $\star$ is a positive involution of $\mathcal{O}$; namely, an anti-automorphism of order 2 such that $\mathrm{tr}_{\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}/\mathbb{R}}(xx^\star) > 0$ for every $x \neq 0$ in $\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}$.
(3) L is an $\mathcal{O}$-lattice; namely, a finitely generated free $\mathbb{Z}$-module L with the structure of an $\mathcal{O}$-module.
(4) $\langle\,\cdot\,,\,\cdot\,\rangle : L \times L \to \mathbb{Z}(1)$ is an alternating pairing satisfying the condition that

$$\langle bx, y\rangle = \langle x, b^\star y\rangle \tag{5.1.1.2}$$

for all $x, y \in L$ and $b \in \mathcal{O}$. Here

$$\mathbb{Z}(1) := \ker(\exp : \mathbb{C} \to \mathbb{C}^\times) = 2\pi i\,\mathbb{Z} \tag{5.1.1.3}$$

is the formal Tate twist of $\mathbb{Z}$.
(5) h_0 is an $\mathbb{R}$-algebra homomorphism

$$h_0 : \mathbb{C} \to \mathrm{End}_{\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}}(L \underset{\mathbb{Z}}{\otimes} \mathbb{R}), \tag{5.1.1.4}$$

satisfying the following conditions:
   (a) For any $z \in \mathbb{C}$ and $x, y \in L \underset{\mathbb{Z}}{\otimes} \mathbb{R}$, we have $\langle h_0(z)x, y\rangle = \langle x, h_0(\overline{z})y\rangle$, where $z \mapsto \overline{z}$ denotes the complex conjugation.
   (b) The (symmetric) $\mathbb{R}$-bilinear pairing $(2\pi i)^{-1}\langle\,\cdot\,, h_0(i)(\,\cdot\,)\rangle$ on $L \underset{\mathbb{Z}}{\otimes} \mathbb{R}$ is positive definite.

What do these mean? Given any integral PEL datum as above, we can write down a real torus

$$A_0 := (L \underset{\mathbb{Z}}{\otimes} \mathbb{R})/L \tag{5.1.1.5}$$

with a complex structure given by $h_0 : \mathbb{C} \to \mathrm{End}_{\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}}(L \underset{\mathbb{Z}}{\otimes} \mathbb{R})$, which can be shown to be an *abelian variety* (not just a complex torus) with a *polarization*

$$\lambda_0 : A_0 \to A_0^\vee \cong (L \underset{\mathbb{Z}}{\otimes} \mathbb{R})/L^\#, \tag{5.1.1.6}$$

where $A_0^\vee$ is the *dual abelian variety* of A_0, with

$$L^\# := \{x \in L \underset{\mathbb{Z}}{\otimes} \mathbb{Q} : \langle x, y\rangle \in \mathbb{Z}(1), \text{ for all } y \in L\} \tag{5.1.1.7}$$

the *dual lattice* of L with respect to $\langle\,\cdot\,,\,\cdot\,\rangle$, and where the λ_0 is induced by the natural inclusion $L \subset L^\#$ (see [Mum70, Ch. I, Sec. 2, and Ch. II, Sec. 9] and [Lan12a, Sec. 2.2]; cf. [Lan13, Sec. 1.3.2]). The assumptions that L is an $\mathcal{O}$-lattice and that the compatibility (5.1.1.2) holds imply that we have an *endomorphism structure*

$$i_0 : \mathcal{O} \to \mathrm{End}_{\mathbb{C}}(A_0) \tag{5.1.1.8}$$

satisfying the *Rosati condition*; namely, for each $b \in \mathcal{O}$, the diagram

$$\begin{tikzcd} A_0 \arrow[r, "\lambda_0"] \arrow[d, "i(b^\star)"'] & A_0^\vee \arrow[d, "(i(b))^\vee"] \\ A_0 \arrow[r, "\lambda_0"'] & A_0^\vee \end{tikzcd} \tag{5.1.1.9}$$

is commutative (cf. [Lan13, Sec. 1.3.3]). Moreover, for each integer $n \geq 1$, we have a *principal level-n structure*

$$(\alpha_{0,n} : L/nL \xrightarrow{\sim} A_0[n],\, \nu_{0,n} : (\mathbb{Z}/n\mathbb{Z})(1) \xrightarrow{\sim} \boldsymbol{\mu}_n), \tag{5.1.1.10}$$

where $A_0[n]$ denotes the n-torsion subgroup of A_0, which is canonically isomorphic to $(\frac{1}{n}L)/L$ because of (5.1.1.5), where $\alpha_{0,n}$ is induced by this canonical isomorphism, and where $\nu_{0,n}$ is induced by the exponential function $\exp : \mathbb{C} \to \mathbb{C}^\times$ (with kernel $\mathbb{Z}(1)$ by definition), so that the diagram

$$\begin{tikzcd} (L/nL) \times (L/nL) \arrow[r, "\langle\,\cdot\,,\,\cdot\,\rangle"] \arrow[d, "\alpha_{0,n} \times \alpha_{0,n}"', "\wr"] & (\mathbb{Z}/n\mathbb{Z})(1) \arrow[d, "\nu_{0,n}", "\wr"'] \\ A_0[n] \times A_0[n] \arrow[r, "\lambda_0\text{-Weil pairing}"'] & \boldsymbol{\mu}_n \end{tikzcd} \tag{5.1.1.11}$$

is commutative (see [Lan12a, Sec. 2.2 and 2.3]; cf. [Lan13, Sec. 1.3.6]), and so that such a level structure is *liftable* to all higher levels n' divisible by n. (The second entry $\nu_{0,n}$ in $(\alpha_{0,n}, \nu_{0,n})$ is often omitted in the notation, although $\nu_{0,n}$ is not determined by $\alpha_{0,n}$ when $\langle\,\cdot\,,\,\cdot\,\rangle$ is not a perfect pairing modulo n.) Furthermore, the homomorphism $h_0 : \mathbb{C} \to \mathrm{End}_{\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}}(L \underset{\mathbb{Z}}{\otimes} \mathbb{R})$ defines a *Hodge decomposition*

$$H_1(A_0, \mathbb{C}) \cong L \underset{\mathbb{Z}}{\otimes} \mathbb{C} = V_0 \oplus \overline{V}_0, \tag{5.1.1.12}$$

where

$$V_0 := \{x \in L \underset{\mathbb{Z}}{\otimes} \mathbb{C} : h_0(z)x = (1 \otimes z)x, \text{ for all } z \in \mathbb{C}\} \tag{5.1.1.13}$$

and

$$\overline{V}_0 := \{x \in L \underset{\mathbb{Z}}{\otimes} \mathbb{C} : h_0(z)x = (1 \otimes \overline{z})x, \text{ for all } z \in \mathbb{C}\}. \tag{5.1.1.14}$$

Then we also have

$$\mathrm{Lie}\, A_0 \cong V_0 \tag{5.1.1.15}$$

as $\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{C}$-modules. The *reflex field* F_0 for h_0 (which is defined in this context without referring to any Shimura datum) is the field of definition of the isomorphism class of V_0 as an $\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{C}$-module, or more precisely the subfield

$$F_0 := \mathbb{Q}(\operatorname{tr}(b|V_0) : b \in \mathcal{O}) \tag{5.1.1.16}$$

of $\mathbb{C}$ (see [Lan13, Sec. 1.2.5]). This is the smallest subfield of $\mathbb{C}$ over which we can formulate a trace or determinantal condition on the Lie algebra of an abelian variety, to ensure that the pullback to $\mathbb{C}$ of its Lie algebra is isomorphic to V_0 as an $\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{C}$-module (see [Lan13, Sec. 1.2.5 and 1.3.4]).

The upshot is that, by combining all of these, we obtain a tuple

$$(A_0, \lambda_0, i_0, (\alpha_{0,n}, \nu_{0,n})), \tag{5.1.1.17}$$

which is an abelian variety with a *polarization*, an *endomorphism structure*, and a *level structure*, such that $\operatorname{Lie} A_0$ satisfies the Lie algebra condition defined by h_0. These additional structures form the so-called *PEL structures*. Thus, writing down an integral PEL datum means writing down such a tuple.

5.1.2. *Smooth PEL moduli problems.* Given any integral PEL datum as in (5.1.1.1), we can define the associated *PEL moduli problems* (at varying levels) over (the category of schemes over) F_0, or over suitable localizations of the ring of integers $\mathcal{O}_{F_0}$. Roughly speaking, they are the smallest moduli problems (over the respective base rings) parameterizing tuples of abelian schemes with additional PEL structures as above, including the one we wrote down in (5.1.1.17). See [Lan13, Sec. 1.4.1] for more details on defining *smooth* PEL moduli problems over localizations of $\mathcal{O}_{F_0}$ of *good* residue characteristics. (A prime number p is good for $(\mathcal{O}, \star, L, \langle\,\cdot\,,\,\cdot\,\rangle, h_0)$ if it satisfies all of the following conditions: p is unramified in $\mathcal{O}$, as in [Lan13, Def. 1.1.1.18]; $p \nmid [L^\# : L]$; $p \neq 2$ whenever $(\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}, \star)$ has a simple factor isomorphic to a matrix algebra over the Hamiltonian numbers $\mathbb{H}$ with its canonical positive involution; and there is no nontrivial level structure at p. See [Lan13, Def. 1.4.1.1].) Note that people often talk about *good reductions* instead of *smoothness* when describing integral models, even though it is not always clear what reductions should mean without some properness assumption.

The moduli problems above are defined using the language of *isomorphism classes*. If we use the language of *isogeny classes* instead, then it is more natural to work with *rational PEL datum*

$$(\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{Q}, \star, L \underset{\mathbb{Z}}{\otimes} \mathbb{Q}, \langle\,\cdot\,,\,\cdot\,\rangle, h_0), \tag{5.1.2.1}$$

which is good for defining PEL moduli problems over F_0 (in characteristic zero) up to isogenies of all possible degrees, or with its p-integral version

$$(\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{Z}_{(p)}, \star, L \underset{\mathbb{Z}}{\otimes} \mathbb{Z}_{(p)}, \langle\,\cdot\,,\,\cdot\,\rangle, h_0), \tag{5.1.2.2}$$

which is good for defining PEL moduli problems over $\mathcal{O}_{F_0,(p)}$ only up to isogenies of degrees prime to p. The latter was done in [Kot92, Sec. 5] and generalized in [Lan13, Sec. 1.4.2], and then compared in [Lan13, Sec. 1.4.3] with the definition by isomorphism classes in [Lan13, Sec. 1.4.3]. (While the comparison was not difficult, the author still found it helpful to have the details recorded in the literature.)

