INTRODUCTION TO SHIMURA VARIETIES

ANDREAS MIHATSCH

Abstract. These are lecture notes for a course on Shimura varieties I am currently teaching at Zhejiang University. Comments are highly welcome and much appreciated.

Contents

1. Introduction

In this first lecture, we will learn, very roughly, what Shimura varieties are and why they are interesting. Everything brought up today will be covered in much more detail later in the course, and it will be perfectly normal that many terms will be new during a first reading. Our goal today is only to get an overview.

1.1. **Why study Shimura varieties?** Shimura varieties combine two interesting properties:

- They are varieties defined over number fields which makes them interesting from a number theory perspective. Most importantly, their étale cohomology groups are representations of Galois groups of number fields.
- Their definition is in terms of connected reductive algebraic groups $G/\mathbb{Q}$. They come equipped with an action of the adelic points $G(\mathbb{A}_f)$, which implies that their étale cohomology groups are also $G(\mathbb{A}_f)$-representations.

Hence, the étale cohomology groups of Shimura varieties are both Galois and $G(\mathbb{A}_f)$-representations. Conjecturally, this two-fold structure is described by the global Langlands correspondence. Conversely, one can use the cohomology of Shimura varieties to prove important cases of this correspondence. This is the main motivation for our course, and our overall aim is to learn about several important ideas in this context.

Date: November 13, 2025.

Let us mention that Shimura varieties are also interesting for other reasons. For example, the study of heights on the Siegel variety plays an important role in Faltings's proof of the Mordell Conjecture [3]. Another example is the Gross–Zagier formula [6], which states an identity between height pairings of complex multiplication points on the modular curve and derivatives of L-functions. It plays a major role in the proof of cases of the Birch–Swinnerton-Dyer Conjecture. Its higher-dimensional generalizations, the arithmetic Gan–Gross–Prasad Conjectures [5,19], are an important topic in current arithmetic geometry research. In a related direction, the Kudla program [9] seeks to establish connections between cycles on Shimura varieties and modular forms or Eisenstein series. The proof of the averaged Colmez conjecture [1,18] has been an application of such ideas.

1.2. **This course.** The first part of our course will be an introduction to Shimura varieties. We will learn how to define them in terms of moduli spaces of abelian varieties and how to relate this definition to the group-theoretic one of Deligne. One of our goals is to obtain familiarity with the adelic formalism which will become important later.

In the second part of the course, we will study the cohomology of Shimura varieties. We will first get to know Matsushima's formula, which expresses the Betti cohomology of compact Shimura varieties in terms of automorphic representations. We will then learn about point counting in characteristic p (Langlands–Kottwitz method). The aim here is to give an orbital integral expression for the number of $\mathbb{F}_{p^n}$-points of the reduction mod p of the Shimura variety.

1.3. **References.** The following two are our main background references.

- The introductory lecture notes by Milne [13]. They focus on the group-theoretic definition of Shimura varieties and the definition of canonical models.
- The first few articles in the lecture notes volume [7]. They provide an introduction to PEL type Shimura varieties. The article of Yihang Zhu [20] is directly related to the material of the second part of the course.

1.4. **Prerequisites.** We will assume as little as possible. The only necessary background is some familiarity with varieties and algebraic number theory.

In the rest of this introduction, we sketch the definition of Shimura varieties and give an outline of the course contents.

1.5. **Shimura data.** Shimura varieties are attached to Shimura data. The formalism starts with a connected reductive group G over $\mathbb{Q}$. For example, G might be one of the following.

- $G = \mathrm{GL}_2$
- $G = \mathrm{GSp}_{2g}$, the general symplectic group in $2g$ variables. Let $J = \begin{pmatrix} & 1_g \\ -1_g & \end{pmatrix}$ be the matrix defining the standard symplectic form on $\mathbb{Q}^{2g}$. Then GSp is defined by
$$\mathrm{GSp}_{2g}(\mathbb{Q}) = \{g \in GL_{2g}(\mathbb{Q}) \mid {}^t g \cdot J \cdot g = c \cdot J \text{ for some } c \in \mathbb{Q}^\times\}. \tag{1.1}$$
It is related to the usual symplectic group Sp_{2g} by the exact sequence
$$1 \longrightarrow \mathrm{Sp}_{2g} \longrightarrow \mathrm{GSp}_{2g} \xrightarrow{c} \mathrm{GL}_1 \longrightarrow 1.$$
The map c is called the *similitude factor*. Note that $\mathrm{GSp}_2 = \mathrm{GL}_2$ and $\mathrm{Sp}_2 = \mathrm{SL}_2$, recovering the previous example.

- $G = \mathrm{U}(V)$, a unitary group. Let $K/\mathbb{Q}$ be an imaginary quadratic extension. (This means that $\mathbb{R} \otimes_{\mathbb{Q}} K \cong \mathbb{C}$.) Let V be an n-dimensional hermitian K-vector space. If V is not positive or negative definite then $\mathrm{U}(V)$ can occur as part of a Shimura datum.

Next, the formalism requires the datum of a homomorphism of real algebraic groups

$$h : \mathbb{C}^\times \longrightarrow G(\mathbb{R}) \tag{1.2}$$

which satisfies certain axioms introduced by Deligne [2]. Such an h is called a *Deligne homomorphism.* If $g \in G(\mathbb{R})$ is a real point of G, then we may conjugate h to define a new Deligne homomorphism,

$$\big(ghg^{-1}\big)(z) := gh(z)g^{-1}.$$

Let $S_h \subset G(\mathbb{R})$ denote the centralizer of h, meaning the subgroup of elements g with $ghg^{-1} = h$. The quotient $X = G(\mathbb{R})/S_h$ is precisely the set of Deligne homomorphisms that are conjugate to h. An important consequence of Deligne's axioms is that X is a finite union of hermitian symmetric domains for $G(\mathbb{R})$. In particular, it is a complex manifold. The pair (G, X) is called a *Shimura datum.*

Example 1.1. Consider $G = \mathrm{GL}_2$. We can embedd $\mathbb{C}$ into $\mathrm{M}_2(\mathbb{R})$ as $\mathbb{R}$-algebra by

$$h(a+bi) := \begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \tag{1.3}$$

If we restrict this embedding to unit groups, then we obtain a Deligne homomorphism $h : \mathbb{C}^\times \to \mathrm{GL}_2(\mathbb{R})$. Its centralizer is precisely $h(\mathbb{C}^\times)$ and the quotient X is the set of complex structures on $\mathbb{R}^2$. Since $\mathbb{C}^\times$ is connected and since $\mathrm{GL}_2(\mathbb{R})$ has two connected components, X has two connected components. We want to give a more explicit description of X.

Recall that $\mathbb{P}^1(\mathbb{C})$ is the space of complex lines in $\mathbb{C}^2$. Clearly, the Lie group $\mathrm{GL}_2(\mathbb{C})$ acts on it by its natural action on $\mathbb{C}^2$. The subgroup $\mathrm{GL}_2(\mathbb{R})$ preserves the real projective line $\mathbb{P}^1(\mathbb{R})$ and hence acts on the complement,

$$\mathrm{GL}_2(\mathbb{R}) \ \circlearrowright \ \mathbb{C}\backslash\mathbb{R}, \quad g \cdot \tau = \frac{a\tau + b}{c\tau + d}. \tag{1.4}$$

The complement $\mathbb{C}\backslash\mathbb{R}$ is the union of the upper and lower half plane which we often denote by $\mathbb{H}^\pm$. As an open subset of $\mathbb{C}$, it is naturally a complex manifold. Let us compute the stabilizer of i:

$$\begin{aligned} i = \frac{ai+b}{ci+d} \quad &\iff \quad -c + di = ai + b \\ &\iff \quad a = d, \quad c = -b. \end{aligned} \tag{1.5}$$

That is, the stabilizer of i is precisely $h(\mathbb{C}^\times)$. Moreover, it is clear that $\mathrm{GL}_2(\mathbb{R})$ acts transitively on $\mathbb{H}^\pm$ because

$$\begin{pmatrix} a & b \\ & 1 \end{pmatrix} \cdot i = ai + b.$$

Hence, we see that

$$X \xrightarrow{\sim} \mathbb{H}^\pm, \quad ghg^{-1} \longmapsto g \cdot i \tag{1.6}$$

as smooth manifolds in a $\mathrm{GL}_2(\mathbb{R})$-equivariant way. We have not defined the complex structure on X, but it is, in fact, given by the complex structure on $\mathbb{H}^\pm$ under (1.6).

Remark 1.2. Some groups, such as GL_n with $n \geq 3$, cannot occur as part of a Shimura datum. For example, the dimension of the symmetric space for $\mathrm{GL}_3(\mathbb{R})$ is

$$\dim \mathrm{SL}_3(\mathbb{R}) - \dim \mathrm{SO}(3) = 8 - 3$$

which is odd and hence cannot be a complex manifold.

1.6. **Shimura varieties over $\mathbb{C}$.** Given a Shimura datum (G, X), one next defines a complex variety in the following way. Let $\mathbb{A}$ denote the ring of adeles of $\mathbb{Q}$, and let $\mathbb{A} = \mathbb{A}_f \times \mathbb{R}$ be its factorization into finite and archimedean part. (We will review these definitions later in the course.) Given an open compact subgroup $K \subset G(\mathbb{A}_f)$, the quotient $G(\mathbb{A}_f)/K$ is a discrete countably infinite set with transitive $G(\mathbb{A}_f)$-action. Hence, the product $X \times G(\mathbb{A}_f)/K$ is a countable union of copies of X. We consider the diagonal action

$$G(\mathbb{Q}) \circlearrowright X \times G(\mathbb{A}_f)/K.$$

If K is small enough then the $G(\mathbb{Q})$-action is free. (The technical term is "neat" and we will get to know it later in the course.) It is also properly discontinuous, so we can form the quotient complex manifold

$$\mathrm{Sh}_K(G, X)(\mathbb{C}) := G(\mathbb{Q})\backslash\big(X \times G(\mathbb{A}_f)/K\big). \tag{1.7}$$

At this point, we have defined the complex points of the *Shimura variety for Shimura datum* (G, X) *and level* K as a complex manifold. The theorem of Baily–Borel states that there is a unique way to endow it with an algebraic structure.

Theorem 1.3 (Baily–Borel, see [13, Corollary 3.16])**.** *There exists a quasi-projective complex variety* $\mathrm{Sh}_K(G, X)_{\mathbb{C}}$ *such that there exists an isomorphism of complex manifolds* $\mathrm{Sh}_K(G, X)_{\mathbb{C}}(\mathbb{C}) \xrightarrow{\sim} \mathrm{Sh}_K(G, X)(\mathbb{C})$*. This variety is unique up to isomorphism.*

Remark 1.4. Simple examples of non-unique algebraic structures on complex manifolds can be found in [8].

Example 1.5. Let us again consider the case $G = \mathrm{GL}_2$ and let us give an example of a connected component of (1.7). Let $\widehat{\mathbb{Z}} = \prod_{p<\infty} \mathbb{Z}_p$ be the subring of integral elements of $\mathbb{A}_f$. For $n \geq 1$, consider the kernel

$$K(n) = \ker\big(\mathrm{GL}_2(\widehat{\mathbb{Z}}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})\big)$$

which is an open compact subgroup of $G(\mathbb{A}_f)$. It is small enough (in the above sense) if $n \geq 3$. The intersection

$$\Gamma(n) := \mathrm{GL}_2(\mathbb{Q}) \cap K(n)$$

is the classical congruence subgroup

$$\Gamma(n) = \left\{ \gamma \in \mathrm{GL}_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \bmod n \right\}.$$

The quotients $\Gamma(n)\backslash\mathbb{H}^+$ and $\Gamma(n)\backslash\mathbb{H}^-$ will be two of the connected components of the complex manifold $\mathrm{Sh}_{K(n)}(\mathrm{GL}_2, \mathbb{H}^{\pm})$.

1.7. **Shimura varieties over number fields.** Finally, one descends $\mathrm{Sh}_K(G, X)$ to a number field. Starting from a Shimura datum (G, X), Deligne defines a number field $E \subset \mathbb{C}$ called the *reflex field*. In a suitable sense, it is the smallest field over which the conjugacy class X is defined.

Example 1.6. Consider the three examples from §1.5.

- If $G = \mathrm{GL}_2$ or more generally $G = \mathrm{GSp}_{2g}$, then the reflex field is $\mathbb{Q}$.
- If $G = U(V)$ is a non-definite unitary group for an imaginary-quadratic field $K/\mathbb{Q}$, then the reflex field is the subfield $E \subset \mathbb{C}$ that is isomorphic to K.

Deligne [2] gave a definition of *canonical model* of $\mathrm{Sh}_K(G, X)_{\mathbb{C}}$ over E. It is a variety $\mathrm{Sh}_K(G, X)$ over $\mathrm{Spec}(E)$ together with an isomorphism

$$\mathbb{C} \otimes_E \mathrm{Sh}_K(G, X) \xrightarrow{\sim} \mathrm{Sh}_K(G, X)_{\mathbb{C}}$$

that satisfies a certain reciprocity law for complex multiplication points. Deligne proves that the canonical model $\mathrm{Sh}_K(G, X)$ is unique up to isomorphism if it exists.

Theorem 1.7 (Borovoi, Milne [11]). *For every Shimura datum, the canonical model exists.*

Definition 1.8. Let (G, X) be a Shimura datum with reflex field E and let $K \subset G(\mathbb{A}_f)$ be a sufficiently small level subgroup. The Shimura variety of level K attached to (G, X) is the canonical model $\mathrm{Sh}_K(G, X)$ from Theorem 1.7.

Remark 1.9. Historically, the study of Shimura varieties started with Shimura in the 1960s. He first considered moduli spaces of abelian varieties with **P**olarization, **E**ndomorphisms, and **L**evel structure (PEL). These are the Shimura varieties defined by *PEL type* Shimura data.

Shimura also studied several non-PEL cases and defined the corresponding Shimura varieties as varieties over number fields. Deligne [2] gave a group-theoretic framework for Shimura's work. His definition in terms of a reciprocity law for complex multiplication points is extrapolated from the Shimura–Taniyama reciprocity law for abelian varieties with complex multiplication. Deligne also constructed the canonical model for abelian type Shimura varieties. The proof of existence in the general case was completed by Milne based on ideas of Borovoi. See here for a short summary of the history by Milne [12, §6].

Example 1.10. Consider the two cases from Example 1.6. The unitary group $U(V)$ has no PEL type Shimura data. For the group GSp_{2g}, there exists a PEL type Shimura datum (GSp_{2g}, X). Consider a principal congruence level subgroup

$$K(n) = \ker\left(\mathrm{GSp}_{2g}(\widehat{\mathbb{Z}}) \longrightarrow \mathrm{GSp}_{2g}(\mathbb{Z}/n\mathbb{Z})\right)$$

with $n \geq 3$. Then the canonical model $\mathrm{Sh}_{K(n)}(\mathrm{GSp}_{2g}, X)$ can be described as a moduli space of principally polarized abelian varieties with level-n-structure. For example, if we look at $\mathbb{C}$-points and specialize to GL_2, then we obtain

$$\mathrm{Sh}_{K(n)}(\mathrm{GL}_2, X)(\mathbb{C}) \xrightarrow{\sim} \{(E, \eta)/\mathbb{C}\}/\sim \tag{1.8}$$

where the right hand side denotes the set of isomorphism classes of pairs (E, η) with

- E an elliptic curve over $\mathbb{C}$,
- $\eta : (\mathbb{Z}/n\mathbb{Z})^{\oplus 2} \xrightarrow{\sim} E[n]$ a choice of basis for the n-torsion.

The datum η is called a *level structure* for E. Proving (1.8) will be one of our first goals.

1.8. **Further topics.** We will say more about this when the time comes. For now, let us start looking at Shimura varieties in detail.

Part 1. The Shimura variety of GL_2

2. The upper half plane

In Example 1.1, we have introduced the action of $\mathrm{GL}_2(\mathbb{Q})$ on the union of upper and lower half plane $\mathbb{H}^\pm = \mathbb{C}\backslash\mathbb{R}$. Recall that it is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}. \tag{2.1}$$

In Example 1.5, we have seen that we are especially interested in actions by subgroups such as $\mathrm{GL}_2(\mathbb{Z})$ and $\Gamma(n)$. Our aim in this section is to give a definition of such *arithmetic subgroups* and to prove properties about their action on $\mathbb{H}^\pm$.

Note that elements of $\mathrm{GL}_2(\mathbb{Z})$ have determinant 1 or -1, and that the elements of determinant -1 interchange upper and lower half plane. So we will focus on the action of $\mathrm{SL}_2(\mathbb{Q})$ on the upper half plane $\mathbb{H} \subset \mathbb{H}^\pm$.

2.1. The fundamental domain.

Let $\mathcal{F}$ be the area defined by

$$\mathcal{F} = \left\{ \tau \in \mathbb{H} \;\middle|\; |\tau| \geq 1 \text{ and } -\frac{1}{2} \leq \mathrm{Re}(\tau) \leq \frac{1}{2} \right\}. \tag{2.2}$$

Its interior $\mathcal{F}^\circ$ is the open subset where $|\tau| > 1$ and $-1/2 \leq \mathrm{Re}(\tau) \leq 1/2$.

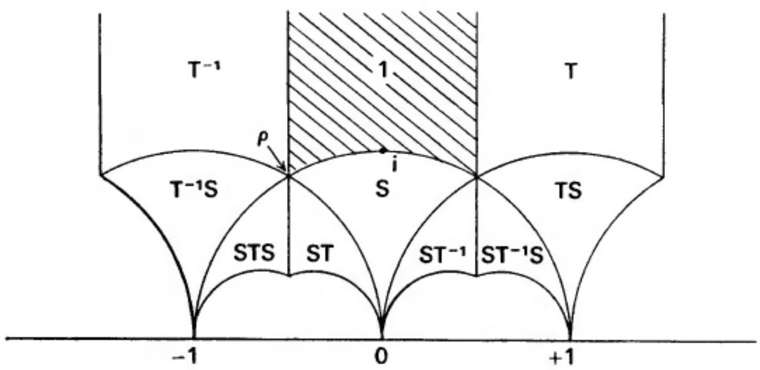

Figure 1. The area $\mathcal{F}$ is depicted in grey. The remaining areas show translates of $\mathcal{F}$ under the action of the elements S and T defined in (2.4). By Proposition 2.1 and Remark 2.2, these translates cover all of $\mathbb{H}$. The picture is taken from [16, §VII].

Proposition 2.1. *The set $\mathcal{F}$ is a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$ on $\mathbb{H}$. That is, it has the following two properties.*

(1) For every $\tau \in \mathbb{H}$, there exists $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma\tau \in \mathcal{F}$.

(2) $\mathcal{F}^\circ \cap \gamma\mathcal{F}^\circ = \emptyset$ whenever $\gamma \notin \{\pm 1\}$.

Proof. Fix $\tau \in \mathbb{H}$ and let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ be any element. By direct computation, we see that

$$\mathrm{Im}(\gamma\tau) = \mathrm{Im}\left(\frac{(a\tau + b)(c\bar{\tau} - d)}{|c\tau + d|^2}\right) = \frac{(ad - bc)\mathrm{Im}(\tau)}{|c\tau + d|^2} = \frac{\mathrm{Im}(\tau)}{|c\tau + d|^2}. \tag{2.3}$$

The denominator $|c\tau + d|^2$ defines a positive definite quadratic form in $(c, d) \in \mathbb{Z}^2$. It hence takes a minimum on the set of (c, d) that occur as the bottom row of an element of $\mathrm{SL}_2(\mathbb{Z})$. (These are precisely the (c, d) with $\gcd(c, d) = 1$.) So we see that $\{\mathrm{Im}(\gamma\tau) \mid \gamma \in \mathrm{SL}_2(\mathbb{Z})\}$ has a maximum.

Let γ be such that $\mathrm{Im}(\gamma\tau)$ is maximal. Consider the two matrices

$$S = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} \tag{2.4}$$

and observe that they act as the very simple transformations

$$S\tau = -\frac{1}{\tau}, \qquad T\tau = \tau + 1. \tag{2.5}$$

In particular, acting with a suitable power T^m, $m \in \mathbb{Z}$, we can translate $\gamma\tau$ to assume it lies in the strip $-1/2 \leq \mathrm{Re}(z) \leq 1/2$. Then also $|\gamma\tau| \geq 1$ because otherwise $\mathrm{Im}(S\gamma\tau) > \mathrm{Im}(\gamma\tau)$ would contradict the maximality of $\mathrm{Im}(\gamma\tau)$. This proves statement (1) of the proposition.

We now prove statement (2). Assume that τ and $\gamma\tau$ both lie in $\mathcal{F}^\circ$, our aim being to show that $\gamma \in \{\pm 1\}$. After possibly replacing the pair (γ, τ) by $(\gamma^{-1}, \gamma\tau)$, we can assume that $\mathrm{Im}(\gamma\tau) \geq \mathrm{Im}(\tau)$. Considering again (2.3), this means that $|c\tau + d|^2 \leq 1$.

Clearly, we now have $c = 0$ because $|c\tau + d| > 1$ for every $c \neq 0$ (use $\tau \in \mathcal{F}^\circ$). This means that γ is of the form

$$\gamma = \pm \begin{pmatrix} 1 & m \\ & 1 \end{pmatrix}$$

for some $m \in \mathbb{Z}$. Since both τ and $\gamma\tau$ have real part in $(-1/2, 1/2)$, the only possibility is $m = 0$. This finishes the proof. $\square$

Remark 2.2. One can show that the matrices S and T from (2.4) generate $\mathrm{SL}_2(\mathbb{Z})$. That is, every element of $\mathrm{SL}_2(\mathbb{Z})$ can be written as a product of the three elements S, T and T^{-1}. The proof is not difficult and can be found in [16, §VII.1, Theorem 2].

2.2. **Arithmetic subgroups of $\mathrm{SL}_2(\mathbb{Q})$.** We now define arithmetic subgroups of $\mathrm{SL}_2(\mathbb{Q})$.

Definition 2.3. (1) For $n \geq 1$, we define the *principal congruence subgroup* $\Gamma(n)$ by

$$\Gamma(n) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma \equiv 1 \bmod n\}.$$

(2) We call a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Q})$ *arithmetic* if it contains a principal congruence group $\Gamma(n)$ with finite index.

The group SL_2 has a very interesting property which will come up again later. Namely, for each $n \geq 1$, the projection map

$$\mathrm{SL}_2(\mathbb{Z}) \longrightarrow \mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z}) \tag{2.6}$$

is surjective. This is not hard to show directly, but also follows from Theorem 3.15 (2) below.

Example 2.4. By the surjectivity we just stated for SL_2, the image of the projection map $\mathrm{GL}_2(\mathbb{Z}) \to \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ is the set of matrices with determinant ± 1. In particular, this projection is not surjective when $n = 5$ or $n \geq 7$.

In the context of Definition 2.3, the surjectivity of (2.6) implies that $\Gamma(n) \trianglelefteq \mathrm{SL}_2(\mathbb{Z})$ is a normal subgroup of index equal to $|\mathrm{SL}_2(\mathbb{Z}/n\mathbb{Z})|$. In particular, if a group Γ contains $\Gamma(n)$ with finite index, then it also contains all $\Gamma(mn)$ with finite index.

Proposition 2.5. *Let Γ be an arithmetic subgroup.*

(1) There exists a lattice $\Lambda \subset \mathbb{Q}^2$ such that $\Gamma \subseteq \mathrm{SL}(\Lambda)$.

(2) More precisely, there exist an integer n and an element $g \in \mathrm{GL}_2(\mathbb{Q})$, $\det(g) > 0$, such that

$$\Gamma(m) \ \subseteq \ g\Gamma g^{-1} \ \subseteq \ \mathrm{SL}_2(\mathbb{Z}).$$

Proof. The two statements are proved by very simple and universal arguments. First, by assumption on Γ, there exists an integer n such that $\Gamma(n) \subseteq \Gamma$ with finite index. Let $\gamma_1, \dots, \gamma_r$ be representatives for the cosets $\Gamma/\Gamma(n)$. Then Γ stabilizes the lattice

$$\Lambda := \sum_{i=1}^{r} \gamma_i \cdot \mathbb{Z}^2.$$

Indeed, since $\gamma\mathbb{Z}^2 = \mathbb{Z}^2$ for every $\gamma \in \Gamma(n)$, we can also write Λ as

$$\Lambda = \sum_{\gamma \in \Gamma} \gamma \cdot \mathbb{Z}^2,$$

and from this second expression the Γ-stability is clear. This means that $\Gamma \subseteq \mathrm{SL}(\Lambda)$ which proves statement (1).

Let $\lambda_1, \lambda_2 \in \Lambda$ be a basis as $\mathbb{Z}$-module. Viewing λ_1 and λ_2 as column vectors, the base change matrix $g = (\lambda_1\ \lambda_2)$ lies in $\mathrm{GL}_2(\mathbb{Q})$ and has the property $g\mathbb{Z}^2 = \Lambda$. Changing λ_1 to $-\lambda_1$ if necessary, we may assume $\det(g) > 0$. Then $\mathrm{SL}_2(\mathbb{Z}) = g^{-1}\,\mathrm{SL}(\Lambda)g$ and hence $g\Gamma g^{-1} \subseteq \mathrm{SL}_2(\mathbb{Z})$.

We still need to show that $g\Gamma g^{-1}$ contains a principal congruence subgroup. This is the content of the next lemma which completes the proof. □

Lemma 2.6. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Q})$ be an arithmetic subgroup and $g \in \mathrm{GL}_2(\mathbb{Q})$. Then $g\Gamma g^{-1}$ is again an arithmetic subgroup.*

Proof. Let d be the least common multiple of all the denominators of all the entries of g and g^{-1}. Then, if $A \in d^2m\mathrm{M}_2(\mathbb{Z})$ is an integer matrix divisible by d^2m, we find $g^{-1}Ag \in m\mathrm{M}_2(\mathbb{Z})$. This shows that $g^{-1}\Gamma(d^2m)g \subseteq \Gamma(m)$ which is equivalent to

$$\Gamma(d^2m) \ \subseteq\ g\Gamma(m)g^{-1}. \tag{2.7}$$

Now, for the given Γ, choose n with $\Gamma(n) \subseteq \Gamma$. Conjugating this relation by g and using (2.7), we find $\Gamma(d^2n) \subseteq g\Gamma g^{-1}$ which proves that $g\Gamma g^{-1}$ is again arithmetic. □

In other words, Proposition 2.5 shows that the arithmetic subgroups in $\mathrm{SL}_2(\mathbb{Q})$ are precisely the $\mathrm{GL}_2(\mathbb{Q})$-conjugates of groups between $\mathrm{SL}_2(\mathbb{Z})$ and some $\Gamma(n)$.

2.3. Stabilizers.

Definition 2.7. We say that an arithmetic subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Q})$ is *neat* if it is torsion free.

Proposition 2.8. *Let Γ be a neat arithmetic subgroup of $\mathrm{SL}_2(\mathbb{Q})$. Then Γ acts with trivial stabilizers on $\mathbb{H}$. That is, if $\gamma\tau = \tau$ for some $\gamma \in \Gamma$ and $\tau \in \mathbb{H}$, then $\gamma = 1$.*

Proof. We have seen in (1.5) that the stabilizer of $i \in \mathbb{H}$ in $\mathrm{GL}_2(\mathbb{R})$ is a copy of $\mathbb{C}^\times$. The unit circle $\mathbb{C}^1 \subset \mathbb{C}^\times$ is compact and equals the intersection $\mathbb{C}^\times \cap \mathrm{SL}_2(\mathbb{R})$. For a general point $\tau \in \mathbb{H}$, we can write $\tau = g \cdot i$ for some $g \in \mathrm{SL}_2(\mathbb{R})$:

$$\begin{pmatrix} 1 & b \\ & 1 \end{pmatrix} \begin{pmatrix} a^{1/2} & \\ & a^{-1/2} \end{pmatrix} \cdot i = ai + b.$$

The stabilizers S_i and S_τ of τ and i in $\mathrm{SL}_2(\mathbb{R})$ are then related by $S_\tau = gS_ig^{-1}$. In this way, we see that for every $\tau \in \mathbb{H}$, the stabilizer $S_\tau \subset \mathrm{SL}_2(\mathbb{R})$ is isomorphic to $\mathbb{C}^1$, in particular compact.

Assume that $\gamma\tau = \tau$, where $\gamma \in \Gamma$ and $\tau \in \mathbb{H}$. This is equivalent to $\gamma \in \Gamma \cap S_\tau$. Since $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ is a discrete subgroup, the intersection $\Gamma \cap S_\tau$ is a discrete subgroup of S_τ. Since the discrete subgroups of $\mathbb{C}^1$ are all finite cyclic (generated by a root of unity), and since Γ is torsion-free by assumption, we see that $\Gamma \cap S_\tau = \{1\}$. Hence $\gamma = 1$, and the proof is complete. □

Example 2.9. The element $-1 \in \mathrm{SL}_2(\mathbb{Z})$ acts trivially on $\mathbb{H}$ because $(-\tau)/(-1) = \tau$ (substitute in (2.1)). The element $\left(\begin{smallmatrix} & -1 \\ 1 & \end{smallmatrix}\right)$, which has order 4, stabilizes the point i because $-1/i = i$.

The next proposition provides a simple criterion for detecting neatness.

Proposition 2.10. *For all $n \geq 3$, the principal congruence subgroup $\Gamma(n)$ is neat. In particular, if $\Gamma \subseteq \Gamma(n)$ is an arithmetic subgroup, then Γ is neat.*

Proof. The minimal polynomial $\Phi_d(T)$ of a primitive d-th root of unity has degree $\varphi(d)$ (Euler φ-function). Recall that $\Phi_d(T)$ is called the d-th cyclotomic polynomial and that

$$T^m - 1 = \prod_{d \mid m} \Phi_d(T)$$

because the roots of $T^m - 1$ are precisely the m-th roots of unity, and each such root of unity is a primitve d-th root of unity for a unique divisor $d \mid m$.

The only values for d such that $\varphi(d) \leq 2$ are 1, 2, 3, 4, and 6. These are precisely the values for d such that $\mathbb{Q}(\zeta_d)$ has degree ≤ 2 over $\mathbb{Q}$.

Let $n \geq 1$ and let $\gamma \in \mathrm{SL}_2(\mathbb{Q})$ be a torsion element, say $\gamma^m = 1$. Then the minimal polynomial of γ divides $T^m - 1$. We know that the minimal polynomial and the characteristic polynomial of a matrix have the same irreducible factors. So the characteristic polynomial $P(T)$ of γ is a product of $\Phi_d(T)$ with $d \mid m$. The only possibilities for $P(T)$ are hence[1]

$$(T-1)^2, \quad (T+1)^2, \quad (T-1)(T+1), \quad T^2+1, \quad T^2+T+1, \text{ and } T^2-T+1. \tag{2.8}$$

If $n \geq 3$ and if γ is integral with $\gamma \equiv 1 \bmod n$, then also $P(T) \equiv (T-1)^2 \bmod n$, leaving $P(T) = (T-1)^2$ as the only possibility. This means that γ is either equal to 1 or $\mathrm{GL}_2(\mathbb{Q})$-conjugate to $\left(\begin{smallmatrix} 1 & 1 \\ & 1 \end{smallmatrix}\right)$ (Jordan normal form). But γ is also a torsion element by assumption, so $\gamma = 1$ is the only possibility. $\square$

Exercise 2.11. Extend the argument of the previous proof to GL_n. That is, given $n \geq 1$, find an integer $m \geq 1$ such that for $\gamma \in \mathrm{GL}_n(\mathbb{Z})$,

$$\gamma \equiv 1 \bmod m \quad \Longrightarrow \quad \gamma \text{ non-torsion}.$$

Conclusion 2.12. In this lecture, we saw the definition of neat arithmetic subgroups of $\mathrm{SL}_2(\mathbb{Q})$. We have seen in Proposition 2.8 that such groups act freely on $\mathbb{H}$. So the quotient $\Gamma\backslash\mathbb{H}$ will be a Riemann surface and the quotient map

$$\mathbb{H} \longrightarrow \Gamma\backslash\mathbb{H} \tag{2.9}$$

a holomorphic covering map in the sense of topology. We have seen in Proposition 2.5 that, in order to study $\Gamma\backslash\mathbb{H}$, we may always assume $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$. Then we can think of $\Gamma\backslash\mathbb{H}$ as being glued from finitely many $\mathrm{SL}_2(\mathbb{Z})$-translates of the fundamental domain $\mathcal{F}$ as in Figure 2.1 along their edges.

3. Adelic double quotients

In this lecture, we study the adelic double quotients $\mathrm{GL}_2(\mathbb{Q})\backslash(\mathbb{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/K)$ and relate them to the quotients $\Gamma\backslash\mathbb{H}$ from the previous lecture. We will first revisit the definition of the adeles and explain the definition of $\mathrm{GL}_2(\mathbb{A}_f)$ as a topological group in more detail. In fact, we will use this opportunity to also study groups of the form $G(\mathbb{A}_f)$ more generally.

[1] The product $(T-1)(T+1)$ cannot actually occur, of course, because $\det(\gamma) = 1$ for $\gamma \in \mathrm{SL}_2(\mathbb{Q})$. This does not affect the argument, though.

3.1. **The adeles.** We begin by defining the ring of *integral adeles*. It is the profinite ring given by[2] $\widehat{\mathbb{Z}} := \lim \mathbb{Z}/n\mathbb{Z}$. The transition maps here are given by the projections $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, whenever $m \mid n$. Concretely, we have

$$\widehat{\mathbb{Z}} = \Big\{ (x_1, x_2, \dots) \in \prod_{n \geq 1} \mathbb{Z}/n\mathbb{Z} \;\Big|\; x_{dn} \equiv x_n \bmod n \text{ for all } d, n \geq 1 \Big\}.$$

Recall that the Chinese remainder theorem identifies $\mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} \prod_p \mathbb{Z}/p^{v_p(n)}\mathbb{Z}$. If we apply this identification to each term of the limit, then we obtain an isomorphism

$$\widehat{\mathbb{Z}} \xrightarrow{\sim} \prod_p \mathbb{Z}_p, \qquad (x_1, x_2, \dots) \longmapsto \big((x_1, x_p, x_{p^2}, \dots)\big)_p. \tag{3.1}$$

We endow each $\mathbb{Z}_p$ with the usual p-adic topology and their product with the product topology. Then (3.1) is an isomorphism of topological rings.