5.1.3. *PEL-type Shimura varieties.* The PEL moduli problems as above are related to the double coset spaces in Section 2 as follows. Firstly, the subtuple $(\mathcal{O}, \star, L, \langle\,\cdot\,,\,\cdot\,\rangle)$ of (5.1.1.1) defines a group scheme G over $\mathbb{Z}$ by

(5.1.3.1)
$$\mathrm{G}(R) := \{(g, r) \in \mathrm{End}_{\mathcal{O}}(L \underset{\mathbb{Z}}{\otimes} R) \times R^\times : \langle gx, gy\rangle = r\langle x, y\rangle, \text{ for all } x, y \in L \underset{\mathbb{Z}}{\otimes} R\},$$

for each ring R. The elaborate conditions for a prime p to be good, in the first paragraph of Section 5.1.2, then imply that $\mathrm{G}(\mathbb{Z}_p)$ is a so-called *hyperspecial* open compact subgroup of $\mathrm{G}(\mathbb{Q}_p)$ when $\mathrm{G}_{\mathbb{Q}_p}$ is connected as an algebraic group over $\mathbb{Q}_p$. Secondly, the remaining entry $h_0 : \mathbb{C} \to \mathrm{End}_{\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{R}}(L \underset{\mathbb{Z}}{\otimes} \mathbb{R})$ defines a homomorphism

$$h_0 : \mathbb{C}^\times \to \mathrm{G}(\mathbb{R}), \tag{5.1.3.2}$$

whose $\mathrm{G}(\mathbb{R})$-conjugacy class defines a manifold

$$\mathcal{D} := \mathrm{G}(\mathbb{R}) \cdot h_0. \tag{5.1.3.3}$$

Then $(\mathrm{G}, \mathcal{D})$ defines a double coset space $\mathsf{X}_{\mathcal{U}}$ as in (2.2.15), for each neat open compact subgroup $\mathcal{U}$ of $\mathrm{G}(\mathbb{A}^\infty)$. When $(\mathrm{G}, \mathcal{D})$ is a Shimura datum (which is, however, not always the case), we say it is a *PEL-type Shimura datum*, and say $\mathsf{X}_{\mathcal{U}}$ is a *PEL-type Shimura variety*. In this case, the reflex field F_0 for h_0 coincides with the reflex field of $(\mathrm{G}, \mathcal{D})$ (see Theorem 2.4.3).

Example 5.1.3.4. Suppose $\mathcal{O} = \mathbb{Z}$; $\star$ is trivial; $L = \mathbb{Z}^{2n}$; $\langle\,\cdot\,,\,\cdot\,\rangle$ is defined by the matrix $2\pi i \begin{pmatrix} & 1_n \\ -1_n & \end{pmatrix}$; and $h_0(e^{i\theta}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, for all $\theta \in \mathbb{R}$. Then $\mathrm{G} \cong \mathrm{GSp}_{2n}$ (see Example 2.1.2), and $\mathcal{D} = \mathrm{G}(\mathbb{R}) \cdot h_0 \cong \mathcal{H}_n^\pm$ (see (3.1.4.6)). This is the important special case of Siegel moduli problems.

Example 5.1.3.5. Suppose $\mathcal{O} = \mathcal{O}_E$ is an order in an imaginary quadratic extension field E of $\mathbb{Q}$, with a fixed isomorphism $E \underset{\mathbb{Q}}{\otimes} \mathbb{R} \cong \mathbb{C}$; $\star$ is the complex conjugation of E over $\mathbb{Q}$; $L = \mathcal{O}^{a+b}$ for some integers $a \geq b \geq 0$; $\langle\,\cdot\,,\,\cdot\,\rangle$ is defined by $2\pi i$ times the composition of the skew-Hermitian pairing defined by some matrix $\begin{pmatrix} & & 1_b \\ & \varepsilon 1_{a-b} & \\ -1_b & & \end{pmatrix}$ with $\mathrm{tr}_{\mathcal{O}/\mathbb{Z}} : \mathcal{O} \to \mathbb{Z}$, where $\varepsilon \in \mathcal{O}$ satisfies $-i\varepsilon \in \mathbb{R}_{>0}$; and $h_0(e^{i\theta}) = \begin{pmatrix} \cos\theta & & -\sin\theta \\ & e^{-i\theta} & \\ \sin\theta & & \cos\theta \end{pmatrix}$, for all $\theta \in \mathbb{R}$. Then $\mathrm{G}(\mathbb{R}) \cong \mathrm{GU}_{a,b} \cong \mathrm{GU}'_{a,b}$, and $\mathcal{D} = \mathrm{G}(\mathbb{R}) \cdot h_0 \cong \mathcal{H}_{a,b}^\pm$ (see (3.2.5.9)).

Example 5.1.3.6. Suppose $\mathcal{O}$ is an order in a quaternion algebra B over $\mathbb{Q}$ such that $B \underset{\mathbb{Q}}{\otimes} \mathbb{R} \cong \mathbb{H}$ and $\mathcal{O}$ is stabilized by the canonical positive involution $\star$; $L = \mathcal{O}^{2n}$; $\langle\,\cdot\,,\,\cdot\,\rangle$ is $2\pi i$ times the composition of the skew-Hermitian pairing defined by the matrix $\begin{pmatrix} & 1_n \\ -1_n & \end{pmatrix}$ with $\mathrm{tr}_{\mathcal{O}/\mathbb{Z}} : \mathcal{O} \to \mathbb{Z}$; and $h_0(e^{i\theta}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, for all $\theta \in \mathbb{R}$. Then the derived group of $\mathrm{G}(\mathbb{R})$ is isomorphic to SO^*_{2n}, and $\mathcal{D} = \mathrm{G}(\mathbb{R}) \cdot h_0$ contains $\mathcal{H}_{\mathrm{SO}^*_{2n}}$ as a connected component (see Section 3.3.1). However, G is not connected as an algebraic group over $\mathbb{Q}$.

These three examples are prototypical for integral PEL data and the associated PEL moduli problems. By the classification of positive involutions (see the review in [Lan13, Prop. 1.2.1.13 and 1.2.1.14] based on Albert's classification, or just the classification over $\mathbb{R}$ in [Kot92, Sec. 2]), up to cocenter (i.e., up to replacing the group with a subgroup containing the derived group), $\mathrm{G}(\mathbb{R})$ factorizes as a product of groups of the form $\mathrm{Sp}_{2n}(\mathbb{R})$, $\mathrm{U}_{a,b}$, or SO^*_{2n} (with varying values of $n \geq 0$ and $a \geq b \geq 0$). Thus, the connected components of $\mathcal{D}$ are products of the Hermitian symmetric domains we have seen in Sections 3.1, 3.2, and 3.3. However, we never see the Hermitian symmetric domains in Sections 3.4, 3.6, and 3.5 as factors, except when it is because of the *accidental* (or *special*) isomorphisms $\mathcal{H}^+_{\mathrm{SO}_{1,2}} \cong \mathcal{H} \cong \mathcal{H}_{1,1}$, $\mathcal{H}^+_{\mathrm{SO}_{2,2}} \cong \mathcal{H} \times \mathcal{H}$, $\mathcal{H}^+_{\mathrm{SO}_{3,2}} \cong \mathcal{H}_2$, $\mathcal{H}^+_{\mathrm{SO}_{4,2}} \cong \mathcal{H}_{2,2}$, and $\mathcal{H}^+_{\mathrm{SO}_{6,2}} \cong \mathcal{H}_{\mathrm{SO}^*_8}$ in low dimensions (see [Hel01, Ch. X, Sec. 6]).

For a PEL moduli problem defined by an integral PEL datum with $\mathcal{O} \underset{\mathbb{Z}}{\otimes} \mathbb{Q}$ a *simple algebra* over $\mathbb{Q}$, the complex Lie group $\mathrm{G}(\mathbb{C})$ factorizes up to cocenter (i.e., again, up to replacing the group with a subgroup containing the derived group) as a product of symplectic, special linear group, or special orthogonal groups, but only one type of them can occur. Accordingly, we say the PEL moduli problem is of type C, type A, or type D. People often also say that a PEL moduli problem of type A is a unitary Shimura variety, even when G violates Condition (3) in Section 2.3 and so $(\mathrm{G}, \mathcal{D})$ is not a Shimura datum. However, by orthogonal Shimura varieties, people might not mean PEL moduli problems of type D. On one hand, this is because there are useful abelian-type Shimura varieties associated with special orthogonal groups, as we shall see in Example 5.2.2.11 below. On the other hand, since the groups G associated with PEL moduli problems of type D are not even connected, the corresponding pairs $(\mathrm{G}, \mathcal{D})$ violate the conditions for being Shimura data in a more fundamental way, and hence there is some ambiguity in what PEL-type Shimura varieties of type D should mean.

In all cases, whether $(\mathrm{G}, \mathcal{D})$ is a Shimura datum or not, when the corresponding (smooth) PEL moduli problem $\mathsf{M}_{\mathcal{U}}$ is defined at level $\mathcal{U}$, there is an open and closed immersion

$$\mathsf{X}_{\mathcal{U}} \hookrightarrow \mathsf{M}_{\mathcal{U}}(\mathbb{C}) \tag{5.1.3.7}$$

(see [Lan12a, Sec. 2.4 and 2.5]), which is generally not an isomorphism, due to the so-called *failure of Hasse's principle* (see [Lan13, Rem. 1.4.3.12] and [Kot92, Sec. 7–8]). More concretely, this is because the definition of $\mathsf{M}_{\mathcal{U}}$ uses only the adelic group $\mathrm{G}(\mathbb{A})$—the level structures use $\mathrm{G}(\mathbb{A}^\infty)$, while the Lie algebra condition uses $\mathrm{G}(\mathbb{R})$. But there can be more than one mutually nonisomorphic algebraic groups $\mathrm{G}^{(1)}, \ldots, \mathrm{G}^{(m)}$ (including G) over $\mathbb{Q}$ such that $\mathrm{G}^{(j)}(\mathbb{A}) \cong \mathrm{G}(\mathbb{A})$, for all j, and such that $\mathsf{M}_{\mathcal{U}}(\mathbb{C})$ is the disjoint union of the $\mathsf{X}^{(j)}_{\mathcal{U}}$'s defined using $\mathrm{G}^{(j)}$'s. (When the pairs $(\mathrm{G}^{(j)}, \mathcal{D})$'s are Shimura data, by checking the action of $\mathrm{Aut}(\mathbb{C}/F_0)$ on the dense subset of the so-called *CM points*, it can be shown that this disjoint union descends to a disjoint union over F_0 of the canonical models of the $\mathsf{X}^{(j)}_{\mathcal{U}}$'s.) In cases of types A and C, such a failure is harmless in practice, because all the $\mathsf{X}^{(j)}_{\mathcal{U}}$'s are isomorphic to each other, even at the level of canonical models over F_0 (see [Kot92, Sec. 8]). However, in cases of type D, such a failure can be quite serious. It can even happen that some of the $\mathsf{X}^{(j)}_{\mathcal{U}}$'s are noncompact, while others are not (see [Lan15, Ex. A.7.2]).

With the tacit running assumption that $\mathcal{U}$ is *neat* (see [Pin89, 0.6] again, or see [Lan13, Def. 1.4.1.8]), $\mathsf{M}_{\mathcal{U}}$ is a (smooth) quasi-projective scheme because it is finite over the base change of some Siegel moduli scheme (see Section 3.1.5), and hence an open and closed subscheme of $\mathsf{M}_{\mathcal{U}}$ serves as an integral model of $\mathsf{X}_{\mathcal{U}}$. (When $\mathcal{U}$ is not neat, we still obtain a Deligne–Mumford stack whose associated coarse moduli space is a quasi-projective scheme.)

Remark 5.1.3.8. The consideration of certain Shimura varieties or connected Shimura varieties as parameter spaces for complex abelian varieties with PEL structures has a long history—see, for example, Shimura's articles [Shi63b] and [Shi66]. The theory over the complex numbers is already quite complicated and heavy in notation, which seems inevitable for anything involving semisimple algebras and positive involutions, but there is an additional subtlety when formulating the PEL structures over the integers. That is, we can no longer talk about the singular homology or cohomology of an abelian variety over a general base field, let alone their relative versions for an abelian scheme over a general base scheme. Instead, we have resorted to the (relative) de Rham and étale homology or cohomology. This is the reason that PEL moduli problems suffer from the failure of Hasse's principle, and that these moduli problems can only be easily defined in mixed characteristics $(0, p)$ when there is no ramification or level at p. (In Section 5.1.5 below, we will summarize the known constructions when there are indeed some ramification and levels at p.)

5.1.4. *Toroidal and minimal compactifications of PEL moduli.* The above PEL moduli $\mathsf{M}_{\mathcal{U}}$ carry *universal* or *tautological* objects, which are abelian schemes with additional PEL structures, essentially by definition.

In the important special case of Siegel moduli problems considered in Example 5.1.3.4, the tautological objects are just principally polarized abelian schemes with level structures (but with no nontrivial endomorphism structures or polarization degrees specified in the moduli problem). By vastly generalizing results due to Mumford (see [Mum72]), Faltings and Chai constructed toroidal compactifications of the Siegel moduli problems by studying the degeneration of polarized abelian schemes into semi-abelian schemes (see [FC90, Ch. III and IV]). Over such toroidal compactifications, they can algebro-geometrically define modular forms, and construct minimal compactifications of Siegel moduli problems as the projective spectra of some graded algebras of modular forms (with inputs of positivity results from [MB85]; see [FC90, Ch. V]). The main difficulty in Faltings and Chai's construction was that they needed to glue together boundary strata parameterizing degenerations of *different ranks*. This was much more complicated than the earlier boundary construction in [Rap78], based on the original results of [Mum72].

The generalization of the results of [FC90, Ch. III–V] to all smooth PEL moduli problems was carried out in [Lan13]. Later in [Lan12a], it was shown that the toroidal and minimal compactifications in [Lan13] (and also in [FC90, Ch. III–V], both based on the theory of degeneration) are indeed compatible with the toroidal and minimal compactifications in [Har89] and [Pin89] (based on the constructions in [BB66] and [AMRT75] using analytic coordinate charts). In other words, toroidal and minimal compactifications of PEL moduli problems provide integral models of toroidal and minimal compactifications of PEL-type Shimura varieties.

Over the integral models of PEL-type Shimura varieties and their toroidal compactifications, by considering the relative cohomology of certain *Kuga families* (which are abelian scheme torsors which are integral models of certain *mixed Shimura varieties* as in Section 4.2.8) and their toroidal compactifications, we can also define automorphic vector bundles and their canonical and subcanonical extensions over these integral models (see [FC90, Ch. VI], [Lan12c, especially Sec. 6], [LS12, Sec. 1–3], and [LS13, Sec. 4–5]), whose global sections and cohomology classes then provide an algebro-geometric definition of modular forms in mixed characteristics, generalizing the theory over $\mathbb{C}$ in Section 4.2.7. See [Lan12b] and [Lan19] again for some surveys on this topic. See also [Hid02] and [Hid04], for example, for some application to the theory of p-adic modular forms, with the assumptions on the integral models of toroidal and minimal compactifications justified by the constructions in [Lan13].

5.1.5. *Integral models not defined by smooth moduli.* Let p be any prime number, which is not necessarily good for the integral PEL datum. (The conditions for being good for the integral PEL datum are the same as the conditions for being good for defining the PEL moduli, except that we ignore the condition of having no nontrivial level structure at p.)

Let $\widehat{\mathbb{Z}}^p := \varprojlim_{p \nmid N} (\mathbb{Z}/N\mathbb{Z})$ and $\mathbb{A}^{\infty,p} := \widehat{\mathbb{Z}}^p \underset{\mathbb{Z}}{\otimes} \mathbb{Q}$, with the superscript "$p$" meaning "away from p". When the level $\mathcal{U} \subset \mathrm{G}(\mathbb{A}^\infty)$ is of the form $\mathcal{U} = \mathcal{U}^p\mathcal{U}_p$, where $\mathcal{U}^p \subset \mathrm{G}(\mathbb{A}^{\infty,p})$ and $\mathcal{U}_p \subset \mathrm{G}(\mathbb{Q}_p)$ such that $\mathcal{U}_p$ is a *parahoric* open compact subgroup of $\mathrm{G}(\mathbb{Q}_p)$—that is, when $\mathcal{U}_p$ is the identity component of the stabilizer of some *multichain of* $\mathbb{Z}_p$*-lattices* in $L \underset{\mathbb{Z}}{\otimes} \mathbb{Q}_p$—there is also a natural way to define a PEL moduli problem parameterizing not just individual abelian schemes with additional PEL structures, but rather *multichains of isogenies* of them. (See [RZ96, Ch. 3 and 6].) This generalizes the idea that, in the case of modular curves, the moduli problem at level $\Gamma_0(p)$ parameterizes degree-p cyclic isogenies of elliptic curves. (See, for example, [KM85, Sec. 3.4 and 6.6], which gave a moduli interpretation of the corresponding construction in [DR73] that was by merely taking normalizations.)

PEL moduli problems at parahoric levels at p are generally far from smooth in mixed characteristics $(0, p)$, in which case we say that the integral models have *bad reductions*. Nevertheless, people have developed a theory of *local models* which describes the singularities of such moduli problems—see, for example, [RZ96], [PR03], [PR05], [PR09], and [PZ13]. (See also [Rap05], [Hai05], and [PRS13] for some surveys.) Beginners should be warned that the forgetful morphisms from a higher parahoric level to a lower one, or more precisely from one such moduli problem defined by some multichain of $\mathbb{Z}_p$-lattices to another one defined by a simpler multichain, is usually just projective but not finite. This is very different from what we have seen in characteristic zero (or in the classical modular curve case in [DR73] and [KM85]).

For higher levels at p, the older ideas in [DR73] by taking normalizations, and in [KM85] by generalizing the so-called *Drinfeld level structures*, are both useful. But they are useful in rather different contexts—exactly which strategy is more useful depends on the intended applications.

The idea of generalizing *Drinfeld level structures* played important roles in [HT01], [Man04], [Man05], and [Man11], and in subsequent works based on them such as [Shi11], where people needed to compute the cohomology of certain nearby cycles and *count points*, and hence would prefer to work with moduli problems even when they are not *flat*. (It was known much earlier—see [CN90, Appendix A]—that moduli problems defined by Drinfeld level structures are generally not even flat over the base rings.) But such moduli problems tend to be very complicated, and might be too cumbersome when it comes to the construction of compactifications, because most known techniques of gluing require at least some noetherian normal base schemes. (The results in [HT01], [Man04], [Man05], [Man11], and [Shi11] were all for compact PEL-type Shimura varieties.)

On the contrary, the idea of taking normalizations produces, essentially by definition, noetherian normal schemes flat over the base rings. It had been overlooked for a long time because of its lack of moduli interpretations, and because no theory of local models was available for them. While these two problems still remain, recent developments beyond the PEL-type cases (see the discussion on the Hodge-type cases below in Section 5.2.3), and also developments in the construction of p-adic modular forms (as in [HLTT16], which was based on [Lan18a] and did not require the integral models to be constructed as moduli problems) motivated people to systematically reconsider the constructions by taking normalizations. In the PEL-type cases, we can construct good compactifications for all such integral models (and their variants)—see [Lan16], [Lan17], and [Lan18b]. Such compactifications are good in the sense that the complications in describing its local structures and its boundary stratifications are completely independent of each other. This is useful in, for example, [LS18b, Sec. 6.3], which generalized the results of [Man05] and [Man11] to all noncompact cases, without having to compactify any moduli problems defined by Drinfeld level structures.

In [LS18a], it is shown that many familiar subsets of the reductions of the integral models of Shimura varieties in characteristic $p > 0$, such as *p-rank strata*, *Newton strata*, *Oort central leaves*, *Ekedahl–Oort strata*, and *Kottwitz–Rapoport strata*, and their pullbacks to the integral models defined by taking normalizations at arbitrarily higher levels at p, admit partial toroidal and minimal compactifications with many properties similar to the whole toroidal and minimal compactifications.

5.1.6. *Langlands–Rapoport conjecture.* In [Lan76] and [Lan77], where the terminology *Shimura varieties* first appeared in the literature, Langlands proposed to study Hasse–Weil zeta functions of Shimura varieties and relate them to the automorphic L-functions, generalizing earlier works for GL_2 and its inner forms by Eichler, Shimura, Kuga, Sato, and Ihara. This is very ambitious—see the table of contents of [LR92] for the types of difficulties that already showed up in the case of Picard modular surfaces. This motivated the *Langlands–Rapoport conjecture*, which gives a detailed description of the points of the reduction modulo p of a general Shimura variety at a hyperspecial level at p—see [LR87], [Pfa96], [Rei97], and [Rap05]. See also [Kot90] for a variant of the conjecture due to Kottwitz, which was proved by himself for all PEL-type Shimura varieties of types A and C in [Kot92].

5.2. **Hodge-type and abelian-type cases.**

5.2.1. *Hodge-type Shimura varieties.*

Definition 5.2.1.1. *A* **Hodge-type Shimura datum** $(\mathrm{G}, \mathcal{D})$ *is a Shimura datum such that there exists a* **Siegel embedding**

$$(\mathrm{G}, \mathcal{D}) \hookrightarrow (\mathrm{GSp}_{2n}, \mathcal{H}_n^{\pm}); \tag{5.2.1.2}$$

namely, an injective homomorphism $\mathrm{G} \hookrightarrow \mathrm{GSp}_{2n}$ *of algebraic groups over* $\mathbb{Q}$ *inducing an (automatically closed) embedding* $\mathcal{D} \hookrightarrow \mathcal{H}_n^{\pm}$, *for some* $n \geq 0$ *(cf. Remarks* 2.4.11 *and* 2.4.13*). A* **Hodge-type Shimura variety** *is a Shimura variety associated with a Hodge-type Shimura datum.*

If $\mathcal{D} = \mathrm{G}(\mathbb{R}) \cdot h_0$ is the conjugacy class of some homomorphism $h_0 : \mathbb{C}^\times \to \mathrm{G}(\mathbb{R})$ as before, then the above means the composition of h_0 with $\mathrm{G}(\mathbb{R}) \hookrightarrow \mathrm{GSp}_{2n}(\mathbb{R})$ lies in the conjugacy class of an analogous homomorphism for $\mathrm{GSp}_{2n}(\mathbb{R})$ as in (3.1.4.7). As explained in Remark 2.4.13, each neat open compact subgroup $\mathcal{U}$ of $\mathrm{G}(\mathbb{A}^\infty)$ is the pullback of some neat open compact subgroup $\mathcal{U}'$ of $\mathrm{GSp}_{2n}(\mathbb{A}^\infty)$ such that the Shimura variety $\mathsf{X}_{\mathcal{U}}$ associated with $(\mathrm{G}, \mathcal{D})$ at level $\mathcal{U}$ is a special subvariety of the Siegel modular variety at level $\mathcal{U}'$. Then the points of $\mathsf{X}_{\mathcal{U}}$ are equipped with polarized abelian varieties with level structures, which are pulled back from the Siegel modular varieties (and hence depend on the choice of the Siegel embedding above).

Example 5.2.1.3. All the PEL-type Shimura data are, essentially by definition, of Hodge type, because of the definition of G in (5.1.3.1), and because of the condition on h_0 in Section 5.1.1. But there are many Hodge-type Shimura data which are not of PEL type. Important examples of this kind are given by *general spin groups* (with associated Hermitian symmetric domains as in Section 3.4.3), which admit Siegel embeddings thanks to the so-called *Kuga–Satake construction* (see [Sat66], [Sat68], and [KS67]; see also the reviews in [Del72, Sec. 3–4] and [MP16, Sec. 1 and 3]).