Definition 3.1. The *ring of finite adeles* is defined by $\mathbb{A}_f := \mathbb{Q} \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}$. Since $\widehat{\mathbb{Z}}$ is torsion-free, we can view it as a subring $\widehat{\mathbb{Z}} \subset \mathbb{A}_f$. We endow $\mathbb{A}_f$ with the topology such that $\widehat{\mathbb{Z}}$ is an open subring.

Let us unravel this definition. First, on the level of rings, $\mathbb{A}_f$ is the ring of fractions x/m with $x \in \widehat{\mathbb{Z}}$ and $m \geq 1$, where the usual rules of arithmetic apply. Using (3.1), we can more explicitly describe it as the subring

$$\mathbb{A}_f = \Big\{ (x_p) \in \prod_p \mathbb{Q}_p \;\Big|\; x_p \in \mathbb{Z}_p \text{ for almost all } p \Big\}.$$

Now we describe the topology. In $\widehat{\mathbb{Z}}$, a neighborhood basis of 0 is given by all the kernels of the projections $\widehat{\mathbb{Z}} \to \mathbb{Z}/n\mathbb{Z}$. These are precisely the ideals $n\widehat{\mathbb{Z}}$. Under the isomorphism (3.1), they are the subsets of the form

$$\prod_{p \in S} p^{m_p}\mathbb{Z}_p \times \prod_{p \notin S} \mathbb{Z}_p$$

where S is a finite set of primes and $(m_p)_{p \in S}$ a tuple of non-negative integers. Such sets forming a neighborhood basis of 0 means that the sets

$$\big\{ x + n\widehat{\mathbb{Z}} \mid x \in \widehat{\mathbb{Z}},\ n \geq 1 \big\} \tag{3.2}$$

give a basis of the topology on $\widehat{\mathbb{Z}}$. Declaring $\widehat{\mathbb{Z}} \subset \mathbb{A}_f$ an open subring then simply means that the sets $n\widehat{\mathbb{Z}}$ also form a neighborhood basis of 0 in $\mathbb{A}_f$. Equivalently, the sets

$$\big\{ x + n\widehat{\mathbb{Z}} \mid x \in \mathbb{A}_f,\ n \geq 1 \big\} \tag{3.3}$$

provide a basis for the topology on $\mathbb{A}_f$.

Definition 3.2. The ring of adeles is defined as the product $\mathbb{A} := \mathbb{A}_f \times \mathbb{R}$ endowed with the product topology.

Proposition 3.3. *The subring $\mathbb{Q} \subset \mathbb{A}$ is discrete.*

Proof. By definitions, the product $U = \widehat{\mathbb{Z}} \times (-1, 1)$ is an open subset of $\mathbb{A}$. The intersection $U \cap \mathbb{Q}$ consists of those rational numbers that lie in $\mathbb{Z} = \widehat{\mathbb{Z}} \cap \mathbb{Q}$ and in the interval $(-1, 1)$. In other words, $U \cap \mathbb{Q} = \{0\}$. Thus, $\{0\} \subset \mathbb{Q}$ is an open subset for the subspace topology. By additive translation invariance of the topology ($\mathbb{A}$ is a topological ring), the same argument applies for all rational numbers. This shows that the subspace topology on $\mathbb{Q}$ is the discrete topology as claimed. $\square$

[2]We use lim and colim to denote the limit and the colimit. In other references, these might be called $\varprojlim$ and $\varinjlim$.

Let $F/\mathbb{Q}$ be a finite extension. The adeles of F can be defined in the same way as for $\mathbb{Q}$. First, we define the integral adeles with profinite topology

$$\widehat{O}_F \;:=\; \lim_{\mathfrak{a}\subseteq O_F} O_F/\mathfrak{a} \;\xrightarrow{\sim}\; \prod_{\mathfrak{p}} O_{F,\mathfrak{p}}. \tag{3.4}$$

The we tensor by $\mathbb{Q}$ over $\mathbb{Z}$, or equivalently by F over O_F, to define the finite adeles:

$$\begin{aligned} \mathbb{A}_{F,f} \;&:=\; \mathbb{Q}\otimes_{\mathbb{Z}}\widehat{O}_F \\ &\xrightarrow{\sim} \Big\{(x_{\mathfrak{p}})\in\prod_{\mathfrak{p}} F_{\mathfrak{p}} \;\Big|\; x_{\mathfrak{p}}\in O_{F,\mathfrak{p}} \text{ for almost all } \mathfrak{p}\Big\}. \end{aligned} \tag{3.5}$$

Again, the topology on $\mathbb{A}_{F,f}$ is defined by declaring $\widehat{O}_F$ to be an open subring. Finally, we define the adeles as the product

$$\mathbb{A}_F \;:=\; \mathbb{A}_{F,f}\times(\mathbb{R}\otimes_{\mathbb{Q}} F) \;\xrightarrow{\sim}\; \mathbb{A}_{F,f}\times\prod_{\sigma:F\to\mathbb{R}}\mathbb{R}\;\times\prod_{\{\sigma,\overline{\sigma}\}:F\to\mathbb{C}}\mathbb{C}. \tag{3.6}$$

Here, the real factors have their real vector space topology, and the last two products are over the real (resp. complex) places of F.

Recall that O_F is a free abelian group of rank equal to $d=[F:\mathbb{Q}]$. Let $\alpha_1,\ldots,\alpha_d$ be a $\mathbb{Z}$-module basis of O_F. Such a choice provides isomorphisms of $\widehat{\mathbb{Z}}$-, $\mathbb{A}_f$-, resp. $\mathbb{A}$-modules

$$\widehat{\mathbb{Z}}^n \xrightarrow{\sim} O_F\otimes_{\mathbb{Z}}\widehat{\mathbb{Z}},\quad \mathbb{A}_f^n\xrightarrow{\sim} F\otimes_{\mathbb{Q}}\mathbb{A}_f,\quad \mathbb{A}^n\xrightarrow{\sim} F\otimes_{\mathbb{Q}}\mathbb{A}. \tag{3.7}$$

We endow $\widehat{\mathbb{Z}}^n$, $\mathbb{A}_f^n$ and $\mathbb{A}^n$ with the product topology and use the isomorphisms in (3.7) to define from this the topology on the three tensor products. This topology is independent of the choice of $\alpha_1,\ldots,\alpha_d$.

Remark 3.4. The previous definition is a general principle. Let R be a topological ring and let M be a finite free R-module. Any choice of R-basis $\alpha_1,\ldots,\alpha_d$ defines an isomorphism $R^d\xrightarrow{\sim} M$ and, in this way, endows M with a topology.

Any two such isomorphisms differ by an element of $\mathrm{GL}_d(R)$. Since the action of every $g\in\mathrm{GL}_d(R)$ on R^d is continuous, the topology is independent of the chosen basis.

Proposition 3.5. *Multiplication defines isomorphisms of topological rings*

$$O_F\otimes_{\mathbb{Z}}\widehat{\mathbb{Z}}\xrightarrow{\sim}\widehat{O}_F,\quad F\otimes_{\mathbb{Q}}\mathbb{A}_f\xrightarrow{\sim}\mathbb{A}_{F,f},\quad F\otimes_{\mathbb{Q}}\mathbb{A}\xrightarrow{\sim}\mathbb{A}_F.$$

Proof. Every ideal $\mathfrak{a}\subseteq O_F$ contains an ideal nO_F with $n\in\mathbb{Z}_{\geq 1}$. So we can rewrite (3.4) as $\widehat{O}_F=\lim O_F/nO_F$. Having chosen $\alpha_1,\ldots,\alpha_d$, we obtain

$$\begin{aligned} \widehat{O}_F \;&=\; \lim\Big(\bigoplus_{i=1}^d \mathbb{Z}/n\mathbb{Z}\cdot\alpha_i\Big) \\ &\xrightarrow{\sim} \bigoplus_{i=1}^d\big(\lim\mathbb{Z}/n\mathbb{Z}\big)\cdot\alpha_i \\ &\xrightarrow{\sim} \bigoplus_{i=1}^d\widehat{\mathbb{Z}}\cdot\alpha_i. \end{aligned}$$

This shows that $O_F\otimes_{\mathbb{Z}}\widehat{\mathbb{Z}}\xrightarrow{\sim}\widehat{O}_F$ as topological rings. The statements for $\mathbb{A}_{F,f}$ and $\mathbb{A}_F$ follow from this. $\square$

Corollary 3.6. *Let $F/\mathbb{Q}$ be a finite extension. Then $F\subset\mathbb{A}_F$ is discrete.*

Proof. Since $\mathbb{Q}$ is discrete in $\mathbb{A}$, we have that $\mathbb{Q}^n$ is discrete in $\mathbb{A}^n$. Choosing a $\mathbb{Q}$-basis $\alpha = (\alpha_1, \dots, \alpha_d)$ for F, we obtain a commutative square of the form

$$\begin{tikzcd} \mathbb{Q}^n \arrow[r, hook] \arrow[d, equal, "\alpha"'] & \mathbb{A}^n \arrow[d, equal, "\alpha"] \\ F \arrow[r, hook] & \mathbb{A}_F. \end{tikzcd}$$

By Proposition 3.5, the right vertical identification is a homeomorphism. Hence we obtain that F is discrete in $\mathbb{A}_F$. □

3.2. **Groups of the form $G(\mathbb{A}_f)$.** Let us formulate the problem more generally.

Question 3.7. Let X be an affine variety[3] over $\mathbb{Q}$ and let R be a topological $\mathbb{Q}$-algebra. We assume that points of R are closed. For example, R could be $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}_p$, $\mathbb{A}_f$ or $\mathbb{A}$. How to define the topological space $X(R)$ in a natural way?

The answer is very simple. Let us write $\mathcal{A}^N = \operatorname{Spec} \mathbb{Q}[t_1, \dots, t_N]$ for affine N-space over $\mathbb{Q}$ to avoid confusion with the adele notation. We endow $\mathcal{A}^N(R) = R^N$ with the product topology.

Let $f_1, \dots, f_m \in \mathbb{Q}[t_1, \dots, t_N]$ be polynomials and let $X = V(f_1, \dots, f_m) \subseteq \mathcal{A}^N$ be their vanishing locus. Then $X(R) \subseteq R^N$ is a closed subset because it equals the intersection $\cap_{i=1}^m f_i^{-1}(0)$, and we endow it with the subspace topology.

Definition 3.8. Let X be an affine $\mathbb{Q}$-variety. Choose a presentation $\varphi : X \xrightarrow{\sim} V(f_1, \dots, f_m)$ as above. The topology on $X(R)$ is defined as the subspace topology with respect to $\varphi(R) : X(R) \hookrightarrow R^N$.

Lemma 3.9. *This topology on $X(R)$ is independent of the choices of N, $(f_1, \dots, f_m)$ and φ.*

Proof. Assume that we are given two affine varieties $V(f_1, \dots, f_{m_1}) \subseteq \mathcal{A}^{N_1}$ as well as $V(g_1, \dots, g_{m_2}) \subseteq \mathcal{A}^{N_2}$. Assume that

$$\varphi : V(f_1, \dots, f_{m_1}) \xrightarrow{\sim} V(g_1, \dots, g_{m_2})$$

is an isomorphism of $\mathbb{Q}$-varieties. Then φ and $\psi = \varphi^{-1}$ lift to morphisms $\Phi : \mathcal{A}^{N_1} \to \mathcal{A}^{N_2}$ and $\Psi : \mathcal{A}^{N_2} \to \mathcal{A}^{N_1}$. The induced maps

$$R^{N_1} \underset{\Psi}{\overset{\Phi}{\rightleftarrows}} R^{N_2}$$

are continuous because they are given by polynomials. Hence their restrictions φ and ψ are continuous as well. Since $\psi = \varphi^{-1}$, this shows that φ is a homeomorphism. □

Example 3.10. Consider the group variety GL_n. One possible presentation as a closed subset of an affine space is given by

$$\begin{aligned} \mathrm{GL}_n \quad &\xrightarrow{\sim} \quad V\big(1 - t \cdot \det\big((t_{ij})_{i,j=1}^n\big)\big) \subset \mathcal{A} \times_{\operatorname{Spec}(\mathbb{Q})} \mathcal{A}^{n^2} \\ g = (t_{ij})_{i,j=1}^n \quad &\longmapsto \quad (\det(g)^{-1}, g). \end{aligned}$$

For example, if $n = 1$, we recover the closed immersion[4]

$$\mathbb{G}_m \hookrightarrow \mathcal{A}^2, \quad t \longmapsto (t^{-1}, t).$$

According to Definition 3.8, the topology on $\mathrm{GL}_n(\mathbb{A}_f)$ is then given as the subspace topology with respect to

$$\mathrm{GL}_n(\mathbb{A}_f) \hookrightarrow \mathbb{A}_f \times \mathrm{M}_n(\mathbb{A}_f), \quad g \longmapsto (\det(g)^{-1}, g).$$

[3] More generally, an affine finite type $\mathbb{Q}$-scheme.

[4] $\mathbb{G}_m$ is just another notation for GL_1. The notation symbolizes *multiplicative group*.

The product $\widehat{\mathbb{Z}} \times \mathrm{M}_n(\widehat{\mathbb{Z}})$ is an open subset on the right hand side. So the intersection

$$\mathrm{GL}_n(\widehat{\mathbb{Z}}) = \mathrm{GL}_n(\mathbb{A}_f) \cap (\widehat{\mathbb{Z}} \times \mathrm{M}_n(\widehat{\mathbb{Z}}))$$

is an open subset of $\mathrm{GL}_n(\mathbb{A}_f)$. (The elements of $\mathrm{GL}_n(\widehat{\mathbb{Z}})$ are precisely those elements of $\mathrm{GL}_n(\mathbb{A}_f) \cap \mathrm{M}_n(\widehat{\mathbb{Z}})$ whose inverse determinant again lies in $\widehat{\mathbb{Z}}$.) As a closed subset of the profinite set $\widehat{\mathbb{Z}} \times \mathrm{M}_n(\widehat{\mathbb{Z}})$, $\mathrm{GL}_n(\widehat{\mathbb{Z}})$ is again profinite. In fact, we have

$$\mathrm{GL}_n(\widehat{\mathbb{Z}}) \xrightarrow{\sim} \lim_{m \geq 1} \mathrm{GL}_n(\mathbb{Z}/m\mathbb{Z})$$

as topological group. The principal congruence subgroups

$$K(m) := \ker\big(\mathrm{GL}_n(\widehat{\mathbb{Z}}) \longrightarrow \mathrm{GL}_n(\mathbb{Z}/m\mathbb{Z})\big)$$

form a neighborhood basis of 1 in $\mathrm{GL}_n(\mathbb{A}_f)$.

Example 3.11. We always view $\mathbb{A}_f^\times$ with the topology coming from $\mathbb{A}_f^\times = \mathbb{G}_m(\mathbb{A}_f)$. Then the inclusion map $\mathbb{A}_f^\times \to \mathbb{A}_f$ is continuous because it is induced from the morphism of varieties $\mathbb{G}_m \to \mathcal{A}$, $t \mapsto t$. But it is not an open immersion. For example, $\widehat{\mathbb{Z}}^\times$ is open in $\mathbb{A}_f^\times$, but not in $\mathbb{A}_f$.

Exercise 3.12. Prove the claim in the previous example. That is, show that none of the open subsets $1 + n\widehat{\mathbb{Z}} \subseteq \widehat{\mathbb{Z}}$, which form a neighborhood basis of $1 \in \widehat{\mathbb{Z}}$, is contained in $\widehat{\mathbb{Z}}^\times$.

Example 3.13. Let G be a general linear algebraic group over $\mathbb{Q}$. There always exist some $N \geq 1$ and a closed immersion $G \hookrightarrow \mathrm{GL}_N$. Then $G(\mathbb{A}_f) \subseteq \mathrm{GL}_N(\mathbb{A}_f)$ has the subspace topology. In particular, the intersections $G(\mathbb{A}_f) \cap K(m)$ with all congruence subgroups form a neighborhood basis of $1 \in G(\mathbb{A}_f)$.

This applies, for example, to the standard representations

$$\mathrm{SL}_2 \hookrightarrow \mathrm{GL}_2, \quad \mathrm{Sp}_{2g} \hookrightarrow \mathrm{GL}_{2g}, \quad \mathrm{GSp}_{2g} \hookrightarrow \mathrm{GL}_{2g}\,.$$

Let V be a quadratic $\mathbb{Q}$-vector space. Then it applies to the closed immersions

$$SO(V) \hookrightarrow \mathrm{GL}(V), \quad O(V) \hookrightarrow \mathrm{GL}(V).$$

Remark 3.14. For local fields k, such as $k \in \{\mathbb{R}, \mathbb{C}, \mathbb{Q}_p\}$, the situation is more straightforward in the following sense. If $X \hookrightarrow Y$ is an open immersion of k-varieties, then $X(k) \to Y(k)$ is an open immersion with respect to the topologies from Definition 3.8. In particular, the topology on $X(k)$ from Definition 3.8 agrees with the subspace topology in $Y(k)$.

This remark applies, for example, to

$$\mathrm{GL}_n(\mathbb{R}) \subset \mathrm{M}_n(\mathbb{R}) \quad \text{and} \quad \mathrm{GL}_n(\mathbb{Q}_p) \subset \mathrm{M}_n(\mathbb{Q}_p).$$

3.3. **General adelic double quotients.** Let us begin with a general theorem which we will not prove.

Theorem 3.15 ([13, Theorem 4.16])**.** *(1) Let $G/\mathbb{Q}$ be a connected reductive algebraic group. Then, for every compact open subgroup $K \subset G(\mathbb{A}_f)$, the double quotient $G(\mathbb{Q})\backslash G(\mathbb{A}_f)/K$ is finite.*

(2, Strong approximation) Let $G/\mathbb{Q}$ be a connected, simply connected semi-simple group of non-compact type. Then $G(\mathbb{Q})$ is dense in $G(\mathbb{A}_f)$. In particular, for every compact open subgroup $K \subset G(\mathbb{A}_f)$,

$$G(\mathbb{A}_f) = \big\{\gamma \cdot k \ \mid \ \gamma \in G(\mathbb{Q}),\ k \in K\big\}.$$

As our first application, we obtain a more concrete description of the adelic double quotients that make up the complex points of a Shimura variety (1.7). Let (G, X) be a Shimura datum and let $K \subset G(\mathbb{A}_f)$ be a level subgroup. In particular, G is a connected reductive group over $\mathbb{Q}$, so Theorem 3.15 (1) applies. So we find finitely many double coset representatives $g_1, \dots, g_r \in G(\mathbb{A}_f)$,

$$G(\mathbb{A}_f) = \bigsqcup_{i=1}^{r} G(\mathbb{Q}) g_i K. \tag{3.8}$$

Each of the sets on the right hand side of (3.8) is $G(\mathbb{Q})$-stable. Moreover, $G(\mathbb{Q})$ acts transitively on the cosets $G(\mathbb{Q})g_iK/K$, and the stabilizer of the coset $g_iK \in G(\mathbb{Q})g_iK/K$ is the subgroup

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

So we obtain

$$\begin{aligned} G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K) &= \bigsqcup_{i=1}^{r} G(\mathbb{Q})\backslash(X \times G(\mathbb{Q})g_iK/K) \\ &\xrightarrow{\sim} \bigsqcup_{i=1}^{r} \Gamma_i\backslash X \times \{g_iK\}. \end{aligned} \tag{3.9}$$

If K is small enough, which we will make precise for GL_2 in a minute, then each Γ_i is torsion-free and acts without stabilizers on X. Each quotient $\Gamma_i\backslash X$ is then a complex manifold in the same way as we saw before in Conclusion 2.12.

Exercise 3.16. Work out (3.9) for yourself. For example, first prove the following variant. Let H be a group acting on sets X and Y. Let $Y = \sqcup_{i\in I} G \cdot y_i$ be the decomposition of Y into orbits and let Γ_i be the stabilizer of y_i in H. Then

$$H\backslash(X \times Y) \xrightarrow{\sim} \bigsqcup_{i\in I} \Gamma_i\backslash X.$$

Specialize to the situation $H = G(\mathbb{Q})$ and $Y = G(\mathbb{A}_f)/K$.

Exercise 3.17. The group SL_n is connected, simply connected, semi-simple and of non-compact type, so $\mathrm{SL}_n(\mathbb{Q}) \subset \mathrm{SL}_n(\mathbb{A}_f)$ is dense (Strong approximation, see Theorem 3.15 (2)). Using this property, show that

$$\mathrm{SL}_n(\mathbb{Z}) \longrightarrow \mathrm{SL}_n(\mathbb{Z}/m\mathbb{Z})$$

is surjective for all $m \geq 1$. In particular, this shows the surjectivity of (2.6).

3.4. **Back to GL_2.** The description in (3.9) is still quite abstract. We now want to make it completely explicit for congruence subgroups of GL_2. Let us begin by studying $\mathbb{G}_m$.

Proposition 3.18. *Let $K(m) = \ker\left(\widehat{\mathbb{Z}}^\times \to (\mathbb{Z}/m\mathbb{Z})^\times\right)$ be the m-th congruence subgroup of $\mathbb{A}_f^\times$. Then there is an isomorphism*

$$\mathbb{Q}_{>0}^\times\backslash\mathbb{A}_f^\times/K(m) \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^\times. \tag{3.10}$$

Proof. Let $x = (x_p)_p \in \mathbb{A}_f^\times$ be an element. Here, the component x_p lies in $\mathbb{Q}_p^\times$, and almost all components x_p even lie in $\mathbb{Z}_p^\times$. For each prime p, let $v_p : \mathbb{Q}_p^\times \to \mathbb{Z}$ denote the valuation normalized by $v_p(p) = 1$. Take the vector of valuations of all the entries of x:

$$(e_p)_p, \quad e_p = v_p(x_p).$$

Only finitely many of the e_p are non-zero. There is a rational number in $\mathbb{Q}_{>0}$ with the same valuations, namely $t = \prod_p p^{e_p}$. So $t^{-1}x$ lies in $\widehat{\mathbb{Z}}^\times$ which shows that every double coset in (3.10) has a representative in $\widehat{\mathbb{Z}}^\times$. Purely formally, we now obtain

$$\mathbb{Q}_{>0}^\times \backslash \mathbb{A}_f^\times / K(m) \xrightarrow{\sim} (\mathbb{Q}_{>0}^\times \cap \widehat{\mathbb{Z}}^\times) \backslash \widehat{\mathbb{Z}}^\times / K(m). \tag{3.11}$$

The rational number t is, in fact, uniquely determined which reflects that $\mathbb{Q}_{>0}^\times \cap \widehat{\mathbb{Z}}^\times = \{1\}$. So (3.11) simplifies to $\widehat{\mathbb{Z}}^\times / K(m)$, which is isomorphic to $(\mathbb{Z}/m\mathbb{Z})^\times$ as claimed. □

We write $\mathrm{GL}_n(\mathbb{Q})_{>0}$ for the subgroup of elements of $\mathrm{GL}_n(\mathbb{Q})$ with positive determinant.

Proposition 3.19. *Let $K \subset \mathrm{GL}_n(\mathbb{A}_f)$ be an open compact subgroup. The determinant map $\det : \mathrm{GL}_n(\mathbb{A}_f) \to \mathbb{A}_f^\times$ induces a bijection*

$$\det : \mathrm{GL}_n(\mathbb{Q})_{>0} \backslash \mathrm{GL}_n(\mathbb{A}_f)/K \xrightarrow{\sim} \mathbb{Q}_{>0}^\times \backslash \mathbb{A}_f^\times / \det(K). \tag{3.12}$$

Proof. The group SL_n is connected, simply connected, semi-simple and of non-compact type, so $\mathrm{SL}_n(\mathbb{Q}) \subset \mathrm{SL}_n(\mathbb{A}_f)$ is dense (Strong approximation, see Theorem 3.15 (2)). We will use this property freely.

Consider the determinant map in (3.12). It is clearly surjective because already the map $\det : \mathrm{GL}_n(\mathbb{A}_f) \to \mathbb{A}_f^\times$ is surjective. So our task is to prove that (3.12) is injective.

The source in (3.12) is only a set, so we cannot argue with kernels. Instead, we consider two elements $g_1, g_2 \in \mathrm{GL}_n(\mathbb{A}_f)$ with the same image, meaning that

$$\det(g_1) \in \mathbb{Q}_{>0}^\times \det(g_2)\det(K). \tag{3.13}$$

Our task is to show that $g_1 \in \mathrm{GL}_n(\mathbb{Q}) g_2 K$.

First, observe that $\det : \mathrm{GL}_n(\mathbb{Q})_{>0} \to \mathbb{Q}_{>0}^\times$ is surjective. So we find elements $h \in \mathrm{GL}_n(\mathbb{Q})_{>0}$ and $k \in K$ such that $\det(g_1) = \det(hg_2k)$. So after replacing g_2 by hg_2k, we may assume $\det(g_1) = \det(g_2)$.

Next, we consider the conjugate group $g_2Kg_2^{-1}$. Strong approximation for SL_n implies that

$$\mathrm{SL}_n(\mathbb{A}_f) = \mathrm{SL}_n(\mathbb{Q}) \cdot (g_2Kg_2^{-1} \cap \mathrm{SL}_n(\mathbb{A}_f)).$$

Hence, there are $h' \in \mathrm{SL}_n(\mathbb{Q})$ and $k' \in K \cap \mathrm{SL}_n(\mathbb{A}_f)$ with

$$g_1g_2^{-1} = h'g_2k'g_2^{-1}.$$

This is equivalent to $g_1 = h'g_2k'$, showing that the double cosets of g_1 and g_2 are equal as claimed. □

Corollary 3.20. *Let $K(m) \subset \mathrm{GL}_2(\mathbb{A}_f)$ be the m-th congruence subgroup. There is a bijection of connected components*

$$\pi_0\big(\mathrm{GL}_2(\mathbb{Q}) \backslash (\mathbb{H}^\pm \times \mathrm{GL}_2(\mathbb{A}_f)/K(m))\big) \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^\times. \tag{3.14}$$

Moreover, the connected components are all of the form $\Gamma\backslash\mathbb{H}$ with $\Gamma = \mathrm{GL}_2(\mathbb{Q})_{>0} \cap gK(m)g^{-1}$ for some element $g \in \mathrm{GL}_2(\mathbb{A}_f)$.

Proof. The two connected components of $\mathbb{H}^\pm$ are interchanged by the elements of negative determinant in $\mathrm{GL}_2(\mathbb{Q})$. Hence, we obtain

$$\begin{aligned}\pi_0\big(\mathrm{GL}_2(\mathbb{Q}) \backslash (\mathbb{H}^\pm \times \mathrm{GL}_2(\mathbb{A}_f)/K(m))\big) &\xrightarrow{\sim} \pi_0\big(\mathrm{GL}_2(\mathbb{Q})_{>0} \backslash (\mathbb{H} \times \mathrm{GL}_2(\mathbb{A}_f)/K(m))\big) \\ &\xrightarrow{\sim} \quad \mathrm{GL}_2(\mathbb{Q})_{>0} \backslash \mathrm{GL}_2(\mathbb{A}_f)/K(m).\end{aligned} \tag{3.15}$$

Here, the second isomorphism simply used that $\mathbb{H}$ is connected. Next, observe that

$$L := \det(K(m)) = \ker(\widehat{\mathbb{Z}}^\times \longrightarrow (\mathbb{Z}/m\mathbb{Z})^\times)$$

is the m-th congruence subgroup in $\mathbb{A}_f^\times$. So, by Proposition 3.19, the determinant allows to rewrite (3.15) as

$$\det : \mathrm{GL}_2(\mathbb{Q})_{>0}\backslash \mathrm{GL}_2(\mathbb{A}_f)/K(m) \xrightarrow{\sim} \mathbb{Q}_{>0}^\times\backslash\mathbb{A}_f^\times/L.$$

By Proposition 3.18, the last expression can be identified with $(\mathbb{Z}/m\mathbb{Z})^\times$ as claimed.

The final statement (each connected component being isomorphic to some $\Gamma\backslash\mathbb{H}$ with Γ of the form $\mathrm{GL}_2(\mathbb{Q})_{>0}\cap gK(m)g^{-1}$) is a special case of the decomposition in (3.9), except that we have already replaced $(\mathbb{H}^\pm, \mathrm{GL}_2(\mathbb{Q}))$ by $(\mathbb{H}, \mathrm{GL}_2(\mathbb{Q})_{>0})$. □

Let us go further and prove a criterion that ensures that all the occurring Γ are torsion free. The arguments will be similar to the ones we saw in §2.3.

Proposition 3.21. *For any $m \geq 3$ and $g \in \mathrm{GL}_2(\mathbb{A}_f)$, the intersection $\Gamma = \mathrm{GL}_2(\mathbb{Q}) \cap gK(m)g^{-1}$ is torsion free.*

Proof. Let γ be an element of $K(m)$. Then, since $K(m) \subseteq \mathrm{GL}_2(\widehat{\mathbb{Z}})$, the characteristic polynomial $P_\gamma(T)$ lies in $\widehat{\mathbb{Z}}[T]$. Since $\gamma \equiv 1 \bmod m$, we even know $P_\gamma(T) \equiv (T-1)^2 \bmod m$. In general, for every $n \geq 1$ and any ring R, the characteristic polynomial of an element from $\mathrm{GL}_n(R)$ is invariant under conjugation. So, in our setting, the same properties hold for $P_\gamma(T)$ for $\gamma \in gK(m)g^{-1}$.

Assume that $\gamma \in \Gamma = \mathrm{GL}_2(\mathbb{Q}) \cap gK(m)g^{-1}$. Then, the characteristic polynomial of γ has rational coefficients, and hence lies in the intersection

$$\mathbb{Q}[T] \cap \big((T-1)^2 + n\widehat{\mathbb{Z}}[T]\big).$$

This means that $P_\gamma(T) \in \mathbb{Z}[T]$ and $P_\gamma(T) \equiv (T-1)^2 \bmod m$.

If γ is a torsion element, then we have already seen during the proof of Proposition 2.10 that $P_\gamma(T)$ comes from the list (2.8). By the congruence condition we just established, the only possibility is $P_\gamma(T) = (T-1)^2$. The matrix $\left(\begin{smallmatrix} 1 & 1 \\ & 1 \end{smallmatrix}\right)$ is not torsion, so cannot be the Jordan normal form of γ. We conclude that $\gamma = 1$, showing that Γ is torsion-free as claimed. □

Conclusion 3.22. Let us come back to the situation of Corollary 3.20. Assume that $m \geq 3$. Then the connected components of

$$\mathrm{GL}_2(\mathbb{Q})\backslash\big(\mathbb{H}^\pm \times \mathrm{GL}_2(\mathbb{A}_f)\big)$$

are in natural bijection with $(\mathbb{Z}/m\mathbb{Z})^\times$. Each connected component is of the form $\Gamma\backslash\mathbb{H}$ for a torsion free arithmetic subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Q})$.

4. Group schemes

Our next aim is to endow the complex manifolds $\mathrm{Sh}_K(\mathrm{GL}_2, \mathbb{H}^\pm)(\mathbb{C})$ with an algebraic structure and to even define them over $\mathbb{Q}$ (see 1.7). This relies on their description as moduli spaces of elliptic curves:

Definition 4.1. Let k be a field. An *elliptic curve* over k is a proper, smooth, connected and 1-dimensional k-group scheme.

Later in the course, we will also consider other Shimura varieties and describe them as moduli spaces of abelian varieties:

Definition 4.2. An *abelian variety* over k is a proper, smooth and connected k-group scheme.

In this lecture, we will first discuss some background on group schemes. This will also be useful for talking about groups like GL_n, GSp_{2g} etc. which we have secretly already considered as group schemes over $\operatorname{Spec}\mathbb{Z}$ or $\operatorname{Spec}\mathbb{Q}$ in previous lectures. In general, group schemes are also an interesting topic in itself and come up in many areas of algebra.

Recommended reading closely related to our course: My lecture notes on moduli spaces of elliptic curves [10]. Parts of our discussion here are taken from [10, §2].

General reference on algebraic groups: Milne's book [14], especially §1 about basic definitions.

4.1. **Basic definitions.** We give the definition over a general base S, but the case to keep in mind is $S = \mathrm{Spec}(k)$ for a field k.

Definition 4.3. Let S be a scheme. A *group scheme* over S is a pair (G, m) that consists of an S-scheme G and an S-scheme morphism (called multiplication morphism)

$$m : G \times_S G \longrightarrow G$$

such that for every S-scheme T, the resulting map on T-valued points

$$m(T) : G(T) \times G(T) \longrightarrow G(T)$$

makes $G(T)$ into a group. We call G *commutative* if $G(T)$ is a commutative group for every T.

Observe that for every morphism $u : T' \to T$ of S-schemes, the diagram

$$\begin{tikzcd} G(T) \times G(T) \arrow[r, "m(T)"] \arrow[d, "u^* \times u^*"'] & G(T) \arrow[d, "u^*"] \\ G(T') \times G(T') \arrow[r, "m(T')"] & G(T') \end{tikzcd} \tag{4.1}$$

commutes which means that $u^* : G(T) \to G(T')$ is a group homomorphism. Furthermore, if (G, m) is a group scheme over S, then the Yoneda Lemma implies the existence of two additional S-scheme morphisms:

$$\begin{aligned} e : S &\longrightarrow G, && \text{(neutral element section)} \\ i : G &\longrightarrow G, && \text{(inversion morphism).} \end{aligned} \tag{4.2}$$

The first one is simply the neutral element $e \in G(S)$ of the group $G(S)$. Given $u : T \to S$, the pullback $u^*(e) = e \circ u \in G(T)$ is the neutral element of $G(T)$. The second one is characterized as the unique morphism that provides the inverse in all the groups $\{G(T)\}_{T \to S}$:

$$i(T) : G(T) \longrightarrow G(T), \quad g \longmapsto g^{-1}.$$

The datum (G, m, e, i) satisfies the group axioms in a scheme sense, meaning that the three diagrams

$$\begin{tikzcd} S \times_S G \arrow[rr, "e \times \mathrm{id}"] \arrow[dr, equal] & & G \times_S G \arrow[dl, "m"] \\ & G, & \end{tikzcd} \tag{4.3}$$

$$\begin{tikzcd} G \times_S G \arrow[r, "\mathrm{id} \times i"] \arrow[d] & G \times_S G \arrow[d, "m"] \\ S \arrow[r, "e"] & G, \end{tikzcd} \qquad \begin{tikzcd} G \times_S G \times_S G \arrow[r, "m \times \mathrm{id}"] \arrow[d, "\mathrm{id} \times m"'] & G \times_S G \arrow[d, "m"] \\ G \times_S G \arrow[r, "m"] & G. \end{tikzcd} \tag{4.4}$$

all commute. In fact, one may also reverse the above logic and obtains the more classical definition of a group scheme over S: It is the same as an S-scheme G together with a morphism $m : G \times_S G \to G$ such that there exist morphisms $e : S \to G$ and $i : G \to G$ such that the diagrams in (4.3) and (4.4) commute. The group scheme (G, m) is commutative if and only if multiplication interchanges with switching the factors in the sense that also the following diagram commutes:

$$\begin{tikzcd} G \times_S G \arrow[rr, "(g{,}h) \mapsto (h{,}g)"] \arrow[dr, "m"'] & & G \times_S G \arrow[dl, "m"] \\ & G. & \end{tikzcd} \tag{4.5}$$

Definition 4.4. Let (G_1, m_1) and (G_2, m_2) be group schemes over S. A group scheme morphism from G_1 to G_2 is a morphism of S-schemes $f : G_1 \to G_2$ such that $m_2 \circ (f \times f) = f \circ m_1$. Equivalently, it is an S-morphism f such that for all $T \to S$, the induced map

$$f(T) : G_1(T) \longrightarrow G_2(T)$$

is a group homomorphism.