The *injectivity* of the homomorphism $\mathrm{G} \hookrightarrow \mathrm{GSp}_{2n}$ means we have a faithful $2n$-dimensional symplectic representation of G. By [DMOS82, Ch. I, Prop. 3.1], the faithfulness of this representation implies that $\mathrm{G}(\mathbb{Q})$ is the subgroup of $\mathrm{GSp}_{2n}(\mathbb{Q})$ fixing some tensors (s_i) in the tensor algebra generated by $\mathbb{Q}^{2n}$ and its dual. But since $\mathbb{C}^\times \stackrel{h_0}{\to} \mathrm{G}(\mathbb{R}) \to \mathrm{GSp}_{2n}(\mathbb{R})$ defines a Hodge structure on $\mathbb{R}^{2n}$, by [DMOS82, Ch. I, Prop. 3.4], these tensors (s_i) are *Hodge tensors*; namely, tensors of weight 0 and Hodge type $(0,0)$ with respect to the induced Hodge structure on the tensor algebra of $\mathbb{R}^{2n}$.

Then the Shimura varieties $\mathsf{X}_{\mathcal{U}}$ associated with $(\mathrm{G}, \mathcal{D})$ can be interpreted as parameter spaces of polarized abelian varieties (over $\mathbb{C}$) equipped with some *Hodge cycles* and level structures. According to Deligne (see [DMOS82, Ch. I, Main Thm. 2.11]), Hodge cycles on complex abelian varieties are *absolute Hodge cycles*. Roughly speaking, this means that the Hodge cycles can be conjugated, as if they were known to be defined by algebraic cycles (as if the Hodge conjecture were true for them), when the underlying abelian varieties are conjugated by an automorphism of $\mathbb{C}$. This defines a canonical action of $\mathrm{Aut}(\mathbb{C}/F_0)$ on the points of $\mathsf{X}_{\mathcal{U}}$ by conjugating the Hodge cycles of the corresponding abelian varieties, which coincides with the action given by the structure of $\mathsf{X}_{\mathcal{U}}$ as the $\mathbb{C}$-points of its canonical model over F_0. Because of this, people can formulate some *moduli interpretation* for the points of the canonical model of $\mathsf{X}_{\mathcal{U}}$ over F_0.

Such a pointwise moduli interpretation has not really helped people write down any useful moduli problem, but they are still important for studying the less direct construction by taking normalizations, to be introduced in Section 5.2.3 below.

5.2.2. *Abelian-type Shimura varieties.*

Definition 5.2.2.1. *An* **abelian-type Shimura datum** $(\mathrm{G}, \mathcal{D})$ *is a Shimura datum such that there exist a Hodge-type Shimura datum* $(\mathrm{G}_1, \mathcal{D}_1)$ *and a central isogeny*

$$\mathrm{G}_1^{\mathrm{der}} \to \mathrm{G}^{\mathrm{der}} \tag{5.2.2.2}$$

between the derived groups which induces an isomorphism

$$(\mathrm{G}_1^{\mathrm{ad}}, \mathcal{D}_1^{\mathrm{ad}}) \xrightarrow{\sim} (\mathrm{G}^{\mathrm{ad}}, \mathcal{D}^{\mathrm{ad}}) \tag{5.2.2.3}$$

between the adjoint quotients. An **abelian-type Shimura variety** *is a Shimura variety associated with an abelian-type Shimura datum.*

By definition, every Hodge-type Shimura datum (resp. Shimura variety) is of abelian type. Essentially by definition, the connected components of an abelian-type Shimura variety are finite quotients of those of some Hodge-type Shimura variety. Thus, results for abelian-type Shimura varieties are often proved by reducing to the case of Hodge-type Shimura variety.

Example 5.2.2.4. The simplest and perhaps the most prominent examples of abelian-type Shimura varieties that are not of Hodge type are the *Shimura curves*, which are associated with Shimura data $(\mathrm{G}, \mathcal{D})$ with $\mathrm{G}(\mathbb{Q}) = B^\times$, where B is a quaternion algebra over a totally real extension F of $\mathbb{Q}$ of degree $d > 1$ such that

$$B \underset{\mathbb{Q}}{\otimes} \mathbb{R} \cong \mathrm{M}_2(\mathbb{R}) \times \mathbb{H}^{d-1}, \tag{5.2.2.5}$$

in which case

$$\mathcal{D} \cong \mathcal{H}^\pm \tag{5.2.2.6}$$

is one-dimensional. More generally, there are abelian-type Shimura data $(\mathrm{G}, \mathcal{D})$ with $\mathrm{G}(\mathbb{Q}) = B^\times$, where B is a quaternion algebra over F such that

$$B \underset{\mathbb{Q}}{\otimes} \mathbb{R} \cong \mathrm{M}_2(\mathbb{R})^a \times \mathbb{H}^b \tag{5.2.2.7}$$

for some $a, b \geq 1$, in which case

$$\mathcal{D} \cong (\mathcal{H}^\pm)^a \tag{5.2.2.8}$$

(cf. [Shi71, Sec. 9.2]). Even more generally, as explained in [Del71b, Sec. 6], with the same B and F (and $a, b \geq 1$), for each integer $n \geq 1$, there are abelian-type Shimura data $(\mathrm{G}, \mathcal{D})$ such that $\mathrm{G}(\mathbb{Q})$ is the group of similitudes over F of a symmetric bilinear pairing over B^n, and such that

$$\mathrm{G}(\mathbb{R}) \cong \mathrm{GSp}_{2n}(\mathbb{R})^a \times \mathrm{H}^b, \tag{5.2.2.9}$$

where H is a real form of $\mathrm{GSp}_{2n}(\mathbb{C})$ with compact derived group (which is $\mathbb{H}^\times$ when $n = 1$), in which case

$$\mathcal{D} \cong (\mathcal{H}_n^\pm)^a. \tag{5.2.2.10}$$

In all these cases, the related Hodge-type Shimura data (as in Definition 5.2.2.1) and their Siegel embeddings (as in Definition 5.2.1.1) can be constructed with some auxiliary choice of imaginary quadratic field extensions of $\mathbb{Q}$ (see [Del71b, Sec. 6], and also [Car86, Sec. 2]). However, none of these are of Hodge type (let alone of PEL type) because the Hodge structures defined by rational representations of G are not *rational Hodge structures* (see [Mil05, Sec. 9, part (c) of the last example]).

Example 5.2.2.11. There are also the abelian-type Shimura varieties associated with Shimura data $(\mathrm{G}, \mathcal{D})$ such that the derived group of G as an algebraic group over $\mathbb{Q}$ is the special orthogonal group defined by some symmetric bilinear form over $\mathbb{Q}$ of signature $(a, 2)$ over $\mathbb{R}$, for some $a \geq 1$, so that $\mathrm{G}(\mathbb{R})^+$ is up to cocenter the group $\mathrm{SO}_{a,2}(\mathbb{R})^+$ studied in Section 3.4.3. (See also the last paragraph of Section 3.4.4.) In this case, the related Hodge-type Shimura data (as in Definition 5.2.2.1) are general spin groups, with their Siegel embeddings (as in Definition 5.2.1.1) given by the Kuga–Satake construction, as in Example 5.2.1.3, mapping G into GSp_{2n} with $n = 2^a$ (see the same references in Example 5.2.1.3). Such Shimura varieties associated with special orthogonal groups are important because they are related to the moduli of interesting algebraic varieties. For example, when $a = 19$ in the above, the corresponding Shimura varieties contain as open subspaces the moduli spaces of *polarized K3 surfaces* with level structures (see, for example, [Riz06, Sec. 6], [MP15, Sec. 2 and 4], and [She17, Sec. 4.2]).

In [Del79, 2.3], based on earlier works due to Satake (see [Sat65]; and see also [Sat67] and [Sat80, Ch. IV]), Deligne analyzed all abelian-type Shimura data $(\mathrm{G}, \mathcal{D})$ that are *adjoint* and *simple* in the sense that G is adjoint and simple as an algebraic group over $\mathbb{Q}$. (See also [Mil13, Sec. 10].) When G is simple over $\mathbb{Q}$, there is up to isomorphism a unique irreducible Dynkin diagram associated with all the simple factors of $\mathrm{G}_{\mathbb{C}}$, or rather $\operatorname{Lie}\mathrm{G}(\mathbb{C})$. (In fact, in this case, $\mathrm{G}^{\mathrm{ad}} \cong \mathrm{Res}_{k/\mathbb{Q}}\mathrm{H}$ for some number field k and some connected semisimple algebraic group algebraic group H over k such that $\mathrm{H}_{\overline{\mathbb{Q}}}$ remains simple over the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ in $\mathbb{C}$; see, for example, [BT65, 6.21(ii)].) We say that the group is of type A_n, B_n, C_n, D_n, E_6, or E_7 accordingly to the type of such a Dynkin diagram (cf. the summary in Section 3.7). In cases of type D_n, for $n \geq 5$, we further single out the special case of type $\mathrm{D}_n^{\mathbb{R}}$ (resp. $\mathrm{D}_n^{\mathbb{H}}$) if all the noncompact simple factors of $\operatorname{Lie}\mathrm{G}(\mathbb{R})$ are isomorphic to $\mathfrak{so}_{2n-2,2}$ (resp. $\mathfrak{so}^*_{2n}$). In cases of type D_4, the situation is more complicated because $\mathfrak{so}_{6,2} \cong \mathfrak{so}^*_8$, but we can still single out some special cases of type $\mathrm{D}_4^{\mathbb{R}}$ and $\mathrm{D}_4^{\mathbb{H}}$. For simplicity, we shall drop the subscript n when it is not used in the discussion. Deligne showed that an adjoint and simple Shimura datum $(\mathrm{G}, \mathcal{D})$ is of abelian type exactly when it is of types A, B, C, $\mathrm{D}^{\mathbb{R}}$, or $\mathrm{D}^{\mathbb{H}}$. In cases of types A, B, C, and $\mathrm{D}^{\mathbb{R}}$, the related Hodge-type Shimura datum $(\mathrm{G}_1, \mathcal{D}_1)$ (as in Definition 5.2.1.1) can be chosen such that the derived group $\mathrm{G}_1^{\mathrm{der}}$ is simply-connected as an algebraic group over $\mathbb{Q}$. However, in the remaining cases of $\mathrm{D}_n^{\mathbb{H}}$, where $n \geq 4$, the related Hodge-type Shimura datum can only be chosen such that the noncompact simple factors of $\mathrm{G}_1^{\mathrm{der}}(\mathbb{R})$ are all isomorphic to SO^*_{2n}. (Although this could have been a more precise statement in terms of algebraic groups, we are content with the simpler statement here in terms of real Lie groups.)

More generally, suppose that $(\mathrm{G}, \mathcal{D})$ is a Shimura datum that is *simple* in the sense that the adjoint quotient G^{ad} is simple. (But we no longer assume that G itself is adjoint.) Then we can still classify the type of $(\mathrm{G}, \mathcal{D})$ by classifying the type of $(\mathrm{G}^{\mathrm{ad}}, \mathcal{D}^{\mathrm{ad}})$ as above, and we have the following possibilities:

(1) All cases of types A, B, C, and $\mathrm{D}^{\mathbb{R}}$ are of abelian type.
(2) Cases of type $\mathrm{D}^{\mathbb{H}}$ can be of abelian type only when $\mathrm{G}^{\mathrm{der}}$ is the quotient of $\mathrm{G}_1^{\mathrm{der}}$ for the Hodge-type datum $(\mathrm{G}_1, \mathcal{D}_1)$ mentioned in the last paragraph. In particular, $\mathrm{G}_1^{\mathrm{der}}(\mathbb{R})$ cannot have factors isomorphic to the spin group cover Spin^*_{2n} of SO^*_{2n}.
(3) Cases of *mixed type* D, namely, those that are neither of type $\mathrm{D}^{\mathbb{R}}$ nor of type $\mathrm{D}^{\mathbb{H}}$, are *not* of abelian type. Such cases exist by a general result due to Borel and Harder (see [BH78, Thm. B]).
(4) Cases of type E_6 and E_7 are never of abelian type.

Thus, roughly speaking, all Shimura varieties associated with symplectic and unitary groups are of abelian type, and none of the Shimura varieties associated with exceptional groups can be of abelian type, while there are some complications for Shimura varieties associated with spin groups or special orthogonal groups.

5.2.3. *Integral models of these Shimura varieties.* The strategy for constructing integral models of Hodge-type Shimura varieties is as follows. Suppose $(\mathrm{G}, \mathcal{D})$ is a Hodge-type Shimura datum with Siegel embedding

$$(\mathrm{G}, \mathcal{D}) \hookrightarrow (\mathrm{GSp}_{2n}, \mathcal{H}_n^{\pm}), \tag{5.2.3.1}$$

and suppose $\mathcal{U}$ is an open compact subgroup of $\mathrm{G}(\mathbb{A}^\infty)$ that is the pullback of some open compact subgroup $\mathcal{U}'$ of $\mathrm{GSp}_{2n}(\mathbb{A}^\infty)$ such that $\mathcal{U}' = \mathcal{U}'^{,p}\mathcal{U}'_p$, where $\mathcal{U}'^{,p}$ is some neat open compact subgroup of $\mathrm{GSp}_{2n}(\mathbb{A}^{\infty,p})$, and where $\mathcal{U}'_p = \mathrm{GSp}_{2n}(\mathbb{Z}_p)$. (In the literature, people have considered the more general setting where $\mathcal{U}'_p$ is the stabilizer in $\mathrm{GSp}_{2n}(\mathbb{Q}_p)$ of some $\mathbb{Z}_p$-lattice, which can nevertheless be reduced to our setting here by modifying the Siegel embedding (5.2.3.1), at the expense of replacing n with $8n$, using "Zarhin's trick" as in [Lan16, Lem. 4.9] or [Lan18a, Lem. 2.1.1.9].) Let $v|p$ be a place of F_0, and let $\mathcal{O}_{F_0,(v)}$ demote the localization of F_0 at v. Then we have a canonical morphism

$$\mathsf{X}_{\mathcal{U},F_0} \to \mathsf{X}_{\mathcal{U}',\mathcal{O}_{F_0,(v)}}, \tag{5.2.3.2}$$

where $\mathsf{X}_{\mathcal{U},F_0}$ abusively denotes the canonical model of $\mathsf{X}_{\mathcal{U}}$ over F_0, and where $\mathsf{X}_{\mathcal{U}',\mathcal{O}_{F_0,(v)}}$ denotes the base change to $\mathcal{O}_{F_0,(v)}$ of the Siegel moduli scheme at level $\mathcal{U}'$. Then we define the integral model

$$\mathsf{X}_{\mathcal{U},\mathcal{O}_{F_0,(v)}} \tag{5.2.3.3}$$

of $\mathsf{X}_{\mathcal{U},F_0}$ over $\mathcal{O}_{F_0,(v)}$ by taking the *normalization* of $\mathsf{X}_{\mathcal{U}',\mathcal{O}_{F_0,(v)}}$ in $\mathsf{X}_{\mathcal{U},F_0}$ under the above canonical morphism.

A priori, this is a very general construction, which also makes sense when $\mathcal{U}$ is just a subgroup of the pullback of $\mathcal{U}'$. This is what we did in the PEL-type cases in [Lan18a] and [Lan16], as reviewed in Section 5.1.5, and similar ideas can be traced back to [DR73].

But when $\mathcal{U}$ is of the form $\mathcal{U}^p\mathcal{U}_p$ such that $\mathcal{U}^p$ is a neat open compact subgroup of $\mathrm{G}(\mathbb{A}^{\infty,p})$ and such that $\mathcal{U}_p$ is a *hyperspecial* maximal open compact subgroup of $\mathrm{G}(\mathbb{Q}_p)$, Vasiu, Kisin, and others (see [Mil92, Rem. 2.15], [Vas99], [Moo98], [Kis10], and [KMP16]) proved that $\mathsf{X}_{\mathcal{U},\mathcal{O}_{F_0,(v)}}$ is *smooth* and satisfies certain *extension properties* making it an *integral canonical model*. (The formulations of the extension properties vary from author to author. See, in particular, [Mil92, Sec. 2], [Moo98, Sec. 3], and [Kis10, (2.3.7)]. An *integral canonical model* then just means an integral model satisfying some formulation of the extension property. This notion is not expected to be useful when the levels at p are not hyperspecial.) They also showed that, roughly speaking, by working with connected components, by taking suitable quotients of integral models of Hodge-type ones by finite groups, and by descent, we also obtain integral models for all abelian-type Shimura varieties at *hyperspecial* levels at p, which are *smooth* and satisfy the same extension properties.

When $p > 2$, and when G splits over a tamely ramified extension of $\mathbb{Q}_p$, Kisin and Pappas proved in [KP18] that all abelian-type Shimura varieties at *parahoric* levels at p have integral models which have *local models* as predicted by the group-theoretic construction in [PZ13], and which satisfy a weaker extension property.

In [Kis17], Kisin proved a variant of the Langlands–Rapoport conjecture (partly following Kottwitz's variant; cf. Section 5.1.6) for all abelian-type Shimura varieties at *hyperspecial* levels at p, when $p > 2$.

5.2.4. *Integral models of toroidal and minimal compactifications.* In [MP19], it is proved that, by taking normalizations over the toroidal and minimal compactifications of Siegel moduli constructed in [FC90] and [Lan13], we also have good integral models of toroidal and minimal compactifications of integral models of Hodge-type Shimura varieties constructed in the previous subsection. The cone decompositions for the toroidal compactifications thus obtained are rather inexplicit pullbacks of certain smooth ones for the Siegel moduli. But it is believed that, by considering normalizations of blowups of the minimal compactifications as in [Lan17], we can obtain toroidal compactifications associated with more general projective and smooth cone decompositions as well.

As in the PEL-type cases in Section 5.1.5, such compactifications are good in the sense that the complications in describing its local structures and its boundary stratifications are completely independent of each other. Using the toroidal compactifications thus constructed, we showed in [LS18b] and [LS18c] that, for each integral model of Hodge-type or abelian-type Shimura varieties constructed in [Kis10], [KMP16], and [KP18], and for any prime ℓ different from the residue characteristics, the ℓ-adic étale nearby cycle cohomology over the special fiber is canonically isomorphic to the ℓ-adic étale cohomology over the (geometric) generic fiber, without resorting to the proper base change theorem (cf. [AGV73, XII, 5.1] and [DK73a, XIII, (2.1.7.1) and (2.1.7.3)]). (In [LS18c], the proof in the case of abelian-type Shimura varieties is achieved by reduction to the case of Hodge-type Shimura varieties.) Intuitively speaking, the special fibers of these integral models have as many points as there should be, at least for studying the ℓ-adic étale cohomology of Shimura varieties.

5.3. **Beyond abelian-type cases.** Although abelian-type Shimura varieties are already very useful and general, as we have seen in Section 5.2.2, they do not cover all possibilities of Shimura varieties. Among those associated with simple Shimura data, we are not just missing those of types E_6 and E_7—we are also missing the majority (depending on one's viewpoint) of type D cases. But these missing cases can still be quite interesting, both for geometric and arithmetic reasons.

For example, in [MS10], Milne and Suh showed that there are many examples of geometrically connected quasi-projective varieties X over number fields F such that the complex analytifications of $\mathsf{X} \underset{F,\sigma}{\otimes} \mathbb{C}$ and $\mathsf{X} \underset{F,\tau}{\otimes} \mathbb{C}$ have nonisomorphic fundamental groups for different embeddings σ and τ of F into $\mathbb{C}$, by considering connected Shimura varieties of mixed type D (which exist by [BH78, Thm. B]).

As another example, we have the following result of R. Liu and X. Zhu's:

Theorem 5.3.1 (see [LZ17, Thm. 1.2])**.** *Suppose $\mathsf{X}_{\mathcal{U}}$ is a Shimura variety associated with a* **general** *Shimura datum $(\mathrm{G}, \mathcal{D})$ (which can certainly be simple of types E_6, E_7, or mixed D) at some level $\mathcal{U}$. Suppose $\mathcal{V}$ is an étale $\mathbb{Q}_p$-local system associated with some finite-dimensional representation of $\mathrm{G}^c = \mathrm{G}/Z_s(\mathrm{G})$, where $Z_s(\mathrm{G})$ is the minimal subtorus $Z_s(\mathrm{G})$ of the center $Z(\mathrm{G})$ such that the torus $Z(\mathrm{G})^\circ/Z_s(\mathrm{G})$ has the same split ranks over $\mathbb{Q}$ and $\mathbb{R}$. Suppose $x : \mathrm{Spec}(F) \to \mathsf{X}_{\mathcal{U}}$ is a closed point of $\mathsf{X}_{\mathcal{U}}$, together with a geometric point $\overline{x} : \mathrm{Spec}(\overline{F}) \to \mathsf{X}_{\mathcal{U}}$ lifting x. Then the stalk*

$\mathcal{V}_{\bar{x}}$, when regarded as a representation of $\mathrm{Gal}(\overline{F}/F)$*, is* **geometric** *in the sense of Fontaine–Mazur* (*see* [FM97, Part I, Sec. 1])*. That is, it is unramified at all but finitely many places of F, and is de Rham at the places of F above p.*

Recall that Fontaine and Mazur conjectured that every irreducible geometric Galois representation *comes from geometry* in the sense that it appears as a subquotient of the étale cohomology of an algebraic variety (see [FM97, Part I, Sec. 1] again). Theorem 5.3.1 would have been less surprising if it only considered abelian-type Shimura varieties with certain assumptions such that the stalks as in the theorem are just the p-adic realizations of some *abelian motives*. But the theorem does consider non-abelian-type Shimura varieties as well, and we do not know any families of motives parameterized by these general Shimura varieties. We hope that there are such families, but this is a difficult problem that is still wide open.

Both results mentioned above are about Shimura varieties or connected Shimura varieties in characteristic zero. But what can we say about their integral models? Certainly, any quasi-projective variety over the reflex field F_0 has some integral model over the ring of integers $\mathcal{O}_{F_0}$ in F_0, but we are not interested in such a general construction. It would be desirable to have a construction such that the pullback to $\mathcal{O}_{F_0,(p)}$ of the integral model at level $\mathcal{U}$ is smooth and satisfies some reasonable extension property, when $\mathcal{U}$ is of the form $\mathcal{U}^p\mathcal{U}_p$ such that $\mathcal{U}^p$ is a neat open compact subgroup of $\mathrm{G}(\mathbb{A}^{\infty,p})$ and such that $\mathcal{U}_p$ is a *hyperspecial* maximal open compact subgroup of $\mathrm{G}(\mathbb{Q}_p)$. Moreover, it would be desirable that the reduction modulo p of such integral models can be described by some variant of the Langlands–Rapoport conjecture (see Section 5.1.6 and the end of Section 5.2.3). These are also difficult and wide open problems, but their nature might be very different from that of the problem (in the previous paragraph) of finding motives parameterized by such Shimura varieties.

5.4. **Beyond Shimura varieties?** There are some exciting recent developments which are not about integral models of Shimura varieties, or not even about Shimura varieties, but it would have been a shame not to mention them at all.