If (G, m) is a *commutative* S-group scheme, then $\mathrm{End}_{S-\mathrm{Grp.Sch.}}(G, m)$ forms a (possibly non-commutative) ring because endomorphisms can be "added" (meaning multiplied in G) and multiplied (meaning composed). Concretely, sum and product of two elements $f, g \in \mathrm{End}(G)$ are given by

$$f + g := m \circ (f, g), \qquad fg := f \circ g.$$

In particular, we can add the identity n times to itself and obtain the n-th power endomorphism $[n] : G \to G$. On each of the groups $G(T)$, it is given by $[n](g) = g^n$. This is even defined for $n \in \mathbb{Z}$ by $[n] \circ i = [-n]$. In total, these give the ring map

$$[\,\cdot\,] : \mathbb{Z} \to \mathrm{End}(G). \tag{4.6}$$

Coming back to general group schemes, we next define kernels. This is straightforward because fiber products exist in the category of S-schemes. (Defining quotients, on the other hand, is tricky. We refer to [10, §13] for some cases.)

Definition 4.5. Let $f : G_1 \to G_2$ be a homomorphism of S-group schemes. Let $e_2 : S \to G_2$ be the neutral element section of G_2. The kernel of f is defined as the fiber product

$$\begin{tikzcd} \ker(f) \arrow[r] \arrow[d] & S \arrow[d, "e_2"] \\ G_1 \arrow[r, "f"] & G_2. \end{tikzcd} \tag{4.7}$$

It is clear from its definition that $\ker(f)$ has the property

$$\ker(f)(T) = \ker\left(f(T) : G_1(T) \longrightarrow G_2(T)\right). \tag{4.8}$$

The multiplication morphism of G_1 restricts to a multplication on $\ker(f)$ which makes $\ker(f)$ into a group scheme:

$$\begin{tikzcd} \ker(f) \times_S \ker(f) \arrow[r, dashed] \arrow[d] & \ker(f) \arrow[d] \\ G \times_S G \arrow[r, "m"] & G. \end{tikzcd} \tag{4.9}$$

Remark 4.6. Recall that if $X \to S$ is a separated morphism, then every section $\sigma : S \to X$ is a closed immersion. Thus, if $G \to S$ is a separated group scheme (e.g. affine or proper), then the neutral element e is a closed immersion. It follows that if in (4.7) $G_2 \to S$ is separated, then $\ker(f) \to G_1$ is a closed immersion.

4.2. **A commutative example: The multiplicative group.** Assume that $S = \operatorname{Spec} R$ is affine. Define $\mathbb{G}_{m,S} = \operatorname{Spec} R[t, t^{-1}]$ which we would like to make into a group scheme over S. Recall that $\operatorname{Spec}(-)$ is an anti-equivalence from R-algebras to affine S-schemes. We define the multiplication map $m : \mathbb{G}_{m,S} \times_S \mathbb{G}_{m,S} \to \mathbb{G}_{m,S}$ as $\operatorname{Spec}(m^*)$ where m^* is

$$\begin{aligned} m^* : R[t, t^{-1}] &\longrightarrow R[t, t^{-1}] \otimes_R R[t, t^{-1}] \\ t &\longmapsto t \otimes t. \end{aligned} \tag{4.10}$$

We next verify that this makes $\mathbb{G}_{m,S}$ into an S-group scheme. For every S-scheme T, we identify

$$\begin{aligned} \mathbb{G}_{m,S}(T) &\xrightarrow{\sim} \mathcal{O}_T(T)^\times \\ [g : T \to \mathbb{G}_{m,S}] &\longmapsto g^*(t). \end{aligned} \tag{4.11}$$

Note that this map is obviously defined; the fact that it is an isomorphism is the adjunction $\operatorname{Mor}_S(T, \operatorname{Spec}(A)) \xrightarrow{\sim} \operatorname{Hom}_R(A, \mathcal{O}_T(T))$. Given two morphisms $g_1, g_2 : T \to \mathbb{G}_{m,S}$, we compute the (dual of the) composition $m \circ (g_1, g_2)$ by

$$\begin{array}{ccccc} R[t, t^{-1}] & \xrightarrow{m^*} & R[t, t^{-1}] \otimes_R R[t, t^{-1}] & \xrightarrow{g_1^* \otimes g_2^*} & \mathcal{O}_T(T) \\ t & \longmapsto & t \otimes t & \longmapsto & g_1^*(t) g_2^*(t). \end{array}$$

Thus we see that the operation $m(T)$ on $\mathbb{G}_{m,S}(T)$ translates to the usual multiplication under (4.11). In particular, $m(T)$ is a group structure for every T, and hence $(\mathbb{G}_{m,S}, m)$ a group scheme.

We can next calculate the neutral element e and the inversion map i from (4.2). Under (4.11), the unit element $1 \in R^\times$ corresponds to

$$e^* : R[t, t^{-1}] \longrightarrow R, \quad t \longmapsto 1.$$

Taking $e = \operatorname{Spec}(e^*)$ gives the neutral element section. The inversion map $i = \operatorname{Spec}(i^*)$ is given by

$$i^* : R[t, t^{-1}] \longrightarrow R[t, t^{-1}], \quad t \longmapsto t^{-1}. \tag{4.12}$$

The n-th power maps are given as $[n] = \operatorname{Spec}([n]^*)$ with

$$[n]^* : R[t, t^{-1}] \longrightarrow R[t, t^{-1}], \quad t \longmapsto t^n. \tag{4.13}$$

Note that (4.12) and (4.13) are compatible in the sense that $i = [-1]$, which is always the case for a commutative group scheme. The next proposition, on the other hand, is very specific to $\mathbb{G}_m$.

Proposition 4.7. *Let S be a connected scheme. Then* $\operatorname{End}(\mathbb{G}_{m,S}) = \mathbb{Z}$.

Proof. We only consider the case $S = \operatorname{Spec}(k)$. The extension to general S can be found in [10, Proposition 2.12].

By definition, a group scheme endomorphism f of $\mathbb{G}_{m,k}$ is the same as $f = \operatorname{Spec}(f^*)$ for a unique k-algebra morphism $f^* : k[t, t^{-1}] \longrightarrow k[t, t^{-1}]$ such that

$$(f^* \otimes f^*) \circ m^* = m^* \circ f^* \tag{4.14}$$

where $m^*(t) = t \otimes t$ is as in (4.10). Giving a k-algebra morphism f^* is equivalent to specifying its image $f^*(t) \in k[t, t^{-1}]^\times$. These units are

$$k[t, t^{-1}]^\times = \{\lambda t^n \mid \lambda \in k^\times, n \in \mathbb{Z}\}.$$

If $f^*(t) = \lambda t^n$, then (4.14) evaluated at t becomes

$$\lambda t^n \otimes \lambda t^n \overset{?}{=} \lambda (t \otimes t)^n \tag{4.15}$$

which holds if and only if $\lambda^2 = \lambda$, meaning $\lambda = 1$. Note that $f^*(t) = t^n$ precisely defines the multiplication-by-n morphism $[n]$ (meaning taking n-th power in this context) and thus $\operatorname{End}(\mathbb{G}_{m,k}) = \mathbb{Z}$ is proved. □

We next determine the kernel $\mu_{n,S} := \ker([n])$. By definition, see (4.7), we need to compute the fiber product

$$\begin{tikzcd} \mu_{n,S} \arrow[r] \arrow[d] & S \arrow[d] \\ \mathbb{G}_{m,S} \arrow[r, "{[n]}"] & \mathbb{G}_{m,S}. \end{tikzcd}$$

Fiber products of affine schemes are computed by tensor products of rings, so we get

$$\begin{aligned} \mu_{n,S} &= \operatorname{Spec}\left(R \otimes_{1 \leftarrow t,\ R[t,t^{-1}],\ t \mapsto t^n} R[t,t^{-1}]\right) \\ &= \operatorname{Spec}\left(R[t]/(t^n - 1)\right). \end{aligned} \tag{4.16}$$

In terms of (4.8) and (4.11), we see

$$\mu_{n,S}(T) = \{\zeta \in \mathcal{O}_T(T)^\times \mid \zeta^n = 1\}. \tag{4.17}$$

That is, $\mu_{n,S}$ is the *group scheme of n-th roots of unity.* Let us assume that $S = \operatorname{Spec}(k)$. We observe the following interesting phenomenon:

Assume that n is prime to $\operatorname{char}(k)$. Then $t^n - 1 \in k[t]$ is a separable polynomial. Hence, $\mu_{n,k} = \operatorname{Spec} k[t]/(t^n - 1)$ is an étale k-scheme. On the other hand, if $p = \operatorname{char}(k) \mid n$, then $t^n - 1$ is not separable and $k[t]/(t^n - 1)$ is not reduced. For example,

$$\begin{aligned} \mu_{p,k} &= \operatorname{Spec} k[t]/(t^p - 1) \\ &= \operatorname{Spec} k[t]/(t-1)^p \\ &\xrightarrow{\sim} \operatorname{Spec} k[\varepsilon]/(\varepsilon^p) \end{aligned}$$

is completely infinitesimal. We have the following general results in this direction.

Theorem 4.8 (Cartier, [14, Corollary 8.38])**.** *Let k be a field of characteristic* 0 *and let G/k be a finite type group scheme. Then G is smooth.*

A morphism $f : X \to S$ is said to be *finite locally free of rank n* if it is finite and if $f_*(\mathcal{O}_X)$ is locally free of rank n as $\mathcal{O}_S$-module.

Theorem 4.9. *Let G be a commutative S-group scheme which is finite locally free of rank n. Assume that $n \in \mathcal{O}_S(S)^\times$. Then G is étale.*

Exercise 4.10. Verify the commutativity of (4.3) and (4.4) for $(\mathbb{G}_m, m, e, i)$.

4.3. **A non-commutative example: GL_n.** Let $S = \operatorname{Spec}(R)$ be affine as before. The (underlying scheme of the) general linear group in n variables over S is defined as

$$\mathrm{GL}_{n,S} = \operatorname{Spec} R\left[t_{ij},\ 1 \le i,j \le n;\ \det((t_{ij})_{i,j})^{-1}\right].$$

For every S-scheme T, we can (exercise) identify $\mathrm{GL}_{n,S}(T)$ with $\mathrm{GL}_n(\mathcal{O}_T(T))$ by

$$\Phi : [g : T \to \mathrm{GL}_{n,S}] \longmapsto (g^*(t_{ij}))_{i,j}. \tag{4.18}$$

We have the usual matrix multiplication on $\mathrm{GL}_n(\mathcal{O}_T(T))$. In terms of (4.18), it comes from the multiplication morphism $m : \mathrm{GL}_{n,S} \times_S \mathrm{GL}_{n,S} \to \mathrm{GL}_{n,S}$ which is given in coordinates by

$$m^*(t_{ij}) = \sum_{k=1}^n t_{ik} \otimes t_{kj}.$$

The pair $(\mathrm{GL}_{n,S}, m)$ is then an S-group scheme. The identity map $e = \operatorname{Spec}(e^*)$ is given by

$$e^* : R[t_{ij}, \det((t_{ij})_{ij})^{-1}] \longrightarrow R, \quad e^*(t_{ij}) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

The inverse of the matrix $(t_{ij})_{ij} \in \mathrm{GL}_n(R[t_{ij}, \det((t_{ij})_{ij})^{-1}])$ has an expression of the form $\det((t_{ij})_{ij})^{-1} \cdot (s_{ij})_{ij}$ where the s_{ij} are polynomials in the variables t_{ij}. (In fact, the s_{ij} are the entries of the adjugate matrix.) Then the inverse morphism $i : \mathrm{GL}_{n,S} \to \mathrm{GL}_{n,S}$ is given in coordinates by

$$i^*(t_{k\ell}) = \det((t_{ij})_{ij})^{-1} s_{k\ell}.$$

Clearly, $\mathrm{GL}_{1,S}$ is the same as the multiplicative group $\mathbb{G}_{m,S}$. For every S-scheme T, we have a determinant morphism $\mathrm{GL}_n(\mathcal{O}_T(T)) \to \mathcal{O}_T(T)^\times$. With respect to our identifications (4.11) and (4.18), these come from the group scheme homomorphism

$$\det : \mathrm{GL}_{n,S} \longrightarrow \mathbb{G}_{m,S}, \quad \det^*(t) = \det((t_{ij})_{ij}). \tag{4.19}$$

Its kernel $\ker(\det)$ is the group subscheme $\mathrm{SL}_{n,S} \subset \mathrm{GL}_{n,S}$. Being a closed subscheme of an affine scheme, it is again affine. It can be described explicitly by

$$\mathrm{SL}_{n,S} = \mathrm{Spec}\left(R[t_{ij},\, 1 \le i, j \le n]/(\det((t_{ij})_{ij}) - 1)\right).$$

4.4. **Linear algebraic groups.** We now specialize to the case of finite type group schemes over a field k. A general classification theorem essentially reduces their study to the affine and the proper case.

Theorem 4.11 (see [14, §8a]). *Let G/k be a connected finite type k-group scheme. Then there exists a unique maximal normal, connected, affine closed group sub-scheme $N \subseteq G$. The quotient G/N is an abelian variety.*

Affine finite type k-group schemes are also called *linear algebraic groups*. The reason for this name is that they can always be realized as a group of linear automorphisms of some vector space. That is, they always embed into some GL_N.

Theorem 4.12 (see [14, Corollary 4.10]). *Let G be an affine finite type k-group scheme. Then there exist an integer n and a closed immersion group scheme morphism $G \to \mathrm{GL}_N$.*

4.5. **Abelian varieties.** We have already defined abelian varieties in Definition 4.2. The main point of this definition is that abelian varieties are proper. This implies that they are necessarily commutative which also explains their name.

Theorem 4.13 ([10, Corollary 3.7]). *Let (A, m) be an abelian variety over k. Then (A, m) is a commutative group scheme.*

5. Elliptic curves

In the previous section, we defined elliptic curves as proper, smooth, 1-dimensional, connected group schemes and stated that they are always commutative (Definition 4.1 and Theorem 4.13). However, this definition does not shed any light on how to actually write down an example of an elliptic curve. For this reason, we want to next learn about two equivalent definitions:

- An elliptic curve over a field k is a pair (E, e) consisting of a proper smooth connected curve E/k of genus 1 and a rational point $e \in E(k)$.
- An elliptic curve over a field k is a smooth cubic curve $E \subset \mathbb{P}^2_k$ that contains the point $[0 : 1 : 0]$.

Passing between these definitions involves the theory of curves and line bundles. A careful discussion with many details can be found in [10, §4 – §7], but some of these details are tangential for our course. So we will give a shorter and more high-level treatment.

5.1. **Cubic curves are elliptic curves.** Our first aim is to construct elliptic curves. Let $h(x) = x^3 + ax + b$ be a monic cubic polynomial (without x^2-term). A polynomial of the form

$$f = y^2 - h(x) \tag{5.1}$$

is called a *simplified Weierstrass equation.* Let

$$F(X, Y, Z) = Y^2 Z - X^3 - aXZ^2 - bZ^3 \tag{5.2}$$

be the homogenization of f, and let $E = V_+(F) \subset \mathbb{P}^2_k$ be its vanishing locus.

Lemma 5.1. *Assume that* $\mathrm{char}(k) \neq 2$ *and that h is separable. Then E is a smooth curve.*

Proof. First observe by direct substitution in (5.2) that, on the level of sets, $E \cap V_+(Z) = \{[0:1:0]\}$. We can thus proceed by checking the Jacobi criterion on $E \cap D_+(Z)$ and for the point $[0:1:0]$.

By definition, we have

$$E \cap D_+(Z) \xrightarrow{\sim} V(y^2 - h(x)) \subset \mathbb{A}^2_k.$$

The Jacobi matrix of the Weierstrass polynomial is the gradient

$$(\partial f/\partial x,\ \partial f/\partial y) = (-h'(x),\ 2y). \tag{5.3}$$

Let $e \in E \cap D_+(Z)$ be an arbitrary point. Let $\kappa(e)$ be the residue field of e and let $(e_1, e_2) \in \kappa(e) \times \kappa(e)$ be the coordinates of e.[5] If $e_2 \neq 0$, then also $2e_2 \neq 0$ by our assumption $\mathrm{char}(k) \neq 2$, meaning $2y$ does not vanish in e. If $e_2 = 0$, however, then $h(e_1) = 0$ since $f(e_1, e_2) = 0$. We have assumed that h is separable, which is equivalent to $h(x)$ and $h'(x)$ being coprime. Thus $h'(e_1) \neq 0$. In summary, we have seen that the gradient (5.3) does not vanish in e.

We now consider the point $[0:1:0]$. An affine chart is given by

$$E \cap D_+(Y) \xrightarrow{\sim} V(z - x^3 - axz^2 - bz^3) \subset \mathbb{A}^2_k.$$

In these coordinates, $[0:1:0]$ maps to $(0,0)$. Moreover, the gradient of that equation is

$$(-3x^2 - az^2,\ 1 - 2axz - bz^2). \tag{5.4}$$

Its second entry does not vanish in $(0,0)$, so the Jacobi criterion holds in $(0,0)$. The proof of the lemma is now complete. $\square$

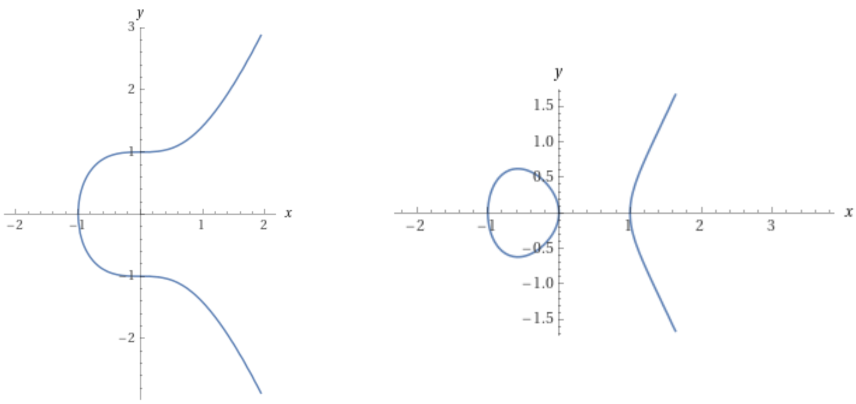

FIGURE 2. The $\mathbb{R}$-points of the two Weierstrass equations $y^2 = x^3 + 1$ and $y^2 = x^3 - x$. Note that $V(y^2 - (x^3 - x)) \subset \mathbb{A}^2_{\mathbb{R}}$ is a connected scheme. Only its $\mathbb{R}$-points when endowed with the real topology are disconnected.

[5]Given a scheme X and a point $x \in X$, we use $\kappa(x) = \mathrm{Quot}(\mathcal{O}_{X,x}/\mathfrak{m}_x)$ to denote the residue field in x.

Theorem 5.2. *Let $E = V_+(F) \subset \mathbb{P}^2_k$ be a smooth cubic curve, and let $O \in E(k)$ be a rational point. Then there exists a unique group scheme structure $+ : E \times_{\mathrm{Spec}(k)} E \to E$ on E with neutral element O. By Theorem 4.13, it is necessarily commutative.*

There are two approaches to this theorem. Today, we will explain the more elementary one, which is to give a geometric construction of $+$ in terms of the geometry of $\mathbb{P}^2$. A beautiful aspect of this construction is that it illustrates why *cubic* curves behave so special. Details on some calculations behind this approach may be found in Silverman's book [17, §III.1-3].

The second approach is based on line bundles, the Riemann–Roch Theorem, and the Yoneda Lemma. It is more conceptual, and some of its aspects will be discussed in more detail later in the course. A reference is [10, §7].

Proof of Theorem 5.2. We will admit the uniqueness part of the theorem, which is a general property of abelian varieties [10, Proposition 3.6]. Thus, the main problem is to construct the addition law.

Lemma 5.3. *Let $F \in k[X, Y, Z]$ be homogeneous of degree 3 without linear factor and let $E = V_+(F)$. Let $L \subset \mathbb{P}^2_k$ be any line. Then E intersects L in three points when counted with multiplicities. More precisely, $E \cap L = \mathrm{Spec}\, A$ for a k-algebra A with $\dim_k(A) = 3$.*

Here, by line we mean a curve of the form $V_+(aX + bY + cZ)$, where $(a, b, c) \neq (0, 0, 0)$.

Proof. After a linear change of coordinates, we may assume that $L = V_+(Z)$. Since F has no linear factor, $Z \nmid F$. Thus $F|_L = F(X, Y, 0)$ is a non-zero homogeneous polynomial of degree 3 and hence has three zeroes (counted with multiplicities) as claimed. □

Construction 5.4. Given $P, Q \in E(k)$, define a line $L \subset \mathbb{P}^2_k$ as follows:

(1) If $P \neq Q$, then let L be the unique line that passes through P and Q.
(2) If $P = Q$, then let L be the tangent line to E in that point.

The definition of the tangent uses the smoothness of E. (In a local chart, take the line perpendicular to the gradient of the equation defining E.) The smoothness of E also implies that F has no linear factor. Hence Lemma 5.3 applies and shows that E and L intersect in three points (counting multiplicities). But two of these points are known to be P and Q which lie in $L(k)$! And if a cubic polynomial has two rational roots, then the third root is rational as well. Thus there exists a unique third rational intersection point $R \in (E \cap L)(k)$. Repeating this construction with O, R instead of P, Q, defines a fourth point $S \in E(k)$.

Definition 5.5. The sum of $P, Q \in E(k)$ is defined as $P + Q := S$.

It is true, but not obvious, that this defines a group structure on $E(k)$. The easy part is to show that O is a neutral element and that every element has an inverse (exercise). It is moreover clear that the operation $(P, Q) \mapsto P + Q$ is commutative. Showing associativity is more tricky, however.

So far, we have defined a commutative group $E(k)$. If K/k is a field extension, then we can apply the above construction to $K \otimes_k E \subset \mathbb{P}^2_K$ and obtain a group structure on $E(K) = (K \otimes_k E)(K)$. We know from algebraic geometry that, given reduced varieties (smooth, for example) X and Y over an algebraically closed field K, a morphism $f : X \to Y$ is uniquely determined by the map $f(K) : X(K) \to Y(K)$ on K-points. So there is at most one morphism $E \times_{\mathrm{Spec}(k)} E \to E$ that induces the above group structures on all the $E(K)$, K/k. Moreover, if it exists, it will satisfy all group axioms because the sets $E(K)$ do (apply the uniqueness to the diagrams (4.3) and (4.4)).

To complete the proof, one carries out Construction 5.4 in indeterminates and sees that it indeed comes from a morphism of varieties. We refer the curious reader to [17, Theorem 3.6]. □

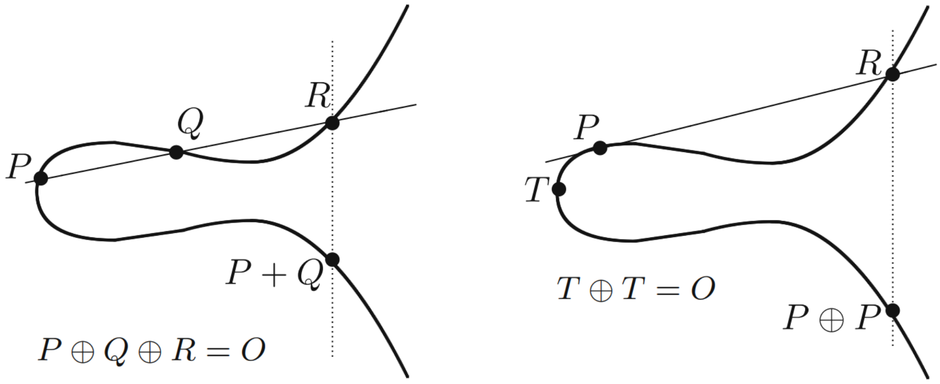

FIGURE 3. The case $P \neq Q$ is shown on the left, the tangent construction when $P = Q$ on the right. The point O here is the point $[0 : 1 : 0]$ at infinity. The vertical dotted lines are the lines through O and R. The picture is taken from [17, §III].

The simplified Weierstrass equations from Lemma 5.1 give simple examples of smooth cubic curves. We will later see that if $\mathrm{char}(k) \neq 2, 3$, then every elliptic curve can be described by $(V_+(F), [0 : 1 : 0])$ for a simplified Weierstrass equation F. In particular, the isomorphism classes of elliptic curves over k can be parametrized by the two coefficients $a, b \in k^2$ of $h(x) = x^3 + ax + b$. (Only those a and b such that h is separable occur, of course.)

5.2. **Elliptic curves have genus 1.** Our next goal is to show that all elliptic curves come from plane cubic curves. For this, we first need to find a way to extract geometric properties of E from the existence of the group structure. This is done using differential forms. Let us begin by recalling their definition.

Definition 5.6. Let R be a ring, A an R-algebra, and M an A-module. An *R-derivation from A to M* is an R-linear map $\delta : A \to M$ such that the Leibniz rule holds: For all $a, b \in A$,

$$\delta(ab) = a\delta(b) + b\delta(a).$$

Lemma 5.7. *There exists a universal R-derivation. That is, there exists an A-module $\Omega^1_{A/R}$ together with an R-derivation $d : A \to \Omega^1_{A/R}$ such that every R-derivation $\delta : A \to M$ factors through a unique A-module homomorphism $\varphi : \Omega^1_{A/R} \to M$. As diagram,*

$$\begin{tikzcd} A \arrow[r, "d"] \arrow[dr, "\forall\ \delta"'] & \Omega^1_{A/R} \arrow[d, "\exists!\ \varphi"] \\ & M. \end{tikzcd} \tag{5.5}$$

The pair $(\Omega^1_{A/R}, d)$ is called the module of Kähler differentials of A over R. It is easy to describe in terms of generators and relations. Let

$$A = R[X_1, \ldots, X_n]/(f_1, \ldots, f_m)$$

be a presentation of A as a quotient of a polynomial ring over R. Consider the free module $\bigoplus_{i=1}^n A\, dX_i$ generated by symbols $dX_1, \ldots, dX_n$. (This is really A^n; the symbols dX_i are just the traditional notation for the standard basis here.) For each $f \in R[X_1, \ldots, X_n]$, we can take the gradient vector

$$df := \frac{\partial f}{\partial X_1} \cdot dX_1 + \ldots + \frac{\partial f}{\partial X_n} \cdot dX_n. \tag{5.6}$$

Then

$$\begin{aligned} \big(A\, dX_1 \oplus \ldots \oplus A\, dX_n\big)/\langle df_1, \ldots, df_m \rangle &\xrightarrow{\sim} \Omega^1_{A/R} \\ dX_i &\xrightarrow{\sim} d(X_i). \end{aligned} \tag{5.7}$$

The key ideas for proving Lemma 5.7 and (5.7) are as follows:

- Since $d : A \to \Omega^1_{A/R}$ is supposed to be universal, the module $\Omega^1_{A/R}$ has to be generated by all derivatives $d(a)$ as A-module.
- Since every element of A is a polynomial in the X_i with R-coefficients, the Leibniz rule allows to write every $d(a)$ as an A-linear combination of the $d(X_i)$. Hence, the $d(X_i)$ already generate $\Omega^1_{A/R}$ as A-module.
- Since the $f_j \in A$ are zero, also the $d(f_j)$ in $\Omega^1_{A/R}$ have to be zero. By the Leibniz rule,

  $$d(f_j) = (\partial f/\partial X_1) \cdot d(X_1) + \ldots (\partial f/\partial X_n) \cdot d(X_n),$$

  which explains the relations $df_1, \ldots, df_m$ in (5.7).

Given an element $g \in A$, there is an isomorphism of $A[g^{-1}]$-modules

$$\Omega^1_{A/R}[g^{-1}] \xrightarrow{\sim} \Omega^1_{A[g^{-1}]/R} \tag{5.8}$$

which is uniquely characterized by sending $d(a)$ to $d(a)$. In other words, the formation of $\Omega^1_{A/R}$ is compatible with localizations. This means that the construction can be glued from rings to schemes.

Definition 5.8. Let $\pi : X \to S$ be a morphism of schemes. The quasi-coherent module with derivation $d : \mathcal{O}_X \to \Omega^1_{X/S}$ is defined as the unique datum (up to isomorphism) that is, locally on affine charts $\mathrm{Spec}(R) \subseteq S$ and $\mathrm{Spec}(A) \subseteq \pi^{-1}(\mathrm{Spec}(R))$, given by $d : A \to \Omega^1_{A/R}$ glued along (5.8).

Kähle differentials are closely related to smoothness, and we next state one form of this relation.

Theorem 5.9 ([10, Theorem 4.18])**.** *Let $\pi : X \to S$ be a morphism that is locally of finite presentation with purely d-dimensional fibers. Then π is smooth if and only if $\Omega^1_{X/S}$ is locally free of rank d as $\mathcal{O}_X$-module.*

Definition 5.10 (Genus of a curve)**.** (1) By curve over a field k, we mean a proper, smooth, geometrically connected and 1-dimensional k-scheme.

(2) Let $C \to \mathrm{Spec}(k)$ be a curve. By Theorem 5.9, $\Omega^1_{C/k}$ is a line bundle on C. Being a coherent sheaf on a proper variety, the space of global sections $\Omega^1_{C/k}(C)$ is a finite-dimensional k-vector space. Its dimension is called the genus of C.

Here, recall that a finite type k-scheme X is said to be geometrically reduced, connected, integral, etc. if the base change $\bar{k} \otimes_k X$ is reduced, connected, integral, etc. An equivalent condition is that for all field extensions K/k, the base change $K \otimes_k X$ has the relevant property.

For example, elliptic curves are geometrically connected because they are connected over k (by definition) and have a rational point (the neutral element).

Theorem 5.11. *Let E be an elliptic curve over a field k. Then E has genus 1.*

Sketch of proof. The key point is that the sheaf of differential forms of a group scheme is generated by invariant forms. The proof of this (see [10, Proposition 5.7]) does not concern us here, we will only state and use the result.

Let $\pi : G \to S$ be a group scheme with neutral element section $e : S \to G$. Recall that quasi-coherent modules can be pulled back under scheme morphisms. So we may first form $e^*(\Omega^1_{G/S})$, a quasi-coherent S-module. Then we may again pull back along π. The statement is that there exists an isomorphism

$$\gamma : \pi^* e^* \Omega^1_{G/S} \xrightarrow{\sim} \Omega^1_{G/S}. \tag{5.9}$$

We now apply (5.9) to our elliptic curve $E \to \mathrm{Spec}(k)$. The pullback $V = e^*(\Omega^1_{E/k})$ is a one-dimensional k-vector space because $\Omega^1_{E/k}$ is a line bundle. Then (5.9) states that

$$\gamma : \mathcal{O}_E \otimes_k V \xrightarrow{\sim} \Omega^1_{E/k}.$$

Choosing a basis vector $\omega \in V$, we have thus obtained an isomorphism $\mathcal{O}_E \xrightarrow{\sim} \Omega^1_{E/k}$. The genus of E is hence $\dim_k \mathcal{O}_E(E)$.

Lemma 5.12. *Let $X \to \mathrm{Spec}(k)$ be a proper k-scheme that is geometrically reduced and geometrically connected. Then $\dim_k \mathcal{O}_X(X) = 1$.*

Proof. The global sections $A = \mathcal{O}_X(X)$ are a finite-dimensional k-algebra. Its formation commutes with base change in the sense that for every field extension K/k, we have

$$K \otimes_k A = \mathcal{O}_{K\otimes_k X}(K \otimes_k X).$$

Hence, if X is geometrically reduced and connected, then $\bar{k} \otimes_k A$ is reduced and has a unique maximal ideal. The residue field is necessarily $\bar{k}$ because $\bar{k}$ is algebraically closed. So $\bar{k} \xrightarrow{\sim} \bar{k} \otimes_k A$. Thus A was one-dimensional to begin with, meaning $k \xrightarrow{\sim} A$. □

Coming back to our elliptic curve $E \to \mathrm{Spec}(k)$, we see that $\mathcal{O}_E(E) = k$, meaning that E has genus 1 as claimed. □

Remark 5.13. The isomorphism in (5.9) is given by extending the value of a differential form on $e(S)$ in the unique way to a left-translation invariant differential form on G. This concept is also commonly used in differential geometry, where one often identifies the Lie algebra $\mathfrak{g}$ of a Lie group G with the space of translation invariant vector fields on G.

For example, the form dt on $\mathbb{R}$ is translation invariant with respect to addition because $d(t + \lambda) = dt$ for all $\lambda \in \mathbb{R}$. The form $t^{-1}dt$ on $\mathbb{R}^\times$ is translation invariant with respect to multiplication because $(t\lambda)^{-1}d(\lambda t) = t^{-1}dt$ for all $\lambda \in \mathbb{R}^\times$.

5.3. **Genus 1 curves as cubics.** We have just shown that every elliptic curve has genus 1. In order to complete the circle of equivalent definitions (a triangle, actually), it is left to realize curves of genus 1 as cubic curves in $\mathbb{P}^2$. Let us first briefly recall a bit of general formalism.

Construction 5.14. Let X be a k-scheme. Giving a morphism $f : X \to \mathbb{P}^n_k$ is the same as giving a line bundle $\mathcal{L}$ on X and a surjection of $\mathcal{O}_X$-modules

$$\ell : \mathcal{O}_X^{\oplus(n+1)} \twoheadrightarrow \mathcal{L}.$$

Namely, on $\mathbb{P}^n_k$, we have the standard line bundle $\mathcal{O}(1)$. It is generated by the $n+1$ global sections $X_0, \dots, X_n$ corresponding to the $n + 1$ coordinates on $\mathbb{P}^n_k$. That is, we have a surjection

$$\mathcal{O}_{\mathbb{P}^n_k}^{\oplus(n+1)} \twoheadrightarrow \mathcal{O}(1), \quad e_i \longmapsto X_i.$$

Given $f : X \to \mathbb{P}^n_k$, we can pull back that surjection and obtain a pair $\mathcal{L} = f^*\mathcal{O}(1)$, $\ell : \mathcal{O}_X^{\oplus(n+1)} \twoheadrightarrow \mathcal{L}$ as desired.