In [Sch15], Scholze constructed *perfectoid Shimura varieties* in all Hodge-type cases, which are limits at the infinite level at p (in a subtle sense) of Shimura varieties at finite levels. Base on these, in [She17], Shen also constructed perfectoid Shimura varieties in all abelian-type cases. Although the constructions in [Sch15] used Siegel moduli schemes over the integers, these perfectoid Shimura varieties are objects (called *adic spaces*) defined over the characteristic zero base field $\mathbb{C}_p$ (the completion of an algebraic closure $\overline{\mathbb{Q}}_p$ of $\mathbb{Q}_p$). But the property of being *perfectoid* (see [Sch12]) makes them much more powerful than integral models at finite levels for studying many questions about torsion in the cohomology of Shimura varieties and related geometric objects—see, for example, [Sch15], [NT16], and [CS17].

Even the concept of Shimura varieties itself has been challenged and generalized. In his Berkeley–MSRI lectures in the Fall of 2014 (see [SW20]), Scholze defined certain moduli spaces of *mixed-characteristic shtukas*, which can be viewed as a local analogue of Shimura varieties, but is much more general than the previously studied local geometric objects such as the so-called *Rapoport–Zink spaces* (see [RZ96]) and can be defined for all reductive linear algebraic groups over $\mathbb{Q}_p$. He proposed to use them to construct *local* Langlands parameters for all reductive linear algebraic groups over $\mathbb{Q}_p$, along the lines of Vincent Lafforgues's construction of global

Langlands parameters for global fields of positive characteristics (see [Laf18]), but in the much fancier category of *diamonds* (see [SW20] and [Sch17]).

Acknowledgements

Some materials in this article are based on introductory lectures given at the National Center for Theoretical Sciences (NCTS) in Taipei in 2014; at the Mathematical Sciences Research Institute (MSRI) in Berkeley in 2014; at the University of Minnesota, Twin Cities, in 2015; and at the summer school sponsored by ETH Zürich in Ascona, Switzerland, in 2016. We heartily thank the organizers for their invitations, and thank the audiences for their valuable feedbacks. We also thank Brian Conrad and Sug Woo Shin, and the anonymous referees, for their helpful comments and suggestions.

References

[AEK05] J. Arthur, D. Ellwood, and R. Kottwitz (eds.), *Harmonic analysis, the trace formula, and Shimura varieties*, Clay Mathematics Proceedings, vol. 4, Proceedings of the Clay Mathematics Institute 2003 Summer School, The Fields Institute, Toronto, Canada, June 2–27, 2003, American Mathematical Society, Providence, Rhode Island, Clay Mathematics Institute, Cambridge, Massachusetts, 2005.

[AGV73] M. Artin, A. Grothendieck, and J.-L. Verdier (eds.), *Théorie des topos et cohomologie étale des schémas* (*SGA 4*)*, Tome 3*, Lecture Notes in Mathematics, vol. 305, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[AMRT75] A. Ash, D. Mumford, M. Rapoport, and Y. Tai, *Smooth compactification of locally symmetric varieties*, Lie Groups: History Frontiers and Applications, vol. 4, Math Sci Press, Brookline, Massachusetts, 1975.

[AMRT10] ———, *Smooth compactification of locally symmetric varieties*, 2nd ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, New York, 2010.

[Art69] M. Artin, *Algebraization of formal moduli: I*, in Spencer and Iyanaga [SI69], pp. 21–71.

[AT83] M. Artin and J. Tate (eds.), *Arithmetic and geometry, Vol. I. arithmetic. Papers dedicated to I. R. Shafarevich on the occasion of his sixtieth birthday*, Progress in Mathematics, vol. 35, Birkhäuser, Boston, 1983.

[Bai58] W. L. Baily, Jr., *Satake's compactification of* V_n, Amer. J. Math. **80** (1958), no. 2, 348–364.

[Bai70] ———, *An exceptional arithmetic group and its eisenstein series*, Ann. Math. (2) **91** (1970), no. 3, 512–549.

[BB66] W. L. Baily, Jr. and A. Borel, *Compactification of arithmetic quotients of bounded symmetric domains*, Ann. Math. (2) **84** (1966), no. 3, 442–528.

[BC79] A. Borel and W. Casselman (eds.), *Automorphic forms, representations and L-functions*, Proceedings of Symposia in Pure Mathematics, vol. 33, Part 2, held at Oregon State University, Corvallis, Oregon, July 11–August 5, 1977, American Mathematical Society, Providence, Rhode Island, 1979.

[BG84] M.-J. Bertin and C. Goldstein (eds.), *Séminaire de théorie des nombres, Paris 1982–1983*, Progress in Mathematics, vol. 51, Birkhäuser, Boston, Basel, Stuttgart, 1984.

[BH78] A. Borel and G. Harder, *Existence of discrete cocompact subgroups of reductive groups over local fields*, J. Reine Angew. Math. **298** (1978), 53–64.

[BHC62] A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. Math. (2) **75** (1962), no. 3, 483–535.

[BJ06] A. Borel and L. Ji, *Compactifications of symmetric and locally symmetric spaces*, Mathematics: Theory & Applications, Birkhäuser, Boston, 2006.

[Bor62] A. Borel, *Ensembles fondamentaux pour les groupes arithmétiques*, Colloq. Théorie des Groupes Algébriques (Bruxelles, 1962), Librairie Universitaire, Louvain; GauthierVillars, Paris, 1962, pp. 23–40.

[Bor63] ———, *Some finiteness properties of adele groups over number fields*, Publ. Math. Inst. Hautes Étud. Sci. **16** (1963), 5–30.

[Bor69] ______, *Introduction aux groupes arithmétiques*, Publications de l'Institut de Mathématique de l'Université de Strasbourg XV, Actualités scientifiques et industrielles, vol. 1341, Hermann, Paris, 1969.

[Bor72] ______, *Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem*, J. Differential Geometry **6** (1972), 543–560.

[Bor87] M. V. Borovoĭ, *The group of points of a semisimple group over a totally real-closed field*, Problems in group theory and homological algebra (Russian), Mathematika, Yaroslav. Gos. Univ., Yaroslavl′, 1987, pp. 142–149.

[Bor91] A. Borel, *Linear algebraic groups*, 2nd ed., Graduate Texts in Mathematics, vol. 126, Springer-Verlag, Berlin, Heidelberg, New York, 1991.

[Bor84] M. V. Borovoĭ, *Langlands' conjecture concerning conjugation of connected Shimura varieties*, Selecta Math. Soviet. **3** (1983/84), no. 1, 3–39.

[BS73] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Comment. Math. Helv. **48** (1973), 436–491, with an appendix by A. Douady et L. Hérault.

[BT65] A. Borel and J. Tits, *Groupes réductifs*, Publ. Math. Inst. Hautes Étud. Sci. **27** (1965), 55–151.

[BW00] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, 2nd ed., Mathematical Surveys and Monographs, vol. 67, American Mathematical Society, Providence, Rhode Island, 2000.

[Car58] H. Cartan, *Prolongement des espaces analytiques normaux*, Math. Ann. (1958), 97–110.

[Car86] H. Carayol, *Sur la mauvaise réduction des courbes de Shimura*, Compositio Math. **59** (1986), no. 2, 151–230.

[CCM91] J. A. Carlson, C. H. Clemens, and D. R. Morrison (eds.), *Complex geometry and Lie theory*, Proceedings of Symposia in Pure Mathematics, vol. 53, Proceedings of a Symposium on Complex Geometry and Lie Theory held at Sundance, Utah, May 26–30, 1989, Springer-Verlag, Berlin, Heidelberg, New York, 1991.

[CHKS18] J. Cogdell, G. Harder, S. Kudla, and F. Shahidi (eds.), *Cohomology of arithmetic groups: On the occasion of Joachim Schwermer's 66th birthday, Bonn, Germany, June 2016*, Springer Proceedings in Mathematics & Statistics, vol. 245, Springer International Publishing AG, 2018.

[CM90a] L. Clozel and J. S. Milne (eds.), *Automorphic forms, Shimura varieties, and L-functions. Volume II*, Perspectives in Mathematics, vol. 11, Proceedings of a Conference held at the University of Michigan, Ann Arbor, July 6–16, 1988, Academic Press Inc., Boston, 1990.

[CM90b] L. Clozel and J. S. Milne (eds.), *Automorphic forms, Shimura varieties, and L-functions. Volume I*, Perspectives in Mathematics, vol. 10, Proceedings of a Conference held at the University of Michigan, Ann Arbor, July 6–16, 1988, Academic Press Inc., Boston, 1990.

[CN90] C.-L. Chai and P. Norman, *Bad reduction of the Siegel moduli scheme of genus two with $\Gamma_o(p)$-level structure*, Amer. J. Math. **112** (1990), no. 6, 1003–1071.

[Con12] B. Conrad, *Weil and Grothendieck approaches to adelic points*, Enseign. Math. (2) **58** (2012), no. 1–2, 61–97.

[CS17] A. Caraiani and P. Scholze, *On the generic part of the cohomology of compact unitary Shimura varieties*, Ann. Math. (2) **186** (2017), no. 3, 649–766.

[CY97] J. Coates and S. T. Yau (eds.), *Elliptic curves, modular forms & Fermat's last theorem*, Proceedings of a conference held in the Institute of Mathematics of the Chinese Univerisity of Hong Kong, International Press, Cambridge, Massachusetts, 1997.

[Del70] P. Deligne, *Equations différentielles à points singuliers réguliers*, Lecture Notes in Mathematics, vol. 163, Springer-Verlag, Berlin, Heidelberg, New York, 1970.

[Del71a] ______, *Formes modulaires et représentations ℓ-adiques*, Séminaire Bourbaki, exposé n° 355 (février 1969), Lecture Notes in Mathematics, vol. 179, Springer-Verlag, Berlin, Heidelberg, New York, 1971, pp. 139–172.

[Del71b] ______, *Travaux de Shimura*, Séminaire Bourbaki, exposé n° 389 (février 1971), Lecture Notes in Mathematics, vol. 244, Springer-Verlag, Berlin, Heidelberg, New York, 1971, pp. 123–165.

[Del72] ______, *La conjecture de Weil pour les surfaces K3*, Invent. Math. **15** (1972), 206–226.

[Del77] P. Deligne (ed.), *Cohomologie étale* (*SGA* $4\frac{1}{2}$), Lecture Notes in Mathematics, vol. 569, Springer-Verlag, Berlin, Heidelberg, New York, 1977.

[Del79] ———, *Variétés de Shimura: Interprétation modulaire, et techniques de construction de modèles canoniques*, in Borel and Casselman [BC79], pp. 247–290.

[DK73a] P. Deligne and N. Katz (eds.), *Groupes de monodromie en géométrie algébrique* (*SGA 7 II*), Lecture Notes in Mathematics, vol. 340, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[DK73b] P. Deligne and W. Kuyk (eds.), *Modular functions of one variable II*, Lecture Notes in Mathematics, vol. 349, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[DMOS82] P. Deligne, J. S. Milne, A. Ogus, and K. Shih, *Hodge cycles, motives, and Shimura varieties*, Lecture Notes in Mathematics, vol. 900, Springer-Verlag, Berlin, Heidelberg, New York, 1982.

[DR73] P. Deligne and M. Rapoport, *Les schémas de modules de courbes elliptiques*, in Deligne and Kuyk [DK73b], pp. 143–316.

[EHLS20] E. Eischen, M. Harris, J.-S. Li, and C. Skinner, *p-adic L-functions for unitary groups*, Forum Math. Sigma **8** (2020), e8, 160 pp.

[Eie09] M. Eie, *Topics in number theory*, Monographs in Number Theory, vol. 2, World Scientific, Singapore, 2009.

[Fal84] G. Faltings, *Arithmetic varieties and rigidity*, in Bertin and Goldstein [BG84], pp. 64–77.

[FC90] G. Faltings and C.-L. Chai, *Degeneration of abelian varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, vol. 22, Springer-Verlag, Berlin, Heidelberg, New York, 1990.

[FK95] J. Faraut and A. Korányi, *Analysis on symmetric cones*, Oxford Mathematical Monographs, Clarendon Press, Oxford University Press, Oxford, 1995.

[FM97] J.-M. Fontaine and B. Mazur, *Geometric Galois representations*, in Coates and Yau [CY97], pp. 190–227.

[FM04] L. Fargues and E. Mantovan, *Variétés de Shimura, espaces de Rapoport–Zink, et correspondances de Langlands locales*, Astérisque, no. 291, Société Mathématique de France, Paris, 2004.

[FM13a] Gavril Farkas and Ian Morrison (eds.), *Handbook of moduli: Volume II*, Advanced Lectures in Mathematics, vol. 25, International Press, Somerville, Massachusetts, Higher Education Press, Beijing, 2013.

[FM13b] Gavril Farkas and Ian Morrison (eds.), *Handbook of moduli: Volume III*, Advanced Lectures in Mathematics, vol. 26, International Press, Somerville, Massachusetts, Higher Education Press, Beijing, 2013.

[Fra98] J. Franke, *Harmonic analysis in weighted L_2-spaces*, Ann. Sci. Ecole Norm. Sup. (4) **31** (1998), 181–279.

[FS98] J. Franke and J. Schwermer, *A decomposition of spaces of automorphic forms, and the Eisenstein cohomology of arithmetic groups*, Math. Ann. **311** (1998), 765–790.

[GHM94] M. Goresky, G. Harder, and R. MacPherson, *Weighted cohomology*, Invent. Math. **116** (1994), 139–213.

[GL97] W. T. Gan and H. Y. Loke, *Modular forms of level p on the exceptional tube domain*, J. Ramanujan Math. Soc. **12** (1997), no. 2, 162–202.

[Gor05] M. Goresky, *Compactifications and cohomology of modular varieties*, in Arthur et al. [AEK05], pp. 551–582.

[GW09] R. Goodman and N. R. Wallach, *Symmetry, representations, and invariants*, Graduate Texts in Mathematics, vol. 255, Springer-Verlag, Berlin, Heidelberg, New York, 2009.

[Hai05] T. J. Haines, *Introduction to Shimura varieties with bad reductions of parahoric type*, in Arthur et al. [AEK05], pp. 583–658.

[Har89] M. Harris, *Functorial properties of toroidal compactifications of locally symmetric varieties*, Proc. London Math. Soc. (3) **59** (1989), 1–22.

[Har90] ———, *Automorphic forms and the cohomology of vector bundles on Shimura varieties*, in Clozel and Milne [CM90a], pp. 41–91.

[Hel01] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Graduate Studies in Mathematics, vol. 34, American Mathematical Society, Providence, Rhode Island, 2001.

[Hid02] H. Hida, *Control theorems of coherent sheaves on Shimura varieties of PEL type*, J. Inst. Math. Jussieu **1** (2002), no. 1, 1–76.

[Hid04] ———, *p-adic automorphic forms on Shimura varieties*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, Heidelberg, New York, 2004.

[Hir64a] H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero: I*, Ann. Math. (2) **79** (1964), no. 1, 109–203.

[Hir64b] ———, *Resolution of singularities of an algebraic variety over a field of characteristic zero: II*, Ann. Math. (2) **79** (1964), no. 2, 205–326.

[Hir66] U. Hirzebruch, *Über Jordan-Algebren und beschränkte symmetrische Gebiete*, Math. Z. **94** (1966), 387–390.

[HLS06] M. Harris, J.-S. Li, and C. M. Skinner, *p-adic L-functions for unitary Shimura varieties. I. Construction of the Eisenstein measure*, Doc. Math. (2006), no. Extra Vol., 393–464.

[HLTT16] M. Harris, K.-W. Lan, R. Taylor, and J. Thorne, *On the rigid cohomology of certain Shimura varieties*, Res. Math. Sci. **3** (2016), article no. 37, 308 pp.

[HT01] M. Harris and R. Taylor, *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies, vol. 151, Princeton University Press, Princeton, 2001.

[Int75] International Colloquium on Discrete Subgroups of Lie Groups and Applications to Moduli, Bombay, 8–15 January 1973, *Discrete subgroups of Lie groups and applications to moduli: Papers presented at the Bombay Colloquium, 1973*, Tata Institute of Fundamental Research Studies in Mathematics, vol. 7, Oxford University Press, Oxford, 1975.

[JPYY12] L. Ji, Y. S. Poon, L. Yang, and S.-T. Yau (eds.), *Fifth International Congress of Chinese Mathematicians*, AMS/IP Studies in Advanced Mathematics, vol. 51, part 1, Proceedings of the Fifth International Congress of Chinese Mathematicians held at Tsinghua University, Beijing, in December 2010, American Mathematical Society, Providence, Rhode Island, and International Press, Cambridge, MA, 2012.

[JYY19] L. Ji, L. Yang, and S.-T. Yau (eds.), *Proceedings of the Seventh International Congress of Chinese Mathematicians, Volume I*, Advanced Lectures in Mathematics, vol. 43, Proceedings of the Seventh International Congress of Chinese Mathematicians held at the Chinese Academy of Sciences, Beijing, in August 2016, International Press, Somerville, Massachusetts, Higher Education Press, Beijing, 2019.

[Kim93] H. H. Kim, *Exceptional modular form of weight* 4 *on an exceptional domain contained in* $\mathbb{C}^{27}$, Rev. Mat. Iberoamericana **9** (1993), no. 1, 139–200.

[Kis10] M. Kisin, *Integral models for Shimura varieties of abelian type*, J. Amer. Math. Soc. **23** (2010), no. 4, 967–1012.

[Kis17] ———, *Mod p points on Shimura varieties of abelian type*, J. Amer. Math. Soc. **30** (2017), no. 3, 819–914.

[KM85] N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies, vol. 108, Princeton University Press, Princeton, 1985.

[KMP16] W. Kim and K. Madapusi Pera, *2-adic integral canonical models*, Forum Math. Sigma **4** (2016), e28, 34 pp.

[Kna01] A. W. Knapp, *Representation theory of semisimple groups: An overview based on examples*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, 2001.

[Kna02] ———, *Lie groups beyond an introduction*, 2nd ed., Progress in Mathematics, vol. 140, Birkhäuser, Boston, 2002.

[Kot90] R. E. Kottwitz, *Shimura varieties and* λ*-adic representations*, in Clozel and Milne [CM90b], pp. 161–209.

[Kot92] ———, *Points on some Shimura varieties over finite fields*, J. Amer. Math. Soc. **5** (1992), no. 2, 373–444.

[KP18] M. Kisin and G. Pappas, *Integral models of Shimura varieties with parahoric level structure*, Publ. Math. Inst. Hautes Étud. Sci. **128** (2018), no. 1, 121–218.

[KS67] M. Kuga and I. Satake, *Abelian varieties attached to polarized* K_3*-surfaces*, Math. Ann. **169** (1967), 239–242.

[Laf18] V. Lafforgue, *Chtoucas pour les groupes réductifs et paramétrisation de Langlands globale*, J. Amer. Math. Soc. **31** (2018), no. 3, 719–891.

[Lan76] R. P. Langlands, *Some contemporary problems with origins in the Jugendtraum*, Mathematical developments arising from Hilbert problems (Proc. Sympos. Pure Math., Vol. XXVIII, Northern Illinois Univ., De Kalb, Ill., 1974), American Mathematical Society, Providence, Rhode Island, 1976, pp. 401–418.

[Lan77] ———, *Shimura varieties and the Selberg trace formula*, Canad. J. Math. **29** (1977), no. 6, 1292–1299.

[Lan12a] K.-W. Lan, *Comparison between analytic and algebraic constructions of toroidal compactifications of PEL-type Shimura varieties*, J. Reine Angew. Math. **664** (2012), 163–228.

[Lan12b] ———, *Geometric modular forms and the cohomology of torsion automorphic sheaves*, in Ji et al. [JPYY12], pp. 183–208.

[Lan12c] ———, *Toroidal compactifications of PEL-type Kuga families*, Algebra Number Theory **6** (2012), no. 5, 885–966.

[Lan13] ———, *Arithmetic compactification of PEL-type Shimura varieties*, London Mathematical Society Monographs, vol. 36, Princeton University Press, Princeton, 2013, errata available on the author's website.

[Lan15] ———, *Boundary strata of connected components in positive characteristics*, Algebra Number Theory **9** (2015), no. 9, 1955–2054, an appendix to the article "Families of nearly ordinary Eisenstein series on unitary groups" by Xin Wan.

[Lan16] ———, *Compactifications of PEL-type Shimura varieties in ramified characteristics*, Forum Math. Sigma **4** (2016), e1, 98 pp.

[Lan17] ———, *Integral models of toroidal compactifications with projective cone decompositions*, Int. Math. Res. Not. IMRN **2017** (2017), no. 11, 3237–3280.

[Lan18a] ———, *Compactifications of PEL-type Shimura varieties and Kuga families with ordinary loci*, World Scientific, Singapore, 2018.

[Lan18b] ———, *Compactifications of splitting models of PEL-type Shimura varieties*, Trans. Amer. Math. Soc. **370** (2018), no. 4, 2463–2515.

[Lan19] ———, *Cohomology of automorphic bundles*, in Ji et al. [JYY19], pp. 303–325.

[Loo88] E. Looijenga, *L^2-cohomology of locally symmetric varieties*, Compositio Math. **67** (1988), 3–20.

[LR87] R. P. Langlands and M. Rapoport, *Shimuravarietäten und Gerben*, J. Reine Angew. Math. **378** (1987), 113–220.

[LR91] E. Looijenga and M. Rapoport, *Weights in the local cohomology of a Baily–Borel compactification*, in Carlson et al. [CCM91], pp. 223–260.

[LR92] R. P. Langlands and D. Ramakrishnan (eds.), *The zeta functions of Picard modular surfaces*, based on lectures delivered at a CRM Workshop in the spring of 1988, Les Publications CRM, Montréal, 1992.

[LS90] J.-P. Labesse and J. Schwermer (eds.), *Cohomology of arithmetic groups and automorphic forms*, Lecture Notes in Mathematics, vol. 1447, Proceedings of a Conference held in Luminy/Marseille, France, May 22–27, 1989, Springer-Verlag, Berlin, Heidelberg, New York, 1990.

[LS04] J.-S. Li and J. Schwermer, *On the Eisenstein cohomology of arithmetic groups*, Duke Math. J. **123** (2004), no. 1, 141–169.

[LS12] K.-W. Lan and J. Suh, *Vanishing theorems for torsion automorphic sheaves on compact PEL-type Shimura varieties*, Duke Math. J. **161** (2012), no. 6, 1113–1170.

[LS13] ———, *Vanishing theorems for torsion automorphic sheaves on general PEL-type Shimura varieties*, Adv. Math. **242** (2013), 228–286.

[LS18a] K.-W. Lan and B. Stroh, *Compactifications of subschemes of integral models of Shimura varieties*, Forum Math. Sigma **6** (2018), e18, 105.

[LS18b] ———, *Nearby cycles of automorphic étale sheaves*, Compos. Math. **154** (2018), no. 1, 80–119.