Conversely, assume that $(\mathcal{L}, \ell)$ is given. Let $s_i = \ell(e_i) \in \mathcal{L}(X)$ be the $n+1$ global sections defined by ℓ. Let $U_i = D(s_i) \subseteq X$ be the open subscheme where s_i is a generator. That is, if we locally trivialize $\mathcal{L}$, say

$$\mathcal{O}_U \cdot s \xrightarrow{\sim} \mathcal{L}|_U, \quad s_i = f_i s, \quad f_i \in \mathcal{O}_U(U),$$

then $U_i \cap U = D(f_i)$ is the locus where f_i is invertible.

Over the open subset U_i, every section of $\mathcal{L}$ is a unique multiple of s_i. So we have defined functions $s_j/s_i \in \mathcal{O}_X(U_i)$ by the identity $s_j = (s_j/s_i) \cdot s_i$. This defines a morphism

$$f_i = \left(\frac{s_0}{s_i}, \ldots, \frac{\widehat{s_i}}{s_i}, \ldots, \frac{s_n}{s_i}\right) : U_i \longrightarrow \mathbb{A}^n.$$

On overlaps $U_i \cap U_j$, we have the (obvious) relation

$$\frac{s_k}{s_i} = \frac{s_j}{s_i} \cdot \frac{s_k}{s_j}.$$

If we spell out how $\mathbb{P}^n_k$ is glued from $n+1$ copies of $\mathbb{A}^n_k$ by the exact same rule of coordinate transformation, then this implies that the f_i glue to a morphism

$$f : X \longrightarrow \mathbb{P}^n_k.$$

A good notation for this morphism is $[s_0 : s_1 : \ldots : s_n]$. Namely, if $x \in X$ is a point then we may view $[s_0(x) : \ldots : s_n(x)] \in \mathbb{P}^n(\kappa(x))$ as follows. Let $s \in \mathcal{L}_x$ be a generator as $\mathcal{O}_{X,x}$-module. Then we may write $s_{i,x} = h_i s$ for unique functions $h_i \in \mathcal{O}_{X,x}$. The tuple $[h_0(x) : \ldots : h_n(x)]$ is a point of $\mathbb{P}^n(\kappa(x))$. Any other generator of $\mathcal{L}_x$ differs from s by a unit, hence the tuple $(h_0(x), \ldots, h_n(x))$ is unique up to $\kappa(x)^\times$, meaning that

$$[s_0(x) : \ldots : s_n(x)] := [h_0(x) : \ldots : h_n(x)]$$

is well-defined.

Exercise 5.15. Verify that the above two constructions $(\mathcal{L}, \ell) \longleftrightarrow (f : X \to \mathbb{P}^n_k)$ are inverse to each other.

Example 5.16. We know that every line bundle on $\mathbb{P}^1_k$ is isomorphic to one of the line bundles $\mathcal{O}(d)$. The integer $d \in \mathbb{Z}$ is its degree. We know that

$$\dim_k(\mathcal{O}(d)(\mathbb{P}^1_k)) = \begin{cases} d+1 & \text{if } d \geq 0 \\ 0 & \text{if } d < 0. \end{cases}$$

If $d \geq 0$, then a basis for the global sections $\mathcal{O}(d)(\mathbb{P}^1_k)$ is given by the monomials

$$X_0^d,\ X_0^{d-1}X_1,\ \ldots,\ X_0X_1^{d-1},\ X_1^d$$

where $X_0, X_1 \in \mathcal{O}(1)(\mathbb{P}^1_k)$ are the coordinates on $\mathbb{P}^1_k$. If $d \geq 0$, then these monomials also generate $\mathcal{O}(d)$ as line bundle. That is, the map

$$\mathcal{O}_{\mathbb{P}^1_k}^{\oplus(d+1)} \twoheadrightarrow \mathcal{O}(d), \quad e_i \longmapsto X_0^{d-i}X_1^i$$

is a surjection of quasi-coherent $\mathcal{O}_{\mathbb{P}^1_k}$-modules. The corresponding morphism $\mathbb{P}^1_k \to \mathbb{P}^d_k$ is called the Veronese map. It is a closed immersion when $d \geq 1$.

Construction 5.14 shows that, if we want to define a morphism $E \to \mathbb{P}^2_k$ from an elliptic curve to the projective plane, then we need to understand line bundles and their global sections on E. Let us begin with some general observations and definitions.

- Let X be a noetherian scheme and $\mathcal{F}$ a coherent $\mathcal{O}_X$-module. (This is the same as $\mathcal{F}$ being quasi-coherent and of finite type.) Then $\mathcal{F}$ is locally free (meaning a vector bundle) if and only if for every $x \in X$, the stalk $\mathcal{F}_x$ is a free $\mathcal{O}_{X,x}$-module.
- Thus, if C is a curve over a field k, then a coherent module $\mathcal{L}$ is a line bundle if and only if for every $x \in X$, the stalk $\mathcal{L}_x$ is free of rank 1 over $\mathcal{O}_{C,x}$.

- By definition, all our curves are smooth, hence normal. So for $x \in C$ closed, the local ring $\mathcal{O}_{C,x}$ is a discrete valuation ring (DVR). By the classification of modules over principal ideal domains (PIDs), a finite type $\mathcal{O}_{C,x}$-module is free if and only if it is torsion-free.

Conclusion 5.17. Let $0 \neq \mathcal{I} \subseteq \mathcal{O}_C$ be an ideal sheaf in $\mathcal{O}_C$. Then $\mathcal{I}$ is stalk-by-stalk torsion-free because it is a subsheaf of torsion-free sheaf $\mathcal{O}_C$, and hence $\mathcal{I}$ is a line bundle.

Definition 5.18 (Degree of a line bundle)**.** (1) Let $\mathcal{I} \subseteq \mathcal{O}_C$ be a non-zero ideal sheaf. Then $Z = V(\mathcal{I}) \subset C$ is a proper closed subscheme. It has to be 0-dimensional, and hence is a finite k-scheme. As such, it is affine, meaning $Z \cong \mathrm{Spec}(A)$ for a finite dimension k-algebra A. The *degree* of $\mathcal{I}$ is defined as $-\dim_k(A)$. More concretely, because each local ring $\mathcal{O}_{C,x}$ is a DVR, we can write

$$Z = \bigsqcup_{i=1}^{r} \mathrm{Spec}(\mathcal{O}_{C,x_i}/\mathfrak{m}_{x_i}^{e_i})$$

for uniquely determined pairwise different closed points $x_1, \dots, x_r \in C$ and exponents $e_1, \dots, e_r \geq 1$. Then

$$\deg(\mathcal{I}) = -\sum_{i=1}^{r} e_i \cdot [\kappa(x_i) : k].$$

(2) Let $\mathcal{L}$ be a line bundle on C. There always exist two ideal sheaves $\mathcal{I}_1, \mathcal{I}_2 \subseteq \mathcal{O}_C$ such that $\mathcal{L} \cong \mathcal{I}_1 \otimes \mathcal{I}_2^{-1}$. We define

$$\deg(\mathcal{L}) := \deg(\mathcal{I}_1) - \deg(\mathcal{I}_2).$$

This does not depend on the choices of $\mathcal{I}_1$ and $\mathcal{I}_2$. In particular, the degree defines a group homomorphism

$$\deg : \mathrm{Pic}(C) \longrightarrow \mathbb{Z}.$$

Motivation 5.19. The degree is a simple numerical invariant of a line bundle on a curve. The following results show that it is extremely helpful when studying global sections of line bundles and hence, by Construction 5.14, maps $C \to \mathbb{P}_k^n$.

Theorem 5.20 (Riemann–Roch)**.** *Let C be a curve of genus g over a field k. Then, for every line bundle $\mathcal{L}$ on C,*

$$\dim \mathcal{L}(C) = \deg(\mathcal{L}) + 1 - g + \dim(\Omega^1_{C/k} \otimes \mathcal{L}^{-1})(C). \tag{5.10}$$

Corollary 5.21. *The degree of $\Omega^1_{C/k}$ is $2g-2$.*

Proof. Apply the Riemann–Roch Theorem 5.20 to $\Omega^1_{C/k}$. We obtain

$$g = \deg(\Omega^1_{C/k}) + 1 - g + 1$$

which we may rearrange as claimed. □

Corollary 5.22. *Let C/k be a curve of genus 1 and let $\mathcal{L}$ be a line bundle of degree $\deg(\mathcal{L}) \geq 1$ on C. Then*

$$\dim \mathcal{L}(C) = \deg(\mathcal{L}).$$

Proof. By Corollary 5.21, $\deg(\Omega^1_{C/k}) = 0$. Since $\deg(\mathcal{L}) \geq 1$, we then have

$$\deg(\Omega^1_{C/k} \otimes \mathcal{L}^{-1}) = 0 - \deg(\mathcal{L}) < 0.$$

Line bundles of negative degree cannot have non-zero global sections, so $(\Omega^1_{C/k} \otimes \mathcal{L}^{-1})(C) = 0$. Evaluating the Riemann–Roch identity (5.10), we find $\dim \mathcal{L}(C) = \deg(\mathcal{L})$ as claimed. □

Theorem 5.23. *Let E be a curve of genus 1 over k such that $E(k) \neq \emptyset$. Then there exists a closed immersion $E \hookrightarrow \mathbb{P}^2_k$ which identifies E with the curve $V_+(F)$ defined by a cubic homogeneous polynomial.*

Proof. Step 1: Construction of a morphism $E \to \mathbb{P}^2_k$. We have seen in Construction 5.14 that, in order to define a morphism $E \to \mathbb{P}^2_k$, our task is to find a line bundle $\mathcal{L}$ on E together with a surjection $\ell : \mathcal{O}_E^3 \twoheadrightarrow \mathcal{L}$.

We now draw inspiration from the example of $\mathbb{P}^1_k$ above. By assumption, there exists a k-rational point $e \in E(k)$. View $\{e\}$ as a reduced closed subscheme of E, and let $\mathcal{I}_e$ be its ideal sheaf. According to Definition 5.18, its degree is -1. So the dual line bundle $\mathcal{M} := \mathcal{I}_e^{-1}$ has degree 1.

The degree of $\mathcal{M}^{\otimes d}$ is d.[6] By Corollary 5.22, this means

$$\dim \mathcal{M}^{\otimes d}(E) = d, \qquad d \geq 1.$$

We are mostly interested in $\mathcal{L} = \mathcal{M}^{\otimes 3}$. For every closed point $y \in E$ we have an ideal sheaf $\mathcal{I}_y$ as before. Its degree is $-[\kappa(y) : k]$, the negative of the residue field extension degree. On the one hand, we may consider $\mathcal{L}$ and $\mathcal{I}_y$ as abstract line bundles. By Riemann–Roch, the dimension of global sections strictly decreases when tensoring with $\mathcal{I}_y$ because the degree goes down:

$$\dim(\mathcal{L} \otimes \mathcal{I}_y)(E) < 3.$$

On the other hand, we can consider the concrete exact sequence

$$0 \longrightarrow \mathcal{I}_y \longrightarrow \mathcal{O}_E \longrightarrow i_*\kappa(y) \longrightarrow 0$$

where $i : \{y\} \to E$ is the inclusion map. Tensoring by $\mathcal{L}$, which is an exact operation because $\mathcal{L}$ is locally free, we get an exact sequence

$$0 \longrightarrow \mathcal{L} \otimes \mathcal{I}_y \longrightarrow \mathcal{L} \longrightarrow i_*\mathcal{L}(y) \longrightarrow 0.$$

Here, $\mathcal{L}(y) := i^*\mathcal{L}$ is our notation for the 1-dimensional $\kappa(y)$ vector space that forms the fiber of $\mathcal{L}$ in y. Taking global sections, we see that

$$(\mathcal{L} \otimes \mathcal{I}_y)(E) \subseteq \mathcal{L}(E)$$

are precisely those global sections that vanish in y.

We conclude that for every closed point $y \in E$, there exists a global section $s \in \mathcal{L}(E)$ that does not vanish in y. This means that $\mathcal{L}$ is generated by its global sections. That is, after choosing a basis s_0, s_1, s_2 for the three-dimensional vector space $\mathcal{L}(E)$, we obtain a surjection

$$\ell : \mathcal{O}_E^{\oplus 3} \longrightarrow\!\!\!\!\!\rightarrow \mathcal{L}, \quad e_i \longmapsto s_i,$$

and hence a morphism $f : E \to \mathbb{P}^2_k$ as in Construction 5.14.

Step 2: f is a closed immersion. We can prove that f is a closed immersion after base change to $\bar{k}$. So from now on, we assume that k is algebraically closed. This helps, because now every closed point $y \in E$ is k-rational and, in particular, $\deg(\mathcal{I}_y) = -1$. Let $y, y' \in E$ be two (possibly equal) closed points. Corollary 5.22 implies that

$$\dim(\mathcal{L} \otimes \mathcal{I}_y)(E) = 2$$
$$\dim(\mathcal{L} \otimes \mathcal{I}_y \otimes \mathcal{I}_{y'})(E) = 1.$$

So, after applying a linear change of coordinates on $\mathbb{P}^2_k$, we may assume that our basis $s_0, s_1, s_2 \in \mathcal{L}(E)$ is chosen with

$$\begin{aligned} s_0 &\in \mathcal{L}(E) \setminus (\mathcal{L} \otimes \mathcal{I}_y)(E), \\ s_1 &\in (\mathcal{L} \otimes \mathcal{I}_y)(E) \setminus (\mathcal{L} \otimes \mathcal{I}_y \otimes \mathcal{I}_{y'})(E). \end{aligned} \tag{5.11}$$

[6] All tensor products during the proof are taken over $\mathcal{O}_E$.

If $y \neq y'$, then this means that

$$[s_0(y) : s_1(y) : s_2(y)] \neq [s_0(y') : s_1(y') : s_2(y')]$$

because s_1 vanishes in y while it does not vanish in y'. We conclude that f is injective at the level of topological spaces. Since f is also closed by the properness of E, it is topologically a closed immersion.

Finally, if $y = y'$, then the above choice of s_1 ensures that it vanishes to first order in y, but not to second order. Translating this to local coordinates (omitted), it is possible to deduce that $[s_0 : s_1 : s_2]$ is injective on the tangent space $(\mathfrak{m}_y/\mathfrak{m}_y^2)^\vee$ in y, which means that f is even schematically a closed immersion near y.

Step 3: Its image is defined by a cubic equation. We do not assume anymore that k is algebraically closed. Recall that $e \in E(k)$ is our given rational point and that $\mathcal{M} = \mathcal{I}_e^{-1}$. Dualizing the descending chain

$$\ldots \subset \mathcal{I}_e^3 \subset \mathcal{I}_e^2 \subset \mathcal{I}_e \subset \mathcal{O}_E,$$

we obtain an ascending chain

$$\mathcal{O}_E \subset \mathcal{M} \subset \mathcal{M}^2 \subset \mathcal{M}^3 \subset \ldots.$$

Proceeding with the same logic as in (5.11), we choose elements

$$\begin{aligned} &1 \in \mathcal{O}_E(E) \\ &\qquad \mathcal{M}(E) = \mathcal{O}_E(E) \text{ by Cor. 5.22} \\ &x \in \mathcal{M}^{\otimes 2}(E) \backslash \mathcal{M}(E) \\ &y \in \mathcal{M}^{\otimes 3}(E) \backslash \mathcal{M}^{\otimes 2}(E). \end{aligned} \tag{5.12}$$

View $1, x, y$ as elements of $\mathcal{L}(E) = \mathcal{M}^{\otimes 3}(E)$. Then they form a basis because y generates $\mathcal{L}$ near e, while x vanishes to first order and 1 to third order in e. We consider the morphism

$$[x : y : 1] : E \longrightarrow \mathbb{P}^2_k.$$

Consider the sections

$$1, x, y, x^2, xy, y^2, x^3 \in \mathcal{M}^{\otimes 6}(E). \tag{5.13}$$

These are seven sections of a six-dimensional vector space (use again Corollary 5.22), and hence there exists a non-trivial linear relation

$$a_0 y^2 + b_0 x^3 + a_1 xy + a_2 x^2 + a_3 y + a_4 x + a_6 = 0. \tag{5.14}$$

Claim: Both a_0 and b_0 are non-zero. The section x is a generator of $\mathcal{M}^{\otimes 2}$ near e; the section y a generator of $\mathcal{M}^{\otimes 3}$ near e. Hence, y^2 and x^3 are both generators of $\mathcal{M}^{\otimes 6}$ near e. Thus either of the set of vectors

$$1, x, y, x^2, xy, y^2, \quad \text{or} \quad 1, x, y, x^2, xy, x^3 \tag{5.15}$$

has the property that the six sections vanish to orders precisely $6, 4, 3, 2, 1, 0$ in the stalk $(\mathcal{M}^{\otimes 6})_e$. Thus, either of the two sets forms a basis for $\mathcal{M}^{\otimes 6}(E)$. It follows that $a_0 b_0 \neq 0$ as claimed.

Conclusion: Identity (5.14) means that the morphism $[x : y : 1]$ factors through the cubic curve

$$V_+(F), \quad F = a_0 Y^2 Z + b_0 X^3 + a_1 XYZ + a_2 X^2 Z + a_3 YZ^2 + a_4 XZ^2 + a_6 Z^3.$$

The linear independence in (5.15) moreover shows that the morphism does not factor through a line or quadric in $\mathbb{P}^2_k$. It follows that F is irreducible and hence $E \xrightarrow{\sim} V_+(F)$ because we already know from Step 2 that $[x : y : 1]$ is a closed immersion. $\square$

Our proof even showed that the affine cubic equation for E may always be chosen in the form (5.14). We can simplify this expression further:

- Scaling y and x by a_0/b_0, we obtain a relation of the form
$$y^2 + (b_1x + b_3)y = x^3 + b_2x^2 + a_4x + a_6.$$
This kind of cubic equation is called a *general Weierstrass equation.*
- If $\mathrm{char}(k) \neq 2$, then we can change y to $y + (b_1x + b_3)/2$ to simplify further to a relation of the form
$$y^2 = x^3 + c_2x^2 + c_4x + c_6.$$
- If $\mathrm{char}(k) \neq 3$, then we may further replace x by $x + c_2/3$ and arrive at the simplified form
$$y^2 = x^3 + ax + b. \tag{5.16}$$

Ultimately, we conclude that every elliptic curve can be defined by a general Weierstrass equation. Outside of characteristics 2 and 3, we may even restrict to simplified Weierstrass equations.

Corollary 5.24. *Let $E \to \mathrm{Spec}(k)$ be a curve of genus 1 and let $e \in E(k)$ be a rational point. Then there exists a unique group scheme structure on E with identity element e.*

Proof. Apply Theorem 5.23 to realize E as a cubic in $\mathbb{P}^2_k$. Then use Theorem 5.2 to endow E with a group scheme structure (in a unique way). □

Remark 5.25. Let $F \in k[X, Y, Z]$ be homogeneous of degree d and such that $V_+(F) \subset \mathbb{P}^2_k$ is smooth. Then $V_+(F)$ is a curve of genus $(d-1)(d-2)/2$.

6. Arithmetic of elliptic curves

For every elliptic curve E, we have a multiplication-by-n homomorphism $[n] : E \to E$ which was defined in (4.6). Let $E[n] := \ker([n])$ be its kernel. Our next goal is to prove that $E[n]$ is always finite of degree n^2. This is not at all obvious as the following two (also 1-dimensional) examples show.

- The n-torsion $\mathbb{G}_m[n]$ of the multiplicative group is the group scheme μ_n of n-th roots of unity. It is finite of order n.
- Let k be a field and let $\mathbb{G}_{a,k} = \mathrm{Spec}\, k[t]$ be the additive group over k. Its group scheme structure a (the additional law) is defined by $a^*(t) = t \otimes 1 + 1 \otimes t$. For a k-scheme T, the T-valued points $\mathbb{G}_{a,k}(T)$ are the additive group $(\mathcal{O}_T(T), +)$. In particular,
$$\mathbb{G}_{a,k}[n] = \begin{cases} \{0\} & \text{if } \mathrm{char}(k) \nmid n \\ \mathbb{G}_{a,k} & \text{if } \mathrm{char}(k) \mid n. \end{cases}$$
For example, if $\mathrm{char}(k) = p$, then $[p]$ equals the 0-map $[0]$.

Our proof of $|E[n]| = n^2$ will be in two steps:

Step 1. First, we study elliptic curves over $\mathbb{C}$ where we can use their description by lattices to prove the statement over $\mathbb{C}$. By extension, the statement then even holds over all fields of characteristic 0.

Step 2. We extend the statement from $\mathbb{C}$ to all fields by using the universal Weierstrass family.

6.1. Analytification of complex varieties.

Recall from §3.2 that we defined a topology on $X(\mathbb{C})$ for every affine complex variety X. This construction can be upgraded to an analytification functor
$$\begin{aligned} \{\text{Smooth } \mathbb{C}\text{-schemes}\} &\longrightarrow \{\text{Smooth complex manifolds}\} \\ X &\longmapsto X(\mathbb{C}). \end{aligned} \tag{6.1}$$

First, if $X \subseteq \mathbb{A}^n_{\mathbb{C}}$ is a smooth affine variety embedded into affine space, then $X(\mathbb{C}) \subseteq \mathbb{C}^n$ has a unique structure as a smooth complex submanifold. Namely, the Jacobi criterion holds in the algebraic sense for X, and so also holds in the analytic sense for $X(\mathbb{C})$. Hence, $X(\mathbb{C})$ is a complex submanifold by the inverse function theorem.

The construction of this manifold structure is functorial: If $\varphi : X \to Y$ is a morphism of smooth affine $\mathbb{C}$-schemes and if $X \subseteq \mathbb{A}^n_{\mathbb{C}}$ and $Y \subseteq \mathbb{A}^m_{\mathbb{C}}$ are embeddings, then there exists an extension of φ to a morphism $\Phi : \mathbb{A}^n_{\mathbb{C}} \to \mathbb{A}^m_{\mathbb{C}}$. Passing to $\mathbb{C}$-points, we obtain a diagram

$$\begin{tikzcd} X(\mathbb{C}) \arrow[r, "\varphi(\mathbb{C})"] \arrow[d, hook] & Y(\mathbb{C}) \arrow[d, hook] \\ \mathbb{C}^n \arrow[r, "\Phi(\mathbb{C})"] & \mathbb{C}^m \end{tikzcd}$$

where $\Phi(\mathbb{C})$ is holomorphic because it is given by polynomials. It follows that $\varphi(\mathbb{C})$ is holomorphic. If φ is an isomorphism, then the same argument applies to φ^{-1} showing that $\varphi(\mathbb{C})$ is biholomorphic. This shows that the complex manifold sturcture on $X(\mathbb{C})$ does not depend on the chosen embedding $X \subseteq \mathbb{A}^n_{\mathbb{C}}$. Moreover, the functoriality allows to glue the construction from the affine to the general case.

Analytification has various nice properties of which we mention a few:

(1) If $X \subseteq \mathbb{P}^n_{\mathbb{C}}$ is a projective variety defined by the vanishing of homogeneous polynomials $F_1, \dots, F_r \in \mathbb{C}[T_0, \dots, T_n]$, then $X(\mathbb{C}) \subseteq \mathbb{P}^n(\mathbb{C})$ is the submanifold defined by the vanishing of the same polynomials.

(2) X is connected if and only if $X(\mathbb{C})$ is connected.

(3) X is proper if and only if $X(\mathbb{C})$ is compact.

(4) Analytification restricts to an equivalence

$$\{\text{Curves over } \mathbb{C}\} \xrightarrow{\sim} \{\text{Compact connected Riemann surfaces}\}. \tag{6.2}$$

This is a non-trivial theorem whose proof requires some functional analysis, see [4, §14]. For curves of genus 1, there is a much simpler proof using the Weierstrass $\wp$-function.

(5) Analytification is a faithful functor. It is fully faithful when restricted to proper smooth $\mathbb{C}$-schemes.

6.2. **Application to abelian varieties.** Let us now consider analytification in the context of abelian varieties. If $A/\mathbb{C}$ is an abelian variety, then $A(\mathbb{C})$ is a compact connected complex manifold (use (2) and (3) above). Moreover, we can analytify the multiplication morphism and obtain a holomorphic map $A(\mathbb{C}) \times A(\mathbb{C}) \to A(\mathbb{C})$. Analytification is functorial, so the group axiom diagrams from (4.3) and (4.4) are still commutative. (In fact, this is simply the statement that the set $A(\mathbb{C})$ is a group which we already knew before.) In this way, $A(\mathbb{C})$ is a compact connected complex Lie group.

We have moreover stated that analytification is fully faithful for proper smooth $\mathbb{C}$-schemes, see (5) above, so for any two abelian varieties A_1, A_2 over $\mathbb{C}$,

$$\mathrm{Hom}_{\mathbb{C}\text{-group scheme}}(A_1, A_2) = \mathrm{Hom}_{\text{complex Lie group}}(A_1(\mathbb{C}), A_2(\mathbb{C})).$$

Theorem 6.1. *Let X be a compact connected complex Lie group of dimension g. Then there exists a lattice $\Lambda \subset \mathbb{C}^g$ and an isomorphism $\mathbb{C}^g/\Lambda \xrightarrow{\sim} X$.*

Proof following [15, p. 1–2]. Consider the action of X on itself by conjugation. It preserves the identity $e \in X$ and hence defines an action of X on the tangent space $V = T_eX$. One can check from the definition of complex Lie group that this defines a holomorphic homomorphism $\mathrm{ad} : X \to GL_{\mathbb{C}}(V)$. By the maximum principle, a holomorphic function on a compact complex manifold is compact. Applying this to each of the coordinates of ad proves that this map is trivial, meaning that X is commutative.

Next, consider the exponential map $\exp : T_eX \to X$. Recall that this map is defined for every complex (or real) Lie group and that it satisfies $\exp(v+w) = \exp(v)\exp(w)$ for all v, w with $[v, w] = 0$. We have already seen that X is commutative, so $[v, w]$ is always 0. It follows that exp is a group homomorphism.

The exponential map is locally biholomorphic. The image of exp hence contains an open neighborhood of e. Any such neighborhood generates X as group because X is connected, so exp is surjective. As exp is biholomorphic near the identity, we find that $X = V/\Lambda$ for a discrete subgroup $\Lambda \subset V$. Any discrete subgroup of a finite-dimensional real vector space with compact quotient is a lattice, which completes the proof. □

Complex Lie groups of the form $X = \mathbb{C}^g/\Lambda$ are called complex tori. (Here, $\Lambda \subset V$ is a lattice.) They always satisfy $X \cong (\mathbb{R}/\mathbb{Z})^{2g}$ as real Lie group, where $g = \dim_{\mathbb{C}}(V)$, but the complex structure is an additional piece of information.

Corollary 6.2. *There is an equivalence of categories*

$$\{\textit{Ell. curves}/\mathbb{C}\} \xrightarrow{\sim} \left\{ \begin{array}{c} \textit{Compact complex Lie groups} \\ \textit{of the form } \mathbb{C}/\Lambda \end{array} \right\}. \tag{6.3}$$

Proof. We stated above that analytification of proper smooth $\mathbb{C}$-varieties is a fully faithful functor. So elliptic curves over $\mathbb{C}$ embed fully faithfully into 1-dimensional compact complex Lie groups. These are all of the form $\mathbb{C}/\Lambda$ by Theorem 6.1. The fullness is (6.2). □

Corollary 6.3. *Let A be a g-dimensional abelian variety over a field k of characteristic 0. Then $A[n]$ is a finite k-group scheme of degree n^{2g}.*

Proof. For simplicity, assume that k can be embedded into $\mathbb{C}$ and fix such an embedding. By definition of the kernel, we have $(\mathbb{C} \otimes_k A)[n] \xrightarrow{\sim} \mathbb{C} \otimes_k A[n]$, so $A[n]$ is finite of degree n^{2g} if and only if $(\mathbb{C} \otimes_k A)[n]$ is finite of such degree. We can hence assume from now on that $k = \mathbb{C}$.

Since $\mathbb{C}$ is algebraically closed, $A[n]$ is finite if and only if the $\mathbb{C}$-points $A[n](\mathbb{C})$ are finite. Moreover, once we know this finiteness, Theorem 4.8 ensures that $A[n]$ is étale. Again since $\mathbb{C}$ is algebraically closed, this is equivalent to $A[n]$ being a disjoint union of copies of $\mathrm{Spec}(\mathbb{C})$. Thus, our proof is complete if we can show that $|A[n](\mathbb{C})| = n^{2g}$.

Recall from (4.8) that the kernel satisfies $A[n](\mathbb{C}) = A(\mathbb{C})[n]$. By Theorem 6.1, $A(\mathbb{C})[n] \xrightarrow{\sim} (n^{-1}\mathbb{Z}/\mathbb{Z})^{\oplus 2g}$, which has order n^{2g} as claimed. □

6.3. **The universal Weierstrass family.** Let S be a scheme. There is a natural definition of elliptic curve over S which extends the case $S = \mathrm{Spec}(k)$. Sometimes, this is also called a *relative elliptic curve*, or a *family of elliptic curves parametrized by S*.

Definition 6.4. An elliptic curve over S is an S-group schemes $E \to S$ which is proper and smooth of relative dimension 1 with connected fibers.

Clearly, if $T \to S$ is a morphism and $E \to S$ an elliptic curve, then the base change $E_T := T \times_S E$ is an elliptic curve over T. In particular, for every point $s \in S$, the fiber $E_s := \mathrm{Spec}(\kappa(s)) \times_S E$ is an elliptic curve over $\kappa(s)$ in our previous sense.

If $E \to S$ is an elliptic curve, then E is fiber by fiber a curve of genus 1 (use Theorem 5.11). Moreover, the identity element defines a section $e : S \to E$. Conversely, we have the following extension of the constructions in §5.

Theorem 6.5. *Let $E \to S$ be proper and smooth with geometrically connected fibers of dimension 1 and genus 1. Let $e : S \to E$ be a section. Then there exists a unique S-group scheme structure $E \times_S E \to E$ with identity element e.*

We apply this theorem to families of cubic equations. Let R be a ring and let $a, b \in R$ be two elements. Consider the homogeneous Weierstrass equation

$$F_{a,b} = Y^2Z - X^3 - aXZ^2 - bZ^3 \quad \in R[X, Y, Z]. \tag{6.4}$$

We take $S = \mathrm{Spec}(R)$ as our base and consider the vanishing locus

$$E_{a,b} := V_+(F_{a,b}) \subset \mathbb{P}^2_S.$$

If k is a field and $s : \mathrm{Spec}(k) \to S$ a k-valued point of S, then we obtain values $s^*(a), s^*(b) \in k$ by specialization. It is clear from the definition that

$$\mathrm{Spec}(k) \times_S E_{a,b} = E_{s^*(a),s^*(b)}.$$

In particular, the fiber of $E_{a,b}$ in s is a cubic curve in $\mathbb{P}^2_k$. We see that $E_{a,b} \to S$ is a projective morphism with 1-dimensional fibers.

Consider for a moment an affine curve $V(f) \subseteq \mathbb{A}^2_S$. Recall that $V(f) \to S$ is smooth if and only if the Jacobi criterion holds, which means that

$$(\partial f/\partial x,\ \partial f/\partial y) \ \in\ R[x, y]dx \oplus R[x, y]dy$$

has rank 1 in each point of $V(f)$. This criterion can be checked fiber by fiber. We conclude that if for all $s \in S$, the specialization

$$E_{a(s),b(s)} \subset \mathbb{P}^2_{\kappa(s)}$$

is a smooth curve, then $E \to S$ is a smooth morphism. Moreover, by Remark 5.25, all these curves have genus 1.

Assume that $E \to S$ is smooth. The shape of (6.4) ensures that the section $[0 : 1 : 0] : S \to \mathbb{P}^2_S$ factors through E. By Theorem 6.5, there is a unique group scheme structure on E with neutral element $[0 : 1 : 0]$. In this way, we have defined an elliptic curve over S.

Construction 6.6 (The universal Weierstrass family)**.** The previous examples all come by specialization from a universal family. Let us, for simplicity, restrict to $\mathbb{Z}[1/6]$-algebras. The discriminant of a polynomial of the form $x^3 + ax + b$ is $4a^3 + 27b^2$, and

$$x^3 + ax + b \text{ is separable} \quad \Longleftrightarrow \quad 4a^3 + 27b^2 \neq 0.$$

Consider the ring

$$R := \mathbb{Z}[1/6][a, b][\Delta^{-1}], \qquad \Delta = 4a^3 + 27b^2,$$

set $S = \mathrm{Spec}(R)$, and let $E_{a,b} \to S$ be as before. We have inverted the discriminant, so for every $s \in S$, the polynomial $x^3 + a(s)x + b(s) \in \kappa(s)[x]$ is separable. By Lemma 5.1, the morphism $E_{a,b} \to S$ is smooth and hence, as just explained, an elliptic curve with identity section $[0 : 1 : 0]$. It is called the *universal Weierstrass family.*[7]

Why the name "universal"? Let T be any $\mathbb{Z}[1/6]$-scheme and let $\alpha, \beta \in \mathcal{O}_T(T)$ be functions such that, for all $t \in T$, the polynomial $x^3 + \alpha(t)x + \beta(t)$ is separable. On the one hand, we have previously defined a Weierstrass elliptic curve $E_{\alpha,\beta} \to T$. On the other hand, (α, β) give rise to a morphism $T \to S$. We find that

$$E_{\alpha,\beta} = T \times_S E_{a,b}$$

which shows that every Weierstrass family comes by pullback from the universal family.

[7]More precisely, it is the universal *simplified* Weierstrass family. The construction can also be carried out for the general Weierstrass equation (5.14) and then includes residue characteristics 2 and 3.

6.4. **Torsion.** We can now use the universal Weierstrass family to extend our knowledge about torsion of elliptic curves from characteristic 0 to all cases.

Theorem 6.7. *Let $E \to S$ be an elliptic curve and let $n \neq 0$. Then $E[n]$ is finite and locally free of rank n^2 over S.*

The proof requires a bit more algebraic geometry which we will leave aside in this course. It suffices for us to have the following summary result.

Proposition 6.8. *Let S be a scheme and let $f : E_1 \to E_2$ be a homomorphism of elliptic curves over S.*

(1) Assume that S is connected and that there exists a point $s \in S$ such that the fiber homomorphism $f(s) : E_1(s) \to E_2(s)$ is 0. Then f is zero.

(2) Assume that f is fiberwise non-zero. Then f is finite and locally free.

Remark 6.9. The properties in Proposition 6.8 are very similar to the ones of $\mathbb{G}_m$ in Proposition 4.7.

Proof of Theorem 6.7. For simplicity, we restrict to $\mathbb{Z}[1/6]$-schemes.

Step 1: The universal Weierstrass curve. The universal Weierstrass curve is defined over $S = \operatorname{Spec} \mathbb{Z}[1/6][a, b, \Delta^{-1}]$. This ring is an integral domain, so S is connected. Moreover, S has points in characteristic zero; for example, every ring homomorphism

$$\mathbb{Z}[1/6][a, b, \Delta^{-1}] \longrightarrow \mathbb{Q}, \quad a, b \longmapsto \alpha, \beta, \quad \Delta(\alpha, \beta) \neq 0$$

defines such a point. By our results from the complex case (Corollary 6.3), we know that over these points, multiplication by n is non-zero and finite locally free of degree n^2. By Proposition 6.8, we see that $[n]$ is finite and locally free of degree n^2 for the whole Weierstrass family.

Step 2: Specialization to specific elliptic curves. Let $E \to T$ be an arbitrary family of elliptic curves with $6 \in \mathcal{O}_T(T)^\times$. For every $t \in T$, the fiber elliptic curve $E(t) \to \operatorname{Spec}(\kappa(t))$ can be defined by a Weierstrass equation (Theorems 5.11 and 5.23, as well as the discussion up to (5.16)). That is, the fibers $E(t)$ all come by pullback from the universal Weierstrass family. Hence, $[n]$ is fiber by fiber finite of degree n^2. By Proposition 6.8, $[n]$ itself is finite and locally free of degree n^2. The kernel $E[n]$ is the pullback of $[n]$ along $S \to E$, and hence finite and locally free of rank n^2 over S as claimed. $\square$

7. The modular curve (algebraically)

In the last few lectures, we have

(1) Defined elliptic curves in terms of group schemes,

(2) Proved that elliptic curves can always be defined by Weierstrass equations (simplified if $\operatorname{char}(k) \neq 2, 3$),

(3) Constructed the universal Weierstrass family, and

(4) Used the universal Weierstrass family to show that $E[n]$ is always of order n^2.

Today, we want to expand on (3) and (4), and define a space that uniquely classifies elliptic curves together with a trivialization of their n-torsion.

7.1. **Moduli spaces.** Let us, for simplicity exclude residue characteristics 2 and 3 throughout the lecture. Consider an elliptic curve E over a field k. By (2) above, we may find $\alpha, \beta \in k$ such that E is isomorphic to the (closure in $\mathbb{P}^2_k$ of the) curve defined by

$$y^2 = x^3 + \alpha x + \beta.$$

The parameters α and β are not unique, however. Indeed, let $\lambda \in k^\times$ and consider the curve

$$y^2 = x^3 + (\lambda^{-4}\alpha)x + (\lambda^{-6}\beta). \tag{7.1}$$

It is isomorphic to the previous curve by the substitution $(x, y) \mapsto (\lambda^2 x, \lambda^3 y)$. This shows that our universal Weierstrass family

$$\mathcal{E} \longrightarrow \operatorname{Spec}(\mathbb{Z}[1/6, a, b, \Delta^{-1}]), \tag{7.2}$$

overparametrizes isomorphism classes of elliptic curves. More precisely, for E/k as above, we find a one-dimensional parameter family

$$\mathbb{G}_{m,k} \longrightarrow \operatorname{Spec}(\mathbb{Z}[1/6, a, b, \Delta^{-1}]), \quad \lambda \longmapsto (\lambda^{-4}\alpha, \lambda^{-6}\beta)$$

over which the relative curve $\mathcal{E}$ (7.2) is fiber by fiber isomorphic to E.

Question 7.1. Is it possible to improve on the construction of the universal Weierstrass family, and construct a family in which every elliptic curves occurs exactly once?

The precise mathematical meaning of this question is as follows.

Question 7.2 (Precise form)**.** Do there exist a scheme $\mathcal{M}$ and an elliptic curve $\mathcal{E} \to \mathcal{M}$ with the following property: For every scheme S and elliptic curve $E \to S$, there exists a *unique* morphism $u : S \to \mathcal{M}$ such that $u^*(\mathcal{E}) \cong E$? Here,

$$u^*(\mathcal{E}) := S \underset{u,\mathcal{M}}{\times} \mathcal{E}$$

denotes the pullback of $\mathcal{E}$ along u. If $(\mathcal{M}, \mathcal{E})$ exists, then we call $\mathcal{M}$ the *moduli space* of elliptic curves and $\mathcal{E}$ the *universal elliptic curve*.

The pair $(\mathcal{M}, \mathcal{E})$ is uniquely determined up to unique isomorphism. Namely, assume that $\mathcal{E}_1 \to \mathcal{M}_1$ and $\mathcal{E}_2 \to \mathcal{M}_2$ are two universal elliptic curves over their respective moduli spaces. By the universal properties, there exist morphisms $u : \mathcal{M}_1 \to \mathcal{M}_2$ and $v : \mathcal{M}_2 \to \mathcal{M}_1$ such that $u^*(\mathcal{E}_2) \cong \mathcal{E}_1$ and $v^*(\mathcal{E}_2) \cong \mathcal{E}_1$. The composition $v \circ u : \mathcal{M}_1 \to \mathcal{M}_1$ satisfies $(v \circ u)^*(\mathcal{E}_1) \cong \mathcal{E}_1$. By the uniqueness part of the universal property of $\mathcal{M}_1$, we find $v \circ u = \mathrm{id}_{\mathcal{M}_1}$. By the same argument, $u \circ v = \mathrm{id}_{\mathcal{M}_2}$. So we see $(\mathcal{M}_1, \mathcal{E}_1) \cong (\mathcal{M}_2, \mathcal{E}_2)$ in a unique way.

Answer 7.3. Assume that $(\mathcal{M}, \mathcal{E})$ exists. Then, in particular, for every field extension $k_0 \subset k$, the map $\mathcal{M}(k_0) \to \mathcal{M}(k)$ from k_0-valued points to k-valued points would be injective. (This is simply a property of schemes.) By the universal property, this would mean that for every k_0/k, the map

$$\begin{array}{ccc} \left\{ \begin{array}{c} \text{Isomorphism classes of} \\ \text{ellipt. curves over } k_0 \end{array} \right\} & \longrightarrow & \left\{ \begin{array}{c} \text{Isomorphism classes of} \\ \text{ellipt. curves over } k \end{array} \right\} \\ E & \longmapsto & k \otimes_{k_0} E \end{array} \tag{7.3}$$

would be injective. However, the next example shows that this map is usually not injective, so $(\mathcal{M}, \mathcal{E})$ cannot exist.

Example 7.4 (Quadratic twists)**.** Consider $\alpha, \beta \in \mathbb{Q}$, a non-square integer $D \neq -1$, and the two cubic equations (over $\mathbb{Q}$)

$$y^2 = x^3 + \alpha x + \beta, \qquad Dy^2 = x^3 + \alpha x + \beta. \tag{7.4}$$

The second equation can be brought into simplified Weierstrass form by substituting Dx and Dy for x and y, which gives

$$y^2 = x^3 + D^{-2}\alpha x + D^{-3}\beta. \tag{7.5}$$

Let us assume $\Delta(\alpha, \beta) \neq 0$ which also implies $\Delta(D^{-2}\alpha, D^{-3}\beta) = D^{-6}\Delta(\alpha, \beta) \neq 0$, so (7.4) defines two elliptic curves E and E_D over $\mathbb{Q}$. The curve E_D is called a *quadratic twist* of E.

On the one hand, E and E_D are clearly isomorphic over $\mathbb{Q}(\sqrt{D})$, because there we have the substitution $y \mapsto \sqrt{D}y$. On the other hand, one can show that E and E_D are not isomorphic over $\mathbb{Q}$. For example one can prove that two simplified Weierstrass equations over a field k of characteristic $\neq 2, 3$

$$y^2 = x^3 + \alpha_1 x + \beta_1, \qquad y^2 = x^3 + \alpha_2 x + \beta_2 \tag{7.6}$$

are isomorphic *if and only if* there exists $\lambda \in k^\times$ with $(\alpha_2, \beta_2) = (\lambda^4\alpha_1, \lambda^6\alpha_2)$. Since we have assumed D to be integral, not a square and $\neq -1$, there is no $\lambda \in \mathbb{Q}^\times$ with $(D^{-2}\alpha, D^{-3}\beta) = (\lambda^{-4}\alpha, \lambda^{-6}\beta)$ (unique prime factorization). In terms of (7.4) and (7.5), this means that E and E_D are not isomorphic.

In summary, we have defined two elliptic curves E and E_D over $\mathbb{Q}$ such that

$$\mathbb{Q}(\sqrt{D}) \otimes_{\mathbb{Q}} E \ncong \mathbb{Q}(\sqrt{D}) \otimes_{\mathbb{Q}} E_D.$$

This shows that (7.3) is not injective.

7.2. **Level structure.** Heuristically, the reason that there is no moduli space of elliptic curves is that elliptic curves have non-trivial automorphisms. For example, every elliptic curve has multiplication by -1 as automorphism, and this is what underlies the quadratic twist construction from Example 7.4.

In fact, this phenomenon is closely related to $\mathrm{GL}_2(\widehat{\mathbb{Z}}) \subset \mathrm{GL}_2(\mathbb{A}_f)$ not being small enough for the adelic double quotient formalism. We later learned in Proposition 3.21 that a simple family of small enough subgroups are the principal congruence subgroups $K(n)$ with $n \geq 3$. In the same spirit, we now introduce a notion of elliptic curve with level-n-structure. These will then have nice moduli spaces.

Example 7.5. Let us first get some intuition by considering the roots of unity. Let k be a field and let $n \geq 1$ be prime to $\mathrm{char}(k)$. Recall that

$$\mu_{n,k} = \operatorname{Spec} k[t]/(t^n - 1).$$

Consider the factorization of $t^n - 1$ into irreducible polynomials over k,

$$t^n - 1 = \prod_{\zeta \in \mu_n(k)} (t - \zeta) \cdot \prod_{i=1}^{r} f_i.$$

Since $t^n - 1$ is separable, the multiplicity of every factor is 1. Moreover, the linear factors correspond to the n-th roots of unity in $k^\times$, denoted by $\mu_n(k)$. The remaining factors $f_1, \ldots, f_r$ are of degree ≥ 2. If we translate this to schemes, we find a disjoint union decomposition

$$\mu_{n,k} = \bigsqcup_{\zeta \in \mu_n(k)} \operatorname{Spec}(k) \sqcup \bigsqcup_{i=1}^{r} \operatorname{Spec}(K_i) \tag{7.7}$$

where $K_i = k[t]/(f_i)$ is some non-trivial field extension of k. For example, we have

$$\mu_{4,\mathbb{Q}} = \bigsqcup_{\zeta \in \{\pm 1\}} \operatorname{Spec}(\mathbb{Q}) \sqcup \operatorname{Spec}(\mathbb{Q}(i)),$$

$$\mu_{4,\mathbb{Q}(i)} = \bigsqcup_{\zeta \in \{\pm 1, \pm i\}} \operatorname{Spec}(\mathbb{Q}(i)),$$

$$\mu_{19,\mathbb{F}_5} = \operatorname{Spec}(\mathbb{F}_5) \sqcup \operatorname{Spec}(\mathbb{F}_{5^{18}}).$$

In general, the union

$$\bigsqcup_{\zeta \in \mu_n(k)} \operatorname{Spec}(k) \subseteq \mu_{n,k}$$

of the connected components corresponding to the k-points itself forms a group scheme. It is a constant group scheme, isomorphic to $\underline{\mu_n(k)}_{\mathrm{Spec}(k)}$.

Definition 7.6. Let S be a scheme and let Γ be a group. The constant group scheme $\underline{\Gamma}_S$ is the S-scheme $\bigsqcup_{\gamma\in\Gamma} S$ together with the S-group scheme structure

$$\underline{\Gamma}_S \times_S \underline{\Gamma}_S \;=\; \bigsqcup_{(\gamma_1,\gamma_2)\in\Gamma\times\Gamma} S \;\longrightarrow\; \underline{\Gamma}_S$$

that reflects the multiplication of Γ: map the copy of S corresponding to (γ_1,γ_2) with id_S to the copy corresponding to $\gamma_1\gamma_2$.

Proposition 7.7. *Let E be an elliptic curve over a field k and let $n \geq 1$ be prime to* $\mathrm{char}(k)$*. Then $E[n](\bar{k})$ is a finite group isomorphic to $(\mathbb{Z}/n\mathbb{Z})^2$.*

Proof. By Theorem 6.7, $E[n]$ is a finite k-group scheme of order n^2. By Theorem 4.9, it is étale over k. So

$$E[n] \xrightarrow{\sim} \bigsqcup_{i=1}^{s} \mathrm{Spec}(K_i)$$

for finite separable field extensions K_i/k with $\sum_{i=1}^s [K_i : k] = n^2$. For every i,

$$\bar{k} \otimes_k K_i \xrightarrow{\sim} \bar{k}^{[K_i:k]},$$

so $E[n](\bar{k})$ is a finite group of order n^2. For every divisor $d \mid n$, the same argument shows that $E[d](\bar{k})$, which equals the d-torsion in $E[n](\bar{k})$, has order d^2. By the classification of finite abelian groups, the only possibility is then $E[n](\bar{k}) \cong (\mathbb{Z}/n\mathbb{Z})^2$. ☐

Example 7.8. Let E be an elliptic curve over a field k and let $n \geq 1$ be prime to $\mathrm{char}(k)$. By the same logic as in Example 7.5, we can decompose $E[n]$ as

$$E[n] = \bigsqcup_{x\in E[n](k)} \mathrm{Spec}(k) \sqcup \big(\mathrm{Rest}\big).$$

Where the rest is the union of all connected components $\mathrm{Spec}(K)$ with $[K : k] \geq 2$. The rational part can also be written as the constant group scheme $\underline{E[n](k)}_{\mathrm{Spec}(k)}$. The group $E[n](k)$ is a subgroup of $E[n](\bar{k})$. So we know from Proposition 7.7 that $E[n](k)$ is isomorphic to a subgroup of $(\mathbb{Z}/n\mathbb{Z})^2$. In general, the group $E[n](k)$ will depend on E, n, and k.

Definition 7.9. Let E be an elliptic curve over a $\mathbb{Z}[1/n]$-scheme S. A *level-n-structure* for E is an isomorphism

$$\eta : \; \underline{(\mathbb{Z}/n\mathbb{Z})^{\oplus 2}}_S \xrightarrow{\sim} E[n].$$

Equivalently, it is the datum of two sections $\eta_1 = \eta(1,0)$ and $\eta_2 = \eta(0,1)$ in $E(S)$ that fiber by fiber induce isomorphisms

$$\eta_s : \; (\mathbb{Z}/n\mathbb{Z})^{\oplus 2} \xrightarrow{\sim} E[n](\kappa(s)).$$

Remark 7.10. We could have formulated Definition 7.9 for every base scheme S, not just those with $n \in \mathcal{O}_S(S)^\times$. However, constant group schemes are clearly étale. Hence, if a level-n-structure η exists, then $E[n]$ is also étale. One can show that

$$E[n] \text{ is étale} \iff n \in \mathcal{O}_S(S)^\times,$$

so the more general definition would simply be empty for S not over $\mathbb{Z}[1/n]$.

7.3. **Back to moduli spaces.** Let (E_1, η_1) and (E_2, η_2) be two elliptic curves with level-n-structure over a scheme S. An isomorphism between these pairs is an isomorphism $\gamma : E_1 \to E_2$ such that $\eta_2 = \gamma \circ \eta_1$. The key point is that adding level structure solves our problem of elliptic curves having automorphisms:

Proposition 7.11 ([10, Proposition 14.8])**.** *Let n be ≥ 3 and let (E, η) be an elliptic curve with level-n-structure over a scheme S. Then the only automorphism of (E, η) is the identity.*

Assume that $(E, \eta)/S$ is an elliptic curve with level-n-structure and that $u : T \to S$ is a morphism. Then we may form the pullback

$$u^*(E, \eta) := (T \times_S E, \mathrm{id}_T \times \eta).$$

In terms of the two basis sections $\eta_1 = \eta(1, 0)$ and $\eta_2 = \eta(0, 1)$, we are considering the pullback $E(S) \to E(T) = (T \times_S E)(T)$.

Theorem 7.12 (The modular curve)**.** *For every integer $n \geq 3$, there exists a moduli space of elliptic curves with level-n-structure.*

That is, there exist a $\mathbb{Z}[1/n]$-scheme $\mathcal{M}_n$, an elliptic curve $\mathcal{E} \to \mathcal{M}_n$, and a level-$n$-structure $\eta \in \mathcal{E}(S)^2$, that together have the following universal property:

For every elliptic curve with level-n-structure (E, η_0) over a scheme S, there exists a unique morphism $u : S \to \mathcal{M}_n$ such that

$$u^*(\mathcal{E}, \eta) \cong (E, \eta_0).$$

Proof idea. Let us focus on $\mathcal{M}_n[1/6]$. The primes 2 and 3 need to be treated by different arguments. We only sketch some ideas and refer the interested reader to [10, §14] for details.

Step 1. Consider the universal Weierstrass family

$$\mathcal{E} \longrightarrow \mathcal{W} = \mathrm{Spec}\,\mathrm{Spec}[1/6, a, b, \Delta^{-1}].$$

Recall that every elliptic curve already occurs (non-uniquely) in this family; our task is to pass to a quotient that has a uniqueness property.

There is a finite étale morphism

$$\mathcal{W}_n \longrightarrow \mathcal{W}[1/n]$$

that parametrizes the level-n-structures on $\mathcal{E}$. That is, giving a morphism $S \to \mathcal{W}_n$ is the same as giving a morphism $u : S \to \mathcal{W}[1/n]$ together with a level-n-structure for $u^*(\mathcal{E})$.

Step 2. Identity (7.1) defines an action of $\mathbb{G}_{m,\mathbb{Z}[1/6]}$ on $\mathcal{W}$. This action can be lifted to an action of $\mathbb{G}_{m,\mathbb{Z}[1/6]}$ on $\mathcal{W}_n$. We define $\mathcal{M}_n$ by taking a quotient

$$\mathcal{M}_n := \mathbb{G}_{m,\mathbb{Z}[1/6n]} \backslash\backslash \mathcal{W}_n.$$

Step 3. We have assumed $n \geq 3$, so the action of $\mathbb{G}_{m,\mathbb{Z}[1/6n]}$ on $\mathcal{W}_n$ is without fixed points by Proposition 7.11. This implies that $\mathcal{W}_n \to \mathcal{M}_n$ is a $\mathbb{G}_m$-torsor which allows to descend the pair $(\mathcal{E}, \eta)$ from $\mathcal{W}_n$ to $\mathcal{M}_n$. □

8. The modular curve (as Shimura variety)

In the very first lecture, we sketched the general definition of Shimura varieties. Let (G, X) be a Shimura datum. The corresponding Shimura variety for small enough level $K \subset G(\mathbb{A}_f)$ starts life as the complex manifold

$$G(\mathbb{Q})\backslash(X \times G(\mathbb{A}_f)/K). \tag{8.1}$$

Next, the theorem of Baily–Borel (Theorem 1.3) states that this manifold has a unique structure as complex variety. That is, there exists a complex variety $\mathrm{Sh}_K(G, X)_{\mathbb{C}}$, unique up to isomorphism, such that the analytification $\mathrm{Sh}_K(G,X)_{\mathbb{C}}(\mathbb{C})$ in the sense of (6.1) is isomorphic to (8.1). Finally, Deligne, Milne, and Borovoi proved that there exists a canonical variety $\mathrm{Sh}_K(G,X)$ over a canonical number field $E \subset \mathbb{C}$ together with an isomorphism

$$\mathbb{C} \otimes_E \mathrm{Sh}_K(G,X) \xrightarrow{\sim} \mathrm{Sh}_K(G,X)_{\mathbb{C}}.$$

The variety $\mathrm{Sh}_K(G,X)$ is then the Shimura variety of level K of (G,X).

Today, our goal is to carry out this construction for GL_2. We have already studied the double quotients $\mathrm{GL}_2(\mathbb{Q})\backslash(\mathbb{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/K)$ in §2 and §3. We have proved that K is small enough if it is contained in a principal level subgroup $K(n)$ with $n \geq 3$ (Proposition 3.21). In the more recent lectures, we have then defined the moduli space of elliptic curves with level-n-structure. Taking its generic fiber, we obtain an algebraic curve $\mathcal{M}_{n,\mathbb{Q}}$ over $\mathbb{Q}$.

Theorem 8.1. *For any $n \geq 3$, the rational curve $\mathcal{M}_{n,\mathbb{Q}}$ is the Shimura variety for GL_2, Deligne homomorphism (1.3), and level subgroup $K(n)$.*

We first explain how to construct the isomorphism

$$\mathcal{M}_n(\mathbb{C}) \xrightarrow{\sim} K(n)\backslash(\mathbb{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/K(n))$$

which, essentially, is also the proof of Theorem 8.1. Afterwards, we will look at group actions and more general level subgroups.

8.1. **Elliptic curves and the upper half plane.** Our first aim is to understand the relation between elliptic curves and the manifold $\mathbb{H}^{\pm}$. Recall from Corollary 6.2 that analytification gives an equivalence

$$\begin{aligned} \{\text{Elliptic curves}/\mathbb{C}\} &\xrightarrow{\sim} \left\{\begin{matrix}\text{Compact complex Lie groups}\\ \text{of the form } \mathbb{C}/\Lambda\end{matrix}\right\} \\ E &\longmapsto E(\mathbb{C}). \end{aligned} \tag{8.2}$$

On the right hand side, we are simply considering $\mathbb{Z}$-lattices $\Lambda \subset \mathbb{C}$. So, in order to parametrize elliptic curves over $\mathbb{C}$ up to isomorphism, we need to parametrize such lattices and understand when two quotients $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ are isomorphic.

Lemma 8.2. *Two lattices $\Lambda_1, \Lambda_2 \subset \mathbb{C}$ define isomorphic elliptic curves if and only if there exists $\lambda \in \mathbb{C}^{\times}$ such that $\Lambda_2 = \lambda\Lambda_1$.*

More generally, the homomorphisms from $\mathbb{C}/\Lambda_1$ to $\mathbb{C}/\Lambda_2$ are given by

$$\mathrm{Hom}(\mathbb{C}/\Lambda_1, \mathbb{C}/\Lambda_2) = \{\lambda \in \mathbb{C} \mid \lambda\Lambda_1 \subseteq \Lambda_2\}.$$

Proof. Assume $\lambda \in \mathbb{C}$ satisfies $\lambda\Lambda_1 \subseteq \Lambda_2$. Multiplication by λ defines a holomorphic group homomorphism $\mathbb{C} \to \mathbb{C}$. By the condition $\lambda\Lambda_1 \subseteq \Lambda_2$, it descends to a homomorphism

$$\lambda : \mathbb{C}/\Lambda_1 \longrightarrow \mathbb{C}/\Lambda_2.$$

If even the equality $\lambda\Lambda_1 = \Lambda_2$ holds, then this map is an isomorphism with inverse defined by λ^{-1}.

Conversely, assume we are given a holomorphic group homomorphism $f : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$. It lifts uniquely to a holomorphic homomorphism on universal covers $\widetilde{f} : \mathbb{C} \to \mathbb{C}$. Writing $\widetilde{f}$ as a power series $\widetilde{f}(z) = \sum_{i\geq 1} a_i z^i$ centered at 0 (and convergent on all of $\mathbb{C}$), the condition

$$\widetilde{f}(z_1 + z_2) = \widetilde{f}(z_1) + \widetilde{f}(z_2)$$

implies that $\widetilde{f}(z) = \lambda z$ for some scalar $\lambda \in \mathbb{C}$. Since $\widetilde{f}$ lifts f, this scalar satisfies $\lambda\Lambda_1 \subseteq \Lambda_2$ as claimed. $\square$

Lemma 8.2 states that

$$\begin{aligned} \{\text{Lattices } \Lambda \subset \mathbb{C}\}/\mathbb{C}^\times &\xrightarrow{\sim} \{\text{Elliptic curves}/\mathbb{C}\}/\cong \\ \Lambda &\longmapsto \mathbb{C}/\Lambda. \end{aligned} \tag{8.3}$$

In order find the relation with $\mathbb{H}^\pm$, we now overparametrize all lattices by considering the set

$$\left\{ (\Lambda, \tau_1, \tau_2) \;\middle|\; \begin{array}{c} \Lambda \subset \mathbb{C} \text{ a lattice} \\ \tau_1, \tau_2 \in \Lambda \text{ a } \mathbb{Z}\text{-basis} \end{array} \right\}. \tag{8.4}$$

It is equipped with an action of $\mathbb{C}^\times \times \mathrm{GL}_2(\mathbb{Z})$ by

$$(\lambda, \gamma) \cdot (\Lambda, \tau_1, \tau_2) := (\lambda\Lambda,\ \lambda(\tau_1, \tau_2) \cdot \gamma^t). \tag{8.5}$$

Concretely, if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, then

$$(\tau_1, \tau_2) \cdot \gamma^t = (a\tau_1 + b\tau_2, c\tau_1 + d\tau_2).$$

On the one hand, the possible choices of a $\mathbb{Z}$-basis for a lattice form a simply transitive $\mathrm{GL}_2(\mathbb{Z})$-orbit. Thus, if we take the quotient by the $\mathrm{GL}_2(\mathbb{Z})$-action, we precisely recover the set of lattices without extra data:

$$\mathrm{GL}_2(\mathbb{Z})\backslash\{(\Lambda, \tau_1, \tau_2)\} \xrightarrow{\sim} \{\Lambda \subset \mathbb{C}\}. \tag{8.6}$$

Note that the $\mathrm{GL}_2(\mathbb{Z})$-action and the action of $\mathbb{C}^\times$ by scaling commute, and that (8.6) is $\mathbb{C}^\times$-equivariant.

On the other hand, we may also first quotient out the $\mathbb{C}^\times$-action. Namely, there exists a unique representative in the orbit $\mathbb{C}^\times \cdot (\Lambda, \tau_1, \tau_2)$ of the form $(\Lambda', \tau, 1)$. It is given by

$$\tau_2^{-1} \cdot (\Lambda, \tau_1, \tau_2) = (\tau_2^{-1}\Lambda, \tau_1/\tau_2, 1).$$

Since τ_1, τ_2 are a basis for a lattice in $\mathbb{C}$, they are also an $\mathbb{R}$-basis of $\mathbb{C}$, and so the ratio τ_1/τ_2 does not lie in $\mathbb{R}$. In this way, we find

$$\begin{aligned} \mathbb{C}^\times\backslash\{(\Lambda,\ \tau_1,\ \tau_2)\} &\xrightarrow{\sim} \mathbb{H}^\pm \\ (\Lambda,\ \tau_1,\ \tau_2) &\longmapsto \tau_1/\tau_2 \\ (\mathbb{Z}\tau + \mathbb{Z},\ \tau,\ 1) &\longleftarrow\!\shortmid\ \tau. \end{aligned} \tag{8.7}$$

We still have a $\mathrm{GL}_2(\mathbb{Z})$-action on the left hand side of (8.7). How does it translate to the right hand side? Given $\tau \in \mathbb{H}^\pm$ and $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we calculate

$$\begin{tikzcd} (\mathbb{Z}\tau + \mathbb{Z},\ \tau,\ 1) \arrow[d, mapsto, "\gamma"'] & \tau \arrow[l, mapsto] \\ (\mathbb{Z}\tau + \mathbb{Z},\ a\tau + b,\ c\tau + d) \arrow[r, mapsto] & \frac{a\tau+b}{c\tau+d}. \end{tikzcd}$$

In other words, we see that the action is given by Moebius transformations. In this way, we have constructed the lower isomorphism in the diagram

$$\begin{tikzcd} \mathbb{C}^\times \times \mathrm{GL}_2(\mathbb{Z})\backslash\{(\Lambda, \tau_1, \tau_2)\} \arrow[r, "\cong"] \arrow[d, "\cong"', "(8.7)"] & \mathbb{C}^\times\backslash\{\Lambda \subset \mathbb{C}\} \arrow[d, "(8.3)"', "\cong"] \\ \mathrm{GL}_2(\mathbb{Z})\backslash\mathbb{H}^\pm \arrow[r, dashed, "\cong"] & \left\{ \begin{array}{l} \text{Isom. classes of} \\ \text{ellipt. curves } /\mathbb{C} \end{array} \right\}. \end{tikzcd} \tag{8.8}$$

The lower horizontal map is simply given by

$$\tau \longmapsto \mathbb{C}/(\mathbb{Z}\tau + \mathbb{Z}).$$

8.2. **The isomorphism between $\mathcal{M}_n(\mathbb{C})$ and $\mathcal{S}_{K(n)}(\mathbb{C})$.** Let us introduce the following shorthand notation: For a level subgroup $K \subset G(\mathbb{A}_f)$, we write

$$\mathcal{S}_K(\mathbb{C}) := \mathrm{GL}_2(\mathbb{Q})\backslash(\mathbb{H}^\pm \times \mathrm{GL}_2(\mathbb{A}_f)/K).$$

Theorem 8.3. *For ever $n \geq 3$, there is an isomorphism*

$$\mathcal{S}_{K(n)}(\mathbb{C}) \xrightarrow{\sim} \mathcal{M}_n(\mathbb{C}).$$

Proof. Step 1: Parametrizing elliptic curves with level structure by lattices. Our first step is to upgrade the upper and right hand side arrows in (8.8) to include level structure.

Let us, from now on, view a basis $\tau_1, \tau_2 \in \Lambda$ as an isomorphism

$$\alpha : \mathbb{Z}^2 \xrightarrow{\sim} \Lambda, \quad e_i \longmapsto \tau_i, \quad i = 1, 2.$$

This helps us keep track of the $\mathrm{GL}_2(\mathbb{Z})$-action because we now have the cleaner expression

$$\gamma \cdot (\Lambda, \alpha) = (\Lambda, \alpha \circ \gamma^t). \tag{8.9}$$

Next, we observe that the n-torsion of $\mathbb{C}/\Lambda$ is given by the quotient $(n^{-1}\Lambda)/\Lambda$. Using multiplication by n, we identify

$$(n^{-1}\Lambda)/\Lambda \xrightarrow{\sim} \Lambda/n\Lambda, \quad n^{-1}\lambda \longmapsto \lambda.$$

Giving a level structure for $\mathbb{C}/\Lambda$ is now the same as giving an isomorphism

$$\eta : (\mathbb{Z}/n\mathbb{Z})^2 \xrightarrow{\sim} \Lambda/n\Lambda.$$

Let us consider the set of triples

$$\left\{ (\Lambda, \alpha, \eta) \;\middle|\; \begin{array}{c} \Lambda \subset \mathbb{C} \text{ a lattice, } \alpha : \mathbb{Z}^2 \xrightarrow{\sim} \Lambda \text{ a basis} \\ \eta : (\mathbb{Z}/n\mathbb{Z})^2 \xrightarrow{\sim} \Lambda/n\Lambda \text{ a level-}n\text{-structure} \end{array} \right\}.$$

It is equipped by an action of $\mathbb{C}^\times \times \mathrm{GL}_2(\mathbb{Z})$ in the same way as (8.5) before:

$$(\lambda, \gamma) \cdot (\Lambda, \alpha, \eta) := (\lambda\Lambda, \lambda\alpha\gamma^t, \lambda\eta).$$

Taking the quotient by $\mathrm{GL}_2(\mathbb{Z})$ is the same as forgetting α, while scaling by $\mathbb{C}^\times$ is the same as passing to the isomorphism class of elliptic curve. So we find

$$\mathbb{C}^\times \times \mathrm{GL}_2(\mathbb{Z})\backslash\{(\Lambda, \alpha, \eta)\} \xrightarrow{\sim} \left\{ \begin{array}{c} \text{Isom. classes of} \\ \text{ellipt. curves with} \\ \text{level-}n\text{-str. } (E, \eta)/\mathbb{C} \end{array} \right\}. \tag{8.10}$$

Step 2: Translate to a statement about $\mathbb{H}^\pm$. As in (8.8), we now want to interchange the order and quotient by $\mathbb{C}^\times$ first. To this end, we first simplify our triples (Λ, α, η). Namely, given such a triple, we may consider the composition

$$\alpha^{-1} \circ \eta := [\alpha^{-1} \bmod n] \circ \eta \in \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z}).$$

Clearly, the two data

$$(\Lambda, \alpha, \eta) \longleftrightarrow (\Lambda, \alpha, \alpha^{-1} \circ \eta) \tag{8.11}$$

are equivalent because we can get one from the other by composing maps. If we act by an element $\gamma \in \mathrm{GL}_2(\mathbb{Z})$ on the left of (8.11), then the matrix on the right transforms by left mulitplication with $\gamma^{t,-1}$ mod n because

$$(\alpha\gamma^t)^{-1} \circ \eta = \gamma^{t,-1} \circ (\alpha^{-1} \circ \gamma). \tag{8.12}$$

This motivates us to renormalize (8.11) as

$$\begin{aligned} \{(\Lambda, \alpha, \eta)\} &\xrightarrow{\sim} \{(\Lambda, \alpha)\} \times \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z}) \\ (\Lambda, \alpha, \eta) &\longmapsto \big((\Lambda, \alpha),\ (\alpha^{-1} \circ \eta)^{t,-1}\big). \end{aligned} \tag{8.13}$$

Under this normalized isomorphism, the $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$-action on the right hand side is given by the straightforward formula

$$\gamma \cdot \big((\Lambda, \alpha),\ g\big) = \big((\Lambda, \alpha\gamma^t),\ \gamma g\big). \tag{8.14}$$

Moreover, the $\mathbb{C}^\times$-scaling does not involve the third entry anymore,

$$\lambda \cdot \big((\Lambda, \alpha),\ g\big) = \big((\lambda\Lambda, \lambda\alpha),\ g\big).$$

This means that (8.7) applies unchanged and we find

$$\begin{aligned} \mathrm{GL}_2(\mathbb{Z})\backslash(\mathbb{H}^\pm \times \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})) &\xrightarrow{\sim} \left\{\begin{matrix}\text{Isom. classes of ellipt.} \\ \text{curves with level-}n\text{-str. } (E,\eta)/\mathbb{C}\end{matrix}\right\}. \\ (\tau, g) &\longmapsto (\mathbb{C}/\Lambda,\ (\tau, 1)\circ g^{t,-1}). \end{aligned} \tag{8.15}$$

Here, $(\tau, 1)\circ g^{t,-1}$ is the map

$$(\mathbb{Z}/n\mathbb{Z})^2 \xrightarrow{\sim} \Lambda/n\Lambda, \quad \begin{pmatrix} e \\ f \end{pmatrix} \longmapsto (\tau,\ 1)\cdot g^{t,-1} \cdot \begin{pmatrix} e \\ f \end{pmatrix}.$$

Step 3: Relate (8.15) to the definition of $\mathcal{S}_{K(n)}(\mathbb{C})$.