[LS18c] ———, *Nearby cycles of automorphic étale sheaves, II*, in Cogdell et al. [CHKS18], pp. 83–106.

[LZ17] R. Liu and X. Zhu, *Rigidity and a Riemann–Hilbert correspondence for p-adic local systems*, Invent. Math. **207** (2017), 291–343.

[Man04] E. Mantovan, *On certain unitary group Shimura varieties*, in *Variétés de Shimura, espaces de Rapoport–Zink, et correspondances de Langlands locales* [FM04], pp. 201–331.

[Man05] ———, *On the cohomology of certain PEL-type Shimura varieties*, Duke Math. J. **129** (2005), no. 3, 573–610.

[Man11] ———, *ℓ-adic étale cohomology of PEL type Shimura varieties with non-trivial coefficients*, WIN–Woman in Numbers, Fields Institute Communications, vol. 60, American Mathematical Society, Providence, Rhode Island, 2011, pp. 61–83.

[MB85] L. Moret-Bailly, *Pinceaux de variétés abéliennes*, Astérisque, vol. 129, Société Mathématique de France, Paris, 1985.

[MFK94] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 34, Springer-Verlag, Berlin, Heidelberg, New York, 1994.

[Mil83] J. S. Milne, *The action of an automorphism of* $\mathbf{C}$ *on a Shimura variety and its special points*, in Artin and Tate [AT83], pp. 239–265.

[Mil90] ———, *Canonical models of (mixed) Shimura varieties and automorphic vector bundles*, in Clozel and Milne [CM90b], pp. 283–414.

[Mil92] ———, *The points on a Shimura variety modulo a prime of good reduction*, in Langlands and Ramakrishnan [LR92], pp. 151–253.

[Mil05] ———, *Introduction to Shimura varieties*, in Arthur et al. [AEK05], pp. 265–378.

[Mil13] ———, *Shimura varieties and moduli*, in Farkas and Morrison [FM13a], pp. 467–548.

[Moo98] B. Moonen, *Models of Shimura varieties in mixed characteristic*, in *Galois Representations in Arithmetic Algebraic Geometry* [ST98], pp. 267–350.

[MP15] K. Madapusi Pera, *The Tate conjecture for K3 surfaces in odd characteristic*, Invent. Math. **201** (2015), no. 2, 625–668.

[MP16] ———, *Integral canonical models for spin Shimura varieties*, Compos. Math. **152** (2016), no. 4, 769–824.

[MP19] ———, *Toroidal compactifications of integral models of Shimura varieties of Hodge type*, Ann. Sci. Ecole Norm. Sup. (4) **52** (2019), no. 2, 393–514.

[MS10] J. S. Milne and J. Suh, *Nonhomeomorphic conjugates of connected Shimura varieties*, Amer. J. Math. **132** (2010), no. 3, 731–750.

[Mum70] D. Mumford, *Abelian varieties*, Tata Institute of Fundamental Research Studies in Mathematics, vol. 5, Oxford University Press, Oxford, 1970, with appendices by C. P. Ramanujam and Yuri Manin.

[Mum72] ———, *An analytic construction of degenerating abelian varieties over complete rings*, Compositio Math. **24** (1972), no. 3, 239–272.

[Mum75] ———, *A new approach to compactifying locally symmetric varieties*, in *Discrete Subgroups of Lie Groups and Applications to Moduli: Papers presented at the Bombay Colloquium, 1973* [Int75], pp. 211–224.

[Mum77] ———, *Hirzebruch's proportionality theorem in the non-compact case*, Invent. Math. **42** (1977), 239–272.

[Mum83] ———, *Tata lectures on theta I*, Progress in Mathematics, vol. 28, Birkhäuser, Boston, 1983.

[NT16] J. Newton and J. A. Thorne, *Torsion Galois representations over CM fields and Hecke algebras in the derived category*, Forum Math. Sigma **4** (2016), e21, 88 pp.

[Pfa96] M. Pfau, *The conjecture of Langlands and Rapoport for certain Shimura varieties of non-rational weight*, J. Reine Angew. Math. **471** (1996), 165–199.

[Pin89] R. Pink, *Arithmetic compactification of mixed Shimura varieties*, Ph.D. thesis, Rheinischen Friedrich-Wilhelms-Universität, Bonn, 1989.

[PR94] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics, vol. 139, Academic Press Inc., New York and London, 1994, translated by Rachel Rowen.

[PR03] G. Pappas and M. Rapoport, *Local models in the ramified case, I. The EL-case*, J. Algebraic Geom. **12** (2003), no. 1, 107–145.

[PR05] ———, *Local models in the ramified case, II. Splitting models*, Duke Math. J. **127** (2005), no. 2, 193–250.

[PR06] G. Prasad and A. S. Rapinchuk, *On the existence of isotropic forms of semi-simple algebraic groups over number fields with prescribed local behavior*, Adv. Math. **207** (2006), no. 2, 646–660.

[PR09] G. Pappas and M. Rapoport, *Local models in the ramified case, III. Unitary groups*, J. Inst. Math. Jussieu **8** (2009), no. 3, 507–564.

[PRS13] G. Pappas, M. Rapoport, and B. Smithling, *Local models of Shimura varieties, I. Geometry and combinatorics*, in Farkas and Morrison [FM13b], pp. 135–217.

[PZ13] G. Pappas and X. Zhu, *Local models of Shimura varieties and a conjecture of Kottwitz*, Invent. Math. **194** (2013), 147–254.

[Rap78] M. Rapoport, *Compactifications de l'espace de modules de Hilbert–Blumenthal*, Compositio Math. **36** (1978), no. 3, 255–335.

[Rap05] ———, *A guide to the reduction modulo p of Shimura varieties*, in Tilouine et al. [TCHV05], pp. 271–318.

[Rei97] H. Reimann, *The semi-simple zeta function of quaternionic Shimura varieties*, Lecture Notes in Mathematics, vol. 1657, Springer-Verlag, Berlin, Heidelberg, New York, 1997.

[Riz06] J. Rizov, *Moduli stacks of polarized $K3$ surfaces in mixed characteristic*, Serdica Math. J. **32** (2006), no. 2–3, 131–178.

[RZ96] M. Rapoport and T. Zink, *Period spaces for p-divisible groups*, Annals of Mathematics Studies, vol. 141, Princeton University Press, Princeton, 1996.

[Sat56] I. Satake, *On the compactification of the Siegel space*, J. Indian Math. Soc. (N.S.) **20** (1956), 259–281.

[Sat60] ———, *On compactifications of the quotient spaces for arithmetically defined discontinuous groups*, Ann. Math. (2) **72** (1960), no. 3, 555–580.

[Sat65] ———, *Holomorphic imbeddings of symmetric domains into a Siegel space*, Amer. J. Math. **87** (1965), no. 2, 425–461.

[Sat66] ———, *Clifford algebras and families of abelian varieties*, Nagoya Math. J. **27** (1966), 435–446.

[Sat67] ———, *Symplectic representations of algebraic groups satisfying a certain analyticity condition*, Acta Math. **117** (1967), 215–279.

[Sat68] ———, *Corrections to "Clifford algebras and families of abelian varieties"*, Nagoya Math. J. **31** (1968), 295–296.

[Sat80] ———, *Algebraic structures of symmetric domains*, Publications of the Mathematical Society of Japan, vol. 14, Princeton University Press, Princeton, 1980.

[Sat01] ———, *Compactifications, old and new*, Sugaku Expositions **14** (2001), no. 2, 175–189.

[SC58] I. Satake and H. Cartan, *Démonstration du théorème fondamental*, Séminaire Henri Cartan 10e année: 1957/1958, exposé 15 (14 avril 1958), Secrétariat mathématique, 11 rue Pierre Curie, Paris 5e, 1958.

[Sch90a] A. J. Scholl, *Motives for modular forms*, Invent. Math. **100** (1990), no. 2, 419–430.

[Sch90b] J. Schwermer, *Cohomology of arithmetic groups, automorphic forms and L-functions*, in Labesse and Schwermer [LS90], pp. 1–29.

[Sch12] P. Scholze, *Perfectoid spaces*, Publ. Math. Inst. Hautes Étud. Sci. **116** (2012), 245–313.

[Sch15] ———, *On torsion in the cohomology of locally symmetric varieties*, Ann. Math. (2) **182** (2015), no. 3, 946–1066.

[Sch17] ———, *Étale cohomology of diamonds*, preprint, 2017.

[Ser56] J.-P. Serre, *Geométrie algébrique et géométrie analytique*, Ann. Inst. Fourier. Grenoble **6** (1955–1956), 1–42.

[She17] X. Shen, *Perfectoid Shimura varieties of abelian type*, Int. Math. Res. Not. IMRN **2017** (2017), no. 21, 6599–6653.

[Shi63a] H. Shimizu, *On discontinuous groups operating on the product of the upper half planes*, Ann. Math. (2) **77** (1963), no. 1, 33–71.

[Shi63b] G. Shimura, *On analytic families of polarized abelian varities and automorphic functions*, Ann. Math. (2) **78** (1963), no. 1, 149–192.

[Shi66] ———, *Moduli and fibre systems of abelian varieties*, Ann. Math. (2) **83** (1966), no. 2, 294–338.

[Shi70] ———, *On canonical models of arithmetic quotients of bounded symmetric domains*, Ann. Math. (2) **91** (1970), no. 1, 144–222.

[Shi71] ———, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan, vol. 11, Princeton University Press, Princeton, 1971.

[Shi78] ———, *The arithmetic of automorphic forms with respect to a unitary group*, Ann. Math. (2) **107** (1978), no. 3, 569–605.

[Shi98] ———, *Abelian varieties with complex multiplication and modular functions*, Princeton University Press, Princeton, 1998.

[Shi11] S. W. Shin, *Galois representations arising from some compact Shimura varieties*, Ann. Math. (2) **173** (2011), no. 3, 1645–1741.

[SI69] D. C. Spencer and S. Iyanaga (eds.), *Global analysis. Papers in honor of K. Kodaira*, Princeton University Press, Princeton, 1969.

[Spr98] T. A. Springer, *Linear algebraic groups*, 2nd ed., Progress in Mathematics, vol. 9, Birkhäuser, Boston, 1998.

[SS90] L. Saper and M. Stern, *L_2-cohomology of arithmetic varieties*, Ann. Math. (2) **132** (1990), no. 1, 1–69.

[ST98] A. J. Scholl and R. L. Taylor, *Galois representations in arithmetic algebraic geometry*, London Mathematical Society Lecture Note Series, vol. 254, Cambridge University Press, Cambridge, New York, 1998.

[SU14] C. Skinner and E. Urban, *The main conjecture for* GL(2), Invent. Math. **195** (2014), no. 1, 1–277.

[SW20] P. Scholze and J. Weinstein, *Berkeley lectures on p-adic geometry*, Annals of Mathematics Studies, vol. 207, Princeton University Press, Princeton, 2020.

[TCHV05] J. Tilouine, H. Carayol, M. Harris, and M.-F. Vignéras (eds.), *Formes automorphes (I): Actes du semestre du Centre émile Borel, printemps 2000*, Astérisque, no. 298, Société Mathématique de France, Paris, 2005.

[Vas99] A. Vasiu, *Integral canonical models of Shimura varieties of preabelian type*, Asian J. Mathematics **3** (1999), 401–518.

[Zuc82] S. Zucker, *L_2 cohomology of warpped products and arithmetic groups*, Invent. Math. **70** (1982), 169–218.

Index

University of Minnesota, Minneapolis, MN 55455, USA
Email address: `kwlan@math.umn.edu`