Lemma 8.4. *For every level subgroup $K \subseteq \mathrm{GL}_2(\widehat{\mathbb{Z}})$, the inclusion map induces an isomorphism*

$$\mathrm{GL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\widehat{\mathbb{Z}})/K \xrightarrow{\sim} \mathrm{GL}_2(\mathbb{Q})\backslash \mathrm{GL}_2(\mathbb{A}_f)/K. \tag{8.16}$$

Proof. The inclusion $\mathrm{GL}_2(\widehat{\mathbb{Z}}) \to \mathrm{GL}_2(\mathbb{A}_f)$ clearly descends to a map on quotients as in (8.16). Assume that for $g_1, g_2 \in \mathrm{GL}_2(\widehat{\mathbb{Z}})$, there exist $h \in \mathrm{GL}_2(\mathbb{Q})$ and $k \in K$ with $g_2 = hg_1k$. Then $h = g_2k^{-1}g_1^{-1}$ lies in $\mathrm{GL}_2(\widehat{\mathbb{Z}}) \cap \mathrm{GL}_2(\mathbb{Q})$ which equals $\mathrm{GL}_2(\mathbb{Z})$. So the map on quotients is injective.

For surjectivity, we recall from Proposition 3.19 that the determinant induces a bijection

$$\det : \mathrm{GL}_2(\mathbb{Q})\backslash \mathrm{GL}_2(\mathbb{A}_f)/K \xrightarrow{\sim} \mathbb{Q}^\times\backslash\mathbb{A}_f^\times/\det(K).$$

We stated that proposition for $\mathrm{GL}_2(\mathbb{Q})_{>0}$, but it also holds for the full matrix group with the same proof. Since

$$\{\pm 1\}\backslash\widehat{\mathbb{Z}}^\times \xrightarrow{\sim} \mathbb{Q}^\times\backslash\mathbb{A}_f^\times,$$

the determinant map restricted to the left hand side of (8.16) is already surjective. This means that (8.16) is also surjective. □

Lemma 8.4 can be used directly in the Shimura variety definition and provides an isomorphism

$$\mathrm{GL}_2(\mathbb{Z})\backslash(\mathbb{H}^\pm \times \mathrm{GL}_2(\widehat{\mathbb{Z}})/K) \xrightarrow{\sim} \mathrm{GL}_2(\mathbb{Q})\backslash(\mathbb{H}^\pm \times \mathrm{GL}_2(\mathbb{A}_f)/K).$$

If we specialize to $K = K(n)$ and use our result from Step 2 (8.15), then we exactly obtain an isomorphism

$$\mathcal{S}_{K(n)}(\mathbb{C}) \xrightarrow{\sim} \mathcal{M}_n(\mathbb{C})$$

as claimed by Theorem 8.3. □

8.3. **The $\mathrm{GL}_2(\widehat{\mathbb{Z}})$-action.** Recall that $K(n)$ is defined as a kernel,

$$K(n) = \ker\big(\,\mathrm{GL}_2(\widehat{\mathbb{Z}}) \longrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})\big).$$

In particular, it is a normal subgroup of $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ with quotient $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$. Let us write $[x, h]$ or $[x, hK]$ for the point$\mathrm{GL}_2(\mathbb{Q})\cdot(x, hK) \in \mathcal{S}_K$. The normality implies that we obtain a right-action of $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ on $\mathcal{S}_{K(n)}$ by

$$\begin{aligned} \mathrm{GL}_2(\widehat{\mathbb{Z}}) \times \mathcal{S}_{K(n)} &\longrightarrow \mathcal{S}_{K(n)} \\ [x, h]\cdot g &= [x, hg]. \end{aligned} \tag{8.17}$$

The point here is that

$$hK(n)\cdot g = hg(g^{-1}K(n)g) = hgK(n)$$

because of the normality. Put differently, the cosets hgK and $hkgK$ are equal for every $k \in K$, which ensures that (8.17) is well-defined. It is clear from the definition that $K(n)$ acts trivially, so this is really a $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$-action.

We now turn to the moduli space $\mathcal{M}_n$. (Recall that $n \geq 3$ is assumed for $\mathcal{M}_n$ to exist.) Given $g \in \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ and an S-valued point $(E, \eta) \in \mathcal{M}_n(S)$, we can define a new S-valued point $(E, \eta \circ g)$. In this way, $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ acts on the right of $\mathcal{M}_n$.

Proposition 8.5. *The isomorphism $\iota : \mathcal{S}_{K(n)} \xrightarrow{\sim} \mathcal{M}_n(\mathbb{C})$ constructed during the proof of Theorem 8.3 is $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$-equivariant in the sense that*

$$\iota([x, h]\cdot g) = \iota([x, h])\cdot g^{t,-1}.$$

Proof. Start with a point $(E, \eta) \in \mathcal{M}_n(\mathbb{C})$. We choose a lattice $\Lambda \subset \mathbb{C}$ and an isomorphism $\mathbb{C}/\Lambda \xrightarrow{\sim} E(\mathbb{C})$. By composition, we obtain a level structure $\eta : (\mathbb{Z}/n\mathbb{Z})^2 \xrightarrow{\sim} \Lambda/n\Lambda$. We arbitrarily choose a basis $\alpha : \mathbb{Z}^2 \xrightarrow{\sim} \Lambda$. In this way, we have represented (E, η) by a triple (Λ, α, η) as on the right hand side of (8.13). A representative in $\mathbb{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/K(n)$ is given by

$$\big[\tau := \mathbb{C}^{\times}(\Lambda, \alpha),\ hK(n) := (\alpha^{-1}\circ\eta)^{t,-1}\big]. \tag{8.18}$$

Here, τ is defined by (8.7) as before. If we substitute $\eta \circ g^{t,-1}$ in (8.18), then we change $hK(n)$ to $hgK(n)$. This is precisely what is claimed by Proposition 8.5. $\square$

Remark 8.6. Note that $g \mapsto g^{t,-1}$ defines a group automorphism of GL_2. The difference between g acting as g or $g^{t,-1}$ in Proposition 8.5 has no mathematical meaning but simply resulted from the choices we made along the way.

8.4. **General level groups.** We have defined the curve $\mathcal{S}_K$ for every small enough level $K \subset \mathrm{GL}_2(\mathbb{A}_f)$. In fact, one can even work without the "small enough" assumption. First, one defines $\mathcal{S}_K$ as set with quotient topology. Then, one proves that there exists a unique Riemann surface structure on $\mathcal{S}_K$ such that the quotient map

$$\mathbb{H}^{\pm} \times \mathrm{GL}_2(\mathbb{A}_f)/K \longrightarrow \mathcal{S}_K \tag{8.19}$$

is holomorphic. The only difference with the case K small enough is that (8.19) might not be a covering map in the sense of topology.

By contrast, we have only defined the algebraic modular curve for principal level subgroups. Our final aim in this section is to define an algebraic curve $\mathcal{M}_K$ for every level K, and to state its comparison with $\mathcal{S}_K$.

Proposition 8.7. *For every level $K \subset \mathrm{GL}_2(\mathbb{A}_f)$, there exist a group element $g \in \mathrm{GL}_2(\mathbb{A}_f)$ and $n \geq 1$ such that*

$$K(n) \subseteq g^{-1}Kg \subseteq \mathrm{GL}_2(\widehat{\mathbb{Z}}).$$

We will use the same argument as during the proof of Proposition 2.5. This argument was based on the notion of $\mathbb{Z}$-lattice in a $\mathbb{Q}$-vector space, and the fact that for any two lattices $\Lambda_1, \Lambda_2 \subset V$, there exists $g \in \mathrm{GL}(V)$ with $\Lambda_2 = g\Lambda_1$.

This statement is more general. Let R be a PID with fraction field K, and let V be a finite dimensional K-vector space (say of dimension m). An R-lattice in V is simply an R-submodule $\Lambda \subset V$ which is isomorphic to R^m. Then, for any two R-lattices $\Lambda_1, \Lambda_2 \subset V$, there exists $g \in \mathrm{GL}(V)$ with $\Lambda_2 = g\Lambda_1$.

The ring $\mathbb{A}_f$ is not a domain, but there is a still a very similar notion of $\widehat{\mathbb{Z}}$-lattice:

Definition 8.8. Let $\mathbb{V}$ be a finite free $\mathbb{A}_f$-module. A $\widehat{\mathbb{Z}}$-submodule $\widehat{\Lambda} \subset \mathbb{V}$ is called a lattice if satisfies the following, equivalent conditions. After choosing a basis $\mathbb{A}_f^m \xrightarrow{\sim} \mathbb{V}$, we have

(1) $\widehat{\Lambda}$ is finitely generated as $\widehat{\mathbb{Z}}$-module and $\mathbb{A}_f \cdot \widehat{\Lambda} = \mathbb{V}$.

(2) There exists an integer $c \geq 1$ such that
$$c \cdot \widehat{\mathbb{Z}}^m \ \subseteq \ \widehat{\Lambda} \ \subseteq \ c^{-1} \cdot \widehat{\mathbb{Z}}^m.$$

(3) There exist finitely many primes S and lattices $\Lambda_p \subset \mathbb{Q}_p^m$, $p \in S$, such that
$$\widehat{\Lambda} = \prod_{p \in S} \Lambda_p \times \prod_{p \notin S} \mathbb{Z}_p^m.$$

(4) There exists an element $g \in \mathrm{GL}_m(\mathbb{A}_f)$ with $\widehat{\Lambda} = g \cdot \widehat{\mathbb{Z}}^m$.

Proof of the equivalence of (1) – (4). Assume statement (1). On the one hand, $\widehat{\Lambda}$ has finitely many generators $\lambda_1, \ldots, \lambda_r$, each of which is an m-tuple of elements from $\mathbb{A}_f$. Recall that every $(x_p)_{p \text{ prime}} \in \mathbb{A}_f$ has only finitely many entries x_p not in $\mathbb{Z}_p$. So only finitely many denominators occur among the entries of the λ_j which means there exists $c_1 \geq 1$ with $\widehat{\Lambda} \subseteq c_1^{-1} \cdot \widehat{\mathbb{Z}}^m$. On the other hand, $\mathbb{A}_f \cdot \widehat{\Lambda} = \mathbb{A}_f^m$ means that each standard basis vector e_i can be written as a linear combination
$$e_i = \sum_{j=1}^{r} a_{ij}\lambda_j, \qquad a_{ij} \in \mathbb{A}_f.$$
Only finitely many denominators occur among the entries of the a_{ij}. So there exists $c_2 \geq 1$ with $c_2(a_{ij}) \in \mathrm{M}_{m\times r}(\widehat{\mathbb{Z}})$. This means $c_2 \cdot \widehat{\mathbb{Z}}^m \subseteq \widehat{\Lambda}$ and we obtain statement (2) with $c = c_1c_2$.

Assume statement (2). Let S be the set of all primes dividing c. For $p \in S$, define Λ_p as the lattice generated by the p-components of all elements of $\widehat{\Lambda}$. With these choices, (3) holds.

Assume statement (3). For every $p \in S$, let $g_p \in \mathrm{GL}_m(\mathbb{Q}_p)$ be such that $g_p \cdot \mathbb{Z}_p^m = \Lambda_p$. Define $g = (g_p)_{p\in S} \times (\mathrm{id}_m)_{p\notin S}$ which lies in $\mathrm{GL}_m(\mathbb{A}_f)$. We obtain $\widehat{\Lambda} = g \cdot \widehat{\mathbb{Z}}^m$ as in statement (4).

Finally, assume statement (4). Since multiplication by g defines an isomorphism $\widehat{\mathbb{Z}}^m \xrightarrow{\sim} \widehat{\Lambda}$, we immediately obtain that $\widehat{\mathbb{Z}}$ is finitely generated as $\widehat{\mathbb{Z}}$-module. Moreover, $g^{-1} \cdot \widehat{\Lambda} = \widehat{\mathbb{Z}}^m$ shows that $\mathbb{A}_f \cdot \widehat{\Lambda}$ equals $\mathbb{A}_f^m$. □

Remark 8.9. Another way to phrase (3) above is to say that a $\widehat{\mathbb{Z}}$-lattice in $\mathbb{A}_f^m$ is the same as a family of $\mathbb{Z}_p$-lattices $\Lambda_p \subset \mathbb{Q}_p$, almost all of which are equal to $\mathbb{Z}_p^m$.

We can now come back to the proof of Proposition 8.7.

Proof. Let $K \subset \mathrm{GL}_2(\mathbb{A}_f)$ be a level subgroup. The intersection of two open subsets is open, so $K' = \mathrm{GL}_2(\widehat{\mathbb{Z}}) \cap K$ is again an open subgroup. Moreover, K as the disjoint union $K = \bigsqcup_{k \in K/K'} kK'$ of its K'-cosets. By compactness of K, finitely many such cosets suffice to cover K from which we obtain that K/K' is finite. Let $k_1, \ldots, k_r$ be a set of coset representatives. Then

$$\widehat{\Lambda} := \sum_{i=1}^{r} k_i \cdot \widehat{\mathbb{Z}}^2$$

is a lattice in $\mathbb{A}_f^2$; it is clearly finitely generated as $\widehat{\mathbb{Z}}$-module and satisfies $\mathbb{A}_f \cdot \widehat{\Lambda} = \mathbb{A}_f^2$. It is also clear that K stabilizes $\widehat{\Lambda}$, meaning $K \subseteq \mathrm{GL}(\widehat{\Lambda})$.

By part (4) of Definition 8.8, there exists $g \in \mathrm{GL}_2(\mathbb{A}_f)$ with $\widehat{\Lambda} = g \cdot \widehat{\mathbb{Z}}^2$. This means that $\mathrm{GL}(\widehat{\Lambda}) = g\, \mathrm{GL}_2(\widehat{\mathbb{Z}}) g^{-1}$ and we find

$$g^{-1} K g \subseteq \mathrm{GL}_2(\widehat{\mathbb{Z}}).$$

Conjugation by an element in a topological group is a homeomorphism, so $g^{-1}Kg$ is again open. We know that the principal congruence subgroups in $\mathrm{GL}_2(\mathbb{A}_f)$ form a neighborhood basis of the identity. So there exists $n \geq 1$ with $K(n) \subseteq g^{-1}Kg$ and the proof is complete. $\square$

Construction 8.10. (1) In light of Proposition 8.5, we use the new convention

$$(E, \eta) \cdot g := (E, \eta \circ g^{t,-1})$$

for our action of $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ on $\mathcal{M}_n$. With this renormalization, $\iota : \mathcal{S}_{K(n)} \xrightarrow{\sim} \mathcal{M}_n(\mathbb{C})$ is $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$-equvariant on the nose.

(2) Let us next come back to the ideas around (8.17). Assume that $K \subset \mathrm{GL}_2(\mathbb{A}_f)$ is a level and $g \in \mathrm{GL}_2(\mathbb{A}_f)$ a group element. Then, multiplication by g on the right induces an isomorphism

$$\begin{aligned} g : \mathcal{S}_K &\xrightarrow{\sim} \mathcal{S}_{g^{-1}Kg} \\ [x, hK] &\longmapsto [x, hKg]. \end{aligned}$$

Note that $hKg = hg(g^{-1}Kg)$ which explains why this is a reasonable definition.

(3) Given K, choose g and $n \geq 3$ such that $K(n) \subseteq g^{-1}Kg \subseteq \mathrm{GL}_2(\widehat{\mathbb{Z}})$. Using the normality of $K(n)$ in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$, we can form the quotient group

$$\mathcal{K} \;=\; K(n) \backslash (g^{-1}Kg).$$

Being a subgroup of $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$, it acts on the right of both $\mathcal{S}_{K(n)}$ and $\mathcal{M}_n$. We *define* $\mathcal{M}_K$ as the quotient curve $\mathcal{M}_{n,\mathbb{Q}}/\mathcal{K}$ (see below). The isomorphism $\iota : \mathcal{S}_{K(n)} \xrightarrow{\sim} \mathcal{M}_n$ descends to quotients, and we obtain the terms in the following diagram

$$\begin{tikzcd} & \mathcal{S}_{K(n)} \arrow[r, "\iota"] \arrow[d, "\mathrm{mod}\ \mathcal{K}"'] & \mathcal{M}_{n,\mathbb{Q}}(\mathbb{C}) \arrow[d, "\mathrm{mod}\ \mathcal{K}"] \\ \mathcal{S}_K \arrow[r, "\cdot g"] & \mathcal{S}_{g^{-1}Kg} \arrow[r, dashed] & \mathcal{M}_K(\mathbb{C}). \end{tikzcd} \tag{8.20}$$

We *define* the isomorphism $\mathcal{S}_K \xrightarrow{\sim} \mathcal{M}_K(\mathbb{C})$ as the composition of the bottom two arrows.

We still have to explain how to define the quotient $\mathcal{M}_{n,\mathbb{Q}}/\mathcal{K}$. Let A be a ring and let Γ be a finite group. A group action of Γ on $X = \mathrm{Spec}(A)$ is the same as a group homomorphism $\Gamma \to \mathrm{Aut}_{\mathrm{Scheme}}(X)$. By the equivalence between rings and affine schemes, this is the same as an action of Γ by ring automorphisms on A. Let

$$A^\Gamma = \{a \in A \mid \gamma a = a \text{ for all } \gamma \in \Gamma\}$$

be the ring of invariant elements. The quotient of $\mathrm{Spec}(A)$ by Γ is defined as

$$\Gamma\backslash X := \mathrm{Spec}(A^\Gamma).$$

This construction has the expected universal property in the category of affine schemes. Namely, assume B is a ring and $f : X \to \mathrm{Spec}(B)$ a Γ-invariant map. This means that $f \circ \gamma = f$ for all $\gamma \in \Gamma$. Translated to rings, it means that $\gamma^* \circ f^* = f^*$ for all γ, which says that $f^* : B \to A$ has image contained in A^Γ. In other words, there exists a unique factorization over the quotient as in the diagram

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d] & \mathrm{Spec}(B). \\ \Gamma\backslash X \arrow[ur, dashed, "\bar{f}"'] & \end{tikzcd} \tag{8.21}$$

We now need the fact that $\mathcal{M}_n$ is an affine $\mathbb{Z}[1/n]$-scheme. That is, there exists a finite type $\mathbb{Z}[1/n]$-algebra A_n with $\mathcal{M}_n = \mathrm{Spec}(A_n)$. The $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$-action on $\mathcal{M}_n$ translates into a $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$-action on A_n, and this is how the previous definitions apply.

For a scheme S, we can consider the map $X(S) \to (\Gamma\backslash X)(S)$ induced by $X \to \Gamma\backslash X$. It is Γ-invariant, so we obtain a natural map

$$\Gamma\backslash X(S) \longrightarrow (\Gamma\backslash X)(S).$$

In general, this map is neither injective nor surjective. However, we have the following result for affine varieties.

Proposition 8.11. *Let $X = \mathrm{Spec}(A)$ be an affine finite type scheme over a field k and let Γ be a finite group acting on X. For every algebraically closed extension K/k, the quotient map $X \to \Gamma\backslash X$ defines a bijection*

$$\Gamma\backslash X(K) \xrightarrow{\sim} (\Gamma\backslash X)(K).$$

In the context of (8.20), Proposition 8.11 ensures that

$$\mathcal{M}_n(\mathbb{C})/\mathcal{K} \xrightarrow{\sim} \mathcal{M}_K(\mathbb{C}). \tag{8.22}$$

This defines the bottom dotted arrow in (8.20) and finally completes our construction. One may check that the construction does not depend on n or g up to natural isomorphism.

8.5. **The classical moduli problems.** We wrap up this section with a description of the classical moduli problems.

Example 8.12 (The j-invariant)**.** Consider $\mathcal{S}_{\mathrm{GL}_2(\widehat{\mathbb{Z}})}$ with its Riemann surface structure from (8.19). By (8.8) (lower arrow), its points are in bijection with the isomorphism classes of elliptic curves over $\mathbb{C}$.

We have now also constructed an algebraic curve $\mathcal{M}_{\mathrm{GL}_2(\widehat{\mathbb{Z}}),\mathbb{Q}}$ over $\mathbb{Q}$ as $\mathcal{M}_n/\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ for $n \geq 3$ (choose $n = 3$ for example). The j-invariant (see [10, §11.1]) defines an isomorphism

$$j : \mathcal{M}_{\mathrm{GL}_2(\widehat{\mathbb{Z}}),\mathbb{Q}} \xrightarrow{\sim} \mathbb{A}^1_{\mathbb{Q}}.$$

On a Weierstrass elliptic curve

$$E : y^2 = x^3 + \alpha x + \beta,$$

it is given by

$$j(E) = 1728 \frac{4\alpha^3}{4\alpha^3 + 27\beta^2}. \tag{8.23}$$

This value only depends on the isomorphism class of E and not on the choice of α and β. Proposition 8.11 states that for every algebraically closed field $K/\mathbb{Q}$, the j-invariant provides a bijection

$$j : \begin{Bmatrix} \text{Isomorphism classes of} \\ \text{elliptic curves over } K \end{Bmatrix} \xrightarrow{\sim} K$$
$$E \longmapsto j(E).$$

In fact, this bijection holds for every algebraically closed field K. Except that in characteristics 2 and 3, one needs to define j by a different formula.

Example 8.13 (Level $K_0(n)$)**.** Consider the level subgroup

$$K_0(n) = \left\{ k \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \;\middle|\; k \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod n \right\}.$$

Note that we always have $-1 \in K_0(n)$, so the intersections $\mathrm{GL}_2(\mathbb{Q}) \cap gK_0(n)g^{-1}$ will never be torsion free. Still, (8.19) defines a Riemann surface $\mathcal{S}_{K_0(n)}$ and Construction 8.10 constructs an algebraic curve $\mathcal{M}_{K_0(n),\mathbb{Q}}$ such that $\mathcal{S}_{K_0(n)} \xrightarrow{\sim} \mathcal{M}_{K_0(n),\mathbb{Q}}(\mathbb{C})$.

The quotient group $K_0(n)/K(n) \subset \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ is the stabilizer of the line spanned by $(1,0)^t$ in $(\mathbb{Z}/n\mathbb{Z})^2$. From (8.22), we obtain a bijection

$$\begin{Bmatrix} \text{Isom. classes of} \\ (E, C) \text{ over } K \end{Bmatrix} \xrightarrow{\sim} \mathcal{M}_{K_0(n)}(\mathbb{C})$$

where, on the left hand side, we consider elliptic curves E with a subgroup $C \subset E[n]$ such that C is cyclic of order n.

Example 8.14 (Level $K_1(n)$)**.** Consider the level subgroup

$$K_1(n) = \left\{ k \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \;\middle|\; k \equiv \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \bmod n \right\}.$$

The quotient $K_1(n)/K(n)$ is the stabilizer of the vector $(1,0)^t \in (\mathbb{Z}/n\mathbb{Z})^2$. One obtains from (8.22) that, for every algebraically closed field $K/\mathbb{Q}$,

$$\begin{Bmatrix} \text{Isom. classes of} \\ (E, P) \text{ over } K \end{Bmatrix} \xrightarrow{\sim} \mathcal{M}_{K_1(n)}(\mathbb{C})$$

where, on the left hand side, we consider elliptic curves E with a point $P \in E[n]$ of exact order n.

Observe that the determinant maps $K_0(n), K_1(n) \to \widehat{\mathbb{Z}}^\times$ are surjective. Proposition 3.19 applies and shows that $\mathcal{S}_{K_0(n)}$ and $\mathcal{S}_{K_1(n)}$ are connected. This allows to define them in terms of SL_2 and the upper half plane. The traditional notation in the literature is

$$\mathcal{Y}_0(n) = \mathcal{S}_{K_0(n)} \quad \text{and} \quad \mathcal{Y}_1(n) = \mathcal{S}_{K_1(n)}.$$

Their traditional definition is

$$\mathcal{Y}_0(n) := \Gamma_0(n)\backslash\mathbb{H}^+, \qquad \mathcal{Y}_1(n) := \Gamma_1(n)\backslash\mathbb{H}^+$$

where

$$\Gamma_0(n) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod n \right\}$$

$$\Gamma_1(n) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \bmod n \right\}.$$

Part 2. General Shimura varieties

9. Linear algebraic groups

Throughout this section, k is a field of characteristic 0 and G a connected linear algebraic group over k. Recall that this means that G is a connected, affine, finite type k-group scheme (Theorem 4.12) and that any such G is automatically smooth over k (Theorem 4.8).

Our goal today is to define reductive groups and to discuss their structure theory. We will also see several important examples like Weil restrictions and unitary groups.

Recommended reading: [14, §19a and §19b] which introduces semi-simple and reductive groups.

Unless indicated otherwise, schemes in this section are over k and products are taken over $\mathrm{Spec}(k)$. By group over k we mean an affine finite type k-group scheme. By subgroup, we mean a closed subgroup scheme. The notation $X_{\bar{k}}$ denotes the base change $\bar{k} \otimes_k X$.

9.1. Quotient groups.

Let $H \subseteq G$ be a subgroup. We call H *normal* if the conjugation morphism

$$\begin{aligned} G \times H &\longrightarrow G \\ (g, h) &\longmapsto ghg^{-1} \end{aligned}$$

has image in H. Equivalently, for every point $g \in G(\bar{k})$, we have $g(H_{\bar{k}})g^{-1} = H_{\bar{k}}$ as closed subschemes of $G_{\bar{k}}$. Yet another way to phrase this condition is to say that $H(\bar{k}) \subset G(\bar{k})$ is normal.

Assume that $H \subseteq G$ is normal. Consider the action of H on G by multiplication

$$\begin{aligned} a : H \times G &\longrightarrow G \\ (h, g) &\longmapsto hg. \end{aligned} \tag{9.1}$$

Since G and H are affine, we can write $G = \mathrm{Spec}(A)$ and $H = \mathrm{Spec}(B)$, $B = A/I$, for a finite type k-algebra A and an ideal $I \subseteq A$. Then (9.1) corresponds to a k-algebra homomorphism

$$a^* : B \otimes_k A \longleftarrow A.$$

Definition 9.1 (Quotients, see [14, §5.c])**.** A function $f \in A$ is *H-invariant* if $a^*(f) = 1 \otimes f$. We denote by $A^H \subseteq A$ the subring of invariant functions. The *quotient* of G by H is the k-scheme

$$G/H := \mathrm{Spec}(A^H).$$

This construction has the following properties:

- A^H is again a finite type k-algebra.
- The comultiplication[8] map $m^* : A \to A \otimes_k A$ restricts to a comultiplication on A^H which makes G/H into a k-group.
- The natural map $G \to G/H$ is a group homomorphism. It is flat and surjective with kernel H.
- Every group scheme homomorphism $\pi : G \to Q$ with $H \subseteq \ker(\pi)$ factors over G/H.
- The quotient construction defines a bijection

$$\begin{aligned} \{\text{Normal subgroups } H \subseteq G\} &\longleftrightarrow \{\text{Surjective flat homomorphisms } G \to Q\} \\ H &\longmapsto [G \to G/H] \\ \ker(\pi) &\longleftarrow [G \xrightarrow{\pi} Q]. \end{aligned} \tag{9.2}$$

[8] *Comultiplication* is the name for the map of rings dual to the multiplication $m : G \times G \to G$.

Remark 9.2. An equivalent and more intuitive definition of H-invariant functions is follows. Every $f \in A$ can be viewed as a function on $G(\bar{k})$ with values in $\bar{k}$. Then, f is H-invariant if $f(hg) = f(g)$ for all $h \in H(\bar{k})$ and $g \in G(\bar{k})$.

Example 9.3. Being of finite type over a field, G has finitely many connected components. Let $G^\circ \subseteq G$ be the connected component containing the identity element. Then G° is normal and G/G° is a finite k-group scheme.

Definition 9.4 (Center of G)**.** The *center of* G is the subgroup $Z(G) \subset G$ such that for all k-schemes S, $Z(G)(S)$ is the center of $G(S)$. It can also be characterized as the unique smooth closed subscheme such that $Z(G)(\bar{k})$ is the center of $G(\bar{k})$.

Example 9.5. The center $Z(G) \subseteq G$ is always a normal subgroup. The quotient $G^{\mathrm{ad}} := G/Z(G)$ is called the *adjoint group of* G.

Example 9.6. (1) The center of GL_n is the torus $\mathbb{G}_m$ of diagonal scalar matrices. The center of SL_n is μ_n.

(2) The *projective general linear group* PGL_n is defined as the adjoint group $\mathrm{PGL}_n := \mathrm{GL}_n / \mathbb{G}_m$. If $n = 2$, an explicit description can be obtained as follows. Recall that

$$\mathrm{GL}_2 = \mathrm{Spec}(A), \quad A = k[x_{11}, x_{12}, x_{21}, x_{22}, \delta^{-1}], \quad \delta = x_{11}x_{22} - x_{12}x_{21}.$$

If we write $\mathbb{G}_m = \mathrm{Spec}\, k[t^{\pm 1}]$, then the multiplication action $a : \mathbb{G}_m \times \mathrm{GL}_2 \to \mathrm{GL}_2$ is given by

$$\begin{aligned} a^*(x_{ij}) &= \ t \otimes x_{ij} \\ a^*(\delta^{-1}) &= t^{-2} \otimes \delta. \end{aligned}$$

This means that a monomial $x_{11}^{m_{11}} x_{12}^{m_{12}} x_{21}^{m_{21}} x_{22}^{m_{22}} \delta^{-m}$ is $\mathbb{G}_m$-invariant if and only if

$$2m = m_{11} + m_{12} + m_{21} + m_{22}.$$

We find that

$$\mathrm{PGL}_2 = \mathrm{Spec}(A^{\mathbb{G}_m}), \quad A^{\mathbb{G}_m} = k\left[\frac{x_{ij}x_{k\ell}}{\delta},\ i,j,k,\ell \in \{1,2\}\right].$$

9.2. **Reductive groups.** Over an algebraically closed field of characteristic 0, every linear algebraic group is a successive extension of copies of $\mathbb{G}_a$, $\mathbb{G}_m$ and *semi-simple groups*. In this sense, the semi-simple groups are the interesting building blocks of the theory.

Groups that arise as successive extension of copies of $\mathbb{G}_m$ and semi-simple groups (no copies of $\mathbb{G}_a$) are called *reductive*. They have a particularly nice representation theory: every representation of a reductive group is semi-simple (see below).

Definition 9.7. (1) G is said to be *solvable* if there exists a sequence of subgroups

$$\{1\} = G_0 \subset G_1 \subset G_2 \subset \cdots \subset G_r = G \tag{9.3}$$

such that for all $0 \leq i < r$, G_i is normal in G_{i+1} and G_{i+1}/G_i commutative.

(2) G is said to be *unipotent* if there exists such a sequence with the additional condition that $G_{i+1}/G_i \cong \mathbb{G}_a$ for all $0 \leq i < r$.

Example 9.8. The subgroup $B \subset \mathrm{GL}_n$ of upper triangular matrices is solvable. For example, if $n = 3$ then

$$\{1\} \subset \left\{ \begin{pmatrix} 1 & 0 & * \\ & 1 & 0 \\ & & 1 \end{pmatrix} \right\} \subset \left\{ \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix} \right\} \subset B$$

defines a sequence as in (9.3). The quotients are $\mathbb{G}_m^3$, $\mathbb{G}_a^2$, and $\mathbb{G}_a$, respectively. The subgroup U of unipotent upper triangular matrices is unipotent. For $n = 3$, a composition series with successive quotients $\mathbb{G}_a$ is given by

$$\{1\} \subset \left\{ \begin{pmatrix} 1 & 0 & * \\ & 1 & 0 \\ & & 1 \end{pmatrix} \right\} \subset \left\{ \begin{pmatrix} 1 & 0 & * \\ & 1 & * \\ & & 1 \end{pmatrix} \right\} \subset \left\{ \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix} \right\} = U.$$

Definition 9.9 (Reductive groups, see [14, §19]). (1) There exists a maximal connected, normal, solvable subgroup $R(G) \subseteq G$ called its *radical*. Maximality means that every connected, normal, solvable subgroup $N \subseteq G$ is contained in $R(G)$.

(2) Similarly, there exists a maximal connected, normal, unipotent subgroup $R_u(G) \subseteq G$, called its *unipotent radical*. Thus, we have defined subgroups

$$R_u(G) \;\subseteq\; R(G) \;\subseteq\; G.$$

(3) G is said to be *reductive* if $R_u(G) = \{1\}$. It is said to be semi-simple if $R(G) = \{1\}$.

In general, for a connected group G/k, the quotient $G/R_u(G)$ will be reductive and $G/R(G)$ will be semi-simple.

Proposition 9.10 (Centers of reductive groups). *Let G be a connected reductive group over k. Then the radical of G agrees with the identity component of its center. That is, $R(G) = Z(G)^\circ$. In particular, G is semi-simple if and only if its center is finite.*

Example 9.11. Consider the group $G = \mathrm{SL}_2$ over k. One may check that $\mathrm{SL}_2(\bar{k})$ has no proper normal subgroups except for the center $\{\pm 1\}$. So we see that $R_u(G) = R(G) = \{1\}$. Note that all conditions in the definitions of $R_u(G)$ and $R(G)$ are important:

(1) The diagonal subgroup $\{\pm 1\}$ is normal and commutative, but not connected.

(2) The subgroup $\left\{ \left(\begin{smallmatrix} 1 & b \\ & 1 \end{smallmatrix}\right) \right\}$ of upper triangular unipotent matrices is unipotent and connected, but not normal.

Remark 9.12. Under our standing assumption that $\mathrm{char}(k) = 0$, all notions introduced so far are compatible with scalar extension to $\bar{k}$. That is, G is solvable/unipotent/semi-simple/reductive if and only if $G_{\bar{k}}$ has that property. Moreover,

$$R(G_{\bar{k}}) = R(G)_{\bar{k}}, \quad R_u(G_{\bar{k}}) = R_u(G)_{\bar{k}}.$$

Reductive groups can be characterized in terms of their representation theory. Let V be a finite-dimensional k-vector space. The k-group $\mathrm{GL}(V)$ is defined as functor on k-algebras by

$$\mathrm{GL}(V)(R) := \mathrm{GL}_R(R \otimes_k V). \tag{9.4}$$

Any choice of basis for V identifies it $\mathrm{GL}_{n,k}$. A *representation* of a k-group G on V is a homomorphism of k-group schemes

$$\rho : G \longrightarrow \mathrm{GL}(V).$$

A representation is said to be *semi-simple* if, for every G-stable subspace $W \subset V$, there exists a G-stable subspace U with $V = W \oplus U$.

Example 9.13. Consider the inclusion map

$$U := \left\{ \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix} \right\} \hookrightarrow \mathrm{GL}_2$$

viewed as representation of U on k^2. Clearly, the line $\ell := k \cdot \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ is U-stable. We claim that it has no complementary U-stable line. Indeed, consider a vector $v = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ with $y \neq 0$. Then v together with

$$\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} \cdot v = \begin{pmatrix} x+y \\ y \end{pmatrix}$$

already generate k^2. This shows that k^2 is not a semi-simple representation of U.

Proposition 9.14 (see [14, Corollary 19.18] for (2))**.** *(1) Every representation of a reductive group is semi-simple.*

(2) A connected k-group G is reductive if and only if there exists an injective semi-simple representation $\rho : G \hookrightarrow \mathrm{GL}(V)$.

Example 9.15. The groups GL_n, SL_n, GSp_{2g} and Sp_{2g} are all reductive.

Let $(V, (\ ,\ ))$ be a finite-dimensional k-vector space together with a non-degenerate symmetric bilinear form. If R is a k-algebra, then we obtain a symmetric bilinear form

$$(\ ,\ )_R : (R \otimes_k V) \times (R \otimes_k V) \longrightarrow R$$

defined on generators by

$$(r_1 \otimes v_1, r_2 \otimes v_2)_R := r_1 r_2 (v_1, v_2).$$

The special orthogonal group is the k-group with functor of points

$$\mathrm{SO}(V)(R) := \big\{ g \in \mathrm{SL}_R(R \otimes_k V) \ \mid \ (gx, gy)_R = (x, y)_R \text{ for all } x, y \in R \otimes_k V \big\}. \tag{9.5}$$

It is a closed subgroup of $\mathrm{GL}(V)$. Just like the groups listed before, it is reductive.

Proof. All listed groups are connected and their standard representation is simple, so Proposition 9.14 applies. □

9.3. **Weil restriction.** We next explain a construction called Weil restriction. Given a finite extension K/k and a K-group H, it allows to define k-group $\mathrm{Res}_{K/k}(H)$ which is "the same as H but viewed over k". It is worth noting that the construction will be very different from just composing the structure map $H \to \mathrm{Spec}(K)$ with $\mathrm{Spec}(K) \to \mathrm{Spec}(k)$. In fact, viewing H over k by this naive procedure would not allow to endow it with a k-group structure. For example, the identity element of H is a K-rational point (and not a k-rational one).

Proposition 9.16. *Let K/k be a finite field extension (or product of such) and let X be an affine, finite type K-scheme. Define a functor on k-algebras by*

$$(\mathrm{Res}_{K/k} X)(R) := X(K \otimes_k R). \tag{9.6}$$

Then $\mathrm{Res}_{K/k} X$ is representable by an affine finite type k-scheme which is called the Weil restriction *of X. The dimensions are related by*

$$\dim(\mathrm{Res}_{K/k} X) = [K : k] \cdot \dim(X).$$

If X is a linear algebraic group over K, then $\mathrm{Res}_{K/k} X$ is naturally a linear algebraic group over k.

Proof. Step 1: The basic construction. Consider first the case of the affine line $X = \mathrm{Spec}\, K[t]$. To define $\mathrm{Res}_{K/k}(X)$ as affine scheme, we are looking for a k-algebra B together with isomorphisms

$$\mathrm{Hom}_k(B, R) \xrightarrow{\sim} \mathrm{Hom}_K(K[t], K \otimes_k R) \tag{9.7}$$

for all k-algebras R, functorially in R. Let $d = [K : k]$ be the degree of K/k and let $\alpha_1, \ldots, \alpha_d$ be a k-basis for K. Giving φ on the right hand side of (9.7) is the same as giving its value $\varphi(t) = \sum_{j=1}^{d} \alpha_j \otimes r_j$ in $K \otimes_k R$. Thus, if we define $B = k[y_1, \ldots, y_d]$, we can construct (9.7) by

$$\begin{aligned} \mathrm{Hom}_k(k[y_1, \ldots, y_d], R) &\xrightarrow{\sim} \mathrm{Hom}_K(K[t], K \otimes_k R) \\ \psi &\longmapsto \varphi(t) := \alpha_1 \otimes \psi(y_1) + \ldots + \alpha_d \otimes \psi(y_d). \end{aligned} \tag{9.8}$$

This shows that $\mathrm{Res}_{K/k}(\mathbb{A}^1_K)$ is representable by $\mathbb{A}^d_k$. By taking products, we constructed an isomorphism

$$\mathbb{A}^{d\cdot n}_k \xrightarrow{\sim} \mathrm{Res}_{K/k}(\mathbb{A}^n_K). \tag{9.9}$$

Step 2: Closed subschemes. Assume now that $X = V(f_1, \ldots, f_m)$ is a closed subscheme of $\mathbb{A}^n_K$. We claim that $\mathrm{Res}_{K/k}(X)$ is a closed subscheme of $\mathrm{Res}_{K/k}(\mathbb{A}^n_K)$, and we are going to exhibit the corresponding equations in $\mathbb{A}^{d\cdot n}_k$ under (9.9).

Let $t_1, \ldots, t_n$ be the coordinates on $\mathbb{A}^n_K$ and let y_{ij}, $1 \le i \le n$, $1 \le j \le d$ denote the corresponding coordinates on $\mathbb{A}^{d\cdot n}_k$. For each $1 \le k \le m$, we substitute $\alpha_1 \otimes y_{i1} + \ldots + \alpha_d \otimes y_{id}$ for t_i to obtain an element in $K \otimes_k k[y_{11}, \ldots, y_{nd}]$:

$$f_k\Big(\sum_{j=1}^{d} \alpha_j \otimes y_{1j},\ \sum_{j=1}^{d} \alpha_j \otimes y_{2j},\ \ldots,\ \sum_{j=1}^{d} \alpha_j \otimes y_{nj}\Big) = \sum_{j=1}^{d} \alpha_j \otimes g_{kj}(y_{11}, \ldots, y_{nd}).$$

Then, if we look at the higher-dimensional variant of (9.8),

$$\begin{aligned} \mathrm{Hom}_k(k[y_{11}, \ldots, y_{nd}], R) &\xrightarrow{\sim} \mathrm{Hom}_K(K[t_1, \ldots, t_n], K \otimes_k R) \\ \psi &\longmapsto \varphi(t_i) := \alpha_1 \otimes \psi(y_{i1}) + \ldots + \alpha_d \otimes \psi(y_{id}), \end{aligned}$$

we find that $\varphi(f_k) = 0$ if and only if $\psi(g_{k1}) = \ldots = \psi(g_{kd}) = 0$. This proves that

$$V(g_{11}, \ldots, g_{md}) \xrightarrow{\sim} \mathrm{Res}_{K/k}\big(V(f_1, \ldots, f_k)\big)$$

which concludes the proof of representability for $\mathrm{Res}_{K/k}(X)$.

Wrap-up: Group structures. It is clear from the functor of points description (9.6) that Weil restriction is functorial and compatible with products. In particular, if we are given a K-group structure $m : X \times_{\mathrm{Spec}(K)} X \to X$, then we obtain a multiplication

$$\mathrm{Res}_{K/k}(X) \times_{\mathrm{Spec}(k)} \mathrm{Res}_{K/k}(X) \longrightarrow \mathrm{Res}_{K/k}(X).$$

The claim about dimensions (at least when k is of characteristic 0) follows from the next proposition. $\square$

Let K/k be a finite separable extension or product of such extensions. For every field extension L/k, we have a natural map

$$K \otimes_k L \longrightarrow \prod_{\varphi \in \mathrm{Hom}_k(K,L)} L$$

which is $a \otimes b \mapsto \varphi(a)b$ in the φ-component. We say that L is a splitting field for K if this map is an isomorphism. This is equivalent to $\mathrm{Hom}_k(K, L)$ having cardinality $[K : k]$. For example, K is a splitting field for K if and only if K is Galois.

Proposition 9.17. *Assume that K/k is a product of finite separable field extensions and L/k a splitting field for k. Then, for X/K as before,*

$$L \otimes_k \mathrm{Res}_{K/k}(X) \xrightarrow{\sim} \prod_{\varphi \in \mathrm{Hom}_k(K,L)} L \otimes_{\varphi, K} X.$$

Proof. Let R be an L-algebra. We write $R|_k$ or $R|_{\varphi,K}$ if we view R as k algebra or as K algebra via $\varphi : K \to L$. Observe that

$$\begin{aligned} K \otimes_k R &= (K \otimes_k L) \otimes_L R \\ &\xrightarrow{\sim} \prod_{\varphi:K\to L} R|_{\varphi,K} \end{aligned}$$

as K-algebras. By definition of the Weil restriction, we have

$$\begin{aligned} (L \otimes_k \mathrm{Res}_{K/k} X)(R) &= (\mathrm{Res}_{K/k} X) X(R|_k) \\ &= X(K \otimes_k R|_k) \\ &\xrightarrow{\sim} \prod_{\varphi:K\to L} X(R|_{\varphi,K}) \\ &= \prod_{\varphi:K\to L} (L \otimes_{\varphi,K} X)(R). \end{aligned}$$

□

Example 9.18 (Deligne torus)**.** The Weil restriction $\mathbb{S} := \mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_m)$ is called the *Deligne torus.* The proof of Proposition 9.16 explains that we can explicitly describe it as follows. First, we have the presentation

$$\mathbb{G}_{m,\mathbb{C}} = \mathrm{Spec}\,\mathbb{C}[x,y]/(f), \quad f = xy - 1.$$

Now we introduce new variables x_1, x_2, y_1, y_2 and substitute

$$x = x_1 + x_2 \cdot i, \quad y = y_1 + y_2 \cdot i$$

into f which yields

$$\underbrace{(x_1y_1 - x_2y_2 - 1)}_{g_1} + \underbrace{(x_1y_2 + x_2y_1)}_{g_2} \cdot i$$

and hence gives

$$\mathbb{S} \;\cong\; \mathrm{Spec}\,A, \quad A = \mathbb{R}[x_1, x_2, y_1, y_2]/(g_1, g_2). \tag{9.10}$$

The group law is given by the multiplication rule for complex numbers of the form $x_1 + x_2 \cdot i$ and $y_1 + y_2 \cdot i$. That is,

$$\begin{aligned} m^*(x_1) &= x_1 \otimes x_1 - x_2 \otimes x_2 \\ m^*(x_2) &= x_1 \otimes x_2 + x_2 \otimes x_1 \end{aligned}$$

and analogously for y_1 and y_2. The equations g_1 and g_2 precisely encode that

$$(x_1 + x_2 \cdot i)(y_1 + y_2 \cdot i) = 1. \tag{9.11}$$

In the ring A, we check that (exercise)

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = 1$$

as we would expect from identity (9.11) for actual complex numbers. We also check

$$\begin{aligned} x_1 &= y_1(x_1^2 + x_2^2) \\ x_2 &= -y_2(x_1^2 + x_2^2). \end{aligned}$$

In this way, we find

$$\begin{aligned} \mathbb{R}[x_1, x_2, (x_1^2 + x_2^2)^{-1}] &\xrightarrow{\sim} A \\ x_i &\longleftrightarrow x_i \\ (-1)^{i+1} x_i (x_1^2 + x_2^2)^{-1} &\longleftarrow y_i. \end{aligned}$$

In other words, for every $\mathbb{R}$-algebra R, we have

$$\mathbb{S}(R) = \{x_1 + x_2 \cdot i \;\mid\; x_1, x_2 \in R,\; x_1^2 + x_2^2 \in R^\times\} \tag{9.12}$$

where i is a symbol with $i^2 = -1$. The multiplication rules are the same as in $\mathbb{C} = \mathbb{R} \oplus \mathbb{R} \cdot i$.

9.4. **Unit groups of algebras.** A more compact way of writing (9.12) is as

$$\mathbb{S}(R) = (\mathbb{C} \otimes_{\mathbb{R}} R)^\times.$$

In this interpretation, it becomes an example of an interesting more general construction. Let $\mathcal{A}$ be a (not necessarily commutative) finite-dimensional k-algebra. Then we define a functor on k-algebras by

$$\underline{\mathcal{A}}^\times(R) := (R \otimes_k \mathcal{A})^\times. \tag{9.13}$$

This is a group valued functor by multiplication in $R \otimes_k \mathcal{A}$.

Proposition 9.19. *The functor $\underline{\mathcal{A}}^\times$ is representable by an affine finite type k-scheme.*

Proof. Choose a k-basis $\alpha_1, \dots, \alpha_d$ for $\mathcal{A}$. The functor $\underline{\mathcal{A}}(R) = R \otimes_k \mathcal{A}$ is representable by $\mathbb{A}^d_k$ via

$$\begin{aligned} \mathbb{A}^d_k(R) &\xrightarrow{\sim} \underline{\mathcal{A}}(R) \\ (r_1, \dots, r_d) &\longmapsto \sum_{i=1}^d r_i \otimes \alpha_i. \end{aligned} \tag{9.14}$$

Our aim is to describe the subfunctor $\underline{\mathcal{A}}^\times \subset \underline{\mathcal{A}}$. We recall the Cayley–Hamilton theorem over a general ring.

Proposition 9.20 (Cayley–Hamilton)**.** *Let R be a ring and let $f \in \mathrm{End}_R(R^d)$ be an endomorphism with characteristic polynomial $P(T) \in R[T]$. Then $P(f) = 0$.*

Corollary 9.21. *Let $T^d + a_{d-1}T^{d-1} + \ldots + a_0$ be the characteristic polynomial of $f \in \mathrm{End}_R(R^d)$. Then f is invertible if and only if $a_0 = \det(f) \in R^\times$, in which case*

$$f^{-1} = -a_0^{-1}\big(f^{d-1} + a_{d-1}f^{d-2} + \ldots + a_1\big).$$

Let us come back to our proof. An element $f \in R \otimes_k \mathcal{A}$ acts by left multiplication on $R \otimes_k \mathcal{A}$ which is a free R-module of rank d. So we may consider its determinant $\det_R(f) \in R$. For example, we may use the basis $1 \otimes \alpha_1, \dots, 1 \otimes \alpha_d$ to first express f as a $(d \times d)$-matrix and then take the determinant. This does not depend on the choice of basis. Coming back to our proof, we have the following application.

Corollary 9.22. *An element $f \in R \otimes_k \mathcal{A}$ lies in $(R \otimes_k \mathcal{A})^\times$ if and only if $\det_R(f) \in R^\times$.*

Proof. If $\det_R(f) \in R^\times$, then $f \in \mathrm{GL}_R(R \otimes_k \mathcal{A})$. By Corollary 9.21, f^{-1} is a polynomial in f, hence again lies in $R \otimes_k \mathcal{A}$. $\square$

So far, we have established that $\underline{\mathcal{A}}^\times(R) \subseteq \underline{\mathcal{A}}(R)$ is the subset of elements whose determinant lies in $R^\times$. We claim that this condition is defined by a principal open subset $D(\det) \subset \underline{\mathcal{A}}$ for a morphism $\det : \mathcal{A} \to \mathbb{A}^1_k$.

In the basis $\alpha_1, \dots, \alpha_d$, multiplication by α_i is given by a $(d \times d)$-matrix $A_i \in M_d(k)$. Then

$$\begin{aligned} \det{}_R\Big(\sum\nolimits_{i=1}^d r_i \otimes \alpha_i\Big) &= \det\Big(\sum\nolimits_{i=1}^d r_i \cdot A_i\Big) \\ &= \delta(r_1, \dots, r_d) \end{aligned}$$

for the (degree d, homogeneous) polynomial

$$\delta(x_1, \dots, x_d) = \det(x_1 A_1 + \ldots + x_d A_d) \in k[x_1, \dots, x_d].$$

This polynomial defines the morphism $\det : \underline{\mathcal{A}} \to \mathbb{A}^1_k$ in terms of the coordinates from (9.14). The proof of Proposition 9.19 is now complete. $\square$

Example 9.23. (1) Let K/k be a finite-dimensional commutative k-algebra (e.g. a field extension). For every k-algebra R, we have

$$\begin{aligned}\mathrm{Res}_{K/k}(\mathbb{G}_{m,K})(R) &= \mathbb{G}_{m,K}(K \otimes_k R) \\ &= (K \otimes_k R)^\times \\ &= \underline{K}^\times(R).\end{aligned}$$

So we see that $\underline{K}^\times = \mathrm{Res}_{K/k}(\mathbb{G}_{m,K})$.

(2) Let us come back to the Deligne torus. We choose the $\mathbb{R}$-basis $\alpha_1 = 1$, $\alpha_2 = i$ of $\mathbb{C}$. In this basis, left multiplication with α_1 and α_2 is given by

$$A_1 = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$$

and

$$\det(x_1 A_1 + x_2 A_2) = \det \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix} = x_1^2 + x_2^2.$$

The description of $\underline{\mathcal{A}}^\times$ during the proof of Proposition 9.19 now directly gives

$$\mathbb{S} = \mathrm{Spec}\, \mathbb{R}[x_1, x_2, (x_1^2 + x_2^2)^{-1}].$$

(3) Let V be a finite-dimensional k-vector space and consider $\mathcal{A} = \mathrm{End}_k(V)$. Then

$$\begin{aligned}\underline{\mathcal{A}}^\times(R) &= (R \otimes_k \mathrm{End}_k(V))^\times \\ &= \mathrm{End}_R(R \otimes_k V)^\times \\ &= \mathrm{GL}_R(R \otimes_k V).\end{aligned}$$

This group scheme is called $\mathrm{GL}(V)$. It is a coordinate independent variant of $\mathrm{GL}_{\dim(V)}$.

(4) Let K/k be a field extension and V a finite-dimensional K-vector space. Then, by definitions,

$$\mathrm{Res}_{K/k} \mathrm{GL}(V) = \underline{\mathrm{End}_K(V)}^\times$$

where $\mathrm{End}_K(V)$ is viewed as k-algebra. In general, if L/k is a field extension, then we immediately see

$$\underline{L \otimes_k \mathcal{A}}^\times = L \otimes_k \underline{\mathcal{A}}^\times.$$

So, if K/k is separable and if L splits K, then we obtain

$$\begin{aligned}L \otimes_k \mathrm{Res}_{K/k} \mathrm{GL}(V) &= \underline{L \otimes_k \mathrm{End}_K(V)}^\times \\ &= \prod_{\varphi : K \to L} \underline{\mathrm{End}_L(L \otimes_{\varphi, K} V)}^\times \\ &\xrightarrow{\sim} (\mathrm{GL}_{\dim(V)})^{[K:k]}\end{aligned}$$

which is in line with Proposition 9.17.

10. Shimura data

Our aim in this section is to explain the following definition and to give several examples.

Definition 10.1 (Shimura datum, [13, Definition 5.5]). A Shimura datum is a pair (G, X) consisting of a connected reductive group $G/\mathbb{Q}$ and a $G(\mathbb{R})$-conjugacy class X of homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$ satisfying the following three axioms. For all $h \in X$:

(SV 1) The image $h(\mathbb{G}_{m,\mathbb{R}})$ is contained in $Z(G_{\mathbb{R}})$. In particular, h mod $Z(G_{\mathbb{R}})$ factors over a homomorphism $\bar{h} : U(1) \to G_{\mathbb{R}}^{\mathrm{ad}}$,

$$\begin{tikzcd} \mathbb{S} \arrow[r, "h"] \arrow[d, "z \mapsto z\bar{z}"'] & G_{\mathbb{R}} \arrow[d] \\ U(1) \arrow[r, "\bar{h}"] & G_{\mathbb{R}}^{\mathrm{ad}}. \end{tikzcd}$$

We require that only the weights $-1, 0$ and 1 occur in the representation $\mathrm{ad} \circ \bar{h}$ of $U(1)$ on $\mathrm{Lie}(G_{\mathbb{R}})$.

(SV 2) Conjugation by $h(i)$ defines a Cartan involution θ on G^{ad}. That is, the Lie group

$$G^{\mathrm{ad},(\theta)}(\mathbb{R}) := \{g \in G^{\mathrm{ad}}(\mathbb{C}) \mid g = h(i) \cdot \bar{g} \cdot h(i)^{-1}\}$$

is compact.

(SV 3) There exists no product decomposition $G^{\mathrm{ad}} = G_1 \times G_2$ with both $G_i \neq \{1\}$ such that one of the components of $\mathrm{ad} \circ h$ is trivial.

It is worth pointing out that if one element $h \in X$ satisfies (SV 1)–(SV 3), then all such elements do. So one could have phrased the definition in terms of a single homomorphism h, and then taken X as its $G(\mathbb{R})$-conjugacy class. As in Milne [13], the formulation here avoided endowing X with a distinguished point.

Axiom (SV 3) is a non-triviality condition. Namely, if $G(\mathbb{R})$ is compact, then any homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ satisfying (SV 2) is central. So we would always get $X = \{\mathrm{pt}\}$. Excluding factors of this form with (SV 3) matters in Deligne's definition of canonical models. However, we will not discuss this aspect any further.

10.1. **The adjoint representation and** $\mathrm{ad}\circ h$**.** We begin with a discussion of Lie algebras and the adjoint representation. Any finite type group scheme G/k has a Lie algebra defined by

$$\mathrm{Lie}(G) := \ker\Big(G\big(k[\varepsilon]/(\varepsilon^2)\big) \longrightarrow G(k)\Big).$$

This is an abelian group even when G is non-commutative.

Example 10.2. (1) The Lie algebra of GL_n is the additive group of $(n \times n)$-matrices,

$$\begin{aligned} \mathrm{Lie}(\mathrm{GL}_n) &= \big(1 + \varepsilon \cdot \mathrm{M}_n(k),\ \mathrm{mult}\big) \\ &\xrightarrow{\sim} \big(\mathrm{M}_n(k),\ \mathrm{add}\big). \end{aligned}$$

The point is that

$$(1 + \varepsilon X)(1 + \varepsilon Y) \equiv 1 + \varepsilon(X + Y) \mod (\varepsilon),$$

so multiplication translates to addition.

(2) The determinant of $1 + \varepsilon X \in \mathrm{M}_n(k[\varepsilon]/(\varepsilon^2))$ is $1 + \varepsilon\,\mathrm{tr}(X)$. So we obtain

$$\mathrm{Lie}(\mathrm{SL}_n) = \mathrm{M}_n(k)^{\mathrm{tr}=0}.$$

(3) For a general linear algebraic group G, we can always find some n and an embedding $\rho : G \to \mathrm{GL}_n$ (Theorem 4.12) and obtain

$$\mathrm{Lie}(G) = \mathrm{Lie}(\mathrm{GL}_n) \cap \rho(G(k[\varepsilon]/(\varepsilon^2))).$$

For example, consider $\mathrm{SO}(k^n, 1_n)$ for a field k with $\mathrm{char}(k) \neq 2$. For vectors $v, w \in (k[\varepsilon]/(\varepsilon^2))^n$ and $X \in \mathrm{M}_n(k)$, we have

$$\begin{aligned}\big((1+\varepsilon X)v,\ (1+\varepsilon X)w\big) &= v^t(1+\varepsilon X)^t(1+\varepsilon X)w \\ &= v^t w + \varepsilon \cdot v^t(X^t + X)w.\end{aligned}$$

This equals $(v, w) = v^t w$ for all v and w if and only if $X^t + X = 0$. So we see that $\mathrm{Lie}(SO(k^n, 1_n)) \subset \mathrm{M}_n(k)$ is the subset of skew-symmetric matrices $X^t = -X$.

For every $\lambda \in k$, there is a scaling map

$$k[\varepsilon]/(\varepsilon^2) \xrightarrow{\sim} k[\varepsilon]/(\varepsilon^2), \quad \varepsilon \longmapsto \lambda\varepsilon.$$

Composition with these maps defines an action of k on $\mathrm{Lie}(G)$ making it into a k-vector space. In the above examples, this is simply the usual k-vector space structure on $\mathrm{M}_n(k)$.

The Lie algebra is also the same as the tangent space to G at 1. In our upcoming applications, we will be in the case $\mathrm{char}(k) = 0$. Then G is necessarily smooth and we have $\dim_k \mathrm{Lie}(G) = \dim(G)$. If $k = \mathbb{R}$ or $\mathbb{C}$, then this algebraic definition agrees with the definition in terms of the tangent space at 1 of the Lie group $G(\mathbb{R})$ (resp. $G(\mathbb{C})$).

Definition 10.3. In general, G acts on $\mathrm{Lie}(G)$ by conjugation. Conjugation by central elements is trivial, so this action factors over the adjoint group. That is, we obtain a representation

$$\mathrm{ad} : G \longrightarrow G^{\mathrm{ad}} \hookrightarrow \mathrm{GL}(\mathrm{Lie}(G))$$

called the *adjoint representation* of G.

Example 10.4. The adjoint representation of GL_n on $\mathrm{Lie}(\mathrm{GL}_n)$ is given by the usual conjugation action of GL_n on $\mathrm{M}_n(k)$. For a general linear algebraic group G, we may always choose an embedding $\rho : G \to \mathrm{GL}_n$ to realize $\mathrm{Lie}(G)$ as a subspace of $\mathrm{M}_n(k)$. Then G acts by conjugation via ρ.

We next explain the representation $\mathrm{ad} \circ \bar{h}$ that occurs in (SV 1). The determinant construction during the proof of Proposition 9.19 restricts to a group homomorphism $\underline{\mathcal{A}}^\times \to \mathbb{G}_m$. Applying this to the Deligne torus $\mathbb{S} = \underline{\mathbb{C}}^\times$ defines an exact sequence

$$1 \longrightarrow U(1) \longrightarrow \mathbb{S} \xrightarrow{N_{\mathbb{C}/\mathbb{R}}} \mathbb{G}_{m,\mathbb{R}} \longrightarrow 1 \tag{10.1}$$

where

$$N_{\mathbb{C}/\mathbb{R}}(x_1 + x_2 \cdot i) = x_1^2 + x_2^2$$

is the norm map from $\mathbb{C}$ to $\mathbb{R}$, but viewed as algebraic morphism $\mathbb{A}^2_{\mathbb{R}} \to \mathbb{A}^1_{\mathbb{R}}$. The kernel $U(1)$ is the $\mathbb{R}$-group scheme unit circle

$$U(1) = \mathrm{Spec}\, \mathbb{R}[x_1, x_2]/(x_1^2 + x_2^2 - 1).$$

Its $\mathbb{R}$-points $U(1)(\mathbb{R})$ are literally the unit circle in $\mathbb{C}^\times$. In addition to the exact sequence (10.1), we also have the sequence

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathbb{G}_m & \longrightarrow & \mathbb{S} & \longrightarrow & U(1) & \longrightarrow & 1 \\ & & & & z & \longmapsto & z/\bar{z}. & & \end{array} \tag{10.2}$$

Note that the composition

$$U(1) \hookrightarrow \mathbb{S} \xrightarrow{N_{\mathbb{C}/\mathbb{R}}} U(1)$$

is $z \mapsto z^2$. The two exact sequences express that $\mathbb{S}$ is almost the product of $\mathbb{G}_{m,\mathbb{R}}$ and $U(1)$, namely

$$\big(\mathbb{G}_{m,\mathbb{R}} \times U(1)\big)/\Delta(\pm 1) \xrightarrow{\sim} \mathbb{S}, \quad (t, z) \longmapsto t \cdot z.$$

Consider now a homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ as in Definition 10.1. If $h(\mathbb{G}_m) \subset Z(G_{\mathbb{R}})$, then the composition $h : \mathbb{S} \to G \to G^{\mathrm{ad}}$ factors through the quotient $U(1)$ in (10.2). If we compose with the adjoint representation, we obtain a representation

$$\mathrm{ad} \circ \bar{h} : U(1) \longrightarrow \mathrm{GL}(\mathrm{Lie}(G_{\mathbb{R}})), \quad z \longmapsto \mathrm{ad}(h(\sqrt{z})). \tag{10.3}$$

Here, the notation $\sqrt{z}$ means pick any inverse image of z under $U(1) \to U(1)$, $z \mapsto z^2$. The second part of (SV 1) is a condition on (10.3) which we explain next.

10.2. **Representation theory of $U(1)$.** We need a classification of the representations of $U(1)$. Let us first note that there is no difference between the representation theory of $U(1)$ as algebraic group and that of $\mathcal{U} := U(1)(\mathbb{R})$ as real Lie group. Every finite-dimensional algebraic representation $\rho : U(1) \to \mathrm{GL}(V)$ gives rise (pass to $\mathbb{R}$-points) to a representations as real Lie group

$$\rho(\mathbb{R}) : \mathcal{U} \longrightarrow \mathrm{GL}(V)(\mathbb{R}).$$

This functor defines an equivalence of categories

$$\mathrm{Rep}^{\mathrm{alg}}(U(1)) \xrightarrow{\sim} \mathrm{Rep}^{\text{Lie group}}(\mathcal{U}).$$

We will work with $\mathcal{U}$ to prove our classification theorem.

Let k be a non-zero integer. Define a two-dimensional real representation (W_k, ρ_k) of $\mathcal{U}$ as $W_k = \mathbb{C}$ and

$$\rho_k : \mathcal{U} \to \mathrm{GL}_{\mathbb{R}}(W_k), \quad \rho_k(z)(v) = z^k v.$$

It is clear that W_k has no $\mathcal{U}$-stable lines and is hence irreducible. It is also clear that the complex conjugation defines an isomorphism

$$W_k \xrightarrow{\sim} W_{-k}, \quad x_1 + x_2 \cdot i \longmapsto x_1 - x_2 \cdot i.$$

Moreover, $W_k \not\cong W_{k'}$ if $k \neq \pm k'$ because the trace of $\rho_k(z)$ is $\mathrm{Re}(z^k)$ and

$$\mathrm{Re}\big(z^k\big) = \mathrm{Re}\big(z^{k'}\big) \text{ for all } z \in \mathcal{U} \quad \Longleftrightarrow \quad k = \pm k'.$$

Finally, we also define $W_0 = \mathbb{R}$ (trivial representation).

Proposition 10.5. *(1) Every finite-dimensional representation of $U(1)$ is a direct sum of irreducible representations.*

(2) Every irreducible such representation is isomorphic to precisely one of the W_k, $k \geq 0$.

Proof. Decomposition into irreducibles. The following argument works more generally for compact Lie groups (Peter–Weyl Theorem). Let V be a finite-dimensional continuous representation of $\mathcal{U}$. If V is not already irreducible, then there exists a proper $\mathcal{U}$-stable subspace $0 \neq W \subset V$. Choose a projection $f_0 : V \twoheadrightarrow W$, meaning a surjection such that $f_0|_W = \mathrm{id}_W$. Define $f : V \to W$ by

$$f(v) = \int_{\mathcal{U}} z^{-1} \cdot f_0(z \cdot v)\ dz$$

where dz is the translation invariant measure on $\mathcal{U}$ of total volume 1. We still have $f|_W = \mathrm{id}$, because

$$\begin{aligned} f(w) &= \int_{\mathcal{U}} z^{-1} \cdot f_0(z \cdot w)\ dz \\ &= \int_{\mathcal{U}} z^{-1} \cdot z \cdot w\ dz \\ &= w. \end{aligned}$$

This shows that f is still a projection to W and hence that $V = W \oplus \ker(f)$. The translation invariance of the measure moreover implies that

$$\begin{aligned} f(gv) &= \int_{\mathcal{U}} z^{-1} f_0(z \cdot gv)\, dz \\ &= \int_{\mathcal{U}} (zg^{-1})^{-1} \cdot f_0((zg^{-1}) \cdot gv)\, dz \\ &= \int_{\mathcal{U}} g \cdot z^{-1} f_0(zv)\, dz \\ &= gf(v), \end{aligned}$$

which shows that f is $\mathcal{U}$-equivariant. So $\ker(f)$ is a $\mathcal{U}$-stable subspace and $V = W \oplus \ker(f)$ is an $\mathcal{U}$-stable direct sum decomposition. By induction on $\dim(V)$, we have proved that V is a direct sum of irreducible representations.

Classification of irreducibles over $\mathbb{C}$. We first consider finite-dimensional complex representations of $\mathcal{U}$. The first part of the proof still applies and shows that every such representation is a direct sum of irreducible ones. Let $\rho : \mathcal{U} \to \mathrm{GL}_{\mathbb{C}}(V)$ be irreducible. The operators $\{\rho(z)\}_{z \in \mathcal{U}}$ all pairwise commute because $\mathcal{U}$ is commutative. So there exists a joint eigenvector $0 \neq v \in V$,

$$\rho(z)(v) = \chi(z)v,$$

with eigenvalues $\chi(z) \in \mathbb{C}$. In particular, $\mathbb{C} \cdot v \subset V$ is $\mathcal{U}$-stable. Since V is irreducible by assumption, $V = \mathbb{C} \cdot v$. In other words, V is defined by a character $\chi : \mathcal{U} \to \mathbb{C}^\times$.

Recall that we are always considering continuous representations. So $\chi(\mathcal{U}) \subset \mathbb{C}^\times$ is a compact subgroup. (Images of compact sets under continuous maps are compact.) Composing with the absolute value map $| \cdot | : \mathbb{C}^\times \to \mathbb{R}_{>0}$, we similarly see that $|\chi(\mathcal{U})| \subset \mathbb{R}_{>0}$ is compact. But the only bounded subgroup of $\mathbb{R}_{>0}$ is $\{1\}$, meaning that $\chi(\mathcal{U}) \subseteq \mathcal{U}$. The characters $\mathcal{U} \to \mathcal{U}$ are precisely the group homomorphisms $\chi_k(z) = z^k$, $k \in \mathbb{Z}$. Taking all these arguments together, we have shown that every complex finite-dimensional representation (V, ρ) has a direct sum decomposition

$$V = \bigoplus_{k \in \mathbb{Z}} V_k \tag{10.4}$$

where V_k is the subspace of v with $\rho(z)(v) = \chi_k(z)v$, and each V_k is isomorphic to $\chi_k^{\oplus \dim(V_k)}$.

Classification of irreducibles over $\mathbb{R}$. Let V be an irreducible $\mathcal{U}$-representation over $\mathbb{R}$ and let $V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V$ viewed as representation over $\mathbb{C}$. There is a complex conjugation on $V_{\mathbb{C}}$ by $\sigma(x \otimes v) = \bar{x} \otimes v$, and $V = (V_{\mathbb{C}})^{\sigma = \mathrm{id}}$. The ($\mathcal{U}$-stable) real subspaces $W \subseteq V$ are in bijection with the ($\mathcal{U}$-stable and) σ-stable complex subspaces $W' \subseteq V_{\mathbb{C}}$, with bijection given by

$$W \longmapsto \mathbb{C} \otimes_{\mathbb{R}} W, \qquad (W')^{\sigma = \mathrm{id}} \longleftarrow W'.$$

Consider now the decomposition

$$V_{\mathbb{C}} = \bigoplus_{k \in \mathbb{Z}} V_k \tag{10.5}$$

obtained before for complex representations. Because of the simple relation

$$\sigma(z^k v) = \bar{z}^k \sigma(v),$$

we have $\sigma(V_k) = V_{-k}$. Thus $(V_k \oplus V_{-k})^{\sigma = \mathrm{id}} \subseteq V$ is a $\mathcal{U}$-stable subspace. Since V was assumed irreducible, there is a unique $k \geq 0$ with $V = V_k + V_{-k}$.

First case: $V_{\mathbb{C}} = V_0$. Then $\mathcal{U}$ acts trivial on $V_{\mathbb{C}}$. Hence V was a trivial 1-dimensional representation by irreducibility, meaning $V \cong W_0$.

Second case: $V_{\mathbb{C}} = V_k \oplus V_{-k}$ *with* $k > 0$. Pick any $0 \neq v \in V_k$. Then $0 \neq \sigma(v) \in V_{-k}$ and $\mathbb{C}v + \mathbb{C}\sigma(v)$ is $\mathcal{U}$-stable and σ-stable. By irreduciblity of V,

$$V = (\mathbb{C}v + \mathbb{C}\sigma(v))^{\sigma=\mathrm{id}}$$

which is isomorphic to W_k. □

Definition 10.6. For a complex representation V as in (10.4), the integers k with $V_k \neq 0$ are called the *weights* of V, and (10.4) is called the *weight decomposition.* For a real representation V we define the weights of V to be those of $V_{\mathbb{C}}$.

We have now discussed all notions that go into (SV 1): We consider homomorphisms $h : \mathbb{S} \to G_{\mathbb{R}}$ such that $\mathrm{ad} \circ h$ factors over the quotient $\mathbb{S} \to U(1)$. We require that the weights of the representation of $U(1)$ on $\mathrm{Lie}(G_{\mathbb{R}})$ are all contained in $\{-1, 0, 1\}$.

10.3. **Cartan involutions.** We call a real linear algebraic group G *compact* if $G(\mathbb{R})$ is a compact Lie group. For example, the special orthogonal group $\mathrm{SO}(V, Q)$ of a vector space V with positive or negative definite quadratic form Q is compact. An example is $U(1) = SO(2)$.

Definition 10.7. (1) In general, an *involution* of an object is an automorphism τ such that $\tau^2 = \mathrm{id}$.

(2) Let G be a connected linear algebra group over $\mathbb{R}$. A *Cartan involution* is an involution $\theta : G \to G$ (automorphism as algebraic group over $\mathbb{R}$) such that

$$G^{(\theta)}(\mathbb{R}) := \{g \in G(\mathbb{C}) \mid \theta(\bar{g}) = g\}$$

is a compact.

Theorem 10.8 ([13, Theorem 1.16])**.** *Let G be a connected algebraic group over $\mathbb{R}$. Cartan involutions exist if and only if G is reductive. Any two Cartan involutions are conjugate by an element of $G(\mathbb{R})$.*

Example 10.9. (1) On $\mathbb{G}_m$, consider the involution $\theta(t) = t^{-1}$. Then

$$\mathbb{G}_m^{(\theta)}(\mathbb{R}) \;=\; \{z \in \mathbb{C}^\times \mid \bar{z}^{-1} = z\} \;=\; \mathcal{U}$$

which is compact.

(2) More generally, on GL_n, consider the involution $\theta(g) = g^{t,-1}$. Note that $(gh)^t = h^t g^t$ and $(gh)^{-1} = h^{-1}g^{-1}$, so the composition of both is really a group automorphism. We find that

$$\mathrm{GL}_n^{(\theta)}(\mathbb{R}) \;=\; \{g \in \mathrm{GL}_n(\mathbb{C}) \mid \bar{g}^t = g^{-1}\} \;=\; U(n)(\mathbb{R})$$

is the unitary group of the standard positive definite hermitian space $(\mathbb{C}^n, 1_n)$ which is compact.

We can now clarify (SV 2). Given $h : \mathbb{S} \to G_{\mathbb{R}}$, define $\theta : G_{\mathbb{R}} \to G_{\mathbb{R}}$ by

$$\theta(g) := h(i)gh(i)^{-1}.$$

The first part of axiom (SV 1) implies that $h(-1)$ is central. So θ is an involution because

$$\theta(\theta(g)) = h(-1)gh(-1) \overset{h(-1)\text{ central}}{=} h(-1)h(-1)g = g.$$

10.4. **Example:** GL_2**.** We continue from Example 1.1. That is, we embed $\mathbb{C}$ into $\mathrm{M}_2(\mathbb{R})$ by

$$h(a+bi) := \begin{pmatrix} a & -b \\ b & a \end{pmatrix}. \tag{10.6}$$

Applying the formalism of unit group schemes from §9.4, we obtain a Deligne homomorphism

$$h : \mathbb{S} \longrightarrow \mathrm{GL}_{2,\mathbb{R}}.$$

It is clear that (SV 3) holds; our task is to verify (SV 1) and (SV 2).

Verification of (SV 1). The image $h(\mathbb{G}_{m,\mathbb{R}})$ is the subgroup of scalar matrices in $\mathrm{GL}_{2,\mathbb{R}}$. These agree with $Z(\mathrm{GL}_{2,\mathbb{R}})$ and there is nothing further to check for the first part. For the second part, let $\sigma \in \mathrm{M}_2(\mathbb{R})$ be an element that satisfies

$$\sigma \cdot h(a+bi) = h(a-bi) \cdot \sigma$$

for all $a+bi \in \mathbb{C}$. For example, we could choose

$$\sigma = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix},$$

but every multiple $h(z)\cdot\sigma$, $z \neq 0$ will work as well because $h(z)$ commutes with $h(\mathbb{C})$. Then we have the vector space decomposition

$$\mathrm{M}_2(\mathbb{R}) = h(\mathbb{C}) \oplus h(\mathbb{C}) \cdot \sigma.$$

Conjugation by an element $h(z)$ is given by

$$\begin{aligned} h(z) \cdot \big(h(w_1) + h(w_2)\sigma\big) \cdot h(z)^{-1} &= h(z)h(w_1)h(z)^{-1} + h(z)h(w_2)\sigma h(z)^{-1} \\ &= h(w_1) + h(z)h(w_2)h(\bar{z})^{-1}\sigma \\ &= h(w_1) + h(z/\bar{z})h(w_2)\sigma. \end{aligned}$$

In particular, we see that $\mathcal{U} \subset \mathbb{S}(\mathbb{R})$ acts as

$$h(z) \cdot \big(h(w_1) + h(w_2)\sigma\big) \cdot h(z)^{-1} = h(w_1) + h(z^2 w_2)\sigma, \quad z \in \mathcal{U}.$$

Thus, in terms of the classification from Proposition 10.5, we have that

$$\mathrm{Lie}(\mathrm{GL}_{2,\mathbb{R}}) \xrightarrow{\sim} W_0^{\oplus 2} \oplus W_2$$

as representation $\mathrm{ad} \circ (h|_{\mathcal{U}})$. We see that $\{\pm 1\} \subset \mathcal{U}$ acts trivially, which we already knew from $h(\mathbb{G}_{m,\mathbb{R}}) \subset Z(G_{\mathbb{R}})$. If we descend $\mathrm{ad} \circ (h|_{\mathcal{U}})$ along the map $\mathcal{U} \to \mathcal{U}$, $z \mapsto z^2$, then we find that

$$\mathrm{ad} \circ \bar{h} \ \cong \ W_0^{\oplus 2} \oplus W_1.$$

That is, all occurring weights are $-1, 0$ and 1, as we needed to show.

Verification of (SV 2): Conjugation by $h(i)$ defines the involution

$$\theta \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} & 1 \\ -1 & \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \tag{10.7}$$

The matrices $g \in \mathrm{GL}_2(\mathbb{C})$ with $\theta(g) = \bar{g}$ are precisely the ones of the form

$$\mathrm{GL}_2^{(\theta)}(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \,\middle|\, a, b \in \mathbb{C} \right\}.$$

This group is not compact; for example, $\{\mathrm{diag}(a, \bar{a})\}$ is a closed subspace isomorphic to $\mathbb{C}$ which is not compact. In particular, θ is not a Cartan involution on $\mathrm{GL}_{2,\mathbb{R}}$. However, axiom (SV 2) only required θ to be a Cartan involution for $\mathrm{PGL}_{2,\mathbb{R}}$. Equivalently, because the map $\mathrm{SL}_{2,\mathbb{R}} \to \mathrm{PGL}_{2,\mathbb{R}}$ is finite flat of degree 2 (its kernel is μ_2), we may check that θ is a Cartan involution of $\mathrm{SL}_{2,\mathbb{R}}$. We see

$$\mathrm{SL}_2^{(\theta)}(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \in \mathrm{GL}_2^{(\theta)}(\mathbb{R}) \;\middle|\; |a|^2 + |b|^2 = 1 \right\}.$$

The condition $|a|^2 + |b|^2 = 1$ defines a compact subset of $\mathbb{C}^2$, so $\mathrm{SL}_2^{(\theta)}(\mathbb{R})$ is compact as was to be shown.

In conclusion, we have shown that (GL_2, X) with $X = \mathrm{GL}_2(\mathbb{R}) \cdot h$ is a Shimura datum.

10.5. **The complex structure on X.** Given $h \in X$, we have a stabilizer

$$K_h := \{g \in G(\mathbb{R}) \mid gh(z)g^{-1} = h(z) \text{ for all } z \in \mathbb{C}\}.$$

This is a closed subgroup of $G(\mathbb{R})$ because of each the conditions $gh(z)g^{-1} = h(z)$ is closed. In particular, K_h is a Lie group (closed subgroup theorem). Recall that the action of $G(\mathbb{R})$ on X is transitive by definition, so we may make X into a smooth manifold by identifying it with the quotient

$$G(\mathbb{R})/K_h \xrightarrow{\sim} X, \quad gK_h \longmapsto g \cdot h.$$

By the following theorem, X has only finitely many connected components.

Theorem 10.10 ([13, Corollary 5.3])**.** *Let V be a variety over $\mathbb{R}$. Then $V(\mathbb{R})$ has only finitely many connected components.*

Recall that a smooth manifold M has a tangent bundle TM. This is the vector bundle $\pi : TM \to M$ whose fiber over $x \in M$ is the tangent space T_xM of M at x. Equivalently, one may take a Grothendieck perspective and consider the sheaf of smooth sections

$$\mathcal{T}_M : U \longmapsto TM(U).$$

That is, for every open $U \subset M$, $TM(U)$ denotes the set of smooth maps $s : U \to TM$ such that $s \circ \pi = \mathrm{id}_U$. Let $\mathcal{C}^\infty(-, \mathbb{R})$ denote the sheaf of smooth real valued maps on M. Then $\mathcal{T}_M$ is a locally free $\mathcal{C}^\infty(-, \mathbb{R})$-module of rank equal to $\dim(M)$. It does not matter which of the two perspectives one takes in the following discussion.

Definition 10.11. An *almost complex structure* on a smooth manifold M is a vector bundle endomorphism $J \in \mathrm{End}_{\mathcal{C}^\infty(-,\mathbb{R})}(\mathcal{T}_M)$ such that $J^2 = -1$.

On any smooth manifold, we also have the sheaf $\mathcal{C}^\infty(-, \mathbb{C})$ of smooth maps to the complex plane. Assume that M is even a complex manifold. We define $\mathcal{T}_M$ as before by viewing M as a smooth manifold. But the complex structure makes $\mathcal{T}_M$ into a locally free $\mathcal{C}^\infty(-, \mathbb{C})$-module of rank equal to $\dim_{\mathbb{C}}(M)$. Multiplication by the constant section $i \in \mathcal{C}^\infty(M, \mathbb{C})$ defines an almost complex structure $J \in \mathrm{End}_{\mathcal{C}^\infty(-,\mathbb{R})}(\mathcal{T}_M)$. In this way, one obtains a fully faithful functor

$$\{\text{Complex manifolds}\} \longleftrightarrow \left\{ \begin{matrix} \text{Smooth manifolds with} \\ \text{almost complex structure} \end{matrix} \right\}$$

where morphisms in the target are smooth maps that are compatible with almost complex structures. An almost complex structure is said to be *integrable* if it comes from a complex structure. Fully faithfulness means that this structure is unique if it exists.

Our next goal is to endow X with an almost complex structure. Let $Z = Z(G(\mathbb{R}))$ be the center of $G(\mathbb{R})$. It is contained in K_h, and we denote by $\mathfrak{z} \subseteq \mathfrak{k}_h \subseteq \mathfrak{g}$ the Lie algebras of $Z \subseteq K_h \subseteq G(\mathbb{R})$. Using axiom (SV 1), we have a weight decomposition

$$\mathfrak{g} = \mathfrak{g}_h^0 \oplus \mathfrak{g}_h^{\pm 1} \tag{10.8}$$

with respect to $\mathrm{ad} \circ \bar{h}$ where at most the weights 0 and ± 1 occur. For every $g \in G(\mathbb{R})$, we have

$$\mathfrak{g}^0_{g\cdot h} = \mathrm{ad}(g)(\mathfrak{g}^0_h), \qquad \mathfrak{g}^{\pm 1}_{g\cdot h} = \mathrm{ad}(g)(\mathfrak{g}^{\pm 1}_h).$$

Lemma 10.12. *There is the equality $\mathfrak{g}^0_h = \mathfrak{k}_h$.*

Proof. The inclusion $\mathfrak{k}_h \subseteq \mathfrak{g}^0_h$ is clear because $h(z)$ centralizes K_h for every $z \in \mathbb{C}^\times$. For the converse, we recall that for every $\gamma \in \mathfrak{g}$, there is a unique group homomorphism

$$\exp_\gamma : \mathbb{R} \to G(\mathbb{R}), \quad t \longmapsto \exp_\gamma(t)$$

such that

$$\frac{d}{dt}\Big|_{t=0} \exp_\gamma(t) = \gamma.$$

It satisfies $\exp_{\mathrm{ad}(g)(\gamma)}(t) = g \exp_\gamma(t) g^{-1}$ by uniqueness. By axiom (SV 1), we have $h(\mathbb{C}^\times) \subseteq Z \cdot h(\mathcal{U})$. So if $\gamma \in \mathfrak{g}^0_h$, then $\mathrm{ad}(h(z))(\gamma) = \gamma$ for every $z \in \mathbb{C}^\times$ and hence

$$\begin{aligned} h(z) \cdot \exp_\gamma(t) h(z)^{-1} &= \exp_{\mathrm{ad}(h(z))(\gamma)}(t) \\ &= \exp_\gamma(t). \end{aligned}$$

This means that $\exp_\gamma$ has image in K_h, and thus $\gamma \in \mathfrak{k}$. $\square$

For every $h \in X$, this constructs an identification

$$\mathfrak{g}^{\pm 1}_h \xrightarrow{\sim} \mathfrak{g}/\mathfrak{k}_h \xrightarrow{\sim} T_h X.$$

By definition, the $\mathcal{U}$-representation $\mathfrak{g}^{\pm 1}_h$ is isomorphic to a direct sum of copies of $W_1 = (\mathbb{C}, \rho_1(z)(v) = zv)$. That is, $J := \mathrm{ad} \circ \bar{h}(i)$ is an endomorphism of $\mathfrak{g}^{\pm 1}_h$ with $J^2 = -1$. This construction, which we have only described for an individual $h \in X$, is compatible as h varies and endows TX with an almost complex structure. This structure is integrable, i.e. comes from a (unique) complex manifold structure on X.

10.6. **The hermitian structure on X.** So far, we have only used axiom (SV 1). Axiom (SV 2) is used to endow X with a hermitian structure. We note upfront that we have a subgroup

$$G(\mathbb{R})/Z \hookrightarrow G^{\mathrm{ad}}(\mathbb{R}) \tag{10.9}$$

which is a union of connected components. Indeed, the quotient map $G \to G^{\mathrm{ad}}$ is smooth and surjective, so (10.9) is submersive near the identity (surjective on tangent spaces), hence a diffeomorphism near the identity because both have the same dimension. So the image of (10.9) contains the identity connected component $G^{\mathrm{ad}}(\mathbb{R})^\circ$. But in general, the map $G(\mathbb{R}) \to G^{\mathrm{ad}}(\mathbb{R})$ is not surjective:

Example 10.13. The adjoint group of SL_n is PGL_n, but the natural map

$$\mathrm{SL}_n(\mathbb{R}) \longrightarrow \mathrm{PGL}_n(\mathbb{R}) = \mathrm{GL}_n(\mathbb{R})/\mathbb{R}^\times$$

is surjective if and only if n is odd. If n is even, then the target has two connected components distinguished by the sign of the determinant. The image consists of the elements $\mathbb{R}^\times \cdot g$ with positive determinant.

Let $\theta : G_\mathbb{R} \to G_\mathbb{R}$ denote conjugation by $h(i)$ (an involution). The centralizer K_h equals

$$G(\mathbb{R}) \cap G^{(\theta)}(\mathbb{R}) = \{g \in G(\mathbb{R}) \mid \theta(g) = g\}.$$

The image K_h/Z in $G^{\mathrm{ad}}(\mathbb{R})$ is then a closed subgroup of

$$G^{\mathrm{ad},(\theta)}(\mathbb{R}) = \{g \in G^{\mathrm{ad}}(\mathbb{C}) \mid \theta(g) = \bar{g}\}$$

which, by axiom (SV 2), is compact. So K_h/Z is a compact subgroup of $G(\mathbb{R})/Z$.

Definition 10.14. A *hermitian pairing* on a complex vector bundle $E \to M$ is a smooth map

$$(\ , \ ) : E \times_M E \longrightarrow \mathbb{C}$$

that is conjugate $\mathbb{C}$-linear in the first coordinate, $\mathbb{C}$-linear in the second, satisfies $(s_1, s_2) = \overline{(s_2, s_1)}$ for all $s_1, s_2 \in E(M)$, and is positive definite in the sense that $(s, s)(x) > 0$ for all $s \in E(M)$ and $x \in M$ with $s(x) \neq 0$. A *hermitian structure* on a complex manifold is a hermitian pairing on its tangent bundle.

Using (SV 2), we can define a hermitian structure on X as follows. Fixing $h \in X$, choose a positive definite hermitian form

$$(\ , \ )_{h,0} : T_hX \times_X T_hX \longrightarrow \mathbb{C}.$$

Let dk denote a translation invariant measure on K_h/Z; such a measure exists and is unique up to scaling by $\mathbb{R}_{>0}$. Using that K_h/Z is compact, we may *average* the hermitian from $(\ , \ )_{h,0}$ to define

$$(v, w)_h := \int_{K_h/Z} (kv, kw)_{h,0} \ dk. \tag{10.10}$$

This form is still hermitian and positive definite. It is additionally K_h-invariant because the measure dk is translation invariant. For every $h' \in X$, there exists an element $g \in G(\mathbb{R})$ with $gh = h'$. Translation by g induces an isomorphism

$$dg : T_hX \xrightarrow{\sim} T_{gh}X.$$

We define a hermitian form $(\ , \ )_{gh}$ on $T_{gh}X$ by

$$(v, w)_{gh} := ((dg)^{-1}v, (dg)^{-1}w).$$

This is well-defined, because g is unique up to $k \in K_h$ and $(\ , \ )_h$ is K_h-invariant. One may check (omitted) that this fiberwise construction defines a hermitian pairing on TX. Moreover, this hermitian structure is $G(\mathbb{R})$-invariant: For every pair of vector fields $s_1, s_2 \in TX(X)$ and every $g \in G(\mathbb{R})$,

$$((dg)(s_1), \ (dg)(s_2)) = g_*(s_1, s_2).$$

By general classification theorems, one can show that every connected component $X^\circ \subseteq X$ is a hermitian symmetric domain. That is, X° is a complex manifold with hermitian structure such that

- For any two points $p, q \in X^\circ$, there exists a holomorphic isometry $f \in \mathrm{Aut}(X^\circ)$ such that $f(p) = q$.
- For every point p, there exists a holomorphic isometry $f \in \mathrm{Aut}(X^\circ)$ with $f(p) = p$ and such that $df : T_pX^\circ \to T_pX^\circ$ is multiplication by -1.

In this context, we also have the following property.

- For every $h \in X$, the subgroup $K_h^\circ := K_h/Z \cap G^{\mathrm{ad}}(\mathbb{R})^\circ$ is a maximal compact subgroup of $G^{\mathrm{ad}}(\mathbb{R})^\circ$. That is, if K' is a compact subgroup with $K_h^\circ \subseteq K'$, then $K_h^\circ = K'$.

We will further develop these topics as needed. For now, we refer the curious reader to [13, §1 – §5].

Example 10.15. Let $h : \mathbb{S} \to \mathrm{GL}_{2,\mathbb{R}}$ be as in (10.6). Then K_h° is the image of the circle $\mathrm{SO}(2) \subset \mathrm{SL}_2(\mathbb{R})$ in $\mathrm{PGL}_2(\mathbb{R})^\circ$. We claim that this is a maximal compact subgroup and that every other maximal compact subgroup is conjugate to it. Since the projection map $\mathrm{SL}_2(\mathbb{R}) \to \mathrm{PGL}_2(\mathbb{R})^\circ$ is 2-to-1, it suffices to show that $\mathrm{SO}(2) \subset \mathrm{SL}_2(\mathbb{R})$ is maximal compact and has the uniqueness up to conjugation property.

Step 1: Every compact subgroup $K \subset \mathrm{SL}_2(\mathbb{R})$ is contained in $\mathrm{SO}(\lambda)$ for some positive definite quadratic form λ on $\mathbb{R}^2$. First note that K is a Lie group by the closed subgroup theorem. Start with any positive definite quadratic form λ_0 on $\mathbb{R}^2$ and average it as in (10.10),

$$\lambda(v,w) = \int_K \lambda_0(v,w)\, dk.$$

Then λ is K-invariant, meaning $K \subseteq \mathrm{SO}(\lambda)$.

Step 2: All $\mathrm{SO}(\lambda)$ *are conjugate.* Since all positive definite quadratic forms on $\mathbb{R}^2$ are isometric, there exists $g \in \mathrm{GL}_2(\mathbb{R})$ with $g\mathrm{SO}(\lambda)g^{-1} = \mathrm{SO}(2)$. If $\det(g) > 0$, then we may scale g by $\sqrt{\det(g)}^{-1}$ to assume $g \in \mathrm{SL}_2(\mathbb{R})$. If $\det(g) < 0$, then we can scale after modifying g with an element from $(2) \setminus \mathrm{SO}(2)$. In this way, we see that $\mathrm{SO}(\lambda)$ and $\mathrm{SO}(2)$ are conjugate in $\mathrm{SL}_2(\mathbb{R})$.

Step 3: $\mathrm{SO}(2)$ *is maximal among compact subgroups of* $\mathrm{SL}_2(\mathbb{R})$. Let $\mathrm{SO}(2) \subseteq K$ be a compact subgroup containing $\mathrm{SO}(2)$. By Steps 1 and 2, $K \subseteq g\mathrm{SO}(2)g^{-1}$ for some $g \in \mathrm{SL}_2(\mathbb{R})$. We necessarily have $\mathrm{SO}(2) = g\mathrm{SO}(2)g^{-1}$ because both are connected Lie subgroups and hence uniquely determined by their Lie algebras as subspaces of $\mathrm{Lie}(\mathrm{SL}_2(\mathbb{R}))$. This shows $\mathrm{SO}(2) = K$.

10.7. **General examples and remarks.** We have now discussed Shimura data in some detail. Let us collect a few easy remarks and examples.

Definition 10.16 (Tori)**.** An algebraic *torus* over a field k is an algebraic group T/k such that $\bar{k} \otimes_k T \cong \mathbb{G}_{m,\bar{k}}^d$ for some $d \geq 0$. In particular, tori are affine, connected, and of finite type.

Tori in characteristic 0 are the same as conncected *commutative* reductive groups.

Example 10.17. If K/k is a finite separable field extension, then $T = \mathrm{Res}_{K/k}(\mathbb{G}_{m,K})$, which agrees with $\underline{K}^\times$, is a torus over k by Proposition 9.17. In general, Weil restrictions of tori over K are tori over k.

To give a concrete example, let $\mathbb{Q}(i)/\mathbb{Q}$, set $T = \underline{\mathbb{Q}(i)}^\times$ and consider the exact sequence

$$1 \longrightarrow T^1 \longrightarrow T \overset{N_{\mathbb{Q}(i)/\mathbb{Q}}}{\longrightarrow} \mathbb{G}_{m,\mathbb{Q}} \longrightarrow 1. \tag{10.11}$$

Both T^1 is the group scheme

$$T^1 = \mathrm{Spec}\, \mathbb{Q}[x,y]/(x^2+y^2-1);$$

it provides one of the many possibilities of defining $U(1)/\mathbb{R}$ over $\mathbb{Q}$. Then T^1 and T are tori over $\mathbb{Q}$. After base change to $\mathbb{Q}(i)$, (10.11) becomes isomorphic to

$$1 \longrightarrow \mathbb{G}_m \overset{t\mapsto(t,t^{-1})}{\longrightarrow} \mathbb{G}_m \times_k \mathbb{G}_m \overset{(x,y)\mapsto xy}{\longrightarrow} \mathbb{G}_m \longrightarrow 1.$$

Example 10.18 (Shimura data for tori)**.** Let $T/\mathbb{Q}$ be a torus and let $h : \mathbb{S} \to T_{\mathbb{R}}$ be *any* morphism. Then $(T, \{h\})$ is a Shimura datum. Indeed (SV 1)–(SV 3) essentially only concern the adjoint group $T^{\mathrm{ad}} = \{1\}$ and are trivially satisfied. Note that (SV 3) explicitly requires factors to be $\neq \{1\}$.

Example 10.19 (Twisting a Shimura datum by central homomorphisms)**.** Let (G, X) be any Shimura datum, and let $h_0 : \mathbb{S} \to Z(G)_{\mathbb{R}}$ be any homomorphism to the center of $G_{\mathbb{R}}$. For any $h \in X$, the product $(h_0 h)(z) := h_0(z)h(z)$ defines a new Deligne homomorphism. Since the axioms (SV 1)–(SV 3) essentially only concern the adjoint group, the pair $(G, h_0 \cdot X)$ defines a new Shimura datum.

In particular, this example shows that a Shimura datum for G is more information than just a X viewed as complex manifold with $G(\mathbb{R})$-action. For example, let (GL_2, X) be our usual Shimura datum for GL_2 (§10.4). Let h_0 be the homomorphism

$$h_0 : \mathbb{S} \overset{N_{\mathbb{C}/\mathbb{R}}}{\longrightarrow} \mathbb{G}_{m,\mathbb{R}} = Z(\mathrm{GL}_{2,\mathbb{R}}).$$

Then $(\mathrm{GL}_2, h_0 \cdot X)$ is a new Shimura datum, but the resulting complex manifold datum is still $\mathrm{GL}_2(\mathbb{R})$ acting on $\mathbb{H}^{\pm}$. In this sense, calling $(\mathrm{GL}_2, \mathbb{H}^{\pm})$ a Shimura datum in the first part of the lecture was an abuse of notation.

Example 10.20 (Passing to the adjoint group)**.** If (G, X) is a Shimura datum and if $T \subseteq Z(G)$ is a subgroup of the center, then $(G/T, (G/T)(\mathbb{R}) \cdot \bar{h})$ for any $h \in X$ is a new Shimura datum. Here $\bar{h} = [G_{\mathbb{R}} \to (G/T)_{\mathbb{R}}] \circ h$ is the composition of h with the quotient map. For example, we always have the adjoint Shimura datum $(G^{\mathrm{ad}}, G^{\mathrm{ad}}(\mathbb{R}) \cdot \bar{h})$. The reason this construction works is that the Shimura variety axioms essentially only depend on the composition $\mathbb{S} \to G_{\mathbb{R}} \to G^{\mathrm{ad}}_{\mathbb{R}}$.

Example 10.21 (SL_2 does not admit a Shimura datum)**.** By the previous example, we obtain a Shimura datum (PGL_2, X) for PGL_2 from our datum for GL_2. But there is no Shimura datum for SL_2, even though the two groups are closely related by $\mathrm{PGL}_2 = \mathrm{SL}_2 / \{\pm 1\}$.

Sketch: Indeed, we have $Z(\mathrm{SL}_2) = \{\pm 1\}$. So any homomorphism $h : \mathbb{S} \to \mathrm{SL}_{2,\mathbb{R}}$ with $h(\mathbb{G}_{m,\mathbb{R}}) \subseteq Z(\mathrm{SL}_{2,\mathbb{R}})$ has to be trivial on $\mathbb{G}_m$ and hence factor through the quotient $q : \mathbb{S} \to U(1)$. But for any $h_0 : U(1) \to \mathrm{SL}_{2,\mathbb{R}}$, if we look at the composition $\mathrm{ad} \circ h_0 \circ q|_{U(1)}$, we will find only weights $\in 4\mathbb{Z}$ in $\mathrm{Lie}(\mathrm{SL}_{2,\mathbb{R}})$. This means that after descending along $U(1) \to U(1)$, $z \mapsto z^2$, we can have only even weights. So (SV 1) can never be satisfied. (Exercise: Fill in the details of this argument and compare with 10.4.)

Example 10.22 (Dimension of symmetric space)**.** Assume that $G = \mathrm{GL}_n$ admits some Shimura datum X. By what was said at the end of §10.6, for every $h \in X$, the group $K_h^\circ := (K_h/\mathbb{R}^\times)^\circ \subset \mathrm{PGL}_n(\mathbb{R})^\circ$ would be a maximal compact subgroup. By the same argument as in Example 10.15, the maximal compact subgroups in $\mathrm{PGL}_n(\mathbb{R})^\circ$ are the conjugates of the image of $\mathrm{SO}(n)$. (Exercise: Check this.) So we see that

$$\dim_{\mathbb{R}}(X) = (n^2 - 1) - \dim_{\mathbb{R}} \mathrm{SO}(n).$$

In Example 10.2, we saw that $\mathrm{Lie}(SO(n))$ is the vector space of skew-symmetric $(n \times n)$-matrices. So its dimension is $n(n-1)/2$. Thus

$$\dim_{\mathbb{R}}(X) = n(n+1)/2 - 1.$$

If $n \equiv 3, 0 \bmod 4$, then this dimension is odd, so X cannot be a complex manifold. In fact, GL_n with $n \geq 3$ can never define a Shimura datum because there is no Cartan involution on PGL_n which is given by conjugation with a group element (compare with Example 10.9).

References

[1] Andreatta, Fabrizio; Goren, Eyal Z.; Howard, Benjamin; Madapusi, Keerthi; *Faltings heights of abelian varieties with complex multiplication*, Ann. of Math. (2) **187** (2018), no. 2, 391–531.

[2] Deligne, Pierre; *Travaux de Shimura*, Séminaire Bourbaki, 23ème année (1970/1971), Exp. No. 389, pp. 123–165. Lecture Notes in Math. **244**, Springer-Verlag, Berlin-New York, 1971.

[3] Faltings, Gerd; *Finiteness theorems for abelian varieties over number fields*, Invent. Math. **73** (1983), no. 3, 349–366.

[4] Forster, O., *Lectures on Riemann surfaces*, Graduate Texts in Mathematics **81**, Springer-Verlag, New York, 1981.

[5] Gan, Wee Teck; Gross, Benedict H.; Prasad, Dipendra; *Symplectic local root numbers, central critical L-values, and restriction problems in the representation theory of classical groups* in Sur les conjectures de Gross et Prasad. I, Astérisque **346** (2012), 1–109.

[6] Gross, Benedict H.; Zagier, Don B.; *Heegner points and derivatives of L-series*, Invent. Math. **84** (1986), no. 2, 225–320.

[7] Haines, Thomas (ed.); Harris, Michael (ed.); *Shimura varieties*, London Math. Soc. Lecture Note Ser., No. 457, Cambridge University Press, Cambridge, 2020.

[8] Jelonek, Zbigniew; *Simple examples of affine manifolds with infinitely many exotic models*, Adv. Math. **284** (2015), 112–121.

[9] Kudla, Stephen; Rapoport, Michael; *Special cycles on unitary Shimura varieties II: Global theory*, J. Reine Angew. Math. **697** (2014), 91–157.

[10] Mihatsch, Andreas; *Lecture notes on moduli spaces of elliptic curves*, https://amihatsch.github.io/ref/EC.pdf.

[11] Milne, James S.; *The action of an automorphism of $\mathbb{C}$ on a Shimura variety and its special points* in *Arithmetic and geometry*, Vol. I, 239–265. Progr. Math. **35**, Birkhäuser Boston, Inc., Boston, MA, 1983.

[12] ———; *Canonical models of Shimura curves*, lecture notes available at https://www.jmilne.org/math/articles/2003a.pdf, 2003.

[13] ———; *Introduction to Shimura varieties*, lecture notes available at https://www.jmilne.org/math/xnotes/svi.html, 2017.

[14] ———; *Algebraic Groups. The theory of group schemes of finite type over a field*, book available at https://www.jmilne.org/math/Books/iAG2017.pdf, 2017.

[15] Mumford, D., *Abelian varieties*, Tata Institute of Fundamental Research, Mumbai, Corrected Reprint, 2012.

[16] Serre, Jean-Pierre; *A course in arithmetic*, Grad. Texts in Math., No. 7, Springer-Verlag, New York-Heidelberg, 1973.

[17] Silverman, Joseph H.; *The arithmetic of elliptic curves*, Grad. Texts in Math. **106**, Springer-Verlag, New York, 1986.

[18] Yuan, Xinyi; Zhang, Shou-Wu; *On the averaged Colmez conjecture*, Ann. of Math. (2) **187** (2018), no. 2, 533–638.

[19] Zhang, Wei; *On arithmetic fundamental lemmas*, Invent. Math. **188** (2012), no. 1, 197–252.

[20] Zhu, Yihang; *Introduction to the Langlands–Kottwitz method* in [7], 115–150.